MATHEMATICAL ANALYSIS AND APPLICATIONS, PART A
ADVANCES IN MATHEMATICS SUPPLEMENTARY STUDIES, VOL. 7A

Une étude des covariances mesurables

PIERRE CARTIER

I.H.E.S.
Bures sur Yvette, France

A LAURENT SCHWARTZ,
À QUI JE DOIS L'ESPOIR TÊTU
D'UN MONDE MEILLEUR, EN
FIDÈLE AMITIÉ

A stochastic process (X_t) admits of a *Karhunen representation* $X_t = \sum_{n=1}^{\infty} f_n(t).Z_n$ with measurable functions f_n and an orthogonal sequence of random variables Z_n if and only if it possesses a measurable version. Its covariance function is defined by $C(t,t') = E[X_t X_{t'}]$ and can be expanded into a series $C(t,t') = \sum_{n=1}^{\infty} f_n(t).f_n(t')$. In this chapter, we study in depth the class of functions of two variables which occur as such covariance functions. We characterize them by various inequalities, both pointwise and integral inequalities. We give a generalization of Mercer's theorem to integral equations with measurable kernels. We study the Hilbert spaces with measurable self-reproducing kernels. Special attention is to be paid to the occurrence of nonseparable Hilbert spaces. Null sets have to be controlled in a very precise way, the main technical difficulty lying in the fact that the diagonal of a square is a null set.

Contents

INTRODUCTION

On ne saurait surestimer l'importance des matrices hermitiennes, et en particulier de celles dont les valeurs propres sont positives. Dans l'étude des équations intégrales, les noyaux symétriques jouent un rôle prédominant,

ISBN 0-12-512801-0

comme l'a montré Schmidt [14] vers 1900. Soit donc K une fonction continue de deux variables réelles, définie dans le produit $I \times I$ où I est un intervalle compact de $\mathbf{R}$, et vérifiant la propriété de symétrie hermitienne $K(y, x) = \overline{K(x, y)}$. Le problème aux valeurs propres correspondant se formule par l'équation

$$\int_I K(x, y) f(y)\, dy = \lambda f(x), \qquad x \in I. \tag{1}$$

Il est assez remarquable que le fait que les valeurs propres soient positives se traduise par une condition ponctuelle: la matrice hermitienne d'éléments $K(x_i, x_j)$ a ses valeurs propres positives quels que soient les nombres distincts $x_1, \ldots, x_n$ dans l'intervalle I.

L'étude de l'équation intégrale (1) permet aussi d'introduire des développements de la forme

$$K(x, y) = \sum_n \lambda_n f_n(x) . \overline{f_n(y)}. \tag{2}$$

Un des résultats culminants de la théorie est le *théorème de Mercer*. Il affirme que lorsque l'équation intégrale (1) n'admet que des valeurs propres positives, on a un développement *uniformément convergent* de la forme (2), où les λ_n sont les valeurs propres, et les f_n les fonctions propres (elles sont continues) normalisées par $\int_I |f_n(t)|^2\, dt = 1$.

Considérons le cas particulier où I est l'intervalle $[0, 1]$ et K est de la forme $K(x, y) = k(x - y)$, la fonction k étant continue de période 1. Les fonctions propres sont les exponentielles $e^{2\pi i n x}$ (n entier de signe quelconque), et le développement (2) n'est autre que le développement en série de Fourier

$$k(t) = \sum_n \lambda_n e^{2\pi i n t}. \tag{3}$$

Par suite, les coefficients de Fourier λ_n sont positifs si et seulement si la matrice d'éléments $k(x_i - x_j)$ est hermitienne à valeurs propres positives quels que soient les nombres réels $x_1, \ldots, x_n$. La fonction k appartient donc à la classe des *fonctions de type positif*, introduite par Bochner vers 1930.

Dans un domaine tout différent, considérons un processus aléatoire (X_t), où la variable t est réelle. Pour simplifier, supposons le processus réel et centré, c'est-à-dire qu'on a $E[X_t] = 0$ pour tout t; supposons aussi que les variables X_t aient des moments du second ordre finis. La covariance du processus est alors définie par $C(t, t') = E[X_t X_{t'}]$. Les propriétés du processus qu'on appelle "du second ordre" sont celles qui ne dépendent que de la covariance, et qui ne font qu'exprimer la géométrie de la "courbe" formée par les points X_t de l'espace de Hilbert des variables aléatoires avec un moment du second ordre fini. Ce point de vue a été exploré vers 1945 par Loéve [11].

La covariance C est une fonction symétrique à valeurs réelles, et, pour chaque système fini $(t_1, \ldots, t_n)$ de valeurs du paramètre t, la matrice symétrique d'éléments $C(t_i, t_j)$ est à valeurs propres positives. Supposons que le processus soit continu en moyenne quadratique, c'est-à-dire que $E[(X_t - X_{t_0})^2]$ tende vers 0 lorsque t tend vers t_0. La covariance C est alors continue, et par application du théorème de Mercer, se développe sous la forme

$$C(t, t') = \sum_n \lambda_n f_n(t) . f_n(t'). \tag{4}$$

Le processus lui-même se représente sous la forme

$$X_t = \sum_n \lambda_n^{1/2} f_n(t) . Z_n, \tag{5}$$

où les variables aléatoires Z_n satisfont aux conditions

$$E[Z_n] = 0, \qquad E[Z_n^2] = 1, \qquad E[Z_m Z_n] = 0 \qquad \text{si} \quad m \neq n \tag{6}$$

("développement de Karhunen"). Lorsque le processus est stationnaire, et admet la période 1, on obtient un développement en série de Fourier aléatoire

$$X_t = \lambda_0^{1/2} Z_0 + 2 \sum_{n=1}^{\infty} \lambda_n^{1/2} [U_n . \cos(2\pi n t) + V_n . \sin(2\pi n t)]. \tag{7}$$

Ces résultats sont le point de départ de l'*analyse harmonique des processus aléatoires*, un sujet très vaste dont le pionnier fut N. Wiener.

Les trois points de vue précédents ont été très développés et généralisés; en particulier, les fonctions de type positif sont un outil fondamental dans l'étude des groupes localement compacts et de leurs représentations (voir, e.g., le livre de Weil [18]). Ce qu'ils ont en commun, c'est la considération d'une certaine classe de fonctions de deux variables, que l'on appelle en général "noyaux de type positif". Comme la terminologie n'est pas tout à fait fixée, et que nous avons en vue les applications probabilistes, nous avons préféré le terme plus bref de "covariance" pour désigner ces fonctions.

Jusqu'à présent, on s'est essentiellement limité aux covariances *continues*. J'y vois au moins deux raisons: en théorie des groupes, les fonctions de type positif considérées sont automatiquement continues, et l'étude des processus aléatoires très discontinus est relativement récente. Cependant, l'introduction de processus aléatoires mesurables conduit à des covariances mesurables. Leur théorie est pleine de pièges; en particulier, la diagonale d'un produit est de mesure nulle dans les cas usuels, et l'on doit donc manier des expressions telles que $C(t, t)$ avec beaucoup de prudence. J'ai fait une chasse attentive aux erreurs de ce genre, mais je ne puis garantir absolument que je les ai toutes dépistées.

