

THE $2d + 4$ SIMPLE QUADRATIC NATURAL EXPONENTIAL FAMILIES ON \mathbb{R}^d

BY M. CASALIS¹

Université Paul Sabatier

The present paper describes all the natural exponential families on \mathbb{R}^d whose variance function is of the form $V(m) = am \otimes m + B(m) + C$, with $m \otimes m(\theta) = \langle \theta, m \rangle m$ and B linear in m . There are $2d + 4$ types of such families, which are built from particular mixtures of families of Normal, Poisson, gamma, hyperbolic on \mathbb{R} and (negative-) multinomial distributions. The proof of this result relies mainly on techniques used in the elementary theory of Lie algebras.

1. Introduction. For an accurate presentation of the simple quadratic natural exponential families, let us introduce some notation.

Let E be a real vector space with finite dimension d , let E^* be its dual and let $E^* \times E \rightarrow \mathbb{R}: (\theta, x) \mapsto \langle \theta, x \rangle$ be the duality bracket. We denote by $\mathcal{L}_s(E^*, E)$ [respectively, $\mathcal{L}_s(E, E^*)$] the space of the symmetric linear operators from E^* to E (resp. from E to E^*), that is, the space of linear operators $V: E^* \rightarrow E$ such that for (α, β) in $(E^*)^2$, $\langle \alpha, V\beta \rangle = \langle \beta, V\alpha \rangle$ [resp. $\psi: E \rightarrow E^*$ such that for (u, v) in E^2 , $\langle \psi u, v \rangle = \langle \psi v, u \rangle$].

For a positive Radon measure on E , we note

$$\begin{aligned} L_\mu: E^* &\rightarrow]0, \infty[: \theta \mapsto \int \exp\langle \theta, x \rangle \mu(dx), \\ \Theta(\mu) &= \text{interior}\{\theta \in E^*; L_\mu(\theta) < +\infty\}, \\ k_\mu &= \log L_\mu \quad \text{on } \Theta(\mu). \end{aligned}$$

L_μ and k_μ are, respectively, the Laplace transform and the cumulant function of μ .

Let $\mathcal{M}(E)$ denote the set of μ such that μ is not concentrated on an affine hyperplane of the space and $\Theta(\mu)$ is not empty. Then, for μ in $\mathcal{M}(E)$, the set of probabilities $F = F(\mu) = \{P(\theta, \mu)(dx) = \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx); \theta \in \Theta(\mu)\}$ is called *the natural exponential family* (NEF) generated by μ . The measure μ is called a *basis* of F .

Since μ is in $\mathcal{M}(E)$, k_μ is strictly convex and real analytic on $\Theta(\mu)$, so that its first derivative $k'_\mu: \Theta(\mu) \rightarrow E$:

$$\theta \mapsto k'_\mu(\theta) = \int xP(\theta, \mu) dx$$

Received May 1994; revised November 1995.

¹Supported by NATO Grant 92-13-47.

AMS 1991 subject classifications. 62E10, 60E10.

Key words and phrases. Variance functions, Morris class.

defines a diffeomorphism from $\Theta(\mu)$ to its range M_F , called the *mean domain* of F . Let $\psi_\mu: M_F \rightarrow \Theta(\mu)$ be its inverse function, and for m in M_F , $P(m, F) = P(\psi_\mu(m), \mu)$. Now the covariance operator of $P(m, F)$ is denoted by $V_F(m)$. Clearly,

$$(1.1) \quad V_F(m) = k''_\mu(\psi_\mu(m)) = [\psi'_\mu(m)]^{-1} \in \mathcal{L}_s(E^*, E).$$

The function $V_F: M_F \rightarrow \mathcal{L}_s(E^*, E)$ is called the *variance function* of F . This function V_F plays an important role in the study of the NEF. Indeed, V_F characterizes the NEF F in the following sense: if F_1 and F_2 are two NEF whose variance functions coincide on a nonempty open set of $M_{F_1} \cap M_{F_2}$, then $F_1 = F_2$.

Several classifications of all the NEF whose variance functions have a given form have appeared in the literature during the past 15 years. For instance, the Morris class describes all the real NEF F such that $V_F(m)$ is a polynomial of degree less than or equal to 2 in the mean m [see Morris (1982)]. Other sets of NEF on \mathbb{R} have also been classified as the Mora class of polynomial variance functions with degree less than or equal to 3 [Mora (1986); see also Letac and Mora (1990)] or various subsets of the class of V_F of the form $P\Delta + Q\sqrt{\Delta}$, where P , Δ and Q are polynomials with degree less than or equal to 1, less than or equal to 2 and less than or equal to 2, respectively [Kokonendji (1993); Letac (1992)].

Several classifications have been realized in higher dimensions, which concern more precisely the extension of the Morris and Mora classes in \mathbb{R}^d . One paper on the subject is Bar-Lev, Bshouty, Enis, Letac, Li Lu and Richards (1994). However, a very different point of view is to consider the quadratic class. It can be defined as the set of NEF F such that

$$(1.2) \quad V_F(m) = A(m, m) + B(m) + C,$$

where $A: E \times E \rightarrow \mathcal{L}_s(E^*, E)$ is bilinear, $B: E \rightarrow \mathcal{L}_s(E^*, E)$ is linear, and C is a constant element of $\mathcal{L}_s(E^*, E)$.

Classifying all these variance functions is an ambitious project and only a few steps have already been performed: Letac (1989) has obtained the subclass $A = 0$ of the products of Normal and Poisson NEF; Casalis (1991) described the homogeneous quadratic case $B = 0$ and $C = 0$, generalizing the gamma NEF on \mathbb{R} [where $V_F(m) = m^2/p$] by the Wishart NEF on symmetric cones. The present paper is one additional step toward this aim; it classifies all the variance functions V_F of the form

$$(1.3) \quad V_F(m) = am \otimes m + B(m) + C,$$

with a in \mathbb{R} and $m \otimes m: \theta \mapsto \langle \theta, m \rangle m$ in $\mathcal{L}_s(E^*, E)$. Such variance functions (1.3) and the corresponding NEF are called *simple quadratic*.

It may seem rather restrictive to keep the quadratic part as “simple” (i.e., of the form “ $a m \otimes m$ ”). Furthermore, one may feel a bit frustrated by the relative simplicity of our results; indeed, all the simple quadratic NEF are basically obtained by combination of conditional distributions of the one-dimensional Morris class, which simply seems to indicate that to produce

really new and genuinely multidimensional distributions with quadratic variance (e.g., the Wishart distributions), nature requires a more involved $a m \otimes m$ quadratic part than $a m \otimes m$. Furthermore, it should be pointed out that, out of the Wishart NEF and their translates, or trivial products of them, we have yet no other examples of quadratic (and not simple) NEF. Recall also that the very classification of the Wishart distributions, that is, the homogeneous quadratic NEF, was a delicate process, involving the von Neumann–Wigner–Jordan classification of Euclidean Jordan algebras. Therefore, many new ideas are probably necessary to continue this program of classification of quadratic variances.

However, the simple quadratic NEF remain quite interesting. Here is a list of situations where they occur naturally.

1. They constitute an important tool in the determination of the so-called Mora class in \mathbb{R}^d . Indeed, Hassairi (1992, 1994) introduces some specific action of the linear group G of $\mathbb{R} \times E$ on the space $\mathcal{L}_s(E, E^*)$. Observing that this action transforms a simple quadratic variance function into a polynomial with degree less than or equal to 3, he defines the Mora class on \mathbb{R}^d as the set of the NEF obtained in this way and describes it entirely using the present list of simple quadratic NEF.

2. As for the real case, one easily checks that the uniformly minimum variance unbiased (UMVU) estimator of the variance $V_F(m)$ written in (1.3) is simply given by $n(n+a)^{-1}V_F(\bar{X}_n)$, where \bar{X}_n is the sample mean of n random variables X_1, \dots, X_n with common distribution $P(m, F)$. This result is not true for general quadratic NEF on \mathbb{R}^d [Casalis (1992a)] and it remains a conjecture even in one dimension that this fact characterizes the simple quadratic class [Letac (1992), page 61].

3. The computation of the UMVU estimator of the generalized variance $\det k'' \mu(\theta)$ is easy to get for the simple quadratic class [see Kokonendji and Seshadri (1996)].

4. Another property is related to the Bayesian theory and states the equality of two conjugate prior distributions families of a simple quadratic NEF. For a given NEF $F(\mu)$, let Π be the family of prior distributions on $\Theta(\mu)$:

$$\pi_{t, m_0}(d\theta) = C_{t, m_0} \exp t\{\langle \theta, m_0 \rangle - k_\mu(\theta)\} \mathbb{1}_{\Theta(\mu)}(\theta) d\theta,$$

where $t > 0$, $m_0 \in M_F$ and C_{t, m_0} a normalizing constant, as considered by Diaconis and Ylvisaker (1979). The family Π is said to be conjugate if the posterior distributions of θ , given X when (θ, X) is $\Pi_{t, m_0}(d\theta)P(\theta, \mu)(dx)$ distributed, still belongs to Π .

