COLLOQUIUM MATHEMATICUM

VOL. 158 2019 NO. 1

CHARACTERIZATION OF THE RIESZ EXPONENTIAL FAMILY ON HOMOGENEOUS CONES

BY

HIDEYUKI ISHI (Nagoya) and BARTOSZ KOŁODZIEJEK (Warszawa)

Abstract. We give a characterization theorem for the Riesz measure and a Wishart exponential family on homogeneous cones through the invariance property of a natural exponential family under the action of the triangular group.

1. Introduction. Following Casalis [3], we consider natural exponential families (NEFs) of probability distributions on a finite-dimensional linear space $\mathbb{E}$ which are invariant under a subgroup G of the general linear group of $\mathbb{E}$. Under a weak assumption on G (see Theorem 2.1), the generating measure of such a family is of the form

$$e^{-\langle\theta_0,x\rangle}\mu_0(dx)$$

for some θ_0 in the dual space $\mathbb{E}^*$, where μ_0 is a G-invariant measure.

We apply Theorem 2.1 to the problem of characterizing the Riesz measure on a homogeneous cone through the invariance property of NEF under the action of the triangular group (Theorem 3.8). Since the NEF generated by the Riesz measure consists of Wishart distributions, Theorem 3.8 provides also a characterization of Wishart distributions on homogeneous cones [1, 11, 27]. We mention that this problem was stated by Letac [25, Section 4], who pointed out that the natural framework for such characterizations is indeed a homogeneous cone.

There are essentially two types of characterizations of NEFs. The first one is connected with the central object of a NEF, that is, the variance function. The aim is then to describe a generating measure of a NEF with given variance function. Much has been done in this direction, but still much more is unknown. A generating measure is known for $\mathbb{E}=\mathbb{R}$ when the variance function is a quadratic polynomial [29] or a cubic polynomial [15, 28] or has some more sophisticated form, like $P\Delta + Q\sqrt{\Delta}$, where P, Δ

2010 *Mathematics Subject Classification*: Primary 60B11; Secondary 62E10.
Key words and phrases: natural exponential families, Riesz measure, Wishart distribution, general affine group.
Received 6 March 2018; revised 7 August 2018.
Published online 14 June 2019.

DOI: 10.4064/cm7548-9-2018

and Q are quadratic polynomials [26]. For $\mathbb{E} = \mathbb{R}^d$ we have to mention [4] with its deep connection between homogeneous quadratic variance functions and Euclidean Jordan algebras, and simple quadratic [5] and simple cubic variance functions [16], to name but a few.

The second type of characterizations of NEFs is through invariance under some group action. We mention some results on NEFs invariant under a given subgroup G of the general affine group: a one-parameter group [7], the group of rotations [33], the Möbius group [23], the identity component $\mathrm{Aut}_0(\mathrm{Sym}_+(N,\mathbb{R}))$ of the linear automorphism group of the cone $\mathrm{Sym}_+(N,\mathbb{R}) \subset \mathrm{Sym}(N,\mathbb{R})$ [25], the triangular group of a simple Euclidean Jordan algebra [13] and its modification [14]. Here $\mathrm{Sym}(N,\mathbb{R})$ stands for the symmetric $N \times N$ matrices with real entries and $\mathrm{Sym}_+(N,\mathbb{R})$ is the cone of positive definite real $N \times N$ matrices. In the last three papers cited, the characterizations were carried out by showing that the variance function (which uniquely determines a NEF) coincides with the variance function of some Riesz measure (or its image by the involution $x \mapsto -x$) on $\mathrm{Sym}_+(N,\mathbb{R})$ [25] and on symmetric cones [13, 14].

In the present paper we solve the problem raised by Letac [25] on homogeneous cones. We use a matrix realization of homogeneous cones, which proves here very useful and is more accessible to the reader who is not familiar with homogeneous cones, clans and T-algebras. We emphasize that homogeneous cones are of great importance in statistics despite their theoretical character. In particular, homogeneous cones appear very naturally in the context of graphical Gaussian models [27]. The aforementioned matrix realization of homogeneous cones allows one to describe many of the colored graphical Gaussian models [17]. The formula for the variance function of $F(\mathcal{R}_{\underline{s}})$ on homogeneous cones is the topic of our joint paper with Piotr Graczyk [12]. Finally, we point out that the present work is closely related to [30, 31], where NEFs and information theory are treated on homogeneous Hessian structures.

The paper is organized as follows. The concept of natural exponential families is introduced in the next section. In Section 3.2 we give an application of Theorem 2.1 to a characterization of the Riesz measure on a homogeneous cone (Theorem 3.8). A crucial role in the proof of the characterization is played by a matrix realization of any homogeneous cone [19], which is explained in Section 3.1. Section 3 ends with some comments.

2. Natural exponential families. In this section we will introduce all the facts about NEFs that will be needed later on. The standard reference on exponential families is [2].

Let $\mathbb{E}$ be a finite-dimensional real linear space and $\mathbb{E}^*$ its dual space. The coupling of $\theta \in E^*$ and $x \in E$ is denoted by $\langle \theta, x \rangle$. Let μ be a positive Radon measure on $\mathbb{E}$. We define its Laplace transform $L_\mu\colon \mathbb{E}^* \to (0,\infty]$ by

$$L_\mu(\theta) = \int_{\mathbb{E}} e^{\langle\theta,x\rangle}\,\mu(dx).$$

We denote by $\Theta(\mu)$ the interior of $\{\theta \in \mathbb{E}^*\colon L_\mu(\theta) < \infty\}$. Hölder's inequality implies that $\Theta(\mu)$ is convex, and the cumulant function

$$k_\mu(\theta) = \log L_\mu(\theta)$$

is convex on $\Theta(\mu)$; moreover, k_μ is strictly convex if and only if μ is not concentrated on any affine hyperplane of $\mathbb{E}$. Let $\mathcal{M}(\mathbb{E})$ be the set of positive Radon measures on $\mathbb{E}$ such that $\Theta(\mu)$ is not empty and μ is not concentrated on any affine hyperplane of $\mathbb{E}$.