Nous utiliserons quatre outils fondamentaux:

(a) *la représentation d'une covariance sous la forme* $C(t,t') = \langle \varphi(t) | \varphi(t') \rangle$, où φ est une application à valeurs dans un espace de Hilbert;

(b) *l'espace de Hilbert* $\mathscr{E}_C$ *de fonctions associé a une covariance*, dont elle est le "noyau reproduisant";

(c) *les opérateurs de Hilbert–Schmidt*;

(d) *les covariances séparables.*

La représentation signalée en (a) est bien naturelle du point de vue des processus. La notion de "noyau reproduisant" n'est pas neuve; elle a été introduite par Aronszajn [1] et appliquée par Bergmann aux fonctions holomorphes de plusieurs variables complexes. Son utilisation systématique dans l'étude des covariances et de leurs développements en série me semble plus originale. Il est inutile de souligner l'importance des opérateurs de Hilbert–Schmidt; dans notre exposé, ils interviennent essentiellement à cause du théorème suivant (démontré au n° 20):

Si $\mathscr{H}$ *est un espace de Hilbert, et* u *une application linéaire continue de* $\mathscr{H}$ *dans un espace* $L^2(T,\mu)$, *alors* u *est de Hilbert-Schmidt si et seulement s'il existe une fonction positive* h *dans* $L^2(T,\mu)$ *telle que l'on ait* $|u(a)| \leq h$ *pour tout vecteur* a *de norme* ≤ 1 *dans* $\mathscr{H}$.

En particulier, on déduit facilement de là une version mesurable du théorème de Mercer.

Une des plaies de la théorie est l'apparition d'espaces de Hilbert non séparables. La notion de covariance séparable est destinée à les éliminer, et il est heureux que la covariance d'un processus mesurable soit automatiquement séparable. Le contre-exemple universel est fourni par les processus monstrueux (X_t) tels que $E[X_t X_{t'}] = 0$ dès que t et t' sont distincts; de tels processus ne peuvent avoir aucune version mesurable.

Nous renvoyons au sommaire le lecteur qui veut se faire une idée rapide du contenu de ce travail, et nous nous contenterons donc d'une brève description. Les parties les plus importantes et les plus nouvelles sont les parties IV à VII. Les propriétés de base des covariances mesurables et séparables sont établies dans IV, en particulier le théorème qui implique que la covariance d'un processus mesurable est mesurable et séparable. La partie V contient les résultats les plus profonds, en particulier une version mesurable du théorème de Mercer, où les ensembles de mesure nulle ont été soigneusement circonscrits, et le théorème selon lequel toute covariance mesurable est égale presque partout à une covariance mesurable et séparable. Dans les parties VI et VII, le thème est la comparaison des inégalités ponctuelles et des inégalités intégrales, avec là aussi le désir de contrôler les ensembles de mesure nulle; une variante du théorème de Dunford-Pettis joue un rôle important.

Les parties I et II sont des préliminaires faciles, où l'on met au point l'outil, et la partie III est un résumé, presque sans démonstration, des propriétés des opérateurs dans un espace de Hilbert que nous utiliserons dans la suite. Enfin, la partie VIII est une série de commentaires et de remarques sur les applications des covariances. C'est ici qu'apparaissent le mieux les limitations de la théorie, en particulier par le fait que nous ne donnons aucun exemple explicite d'une covariance non continue; l'étude des covariances est le chapitre de l'*Analyse Hilbertienne* qui s'occupe des espaces de Hilbert de fonctions où la convergence en norme implique la convergence en tout point. Pour aller plus loin, il faut introduire des espaces de fonctions beaucoup plus discontinues, tels que les espaces de Sobolev.

Dans la théorie générale, un résultat important a été omis. C'est le théorème de Grothendieck [8] selon lequel toute covariance continue C définie sur un espace compact se développe sous la forme

$$C(t,t') = \sum_n \lambda_n f_n(t) . \overline{f_n(t')} \tag{8}$$

avec $|f_n(t)| \leq 1$ et $\sum_n |\lambda_n| \leq sh(\pi/2) \sup_{t,t'} |C(t,t')|$. Par différence avec les résultats précédents, on ne suppose pas les coefficients λ_n positifs. Je n'aurais guère pu que recopier un article ancien de moi-même [3], et pour les progrès récents dans cette direction, je préfère renvoyer aux travaux de Krivine [10].

Conventions

On note **R** le corps des nombres réels. Nous convenons que 0 est un nombre positif, en réservant le vocable "strictement positif" aux nombres positifs non nuls.

On note **C** le corps des nombres complexes. Si z est un nombre complexe, on note respectivement $\mathrm{Re}\, z$, $\mathrm{Im}\, z$ et $\bar{z}$ sa partie réelle, sa partie imaginaire et son conjugué.

Dans un espace de Hilbert, on note $\langle u|v\rangle$ le produit scalaire, et $\|u\| = \langle u|u\rangle^{1/2}$ la norme. On convient que $\langle u|v\rangle$ est fonction linéaire de la *deuxième* variable v.

On note $A\backslash B$ la différence des ensembles. Si μ est une mesure, on dit qu'un ensemble A est μ-négligeable si l'on a $\mu(\mathrm{A}) = 0$; une fonction f est dite μ-négligeable si l'ensemble $\{f \neq 0\}$ est μ-négligeable. La notation $L^p = L^p(T, \mathscr{A}, \mu)$ est la notation usuelle de la théorie de Lebesgue; en particulier, la norme dans L^p est définie par

$$\|f\|_p = \left\{ \int_T |f(t)|^p \, d\mu(t) \right\}^{1/p}$$

pour $p \geq 1$ fini. Dans l'espace de Hilbert L^2, le produit scalaire est noté $\langle f \,|\, g \rangle$ ou parfois $\langle f \,|\, g \rangle_2$; il est défini par la formule

$$\langle f \,|\, g \rangle = \int_T \overline{f(t)} g(t) \, d\mu(t).$$

Soient T_1 et T_2 deux ensembles, munis respectivement de tribus $\mathscr{A}_1$ et $\mathscr{A}_2$ de parties. On note $\mathscr{A}_1 \otimes \mathscr{A}_2$ la tribu de parties de $T_1 \times T_2$ engendrée par les ensembles de la forme $B_1 \times B_2$ avec $B_1 \in \mathscr{A}_1$ et $B_2 \in \mathscr{A}_2$. De même, si μ_i est une mesure positive sur la tribu $\mathscr{A}_i$, on note $\mu_1 \otimes \mu_2$ la mesure produit sur la tribu $\mathscr{A}_1 \otimes \mathscr{A}_2$; elle est caractérisée par la relation

$$(\mu_1 \otimes \mu_2)(B_1 \times B_2) = \mu_1(B_1) \, . \, \mu_2(B_2).$$

I. Propriétés générales des covariances

1. Soit $A = (a_{ij})$ une matrice à éléments complexes d'ordre n; supposons-la *hermitienne*, c'est-à-dire qu'on ait $\overline{a}_{ij} = a_{ji}$, et associons-lui la forme hermitienne

$$q(\xi) = \sum_{i,j=1}^{n} a_{ij} \overline{\xi}_i \xi_j, \tag{1}$$

où $\xi = (\xi_1, \ldots, \xi_n)$ parcourt l'espace $\mathbf{C}^n$ des vecteurs à n coordonnées complexes. On dit classiquement que la matrice A est *hermitienne positive* si l'on a $q(\xi) \geq 0$ pour tout vecteur ξ dans $\mathbf{C}^n$.

