

ANNALES DE L'I. H. P., SECTION B

C. GUILBART

Produits scalaires sur l'espace des mesures

Annales de l'I. H. P., section B, tome 15, n° 4 (1979), p. 333-354

http://www.numdam.org/item?id=AIHPB_1979__15_4_333_0

© Gauthier-Villars, 1979, tous droits réservés.

L'accès aux archives de la revue « *Annales de l'I. H. P., section B* » (<http://www.elsevier.com/locate/anihpb>) 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/>

Produits scalaires sur l'espace des mesures

par

C. GUILBART

Mathématiques, Université de Lille I, 59650 Villeneuve-d'Ascq

SUMMARY. — We first prove that the concept of semi-scalar product on the space of all signed bounded measures is closely related to the concept of reproducing kernel on the basis space. Then we give a characterization for scalar products inducing weak topology on the space of all positive measures.

We finally give some examples of scalar products and a Glivenko-Cantelli theorem.

1. INTRODUCTION

Il est devenu classique d'utiliser la convergence faible des mesures en calcul des probabilités et en statistique mathématique : ainsi, depuis Laplace, de nombreux auteurs ont étudié cette convergence et ont appliqué les résultats obtenus à la résolution de problèmes très variés (comportement asymptotique des marches aléatoires, construction de test, etc.). Une des difficultés de ce genre d'étude vient du fait que les métriques compatibles avec la topologie faible sont souvent peu maniables : c'est le cas de la distance de P. Lévy et Yu. V. Prokhorov. En outre, ces distances ne sont pas prolongeables en normes sur l'espace des mesures bornées à signes ; et comme dans de nombreuses méthodes d'estimations on utilise simultanément la topologie faible et des « projections » sur des sous-espaces de dimensions finies, il nous a semblé intéressant d'étudier les produits scalaires sur l'espace des mesures bornées à signe et en particulier le problème de l'existence de normes hilbertiennes dont la trace sur l'espace des mesures positives soit la topologie faible.

Tout d'abord nous rappelons et établissons quelques propriétés élémentaires des espaces à noyaux reproduisants.

Ensuite, étant donné un espace mesurable Ω et un noyau reproduisant mesurable K défini sur $\Omega \times \Omega$, nous définissons un pseudo-produit scalaire $\langle \cdot, \cdot \rangle_K$ sur l'espace \mathcal{M}_Ω des mesures bornées à signe, en plongeant (théorème de plongement) \mathcal{M}_Ω dans l'espace de Hilbert H_K admettant K comme noyau reproduisant. Pour ce faire nous posons :

$$\phi(\mu) = \int K(\cdot, y) d\mu(y), \mu \in \mathcal{M}_\Omega$$

et

$$\langle \mu, \nu \rangle_K = [\phi(\mu), \phi(\nu)]_{H_K} \quad ; \quad \mu, \nu \in \mathcal{M}_\Omega$$

($[\cdot, \cdot]_{H_K}$ désignant le produit scalaire sur H_K) puis nous montrons que $\langle \mu, \nu \rangle_K = \int K d\mu \otimes \nu$; $\mu, \nu \in \mathcal{M}_\Omega$, ce qui entraîne en particulier $\langle \delta_x, \delta_y \rangle = K(x, y)$; $x, y \in \Omega$. Nous montrons ensuite (théorème de réintégration) qu'une condition nécessaire et suffisante pour qu'un produit scalaire $\langle \cdot, \cdot \rangle$ se réintègre

$$(i. e. : \forall \mu, \nu \in \mathcal{M}_\Omega, \langle \mu, \nu \rangle = \int \langle \delta_x, \delta_y \rangle d\mu \otimes \nu(x, y))$$

est que δ_ω soit un estimateur sans biais de la loi. Nous donnons à la suite un exemple simple de produit scalaire ne se réintégrant pas. Nous montrons enfin (théorème de caractérisation) que tout produit scalaire $\langle \cdot, \cdot \rangle$ sur \mathcal{M}_Ω , possédant certaines propriétés de régularités, admet une représentation de la forme $\langle \mu, \nu \rangle = [h(\phi(\mu)), h(\phi(\nu))]_{H_K}$ où h est linéaire et $K = \langle \delta_\cdot, \delta_\cdot \rangle$.

Le résultat fondamental de 5) (théorème de faible comparaison) peut se résumer de la façon suivante : si Ω est un espace métrique compact, les produits scalaires qui induisent la topologie faible sur \mathcal{M}_Ω^+ (espace des mesures positives bornées) sont ceux qui se réintègrent et dont le noyau associé $K(x, y) = \langle \delta_x, \delta_y \rangle$ est continu. Nous montrons que ce résultat ne s'étend pas au cas localement compact, mais que, dès que Ω est métrique séparable on peut toujours affirmer (théorème d'existence) l'existence de noyaux qui induisent la topologie faible sur \mathcal{M}_Ω^+ , mais pratiquement jamais sur \mathcal{M}_Ω tout entier (théorème de finitude). Nous montrons aussi (théorème du précompactifié) que bien souvent on peut choisir ces noyaux de telle façon que \mathcal{P}_Ω soit précompact.

Nous établissons une méthode générale pour construire de tels noyaux, nous donnons des exemples très simples lorsque $\Omega = \mathbb{R}^n$, et le lemme de majoration des normes produits montre toute la commodité de ces produits scalaires lorsque l'on doit travailler sur des espaces produits.

Nous terminons avec le « théorème de Glivenko-Cantelli » qui nous donne une bonne majoration de la vitesse de convergence de la loi empirique vers la loi théorique.

2. ÉTUDE HILBERTIENNE

Dans la première partie de cet article, nous allons utiliser la théorie des noyaux reproduisants; théorie qui a été introduite en 1943 par N. Aronsjajn dans [1], a été complétée dans [2] et largement utilisée depuis dans [7], [10], [5], [4], [6]... (dans [12], l'auteur a approfondi les résultats de [1] et [2]).

Soit Ω un espace abstrait, d'éléments x, y, z, \dots . Une fonction réelle K sur $\Omega \times \Omega$ est dite un noyau reproduisant si $K(x, y) = K(y, x)$ (symétrie) et si $\forall x_i, a_i \in \mathbb{R}, 1 \leq i \leq J : \sum_{1 \leq i, j \leq J} a_i a_j K(x_i, x_j) \geq 0$ (semi-définie positive).

Un espace d'applications réelles sur Ω d'éléments f, g, \dots est dit de Hilbert à noyau reproduisant K , s'il contient toutes les $f_y = K(\cdot, y)$, s'il est de Hilbert pour un produit scalaire (f, g) et si $(f, K(\cdot, y)) = f(y)$ (K est reproduisant).

THÉORÈME 1.2. — *Si H est K reproduisant; K est symétrique semi-défini positif, et H_0 , espace engendré par les combinaisons linéaires finies de $K(\cdot, y)$ est dense dans H . Réciproquement, à tout noyau reproduisant K on peut faire correspondre un Hilbert K -reproduisant, fermeture pour*

$$(K(\cdot, y), K(\cdot, z)) = K(y, z)$$

de H_0 (nous considérerons que cet espace K -reproduisant est unique et le noterons H_K , la non-unicité venant uniquement du choix arbitraire d'un représentant dans les différentes classes d'équivalence). Enfin, si K est borné, la K topologie est plus fine que la norme de la convergence uniforme.

Preuve. — Si H est K reproduisant, K est unique, symétrique et semi-défini positif. Si $(K(\cdot, y), f) = 0$ pour tout y de Ω , on a $f = 0$; $(K(\cdot, y), y \in \Omega)$ est donc total dans H ; en d'autres termes, H_0 est dense dans H .

Réciproque. — Si K est un noyau reproduisant, H_0 est un préhilbertien ayant la propriété de reproduction... (voir [1], p. 143-146).

Comme $|f(y)| \leq |(f, K(\cdot, y))| \leq \|f\|_K \cdot K(y, y)^{1/2}$ (inégalité de Schwartz) une suite convergente dans H_K converge simplement, uniformément si K est borné. ■

PROPOSITION 2.2. — H , Hilbert de fonctions réelles, est reproduisant si et seulement si $\forall y \in \Omega, f \in H; f \rightarrow f(y)$ est continue de H dans \mathbb{R} (poser $K(\cdot, y) = g_y$ définie par $f(y) = (f, g_y), \forall f \in H$).

THÉORÈME DE SÉPARABILITÉ 3.2. — H , Hilbert K -reproduisant borné, est séparable si et seulement s'il existe une suite de fonctions (f_i) telle que

$$K = \sum_i f_i \otimes f_i$$

(série convergent simplement sur $\Omega \times \Omega$); alors les f_i sont bornées et la convergence est uniforme en x pour tout y fixé (dite par la suite convergence simple uniforme).

Preuve. — Dans le cas H séparable, on prend (f_i) base orthonormale de H ; alors $(f_i, K(\cdot, y)) = f_i(y)$, de sorte que $K(\cdot, z) = \sum_{i=1}^{\infty} f_i(y) f_i(\cdot)$, série convergente dans H et donc simplement uniformément puisque K est borné. Les f_i sont alors bornées par $\|K\|_{\infty}$.

