

SUR CERTAINES ÉQUATIONS FONCTIONNELLES LIÉES A LA LOI NORMALE

Autor(en): **Fuchs, Aimé / Letta, Giorgio**

Objektyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **25 (1979)**

Heft 1-2: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **20.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-50371>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

SUR CERTAINES ÉQUATIONS FONCTIONNELLES LIÉES A LA LOI NORMALE

par Aimé FUCHS et Giorgio LETTA

Dans le présent article nous nous proposons de donner des démonstrations élémentaires de certaines caractérisations de la loi normale. Il s'agit de démonstrations plus simples que celles qui se trouvent couramment dans la littérature; en outre elles ne font appel à aucune hypothèse d'existence de moments. L'outil qui est à la base de ces démonstrations est constitué par les deux équations fonctionnelles

$$\begin{aligned}\varphi(t) &= (\varphi(t/\sqrt{c}))^c \quad (c \text{ constante réelle } > 1), \\ (\varphi(t))^2 &= \varphi(t+u)\varphi(t-u)\theta(u).\end{aligned}$$

La dernière de ces équations (qui, comme nous le verrons, se ramène aussitôt à la première) permet notamment de démontrer très simplement le théorème de Bernstein-Darmois. En revanche la démonstration du théorème de Skitovitch-Darmois semble échapper aux méthodes élémentaires de cet article.

1. L'ÉQUATION FONCTIONNELLE $\varphi(t) = (\varphi(t/\sqrt{c}))^c$.

(1.1) LEMME. Soient μ une loi de probabilité sur \mathbf{R} (n'admettant pas nécessairement un moment du second ordre fini), φ sa fonction caractéristique. Alors on a

$$\lim_{t \rightarrow 0} \frac{1 - \Re\varphi(t)}{t^2} = \frac{1}{2} \int_{\mathbf{R}} x^2 d\mu(x).$$

Démonstration. Posons, pour tout nombre réel t non nul,

$$I_t = \frac{1 - \Re\varphi(t)}{t^2} = \int_{\mathbf{R}} \frac{1 - \cos tx}{t^2} d\mu(x).$$

Le quotient $\frac{1 - \cos tx}{t^2}$ est compris entre 0 et $\frac{1}{2}x^2$; en outre il converge vers $\frac{1}{2}x^2$ lorsque t tend vers 0. On a donc

$$\frac{1}{2} \int_{\mathbf{R}} x^2 d\mu(x) \leq \liminf_{t \rightarrow 0} I_t \leq \limsup_{t \rightarrow 0} I_t \leq \frac{1}{2} \int_{\mathbf{R}} x^2 d\mu(x),$$

où la première inégalité résulte du lemme de Fatou, les autres étant évidentes.

Remarque. Il résulte de ce lemme que s'il existe une suite (t_n) de nombres réels non nuls, tendant vers 0 et telle que l'on ait $\varphi(t_n) = 1$ pour tout n , alors la loi μ a un moment du second ordre nul, c'est-à-dire qu'elle est dégénérée.

(1.2) THÉORÈME. Soient μ une loi de probabilité sur \mathbf{R} , φ sa fonction caractéristique, c un nombre réel, avec $c > 1$. Alors les deux propriétés suivantes sont équivalentes :

- (a) μ est normale et centrée ;
- (b) φ est réelle et vérifie l'équation fonctionnelle

$$\varphi(t) = (\varphi(t/\sqrt{c}))^c$$

pour tout nombre réel t assez petit.

Démonstration. Il suffit de démontrer l'implication (b) \Rightarrow (a). Supposons donc la propriété (b) vérifiée. On a alors, pour tout nombre réel t assez petit et pour tout entier $n \geq 1$:

$$\varphi(t) > 0, \quad \varphi\left(tc^{-\frac{n}{2}}\right) = (\varphi(t))^{c^n}.$$

Par conséquent, si l'on désigne par σ^2 le moment du second ordre de μ (non supposé *a priori* fini!), le lemme précédent entraîne (pour tout t assez petit et non nul)

$$\begin{aligned} \frac{1}{2} \sigma^2 &= \lim_{n \rightarrow \infty} \frac{1 - \varphi\left(tc^{-\frac{n}{2}}\right)}{\left(tc^{-\frac{n}{2}}\right)^2} = \frac{1}{t^2} \lim_{n \rightarrow \infty} c^n [1 - (\varphi(t))^{c^{-n}}] \\ &= -\frac{1}{t^2} \text{Log } \varphi(t) < +\infty, \end{aligned}$$

d'où $\varphi(t) = \exp\left(-\frac{1}{2} \sigma^2 t^2\right)$.

On voit ainsi que la fonction caractéristique φ coïncide, dans un voisinage de l'origine, avec la fonction caractéristique d'une loi normale centrée;

d'après un résultat connu, cette coïncidence a donc lieu partout, et le théorème est démontré.

Dans le cas particulier où le nombre c est un entier, on peut supprimer, dans l'énoncé de la condition (b) du théorème précédent, l'hypothèse que φ est réelle. On a en effet le corollaire suivant :

(1.3) COROLLAIRE. Soient μ une loi de probabilité sur \mathbf{R} , φ sa fonction caractéristique, n un entier, avec $n > 1$. Alors les deux propriétés suivantes sont équivalentes :

- (a) μ est normale et centrée ;
- (b) φ vérifie l'équation fonctionnelle

$$\varphi(t) = (\varphi(t/\sqrt{n}))^n$$

pour tout nombre réel t assez petit.

Démonstration. Il suffit de démontrer l'implication (b) \Rightarrow (a). Supposons donc que la fonction caractéristique φ (non supposé *a priori* réelle) possède la propriété (b); il en est alors de même de la fonction caractéristique réelle $\varphi \bar{\varphi}$; d'après le théorème précédent, celle-ci est donc la fonction caractéristique d'une loi normale. Mais alors, d'après le théorème classique de Cramér-Lévy, la fonction φ possède elle-même cette propriété; en d'autres termes, la loi μ est normale. Une simple vérification suffit ensuite pour montrer que μ est aussi centrée.

