

CÉLESTIN CLOTAIRE KOKONENDJI

**Sur les familles exponentielles naturelles
réelles de grand-Babel**

Annales de la faculté des sciences de Toulouse 6^e série, tome 4, n^o 4
(1995), p. 763-800

http://www.numdam.org/item?id=AFST_1995_6_4_4_763_0

© Université Paul Sabatier, 1995, tous droits réservés.

L'accès aux archives de la revue « Annales de la faculté des sciences de Toulouse » (<http://picard.ups-tlse.fr/~annales/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Sur les familles exponentielles naturelles réelles de grand-Babel^(*)

CÉLESTIN CLOTAIRE KOKONENDJI⁽¹⁾

RÉSUMÉ. — Un grand nombre de familles exponentielles naturelles (f.e.n.) sur $\mathbb{R}$ dans la littérature ont des fonctions variances de la forme $R\Delta + Q\sqrt{\Delta}$ où R , Q et Δ sont des polynômes de degrés ≤ 1 , ≤ 2 , et ≤ 2 respectivement. Cette classe de f.e.n., appelée “grand-Babel” a été introduite par G. Letac comme une généralisation raisonnable des classes de Morris et de Mora. Elle est fermée par réciprocity et contient les classes de Babel et de Seshadri. Ce travail est consacré à la caractérisation suivante de la classe grand-Babel : pour tout polynôme T de degré ≤ 4 , considérons la courbe gauche $\mathcal{C}_T$ de $\mathbb{R}^3 \lambda \mapsto u(\lambda)$ de vecteur tangent $(\lambda^2, \lambda, 1)/T(\lambda)$. Soit Δ donné; nous montrons qu’il existe au plus une f.e.n. F de fonction variance $R\Delta + Q\sqrt{\Delta}$, qui admette une famille conjuguée au sens bayésien, et qui soit une famille exponentielle générale (f.e.g.) dont la f.e.n. associée soit concentrée sur $\mathcal{C}_T$. Nous montrons aussi qu’il y a une correspondance quasi bijective entre T et (R, Q) . Tout cela est précédé de considérations géométriques, d’exemples de f.e.n. de grand-Babel et de techniques probabilistes de construction de ces exemples.

ABSTRACT. — It is shown that a large number of natural exponential families (NEF) on $\mathbb{R}$ have variance functions of the form $R\Delta + Q\sqrt{\Delta}$ where R , Q and Δ are polynomials of degree ≤ 1 , ≤ 2 and ≤ 2 respectively. This class of NEF, called “grand-Babel” by G. Letac, generalizes the Morris and the Mora classes, contains the Babel and the Seshadri classes, and is closed by inversion. This paper characterizes the grand-Babel class as follows: for all polynomials T of degree ≤ 4 , consider the curve $\mathcal{C}_T$ of $\mathbb{R}^3 \lambda \mapsto u(\lambda)$ with $u'(\lambda) = (\lambda^2, \lambda, 1)/T(\lambda)$. Given Δ , we show that there exists at most one NEF with variance function $R\Delta + Q\sqrt{\Delta}$ which admits a conjugate family in the Bayesian sense and which is a general exponential family such that its associated NEF is concentrated on $\mathcal{C}_T$. We show also that there is a quasi-bijection between T and (R, Q) . This characterization is preceded by some examples, a probabilistic construction and a geometric study of the class.

(*) Reçu le 16 septembre 1994

(1) Université de Pau et des Pays de l’Adour, Dépt. de Statistiques et Traitements Informatiques des données, Av. de l’Université, F-64000 Pau (France)
E-mail : celestin.kokonendji@univ-pau.fr

1. Introduction

Nous commençons par rappeler le vocabulaire sur $\mathbb{R}$ des familles exponentielles naturelles et de leurs fonctions variances. Sur $\mathbb{R}^n$ ou un espace E approprié, on peut consulter [4] ou [22].

Soit $\mathcal{M} = \mathcal{M}(\mathbb{R})$ l'ensemble des mesures μ positives sur $\mathbb{R}$ (éventuellement non bornées) non concentrées en un point et telles que si la transformée de Laplace $L_\mu : \mathbb{R} \rightarrow]0, +\infty]$ est

$$L_\mu(\theta) := \int_{\mathbb{R}} \exp(\theta x) \mu(dx),$$

alors l'intérieur $\Theta(\mu)$ de l'intervalle $\{\theta \mid L_\mu(\theta) < +\infty\}$ est non vide. Pour une mesure μ (non-nécessairement une probabilité) fixée dans $\mathcal{M}$, l'ensemble $F = F(\mu)$ des probabilités sur $\mathbb{R}$:

$$P(\theta, \mu)(dx) = \exp(\theta x) \frac{\mu(dx)}{L_\mu(\theta)}, \quad (1.1)$$

quand θ varie dans $\Theta(\mu)$, est appelé la *famille exponentielle naturelle* (f.e.n.) engendrée par μ . Il est clair que si μ et μ' sont dans $\mathcal{M}$, alors les deux ensembles $F(\mu)$ et $F(\mu')$ coïncident si et seulement s'il existe (a, b) dans $\mathbb{R}^2$ tels que

$$\mu'(dx) = \exp(ax + b) \mu(dx).$$

Étant donnée une f.e.n. F , l'ensemble $\mathcal{B}_F = \{\mu \in \mathcal{M} \mid F = F(\mu)\}$ désigne l'ensemble des générateurs (ou bases) possibles de F . Évidemment $F \subset \mathcal{B}_F$.

Notant pour μ dans $\mathcal{M}$, $k_\mu := \text{Log } L_\mu$ la transformée cumulée de μ ; k_μ est analytique et strictement convexe dans $\Theta(\mu)$. De plus, pour θ dans $\Theta(\mu)$, la moyenne et la variance de $P(\theta, \mu)$ sont simplement $k'_\mu(\theta)$ et $k''_\mu(\theta)$ respectivement.

Définissons maintenant la fonction variance d'une f.e.n. Tout d'abord, il est raisonnable d'appeler l'intervalle ouvert $M_F = k'_\mu(\Theta(\mu))$ domaine des moyennes de F ; il dépend uniquement de F et non d'un μ particulier $\mathcal{B}_F$. Ensuite, si $\psi_\mu : M_F \rightarrow \Theta(\mu)$ est la réciproque de la fonction croissante k'_μ et si on définit, pour m dans M_F ,

$$P(m, F)(dx) = P(\psi_\mu(m), \mu)(dx)$$

alors l'application $M_F \rightarrow F$, $m \mapsto P(m, F)$ est une bijection; c'est la *paramétrisation de la f.e.n. F par la moyenne*. Enfin, la fonction

$$V_F : M_F \rightarrow]0, +\infty[, \quad m \mapsto \int_{\mathbb{R}} (x - m)^2 P(m, F)(dx)$$

est appelée la *fonction variance de F* ; il est alors facile de vérifier qu'on a

$$V_F(m) = k''_{\mu}(\psi_{\mu}(m)) = \frac{1}{\psi'_{\mu}(m)}, \quad \forall m \in M_F. \quad (1.2)$$

La fonction variance joue un rôle important en théorie de familles exponentielles. Depuis que C.-N. Morris (1982) a montré qu'elle caractérise la f.e.n. (i.e. $M_F = M_{F_1}$, $V_F = V_{F_1}$ si et seulement si $F = F_1$), la fonction variance présente généralement une forme plus simple que les mesures génératrices de la f.e.n. ou leurs transformées de Laplace. Cette observation a conduit (en commençant par C.-N. Morris) à chercher à décrire les f.e.n. ayant une fonction variance de forme donnée.

Ainsi, on a tout d'abord la classe de Morris [27], notée $\mathcal{M}_2$, des f.e.n. de fonction variance quadratique qui compte six types bien connus : normal, Poisson, binomiale, binomiale négative, gamma et cosinus hyperbolique. Ensuite, nous avons la classe de Mora [25] (voir aussi [24]) notée $\mathcal{M}_3$ des f.e.n. de fonction variance strictement cubique qui comporte également six types : gaussienne-inverse, Ressel-Kendall, Arcsinus strict, Abel, Takacs, Arcsinus large. Ces douze types de $\mathcal{M}_2 \cup \mathcal{M}_3$ sont des exemples fondamentaux de f.e.n. réelles.

Notons que pour des raisons discutées ailleurs (e.g. [24]), il n'est pas très intéressant de classer les f.e.n., dont la fonction variance est un polynôme de degré ≥ 4 . Signalons par contre, que les f.e.n. de fonctions variances puissances (i.e. $V_F(m) = Am^{\alpha}$, $(A, \alpha) \in \mathbb{R}^2$) et exponentielles (i.e. $V_F(m) = e^m$) ont été entièrement décrites par B. Jorgensen [13] (voir aussi Bar-Lev et Enis [2]) et [14, § 2.8] respectivement.

Cependant, dans la discussion de l'article de Jorgensen [13], G. Letac a mentionné l'importance des f.e.n. de fonction variance de la forme $P_1 + Q\sqrt{\Delta}$ où P_1 , Q et Δ sont des polynômes de degrés ≤ 3 , ≤ 2 et ≤ 2 . Mais peu de temps après, il a été observé dans Bar-Lev, Bshouty et Enis [1] et dans Letac [21] que cette dernière classe de f.e.n. s'avère trop vaste pour être étudiée, et il existe une généralisation plus petite et plus réaliste des classes $\mathcal{M}_2$ et $\mathcal{M}_3$. D'où la définition suivante.

DÉFINITION 1.1. — Une f.e.n. F sur $\mathbb{R}$ est dite de grand-Babel s'il existe trois polynômes R , Q et Δ de degrés ≤ 1 , ≤ 2 et ≤ 2 respectivement tels que V_F soit la restriction sur M_F de

$$R\Delta + Q\sqrt{\Delta}. \quad (1.3)$$

De plus, on dira que F est dégénérée si Δ est le carré d'un polynôme de degré ≤ 1 . Elle est dite non dégénérée dans le cas contraire. (Notons que le mot "Babel" est formé des premières lettres des noms des auteurs de [1] et de [21] cités ci-dessus.)

Cette classe de grand-Babel contient évidemment les éléments de $\mathcal{M}_2 \cup \mathcal{M}_3$ qui sont évidemment les seuls dégénérés dans la classe. Elle contient naturellement les dix-huit types de la classe de Babel (d'où vient le nom "grand-Babel") décrits par Letac ([21] ou [22]) et caractérisés par des fonctions variances de la forme (1.3), avec $R(m) = b$ (une constante réelle), $Q(m) = am + c$ et $\Delta(m)$ un polynôme de degré ≤ 2 . Notons que la classe de Babel (qui est à l'origine une extension de $\mathcal{M}_2$) admet une sous-classe remarquable notée $\mathcal{S}_2$ et formée des trois types "Hermite, Laguerre, non central χ^2 " dont les fonctions variances respectives peuvent se mettre sous la forme

$$\sqrt{\Delta}P(\sqrt{\Delta}), \quad (1.4)$$

où Δ est un polynôme de degré 1 et P un polynôme de degré ≤ 2 . Enfin la classe de grand-Babel contient aussi les types (SIG, SRK, SAb, STa, SSA et SLA) de la classe de Seshadri et notée $\mathcal{S}_3$ ([16] ou [17]). Cette classe de Seshadri est définie par les fonctions variances de la forme (1.4) avec P de degré 3 et Δ de degré 1. La transformation de Lindsay de $\mathcal{M}_3$ ([18] ou [19]) prend ses valeurs dans $\mathcal{S}_3$. Les fonctions variances des éléments de $\mathcal{S}_3$ sont donc de la forme (1.3) avec R et Δ de degré 1 et Q de degré ≤ 1 .

Une propriété importante de la classe de grand-Babel est qu'elle est fermée par "réciprocité". En effet, rappelons d'abord l'action de cette transformation sur les f.e.n. réelles introduites par Letac [20].

Soit F et F_1 deux f.e.n. sur $\mathbb{R}$. Le couple (F, F_1) est dit couple réciproque s'il existe un couple (μ, μ_1) dans $\mathcal{B}_F \times \mathcal{B}_{F_1}$ (appelé aussi "réciproque" en tant qu'élément de $\mathcal{M} \times \mathcal{M}$) tels que :

- i) les deux intervalles ouverts $\Theta^+(\mu) := \{\theta \in \Theta(\mu) \mid k'_\mu(\theta) > 0\}$ et $\Theta^+(\mu_1)$ sont non vides;
- ii) la fonction $-k_\mu : \Theta^+(\mu) \rightarrow \Theta^+(\mu_1)$ est une bijection de fonction réciproque $-k_{\mu_1}$.

De plus, on a la caractérisation suivante en terme de fonction variance : le couple (F, F_1) est réciproque si et seulement si :

- $M_F^+ := M_F \cap]0, +\infty[$ et $M_{F_1}^+$ sont non vides,
- l'application $m \mapsto 1/m$ restreinte à M_F^+ est une bijection dans $M_{F_1}^+$,
- pour tout m dans M_F^+ , on a

$$V_{F_1}(m) = m^3 V_F \left(\frac{1}{m} \right). \quad (1.5)$$

On en déduit que si une f.e.n. F est telle que $V_F = R\Delta + Q\sqrt{\Delta}$ avec $\deg R \leq 1$, $\deg Q \leq 2$ et $\deg \Delta \leq 2$ et s'il existe une f.e.n. F_1 telle que (F, F_1) soit un couple réciproque, alors on a :

$$V_{F_1} = R_1\Delta_1 + Q_1\sqrt{\Delta_1},$$

avec

$$R_1(m) = mR \left(\frac{1}{m} \right), \quad Q_1(m) = m^2 Q \left(\frac{1}{m} \right) \quad \text{et} \quad \Delta_1(m) = m^2 \Delta \left(\frac{1}{m} \right)$$

de degrés ≤ 1 , ≤ 2 et ≤ 2 respectivement.