Soient $\lambda_1, \ldots, \lambda_n$ les valeurs propres de A, éventuellement répétées selon leur multiplicité. Elles sont réelles, et il existe une matrice unitaire $U = (u_{ij})$ telle que l'on ait $A = U^{-1} \Lambda U$, où Λ est la matrice diagonale d'éléments $\lambda_1, \ldots, \lambda_n$. On a

$$a_{ij} = \sum_{k=1}^{n} \lambda_k \overline{u}_{ki} u_{kj}, \tag{2}$$

et donc

$$q(\xi) = \sum_{k=1}^{n} \lambda_k \left| \sum_{i=1}^{n} u_{ki} \xi_i \right|^2 . \tag{3}$$

Comme la matrice U est inversible, la matrice A est hermitienne positive si et seulement si l'on a $\sum_{k=1}^{n} \lambda_k |\eta_k|^2 \geq 0$ quels que soient les nombres complexes $\eta_1, \ldots, \eta_n$. Autrement dit, *la matrice hermitienne A est positive si et seulement si ses valeurs propres $\lambda_1, \ldots, \lambda_n$ sont positives*; cette propriété sert parfois de définition.

2. Soient T un ensemble et C une fonction à valeurs complexes sur $T \times T$. On dira que C est une *covariance* sur T si, quels que soient les points $t_1, \ldots, t_n$ distincts de T, la matrice d'éléments $C(t_i, t_j)$ est hermitienne positive. Autrement dit, on impose les deux conditions suivantes:

(a) *on a* $C(t', t) = \overline{C(t, t')}$ *quels que soient* t, t' *dans* T;
(b) *quels que soient les points distincts* $t_1, \ldots, t_n$ *de* T, *et les nombres complexes* $\xi_1, \ldots, \xi_n$, *on a* $\sum_{i,j=1}^n C(t_i, t_j)\bar{\xi}_i \xi_j \geq 0$.

En fait, l'hypothèse (a) est conséquence de (b). En effet, le cas particulier $n = 1$, $t_1 = t$ et $\xi_1 = 1$ de l'hypothèse (b) s'écrit

$$C(t, t) \geq 0 \qquad \text{pour tout } t \text{ dans } T. \tag{4}$$

Considérons maintenant le cas $n = 2$ et faisons successivement (ξ_1, ξ_2) égal à $(1, 1)$ et à $(1, i)$; on en déduit que les nombres

$$C(t_1, t_1) + C(t_2, t_2) + C(t_1, t_2) + C(t_2, t_1)$$
$$C(t_1, t_1) + C(t_2, t_2) + i(C(t_1, t_2) - C(t_2, t_1))$$

sont réels, et comme les nombres $C(t_1, t_1)$ et $C(t_2, t_2)$ sont aussi réels, on conclut que $C(t_2, t_1)$ est le conjugué de $C(t_1, t_2)$.

Utilisons le critère classique qui caractérise les matrices hermitiennes positives par la positivité des mineurs principaux (voir Gantmacher [6]); on voit que les covariances sont les fonctions C satisfaisant à la condition (a) ci-dessus et à la suivante:

(b′) *quels que soient l'entier* $n \geq 1$ *et les points* $t_1, \ldots, t_n$ *de* T, *le déterminant*

$$D_n(t_1, \ldots, t_n) = \det(C(t_i, t_j))_{1 \leq i, j \leq n} \tag{5}$$

est positif.

Le cas $n = 1$ redonne l'inégalité (4); le cas $n = 2$ est particulièrement important:

$$|C(t, t')|^2 \leq C(t, t) . C(t', t'). \tag{6}$$

3. Soit E un espace vectoriel complexe. On appelle d'habitude *forme sesquilinéaire* sur E toute application Γ de $E \times E$ dans $\mathbf{C}$ telle que $\Gamma(f, g)$ et $\overline{\Gamma(g, f)}$ soient des fonctions linéaires de g pour tout f fixé. On dit que Γ est *hermitienne* si l'on a

$$\Gamma(g, f) = \overline{\Gamma(f, g)} \qquad \text{pour } f, g \text{ dans } E. \tag{7}$$

On a la *formule de polarisation*

$$3\Gamma(f, g) = \sum_j j^{-1} \Gamma(f + jg, f + jg), \tag{8}$$

où j parcourt les solutions de l'équation $j^3 = 1$; par suite, Γ est hermitienne si et seulement si $\Gamma(f,f)$ est réel pour tout f dans E. On dit que Γ est *positive* si $\Gamma(f,f)$ est positif pour tout f dans E; une forme sesquilinéaire positive est donc hermitienne.

Soit Γ une forme sesquilinéaire positive. On a l'inégalité de Cauchy–Schwarz

$$|\Gamma(f,g)|^2 \leq \Gamma(f,f) . \Gamma(g,g) \qquad \text{pour } f, g \text{ dans } E. \tag{9}$$

Par suite, un élément f de E satisfait à la relation $\Gamma(f,f) = 0$ si et seulement si l'on a $\Gamma(f,g) = 0$ pour tout g dans E. L'ensemble N des éléments f de E tels que $\Gamma(f,f) = 0$ est donc un sous-espace vectoriel de E, et il existe une forme sesquilinéaire positive Γ_0 sur l'espace quotient $E_0 = E/N$ caractérisée par la relation

$$\Gamma(f,g) = \Gamma_0(f + N, g + N) \qquad \text{pour } f, g \text{ dans } E. \tag{10}$$

Par construction, on a $\Gamma_0(u,u) > 0$ pour tout u non nul dans E_0. Le procédé de complétion fournit un espace de Hilbert $\mathscr{H}$ contenant E_0 comme sous-espace vectoriel dense, tel que $\Gamma_0(u,v) = \langle u|v\rangle$ pour u, v dans E_0; notons π l'application linéaire de E dans $\mathscr{H}$ qui associe à f la classe $f + N$. En résumé, on a établi le résultat suivant:

LEMME. *Soient E un espace vectoriel complexe et Γ une forme sesquilinéaire positive sur E. Il existe un espace de Hilbert $\mathscr{H}$ et une application linéaire π de E dans $\mathscr{H}$ satisfaisant aux conditions suivantes*:

(a) *l'image $\pi(E)$ de π est dense dans $\mathscr{H}$;*
(b) *on a $\Gamma(f,g) = \langle\pi(f)|\pi(g)\rangle$ pour f, g dans E.*

4. Nous pouvons faire le lien entre les covariances et les espaces de Hilbert.

LEMME. *Soit C une fonction à valeurs complexes sur $T \times T$. Pour que C soit une covariance, il faut et il suffit qu'il existe un espace de Hilbert $\mathscr{H}$ et une application φ de T dans $\mathscr{H}$ tels que l'on ait*

$$C(t,t') = \langle\varphi(t)|\varphi(t')\rangle \qquad \textit{pour } t, t' \textit{ dans } T. \tag{11}$$

Si C satisfait à la relation précédente, on a

$$\sum_{i,j=1}^{n} C(t_i,t_j)\bar{\xi}_i\xi_j = \left\|\sum_{i=1}^{n} \xi_i\varphi(t_i)\right\|^2 \tag{12}$$

pour $t_1, \ldots, t_n$ dans T et $\xi_1, \ldots, \xi_n$ dans $\mathbf{C}$; donc C est une covariance.

