HAL open science

Associated natural exponential families and elliptic functions

Gérard Letac

To cite this version:

Gérard Letac. Associated natural exponential families and elliptic functions. Conference on Probability, statistics and their applications, Jun 2015, Aarhus, Denmark. hal-01977671

HAL Id: hal-01977671

https://hal.science/hal-01977671

Submitted on 10 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Associated natural exponential families and elliptic functions

Gérard Letac

To Ole Barndorff-Nielsen for his 80th birthday.

Abstract This paper studies the variance functions of the natural exponential families (NEF) on the real line of the form $(Am^4+Bm^2+C)^{1/2}$ where m denoting the mean. Surprisingly enough, most of them are discrete families concentrated on $\lambda\mathbb{Z}$ for some constant λ and the Laplace transform of their elements are expressed by elliptic functions. The concept of association of two NEF is an auxilliary tool for their study: two families F and G are associated if they are generated by symmetric probabilities and if the analytic continuations of their variance functions satisfy $V_F(m)=V_G(m\sqrt{-1})$. We give some properties of the association before its application to these elliptic NEF. The paper is completed by the study of NEF with variance functions $m(Cm^4+Bm^2+A)^{1/2}$. They are easier to study and they are concentrated on $a\mathbb{N}$.

Primary: 62E10 , 60G51
Secondary: 30E10

Key words: Variance functions, exponential dispersion models, function $\wp$ of Weierstrass.

1 Foreword

Ole and I met for the first time in the Summer School of Saint Flour in 1986. Having been converted to statistics by V. Seshadri two years before, I had learnt about exponential families through Ole's book (1978) and I had fought with cuts and steepness. Marianne Mora had just completed her thesis in Toulouse and was one of

Gérard Letac
Equipe de Statistique et Probabilités, Université de Toulouse,
118 route de Narbonne, 31062 Toulouse, France,
gerard.letac@math.univ-toulouse.fr

the Saint Flour participants: she was the first from Toulouse to make the pilgrimage to Aarhus the year after, followed by many others researchers from Toulouse: Evelyne Bernadac, Célestin Kokonendji, Dhafer Malouche, Muriel Casalis, Abdelhamid Hassairi, Angelo Koudou and myself. Over the years all of us were in contact with the everflowing ideas of Ole. During these Aarhus days (and Ole's visits to Toulouse) we gained a better understanding of the Lévy processes, of generalized inverse Gaussian distributions and their matrix versions, of differential geometry applied to statistics. Among all the topics which have interested Ole, the choice today is the one for which he may be the least enthusiastic (see the discussion of Letac 1991), namely the classification of exponential families through their variance functions: Ole thought correctly that although the results were satisfactory for the mind, one could not see much real practical applications: in the other hand Mendeleiev is universally admired for its prophetic views of chemical elements which had not been yet discovered. Descriptions of natural exponential families with more and more sophisticated variance functions V have been done: when V is a second degree polynomial in the mean (Morris 1982), a power (Tweedie 1984, Jørgensen 1987), a third degree polynomial (Letac-Mora 1990), the Babel class $P+Q\sqrt{R}$ where polynomials P,Q,R have degrees not bigger than $2,1,2$ (Letac 1992, Jørgensen 1997). This is for univariate NEF: even more important works have been done for multivariate NEF, but the present paper will confine to one dimensional distributions only.

Having forgotten variance functions during the last twenty years and having turned to random matrices and Bayesian theory, our interest for the topic has been rejuvenated by the paper by S. Bar-Lev and F. Van de Duyn Schouten (2004). The authors consider exponential dispersion models G such that one of the transformations

$$T(P)(dx)=\frac{xP(dx)}{\int_{\mathbb{R}}xP(dx)},\ \ T^2(P)(dx)=\frac{x^2P(dx)}{\int_{\mathbb{R}}x^2P(dx)}.$$

maps G into itself or one of its translates. For $T(P)$ they obtain exactly the exponential dispersion models concentrated on the positive line with quadratic variance functions: gamma, Poisson, negative binomial and binomial and no others. For $T^2(P)$ they obviously obtain the previous ones, but they observe that new variance functions appear, in particular $(m^4+c^2)^{1/2}$, without being able to decide whether these natural exponential families (NEF) exist after all (it should be mentioned here that their formula (11) is not correct and this fact greatly invalidates their paper). As a result, our initial motivation was to address this particular question: is $(m^4+c^2)^{1/2}$ a variance function? As we shall see, the answer is yes for a discrete set of c. To solve this particular problem, we have to design methods based on elliptic functions, and these methods appear to have a wider domain of applicability. For this reason, the aim of the present paper is the classification of the variance functions of the form $(Am^4+Bm^2+C)^{1/2}$ and their reciprocals in the sense of Letac-Mora (1990), namely the variance functions of the form $m(Cm^4+Bm^2+A)^{1/2}$.

Section 2 recalls general facts and methods for dealing with NEF's. Section 3 opens a long parenthesis on pairs of associated NEF: if F and G are NEF generated by symmetric probabilities, we say that they are associated roughly if we can write

$V_F(m) = f(m^2)$ and $V_G(m) = f(-m^2)$. This definition seems to be a mere curiosity of distribution theory, but appears to be illuminating when applied to our elliptic NEF. Since this concept of association has several interesting aspects, I have provided here several detailed examples (Section 3.1) that the reader should skip if he is only interested in elliptic NEF.

Section 4 rules out trivial values of the parameters for elliptic NEF. Sections 5-7 investigate the various cases according to the parameters (A,B,C), Section 8 considers the reciprocal families of the previous elliptic NEF: these reciprocal NEF are interesting distributions of the positive integers. Section 9 makes brief comments on the variance functions $(\alpha m+\beta)\sqrt{P(m)}$ where P is an arbitrary polynomial of degree ≤ 4, actually a complete new field of research. While the statements of the present paper are understandable without knowledge of elliptic functions, the proofs of Sections 5-7 make heavy use of them, and we shall constantly refer to the magnificent book by Sansone and Gerretsen (1960) that we frequently quote by SG.

2 Retrieving an NEF from its variance function

The concept of exponential family is obviously the backbone of Ole's book (Barndorff -Nielsen (1978)) or of Brown (1986), but the notations for the simpler object called natural exponential family are rather to be found in Morris (1982), Jørgensen (1987) and Letac and Mora (1990).

If μ is a positive non Dirac Radon measure on $\mathbb{R}$ consider its Laplace transform

$$L_\mu(\theta) = \int_{-\infty}^{\infty} e^{\theta x}\mu(dx) \leq \infty.$$

Assume that the interior $\Theta(\mu)$ of the interval $D(\mu) = \{\theta \in \mathbb{R}; L_\mu(\theta) < \infty\}$ is not empty. The set of positive measures μ on $\mathbb{R}$ such that $\Theta(\mu)$ is not empty and such that μ is not concentrated on one point is denoted by $\mathscr{M}(\mathbb{R})$. We denote by $\mathscr{M}_1(\mathbb{R})$ the set of probabilities μ contained in $\mathscr{M}(\mathbb{R})$. In other terms, the elements of $\mathscr{M}_1(\mathbb{R})$ are the probability laws on $\mathbb{R}$ which have a non trivial Laplace transform.

Write $k_\mu = \log L_\mu$. Then the family of probabilities

$$F = F(\mu) = \{P(\theta,\mu)\ ;\ \theta \in \Theta(\mu)\}$$

where

$$P(\theta,\mu)(dx) = e^{\theta x - k_\mu(\theta)}\mu(dx)$$

is called the natural exponential family generated by μ. Two basic results are

$$m = k'_\mu(\theta) = \int_{-\infty}^{\infty} xP(\theta,\mu)(dx)$$

and the fact that k'_μ is increasing (or that k_μ is convex). The set $k'_\mu(\Theta(\mu)) = M_F$ is called the domain of the means. We denote by $\psi_\mu : M_F \to \Theta(\mu)$ the reciprocal

function of $m = k'_\mu(\theta)$. Thus $F = F(\mu)$ can be parametrized by M_F by the map from M_F to F which is

$$m \mapsto P(\psi_\mu(m), \mu) = P(m, F).$$

In other terms an element of $F(\mu)$ can be identified by the value of the mean m rather than by the value of the canonical parameter θ. One can prove that the variance $V_F(m)$ of $P(m,F)$ is

$$V_F(m) = k''_\mu(\psi_\mu(m)) = \frac{1}{\psi'_\mu(m)}. \tag{1}$$

The map $m \mapsto V_F(m)$ from M_F to $(0,\infty)$ is called the variance function and characterizes F. The Jørgensen set of μ is the set $\Lambda(\mu)$ of positive numbers t such that there exists a positive measure μ_t such that $\Theta(\mu_t) = \Theta(\mu)$ and such that $L_{\mu_t} = (L_\mu)^t$. Obviously, $\Lambda(\mu)$ is an additive semigroup which contains all positive integers. If $t \in \Lambda(\mu)$ we denote $F_t = F(\mu_t)$ and it is easily checked that $M_{F_t} = tM_F$ and that

$$V_{F_t}(m) = tV_F\left(\frac{m}{t}\right). \tag{2}$$

The union $G = G(\mu) = \cup_{t\in\Lambda(\mu)} F(\mu_t)$ is called the exponential dispersion model generated by μ.

If F is a NEF and if $h(x) = ax + b$ (with $a \neq 0$) then the family $h(F)$ of images of elements of F by h is still a NEF with $M_{h(F)} = h(M_F)$ and

$$V_{h(F)}(m) = a^2 V_F\left(\frac{m-b}{a}\right). \tag{3}$$

In spite of the similarity between (2) and (3), the last formula is much more useful for dealing with a NEF F which is known only by its variance function: the reason is that the Jørgensen set $\Lambda(F)$ of F is unknown in many circumstances. In fact $\Lambda(F)$ is a closed additive semi group of $[0,\infty)$ which can be rather complicated (see Letac, Malouche and Maurer (2002) for an example). In the other hand an affinity is always defined, and this fact can be use to diminish the number of parameters of a family of variance functions. For instance, if we consider the variance function $\sqrt{Am^4 + Bm^2 + C}$ such that $C > 0$, without loss of generality we could assume that $C = 1$ by using the dilation $x \mapsto x/\sqrt{C}$.

An important fact for the sequel is that V_F is real analytic, that means that for any $m_0 \in M_F$ there exists a positive number r such that for $m_0 - r < m < m_0 + r$ we have

$$V_F(m) = \sum_{n=0}^{\infty} \frac{(m-m_0)^n}{n!} V_F^{(n)}(m_0)$$

which implies that V_F is analytically extendable to a connected open set of the complex plane containing the real segment M_F. If $\mu \in \mathscr{M}_1(\mathbb{R})$, the Laplace transform L_μ defined on the open interval $\Theta(\mu)$ is extendable analytically in a unique way to the strip $\Theta(\mu) + i\mathbb{R}$ of the complex plane. This extension is also denoted L_μ and $\theta \mapsto L_\mu(i\theta)$ is the Fourier transform of the probability μ. The function

$k_\mu(\theta) = \log L_\mu(\theta)$ could be also extendable to an analytic function on the same strip, but it would be a multivalued function if $L_\mu(\theta)$ has zeros in the strip.

