The Annals of Probability
2000, Vol. 28, No. 3, 1371–1383

AN INDEPENDENCE PROPERTY FOR THE PRODUCT OF GIG AND GAMMA LAWS[1]

BY GÉRARD LETAC AND JACEK WESOŁOWSKI

Université Paul Sabatier and Politechnika Warszawska

Matsumoto and Yor have recently discovered an interesting transformation which preserves a bivariate probability measure which is a product of the generalized inverse Gaussian (GIG) and gamma distributions. This paper is devoted to a detailed study of this phenomenon. Let X and Y be two independent positive random variables. We prove (Theorem 4.1) that $U = (X+Y)^{-1}$ and $V = X^{-1} - (X+Y)^{-1}$ are independent if and only if there exists $p, a, b > 0$ such that Y is gamma distributed with shape parameter p and scale parameter $2a^{-1}$, and such that X has a GIG distribution with parameters $-p$, a and b (the direct part for $a = b$ was obtained in Matsumoto and Yor). The result is partially extended (Theorem 5.1) to the case where X and Y are valued in the cone V_+ of symmetric positive definite (r, r) real matrices as follows: under a hypothesis of smoothness of densities, we prove that $U = (X+Y)^{-1}$ and $V = X^{-1} - (X+Y)^{-1}$ are independent if and only if there exists $p > (r-1)/2$ and a and b in V_+ such that Y is Wishart distributed with shape parameter p and scale parameter $2a^{-1}$, and such that X has a matrix GIG distribution with parameters $-p$, a and b. The direct result is also extended to singular Wishart distributions (Theorem 3.1).

1. Introduction. If p, a and b are positive numbers, consider two independent positive random variables X and Y with respective distributions

$$\mu_{-p,a,b}(dx) = C_1 x^{-p-1} \exp\left(-\tfrac{1}{2}(ax + bx^{-1})\right) \mathbf{1}_{(0,+\infty)}(x)\,dx, \tag{1.1}$$

$$\gamma_{p,2a^{-1}}(dy) = C_2 y^{p-1} \exp\left(-\tfrac{1}{2}ay\right) \mathbf{1}_{(0,+\infty)}(y)\,dy. \tag{1.2}$$

Thus, the distribution of X is the generalized inverse Gaussian distribution (GIG) with parameters $(-p, a, b)$ and the distribution of Y is the gamma distribution with shape parameter p and scale parameter $2a^{-1}$.

Denote $U = (X+Y)^{-1}$ and $V = X^{-1} - (X+Y)^{-1}$. The present paper is motivated by the recent observation due to Matsumoto and Yor (1998) (see Proposition 9.1 in their paper) in the case $a = b$, which is that $(X/Y, 1/X) =^d (X/Y, X+Y)$ (where $=^d$ means identically distributed). The fact that $1/X =^d X+Y$ has already been used by Letac and Seshadri (1983) to characterize the GIG law as the distribution of a random continued fraction whose entries are i.i.d. gamma random variables.

In terms of U and V, the above result of Matsumoto and Yor can be reformulated by saying that $(X/Y, 1/X) =^d (U/V, 1/U)$, or, applying the

Received June 1999; revised January 2000.
[1]Supported by the Polonium project 98031.
AMS 1991 *subject classifications*. Primary 60E10; secondary 62E10.
Key words and phrases. Generalized inverse Gaussian distributions, Wishart distributions, Matsumoto–Yor property.

transformation $(x, y) \mapsto (1/y, xy)$, by saying that $(X, Y) =^d (U, V)$. Under this last form, the result can be extended to the case where a and b are not necessarily equal. This extension is given by the following statement: U and V are independent with respective distributions $\mu_{-p,b,a}(dx)$ and $\gamma_{p,2b^{-1}}(dy)$. We skip its proof, which is easily obtained by computing a Jacobian: we let $r = 1$ in the first proof of Theorem 3.1 below to obtain the complete details.

What is especially remarkable in the above statement is the independence of U and V. The first aim of the present paper is to prove the converse: if U and V are independent, then X and Y are, respectively, GIG and gamma. This is proved in Theorem 4.1 below, using a simple differential equation to characterize the gamma distribution, but relying on the characterization of the GIG distribution by Letac and Seshadri (1983) to get the second part of this simultaneous characterization. We thank V. Seshadri for several comments about this problem.