Besides Π , Consonni and Veronese (1992) consider (at least on \mathbb{R}) two other families of prior distributions on M_F . The first one, $\tilde{\Pi}$, is defined by a similar construction as for Π , that is, for suitable (t, m_0) ,

$$\tilde{\pi}_{t, m_0}(dm) = \tilde{C}_{t, m_0} \exp t\{\langle \psi_\mu(m), m_0 \rangle - k_\mu(\psi_\mu(m))\} \mathbb{1}_{M_F}(m) dm,$$

while the second one, Π^* , is just the set of the images $k'_\mu(\pi_{t,m_0})$ of π_{t,m_0} of Π by k'_μ . Then, considering real NEF, Consonni and Veronese state that $\tilde{\Pi} = \Pi^*$ if and only if the initial NEF $F(\mu)$ is in the Morris class.

On \mathbb{R}^d , the situation is much more complicated. Here, the equality $\tilde{\Pi} = \Pi^*$ is still satisfied by the simple quadratic and the Wishart NEF. Actually, up to now, we have not been able to determine the whole class of NEF for which $\tilde{\Pi} = \Pi^*$. The only available criteria are that there is equivalence between the three following facts:

- (i) $\tilde{\Pi} = \Pi^*$.
- (ii) There exists (B, b, c) in $E \times \mathbb{R}^2$, such that for all m in M_F , $\det V_F(m) = \exp\{\langle \psi_\mu(m), B \rangle + bk_\mu(m) + c\}$.
- (iii) There exists (B, b) in $E \times \mathbb{R}$ such that for all m in M_F and any basis $(e_i)_{i=1}^d$ of E with dual basis $(e_i^*)_{i=1}^d$, we have

$$\sum_{i=1}^d [V'_F(m)e_i]e_i^* = B + bm.$$

A similar statement has been independently established by Gutierrez-Peña and Smith (1995). These authors have also extended Consonni and Veronese's problem to any parametrization $\Theta(\mu) \rightarrow \Lambda: \theta \mapsto \lambda = \varphi(\theta)$ of $F(\mu)$.

5. Finally, different characterizations of the Morris class involving orthogonal polynomials are due to Feinsilver (1986), Meixner (1934) and Shanbhag (1979) [see also Letac (1992) for a presentation in terms of NEF]. They have been recently extended to \mathbb{R}^d by Pommeret (1995). For this, the simple quadratic class is the natural object to replace Morris class. The whole quadratic class can also be characterized similarly but with weaker hypothesis on polynomials. Using these orthogonal polynomials, Feinsilver (1991) deduced an interesting correspondence between the simple quadratic class and three Lie algebras with finite dimension. In particular, the Gaussian families are put in one-to-one correspondence with self-adjoint operators of the Heisenberg–Weyl algebra, the Poisson families with those of the oscillator algebra and the simple quadratic NEF with a nonnull quadratic part with those of $SL(d + 1)$ [see also Pommeret (1995)]. Such work done with other Lie algebras would enable us to get new quadratic NEF.

We now come to our results. Section 2 presents $2d + 4$ particular NEF on \mathbb{R}^d . They are important because when we take affinities and powers of each of them, we actually get all simple quadratic NEF; this is the essence of our Theorem 2.1, the main theorem of the present paper. Its proof is long and technical and is given in Section 3. However, the discussion relies only on algebraic arguments from the following three simple necessary conditions:

- (i) $\langle \alpha, V_F(m)\beta \rangle = \langle \beta, V_F(m)\alpha \rangle$ (symmetry of V_F),
- (ii) $V'_F(m)(V_F(m)\alpha)\beta = V'_F(m)(V_F(m)\beta)\alpha$,
- (iii) $V_F(m)$ is positive definite on M_F .

(1.4)

Condition (ii) comes from the symmetry of $\psi''_\mu(m) = (V_F(m)^{-1})'$ [see (1.1)] as a Hessian operator. We have gathered in Appendixes A and B some delicate points of our arguments.

2. The $(2d + 4)$ types of simple quadratic NEF.

2.1. We begin first by defining what we call a type. Let $\text{GA}(E)$ denote the affine group of E and let φ be in $\text{GA}(E)$, $\varphi: E \rightarrow E: x \mapsto gx + v$, with linear part g in the linear group $\text{GL}(E)$ and translation vector v . If F is a NEF on E , then $\varphi(F)$ defined as the set of images by φ of each probability of F is still a NEF characterized by

$$(2.1) \quad \begin{aligned} M_{\varphi(F)} &= \varphi(M_F) \\ V_{\varphi(F)}(m) &= gV_F(\varphi^{-1}m)g^t, \end{aligned}$$

where g^t denotes the transpose of g acting on E^* . All the NEF $\varphi(F)$ with φ varying in $\text{GA}(E)$ are called the *affinities of F* .

Now, if μ generates F and if p is a positive real number such that $(L_\mu)^p$ is still the Laplace transform L_{μ_p} of some μ_p in $\mathcal{M}(E)$, then the NEF $F_p = \tilde{F}(\mu_p)$ is called the *p th convolution power of F* . (Note that p is not necessarily an integer). F_p is characterized by

$$\begin{aligned} M_{F_p} &= pM_F \\ V_{F_p}(m) &= pV_F\left(\frac{m}{p}\right). \end{aligned}$$

The set of possible p is called the Jorgensen set of F .

Now, two NEF, F and F' , are said to be of *the same type* if there exist an affinity φ in $\text{GA}(E)$ and a positive real number p such that $F_p = \varphi(F')$.

Note that any affinity or convolution power of a simple quadratic NEF is simple quadratic, too. Hence, to describe the class of simple quadratic NEF in \mathbb{R}^d , we only have to specify one NEF for each of its different types. We will say that this NEF generates the corresponding type.

2.2. We now present $2d + 4$ simple quadratic NEF in \mathbb{R}^d . For each of them, we specify the variance-function and a basis (or occasionally its Laplace transform). These NEF generate $2d + 4$ different types. The first ones, $d + 1$, correspond to affine variance functions, that is, $V(m) = B(m) + C$, and have already been determined by Letac (1989). The last ones, $d + 3$, correspond to simple quadratic variance functions $V_F(m) = am \otimes m + B(m) + C$ with $a \neq 0$. When $d = 2$, we find again the five types given in Casalis (1992b). It was proved there that these eight NEF (five corresponding to the case $a \neq 0$ and three to the case $a = 0$) generate the whole class of simple quadratic NEF in \mathbb{R}^2 . The main result of the present paper is to prove the following theorem for \mathbb{R}^d .

The variance function of $F(\nu)$ is given on $M_{F(\nu)} = \{m \in \mathbb{R}^d; \forall j, m_j > 0\}$ by

$$V_{F(\nu)}(m) = m \otimes m + \text{diag}(m_1, \dots, m_d).$$

For $p > 0$, the p th power ν_p of ν generates the family of the negative-multinomial distributions with parameter p on \mathbb{R}^d . $F(\nu)$ is the $(\text{NM-ga})_d$ family.

To define the $(\text{NM-ga})_{d-1}$ family, we consider the following mixture of a $(d-1)$ -dimensional negative-multinomial family and a gamma family on \mathbb{R} . Let $\nu^{(d-1)}$ denote the measure given in (2.2) on \mathbb{R}^{d-1} , and for $p > 0$ let γ_p be the following measure on \mathbb{R} :

$$(2.4) \quad \gamma_p(dx) = \frac{1}{\Gamma(p)} x^{p-1} \mathbb{1}_{(0, \infty)}(x) dx.$$

We then introduce

$$\mu^{(d-1)}(dx_1, \dots, dx_d) = \nu^{(d-1)}(dx_1, \dots, dx_{d-1}) \gamma_{\sum_{i=1}^{d-1} x_i + 1}(dx_d)$$

with Laplace transform on

$$\Theta(\mu^{(d-1)}) = \left\{ \theta \in \mathbb{R}^d; \sum_{i=1}^{d-1} \exp(\theta_i) + \theta_d < 0 \right\},$$

$$L_{\mu^{(d-1)}}(\theta) = \left(-\theta_d - \sum_{i=1}^{d-1} \exp(\theta_i) \right)^{-1}.$$

The variance function of $F = F(\mu^{(d-1)})$ is defined on $M_F = (0, \infty)^d$ by

$$V_F(m) = m \otimes m + \text{diag}(m_1, \dots, m_{d-1}, 0).$$

The powers F_p of F are generated by the measures

$$\mu_p^{(d-1)}(dx_1, \dots, dx_d) = \nu_p^{(d-1)}(dx_1, \dots, dx_{d-1}) \gamma_{\sum_{i=1}^{d-1} x_i + p}(dx_d)$$

for all $p > 0$. They are composed by the distributions of the random variables (X_1, \dots, X_d) , where (X_1, \dots, X_{d-1}) has a negative-multinomial distribution with parameter p and X_d , conditionally on (X_1, \dots, X_{d-1}) , has a gamma distribution with parameter $\sum_{i=1}^{d-1} X_i + p$.