To conclude this section, recall the four steps allowing us to pass from the variance function V_F of a NEF F to a measure μ such that $F = F(\mu)$.

1. Writing $d\theta = \psi'_\mu(m)dm = \frac{dm}{V_F(m)}$, we compute $\theta = \psi_\mu(m)$ as a function of m by a quadrature;
2. we deduce from this the parameter m as a function $m = k'_\mu(\theta)$ (this is generally a difficult point);
3. we compute $k_\mu(\theta)$ by a second quadrature and obtain $L_\mu = e^{k_\mu}$;
4. we use dictionary, creativity, or inversion Fourier formulas to retrieve μ from its Laplace transform.

We keep these four steps in mind for dealing with $V_F(m) = \sqrt{Am^4 + Bm^2 + C}$ in the sequel. It is worthwhile to sketch here an example with

$$V(m) = \sqrt{1+4m^4}.$$

For $0 < w < 1$ we do the change of variable and perform the first step:

$$m = \frac{\sqrt{1-w^4}}{2w},\ w^2 = -2m^2 + \sqrt{1+4m^4},\ d\theta = \frac{dw}{\sqrt{1-w^4}},\ \theta = \int_{w(m)}^1 \frac{dw}{\sqrt{1-w^4}}.$$

The second step introduces a function $C(\theta)$ defined on the interval $[0,K]$ where $K = \int_0^1 \frac{dw}{\sqrt{1-w^4}} = 1.3098..$ as

$$\theta = \int_{C(\theta))}^1 \frac{dw}{\sqrt{1-w^4}}. \tag{4}$$

Since $w(m) = C(\theta)$, up to the knowledge of $C(\theta)$, and taking derivative of both sides of (4), the second step is performed since

$$m = k'(\theta) = \frac{\sqrt{1-C(\theta)^4}}{2C(\theta)} = -\frac{C'(\theta)}{2C(\theta)}.$$

The third step is easy and we get $k(\theta) = -\frac{1}{2}\log C(\theta)$ and the Laplace transform $L(\theta) = \frac{1}{\sqrt{C(\theta)}}$. The fourth step needs to be explicit about $C(\theta)$ and the theory of elliptic functions becomes necessary: details about this particular example are in Theorem 4.1 when doing $k^2 = -1$. The function L will be the Laplace transform of a discrete distribution concentrated on set on numbers of the form n/a where n is a relative integer and where a is the complicated number $\frac{2K}{\pi} = 0.8338...$ If we use formula (3) we get the following surprizing result: the function $\sqrt{a^4+4m^4}$ is the variance function of a NEF concentrated on $\mathbb{Z}$.

3 Associated natural exponential families

The source of this concept is the pair of identities (5) below: if

$$\mu(dx) = \frac{dx}{2\cosh\frac{\pi x}{2}}$$

and $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$ is the symmetric Bernoulli distribution then

$$\int_{-\infty}^{+\infty} e^{\theta x}\mu(dx) = \frac{1}{\cos\theta} \quad (\text{for } |\theta| < \pi/2), \quad \int_{-\infty}^{+\infty} e^{i\theta x}\nu(dx) = \cos\theta, \tag{5}$$

which could be as well presented by reversing the roles of Fourier and Laplace transforms:

$$\int_{-\infty}^{+\infty} e^{i\theta x}\mu(dx) = \frac{1}{\cosh\theta}, \int_{-\infty}^{+\infty} e^{\theta x}\nu(dx) = \cosh\theta. \tag{6}$$

This is an example of what we are going to call an associated pair (μ,ν) of probabilities on $\mathbb{R}$. Here is the definition:

Definition 3.1 Let μ and ν be in $\mathscr{M}_1(\mathbb{R})$ such that μ and ν are symmetric. We say that (μ,ν) *is an associated pair* if for all $\theta \in \Theta(\mu)$ the Fourier transform of ν is $1/L_\mu(\theta)$. In other terms for $\theta \in \Theta(\mu)$ we have

$$\int_{-\infty}^{+\infty} e^{\theta x}\mu(dx) = L_\mu(\theta), \int_{-\infty}^{+\infty} e^{i\theta x}\nu(dx) = \frac{1}{L_\mu(\theta)}. \tag{7}$$

The corresponding natural exponential families $F = F(\mu)$ and $G = F(\nu)$ are also said to be associated.

We describe now the easy consequences of this definition:

Proposition 3.1 Let (μ,ν) in $\mathscr{M}_1(\mathbb{R})$ be an associated pair. Then

1. (Symmetry) The pair (ν,μ) is also associated;
2. (Uniqueness) If (μ,ν_1) is also associated, then $\nu_1 = \nu$.
3. (Convolution) If (μ',ν') is an associated pair then $(\mu * \mu', \nu * \nu')$ is also an associated pair.
4. (Zeros) Denote $z_\mu = \inf\{\theta > 0\ ;\ L_\mu(i\theta) = 0\}$. Then $\Theta(\nu) = (-z_\mu, z_\mu)$.
5. (Variance functions) Consider the associated pair $F = F(\mu)$ and $G = F(\nu)$ of NEF. If V_F and V_G are extended as analytic functions to the complex plane in a neighborhood of zero, then $V_F(m) = V_G(im)$.

Comments

1. Clearly, since the Fourier transform $\frac{1}{L_\mu(\theta)}$ is real, the probability ν must be symmetric.
2. Symmetry of ν implies that $\Theta(\nu)$ is a symmetric interval, as well as the mean domain of $F(\nu)$.
3. Because of the uniqueness of Part 2, we shall also write μ^* for indicating that (μ,μ^*) is an associated pair. In this case μ^* is called the associated probability to μ (when it exists). We also observe that

$$(\mu^*)^* = \mu, \;\; (\mu * \mu')^* = \mu^* * (\mu')^*.$$

4. It is not correct to think that if μ is in $\mathscr{M}_1(\mathbb{R})$ then μ^* always exists. An example is given by the first Laplace distribution (also called the bilateral exponential distribution)

$$\mu(dx) = \frac{1}{2}e^{-|x|}dx, \; \Theta(\mu) = (-1,1), \; L_\mu(\theta) = \frac{1}{1-\theta^2}.$$

Suppose that $\nu = \mu^*$ exists. Then its Fourier transform on $(-1,1)$ is $1-\theta^2$. This implies that its Laplace transform is $L_\nu(\theta) = 1+\theta^2$ and therefore $\Theta(\nu) = \mathbb{R}$. But if $k_\nu = \log L_\mu$ it is easy to see that the sign of $k_\nu''(\theta)$ is the sign of $1-\theta^2$, which implies that k_ν is not convex, a contradiction.
A more complicated example is given by

$$\mu_t(dx) = \frac{2^{t-2}}{\pi\Gamma(t)}\left|\Gamma(\frac{t+ix}{2})\right|^2 dx \tag{8}$$

for $t>0$. We will see in Section 3.1 that $\mu_t \in \mathscr{M}_1(\mathbb{R})$ satisfies $\Theta(\mu_t) = (-\frac{\pi}{2},\frac{\pi}{2})$ and

$$\frac{2^{t-2}}{\pi\Gamma(t)}\int_{-\infty}^{+\infty} e^{\theta x}|\Gamma(\frac{t+ix}{2})|^2 dx = \frac{1}{(\cos\theta)^t}. \tag{9}$$

If t is not an integer, then $\mu_t^* = \nu$ does not exist (Proposition 3.2). An obvious case is $t = 1/2$: if X,Y are iid such that $\Pr(X+Y=\pm 1) = 1/2$ then $\Pr(X = \pm 1/2) = 1/4$ and $\Pr(X+Y=0) \geq 1/16 > 0$, a contradiction.
5. In Definition 3.1, suppose that we relax the constraint on ν to have a Laplace transform. Consider the example

$$\mu(dx) = \frac{dx}{2\cosh(\pi x)/2}$$

with Laplace transform $1/\cos\theta$ defined on $\Theta(\mu) = (-\frac{\pi}{2},\frac{\pi}{2})$. A possible associated ν is the Bernoulli $\frac{1}{2}(\delta_{-1}+\delta_1)$ which satisfies $\int_{-\infty}^{+\infty} e^{i\theta x}\nu(dx) = \cos\theta$ in particular on $|\theta| < \pi/2$) However it is not excluded that there exists other probabilities ν fulfilling the same property on $|\theta| < \pi/2$). Imposing $\nu \in \mathscr{M}_1(\mathbb{R})$ rules out this phenomenon, from Part 2 of Proposition 3.1.

6. Here is the simplest example illustrating Part 5 of Proposition 3.1. We use once more the associated pair (5). In this case $M_F = \mathbb{R}$, $a_F = \infty$, $V_F(m) = 1 + m^2$, $M_G = (-1,1)$, $a_G = 1$, $V_G(m) = 1 - m^2$. See Proposition 3.5 below.
7. SELF ASSOCIATED PAIRS AND NEF: A trivial example is $\mu = N(0,1)$ since $\mu = \mu^*$. More generally, V_F is a function of m^4 if and only if $\mu = \mu^*$. An other important example will be found in Theorem 5.1 below, which is $V_F(m) = \sqrt{1+4m^4}$. Note that the symmetry of μ is essential: if $V_F(m) = m^4$, with $M_F = (0,\infty)$, we have $V_F(m) = V_F(im)$ but the concept of association does not make sense here.
8. This Part 5 provides also a way to decide quickly from the examination of the variance function that μ^* does not exist. For instance, if $\mu \sim X - Y$ where X and Y are iid with the Poisson distribution of mean 1, then $F = F(\mu)$ has variance function $V_F(m) = \sqrt{1+m^2}$. Would $G = F(\mu^*)$ exist, its variance function would be $V_G(m) = \sqrt{1-m^2}$. The domain of the mean of G would be $(-1,1)$, from the principle of analytic continuation of variance functions (Theorem 3.1 in Letac and Mora (1990)). However on around the point $m = 1$, the function V_G would be equivalent to $2(1-m)^{1/2}$. This is forbidden by the principle of Jørgensen, Martinez and Tsao (1993): this principle says that if $M_G = (a,b)$ with $b < \infty$ and if
$$V_G(m) \sim_{m\to b} A \times (b-m)^p \tag{10}$$
then $p \notin (0,1)$.
Similarly consider the variance function $V_F(m) = (1+m^2)^{3/2}$ defined on $M_F = \mathbb{R}$. One can consult Letac(1991) chapter 5 example 1.2 for a probabilistic interpretation. It is generated by a μ such that $\Theta(\mu) = (-1,1)$ and $k_\mu = \sqrt{1-\theta^2} - 1$. For seeing that $V_G(m) = (1-m^2)^{3/2}$ cannot be a variance function we observe the following. If $\nu = \mu^*$ exists then
$$k_\nu(\theta) = \sqrt{1+\theta^2} - 1.$$
Therefore, by using the principle of maximal analytic continuation (see Proposition 3.2 below), we have $\Theta(\nu) = \mathbb{R}$. As a consequence $L_\nu(\theta) = e^{\sqrt{1+\theta^2}-1}$ is an entire function, which is clearly impossible.