A second aim is to extend the Matsumoto–Yor result to the symmetric matrices, as suggested by these authors in their Section 11. Actually, a natural frame for considering the GIG distribution and the gamma distribution is rather the cone of positive definite matrices (and more generally a symmetric cone). For gamma, it is a classical fact that the Wishart distributions are the natural extension of the one-dimensional gamma distributions: see Massam (1994), Casalis and Letac (1996) and Letac and Massam (1998). For the matrix GIG laws, an important paper is the monumental work of Bernadac (1995), who has extended the characterization of Letac and Seshadri (1983) to random continued fractions on these cones [by considering $Y_1, Y_1 + Y_2^{-1}, Y_1 + (Y_2 + Y_3^{-1})^{-1}, \dots$ where the random matrices Y_j are independent and Wishart distributed]. For a recent work on the GIG distributions and some bibliography about it, see Butler (1998). We prove the Matsumoto–Yor result mentioned above and its extension for different a and b in the space of symmetric matrices (Theorem 3.1). We also offer a partial converse in Theorem 5.1 with a hypothesis of smoothness. Clearly, this is not the best result, since Theorem 3.1 shows that the direct result is also true with singular Wishart distributions. Section 5 contains also some comments about the difficulty in extending the proof of Theorem 4.1 to symmetric matrices. Finally, Theorems 3.1 and 5.1 are more generally true for any of the five types of symmetric cones [see Faraut and Koranyi (1994) for a detailed study of the five types]. We have, however, refrained from giving the proof in the framework of the Euclidean Jordan algebras: the chosen notation is probably enough to allow their aficionados to do it by themselves. To keep the paper easier to read, Section 2 proves everything needed here about the GIG and the Bessel functions on symmetric matrices. Readers interested only in the real case must rush to Section 4.

2. The generalized inverse Gaussian distribution. Let us describe first the generalized inverse Gaussian distribution (GIG) on the cone V_+ of symmetric positive definite (r, r) real matrices. We equip the linear space V of real symmetric (r, r) matrices with the Euclidean structure defined by the scalar product $\langle a, b\rangle = \operatorname{trace} ab$. This induces on V the Lebesgue measure dx which gives mass 1 to the unit cube of V. For p real and $s \in V_+$, define $K_p(s) \in (0, +\infty]$ by

$$K_p(s) = \frac{1}{2}\int_{V_+} (\det x)^{p-(1/2)(r+1)} \exp\bigl(-\tfrac{1}{2}\langle s, x + x^{-1}\rangle\bigr)\, dx. \tag{2.1}$$

Similarly, for p real and a and b in V_+, define $K_p(a, b)$ by

$$K_p(a, b) = \frac{1}{2}\int_{V_+} (\det x)^{p-(1/2)(r+1)} \exp\bigl(-\tfrac{1}{2}(\langle a, x\rangle + \langle b, x^{-1}\rangle)\bigr)\, dx. \tag{2.2}$$

The functions K_p belong to the Bessel-like functions which have been introduced by Herz (1956). The only properties of $K_p(s)$ and $K_p(a, b)$ that we shall need are the following.

PROPOSITION 2.1.

(i) *$K_p(s)$ is finite for all $(p, s) \in \mathbf{R} \times V_+$ and $K_p(s) = K_{-p}(s)$.*

(ii) *$K_p(a, b)$ is finite for all $(p, a, b) \in \mathbf{R} \times V_+ \times V_+$ and $K_p(a, b) = K_{-p}(b, a)$.*

(iii) *If $s(a, b) = \sqrt{b^{1/2}ab^{1/2}}$, then*

$$K_p(a, b) = (\det a)^{-p/2}(\det b)^{p/2} K_p(s(a, b)). \tag{2.3}$$

(iv) *For any $(p, a, b) \in \mathbf{R} \times V_+ \times V_+$*

$$K_p(a, b)(\det a)^p = K_p(b, a)(\det b)^p. \tag{2.4}$$

For convenience we give a proof of the proposition which relies on the following lemma.

LEMMA 2.2. *Let c be a real (r, r) matrix and let c^* be the transposed matrix. Denote by g_c the endomorphism of the linear space V of real symmetric (r, r) matrices defined by $x \mapsto cxc^*$. Then the absolute value of the determinant of g_c is $|\det c|^{r+1}$.*

PROOF. Suppose first that c is diagonal, with $c = \operatorname{diag}(c_1, \dots, c_r)$. For $1 \le i \le j \le r$, define $e_{ij} \in V$ as the matrix whose entries (i, j) and (j, i) are 1 and other entries 0. If $x = (x_{ij})$, clearly $g_c(x) = (c_i x_{ij} c_j)$, thus the e_{ij} are eigenvectors of g_c associated to the eigenvalues $c_i c_j$. Thus if c is diagonal,

$$\det g_c = \prod_{1\le i\le j\le r} c_i c_j = (c_1 \cdots c_r)^{r+1} = (\det c)^{r+1}.$$

Suppose now that c is an orthogonal matrix. Then, with the Euclidean structure on V given by the scalar product $\langle a, b\rangle = \operatorname{trace} ab$, since $cc^* = I_r$ and $\operatorname{trace} xy = \operatorname{trace} yx$, it is easily verified that

$$\langle g_c(a), g_c(b)\rangle = \operatorname{trace} cac^*cbc^* = \operatorname{trace} ab = \langle a, b\rangle.$$

This means that g_c is an orthogonal transformation of V and has determinant ± 1. Hence $|\det g_c| = 1$ and the lemma is still true in this case.