For $\mu \in \mathcal{M}(\mathbb{E})$ we define the *natural exponential family (NEF) generated by* μ (denoted by $F(\mu)$) as the set of probability measures of the form

$$P(\theta,\mu)(dx) = e^{\langle\theta,x\rangle - k_\mu(\theta)}\mu(dx), \qquad \theta \in \Theta(\mu).$$

Note that $F(\mu) = F(\mu')$ if and only if $\mu'(dx) = e^{\langle a,x\rangle+b}\mu(dx)$ for some $a \in \mathbb{E}^*$ and $b \in \mathbb{R}$.

We will now describe the action of elements of the general linear group $\mathrm{GL}(\mathbb{E})$ on a NEF. The identity element of $\mathrm{GL}(\mathbb{E})$ will be denoted by Id. Let $F = F(\mu)$ be a NEF on $\mathbb{E}$. Let $g_*\mu$ denote the image measure of μ by g and let $g.F(\mu)$ stand for the family of image measures $g_*P(\theta,\mu)$ $(\theta \in \Theta(\mu))$. Then, for any $g \in \mathrm{GL}(\mathbb{E})$, we have $g.F(\mu) = F(g_*\mu)$.

We say that a measure μ_0 is *invariant under a subgroup* G of $\mathrm{GL}(\mathbb{E})$ if for all $g \in G$ there exists a constant $c_g > 0$ such that $\mu_0(gA) = c_g\mu_0(A)$ for any measurable set $A \subset \mathbb{E}$. This condition is equivalent to

$$L_{\mu_0}(g^*\theta) = c_g^{-1}L_{\mu_0}(\theta), \qquad \theta \in \Theta(\mu_0). \tag{2.1}$$

Note that the correspondence $G \ni g \mapsto c_g \in \mathbb{R}$ is a character of the group G.

Let $\mu \in \mathcal{M}(\mathbb{E})$. Observe that the condition $g.F(\mu) = F(\mu)$ implies that for any $g \in G$ there exist $a(g) \in \Theta(\mu) \subset \mathbb{E}^*$ and $b(g) \in \mathbb{R}$ such that

$$g_*\mu(dx) = e^{\langle a(g),x\rangle + b(g)}\mu(dx). \tag{2.2}$$

Casalis [3, Theorem 2.2] showed that the functions a and b satisfy the following system of equations: for any $(g,g') \in G^2$,

$$a(gg') = (g^*)^{-1}a(g') + a(g), \qquad b(gg') = b(g) + b(g').$$

Assume that G contains $c\,\mathrm{Id}$ for some $c \neq 1$. Then, for any $g \in G$,

$$a(c\,\mathrm{Id}\,g) = \frac{1}{c}a(g) + a(c\,\mathrm{Id}), \qquad a(g\,c\,\mathrm{Id}) = (g^*)^{-1}a(c\,\mathrm{Id}) + a(g).$$

Equating the right hand sides of the above formulas, we arrive at

$$a(g) = \theta_0 - (g^*)^{-1}\theta_0, \quad g \in G, \quad \text{with} \quad \theta_0 = \frac{c}{1-c}a(c\,\mathrm{Id}).$$

Define $\mu_0(dx) = e^{\langle\theta_0,x\rangle}\mu(dx)$. Then (2.2) implies that μ_0 is G-invariant. Thus, we obtain the following

THEOREM 2.1. *Let G be a subgroup of $\mathrm{GL}(\mathbb{E})$ and let $F = F(\mu)$ be a G-invariant NEF on $\mathbb{E}$, that is, $g.F = F$ for any $g \in G$. If G contains $c\,\mathrm{Id}$ for some $c \neq 1$, then there exist $\theta_0 \in \mathbb{E}^*$ and a G-invariant measure μ_0 such that*

$$\mu(dx) = e^{-\langle\theta_0,x\rangle}\mu_0(dx).$$

In that case also $F = F(\mu_0)$.

3. Characterization of the Riesz measure on homogeneous cones

3.1. Matrix realization of homogeneous cones. Let V be a real linear space and Ω a regular open convex set in V containing no line. The cone Ω is said to be *homogeneous* if the linear automorphism group $G(\Omega) = \{g \in \mathrm{GL}(V) \colon g\Omega = \Omega\}$ acts transitively on Ω, that is, for any x and y in Ω there exists $g \in G(\Omega)$ such that $y = gx$.

We will now give a useful representation of homogeneous cones following [21, Section 3]. For a symmetric matrix $x \in \mathrm{Sym}(N,\mathbb{R})$, we denote by $\underset{\vee}{x}$ the lower triangular matrix of size N defined by

$$(\underset{\vee}{x})_{ij} = \begin{cases} x_{ij} & \text{if } i > j, \\ x_{ii}/2 & \text{if } i = j, \\ 0 & \text{if } i < j. \end{cases}$$

Then $x = \underset{\vee}{x} + \hat{x}$, where $\hat{x} = \underset{\vee}{x}^\top$ is the transpose of $\underset{\vee}{x}$. For $x, y \in \mathrm{Sym}(N,\mathbb{R})$, we define

$$x \bigtriangleup y := \underset{\vee}{x}y + y\hat{x} \in \mathrm{Sym}(N,\mathbb{R}).$$

Then $(\mathrm{Sym}(N,\mathbb{R}), \bigtriangleup)$ is a non-associative algebra with unit element I_N. Let $\mathcal{Z}$ be a subalgebra of $(\mathrm{Sym}(N,\mathbb{R}), \bigtriangleup)$ and $H_{\mathcal{Z}}$ be the set of lower triangular matrices obtained from elements of $\mathcal{Z}$ with positive diagonal entries, that is, $H_{\mathcal{Z}} := \{\underset{\vee}{x} \colon x \in \mathcal{Z} \text{ and } x_{ii} > 0\}$.