Réciproquement, supposons que C soit une covariance. Soit E l'espace vectoriel des fonctions $f : T \to \mathbf{C}$ telles que l'ensemble $\{f \neq 0\}$ soit fini. On définit une forme sesquilinéaire Γ sur E par la formule

$$\Gamma(f, g) = \sum_{t, t' \in T} C(t, t')\overline{f(t)}g(t') \tag{13}$$

et par définition d'une covariance, on a $\Gamma(f, f) \geq 0$ pour tout f dans E. Appliquons le lemme du n° 3. Pour tout point t de T, soit ε_t la fonction égale à 1 en t et nulle ailleurs; si l'on pose $\varphi(t) = \pi(\varepsilon_t)$, on a

$$\langle \varphi(t) | \varphi(t') \rangle = \Gamma(\varepsilon_t, \varepsilon_{t'}) = C(t, t') \qquad \text{pour } t, t' \text{ dans } T.$$

5. Soit C une covariance sur T. Choisissons $\mathscr{H}$ et φ de manière à satisfaire à la relation (11), ce que nous exprimerons en disant que la covariance C est *associée* à $\mathscr{H}$ et φ. Soit $\mathscr{F}$ le plus petit sous-espace vectoriel fermé de $\mathscr{H}$ contenant l'image de φ. A tout élément u de $\mathscr{F}$, on associe la fonction Hu sur T par la formule

$$Hu(t) = \langle \varphi(t) | u \rangle. \tag{14}$$

Si Hu est nulle, le vecteur u de $\mathscr{F}$ est orthogonal à l'image de φ, donc à $\mathscr{F}$ tout entier, et l'on a $u = 0$. Par suite, l'application $u \mapsto Hu$ est un isomorphisme de $\mathscr{F}$ sur un espace vectoriel $\mathscr{E}_C$ de fonctions à valeurs complexes sur T. Au moyen de cet isomorphisme, on transporte le produit scalaire de $\mathscr{F}$ en un produit scalaire sur $\mathscr{E}_C$, noté $\langle f | g \rangle_C$; on pose $\|f\|_C = \langle f | f \rangle_C^{1/2}$. Par définition, on a donc

$$\langle Hu | Hv \rangle_C = \langle u | v \rangle \qquad \text{pour } u, v \text{ dans } \mathscr{F}. \tag{15}$$

Le lemme suivant montre que l'espace de Hilbert $\mathscr{E}_C$ ne dépend que de la covariance C, mais non des données auxiliaires $\mathscr{H}$ et φ; on dit parfois que $\mathscr{E}_C$ est *l'espace de noyau reproduisant C.*

LEMME. *Soient C une covariance sur T, f une application de T dans* $\mathbf{C}$ *et M une constante positive. Pour que f appartienne à $\mathscr{E}_C$ et satisfasse à $\|f\|_C \leq M$, il faut et il suffit que la fonction D_M définie par*

$$D_M(t, t') = M^2 . C(t, t') - f(t) . \overline{f(t')} \tag{16}$$

soit une covariance sur T.

Par construction, pour que f appartienne à $\mathscr{E}_C$ et satisfasse à $\|f\|_C \leq M$, il faut et il suffit qu'il existe un élément u de norme $\leq M$ dans $\mathscr{F}$ tel que $\langle u | \varphi(t) \rangle = \overline{f(t)}$ pour tout $t \in T$. Comme un espace de Hilbert est son propre dual, ceci signifie qu'il existe sur $\mathscr{F}$ une forme linéaire de norme $\leq M$ qui prenne la valeur $\overline{f(t)}$ en l'élément $\varphi(t)$ pour tout t dans T. D'après un critère élémentaire et classique, cela signifie que l'on a

$$\left|\sum_{i=1}^{n} \xi_i \overline{f(t_i)}\right| \leq M.\left\|\sum_{i=1}^{n} \xi_i \varphi(t_i)\right\| \tag{17}$$

quels que soient $t_1, \ldots, t_n$ dans T, et $\xi_1, \ldots, \xi_n$ dans $\mathbf{C}$. Le lemme résulte alors de la relation évidente

$$\sum_{i,j=1}^{n} D_M(t_i, t_j)\bar{\xi}_i \xi_j = M^2.\left\|\sum_{i=1}^{n} \xi_i \varphi(t_i)\right\|^2 - \left|\sum_{i=1}^{n} \xi_i \overline{f(t_i)}\right|^2. \tag{18}$$

6. Le problème se pose maintenant de caractériser les espaces de Hilbert de fonctions admettant un noyau reproduisant.

Soit d'abord C une covariance. Je dis que la fonction $C_t : t' \mapsto C(t', t)$ appartient à $\mathscr{E}_C$, et qu'elle satisfait aux relations

$$f(t) = \langle C_t | f \rangle_C \qquad \text{pour} \quad f \in \mathscr{E}_C \quad \text{et} \quad t \in T, \tag{19}$$

$$C(t, t') = \langle C_t | C_{t'} \rangle_C \qquad \text{pour} \quad t, t' \text{ dans } T. \tag{20}$$

En effet, les relations (11) et (14) montrent que l'on a $C_t = Hu$ avec $u = \varphi(t)$, d'où $C_t \in \mathscr{E}_C$; par ailleurs, si $f \in \mathscr{E}_C$ est de la forme Hu avec $u \in \mathscr{F}$, on a

$$f(t) = \langle \varphi(t) | u \rangle = \langle H\varphi(t) | Hu \rangle_C = \langle C_t | f \rangle_C$$

d'après la définition (14) de Hu et la définition (15) du produit scalaire dans $\mathscr{E}_C$; enfin, d'après la définition même de C_t, la formule (20) est le cas particulier $f = C_{t'}$ de la formule (19).

On remarquera que la formule (19) caractérise l'élément C_t de $\mathscr{E}_C$; de plus, si un élément de $\mathscr{E}_C$ est orthogonal à tous les C_t, il est nul d'après (19), et par conséquent, l'ensemble des combinaisons linéaires de la forme $\xi_1 . C_{t_1} + \cdots + \xi_n . C_{t_n}$ est dense dans $\mathscr{E}_C$.

LEMME. *Soit $\mathscr{E}$ un espace vectoriel de fonctions complexes sur T, muni d'un produit scalaire $\langle f | g \rangle$ qui en fasse un espace de Hilbert. Les conditions suivantes sont équivalentes*:

(a) *il existe une covariance C sur T telle que $\mathscr{E} = \mathscr{E}_C$ et $\langle f | g \rangle = \langle f | g \rangle_C$ pour f et g dans $\mathscr{E}$*;

(b) *il existe une fonction h sur T à valeurs positives finies, telle que l'on ait*

$$|f(t)| \leq h(t) \tag{21}$$

pour tout point t de T et tout élément f de norme ≤ 1 dans $\mathscr{E}$.

De plus, s'il existe une covariance C comme dans (a), *elle est unique.*

Les formules (19) et (20) fournissent une caractérisation de la covariance C en seuls termes de l'espace $\mathscr{E}_C$ et de son produit scalaire $\langle f \mid g \rangle_C$, d'où l'unicité. Par ailleurs, ces mêmes formules et l'inégalité de Cauchy–Schwarz entraînent

$$|f(t)| \leq \|f\|_C . C(t,t)^{1/2} \qquad \text{pour} \quad f \in \mathscr{E}_C \quad \text{et} \quad t \in T, \tag{22}$$

avec égalité seulement si f est proportionnelle à C_t. Par suite, (a) entraîne (b).