Le corollaire (1.3) peut être reformulé, dans le langage des variables aléatoires, sous la forme suivante :

(1.4) COROLLAIRE. Soient n un entier, avec $n > 1$, et (X_1, \dots, X_n) un système de n variables aléatoires réelles indépendantes, de même loi μ . Alors les deux propriétés suivantes sont équivalentes :

- (a) μ est normale et centrée ;
- (b) la loi de $(X_1 + \dots + X_n) / \sqrt{n}$ coïncide avec μ . ¹⁾

Le corollaire (1.4) est susceptible d'une application intéressante: il permet en effet de démontrer rapidement le théorème de Maxwell sur la répartition des vitesses dans un gaz parfait.

¹⁾ La propriété (b) est évidemment vérifiée si μ est une loi stable *stricte* d'exposant caractéristique $\alpha = 2$ (au sens de [3], page 166). Le corollaire énoncé contient donc le résultat bien connu, selon lequel une telle loi est normale et centrée.

(1.5) THÉORÈME (Maxwell). Soient n un entier, avec $n > 1$, et X un vecteur aléatoire, à valeurs dans \mathbf{R}^n , dont les composantes X_1, \dots, X_n sont indépendantes. Alors les deux propriétés suivantes sont équivalentes :

- (a) les variables aléatoires X_1, \dots, X_n ont une même loi, qui est normale et centrée ;
- (b) la loi du vecteur aléatoire X est invariante par rapport à toute rotation de \mathbf{R}^n autour de l'origine.

Démonstration. Il suffit de démontrer l'implication (b) \Rightarrow (a). Supposons donc la propriété (b) vérifiée. Pour tout indice $i = 1, \dots, n$, il existe une rotation autour de l'origine qui transforme X en un vecteur aléatoire dont la première composante est égale à X_i . (Pour avoir une telle rotation, il suffit de prendre une matrice orthogonale, de déterminant égal à 1, admettant pour première ligne le i -ème vecteur de la base canonique de \mathbf{R}^n .) On voit ainsi que chacune des composantes de X a même loi que X_1 . Il existe d'autre part une rotation autour de l'origine qui transforme X en un vecteur aléatoire dont la première composante est égale à $(X_1 + \dots + X_n) / \sqrt{n}$. (Pour avoir une telle rotation, il suffit de prendre une matrice orthogonale, de déterminant égal à 1, admettant pour première ligne le vecteur $(1/\sqrt{n}, \dots, 1/\sqrt{n})$.) La conclusion résulte du corollaire (1.4).

En calquant le raisonnement utilisé lors de la démonstration du Théorème (1.2), on peut obtenir une caractérisation analogue pour la loi de Cauchy :

(1.6) THÉORÈME. Soient μ une loi de probabilité sur \mathbf{R} , φ sa fonction caractéristique, c un nombre réel, avec $c > 1$. Alors les deux propriétés suivantes sont équivalentes :

- (a) μ est, soit la loi dégénérée à l'origine, soit une loi de Cauchy symétrique ;
- (b) φ est réelle, possède une dérivée à droite à l'origine (non supposée a priori finie) et vérifie l'équation fonctionnelle

$$\varphi(t) = (\varphi(t/c))^c$$

pour tout nombre réel t .

Démonstration. Il suffit de démontrer l'implication (b) \Rightarrow (a). Supposons donc la propriété (b) vérifiée. Commençons par montrer que φ ne peut pas s'annuler : si en effet φ s'annulait en t_0 , elle s'annulerait en t_0/c , donc

en t_0/c^n pour tout $n \geq 1$. La fonction φ étant continue, il en résulterait $\varphi(0) = \lim_{n \rightarrow \infty} \varphi(t_0/c^n) = 0$, ce qui est impossible. Remarquons ensuite qu'on a, pour tout nombre réel t et pour tout entier $n \geq 1$,

$$\varphi(t/c^n) = (\varphi(t))^{1/c^n}.$$

Par conséquent, si l'on désigne par $-a$ ($a \geq 0$) la dérivée à droite de φ à l'origine (non supposée *a priori* finie), on a, pour tout $t > 0$,

$$\begin{aligned} a &= \lim_{n \rightarrow \infty} \frac{1 - \varphi(tc^{-n})}{tc^{-n}} = \frac{1}{t} \lim_{n \rightarrow \infty} c^n [1 - (\varphi(t))]^{c^{-n}} \\ &= -\frac{1}{t} \text{Log } \varphi(t) < +\infty, \end{aligned}$$

d'où $\varphi(t) = \exp(-at)$.

Puisque d'autre part φ est réelle, donc paire, on a, pour tout nombre réel t ,

$$\varphi(t) = \exp(-a|t|),$$

ce qui démontre le théorème.

On remarquera que le cas d'une loi dégénérée à l'origine correspond au cas où a est nul, c'est-à-dire au cas où φ est dérivable à l'origine.

Dans le cas particulier où c est un entier, le Théorème (1.6) peut être reformulé, dans le langage des variables aléatoires, sous la forme suivante:

(1.7) COROLLAIRE. Soient n un entier, avec $n > 1$, et (X_1, \dots, X_n) un système de n variables aléatoires réelles indépendantes, de même loi μ . Si la fonction caractéristique φ de μ est réelle et possède une dérivée à droite à l'origine (non supposée *a priori* finie), alors les deux propriétés suivantes sont équivalentes :

- (a) μ est, soit la loi dégénérée à l'origine, soit une loi de Cauchy symétrique ;
- (b) la loi de $(X_1 + \dots + X_n)/n$ coïncide avec μ .

Remarque. Si, étant donnée la loi μ symétrique et non dégénérée, au lieu de supposer la propriété (b) vérifiée pour un n fixé (ce qui est notre cas), on la supposait vérifiée pour *tout* entier n , ceci équivaudrait à supposer que μ est une loi stable stricte d'exposant caractéristique $\alpha = 1$ (au sens de [3], page 166). D'après le théorème de représentation des lois stables de P. Lévy, la loi μ serait alors une loi de Cauchy, et ceci indépendamment de toute hypothèse de dérivabilité de φ à l'origine.