Define $\Omega_{\mathcal{Z}} = \{x \in \mathcal{Z} \colon x \text{ is positive definite}\}$ and consider for any $T \in H_{\mathcal{Z}}$ the linear operators $\rho(T) \colon \mathcal{Z} \to \mathcal{Z}$, $x \mapsto \rho(T)x = TxT^\top$. It can be shown that $\rho(H_{\mathcal{Z}})$ acts on $\Omega_{\mathcal{Z}}$ transitively, which means that $\Omega_{\mathcal{Z}}$ is a homogeneous cone [22, Theorem 3]. Furthermore, we have the following

THEOREM 3.1 ([20]). *For a homogeneous cone $\Omega \subset V$, there exists a subalgebra $\mathcal{Z} \subset \mathrm{Sym}(N,\mathbb{R})$ and a linear isomorphism $\phi \colon V \to \mathcal{Z}$ such that $\phi(\Omega) = \Omega_{\mathcal{Z}}$.*

We say that $\mathcal{Z} \subset \mathrm{Sym}(N,\mathbb{R})$ *admits a normal block decomposition* if there exists a partition $N = n_1 + \cdots + n_r$ and subspaces $\mathcal{V}_{lk} \subset \mathrm{Mat}(n_l, n_k, \mathbb{R})$,

$1 \le k < l \le r$, such that $\mathcal{Z}$ is the set of symmetric matrices of the form

$$\begin{pmatrix} X_{11} & X_{21}^\top & \cdots & X_{r1}^\top \\ X_{21} & X_{22} & & X_{r2}^\top \\ \vdots & & \ddots & \\ X_{r1} & X_{r2} & & X_{rr} \end{pmatrix} \qquad \begin{pmatrix} X_{ll} = x_{ll} I_{n_l},\ x_{ll} \in \mathbb{R},\ 1 \le l \le r \\ X_{lk} \in \mathcal{V}_{lk},\ 1 \le k < l \le r \end{pmatrix}.$$

We will write $\mathcal{Z}_{\mathcal{V}}$ for this space.

THEOREM 3.2.

(i) ([22, Theorem 2]) *Let $\mathcal{Z}$ be a subalgebra of* $(\mathrm{Sym}(N,\mathbb{R}), \triangle)$ *with $I_N \in \mathcal{Z}$. Then there exists a permutation matrix w such that $w\mathcal{Z}w^\top$ admits a normal block decomposition.*

(ii) ([22, Proposition 2]) *$\mathcal{Z}_{\mathcal{V}}$ is a subalgebra of* $(\mathrm{Sym}(N,\mathbb{R}), \triangle)$ *if and only if the subspaces $\{\mathcal{V}_{lk}\}_{1\le k<l\le r}$ satisfy the following conditions:*

(V1) *$A \in \mathcal{V}_{lk}$, $B \in \mathcal{V}_{ki} \Rightarrow AB \in \mathcal{V}_{li}$ for any $1 \le i < k < l \le r$,*
(V2) *$A \in \mathcal{V}_{li}$, $B \in \mathcal{V}_{ki} \Rightarrow AB^\top \in \mathcal{V}_{lk}$ for any $1 \le i < k < l \le r$,*
(V3) *$A \in \mathcal{V}_{lk} \Rightarrow AA^\top \in \mathbb{R}I_{n_l}$ for any $1 \le k < l \le r$.*

If $\mathcal{Z} = \mathcal{Z}_{\mathcal{V}}$ we shall write $\Omega_{\mathcal{V}}$, $H_{\mathcal{V}}$ for $\Omega_{\mathcal{Z}}$ and $H_{\mathcal{Z}}$, respectively.

Condition (V3) allows us to define an inner product on $\mathcal{V}_{lk}$, $1 \le k < l \le r$, by

$$AA^\top = (A|A)I_{n_l}, \qquad A \in \mathcal{V}_{lk}.$$

We then define the *standard inner product* on $\mathcal{Z}_{\mathcal{V}}$ by

$$\langle x, y\rangle := \sum_{k=1}^r x_{kk}y_{kk} + 2\sum_{1\le k<l\le r} (X_{lk}|Y_{lk}), \qquad x, y \in \mathcal{Z}_{\mathcal{V}}.$$

We identify $\mathcal{Z}_{\mathcal{V}}^*$ with $\mathcal{Z}_{\mathcal{V}}$ using this inner product. Note that $\langle\cdot,\cdot\rangle$ coincides with the trace inner product only if $n_1 = \cdots = n_r = 1$.

Define a one-dimensional representation of $H_{\mathcal{V}}$ by

$$\chi_{\underline{s}}(T) := \prod_{k=1}^r t_{kk}^{2s_k},$$

where $\underline{s} = (s_1, \ldots, s_r) \in \mathbb{C}^r$. We have (see e.g. [8, Lemma 2.4])

LEMMA 3.3. *Let χ be a one-dimensional representation of $H_{\mathcal{V}}$. Then there exists $\underline{s} \in \mathbb{C}^r$ such that $\chi = \chi_{\underline{s}}$.*

This fact will be important later on.

For any open convex cone Ω we define the *dual cone* of Ω by

$$\Omega^* = \{\xi \in V^* \colon \langle \xi, x\rangle > 0\ \forall x \in \bar{\Omega} \setminus \{0\}\},$$

where V^* is the dual space of V. If Ω is homogeneous, then so is Ω^*. Let $\Omega_{\mathcal{V}}^*$ denote the dual cone of $\Omega_{\mathcal{V}}$.

For $T \in H_{\mathcal{V}}$, we denote by $\rho^*(T)$ the adjoint operator of $\rho(T) \in \mathrm{GL}(\mathcal{Z}_{\mathcal{V}})$ defined by $\langle \rho^*(T)\xi, x\rangle = \langle \xi, \rho(T)x\rangle$ for $x, \xi \in \mathcal{Z}_{\mathcal{V}}$. Then we see from [32, Chapter 1, Proposition 9] that for any $\xi \in \Omega_{\mathcal{V}}^*$, there exists a unique $T \in H_{\mathcal{V}}$ such that $\xi = \rho^*(T)I_N$.