Réciproquement, sous l'hypothèse (b), l'application $f \mapsto f(t)$ est une forme linéaire continue sur l'espace de Hilbert $\mathscr{E}$ pour tout $t \in T$. Il existe donc un élément $\varphi(t)$ de $\mathscr{E}$ tel que l'on ait $f(t) = \langle \varphi(t) \mid f \rangle$ pour $f \in \mathscr{E}$ et $t \in T$; soit C la covariance associée à l'application $t \mapsto \varphi(t)$ de T dans $\mathscr{E}$. Reprenons les notations du n° 5 en y faisant $\mathscr{H} = \mathscr{E}$; si un élément f de $\mathscr{E}$ est orthogonal à l'image de φ, on a $f = 0$ d'après la définition des $\varphi(t)$; on a donc $\mathscr{F} = \mathscr{E}$. D'après la formule (14), on a $Hu = u$ pour tout $u \in \mathscr{E}$, d'où $\mathscr{E} = \mathscr{E}_C$ et $\langle f \mid g \rangle = \langle f \mid g \rangle_C$ pour f, g dans $\mathscr{E}$. Par suite, (b) entraîne (a).

7. Considérons toujours l'ensemble T et introduisons un espace de Hilbert $\mathscr{H}$ muni d'une base orthonormale $(e_\alpha)_{\alpha \in A}$. Se donner une application φ de T dans $\mathscr{H}$ revient à se donner une famille $(f_\alpha)_{\alpha \in A}$ d'applications de T dans $\mathbf{C}$ telle que l'on ait

$$\sum_{\alpha \in A} |f_\alpha(t)|^2 < +\infty \qquad \text{pour tout } t \text{ dans } T. \tag{23}$$

La correspondance entre φ et les f_α s'exprime par les formules

$$f_\alpha(t) = \langle \varphi(t) \mid e_\alpha \rangle, \tag{24}$$

$$\varphi(t) = \sum_{\alpha \in A} \overline{f_\alpha(t)} . e_\alpha . \tag{25}$$

S'il en est ainsi, la covariance sur T associée à $\mathscr{H}$ et φ est définie par la formule

$$C(t,t') = \sum_{\alpha \in A} f_\alpha(t) . \overline{f_\alpha(t')}. \tag{26}$$

De plus, pour tout α dans A, la fonction D_1 définie par

$$D_1(t,t') = C(t,t') - f_\alpha(t) . \overline{f_\alpha(t')}$$

est égale à $\sum_{\beta \neq \alpha} f_\beta(t) . \overline{f_\beta(t')}$, donc est encore une covariance. Compte tenu du lemme du n° 5, on a donc établi le résultat suivant:

LEMME. *Si la famille de fonctions* $(f_\alpha)_{\alpha \in A}$ *sur* T *satisfait à la condition* (23), *la somme dans* (26) *converge absolument dans* $T \times T$, *et définit une covariance*

C sur T. Pour tout $\alpha \in A$, la fonction f_α est un élément de norme ≤ 1 de l'espace de Hilbert $\mathscr{E}_C$.

Particularisons au cas où $\mathscr{H} = \mathscr{E}$ est un espace de fonctions sur T satisfaisant à la condition (b) du lemme du n° 6. On a vu qu'il existait une covariance C sur T telle que $\mathscr{E}$ soit l'espace de noyau reproduisant C. Alors, pour toute base orthonormale $(f_\alpha)_{\alpha \in A}$ de $\mathscr{E}$, on a $\sum_{\alpha \in A} |f_\alpha(t)|^2 < +\infty$ pour tout point t de T, et la covariance C est donnée explicitement par la formule (26).

8. Illustrons ce qui précède par l'exemple classique des polynômes orthogonaux. Soit μ une mesure positive sur la droite réelle $\mathbf{R}$ telle que l'intégrale $\int_{\mathbf{R}} |x|^n \, d\mu(x)$ soit finie pour tout entier $n \geq 0$. Soit $(p_n)_{n \geq 0}$ une famille de polynômes orthogonaux correspondant à μ; on a donc

$$\int_{\mathbf{R}} p_m(x) . p_n(x) \, d\mu(x) = 0 \qquad \text{si} \quad m \neq n, \tag{27}$$

et p_n est un polynôme à coefficients réels de degré égal à n.

Fixons un entier $n \geq 0$, et supposons que la mesure μ ne soit pas localisée sur un ensemble fini ayant au plus n points. Soit $\mathscr{E}$ l'espace vectoriel formé des polynômes à coefficients complexes, de degré au plus égal à n. On définit un produit scalaire sur $\mathscr{E}$ par la formule

$$\langle p | q \rangle = \int_{\mathbf{R}} \overline{p(x)} . q(x) \, d\mu(x). \tag{28}$$

Alors $\mathscr{E}$ est un espace de Hilbert de dimension finie $n + 1$, admettant une base orthogonale $(p_0, p_1, \ldots, p_n)$.

Avec les notations classiques, on pose

$$h_\nu = \int_{\mathbf{R}} p_\nu(x)^2 \, d\mu(x) \qquad \text{pour} \quad 0 \leq \nu \leq n, \tag{29}$$

$$K_n(x, y) = \sum_{\nu=0}^{n} h_\nu^{-1} p_\nu(x) . \overline{p_\nu(y)} \qquad \text{pour} \quad x, y \text{ dans } \mathbf{C}. \tag{30}$$

Rappelons aussi que, si k_n est le coefficient de x^n dans $p_n(x)$ et k_{n+1} celui de x^{n+1} dans $p_{n+1}(x)$, on a la formule de Christoffel–Darboux:

$$K_n(x, y) = (k_n/k_{n+1}h_n) . (p_{n+1}(x)\overline{p_n(y)} - p_n(x)\overline{p_{n+1}(y)})/(x - y) \tag{31}$$

où x et y sont deux nombres complexes distincts. Enfin, on a

$$p(x) = \int_{\mathbf{R}} K_n(x, y) . p(y) \, d\mu(y) \tag{32}$$

pour tout nombre complexe x et tout polynôme p de degré $\leq n$.

Interprétons les polynômes comme des applications de $\mathbf{C}$ dans $\mathbf{C}$. Ce qui précède montre que K_n est une covariance sur $\mathbf{C}$, et que $\mathscr{E}$ est l'espace de

Hilbert associé à cette covariance. Soit x_0 un nombre complexe. D'après la formule (22), on a

$$|p(x_0)| \leq K_n(x_0, x_0)^{1/2} \tag{33}$$

pour tout polynôme p de degré $\leq n$ tel que $\int_{\mathbf{R}} |p(x)|^2 \, d\mu(x) \leq 1$, avec égalité si et seulement si l'on a

$$p(x) = u.[K_n(x_0, x_0)]^{-1/2} K_n(x, x_0) \tag{34}$$

avec une constante u de module 1 convenable (voir Szegö [17] pour tout ce n°).

II. Covariances continues

Dans toute cette partie, on note T un espace topologique.