2. UNE GÉNÉRALISATION DE L'ÉQUATION FONCTIONNELLE DU § 1.

Si l'on regarde attentivement la démonstration du Théorème (1.5) de Maxwell, on s'aperçoit que, en ce qui concerne l'implication (b) \Rightarrow (a), on n'a exploité l'hypothèse d'invariance par rotation que pour des rotations bien particulières, à savoir :

- (a) des rotations dont la matrice admet comme lignes des vecteurs de la base canonique, éventuellement multipliés par -1 ;
- (b) des rotations dont la matrice admet comme première ligne le vecteur $(1/\sqrt{n}, \dots, 1/\sqrt{n})$.

Les rotations du type (a) nous servaient en effet à démontrer préalablement l'égalité des lois des variables aléatoires X_1, \dots, X_n . Une seule rotation du type (b) suffisait ensuite à démontrer la nature gaussienne de la loi commune.

Dans le cas du plan, en se plaçant d'emblée dans l'hypothèse de l'égalité des lois, on peut même énoncer le résultat suivant, faisant intervenir une seule rotation.

(2.1) THÉORÈME. Soit (X, Y) un vecteur aléatoire, à valeurs dans \mathbf{R}^2 , dont les composantes sont indépendantes et de même loi μ . Désignons par (X', Y') le vecteur aléatoire

$$((X - Y)/\sqrt{2}, (X + Y)/\sqrt{2})$$

(transformé de (X, Y) par une rotation de $\pi/4$ autour de l'origine). Si l'une des composantes de (X', Y') suit la loi μ , alors μ est normale et centrée. ¹⁾

Démonstration. Supposons d'abord que la composante $Y' = (X + Y)/\sqrt{2}$ suit la loi μ ; le résultat est alors une conséquence du Corollaire (1.4) pour $n = 2$.

Supposons ensuite que la composante $X' = (X - Y)/\sqrt{2}$ suit la loi μ . On voit alors que μ est une loi symétrique; donc $-Y$ suit la même loi que Y , et l'on est ramené au cas précédent.

De façon plus générale, on peut se poser le problème suivant:

¹⁾ Il en résulte accessoirement que l'autre composante suit également la loi μ et que les deux composantes sont indépendantes.

Soit X un vecteur aléatoire, à valeurs dans \mathbf{R}^n (avec $n > 1$), dont les composantes sont indépendantes et de même loi μ . Considérons le transformé Y de X par une rotation autour de l'origine, dont on désignera par $A = (a_{ij})$ la matrice. Supposons que cette rotation ne soit pas du type (a) ci-dessus, c'est-à-dire qu'il existe un indice i tel que chacun des éléments a_{i1}, \dots, a_{in} de la i -ème ligne de A soit différent de 1 et de -1 . Si, pour un tel indice i , la i -ème composante de Y , c'est-à-dire la variable aléatoire

$$Y_i = a_{i1}X_1 + \dots + a_{in}X_n,$$

suit la loi μ , peut-on affirmer que μ est normale et centrée?

En faisant intervenir la fonction caractéristique φ de la loi μ , le problème peut être reformulé ainsi:

Soit (a_1, \dots, a_n) un système de n nombres réels satisfaisant aux conditions

$$a_1^2 + \dots + a_n^2 = 1, \quad |a_i| < 1 \text{ pour tout } i.$$

Si φ vérifie l'équation fonctionnelle

$$(2.1) \quad \varphi(t) = \varphi(a_1 t) \dots \varphi(a_n t),$$

peut-on affirmer que μ est normale et centrée?

La réponse est évidemment affirmative dans le cas particulier où les a_i sont égaux, car l'équation fonctionnelle (2.1) se réduit alors à l'équation fonctionnelle $\varphi(t) = (\varphi(t/\sqrt{n}))^n$ déjà étudiée au paragraphe précédent. Dans le cas général, la réponse est encore affirmative, mais elle exige une démonstration beaucoup plus délicate, fondée sur la théorie des lois limites de Kolmogorov: cf. [5], Théor. 7.2.1, pag. 161, où le problème est généralisé au cas d'une suite infinie de variables aléatoires (indépendantes et de même loi). ¹⁾

3. L'ÉQUATION FONCTIONNELLE $(\varphi(t))^2 = \varphi(t+u) \varphi(t-u) \theta(u)$.

Le théorème qui suit est une conséquence immédiate de (1.2). Il nous servira à démontrer de façon très simple le théorème de Bernstein-Darmois.

(3.1) THÉORÈME. Soient μ une loi de probabilité sur \mathbf{R} , φ sa fonction caractéristique. Alors les deux propriétés suivantes sont équivalentes :

¹⁾ A signaler que l'énoncé du théorème de [5] cité ci-dessus est incorrect: en effet, il est faux si l'on prend $a_1 = \alpha$ et $a_i = 0$ pour tout $i \geq 2$. Pour le rendre correct, il suffit d'ajouter l'hypothèse: $|a_i| < |\alpha|$ pour tout i .

(a) μ est normale ;

(b) il existe un voisinage V de l'origine dans \mathbf{R} et une application θ de V dans \mathbf{C} telle que l'on ait

$$(3.2) \quad (\varphi(t))^2 = \varphi(t+u)\varphi(t-u)\theta(u)$$

pour tout couple t, u d'éléments de V .

Démonstration. Il suffit de démontrer l'implication (b) \Rightarrow (a). Supposons donc la propriété (b) vérifiée. Quitte à remplacer φ par $\varphi\bar{\varphi}$, on peut supposer φ réelle. L'équation (3.2) fournit alors, pour $t = 0$, $1 = (\varphi(u))^2\theta(u)$, de sorte qu'elle peut s'écrire sous la forme équivalente

$$(\varphi(t)\varphi(u))^2 = \varphi(t+u)\varphi(t-u).$$

En particulier, pour $t = u = s/2$, on trouve $(\varphi(s/2))^4 = \varphi(s)$. La conclusion résulte alors du Théorème (1.2) pour $c = 4$.