DEFINITION 3.4. Let $\Delta_{\underline{s}}^*\colon \Omega_{\mathcal{V}}^* \to \mathbb{C}$ be the function given by

$$\Delta_{\underline{s}}^*(\xi) = \Delta_{\underline{s}}^*(\rho^*(T)I_N) := \chi_{\underline{s}^*}(T), \quad \text{where} \quad \underline{s}^* = (s_r, \dots, s_1) \in \mathbb{C}^n.$$

EXAMPLE 3.5. Let

$$\mathcal{Z}_{\mathcal{V}} := \left\{ \begin{pmatrix} x_{11} & 0 & x_{31} \\ 0 & x_{22} & x_{32} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} : x_{11}, x_{22}, x_{33}, x_{31}, x_{32} \in \mathbb{R} \right\}.$$

Conditions (V1)–(V3) are satisfied and we have $n_1 = n_2 = n_3 = 1$, $N = r = 3$. Then

$$\Omega_{\mathcal{V}} = \mathcal{Z}_{\mathcal{V}} \cap \mathrm{Sym}_+(3, \mathbb{R}) = \{x \in \mathcal{Z}_{\mathcal{V}}\colon x_{11} > 0,\ x_{22} > 0,\ \det x \in \mathbb{R}_{>0}\}$$

and the dual cone is

$$\Omega_{\mathcal{V}}^* = \{\xi \in \mathcal{Z}_{\mathcal{V}}\colon \xi_{33} > 0,\ \xi_{11}\xi_{33} > \xi_{31}^2,\ \xi_{22}\xi_{33} > \xi_{32}^2\}.$$

The cone $\Omega_{\mathcal{V}}^*$ is called the *Vinberg cone*, while $\Omega_{\mathcal{V}}$ is the *dual Vinberg cone.* The cones $\Omega_{\mathcal{V}}^*$ and $\Omega_{\mathcal{V}}$ are the lowest-dimensional non-symmetric homogeneous cones. Moreover, for $\xi \in \Omega_{\mathcal{V}}^*$, we have

$$\Delta_{\underline{s}}^*(\xi) = \left(\frac{\xi_{11}\xi_{33} - \xi_{31}^2}{\xi_{33}}\right)^{s_3} \left(\frac{\xi_{22}\xi_{33} - \xi_{32}^2}{\xi_{33}}\right)^{s_2} \xi_{33}^{s_1}.$$

We see that for any $S \in H_{\mathcal{V}}$ and $\xi = \rho^*(T)I_N \in \Omega_{\mathcal{V}}^*$ we have

$$\Delta_{\underline{s}}^*(\rho^*(S)\xi) = \chi_{\underline{s}^*}(TS) = \chi_{\underline{s}^*}(T)\chi_{\underline{s}^*}(S) = \Delta_{\underline{s}}^*(\xi)\Delta_{\underline{s}}^*(\rho^*(S)I_N), \tag{3.1}$$

and since any character of $H_{\mathcal{V}}$ is of the form $\chi_{\underline{s}^*}$, property (3.1) characterizes Δ^* (see [24]). The function Δ^* is sometimes termed a *generalized power function.* Its importance is emphasized by the following result [10, 18].

THEOREM 3.6. *There exists a positive measure $\mathcal{R}_{\underline{s}}$ on $\mathcal{Z}_{\mathcal{V}}$ with Laplace transform $L_{\mathcal{R}_{\underline{s}}}(-\theta) = \Delta_{-\underline{s}^*}^*(\theta)$ for $\theta \in \Omega_{\mathcal{V}}^*$ if and only if $\underline{s} \in \Xi := \bigsqcup_{\underline{\varepsilon} \in \{0,1\}^r} \Xi(\underline{\varepsilon})$, where*

$$\Xi(\underline{\varepsilon}) := \left\{ \underline{s} \in \mathbb{R}^r : \begin{matrix} s_k > p_k(\underline{\varepsilon})/2 \ \text{ if } \varepsilon_k = 1 \\ s_k = p_k(\underline{\varepsilon})/2 \ \text{ if } \varepsilon_k = 0 \end{matrix} \right\}, \qquad p_k(\underline{\varepsilon}) = \sum_{i<k} \varepsilon_i \dim \mathcal{V}_{ki}.$$

The measure $\mathcal{R}_{\underline{s}}$ with $\underline{s} \in \Xi$ has support in $\overline{\Omega_{\mathcal{V}}}$ and is called the *Riesz measure*, while the set Ξ is called the *Gindikin–Wallach set.*

In order to define the NEF generated by $\mathcal{R}_{\underline{s}}$ we have to know whether $\mathcal{R}_{\underline{s}} \in \mathcal{M}(\mathcal{Z}_{\mathcal{V}})$, that is, whether $\mathcal{R}_{\underline{s}}$ is not concentrated on an affine hyperplane of $\mathcal{Z}_\mathcal{V}$. The following result is a generalization of [13, Theorem 3.1]. For $\underline{\varepsilon} \in \{-1, 0, 1\}^r$, define

$$E_{\underline{\varepsilon}} := \begin{pmatrix} \varepsilon_1 I_{n_1} & & \\ & \ddots & \\ & & \varepsilon_r I_{n_r} \end{pmatrix} \in \mathcal{Z}_\mathcal{V}.$$

THEOREM 3.7. *The support of $\mathcal{R}_{\underline{s}}$ is not concentrated on any affine hyperplane in $\mathcal{Z}_\mathcal{V}$ if and only if $s_k > 0$ for all $k = 1, \ldots, r$.*

Proof. We write $\mathcal{O}_{\underline{\varepsilon}}$ for the $\rho(H_\mathcal{V})$-orbit in $\mathcal{Z}_\mathcal{V}$ through $E_{\underline{\varepsilon}}$. Note that $\mathcal{O}_{(1,\ldots,1)} = \Omega_\mathcal{V}$. It is shown in [18, Theorem 6.2] that if $\underline{s} \in \Xi(\underline{\varepsilon})$, then $\mathcal{R}_{\underline{s}}$ is a positive measure on $\mathcal{O}_{\underline{\varepsilon}}$, so that the support of $\mathcal{R}_{\underline{s}}$ coincides with the closure $\overline{\mathcal{O}_{\underline{\varepsilon}}}$. In particular, if for any $k = 1, \ldots, r$,

$$s_k > p_k(1, \ldots, 1)/2 = \frac{1}{2} \sum_{i<k} \dim \mathcal{V}_{ki}, \tag{3.2}$$

then $\underline{s} \in \Xi(1, \ldots, 1)$ and $\mathcal{R}_{\underline{s}}$ is a regular measure on the cone $\Omega_\mathcal{V} = \rho(H_\mathcal{V}) I_N$.