In the general case, we use polar decomposition and we write $c = udv$ where u and v are orthogonal (r, r) matrices and d is diagonal with nonnegative coefficients. We use also the fact that $g_{cc'}(x) = g_c(g_{c'}(x))$, which implies that $\det g_{cc'} = \det g_c \det g_{c'}$. Thus with the polar decomposition,

$$|\det g_c| = |\det g_u|\, |\det g_d|\, |\det g_v| = |\det d|^{r+1} = |\det c|^{r+1}. \qquad \square$$

PROOF OF PROPOSITION 2.1. (i) If $p > (r-1)/2$, then $K_p(s)$ is less than the integral

$$\frac{1}{2}\int_{V_+} (\det x)^{p-(1/2)(r+1)} \exp\left(-\tfrac{1}{2}\langle s, x\rangle\right) dx,$$

which is known to be convergent from the theory of Wishart distributions [see, e.g., Muirhead (1982)]. On the other hand if $p < -(r-1)/2$ then $K_p(s)$ is less than

$$\begin{aligned}\frac{1}{2}\int_{v_+} &(\det x)^{p-(1/2)(r+1)} \exp\left(-\tfrac{1}{2}\langle s, x^{-1}\rangle\right) dx\\ &= \frac{1}{2}\int_{V_+} (\det x)^{-p-(1/2)(r+1)} \exp\left(-\tfrac{1}{2}\langle s, x\rangle\right) dx,\end{aligned}$$

since the differential of $x \mapsto x^{-1}$ is the linear endomorphism of V defined by $h \mapsto -x^{-1}hx^{-1}$ whose determinant is $|\det x|^{-r-1}$ from Lemma 2.2.

Since for fixed s in V_+ the function $p \mapsto K_p(s)$ is a Laplace transform of a positive measure, its domain of existence is an interval. Since this interval contains the two half lines $((r-1)/2, +\infty)$ and $(-\infty, -(r-1)/2)$, the interval is $\mathbf{R}$.

The formula $K_p(s) = K_{-p}(s)$ is proved easily by the change of variable $x = y^{-1}$ (which has been done in the last formula) in K_p and applying again Lemma 2.2.

(ii) It follows, by applying step by step the argument used in the proof of (1).

(iii) Denote $s = s(a, b)$ for a while. We have

$$\begin{aligned}\langle a, x\rangle + \langle b, x^{-1}\rangle &= \langle b^{-1/2}s^2b^{-1/2}, x\rangle + \langle b, x^{-1}\rangle\\ &= \langle s, s^{1/2}b^{-1/2}xb^{-1/2}s^{1/2}\rangle + \langle s, (s^{1/2}b^{-1/2}xb^{-1/2}s^{1/2})^{-1}\rangle.\end{aligned}$$

We now make the change of variable $y = s^{1/2}b^{-1/2}xb^{-1/2}s^{1/2}$ in the integral $K_p(a, b)$. The determinant of $s^{1/2}b^{-1/2}$ is $(\det s^2)^{1/4}(\det b)^{-1/2} = (\det a)^{1/4}(\det b)^{-1/4}$, and the application of Lemma 2.2 leads to the desired formula (2.3).

(iv) To prove (2.4) we start from (2.3) and subsequently apply (i), then again (2.3) and finally (ii):

$$\begin{aligned}K_p(a, b) &= (\det a)^{-p/2}(\det b)^{p/2}K_{-p}(s(a, b)) = (\det a)^{-p}(\det b)^{p}K_{-p}(a, b)\\ &= (\det a)^{-p}(\det b)^{p}K_{-p}(a, b) = (\det a)^{-p}(\det b)^{p}K_p(b, a). \qquad \square\end{aligned}$$

We are now in a position to define the GIG distribution. For p in $\mathbf{R}$, a and b in V_+ this is the probability measure on V_+ defined by

$$\begin{aligned}\mu_{p,a,b}(dx) &= \frac{1}{K_p(a,b)}(\det x)^{p-(1/2)(r+1)} \\ &\quad\times \exp\left(-\frac{1}{2}(\langle a, x\rangle + \langle b, x^{-1}\rangle)\right)\mathbf{1}_{V_+}(x)\,dx.\end{aligned} \tag{2.5}$$

We point out right now the following remarkable integral.

PROPOSITION 2.3. *For θ and σ in V such that $a-2\theta$ and $b-2\sigma$ are in V_+, we have*

$$\int_{V_+} \exp(\langle\theta, x\rangle + \langle\sigma, x^{-1}\rangle)\mu_{p,a,b}(dx) = \frac{K_p(a-2\theta, b-2\sigma)}{K_p(a,b)}. \tag{2.6}$$

Furthermore, X has the distribution $\mu_{p,a,b}(dx)$ if and only if X^{-1} has distribution the $\mu_{-p,b,a}(dx)$.