(3.3) *Remarque.* Le cas d'une loi μ dégénérée correspond à celui où l'équation (3.2) est vérifiée avec une fonction θ identiquement égale à 1.

4. LE THÉORÈME DE BERNSTEIN-DARMOIS

Soit (X_1, X_2) un couple de variables aléatoires réelles, de même loi. Si le couple $(X_1 + X_2, X_1 - X_2)$ est formé de variables aléatoires indépendantes, alors la loi commune de X_1 et de X_2 est normale: c'est le théorème de Bernstein-Darmonis sous sa forme primitive. Il fut d'abord démontré par S. Bernstein [1] avec l'hypothèse que la loi commune de X_1 et de X_2 possède des moments finis jusqu'à l'ordre 4. Plus tard G. Darmonis [2] réussit à généraliser ce résultat, tout en s'affranchissant de l'hypothèse concernant l'existence des moments. Il employa à cet effet une technique de différences finies, qui lui permit également de démontrer une généralisation ultérieure, bien plus profonde, connue sous le nom de théorème de Skitovitch-Darmonis (cf. [6]).

Nous présentons ci-dessous le théorème de Bernstein-Darmonis, que nous démontrons à l'aide de l'équation fonctionnelle du paragraphe 3.

(4.1) THÉORÈME. Soit X un vecteur aléatoire à valeurs dans \mathbf{R}^2 , dont les composantes X_1, X_2 sont des variables aléatoires indépendantes. Soit $A = (a_{ij})$ une matrice réelle $(2, 2)$, et supposons que les composantes du

vecteur aléatoire $Y = AX$, c'est-à-dire les deux variables aléatoires réelles Y_1, Y_2 définies par

$$\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 \\ Y_2 = a_{21}X_1 + a_{22}X_2 \end{cases},$$

soient elles aussi indépendantes. Alors, pour chaque indice i tel que la i -ème colonne de A soit formée d'éléments non nuls, la loi de X_i est normale (éventuellement dégénérée).

Pour rendre plus claire la démonstration, nous commencerons par démontrer un lemme préliminaire:

(4.2) LEMME. Soit X un vecteur aléatoire à valeurs dans \mathbf{R}^2 , de composantes X_1, X_2 , et soit Φ sa fonction caractéristique. Les deux propriétés suivantes sont alors équivalentes :

- (a) (X_1, X_2) est un couple de variables aléatoires indépendantes ;
- (b) pour tout système de scalaires s_1, s_2, u, v , on a

$$\Phi(s_1, s_2)\Phi(s_1 + u, s_2 + v) = \Phi(s_1 + u, s_2)\Phi(s_1, s_2 + v).$$

Démonstration du lemme.

Il suffit de démontrer (b) \Rightarrow (a). Si l'on désigne par φ_i la fonction caractéristique de X_i , l'hypothèse (b) fournit (pour $s_1 = s_2 = 0$)

$$\Phi(u, v) = \Phi(u, 0)\Phi(0, v) = \varphi_1(u)\varphi_2(v),$$

c'est-à-dire l'indépendance du couple (X_1, X_2) .

Démonstration du théorème. Supposons, pour fixer les idées, que la première colonne de A soit formée d'éléments non nuls, et montrons que la loi de X_1 est normale. Quitte à multiplier chacune des lignes de A par un scalaire convenable, on pourra supposer

$$(4.3) \quad a_{11} = a_{21} = 1.$$

1) Supposons d'abord que la matrice A soit singulière. En vertu de notre hypothèse on a alors $Y_1 = Y_2$. Par conséquent Y_1 est indépendante d'elle-même, c'est-à-dire p.s. égale à une constante:

$$Y_1 = X_1 + a_{12}X_2 = c \quad \text{p.s.};$$

il en résulte

$$X_1 = c - a_{12}X_2 \quad \text{p.s.},$$

de sorte que X_1 est également indépendante d'elle-même, c'est-à-dire p.s. égale à une constante.

2) Supposons maintenant que la matrice A ne soit pas singulière, et désignons par φ_i la fonction caractéristique de X_i et par Φ celle de X :

$$(4.4) \quad \Phi(t_1, t_2) = \varphi_1(t_1) \varphi_2(t_2).$$

Désignons en outre par ψ la fonction caractéristique du vecteur aléatoire $Y = AX$, c'est-à-dire la fonction définie par

$$(4.5) \quad \psi(s_1, s_2) = \Phi(s_1 + s_2, s_1 a_{12} + s_2 a_{22}).$$

En appliquant le lemme précédent au couple de variables aléatoires indépendantes (Y_1, Y_2) , on trouve, pour tout système de scalaires s_1, s_2, u ,

$$\psi(s_1, s_2) \psi(s_1 + u, s_2 - u) = \psi(s_1 + u, s_2) \psi(s_1, s_2 - u).$$

Grâce à (4.5), cette relation s'écrit, en fonction de Φ ,

$$\begin{aligned} & \Phi(s_1 + s_2, s_1 a_{12} + s_2 a_{22}) \Phi(s_1 + s_2, s_1 a_{12} + s_2 a_{22} + u a_{12} - u a_{22}) \\ &= \Phi(s_1 + s_2 + u, s_1 a_{12} + s_2 a_{22} + u a_{12}) \Phi(s_1 + s_2 - u, s_1 a_{12} + s_2 a_{22} - u a_{22}). \end{aligned}$$

Etant donné le scalaire t , choisissons maintenant s_1, s_2 de façon à satisfaire aux conditions

$$\begin{cases} s_1 + s_2 = t \\ s_1 a_{12} + s_2 a_{22} = 0 \end{cases}$$

(ce qui est possible, car la matrice A n'est pas singulière). La relation précédente devient alors

$$\Phi(t, 0) \Phi(t, u(a_{12} - a_{22})) = \Phi(t + u, u a_{12}) \Phi(t - u, -u a_{22}),$$

c'est-à-dire, grâce à (4.4),

$$(4.6) \quad \begin{aligned} & \varphi_1(t) \varphi_2(0) \varphi_1(t) \varphi_2(u(a_{12} - a_{22})) \\ &= \varphi_1(t + u) \varphi_2(u a_{12}) \varphi_1(t - u) \varphi_2(-u a_{22}). \end{aligned}$$

Or, si $|u|$ est assez petit, on a $\varphi_2(u(a_{12} - a_{22})) \neq 0$, de sorte que la relation précédente peut s'écrire sous la forme

$$(4.7) \quad (\varphi_1(t))^2 = \varphi_1(t + u) \varphi_1(t - u) \theta(u).$$

Il en résulte, grâce à (3.1), que φ_1 est la fonction caractéristique d'une loi normale.