Ce résultat permettrait la fabrication d'autres éléments de grand-Babel directement à partir de ceux de Babel (ex. 2.7) ou de Seshadri. Notons enfin que ce concept de réciprocité possède quelques interprétations probabilistes ([20], [22] ou [24]). Dans la section 2, nous donnerons (voire rappellerons) quelques exemples de f.e.n. de grand-Babel. Dans la section 3, nous donnerons quelques constructions probabilistes d'éléments de grand-Babel à partir des éléments dégénérés de $\mathcal{M}_2$ et de $\mathcal{M}_3$. Ces exemples sont toutefois insuffisants pour pouvoir décrire tous les types de la classe de grand-Babel. Mais avant de montrer dans la section 6 le principal résultat de ce travail, qui est une caractérisation de la classe grand-Babel, nous introduisons d'abord dans la section 4 la notion de "caste" qui sera suivie dans la section 5 de considérations géométriques associées à cette notion.

2. Exemples

Rappelons d'abord l'outil indispensable pour le calcul de la fonction variance d'une f.e.n. de grand-Babel ayant une transformée de Laplace explicite.

THÉOREME 2.1. — Soit A , B et C trois fonctions analytiques réelles sur l'intervalle ouvert M , et notons

$$(\tilde{A}, \tilde{B}, \tilde{C}) = (BC' - CB', CA' - AC', AB' - BA'),$$

$$D = B^2 - 4AC \quad \text{et} \quad \tilde{D} = \tilde{B}^2 - \tilde{A}\tilde{C}.$$

Supposons que A et $\tilde{D}$ n'ont pas de zéro sur M , et que D est strictement positive sur M . Si $\chi : M \rightarrow \mathbb{R}$ est une fonction continue telle que :

$$A\chi^2 + B\chi + C = 0 \quad \text{sur } M, \quad (2.1)$$

alors il existe ε dans $\{-1, 1\}$ tel que $\chi = (-B + \varepsilon\sqrt{D})/2A$, et χ est une fonction analytique. De plus si χ' n'a pas de zéro sur M , nous avons :

$$i) \quad 1/\chi' = (\tilde{C}D + (2A\tilde{B} + B\tilde{C})\varepsilon\sqrt{D})/2\tilde{D},$$

$$ii) \quad \chi/\chi' = (\tilde{B}D - (2C\tilde{C} + B\tilde{B})\varepsilon\sqrt{D})/2\tilde{D},$$

$$iii) \quad \chi^2/\chi' = (\tilde{A}D - (2C\tilde{B} + B\tilde{A})\varepsilon\sqrt{D})/2\tilde{D}.$$

Démonstration. — Voir [17, lemme 2.1].

Comme exemples d'utilisation, citons les propositions 2.2 et 2.3 de [17]. C'est un résultat très important par la suite et plus général que celui donné dans le cas Babel ([22, chap. 5, th. 3.1]).

Donnons maintenant quelques exemples d'éléments de grand-Babel.

Exemple 2.2 [15, sect. 6.2]. — Soit $a > 0$ et

$$\mu = \frac{1}{2}\delta_{-1} + a\delta_0 + \frac{1}{2}\delta_1.$$

Alors $\Theta(\mu) = \mathbb{R}$, $L_\mu(\theta) = a + \text{ch } \theta$, $M_F =]-1, 1[$ et

$$V_F(m) = \frac{1}{1-a^2} \left((a^2-1)m^2 + 1 - a\sqrt{(a^2-1)m^2 + 1} \right) \quad \text{si } a \neq 1$$

et

$$V_F(m) = \frac{1}{2}(1-m^2) \quad \text{si } a = 1.$$

On remarque évidemment que, si $a = 1$ alors $F = F(\mu)$ est du type binomial de $\mathcal{M}_2$, si $0 < a < 1$ alors F est du type "trinomial" de Babel

[22, chap. 5, sect. 6], et si $a > 1$ alors F est du type “convolution de deux binomiales” de Babel [22, chap. 5, sect. 7, (6)].

Exemple 2.3. — Soit $\mu = \sum_{n=0}^{\infty} \mu_n \delta_n$ dans $\mathcal{M}$ définie par sa fonction génératrice

$$\exp(1 - \sqrt{1 - z}) = \sum_{n=0}^{\infty} \mu_n z^n.$$

Alors $\Theta(\mu) =] - \infty, 0[$, $k_{\mu}(\theta) = 1 - \sqrt{1 - e^{\theta}}$, $M_F =] 0, +\infty[$ et

$$V_F(m) = m(m^2 + 1) + 3m^2 \sqrt{m^2 + 1}.$$

Une interprétation probabiliste de $F = F(\mu)$ est la suivante : soit (X, Y) une variable aléatoire (v.a.) de $\mathbb{R}^2$ telle que Y suit une loi gaussienne inverse et $\text{Loi}(X | Y)$ soit de Poisson de moyenne Y . Alors, on voit aisément que $\text{Loi}(X) \in F$ par le calcul de $\mathbb{E}(z^X)$.

Exemple 2.4 [8, corol. 5]. — Soit $B = (B_t)_{t \geq 0}$ un mouvement brownien standard. Notons :

$$\begin{aligned} \overline{B}_t &= \sup_{0 \leq s \leq t} B_s, \quad \tau = \inf\{t > 0 \mid B_t = \overline{B}_t = t\}, \\ \eta &= \int_0^{+\infty} \mathbf{1}_{\{B_t > t\}} dt \quad \text{et} \quad \xi = \inf\{t > 0 \mid \overline{B}_t \leq t\}. \end{aligned}$$

Alors

$$\mathbb{E}(e^{\theta\tau} \mid \tau < \infty) = \frac{2}{1 + \sqrt{1 - 2\theta}} = \mathbb{E}(e^{\theta\eta}) = \mathbb{E}(e^{\theta\xi})$$

pour $\theta < 1/2$. Par conséquent, si μ est la loi commune de η , ξ et τ ($\tau < \infty$) alors $M_{F(\mu)} =] 0, +\infty[$ et

$$V_{F(\mu)}(m) = \frac{1}{2} \left(m(m^2 + 4m - 4) + m^2 \sqrt{m^2 + 4m - 4} \right).$$

On peut consulter l'article de Imhof ([10, lemme 1] avec $\delta = 1$) pour la densité de μ .

Exemple 2.5 [12, (2.5)]. — Jorgensen a considéré la distribution

$$\mu(dx) = \frac{\lambda K_1(\lambda \sqrt{1 + x^2})}{\pi \sqrt{1 + x^2}} \exp\left(\lambda(xa + \sqrt{1 - a^2})\right) dx,$$

$\lambda \geq 0$, $|a| < 1$ et

$$K_1(\alpha) = \int_0^\infty \exp\left(-\frac{\alpha}{2}\left(y + \frac{1}{y}\right)\right) dy$$

est la fonction de Bessel modifiée de 3^{ième} espèce; μ est un cas spécial de la “distribution hyperbolique généralisée” de Barndorff-Nielsen [4]. Alors $\Theta(\mu) =]-1, 1[$, $k_\mu(\theta) = \sqrt{1-\theta^2}$, $M_F = \mathbb{R}$ et

$$V_F(m) = (1+m^2)^{3/2}.$$

Notons que $F = F(\mu)$ est sa propre réciproque au sens (1.5).

Interprétation probabiliste : soit B_1 et B_2 deux mouvements browniens standards indépendants, et soit

$$\tau = \inf\{t > 0 \mid B_2(t) + t + 1 = 0\} \quad \text{et} \quad +\infty \text{ si cet ensemble est vide.}$$

Alors

$$\text{Loi}(B_1(\tau) \mathbf{1}_{\tau < +\infty}) \in F = F(\mu).$$

Exemple 2.6 [9]. — Soit $X = (X_n)_{n=1}^\infty$ une suite de v.a. i.i.d. de Bernoulli telle que pour tout

$$n \geq 1, \quad P[X_n = \pm 1] = \frac{1}{2},$$

et soit $N = (N(t))_{t \geq 0}$ un processus de Poisson standard, indépendant de X . Si on note $S(t) \equiv X_1 + \cdots + X_{N(t)}$ (avec $S(t) = 0$ si $N(t) = 0$) et $\mu_1 = \text{Loi } S(1)$. Alors pour tout $p > 0$ la puissance $p^{\text{ième}}$ de μ_1 est

$$\mu_p(dx) = \sum_{k \in \mathbb{Z}} e^{-p} I_k(p) \delta_k(dx),$$

où

$$I_r(t) = \sum_{j=0}^{\infty} \frac{1}{j! \Gamma(j+r+1)} \left(\frac{t}{2}\right)^{2j+r} = I_{-r}(t)$$

est la fonction de Bessel modifiée avec $r \in \mathbb{R}_+$, et donc $\Theta(\mu_p) = \mathbb{R}$, $k_{\mu_p}(\theta) = p(\text{ch } \theta - 1)$, $M_{F(\mu_p)} = \mathbb{R}$ et

$$V_{F(\mu_p)}(m) = p \sqrt{1 + \left(\frac{m}{p}\right)^2}.$$

De plus, si ν_1 est la loi de $\tau_1 = \inf\{t > 0 \mid S(t) = 1\}$ restreint à $\tau_1 < +\infty$ alors pour tout $p > 0$ la puissance $p^{\text{ième}}$ de ν_1 est

$$\nu_p(dx) = p \frac{e^{-x}}{x} I_p(x) \mathbf{1}_{x>0} dx,$$

et donc

$$\Theta(\nu_p) =]-\infty, -1[, \quad k_{\nu_p}(\theta) = p \operatorname{Arg} \operatorname{ch} \theta,$$

$$M_{F(\nu_p)} =]0, +\infty[$$

et

$$V_{F(\nu_p)}(m) = m^2 \sqrt{1 + (pm)^2}.$$

Il est clair que $F(\mu_p)$ et $F(\nu_p)$ sont réciproques au sens (1.5). C'est un exemple typique d'interprétation probabiliste de réciprocité donnée par Letac ([20], [22] ou [24]).

Exemple 2.7. — (Construction par réciprocité : une application du théorème 5.3 de Letac et Mora [24].) Soit $\mu_t = \exp(t\gamma_1)$ dans $\mathcal{M}$ avec $t > 0$ et $\gamma_1(dx) = \mathbf{1}_{x>0} dx$. On a :

$$\Theta(\mu_t) =]-\infty, 0[, \quad k_{\mu_t}(\theta) = \frac{t}{-\theta}, \quad M_{F(\mu_t)} =]0, +\infty[$$

et

$$V_{F(\mu_t)}(m) = 2m \sqrt{\frac{m}{t}}.$$

Soit $X = (X(t))_{t \geq 0}$ le processus à accroissements indépendants et stationnaires gouverné par le semi-groupe de probabilités $(P(\theta, \mu_t))_{t \geq 1}$, avec $\theta < 0$, et soit

$$\tau_1 = \inf\{t \mid t - X(t) = 1\} \quad \text{et} \quad +\infty \text{ si cet ensemble est vide.}$$

Alors pour tout $p > 0$, on a :

- i) la loi ν_p de l'image de τ_1 par $x \mapsto x - p$ est la probabilité

$$\nu_p(dx) = e^{p/\theta} \delta_0(dx) + p \exp\left(\theta x + \frac{x+p}{\theta}\right) f(x^2 + px) \mathbf{1}_{x>0} dx$$

avec

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{(n!)^2(n+1)}$$

et $\theta \leq -1$ (notons que si $-1 < \theta < 0$, ν_p n'est pas une loi de probabilité);

ii) la transformée de Laplace de ν_p pour $\theta < 0$ est

$$L_{\nu_p}(s) = \exp \left\{ -\frac{p}{2} \left(s + \theta - \frac{1}{\theta} + \sqrt{\left(\theta - \frac{1}{\theta} \right)^2 + 2s \left(\theta + \frac{1}{\theta} \right) + s^2} \right) \right\}$$

pour tout $s > p(1 + \theta)$;

iii) la fonction variance est donnée sur $M_{F(\nu_p)} =]0, +\infty[$ par

$$V_{F(\nu_p)}(m) = \frac{2}{\sqrt{p}} m^{3/2} \left(1 + \frac{m}{p} \right)^{\frac{3}{2}}.$$

Remarques. — Pour i), on a utilisé un résultat dû à Zolotarev et simplement énoncé dans [24, th. 5.4, (5.17)]; pour ii), on s'est servi de la définition de couple réciproque de mesures et on en déduit que

$$P[\tau_1 < +\infty] = \begin{cases} 1 & \text{si } \theta \leq -1 \\ \exp \left(\frac{1}{\theta} - \theta \right) & \text{si } -1 < \theta < 0; \end{cases}$$

et pour iii), on a utilisé (1.5) puis la transformation affine (2.2) ci-dessous. Notons que ce genre d'application du théorème 5.3 (et éventuellement du théorème 5.6) de Letac et Mora [24] permet de construire des éléments de la classe de grand-Babel à partir de ceux bien connus des sous-classes non dégénérées de Babel et de Seshadri.

Nous terminons cette section en rappelant deux transformations élémentaires sur les f.e.n. (e.g. [22] ou [24]).

a) Affinité. — Soit F une f.e.n. et $\mathcal{A} : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto a_1 x + a_0$ avec $(a_1, a_0) \in \mathbb{R}^* \times \mathbb{R}$. Alors l'image $\mathcal{A}(F)$ des éléments de F par $\mathcal{A}$ est une f.e.n. telle que $M_{\mathcal{A}(F)} = \mathcal{A}(M_F)$ et

$$V_{\mathcal{A}(F)}(m) = a_1^2 V_F \left(\frac{m - a_0}{a_1} \right). \quad (2.2)$$

b) Puissance. — Soit $\mu_1 \in \mathcal{M}$ et $F_1 = F(\mu_1)$. L'ensemble de Jorgensen Λ_{μ_1} ou Λ_{F_1} est l'ensemble des réels $p > 0$ tels qu'il existe μ_p (notée aussi μ_1^{*p}) dans $\mathcal{M}$ avec $\Theta(\mu_p) = \Theta(\mu_1)$ et $k_{\mu_p}(\theta) = p k_{\mu_1}(\theta)$. Alors pour tout p dans Λ_{F_1} la f.e.n. $F_p = F(\mu_p)$ est telle que $M_{F_p} = p M_{F_1}$ et

$$V_{F_p}(m) = p V_{F_1} \left(\frac{m}{p} \right). \quad (2.3)$$

Notons que $\Lambda_{\mu_1} \cup \{0\}$ est un semi-groupe additif fermé de $]0, +\infty[$ et que la convolution $\mu_p \star \mu_q$ est égale à μ_{p+q} . De plus, on a toujours $\mathbb{N}^* \subset \Lambda_{\mu_1}$. En particulier si $\Lambda_{\mu_1} =]0, +\infty[$, μ_1 ou F_1 est dit indéfiniment divisible.

