PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 30, Fasc. 1 (2010), pp. 121–139

BIVARIATE NATURAL EXPONENTIAL FAMILIES WITH LINEAR DIAGONAL VARIANCE FUNCTIONS*

BY

JOANNA CHACHULSKA (WARSZAWA)

Abstract. It is well known that natural exponential families (NEFs) are uniquely determined by their variance functions (VFs). However, there exist examples showing that even an incomplete knowledge of a matrix VF can be sufficient to determine a multivariate NEF. Following such an idea, in this paper a complete description of bivariate NEFs with linear diagonal of the matrix VF is given. As a result we obtain the families of distributions with marginals that are some combinations of Poisson and normal distributions. Furthermore, the characterization extends (in two-dimensional case) the classification of NEFs with linear matrix VF obtained by Letac [11]. The main result is formulated in terms of regression properties.

2000 AMS Mathematics Subject Classification: Primary: 62E10, 44A10; Secondary: 60E05, 60E10.

Key words and phrases: Natural exponential families, variance functions, Laplace transforms.

1. INTRODUCTION

1.1. Natural exponential families. Let μ be a positive measure on $\mathbb{R}^n$. Its Laplace transform is defined as

$$L_\mu(\theta) = \int_{\mathbb{R}^n} \exp\langle\theta, x\rangle \mu(dx) \tag{1.1}$$

and $\Theta(\mu) = \mathrm{Int}\{\theta \in \mathbb{R}^n\colon\ L_\mu(\theta) < \infty\}$. We write k_μ for the cumulant function of μ, that is, $k_\mu(\theta) = \log L_\mu(\theta)$ for $\theta \in \Theta(\mu)$. We will denote by $\mathcal{M}$ the set of measures μ such that $\Theta(\mu) \neq \emptyset$ and μ is not concentrated on an affine hyperplane of $\mathbb{R}^n$. If $\mu \in \mathcal{M}$, the set of probabilities

$$F(\mu) = \{P(\theta, \mu)(dx) :\ \theta \in \Theta(\mu)\}, \tag{1.2}$$

where $P(\theta, \mu)(dx) = \exp\{\langle\theta, x\rangle - k_\mu(\theta)\}\mu(dx)$, is called the *natural exponential family* (NEF) generated by μ, see [12].

* Research supported by MNiSW grant N20104332/3272.

Since k_μ is strictly convex on $\Theta(\mu)$, its derivative $k'_\mu : \Theta(\mu) \to M_F$ defines a diffeomorphism; M_F is called the *domain of means* of μ. Let us denote by $\Psi_\mu : M_F \to \Theta(\mu)$ the inverse function of k'_μ. The mapping $\mathbf{m} \to P\big(\Psi_\mu(\mathbf{m}), \mu\big)$ is bijective; it is a parametrization by the mean of the NEF F. We define the variance function (VF) of the NEF by

$$V_F(\mathbf{m}) = k''_\mu\big(\Psi_\mu(\mathbf{m})\big) = [\Psi'_\mu(\mathbf{m})]^{-1}, \tag{1.3}$$

and using the matrix notation:

$$V_F(\mathbf{m}) = \left[\frac{\partial^2 k_\mu(\theta)}{\partial\theta_i\partial\theta_j}\Big|_{\theta=\Psi_\mu(\mathbf{m})}\right]_{i,j=1}^{n}, \qquad \mathbf{m} \in M_F. \tag{1.4}$$

It is well known that the mapping $\mathbf{m} \to V_F(\mathbf{m})$ on M_F characterizes NEF uniquely. It depends only on the family F, not on a particular measure used to generate it. Using this fact Letac [11] described all NEFs with linear VF: $V_F(\mathbf{m}) = B\mathbf{m} + C$, where $B : \mathbb{R}^n \to \mathbb{S}_n$ is a linear operator, $C \in \mathbb{S}_n$, and $\mathbb{S}_n$ is the space of $n \times n$ symmetric real matrices. This result was extended by Casalis [3], [4], where a simple quadratic matrix VF was considered: $V_F(\mathbf{m}) = a\mathbf{m} \otimes \mathbf{m} + B\mathbf{m} + C$, $\mathbf{m} \otimes \mathbf{m} = [m_i m_j]_{i,j=1}^n$. Hassairi and Zarai [7] generalized Casalis' result to cubic NEFs, i.e. NEFs with cubic variance functions. On the other hand, there were successful attempts to identify the NEFs by incomplete knowledge of the variance function. Kokonendji and Seshadri [9] gave a characterization of the Gaussian law in $\mathbb{R}^n$ based on the fact that $\det V(\mathbf{m}) = \mathrm{const}$. Kokonendji and Masmoudi [8] characterized Poisson-Gaussian families by generalized variance extending the result given in [9] for the Gaussian families. Letac and Wesołowski [13] classified NEFs with VF of the type $p^{-1}\mathbf{m} \otimes \mathbf{m} - \varphi(\mathbf{m})M_\nu$, where M_ν is a symmetric matrix associated with a quadratic form ν, and $\mathbf{m} \mapsto \varphi(\mathbf{m})$ an unknown real function. The diagonal family of NEFs in $\mathbb{R}^n$:

$$\operatorname{diag} V(\mathbf{m}) = \big(f_1(m_1), \ldots, f_n(m_n)\big), \qquad \mathbf{m} = (m_1, \ldots, m_n), \tag{1.5}$$

was considered by Bar-Lev et al. [2]. This class was entirely characterized only by the diagonal of VF and it was shown that f_i, $i = 1, 2, \ldots, n$, have to be polynomials of degree at most two.

It is of interest to describe other NEFs with $f_i = f_i(m_1, \ldots, m_n)$, $i = 1, \ldots, n$, in (1.5). In general it seems to be a rather difficult task. In this paper we solve this problem for $n = 2$ and $f_i(\mathbf{m})$, $i = 1, 2$, being affine functions of $\mathbf{m} = (m_1, m_2)$.

Since the above condition on the diagonal of VF can be formulated using regression properties, this identification of NEFs can be considered in the framework of regression characterizations. Some of the papers relevant to the subject are Laha and Lukacs [10], Fosam and Shanbhag [5], Gordon [6], Bar-Lev [1].

This condition on the diagonal of the VF leads to a system of linear PDEs for the cumulant transform. Following a routine technique we conclude that a general solution of this system is of the form

$$k(\theta) = \sum_{\psi\in\Psi} Q_\psi(\theta)e^{\langle\psi,\theta\rangle}, \tag{1.6}$$

where Q_ψ are certain polynomials and $\Psi \subset \mathbb{C}^2$ is a finite set (see Appendix 6.1). The basic difficulty we deal with in this paper is to identify all probabilistic solutions among functions (1.6), that is to decide whether $L = e^k$ for k given by (1.6) is the Laplace transform of a probability measure. The crucial problem is to find the admissible forms of polynomials Q_ψ, $\psi \in \Psi$. This is the basic reason why we restrict ourselves to the case $n = 2$. Though in the case $n > 2$ the cumulant transform obtained as a solution of the system of PDEs is also of the form (1.6), the number of elements of the set Ψ does not allow to translate the methods we developed for $n = 2$ to this case; see Section 5.

The paper is organized as follows. Section 2 contains the main result (Theorem 2.1). The third section collects auxiliary propositions useful in the proof of Theorem 2.1. For the sake of clarity the detailed proof will be divided into two parts. The main part will be presented in the fourth section and the rest of the proof will be stated in Appendix 6.3. Section 5 is intended to explain the difficulties of analogous investigations in higher dimensions.

1.2. Variance functions and regression conditions. The condition that the diagonal of the matrix VF is an affine function, that is

$$\operatorname{diag} V(m_1, m_2) = (am_1 + bm_2 + e, cm_1 + dm_2 + f),$$

in terms of $k_\mu = k$ takes the form

$$\frac{\partial^2 k}{\partial\theta_1^2} = a\frac{\partial k}{\partial\theta_1} + b\frac{\partial k}{\partial\theta_2} + e, \tag{1.7}$$

$$\frac{\partial^2 k}{\partial\theta_2^2} = c\frac{\partial k}{\partial\theta_1} + d\frac{\partial k}{\partial\theta_2} + f, \tag{1.8}$$

where $a, b, c, d, e, f \in \mathbb{R}$.