9. Soient φ une application de T dans un espace de Hilbert $\mathscr{H}$, et C la covariance associée. On a donc par définition

$$C(t, t') = \langle \varphi(t) | \varphi(t') \rangle \qquad \text{pour } t, t' \text{ dans } T. \tag{1}$$

On en déduit

$$\|\varphi(t) - \varphi(t')\| = d_C(t, t') \tag{2}$$

où le nombre réel positif $d_C(t, t')$ ne dépend que de la covariance C, puisque l'on a

$$d_C(t, t')^2 = C(t, t) + C(t', t') - 2 \operatorname{Re} C(t, t'). \tag{3}$$

Lemme. *L'application φ de T dans $\mathscr{H}$ est continue si et seulement si la fonction C est continue sur $T \times T$.*

Le produit scalaire est une application continue de $\mathscr{H} \times \mathscr{H}$ dans $\mathbf{C}$, donc la continuité de φ entraîne celle de C. Réciproquement, supposons que C soit continue; d'après la formule (3), d_C est une fonction continue sur $T \times T$. Donnons-nous un point t_0 de T et un nombre réel $\varepsilon > 0$. Vu la continuité de d_C, il existe un voisinage U de t_0 dans T tel que l'on ait $d_C(t, t_0) < \varepsilon$ pour tout $t \in U$. On a donc $\|\varphi(t) - \varphi(t_0)\| < \varepsilon$ pour tout t dans U, d'où la continuité de φ.

10. Considérons maintenant une covariance C sur T, et appliquons ce qui précède à l'espace de Hilbert $\mathscr{H} = \mathscr{E}_C$ et à $\varphi(t) = C_t$. On a donc

$$d_C(t, t') = \|C_t - C_{t'}\|_C, \tag{4}$$

d'où les relations

$$d_C(t,t) = 0, \tag{5}$$

$$d_C(t,t') = d_C(t',t), \tag{6}$$

$$d_C(t,t'') \leq d_C(t,t') + d_C(t',t'') \tag{7}$$

(pour t, t' et t'' dans T). Il s'en faut de peu que d_C ne soit une distance sur T; simplement, on peut avoir $d_C(t,t') = 0$ même lorsque t et t' sont distincts. Pour tout élément f de $\mathscr{E}_C$, on a $f(t) - f(t') = \langle C_t - C_{t'} | f \rangle_C$ et l'inégalité

$$|f(t) - f(t')| \leq \|f\|_C . d_C(t,t') \tag{8}$$

s'obtient par application de l'inégalité de Cauchy–Schwarz.

11. Lemme. *Pour que la covariance C soit continue sur $T \times T$, il faut et il suffit que $\mathscr{E}_C$ se compose de fonctions continues, et que toute partie bornée de $\mathscr{E}_C$ soit équicontinue.*

Supposons d'abord que C soit continue sur $T \times T$. Alors la fonction d_C est continue sur $T \times T$, et l'on a $d_C(t,t) = 0$ et $d_C(t,t') \geq 0$. L'inégalité (8) fournit un module de continuité pour les fonctions de $\mathscr{E}_C$, uniforme sur toute partie de $\mathscr{E}_C$ bornée en norme.

Réciproquement, supposons que l'ensemble B des fonctions f de $\mathscr{E}_C$ telles que $\|f\|_C \leq 1$ soit équicontinu, et montrons que l'application $t \mapsto C_t$ de T dans $\mathscr{E}_C$ est continue. Soient $t_0 \in T$ et $\varepsilon > 0$; par hypothèse, il existe un voisinage U de t_0 dans T tel que l'on ait $|f(t) - f(t_0)| \leq \varepsilon$ pour $t \in U$ et $f \in B$. On a alors

$$\|C_t - C_{t_0}\| = \sup_{f \in B} |\langle C_t - C_{t_0} | f \rangle| = \sup_{f \in B} |f(t) - f(t_0)| \leq \varepsilon$$

pour tout t dans U, d'où notre assertion. Comme on a $C(t,t') = \langle C_t | C_{t'} \rangle_C$, la continuité de l'application $t \mapsto C_t$ entraîne celle de C d'après le lemme du n° 9.

12. Gardons les notations précédentes et supposons C continue. D'après le lemme du n° 5, une fonction f sur T appartient à B si et seulement l'on a les inégalités

$$\left| \sum_{i=1}^{n} \xi_i \overline{f(t_i)} \right|^2 \leq \sum_{i,j=1}^{n} C(t_i, t_j) \bar{\xi}_i \xi_j \tag{9}$$

où $t_1, \ldots, t_n$ sont des points de T et $\xi_1, \ldots, \xi_n$ des nombres complexes. Par suite, B est fermé pour la topologie de la convergence simple dans T. De plus, on a $|f(t)| \leq C(t,t)^{1/2}$ pour $f \in B$ et $t \in T$, et la fonction $t \mapsto C(t,t)^{1/2}$ est continue, donc bornée sur toute partie compacte de T. Enfin, le lemme précédent montre que B est un ensemble équicontinu de fonctions. Par application du théorème d'Ascoli, on obtient donc le résultat suivant:

LEMME. *Supposons la covariance C continue sur $T \times T$. Alors l'ensemble B des fonctions f de $\mathscr{E}_C$ telles que $\|f\|_C \leq 1$ est compact pour la topologie de la convergence compacte sur T. De plus, la convergence en norme dans $\mathscr{E}_C$ est plus forte que la convergence compacte sur T.*

13. Venons-en aux propriétés de développement en série des covariances continues.

THÉORÈME. *Supposons que l'espace T soit compact et que C soit une covariance continue sur T. Soit $(f_\alpha)_{\alpha \in A}$ une famille de fonctions sur T, telle que l'on ait*

$$C(t, t') = \sum_{\alpha \in A} f_\alpha(t) . \overline{f_\alpha(t')}, \tag{10}$$

la série convergeant absolument sur $T \times T$. Alors:

(a) *Les fonctions f_α sont continues sur T.*
(b) *L'ensemble des indices $\alpha \in A$ tels que $f_\alpha \neq 0$ est dénombrable.*
(c) *La série* (10) *converge uniformément dans $T \times T$.*

D'après le lemme du n° 7, chaque fonction f_α appartient à $\mathscr{E}_C$, donc est continue d'après le lemme du n° 11. Ceci prouve (a).

Comme la fonction d_C est continue, et que T est compact, il existe pour tout entier $n \geq 1$ une partie finie D_n de T telle que, pour tout $t \in T$, il existe $s \in D_n$ avec $d_C(t, s) \leq 2^{-n}$. Compte tenu de l'inégalité (8), toute fonction appartenant à $\mathscr{E}_C$ qui est nulle en tout point de l'ensemble dénombrable $D = \bigcup_{n=1}^{\infty} D_n$ est nulle. Pour chaque point s de D, on a $\sum_{\alpha \in A} |f_\alpha(s)|^2 < +\infty$, et par conséquent, il existe une partie dénombrable $A(s)$ de A telle que l'on ait $f_\alpha(s) = 0$ pour α dans $A \backslash A(s)$. L'ensemble $B = \bigcup_{s \in D} A(s)$ est une partie dénombrable de A, et l'on a $f_\alpha(s) = 0$ pour $s \in D$ et $\alpha \in A \backslash B$. On a finalement $f_\alpha = 0$ pour tout indice α dans $A \backslash B$, d'où (b).