Réciproque. — Soit ϕ_i la forme linéaire sur H_0 définie par prolongement linéaire de $\phi_i(K(\cdot, y)) = f_i(y)$. Majorons sa norme :

$$\left[\phi_i \left(\sum_{j=1}^p a_j K(\cdot, y_j) \right) \right]^2 = \left[\sum_{j=1}^p a_j f_i(y_j) \right]^2 \leq \sum_{j,j',t} a_j a_{j'} f_t(y_j) f_t(y_{j'})$$

(somme finie de séries convergentes, on peut donc intervertir les sommations) d'où

$$\left[\phi_i \left(\sum_{j=1}^p a_j K(\cdot, y_j) \right) \right]^2 \leq \left(\sum_{j=1}^p a_j K(\cdot, y_j), \sum_{j=1}^p a_j K(\cdot, y_j) \right)_{H_K}.$$

ϕ_i étant de norme moindre que 1, elle est représentable par un élément de H ; la formule reproduisante dit alors que c'est f_i . De même on trouve que $\sum_i f_i(x) f_i(\cdot)$ a, comme opérateur sur H_0 , une norme moindre que $K(x, x)^{1/2}$, et la série du théorème est convergente dans H . Les f_i engendrent donc les $K(\cdot, y)$, et de ce fait sont totales dans H . ■

3. MESURABILITÉS

On suppose K borné et strictement positif sur la diagonale, dans toute la suite de ce travail.

H_K étant fermeture, pour une norme plus fine que l'uniforme, des $K(\cdot, y)$ on peut énoncer :

Si Ω a une structure mesurable et si $\forall y \in \Omega, K(\cdot, y)$ est mesurable, H_K est composé de *mesurables*.

Si Ω est topologique, si chaque point admet une base dénombrable de voisinages et si pour tout y de $\Omega, K(\cdot, y)$ est continu, H_K est composé de continues. Même chose pour la continuité uniforme.

Si K est *continu* et Ω *séparable*, H_K est séparable ((y_n) étant dense, les $K(\cdot, y_n)$ sont K -denses dans les $K(\cdot, y)$ qui sont totales dans H_K); le développement du théorème de séparabilité est alors composé de continues (resp. uniformément continues). La réciproque de cette dernière affirmation est vraie dans le cas séparable (convergence simple uniforme).

Enfin dans le cas abstrait, si H_K est séparable et si $(K(\cdot, y), y \in \Omega)$ sépare les points, on peut munir Ω d'une topologie métrisable séparable rendant K uniformément continu (prendre la structure uniforme rendant uniformément continues les f_i).

4. MESURES

Désormais, Ω est muni d'une structure mesurable. On appelle \mathcal{M}_Ω l'espace des mesures signées bornées, \mathcal{M}_Ω^+ les mesures positives bornées, \mathcal{P}_Ω les probabilités, \mathcal{M}_0 les combinaisons linéaires finies de mesures de Dirac $\delta_x, \mathcal{P}_0 = \mathcal{P}_\Omega \cap \mathcal{M}_0$.

Un (pseudo)-produit scalaire sur \mathcal{M}_Ω est noté $\langle \cdot, \cdot \rangle$ et l'on pose $\langle \delta_x, \delta_y \rangle = K(x, y)$.

LEMME 1.4. — Si K est mesurable borné, l'application linéaire de $(\mathcal{M}_0, \langle \cdot, \cdot \rangle)$ sur $(H_0, (\cdot, \cdot)_{H_K})$ définie par $\phi(\delta_y) = K(\cdot, y)$ est une isométrie entre ces deux espaces.

THÉORÈME DE PLONGEMENT 2.4. — On peut alors prolonger ϕ en une application linéaire de \mathcal{M}_Ω dans H_K par $\phi(\mu) = \int K(\cdot, y)d\mu(y)$; on note alors $(\mu, \nu)_K$ le pseudo-produit scalaire défini sur \mathcal{M}_Ω par $(\mu, \nu)_K = (\phi(\mu), \phi(\nu))_{H_K}$.

Preuve. — Pour μ fixée, la forme linéaire ϕ_μ définie sur H_0 par prolongement de $\phi_\mu(K(\cdot, x)) = \int K(y, x)d\mu(y)$ est continue (en effet, K étant

borné, il existe une constante $b > 0$ telle que $\forall f \in H_K, \|f\|_U \leq b \|f\|_{H_K}$; on a alors

$$\left| \phi_\mu \left(\sum_{i=1}^n a_i K(\cdot, x_i) \right) \right| \leq \left\| \sum_{i=1}^n a_i K(\cdot, x_i) \right\|_U \cdot \|\mu\|_{\text{var}}$$

$$\leq b \left\| \sum_{i=1}^n a_i K(\cdot, x_i) \right\|_{H_K} \cdot \|\mu\|_{\text{var}};$$

elle se prolonge donc en une forme linéaire continue sur H_K tout entier; il existe donc un élément h de H_K tel que $\forall x \in \Omega,$

$$\int K(x, y) d\mu(y) = (h, K(x, \cdot))_{H_K} = h(x);$$

$\int K(\cdot, y) d\mu(y)$ appartient donc à H_K . ■

Remarques. — Pour ce (pseudo)-produit scalaire, \mathcal{M}_0 est dense dans \mathcal{M}_Ω .

— On a $(\mu, \nu)_K = \int K d\mu \otimes \nu$ (en effet, pour μ fixée dans \mathcal{M}_Ω , on a :

$$\left| \int K d\mu \otimes \nu \right| \leq \|\phi(\nu)\|_U \times \|\mu\|_{\text{var}} \leq b \|\phi(\nu)\|_{H_K} \cdot \|\mu\|_{\text{var}};$$

d'où $\nu \rightarrow \int K d\mu \otimes \nu$ et $\nu \rightarrow (\nu, \mu)_K$ sont deux formes linéaires continues sur $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$ qui coïncident sur \mathcal{M}_0 ; elles sont donc égales.

— $(\cdot, \cdot)_K$ sépare les mesures de Dirac dès que $K(\cdot, x) = K(\cdot, z) \Rightarrow x = z$;

il sépare les mesures dès que $K = \sum_{i=1}^\infty c_i 1_{A_i} \otimes 1_{A_i}$, où (c_i) est une série

convergente de termes strictement positifs, (A_i) engendre la tribu de Ω (de même en remplaçant les A_i par des f_i continues bien choisies dans le cas topologique).

— Si $(\cdot, \cdot)_K$ est non dégénéré, ϕ est une isométrie et $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$ peut être considéré comme un sous-espace de H_K .

LEMME 3.4. — Soient (Ω, \mathcal{B}) un espace mesurable, K un noyau reproduisant $\mathcal{B} \otimes \mathcal{B}$ mesurable borné et $T_1 : (\Omega, \mathcal{B}) \rightarrow (\mathcal{M}_\Omega, (\cdot, \cdot)_K)$. Alors :

$$\omega \rightarrow \delta_\omega$$

① T_1 est bornée et faiblement mesurable.

② Si $(\cdot, \cdot)_K$ est séparable, T_1 est mesurable.

③ Si \mathcal{B} est la tribu borélienne d'une topologie (où tout point admet une base dénombrable de voisinages) rendant K continu, T_1 est continu.

④ Si de plus cette topologie est d -métrisable et $K - d \otimes d$ -uniformément continu, T_1 est uniformément continu.

Preuve. — ① Soit e une forme linéaire continue sur $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$; il existe alors une suite (μ_n) de \mathcal{M}_0 telle que $\forall \omega \in \Omega, (\delta_\omega, \mu_n)_K \rightarrow e(\delta_\omega)$. Comme $(\delta_\omega, \mu_n)_K = \int K(\omega, x) d\mu_n(x) = \phi(\mu_n)(\omega)$, $e \circ \delta$ est mesurable.

② et ③ découlent du fait que :

$$T_1^{-1}(B(\mu, a)) = \left\{ \omega \in \Omega : K(\omega, \omega) - 2\phi(\mu)(\omega) < a^2 - \int K d\mu \otimes \mu \right\}$$

④ $\| \delta\omega_1 - \delta\omega_2 \|_K^2 = K(\omega_1, \omega_1) + K(\omega_2, \omega_2) - 2K(\omega_1, \omega_2)$.

THÉORÈME DE RÉINTÉGRATION 4.4. — Soient (Ω, \mathcal{B}) un espace mesurable et $\langle \cdot, \cdot \rangle$ un produit scalaire sur \mathcal{M}_Ω , où $\langle \delta x, \delta y \rangle$ est $\mathcal{B} \otimes \mathcal{B}$ -mesurable et borné. Pour que $\langle \cdot, \cdot \rangle$ se réintègre

$$\left(i. e. : \forall \mu, \nu \in \mathcal{M}_\Omega, \langle \mu, \nu \rangle = \int \langle \delta x, \delta y \rangle d\mu \otimes \nu(x, y) \right)$$

il faut et il suffit que $T_1 : \omega \rightarrow \delta_\omega$ soit un estimateur sans biais d'ordre (1) de la loi.