The $d-1$ other $(\text{NM-ga})_k$ families admit a Gaussian part in addition. Let k be in $\{0, 1, \dots, d-2\}$. We still denote $\nu^{(k)}$ and γ_p the measures on \mathbb{R}^k and \mathbb{R} , respectively, given by (2.2) and (2.4). Let $\lambda_p^{(d-k-1)}$ be the normal distribution on \mathbb{R}^{d-k-1} with mean 0 and covariance pI_{d-k-1} . Then we put, if $k \geq 1$,

$$\mu^{(k)}(dx_1, \dots, dx_d) = \nu^{(k)}(dx_1, \dots, dx_k) \gamma_{\sum_{i=1}^k x_i + 1}(dx_{k+1})$$

$$\times \lambda_{x_{k+1}}^{(d-k-1)}(dx_{k+2}, \dots, dx_d)$$

and

$$\mu^{(0)}(dx_1, \dots, dx_d) = \gamma_1(dx_1) \lambda_{x_1}^{(d-1)}(dx_2, \dots, dx_d).$$

Then we have

$$\Theta(\mu^{(k)}) = \left\{ \theta \in \mathbb{R}^d; \theta_{k+1} + \frac{1}{2} \sum_{i=k+2}^d \theta_i^2 + \sum_{i=1}^k \exp(\theta_i) < 0 \right\},$$

$$L_{\mu^{(k)}}(\theta) = \left(-\theta_{k+1} - \frac{1}{2} \sum_{i=k+2}^d \theta_i^2 - \sum_{i=1}^k \exp(\theta_i) \right)^{-1},$$

$$M_{F(\mu^{(k)})} = (0, \infty)^{k+1} \times \mathbb{R}^{d-k-1},$$

$$V_{F(\mu^{(k)})}(m) = m \otimes m + \text{diag}(m_1, \dots, m_k, 0, m_{k+1}, \dots, m_{k+1}).$$

(Note that $a \neq 0$ and $C = 0$ for these NEF).

Here again, the powers F_p of $F(\mu^{(k)})$ are generated by

$$\mu_p^{(k)}(dx_1, \dots, dx_n) = \nu_p^{(k)}(dx_1, \dots, dx_k) \gamma_{\sum_{i=1}^k X_i + p}(dx_{k+1})$$

$$\times \lambda_{x_{k+1}}^{(d-k-1)}(dx_{k+1}, \dots, dx_n)$$

for all $p > 0$. They are composed by the distributions of (X_1, \dots, X_d) , where (X_1, \dots, X_k) has a negative-multinomial distribution with parameter p , X_{k+1} given (X_1, \dots, X_k) is gamma distributed with parameter $\sum_{i=1}^k X_i + p$, and (X_{k+2}, \dots, X_d) given (X_1, \dots, X_{k+1}) are $d - k - 1$ real independent Gaussian variables with mean 0 and variance X_{k+1} .

Note that the cases where the three negative-multinomial, gamma and Gaussian families are mixed do not appear in \mathbb{R}^2 . The family $F(\mu^{(1)})$ on \mathbb{R}^2 appears the first time in the paper of Bar-Lev, Bshouty, Enis, Letac, Li Lu and Richards (1994) as one of the NEF whose margins are in two different Morris families.

We now describe two isolated types.

(c) *The multinomial type M.* We take again the representation of the multinomial distributions from Letac (1989) and refer the reader to Ratnaparkhi (1988) for a bibliography about them.

Let $\{e_1, \dots, e_d\}$ denote the canonical basis of \mathbb{R}^d and let e_0 be the null vector. Then the measure $\mu(dx) = \sum_{i=0}^d \delta_{e_i}$ on \mathbb{N}^d generates a NEF F with variance function on $M_F = \{m \in \mathbb{R}^d, \forall j, m_j > 0, \sum_{j=1}^d m_j < 1\}$ equal to

$$V_F(m) = -m \otimes m + \text{diag}(m_1, \dots, m_d).$$

For any p in $\mathbb{N} \setminus \{0\}$, the p th power F_p of F is the set of probabilities $P(m, F_p)$ defined by

$$(2.5) \quad P(m, F_p)(n_1 e_1 + \dots + n_d e_d)$$

$$= \binom{p}{n_0, n_1, \dots, n_d} \left(1 - \frac{\sum m_j}{p} \right)^{n_0} \prod_{j=1}^d \left(\frac{m_j}{p} \right)^{n_j},$$

where n_0, n_1, \dots, n_d are positive integers with sum p .

(d) *The hyperbolic type H.* Similar to the $(\text{NM-ga})_{d-1}$ type, this last case is built from the following mixture of a negative-multinomial family on \mathbb{R}^{d-1}

variable with a negative-multinomial law $P(m, F)$ given by (2.3), then, for any $k = 1, \dots, d - 1$, (X_1, \dots, X_k) still has a negative-multinomial distribution [still given by (2.3) on replacing d by k], while X_{k+1} conditionally on (X_1, \dots, X_k) has a negative-binomial distribution with Jorgensen parameter $1 + \sum_{i=1}^k X_i$. The same remark holds true for a multinomial distribution $P(m, F_p)$ given in (2.5): (X_1, \dots, X_k) has the distribution $P((m_1, \dots, m_k), F_p)$ in \mathbb{R}^k [still given by (2.5)], while X_{k+1} conditionally on (X_1, \dots, X_k) has a binomial distribution with Jorgensen parameter $p - \sum_{i=1}^k X_i$. Hence, any simple quadratic distribution presented in this section has the following remarkable property: if (X_1, \dots, X_d) is so distributed, then the law of X_1 belongs to a Morris family and for any $k = 2, \dots, d$, the law of X_k conditionally on (X_1, \dots, X_{k-1}) is also a Morris distribution with Jorgensen parameter depending on an affinity of (X_1, \dots, X_{k-1}) . Nevertheless, such mixtures do not always give simple quadratic NEF.

3. The classification of the simple quadratic variance functions.

This section is entirely devoted to the proof of Theorem 2.1 giving all the types of simple quadratic variance functions $V_F(m) = am \otimes m + B(m) + C$ on \mathbb{R}^d .

The case $a = 0$ has already been developed by Letac (1989) and yields the $(d + 1)$ Poisson–Gaussian types $(PG)_k$. Therefore we only consider the case $a \neq 0$. Via the following lemma, whose proof is reported in Appendix A, the problem is reduced to the case where $C = 0$.

LEMMA 3.1. *Let F be a NEF on \mathbb{R}^d with variance function $V_F(m) = am \otimes m + B(m) + C$ with $a \neq 0$ and let \tilde{V}_F be the polynomial function on \mathbb{R}^d defined by V_F . Then, if $d \geq 2$, there exists m_0 in \mathbb{R}^d such that $\tilde{V}_F(m_0) = 0$.*

Consequently, if τ_{-m_0} denotes the translation of vector $-m_0$, then from (2.1), the variance function of $\tau_{-m_0}(F)$ clearly is

$$V_{\tau_{-m_0}(F)}(m) = am \otimes m + \tilde{B}(m)$$

with

$$\tilde{B}(m) = a(m \otimes m_0 + m_0 \otimes m) + B(m).$$

(In Section 2 we have written the $d + 3$ variance functions corresponding to $a \neq 0$ with a C null. For the hyperbolic type, we have done the translation $\tau_{e_{d-1}}(\cdot)$.)

Theorem 2.1 is now reduced to the following statement:

PROPOSITION 3.2. *Let F be a NEF on E with variance function*

$$(3.1) \quad V_F(m) = am \otimes m + B(m) \quad (a \neq 0).$$

Then F belongs to the M , H or $(NM\text{-}ga)_k$ types.

PROOF. The proof is divided into several steps. In the first one, we introduce some linear endomorphisms $\{Q(\alpha), \alpha \in E^*\}$ of $\mathcal{L}(E^*)$ in a one-to-

one correspondence with the operators $\{B(m), m \in E\}$ of (3.1), so that knowing $Q: \alpha \mapsto Q(\alpha)$ will be equivalent to knowing B and hence, given the real number a , to knowing $V: m \mapsto V(m) = am \otimes m + B(m)$. We also write the necessary conditions (1.4) satisfied by V in terms of Q and then deduce the different possible forms of the $Q(\alpha)$.

In order to classify the functions V in types, we have to simplify the functions Q as much as possible by some action of affinities on V . Hence, in a second step, we translate the action of affinities on the operators $Q(\alpha)$ (Lemma 3.4).

In a third step, we examine each possible form of $Q(\alpha)$ separately, simplify it and finally recognize the corresponding V and hence F .

First step. Let F be a NEF with variance function of the form (3.1). Then the three conditions (1.4) satisfied by V_F can simply be written

$$(3.2) \quad \begin{aligned} & \text{(i) } \langle \alpha, B(m)\beta \rangle = \langle \beta, B(m)\alpha \rangle, \\ & \text{(ii) } B(B(m)\alpha)\beta = B(B(m)\beta)\alpha, \\ & \text{(iii) for } m \text{ in } M_F, V_F(m) \text{ is positive-definite.} \end{aligned}$$

Let us define the unique linear map $Q: E^* \rightarrow \mathcal{L}(E^*)$ such that for (α, β) in $(E^*)^2$ and m in E ,

$$(3.3) \quad \langle Q(\alpha)\beta, m \rangle = \langle \beta, B(m)\alpha \rangle,$$

that is, $Q^t(\alpha)m = B(m)\alpha$.

For example, the $Q(\alpha)$ corresponding to the $d + 3$ simple quadratic NEF with $a \neq 0$ presented in Section 2 are as follows.