Proof of Proposition 3.1. 1) Suppose that ν is in $\mathscr{M}_1(\mathbb{R})$. Then the knowledge of the Fourier transform of ν on the interval $\Theta(\mu)$ gives the knowledge of the Laplace transform L_ν on $\Theta(\nu)$. Now the Fourier transform of μ restricted to $\Theta(\nu)$ is $L_\mu(i\theta) = 1/L_\nu(i\theta)$ from the relation (7) extended by analyticity.

2) If ν_1 exists, its Fourier transform coincides with the Fourier transform of ν on the interval $\Theta(\mu)$. By analyticity, the two coincide everywhere and $\nu = \nu_1$.

3) is obvious.

4) Since the Fourier transform of ν restricted to $\Theta(\mu)$ is $1/L_\mu(\theta)$ then in a neighborhood of $\theta = 0$, the Laplace transform of ν satisfies $L_\nu(\theta) = 1/L_\mu(i\theta)$. Now we use the following result:

Proposition 3.2. (Principle of maximal analyticity) If $\nu \in \mathscr{M}(\mathbb{R})$ and if $\Theta(\nu) = (a,b)$ suppose that there exists $(a_1,b_1) \supset (a,b)$ and a real analytic function f on (a_1,b_1) which is strictly positive and such that $f(\theta) = L_\nu(\theta)$ for $a < \theta < b$. Then $a = a_1$ and $b = b_1$.

Proof. Use the method of proof of Theorem 3.1 of Letac and Mora (1990) or Kawata (1972), chapter 7.

We now return to the proof of Proposition 3.1, Part 4). Write $\Theta(\nu) = (-b,b)$. Clearly $b > z_\mu$ is impossible since it would imply that $L_\nu(z_\mu)$ would be finite, a contradiction with $L_\mu(z_\mu) = 0$. We apply Proposition 2.2 to the present ν, to $(a_1,b_1) = (-z_\mu, z_\mu)$ and to the positive analytic function on this interval $f(\theta) = 1/L_\mu(\theta)$. As a consequence $b = b_1 = z_\mu$ and the result 4) is proved.

5) Consider the functions L_μ and L_ν. They are analytic on the strips $\Theta(\mu) \times i\mathbb{R}$ and $\Theta(\nu) \times i\mathbb{R}$, and from Part 4) $\Theta(\mu) + i\Theta(\nu)$ is the open square with vertices $\pm z_\nu \pm i z_\nu$. Let Z be the set of zeros of the analytic function $\theta \mapsto L_\mu(\theta)$ restricted to the square $\Theta(\mu) + i\Theta(\nu)$. From the principle of isolated zeros, Z contains only a finite number of points in the compact set $[-a,a] \times [-b,b]$ when $a < z_\nu$ and $b < z\mu$. Also Z has no zeros on the set $S = (-z_\nu, z_\nu) \cup (-iz_\mu, iz_\mu)$. Consider now the part Z_{++} contained in the first quadrant, and its closed convex hull C_{++}. Similarly consider $C_{\pm,\pm}$, the closed set $C = C_{++} \cup C_{+-} \cup C- + \cup C_{--}$ and the open set $U = \Theta(\mu) + i\Theta(\nu) \setminus C$. Then U is a simply connected set and is a neighborhood of S.

We are in position to define $\log L_\mu = k_\mu$ on the open set U as an analytic function. On this set U we have

$$k_\mu(\theta) = -k_\nu(i\theta),\ k'_\mu(\theta) = -ik_\nu(i\theta),\ k''_\mu(\theta) = k''_\nu(i\theta). \tag{11}$$

Since

$$V_F(k'_\mu(\theta)) = k''_\mu(\theta),\ V_G(k'_\nu(\theta)) = k''_\nu(\theta)$$

we get finally

$$V_G(ik'_\mu(i\theta)) = k''_\mu(i\theta)$$

and this is saying that for m in the open set $k'_\mu(U)$ we have $V_F(m) = V_G(im)$, which is the desired result.

Proposition 3.3. (Convolution of Bernoulli's). Let $(a_n)_{n=1}^\infty$ be a real sequence such that $\sum_{n=1}^\infty a_n^2 < \infty$. Let $(X_n)_{n=1}^\infty$ and $(Y_n)_{n=1}^\infty$ be two iid sequences such that

$$X_n \sim \frac{1}{2\cosh(\pi x/2)},\ \ Y_n \sim \frac{1}{2}(\delta_{-1} + \delta_1).$$

Then the distributions μ of $\sum_{n=1}^\infty a_n X_n$ and ν of $\sum_{n=1}^\infty a_n Y_n$ are associated.

Proof. Easy, from (5) and Part 3) of Proposition 2.1. Note that for $a_n = 1/3^n$ then ν is the purely singular Cantor distribution on $(-1/2, 1/2)$, while μ has a density.

3.1 Examples of associated probabilities

Here are 3 groups of examples. It can be observed that they offer three different generalizations of (5). We start with the classical formula for $t>0$ correct for $\theta \in (-t,t)$:

$$\int_{-\infty}^{+\infty} e^{x\theta}\frac{dx}{(\cosh x)^t} = \frac{2^{t-1}}{\Gamma(t)}\Gamma(\frac{t+\theta}{2})\Gamma(\frac{t-\theta}{2}) \tag{12}$$

with $\Theta(\mu_t)=(-t,t)$. In particular using the duplication formula $\sqrt{\pi}\Gamma(t)=2^{t-1}\Gamma(\frac{t}{2})\Gamma(\frac{t+1}{2})$ we get the Laplace transform of the probability α_t below:

$$\alpha_t(dx)=\frac{\Gamma(\frac{t+1}{2})}{\sqrt{\pi}\Gamma(\frac{t}{2})}\times\frac{dx}{(\cosh x)^t},\quad L_{\alpha_t}(\theta)=\frac{1}{\Gamma(\frac{t}{2})^2}\times\Gamma(\frac{t+\theta}{2})\Gamma(\frac{t-\theta}{2}) \tag{13}$$

with $\Theta(\alpha_t)=(-t,t)$. It is worthwhile to mention that if X and Y are iid with distribution

$$\beta(\frac{t}{2},1)(dx)=\frac{t}{2}x^{\frac{t}{2}-1}\mathbf{1}_{(0,1)}(x)dx$$

and if $U=\sqrt{X/Y}$ then $\log U\sim\alpha_t$.

Formula (12) is easily proven by the change of variable $u=e^{2x}$ and the formula $\int_0^\infty \frac{u^{p-1}du}{(1+u)^{p+q}}=B(p,q)$ for $p,q>0$. The Fourier version of (12) is

$$\int_{-\infty}^{+\infty} e^{ix\theta}\frac{dx}{(\cosh x)^t} = \frac{2^{t-1}}{\Gamma(t)}\left|\Gamma(\frac{t+i\theta}{2})\right|^2 \tag{14}$$

leading by Fourier inversion to

$$\frac{2^{t-1}}{2\pi\Gamma(t)}\int_{-\infty}^{+\infty} e^{i\theta x}\left|\Gamma(\frac{t+ix}{2})\right|^2 dx=\frac{1}{(\cosh\theta)^t} \tag{15}$$

and by analyticity to (9). For a while, let us specialize these formulas to $t=2p-1$ and to $t=2p$ where p is a positive integer. From the complements formula $\Gamma(z)\Gamma(1-z)=\pi/\sin(\pi z)$ and $\Gamma(z+1)=z\Gamma(z)$ we have for $t=1,2$

$$\Gamma(\frac{1+\theta}{2})\Gamma(\frac{1-\theta}{2})=\frac{\pi}{\cos\frac{\pi\theta}{2}},\quad \Gamma(1+\frac{\theta}{2})\Gamma(1-\frac{\theta}{2})=\frac{\pi\theta}{2\sin\frac{\pi\theta}{2}}$$

and more generally

$$\Gamma(\frac{2p+1+\theta}{2})\Gamma(\frac{2p+1-\theta}{2}) = \frac{1}{2^p}(1-\theta^2)(9-\theta^2)\dots((2p-1)^2-\theta^2)\times\frac{\pi}{\cos\frac{\pi\theta}{2}} \tag{16}$$

$$\Gamma(p+\frac{\theta}{2})\Gamma(p-\frac{\theta}{2}) = \frac{1}{2^p}(4-\theta^2)(16-\theta^2)\dots(4p^2-\theta^2)\times\frac{\pi\theta}{2\sin\frac{\pi\theta}{2}} \tag{17}$$

Proposition 3.4. If α_t is defined by (13) then α_t^* exists if and only if $t\geq 1$. In

particular $\alpha_1^* = \frac{1}{2}(\delta_{-\pi/2} + \delta_{\pi/2})$ is a Bernoulli distribution and for $t > 1$ we have

$$\alpha_t^*(dx) = \frac{\Gamma(t)}{\sqrt{\pi}\Gamma(\frac{t-1}{2})}(\cos x)^{t-2} 1_{(-\pi/2,\pi/2)}(x)dx.$$

In particular for $t = 2p+1$ and $t = 2p+2$ where p is an non negative integer, then (16) and (17) give $(\varphi_t)^{-1}$ when φ_t is the Fourier transform of α_t^*.

Comments. For this example, the explicit calculation of the variance functions of $F = F(\alpha_t)$ and $G = F(\alpha_t^*)$ is not possible. For instance if $t = 2$ the probability α_2^* is the uniform distribution on the segment $(-\pi/2, \pi/2)$. In this case $L_{\alpha_2^*}(\theta) = \frac{\sinh(\pi t/2)}{\pi t/2}$: no way to compute $\theta = \psi_{\alpha_2^*}(m)$ in a close formula when

$$m = k'_{\alpha_2^*}(\theta) = \frac{\pi}{2}\left(\text{cotanh}(\frac{\pi\theta}{2}) - \frac{2}{\pi\theta}\right).$$

Shanbhag (1979) and, in their Proposition 4, Barlev and Letac (2012), have other proofs of the 'only if' condition of existence of α_t^*.

Proof. For $t > 1$ we just rely on entry 3. 631, 9 of Gradshteyn and Ryzhik (2007). If $t < 1$ we show that α_t^* does not exist by showing that $\theta \mapsto k''_{\alpha_t}(i\theta)$ is not positive. We obtain

$$k''_{\alpha_t}(i\theta) = \sum_{n=0}^{\infty} \frac{(n+\frac{t}{2})^2 - \frac{\theta^2}{4}}{[(n+\frac{t}{2})^2 + \frac{\theta^2}{4}]^2}.$$

and a careful calculation shows that

$$\lim_{\theta\to\infty} \theta^2 k''_{\alpha_t}(i\theta) = 2(t-1)$$

If $t < 1$ then $\theta \mapsto k''_{\alpha_t}(i\theta)$ cannot be positive for all $\theta \in \mathbb{R}$, and this ends the proof.

Proposition 3.5. If μ_t is defined by (9) then μ_t^* exists if and only if t is a positive integer N. In this case μ_N^* is the image of the binomial distribution $B(N, 1/2)$ by $x \mapsto 2x - N$.