Pour prouver (c), on peut donc supposer que A est l'ensemble des entiers $\alpha \geq 1$. Chacune des fonctions f_α est continue d'après (a), et il en est de même de $\sum_{\alpha \in A} |f_\alpha|^2$ puisqu'on a $C(t, t) = \sum_{\alpha \in A} |f_\alpha(t)|^2$ et que la fonction C est continue. D'après le lemme de Dini, la série $\sum_{\alpha \in A} |f_\alpha|^2$ converge uniformément sur T. Soit $\varepsilon > 0$; il existe donc un entier $n \geq 1$ tel que l'on ait $\sum_{\alpha \geq n} |f_\alpha(t)|^2 \leq \varepsilon$ pour tout $t \in T$. D'après l'inégalité de Cauchy–Schwarz, on a donc

$$\left| \sum_{\alpha \geq n} f_\alpha(t) . \overline{f_\alpha(t')} \right|^2 \leq \sum_{\alpha \geq n} |f_\alpha(t)|^2 . \sum_{\alpha \geq n} |f_\alpha(t')|^2 \leq \varepsilon^2$$

quels que soient t et t' dans T. Ceci prouve (c).

14. Nous allons montrer par un exemple comment la continuité de la covariance et l'uniformité de la convergence de la série (10) sont étroitement liées. Nous allons en effet construire une covariance C sur l'intervalle $[0,1]$ de $\mathbf{R}$ qui a les deux propriétés suivantes:

(a) *la fonction C est continue en tout point du carré $[0,1] \times [0,1]$, à l'exception du point* $(0,0)$;
(b) *toute fonction appartenant à $\mathscr{E}_C$ est continue sur l'intervalle* $[0,1]$.

Il n'existe alors aucun développement de la forme $C(t,t') = \sum_{n=1}^{\infty} f_n(t)\overline{f_n(t')}$ qui converge uniformément sur $[0,1] \times [0,1]$. En effet, s'il en était ainsi, chaque fonction f_n appartiendrait à $\mathscr{E}_C$(n° 7), donc serait continue [propriété (b)] et la fonction C serait continue, contrairement à la propriété (a).

Soit $\mathscr{H}$ un espace de Hilbert, muni d'une base orthonormale $(e_n)_{n\geq 0}$. On considère l'application h de $[0,1]$ dans $\mathscr{H}$ admettant les valeurs suivantes:

$$h(0) = 0, \qquad h(2^{-2n}) = 0, \qquad h(2^{-2n-1}) = e_n \qquad \text{pour tout} \quad n \geq 0 \quad (11)$$

et qui est linéaire dans chacun des intervalles $[2^{-m-1}, 2^{-m}]$ (pour $m \geq 0$). La covariance C est la covariance associée à h.

Il est immédiat que l'application h est continue en tout point non nul de $[0,1]$, donc C est continue en tout point (t,t') de $[0,1] \times [0,1]$ pour lequel t et t' sont non nuls. Soit t un point non nul de l'intervalle $[0,1]$, et soit n un entier positif tel que $2^{-2n} < t$; alors la fonction C est identiquement nulle dans le voisinage $[2^{-2n}, 1] \times [0, 2^{-2n}]$ du point $(t,0)$, donc la fonction C est continue en ce point. Pour une raison analogue, la fonction C est continue en tout point $(0,t)$ avec $t \neq 0$. Enfin, on a $C(0,0) = 0$ et $C(2^{-2n-1}, 2^{-2n-1}) = 1$ pour tout entier $n \geq 0$, donc la fonction C n'est pas continue au point $(0,0)$. Ceci établit la propriété (a).

Il est clair que $\mathscr{H}$ est le seul sous-espace vectoriel fermé de lui-même contenant l'image de h. La formule $Hu(t) = \langle h(t) | u \rangle$ définit donc un isomorphisme $u \mapsto Hu$ de $\mathscr{H}$ sur $\mathscr{E}_C$. Il résulte de la formule (11) que l'espace $\mathscr{E}_C$ se compose des fonctions f sur $[0,1]$ telles que $f(0) = f(2^{-2n}) = 0$ pour tout $n \geq 0$, qui sont linéaires dans chacun des intervalles $[2^{-m-1}, 2^{-m}]$ (pour $m \geq 0$), et telles que $\sum_{n=0}^{\infty} |f(2^{-2n-1})|^2$ soit fini. Il est clair qu'une telle fonction est continue en dehors de 0; comme on a $|f(t)| \leq \sup_{n\geq m} |f(2^{-2n-1})|$ pour tout t dans l'intervalle $[0, 2^{-2m}]$, la fonction f est aussi continue en 0. Ceci établit la propriété (b).

15. Dans ce n°, on suppose que l'espace T est *localement compact* (donc séparé); on note $\mathscr{C}_c(T)$ l'espace vectoriel des fonctions continues (complexes) à support compact sur T. Soient $\mathscr{A}$ la tribu borélienne de T et μ une mesure positive définie sur $\mathscr{A}$; on suppose que le nombre $\mu(K)$ est fini pour toute

partie compacte K de T. Alors les éléments de $\mathscr{C}_c(T)$ sont des fonctions μ-intégrables.

THÉORÈME. *Soit C une fonction continue sur $T \times T$.*

(a) *Si C est une covariance, l'intégrale*

$$C[u] = \int_T \int_T C(t, t')\overline{u(t)}u(t')\, d\mu(t)\, d\mu(t') \tag{12}$$

est positive pour toute fonction u dans $\mathscr{C}_c(T)$.

(b) *Réciproquement, supposons que $\mu(U)$ soit non nul pour tout ouvert U non vide de T, et que l'on ait $C[u] \geq 0$ pour toute fonction $u \in \mathscr{C}_c(T)$. Alors C est une covariance sur T.*

Supposons que C soit une covariance. Soit u une fonction continue sur T, nulle hors d'une partie compacte K de T. Comme la restriction de C à $K \times K$ est une covariance, le théorème du n° 13 montre que C se représente sous forme d'une série

$$C(t, t') = \sum_{n=1}^{\infty} f_n(t) . \overline{f_n(t')} \tag{13}$$

qui converge uniformément sur $K \times K$, les fonctions f_n étant continues sur K. On peut donc intégrer terme à terme, d'où

$$\begin{aligned} C[u] &= \int_K \int_K C(t, t')\overline{u(t)}u(t')\, d\mu(t)\, d\mu(t') \\ &= \sum_{n=1}^{\infty} \int_K \int_K f_n(t)\overline{f_n(t')}\,\overline{u(t)}u(t')\, d\mu(t) d\mu(t') \\ &= \sum_{n=1}^{\infty} \left| \int_K \overline{u(t)} f_n(t)\, d\mu(t) \right|^2 ; \end{aligned}$$

on a donc $C[u] \geq 0$.