The proof follows immediately from Proposition 2.1(ii) with a change of the variable $x \mapsto x^{-1}$. □

3. An independence property.

For p in the following set:

$$\Lambda = \left\{\frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \ldots, \frac{r-1}{2}\right\} \cup \left(\frac{r-1}{2} + \infty\right)$$

and a in V_+, define the Wishart distribution $\gamma_{p,2a^{-1}}$ on the closed cone $\overline{V}_+$ of positive symmetric matrices by its Laplace transform,

$$\int_{\overline{V}_+} e^{\langle\theta, y\rangle}\gamma_{p,2a^{-1}}(dy) = \frac{(\det a)^p}{(\det(a-2\theta))^p}. \tag{3.1}$$

If p is in the singular part of Λ, namely $\{\frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \ldots, (r-1)/2\}$, then $\gamma_{p,2a^{-1}}$ is concentrated on the boundary of V_+ which consists of the singular positive symmetric matrices. If $p > (r-1)/2$ then

$$\begin{aligned}\gamma_{p,2a^{-1}}(dy) &= \frac{1}{\Gamma_r(p)}\left(\det\frac{a}{2}\right)^p (\det y)^{p-(1/2)(r+1)} \\ &\quad\times \exp\left(-\frac{1}{2}\langle a, y\rangle\right)\mathbf{1}_{V_+}(y)\,dy,\end{aligned} \tag{3.2}$$

where $\Gamma_r(p)$ is the multivariate gamma function; see Muirhead (1982), page 61.

THEOREM 3.1. *Let p be in Λ and a and b in V_+. Let X and Y be independent random variables in V_+ and $\overline{V}_+$ with respective distributions $\mu_{-p,a,b}$ and $\gamma_{p,2a^{-1}}$. Then the random variables $U = (X+Y)^{-1}$ and $V = X^{-1} - (X+Y)^{-1}$ are independent with respective distributions $\mu_{-p,b,a}$ and $\gamma_{p,2b^{-1}}$.*

We give two proofs. The first one is only a partial one since it is restricted to the absolutely continuous case $p > (r-1)/2$ and consequently uses the Jacobian. We give it because it is rather representative of the method that we shall use in the proof of the converse in Theorem 5.1. The second proof gives the result in full generality and is based on applying Laplace transforms.

PROOF OF THEOREM 3.1 (partial, by the Jacobian). Let $f\colon V_+\times V_+ \to V_+\times V_+$ be defined by

$$f(x, y) = (u, v) = ((x+y)^{-1}, x^{-1} - (x+y)^{-1}).$$

Note that f is involutive, that is, $f(f(x, y)) = (x, y)$. Hence $(x, y) = ((u+v)^{-1}, u^{-1} - (u+v)^{-1})$, and it becomes easy to compute $dxdy$ with respect to $dudv$. For this we use once again the fact that the differential of $x \mapsto x^{-1}$ is $h \mapsto -x^{-1}hx^{-1}$, or with more convenient notations $dx \mapsto -x^{-1}(dx)x^{-1}$. Thus we get

$$dx = -(u+v)^{-1}(du)(u+v)^{-1} - (u+v)^{-1}(dv)(u+v)^{-1},$$
$$dy = -u^{-1}(du)u^{-1} + (u+v)^{-1}(du+dv)(u+v)^{-1}.$$

For simplicity let us introduce the notation

$$\mathbb{P}(x)\colon V \to V, \qquad h \mapsto \mathbb{P}(x)(h) = xhx.$$

Thus the differential of the map $(u, v) \mapsto (x, y)$ from $V_+ \times V_+$ to itself is a linear endomorphism of $V \times V$ which is conveniently written by blocks,

$$\begin{pmatrix} du \\ dv \end{pmatrix} \mapsto \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{bmatrix} -\mathbb{P}(x) & -\mathbb{P}(x) \\ \mathbb{P}(x) - \mathbb{P}(x+y) & \mathbb{P}(x) \end{bmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \tag{3.3}$$

The Jacobian that we want to compute is the determinant of this linear map. Adding the second row to the first shows that the absolute value of the Jacobian is $\det(\mathbb{P}(x+y))\det(\mathbb{P}(x))$. Recall from Lemma 2.2 that $\det(\mathbb{P}(x)) = (\det x)^{r+1}$. Consequently, the absolute value of the Jacobian equals $(\det x \times \det(x+y))^{r+1} = (\det u \det(u+v))^{-(r+1)}$.

Now it suffices to check the factorization property for the densities of X, Y, U, V denoted respectively by f_X, f_Y, f_U, f_V,

$$f_U(u)f_V(v) = (\det u \det(u+v))^{-(r+1)} f_X((u+v)^{-1}) \times f_Y(u^{-1} - (u+v)^{-1}) \tag{3.4}$$

for any $(u, v) \in V_+ \times V_+$. This is quite a standard calculation in the univariate case. Since we use matrix variates some details will be presented.