(a) For the $(\text{NM-ga})_k$ NEF,

if $k = d$, $Q(\alpha) = \text{diag}(\alpha_1, \dots, \alpha_d)$;

$$\text{if } 0 \leq k \leq d - 1, \quad Q(\alpha) = \begin{pmatrix} \alpha & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & \alpha_k & & & & & & \\ & & & & 0 & & & & & \\ & & & & & \alpha_{k+2} & \cdots & & \alpha_d & \\ & & & & & 0 & \cdots & & 0 & \\ & & & & & \cdot & \cdots & & \cdot & \\ 0 & & & & & 0 & \cdots & & 0 & \end{pmatrix}.$$

(b) For the M.NEF: $Q(\alpha) = \text{diag}(\alpha_1, \dots, \alpha_d)$.

(c) For the H.NEF,

$$Q(\alpha) = \begin{pmatrix} \alpha_1 & & & & & & & & & 0 \\ & \ddots & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & \alpha_{d-2} & & & & & & \\ & & & & -\alpha_{d-1} & & \alpha_d & & & \\ 0 & & & & -\alpha_d & & -\alpha_{d-1} & & & \end{pmatrix}.$$

From (3.2)(iii), we derive the following necessary condition.

LEMMA 3.3. Let V_F be a variance function such that $V_F(m) = am \otimes m + B(m)$ and let Q be the linear endomorphism of E^* defined in (3.3). If H is a subspace of E^* such that for all α in H , we have $Q(\alpha)\alpha = 0$, then H has dimension 0 or 1.

PROOF. For any m in M_F and α in H , we have $\langle Q(\alpha)\alpha, m \rangle = \langle \alpha, B(m)\alpha \rangle = 0$, thus $\langle \alpha, V_F(m)\alpha \rangle = a\langle \alpha, m \rangle^2$. Let m^\perp denote the subspace of E^* orthogonal to m . Then if the dimension of H is greater than or equal to 2, the intersection $m^\perp \cap H$ is not reduced to $\{0\}$ and for any α in $m^\perp \cap H$, $\langle \alpha, V_F(m)\alpha \rangle = 0$. Hence V_F is not positive-definite on M_F , which contradicts (3.2)(iii). \square

Written with Q , (3.2)(i) and (ii) become

$$(3.4) \quad \begin{aligned} & \text{(i) } Q(\alpha)\beta = Q(\beta)\alpha, \\ & \text{(ii) } Q(\alpha)Q(\beta) = Q(\beta)Q(\alpha). \end{aligned}$$

Thus, the vector space $Q = \{Q(\alpha), \alpha \in E^*\}$ generates a commutative subalgebra of $\mathcal{L}(E^*)$. Using the theory of nilpotent Lie algebras [Dynkin (1950), Theorem II, page 380], we can split the space E^* into a direct sum of r subspaces E_i^* of dimension n_i , $i = 1, \dots, r$, invariant under the endomorphisms $Q(\alpha)$ and for which there exists r_0 in $\{0, 1, \dots, r\}$ such that the following hold.

1. If $0 \leq i \leq r_0$, there exist a basis $e_i^* = (e_{ij}^*)_{j=1, \dots, n_i}$ of E_i^* and vectors d^i and w_{jk}^i , $1 \leq j < k \leq n_i$, of E such that the matrix in e_i^* of the restriction of $Q(\alpha)$ to E_i^* is exactly

$$(3.5) \quad Q_i(\alpha) = \begin{pmatrix} \langle \alpha, d^i \rangle & & \langle \alpha, w_{jk}^i \rangle \\ & \ddots & \\ 0 & & \langle \alpha, d^i \rangle \end{pmatrix}.$$

2. If $r_0 + 1 \leq i \leq r$, then $n_i/2 = p_i \in \mathbb{N}$ and there exist a basis $e_i^* = (e_{ij}^*)_{j=1, \dots, n_i}$ of E_i^* , and vectors $a^i, b^i, s_{jk}^i, t_{jk}^i$, $1 \leq j < k < p_i$, such that, writing

$$(3.6) \quad D^i = \begin{pmatrix} \langle \alpha, a^i \rangle & \langle \alpha, b^i \rangle \\ -\langle \alpha, b^i \rangle & \langle \alpha, a^i \rangle \end{pmatrix} \quad \text{and} \quad W_{jk}^i = \begin{pmatrix} \langle \alpha, s_{jk}^i \rangle & \langle \alpha, t_{jk}^i \rangle \\ -\langle \alpha, t_{jk}^i \rangle & \langle \alpha, s_{jk}^i \rangle \end{pmatrix},$$

the matrix in e_i^* of the restriction of $Q(\alpha)$ to E_i^* is

$$(3.7) \quad Q_i(\alpha) = \begin{pmatrix} D^i & & W_{jk}^i \\ & \ddots & \\ 0 & & D^i \end{pmatrix}.$$

(See also Appendix B.)

Moreover, the relations (3.4) are equivalent to

$$(3.8) \quad \begin{aligned} & \text{(i) } Q_i(\alpha)\beta = Q_i(\beta)\alpha \text{ for } (\alpha, \beta) \in (E_i^*)^2, \\ & \text{(ii) } Q_i(\alpha) = 0 \text{ if } \alpha \in \bigoplus_{j \neq i} E_j^*, \\ & \text{(iii) } Q_i(\alpha)Q_i(\beta) = Q_i(\beta)Q_i(\alpha). \end{aligned}$$

The second equality, (3.8)(ii), clearly implies that the vectors d^i, w_{jk}^i or a^i, b^i, s_{jk}^i and t_{jk}^i are vectors of the dual space E_i of E_i^* . We note by $e_i = (e_{ij})_{j=1, \dots, n_i}$ the dual basis of e_i^* .

Second step. To simplify the matrices (3.5), (3.6) and (3.7), we can only work with affinities which preserve the types of NEF. Therefore, we have to know their action on the operator Q .

LEMMA 3.4. *Let F be a NEF on E with variance function given by (3.1). Let g be in $GL(E)$ and let τ_{-m_0} be the translation of E of vector $-m_0$ such that $V_F(m_0) = 0$. Then the variance functions of $g(F)$ and $\tau_{-m_0}(F)$ are also of the form (3.1). Moreover, the associated operators Q_g and Q_{m_0} satisfy*

$$\begin{aligned} Q_g(\alpha) &= (g^t)^{-1} Q(g^t \alpha) g^t, \\ Q_{m_0}(\alpha) &= Q(\alpha) + a \langle \alpha, m_0 \rangle \text{id} + a \alpha \otimes m_0. \end{aligned}$$

PROOF. The result follows from an obvious calculation using (2.1) and (3.3). \square

Therefore, given a NEF F with variance function (3.1) and Q as given in (3.3), the only translations that we are allowed to use to simplify Q are the translations of vectors $-m_0$ such that $V_F(m_0) = 0$. The following lemma describes the set of such vectors.

LEMMA 3.5. *Let be $V(m) = am \otimes m + B(m)$ and let Q be as defined in (3.3). Then we have the following statements:*

- (i) *If $V(m_0) = 0$, the hyperplane $H_{m_0} = \{\alpha \in E^*; \langle \alpha, m_0 \rangle = 0\}$ is stable under Q .*
- (ii) *Conversely, let H be a hyperplane of E^* invariant by the $Q(\alpha)$. Let be $e \in H^\perp = \{x \in E; \forall \alpha \in H \langle \alpha, x \rangle = 0\}$ with $e \neq 0$ and $e^* \in E^*$ such that $\langle e^*, e \rangle = 1$. Then we have $V(-(\langle Q(e^*)e^*, e \rangle/a)e) = 0$.*

PROOF. Here again, the result comes from the definitions through a simple calculation. \square

Third step. We will now study and simplify each form of Q_i given in (3.5) and (3.7) separately by affinities. Observe that a linear operator acting on a

E_i^* only (and hence equal to the identity on $\bigoplus_{j \neq i} E_j^*$) does not change the Q_j for $j \neq i$. However, this is no longer true for a translation and the whole matrix Q has to be modified in that case. We thus successively examine the three following cases:

1. Q_i has the form (3.7);
2. Q_i has the form (3.5) with $d^i \in E_i$;
3. Q_i has the form (3.5) with $d^i = 0$.

CASE 1. We consider the situation where Q_i has the form (3.7).

LEMMA 3.6. *Suppose that $n_i = 2p$ is even and that Q_i has the form (3.7) in the basis $e_i^* = (f_1^*, f_1'^*, f_2^*, f_2'^*, \dots, f_p^*, f_p'^*)$. Let $\{e_i = (f_1, f_1', f_2, f_2', \dots, f_p, f_p')\}_i$ be the dual basis of $\{e_i^*\}_i$. Then $p = 1$ and there exists a linear operator g of E acting only on E_i^* such that*

$$(3.9) \quad (Q_g)_i(\alpha) = \begin{pmatrix} -\langle \alpha, f_1 \rangle & \langle \alpha, f_1' \rangle \\ -\langle \alpha, f_1' \rangle & -\langle \alpha, f_1 \rangle \end{pmatrix}.$$

PROOF. We have only to consider E_i^* . Then, with the notations of (3.6) and (3.7), writing

$$\alpha^i = \sum_{k=1}^p (a_k f_k + a'_k f'_k), \quad s_{jk}^i = \sum_{l=1}^p (s_{jk,l} f_l + s'_{jk,l} f'_l)$$

and with similar notation for b^i and t_{jk}^i , we get from (3.8), for $p \geq 2$,

$$\begin{aligned} Q(f_1^*)f_2^* &= Q(f_2^*)f_1^* \rightarrow a_1 = b_1 = 0 \\ Q(f_1'^*)f_2'^* &= Q(f_2'^*)f_1'^* \rightarrow a'_1 = b'_1 = 0. \end{aligned}$$