On the basis of the conditions above one can provide examples of NEFs which do not belong to the diagonal family (see [2]) and with the VFs that are not affine functions of m_1 and m_2 (see [7]). We present an example of such a measure by giving its Laplace transform.

EXAMPLE 1.1. Let ν be the distribution of (X_1, X_2) with

$$\begin{aligned} L(\theta_1, \theta_2) &= \\ &= \exp\left(\frac{e}{2}\theta_1^2 + D_1\big(\exp(\theta_1+\theta_2) - 1\big) + D_2\big(\exp(-\theta_1+\theta_2) - 1\big)\right), \end{aligned} \tag{1.9}$$

where $e/2 > 0$. The diagonal of the VF generated by ν has the form

$$\operatorname{diag} V(m_1, m_2) = (m_2 + e, m_2),$$

and hence it is not a matrix VF of the diagonal family. The off-diagonal entry of V:

$$\frac{\partial^2 k}{\partial\theta_1 \partial\theta_2} = \frac{\partial k}{\partial\theta_1} - \theta_1,$$

is not an affine function of $\partial k/\partial\theta_1 (= m_1)$ and $\partial k/\partial\theta_2 (= m_2)$.

We now proceed to present (1.7) and (1.8) in terms of regression conditions. For any NEF $F(\mu)$ with a variance function V and $\theta_0 \in \operatorname{Int}\Theta(\mu)$ one can construct a random variable X with $P(\theta_0, \mu)(dx)$ as its distribution. It follows that $0 \in \operatorname{Int}\Theta_X = \operatorname{Int}\{\theta : \mathbb{E}e^{\langle \mathbf{X}, \theta\rangle} < \infty\}$. Let $k(\theta) = \log \mathbb{E}\exp\langle \mathbf{X}, \theta\rangle = \log L(\theta)$. Then

$$\frac{\partial k}{\partial\theta_i} = \frac{1}{L}\frac{\partial L}{\partial\theta_i}, \tag{1.10}$$

$$\frac{\partial^2 k}{\partial\theta_i^2} = -\frac{1}{L^2}\left(\frac{\partial L}{\partial\theta_i}\right)^2 + \frac{1}{L}\left(\frac{\partial^2 L}{\partial\theta_i^2}\right), \qquad i = 1, 2. \tag{1.11}$$

REMARK 1.1. *Let* $\mathbf{X} = (X_1, X_2)$ *and* $\mathbf{Y} = (Y_1, Y_2)$ *be two identically distributed independent random vectors such that* $\mathbf{X} \sim P(\theta_0, \mu)$, $\mathbb{E}\exp\langle \mathbf{X}, \theta\rangle < \infty$, $\theta \in \Theta_X$ *and* $\operatorname{Int}\Theta_X \ni 0$, *where* $\Theta_X = \Theta(\mu) - \theta_0$. *Then the following conditions are equivalent to* (1.7) *and* (1.8), *respectively*:

$$\mathbb{E}\big((X_1 - Y_1)^2 | \mathbf{X} + \mathbf{Y}\big) = 2e + a(X_1 + Y_1) + b(X_2 + Y_2), \tag{1.12}$$

$$\mathbb{E}\big((X_2 - Y_2)^2 | \mathbf{X} + \mathbf{Y}\big) = 2f + c(X_1 + Y_1) + d(X_2 + Y_2). \tag{1.13}$$

Let us write the conditions equivalent to (1.12) and (1.13):

$$\mathbb{E}(X_1^2 - X_1Y_1)\exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle) =$$
$$a\mathbb{E}\big(X_1 \exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle)\big) + b\mathbb{E}\big(X_2 \exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle)\big) + e\mathbb{E}\exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle),$$
$$\mathbb{E}(X_2^2 - X_2Y_2)\exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle) =$$
$$c\mathbb{E}\big(X_1 \exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle)\big) + d\mathbb{E}\big(X_2 \exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle)\big) + f\mathbb{E}\exp(\langle\theta, \mathbf{X} + \mathbf{Y}\rangle).$$

By the argument of independence, an equivalent formulation of the above is:

$$\frac{\partial^2 L}{\partial\theta_1^2} \cdot L - \left(\frac{\partial L}{\partial\theta_1}\right)^2 = a\frac{\partial L}{\partial\theta_1} \cdot L + b\frac{\partial L}{\partial\theta_2} \cdot L + eL^2,$$
$$\frac{\partial^2 L}{\partial\theta_2^2} \cdot L - \left(\frac{\partial L}{\partial\theta_2}\right)^2 = c\frac{\partial L}{\partial\theta_1} \cdot L + d\frac{\partial L}{\partial\theta_2} \cdot L + fL^2.$$

Using (1.10) and (1.11) we can see that (1.7) and (1.8) are other equivalent formulations of (1.12) and (1.13).

REMARK 1.2. *The following conditions are equivalent to* (1.12) *and* (1.13), *respectively*:

$$\mathbb{E}(X_1^2 - X_1Y_1 - aX_1 - bX_2 - e|\mathbf{X}+\mathbf{Y}) = 0, \tag{1.14}$$

$$\mathbb{E}(X_2^2 - X_2Y_2 - cX_1 - dX_2 - f|\mathbf{X}+\mathbf{Y}) = 0. \tag{1.15}$$

REMARK 1.3. *In* (1.7) *and* (1.8) *constants* e *and* f *can be eliminated only in the case when* $ad - bc \neq 0$.

In such a case it is sufficient to take an appropriate Dirac measure and consider its convolution with the distribution of $\mathbf{X}$. The question is whether there exist constants α and β such that $\tilde{k}(\theta_1, \theta_2) = k(\theta_1, \theta_2) + \alpha\theta_1 + \beta\theta_2$ satisfies homogeneous PDE. Thus the necessary conditions for α and β are

$$a\alpha + b\beta = e \quad \text{and} \quad c\alpha + d\beta = f.$$

These equations have the solution if $ad - bc \neq 0$.

2. MAIN RESULT

We can now formulate our main result. The following characterization based on the regression property provides (also through Remark 1.1) a classification of measures generating NEFs with affine diagonal of a matrix variance function.

THEOREM 2.1. *Let* $\mathbf{X} = (X_1, X_2)$ *and* $\mathbf{Y} = (Y_1, Y_2)$ *be two identically distributed independent random vectors such that* $\mathbb{E}\exp(\langle \mathbf{X}, \theta\rangle) < \infty$ *for* $\theta \in \Theta$, $\operatorname{Int}\Theta \ni 0$ *and* (1.12), (1.13) *hold. Then the distribution of* $\mathbf{X}$ *is determined by*

$$X_1 = \sum_{i=1}^{3} \varepsilon_i Z_i + \varepsilon_4 Z_4 + \xi_1 G_1 + \xi_2 G_2 + s_1, \tag{2.1}$$

$$X_2 = \sum_{i=1}^{3} \omega_i Z_i + \omega_4 Z_5 + \rho_1 G_1 + \rho_2 G_3 + s_2, \tag{2.2}$$

where Z_i, $i = 1, \ldots, 5$, G_j, $j = 1, 2, 3$, *are independent random variables such that* Z*'s have Poisson and* G*'s standard normal distributions*, $s_1, s_2 \in \mathbb{R}$. *Coefficients* $\varepsilon_i, \omega_i, \xi_j, \rho_j \in \mathbb{R}$, $i = 1, \ldots, 4$, $j = 1, 2$, *depend on parameters* $a, \ldots, f$.

REMARK 2.1. *In Theorem* 2.1 *only the following ten cases are possible*:

I. *If* $ad - bc \neq 0$, *then* $\xi_1 = \xi_2 = \rho_1 = \rho_2 = 0$ (*there is no Gaussian part in* (2.1) *and* (2.2)).

1. *If* $bc \neq 0$, *then*

$$\varepsilon_4 = \omega_4 = 0, \quad s_1 = \frac{ed - fb}{bc - ad}, \quad s_2 = \frac{af - ec}{bc - ad},$$

and if there exist j non-zero different real roots $(\alpha_i,\ i=1,\dots,j,\ j\leqslant 3)$ *of the equation*

$$\alpha^3-2a\alpha^2+\alpha(a^2-bd)+b(ad-bc)=0,$$

then $\varepsilon_i=\alpha_i$, $\omega_i=(\alpha_i^2-a\alpha_i)/b$ *for* $i=1,\dots,j$, *and* $\varepsilon_i=\omega_i=0$ *for* $i=j+1,\dots,3$.