Preuve. — *Condition nécessaire.* — Si l'on pose $K(x, y) = \langle \delta x, \delta y \rangle$ on a alors $\langle \cdot, \cdot \rangle = (\cdot, \cdot)_K$; comme $\langle \mu, \delta y \rangle = \int_\Omega \langle \delta x, \delta y \rangle d\mu(x)$, pour tout ν de \mathcal{M}_0 on a $\langle \mu, \nu \rangle = \int_\Omega \langle \delta x, \nu \rangle d\mu(x)$; \mathcal{M}_0 étant dense dans $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$ (donc dans $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle)$), pour toute forme linéaire $\langle \cdot, \cdot \rangle$ -continue, e , il existe une suite (ν_n) de \mathcal{M}_0 telle que $\forall \nu \in \mathcal{M}_\Omega, \langle \nu, \nu_n \rangle \rightarrow e(\nu)$; on a alors (pour μ fixée) $\int_\Omega \langle \delta_\omega, \nu_n \rangle d\mu(\omega) \rightarrow \int_\Omega e(\delta_\omega) d\mu(\omega)$ (théorème de convergence dominée) et $\int_\Omega \langle \delta_\omega, \nu_n \rangle d\mu(\omega) = \langle \mu, \nu_n \rangle$; d'où

$$e(\mu) = \int_\Omega e \circ \delta_\omega d\mu(\omega).$$

Condition suffisante. — T_1 étant un estimateur sans biais de la loi, le théorème de Fubini et la définition de l'intégrale de Pettis nous permettent alors d'écrire :

$$\begin{aligned} \int_{\Omega^2} \langle \delta x, \delta y \rangle d\mu \otimes \nu(x, y) &= \int_\Omega \left(\int_\Omega \langle \delta x, \delta y \rangle d\mu(x) \right) d\nu(y) \\ &= \left\langle \int_\Omega \delta x d\mu(x), \int_\Omega \delta y d\nu(y) \right\rangle = \langle \mu, \nu \rangle \quad \blacksquare \end{aligned}$$

Remarque. — Si K est un noyau reproduisant $\mathcal{B} \otimes \mathcal{B}$ -mesurable, borné, non dégénéré (i. e. : $\int K d\mu \otimes \mu = 0 \Rightarrow \mu = 0$), et si l'on munit \mathcal{M}_Ω de $(\cdot, \cdot)_K$, $T_1 : \omega \rightarrow \delta_\omega$ est un estimateur sans biais d'ordre 1 de la loi.

Remarque. — Tout produit scalaire sur $\mathcal{M}(\mathcal{B})$ ne se réintègre pas; nous allons donner un exemple de produit scalaire séparable sur $\mathcal{M}_{[0,1]}$ ne se réintégrant pas, et où $\langle \delta x, \delta y \rangle$ est continu et borné.

EXEMPLE 5.4. — Si l'on note \mathcal{M}_D (resp. \mathcal{M}_a) l'ensemble des mesures diffuses (resp. discrètes) sur $[0, 1]$, on a $\mathcal{M}_{[0,1]} = \mathcal{M}_D \otimes \mathcal{M}_a$ (somme directe) [en effet, toute mesure sur $[0, 1]$ se décompose d'une façon unique en la somme d'une mesure diffuse et d'une mesure discrète].

D'après le théorème d'existence (3.5) (que nous démontrerons par la suite) on peut exhiber un noyau reproduisant K borné, continu et non dégénéré (i. e. : $\int K d\mu \otimes \mu = 0 \Rightarrow \mu = 0$) sur $[0, 1]$.

Si l'on note h la projection parallèle à \mathcal{M}_D et si l'on pose

$$\langle \mu, \nu \rangle = (\mu, \nu)_K + (h(\mu), h(\nu))_K,$$

$\langle \cdot, \cdot \rangle$ est non dégénéré et l'on a :

① $\langle \delta x, \delta y \rangle = 2(\delta x, \delta y)_K = 2K(x, y)$ est continu et borné.

② $\langle \cdot, \cdot \rangle$ est séparable [en effet, si l'on note $\langle \mu, \nu \rangle_0 = (h(\mu), h(\nu))_K$, \mathcal{M}_0 est dense dans $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle_0)$; $(\mathcal{M}_0, \langle \cdot, \cdot \rangle_0)$ étant séparable, $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle_0)$ l'est; et comme $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$ est séparable, $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle)$ l'est].

③ $\langle \cdot, \cdot \rangle$ ne se réintègre pas [en effet, si μ_0 est non nulle et diffuse on a : $\langle \mu_0, \mu_0 \rangle = (\mu_0, \mu_0)_K \neq 0$ et

$$\int \langle \delta x, \delta y \rangle d\mu_0 \otimes \mu_0 = 2(\mu_0, \mu_0)_K \Big].$$

④ \mathcal{M}_0 n'est pas dense dans $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle)$. [En effet, si μ_0 est non nulle et diffuse et s'il existe une suite (ν_n) de \mathcal{M}_0 convergent vers μ_0 pour $\langle \cdot, \cdot \rangle$, on a ν_n tend vers μ_0 pour $(\cdot, \cdot)_K$ qui est moins fin que $\langle \cdot, \cdot \rangle$. Ce qui nous conduit à :

$$\begin{aligned} \cdot 2(\nu_n, \nu_n)_K &= \langle \nu_n, \nu_n \rangle \rightarrow \langle \mu_0, \mu_0 \rangle = (\mu_0, \mu_0)_K \\ \cdot &(\nu_n, \nu_n)_K \rightarrow (\mu_0, \mu_0)_K \neq 0 \end{aligned}$$

qui est impossible].

⑤ $T_1 : \omega \rightarrow \delta_\omega$ n'est pas un estimateur sans biais de la loi [on peut conclure directement à l'aide du théorème de réintégration. Mais on peut aussi le constater directement. Supposons qu'il soit sans biais; on a alors

pour μ_0 diffuse non nulle : $\left\langle \delta x, \int \delta_\omega d\mu_0(\omega) \right\rangle = (\delta x, \mu_0)_K$, et en utilisant la définition de l'intégrale de Pettis

$$\left\langle \delta x, \int \delta_\omega d\mu_0(\omega) \right\rangle = \int \langle \delta x, \delta_\omega \rangle d\mu_0(\omega) = 2 \int (\delta x, \delta_\omega)_K d\mu_0(x) = 2(\delta x, \mu_0)_K$$

qui est contradictoire, dès que $(\delta x, \mu_0)_K \neq 0$.

Il n'existe donc pas d'estimateurs sans biais de la loi pour ce produit scalaire.

Le théorème suivant caractérise les produits scalaires sur \mathcal{M}_Ω , où \mathcal{M}_0 est dense dans \mathcal{M}_Ω (Ce théorème ne s'applique donc pas à l'exemple (5.C.I)).

THÉORÈME DE CARACTÉRISATION 6.4. — (Ω, \mathcal{B}) étant un espace mesurable; si $\langle \cdot, \cdot \rangle$ est un pseudo-produit scalaire sur \mathcal{M}_Ω , borné sur \mathcal{P}_0 , tel que \mathcal{M}_0 est dense dans \mathcal{M}_Ω et $K(x, y) = \langle \delta x, \delta y \rangle$ est borné; il existe une application linéaire h unique de $L_K = \left\{ \phi(\mu) = \int K(\cdot, x) d\mu(x), \mu \in \mathcal{M}_\Omega \right\}$ dans H_K telle que $\langle \mu, \nu \rangle = (h(\phi(\mu)), h(\phi(\nu)))_{H_K}$.

Preuve. — μ étant un élément de \mathcal{M}_Ω , il existe une suite (ν_n) de \mathcal{M}_0 convergent vers μ dans $(\mathcal{M}_\Omega, \langle \cdot, \cdot \rangle)$; et comme

$$\langle \nu_n - \nu_p, \nu_n - \nu_p \rangle = (\phi(\nu_n) - \phi(\nu_p), \phi(\nu_n) - \phi(\nu_p))_{H_K}$$

la suite $(\phi(\nu_n))$ est une suite de Cauchy de H_K ; il existe donc un élément f unique de H_K tel que $(\phi(\nu_n) \xrightarrow{H_K} f)$; si l'on pose $h(\phi(\mu)) = f$, h est linéaire de L_K dans H_K , et comme $\forall \nu \in \mathcal{M}_\Omega, \|h(\phi(\nu))\|_{H_K} = \langle \nu, \nu \rangle^{1/2}$, on a :

$$\langle \mu, \nu \rangle = \frac{1}{4} (\langle \mu + \nu, \mu + \nu \rangle - \langle \mu - \nu, \mu - \nu \rangle) = (h(\phi(\mu)), h(\phi(\nu)))_{H_K} \quad \blacksquare$$

5. RELATIONS ENTRE PRODUITS SCALAIRES ET TOPOLOGIE FAIBLE

(Ω, d) étant un espace métrique, si l'on munit l'espace, C_b , des fonctions réelles continues bornées de la norme de la convergente uniforme; C_b est un espace de Banach et \mathcal{M}_Ω est un sous-espace du dual topologique, C_b^* , de C_b .

DÉFINITION. — On appelle *topologie faible* (ou encore étroite) de \mathcal{M}_Ω (que l'on note \mathcal{F}_d) la trace (sur \mathcal{M}_Ω) de la topologie affaiblie $\sigma(C_b^*, C_b)$.