Since the joint density of (X, Y) on $V_+ \times V_+$ is proportional to

$$(\det x)^{-p-(1/2)(r+1)}(\det y)^{p-(1/2)(r+1)} \exp\Bigl(-\tfrac{1}{2}\langle a, x+y\rangle - \tfrac{1}{2}\langle b, x^{-1}\rangle\Bigr),$$

we replace x and y by their expression in u and v. The argument of the exponential becomes

$$-\tfrac{1}{2}\langle a, u^{-1}\rangle - \tfrac{1}{2}\langle b, u+v\rangle = -\tfrac{1}{2}(\langle a, u^{-1}\rangle + \langle b, u\rangle) - \tfrac{1}{2}\langle b, v\rangle.$$

To compute the remainder of the expression of the density, suitably multiplied by the Jacobian, observe first that

$$\det(u^{-1}-(u+v)^{-1}) = \det(u^{-1}u(u^{-1}-(u+v)^{-1})(u+v)(u+v)^{-1})$$
$$= \det(u^{-1}v(u+v)^{-1}) = \frac{\det v}{\det u \, \det(u+v)}.$$

Finally, after simplification, the remaining part of the expression becomes

$$(\det u)^{-p-(1/2)(r+1)}(\det v)^{p-(1/2)(r+1)}$$

and the product of this and of the exponential yields the required density. □

REMARK. Observe that the above factorization property is not valid if $\mathscr{L}(X) = \mu_{-p_1,a,b}$ and $\mathscr{L}(Y) = \gamma_{p_2,2a^{-1}}$ with $p_1 \neq p_2$, consequently U and V are not independent in this case.

Before we proceed to the general proof of Theorem 3.1, we give an auxiliary result, which will be used later.

LEMMA 3.2. *Let (Z,T) be a random variable on $V_+ \times \overline{V}_+$. For (θ,σ) in $V_+ \times V_+$, define*

$$M_{Z,T}(\theta,\sigma) = \mathbb{E}(\exp(\langle\theta, Z+T\rangle + \langle\sigma, Z^{-1}\rangle)).$$

Then $M_{Z,T}$ determines the distribution of (Z,T).

PROOF. If (Z_1,T_1) is valued in $V_+ \times \overline{V}_+$, suppose that on $V_+ \times V_+$ we have

$$M_{Z,T} = M_{Z_1,T_1}.$$

Then, since the Laplace transforms coincide, we have

$$\mathscr{L}(Z+T, Z^{-1}) = \mathscr{L}(Z_1+T_1, Z_1^{-1}).$$

Taking the images of these random variables by $g(x,y) = (y^{-1}, x-y^{-1})$ mapping $V_+ \times V_+$ into $V_+ \times V$ we get $\mathscr{L}(Z,T) = \mathscr{L}(Z_1,T_1)$.

PROOF OF THEOREM 3.1. Observe that from the definition of U and V we have

$$M_{U,V}(\theta,\sigma) = \mathbb{E}\Big(\exp(\langle\sigma, X\rangle + \langle\theta, X^{-1}\rangle)\Big)\,\mathbb{E}(\exp(\langle\sigma, Y\rangle)).$$

Consequently, by (2.6), (3.1) and the assumptions of the theorem, one gets

$$M_{U,V}(\theta,\sigma) = \frac{K_{-p}(a-2\sigma, b-2\theta)(\det a)^p}{K_{-p}(a,b)(\det(a-2\sigma))^p}.$$

Now, to conclude the proof, using Lemma 3.2, it suffices to introduce a random variable (U_1,V_1) with distribution $\mu_{-p,b,a} \otimes \gamma_{p,2b^{-1}}$ and to show that

$M_{U,V} = M_{U_1,V_1}$. Using again (2.6) and (3.1) one finds easily that

$$M_{U_1,V_1}(\theta,\sigma) = M_{U_1,U_1^{-1}}(\theta,\sigma)\mathbb{E}(\exp(\langle\theta, V_1\rangle)) = \frac{K_{-p}(b-2\theta, a-2\sigma)(\det b)^p}{K_{-p}(b,a)(\det(b-2\theta))^p}.$$

The final result follows immediately from (2.4). □

4. Characterization on the real line. We are unable to provide the converse of Theorem 3.1 in full generality in the matrix case (see Section 5, where also some of the reasons are indicated). However, in the important univariate case the complete converse is available. This is the subject of the present section.

THEOREM 4.1. *Let X and Y be real positive independent random variables. Assume that X or Y is non-Dirac and that $U = (X+Y)^{-1}$ and $V = X^{-1}-(X+Y)^{-1}$ are independent. Then there exist $p, a, b > 0$ such that $\mathscr{L}(X) = \mu_{-p,a,b}$ and $\mathscr{L}(Y) = \gamma_{p,2a^{-1}}$.*

PROOF. Observe first that if any one of X and Y is non-Dirac then all four random variables X, Y, U, V are non-Dirac also.