2. *If* $b\neq 0, c=0$, *then* $\varepsilon_4=a, \varepsilon_1=\omega_1=\omega_4=0, s_2=-f/d, s_1=(bf-ed)/ad$, *and*
if $a^2+4bd>0$, *then*

$$\varepsilon_2=\frac{a-\sqrt{a^2+4bd}}{2},\qquad \varepsilon_3=\frac{a+\sqrt{a^2+4bd}}{2},\qquad \omega_2=\omega_3=d,$$

if $a^2+4bd=0$, *then* $\varepsilon_2=a/2, \varepsilon_3=0, \omega_2=d, \omega_3=0$,
if $a^2+4bd<0$, *then* $\varepsilon_2=\varepsilon_3=0, \omega_2=\omega_3=0$.

3. *If* $b=0, c\neq 0$, *then* $\omega_4=d, \varepsilon_4=\omega_1=\varepsilon_1=0, s_1=-e/a, s_2=(ce-fa)/da$, *and*
if $d^2+4ac>0$, *then*

$$\omega_2=\frac{d-\sqrt{d^2+4ac}}{2},\qquad \omega_3=\frac{d+\sqrt{d^2+4ac}}{2},\qquad \varepsilon_2=\varepsilon_3=a,$$

if $d^2+4ac=0$, *then* $\omega_2=d/2, \omega_3=0, \varepsilon_2=a, \varepsilon_3=0$,
if $d^2+4ac<0$, *then* $\omega_2=\omega_3=0, \varepsilon_2=\varepsilon_3=0$.

4. *If* $b=0$, $c=0$, *then* $\varepsilon_1=\varepsilon_4=a$, $\omega_1=\omega_4=d$, $\varepsilon_2=\varepsilon_3=\omega_2=\omega_3=0$, $s_1=-e/a, s_2=-f/d$.

II. *If* $ad-bc=0$, *then at least one of the parameters* $\xi_1,\xi_2,\rho_1,\rho_2$ *is different from zero* (*there exists a Gaussian part in* (2.1) *and* (2.2)).

1. *If* $b=d=0$, $ac\neq 0$, *then* $(af-ec)/a\geqslant 0$ *and* $\rho_2=\sqrt{(af-ec)/a}$, $\rho_1=\xi_1=\xi_2=0$, $s_1=-e/a$, $s_2\in\mathbb{R}$, $\varepsilon_3=\varepsilon_4=\omega_3=\omega_4=0$,
if $ac>0$, *then* $\varepsilon_1=\varepsilon_2=a, \omega_1=\sqrt{ca}, \omega_2=-\sqrt{ca}$,
if $ac<0$, *then* $\varepsilon_1=\varepsilon_2=\omega_1=\omega_2=0$.

2. *If* $bc\neq 0$ *and* $db\neq a^2$, *then* $[b(ed-fb)]/(db-a^2)\geqslant 0$ *and* $\varepsilon_4=\omega_4=\xi_2=\rho_2=\varepsilon_3=\omega_3=0$,

$$\xi_1=\sqrt{\frac{b(ed-fb)}{bd-a^2}},\qquad \rho_1=-\frac{a}{b}\sqrt{\frac{b(ed-fb)}{bd-a^2}},$$

$$s_1\in\mathbb{R},\qquad s_2=\frac{1}{b}\left(\frac{b(ed-fb)}{bd-a^2}-as_1-e\right),$$

and if there exist j non-zero different roots $(\alpha_i,\ i=0,\dots,j,\ j\leqslant 2)$ *of the equation* $\alpha^2-2\alpha a+a^2-bd=0$, *then* $\varepsilon_i=\alpha_i$, $\omega_i=(\alpha_i^2-a\alpha_i)/b$ *for* $i=0,\dots,j, j\leqslant 2$, *and* $\varepsilon_i=\omega_i=0$ *for* $i=j+1,\dots,2$.

3. *If* $bc \neq 0$ *and* $db = a^2$, *then* $ed - fb = 0$ *and* $\varepsilon_4 = \omega_4 = \xi_2 = \rho_2 = \varepsilon_3 = \omega_3 = 0$,

$$\xi_1 \in \mathbb{R}_+, \quad \rho_1 = -\frac{a}{b}\xi_1, \quad s_1 \in \mathbb{R}, \quad s_2 = \frac{1}{b}\left(\frac{b(ed - fb)}{bd - a^2} - as_1 - e\right),$$

and if there exist j *non-zero different roots* $(\alpha_i,\ i = 0, \ldots, j,\ j \leqslant 2)$ *of the equation* $\alpha^2 - 2\alpha a + a^2 - bd = 0$, *then* $\varepsilon_i = \alpha_i$, $\omega_i = (\alpha_i^2 - a\alpha_i)/b$ *for* $i = 0, \ldots, j,\ j \leqslant 2$, *and* $\varepsilon_i = \omega_i = 0$ *for* $i = j + 1, \ldots, 2$.

4. *If* $a = c = 0$, $bd \neq 0$, *then* $(ed - fb)/d \geqslant 0$ *and* $\xi_1 = \sqrt{(ed - fb)/d}$, $\rho_1 = \rho_2 = \xi_2 = 0$, $s_1 \in \mathbb{R}$, $s_2 = -f/d$, $\varepsilon_3 = \varepsilon_4 = \omega_3 = \omega_4 = 0$,
if $bd > 0$, *then* $\varepsilon_1 = \sqrt{bd}$, $\varepsilon_2 = -\sqrt{bd}$, $\omega_1 = \omega_2 = d$,
if $bd < 0$, *then* $\varepsilon_1 = \varepsilon_2 = \omega_1 = \omega_2 = 0$.

5. *If* $a = b = 0$, *then* $e \geqslant 0$ *and* $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = 0$, $\omega_1 = \omega_2 = \omega_3 = 0$, $\xi_2 = 0$,
if $cd \neq 0$, *then* $\xi_1 = \sqrt{e}$, $\omega_4 = d$,

$$\rho_1 = -\frac{c}{d}\sqrt{e}, \quad \rho_2 = 0, \quad s_1 \in \mathbb{R}, \quad s_2 = \frac{1}{d}\left(\frac{ec^2}{d^2} - cs_1 - f\right),$$

if $c = 0$, $d \neq 0$, *then* $\xi_1 = \sqrt{e}$, $\omega_4 = d$, $s_1 \in \mathbb{R}$, $s_2 = -f/d$, $\rho_1 = \rho_2 = 0$,
if $c \neq 0$, $d = 0$ *and* $e = 0$, *then* $\omega_4 = \rho_1 = \xi_1 = 0$, $\rho_2 = \sqrt{cs_1 + f}$, $s_1 \geqslant -f/c$, $s_2 \in \mathbb{R}$.

6. *If* $c = d = 0$, *then* $f \geqslant 0$ *and* $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \omega_1 = \omega_2 = \omega_3 = \omega_4 = 0$, $\rho_2 = 0$,
if $ab \neq 0$, *then* $\rho_1 = -\sqrt{f}$, $\varepsilon_4 = a$,

$$\xi_1 = \frac{b}{a}\sqrt{f}, \quad \xi_2 = 0, \quad s_1 = \frac{1}{a}\left(\frac{fb^2}{a^2} - bs_2 - e\right), \quad s_2 \in \mathbb{R},$$

if $a \neq 0$, $b = 0$, *then* $\rho_1 = \sqrt{f}$, $\varepsilon_4 = a$, $s_1 = -e/a$, $s_2 \in \mathbb{R}$, $\xi_1 = \xi_2 = 0$,
if $a = 0$, $b \neq 0$ *and* $f = 0$, *then* $\varepsilon_4 = \xi_1 = \rho_1 = 0$, $s_1 \in \mathbb{R}$, $\xi_2 = \sqrt{bs_2 + e}$, $s_2 \geqslant -e/b$.