NOTATIONS. — Nous écrivons indifféremment $(\mu_n \Rightarrow \mu)$ ou $(\mu_n \xrightarrow{\mathcal{F}_d} \mu)$ pour dire que μ_n tend faiblement vers μ .

Remarque. — On a $(\mu_n \xrightarrow{\mathcal{F}_d} \mu) \Leftrightarrow (\forall f \in C_b, \int f d\mu_n \rightarrow \int f d\mu)$.

Remarque. — La trace sur \mathcal{M}_Ω de la topologie forte de C_b^* n'est autre que la norme en variation totale.

En effet $\forall \mu \in \mathcal{M}_\Omega$, on a

$$\|\mu\|_{C_b^*} = \sup_{f \in C_b} \left(\left| \int \frac{f}{\|f\|_\infty} d\mu \right| \right) \leq \int 1 d\mu^+ + \mu^- = \|\mu\|_{\text{var}}$$

Maintenant comme il existe ⁽¹⁾ A tel que $\|\mu\|_{\text{var}} = \int 1_A - 1_{A^c} d\mu$ et comme μ est régulière (Ω est métrique) on a : $\|\mu\|_{\text{var}} \leq \|\mu\|_{C_b^*}$.

Nous nous placerons désormais dans le cas où (Ω, d) est séparable.

THÉORÈME DE FAIBLE COMPARAISON 1.5. — Soient (Ω, d) un espace métrique séparable et $\langle \cdot, \cdot \rangle$ un pseudo-produit scalaire sur \mathcal{M}_Ω . Si $\langle \delta x, \delta y \rangle$ est bornée, on peut énoncer :

1) Pour que $\langle \cdot, \cdot \rangle$ soit moins fin que la topologie faible sur \mathcal{M}_Ω^+ il faut et il suffit que $\langle \cdot, \cdot \rangle$ se réintègre et que $\langle \delta x, \delta y \rangle$ soit $d \otimes d$ -continu.

2) Si $\langle \cdot, \cdot \rangle$ vérifie 1), il existe une suite (f_i) uniformément bornée de C_b , telle que $\langle \delta x, \delta y \rangle = \sum_{i=1}^{\infty} f_i(x) f_i(y)$ (la convergence étant simple uniforme).

3) Si de plus $\langle \cdot, \cdot \rangle$ est non dégénéré, pour tout compact A de Ω , $(f_i|_A)$ est totale dans $U_b(A)$ et $\langle \cdot, \cdot \rangle$ a pour trace sur \mathcal{M}_A^+ la topologie faible. ($U_b(A)$ étant l'ensemble des fonctions réelles, bornées et uniformément continues sur A).

Preuve. — 1) a) Condition nécessaire : $\langle \cdot, \cdot \rangle$ étant moins fin sur \mathcal{M}_Ω^+ que \mathcal{F}_d on a $(x_n \rightarrow x, y_n \rightarrow y) \Rightarrow (\langle \delta x_n, \delta y_n \rangle \rightarrow \langle \delta x, \delta y \rangle)$ et $\langle \delta x, \delta y \rangle$ est continu. μ étant un élément de \mathcal{P}_Ω il existe $(\mu_n) \subset \mathcal{P}_0$ telle que $\mu_n \Rightarrow \mu$ ([1] p. 44). Alors

$$\langle \mu_n, \mu_n \rangle \rightarrow \langle \mu, \mu \rangle, \mu_n \otimes \mu_n \rightarrow \mu \otimes \mu, \int K d\mu_n \otimes \mu_n \rightarrow \int K d\mu \otimes \mu.$$

Finalement $\langle \mu, \mu \rangle = \int K d\mu \otimes \mu$ et $\langle \cdot, \cdot \rangle$ se réintègre puisque deux pseudo-produits scalaires ayant même norme sur \mathcal{P}_Ω sont égaux.

(1) Voir [9], page 99.

b) *Condition suffisante.* — Comme $K(x, y) = \langle \delta x, \delta y \rangle$ on a $\langle \mu, \nu \rangle = \int K d\mu \otimes \nu$ avec K continu et borné. D'où si $(\mu_n \xrightarrow{n \rightarrow \infty} \mu)$ dans \mathcal{P}_Ω :

$$\int K d(\mu_n - \mu) \otimes (\mu_n - \mu) \rightarrow \int K d\mu \otimes \mu + \int K d\mu \otimes \mu - 2 \int K d\mu \otimes \mu = 0.$$

Supposons maintenant que $(\mu_n \xrightarrow{n \rightarrow \infty} \mu)$ dans \mathcal{M}_Ω^+ ; si $\mu = 0$ on a $\mu_n(\Omega) \rightarrow 0$ et $\int K d\mu_n \otimes \mu_n \xrightarrow{n \rightarrow \infty} 0$; si $\mu \neq 0$, on a

$$\left(\frac{\mu_n}{\mu_n(\Omega)} \Rightarrow \frac{\mu}{\mu(\Omega)} \right) \text{ qui implique } \left(\frac{\mu_n}{\mu_n(\Omega)} \xrightarrow{\langle \cdot, \cdot \rangle} \frac{\mu}{\mu(\Omega)} \right) \text{ et } (\mu_n \xrightarrow{\langle \cdot, \cdot \rangle} \mu).$$

2) $(\mathcal{P}_\Omega, \mathcal{F}_d)$ étant séparable, $(\mathcal{P}_\Omega, (\cdot, \cdot)_K)$ l'est et de ce fait $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$ l'est; 2) découle alors directement du théorème de séparabilité (2.2) (puisque H_K est séparable).

3) En vertu de 2), il existe une suite (f_i) de C_b telle que

$$K(x, y) = \langle \delta x, \delta y \rangle = \sum_{i=1}^{\infty} f_i(x) f_i(y).$$

A étant compact dans Ω , le théorème classique de Dini nous permet d'affirmer que la convergence de cette série est uniforme sur la diagonale de $A \times A$ et l'inégalité de Schwartz nous permet d'écrire que la convergence

de la série $\sum_{i=1}^{\infty} f_i \otimes f_i$ est uniforme sur $A \times A$.

Si (f_i/A) n'est pas totale dans $C_b(A)$, on peut exhiber g dans $C_b(A) - \overline{S_p((f_i))}(\overline{S_p((f_i))})$ étant l'adhérence uniforme de $S_p((f_i))$ enveloppe vectorielle de la suite (f_i) ; il existe alors (corollaire du théorème de Hahn-Banach) une forme linéaire continue, F , nulle sur $\overline{S_p((f_i))}$ et non nulle en g . Le théorème de Riesz nous permet d'affirmer qu'il existe $\mu \in \mathcal{M}_A$ telle que $\forall f \in C_b(A), F(f) = \int_A f(x) d\mu(x)$; alors si pour tout B de $\mathcal{B}(d)$ on pose $\bar{\mu}(B) = \mu(A \cap B)$, $\bar{\mu} \in \mathcal{M}_\Omega$ et l'on a $\bar{\mu} \neq 0$ et $\langle \bar{\mu}, \bar{\mu} \rangle = 0$ qui est contraire à $\langle \cdot, \cdot \rangle$ non dégénéré. (f_i/A) étant totale dans $C_b(A)$ la topologie trace de $\langle \cdot, \cdot \rangle$ sur \mathcal{P}_A est la topologie faible.

Supposons maintenant $(\mu_n) \cup \{ \mu \} \subset \mathcal{M}_A^+$; (f_i/A) étant totale dans $C_b(A)$, il existe une combinaison linéaire h_0 des f_i telle que

$$\| h_0/A - 1_A \|_U < \frac{1}{2};$$

on a alors $h_0/A \geq \frac{1}{2} 1_A$, $\int h_0/A d\mu_n \geq \frac{1}{2} \mu_n(A)$ et $(\|\mu_n\|_{\text{var}})$ est bornée; on a donc

$$(\mu_n \xrightarrow{\langle \dots \rangle} \mu) \Rightarrow (\mu_n(A) \rightarrow \mu(A)).$$

Si $\mu = 0$, on a

$$(\mu_n \Rightarrow \mu) \Leftrightarrow (\mu_n \xrightarrow{\langle \dots \rangle} \mu)$$

Si $\mu \neq 0$, on a :

$$(\mu_n \Rightarrow \mu) \Leftrightarrow \left(\frac{\mu_n}{\mu_n(A)} \Rightarrow \frac{\mu}{\mu(A)} \right) \Leftrightarrow \left(\frac{\mu_n}{\mu_n(A)} \xrightarrow{\langle \dots \rangle} \frac{\mu}{\mu(A)} \right) \Leftrightarrow (\mu_n \xrightarrow{\langle \dots \rangle} \mu) \quad \blacksquare$$

Remarque. — Si l'on se suppose placé dans les conditions de l'assertion (3) avec en plus Ω compact, on peut énoncer :

- . \mathcal{F}_d et $\langle \dots \rangle$ ont même topologie trace sur \mathcal{M}_Ω^+ .
- . \mathcal{F}_d est plus fine que $\langle \dots \rangle$ sur \mathcal{M}_Ω tout entier. En effet, si $\mathcal{F}_n \xrightarrow{\mathcal{F}_d} \mu$, $(\mu_n) \cup \{\mu\}$ est compact pour \mathcal{F}_d , donc borné pour \mathcal{F}_d (prop. 3, p. 6, chap. III, § 2, n° 3 de [3]), donc borné pour la topologie forte (prop. (1), p. 86, chap. IV, § 3, n° 3 de [3]), et de ce fait borné pour la norme de la variation totale. Ce qui nous permet de conclure puisque $\sum_i f_i \otimes f_i$ converge uniformément) [Cette remarque est vraie pour Ω non compact].