3. Visite de l'origine de processus gouvernés par les lois de Morris-Mora

Dans cette section, nous allons montrer deux résultats qui sont en fait deux techniques de construction probabiliste d'éléments de grand-Babel non dégénérés à partir de ceux de Mora $\mathcal{M}_3$.

Nous allons d'abord énoncer plus généralement chacun des résultats qui seront suivis des remarques d'application concrètes. Ensuite, nous donnerons des démonstrations plus loin. Nous noterons ici pour les v.a. X ,

$$L_X(\theta) = \mathbf{E}(e^{\theta X})$$

la transformée de Laplace de X ; et donc $k_X(\theta) = \text{Log } L_X(\theta)$ la transformée de cumulant de X , $F(X) = F(\text{Loi}(X))$ la f.e.n. engendrée par la Loi(X) est parfois $V_X = V_{F(X)}$ la fonction variance de $F(X)$.

PROPOSITION 3.1. — *Soit μ_1 une loi de probabilité sur $\mathbb{R}_+$ et indéfiniment divisible et soit $X = (X_t)_{t \geq 0}$ le processus de Lévy gouverné par μ_1 . Notons, pour tout $i = 0$ ou 1 ,*

$$\tau_i = \inf\{t > 0 \mid t - X_t = i\},$$

avec $\tau_i = +\infty$ si cet ensemble est vide. Alors

$$\frac{1}{1 - L_{\tau_0}(\theta)} = k'_{\tau_1}(\theta), \quad \forall \theta \in \Theta^+(\tau_1). \quad (3.1)$$

De plus, si μ_1 est telle que $V_{F(\mu_1)}(m) = m^2(a + bm)$ avec $M_{F(\mu_1)} =]0, +\infty[$ et $(a, b) \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$, alors on a

$$V_{F(\tau_0)}(m) = \frac{m}{2(a+b)} \left(\Delta(m) + (m-a)\sqrt{\Delta(m)} \right), \quad \forall m \in]0, +\infty[,$$

où $\Delta(m) = m^2 + 2(a+2b)m + a^2$.

Remarques. — Les f.e.n. de $\mathcal{M}_2 \cup \mathcal{M}_3$ (la classe Morris-Mora) engendrées par une telle μ_1 de la proposition 3.1 sont de trois types et appelées

- gamma ($a > 0$, $b = 0$ et donc $F(\tau_0)$ est du type Ressel-Kendall de $\mathcal{M}_3$);
- gaussienne inverse ($a = 0$, $b > 0$);
- Ressel-Kendall ($a > 0$, $b > 0$).

PROPOSITION 3.2. — Soit μ_1 une probabilité concentrée sur $\mathbb{N}$ telle que $\mu_1(\{0\}) > 0$, et soit $X = (X_n)_{n=1}^\infty$ une suite de v.a. i.i.d. de loi μ_1 . Notons $S_0 = 0$, $S_n = \sum_{k=1}^n (1 - X_k)$ et pour tout $i = 0$ ou 1 :

$$T_i = \inf\{n > 0 \mid S_n = i\},$$

avec $T_i = +\infty$ si cet ensemble est vide. Alors

$$\frac{1}{1 - L_{T_0}(\theta)} = k'_{T_1}(\theta), \quad \forall \theta \in \Theta^+(T_1). \quad (3.2)$$

De plus, si μ_1 est telle que :

i) $V_{F(\mu_1)}(m) = m(1 + am)$, avec $a \geq 0$ et $M_{F(\mu_1)} =]0, +\infty[$, alors

$$V_{F(T_0)}(m) = \frac{m}{a+1}(m-1)(m+a), \quad \forall m \in]1, +\infty[;$$

ii) $V_{F(\mu_1)}(m) = m(1 + 2\beta m + \alpha m^2)$, avec $\alpha > 0$, $1 + 2\beta > 0$ et $M_{F(\mu_1)} =]0, +\infty[$, alors

$$V_{F(T_0)}(m) = C_1 m \left(\Delta(m) + (m + C_2) \sqrt{\Delta(m)} \right), \quad \forall m \in]1, +\infty[,$$

avec $\Delta(m) = m^2 + 4(\alpha + \beta)m + 4(\beta^2 - \alpha)$, $C_1 = 1/(2(1 + 2\beta + \alpha))$
et $C_2 = -2(1 + \beta)$.

Remarques

A) Les f.e.n. engendrées par une telle μ_1 de i) sont de deux types dans $\mathcal{M}_2$ (classe de Morris [27]) : Poisson ($a = 0$) et binomiale négative ($a > 0$). Ainsi, la v.a. T_0 fournit une interprétation probabiliste des deux types correspondants de $\mathcal{M}_3$ (classe de Mora [24]) : Abel ($a = 0$) et Takaçs ($a > 0$).

B) Les f.e.n. $F(\mu_1)$ de ii) sont au nombre de quatre types dans $\mathcal{M}_3$: Abel, Takaçs, Arcsinus stricte ($\beta = 0$) et Arcsinus large.

Démonstration de la proposition 3.1

Notons $\overline{X}_t = t - X_t$. Puisque les lois de $\overline{X}_1$ et de τ_1 forment un couple réciproque (d'après [20, th. 2.1], [24, th. 5.3]), on a

$$-k_{\overline{X}_1}(-k_{\tau_1}(\theta)) = \theta, \quad \forall \theta \in \Theta^+(\tau_1).$$

Soit $T = \inf\{t > 0 \mid \overline{X}_t > 0\}$ avec $T = +\infty$ si cet ensemble est vide. Prabhu [28, th. 1] a montré que pour tout $\theta \in \Theta^+(\tau_1)$ et $\alpha \in \Theta^+(X_1)$:

$$\mathbb{E}\left(e^{\theta T + \alpha \overline{X}_T}\right) = 1 - \frac{k_{\overline{X}_1}(-k_{\tau_1}(\theta)) - k_{\overline{X}_1}(\alpha)}{-k_{\tau_1}(\theta) - \alpha}.$$

Donc, pour $\theta \neq \theta'$ dans $\Theta^+(\tau_1) \cap \Theta^+(X_1)$:

$$\begin{aligned} \mathbb{E}(e^{\theta T + \theta'(\tau_0 - T)}) &= 1 - \frac{k_{\overline{X}_1}(-k_{\tau_1}(\theta)) - k_{\overline{X}_1}(-k_{\tau_1}(\theta'))}{-k_{\tau_1}(\theta) + k_{\tau_1}(\theta')} \\ &= 1 - \frac{-\theta + \theta'}{-k_{\tau_1}(\theta) + k_{\tau_1}(\theta')}. \end{aligned}$$

En faisant tendre θ' vers θ , on trouve

$$\mathbb{E}(e^{\theta \tau_0}) = 1 - \frac{1}{k'_{\tau_1}(\theta)}, \quad \forall \theta \in \Theta^+(\tau_1).$$

D'où, la formule (3.1) est montrée; et donc, il existe un intervalle ouvert non vide Θ tel que $\Theta \subset \Theta(\tau_0) \cap \Theta(\tau_1)$.

Par conséquent, puisque

$$k'_{\tau_0}(\theta) = \frac{k''_{\tau_1}(\theta)}{k'_{\tau_1}(\theta)(k'_{\tau_1}(\theta) - 1)}, \quad \forall \theta \in \Theta,$$

on a, en posant $\overline{m} = k'_{\tau_0}(\theta)$ et $m = k'_{\tau_1}(\theta)$ pour tout $\theta \in \Theta$,

$$\overline{m} = \frac{V_{\tau_1}(m)}{m(m-1)}; \tag{3.3}$$

et donc, on peut utiliser ici la formule de l'action des transformations des f.e.n. sur les variances ([18, propo. 1.1] ou [19, th. 2.1]) :

$$V_{\tau_0}(\overline{m}) = V_{\tau_1}(m) \frac{d\overline{m}}{dm}. \tag{3.4}$$

Or, par (3.3), on a successivement

$$\begin{aligned}\frac{d\overline{m}}{dm} &= \frac{1}{m(m-1)} (V'_{\tau_1}(m) + (1-2m)\overline{m}) \\ &= \frac{1}{m(m-1)} (m-1) \left(\frac{d}{dm} \left(\frac{V_{\tau_1}(m)}{m-1} \right) - \overline{m} \right).\end{aligned}$$

Donc, en notant $U_1(m) = V_{\tau_1}(m)/(m-1)$ et en réutilisant (3.3), la formule (3.4) devient :

$$V_{\tau_0}(\overline{m}) = \overline{m}(m-1)(U'_1(m) - \overline{m}). \quad (3.5)$$

Supposons maintenant que $V_{X_1}(m) = V_{F(\mu_1)}(m) = m^2(a + bm)$ sur $M_{X_1} =]0, +\infty[$. Puisque $\overline{X}_1$ est l'image de X_1 par $x \mapsto 1-x$ et $(\overline{X}_1, \tau_1)$ est un couple réciproque, on a respectivement par (2.2) et (1.5)

$$\begin{aligned}V_{\tau_1}(m) &= m^3 V_{X_1} \left(\frac{m-1}{m} \right), \quad \forall m \in]0, 1[\\ &= m(m-1)^2 \left(a + b \frac{m-1}{m} \right), \quad \forall m \in]0, 1[.\end{aligned}$$

Donc, (3.3) est équivalent ici à

$$m^2(a+b) - m(\overline{m} + a + 2b) + b = 0 \quad (3.6)$$

et permet de vérifier aussi que $\overline{m} \in]0, +\infty[$ (car $(a, b) \in \mathbb{R}_+^2 \setminus \{(0, 0)\}$); la dérivée de $U_1(m) = V_{\tau_1}(m)/(m-1)$ devient ici :

$$U'_1(m) = 2m(a+b) - (a+2b). \quad (3.7)$$

Ainsi, la plus grande solution de (3.6) fournit $m \in]0, 1[$ en fonction de $\overline{m}$ par

$$m = \frac{\overline{m} + a + 2b + \sqrt{\Delta(\overline{m})}}{2(a+b)} = m(\overline{m}),$$

où

$$\Delta(\overline{m}) = \overline{m}^2 + 2\overline{m}(a+2b) + a^2. \quad (3.8)$$

(Notons que si $b = 0$, $\Delta(\overline{m}) = (\overline{m} + a)^2$ correspond au cas où X_1 suit une loi gamma et fournit τ_0 qui suit une loi Ressel-Kendall). Finalement en remplaçant m par $m(\overline{m})$ dans (3.7) puis dans (3.5) on vérifie facilement que

$$V_{\tau_0}(\overline{m}) = \overline{m} \left(\frac{\overline{m} - a + \sqrt{\Delta(\overline{m})}}{2(a+b)} \right) \sqrt{\Delta(\overline{m})},$$

avec $\Delta(\overline{m})$ donné en (3.8) et pour tout $\overline{m}$ dans $]0, +\infty[$. Ce qui achève la démonstration de la proposition 3.1. $\square$

Démonstration de la proposition 3.2

Nous nous contenterons de montrer la formule (3.2). Pour le reste, il suffit d'imiter la fin de la démonstration de la proposition 3.1 (i.e. appliquer (3.5) au cas i) et ii)).

Par le théorème du renouvellement, on a :

$$\frac{1}{1 - \mathbb{E}(e^{\theta T_0})} = \sum_{n=0}^{\infty} e^{\theta n} P[S_n = 0], \quad \forall \theta \in \Theta(T_0). \quad (3.9)$$

Pour tout $i = 0, 1, 2, \dots$, notons $\overline{X}_i = 1 - X_i$ et

$$g(z) = \mathbb{E}(z^{X_i}), \quad |z| < 1 \quad (3.10)$$

la fonction génératrice des v.a. i.i.d. X_i de loi μ_1 . On a donc

$$P[S_n = 0] = \frac{1}{n!} \left(\frac{d^n}{dz^n} (g(z))^n \right)_{z=0}.$$

Rappelons alors que la formule d'inversion de Lagrange (e.g. [7]) affirme :

- 1) qu'il existe un disque D centré en 0 et une fonction (implicite)
 $h : D \rightarrow \mathbb{C}$ analytique telle que $h(\omega) = \omega g(h(\omega))$;
- 2) que si f est analytique au voisinage de 0 alors

$$f(h(\omega)) = f(0) + \sum_{n=1}^{\infty} \frac{\omega^n}{n!} \left(\frac{d^{n-1}}{dz^{n-1}} f'(z)(g(z))^n \right)_{z=0}.$$

Appliquons cela à f telle que $f'(z) = 1/g(z)$ et $f(0) = 0$. On trouve donc

$$\begin{aligned} f(h(\omega)) &= \sum_{n=1}^{\infty} \frac{\omega^n}{n!} \left(\frac{d^{n-1}}{dz^{n-1}} (g(z))^{n-1} \right)_{z=0} \\ &= \sum_{n=0}^{\infty} \frac{\omega^{n+1}}{n+1} P[S_n = 0]. \end{aligned}$$

Puisqu'on a successivement

$$\frac{\omega h'(\omega)}{h(\omega)} = \frac{h'(\omega)}{g(h(\omega))} = f'(h(\omega)) h'(\omega) = \sum_{n=0}^{\infty} \omega^n P[S_n = 0],$$

on déduit alors en posant $\omega = e^\theta$ que (3.9) est équivalente à

$$\frac{1}{1 - \mathbb{E}(\omega^{T_0})} = \frac{\omega h'(\omega)}{h(\omega)}, \quad (3.11)$$

où h est l'unique fonction analytique au voisinage de 0 satisfaisant $h(0) = 0$, $h(\omega) = \omega g(h(\omega))$ avec g comme dans (3.10).