7. *If* $a = b = c = d = 0$, *then* $\xi_1 = \gamma_1$, $\rho_1 = \gamma_2$, $\xi_2 = \sqrt{e - \gamma_1^2}$, $\rho_2 = \sqrt{f - \gamma_2^2}$, $0 < \gamma_1 < \sqrt{e}$, $0 < \gamma_2 < \sqrt{f}$.

3. AUXILIARY FACTS

PROPOSITION 3.1. *If* $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy the conditions given in Theorem* 2.1, *then* $\{\theta \colon \mathbb{E}\exp(\langle \mathbf{X}, \theta\rangle) < \infty\} = \mathbb{R}^2$.

P r o o f. By the theory of differential equations the general solution of (1.7) and (1.8) has the following form (see Appendix 6.1):

$$k(\theta) = \sum_{\psi \in \Psi} Q_\psi(\theta) \exp(\langle \psi, \theta \rangle), \tag{3.1}$$

where $\Psi \subset \mathbb{C}^2$ is a finite set and $Q_\psi(\theta)$ are polynomials in θ, i.e. (3.1) is the cumulant transform of $\mathbf{X} = (X_1, X_2)$ at least in a neighborhood of the origin. For (θ_1, θ_2) arbitrarily chosen in $\mathbb{R}^2$ there exist a, b such that the point (θ_1, θ_2) belongs to the line (at, bt), $t \in \mathbb{R}$. Let us consider a random variable $U(a, b) = aX_1 + bX_2$. Its Laplace transform $M_U^{a,b}$ is well defined in the neighborhood of the origin. We can extend its domain to the real line, since $M_U^{a,b}(\theta) = e^{k(at,bt)}$, k given in (3.1), is an analytic function on $\mathbb{R}$. Thus $\mathbb{E}\exp(\theta_1 X_1 + \theta_2 X_2) = M_U^{a,b}(t_0) < \infty$ for t_0 defined by $(\theta_1, \theta_2) = (at_0, bt_0)$. ∎

PROPOSITION 3.2. *If the cumulant generating function k satisfies* (1.7) *and* (1.8) *for all $(\theta_1, \theta_2) \in \mathbb{R}^2$, then*

$$(3.2)\quad k(\theta_1,\theta_2) = A_0 + \sum_{1\leqslant i\leqslant j\leqslant 2} A_{ij}\theta_i\theta_j + \sum_{i=1}^{2} S_i\theta_i + \sum_{i,j=1}^{2} B_{ij}\exp(\lambda_{ij}\theta_i) + \sum_{i=1}^{3} D_i\exp(\alpha_i\theta_1 + \beta_i\theta_2),$$

where A_{ij}, S_i, B_{ij}, D_i, λ_{ij}, α_i, β_j are some real constants.

Proof. Suppose first that $b \neq 0$ (the case of $b = 0$ will be treated separately). Then from (1.7) and (1.8) (see Appendix 6.1) we have

$$(3.3)\qquad \frac{\partial^4 k}{\partial\theta_1^4} - 2a\frac{\partial^3 k}{\partial\theta_1^3} - (bd - a^2)\frac{\partial^2 k}{\partial\theta_1^2} + b(ad - bc)\frac{\partial k}{\partial\theta_1} + b(ed - fb) = 0.$$

It leads to the characteristic equation:

$$(3.4)\qquad \lambda^4 - 2a\lambda^3 - (bd - a^2)\lambda^2 + b(ad - bc)\lambda = 0.$$

Let us discuss possible solutions of (3.4) knowing that k is the cumulant transform of a probability measure. We consider only the cases that provide non-polynomial real functions or real polynomials of order greater than 2.

1. Complex solutions of (3.4). Suppose that there exists a complex solution $\lambda_1 = \lambda_{11} + i\lambda_{12}$. Thus k is of the form

$$(3.5)\quad k(\theta_1,\theta_2) = A_0(\theta_2) + \exp(\lambda_{11}\theta_1)\big(A_1(\theta_2)\cos(\lambda_{12}\theta_1) + A_2(\theta_2)\sin(\lambda_{12}\theta_1)\big) + A_3(\theta_2)\exp(\lambda_2\theta_1),$$

where $A_i(\cdot)$, $i = 0, 1, 2, 3$, are complex functions of θ_2. Let us fix θ_2. Then $k(\theta) = k(\theta, \theta_2)$ is (up to an additive constant) the cumulant transform of a univariate probability measure. Consider the corresponding characteristic function:

$$\phi(t) = A\exp\big\{A_0(\theta_2) + A_3(\theta_2)\exp(\lambda_2 it) + \exp(\lambda_{11}it)\big(A_1(\theta_2)\cos(\lambda_{12}it) + A_2(\theta_2)\sin(\lambda_{12}it)\big)\big\},$$

where $A > 0$ is a normalizing constant. We can rewrite the formula as follows:

$$\phi(t) = A\exp\Bigg\{A_0(\theta_2) + A_3(\theta_2)\exp(\lambda_2 it) + \exp(\lambda_{11}it)$$
$$\times\bigg(A_1(\theta_2)\frac{\exp(-\lambda_{12}t)+\exp(\lambda_{12}t)}{2} + A_2(\theta_2)\frac{\exp(-\lambda_{12}t)-\exp(\lambda_{12}t)}{2}(-i)\bigg)\Bigg\}.$$

The absolute value of $\phi(t)$ is

$$|\phi(t)| = \alpha\bigg|\exp\bigg\{A_1(\theta_2)\cos(\lambda_{11}t)\bigg(\frac{\exp(-\lambda_{12}t)+\exp(\lambda_{12}t)}{2}\bigg)$$
$$+ A_2(\theta_2)\sin(\lambda_{11}t)\bigg(\frac{\exp(-\lambda_{12}t)-\exp(\lambda_{12}t)}{2}\bigg)\bigg\}\bigg|,$$

where $\alpha > 0$. Note that $|\phi(t)|$ is unbounded if $A_1(\theta_2)$ or $A_2(\theta_2)$ are different from zero. By Proposition 3.2, θ_2 can be chosen arbitrarily in $\mathbb{R}$, so $A_1 \equiv A_2 \equiv 0$.

2. Multiple roots of (3.4) different from zero. Suppose that there exists double root (necessarily real) at $\lambda_1 \neq 0$. The case of triple root can be treated analogously. Here we have

$$k(\theta_1, \theta_2) = A_0(\theta_2) + A_1(\theta_2)\theta_1\exp(\lambda_1\theta_1) + A_2(\theta_2)\exp(\lambda_1\theta_1)$$
$$+ A_3(\theta_2)\exp(\lambda_2\theta_1),$$

where $A_i(\cdot)$, $i = 0, 1, 2, 3$, are complex functions of θ_2. Our aim is to show that $A_1 \equiv 0$. As in the previous case, we fix θ_2 and consider the univariate characteristic function

$$\phi(t) = A\exp\big[A_0(\theta_2) + A_1(\theta_2)\big(it\cos(\lambda_1 t) - t\sin(\lambda_1 t)\big)$$
$$+ A_2(\theta_2)\exp(\lambda_1 it) + A_3(\theta_2)\exp(\lambda_2 it)\big].$$

We examine whether $|\phi(\cdot)|$ is bounded. There is a positive constant M such that

$$|\phi(t)| = M\exp\big[A_1(\theta_2)\big(-t\sin(\lambda_1 t)\big)\big].$$

Thus $A_1 \equiv 0$, otherwise $|\phi(\cdot)|$ is unbounded.

3. Triple root of (3.4) at zero. Suppose that there exists triple root at zero. In this case

$$k(\theta_1, \theta_2) = A_0(\theta_2) + A_1(\theta_2)\theta_1 + A_2(\theta_2)\theta_1^2 + A_3(\theta_2)\exp(\lambda_1\theta_1) + C(\theta_2)\theta_1^3,$$

where $C(\cdot)$, $A_i(\cdot)$, $i = 0, 1, 2, 3$, are complex functions of θ_2. We are going to show that $C \equiv 0$. By fixing θ_2 we obtain the univariate Laplace transform:

$$L(\theta) = A\exp\big[A_1(\theta_2)\theta + A_2(\theta_2)\theta^2 + C(\theta_2)\theta^3 + A_3(\theta_2)\big(\exp(\gamma_1\theta) - 1\big)\big],$$

where $A > 0$ is a normalizing constant, and the characteristic function has the form

(3.6)
$$\phi(t) = A\exp\Big[A_1(\theta_2)it + A_2(\theta_2)(it)^2 + C(\theta_2)(it)^3 + A_3(\theta_2)\Big(\exp\big(\gamma_1(it)\big) - 1\Big)\Big].$$

Now we use the following extension of the Marcinkiewicz theorem (Theorem 2 in [12]).