. Ω étant compact, on a $C_b^* = \mathcal{M}_\Omega$ (théorème de Riesz); toute boule en variation $\|\mu\|_{\text{var}} \leq p$ est alors compacte pour \mathcal{F}_d (prop. 1, p. 112, chap. IV, § 5, n° 2 de [3]) puisque $\|C_b^*\| = \|\text{var}\|$; la topologie de $\langle \dots \rangle$ étant moins fine que \mathcal{F}_d et séparée, lui est donc égale sur cette boule (qui est donc compacte pour $\langle \dots \rangle$).

- . \mathcal{P}_Ω étant fermé dans la boule $\|\mu\|_{\text{var}} \leq 1$ est aussi compact.
- . $(\mathcal{M}_\Omega^+, \langle \dots \rangle)$ est complet. En effet, si (μ_n) est de Cauchy dans $(\mathcal{M}_\Omega^+, \langle \dots \rangle)$, $(\mu_n(\Omega))$ est bornée; (μ_n) est donc incluse dans une boule en variation qui est compacte dans $(\mathcal{M}_\Omega, \langle \dots \rangle)$; ce qui nous permet de conclure.

. Par contre dès que \mathcal{M}_Ω est de dimension infinie, $(\mathcal{M}_\Omega, \langle \dots \rangle)$ n'est pas complet; en effet, il l'est pour $\|C_b^*\| = \|\text{var}\|$ qui est strictement plus fine (remarque 2, p. 111, chap. IV, § 5, n° 1 de [3]) que \mathcal{F}_d , donc que $\langle \dots \rangle$.

Le théorème ci-dessous montre qu'en général il est inutile de chercher un produit scalaire dont la topologie associée se confonde avec la topologie faible sur \mathcal{M}_Ω tout entier. Par contre dès que l'on se limite à \mathcal{M}_Ω^+ le « théorème d'existence » (démontré un peu plus loin) résout positivement cette question.

THÉORÈME DE FINITUDE 2.5. — (Ω, d) étant un espace métrique séparable; une condition nécessaire et suffisante pour qu'il existe un noyau reproduisant K dont la topologie associée se confonde avec la topologie faible sur \mathcal{M}_Ω tout entier est que Ω soit de cardinal fini.

Preuve. — La condition nécessaire découle directement du fait que $(\mathcal{M}_\Omega, \mathcal{F}_d)$ est métrisable si et seulement si Ω est de cardinal fini.

La condition suffisante est immédiate.

Remarque. — (Ω, d) étant séparable, si K est continu et si $K = \sum_i f_i \otimes f_i$ (la convergence étant uniforme), \mathcal{F}_d est plus fine que $(\cdot, \cdot)_K$ sur \mathcal{M}_Ω tout entier.

Remarque. — Le théorème de faible comparaison ne peut être amélioré; (Ω, d) étant séparable et $\langle \cdot, \cdot \rangle$ étant un produit scalaire (non dégénéré) vérifiant l'assertion (1) du théorème de faible comparaison, la topologie trace de $\langle \cdot, \cdot \rangle$ sur \mathcal{M}_Ω^+ peut être strictement moins fine que \mathcal{F}_d . Il suffit pour cela de prendre $\Omega = \mathbb{R}$, de choisir (pour tout $i \in \mathbb{N}^*$) une suite $f_{i,j}$ de fonctions continues, bornées par 1, nulles sur $] - \infty, - (i + 1) [\cup] i + 1, + \infty [$, telle que la suite des restrictions $(f_{i,j} / [-i, +i])$ soit totale dans

$C_b([-i, +i])$ et de poser $K = \sum_{i,j=1}^\infty \frac{1}{2^{i+j}} f_{i,j} \otimes f_{i,j}$; $(f_{i,j}, (i, j) \in \mathbb{N}^* \times \mathbb{N}^*)$

étant totale dans l'ensemble des fonctions continues à support compact, $(\cdot, \cdot)_K$ est non dégénéré et vérifie l'assertion 1) du théorème de faible comparaison; et pourtant (δ_n) est une suite tendant vers zéro pour $(\cdot, \cdot)_K$ alors qu'elle ne converge pas vers 0 pour la métrique de Prokorov.

Remarque. — (Ω, d) étant séparable si $\langle \cdot, \cdot \rangle$ a pour trace \mathcal{F}_d sur \mathcal{M}_Ω^+ (resp. \mathcal{P}_Ω), $\langle \cdot, \cdot \rangle$ est non dégénéré dès que Ω s'écrit comme une réunion dénombrable de compact.

Remarque. — Si $K = \sum f_i \otimes f_i$ la convergence étant uniforme en (x, y) , \mathcal{F}_d est plus fine que $(\cdot, \cdot)_K$ sur \mathcal{M}_Ω tout entier.

COROLLAIRE 3.5. — (Ω, d) étant séparable; si K est un noyau reproduisant borné, $\mathcal{B}(d) \otimes \mathcal{B}(d)$ -mesurable et si H_K est séparable, l'ensemble des combinaisons convexes de mesures de Dirac est dense dans $(\mathcal{P}_\Omega, (\cdot, \cdot)_K)$.

Preuve. — L'ensemble des combinaisons convexes de mesures de Dirac étant dense dans $(\mathcal{P}_\Omega, \mathcal{F}_d)$, la proposition ci-dessus nous permet de conclure. ■

THÉORÈME D'EXISTENCE 4.5. — Si (Ω, d) est séparable, il existe un noyau reproduisant, K , non dégénéré (i. e. : $\int K d\mu \otimes \mu = 0 \Rightarrow \mu = 0$) tel que la topologie trace de $(\cdot, \cdot)_K$ sur \mathcal{M}_Ω^+ soit la topologie faible.

Preuve. — Il existe alors une métrique d' totalement bornée ⁽²⁾ équivalente à d ; l'espace des applications réelles, bornées, d' -uniformément continues muni de la norme de convergence uniforme (relative à d') étant séparable, il existe une suite (f_i) bornée par 1 et totale dans l'ensemble de ces applications; si l'on pose $K = \sum_{i=1}^{\infty} \frac{1}{2^i} f_i \otimes f_i$, $(\cdot, \cdot)_K$ est moins fin sur \mathcal{M}_Ω^+ que la topologie faible (théorème de faible comparaison). Maintenant si $\mu_n \xrightarrow[n \rightarrow \infty]{\langle \cdot, \cdot \rangle} \mu$ dans \mathcal{P}_Ω , on a par linéarité et limite uniforme : $\int f d\mu_n \rightarrow \int f d\mu$ pour toute application f d' -uniformément continue et bornée; le théorème de Billingsley ⁽³⁾ nous permet alors de conclure pour \mathcal{P}_Ω ; on passe alors de \mathcal{P}_Ω à \mathcal{M}_Ω^+ comme dans le théorème de faible comparaison. Enfin, si $(\mu, \mu)_K = 0$, on a $(\int f_i d\mu = 0, \forall i \in \mathbb{N}^*)$ qui est équivalent à $(\int f d\mu = 0, \forall f$ d' -uniformément continue), d'où $\mu = 0$. ■

Remarque. — \mathcal{F}_d est alors plus fine que $(\cdot, \cdot)_K$ sur \mathcal{M}_Ω tout entier.

K étant $d \otimes d$ continu, si (x_n) est dense dans (Ω, d) , (δx_n) est totale dans $(\mathcal{M}_\Omega, (\cdot, \cdot)_K)$.

Remarque. — (Ω, d) étant séparable; si Ω_1 est un borélien de Ω tel que $\overline{\Omega_1} = \Omega$ (resp. $\overline{\Omega_1}$ est compact, Ω_1 est convexe et Ω est de Hilbert), et si $\langle \cdot, \cdot \rangle$ est un produit scalaire sur \mathcal{M}_Ω dont la topologie sur \mathcal{M}_Ω^+ est \mathcal{F}_d ; $\langle \mu/C_1, \mu/C_1 \rangle$ admet \mathcal{F}_d pour trace sur $\mathcal{M}_{\Omega_1}^+$.