Hence, if H is the subspace generated by f_1^* and $f_1'^*$, we have $Q(\alpha)\beta = 0$ for (α, β) in H^2 . From Lemma 3.3, Q does not yield a variance function. Hence $p = 1$. Now, since $Q(\alpha)\beta = Q(\beta)\alpha$, we also get $a'_1 = b_1$, $b'_1 = -a_1$ and $(a_1, b_1) \neq (0, 0)$. The linear transformation g of matrix

$$\frac{1}{a_1^2 + b_1^2} \begin{pmatrix} -a_1 & -b_1 \\ b_1 & -a_1 \end{pmatrix}$$

in the basis (f_1, f_1') and equal to the identity on $\bigoplus_{j \neq i} E_j$ simplifies $Q(\alpha)$ into $Q_g(\alpha)$ given in (3.9). \square

PROPOSITION 3.7. *Let F be a NEF with variance function (3.1), let Q be the associated operator defined by (3.3) and let $E^* = \bigoplus_{i=1}^r E_i^*$ be the decomposition of E^* into invariant subspaces. Then, there exists at most one E_i^* such that Q_i has the form (3.7).*

PROOF. If it is not the case, suppose that on E_1^* and E_2^* , Q_1 and Q_2 have the form (3.7). Then $n_1 = n_2 = 2$ and, after a linear transformation, Q_1 and Q_2 can be written as in (3.9) in the basis $(f_1^*, f_1'^*)$ and $(f_2^*, f_2'^*)$. Then the corresponding function $V(m) = (V_{ij}(m)) = am \otimes m + B(m)$ satisfies (with obvious notations), for $i = 1, 2$, $V_{ii}(m) = am_i^2 - m_i$, $V_{i'i'}(m) = am_i m_i' - m_i'$, $V_{i'i}(m) = am_i'^2 + m_i$ and $V_{12}(m) = am_1 m_2$, so that the principal minors

$$\begin{aligned}
 V_{ii}(m) &= (am_i - 1)m_i \quad (i = 1, 2), \\
 \begin{vmatrix} V_{ii} & V_{i'i'} \\ V_{i'i'} & V_{i'i'} \end{vmatrix} &= (am_i - 1)(m_i^2 + m_i'^2) \quad (i = 1, 2), \\
 \begin{vmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{vmatrix} &= m_1 m_2 (1 - am_1 - am_2)
 \end{aligned}$$

are not all positive. So V is positive-definite on no open subset of E and, hence, V is not a variance function. \square

CASE 2. We now examine the triangular form (3.5). Let us begin with the following lemma.

LEMMA 3.8. *Suppose that in the basis $e_i^* = (e_{ij}^*)_{j=1, \dots, n_i}$, the restriction of Q to E_i^* is the triangular matrix*

$$Q_i(\alpha) = \begin{pmatrix} \langle \alpha, d^i \rangle & & \langle \alpha, w_{jk}^i \rangle \\ & \ddots & \\ 0 & & \langle \alpha, d^i \rangle \end{pmatrix},$$

where d^i and w_{jk}^i are in the dual space E_i . Then (i) $d^i = d_i e_{in_i}$ and (ii) if $d^i \neq 0$, there exist a linear operator g acting on E_i (and then equal to the identity on $\bigoplus_{j \neq i} E_j$) and vectors $(\tilde{w}_{jk}^i)_{1 \leq j < k \leq n_i}$ of E_i such that

$$(Q_g)_i(\alpha) = \begin{pmatrix} \alpha_{in_i} & & \langle \alpha, \tilde{w}_{jk}^i \rangle \\ & \ddots & \\ 0 & & \alpha_{in_i} \end{pmatrix}.$$

PROOF. (i) comes from the relation $Q(e_{ij}^*)e_{ik}^* = Q(e_{ik}^*)e_{ij}^*$ and (ii) is obtained with $g = \text{id}_{\bigoplus_{j \neq i} E_j} + (1/d_i)\text{id}_{E_i}$. \square

PROPOSITION 3.9. *Let F be a NEF with variance function (3.1), let Q be the associated operator defined in (3.3) and let $E^* = \bigoplus_i E_i^*$ be the decomposition of E into invariant subspaces. Then the following hold.*

(i) *There exists at most one E_i^* with dimension greater than 1 such that Q_i has the triangular form (3.5).*

(ii) If there exists such a E_i^* , then for any other Q_j ($j \neq i$) of the form (3.5), $d^j \neq 0$.

PROOF. (i) If it is not the case, suppose that E_1^* and E_2^* are subspaces of dimension greater than 1 such that Q_1 and Q_2 have the form (3.5) in the basis $(e_{1j}^*)_{j=1, \dots, n_1}$ and $(e_{2j}^*)_{j=1, \dots, n_2}$. Then from Lemma 3.8, for $i = 1, 2$, $Q_i(e_{i1}^*) = 0$. Hence, if H is the subspace generated by e_{11}^* and e_{21}^* , we have $Q(\alpha) = 0$ on H . From Lemma 3.3, Q does not yield a variance function.

(ii) If Q_j has the form (3.5) with $d^j = 0$, then $Q(e_{i1}^*) = Q(e_{j1}^*) = 0$ and we conclude by using Lemma 3.3, with H generated by e_{i1}^* and e_{j1}^* . \square

To summarize, if there exists a subspace E_i^* with dimension $n_i > 1$ such that Q_i has the triangular form (3.5), then, up to a linear transformation, there exist a basis (e_{ij}^*) of E^* , a scalar ε_r equal to 1 or 0 and vectors w_{jk}^r of E_r such that the matrix of Q in the basis of $e^* = (e_{ij}^*)$ is exactly

$$[Q(\alpha)]_{e^*} = \begin{pmatrix} Q_1(\alpha) & & 0 \\ & \ddots & \\ 0 & & Q_r(\alpha) \end{pmatrix}$$

with

$$Q_1(\alpha) = (\langle \alpha, e_{11} \rangle) \quad \text{or} \quad \begin{pmatrix} -\langle \alpha, e_{11} \rangle & \langle \alpha, e_{12} \rangle \\ -\langle \alpha, e_{12} \rangle & -\langle \alpha, e_{11} \rangle \end{pmatrix},$$

$$Q_i(\alpha) = (\langle \alpha, e_{i1} \rangle) \quad \text{for } i = 2, \dots, r - 1,$$

$$Q_r(\alpha) = \begin{pmatrix} \varepsilon_r \langle \alpha, e_{rn_r} \rangle & & \langle \alpha, w_{jk}^r \rangle \\ & \ddots & \\ 0 & & \varepsilon_r \langle \alpha, e_{rn_r} \rangle \end{pmatrix}.$$

In particular, the hyperplane $H = \{\alpha \in E^*; \langle \alpha, e_{rn_r} \rangle = 0\}$ is invariant under the $Q(\alpha)$. From Lemma 3.5, the vector

$$m_0 = -\frac{\langle Q(e_{rn_r}^*)e_{rn_r}^*, e_{rn_r} \rangle}{\alpha} e_{rn_r} = -\frac{\varepsilon_r}{\alpha} e_{rn_r}$$

satisfies $V(m_0) = 0$. If $m_0 \neq 0$, that is, $\varepsilon_r = 1$, by the translation τ_{-m_0} , Q is changed into the operator \tilde{Q} defined in Lemma 3.4 by

$$\tilde{Q}(\alpha) = Q(\alpha) - \alpha_{rn_r} \text{id} - \alpha \otimes e_{rn_r},$$

and having the matrix form in the basis e^* :

$$[\tilde{Q}(\alpha)]_{e^*} = \begin{pmatrix} \tilde{Q}_1(\alpha) & & -\langle \alpha, e_{11} \rangle & & \\ & & \vdots & & \\ & \ddots & & -\langle \alpha, e_{r-1, n_{r-1}} \rangle & \\ 0 & & & & \tilde{Q}_r(\alpha) \end{pmatrix},$$

where

$$\tilde{Q}_i(\alpha) = Q_i(\alpha) - \alpha_{rn_r} \text{id}_{E_i^*} \text{ for } i \neq r$$

and

$$\begin{aligned} \tilde{Q}_r(\alpha) &= Q_r(\alpha) - \alpha_{rn_r} \text{id}_{E_r^*} - {}^t(\alpha_{r1}, \dots, \alpha_{rn_r}) \otimes e_{rn_r} \\ &= \begin{pmatrix} 0 & \langle \alpha, w_{jk}^r \rangle & & -\alpha_{r1} + \langle \alpha, w_{1, n_r}^r \rangle \\ & 0 & & -\alpha_{r2} + \langle \alpha, w_{2, n_r}^r \rangle \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 - \alpha_{rn_{r-1}} + \langle \alpha, w_{n_r-1, n_r}^r \rangle \\ 0 & 0 & \cdot & 0 - \alpha_{rn_r} \end{pmatrix}. \end{aligned}$$