Now we show the 'if' part of the statement. Assume that $\underline{s} \in \Xi \cap \mathbb{R}^r_{>0}$. In view of (3.2), we see that there exists a positive integer m such that $m\underline{s} \in \Xi(1, \ldots, 1)$. Then $\mathcal{R}_{m\underline{s}}$ is a regular measure, because it equals the convolution measure $\mathcal{R}_{\underline{s}} * \cdots * \mathcal{R}_{\underline{s}}$ (m times). It follows that the support of $\mathcal{R}_{\underline{s}}$ is not concentrated on any affine hyperplane in $\mathcal{Z}_\mathcal{V}$.

Next we show the 'only if' part. It suffices to show that if $\underline{s} \in \Xi(\underline{\varepsilon})$ with $s_k = 0$ for some k, then $\operatorname{supp} \mathcal{R}_{\underline{s}} = \overline{\mathcal{O}_{\underline{\varepsilon}}} \subset (\mathbb{R}E_k)^\perp := \{x \in \mathcal{Z}_\mathcal{V} : x_{kk} = 0\}$. Recalling the definition of $\Xi(\underline{\varepsilon})$, we see that

$$\varepsilon_k = 0 \quad \text{and} \quad p_k(\underline{\varepsilon}) = \sum_{i<k} \varepsilon_i \dim \mathcal{V}_{ki} = 0,$$

and the latter equality implies

$$\mathcal{V}_{ki} = \{0\} \quad \text{if } \varepsilon_i = 1.$$

Therefore, for any $x = \rho(T) E_{\underline{\varepsilon}} = T E_{\underline{\varepsilon}} T^\top \in \mathcal{O}_{\underline{\varepsilon}}$ with $T \in H_\mathcal{V}$, we have

$$x_{kk} = \varepsilon_k (t_{kk})^2 + \sum_{i<k} \varepsilon_i \|T_{ki}\|^2 = 0,$$

which means that $\mathcal{O}_{\underline{\varepsilon}} \subset (\mathbb{R}E_k)^\perp$. Hence $\operatorname{supp} \mathcal{R}_{\underline{s}} \subset (\mathbb{R}E_k)^\perp$ and the proof is complete. ∎

For $\underline{\varepsilon} \in \{-1, 1\}^r$, consider $\mathcal{O}^*_{\underline{\varepsilon}} := \rho^*(H_\mathcal{V}) E_{\underline{\varepsilon}}$. The set

$$\bigsqcup_{\underline{\varepsilon} \in \{-1,1\}^r} \mathcal{O}^*_{\underline{\varepsilon}} \tag{3.3}$$

is dense in $\mathcal{Z}_\mathcal{V}$, and the $\mathcal{O}^*_{\underline{\varepsilon}}$ are the only open orbits of $\rho^*(H_\mathcal{V})$ (see [9, p. 77]).

3.2. Characterization of the Riesz measure on a homogeneous cone. In the following section we will give an application of Theorem 2.1 to a characterization of the Riesz measure on a homogeneous cone. We generalize the results of [13], where the characterization of the Riesz measure through invariance of NEFs on simple Euclidean algebras was considered.

We say that a subalgebra of $(\mathrm{Sym}(N,\mathbb{R}),\triangle)$ is *irreducible* if it is not equal to a direct sum of two non-trivial ideals.

THEOREM 3.8. *Let* $\mathbb{E}=\mathcal{Z}_{\mathcal{V}}$ *be an irreducible subalgebra of* $(\mathrm{Sym}(N,\mathbb{R}),\triangle)$ *that admits a normal block decomposition and let* $\mu\in\mathcal{M}(\mathbb{E})$. *Assume that* $F(\mu)$ *is a NEF invariant under* $G=\rho(H_{\mathcal{V}})$. *Then there exist* $\theta_0\in\mathcal{Z}_{\mathcal{V}}$, $a_0\in\mathbb{R}$ *and* $\underline{s}\in\Xi\cap\mathbb{R}^r_{>0}$ *such that*

$$\mu(dx)=e^{a_0-\langle\theta_0,x\rangle}\mathcal{R}_{\underline{s}}(dx)\quad \textit{or}\quad \mu(dx)=e^{a_0-\langle\theta_0,x\rangle}\mathcal{R}_{\underline{s}}(-dx).$$

Thanks to Theorem 3.1, Theorem 3.8 provides a characterization of Riesz measures on homogeneous cones (note that the support of the measure being characterized is a homogeneous cone).