D'autre part, on a par (3.10)

$$e^{k_{\overline{X}_1}(y)} = e^y g(e^{-y}), \quad \forall y \in \Theta(\overline{X}_1)$$

et, puisque $(\overline{X}_1, T_1)$ forme un couple réciproque ([20, th. 2.2] ou [24, th. 5.6]), on a

$$-k_{\overline{X}_1}(-k_{T_1}(\theta)) = \theta, \quad \forall \theta \in \Theta^+(T_1).$$

Donc pour tout θ dans $\Theta^+(T_1)$, on a

$$g\left(e^{k_{T_1}(\theta)}\right) e^{-k_{T_1}(\theta)} = e^{-\theta}, \quad (3.12)$$

Autrement dit, si $h(e^\theta) = e^{k_{T_1}(\theta)}$ et $\omega = e^\theta$ alors (3.12) est équivalente à $\omega g(h(\omega)) = h(\omega)$ et on vérifie facilement que

$$\frac{\omega h'(\omega)}{h(\omega)} = k'_{T_1}(\theta). \quad (3.13)$$

Finalement en regardant (3.11) et (3.13), on obtient la formule (3.2). $\square$

4. Castes

Dans cette section, nous donnons la définition étendue à la classe de grand-Babel de la notion de “castes” introduite par Letac [22] et nous montrons quelques propriétés élémentaires. Nous commençons par rappeler comment on recouvre une f.e.n. à partir de sa fonction variance.

Si V_F est une fonction variance définie sur un intervalle M_F , il y a quatre étapes :

- i) calcul d'une primitive $\psi_\mu(m) = \int dm/V_F(m)$, où μ est dans $\mathcal{B}_F = \{\mu \in \mathcal{M} \mid F = F(\mu)\}$ (inverse de (1.2));

- ii) calcul de la fonction réciproque k'_μ de ψ_μ ;
- iii) calcul d'une primitive $\text{Log}(L_\mu(\theta)) = \int k'_\mu(\theta) d\theta$;
- iv) reconnaissance du type de la f.e.n. engendrée par une telle μ .

Supposons maintenant que $V_F = R\Delta + Q\sqrt{\Delta}$ soit une fonction variance de grand-Babel définie sur l'intervalle M_F .

La première étape de l'inversion ci-dessus est ici le calcul d'une intégrale abélienne standard. Ceci conduit donc à la subdivision des f.e.n. de grand-Babel suivante.

DÉFINITION 4.1. — *Soit F une f.e.n. de grand-Babel de fonction variance $V_F(m) = R(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$ sur M_F . Alors : F est dite de caste m (resp. $m^2 + 1$, $m^2 - 1$ et $1 - m^2$) s'il existe une affinité $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ telle que*

$$\Delta(\varphi(m)) = m \quad (\text{resp. } m^2 + 1, m^2 - 1 \text{ et } 1 - m^2) ;$$

F est dite de caste 1 si F est du type normal ou du type cosinus hyperbolique ; F est dite de caste m^2 si F est de la classe de Morris-Mora et non de caste 1.

Les castes 1 et m^2 sont dites dégénérées, toutes les quatre autres sont dites non dégénérées.

Citons comme exemples des f.e.n. des grand-Babel de :

- caste $1 - m^2$: exemple 2.2 avec $0 < a < 1$;
- caste $m^2 - 1$: exemple 2.4, exemple 2.7 iii) ; proposition 3.1 avec $a = 0$ et $b > 0$; proposition 3.2 ii) avec $\alpha = \beta^2$;
- caste $m^2 + 1$: exemple 2.2 avec $a > 1$; exemple 2.3 ; exemple 2.5 ; exemple 2.6 ; proposition 3.1 avec $a > 0$ et $b > 0$;
- caste m^2 : les six types de $\mathcal{M}_3$, les types Poisson, gamma, binomiale et binomiale négative de $\mathcal{M}_2$;
- caste m : les neuf types de $\mathcal{S}_2 \cup \mathcal{S}_3$;
- caste 1 : les types normal et cosinus hyperbolique de $\mathcal{M}_2$.

Rappelons que M_F (de la f.e.n. F telle que $V_F = R\Delta + Q\sqrt{\Delta}$) est inclus dans $I_\Delta = \{m \in \mathbb{R} \mid \Delta(m) > 0\}$.

Nous décrivons maintenant l'influence des trois transformations classiques des f.e.n. sur les castes : affinité – puissance – réciprocity.

PROPOSITION 4.2

- i) Chacune des six castes de la classe grand-Babel est fermée par affinité et par puissance.
- ii) Les trois groupes de castes suivantes sont fermés par réciprocité : castes 1 et m^2 , caste $m^2 + 1$ et castes m , $m^2 - 1$ et $1 - m^2$.

Démonstration. — Elle est acquise pour les castes 1 et m^2 (e.g. [23]) et nous la montrons pour les castes non dégénérées.

i) Soit F une f.e.n. de grand-Babel de caste donnée, c'est-à-dire il existe une affinité $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ telle que si $V_F = R\Delta + Q\sqrt{\Delta}$ alors $\Delta(\varphi(m))$ définit la caste.

Affinité. Soit $\mathcal{A} : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto a_1x + a_0$ avec $a_1 \neq 0$ (donc $\mathcal{A}^{-1}$ existe). Par (2.2), il est facile de voir que si $V_{\mathcal{A}(F)} = R_{\mathcal{A}}\Delta_{\mathcal{A}} + Q_{\mathcal{A}}\sqrt{\Delta_{\mathcal{A}}}$, avec

$$\Delta_{\mathcal{A}}(m) = \Delta\left(\frac{m - a_0}{a_1}\right) = \Delta(\mathcal{A}^{-1}(m)),$$

alors il existe $\varphi_1 = \mathcal{A} \circ \varphi$ telle que $\Delta_{\mathcal{A}}(\varphi_1(m)) = \Delta(\varphi(m))$.

Puissance. Soit p dans Λ_F et soit $h_p : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto px$ (donc $h_p^{-1} = h_{1/p}$ existe car $p \neq 0$). Par (2.3) et ce qui précède, il est immédiat que si $V_{F_p} = R_p\Delta_p + Q_p\sqrt{\Delta_p}$, avec

$$\Delta_p(m) = \Delta\left(\frac{m}{p}\right) = \Delta(h_p^{-1}(m)),$$

alors il existe $\varphi_2 = h_p \circ \varphi$ telle que $\Delta_p(\varphi_2(m)) = \Delta(\varphi(m))$.

ii) C'est une conséquence immédiate du théorème 5.2. $\square$

Nous terminons cette section par les observations suivantes.

- 1) Les castes 1, m , m^2 formées d'éléments de $\mathcal{M}_2$, $\mathcal{M}_3$, $\mathcal{S}_2$, $\mathcal{S}_3$ et $\mathcal{S}_4$ (où $\mathcal{S}_4$ est définie d'une manière analogue à $\mathcal{S}_3$) sont bien connues maintenant à l'exception de $\mathcal{S}_4$.
- 2) La caste $1 - m^2$ est rare dans la littérature. Le seul type à ma connaissance est le type trimonial de Babel : exemple 2.2 avec $0 < a < 1$.
- 3) Les castes $m^2 + 1$ et $m^2 - 1$ dont les éléments sont assez nombreux dans la nature pourraient être décrites par le concept de réciprocité (ex. 2.7).

5. Géométrie des castes

Dans cette section, nous donnons la structure géométrique dans $\mathbb{R}^3$ (théorème 5.2) des castes de grand-Babel; c'est un résultat préliminaire au principal (théorème 6.2) de ce travail.

Soit Δ un polynôme de degré ≤ 2 positif sur un intervalle ouvert non vide I_Δ tel qu'il existe une f.e.n. F de grand-Babel ayant $V_F = R\Delta + Q\sqrt{\Delta}$ sur M_F . Δ est donc une forme quadratique définie positive sur M_F . Ceci conduit, à l'aide de diverses techniques de calcul de fonction variance de grand-Babel (théorème 2.1), à associer à Δ l'image d'une droite de $\mathbb{R}^3$ par une forme quadratique.

Considérons donc l'espace $\mathbb{R}^3$ muni de la structure euclidienne canonique orientée (i.e. $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ est une base directe orthonormée) et définissons les deux formes quadratiques q et $\tilde{q}$ sur $\mathbb{R}^3$ par : pour tout $\vec{x} = (a, b, c)$, on a

$$q(\vec{x}) = b^2 - 4ac \quad \text{et} \quad \tilde{q}(\vec{x}) = b^2 - ac. \quad (5.1)$$

(Les signatures de $q(\vec{x}) = b^2 + (a - c)^2 - (a + c)^2$ et de $\tilde{q}(\vec{x}) = b^2 + ((a - c)/2)^2 - ((a + c)/2)^2$ sont $(+, +, -)$.)

Notons alors par $q(\cdot; \cdot)$ la forme polaire de q , on peut vérifier directement que

$$4\tilde{q}(\vec{x} \wedge \vec{y}) = (q(\vec{x}; \vec{y}))^2 - q(\vec{x})q(\vec{y}), \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^3. \quad (5.2)$$

Cette égalité fondamentale (5.2) qui relie les deux formes quadratiques q et $\tilde{q}$ est semblable à l'aire du parallélogramme de côtés $\vec{x}, \vec{y}$ dans $\mathbb{R}^3$:

$$\|\vec{x} \wedge \vec{y}\|^2 = \|\vec{x}\|^2 \|\vec{y}\|^2 - \langle \vec{x}, \vec{y} \rangle^2$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire canonique de $\mathbb{R}^3$.

Plus généralement, si q est une forme quadratique sur un espace euclidien orienté E de dimension 3, il existe une unique forme quadratique $\tilde{q}$ sur E telle que

$$\tilde{q}(\vec{x} \wedge \vec{y}) = (q(\vec{x}, \vec{y}))^2 - q(\vec{x})q(\vec{y}).$$

Une manière expéditive de le voir est de choisir une base orthonormale directe de E telle que l'endomorphisme symétrique de E associé à q soit diagonal dans cette base. Si cette diagonale est $(\lambda_1, \lambda_2, \lambda_3)$ l'endomorphisme

symétrique associé à $\tilde{q}$ sera diagonal dans la même base, avec pour diagonale $(-\lambda_2\lambda_3, -\lambda_3\lambda_1, -\lambda_1\lambda_2)$. Pour le voir il suffit d'observer l'identité :

$$\begin{aligned} & -\lambda_2\lambda_3(x_2y_3 - x_3y_2)^2 - \lambda_3\lambda_1(x_3y_1 - x_1y_3)^2 - \lambda_1\lambda_2(x_1y_2 - x_2y_1)^2 = \\ & = (\lambda_1x_1y_1 + \lambda_2x_2y_2 + \lambda_3x_3y_3)^2 + \\ & - \left(\lambda_1x_1^2 + \lambda_2x_2^2 + \lambda_3x_3^2 \right) \left(\lambda_1y_1^2 + \lambda_2y_2^2 + \lambda_3y_3^2 \right). \end{aligned}$$

Revenons aux formes quadratiques q et $\tilde{q}$ définies sur $\mathbb{R}^3$ par (5.1). Remarquons d'abord qu'il est facile de vérifier directement que la transformation

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (\vec{e}_0, \vec{e}_1) \mapsto (q(\vec{e}_0), q(\vec{e}_0; \vec{e}_1), q(\vec{e}_1))$$

est surjective (même pour $\vec{e}_0$ et $\vec{e}_1$ linéairement indépendants). Donc, étant donné le polynôme Δ , il existe toujours $(\vec{e}_0, \vec{e}_1)$ *indépendant* dans $\mathbb{R}^3$ tel que

$$\Delta(m) = q(\vec{e}_0 + m\vec{e}_1), \quad \forall m \in \mathbb{R};$$

notons alors $\mathcal{B}_\Delta$ l'ensemble des générateurs possibles $(\vec{e}_0, \vec{e}_1)$ de Δ . Pour tout $(\vec{e}_0, \vec{e}_1)$ dans $\mathcal{B}_\Delta$ avec $(\vec{e}_0 \wedge \vec{e}_1) = (\alpha, \beta, \gamma)$, on définit les quantités suivantes :

$$\hat{d} := \tilde{q}(\vec{e}_0 \wedge \vec{e}_1) \quad \text{et} \quad \hat{r} := \begin{cases} \beta^2/\alpha\gamma & \text{si } \alpha\gamma \neq 0 \\ \infty & \text{sinon.} \end{cases} \quad (5.3)$$

La proposition suivante montre la dépendance par rapport à Δ de ces dernières quantités $\hat{d}$ et $\hat{r}$.

PROPOSITION 5.1. — *Soit $\Delta(m) = d_0 + d_1m + d_2m^2$ avec $d_i \in \mathbb{R}$, $i = 0, 1, 2$. Pour tout $(\vec{e}_0, \vec{e}_1)$ dans $\mathcal{B}_\Delta$, on a :*

- i) $\tilde{q}(\vec{e}_0 \wedge \vec{e}_1) = (1/16)q(d_0, d_1, d_2)$, c'est-à-dire $\hat{d} = \hat{d}_\Delta$ dépend uniquement de Δ et non d'un couple particulier $(\vec{e}_0, \vec{e}_1)$ de $\mathcal{B}_\Delta$;
- ii) $\hat{r} = \hat{r}_{\vec{e}_0, \vec{e}_1}$ dépend uniquement du plan $\mathcal{H}_{\vec{e}_0, \vec{e}_1}$ engendré par $(\vec{e}_0, \vec{e}_1)$ de $\mathcal{B}_\Delta$.