A titre d'exemple, nous analyserons l'énoncé (4.1) dans deux *cas particuliers*.

a) Supposons d'abord

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

c'est-à-dire $Y_1 = X_1$, $Y_2 = X_1 + X_2$. Si chacun des couples (X_1, X_2) , (Y_1, Y_2) est formé de variables aléatoires indépendantes, le Théorème (4.1) permet d'affirmer que la variable aléatoire X_1 est normale (en revanche, on ne peut rien affirmer sur X_2). On peut d'ailleurs préciser que la variable aléatoire X_1 est dégénérée. Il suffit pour cela de remarquer que, dans le cas présent, l'équation (4.6) se réduit à la forme (4.7) avec $\theta(u) = 1$ (cf. (3.3)).

b) Supposons ensuite

$$A = \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix},$$

c'est-à-dire $Y_1 = X_1 \cos \omega - X_2 \sin \omega$, $Y_2 = X_1 \sin \omega + X_2 \cos \omega$. Si chacun des couples (X_1, X_2) , (Y_1, Y_2) est formé de variables aléatoires indépendantes, *et si ω n'est pas un multiple entier de $\pi / 2$* , le Théorème (4.1) permet d'affirmer que chacune des variables aléatoires X_1, X_2 est normale (en revanche, si ω est un multiple entier de $\pi / 2$, on ne peut rien affirmer, ni sur X_1 ni sur X_2). On reconnaîtra ici un résultat ayant des analogies avec le Théorème (2.1).

Remarquons enfin que l'énoncé du théorème de Bernstein-Darmois tel qu'il figure dans [3] (pag. 77 et pag. 499) est incorrect. En effet il entraîne notamment que, dans les hypothèses du cas particulier a) ci-dessus, la variable aléatoire X_2 est normale, ce qui est manifestement faux (il suffit, pour s'en convaincre, de prendre X_1 constante et X_2 non normale).

5. LE THÉORÈME DE SKITOVITCH-DARMOIS

Voici l'énoncé du théorème de Skitovitch-Darmois mentionné au paragraphe précédent:

(5.1) THÉORÈME. *Soit X un vecteur aléatoire, à valeurs dans \mathbf{R}^n , dont les composantes X_1, \dots, X_n sont indépendantes. Considérons les deux variables aléatoires Y_1, Y_2 définies par les relations*

$$\begin{cases} Y_1 = a_1 X_1 + \dots + a_n X_n \\ Y_2 = b_1 X_1 + \dots + b_n X_n \end{cases}.$$

où les coefficients a_i, b_i sont des scalaires donnés. Si les variables aléatoires Y_1, Y_2 sont elles mêmes indépendantes, alors, pour chaque indice i tel que $a_i b_i \neq 0$, la loi de X_i est normale (éventuellement dégénérée).

Ce théorème est bien plus profond que celui de Bernstein-Darmois (auquel il se réduit pour $n = 2$). Il permet en effet de démontrer directement le théorème de Cramér-Lévy, ainsi que certains résultats de α -factorisation.¹⁾

Il n'est donc pas étonnant que toutes les démonstrations connues du théorème de Skitovitch-Darmois fassent appel à des résultats profonds de la théorie des fonctions caractéristiques analytiques (théorème de Marcinkiewicz, théorèmes de α -factorisation). Voici en revanche une généralisation du théorème de Bernstein-Darmois qu'on peut obtenir par nos méthodes d'équations fonctionnelles.

(5.1) THÉORÈME. Soit X un vecteur aléatoire à valeurs dans \mathbf{R}^n ($n \geq 2$), dont les composantes X_1, \dots, X_n sont des variables aléatoires indépendantes. Soit $A = (a_{ij})$ une matrice réelle (n, n) non singulière, et supposons que les composantes du vecteur aléatoire $Y = AX$ sont également indépendantes.

Alors, pour chaque indice i tel que la i -ème colonne de A possède au moins deux éléments non nuls, la loi de X_i est normale (éventuellement dégénérée).

Pour rendre plus claire la démonstration, nous commencerons par démontrer un lemme préliminaire:

(5.2) LEMME. Soit X un vecteur aléatoire à valeurs dans \mathbf{R}^n ($n \geq 2$), dont les composantes X_1, \dots, X_n (par rapport à la base canonique (e_1, \dots, e_n) de \mathbf{R}^n) sont des variables aléatoires indépendantes²⁾. Alors la fonction caractéristique Φ de X vérifie la relation

$$\Phi(s) \Phi(s + ue_1 + ve_2) = \Phi(s + ue_1) \Phi(s + ve_2)$$

pour tout vecteur s de \mathbf{R}^n et pour tout couple (u, v) de scalaires.

¹⁾ Voici, par exemple, comment on peut en déduire le théorème de Cramér-Lévy (cf. [4], pag. 193). Soit λ, μ un couple de lois de probabilité sur \mathbf{R} , dont le produit de convolution est une loi normale; prenons un système de quatre variables aléatoires indépendantes X, X', Y, Y' tel que les lois de X et de X' coïncident avec λ et celles de Y et de Y' avec μ . Les deux variables aléatoires $(X + Y) + (X' + Y')$, $(X + Y) - (X' + Y')$ sont alors normales, non corrélées, donc indépendantes. En appliquant le théorème de Skitovitch-Darmois pour $n = 4$, on en déduit que chacune des lois λ, μ est normale.