Comments. The most interesting particular case corresponds to $t = 2$ since in this case we meet the uniform distribution on a segment with the associated pair

$$\mu_2(dx) = \frac{x}{4\sinh(\pi x/2)}dx,\ (\mu_2)^*(dy) = \frac{1}{2}1_{(-1,1)}(y)dy.$$

This is also an illustration of Proposition 2.3 applied to $a_n = 1/2^n$ since $\sum_{n=1}^{\infty} \frac{Y_n}{2^n}$ is uniform on (-1,1) when $(Y_n)_{n=0}^{\infty}$ is an iid sequence of symmetric Bernoulli random variables. For this example, the explicit calculation of the variance functions of $F = F(\mu_t)$ and $G = F(\mu_t^*)$ gives

$$V_F(m) = t + \frac{m^2}{t},\ V_G(m) = N - \frac{m^2}{N}.$$

Proof of Proposition 3.5. $\Leftarrow$ is obvious. To prove $\Rightarrow$ suppose that there exists a positive integer n_0 such that $n_0 - 1 < t < n_0$ and suppose that μ_t^* exists. Taking the image τ of μ_t^* by the map $x \mapsto x' = x - t$, choosing $\theta > 0$ and denoting $z = e^{-2\theta}$ we get

$$\int_{-\infty}^{+\infty} e^{\theta x'} \tau(dx') = \int_{-\infty}^{+\infty} e^{\theta(x-t)} \mu_t^*(dx) = \frac{1}{2^t} \sum_{n=0}^{\infty} \frac{t(t-1)\dots(t-n+1)}{n!} z^n.$$

Since $t(t-1)\dots(t-n+1) < 0$ when $n = n_0 + 1$ this shows that $\tau(\{-2n_0 - 2\}) < 0$, a contradiction.

The third example is obtained by considering the Babel class of NEF, namely the set of exponential families such that the variance function has the form $V_F = P\Delta + Q\sqrt{\Delta}$ where Δ, P and Q are polynomials with respective degrees less or equal to $2, 1, 2$. Looking for possible pairs (F, G) in this class such that $V_F(m) = V_G(im)$ and such that F and G are generated by associated distributions (μ, ν) -and therefore symmetric- implies that $\Delta(m) = Am^2 + C$, P is a constant and $Q(m) = A'm^2 + C'$. The case $C = 0$ is excluded since the domain of the mean M_F and M_G are symmetric interval and V_F and V_G are real analytic on them. As a consequence either F or G must be such that $\Delta(m) = 1 - m^2$ (up to affinities). But there is only one type of NEF in the Babel class such that $\Delta(m) = 1 - m^2$ and it is generated by the trinomial distributions defined for $0 < a < 1$ by

$$\mu_a = \frac{1}{a+1}(a\delta_0 + \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1) \tag{18}$$

and their entire powers of convolution. Of course the limit cases are related to Bernoulli, since

$$\mu_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1,\ \mu_1 = (\frac{1}{2}\delta_{-1/2} + \frac{1}{2}\delta_{1/2}) * (\frac{1}{2}\delta_{-1/2} + \frac{1}{2}\delta_{1/2}).$$

Proposition 3.6. If μ_a is defined by (18) with $a \in (0,1)$ then μ_a^* exists and is

$$\mu_a^* = \tau_b * \tau_{-b}.$$

where $a = \cos 2b$ with $0 < b < \pi/4$ and

$$\tau_{\pm b}(x) = \frac{\cos b}{\cosh \frac{\pi x}{4}} e^{\pm bx} dx.$$

Proof. We have

$$L_{\mu_a}(\theta) = \frac{a+\cosh\theta}{a+1},\ V_{F(\mu_a)} = \frac{1}{1-a^2} - m^2 - \frac{a}{\sqrt{1-a^2}}\sqrt{\frac{1}{1-a^2} - m^2}.$$

Therefore, if μ_a^* does exist it must satisfy

$$L_{\mu_a^*}(\theta) = \frac{a+1}{a+\cos\theta},\ V_{F(\mu_a^*)} = \frac{1}{1-a^2} + m^2 - \frac{a}{\sqrt{1-a^2}}\sqrt{\frac{1}{1-a^2} + m^2}$$

with $\Theta(\mu_a)^* = (-z_{\mu_a}, z_{\mu_a})$ where z_{μ_a} is the smallest positive solution of $\cos\theta = -a$. Such a μ_a^* actually exists. To see this we write $a = \cos 2b$ with $0 < b < \pi/4$ and by simple trigonometry and the help of formula (6):

$$\frac{\cos 2b + 1}{\cos 2b + \cos\theta} = \frac{\cos b}{\cos(\frac{\theta}{2} - b)} \times \frac{\cos b}{\cos(\frac{\theta}{2} + b)} = L_{\tau_b}(\theta) L_{\tau_{-b}}(\theta)$$

where

$$\tau_{\pm b}(x) = \frac{\cos b}{\cosh\frac{\pi x}{4}} e^{\pm bx} dx.$$

4 Discussion and easy cases for $(Am^4 + Bm^2 + C)^{1/2}$

In this section we recall known and not so well known results about a few particular cases. The cases where only one of the three numbers A, B, C is not zero are classical: we get respectively the gamma, Poisson or normal case. We now investigate three more interesting particular cases (they are all described in Letac 1992 as elements of the Babel class).

4.1 The case $A = 0$

The useful results are contained in the following proposition:

Proposition 4.1. Let $t > 0$. Let N_1 and N_2 be two independent standard Poisson random variables with expectation $t/2$. Then the exponential family F_t with domain of the means $\mathbb{R}$ and variance function $(m^2 + t^2)^{1/2}$ exists and is generated by the distribution μ_t of $N_1 - N_2$. Furthermore

$$\mu_t(dx) = \sum_{n\in\mathbb{Z}} e^{-t} I_{|n|}(t) \delta_n(dx)$$

where

$$I_x(t) = \sum_{n=0}^{\infty} \frac{1}{n!\Gamma(n+x+1)} \left(\frac{t}{2}\right)^{2n+x}.$$

Proof. Since $\mathbb{E}(e^{\theta(N_1-N_2)} = e^{t(\cosh\theta-1)}$ we get that $\Theta(\mu_t) = \mathbb{R}$ and that

$$k_{\mu_t}(\theta) = t(\cosh\theta - 1),\ k'_{\mu_t}(\theta) = t\sinh\theta,\ \ k''_{\mu_t}(\theta) = t\cosh\theta = (k'_{\mu_t}(\theta)^2 + t^2)^{1/2}.$$

Thus $V_{F(\mu_t)}(m) = (m^2+t^2)^{1/2}$ as desired, and the domain of the means is $\mathbb{R}$. A consequence of this proposition and of 3 and 2 is that $(Bm^2+C)^{1/2}$ is always a variance function for B and $C > 0$.

4.2 The case $C = 0$

Proposition 4.2. Let $t > 0$. Then the exponential family F_t with domain of the means $\mathbb{R}$ and variance function $m(1+\frac{m^2}{t^2})^{1/2}$ exists. In particular F_1 is generated by $\mu_1 = \delta_0 + 2\sum_{n=1}^{\infty}\delta_n$. More specifically, P is in F_1 if and only if there exists $q \in (0,1)$ such that P is the convolution of the Bernoulli distribution $\frac{1}{1+q}\delta_0 + \frac{q}{1+q}\delta_1$ with the geometric distribution $(1-q)\sum_{n=0}^{\infty} q^n\delta_n$.

Proof. Writing for $\theta < 0$ $L_{\mu_1}(\theta) = \frac{1+e^\theta}{1-e^\theta}$ it is easily seen that it generates a natural exponential family with domain of the means $(0,\infty)$ and variance function $m(1+m^2)^{1/2}$. The only non trivial point of the proposition is the fact that the elements of F_1 are infinitely divisible. For this we write

$$k_{\mu_1}(\theta) = \sum_{n=1}^{\infty}\frac{1}{n}(1+(-1)^n)e^{n\theta}.$$

Since the coefficient $\frac{1}{n}(1+(-1)^n)$ of $e^{n\theta}$ is ≥ 0 the result is proved (although it is difficult to compute μ_t explicitly when t is not an integer.

A consequence of this proposition is that $(Am^4+Bm^2)^{1/2}$ is a variance function for A and $B > 0$ with domain of the means $(0,\infty)$.

4.3 The case $B^2 - 4AC = 0$

Here is a well known fact (see Morris (1982)):

Proposition 4.3. Let $t > 0$. The natural exponential family F_t with domain of the means $\mathbb{R}$ and variance function $t(1+\frac{m^2}{t^2})$ is generated by the probability μ_t defined by (8).

This rules the case $B^2 - 4AC = 0$ such that Ax^2+Bx+C has a negative double root with $A > 0$.

Proposition 4.4. Let $t > 0$. The natural exponential family F_t with domain of the means (t, ∞) and variance function $\frac{t}{2}(\frac{m^2}{t^2} - 1)$ exists. In particular F_1 is generated by $\mu_1 = \sum_{n=1}^{\infty} n\delta_n$.

Proof. We do not give the details about μ_1 which are standard. Since the elements of F_1 are negative binomial distributions shifted by 1, they are still infinitely divisible and F_t does exist for all $t > 0$.

This rules out the case $B^2 - 4AC = 0$ such that $Ax^2 + Bx + C$ has a positive double root x_0 with $A > 0$ and domain of the means $(\sqrt{x_0}, \infty)$.

Proposition 4.5. Let $N > 0$ be an integer. The natural exponential family F_t with domain of the means $(-N, N)$ and variance function $\frac{N}{2}(1 - \frac{m^2}{N^2})$ exists. It is generated by $(\delta_1 + \delta_{-1})^{*N}$.

Proof. This is an easy and classical fact.

4.4 $Ax^2 + Bx + C$ *cannot have simple roots on* $(0, \infty)$

We discard some values of (A, B, C). Suppose that $Ax^2 + Bx + C$ has a positive simple root $x_0 > 0$. Then $(Am^4 + Bm^2 + C)^{1/2}$ cannot be a variance function. For by the principle of maximal analyticity, the domain of the means will have $m_0 = \sqrt{x_0}$ has boundary point. Since x_0 is a simple root, then the variance function around m_0 will be equivalent to $k|m - m_0|^{1/2}$ for some positive constant k. But this is forbidden by the principle of Jørgensen, Martinez and Tsao (1994) mentioned in (10).