Remarque. — (Ω, d) étant séparable, A et B deux boréliens de Ω , $\langle \cdot, \cdot \rangle_A$ et $\langle \cdot, \cdot \rangle_B$ deux produits scalaires sur \mathcal{M}_A et \mathcal{M}_B ; si pour tout $\mu, \nu \in \mathcal{M}_{\Omega_1}$ ($\Omega_1 = A \cup B$) on pose $\langle \mu, \nu \rangle = \langle \mu/A, \mu/A \rangle_A + \langle \mu/B, \nu/B \rangle_B$, on peut énoncer :

- 1) $\langle \cdot, \cdot \rangle$ est un produit scalaire sur \mathcal{M}_{Ω_1} .
- 2) Si $A \cap B = \emptyset$ et si $\langle \cdot, \cdot \rangle_A$ (resp. $\langle \cdot, \cdot \rangle_B$) est plus fin que \mathcal{F}_d sur \mathcal{M}_A^+ (resp. \mathcal{M}_B^+), il est de même pour $\langle \cdot, \cdot \rangle$ sur $\mathcal{M}_{\Omega_1}^+$.
- 3) Si $d(A, B) > 0$ et si $\langle \cdot, \cdot \rangle_A$ (resp. $\langle \cdot, \cdot \rangle_B$) admet \mathcal{F}_d pour trace sur \mathcal{M}_A^+ (resp. \mathcal{M}_B^+); $\langle \cdot, \cdot \rangle$ admet \mathcal{F}_d pour trace sur $\mathcal{M}_{\Omega_1}^+$.

⁽²⁾ d' totalement bornée est identique à (Ω, d') précompact.

⁽³⁾ Voir [11] par exemple.

Remarque. — (Ω, d) étant séparable, localement compact et dénombrable à l'infini; si (f_i) est totale dans l'ensemble des fonctions continues à support compact et si $K = 1 + \sum_{i=1}^{\infty} f_i \otimes f_i$ est borné, $(\cdot, \cdot)_K$ admet \mathcal{F}_d pour trace sur \mathcal{M}_{Ω}^+ .

THÉORÈME DE PRÉCOMPACTITÉ 4.5. — *Si Ω est un espace métrique séparable, tel qu'il existe une métrique équivalente d rendant Ω précompact et borélien de son complété; on peut exhiber un noyau reproduisant K tel que :*

- 1) $\|\cdot\|_K$ a pour trace sur \mathcal{M}_{Ω}^+ la topologie faible,
- 2) $(\mathcal{P}_{\Omega}, \|\cdot\|_K)$ est précompact.

Preuve. — Soient (f_i) une suite uniformément bornée, totale dans $U_b(\Omega, d)$ et (α_i) une série strictement positive et convergente. Si l'on pose

$$K = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i,$$

- 1) découle directement de la démonstration du théorème d'existence.
- 2) Le complété $(\bar{\Omega}, \bar{d})$ de (Ω, d) est compact, (\bar{f}_i) est totale dans $U_b(\bar{\Omega}, \bar{d})$,

et si l'on pose $\bar{K} = \sum_{i=1}^{\infty} \alpha_i \bar{f}_i \otimes \bar{f}_i$, $\|\cdot\|_{\bar{K}}$ a pour trace sur $\mathcal{M}_{\bar{\Omega}}^+$ la topologie faible.

$(\mathcal{P}_{\bar{\Omega}}, \mathcal{F}_{\bar{d}})$ étant compact, $(\mathcal{P}_{\bar{\Omega}}, \|\cdot\|_{\bar{K}})$ l'est. Maintenant si pour tout $(\mu, B) \in \mathcal{M}_{\Omega}^+ \times \mathcal{B}(\bar{d})$ on pose $\mu(B) = \bar{\mu}(B \cap \Omega)$, l'application

$$\begin{aligned} \psi : (\mathcal{P}_{\Omega}, \|\cdot\|_K) &\rightarrow (\mathcal{P}_{\bar{\Omega}}, \|\cdot\|_{\bar{K}}) \\ \mu &\rightarrow \bar{\mu} \end{aligned}$$

est une isométrie de \mathcal{P}_{Ω} sur $\psi(\mathcal{P}_{\Omega})$; $(\mathcal{P}_{\Omega}, \|\cdot\|_K)$ est donc précompact. ■

Remarque. — Il suffit en vertu de la démonstration du théorème de précompactité de regarder la démonstration du théorème de construction (ci-dessous) pour s'apercevoir que les produits scalaires exhibés sur \mathbb{R}^n (dans le paragraphe suivant) sont tels que \mathcal{P}_{Ω} est précompact.

6. CONSTRUCTION

Nous allons nous limiter à donner des produits scalaires sur $\mathcal{M}_{\mathbb{R}}, \mathcal{M}_{\mathbb{R}^n}, \mathcal{M}_A$ (où $A \subset \mathbb{R}$ (resp. $A \subset \mathbb{R}^n$)) dont la topologie trace sur $\mathcal{M}_{\mathbb{R}}^+, \mathcal{M}_{\mathbb{R}^n}^+, \mathcal{M}_A^+$ est égale à (resp. moins fine que) \mathcal{F}_d .

THÉOREME DE CONSTRUCTION 1.6. — ψ étant un homéomorphisme de \mathbb{R} sur $\psi(\mathbb{R}) \subset]-a_\psi, a_\psi[$ (avec $a_\psi < +\infty$), si $(\alpha_{p(n)})$ est une suite strictement positive et si

$$K(x, y) = \sum_{n=0}^{\infty} \alpha_{p(n)} (\psi(x))^{p(n)} (\psi(y))^{p(n)}$$

(resp : $K(x, y) = \sum_{n=0}^{\infty} \alpha_{n(p)} e^{-p(n)\psi(x)} e^{-p(n)\psi(y)}$) [où $p(n)$ est une suite croissante

de \mathbb{N} telle que $(x^{p(n)}, n \in \mathbb{N}^*)$, (resp : $(e^{-p(n)x}, n \in \mathbb{N}^*)$) soit totale dans $U_b([-a_\psi, a_\psi])$ (*) est borné, le produit scalaire $(\cdot, \cdot)_K$ a pour topologie trace, sur $\mathcal{M}_\mathbb{R}^+$, la topologie faible.

Preuve. — Les deux démonstrations étant identiques, nous établirons uniquement celle relative au premier noyau. \mathbb{R} étant muni de la métrique $d(x, y) = |\psi(x) - \psi(y)|$ est précompact (puisque ψ est une isométrie de (\mathbb{R}, d) sur $(\psi(\mathbb{R}), |\cdot|)$). Si l'on pose $f_{p(n)}(x) = x^{p(n)}$, $(f_{p(n)})$ est totale dans $U_b(\psi(\mathbb{R}))$, donc dans $U_b(\psi(\mathbb{R}))$; g étant d -uniformément continue, $g \circ \psi^{-1}$ est uniformément continue sur $\psi(\mathbb{R})$; il existe donc $\lambda_1, \dots, \lambda_k, f_{p(n_1)}, \dots, f_{p(n_k)}$ tels que $\|g \circ \psi^{-1} - \sum \lambda_i f_{p(n_i)}\|_u < \varepsilon$ (dans (\mathbb{R}, d)); $(\psi^{p(n)})$ est donc totale dans $U_b(\mathbb{R}, d)$; si $(\mu_n \xrightarrow[n \rightarrow \infty]{\mathcal{F}_d} \mu)$ dans $\mathcal{P}_\mathbb{R}$, on a par linéarité et limite uniforme $(\forall f \in U_b(\mathbb{R}, d), \int f d\mu_n \rightarrow \int f d\mu)$; d'où $(\forall f \in C_b(\mathbb{R}), \int f d\mu_n \rightarrow \int f d\mu)$ (théorème de Billingsley (*)); le théorème de faible comparaison nous permet alors d'affirmer que \mathcal{F}_d et $(\cdot, \cdot)_K$ ont même trace sur $\mathcal{P}_\mathbb{R}$, et l'on passe de $\mathcal{P}_\mathbb{R}$ à $\mathcal{M}_\mathbb{R}^+$ comme à la fin du théorème de faible comparaison. $(\cdot, \cdot)_K$ est bien sûr dégénéré. ■

Remarque. — Si la convergence des séries (ci-dessus) est uniforme, la topologie faible est plus fine que $(\cdot, \cdot)_K$ sur $\mathcal{M}_\mathbb{R}$ tout entier. En faisant varier ψ et (α_n) dans le théorème ci-dessus on obtient les noyaux suivants :

$$K_1(x, y) = \sum_{i=0}^{\infty} \frac{1}{2^i} \left(\frac{x}{1+|x|} \right)^i \left(\frac{y}{1+|y|} \right)^i = \frac{1}{1 - \frac{1}{2} \frac{xy}{(1+|x|)(1+|y|)}}$$

(*) On peut, par exemple, prendre $(p(n) = n)$ ou $(p(0) = 0$ et $p(n) = 2n + 1, n \in \mathbb{N})$. Maintenant si l'on remplace « $\psi(x)$ » par « $a\varphi + \psi(x)$ » il suffit (en vertu du théorème de Müntz, p. 22, de [8]) de prendre pour $(p(n))$ une suite croissante telle que $p(0) = 0$ et

$$\sum_{n>0} \frac{1}{p(n)} \text{ diverge.}$$

(*) Voir [11] (par exemple).