Let us fix $\alpha \geq 0$, $\theta < 0$ and $\sigma < 0$. Denote for simplicity,

$$A = \exp(\sigma X + \theta X^{-1}), \qquad B = \exp(\sigma U^{-1} + \theta U).$$

Using the independence, the fact that $Y/X = V/U$ and noting that all the expectations exist, we have

$$\mathbb{E}(Y^\alpha e^{\sigma Y})\mathbb{E}(X^{-\alpha}A) = \mathbb{E}(V^\alpha e^{\theta V})\mathbb{E}(U^{-\alpha}B). \tag{4.1}$$

Take the logarithm of both sides of the above equality and apply $\partial^2/\partial\theta\partial\sigma$ to them. We get

$$\frac{\mathbb{E}(X^{-\alpha-1}A)\mathbb{E}(X^{-\alpha+1}A)}{(\mathbb{E}(X^{-\alpha}A))^2} = \frac{\mathbb{E}(U^{-\alpha-1}B)\mathbb{E}(U^{-\alpha+1}B)}{(\mathbb{E}(U^{-\alpha}B))^2}. \tag{4.2}$$

We now put $\alpha = 1$ in (4.2). Applying (4.1) successively to $\alpha = 0$, $\alpha = 1$ and $\alpha = 2$ and carrying it in (4.2) we get for all $\theta < 0$ and $\sigma < 0$,

$$\frac{\mathbb{E}(Y^2 e^{\sigma Y})\mathbb{E}(e^{\sigma Y})}{(\mathbb{E}(Ye^{\sigma Y}))^2} = \frac{\mathbb{E}(V^2 e^{\theta V})\mathbb{E}(e^{\theta V})}{(\mathbb{E}(Ve^{\theta V}))^2}. \tag{4.3}$$

From the principle of separation of variables, the two sides of (4.3) are constant. Since Y is not Dirac, the left-hand side of (4.2) is a number $1 + p > 1$. One gets easily from this that the two random variables Y and V are gamma distributed, with the same shape parameter p: if $L_Y(\sigma) = \mathbb{E}(e^{\sigma Y})$, just solve the second order equation $L''_Y L_Y = (p+1)L'^2_Y$ and do the same for V. We write $\mathscr{L}(Y) = \gamma_{p,2a^{-1}}$ and $\mathscr{L}(V) = \gamma_{p,2b^{-1}}$ for some a and $b > 0$.

To show that X is GIG, we now rely on the characterization of these distributions given in Letac and Seshadri (1983). Introduce a random variable Y' independent of X and Y with distribution $\gamma_{p,\,2b^{-1}}$. Then we have

$$X = \frac{1}{V+U} \stackrel{d}{=} \frac{1}{Y'+U} = \frac{1}{Y' + \frac{1}{Y+X}}.$$

The result of the quoted paper implies that X has distribution $\mu_{-p,\,a,\,b}$.

(This part can be also proved directly by repeating the argument used in the first part of the proof of Theorem 5.1; see Section 5.) □

5. Characterization on symmetric matrices. In the general case of random symmetric positive definite matrices a partial converse to Theorem 3.1 is obtained, with the proof depending on the assumed smoothness of densities.

THEOREM 5.1. *Let X and Y be two independent random variables valued in the cone V_+ of symmetric positive definite real (r, r) matrices, with strictly positive densities of class C^2. Let us assume that $U = (X+Y)^{-1}$ and $V = X^{-1} - (X+Y)^{-1}$ are independent. Then there exist $p > (r-1)/2$, and a, b in V_+ such that $\mathscr{L}(X) = \mu_{-p,\,a,\,b}$ and $\mathscr{L}(Y) = \gamma_{p,\,2a^{-1}}$.*

Before we give the proof of this theorem we present two results which will be used in the proof. The first is in the linear algebra framework.

PROPOSITION 5.2. *The space V being equipped with its Euclidean structure, denote by $L_s(V)$ the linear space of symmetric endomorphisms of V. If B in $L_s(V)$ commutes with $\mathbb{P}(x)$ for all x in V_+, then B is a multiple of the identity* id_V.

PROOF. Recall that $\mathbb{P}(x) \in L_s(V)$ is defined by $\mathbb{P}(x)(v) = xvx$ for any $v \in V$. We denote by e the identity in V. By the assumptions it follows that

$$B(xhx) = xB(h)x \tag{5.1}$$

for any $x \in V_+$ and any $h \in V$. Putting then $h = e$ in (5.1) we get $B(x^2) = xcx$, for $c = B(e) \in V$. Consequently, the polarization procedure leads to

$$xcy + ycx = B(xy + yx)$$

for any $x, y \in V_+$. Upon inserting in this formula $y = e$, for any $x \in V_+$, we arrive at

$$B(x) = (cx + xc)/2. \tag{5.2}$$

Carrying it back to (5.1) and inserting then $h = c$ one gets

$$(xc - cx)cx - xc(xc - cx) = 0$$

for any $x \in V_+$. Consequently, since $\mathrm{trace}(cx^2c) = \mathrm{trace}(xc^2x)$, we get

$$\mathrm{trace}((xc - cx)(xc - cx)^*) = 0.$$

Hence for any $x \in V_+$

$$xc = cx, \tag{5.3}$$

which, via (5.2), leads to $B(x) = cx$ for any $x \in V_+$.