4.5 The splitting of the elliptic variances in three cases

The only cases that we are left to consider in order to have a classification of the variance functions of the form $(Am^4 + Bm^2 + C)^{1/2}$ are now the cases where $Ax^2 + Bx + C$ is strictly positive on $[0, \infty)$ and has no double negative root. Of course this implies that $A > 0$ and $C > 0$. To simplify the matters, we choose $C = 1$ and we introduce the function $V(m) = (Am^4 + Bm^2 + 1)^{1/2}$ and, for $t > 0$, the function $V_t(m) = tV(m/t)$. A simple analysis shows that $Ax^2 + Bx + 1$ has no positive roots and no double negative roots if and only if there exists a non zero real number a and a positive number b such that

$$Ax^2 + Bx + 1 = (1 + ax)^2 + 2b^2x.$$

Let us insist of the fact that a can be negative. Finally we introduce a complex number k through its square in order to use the standard notations of elliptic functions:

$$k^2 = 1 + \frac{2a}{b^2}$$

This leads to three cases

1. The case $-1 \le k^2 < 0$. It corresponds to the fact that $P(x) = (1+ax)^2 + 2b^2x$ has no roots and that the minimum of P on $[0,\infty)$ is reached on 0.
2. The case $k^2 < -1$. Here P has no roots and reaches its minimum on $[0,\infty)$ at $-b^2(a+b^2)/a^2$.
3. The case $k^2 > 0$. Here P has two distinct negative real roots. Taking $A = 1$ instead of $C = 1$ and $P(x) = (x+a^2)(x+b^2)$ is convenient.

We investigate these cases in the next three sections.

5 The elliptic cases: The case $-1 \le k^2 < 0$

We write $k^2 = -1 + p$ with $0 \le p < 1$ and we introduce the following two constants:

$$\begin{aligned} K &= \int_0^1 (1-x^2)^{-1/2}(2-p-x^2)^{-1/2}dx \\ K' &= \int_0^1 (1-x^2)^{-1/2}(1+(1-p)x^2)^{-1/2}dx. \end{aligned} \tag{19}$$

Here is our first serious result:

Theorem 5.1. Suppose that $k^2 = 1 + \frac{2a}{b^2} = -1 + p \in [-1,0)$. For $b = \sqrt{2}$ and $a = -2 + p$ there exists a natural exponential family G_t with domain of the means $\mathbb{R}$ and variance function

$$t\sqrt{(1+a\frac{m^2}{t^2})^2 + 2b^2\frac{m^2}{t^2}}$$

when t is a multiple of a. It is concentrated on $\frac{\pi}{2K}\mathbb{Z}$. The family $G_{|a|}$ is generated by a symmetric probability measure $\mu_{|a|}$ which is the convolution of the Bernoulli distribution $\frac{1}{2}(\delta_{-\frac{\pi}{2K}} + \delta_{\frac{\pi}{2K}})$ by an infinitely divisible distribution $\alpha_{|a|}$ concentrated on $\frac{\pi}{K}\mathbb{Z}$. We denote $q = e^{-\pi K'/K}$ and for a positive integer ν we denote

$$c_\nu = c_{-\nu} = \frac{q^\nu - (-1)^\nu q^{2\nu}}{1-q^{2\nu}} > 0.$$

Then the Laplace transform of α_t is

$$\int_{-\infty}^{\infty} e^{\theta x}\alpha_t(dx) = \exp\left(\frac{t}{|a|}\sum_{\nu\in\mathbb{Z}\setminus\{0\}} c_\nu(e^{\frac{\nu\pi\theta}{K}} - 1)\right).$$

Finally the characteristic function of $\mu_{2|a|}$ is $\frac{1}{\wp(s+K)-\frac{p}{3}}$ where $\wp$ is the elliptic Weierstrass function satisfying

$$\wp'^2 = 4(\wp - 1 + \frac{2p}{3})(\wp - \frac{p}{3})(\wp + 1 - \frac{p}{3})$$

which is doubly periodic with primitive periods $2K$ and $2iK'$. In particular it has zeros and G_t cannot be infinitely divisible.

Comments. Doing $b = \sqrt{2}$ is not really a restriction. Using the formula $a^2 V_F(m/a)$ for the image of F by $x \mapsto ax$ gives the description of F for an arbitrary $b > 0$.

Proof. We apply the standard procedure for computing the Laplace transform of a generating measure when the variance function is given. We shall use the following change of variable $u^2 = (1+am^2)^2 + 2b^2m^2$ for $u \geq 1$. This implies that

$$m^2 = \frac{1}{a^2}[-a - b^2 + \sqrt{b^4 + 2ab^2 + a^2u^2}].$$

We consider now the new change of variable

$$u = \frac{1}{2}((2 + \frac{b^2}{a})w^2 - \frac{b^2}{aw^2}) = \frac{b^2}{2a}(k^2w^2 - \frac{1}{w^2})$$

with $0 < w < 1$. This choice is designed in order to have $b^4 + 2ab^2 + a^2u^2 = b^4k^2 + a^2u^2$ transformed in a perfect square of a rational function of w :

$$\sqrt{b^4 + 2ab^2 + a^2u^2} = \frac{a}{2}((2 + \frac{b^2}{a})w^2 + \frac{b^2}{aw^2}) = \frac{b^2}{2}(k^2w^2 + \frac{1}{w^2})$$

This leads to

$$m^2 = \frac{b^2}{2a^2w^2}(1-w^2)(1-k^2w^2) \tag{20}$$

but also to a surprising result

$$\begin{aligned} a + b^2 + a^2m^2 &= (a + \frac{b^2}{2})w^2 + \frac{b^2}{2w^2} = \frac{b^2}{2}(k^2w^2 + \frac{1}{w^2}) \\ du &= [(a + \frac{b^2}{2})w^2 + \frac{b^2}{2w^2}]\frac{2}{aw}dw = \frac{b^2}{a}(k^2w^2 + \frac{1}{w^2})\frac{dw}{w} \\ \frac{du}{a + b^2 + a^2m^2} &= \frac{2}{aw}dw \end{aligned} \tag{21}$$

Recall that $a < 0$ and that $w \mapsto u$ is decreasing. Thus we get, gathering (21) and (20)

$$d\theta = \frac{1}{V(m)} \times dm = \frac{1}{u} \times \frac{2udu}{4m(a+b^2+a^2m^2)} = \frac{du}{2m(a+b^2+a^2m^2)} = \frac{dw}{awm}$$
$$= -\frac{\sqrt{2}}{b} \frac{dw}{\sqrt{(1-w^2)(1-k^2w^2)}} \tag{22}$$

We introduce the function $\theta \mapsto C(\theta)$ by

$$\theta = \int_{C(\theta)}^{1} \frac{dw}{\sqrt{(1-w^2)(1-k^2w^2)}}$$

Thus $C(0) = 1$ and the function C is defined on $[0,K']$. Actually, we have $C(\theta) = \mathrm{sn}\,(K'-\theta)$. In $(0,K')$ it satisfies $C'(\theta) = -(1-C(\theta)^2)^{1/2}(1-k^2C(\theta)^2)^{1/2}$. Now we can write

$$\theta = \int_0^m \psi'_\mu(x)dx = \int_{w(m)}^{1} \frac{dw}{(1-w^2)^{1/2}(1-k^2w^2)^{1/2}}.$$

Thus $w(m) = C(\theta)$ and from (20)

$$m = k'_{\mu_t}(\theta) = m(C(\theta)) = \frac{1}{|a|C(\theta)}(1-C(\theta)^2)^{1/2}(1-k^2C(\theta)^2)^{1/2} = \frac{C'(\theta)}{aC(\theta)}.$$

Thus finally we get the Laplace transform of μ_t as

$$L_{\mu_t}(\theta) = \frac{1}{(C(\theta))^{1/|a|}}.$$

We observe that $\theta \mapsto C(\theta)$ has an analytic continuation to the whole complex plane. We now consider its restriction $c(s) = C(is)$ to the imaginary line. It satisfies the differential equation

$$c'(s)^2 = (c(s)^2-1)(1-k^2c(s)^2) \tag{23}$$

with the initial condition $c(0) = 1$. Now introduce the function $s \mapsto f(s) = -k^2c^2(s)$. It satisfies the differential equation

$$f'(s)^2 = 4f^3(s) + 4(1+k^2)f^2(s) + 4f(s) = 4(f(s)+1)f(s)(f(s)+k^2)$$

(just multiply (23) by c^2 to reach this result). From now it is convenient to write

$$k^2 = 1+a = p-1$$

with $p \in [0,1)$. Then writing $f(s) = -\frac{p}{3} + h(s)$ we get $h'(s)^2 = 4h(s)^3 - g_2h(s) - g_3$ with

$$g_2 = 4(1-p+\frac{2p^2}{3}), \quad g_3 = -\frac{4p}{3}(1-p+\frac{2p^2}{9})$$

Thus h satisfies the differential equation of the $\wp$ function of Weierstrass for the parameters g_2 and g_3 (see SG 247). We can also write $h'^2(s) = 4(h-e_1)(h-e_2)(h-$

e_3) with $e_1 = 1 - \frac{2p}{3} > e_2 = \frac{p}{3} > e_3 = -1 + \frac{p}{3}$ with discriminant

$$\Delta = g_2^3 - 27g_3^2 = [4(1-p)(2-p)]^2.$$

Thus (see SG page 279 and page 283) the function $\wp$ has periods $2K = 2\omega$ and $2iK' = 2\omega'$

$$\begin{array}{ccccc}
2iK' & \text{———} & K+2iK' & \text{———} & 2K+2iK' \\
| & & | & & | \\
iK' & \text{———} & K+iK' & \text{———} & 2K+iK' \\
| & & | & & | \\
0 & \text{———} & K & \text{———} & 2K
\end{array}$$

$$\omega = \frac{1}{\sqrt{2-p}} \int_0^1 \frac{dw}{(1-w^2)^{1/2}(1-\frac{1}{2-p}w^2)^{1/2}},$$
$$\omega' = \frac{i}{\sqrt{2-p}} \int_0^1 \frac{dw}{(1-w^2)^{1/2}(1-\frac{1-p}{2-p}w^2)^{1/2}} = iK'.$$

The last equality $\omega' = iK'$ is obtained from the changes of variable $w = u^{1/2}$, $u = 1 - v$ and $v = t^2$. We have $h(s) = \wp(s+C)$ for some constant C. Now, since the variance function $t(1 + a\frac{m^2}{t^2})^{1/2})^2 + \frac{4m^4}{t^2})^{1/2}$ is symmetric, there exists a symmetric measure which generates it and without loss of generality we assume that the characteristic function $s \mapsto \frac{1}{f(s)^{t/2|a|}}$ is real. Thus we have to take C such that $c(0) = 1$ or $f(0) = 1 - p$ or $h(0) = 1 - \frac{2p}{3} = e_1$ or $\wp(C) = e_1$. Hence from SG page 279 $C = K$.