$$K_2(x, y) = \sum_{i=0}^{\infty} \frac{1}{n!} \left(\frac{x}{1+|x|} \right)^n \left(\frac{y}{1+|y|} \right)^n = e^{\left(\frac{x}{1+|x|}\right)\left(\frac{y}{1+|y|}\right)}$$

$$K_3(x, y) = \sum_{i=0}^{\infty} \frac{1}{10^n} \exp\left(-n \frac{x}{1+|x|}\right) \exp\left(-n \frac{y}{1+|y|}\right) = \left(1 - \frac{1}{10} \exp\left(-\frac{x}{1+|x|}\right) \exp\left(-\frac{y}{1+|y|}\right)\right)^{-1}$$

$$K_4(x, y) = \sum_{i=0}^{\infty} \frac{1}{n!} \exp\left(-\frac{nx}{1+|x|}\right) \exp\left(-\frac{ny}{1+|y|}\right) = \exp\left(\exp\left(-\frac{x}{1+|x|}\right) \exp\left(-\frac{y}{1+|y|}\right)\right)$$

$$K_5(x, y) = 1 + \sum_{i=1}^{\infty} \frac{n}{n!} \left(\frac{x}{1+|x|} \right)^n \left(\frac{y}{1+|y|} \right)^n = 1 + \frac{xy}{(1+|x|)(1+|y|)} e^{\left(\frac{x}{1+|x|}\right)\left(\frac{y}{1+|y|}\right)}$$

$$K_6(x, y) = \sum_{i=0}^{\infty} \frac{1}{2^n} \left(\frac{x}{|x|} e^{-|x|} \right)^n \left(\frac{y}{|y|} e^{-|y|} \right)^n = \frac{1}{1 - \frac{1}{2} \frac{xy}{|xy|} e^{-|x||y|}}$$

$$K_7(x, y) = \sum_{i=0}^{\infty} \frac{1}{3^n} (\text{Arctg}(x))^n (\text{Arctg}(y))^n = \frac{1}{1 - \frac{1}{3} \text{Arctg}(x) \text{Arctg}(y)}$$

⋮ ⋮ ⋮

Lorsque Ω est borné dans \mathbb{R} , on peut prendre pour noyau $K(x, y) = e^{xy}, \dots$, (voir [6 bis], p. 93).

PROPOSITION 2.6. — Si l'on pose $K(x, y) = e^{-\frac{1}{2}(x-y)^2}$, $\int K d\mu \otimes \mu$ est un produit scalaire sur $\mathcal{M}_{\mathbb{R}}$ dont la trace sur $\mathcal{M}_{\mathbb{R}}^+$ est moins fine que la topologie faible, et dont la trace sur $\mathcal{M}_{\mathbb{A}}^+$ est, pour tout compact A , la topologie faible.

Preuve. — μ étant un élément de $\mathcal{M}_{\mathbb{R}}$, si

$$\int e^{-1/2(x-y)^2} d\mu \otimes \mu(x, y) = 0,$$

on a :

$$\int \exp(iu(x-y)) \exp\left(-\frac{1}{2}u^2\right) du d\mu \otimes \mu(x, y) = 0$$

(puisque $e^{-1/2(x-y)^2} = \int \exp(iu(x-y)) \times \exp\left(-\frac{1}{2}u^2\right) du$ [théorème de

réciprocité de Fourier]), d'où

$$\int \exp\left(-\frac{1}{2}u^2\right) |f_\mu(u)|^2 du = 0,$$

f_μ étant la transformée de Fourier de μ ; f_μ est donc identiquement nulle, et de ce fait μ aussi. ■

Remarque. — Ce que l'on vient de faire pour $e^{-1/2(x-y)^2}$ s'étend immédiatement à $e^{-|x-y|}$, etc.

L'intérêt de ce produit scalaire réside dans sa simplicité et dans le fait que toute mesure de Dirac a la même norme.

Si (x_i) est dense dans \mathbb{R} , $\sum_{i=1}^{\infty} \frac{1}{2^i} 1_{[-\infty, x_i]}(x) 1_{[-\infty, x_i]}(y)$ est un noyau dont le produit scalaire est plus fin sur $\mathcal{M}_{\mathbb{R}}^+$ que la topologie faible.

LEMME DES PRODUITS 3.6. — ψ étant un homéomorphisme de \mathbb{R} sur $\psi(\mathbb{R}) =]a\psi, b\psi[$ (avec $\max(|a\psi|, |b\psi|) < +\infty$), si (α_n) est une série strictement positive et si

$$K(x, y) = \sum_{n=0}^{\infty} \alpha_n (\psi(x))^n (\psi(y))^n \left(\text{resp. } K(x, y) = \sum_{n=0}^{\infty} \alpha_n e^{-n\psi(x)} e^{-n\psi(y)} \right)$$

est borné, $(\cdot, \cdot)_{\mathbb{K}^{\otimes m}}$ est un produit scalaire sur $\mathcal{M}_{\mathbb{R}^m}$ dont la topologie trace sur $\mathcal{M}_{\mathbb{R}^m}^+$ est la topologie faible.

Preuve. — ψ se prolonge de façon unique en une bijection $\bar{\psi}$ de $\bar{\mathbb{R}}$ sur $]a\psi, b\psi[$, et si l'on munit $\bar{\mathbb{R}}$ de la métrique $\bar{d}(x, y) = |\bar{\psi}(x) - \bar{\psi}(y)|$, $\bar{\psi}$ est une isométrie de $\bar{\mathbb{R}}$ sur $]a\psi, b\psi[$ et $(\bar{\mathbb{R}}, \bar{d})$ est un complété de (\mathbb{R}, d) . $((\bar{\psi}^{\otimes m})^n, n \in \mathbb{N})$ est alors (théorème de Stone-Weierstrass) totale dans $U_b(\bar{\mathbb{R}}^m, \bar{d}^{\otimes m})$, et de ce fait $((\psi^{\otimes m})^n, n \in \mathbb{N})$ est totale dans $U_b(\mathbb{R}^m, d^{\otimes m})$. On conclut alors comme dans le théorème de construction. ■

D'où

$$K_1^{\otimes m}((x_1, \dots, x_m), (y_1, \dots, y_m)) = \frac{1}{1 - \frac{1}{2} \frac{x_1 y_1}{(1 + |x_1|)(1 + |y_1|)}} \times \dots \times \frac{1}{1 - \frac{1}{2} \frac{x_m y_m}{(1 + |x_m|)(1 + |y_m|)}},$$

$$K_2^{\otimes m}((x_1, \dots, x_m), (y_1, \dots, y_m)) = e^{\frac{x_1 y_1}{(1 + |x_1|)(1 + |y_1|)}} \times \dots \times e^{\frac{x_m y_m}{(1 + |x_m|)(1 + |y_m|)}}, \dots$$

sont des noyaux reproduisant sur \mathbb{R}^m dont les produits scalaires associés $(\cdot, \cdot)_{\mathbb{K}_1^{\otimes m}}$ ont pour topologie trace sur $\mathcal{M}_{\mathbb{R}^m}^+$, la topologie faible.

Remarque. — Lorsque $\Omega = \mathbb{R}^+$ (resp. \mathbb{R}, \mathbb{R}^n) et lorsque l'on munit \mathcal{M}_Ω du produit de convolution usuel

$$(i. e. : \mu * \nu = (\mu \otimes \nu)_S \quad \text{où } S : (x, y) \rightarrow x + y);$$

il existe (voir [6], p. 82-86) des noyaux reproduisants K tels que :

- . $(\mathcal{M}_\Omega, *, \| \cdot \|_K)$ est une algèbre préhilbertienne.
- . La trace de $\| \cdot \|_K$ sur \mathcal{M}_Ω^+ est la topologie faible.

(bK_3 et cK_4 vérifient en particulier cette propriété — [b et c étant des constantes bien choisies]).

* Se prolonge alors naturellement à H_K et $(H_K, *, \| \cdot \|_{H_K})$ est alors une algèbre hilbertienne, on peut faire la même remarque pour $\Omega = [0, 1]$ lorsque l'on munit \mathcal{M}_Ω du produit de convolution $\mu \hat{*} \nu = (\mu \otimes \nu)_\pi$, où $\pi(x, y) = x \times y$.

Le lemme suivant nous donne l'avantage indéniable de pouvoir travailler aisément sur les espaces produits : il permet en particulier de construire en toute généralité des tests d'indépendance convergents, de régions critiques très simples ⁽⁶⁾.

LEMME DE MAJORATION DES NORMES PRODUITS 4.6. — *Soient (Ω, \mathcal{B}) un espace mesurable et K un noyau reproduisant $\mathcal{B} \otimes \mathcal{B}$ mesurable, uniformément borné par 1. Alors :*

- 1) $\| \mu_1 \otimes \dots \otimes \mu_n - \nu_1 \otimes \dots \otimes \nu_n \|_{\mathbb{K}^{\otimes n}}^2 \leq 2(\| \mu_1 - \nu_1 \|_K + \dots + \| \mu_n - \nu_n \|_K)$
- 2) $\| \mu_1 \otimes \dots \otimes \mu_n \|_{\mathbb{K}^{\otimes n}} \leq \| \mu_1 \|_K \times \dots \times \| \mu_n \|_K$.