Proof of Theorem 3.8. We have $\mathbb{R}_{>0}I_N\subset H_{\mathcal{V}}$ and so $c\,\mathrm{Id}\in\rho(H_{\mathcal{V}})$ for all $c>0$. Theorem 2.1 implies that there exists θ_0 such that

$$\mu(dx)=e^{-\langle\theta_0,x\rangle}\mu_0(dx),$$

where (by (2.1))

$$L_{\mu_0}(\rho^*(T)\theta)=\chi(T)L_{\mu_0}(\theta),\quad (\theta,T)\in\Theta(\mu_0)\times H_{\mathcal{V}}, \tag{3.4}$$

with a certain character χ of $H_{\mathcal{V}}$. We will determine L_{μ_0} and $\Theta(\mu_0)$. The proof is split into four steps:

(i) There exists $\underline{\varepsilon}\in\{-1,1\}^r$ such that $E_{\underline{\varepsilon}}\in\Theta(\mu_0)$. Moreover, if $E_{\underline{\varepsilon}},E_{\underline{\varepsilon}'}\in\Theta(\mu_0)$, then $\underline{\varepsilon}=\underline{\varepsilon}'$.
(ii) $\Theta(\mu_0)=\mathcal{O}^*_{\underline{\varepsilon}}$ for some $\underline{\varepsilon}\in\{-1,1\}^r$.
(iii) If $\Theta(\mu_0)=\mathcal{O}^*_{\underline{\varepsilon}}$ then $\underline{\varepsilon}=(-1,\ldots,-1)$ or $\underline{\varepsilon}=(1,\ldots,1)$.
(iv) μ is of the postulated form.

First step. By definition, the set $\Theta(\mu_0)$ is open and non-empty. Since $\rho^*(T)\Theta(\mu_0)=\Theta(\mu_0)$ for any $T\in H_{\mathcal{V}}$ and the set (3.3) is dense in $\mathcal{Z}_{\mathcal{V}}$, we have $E_{\underline{\varepsilon}}\in\mathcal{O}^*_{\underline{\varepsilon}}\subset\Theta(\mu_0)$ for some $\underline{\varepsilon}\in\{-1,1\}^r$.

Assume now $E_{\underline{\varepsilon}},E_{\underline{\varepsilon}'}\in\Theta(\mu_0)$ and $\underline{\varepsilon}\neq\underline{\varepsilon}'$. Since $\Theta(\mu_0)$ is convex, we know that $\sigma:=\frac{1}{2}(E_{\underline{\varepsilon}}+E_{\underline{\varepsilon}'})\in\Theta(\mu_0)$. Set $I_0:=\{i\in\{1,\ldots,r\}\colon(\varepsilon_i+\varepsilon_i')/2=0\}$. Then I_0 is not empty. Define

$$H_0:=\{T\in H_{\mathcal{V}}\colon t_{ii}=1\text{ for any }i\notin I_0\}.$$

If $T\in H_0$ is diagonal, then

$$\rho^*(T)\sigma=\sigma.$$

On the other hand, by (3.4) we obtain

$$L_{\mu_0}(\rho^*(T)\sigma) = \chi(T)L_{\mu_0}(\sigma),$$

which implies that $\chi(T) = 1$ for any $T \in H_0$ (note that $\chi(T)$ depends on T only through diagonal elements; see Lemma 3.3). Further, this implies that for any diagonal $T = \operatorname{diag}(t_{11}, \ldots, t_{rr}) \in H_0$,

$$L_{\mu_0}(\rho^*(T)E_{\underline{\varepsilon}}) = L_{\mu_0}(E_{\underline{\varepsilon}}),$$

where $\rho^*(T)E_{\underline{\varepsilon}} = \operatorname{diag}(t_{11}\varepsilon_1, \ldots, t_{rr}\varepsilon_r)$. This contradicts the assumption that μ_0 is not supported on any affine hyperplane of $\mathcal{Z}_{\mathcal{V}}$. Indeed, the function $(0,\infty) \ni t_{ii} \mapsto \int \exp(t_{ii}\varepsilon_i x_{ii})\,\mu_0(dx)$ is constant if and only if the support of μ_0 is contained in the subspace of $\mathcal{Z}_{\mathcal{V}}$ whose (i,i)-components are zero for any $i \in I_0$.

Second step. Since $\Theta(\mu_0)$ contains only one open orbit, we have

$$\operatorname{int}(\Theta(\mu_0) \setminus \mathcal{O}^*_{\underline{\varepsilon}}) = \emptyset.$$

This implies that

$$\Theta(\mu_0) \setminus \overline{\mathcal{O}^*_{\underline{\varepsilon}}} = \emptyset, \quad \text{thus} \quad \mathcal{O}^*_{\underline{\varepsilon}} \subset \Theta(\mu_0) \subset \overline{\mathcal{O}^*_{\underline{\varepsilon}}},$$

which proves the claim, since $\Theta(\mu_0)$ is open.

Third step. We will show that $\mathcal{O}^*_{\underline{\varepsilon}}$ is not convex unless ε_k, $1 \le k \le r$, are all 1 or all -1. Define

$$I^+(\underline{\varepsilon}) := \big\{i \in \{1, \ldots, r\} : \varepsilon_i = 1\big\} \quad \text{and} \quad I^-(\underline{\varepsilon}) := \{1, \ldots, r\} \setminus I^+(\underline{\varepsilon}).$$

Suppose that $I^+(\underline{\varepsilon})$ and $I^-(\underline{\varepsilon})$ are non-empty. Then there exist $k \in I^+(\underline{\varepsilon})$ and $l \in I^-(\underline{\varepsilon})$ such that $\mathcal{V}_{lk}$ is not $\{0\}$ (without loss of generality we may assume that $l > k$). If not, $\mathcal{Z}_{\mathcal{V}}$ has the block form

$$w\left(\begin{array}{c|c} \mathcal{Z}_+ & 0 \\ \hline 0 & \mathcal{Z}_- \end{array}\right)w^\top$$

for some permutation matrix w. This contradicts the assumption that $\mathcal{Z}_{\mathcal{V}}$ is irreducible.