Since $s \mapsto \wp(s)$ has periods $2K$ and $2iK'$ and since $\wp(K) = 1 - \frac{2p}{3} = e_1$ we have

$$f(s) = \wp(s+K) - \frac{p}{3}.$$

Note that $s \mapsto \wp(s)$ has no real zeros, only poles on multiples of $2K$ and is periodic. See the picture in SG page 280. Thus

$$s \mapsto \left(\frac{1-p}{f(s)}\right)^{t/2|a|}$$

is $2K$ periodic and has zeros on odd multiples of K. Since it is $2K$ periodic, this implies that it is the characteristic function of a symmetric probability concentrated on multiples of π/K.

$$\frac{1-p}{f(s)}^{t/2|a|} = \sum_{\nu\in\mathbb{Z}} p_\nu(t) e^{i\nu\frac{\pi}{K}} = p_0(t) + 2\sum_{\nu=1}^{\infty} p_\nu(t)\cos\frac{\pi\nu}{K}.$$

We are going to consider $-\frac{t}{2|a|}\log f(s)$ and to compute its Fourier series. For this we use formula 5.8-22 in SG page 263 applied to $\alpha = 2$ which gives here since $e_2 = p/3$

$$f(2Kz)^{1/2} = \sqrt{\wp(2Kz+K) - \frac{p}{3}} = \frac{\vartheta_3(z+\frac{1}{2})}{C\vartheta_1(z+\frac{1}{2})} \tag{24}$$

where C is some constant and where the q occurring in the Theta functions is given by SG page 261 by $q = e^{i\pi\tau}$ with $\tau = iK'/K$. Thus $q = e^{-\frac{K'}{K}\pi}$ here. Forgetting the factor $t/|a|$ we have

$$-\frac{1}{2}\log f(s) = \log C + \log\vartheta_1(\frac{s}{2K}+\frac{1}{2}) - \log\vartheta_3(\frac{s}{2K}+\frac{1}{2}).$$

Consider the derivative of this function:

$$-\frac{1}{2}(\log f(s))' = \frac{1}{2K}\frac{\vartheta_1'(\frac{s}{2K}+\frac{1}{2})}{\vartheta_1(\frac{s}{2K}+\frac{1}{2})} - \frac{1}{2K}\frac{\vartheta_3'(\frac{s}{2K}+\frac{1}{2})}{\vartheta_3(\frac{s}{2K}+\frac{1}{2})}.$$

Now we use formulas about ϑ_j'/ϑ_j given in SG page 274. They are

$$\begin{aligned}
\frac{\vartheta_1'(z)}{\vartheta_1(z)} &= \pi\tfrac{\cos\pi z}{\sin\pi z} + 4\pi\sum_{\nu=1}^{\infty}\frac{q^{2\nu}}{1-q^{2\nu}}\sin 2\nu\pi z\\
\frac{\vartheta_3'(z)}{\vartheta_3(z)} &= \qquad\qquad 4\pi\sum_{\nu=1}^{\infty}\frac{(-1)^\nu q^\nu}{1-q^{2\nu}}\sin 2\nu\pi z
\end{aligned}$$

In these expressions we replace z by $\frac{s}{2K}+\frac{1}{2}$ and we get

$$\begin{aligned}
\frac{\vartheta_1'(\frac{\pi s}{2K}+\frac{\pi}{2})}{\vartheta_1(\frac{\pi s}{2K}+\frac{\pi}{2})} &= -\pi\tfrac{\sin\frac{\pi s}{2K}}{\cos\frac{\pi s}{2K}} + 4\pi\sum_{\nu=1}^{\infty}\frac{(-1)^\nu q^{2\nu}}{1-q^{2\nu}}\sin\frac{\pi s}{K}\\
\frac{\vartheta_3'(\frac{\pi s}{2K}+\frac{\pi}{2})}{\vartheta_3(\frac{\pi s}{2K}+\frac{\pi}{2})} &= \qquad\qquad 4\pi\sum_{\nu=1}^{\infty}\frac{q^\nu}{1-q^{2\nu}}\sin\frac{\pi s}{K}
\end{aligned}$$

Thus

$$\begin{aligned}
-\frac{K}{\pi}(\log f(s))' &= -\frac{\sin\frac{\pi s}{2K}}{\cos\frac{\pi s}{2K}} + 4\sum_{\nu=1}^{\infty}\frac{(-1)^\nu q^{2\nu}-q^\nu}{1-q^{2\nu}}\sin 2\frac{\pi s}{2K},\\
-\frac{1}{2}\log f(s) &= C_1 + \log\cos\frac{\pi s}{2K} + 2\sum_{\nu=1}^{\infty}\frac{q^\nu-(-1)^\nu q^{2\nu}}{1-q^{2\nu}}\cos 2\nu\frac{\pi s}{2K},\\
-\frac{1}{2}\log f(s) &= C_1 + \log\cos\frac{\pi s}{2K} + \sum_{\nu\in\mathbb{Z}\setminus\{0\}}\frac{q^{|\nu|}-(-1)^\nu q^{2|\nu|}|}{1-q^{2|\nu|}}\exp 2\nu i\pi\frac{s}{2K},
\end{aligned}$$

where C_1 is a constant such that $f(0) = 1-p$. Thus

$$\frac{1}{c(s)^{t/|a|}} = C_2(\cos\frac{\pi s}{2K})^{t/|a|}\exp\left[\sum_{\nu\in\mathbb{Z}\setminus\{0\}}\frac{q^{|\nu|}-(-1)^\nu q^{2|\nu|}}{1-q^{2|\nu|}}\exp 2\nu i\pi\frac{s}{2K}\right]$$

The theorem is proved.

6 The elliptic cases: $k^2 < -1$

This case is more complicated when treated by the retrieving method of Section 2. A reason is the fact that the function $m \mapsto V(m)$ is not convex. More specifically, if $P(x) = (1+ax)^2 + 2b^2x$ is used to define $V(m) = \sqrt{P(m^2)}$ the case $k^2 = 1 + \frac{2a}{b^2} < -1$ correspond to the case where P' has a positive root. Here we shall rather use the method of associated NEF, but no new interesting distributions will appear, as shown by the following result:

Theorem 6.1. If $k^2 = 1 + \frac{2a}{b^2} < -1$ then $\sqrt{(1+am^2)^2 + 2b^2m^2}$ is not a variance function.

Proof. Suppose that $\sqrt{(1+am^2)^2 + 2b^2m^2}$ is the variance function of some NEF F_1. Let us assume first that the associated F_2 exists. As a consequence the variance function of F_2 is

$$V_{F_2}(m) = \sqrt{(1-am^2)^2 - 2b^2m^2} = \sqrt{(1+am^2)^2 + 2(2a-b^2)m^2}.$$

Like in Theorem 5.1 without loss of generality we may assume that $2a - b^2 = 2$, and Theorem 5.1 gives us a detailed description of F_2. If μ_2 is the symmetric probability generating F_2 we have seen that $L_\mu(\theta) = C(\theta)^{-1/|a|}$ where $C(\theta) = \mathrm{sn}\,(K' - \theta)$ with K and K' defined by (19). As a consequence, if μ_1 is the symmetric probability generating F_1, then from Proposition 3.1 its Fourier transform is

$$\varphi_{\mu_1}(s) = C(s)^{1/|a|}.$$

Now we use the fact that the function C is doubly periodic with periods $2K$ and $2iK'$. This implies that the Fourier transform $\varphi_{\mu_1}(s)$ has period $2K'$ which means that μ_2 is concentrated on a coset of the group $\mathbb{Z}/(2K')$. We are going to use this to see that $s \mapsto C(s)^{-1/|a|}$ is also a Fourier transform of a probability and this will obviously contradict the existence of μ_2. For this, we have to understand the $2K'$ periodicity of C by coming back to formula (24) which shows that $C(2Ks)$ is the power of a function of the form

$$\frac{\vartheta_3(z+\frac{1}{2})}{C\vartheta_1(z+\frac{1}{2})}.$$

Now we use the Jacobi imaginary transformation (see SG pages 269-272). In our particular case this Jacobi transformation is the following. Denote $\tau = iK'/K$. Then $-1/\tau = iK/K'$. The theta functions ϑ_1 and ϑ_3 are implicitely functions of τ and it is sometimes useful to write $\vartheta_i(z|\tau)$ instead of $\vartheta_i(z)$. Formulas linking $\vartheta_i(z|\tau)$ and $\vartheta_i(z|-1/\tau)$ are known (see 5.10-9 in SG). These formulas show, by the magic of the Jacobi transformation, that $C(s)^{-1/|a|}$ is the Fourier transform of the probability on $\mathbb{Z}/K'$ which is obtained from Theorem 5.1 simply by exchanging the roles of K and K'. We therefore obtain the desired contradiction.

The last task is to get rid of the hypothesis that F_1 has an associated NEF. If $F_1 = F(\mu_1)$ exists with $V_{F_1}(m) = \sqrt{(1+am^2)^2 + 2b^2m^2}$ then playing with the affine and the Jørgensen transformations (2) and (3) it is possible to find a positive number t such that $\sqrt{(1-a\frac{m}{t}^2)^2 - 2b^2\frac{m}{t}^2}$ is the variance function of some NEF F_2. This $F_2 = F(\mu_2)$ is necessarily of the type considered in Theorem 5.1 (namely with $1 < k^2 < 0$). We have seen that in this case the Laplace transform of μ_2 is a negative power of $C(\theta) = \mathrm{sn}\,(\theta - K')$. Therefore the Fourier transform of μ_1 is a positive power of $C(s)$. However, a negative power of $C(s)$ was also the Fourier transform of a probability: we get a contradiction and this ends the proof.

7 The elliptic cases: $k^2 > 0$

In this section we study the variances of the form $V(m) = \sqrt{(m^2+a^2)(m^2+b^2)}$ where $0 < a < b$.

Theorem 7.1 Let $0 < a < b$ such that $b^2 - a^2 = 2^{4/3}$. There exists an natural exponential family with variance function $2\sqrt{(\frac{m^2}{4}+a^2)(\frac{m^2}{4}+b^2)}$. It is generated by a symmetric discrete distribution concentrated on the group $\frac{\pi}{K}\mathbb{Z}$ where K i a constant given below by (25).

Proof. We have $\theta = \int_0^m \frac{dt}{\sqrt{(t^2+a^2)(t^2+b^2)}}$. We do the changes of variable $t \mapsto u \mapsto v \mapsto w$ defined for $u > ab$, $v > \frac{2ab}{b^2-a^2}$, $w > \sqrt{\frac{b+a}{b-a}}$.

$$u^2 = (t^2+a^2)(t^2+b^2), \;\; u = \frac{b^2-a^2}{2}v, \;\; v = \frac{1}{2}(w^2 - \frac{1}{w^2}).$$

Since $udu = 2t(2t^2+b^2+a^2)dt$ and $dv = (w^2 + \frac{1}{w^2})\frac{dw}{w}$ we get

$$\theta = \frac{b^2-a^2}{4}\int_{\frac{2ab}{b^2-a^2}}^{\frac{2\sqrt{(m^2+a^2)(m^2+b^2)}}{b^2-a^2}} \frac{dv}{t(2t^2+b^2+a^2)}.$$

We have also

$$\begin{aligned} t^2 &= -\frac{b^2+a^2}{2} + \frac{b^2-a^2}{2}\sqrt{v^2+1} = -\frac{b^2+a^2}{2} + \frac{b^2-a^2}{4}(w^2+\frac{1}{w^2}) \\ &= \frac{b^2-a^2}{4w^2}(w^2-\frac{b+a}{b-a})(w^2-\frac{b-a}{b+a}). \end{aligned}$$