THEOREM 3.1. *Let $P_m(t) = \sum_{v=0}^{m} c_v t^v$ be a polynomial of degree m. The function*

$$f(t) = \exp[\lambda_1(e^{it} - 1) + \lambda_2(e^{-it} - 1) + P_m(t)] \tag{3.7}$$

is a characteristic function if and only if $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$, $m \leqslant 2$ and $P_2(t) = a_1(it) - a_2 t^2$, where a_1 and a_2 are real and $a_2 \geqslant 0$.

If we put $\lambda_2 = 0$ in (3.7), Theorem 3.1 implies $C(\theta_2) = 0$ and $A_4(\theta_2) \geqslant 0$. As a consequence of the discussion of the cases 1–3, k has the following form:

$$k(\theta_1, \theta_2) = A_0(\theta_2) + A_1(\theta_2)\theta_1 + A_2(\theta_2)\theta_1^2 + \sum_{i=1}^{3} A_{3i}(\theta_2)\exp(\lambda_i\theta_1). \tag{3.8}$$

As a general solution for $b > 0$, applying (3.8) to (1.7) and (1.8), we obtain k given by (3.2).

Now, what is left to do is to analyze the case of $b = 0$. In this case the general solution of (1.7) and (1.8) is

$$k(\theta_1, \theta_2) = A_0(\theta_2) + A_1(\theta_2)\theta_1 + A_2(\theta_2)\exp(\theta_1 a) + A_3(\theta_2)\theta_1^2,$$

where $A_i(\cdot)$, $i = 0, 1, 2, 3$, satisfy

$$\begin{aligned}
A_0(\theta_2) &= A_{00} + A_{01}\theta_2 + A_{02}\theta_2^2 + A_{03}\exp(\delta_0\theta_2),\\
A_1(\theta_2) &= A_{10} + A_{11}\theta_2,\\
A_2(\theta_2) &= A_{20} + A_{23}\exp(\delta_1\theta_2),\\
A_3(\theta_2) &= A_{30},
\end{aligned}$$

and A_{ij}, $i, j = 0, 1, 2, 3$, δ_l, $l = 0, 1$, are some real constants. Thus k is of the form given by (3.2) also in this case. ∎

PROPOSITION 3.3. *If k given in (3.2) is the cumulant transform of a probability measure, then*

- $A_{11}, A_{22}, 4A_{11}A_{22} - A_{12}^2 \geqslant 0$;
- $\sum_{1\leqslant i\leqslant j\leqslant 2} A_{ij}\theta_i\theta_j$ *provides the Gaussian parts in* (3.2).

P r o o f. Let us define functions g and h:

$$g(\theta_1,\theta_2)=\sum_{1\leqslant i\leqslant j\leqslant 2}A_{ij}\theta_i\theta_j,\qquad h(\theta_1,\theta_2)=k(\theta_1,\theta_2)-g(\theta_1,\theta_2).$$

Since k is the cumulant transform of a probability measure, $\varphi(t)=e^{k(it,0)}$ is the characteristic function of its marginal distribution. Hence $|\varphi|$ is bounded on $\mathbb{R}$.

Suppose that $A_{11}<0$. Then $g(it,0)=-A_{11}t^2$ and the function

$$|\varphi(t)|=|\exp(-A_{11}t^2)|\big|\exp\big(h(it,0)\big)\big|$$

is unbounded on $\mathbb{R}$. This contradicts the fact that φ is a characteristic function, and hence $A_{11}\geqslant 0$. Analogously one can prove that $A_{22}\geqslant 0$.

Let us show that $4A_{11}A_{22}-A_{12}^2\geqslant 0$. Suppose that $A_{11}>0$. Then

$$g(\theta_1,\theta_2)=A_{11}\left(\theta_1+\frac{A_{12}}{2A_{11}}\theta_2\right)^2+\left(A_{22}-\frac{A_{12}^2}{4A_{11}}\right)\theta_2^2.$$

We define the characteristic function

$$\psi(t)=\exp\left\{k\left(-\frac{A_{12}}{2A_{11}}it,it\right)\right\}.$$

Then

$$g\left(-\frac{A_{12}}{2A_{11}}it,it\right)=-\left(A_{22}-\frac{A_{12}^2}{4A_{11}}\right)t^2;$$

hence ψ is unbounded on $\mathbb{R}$ if $4A_{11}A_{22}-A_{12}^2<0$.

If $A_{11}=A_{22}=0$, we have to show that $A_{12}=0$. In such a case we define a characteristic function $\eta(t)=\exp\left[k(it,it)\right]$. η is bounded on $\mathbb{R}$ iff $A_{12}=0$. ■

4. PROOF OF THEOREM 2.1

Now our aim is to show that the polynomials in (3.2) satisfy the conditions given in Theorem 2.1. Using the formula for k given by (3.2), it suffices to prove that coefficients B_{ij} and D_i are nonnegative.

Applying (3.2) to (1.7) and (1.8), we obtain a system of relations between coefficients: A_{ij}, S_i, B_{ij}, D_i, λ_{ij}, α_i, β_j (see Appendix 6.2).

In the proof we discuss various cases based on relations between coefficients $a,\ldots,d$.

• $ad-bc\neq 0, bc\neq 0$. From (6.11)–(6.14) in Appendix 6.2 we see that $A_{12}=A_{11}=A_{22}=0$. By (6.19)–(6.22) we have $B_{11}=B_{12}=B_{21}=B_{22}=0$. It leads to k of the form

$$k(\theta_1,\theta_2)=A_0+\sum_{i=1}^{2}S_i\theta_i+\sum_{i=1}^{3}D_i\exp\left(\alpha_i\theta_1+\frac{\alpha_i^2-a\alpha_i}{b}\theta_2\right), \tag{4.1}$$

where $\alpha_i \in \mathbb{R}$, $i = 1, 2, 3$. With no loss of generality we can assume that $\alpha_1 < \alpha_2 < \alpha_3$.

Suppose that $\alpha_1 > 0$. An argument of convexity of k implies that $D_3 \geqslant 0$. Indeed, let us consider

$$\frac{\partial^2 k}{\partial\theta_1^2} = \sum_{i=1}^{3} D_i \alpha_i^2 \exp\left(\alpha_i\theta_1 + \frac{\alpha_i^2 - a\alpha_i}{b}\theta_2\right). \tag{4.2}$$

Since $\partial^2 k/\partial\theta_1^2 \geqslant 0$ for all $\theta_1 \in \mathbb{R}$, we have $D_3 \geqslant 0$.

We now proceed to show that $D_1, D_2 \geqslant 0$. Consider the Laplace transform

$$\tilde{L}(\theta_1, \theta_2) = A \sum_{j=0}^{\infty} \frac{1}{j!}\left[\sum_{i=1}^{3} D_i \exp\left(\alpha_i\theta_1 + \frac{\alpha_i^2 - a\alpha_i}{b}\theta_2\right)\right]^j,$$

where $A > 0$, that is, $\tilde{L}$ is the Laplace transform of a discrete probability measure in $\mathbb{R}^2$. Note that D_1 as the unique coefficient at $\exp\big(\alpha_1\theta_1 + [(\alpha_1^2 - a\alpha_1)/b]\theta_2\big)$ is nonnegative, and AD_1 is the mass of the point $\big(\alpha_1, (\alpha_1^2 - a\alpha_1)/b\big)$.

To deal with D_2 it is sufficient to show that there is no $l > 1$ such that

$$\alpha_2 = l\alpha_1, \tag{4.3}$$

$$\frac{\alpha_2^2 - a\alpha_2}{b} = l\frac{\alpha_1^2 - a\alpha_1}{b}. \tag{4.4}$$

Since $\alpha_1 \geqslant 0$, (4.3) and (4.4) imply $l = 1$, which contradicts the assumption that $\alpha_1 < \alpha_2$. Then D_2 is the unique coefficient at $\exp\big(\alpha_2\theta_1 + (\alpha_2^2 - a\alpha_2)/b\big)$, and thus $D_2 \geqslant 0$.