²⁾ En fait, il suffirait de supposer que le triplet formé de X_1, X_2 et du système (X_3, \dots, X_n) est indépendant.

Démonstration du lemme. Désignons par φ_i la fonction caractéristique de X_i et par s_i la i -ème composante du vecteur s (par rapport à la base canonique de \mathbf{R}^n). On a alors

$$\Phi(s) = \varphi_1(s_1) \varphi_2(s_2) \prod_{i=3}^n \varphi_i(s_i),$$

$$\Phi(s + ue_1 + ve_2) = \varphi_1(s_1 + u) \varphi_2(s_2 + v) \prod_{i=3}^n \varphi_i(s_i),$$

$$\Phi(s + ue_1) = \varphi_1(s_1 + u) \varphi_2(s_2) \prod_{i=3}^n \varphi_i(s_i),$$

$$\Phi(s + ve_2) = \varphi_1(s_1) \varphi_2(s_2 + v) \prod_{i=3}^n \varphi_i(s_i),$$

d'où la conclusion.

Démonstration du théorème. Supposons que la première colonne de A possède deux éléments non nuls, et montrons que la loi de X_1 est normale. On pourra supposer, par exemple, $a_{11} \neq 0$ et $a_{21} \neq 0$. Quitte à multiplier chacune des deux premières lignes de A par un scalaire convenable, on pourra même supposer

$$(5.3) \quad a_{11} = a_{21} = 1.$$

Désignons par L l'application linéaire de \mathbf{R}^n dans lui-même dont la matrice, par rapport à la base canonique de \mathbf{R}^n , est la transposée de A . On a donc, pour tout i ,

$$(5.4) \quad Le_i = \mathbf{a}_i,$$

où \mathbf{a}_i désigne le vecteur (a_{i1}, \dots, a_{in}) .

Désignons en outre par φ_i la fonction caractéristique de X_i et par Φ celle de X :

$$(5.5) \quad \Phi(t_1, \dots, t_n) = \varphi_1(t_1) \dots \varphi_n(t_n).$$

Le vecteur aléatoire $Y = AX$ admet alors comme fonction caractéristique la fonction composée $\psi = \Phi \circ L$. En lui appliquant le lemme précédent, on trouve, pour tout scalaire u et pour tout vecteur s de \mathbf{R}^n ,

$$\psi(s) \psi(s + ue_1 - ue_2) = \psi(s + ue_1) \psi(s - ue_2),$$

c'est-à-dire

$$\Phi(Ls) \Phi(Ls + u\mathbf{a}_1 - u\mathbf{a}_2) = \Phi(Ls + u\mathbf{a}_1) \Phi(Ls - u\mathbf{a}_2).$$

Etant donné le scalaire t , appliquons cette relation au vecteur s déterminé

par $Ls = te_1$ (un tel vecteur existe puisque la matrice de L est régulière). On obtient

$$\Phi(te_1)\Phi(te_1 + u(a_1 - a_2)) = \Phi(te_1 + ua_1)\Phi(te_1 - ua_2),$$

c'est-à-dire, grâce à (5.5) et (5.3),

$$(5.6) \quad \begin{aligned} \varphi_1(t)\varphi_1(t) \prod_{i=2}^n \varphi_i(u(a_{1i} - a_{2i})) \\ = \varphi_1(t+u) \prod_{i=2}^n \varphi_i(ua_{1i}) \varphi_1(t-u) \prod_{i=2}^n \varphi_i(-ua_{2i}). \end{aligned}$$

Or, si $|u|$ est assez petit, on a $\varphi_i(u(a_{1i} - a_{2i})) \neq 0$ pour tout i , de sorte que la relation précédente peut s'écrire sous la forme

$$(5.7) \quad (\varphi_1(t))^2 = \varphi_1(t+u)\varphi_1(t-u)\theta(u).$$

Il en résulte, grâce à (3.1), que φ_1 est la fonction caractéristique d'une loi normale.

6. RETOUR SUR LES ÉQUATIONS FONCTIONNELLES DES PARAGRAPHERS PRÉCÉDENTS

Dans l'étude des équations fonctionnelles des paragraphes précédents, nous nous étions systématiquement bornés à rechercher les solutions dans l'ensemble des fonctions caractéristiques. Nous nous proposons maintenant d'étudier ces mêmes équations (ou des équations analogues) dans l'ensemble de toutes les fonctions complexes continues définies dans \mathbf{R} .

Nous commencerons par l'équation fonctionnelle

$$(6.1) \quad (f(t) |f(u)|)^2 = f(t+u)f(t-u),$$

qui se réduit manifestement à l'équation (3.2) dans le cas où f est une fonction caractéristique.

(6.2) THÉORÈME. *Soit f une application continue de \mathbf{R} dans \mathbf{C} , telle que $f(0) = 1$.*

Alors les deux propriétés suivantes sont équivalentes :

(a) *f vérifie la relation (6.1) pour tout couple t, u de nombres réels ;*

(b) il existe deux nombres réels a, b tels que l'on ait

$$f(t) = \exp (at^2 + ibt)$$

pour tout nombre réel t .

Nous démontrerons d'abord le lemme suivant:

(6.3) LEMME. Soit H une partie fermée de \mathbf{R} possédant les propriétés suivantes :

- (a) 0 appartient à H ;
- (b) si t appartient à H , alors $-t$ appartient à H ;
- (c) $2t$ appartient à H si et seulement si t appartient à H ;
- (d) si $t + u, t - u$ appartiennent à H , alors t appartient à H .

Dans ces conditions on a, soit $H = \{0\}$, soit $H = \mathbf{R}$.

Démonstration du Lemme. Supposons que l'ensemble H possède un élément t_0 non nul. D'après (b), on pourra supposer t_0 positif. En vertu de (b) et de (c), il suffira de montrer que H contient l'intervalle $[0, t_0]$. Or il résulte de (c) que H contient tout nombre de la forme $t_0/2^n$ avec $n \geq 0$. La propriété (d) entraîne alors que H contient aussi tout nombre de la forme $kt_0/2^n$ avec $0 \leq k \leq 2^n$. Puisque H est fermé, il contient l'intervalle $[0, t_0]$, ce qui achève la démonstration du Lemme.