Clearly, E_1^*, \dots, E_{r-1}^* are still invariant subspaces on which $Q(\alpha)$ has either complex eigenvalues or the simple eigenvalues $\langle \alpha, e_{i1} - e_{rn_r} \rangle$. The other eigenvalues of $Q(\alpha)$ are 0 and $-\langle \alpha, e_{r, n_r} \rangle = -\alpha_{rn_r}$. From Appendix B, E^* can be split as a direct sum $\bigoplus_{i=1}^{r+1} \tilde{E}_i^*$ of $(r + 1)$ invariant subspaces such that $\tilde{E}_i^* = E_i^*$ for $i = 1, \dots, r - 1$, \tilde{E}_r^* is a real line and on \tilde{E}_{r+1}^* , $Q(\alpha)$ has the unique eigenvalue 0. Therefore, in a suitable basis of E^* , that is, $e^* = (e_1^*, \dots, e_r^*, e_{r+1}^*, \dots, e_n^*)$ or $e^* = (e_1^*, e_1'^*, e_2^*, e_3^*, \dots, e_n^*)$, we have

$$[Q(\alpha)]_{e^*} = \begin{pmatrix} \hat{Q}_1(\alpha) & 0 \\ 0 & \hat{Q}_{r+1}(\alpha) \end{pmatrix}$$

with

$$\begin{aligned} \hat{Q}_1(\alpha) &= (\alpha_1) \text{ or } \begin{pmatrix} -\alpha_1 & \alpha_1' \\ -\alpha_1' & \alpha_1 \end{pmatrix}, \\ (3.10) \quad \hat{Q}_i(\alpha) &= (\alpha_i) \text{ for } i = 2, \dots, r - 1 \\ \hat{Q}_{r+1}(\alpha) &= \begin{pmatrix} 0 & (\langle \alpha, w_{jk}^{r+1} \rangle) \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

CASE 3. We now restrict our attention to the only subspace \tilde{E}_{r+1}^* , on which $Q(\alpha)$ is given by the matrix $\hat{Q}_{r+1}(\alpha)$ in (3.10).

To simplify the notation, we write E^* and $Q(\alpha)$ for \tilde{E}_{r+1}^* and $\hat{Q}_{r+1}(\alpha)$, respectively. We will prove by induction that there exists a linear transformation g of E changing Q into Q_g , with matrix

$$(3.11) \quad [Q_g(\alpha)]_{e^*} = \begin{pmatrix} 0 & \alpha_2 & \cdots & \alpha_n \\ 0 & 0 & \cdots & 0 \\ \cdot & & & \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Write

$$[Q(\alpha)]_{e^*} = \begin{pmatrix} 0 & & \langle \alpha, w_{ij} \rangle \\ & \cdot & \\ 0 & & 0 \end{pmatrix},$$

with $w_{ij} = \sum_{k=1}^n w_{ij,k} e_k$. The relations $Q(e_k^*)e_j^* = Q(e_j^*)e_k^*$ imply here that

$$(3.12) \quad w_{ij,k} = \begin{cases} w_{ik,j}, \\ 0, & \text{if } i \geq \inf\{j, k\}. \end{cases}$$

We will kill the $w_{ij,j}$ for $j > 1$ successively by induction.

Step 1. $j = 2$. Then $w_{12,2} \neq 0$. If not, on the subspace generated by e_1^* and e_2^* , $Q(\alpha)\beta = 0$, which is impossible from Lemma 3.5. Let us denote $\varepsilon_1 = \text{sign}(w_{12,2})$ and consider the transformation g_1 such that $g_1(e_1) = \varepsilon_1 e_1$,

$$g_1 e_2 = \frac{\varepsilon_1 e_2}{|w_{12,2}|^{1/2}} - \sum_{j=3}^r \frac{w_{12,j}}{w_{12,2}} e_j,$$

and for $i \geq 3$, $g_1(e_i) = e_i$. Then a simple calculation gives

$$\begin{aligned} Q_1(\alpha) &= (g_1^t)^{-1}(\alpha)Q(g_1^t \alpha)g_1^t \\ &= \begin{pmatrix} 0 & & \langle \alpha, w_{ij}^{(1)} \rangle \\ & \cdot & \\ 0 & & 0 \end{pmatrix} \end{aligned}$$

with $w_{12}^{(1)} = e_2$.

Step 2. Suppose that there exists a linear transformation $g_{(k)}$ of E such that

$$(3.13) \quad \begin{aligned} Q_k(\alpha) &= (g_{(k)}^t)^{-1}Q(g_{(k)}^t \alpha)g_{(k)}^t \\ &= \begin{pmatrix} 0 & & \langle \alpha, w_{ij}^{(k)} \rangle \\ & \cdot & \\ 0 & & 0 \end{pmatrix} \end{aligned}$$

with $w_{1j} = e_j$ for $j \leq k$ and $w_{ij} = 0$ for $i \geq 2$ and $j \leq k$. Then we will prove the following:

1. for $i \geq 2$, $w_{ik+1} = 0$;
2. $w_{1k+1,k+1} \neq 0$;

3. if $\varepsilon_{k+1} = \text{sign}(w_{1k+1, k+1})$ and g_{k+1} is the transformation such that for $i \leq k$, $g_{k+1}(e_i) = \varepsilon_{k+1}e_i$,

$$g_{k+1}(e_{k+1}) = \frac{\varepsilon_{k+1}e_{k+1}}{|w_{1k+1, k+1}|^{1/2}} - \sum_{j \geq k+2} \frac{w_{1k+1, j}}{w_{1k+1, k+1}}e_j,$$

and for $j \geq k + 2$, $g_{k+1}(e_j) = e_j$, then the property (3.13) for $(k + 1)$ is obtained with $g_{(k+1)} = g_{k+1}g_{(k)}$.

To prove (1), we only write the equality $Q(\beta)Q(e_i^*)e_{k+1}^* = Q(e_i^*)Q(\beta)e_{k+1}^*$, for any β . For (2), we have $Q(e_{k+1}^*)e_{k+1}^* = w_{1k+1, k+1}e_1^*$. If $w_{1k+1, k+1} = 0$, then for any α, β in the subspace generated by e_1^* and e_{k+1}^* , $Q(\alpha)\beta = 0$. From Lemma 3.5 this is impossible. From this, a simple calculation yields the property (3.13) for $(k + 1)$. Consequently, for $k = n$, we get

$$Q_n(\alpha) = \begin{pmatrix} 0 & \alpha_2 & \cdots & \alpha_n \\ 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

4. Conclusion. Here up to affinities we summarize the only operators Q which yield a symmetric nondegenerate function $V(m) = am \otimes m + B(m)$ satisfying the relation of symmetry (1.4)(ii). The last property of positive-definiteness is now used to exclude some cases. Thus we get the following.

Case 1. $Q(\alpha) = \text{diag}(\alpha_1, \dots, \alpha_d)$ leads to $V(m) = am \otimes m + \text{diag}(m_1, \dots, m_d)$, which is the variance function of a multinomial or negative-multinomial family.

Case 2.

$$Q(\alpha) = \begin{pmatrix} \alpha_1 & & & & & \\ & \cdot & & & & \\ & & \alpha_k & & & \\ & & & 0 & \alpha_{k+2} & \cdots & \alpha_d \\ & & & 0 & 0 & \cdots & 0 \\ & & & 0 & & & 0 \end{pmatrix}$$

leads to $V(m) = am \otimes m + \text{diag}(m_1, \dots, m_k, 0, m_{k+1}, \dots, m_{k+1})$, which is the variance function of a (NM-ga)_k family (with $0 \leq k \leq d - 1$).

Case 3.

$$Q(\alpha) = \begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & & \\ & & \alpha_{n-2} & & & \\ 0 & & & -\alpha_{n-1} & \alpha_n & \\ & & & -\alpha_n & -\alpha_{n-1} & \end{pmatrix}$$

Let us accept Lemma B for a while.

PROOF OF LEMMA A. We proceed by induction on the dimension of E .

Step 1. Suppose that $d = 2$ and $c \neq 0$ (if $c = 0$, $m_0 = 0$ is a solution). From Lemma B, there exists u such that u is an eigenvector of $B(u)$. Then put $e_1 = u/\|u\|$ and e_2 such that $e = (e_1, e_2)$ is an orthonormal basis of E . From (A.1) and the symmetry of $B(e_2)$, we can write

$$[B(e_1)]_e = \begin{pmatrix} a_1 & 0 \\ 0 & b \end{pmatrix}, \quad [B(e_2)]_e = \begin{pmatrix} 0 & b \\ b & a_2 \end{pmatrix}.$$

Moreover, from (A.1)(ii), $[B(e_1), B(e_2)] = -ce_2$ implies that $b^2 - a_1b + c = 0$. Since $c \neq 0$, we have also $b \neq 0$ and $B(e_1) = (a_1 - b)e_1 \otimes e_1 + b \text{id}_E$. We then easily check that $V(-(c/b)e_1) = 0$.