The same conclusions can be drawn for the case $\alpha_3 \leqslant 0$, if we consider the Laplace transform $\tilde{L}(-\theta_1, \theta_2)$.

Now we are left with $\alpha_1 < 0 < \alpha_2 < \alpha_3$. By (4.2), the argument of convexity assures that $D_1 \geqslant 0$ and $D_3 \geqslant 0$. Since there are no $m, r > 1$ such that

$$\alpha_2 = r\alpha_1 + m\alpha_2,$$
$$\alpha_2^2 = r\alpha_1^2 + m\alpha_2^2,$$

D_2 is the unique coefficient at $\exp\big(\alpha_2\theta_1 + [(\alpha_2^2 - a\alpha_2)/b]\theta_2\big)$, and thus $D_2 \geqslant 0$.

- $ad - bc = 0$, $bc \neq 0$. Since

$$A_{22} = \frac{a^2}{b^2}A_{11}, \quad A_{12} = -2\frac{a}{b}A_{11} \quad \text{and} \quad 4A_{22}A_{11} - A_{12}^2 = 0,$$

the assumption leads to

$$k(\theta_1,\theta_2) = A_0 + A_{11}\Big(\theta_1 - \frac{a}{b}\theta_2\Big)^2 + D_1 \exp\left(\alpha_1\theta_1 + \frac{\alpha_1^2 - a\alpha_1}{b}\theta_2\right) + D_2 \exp\left(\alpha_2\theta_1 + \frac{\alpha_2^2 - a\alpha_2}{b}\theta_2\right) + S_1\theta_1 + S_2\theta_2.$$

Consider θ_1 such that

$$\alpha_1\theta_1 + \frac{\alpha_1^2 - a\alpha_1}{b}\theta_2 = 0.$$

Then we can build a cumulant function $k(\theta) = k\big(-[(\alpha_1 - a)/b]\theta, \theta\big)$:

$$k(\theta) = A_0 + D_1 + A_{11}\Big(\frac{\alpha_1}{b}\theta\Big)^2 + D_2 \exp\left(\frac{\alpha_2^2 - \alpha_2\alpha_1}{b}\theta\right).$$

Applying Theorem 3.1 we get $D_2 \geqslant 0$. Analogously one can prove that $D_1 \geqslant 0$. As a result we obtain

$$k(\theta_1,\theta_2) = \frac{b(ed - fb)}{2(bd - a^2)}\Big(\theta_1 - \frac{a}{b}\theta_2\Big)^2 + D_1 \exp\left(\alpha_1\theta_1 + \frac{\alpha_1^2 - a\alpha_1}{b}\theta_2\right) + D_2 \exp\left(\alpha_2\theta_1 + \frac{\alpha_2^2 - a\alpha_2}{b}\theta_2\right) + S_1\theta_1 + \frac{1}{b}\left(\frac{b(ed - fb)}{bd - a^2} - aS_1 - e\right)\theta_2.$$

- $ad - bc \neq 0$, $b \neq 0$, $c = 0$. If there exists $D_i \neq 0$ for $i = 1, 2, 3$, from (6.23) and (6.24) we obtain:

$$\beta_i = 0 \text{ and } \alpha_i = a \quad \text{or} \quad \beta_i = d \text{ and } \alpha_i^2 - a\alpha_i - db = 0.$$

We analyze separately particular cases.

- $a > 0$. Let $a^2 + 4bd > 0$. In this case k has the form

$$k(\theta_1,\theta_2) = B_{11}\big(\exp(a\theta_1) - 1\big) + D_1\big(\exp(a\theta_1 + d\theta_2) - 1\big) + D_2\left(\exp\left(\frac{a - \sqrt{a^2 + 4bd}}{2}\theta_1 + d\theta_2\right) - 1\right) + D_3\left(\exp\left(\frac{a + \sqrt{a^2 + 4bd}}{2}\theta_1 + d\theta_2\right) - 1\right) + \frac{bf - ed}{ad}\theta_1 - \frac{f}{d}\theta_2. \tag{4.5}$$

We want to show that $B_{11} \geqslant 0$, $D_1 \geqslant 0$, $D_2 \geqslant 0$, $D_3 \geqslant 0$. Since $ad \neq 0$, putting (4.5) into (1.7) yields $D_1 = 0$.

Consider $\partial^2 k/\partial\theta_1^2$. Since θ_2 can be arbitrarily chosen in $\mathbb{R}$, one can easily verify that $B_{11} \geqslant 0$. The argument of convexity also assures that $D_3 \geqslant 0$. To deal with D_2 let us write an expansion of the Laplace transform:

$$(4.6)\quad L(\theta_1,\theta_2) = \sum_{j=0}^{\infty} \frac{1}{j!}\Bigg[B_{11}\exp(a\theta_1) + D_2\exp\left(\frac{a-\sqrt{a^2+4bd}}{2}\theta_1 + d\theta_2\right) + D_3\exp\left(\frac{a+\sqrt{a^2+4bd}}{2}\theta_1 + d\theta_2\right)\Bigg]^j.$$

Note that D_2 is the unique coefficient at $\exp\big([(a-\sqrt{a^2+4bd})/2]\theta_1 + d\theta_2\big)$ in (4.6), and hence $D_2 \geqslant 0$.

If $a^2+4bd = 0$, then

$$k(\theta_1,\theta_2) = = B_{11}\big(\exp(a\theta_1)-1\big) + D_2\left(\exp\left(\frac{a}{2}\theta_1 + d\theta_2\right) - 1\right) + \frac{bf-ed}{ad}\theta_1 - \frac{f}{d}\theta_2.$$

Convexity of k implies that $B_{11} \geqslant 0$ and $D_2 \geqslant 0$.

If $a^2+4bd < 0$, then

$$k(\theta_1,\theta_2) = B_{11}\big(\exp(a\theta_1)-1\big) + \frac{bf-ed}{ad}\theta_1 - \frac{f}{d}\theta_2.$$

- $a<0$. One can apply similar arguments to those in the previous case to show that $B_{11}, D_2, D_3 \geqslant 0$.
- $ad-bc \neq 0$, $b=0$, $c\neq 0$. In this case the analysis is similar and we skip it.
- $ad-bc = 0$, $b=d=0$, $a\neq 0$, $c\neq 0$. Suppose that $ac>0$. Then

$$(4.7)\quad k(\theta_1,\theta_2) = A_0 + S_1\theta_1 + S_2\theta_2 + A_{12}\theta_1\theta_2 + A_{22}\theta_2^2 + A_3\theta_2^3 + D_1\exp(\theta_1 a + \theta_2\sqrt{ca}) + D_2\exp(\theta_1 a - \theta_2\sqrt{ca}).$$

Without any loss of generality we assume that $S_1 = S_2 = 0$. Let us fix θ_1 and consider the Laplace transform

$$L(\theta) = \exp\big(A_0 + A_{12}\theta\theta_1 + A_{22}\theta^2 + A_3\theta^3 + D_1\exp(\theta_1 a + \theta\sqrt{ca}) + D_2\exp(\theta_1 a - \theta\sqrt{ca})\big).$$

Theorem 3.1 assures that $A_3 = 0$ and D_1, $D_2 \geqslant 0$. From Proposition 3.3 we have $A_{12} = 0$, and k is of the form

$$k(\theta_1,\theta_2) = \frac{af-ec}{2a}\theta_2^2 + D_1[\exp(\theta_1 a + \theta_2\sqrt{ca}) - 1] + D_2[\exp(\theta_1 a - \sqrt{ca}\theta_2) - 1] - \frac{e}{a}\theta_1 + s_2\theta_2,$$

where D_1, $D_2 \geqslant 0$.