Thus there exists $\mathcal{V}_{lk} \neq \{0\}$. We have $\varepsilon_k = 1$ and $\varepsilon_l = -1$. For $v \in \mathcal{V}_{lk}$ let $T(v) \in H_{\mathcal{V}}$ be such that $T_{lk}(v) = v$, $t_{ii}(v) = 1$ for $i = 1, \ldots, r$ and $T_{ji}(v) = 0$ for all $(i,j) \neq (k,l)$, $1 \le i < j \le r$. Take any $v \in \mathcal{V}_{lk}$ with $(v|v) = 2$. Then

$$\tfrac{1}{2}\big(\rho^*(T(v))E_{\underline{\varepsilon}} + \rho^*(T(-v))E_{\underline{\varepsilon}}\big) = E_{\underline{\varepsilon}'}, \tag{3.5}$$

where $\varepsilon'_i = \varepsilon_i$ for $i \neq k$ and $\varepsilon'_k = -1$. We will use the definition of $\rho^*(T)$, but there is another natural approach; see Remark 3.9 after the proof. For any $x \in \mathcal{Z}_{\mathcal{V}}$ consider the matrix

$$x_v := \tfrac{1}{2}\big(\rho(T(v))x + \rho(T(-v))x\big) = \tfrac{1}{2}\big(T(v)xT(v)^\top + T(-v)xT(-v)^\top\big).$$

It may be verified by direct calculation that the matrices x and x_v differ only in their (l,l)-components, which in the latter case equals

$$(vv^\top x_{kk} + x_{ll})I_{n_l} = (2x_{kk} + x_{ll})I_{n_l}.$$

Thus,

$$\begin{aligned}\left\langle \tfrac{1}{2}\big(\rho^*(T(v))E_{\underline{\varepsilon}} + \rho^*(T(-v))E_{\underline{\varepsilon}}\big), x\right\rangle &= \left\langle E_{\underline{\varepsilon}}, \tfrac{1}{2}\big(\rho(T(v))x + \rho(T(-v))x\big)\right\rangle \\ &= \sum_{i\notin\{k,l\}} x_{ii}\varepsilon_i + x_{kk} + (2x_{kk}+x_{ll})(-1) = \left\langle E_{\underline{\varepsilon}'}, x\right\rangle.\end{aligned}$$

But $\Theta(\mu_0)$ is convex, so (3.5) implies $E_{\underline{\varepsilon}'} \in \Theta(\mu_0)$. This contradicts (i). Thus $I^+(\underline{\varepsilon}) = \{1,\ldots,r\}$ or $I^-(\underline{\varepsilon}) = \{1,\ldots,r\}$, and so $-I_N$ or I_N belongs to $\Theta(\mu_0)$. Finally, by (ii) and the fact that $\rho^*(H_{\mathcal{V}})I_N = \Omega^*_{\mathcal{V}}$,

$$\Theta(\mu_0) = -\Omega^*_{\mathcal{V}} \quad \text{or} \quad \Theta(\mu_0) = \Omega^*_{\mathcal{V}}.$$

Fourth step. Let us first consider the case $\Theta(\mu_0) = -\Omega^*_{\mathcal{V}}$. Putting $\theta = \rho^*(S)(-I_N)$ for $S \in H_{\mathcal{V}}$, we obtain

$$L_{\mu_0}(-\rho^*(ST)I_N) = \chi(T)L_{\mu_0}(-\rho^*(S)I_N), \quad (S,T) \in H^2_{\mathcal{V}},$$

which implies that for $a_0 = \log L_{\mu_0}(-I_N)$,

$$L_{\mu_0}(-\rho^*(T)I_N) = e^{a_0}\chi(T)$$

and $\chi(T)$ is a one-dimensional representation of $H_{\mathcal{V}}$. Thus, by Lemma 3.3, there exists $\underline{s} \in \Xi$ such that $\chi = \chi_{-\underline{s}}$ and then, by the definition of $\Delta^*_{\underline{s}^*}$,

$$L_{\mu_0}(-\theta) = e^{a_0}\Delta^*_{-\underline{s}^*}(\theta), \quad \theta \in \Omega^*_{\mathcal{V}} = -\Theta(\mu_0).$$

By Theorems 3.6 and 3.7 we see that $\underline{s} \in \Xi \cap \mathbb{R}^r_{>0}$, and that $\mu_0(dx) = e^{a_0}\mathcal{R}_{\underline{s}}(dx)$ and $\Theta(\mu_0) = -\Omega^*_{\mathcal{V}}$.

When $\Theta(\mu_0) = \Omega^*_{\mathcal{V}}$, one shows similarly that $\mu_0(dx) = e^{a_0}\mathcal{R}_{\underline{s}}(-dx)$. ■

Remark 3.9. To show (3.5) we could switch to a matrix realization of the dual cone $\Omega^*_{\mathcal{V}}$, where (under a suitable linear isomorphism) $\rho^*(T)$ is just left multiplication by some upper triangular matrix and its transpose is right multiplication.

3.3. Comments. (1) In [14], the authors considered a characterization of the Riesz measure $\mathcal{R}_{\underline{s}}$ through the invariance property of a NEF on a simple Euclidean algebra $\mathbb{E}$ under some subgroup G of $\mathrm{GL}(\mathbb{E})$. This subgroup was carefully chosen in order to ensure that some components of the vector $\underline{s} \in \Xi$ are equal. Taking the ordinary triangular group $\rho(H_{\mathcal{V}})$ imposes no additional conditions on these components. On the other hand, if one considers the invariance of NEF under the identity component of $G(\Omega)$ (this is in principle what Letac did for symmetric matrices in [25], but it is true on homogeneous cones also; see comment (4) below), then all s_i have to be equal, $s_i = p$ for $1 \le i \le r$. Then $\Delta^*_{\underline{s}^*} = \det^p$ and p belongs to a set Λ called the *Jørgensen set* (see [6]).

(2) Elements of $F(\mathcal{R}_{\underline{s}})$ are actually the Wishart distributions on homogeneous cones introduced in [1] (the subcase of $s_i = p$ for $1 \le i \le r$) and in [11].

(3) It should be stressed that our approach is very different from [13, 25], where the characterization of NEFs was proved by showing that the variance function of a $\rho(H_{\mathcal{V}})$-invariant NEF coincides with the one of $F(\mathcal{R}_{\underline{s}})$ for some $\underline{s} \in \Xi$. In the present paper we need not even know what is the variance function of the Riesz measure on a homogeneous cone. We perceive our approach as less technical and more natural.