Preuve. — 1) Pour $n = 2$ on a :

$$\begin{aligned} \| \mu_1 \otimes \mu_2 - \nu_1 \otimes \nu_2 \|_{\mathbb{K}^{\otimes 2}}^2 &= \int_{\Omega^2 \times \Omega^2} K(x_1, y_1) \times K(x_2, y_2) d(\mu_1 \otimes \mu_2 - \nu_1 \otimes \nu_2) \\ &\quad \otimes (\mu_1 \otimes \mu_2 - \nu_1 \otimes \nu_2)(x_1, x_2)(y_1, y_2) \\ &= (\mu_1, \mu_1)_K \times (\mu_2, \mu_2)_K - (\mu_1, \nu_1)_K \times (\mu_2, \nu_2)_K \\ &\quad + (\mu_1, \nu_1)_K \times (\nu_2, \nu_2)_K - (\mu_1, \nu_1)_K \times (\mu_2, \nu_2)_K \\ &\leq |(\mu_1, \mu_1)_K| |(\mu_2, \mu_2)_K - (\mu_2, \nu_2)_K| + |(\mu_2, \nu_2)_K| |(\mu_1, \mu_1)_K - (\mu_1, \nu_1)_K| \\ &\quad + |(\nu_1, \nu_1)_K| |(\nu_2, \nu_2)_K - (\mu_2, \nu_2)_K| + |(\mu_2, \nu_2)_K| |(\nu_1, \nu_1)_K - (\mu_1, \nu_1)_K| \\ &\leq |(\mu_2, \mu_2 - \nu_2)_K| + |(\mu_1, \mu_1 - \nu_1)_K| + |(\nu_2, \nu_2 - \mu_2)_K| + |(\nu_1, \mu_1 - \nu_1)_K| \\ &\leq 2(\| \mu_2 - \nu_2 \|_K + \| \mu_1 - \nu_1 \|_K). \end{aligned}$$

⁽⁶⁾ Cf. [6].

Pour $n = 3$ on a :

$$\begin{aligned} \|\mu_1 \otimes \mu_2 \otimes \mu_3 - \nu_1 \otimes \nu_2 \otimes \nu_3\|_{\mathbb{K}^{\otimes 3}}^2 &= (\mu_1, \mu_1)_{\mathbb{K}} \times (\mu_2, \mu_2)_{\mathbb{K}} \times (\mu_3, \mu_3)_{\mathbb{K}} \\ &\quad - (\mu_1, \nu_1)_{\mathbb{K}} \times (\mu_2, \nu_2)_{\mathbb{K}} (\mu_3, \nu_3) + (\nu_1, \nu_1)_{\mathbb{K}} + (\nu_2, \nu_2)_{\mathbb{K}} \times (\nu_3, \nu_3)_{\mathbb{K}} \\ &\quad - (\mu_1, \nu_1)_{\mathbb{K}} \times (\mu_2, \nu_2)_{\mathbb{K}} (\mu_3, \nu_3)_{\mathbb{K}} \\ &\leq |(\mu_1, \mu_1)_{\mathbb{K}} [(\mu_2, \mu_2)_{\mathbb{K}} \times (\mu_3, \mu_3)_{\mathbb{K}} - (\mu_2, \nu_2)_{\mathbb{K}} (\mu_3, \nu_3)_{\mathbb{K}}]| \\ &\quad + |(\mu_2, \nu_2)_{\mathbb{K}} \times (\mu_3, \nu_3)_{\mathbb{K}} (\mu_1, \nu_1)_{\mathbb{K}} - (\mu_1, \nu_1)_{\mathbb{K}}| + \dots \\ &\leq |(\mu_2, \mu_2)_{\mathbb{K}} \times (\mu_3, \mu_3)_{\mathbb{K}} - (\mu_2, \nu_2)_{\mathbb{K}} (\mu_3, \nu_3)_{\mathbb{K}}| + [(\mu_1, \mu_1 - \nu_1)_{\mathbb{K}}] + \dots \\ &\leq \|\mu_2 - \nu_2\|_{\mathbb{K}} + \|\mu_3 - \nu_3\|_{\mathbb{K}} + \|\mu_1 - \nu_1\|_{\mathbb{K}} + \dots \end{aligned}$$

(d'après la démonstration pour $n = 2$).

On conclut alors par récurrence en s'inspirant du passage de $n = 2$ à $n = 3$.

2) Immédiat. ■

Remarque. — Le théorème de Fubini généralisé nous permet de calculer très simplement la distance (associée à une norme produit du type ci-dessus), lorsque l'on a affaire à des observations dépendantes.

Le théorème cité dans le paragraphe ci-dessous est fondamental quant aux applications à la théorie de la Statistique.

7. THÉORÈME DE GLIVENKO-CANTELLI

DÉFINITION. — (Ω, \mathcal{B}) étant un espace mesurable et \mathbb{K} un noyau reproduisant défini sur Ω , nous dirons que \mathbb{K} est un *noyau normal* s'il existe une suite (f_i) de v. a. r. bornées par 1 et une série (α_i) convergente de termes positifs, telle que $\mathbb{K}(x, y) = \sum_{i=1}^{\infty} \alpha_i f_i(x) f_i(y)$.

. \mathbb{K} étant un noyau reproduisant normal, il existe une infinité de séries (α_i) et une infinité de suites (f_i) telles que $\mathbb{K} = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$; mais pour simplifier la rédaction, nous ne considérerons que la série (α_i) et la suite (f_i) apparaissant dans l'écriture $\mathbb{K} = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$.

Sous ces conditions, $\mathbb{K} = \sum_{i=1}^{\infty} \alpha_i f_i \otimes f_i$ étant un noyau normal et $a > 0$ étant donné, nous désignerons par $n_0(a)$ le plus petit entier tel que

$$\sum_{i=n_0(a)+1}^{\infty} \alpha_i < \frac{a^2}{2}$$

si les f_i sont positives, ou tel que $\sum_{i=n_0(a)+1} \alpha_i < \frac{a^2}{8}$ sinon. Sous ces conditions on démontre le :

THÉORÈME DE GLIVENKO-CANTELLI 1.7. — Soient $K = \sum \alpha_i f_i \otimes f_i$ un noyau normal sur E et (X_n) une suite de v. a. indépendantes définies sur (Ω, \mathcal{A}, P) telle que :

$$\forall n > 0, \forall i \in \mathbb{N} : \alpha_i \neq 0, \int f_i dP_{X_n} = \int f_i dP_{X_1}$$

On a alors :

$$\begin{aligned} \text{a) } P \left(\left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)} - P_{X_1} \right\|_K \geq a \right\} \right) &\leq 2n_0(a) \exp \left(- \frac{na^2}{4 \sum_{i=1}^{\infty} \alpha_i} \right) \\ \text{b) } P \left(\bigcup_{n \geq N_1} \left\{ \omega \in \Omega : \left\| \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)} - P_{X_1} \right\|_K \geq a \right\} \right) &\leq \frac{2n_0(a)}{1 - e^{-a^2/4\sum\alpha_i}} \exp \left(- \frac{a^2 N_1}{4\sum\alpha_i} \right). \end{aligned}$$

Pour certains noyaux, ce théorème est encore vrai lorsque (X_n) est une martingale.

Nous obtenons des résultats un peu moins forts lorsque les (X_i) sont fortement mélangées ou vérifient le principe d'invariance.

BIBLIOGRAPHIE

- [1] N. ARONSZAJN, La théorie des noyaux reproduisants et applications, dans *Proceedings of Cambridge philosophical Society*, vol. 39, 1943, p. 133-153.
- [2] N. ARONSZAJN, Theory of reproducing kernels, dans *Transactions of the American mathematical Society*, n° 68, 1950, p. 307-404.
- [3] N. BOURBAKI, *Espace vectoriel topologique*, Chap. III et IV (1^{re} édition).
- [4] D. BOSQ, *Sur l'estimation d'un paramètre à valeurs dans un espace de Hilbert*. Publ. Int. de l'U. E. R. de Math. Pures et Appliquées de Lille I, n° 70, janvier 1976.
- [5] M. DUC-JACQUET, *Approximation des fonctionnelles linéaires sur les espaces hilbertiens auto-reproduisants*. Thèse, Université Scientifique et Médicale de Grenoble, 1973.
- [6] C. GUILBART, *Étude des produits scalaires sur l'espace des mesures. Estimation par projections. Tests à noyaux*. Thèse de doctorat d'État soutenue à Lille I (le 11 octobre 1978).
- [6 bis] C. GUILBART, *Étude des produits scalaires sur l'espace des mesures et applications à l'estimation*. Publ. Int. de l'U. E. R. de Mathématiques Pures et Appliquées de Lille I, n° 139, 1978.

- [7] R. MOCHÉ, *Décantation et séparation asymptotiques uniformes...* Thèse d'État, Université des Sciences et Techniques de Lille I, soutenue le 10 janvier 1977.
- [8] MÜNTZ, *Ueber den Approximationssatz von Weierstrass*, *Mathematische Abhandlungen* (Schwarzes Festzchrift), Berlin, 1914, p. 302-312.
- [9] J. NEVEU, *Bases mathématiques du calcul des probabilités*, Masson, Paris, 1964.
- [10] J. NEVEU, *Processus aléatoires gaussiens*. Séminaire Math. Sup. (Les Presses de l'Université de Montréal).
- [11] K. R. PARTHASARATHY, *Probability measures on metric spaces*. Academic press, 1967.
- [12] L. SCHWARTZ, Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *J. Anal. Math. of Jerusalem*, XIII, 1964, p. 115-256.

(Manuscrit reçu le 9 octobre 1979).