If $ac < 0$, then

$$k(\theta_1, \theta_2) = \frac{af - ec}{2a}\theta_2^2 - \frac{e}{a}\theta_1 + s_2\theta_2.$$

5. PROBLEMS IN HIGHER DIMENSIONS

One can try to generalize the result given above to $\mathbb{R}^n$ for $n > 2$. Analogously to the 2-dimensional case, by considering a characteristic equation we obtain the general formula of the Laplace transform L:

$$\begin{aligned}(5.1)\quad \log L(\theta_1, \ldots, \theta_n) = &\sum_{k_1,\ldots,k_n=0}^{n} \sum_{i_1,\ldots,i_n=0}^{n} [A_{i_1,\ldots,i_n,k_1,\ldots,k_n}\theta_{i_1}^{k_1}\ldots\theta_{i_n}^{k_n} \\ &\times \exp(\lambda_{1,i_1,\ldots,i_n}\theta_{i_1} + \ldots + \lambda_{n,i_1,\ldots,i_n}\theta_{i_n})] \\ + \sum_{k_1,\ldots,k_n=0}^{n} \sum_{i_1,\ldots,i_n=0}^{n} \prod_{l=0}^{n} &\big[\big(B_{l,1,i_1,i_2,\ldots,i_n}\cos(\delta_{l,i_1,i_2,\ldots,i_n}\theta_{i_1}) + B_{l,2,i_1,i_2,\ldots,i_n} \\ \times \sin(\gamma_{l,i_1,i_2,\ldots,i_n}\theta_{i_2})\big) &\cdot \exp(\rho_{l,1,i_1,\ldots,i_n}\theta_{i_1} + \ldots + \rho_{l,n,i_1,\ldots,i_n}\theta_{i_n})\big]\theta_{i_1}^{k_1}\ldots\theta_{i_n}^{k_n},\end{aligned}$$

where $A_{i_1,\ldots,i_n,k_1,\ldots,k_n}$, $B_{l,j,i_1,\ldots,i_n}$, $\lambda_{m,i_1,\ldots,i_n}$, $\delta_{l,i_1,\ldots,i_n}$, $\gamma_{l,i_1,\ldots,i_n}$, $\rho_{l,m,i_1,\ldots,i_n} \in \mathbb{R}$ for $0 \leqslant i_1, \ldots, i_n, k_1, \ldots, k_n \leqslant n$, $1 \leqslant l, m \leqslant n$, $1 \leqslant j \leqslant 2$. There arises a problem while one tries to eliminate non-probabilistic solutions. Namely, in order to obtain the Laplace transforms of probability measures (from (5.1)), it is necessary to put some restrictions on the coefficients.

Even in $\mathbb{R}^3$ the problem of identifying a measure from the cumulant transform seems to be difficult as the example below indicates.

EXAMPLE 5.1. Let the diagonal of VF be of the form

$$(5.2)\qquad \operatorname{diag} V(m_1, m_2, m_3) = (m_1, m_1, m_1 + m_2).$$

From (5.2), by solving the system of partial differential equations, we obtain the cumulant transform

$$\begin{aligned}(5.3)\quad &k(\theta_1, \theta_2, \theta_3) = \\ &= \exp(\theta_1 - \theta_3)(A_1 + A_3\theta_3 - 2A_3\theta_2) + \exp(\theta_1 + \theta_3)(A_2 + A_4\theta_3 - 2A_4\theta_2) \\ &\quad + 1 - A_2 - A_3 + B_2\theta_3 + B_3\theta_3^2 + B_4\theta_3^3 + 2B_3\theta_3 + 6B_4\theta_2\theta_3.\end{aligned}$$

But we do not know the sign of each of A_i, B_i, $i = 1, 2, 3, 4$ (which is necessary to identify the measure). To fix the signs we need to cope with a combination of exponential functions and polynomials of order greater than two. To this end, it would be required a kind of generalization of the Marcinkiewicz theorem which goes beyond Theorem 3.1. Unfortunately, such a generalization is not known, at least to us, and consequently we are not able to give a complete description corresponding to the cumulant generating function of the form (5.3).

On the other hand, there exists a natural exponential family with VF of the form (5.2). If we put $A_3 = A_4 = B_4 = 0$, and we assume that $A_1 \geqslant 0$, $A_2 \geqslant 0$, $B_3 \geqslant 0$, then (5.3) is the cumulant transform of a probability measure, namely, the combination of Poisson and Gaussian distributions. The NEF generated by this measure is an example of the family which does not belong to the class with linear matrix VF (described in [7]). It can be easily checked if we take $\partial^2 k/\partial\theta_1\theta_3$ which is not an affine function of $\partial k/\partial\theta_1$, $\partial k/\partial\theta_2$, $\partial k/\partial\theta_3$.

Acknowledgments. I wish to thank J. Wesołowski for many conversations and useful suggestions. I would like also to thank G. Letac for insightful comments.

6. APPENDIX

6.1. Solution of (1.7) and (1.8). Since we deal with a non-homogeneous system of partial differential equations, the solution does not always exist (it depends on the relations between parameters $a, \ldots, f$). Our aim is to provide a general form of k satisfying (1.7) and (1.8). Let us consider separately two cases.

• $b \neq 0$. In such a case, as a consequence of (1.7) and (1.8), we obtain

$$\frac{\partial^4 k}{\partial\theta_1^4} - 2a\frac{\partial^3 k}{\partial\theta_1^3} - (bd - a^2)\frac{\partial^2 k}{\partial\theta_1^2} + b(ad - bc)\frac{\partial k}{\partial\theta_1} + b(ed - fb) = 0. \tag{6.1}$$

Hence we want to solve the fourth order ordinary differential equation (ODE). The general theory of ODEs with constant coefficients assures that the solution can be written in the form

$$k(\theta_1, \theta_2) = \sum_{\lambda\in\Lambda_1} A_\lambda(\theta_1, \theta_2)\exp(\lambda\theta_1), \tag{6.2}$$

where $\Lambda_1 \subset \mathbb{C}$ is a set of the solutions of the characteristic equation

$$\lambda^4 - 2a\lambda^3 - (bd - a^2)\lambda^2 + b(ad - bc)\lambda = 0, \tag{6.3}$$

and

$$A_\lambda(\theta_1, \theta_2) = \sum_{i=0}^{3} a_{\lambda i}(\theta_2)\theta_1^i, \tag{6.4}$$

where $a_{\lambda i}(\cdot)$, $i = 0, 1, 2, 3$, $\lambda \in \Lambda$, are complex functions.

In order to find the explicit form of functions $a_{\lambda i}(\cdot)$ we apply (6.2) and (6.4) to (1.8). It is easy to verify that $a_{\lambda i}(\cdot)$ also satisfy ODEs with constant coefficients (of order at most two). Thus we can present complex functions $a_{\lambda i}(\cdot)$ as

$$a_{\lambda i}(\theta_2) = \sum_{\gamma\in\Upsilon_{\lambda i}} q_{\lambda i}(\theta_2)\exp(\gamma\theta_2), \tag{6.5}$$

where $\Upsilon_{\lambda i} \subset \mathbb{C}$ is a finite set that depends on λ and i ($\Upsilon_{\lambda i}$ is the set of all the solutions of a characteristic equation derived from the ODE satisfied by $a_{\lambda i}(\cdot)$), and $q_{\lambda i}(\cdot)$ are real polynomials of degree at most two.

It follows from the above that the general solution of (1.7) and (1.8) can be written in the form

$$k(\theta_1, \theta_2) = \sum_{\psi \in \Psi} Q_\psi(\theta) \exp(\langle \theta, \psi \rangle), \tag{6.6}$$

where $\Psi \subset \Lambda_1 \times \Lambda_2$ (coordinates are treated symmetrically). Here Λ_2 (analogously to Λ_1) stands for the set of all the solutions of the characteristic equation derived from the following fourth order ODE:

$$\frac{\partial^4 k}{\partial \theta_2^4} - 2d\frac{\partial^3 k}{\partial \theta_2^3} - (ac - d^2)\frac{\partial^2 k}{\partial \theta_2^2} + c(ad - bc)\frac{\partial k}{\partial \theta_2} + c(fa - ec) = 0. \tag{6.7}$$

• $b = 0$. The same conclusions can be drawn for $b = 0$. In such a case (1.7) leads to the following characteristic equation:

$$\lambda^2 - a\lambda = 0. \tag{6.8}$$

Hence the solution of (1.7) has the form given by (6.2), where $\Lambda_1 = \{0, a\}$. Therefore the general solution of (1.7) and (1.8) can be written as in (6.6).