Step 2. Suppose now Lemma A proved for any E of dimension k with $2 \leq k \leq d$. We will prove it for an E of dimension $d + 1$. Let us introduce $e_0 = u/\|u\|$, where u satisfies the conclusion of Lemma B and let us complete e_0 in an orthonormal basis $e = (e_i)_{i=0}^d$ of E which diagonalizes $B(e_0)$. We write $[B(e_0)]_e = \text{diag}(a_1, b_1, \dots, b_d)$. Then, from (A.1), we have successively, for $i, j \geq 1$, $[B(e_0), B(e_i)]_{e_0} = -ce_i$ and $[B(e_i), B(e_j)]_{e_0} = 0$, which imply, respectively,

$$(A.2) \quad b_i^2 - a_1b_i + c = 0 \quad \text{for } i \geq 1,$$

$$(A.3) \quad (b_j - b_i)B(e_i)e_j = 0 \quad \text{for } i, j \geq 1.$$

Hence, if $b_i = b$ for any $i \geq 1$, then $B(e_0) = (a_1 - b)e_0 \otimes e_0 + b \text{id}_E$ and $V(-(c/b)e_0) = 0$ comes from (A.2). If the b_i are not all equal, up to a permutation of e_1, \dots, e_d , we can write

$$(A.4) \quad [B(e_0)]_e = \begin{pmatrix} a_1 & & 0 \\ & bI_k & \\ 0 & & (a_1 - b)I_{d-k} \end{pmatrix}$$

with $a_1 - b \neq b$ and $|a_1 - b| \leq |b|$. Let E_k be the vector space generated by e_1, \dots, e_k , let π_k be the orthogonal projection onto E_k and, for m in E_k , define $\tilde{B}(m) = \pi_k B(m) \pi_k$. It is easy to verify that for (x, y) in E_k^2 , $\tilde{B}(x)y = B(x)y - b\langle x, y \rangle e_0$, so that

$$\tilde{B}(x)y = \tilde{B}(y)x$$

$$[\tilde{B}(x), \tilde{B}(y)] = \tilde{c}\{x \otimes y - y \otimes x\} \quad \text{with } \tilde{c} = c - b^2.$$

Now, for \tilde{m} in E_k , let us write $m_0 = -be_0 + \tilde{m}$. Then from (A.2) and (A.3) we get $V(m_0)e_i = \tilde{V}(\tilde{m})e_i = (\tilde{m} \otimes \tilde{m} + \tilde{B}(\tilde{m}) + \tilde{c} \text{id}_{E_k})e_i$. Therefore, $[V(m_0)]_e$ is reduced to a nonnull $k \times k$ diagonal block,

$$[V(m_0)]_e = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \tilde{V}(\tilde{m}) & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If $k = 1$, the equation $\tilde{V}(\tilde{m}) = \tilde{m}^2 + a_2\tilde{m} + \tilde{c} = 0$ admits a real root \tilde{m}_0 , since $\tilde{c} = b((a_1 - b) - b) < 0$ from (A.4). Hence there exists $m_0 = -be_0 + \tilde{m}_0$ such that $V(m_0) = 0$. If $k > 1$, we use the induction hypothesis and we conclude as before. \square

We now prove Lemma B.

PROOF OF LEMMA B. Under (A.1) axioms, we first prove the equivalence between the following two properties.

- (A.5) (i) There exists a nonnull vector u which is an eigenvector of $B(u)$.
 (ii) There exist two nonnull orthogonal vectors u, v such that v is an eigenvector of $B(u)$.

We will prove (A.5)(ii) later.

The implication (i) \Rightarrow (ii) is trivial, since $B(u)$, as a symmetric operator, is diagonalizable in a orthonormal basis of E . For the converse implication (ii) \Rightarrow (i), let us consider u as given in (A.5)(ii) and note $E = \bigoplus_{\lambda} \ker(B(u) - \lambda \text{id})$, the orthogonal decomposition of E into the eigenspaces $\text{Ker}(B(u) - \lambda \text{id})$ of $B(u)$.

Note first that if $\text{Ker}(B(u) - \lambda \text{id}) \cap u^{\perp} \neq \{0\}$, then $\text{Ker}(B(u) - \lambda \text{id}) \subset u^{\perp}$. Indeed, if $(v, w) \in (\text{Ker}(B(u) - \lambda \text{id}))^2$ and if $v \in u^{\perp}$, then $[B(u), B(v)]w = B(u)B(v)w - \lambda B(v)w = B(u)B(w)v - \lambda B(w)v = [B(u), B(w)]v$. Using (A.1)(ii) and $\langle u, v \rangle = 0$, we get $\langle u, w \rangle = 0$.

Let now E_1 and E_2 denote the sums of the eigenspaces of $B(u)$ which are, respectively, not orthogonal to u and orthogonal to u . Under the hypothesis (A.5)(ii), we have $E_2 \neq \{0\}$. Moreover, E_1 and E_2 are stable under the $B(x)$, $x \in E_1$. Indeed, for w in $\text{Ker}(B(u) - \lambda \text{id}) \subset E_2$ and for x in E_1 , using (A.1)(ii), we have $0 = [B(u), B(x)]w = (B(u) - \lambda \text{id})B(x)w$ so that $B(x)w \in \text{Ker}(B(u) - \lambda \text{id}) \subset E_2$. From the symmetry of $B(x)$, E_1 is also invariant.

Now, from (A.1)(ii), the restriction to E_2 of the $B(x)$, $x \in E_1$, commute. Consequently, there exists an eigenvector v in E_2 which is common to all $B(x)$, $x \in E_1$, and a vector a in E_1 such that

$$(A.6) \quad \forall x \in E_1, \quad B(x)v = \langle a, x \rangle v.$$

Now, since $B(x)x \in E_1$, the relation $[B(x), B(v)]x = -c\langle x, x \rangle v$ given by (A.1)(ii) yields $\langle a, x \rangle^2 - \langle a, B(x)x \rangle + c\langle x, x \rangle = 0$ and this is equivalent to

$$(A.6') \quad (a \otimes a - B(a) + c \text{id})|_{E_1} = 0.$$

Applying (A.6') to $x = a$, we finally get $B(a)a = (\langle a, a \rangle + c)a$. Thus (ii) \Rightarrow (i) is demonstrated.

We now prove the statement (A.5)(ii). Let us fix any unitary vector e_1 of E and choose (e_2, \dots, e_d) an orthonormal basis of e_1^{\perp} which diagonalizes the

restriction of $B(e_1)$ to e_1^\perp . In such a basis $e = (e_i)_{i=1}^d$, we write

$$(A.7) \quad [B(e_1)]_e = \begin{pmatrix} a_1 & c_2 & \cdots & c_d \\ c_2 & b_2 & & 0 \\ \vdots & & \ddots & \\ c_d & 0 & & b_d \end{pmatrix}.$$

If one c_i is 0, then $B(e_1)e_i = b_i e_i$: this is (A.5)(ii) with $u = e_1$ and $v = e_i$. If one b_i is 0, then $B(e_1)e_1 = c_i e_i$, which is (A.5)(ii), again with $u = e_i$ and $v = e_1$. If there exist i, j such that $i \neq j$ and $b_i = b_j$, then we still get (A.5)(ii) with $u = e_1$ and $v = c_j e_i - c_i e_j$. We now suppose that

$$(A.8) \quad c_1, \dots, c_d \neq 0, \quad b_1, \dots, b_d \neq 0 \quad \text{and for } i \neq j, \quad b_i \neq b_j.$$

Let us introduce $P(\lambda) = \det(B(e_1) - \lambda \text{id}_E)$, the characteristic polynomial of $B(e_1)$ and $D(\lambda)$, the matrix of cofactors of $B(e_1) - \lambda \text{id}_E$ in the basis e . We shall prove that there exists $u = \sum_{i=1}^d u_i e_i$ such that

$$(A.9) \quad \begin{aligned} & \text{(i) } B(e_1) - u_1 \text{id}_E \text{ is invertible.} \\ & \text{(ii) } (B(e_1) - u_1 \text{id}_E)u = ce_1. \end{aligned}$$

In this case, we will have $V(-u)e_1 = (u \otimes u - B(u) + c \text{id}_E)e_1 = 0$ and from (A.1), for $i \geq 2$,

$$(A.10) \quad (B(e_1) - u_1 \text{id}_E)V(-u)e_i = (B(e_i) - u_i \text{id}_E)V(-u)e_1 = 0,$$

so that $V(-u)e_i = 0$ from (A.9)(i). Consequently, for any v , $V(-u)v = 0$, and this gives (A.5)(i), since $B(u)u = (\langle u, u \rangle + c)u$, as well as (A.5)(ii), since $B(u)v = cv$ for any v orthogonal to u .