Démonstration du théorème. Il suffit de démontrer l'implication (a) \Rightarrow (b). Supposons donc la propriété (a) vérifiée.

1) Plaçons-nous d'abord dans le cas où la fonction f est réelle et positive. La relation (6.1) devient alors

$$(f(t)f(u))^2 = f(t+u)f(t-u).$$

Puisque f est continue et non nulle à l'origine, il existe un nombre réel t_0 non nul tel que l'on ait $f(t_0) > 0$. Désignons par a la constante réelle déterminée par la condition $f(t_0) = \exp (at_0^2)$, et posons, pour tout nombre réel t ,

$$g(t) = \frac{f(t)}{\exp (at^2)}.$$

On a alors $g(0) = 1, g(t_0) = 1$; en outre la fonction g vérifie la même équation fonctionnelle que f :

$$(g(t)g(u))^2 = g(t+u)g(t-u).$$

En particulier:

$$(g(u))^2 = g(u)g(-u),$$

$$(g(t))^4 = g(2t).$$

Il en résulte que l'ensemble fermé $H = \{t: g(t) = 1\}$ vérifie les hypothèses du Lemme (6.3). Comme d'autre part H contient le nombre réel non nul t_0 , il coïncide avec \mathbf{R} , ce qui revient à dire que l'on a $f(t) = \exp(at^2)$ pour tout t .

2) Supposons ensuite $|f| = 1$. L'équation (6.1) devient alors

$$(6.4) \quad (f(t))^2 = f(t+u)f(t-u).$$

Puisque la fonction f est continue et qu'elle prend la valeur 1 à l'origine, il existe un intervalle ouvert I , centré à l'origine, dans lequel f ne prend jamais la valeur -1 . Pour tout élément t de cet intervalle, désignons par $\theta(t)$ la détermination principale de l'argument de $f(t)$. La fonction θ est alors continue, nulle à l'origine et vérifie l'équation $2\theta(t) = \theta(t+u) + \theta(t-u)$ pour tout couple t, u de nombres réels tels que $t+u$ et $t-u$ appartiennent à I .

En d'autres termes, θ est une fonction linéaire affine: il existe donc une constante réelle b telle que l'on ait $\theta(t) = bt$ (c'est-à-dire $f(t) = \exp(ibt)$) pour tout élément t de I . Si l'on pose

$$g(t) = \frac{f(t)}{\exp(ibt)},$$

la fonction g vérifie, elle aussi, l'équation (6.4), donc en particulier l'équation $(g(t))^2 = g(2t)$. Puisque d'autre part elle est identiquement égale à 1 sur I , il en résulte qu'elle est égale à 1 partout. En d'autres termes, on a

$$f(t) = \exp(ibt).$$

3) Plaçons-nous enfin dans le cas général. Puisque f vérifie l'équation fonctionnelle (6.1), il en est de même de la fonction réelle positive $|f|$. Celle-ci a donc (d'après 1)) la forme:

$$|f(t)| = \exp(at^2).$$

On peut alors considérer le rapport $f/|f|$; c'est une fonction de module égal à 1, vérifiant la même équation fonctionnelle que f . Par conséquent on a, d'après 2),

$$\frac{f(t)}{|f(t)|} = \exp(ibt).$$

Les deux dernières relations montrent que f a la forme désirée.

Remarque. De façon plus générale, on peut montrer que, si f est une application continue de \mathbf{R} dans \mathbf{C} vérifiant l'équation fonctionnelle (6.1) (mais non assujettie à la condition $f(0) = 1$), alors f est, soit identiquement nulle, soit de la forme

$$(6.5) \quad f(t) = \exp(at^2 + ibt + ic),$$

où a, b, c sont des constantes réelles.

En effet, si l'on suppose $f(0) = 0$, l'équation (6.1) (pour $t = u$) montre que f est identiquement nulle. Supposons donc $f(0) \neq 0$. La même équation (pour $t = u = 0$) fournit alors $|f(0)| = 1$. Par conséquent la fonction $f/f(0)$ vérifie encore l'équation fonctionnelle (6.1); puisqu'elle prend la valeur 1 à l'origine, le théorème précédent entraîne qu'elle est de la forme $\exp(at^2 + ibt)$. Si donc on pose $f(0) = \exp(ic)$, on obtient

$$f(t) = f(0) \exp(at^2 + ibt) = \exp(at^2 + ibt + ic).$$

On remarquera qu'inversement la fonction (6.5) vérifie l'équation fonctionnelle (6.1) quelles que soient les constantes réelles a, b, c ; cependant elle n'est une fonction caractéristique que dans le cas où la constante a est négative et la constante c est un multiple entier de 2π .

Occupons-nous à présent des deux autres équations fonctionnelles rencontrées au paragraphe 1, à savoir