Démonstration

i) Puisque $\Delta(m) = q(\vec{e}_0 + m\vec{e}_1) = q(\vec{e}_0) + 2q(\vec{e}_0; \vec{e}_1)m + q(\vec{e}_1)m^2$, on a

$$q(d_0, d_1, d_2) = 4 \left\{ (q(\vec{e}_0; \vec{e}_1))^2 - q(\vec{e}_0)q(\vec{e}_1) \right\};$$

et donc par la formule (5.2) avec $\vec{x} = \vec{e}_0$ et $\vec{y} = \vec{e}_1$, on obtient le résultat.

ii) Pour tout $(\vec{e}'_0, \vec{e}'_1)$ indépendant dans $\mathcal{H}_{\vec{e}_0, \vec{e}_1}$, il existe $\lambda \neq 0$ tel qu'on ait $\vec{e}'_0 \wedge \vec{e}'_1 = \lambda(\vec{e}_0 \wedge \vec{e}_1)$; d'où $\hat{r}_{\vec{e}'_0, \vec{e}'_1} = \hat{r}_{\vec{e}_0, \vec{e}_1}$. $\square$

Exemple. — Soit $\Delta(m) = m^2 + 1$. Si on prend $\vec{e}_0 = (1, 0, -1/4)$ et $\vec{e}_1 = (1, \sqrt{2}, 1/4)$, on vérifie que

$$(\vec{e}_0, \vec{e}_1) \in \mathcal{B}_\Delta, \quad \vec{e}_0 \wedge \vec{e}_1 = \left(\frac{\sqrt{2}}{3}, \frac{-1}{2}, \sqrt{2} \right),$$

$$\hat{d} = -\frac{1}{4} \quad \text{et} \quad \tilde{r}_{\vec{e}_0, \vec{e}_1} = \frac{1}{2}.$$

Mais si on considère $\vec{e}''_0 = (1, 1, 0)$ et $\vec{e}''_1 = (1, 1 + \sqrt{2}, (1 + 2\sqrt{2})/2)$, on aura alors

$$(\vec{e}''_0, \vec{e}''_1) \in \mathcal{B}_\Delta, \quad \vec{e}''_0 \wedge \vec{e}''_1 = \left(\frac{1 + \sqrt{2}}{2}, \frac{1 + \sqrt{2}}{2}, \sqrt{2} \right) \neq \lambda(\vec{e}_0 \wedge \vec{e}_1),$$

$$\hat{d} = -\frac{1}{4}$$

et enfin

$$\hat{r}_{\vec{e}''_0, \vec{e}''_1} = \frac{2 + \sqrt{2}}{2} \neq \hat{r}_{\vec{e}_0, \vec{e}_1}.$$

Pour énoncer enfin le principal résultat de cette section, notons par :

- $\mathcal{C} = \{\vec{x} \in \mathbb{R}^3 \mid q(\vec{x}) = 0\}$ le cône isotrope de q ,
- $\mathcal{C}_{\text{ext}} = \{\vec{x} \in \mathbb{R}^3 \mid q(\vec{x}) > 0\}$ l'extérieur de $\mathcal{C}$,
- $\mathcal{C}_{\text{Int}}^+ = \{\vec{x} = (a, b, c) \in \mathbb{R}^3 \mid q(\vec{x}) < 0 \text{ et } a + c > 0\}$;
- $\mathcal{C}_{\text{Int}}^- = \{\vec{x} = (a, b, c) \in \mathbb{R}^3 \mid q(\vec{x}) < 0 \text{ et } a + c < 0\}$ les deux parties intérieures de $\mathcal{C}$.

Nous désignerons aussi par $\mathcal{D}_{\vec{e}_0, \vec{e}_1}$ la droite $\{\vec{e}_0 + m\vec{e}_1 \mid m \in \mathbb{R}\}$, $\mathcal{H}_{\vec{e}_0, \vec{e}_1}$ le plan engendré par $(\vec{e}_0, \vec{e}_1)$ et passant par l'origine, et $\mathcal{B}_\Delta^*$ l'ensemble des $(\vec{e}_0, \vec{e}_1)$ de $\mathcal{B}_\Delta$ tels que si $(\alpha, \beta, \gamma) = \vec{e}_0 \wedge \vec{e}_1$ alors $\alpha\beta\gamma \neq 0$ (dans ce cas la quantité $\hat{r}$ définie en (5.3) est un réel non nul).

THÉORÈME 5.2. — *Soit Δ un polynôme de degré ≤ 2 et soit $(\vec{e}_0, \vec{e}_1)$ dans $\mathcal{B}_\Delta^*$ auquel on associe les quantités réelles $\hat{d}$ et $\hat{r}$ définies en (5.3). Notons $\mathcal{D} = \mathcal{D}_{\vec{e}_0, \vec{e}_1}$ et $\mathcal{H} = \mathcal{H}_{\vec{e}_0, \vec{e}_1}$. Alors :*

A) il y a équivalence entre :

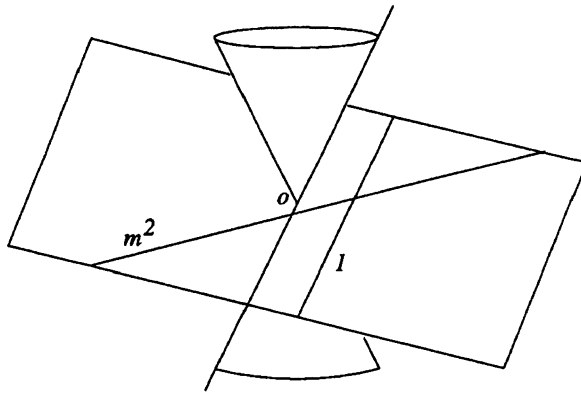
- i) Δ est dégénéré,
- ii) $\hat{d} = 0$,
- iii) $\hat{r} = 1$,
- iv) $\mathcal{H}$ est tangent à $\mathcal{C}$,
- v) ou bien $\mathcal{D}$ est parallèle à la génératrice $\mathcal{G}$ de $\mathcal{C}$ contenue dans $\mathcal{H}$ et Δ est de caste 1,
ou bien $\mathcal{D}$ coupe la génératrice $\mathcal{G}$ et Δ est de caste m^2 ;

B) il y a équivalence entre :

- i) Δ est de caste $m^2 + 1$,
- ii) $\hat{d} < 0$,
- iii) $0 < \hat{r} < 1$,
- iv) $\mathcal{H} \setminus \{0\} \subset \mathcal{C}_{\text{ext}}$,
- v) $\mathcal{D} \subset \mathcal{C}_{\text{ext}}$;

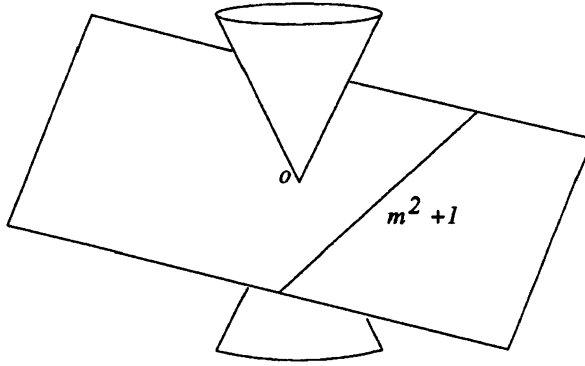
C) il y a équivalence entre :

- i) Δ est non dégénéré et n'est pas de caste $m^2 + 1$,
- ii) $\hat{d} > 0$
- iii) $\hat{r} \in]-\infty, 0[\cup]1, +\infty[$,

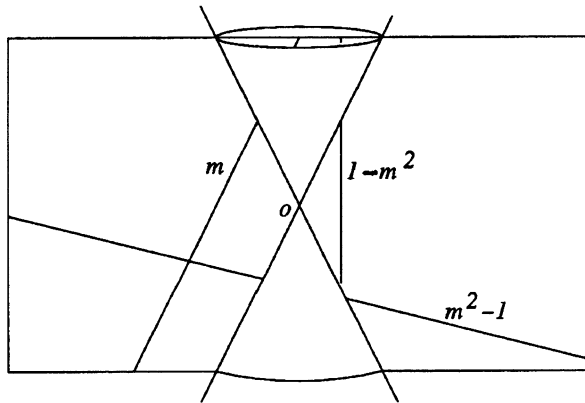


A)

- iv) $\mathcal{H}$ coupe symétriquement $\mathcal{C}_{\text{Int}}$ en deux droites $\mathcal{G}_1$ et $\mathcal{G}_2$ de $\mathcal{C}$,
v) ou bien $\mathcal{D}$ est parallèle à $\mathcal{G}_1$ ou $\mathcal{G}_2$, et Δ est de caste m ;
ou bien $\mathcal{D}$ coupe $\mathcal{G}_1$ et $\mathcal{G}_2$ avec $\mathcal{D} \cap \mathcal{C}_{\text{Int}}^+ \neq \emptyset$ ou $\mathcal{D} \cap \mathcal{C}_{\text{Int}}^- \neq \emptyset$, et Δ est de caste $m^2 - 1$;
ou bien $\mathcal{D}$ coupe $\mathcal{G}_1$ et $\mathcal{G}_2$ avec $\mathcal{D} \cap \mathcal{C}_{\text{Int}}^+ \neq \emptyset$ et $\mathcal{D} \cap \mathcal{C}_{\text{Int}}^- \neq \emptyset$, et Δ est de caste $1 - m^2$;
D) si $\Delta_1(m) = m^2 \Delta(1/m)$, on a $(\vec{e}_1, \vec{e}_0) \in \mathcal{B}_{\Delta_1}^*$.



B)



C)

Remarque. — Le point D) du théorème 5.2 permet donc de déduire facilement (à travers les exemples de figures ci-dessus) la proposition 4.2 ii).

Démonstration du théorème 5.2. — Notons $\vec{e}_i = (a_i, b_i, c_i)$, $i = 0, 1$ et $\vec{e}_0 \wedge \vec{e}_1 = (\alpha, \beta, \gamma)$.

A) i) $\Leftrightarrow$ ii) : Immédiate d'après la proposition 5.1 i).

ii) $\Leftrightarrow$ iii) : Immédiate d'après l'égalité $\hat{d}/(\alpha\gamma) = \hat{r} - 1$, car $\alpha\gamma \neq 0$.

ii) $\Rightarrow$ iv) : Supposons $\hat{d} = 0$, c'est-à-dire $\beta^2 = \alpha\gamma$. Pour montrer que $\mathcal{H}$ est tangent à $\mathcal{C}$, il suffit de montrer que la forme quadratique q restreint à $\mathcal{H}$ est dégénérée de signature $(+, 0)$. En effet, soit $\vec{x} = (x_1, x_2, x_3)$ dans $\mathcal{H} \cap \mathcal{C}$, c'est-à-dire $\langle \vec{x}, \vec{e}_0 \wedge \vec{e}_1 \rangle = 0$ et $q(\vec{x}) = 0$, ou encore

$$\alpha x_1 + \beta x_2 + \gamma x_3 = 0 \quad (*)$$

$$x_2^2 - 4x_1x_3 = 0. \quad (**)$$

Puisque $\beta \neq 0$, $(*)$ implique que

$$x_2 = \frac{-1}{\beta} (\alpha x_1 + \gamma x_3).$$

Et donc, par $(**)$, la forme quadratique q restreint à $\mathcal{H}$ (d'une manière générale) est

$$\begin{aligned} q|_{\mathcal{H}}(\vec{x}) &= x_2^2 - 4x_1x_3 \\ &= \frac{1}{\beta^2} (\alpha^2 x_1^2 + 2(\alpha\gamma - 2\beta^2)x_1x_3 + \gamma^2 x_3^2) \\ &= \frac{1}{\beta^2} \begin{pmatrix} x_1 & x_3 \end{pmatrix} \begin{pmatrix} \alpha^2 & \alpha\gamma - 2\beta^2 \\ \alpha\gamma - 2\beta^2 & \gamma^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} \end{aligned}$$

et de déterminant est

$$\begin{aligned} \beta^2 \det[q|_{\mathcal{H}}] &= 4(\alpha\gamma - \beta^2) \\ &= -4\tilde{q}(\vec{e}_0 \wedge \vec{e}_1). \end{aligned}$$

D'où, sous l'hypothèse $\beta^2 = \alpha\gamma$, la forme quadratique $q|_{\mathcal{H}}$ est dégénérée de signature $(+, 0)$.

iv) $\Rightarrow$ v) : Supposons $\mathcal{H}$ tangent à $\mathcal{C}$. Puisque

$$\mathcal{D} = \{\vec{e}_0 + m\vec{e}_1 \mid m \in \mathbb{R}\} \subset \mathcal{H}$$

avec $\Delta(m) = q(\vec{e}_0 + m\vec{e}_1) = (d_0 + md_1)^2$, on a donc :

- ou bien $d_1 = 0$, c'est-à-dire $q(\vec{e}_1) = 0$; d'où, $\mathcal{D}$ est parallèle à la tangente $\mathcal{G}$ à $\mathcal{C}$ contenue dans $\mathcal{H}$;
- ou bien $d_1 \neq 0$, c'est-à-dire $q(\vec{e}_1) > 0$; d'où, $\mathcal{D}$ coupe $\mathcal{G}$.

v) $\Rightarrow$ ii) : Sous l'hypothèse v) Δ est dégénéré. Donc, la proposition 5.1 i) implique que $\hat{d} = \hat{d}_\Delta = 0$.

B) i) $\Leftrightarrow$ ii) : Immédiate d'après la proposition 5.1 i).

ii) $\Leftrightarrow$ iii) : Si $\hat{d} = \beta^2 - \alpha\gamma < 0$, c'est-à-dire $0 < \beta^2 < \alpha\gamma$, on a évidemment $\hat{r} = \beta^2/(\alpha\gamma) \in]0, 1[$; et, réciproquement.

ii) $\Leftrightarrow$ iv) : Puisque d'après (5.4) la forme quadratique q restreint à $\mathcal{H} \setminus \{0\}$ est de signature $(+, +)$ si et seulement si $-\hat{r} = \alpha\gamma - \beta^2 > 0$, on en déduit cette équivalence.

iv) $\Leftrightarrow$ v) : Triviale (car $\mathcal{D} \subset \mathcal{H} \setminus \{0\}$ et $\mathcal{H} = (0, \mathcal{D})$).