To get (A.9), consider the equation $(B(e_1) - u_1 \text{id}_E)u = ce_1$ and apply the matrix $D(u_1)$ to each of its members. We get

$$P(u_1)u = cD(u_1)e_1.$$

Taking the coordinate with respect to e_1 yields

$$-P(u_1)u_1 + c \langle D(u_1)e_1, e_1 \rangle = 0.$$

From (A.7), it is easy to see that

$$(A.11) \quad \begin{aligned} & \text{(i) } \langle D(\lambda)e_1, e_1 \rangle = \prod_{i=2}^d (b_i - \lambda), \\ & \text{(ii) } P(\lambda) = (a_1 - \lambda) \prod_{i=2}^d (b_i - \lambda) - \sum_{i=2}^d c_i^2 \prod_{k \neq i} (b_k - \lambda) \\ & \quad = \prod_{i=2}^d (b_i - \lambda) \cdot \left(a_1 - \lambda - \sum_{i=2}^d \frac{c_i^2}{b_i - \lambda} \right). \end{aligned}$$

Hence, u_1 is a root of the following polynomial of degree $d + 2$:

$$Q(\lambda) = -\lambda P(\lambda) + c \prod_{i=2}^d (b_i - \lambda),$$

which can also be written

$$(A.12) \quad Q(\lambda) = \prod_{i=2}^d (b_i - \lambda) \left(\lambda^2 - \alpha_1 \lambda + c - \sum_{i=2}^d c_i^2 - \sum_{i=2}^d \frac{c_i^2 b_i}{\lambda - b_i} \right).$$

It is now easy to prove that under the hypothesis (A.8) the rational function $\lambda \mapsto \lambda^2 - \alpha_1 \lambda + c - \sum_{i=2}^d c_i^2 - \sum_{i=2}^d (c_i^2 b_i)/(\lambda - b_i)$ has real roots in $\mathbb{R} \setminus \{b_2, \dots, b_d\}$. Let u_1 be such a root. Then u_1 is not an eigenvalue of $B(e_1)$. If not, $Q(u_1) = P(u_1) = 0$ and from (A.12), $u_1 \in \{b_2, \dots, b_d\}$, which is impossible from (A.11). Hence, $B(e_1) - u_1 \text{id}_E$ is invertible and $u = c(B(e_1) - u_1 \text{id}_E)^{-1} e_1$, which is well defined, satisfies (A.9). This concludes the proof of Lemma B. \square

APPENDIX B

Let E^* be a real linear space with dimension d , with dual space E and let $Q: E^* \rightarrow \mathcal{L}(E^*)$ be a linear operator such that

$$(B.1) \quad \begin{aligned} Q(\alpha)\beta &= Q(\beta)\alpha \\ Q(\alpha)Q(\beta) &= Q(\beta)Q(\alpha) \end{aligned}$$

We denote by \hat{E} , \hat{E}^* and $\hat{Q}: \hat{E}^* \rightarrow \mathcal{L}(\hat{E}^*)$, respectively, the complexified vector spaces of E, E^* and the complexified operator of Q . Then the set $Q = \{\hat{Q}(\alpha); \alpha \in E^*\}$ generates a nilpotent subalgebra of $\mathcal{L}(\hat{E}^*)$. By Theorem II, page 380 of Dynkin (1950), we can decompose \hat{E}^* into a direct sum of subspaces (\hat{F}_k) invariant under the $\hat{Q}(\alpha)$ and such that $\hat{Q}(\alpha)$ has only one (complex) eigenvalue on each \hat{F}_k . Since \hat{Q} is linear in α , there exist vectors d^j and $w_{s,t}^j$ of \hat{E} and a suitable basis of \hat{E}^* such that, in this basis, $\hat{Q}(\alpha)$ has the matrix

$$[\hat{Q}(\alpha)] = \begin{pmatrix} \hat{Q}_1(\alpha) & & 0 \\ & \ddots & \\ 0 & & \hat{Q}_r(\alpha) \end{pmatrix}$$

with

$$\hat{Q}_j(\alpha) = \begin{pmatrix} \langle \alpha, d^j \rangle & & \langle \alpha, w_{s,t}^j \rangle \\ & \ddots & \\ 0 & & \langle \alpha, d^j \rangle \end{pmatrix}.$$

Since $\hat{Q}(\alpha)$ is the complexified operator of $Q(\alpha)$, it is known that if $\langle \alpha, d^k \rangle$ is a real number, \hat{F}_k is the complexified characteristic subspace of F_k . In that case, the restriction of $Q(\alpha)$ to F_k can be written

$$\begin{pmatrix} \langle \alpha, d \rangle & & \langle \alpha, w_{st} \rangle \\ & \ddots & \\ 0 & & \langle \alpha, d \rangle \end{pmatrix}$$

with d, w_{st} in E as given in (3.7).

If $\langle \alpha, d^k \rangle$ is not real, the conjugate number $\overline{\langle \alpha, d^k \rangle} = \langle \alpha, \overline{d^k} \rangle$ for α in E^* is still an eigenvalue of $\hat{Q}(\alpha)$, say $\langle \alpha, d^j \rangle$. The corresponding characteristic subspaces \hat{F}_k and \hat{F}_j are conjugate, too, and hence have the same dimension. It is then possible to choose the basis $f_k = (f_{ks})_s$ and $f_j = (f_{js})_s$ of \hat{F}_k and \hat{F}_j to get conjugate submatrices $\hat{Q}_k(\alpha)$ and $\hat{Q}_j(\alpha)$. Choosing $e_{ks} = \frac{1}{2}(f_{ks} + f_{js})$ and $e_{js} = (1/2i)(f_{ks} - f_{js})$ yields a basis $(e_{k1}, e_{j1}, e_{k2}, e_{j2}, \dots, e_{ks}, e_{js})$ of the real subspace $F_k \oplus F_j$ in which the restriction of $Q(\alpha)$ can be written

$$\begin{pmatrix} D & & W_{st} \\ & \ddots & \\ 0 & & D \end{pmatrix}$$

for which there exists a, b, u_{st}, v_{st} in E such that

$$D = \begin{pmatrix} \langle \alpha, a \rangle & \langle \alpha, b \rangle \\ -\langle \alpha, b \rangle & \langle \alpha, a \rangle \end{pmatrix} \text{ and } W_{st} = \begin{pmatrix} \langle \alpha, u_{st} \rangle & \langle \alpha, v_{st} \rangle \\ -\langle \alpha, v_{st} \rangle & \langle \alpha, u_{st} \rangle \end{pmatrix}$$

as given in (3.8). \square

REFERENCES

BAR-LEV, S. K., BSHOUTY, D., ENIS, P., LETAC, G., LI LU, I. and RICHARDS, D. (1994). The diagonal natural exponential families in \mathbb{R}^2 and their classification. *J. Theoret. Probab.* **7** 883–929.

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

CASALIS, M. (1992a). Un estimateur de la variance pour une famille exponentielle naturelle à fonction-variance quadratique. *C. R. Acad. Sci. Paris Sér. I Math.* **314** 143–146.

CASALIS, M. (1992b). Les familles exponentielles sur \mathbb{R}^2 de fonction-variance $V(m) = am \otimes m + B(m) + C$. *C. R. Acad. Sci. Paris Sér. I Math.* **314** 635–638.

CONSONNI, G. and VERONESE, P. (1992). Conjugate priors for exponential families having quadratic variance functions. *J. Amer. Statist. Assoc.* **87** 1123–1127.

DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281.

DYNKIN, E. (1950). The structure of semi-simple Lie algebras. *Uspehi Mat. Nauk (N.S.)* **2** 59–127. *Am. Math. Soc. Transl.* **17** (1950).

FEINSILVER, P. (1986). Some classes of orthogonal polynomials associated with martingales. *Proc. Amer. Math. Soc.* **98** 298–302.

FEINSILVER, P. (1991). Orthogonal polynomials and coherent states. *Symmetries in Science* **5** 159–172. Plenum, New York.

GUTIERREZ-PEÑA, E. and SMITH, A. F. M. (1995). Conjugate parametrizations for natural exponential families. *J. Amer. Statist. Assoc.* **90** 1347–1356.

- HASSAÏRI, A. (1992). La classification des familles exponentielles naturelles sur \mathbb{R}^n par l'action du groupe linéaire de \mathbb{R}^{n+1} . *C. R. Acad. Sci. Paris Sér. I Math.* **315** 207–210.
- HASSAÏRI, A. (1994). Classification des familles exponentielles naturelles dans \mathbb{R}^d de variance cubique du type Mora. Doctorat d'Etat, Univ. Paul Sabatier, Toulouse.
- KOKONENDJI, C. (1993). Familles exponentielles naturelles réelles de fonction-variance en $R\Delta + Q\sqrt{\Delta}$. Doctorat de l'Univ. Paul Sabatier, Toulouse.
- KOKONENDJI, C. and SESHADRI, V. (1996). On the determinant of the second derivative of the Laplace transform. *Ann. Statist.* **24** 1813–1827.
- LETAC, G. (1989). Le problème de la classification des familles exponentielles naturelles de \mathbb{R}^d ayant une fonction-variance quadratique. *Probability Measures on Groups IX. Lecture Notes in Math.* **1306** 194–215. Springer, Berlin.
- LETAC, G. (1992). Lectures on natural exponential families and their variance-functions. *Mono-graf. Mat.* **5**.
- LETAC, G. and MORA, M. (1990). Natural real exponential families with cubic variance-functions. *Ann. Statist.* **18** 1–37.
- MEIXNER, J. (1934). Orthogonal Polynomsysteme mit einer besonderen Gestalt der erzeugenden Function. *J. London Math. Soc.* **9** 6–13.
- MORA, M. (1986). Familles exponentielles et fonctions-variance. Thèse de 3ème cycle de l'Univ. Paul Sabatier, Toulouse.
- MORRIS, C. N. (1982). Natural exponential families with quadratic variance-function. *Ann. Statist.* **10** 65–80.
- POMMERET, D. (1995). Polynômes orthogonaux associés aux familles exponentielles naturelles. Thèse de l'Univ. Paul Sabatier, Toulouse.
- RATNAPARKHI, M. V. (1985). Multinomial distributions. In *Encyclopedia of Statistical Sciences* **5** 659–665. Wiley, New York.
- SHANBHAG, D. N. (1979). Diagonality of the Bhattacharyya matrix as a characterization. *Theory Probab. Appl.* **24** 430–433.

LABORATOIRE DE STATISTIQUE ET PROBABILITÉS
CNRS-URA D0745
UNIVERSITÉ PAUL SABATIER
118, ROUTE DE NORBONNE
31062 TOULOUSE CEDEX
FRANCE