6.2. Relations between coefficients of (3.2). We have

$$2A_{11} = aS_1 + bS_2 + e, \tag{6.9}$$
$$2A_{22} = cS_1 + dS_2 + f, \tag{6.10}$$
$$0 = 2A_{11}a + A_{12}b, \tag{6.11}$$
$$0 = 2A_{11}c + A_{12}d, \tag{6.12}$$
$$0 = A_{12}a + 2A_{22}b, \tag{6.13}$$
$$0 = A_{12}c + 2A_{22}d, \tag{6.14}$$
$$\lambda_{11}^2 B_{11} = aB_{11}\lambda_{11}, \tag{6.15}$$
$$\lambda_{12}^2 B_{12} = aB_{12}\lambda_{12}, \tag{6.16}$$
$$\lambda_{21}^2 B_{21} = dB_{21}\lambda_{21}, \tag{6.17}$$
$$\lambda_{22}^2 B_{22} = dB_{22}\lambda_{22}, \tag{6.18}$$
$$b\lambda_{21}B_{21} = 0, \tag{6.19}$$
$$b\lambda_{22}B_{22} = 0, \tag{6.20}$$
$$cB_{11}\lambda_{11} = 0, \tag{6.21}$$
$$cB_{12}\lambda_{12} = 0, \tag{6.22}$$
$$D_i\beta_i^2 = cD_i\alpha_i + dD_i\beta_i, \tag{6.23}$$
$$D_i\alpha_i^2 = aD_i\alpha_i + bD_i\beta_i, \quad i = 1, 2, 3. \tag{6.24}$$

6.3. Proof of Theorem 2.1 and the list of remaining cumulant transforms.

- $ad - bc \neq 0$, $b = 0$, $c = 0$. It leads to the cumulant function

$$k(\theta_1, \theta_2) = \frac{f}{d}\theta_2 - \frac{e}{a}\theta_1 + B_{11}[\exp(a\theta_1) - 1] + B_{21}[\exp(d\theta_2) - 1] \\ + D_1[\exp(\theta_1 a + \theta_2 d) - 1].$$

Convexity of k assures that B_{11}, B_{21} and D_1 are nonnegative.

- $ad - bc = 0$, $a = c = 0$, $b \neq 0$, $d \neq 0$. If $bd > 0$ and $(ed - fb)/d \geqslant 0$, then

$$k(\theta_1, \theta_2) = \frac{ed - fb}{2d}\theta_1^2 + D_1[\exp(\theta_1\sqrt{bd} + \theta_2 d) - 1] \\ + D_2[\exp(-\theta_1\sqrt{bd} + \theta_2 d) - 1] - \frac{f}{d}\theta_2 + s_1\theta_1,$$

where $D_1, D_2 \geqslant 0$.

If $bd < 0$ and $(ed - fb)/d \geqslant 0$, then

$$k(\theta_1, \theta_2) = \frac{ed - fb}{2d}\theta_1^2 - \frac{f}{d}\theta_2 + s_1\theta_1.$$

- $ad - bc = 0$, $b = c = a = 0$, $d \neq 0$. Then

$$k(\theta_1, \theta_2) = B_{21}\big(\exp(d\theta_2) - 1\big) + s_1\theta_1 - \frac{f}{d}\theta_2 + e\frac{\theta_1^2}{2},$$

where $B_{21} \geqslant 0$.

- $b = c = d = a = 0$. If $e \geqslant 0$ and $f \geqslant 0$, then

$$k(\theta_1, \theta_2) = \frac{e}{2}\theta_1^2 + \frac{f}{2}\theta_2^2 + S_1\theta_1 + S_2\theta_2 + A_{12}\theta_1\theta_2,$$

where $|A_{12}| \leqslant \sqrt{ef}$.

- $ad - bc = 0$, $a = b = 0$, $cd \neq 0$. If $e \geqslant 0$, then

$$k(\theta_1, \theta_2) = \frac{e}{2}\left(\theta_1 - \frac{c}{d}\theta_2\right)^2 + B_{21}\exp(d\theta_2) + S_1\theta_1 + \left(\frac{c^2 e}{d^3} - \frac{c}{d}S_1 - \frac{f}{d}\right)\theta_2,$$

where $B_{21} \geqslant 0$.

- $ad - bc = 0$, $a = b = d = 0$, $c \neq 0$ and $e = 0$. If $cs_1 \geqslant -f$, then

$$k(\theta_1, \theta_2) = \frac{cs_1 + f}{2}\theta_2^2 + s_1\theta_1 + s_2\theta_2 + f.$$

- $ad - bc = 0$, $b = c = d = 0$, $a \neq 0$. If $f \geqslant 0$, then

$$k(\theta_1, \theta_2) = B_{11}\big(\exp(a\theta_1) - 1\big) - \frac{e}{a}\theta_1 + s_2\theta_2 + \frac{f}{2}\theta_2^2,$$

where $B_{11} \geqslant 0$.

- $ad - bc = 0$, $c = d = 0$, $ab \neq 0$. If $f \geqslant 0$, then

$$k(\theta_1, \theta_2) = \frac{f}{2}\left(\frac{b}{a}\theta_1 - \theta_2\right)^2 + B_{11}\exp(a\theta_1) + \left(f\frac{b^2}{a^3} - \frac{b}{a}s_2 - \frac{e}{a}\right)\theta_1 + s_2\theta_2 + e,$$

where $B_{11} \geqslant 0$.

- $ad - bc = 0$, $a = c = d = 0$, $b \neq 0$ and $f = 0$. If $bs_2 \geqslant -e$, then

$$k(\theta_1, \theta_2) = \frac{bs_2 + e}{2}\theta_1^2 + s_1\theta_1 + s_2\theta_2 + e.$$

REFERENCES

[1] S. K. Bar-Lev, *Methods of constructing characterizations by constancy of regression on the sample mean and related problems for NEF's*, Math. Methods Statist. 16 (2) (2007), pp. 96–109.

[2] S. K. Bar-Lev, D. Bshouty, P. Enis, G. Letac, Lu I-Li and D. Richards, *The diagonal multivariate natural exponential families and their classification*, J. Theoret. Probab. 7 (1994), pp. 883–929.

[3] M. Casalis, *Les familles exponentielles sur* $\mathbb{R}^2$ *de fonction-variance* $V(m) = am \otimes m + B(m) + C$, C. R. Acad. Sci. Paris 314 (1992), pp. 635–638.

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

[5] E. B. Fosam and D. N. Shanbhag, *An extended Laha–Lukacs characterization result based on a regression property*, J. Statist. Plann. Inference 63 (1997), pp. 173–186.

[6] F. S. Gordon, *Characterizations of populations using regression properties*, Ann. Statist. 1 (1973), pp. 114–126.

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

[8] C. Kokonendji and A. Masmoudi, *A characterization of Poisson–Gaussian families by generalized variance*, Bernoulli 12 (2) (2006), pp. 371–379.

[9] C. Kokonendji and V. Seshadri, *On the determinant of the second derivative of a Laplace transform*, Ann. Statist. 24 (1996), pp. 1813–1827.

[10] R. G. Laha and E. Lukacs, *On a problem connected with quadratic regression*, Biometrika 47 (1960), pp. 335–343.

[11] G. Letac, *Le problème de la classification des familles exponentielles naturelles de* $\mathbf{R}^d$ *ayant une fonction-variance quadratique*, in: *Probability Measures on Groups IX*, Lecture Notes in Math. No 1379 (1989).

[12] G. Letac, *Lectures on natural exponential families and their variance functions*, Monograph No 50, IMPA, Rio de Janeiro 1992.

[13] G. Letac and J. Wesołowski, *Laplace transforms which are negative powers of quadratic polynomials*, Trans. Amer. Math. Soc. 360 (12) (2008), pp. 6475–6496.

[14] E. Lukacs, *Some extensions of a theorem of Marcinkiewicz*, Pacific J. Math. 12 (1962), pp. 58–67.

Faculty of Mathematics and Information Science
Warsaw University of Technology
00-661 Warsaw, Poland
E-mail: Joanna.Matysiak@mini.pw.edu.pl

Received on 17.6.2008;
revised version on 24.6.2009