C) i) $\Leftrightarrow$ ii) : Immédiate d'après la proposition 5.1 i).

ii) $\Leftrightarrow$ ii) : En effet, puisque $\alpha\beta\gamma \neq 0$, on a

$$\hat{d} = \beta^2 - \alpha\gamma > 0 \quad \Leftrightarrow \quad \begin{cases} \beta^2 > \alpha\gamma > 0 \Leftrightarrow \hat{r} = \frac{\beta^2}{\alpha\gamma} \in]1, +\infty[\\ \text{ou} \\ \beta^2 > \alpha\gamma \text{ et } \alpha\gamma < 0 \Leftrightarrow \hat{r} = \frac{\beta^2}{\alpha\gamma} \in]-\infty, 0[\end{cases}$$

D'où l'équivalence montrée.

ii) $\Leftrightarrow$ iv) : C'est acquis car cette équivalence se traduit comme suit : la forme quadratique $q|_{\mathcal{H}}$ est de signature $(+, -)$ si et seulement si $-\hat{d} = \alpha\gamma - \beta^2 < 0$ (voir formule (5.4)).

iv) $\Leftrightarrow$ v) : Cette équivalence est immédiate après avoir observé les trois points suivants, avec

$$\Delta(m) = q(\vec{e}_0 + \vec{e}_1) = d_0 + d_1m + d_2m^2.$$

- 1) Δ est de caste $m \Leftrightarrow d_2 = 0$ et $\hat{d} > 0 \Leftrightarrow \vec{e}_1 \in \mathcal{C}$. Bien entendu on peut avoir en plus $d_0 = 0$, mais dans ce cas $\vec{e}_0 \in \mathcal{C}$ n'est pas colinéaire à $\vec{e}_1$.

- 2) Δ est de caste $m^2 - 1 \Leftrightarrow d_2 > 0$ avec $\hat{d} > 0 \Leftrightarrow \vec{e}_1 \in \mathcal{C}_{\text{ext}}$ avec $\mathcal{H} \cap \mathcal{C}_{\text{Int}} \neq \emptyset \Leftrightarrow \mathcal{D}$ coupe une seule nappe ($\mathcal{C}_{\text{Int}}^-$ ou $\mathcal{C}_{\text{Int}}^+$) de $\mathcal{C}$.
- 3) Δ est de caste $1 - m^2 \Leftrightarrow d_2 < 0$ avec $d > 0 \Leftrightarrow \vec{e}_1 \in \mathcal{C}_{\text{Int}}$ avec $\mathcal{H} \cup \mathcal{C}_{\text{ext}} \neq \emptyset \Leftrightarrow \mathcal{D}$ coupe les deux nappes de $\mathcal{C}$.

D) Trivial. $\square$

6. Théorème de caractérisation

Cette section est consacrée à la caractérisation (théorème 6.2) des f.e.n. de grand-Babel à Δ fixé. Pour cela nous avons besoin de préliminaires en théorie bayésienne des familles exponentielles. Commençons par la définition d'une famille exponentielle générale sur $\mathbb{R}^n$.

Soit un espace mesuré (Ω, ξ, ν) et une application $t : \Omega \rightarrow \mathbb{R}^n$ mesurable telle que la mesure image $t_*\nu = \rho$ (de ν par t) est un élément de $\mathcal{M}(\mathbb{R}^n)$. Alors l'ensemble $F(t, \nu)$ des probabilités sur Ω :

$$P(\theta, t, \nu)(d\omega) = \exp \left\{ \langle \theta, t(\omega) \rangle \right\} \frac{\nu(d\omega)}{L_\rho(\theta)}, \quad (6.1)$$

quand θ varie dans $\Theta(\rho)$, est appelé la *famille exponentielle générale* (f.e.g.) engendrée par (t, ν) . Dans ce cas, $F(\rho) = F(t_*\nu)$ est appelée la *f.e.n. associée* à la f.e.g. $F(t, \nu)$.

6.1. Théorie bayésienne

On considère deux espaces métriques séparables munis de leur tribu de Borel $(\Lambda, \sigma(\Lambda))$ et $(\mathcal{X}, \sigma(\mathcal{X}))$. Un *noyau de transition* $K = (K_\lambda)_{\lambda \in \Lambda}$ de Λ vers $\mathcal{X}$ est une application $(\lambda, S) \mapsto K_\lambda(S)$ de $\Lambda \times \mathcal{X} \rightarrow [0, 1]$ telle que :

- i) $\lambda \mapsto K_\lambda(S)$ soit $\sigma(\Lambda)$ -mesurable,
- ii) $S \mapsto K_\lambda(S)$ soit une probabilité.

On munit l'espace $(\Lambda, \sigma(\Lambda))$ d'une probabilité Π dite *a priori*, ce qui équipe l'espace produit $(\Lambda \times \mathcal{X}, \sigma(\Lambda) \otimes \sigma(\mathcal{X}))$ d'une probabilité P_Π définie par

$$P_\Pi(d\lambda, ds) = \Pi(d\lambda)K_\lambda(ds). \quad (6.2)$$

En d'autres termes si (X, Y) est une v.a de $\Lambda \times \mathcal{X}$ telle que $P_\Pi = \text{Loi}(X, Y)$, alors $\text{Loi}(X) = \Pi$ et $\text{Loi}(Y \mid X = \lambda) = K_\lambda$.

L'idée bayésienne consiste alors, en partant de la loi de probabilité $P_\Pi = \Pi \otimes K$ définie en (6.2) sur un produit $\Lambda \times \mathcal{X}$ de deux espaces mesurables, à *désintégrer* ce P_Π suivant $\mathcal{X}$, c'est-à-dire à introduire :

- la mesure image $\mathcal{W}_\Pi$ de P_Π par $(\lambda, s) \mapsto s$,
- un noyau de transition $\tilde{\Pi} = (\tilde{\Pi}_s)_{s \in \mathcal{X}}$ de $\mathcal{X}$ vers Λ tel que

$$P_\Pi(d\lambda, ds) = \mathcal{W}_\Pi(ds) \tilde{\Pi}_s(d\lambda). \quad (6.3)$$

Ce noyau de transition $\tilde{\Pi}$ est unique dans le sens suivant : si $\tilde{\Pi}' = (\tilde{\Pi}'_s)_{s \in \mathcal{X}}$ a la même propriété, alors $\tilde{\Pi}'_s = \tilde{\Pi}_s$ presque sûrement par rapport à $\mathcal{W}_\Pi$. La probabilité $\tilde{\Pi}_s(d\lambda)$ est appelée la *loi a posteriori* sur Λ (en ayant observé s de $\mathcal{X}$).

Après ces notions de probabilités a priori et probabilités a posteriori sur l'espace des paramètres Λ , définissons ce qu'est une "famille a priori conjuguée". En statistique bayésienne, c'est un critère de sélection judicieux des lois a priori (étant donné un noyau de transition).

DÉFINITION 6.1. — Soit $K = (K_\lambda)_{\lambda \in \Lambda}$ un noyau de transition de Λ vers $\mathcal{X}$ et soit $\mathcal{F}$ une famille de probabilités a priori Π sur Λ . Alors $\mathcal{F}$ sera dite conjuguée à K si pour tout Π dans $\mathcal{F}$ et pour toute observation s dans $\mathcal{X}$, la probabilité a posteriori $\tilde{\Pi}_s$ est encore dans $\mathcal{F}$ presque sûrement par rapport à $\mathcal{W}_\Pi$.

6.2. Deux exemples de famille conjuguée à une famille exponentielle

- En pratique on part d'un modèle exponentiel naturel $F = F(\mu)$ sur $\mathbb{R}^n$ comme noyau de transition, et on se contente de considérer pour ce premier exemple la famille $\mathcal{F}$ des lois a priori Π sur $\Theta(m)$ suivante :

α) on se donne une mesure positive $\nu(d\theta)$ sur $\Theta(\mu)$ telle que si $t : \theta \rightarrow (\theta, -k_\mu(\theta))$, alors $\rho := t_*\nu \in \mathcal{M}(\mathbb{R}^{n+1})$; (ainsi la f.e.n. associée $F(\rho)$ est concentrée sur le graphe de t restreint à $\Theta(\mu)$);

β) on définit alors pour tout $(p, x) \in \mathbb{R} \times \mathbb{R}^n$ avec $(p, x) \in \Theta(\rho)$:

$$\Pi(d\theta) = \Pi_{p,x}(d\theta) = K(p, x) \exp\{\langle \theta, x \rangle - pk_\mu(\theta)\} \nu(d\theta), \quad (6.4)$$

où $K(p, x)$ est la constante de normalisation; on remarque que $\{\Pi_{p,x} \mid (p, x) \in \Theta(\rho)\}$ est une f.e.g. avec $\Omega = \Theta(\rho)$ et $t(\theta) = (\theta, -k_\mu(\theta))$.

Diaconis et Ylvisaker [6] prennent $\nu(d\theta) = \mathbf{1}_{\Theta(\mu)}(\theta) d\theta$ et montrent en plus que si $F = F(\mu)$ est *steep* (i.e. le domaine des moyennes M_F de F est égal à l'intérieur de l'enveloppe convexe fermée du support de F), alors

$$\Theta(\rho) \supset \{(p, x) \in \mathbb{R} \times \mathbb{R}^n \mid p > 0 \text{ et } x \in M_F\}.$$

Autre choix : Jeffreys [11] a considéré $\nu(d\theta) = \sqrt{k''_\mu(\theta)} d\theta$ si $n = 1$.

On observe alors que pour tout ν satisfaisant (α)

$$\mathcal{F} = \mathcal{F}(\nu, \mu) = \{\Pi_{p,x} \mid (p, x) \in \Theta(t_*\nu)\}$$

est conjuguée à $F = F(\mu)$. En effet, puisque pour tout $\Pi = \Pi_{p,x} \in \mathcal{F}$ et pour toute observation $Y \in \mathbb{R}^n$

$$\begin{aligned} \Pi_{p,x}(d\theta)P_\Pi(\theta, \mu)(dy) &= \\ &= K(p, x) \exp\{\langle \theta, x + y \rangle - (p + 1)k_\mu(\theta)\} \nu(d\theta)\mu(dy). \end{aligned}$$

On en déduit facilement que

$$\tilde{\Pi}_y = \Pi_{p+1, x+y}$$

est dans $\mathcal{F}$ presque sûrement par rapport à

$$\mathcal{W}_\Pi(dy) = \frac{K(p, x)}{K(p + 1, x + y)} \mu(dy).$$

• Le second exemple est celui qui est important pour la caractérisation des f.e.n. de grand-Babel.

Soit un polynôme T de degré ≤ 4 et un intervalle ouvert $J \subset \mathbb{R}$ ne contenant pas de zéros de T . Soit

$$\vec{u} : \lambda \mapsto \vec{u}(\lambda) = (u_2(\lambda), u_1(\lambda), u_0(\lambda))$$

l'application de J dans $\mathbb{R}^3$ telle que le graphe de $\vec{u}$ soit la courbe gauche $\mathcal{C}_T$ de $\mathbb{R}^3$ de vecteur tangent $\vec{u}'(\lambda) = (\lambda^2, \lambda, 1)/T(\lambda)$. L'application $\vec{u}$ est définie à une constante additive près. Soit également une mesure positive ν sur J telle que $\rho = \vec{u}_*\nu \in \mathcal{M}(\mathbb{R}^3)$ et considérons la f.e.g. $\mathcal{F}$ de probabilités a priori Π sur J définies par

$$\Pi(d\lambda) = \Pi_{\vec{A}}(d\lambda) = K(\vec{A}) \exp\left\{\langle \vec{u}(\lambda), \vec{A} \rangle\right\} \nu(d\lambda), \quad (6.5)$$

où $\vec{A} \in \Theta(\rho)$ et où $K(\vec{A})$ est la constante de normalisation. Une écriture plus détaillée de (6.5) avec $\vec{A} = (a, b, c)$ est

$$\begin{aligned}\Pi(d\lambda) &= \Pi_{a,b,c}(d\lambda) \\ &= K(a, b, c) \exp\{au_2(\lambda) + bu_1(\lambda) + cu_0(\lambda)\} \nu(d\lambda).\end{aligned}$$

Comme exemples de polynômes T remarquables associés à $\nu(d\lambda) = d\lambda$, citons :

1. La famille gaussienne inverse généralisée

On prend $T(\lambda) = \lambda^2$ et $J =]0, +\infty[$. On voit alors que la famille $\mathcal{F}$ de probabilité a priori $\Pi = \Pi_{a,b,c}$ avec

$$\Pi_{a,b,c}(d\lambda) = K(a, b, c) \lambda^b e^{(a\lambda - \frac{c}{\lambda})} \mathbf{1}_{]0, +\infty[}(\lambda) d\lambda$$

est la famille gaussienne inverse généralisée si $a < 0$ et $c > 0$. Rappelons que la gaussienne inverse classique est obtenue pour $b = -3/2$.

2. Les familles hypergéométriques

On prend $T(\lambda) = \lambda(1 - \lambda)(1 - \lambda z)$ avec $z < 1$ et $J =]0, 1[$. On vérifie que $\mathcal{F}$, définie par

$$\Pi_{a,b,c}(d\lambda) = K(a, b, c) \lambda^{b-1} (1 - \lambda)^{c-b-1} (1 - \lambda z)^{-a} \mathbf{1}_{]0,1[}(\lambda) d\lambda,$$

est une famille hypergéométrique avec $a > 0$ et $c > b$.

3. Autres exemples

$T(\lambda) = \lambda^2(\lambda^2 - 1)$ sur $J =]1, +\infty[$ fournit une famille $\mathcal{F}$ de distributions simplex due à Seshadri, $T(\lambda) = \lambda^3$ et $T(\lambda) = \lambda$ sur $J =]0, +\infty[$ donnent respectivement des exemples dus à Larcher et Haphen [26].