Thus $2t^2+a^2+b^2 = \frac{b^2-a^2}{2}(w^2+\frac{1}{w^2})$ which implies that $\frac{dv}{2t^2+b^2+a^2} = \frac{b^2-a^2}{2}\frac{dw}{w}$. Denoting for simplification $r = \frac{b+a}{b-a} > 1$ we get

$$\theta = \frac{(b^2-a^2)^{3/2}}{4}\int_{\sqrt{r}}^{w(m)} \frac{dw}{\sqrt{(w^2-r)(w^2-r^{-1})}}$$

where $m \mapsto w(m) > \sqrt{r}$ is defined by

$$m^2 = \frac{b^2-a^2}{4w^2(m)}(w^2(m)-r)(w^2(m)-r^{-1}).$$

Now for simplification let us assume that $\frac{(b^2-a^2)^{3/2}}{4} = 1$. Introduce the function $\theta \mapsto C(\theta)$ from $(0,\infty)$ to $(\sqrt{r},\infty)$ defined by

$$\theta = \int_{\sqrt{r}}^{C(\theta)} \frac{dw}{\sqrt{(w^2-r)(w^2-r^{-1})}}.$$

Thus we have $w(m) = w(k'(\theta)) = C(\theta)$. This function C satisfies the differential equation

$$C' = \sqrt{(C^2-r)(C^2-r^{-1})}, \quad \frac{C'(\theta)}{C(\theta)} = \frac{1}{C(\theta)}\sqrt{(C^2(\theta)-r)(C^2(\theta)-r^{-1})} = \frac{2}{\sqrt{b^2-a^2}}k'(\theta).$$

Thus the Laplace transform of μ is $L(\theta) = C(\theta)^c$ where $c = \frac{\sqrt{b^2-a^2}}{2}$. We now imitate the procedure used in Theorem 4.1: we consider the Fourier transform $c(s) = C(is)$ for $s \in \mathbb{R}$, which satisfies

$$c'(s)^2 = -(c^2(s)-r)(c^2(s)-r^{-1}),$$

then $f(s) = c^2(s)$ which satisfies $f'^2 = 4c^2c'^2 = -4f(f-r)(f-r^{-1})$, then $h(s) = \frac{1}{3}(r+r^{-1}) - f(s)$ which satisfies

$$\begin{aligned} h'(s)^2 &= 4h(s)^3 - g_2h(s) - g_3 \\ &= 4[h(s)-\frac{1}{3}(r+r^{-1})][h(s)-\frac{1}{3}(r-2r^{-1})][h(s)-\frac{1}{3}(r^{-1}-2r)] \\ &= 4(h(s)-e_1)(h(s)-e_2)(h(s)-e_3) \end{aligned}$$

with $g_2 = \frac{4}{3}(r^2+r^{-2}-1)$, $g_3 = \frac{4}{27}(3r+3r^{-1}-2r^3-2r^{-3})$ and

$$e_1 = \frac{1}{3}(r+r^{-1}) \ > \ e_2 = \frac{1}{3}(r-2r^{-1}) \ > \ e_3 = \frac{1}{3}(r^{-1}-2r)$$

Hence for some complex constant C we have $h(s) = \wp(s+C)$ with periods $2K$ and $2iK'$ defined by

$$K = \int_{e_1}^{\infty} \frac{dx}{\sqrt{4x^3 - g_2 x - g_3}}, \ K' = \int_{-\infty}^{e_3} \frac{dx}{\sqrt{-4x^3 + g_2 x + g_3}} \tag{25}$$

(see Whittaker and Watson (1927) Example 1 page 444). Now to determine the constant C one observes that $f(s) = \frac{1}{3}(r+r^{-1}) - \wp(s+C)$ is real since this is the Fourier transform of a symmetric measure. Furthermore $f(0) = c(0)^2 = r$. Thus $\wp(C) = \frac{1}{3}(r^{-1}-2r) = e_3$ which implies $C = iK'$. Now we use the formula (see SG formula 5.8-22)

$$\sqrt{\wp(z) - e_j} = \frac{\pi}{2K}\left[\frac{\vartheta_1'(0)\vartheta_{j+1}(\frac{z}{2K})}{\vartheta_{j+1}(0)\vartheta_1(\frac{z}{2K})}\right]$$

that we shall use for writing

$$c^2(s) = f(s) = \frac{1}{3}(r+r^{-1}) - h(s) = e_1 - \wp(s+iK') = \frac{\pi^2}{4K^2}\left|\frac{\vartheta_1'(0)\vartheta_2(\frac{s}{2K}+i\frac{K'}{2K})}{\vartheta_2(0)\vartheta_1(\frac{s}{2K}+i\frac{K'}{2K})}\right|^2.$$

Let us introduce the notation $q = e^{-\pi K/K'}$ With it, ϑ_1 and ϑ_2 are given by

$$\vartheta_1(z) = 2Cq^{1/4}\sin\pi z\prod_{\nu=1}^{\infty}(1-2q^{2\nu}\cos 2\pi z + q^{4\nu})$$
$$\vartheta_2(z) = 2Cq^{1/4}\cos\pi z\prod_{\nu=1}^{\infty}(1+2q^{2\nu}\cos 2\pi z + q^{4\nu})$$

where $C = \prod_{\nu=1}(1-q^{2\nu})$ (see SG pages 268-9). Let us give a simpler presentation of $c^2(s)$: using $z = \frac{s}{2K} + i\frac{K'}{2K}$ and $u = e^{\frac{i\pi s}{2K}}$ we introduce the following symbols for $\nu = 1,2,\ldots$

$$\varphi_\nu(u) = \frac{u^4+u^{-4}}{q^{2\nu}+q^{-2\nu}+q^2+q^{-2}}, \ \varphi_0(u) = \frac{u^2+u^{-2}}{q+q^{-1}}.$$

We get

$$\left|\frac{\cos\pi z}{\sin\pi z}\right|^2 = \frac{1+\varphi_0(u)}{1-\varphi_0(u)}, \ \left|\frac{1+2q^{2\nu}\cos 2\pi z + q^{4\nu}}{1-2q^{2\nu}\cos 2\pi z + q^{4\nu}}\right|^2 = \frac{1+\varphi_\nu(u)}{1-\varphi_\nu(u)}.$$

and finally the elegant formula

$$c^2(s) = f(s) = C\prod_{\nu=0}^{\infty}\frac{1+\varphi_\nu(u)}{1-\varphi_\nu(u)}$$

where the constant C is such that $f(0)=1$. The last step is the formula, correct for $|Z|<1$:

$$\frac{1+Z}{1-Z}=1+2\sum_{n=1}^{\infty}Z^n=\sum_{n\in\mathbb{Z}}Z^{|n|}.$$

Replacing Z by $\varphi_\nu(u)$ we see that $f_\nu(s)=C_\nu\frac{1+\varphi_\nu(u)}{1-\varphi_\nu(u)}$ where C_ν is such that $f_\nu(0)=1$ is the characteristic function of a probability distribution concentrated on the additive group $\frac{2\pi}{K}\mathbb{Z}$ for $\nu\geq 1$ and on the additive group $\frac{\pi}{K}\mathbb{Z}$ for $\nu=0$. As a result f is the characteristic function of a symmetric discrete distribution on the group $\frac{\pi}{K}\mathbb{Z}$.

Comments. Of course the restriction $b^2-a^2=2^{4/3}$ is not important and can be generalized by a dilation. In the other hand, finding the Jørgensen set of these families is a difficult question. It should also be mentioned that the characteristic function $f_\nu(s)$ above has the form $\frac{1-c_\nu}{1-c_\nu}\frac{1+c_\nu\cos s}{1-c_\nu\cos s}$ after dilation $x\mapsto\frac{2\pi}{K}x$ if $\nu\neq 0$ or $x\mapsto\frac{\pi}{K}x$. If the Poisson kernel distribution on $\mathbb{Z}$ of parameter $r\in(0,1)$ is defined by $p_n=\frac{1-r}{1+r}r^{|n|}$, then $\frac{1-c_\nu}{1-c_\nu}\frac{1+c_\nu\cos s}{1-c_\nu\cos s}$ can be seen as a mixing of a Dirac mass on zero and of Poisson kernel distribution with parameter $r=c_\nu$.

8 The family F

Theorem 8.1. Let $x>0$. The NEF F_x with domain of the means $(0,\infty)$ and variance function

$$V_{F_x}(m)=m(1+\frac{4m^4}{x^4})^{1/2}$$

is generated by a positive measure on $\mathbb{N}$ which is $\nu_x(dt)=\sum_{n=0}^{\infty}\frac{p_n(x)}{n!}\delta_n(dt)$ with generating function

$$f_x(z)=\sum_{n=0}^{\infty}\frac{p_n(x)}{n!}z^n=e^{x\int_0^z\frac{dw}{(1-w^4)^{1/2}}}. \tag{26}$$

which satisfies

$$(1-z^4)f_x''(z)-2z^3f_x'(z)-x^2f_x(z)=0.$$

The total mass of ν_x is $\exp(x\frac{1}{4}B(\frac{1}{2},\frac{1}{4}))$. The polynomials p_n are given by $p_n(x)=x^n$ for $n=0,1,2,3,4$, $p_5(x)=x^5+12x$ and for $n\geq 2$

$$p_{n+2}(x)=x^2p_n(x)+n(n-1)^2(n-2)p_{n-2}(x).$$

Proof. The proof of the first formula is a routine calculation for exponential families concentrated on $\mathbb{N}$, but we give details. We use successively the change of variables $u=2m^2/x^2$ and $u=\sinh v$.

$$d\theta = \frac{dm}{V(m)} = \frac{4mdm}{4m^2\sqrt{1+4\frac{m^4}{x^4}}} = \frac{du}{u\sqrt{1+u^2}} = \frac{dv}{2\sinh v} = \frac{1}{2}(\frac{1}{e^v-1} - \frac{1}{e^v+1})e^v dv$$

Denoting $z = e^\theta$ we get

$$z^2 = \frac{e^v-1}{e^v+1},\ e^v = \frac{1+z^2}{1-z^2},\ e^{-v} = \frac{1-z^2}{1+z^2},\ u = \sinh v = \frac{2z^2}{1-z^4},\ m = \frac{xz}{\sqrt{1-z^4}}.$$

Thus $k'_\mu(\theta) = x\frac{e^\theta}{\sqrt{1-e^{4\theta}}}$ and this leads to the result 26.

The trick to obtain the differential equation for f_x is to write $(1-z^4)^{1/2}f'_x = xf_x$, then to differentiate with respect to z and then to multiply both sides of the result by $(1-z^4)^{1/2}$. Then the differential equation leads to the equality

$$\sum_{n\geq 0}(n+1)(n+2)\frac{p_{n+2}(x)}{(n+2)!}z^n - \sum_{n\geq 4}(n-3)(n-2)\frac{p_{n-2}(x)}{(n-2)!}z^n$$
$$-2\sum_{n\geq 3}(n-2)\frac{p_{n-2}(x)}{(n-2)!}z^n - x^2\sum_{n\geq 0}\frac{p_n(x)}{n!}z^n = 0.$$

Using $f_x(0) = p_0(x) = 1$ and $f'_x(0) = p_1(x) = x$ we get $p_n(x) = x^n$ for $0 \leq n \leq 4$ and for $n \geq 4$ we have

$$(n+1)(n+2)\frac{p_{n+2}(x)}{(n+2)!} - (n-3)(n-2)\frac{p_{n-2}(x)}{(n-2)!} - 2(n-2)\frac{p_{n-2}(x)}{(n-2)!} - x^2\frac{p_n(x)}{n!} = 0,$$

$$(n+1)(n+2)\frac{p_{n+2}(x)}{(n+2)!} - (n-1)(n-2)\frac{p_{n-2}(x)}{(n-2)!} - x^2\frac{p_n(x)}{n!} = 0.$$

Now we multiply both sides by $n!$ and we use the definition of p_n for getting

$$p_{n+2}(x) = x^2p_n(x) + n(n-1)^2(n-2)p_{n-2}(x).$$

Checking the correctness of this equality for $n = 2,3$ is easy.