$$(6.6) \quad \varphi(t) = (\varphi(t/\sqrt{c}))^c,$$

$$(6.7) \quad \varphi(t) = (\varphi(t/c))^c,$$

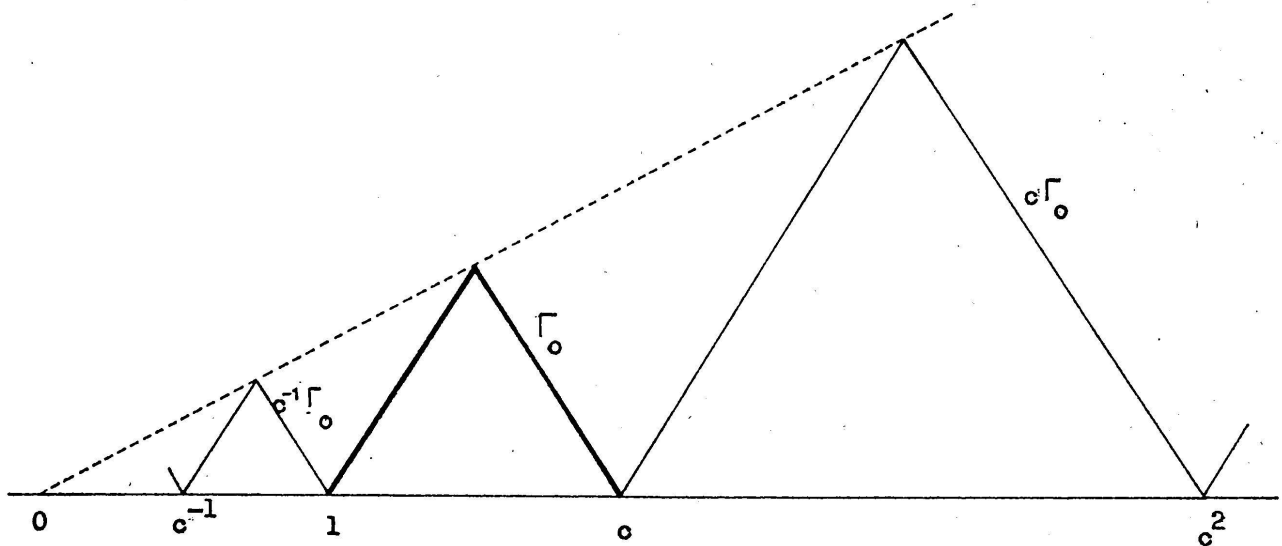
où c est une constante réelle, avec $c > 1$.

Dans chacune de ces deux équations la fonction inconnue φ sera supposée réelle, paire et strictement positive sur \mathbf{R} . En considérant la fonction $f(t) = \text{Log } \varphi(\sqrt{|t|})$ (resp. $f(t) = \text{Log } \varphi(t)$), l'équation (6.6) (resp. (6.7)) se réduit à

$$(6.8) \quad f(t) = cf(t/c).$$

Dans cette dernière équation la fonction inconnue f est réelle, paire, nulle à l'origine: il suffit donc de s'intéresser à sa restriction à l'intervalle $]0, +\infty[$. Or, si l'on désigne par Γ le graphe de cette restriction, l'équa-

tion (6.8) exprime le fait que Γ est invariant par rapport à l'homothétie de centre 0 et de rapport c , c'est-à-dire que l'on a $c\Gamma = \Gamma$. Pour avoir une solution, il suffit donc de se donner un graphe *arbitraire* Γ_0 sur l'intervalle $]1, c]$ et de poser ensuite $\Gamma = \bigcup_{n \in \mathbb{Z}} c^n \Gamma_0$. La figure ci-dessous fournit un exemple.¹⁾



Il est évident qu'en choisissant convenablement Γ_0 , on peut obtenir des solutions de classe C^∞ sur $]0, +\infty[$ et tendant vers 0 à l'origine.

On peut d'ailleurs remarquer que chacune des équations (6.6), (6.7), (6.8) est de la forme $f = h \circ f \circ g^{-1}$ (où f est la fonction inconnue). Le résultat précédent est alors un cas particulier de la proposition suivante:

(6.9) PROPOSITION. Soient T, U deux ensembles, g une bijection de T sur lui-même, h une bijection de U sur lui-même. Considérons l'équation fonctionnelle

$$(6.10) \quad f = h \circ f \circ g^{-1},$$

où la fonction inconnue f est une application de T dans U .

Supposons qu'il existe une partie T_0 de T telle que $(g^n(T_0))_{n \in \mathbb{Z}}$ constitue une partition de T . Dans ces conditions, pour toute application f_0 de T_0 dans U , il existe une application unique f de T dans U qui prolonge f_0 et qui vérifie l'équation fonctionnelle (6.10).

Démonstration. Unicité. Soit f une solution de (6.10). On a alors, pour tout entier n ,

$$f = h^n \circ f \circ g^{-n}$$

En particulier, si f prolonge f_0 , on a, pour tout élément t de $g^n(T_0)$,

¹⁾ L'idée de cette construction nous a été suggérée par Ph. Artzner.

$$f(t) = h^n [f_0(g^{-n}(t))].$$

Puisque $(g^n(T_0))_{n \in \mathbb{Z}}$ est une partition de T , la formule précédente détermine univoquement f sur l'ensemble T tout entier.

Existence. Désignons par Γ_0 le graphe de f_0 et par H la bijection de $T \times U$ sur lui-même définie par $H(t, u) = (g(t), h(u))$. L'ensemble

$$\Gamma = \bigcup_{n \in \mathbb{Z}} H^n(\Gamma_0)$$

est alors le graphe d'une application f de T dans U . Puisque Γ contient Γ_0 , f prolonge f_0 . On a en outre $H(\Gamma) = \Gamma$, et cette relation montre que f vérifie l'équation fonctionnelle (6.10).

BIBLIOGRAPHIE

- [1] BERNSTEIN, S. N. On a property characterizing the Gaussian Law. *Tr. Leningrad Politechn. Inst.*, 3 (1941), pp. 3-20.
- [2] DARMOIS, G. Sur une propriété caractéristique de la loi de probabilité de Laplace. *C. R. Acad. Sc. Paris*, 232 (1951), 1999-2000.
- [3] FELLER, W. *An Introduction to Probability Theory and Its Applications*. Vol. II, Wiley (1966).
- [4] HENNEQUIN, P. L. et A. TORTRAT. *Théorie des probabilités et quelques applications*. Masson et Cie (1965).
- [5] LUKACS, *Stochastic convergence*. Second edition, Academic Press (1975).
- [6] SKITOVITCH, V. P. Linear forms in independent random variables and the normal distribution law. *Izvestia AN SSSR, Ser. Mat.*, 18 (1954), pp. 185-200.

(Reçu le 26 août 1977)

Aimé Fuchs
Giorgio Letta

Université de Strasbourg
Département de Mathématique
7, rue René-Descartes
F-67084 Strasbourg

Vide-leer-empty