Ensuite on considère (comme noyau de transition) une f.e.n. sur $\mathbb{R}$ $F = F(\mu) = \{P(\lambda, \mu) \mid \lambda \in J\}$ telle qu'il existe une droite $\mathcal{D} = \mathcal{D}_{\vec{e}_0, \vec{e}_1} = \{\vec{e}_0 + x\vec{e}_1 \mid x \in \mathbb{R}\}$ de $\mathbb{R}^3$ vérifiant ceci : si $\rho = \varphi_* \mu$ avec $\mu \in \mathcal{B}_F$ et $\varphi : \mathbb{R} \rightarrow \mathcal{D}$, $x \mapsto \vec{e}_0 + x\vec{e}_1$, on a alors

$$\int_{\mathcal{D}} \exp\{au_2(\lambda) + bu_1(\lambda) + cu_0(\lambda)\} \rho(da, db, dc) = 1, \quad \forall \lambda \in J,$$

ce qui est équivalent à

$$\int_{\mathcal{D}} e^{\langle \vec{u}(\lambda), \vec{A} \rangle} \rho(d\vec{A}) = 1, \quad \forall \lambda \in J$$

ou encore

$$\int_{\mathbb{R}} e^{\langle \vec{u}(\lambda), \vec{e}_0 + x\vec{e}_1 \rangle} \mu(dx) = 1, \quad \forall \lambda \in J. \quad (6.6)$$

Dans ce cas la famille $\mathcal{F}$ des probabilités a priori (6.5) est conjuguée à $F = F(\mu)$, où μ satisfait (6.6). En effet, puisque pour tout $\Pi = \Pi_{\vec{A}} \in \mathcal{F}$ et pour toute observation $x \in \mathbb{R}$:

$$\begin{aligned} \Pi_{\vec{A}}(d\lambda) P_{\Pi}(\lambda, \mu)(dx) &= K(\vec{A}) e^{\langle \vec{u}(\lambda), \vec{A} \rangle} e^{x\theta(\lambda) - k_{\mu}(\theta(\lambda))} \nu(d\lambda) \mu(dx) \\ &= K(\vec{A}) e^{\langle \vec{u}(\lambda), \vec{A} + \vec{e}_0 + x\vec{e}_1 \rangle} \nu(d\lambda) \mu(dx) \end{aligned}$$

(car (6.6) implique que $\theta(\mu) = \langle u(\lambda), \vec{e}_1 \rangle$ et $-k_{\mu}(\theta(\lambda)) = \langle \vec{u}(\lambda), \vec{e}_0 \rangle$). On en déduit aisément que la probabilité a posteriori correspondante

$$\tilde{\Pi}_x = \Pi_{\vec{A} + \vec{e}_0 + x\vec{e}_1}$$

est dans $\mathcal{F}$ presque sûrement par rapport à

$$\mathcal{W}_{\Pi}(dx) = \frac{K(\vec{A})}{K(\vec{A} + \vec{e}_0 + x\vec{e}_1)} \mu(dx).$$

Cet exemple est dû essentiellement à G. Letac qui a étudié avec S. K. Bar-Lev et P. Enis un problème plus général. Ainsi, on peut consulter le travail original [3] (ou encore [22]) sur le fait qu'étant donné $(T; \vec{e}_0, \vec{e}_1)$, l'ensemble des $\mu \in \mathcal{M} = \mathcal{M}(\mathbb{R})$ satisfaisant (6.6) est souvent vide.

6.3. Caractérisation des f.e.n. de grand-Babel

Voici maintenant notre résultat principal (nous gardons les notations introduites à la section 5) :

THÉORÈME 6.2. — *Soit T un polynôme de degré ≤ 4 , J un intervalle ne contenant pas de zéro de T et $\vec{u} : J \rightarrow \mathbb{R}^3$ telle que sa dérivée soit*

$$\vec{u}'(\lambda) = \left(\frac{\lambda^2}{T(\lambda)}, \frac{\lambda}{T(\lambda)}, \frac{1}{T(\lambda)} \right).$$

Soit $(\vec{e}_0, \vec{e}_1)$ un couple de vecteurs indépendants dans $\mathbb{R}^3$ tels que les quantités $\hat{d}$ et $\hat{r}$ correspondantes et définies en (5.3) soient respectivement dans $\mathbb{R} \setminus \{0\}$ et $\mathbb{R}$. On note $\Delta(m) = \Delta_{\vec{e}_0, \vec{e}_1}(m) = q(\vec{e}_0 + m\vec{e}_1)$ avec q défini en (5.1).

Supposons qu'il existe μ dans $\mathcal{M}$ vérifiant

$$\int_{\mathbb{R}} \exp \{ \langle \vec{e}_0 + x\vec{e}_1, \vec{u}(\lambda) \rangle \} \mu(dx) = 1 \quad (6.6)$$

pour tout λ dans J . Alors $F = F(\mu)$ est une f.e.n. de grand-Babel non dégénérée avec $V_F = R\Delta + Q\sqrt{\Delta}$ sur M_F .

Inversement, si F est une f.e.n. de grand-Babel non dégénérée avec $V_F = R\Delta + Q\sqrt{\Delta}$ sur M_F et si on considère $(\vec{e}_0, \vec{e}_1)$ dans $\mathcal{B}_\Delta^*$ tels que de plus la quantité associée $\hat{r}$ n'appartienne pas à $\{2, (3 \pm \sqrt{5})/2\}$, alors il existe un unique polynôme T de degré ≤ 4 tel qu'il existe μ dans $\mathcal{B}_F$ vérifiant (6.6).

Avant de faire la démonstration, voici une série de cinq remarques.

Remarques

1) Un résultat analogue pour caractériser les f.e.n. de grand-Babel *dégénérées* (i.e. de Morris-Mora) est donné par Letac [22, th. 4.2]. Il suppose que le polynôme T est de degré ≤ 3 et, puisque c'est commode de travailler dans un plan pour le cas dégénéré (voir théorème 5.2. A)), il considère le plan euclidien orienté $\mathbb{R}^2$ et donc $\vec{u} : J \rightarrow \mathbb{R}^2$ est tel que

$$\vec{u}'(\lambda) = \left(\frac{\lambda}{T(\lambda)}, \frac{1}{T(\lambda)} \right), \quad \forall \lambda \in J.$$

2) Une interprétation de la réciprocité des f.e.n. de grand-Babel dans ce cadre de caractérisation est donnée par le résultat suivant.

COROLLAIRE 6.3. — Soit F une f.e.n. de grand-Babel. Si $(T; \vec{e}_0, \vec{e}_1)$ définit F comme dans le théorème 6.2, et si F_1 est la f.e.n. réciproque de F , alors $(T; \vec{e}_1, \vec{e}_0)$ définit F_1 .

Démonstration du corollaire 6.3. — Soit (μ, μ_1) dans $\mathcal{B}_F \times \mathcal{B}_{F_1}$ un couple réciproque tel que (6.6) soit satisfaite pour μ avec $(T; \vec{e}_0, \vec{e}_1)$. On a, pour tout λ dans J ,

$$-k_\mu(\langle \vec{e}_1, \vec{u}(\lambda) \rangle) = \langle \vec{e}_0, \vec{u}(\lambda) \rangle.$$

Puisque $-k_\mu(-k_{\mu_1}(\theta)) = \theta$ pour des θ dans $\Theta^+(\mu)$, on en déduit (aussi à l'aide du théorème 5.2 D et par prolongement analytique) que

$$-k_{\mu_1}(\langle \vec{e}_0, \vec{u}(\lambda) \rangle) = \langle \vec{e}_1, \vec{u}(\lambda) \rangle, \quad \forall \lambda \in J.$$

Autrement dit, (6.6) est satisfaite pour μ_1 avec $(T; \vec{e}_1, \vec{e}_0)$. $\square$

3) $\vec{u}(\lambda)$ est définie à partir de $T(\lambda)$ à une constante additive $\vec{C} \in \mathbb{R}^3$ près; cette constante change μ (s'il existe) en μ' , mais on a $F(\mu) = F(\mu')$. En effet, puisque μ et μ' de $\mathcal{M}$ vérifient (6.6), c'est-à-dire pour tout $\lambda \in J$:

$$\begin{aligned} 1 &= \int_{\mathbb{R}} \exp \{ \langle \vec{e}_0 + x\vec{e}_1, \vec{u}(\lambda) \rangle \} \mu(dx) \\ &= \int_{\mathbb{R}} \exp \{ \langle \vec{e}_0 + x\vec{e}_1, \vec{u}(\lambda) + \vec{C} \rangle \} \mu'(dx). \end{aligned}$$

Il est clair que pour tout $\lambda \in J$:

$$\int_{\mathbb{R}} \exp \{ \langle \vec{e}_0 + x\vec{e}_1, \vec{u}(\lambda) \rangle \} \mu'(dx) = \frac{e^{-\langle \vec{e}_0, \vec{C} \rangle}}{\mu'(\mathbb{R})L_{\mu'}(\langle \vec{e}_1, \vec{C} \rangle)}$$

est une constante réelle positive; il existe donc $a, b \in \mathbb{R}$ tels que $\mu'(dx) = e^{ax+b}\mu(dx)$ et donc $F(\mu) = F(\mu')$.

4) Changer

$$\left(\frac{\lambda^2}{T(\lambda)}, \frac{\lambda}{T(\lambda)}, \frac{1}{T(\lambda)} \right) \quad \text{en} \quad \left(\frac{B_2(\lambda)}{T(\lambda)}, \frac{B_1(\lambda)}{T(\lambda)}, \frac{B_0(\lambda)}{T(\lambda)} \right),$$

où B_2, B_1, B_0 forment une autre base de l'espace des polynômes de degré ≤ 2 , revient à faire une transformation linéaire de $\mathbb{R}^3$ et à modifier $\vec{e}_0$ et $\vec{e}_1$; cela ne fournit donc pas une généralisation du théorème 6.2. On décide de fixer le polynôme T et la courbe gauche $\mathcal{C}_T = \vec{u}(J)$, et de faire bouger $\vec{e}_0$ et $\vec{e}_1$. Pour ce fait, la partie inverse du théorème 6.2, telle qu'elle est énoncée, nous permet d'échapper aux trois valeurs nuisibles $2, (3 \pm \sqrt{5})/2$ de $\hat{r}$ grâce à la proposition 5.1 ii).

5) Dans ce théorème 6.2 de caractérisation de la classe de grand-Babel, il est clair que Δ dépend uniquement de $(\vec{e}_0, \vec{e}_1)$ étudié en détail dans la section 5; et donc, il est raisonnable que les polynômes R et Q à cinq paramètres dépendent du polynôme T à cinq paramètres aussi. Ainsi, la partie inverse du théorème 6.2 ne fait que mesurer la qualité de la correspondance entre les polynômes T et (R, Q) . Autrement dit, le sens inverse du théorème 6.2 montre dans quelles conditions le procédé (du sens) direct pouvait atteindre toutes les f.e.n. de grand-Babel.

Démonstration du théorème 6.2

Sens direct. Notons $\theta_i(\lambda) := \langle \vec{e}_i, \vec{u}(\lambda) \rangle$, $i = 0, 1$. Il est clair que les $\theta_i(\lambda)$, $i = 0, 1$ sont analytiques réelles sur J ; et donc (6.6) implique que

$$k_\mu(\theta_1(\lambda)) = -\theta_0(\lambda)$$

est analytique réelle sur $\Theta(\mu) \supset \theta_1(J)$. Par une première dérivation de k_μ , la moyenne $m(\lambda) = k'_\mu(\theta_1(\lambda))$ donnée par

$$m(\lambda) = -\frac{\theta'_0(\lambda)}{\theta'_1(\lambda)}$$

est analytique réelle sur $J' = \{\lambda \in J \mid \theta_1(\lambda) \in \Theta(\mu)\}$; et par une seconde dérivation de k_μ la fonction variance $V_{F(\mu)}(m(\lambda)) = k''_\mu(\theta_1(\lambda))$ restreinte à $\widetilde{M} = m(J')$ est donnée par :

$$V_{F(\mu)}(m(\lambda)) = \frac{\theta'_0(\lambda)\theta''_1(\lambda) - \theta''_0(\lambda)\theta'_1(\lambda)}{(\theta'_1(\lambda))^3}. \quad (6.7)$$

Si nous désignons par $(a_i, b_i, c_i) := \vec{e}_i$, $i = 0, 1$, les coordonnées des $\vec{e}_i$ dans la base canonique de $\mathbb{R}^3$, on a pour tout $\lambda \in J$:

$$\theta'_i(\lambda) = \frac{P_i(\lambda)}{T(\lambda)} \quad \text{avec} \quad P_i(\lambda) = a_i\lambda^2 + b_i\lambda + c_i \quad i = 0, 1,$$

et

$$m = m(\lambda) = -\frac{P_0(\lambda)}{P_1(\lambda)}. \quad (6.8)$$

Donc, (6.7) devient alors

$$V_{F(\mu)}(m(\lambda)) = \frac{P_0(\lambda)P'_1(\lambda) - P'_0(\lambda)P_1(\lambda)}{(P_1(\lambda))^3} T(\lambda),$$

et à l'aide de (6.8) on peut encore écrire

$$V_{F(\mu)}(m) = -(P'_0 + mP'_1) \left(\frac{P_1(\lambda)}{\lambda} \right)^{-2} \frac{T(\lambda)}{\lambda^2}. \quad (6.9)$$

Soit $(\alpha, \beta, \gamma) := \vec{e}_0 \wedge \vec{e}_1$ et soit $A(m) := a_0 + ma_1$, $B(m) := b_0 + b_1m$, $C(m) := c_0 + mc_1$ une paramétrisation affine sur $\mathbb{R}$ de la droite $\mathcal{D} = \{\vec{e}_0 + m\vec{e}_1 \mid m \in \mathbb{R}\}$. Il est alors facile de vérifier que $\Delta = \Delta_{\vec{e}_0, \vec{e}_1}$ s'écrit

$$\Delta(m) = B^2(m) - 4A(m)C(m), \quad \forall m. \quad (6.10)$$

L'orthogonalité de $\vec{e}_0 \wedge \vec{e}_1$ avec $\mathcal{D}$ se traduit par

$$\alpha A(m) + \beta B(m) + \gamma C(m) = 0, \quad \forall m, \quad (6.11)$$

et la formule (6.8) devient

$$\lambda^2 A(m) + \lambda B(m) + C(m) = 0, \quad \forall \lambda \in J'. \quad (6.12)$$