Now insert in (5.3) $x = x_{ij} = \mathrm{id}_V + \epsilon e_{ij}$ for any $i, j \in \{1, 2, \dots, r\}$, where e_{ij} are defined in the proof of Lemma 2.2, and $\epsilon > 0$ is small enough to have $x_{ij} \in V_+$. Then it follows immediately that $c_{ik} = c_{jk} = 0$ for any $k \neq i, j$ and $c_{ii} = c_{jj}$. Since the above observation is valid for any i and j it follows that c is a multiple of e, which ends the proof. □

The next result presents the solution of a second-order differential equation, which while being of independent interest, will be also a tool in proving Theorem 5.1.

PROPOSITION 5.3. *Let g: $V_+ \to \mathbf{R}$ be a C^2 function such that there exists a linear endomorphism B of V such that for all y in V_+ one has*

$$g''(y)\mathbb{P}(y) = B.$$

Then there exists λ in $\mathbb{R}$ such that $B = \lambda \mathrm{id}_V$, and there exist a in V and C in $\mathbb{R}$ such that

$$g(y) = C - \tfrac{1}{2}\langle a, y\rangle - \lambda \log \det y.$$

PROOF. Recall first that the differential of $y \mapsto \log \det y$ is y^{-1}, and that the differential of $y \mapsto y^{-1}$ is $-\mathbb{P}(y^{-1})$. Since $g''(y)$ is in the space $L_s(V)$, putting in the differential equation $y = e$ shows that $B = g''(e)$ is also symmteric. Thus since $\mathbb{P}(y^{-1}) = \mathbb{P}(y)^{-1}$, we have

$$\begin{aligned} B\mathbb{P}(y) &= g''(y^{-1})\mathbb{P}(y^{-1})\mathbb{P}(y) = g''(y^{-1}) = (g''(y^{-1}))^* \\ &= (g''(y^{-1})\mathbb{P}(y^{-1})\mathbb{P}(y))^* = (B\mathbb{P}(y))^* = \mathbb{P}(y)B^* = \mathbb{P}(y)B. \end{aligned}$$

Now Proposition 5.2 implies that $B = \lambda \mathrm{id}_V$. Thus $g''(y) = \lambda \mathbb{P}(y^{-1})$, the differential of $y \mapsto g'(y) + \lambda y^{-1}$ is zero on the connected open set V_+ and this implies the existence of a in V such that $g'(y) + \lambda y^{-1} = -a$. In turn, the differential of $y \mapsto g(y) - \lambda \log \det y - \frac{1}{2}\langle a, y\rangle$ on V_+ is zero and the function is a constant C. This ends the proof. □

Now we are ready to prove the main result of this section.

PROOF OF THEOREM 5.1. Let us assume first that $\mathscr{L}(Y) = \gamma_{p_1, 2a^{-1}}$ and $\mathscr{L}(V) = \gamma_{p_2, 2b^{-1}}$. Change now (u, v) into (v, u) in (3.4). The resulting equation compared with (3.4) leads at once to

$$\begin{aligned} &f_U(u)f_V(v)(\det u)^{r+1}f_Y(v^{-1} - (u+v)^{-1}) \\ &\quad = f_U(v)f_V(u)(\det v)^{r+1}f_Y(u^{-1} - (u+v)^{-1}) \end{aligned}$$

for any $(u, v) \in V_+ \times V_+$. Now inserting the proper versions of (3.2) for f_Y and f_V in the above identity we get (after fixing the $v = v_0$) for any $u \in V_+$,

$$f_U(u)(\det u)^{2p_2 - p_1 + (r+1)/2} e \to \exp(\langle b,\ u\rangle + \langle a,\ u^{-1}\rangle)/2 = \text{const.}$$

By comparing with (2.5), we see that $\mathscr{L}(U) = \mu_{p_1 - 2p_2, b, a}$. As a dual $\mathscr{L}(X) = \mu_{p_2 - 2p_1, a, b}$. Now, by the remark following the proof of Theorem 3.1 we get $2p_1 - p_2 = p_1$, and finally $p_1 = p_2 = p$.

Hence to conclude the proof it suffices to show that Y is Wishart, since by duality between Y and V, it will follow immediately that V is also Wishart.

Upon taking logs of both sides of (3.4) we arrive at the identity

$$F(x(u, v)) + g(y(u, v)) = h_1(u) + h_2(v),$$

for any $(u, v) \in V_+ \times V_+$, where all the functions in the expression are of the class C^2, $g = \log(f_Y)$ and the exact forms of the remaining functions F, h_1 and h_2 are not important. Consequently, we get

$$\frac{\partial^2}{\partial v \partial u}(F(x, (u, v)) + g(y(u, v))) = 0, \tag{5.4}$$

where the partial differentiation symbol means taking derivatives with respect to the corresponding matrix variable.