(4) We can rephrase our result in terms of homogeneous cones as follows. Let $\Omega \subset \mathbb{E}$ be an irreducible homogeneous cone, and $G \subset \mathrm{GL}(\mathbb{E})$ a linear algebraic group acting on Ω transitively. Then any G-invariant NEF is generated by a Riesz distribution on Ω or $-\Omega$ because G contains a triangular subgroup isomorphic to the group $\rho(H_{\mathcal{V}})$ discussed in the present paper (cf. [32, Chapter 1, Section 9]).

Acknowledgements. We would like to thank the anonymous referee for helpful comments.

H. Ishi was partially supported by JSPS KAKENHI Grant Number 16K05174 and JST PRESTO. B. Kołodziejek was partially supported by NCN Grant No. 2012/05/B/ST1/00554.

REFERENCES

[1] S. A. Andersson and G. G. Wojnar, *Wishart distributions on homogeneous cones*, J. Theoret. Probab. 17 (2004), 781–818.

[2] O. Barndorff-Nielsen, *Information and Exponential Families in Statistical Theory*, Wiley, Chichester, 2014.

[3] M. Casalis, *Familles exponentielles naturelles sur* $\mathbb{R}^d$ *invariantes par un groupe*, Int. Statist. Rev. 59 (1991), 241–262.

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

[5] M. Casalis, *The* $2d + 4$ *simple quadratic natural exponential families on* $\mathbf{R}^d$, Ann. Statist. 24 (1996), 1828–1854.

[6] M. Casalis and G. Letac, *Characterization of the Jørgensen set in generalized linear models*, Test 3 (1994), 145–162.

[7] P. S. Eriksen, $(k, 1)$ *exponential transformation models*, Scand. J. Statist. 11 (1984), 129–145.

[8] G. Garrigós, *Generalized Hardy spaces on tube domains over cones*, Colloq. Math. 90 (2001), 213–251.

[9] S. G. Gindikin, *Analysis in homogeneous domains*, Russian Math. Surveys 19 (1964), 1–89.

[10] S. G. Gindikin, *Invariant generalized functions in homogeneous domains*, Funktsional. Anal. i Prilozhen. 9 (1975), no. 1, 56–58 (in Russian).

[11] P. Graczyk and H. Ishi, *Riesz measures and Wishart laws associated to quadratic maps*, J. Math. Soc. Japan 66 (2014), 317–348.

[12] P. Graczyk, H. Ishi and B. Kołodziejek, *Wishart laws and variance function on homogeneous cones*, arXiv:1802.02352 (2018).

[13] A. Hassairi and S. Lajmi, *Riesz exponential families on symmetric cones*, J. Theoret. Probab. 14 (2001), 927–948.

[14] A. Hassairi and S. Lajmi, *Classification of Riesz exponential families on a symmetric cone by invariance properties*, J. Theoret. Probab. 17 (2004), 521–539.

[15] A. Hassairi and M. Zarai, *Characterization of the cubic exponential families by orthogonality of polynomials*, Ann. Probab. 32 (2004), 2463–2476.

[16] A. Hassairi and M. Zarai, *Characterization of the simple cubic multivariate exponential families*, J. Funct. Anal. 235 (2006), 69–89.

[17] S. Højsgaard and S. L. Lauritzen, *Graphical Gaussian models with edge and vertex symmetries*, J. Roy. Statist. Soc. Ser. B Statist. Methodol. 70 (2008), 1005–1027.

[18] H. Ishi, *Positive Riesz distributions on homogeneous cones*, J. Math. Soc. Japan 52 (2000), 161–186.

[19] H. Ishi, *On symplectic representations of normal j-algebras and their application to Xu's realizations of Siegel domains*, Differential Geom. Appl. 24 (2006), 588–612.

[20] H. Ishi, *Representation of clans and homogeneous cones*, Vestn. Tambov. Univ. 16 (2011), 1669–1675.

[21] H. Ishi, *Homogeneous cones and their applications to statistics*, in: Modern Methods of Multivariate Statistics, Travaux en Cours 82, Hermann, 2014, 135–154.

[22] H. Ishi, *Matrix realization of a homogeneous cone*, in: Geometric Science of Information, Lecture Notes in Computer Sci. 9389, Springer, 2015, 248–256.

[23] J. L. Jensen, *On the hyperboloid distribution*, Scand. J. Statist. 8 (1981), 193–206.

[24] B. Kołodziejek, *Multiplicative Cauchy functional equation on symmetric cones*, Aequationes Math. 89 (2015), 1075–1094.

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

[26] G. Letac, *Lectures on Natural Exponential Families and their Variance Functions*, Monografías de Matemática 50, Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1992.

[27] G. Letac and H. Massam, *Wishart distributions for decomposable graphs*, Ann. Statist. 35 (2007), 1278–1323.

[28] G. Letac and M. Mora, *Natural real exponential families with cubic variance functions*, Ann. Statist. 18 (1990), 1–37.

[29] C. N. Morris, *Natural exponential families with quadratic variance functions*, Ann. Statist. 10 (1982), 65–80.

[30] H. Shima, *Homogeneous Hessian manifolds*, Ann. Inst. Fourier (Grenoble) 30 (1980), 91–128.

[31] H. Shima, *The Geometry of Hessian Structures*, World Sci., Hackensack, NJ, 2007.

[32] È. B. Vinberg, *The theory of homogeneous convex cones*, Trudy Moskov. Mat. Obshch. 12 (1963), 303–358 (in Russian).

[33] G. S. Watson, *Statistics on Spheres*, Wiley, New York, 1983.

Hideyuki Ishi
Graduate School of Mathematics
Nagoya University
Furo-cho
Nagoya 464-8602, Japan
and
JST PRESTO, 4-1-8 Honcho
Kawaguchi 332-0012, Japan
E-mail: hideyuki@math.nagoya-u.ac.jp

Bartosz Kołodziejek
Faculty of Mathematics and Information Science
Warsaw University of Technology
Koszykowa 75
00-662 Warszawa, Poland
E-mail: b.kolodziejek@mini.pw.edu.pl