Par conséquent, pour ε fixé dans $\{-1, 1\}$ et deux intervalles ouverts $J_1 \subset J$ et $M_1 \subset \widehat{M}$ tels que

$$\lambda = \lambda(m) = \frac{-B + \varepsilon\sqrt{\Delta}}{2A}$$

définit une bijection de M_1 sur J_1 , on a

$$\lambda^{-1} = \frac{-B - \varepsilon\sqrt{\Delta}}{2C}$$

et donc

$$\begin{aligned} \frac{P_1(\lambda)}{\lambda} &= a_1\lambda + b_1 + c_1\lambda^{-1} \\ &= \frac{\gamma C - \alpha A + \varepsilon\beta\sqrt{\Delta}}{2AC}. \end{aligned}$$

En introduisant $\tilde{d} = \beta^2 - \alpha\gamma$, qui est $\neq 0$ par hypothèse, on peut vérifier directement à partir de (6.10) et (6.11) qu'on a

$$(\gamma C - \alpha A)^2 - \beta^2 \Delta = 4\tilde{d}AC.$$

Ainsi, on trouve finalement

$$\left(\frac{P_1(\lambda)}{\lambda}\right)^{-1} = \frac{1}{2\tilde{d}}(\gamma C - \alpha A - \varepsilon\beta\sqrt{\Delta}). \quad (6.13)$$

Par ailleurs, on peut vérifier successivement que

$$\begin{aligned} (P'_0 + mP'_1)^2 - \Delta &= (2\lambda A + B)^2 - \Delta \\ &= 4A(\lambda^2 A + \lambda B + C) \\ &= 0. \end{aligned}$$

D'où,

$$P'_0 + mP'_1 = \varepsilon\sqrt{\Delta}.$$

Finalement, à partir de la dernière égalité ci-dessus et du carré de (6.13), l'expression (6.9) de $V_{F(\mu)}(m)$ devient avec $T(\lambda)/\lambda^2 = t_2\lambda^2 + t_1\lambda + t_0 + t_{-1}\lambda^{-1} + t_{-2}\lambda^{-2}$:

$$\begin{aligned} V_{F(\mu)}(m) &= \frac{1}{4\hat{d}^2} \left(-2\beta(\gamma C - \alpha A)\Delta + ((\gamma C - \alpha A)^2 + \beta^2\Delta)\varepsilon\sqrt{\Delta} \right) \times \\ &\quad \times \left(t_2\lambda^2 + t_1\lambda + t_0 + t_{-1}\lambda^{-1} + t_{-2}\lambda^{-2} \right). \end{aligned} \quad (6.14)$$

Il faut maintenant montrer que les coefficients de $t_2, t_1, t_0, t_{-1}, t_{-2}$ dans (6.14) sont de la forme $R\Delta + Q\sqrt{\Delta}$ où R et Q sont des polynômes en m de degrés ≤ 1 et ≤ 2 respectivement.

Notons que le coefficient de t_0 est de la forme requise (rappelons qu'on a explicitement la forme des fonctions variances reliées à la famille gaussienne inverse généralisée avec $T(\lambda) = \lambda^2$). Il suffit alors de faire les calculs pour les coefficients de t_2 et t_1 , puis on déduit en interchangeant A et C , ε et $-\varepsilon$ pour obtenir les coefficients de t_{-2} et t_{-1} .

Puisque $\alpha\beta\gamma \neq 0$ (par hypothèse), nous simplifions la notation en introduisant

$$a := \frac{\alpha}{\beta} A \quad \text{et} \quad c := \frac{\gamma}{\beta} C.$$

Ainsi $-B = a + c$ puisque $\hat{r} = \beta^2/(\alpha\gamma)$, on a

$$\Delta = (a + c)^2 - 4\hat{r}ac \quad \text{et} \quad \lambda = \frac{\alpha}{2\beta} \frac{a + c + \varepsilon\sqrt{\Delta}}{a}.$$

Tout calcul fait, on obtient à une constante multiplicative près les coefficients de $t_2, t_1, t_0, t_{-1}, t_{-2}$ comme polynômes de a et c suivants ;

$$\begin{aligned} \frac{\alpha^2}{4\hat{d}^2} t_2 &: 2(a + (1 - 2\hat{r})c)\Delta + (a^2 + 2(1 - 2\hat{r})ac + (1 - 2\hat{r} + 2\hat{r}^2)c^2) \varepsilon\sqrt{\Delta} \\ \frac{\alpha\beta}{2\hat{d}^2} t_1 &: (a - \hat{r}c)\Delta + (a^2 + (1 - 3\hat{r})ac + \hat{r}c^2) \varepsilon\sqrt{\Delta} \\ \frac{\beta^2}{2\hat{d}^2} t_0 &: (a - c)\Delta + (a^2 - 2\hat{r}ac + c^2)\varepsilon\sqrt{\Delta} \\ \frac{\alpha\beta}{2\hat{d}^2} t_{-1} &: (-\hat{r}a + c)\Delta - (\hat{r}a^2 + (1 - 3\hat{r})ac + c^2) \varepsilon\sqrt{\Delta} \\ \frac{\alpha^2}{4\hat{d}^2} t_{-2} &: 2((1 - 2\hat{r})a + c)\Delta - ((1 - 2\hat{r} + \hat{r}^2)a^2 + 2(1 - 2\hat{r})ac + c^2) \varepsilon\sqrt{\Delta}. \end{aligned}$$

Ce qui achève la démonstration du sens direct.

Inversement. Supposons que $V_F = R\Delta + Q\sqrt{\Delta}$ soit une fonction variance d'une f.e.n. F de grand-Babel non dégénérée. Soit $(\vec{e}_0, \vec{e}_1)$ dans $\mathcal{B}_\Delta^*$ tel que $\hat{r}$ soit différent de 2 et $(3 \pm \sqrt{5})/2$. Il s'agit de mesurer la correspondance à caractère biunivoque de ce procédé (sens direct du théorème 6.2) entre T et (R, Q) . Cela nous conduit tout d'abord à montrer que :

- i) $a = (\alpha/\beta)(a_0 + ma_1)$ et $c = (\gamma/\beta)(c_0 + mc_1)$ constitue une base de l'espace $\mathcal{P}_1$ des polynômes en m de degré ≤ 1 ,
- ii) (a^2, ac, c^2) est une base de l'espace $\mathcal{P}_2$ de polynômes en m de degré ≤ 2 .

En effet, i) est acquis car

$$\det(a, c) = \begin{vmatrix} \frac{\alpha}{\beta} a_0 & \frac{\gamma}{\beta} c_0 \\ \frac{\alpha}{\beta} a_1 & \frac{\gamma}{\beta} c_1 \end{vmatrix} = -\frac{\alpha\gamma}{\beta} \neq 0.$$

Pour ii), on a besoin du lemme suivant.

LEMME 6.4. — Soit A et B deux vecteurs de l'espace $\mathcal{P}_1$ des polynômes de degré ≤ 1 et soit $n \in \mathbb{N}^*$. Si A et B sont indépendants dans $\mathcal{P}_1$ alors $A^n, A^{n-1}B, \dots, A^{n-k}B^k, \dots, B^n$ le sont dans l'espace $\mathcal{P}_n$ des polynômes de degré $\leq n$.

Démonstration. — Si on considère la base canonique $(1, m, m^2, \dots, m^n)$ de l'espace $\mathcal{P}_n$, le résultat énoncé est immédiat par une vérification (e.g. par récurrence sur $n \in \mathbb{N}^*$) de la jolie formule :

$$\det(A^n, A^{n-1}B, \dots, AB^{n-1}, B^n) = (\det(A, B))^{\frac{n(n+1)}{2}}. \quad \square$$

Revenons à ii). Puisque a et c sont indépendants dans $\mathcal{P}_1$ d'après i), le lemme 6.4 s'applique avec $A = a$, $B = c$ et $n = 2$. Ainsi découle ii), c'est-à-dire (a^2, ac, c^2) forme une base de l'espace $\mathcal{P}_2$ des polynômes en m de degrés ≤ 2 . Par conséquent, la matrice :

$$G = \begin{pmatrix} 2 & 2(1 - 2\hat{r}) & \varepsilon & \varepsilon 2(1 - 2\hat{r}) & \varepsilon(2\hat{r}^2 - 2\hat{r} + 1) \\ 1 & -\hat{r} & \varepsilon & \varepsilon(1 - 3\hat{r}) & \varepsilon\hat{r} \\ 1 & -1 & \varepsilon & -\varepsilon 2\hat{r} & \varepsilon \\ -\hat{r} & 1 & -\varepsilon\hat{r} & \varepsilon(3\hat{r} - 1) & -\varepsilon \\ 2(1 - 2\hat{r}) & 2 & \varepsilon(2\hat{r} - \hat{r}^2 - 1) & \varepsilon 2(2\hat{r} - 1) & -\varepsilon \end{pmatrix}$$

issue de la formule (6.15) a pour déterminant

$$\det G = -8\varepsilon(\hat{r} - 1)^4(\hat{r} - 2) \left(\hat{r} - \frac{3 - \sqrt{5}}{2} \right) \left(\hat{r} - \frac{3 + \sqrt{5}}{2} \right).$$

Puisque par hypothèse on a aussi $\hat{r} \neq 1$ à cause de la non-dégénérescence de Δ (théorème 5.2 A), la matrice G est inversible. D'où, l'inversion du procédé $T \rightarrow (R, Q)$. Ceci achève la démonstration du théorème 6.2. $\square$

Références

- [1] BAR-LEV (S. K.), BSHOUTY (D.) et ENIS (P.) . — *Variance functions with meromorphic means*, Ann. of Proba. **19** (1991), pp. 1349-1366.
- [2] BAR-LEV (S. K.) et ENIS (P.) . — *Reproducibility and natural exponential families with power variance functions*, Anna. Statist. **14** (1987), pp. 1507-1522.
- [3] BAR-LEV (S. K.), ENIS (P.) et LETAC (G.) . — *Sampling models with admit a given general exponential family as conjugate family of priors*, Ann. Statist. (à paraître) (1994).
- [4] BARNDORFF-NIELSEN (O. E.) . — *Information and exponential families in statistical theory*, Wiley, New-York (1978).
- [5] BARNDORFF-NIELSEN (O. E.) . — *Hyperbolic distributions and distributions on hyperbolae*, Scand. J. Statist. **5** (1978), pp. 151-157.
- [6] DIACONIS (P.) et YLVISAKER (D.) . — *Conjugate priors for exponential families*, Ann. Statist. **7** (1979), pp. 269-281.
- [7] DIEUDONNÉ (J.) . — *Infinitesimal calculus*, Houghton Mifflin, Boston (1971).
- [8] DONEY (R. A.) . — *Hitting probabilities for spectrally positive Lévy processes*, J. London Math. Soc. **44** (1991), pp. 566-576.
- [9] FELLER (W.) . — *An introduction to probability theory and its applications*, Vol. 2, Wiley, New-York (1966).
- [10] IMHOF (J. P.) . — *On the time spent above a level by a Brownian motion*, Adv. Appl. Prob. **18** (1986), pp. 1017-1018.
- [11] JEFFREYS (H.) . — *Theory of probability*, Oxford University Press (1948).
- [12] JORGENSEN (B.) . — *Some properties of exponential dispersion models*, Scand. J. Statist. **13** (1986), pp. 187-198.
- [13] JORGENSEN (B.) . — *Exponential dispersion models*, J. Roy. Statist. Soc. Ser. **B49** (1987), pp. 127-162.
- [14] JORGENSEN (B.) . — *The theory of exponential dispersion models and analysis of deviance*. I.M.P.A. Rio de Janeiro (1992).
- [15] JORGENSEN (B.), LETAC (G.) et SESHADRI (V.) . — *Sur une propriété des familles exponentielles naturelles de variance quadratique*. Canadian Jour. Statist. **17** (1989), pp. 1-8.

- [16] KOKONENDJI (C. C.) . — *Caractérisation de fonctions variance de Seshadri des familles exponentielles sur $\mathbb{R}$* , C.R. Acad. Sci. Paris. **314**, série I (1992), pp. 1063-1068.
- [17] KOKONENDJI (C. C.) . — *Exponential families with variance functions in $\sqrt{\Delta P}(\sqrt{\Delta})$: Seshadri's class*, Test **3**, n° 2 (1994), pp. 99-148.
- [18] KOKONENDJI (C. C.) et SESHADRI (V.) . — *La méthode de Lindsay appliquée à la construction de familles exponentielles de fonctions variances de degré 4 en $\sqrt{m}$* , C.R. Acad. Sci. Paris **314**, série I (1992), pp. 305-308.
- [19] KOKONENDJI (C. C.) et SESHADRI (V.) . — *The Lindsay transform of natural exponential families*, Canadian Jour. Statist. **22**, n° 2 (1994), pp. 259-272.
- [20] LETAC (G.) . — *La réciprocité des familles exponentielles naturelles sur $\mathbb{R}$* , C.R. Acad. Sci. Paris, **303**, série I (1986) pp. 61-64.
- [21] LETAC (G.) . — *The classification of natural exponential families by their variances functions*, Proceeding of the 48th session of the International Statistical Institute, Vol. LIV, Book **3** (1991).
- [22] LETAC (G.) . — *Lectures on natural exponential families and their variance functions*, I.M.P.A. Rio de Janeiro (1992).
- [23] LETAC (G.) et MORA (M.) . — *Sur les fonctions-variances des familles exponentielles sur $\mathbb{R}$* , C.R. Acad. Sci. Paris **302**, série I (1986), pp. 551-554.
- [24] LETAC (G.) et MORA (M.) . — *Natural real exponential families with cubic variance functions*, Ann. Statist. **18** (1990), pp. 1-37.
- [25] MORA (M.) . — *Classification des fonctions-variances cubiques des familles exponentielles sur $\mathbb{R}$* , C.R. Acad. Sci. Paris **302**, série I (1986), pp. 587-590.
- [26] MORLAT (G.) . — *Les lois de probabilité de Halphen*, Revue de Statist. Appli. **4**, n° 3 (1956), pp. 21-46.
- [27] MORRIS (C.-N.) . — *Natural exponential families with quadratic variance function*, Ann. Statist. **10** (1982), pp. 65-80.
- [28] PRABHU (N. U.) . — *Ladder variables for a continuous time stochastic process*, Z. Wahrscheinlichkeitstheorie verw. Geb. **16** (1970), pp. 157-164.