We are now in a position to compute the left-hand side of (5.4). To this end we will rely on (3.3). For convenience, we evaluate the differential with respect to u in h and to v in $k_1 = x^{-1}kx^{-1}$. Let us concentrate first on g. Then

$$\frac{\partial}{\partial v}(g(y(u, v)))(k_1) = g'(y)(\mathbb{P}(x)(k_1)) = g'(y)(k)$$

for any $k \in V$. Further,

$$\frac{\partial}{\partial u}(g'(y)(k))(h) = g''(y)((\mathbb{P}(x) - \mathbb{P}(x + y))(k), h)$$

for any $(k, h) \in V \times V$. On the other hand the second derivative of $F(x(u, v))$, evaluated at the same points, is a bilinear form, say $B(x)(k, h)$ in $(k, h) \in V \times V$, which depends only on x [see again at (3.3)]. Consequently for any $(k, h) \in V \times V$ we get

$$B(x)(k, h) = g''(y)((\mathbb{P}(x + y) - \mathbb{P}(x))(k), h)$$

for any $(x, y) \in V_+ \times V_+$. Observe now that the right-hand side of the above identity has a limit as $x \to 0$, hence the left-hand side also converges, and we have with $B = B(0)$ that for any $(k, h) \in V \times V$,

$$B(k, h) = g''(y)(\mathbb{P}(y)(k), h).$$

Now recall that since V is Euclidean, the space of (not necessarily symmetric) bilinear forms on V is isomorphic to the space $L(V)$ of endomorphisms of V by $a \mapsto C_a$ where $C_a(h,k) = \langle h, a(k)\rangle$. For this reason, we use the traditional abuse of notation ($a = C_a$), writing

$$\langle h, B(k)\rangle = \langle h, g''(y)(\mathbb{P}(y)(k))\rangle$$

for any $(k,h) \in V \times V$. Consequently, treating now B and $g''(y)\mathbb{P}(y)$ as elements of $L(V)$, we obtain the identity

$$g''(y)\mathbb{P}(y) = B$$

for any $y \in V_+$.

We now apply Proposition 5.3 to claim that there exist a in V and scalars C_2 and λ such that

$$g(y) = -\lambda \log \det y - \tfrac{1}{2}\langle a, y\rangle + C_2,$$

which implies that Y is a Wishart random matrix. □

COMMENTS. Extending the proof of Theorem 4.1 to symmetric matrices does not seem to be an easy task. It can be compared to the extension to Wishart distributions of the characterization by Lukacs (1955) of the gamma distribution. An obscure proof appears in Olkin and Rubin (1962) and neater ones in Casalis and Letac (1996) and in Letac and Massam (1998). While the analog of (4.1) is natural [replace Y^α by $(\det Y)^\alpha$], obtaining the analog of (4.2) by differentiating is difficult. There is some hope in the fact that a key step in the proof of Theorem 4.1 is the differential equation $L''_Y L_Y = (p+1)L'^2_Y$. Introducing the cumulant generating function $k_Y = \log L_Y$, it is rewritten as $k''_Y = pk'^2_Y$, which says nothing else than that the variance function of the natural exponential family generated by the distribution of Y is $V_F(m) = m^2/p$, thus characterizing the gamma distributions with shape parameter p. This provides a hope for an extension to Wishart, along the lines of the extension of the Lukacs result mentioned above. Finally, to have the characterization even in the singular case, one cannot rely on the analog of the Letac–Seshadri characterization of the GIG distribution given by Bernadac (1995), since she proves it only for the nonsingular case $p > (r-1)/2$, with the help of difficult algebraic identities related to continued fractions in symmetric matrices and Jordan Algebras.

REFERENCES

BERNADAC, E. (1995). Random continued fractions and inverse Gaussian distribution on a symmetric cone. *J. Theoret. Probab.* **8** 221–260.

BUTLER, R. W. (1998). Generalized inverse Gaussian distributions and their Wishart connexions. *Scand. J. Statist.* **25** 69–75.

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

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

HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. of Math.* **61** 474–523.

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

LETAC, G. and SESHADRI, V. (1983). A characterization of the generalized inverse Gaussian distribution by continued fractions. *Z. Wahrsch. Verw. Gebiete* **62** 485–489.

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

MASSAM, H. (1994). An exact decomposition theorem and a unified view of some related distributions for a class of exponential transformation models on symmetric cones. *Ann. Statist.* **22** 369–394.

MATSUMOTO, H. and YOR, M. (1998). An analogue of Pitman's $2M - X$ theorem for exponential Wiener functional, Part II: the role of the generalized inverse Gaussian laws. *Nagoya Math. J.* To appear.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

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

LABORATOIRE DE STATISTIQUE
ET PROBABILITÉS
UNIVERSITÉ PAUL SABATIER
31062, TOULOUSE
FRANCE
E-MAIL: letac@cict.fr

INSTYTUT MATEMATYKI
POLITECHNIKA WARSZAWSKA
PLAC POLITECHNIKI 1
00-661, WARSZAWA
POLAND
E-MAIL: wesolo@alpha.im.pw.edu.pl