Remarks. It is easy to check that if $n = 4q+r$ with $r = 0,1,2,3$ then there exists a monic polynomial $P_{q,r}$ of degree q such that

$$p_n(x) = x^rP_{q,r}(x^4).$$

For instance $P_{0,r}(z) = 1,\ P_{1,0}(z) = z,\ P_{1,1}(z) = z+12,\ P_{1,2}(z) = z+72,$
$P_{1,3}(z) = z+252, P_{2,0}(z) = z^2+672z,\ P_{2,1}(z) = z^2+1512z+1260.$

We now extend Theorem 8.1 to a more general variance function, without being so specific about calculation of the corresponding distribution. This variance function for $x = 1$ is the reciprocal variance function for $t = 1$ of V_{G_t} where

$$V_{G_t}(m) = \sqrt{1+2p\frac{m^2}{t^2}+(2-p)^2\frac{m^4}{t^4}}.$$

Theorem 8.2. Let $p \in [0,1)$. Let $x > 0$. The NEF F_x with domain of the means $(0,\infty)$ and variance function

$$V_{F_x}(m) = m\left[1+2p\frac{m^2}{x^2}+\frac{(2-p)^2m^4}{x^4}\right]^{1/2}$$

is generated by a positive measure on $\mathbb{N}$ with generating function

$$f_x(z) = \exp\left[x\sqrt{\frac{2}{2-p}}\int_0^z (1+qw^2)^{1/2}(1-w^4)^{-1/2}dw\right]. \tag{27}$$

where $q = p/(2-p)^2$.

Proof. It is convenient to denote $c = p/(2-p)$ and to observe that

$$0 \le c < 1\,, \quad \sqrt{1-c^2} = \frac{2\sqrt{1-p}}{2-p}\,, \quad \frac{c\pm 1}{\sqrt{1-c^2}} = \pm\frac{1}{\sqrt{1-p}}.$$

We use successively the change of variables $u = (2-p)m^2/x^2$ and $u = \sqrt{1-c^2}\sinh v - c$.

$$d\theta = \frac{dm}{V(m)} = \frac{4mdm}{4m^2\sqrt{1+2p\frac{m^2}{x^2}+\frac{(2-p)^2m^4}{x^4}}} = \frac{du}{u\sqrt{1+2cu+u^2}} = \frac{dv}{2\sqrt{1-c^2}\sinh v - 2c}$$

$$= \frac{e^v dv}{\sqrt{1-c^2}e^{2v}-2ce^v-\sqrt{1-c^2}} = \frac{1}{2}\left[\frac{1}{e^v-\frac{1}{\sqrt{1-p}}}-\frac{1}{e^v+\frac{1}{\sqrt{1-p}}}\right]e^v dv.$$

Denoting $z = e^\theta$ we get

$$z^2 = \frac{e^v-\frac{1}{\sqrt{1-p}}}{e^v+\frac{1}{\sqrt{1-p}}},\ e^v = \frac{1}{\sqrt{1-p}}\frac{1+z^2}{1-z^2},\ e^{-v} = \sqrt{1-p}\frac{1-z^2}{1+z^2}$$

$$\sqrt{1-c^2}\sinh v = \frac{4z^2+p(1-z^2)^2}{(2-p)(1-z^4)},\ u = \frac{2z^2}{1-z^4}(1+cz^2),\ m^2 = \frac{2}{2-p}\frac{z^2}{1-z^4}(1+qz^2).$$

Thus

$$m = k'_\mu(\theta) = x\sqrt{\frac{2}{2-p}}\frac{(1+qe^{2\theta})^{1/2}}{(1-e^{4\theta})^{1/2}}e^\theta$$

and this leads to the result (27).

It remains to prove that the Taylor expansion of $z \mapsto f_x(z)$ defined by (27) has positive coefficients. For this it is enough to prove that the argument of the expo-

nential

$$z \mapsto \int_0^z (1+qw^2)^{1/2}(1-w^4)^{-1/2}dw$$

has positive coefficients. It is enough to prove that $z \mapsto (1+qz^2)^{1/2}(1-z^4)^{-1/2}$ has positive coefficients. It is enough to prove that $z \mapsto (1+qz)^{1/2}(1-z^2)^{-1/2}$ has positive coefficients. It is enough to prove that

$$z \mapsto \log[(1+qz)(1-z^2)^{-1}] = \sum_{n=1}^{\infty} a_n z^n$$

has positive coefficients. But this very last point is easy to check since $0 \leq q < 1$ and since a_n is computable: for odd n then $a_n = q^n/n > 0$ and for even $n = 2p$ we have

$$a_n = \frac{1}{p} - \frac{q^{2p}}{2p} > 0.$$

The theorem is proved.

9 Conclusion: general elliptic variances

It seems that the present paper is only scratching the surface of an interesting theory. Indeed, consider the set of variance functions of the form

$$V_F(m) = (\alpha m + \beta)\sqrt{P(m)} \tag{28}$$

where P is a polynomial with degree ≤ 4. The present paper has considered only the cases $P(m) = Am^4 + Bm^2 + C$. Recall a definition appearing in Hassairi (1992) and Barlev, Bshouty and Letac (1994). We say that two NEF F_1 and F_2 on the real line belong to the same orbit if there exists a Moebius transformation $y = (ax+b)/(cx+d)$ such that $ad - bc = 1$ and such that on a suitable interval for m we have

$$V_{F_1}(m) = (cm+d)^3 V_{F_2}\left(\frac{am+b}{cm+d}\right).$$

The most celebrated pair (F_1, F_2) is the set of normal distributions with variance 1 and the set of inverse Gaussian distributions, with variance m^3 on $(0,\infty)$. The pair $(\sqrt{4+m^4},\ m\sqrt{1+4m^4})$ offers another example. Roughly saying that F_1 and F_2 belong to the same orbit means the following: suppose that F_1 and F_2 are generated by μ_1 and μ_2 and let us draw in $\mathbb{R}^2$ the curves C_1 and C_2 which are the representative curves of the convex functions k_{μ_1} and k_{μ_2} (in the case of the pair normal-inverse Gaussian, they are a parabola and a piece of parabola). Then F_1 and F_2 are in the same orbit if and only if there exists an affine transformation of the plane $\mathbb{R}^2$ which maps a part of C_1 onto a part of C_2. This affine transformation can be described in terms of the coefficients (a,b,c,d) of the Moebius transformation. A very sat-

isfactory fact observed in Hassairi (1992) is that the quadratic and cubic NEF are split in 4 orbits, respectively generated by the normal, the Poisson, the hyperbolic and the Ressel-Kendall distributions. Now we remark that if F has the form (28) and if G is in the orbit of F then necessarily $V_G(m) = (\alpha_1 m + \beta_1)\sqrt{P_1(m)}$ where the polynomial P_1 has also degree ≤ 4. Therefore we are facing the problem of a whole classification of this set (28) of variance functions into orbits. This implies a mastering of the elliptic curves $y^2 = P(x)$ and the use of beautiful mathematics. The theory of exponential families expanded by Ole Barndorff Nielsen forty years ago is still hiding many secrets.

10 References

1. BARNDORFF-NIELSEN, O. (1978) *'Information and Exponential Families in Statistical Theory'*. Wiley, New York.
2. BAR-LEV, S.K., ENIS, P. AND LETAC, G. (1994) 'Sampling models which admit a given exponential family as conjugate family of priors.' *Ann. Statist.* **22**, 1555-1586.
3. BAR-LEV, S.K. AND VAN DE DUYN SHOUTEN, F.A. (2004) 'A note on exponential dispersion models which are invariant under length-biased sampling.' *Statist. and Probab. Lett.* **70**, 275-284.
4. BAR-LEV, S. K. AND LETAC, G. (2012) 'Increasing hasard rate of mixtures for natural exponential families.' *Adv. Appl. Probab.* **44**, 373-390.
5. GRADSHTEYN I.S., AND RYZHIK, I.M. (2007) *'Table of Integrals, Series and Products'*. Edited by A. Jeffrey and D. Zwillinger, Academic Press,, New York, 7th Edition
6. BROWN, L. D. (1978) *'Fundamentals of Statistical Exonential Families with Applications in Statistical Decision Theory'*. Institute of Methamatical Statistics, Hayward, CA.
7. HASSAIRI, 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. Sc. Paris Série I.* **315**, 207-210.
8. JØRGENSEN, B. (1987) 'Exponential dispersion models.' *J. Roy. Statist. Soc. Ser. B* **49**, 127-162.
9. JØRGENSEN, B., MARTINEZ, J.R. AND TSAO, M. (1993) 'Asymptotic Behavior of the Variance Function.' *Scand. J. Statist.* **21**, 223-243.
10. JØRGENSEN, B. (1997) *'The Theory of Dispersion Models.'* Chapman and Hall, London.
11. KAWATA, T. (1972) *'Fourier Analysis in Probability Theory'*. Academic Press, New York.
12. LETAC, G. AND MORA, M. (1990) 'Natural exponential families with cubic variances.' *Ann. Statist.* **18**, 1-37.
13. LETAC, G. (1991) 'The classification of the natural exponential families by their variance functions.' *Proceedings of the ISI Cairo Congress.* **3**, 271-298.

14. LETAC, G. (1992) '*Lectures on natural exponential families and their variance functions.*', 128 pages, Monografias de matemática, **50**, IMPA, Rio de Janeiro.
15. LETAC, G., MALOUCHE, D. AND MAURER, S. (2002) 'The real powers of the convolution of a negative binomial and a Bernoulli distribution.' *Proc. Amer. Math. Soc.*, **130**, 2107-2114.
16. MORRIS, C.N. (1982) 'Natural exponential families with quadratic variance functions.' *Ann. Statist.*, **10**, 65-80.
17. SANSONE, G. AND GERRETSEN, J. (1960) '*Lectures on the Theory of Functions of a Complex variable', Vol.1.* P. Noordhof Ltd, Groningen.
18. SHANBHAG, D.N. (1979). Diagonality of the Bhattacharyya matrix as a characterization, *Theory Prob. Appl.*, **23**, 430-433.
19. TWEEDIE, M.C.K. (1984) 'An index which distinguishes between some important exponential families' In *Statistics: Applications and New Directions. Proc. Indian Institute Golden Jubilee Internat. Conf.* (J.K. Ghosh and J. Roy, eds) 579-604. Indian Statistical Institute, Calcutta.
20. WHITTAKER, E.T. AND WATSON, G.N. (1927) '*A Course in Modern Analysis'*. Fourth edition. Cambridge University Press.