Réciproquement, faisons les hypothèses de (b). Soient $t_1, \ldots, t_n$ des points distincts de T et $\xi_1, \ldots, \xi_n$ des nombres complexes. Soit $\varepsilon > 0$; comme C est continue, on peut trouver des voisinages ouverts U_1 pour $t_1, \ldots, U_n$ pour t_n, tels que C oscille d'au plus ε sur chacun des ensembles $U_i \times U_j$. Quitte à rétrécir les ensembles U_i, on peut les supposer deux à deux disjoints. Il existe alors, pour chaque i, une fonction continue u_i sur T, positive, non identiquement nulle, à support compact contenu dans U_i. L'ensemble des points t de T tels que $u_i(t) > 0$ est ouvert et non vide, donc n'est pas μ-négligeable; quitte à multiplier u_i par une constante $c_i > 0$, on peut supposer que l'on a

$$\int_T u_i(t)\, d\mu(t) = 1 \qquad \text{pour} \quad 1 \leq i \leq n. \tag{14}$$

Posons $u = \xi_1 u_1 + \cdots + \xi_n u_n$. Alors $\overline{u(t)}u(t')$ est nul, sauf si (t, t') appartient à l'un des ensembles $U_i \times U_j$, auquel cas on a

$$\overline{u(t)}u(t') = \bar{\xi}_i \xi_j \overline{u_i(t)} u_j(t'), \tag{15}$$

$$|C(t, t') - C(t_i, t_j)| \leq \varepsilon. \tag{16}$$

On a alors

$$C[u] = \sum_{i,j=1}^{n} \bar{\xi}_i \xi_j \int_{U_i} \int_{U_j} C(t, t') \overline{u_i(t)} u_j(t')\, d\mu(t)\, d\mu(t'). \tag{17}$$

Compte tenu des formules (14), (16), et (17), on a finalement

$$\left| C[u] - \sum_{i,j=1}^{n} \bar{\xi}_i \xi_j C(t_i, t_j) \right| \leq \varepsilon . \left(\sum_{i=1}^{n} |\xi_i| \right)^2 . \tag{18}$$

Comme $C[u]$ est positif et que ε est arbitrairement petit, on a

$$\sum_{i,j=1}^{n} \bar{\xi}_i \xi_j C(t_i, t_j) \geq 0, \tag{19}$$

donc C est une covariance.

III. Opérateurs dans les espaces de Hilbert

Les notions dont nous avons besoin sont tout à fait classiques, et il existe d'excellents exposés (voir, e.g., Bourbaki [2]). Nous nous contenterons donc d'un bref résumé.

16. Soient $\mathscr{E}$ et $\mathscr{F}$ deux espaces de Hilbert. On appelle *opérateur* de $\mathscr{E}$ dans $\mathscr{F}$ toute application linéaire continue de $\mathscr{E}$ dans $\mathscr{F}$. Si v est un tel opérateur, il existe un opérateur v^* de $\mathscr{F}$ dans $\mathscr{E}$, appelé l'*adjoint* de v, et caractérisé par la relation

$$\langle v(x) | y \rangle = \langle x | v^*(y) \rangle \qquad \text{pour} \quad x \in \mathscr{E}, \quad y \in \mathscr{F}. \tag{1}$$

On dit que l'opérateur u dans $\mathscr{E}$ est *hermitien* si l'on a $u = u^*$; si l'on a de plus $\langle x | u(x) \rangle \geq 0$ pour tout x dans $\mathscr{E}$, on dit que u est *hermitien positif*.

17. Soit u un opérateur hermitien positif dans $\mathscr{E}$. Le nombre

$$\mathrm{Tr}(u) = \sum_{\alpha \in A} \langle e_\alpha | u(e_\alpha) \rangle \tag{2}$$

(positif ou égal à $+\infty$) est indépendant de la base orthonormale $(e_\alpha)_{\alpha \in A}$ de $\mathscr{E}$. On l'appelle la *trace* de u.

18. *Le théorème de décomposition spectrale* pour les opérateurs de trace finie (D.Hilbert, E.Schmidt) s'énonce comme suit:

Soient $\mathscr{E}$ un espace de Hilbert, u un opérateur hermitien positif de trace finie dans $\mathscr{E}$ et $\mathscr{E}_0$ le noyau de u (ensemble des vecteurs x de $\mathscr{E}$ tels que $u(x) = 0$). Soit $\mathscr{E}_+$ le sous-espace de $\mathscr{E}$ orthogonal à $\mathscr{E}_0$. Il existe alors une base orthonormale $(e_n)_{n\in I}$ de $\mathscr{E}_+$ et une suite décroissante $(\lambda_n)_{n\in I}$ de nombres réels strictement positifs qui satisfont aux conditions suivantes:

(a) on a $u(e_n) = \lambda_n e_n$ pour tout $n \in I$;
(b) ou bien I est un intervalle fini $[0, p]$ de l'ensemble des nombres entiers, ou bien I se compose de tous les entiers positifs, et la suite $(\lambda_n)_{n\geq 0}$ tend vers 0.

Sous les hypothèses précédentes, on a

$$u(x) = \sum_{n\in I} \lambda_n \langle e_n | x\rangle . e_n \qquad \text{pour } x \text{ dans } \mathscr{E}, \tag{3}$$

$$\mathrm{Tr}(u) = \sum_{n\in I} \lambda_n. \tag{4}$$

19. Soient $\mathscr{E}$ et $\mathscr{F}$ deux espaces de Hilbert, et v un opérateur de $\mathscr{E}$ dans $\mathscr{F}$. Soient $(e_\alpha)_{\alpha\in A}$ une base orthonormale de $\mathscr{E}$, et $(f_\beta)_{\beta\in B}$ une base orthonormale de $\mathscr{F}$. On a les égalités

$$\mathrm{Tr}(v^*v) = \sum_{\alpha\in A} \|v(e_\alpha)\|^2 = \sum_{\alpha\in A}\sum_{\beta\in B} |\langle v(e_\alpha)| f_\beta\rangle|^2 = \sum_{\beta\in B} \|v^*(f_\beta)\|^2 = \mathrm{Tr}(vv^*).$$

Si tous ces nombres sont finis, on dit que v est un *opérateur de Hilbert–Schmidt.*

La structure d'un opérateur de Hilbert–Schmidt $v\colon \mathscr{E} \to \mathscr{F}$ se décrit ainsi: Notons $\mathscr{E}_0$ le noyau de v et $\mathscr{F}_0$ celui de v^*; notons aussi $\mathscr{E}_+$ l'orthogonal de $\mathscr{E}_0$ dans $\mathscr{E}$ et $\mathscr{F}_+$ celui de $\mathscr{F}_0$ dans $\mathscr{F}$. Alors $\mathscr{E}_0$ est aussi le noyau de v^*v, et $\mathscr{F}_0$ celui de vv^*. Conformément au n° 18, choisissons une base orthonormale $(e_n)_{n\in I}$ de $\mathscr{E}_+$ et des nombres réels $\lambda_n > 0$ tels que $v^*v(e_n) = \lambda_n e_n$ pour tout $n \in I$. Posons $\mu_n = \lambda_n^{1/2}$ et $f_n = \mu_n^{-1} v(e_n)$ pour tout $n \in I$. Alors $(f_n)_{n\in I}$ est une base orthonormale de $\mathscr{F}_+$ et l'on a

$$v(x) = \sum_{n\in I} \mu_n \langle e_n | x\rangle . f_n \qquad \text{pour } x \text{ dans } \mathscr{E}, \tag{5}$$

$$v^*(y) = \sum_{n\in I} \mu_n \langle f_n | y\rangle . e_n \qquad \text{pour } y \text{ dans } \mathscr{F}. \tag{6}$$

De plus, on a

$$\mathrm{Tr}(vv^*) = \mathrm{Tr}(v^*v) = \sum_{n\in I} \mu_n^2. \tag{7}$$

20. L'exemple suivant d'opérateur de Hilbert–Schmidt est crucial pour la suite de ce travail.

THÉORÈME. *Soient $(T, \mathscr{A}, \mu)$ un espace mesuré et $\mathscr{H}$ un espace de Hilbert. Pour qu'un opérateur v de $\mathscr{H}$ dans $L^2 = L^2(T, \mathscr{A}, \mu)$ soit de Hilbert–Schmidt, il faut et il suffit qu'il existe une fonction positive $h \in L^2$ telle que l'on ait $|v(x)| \leq h$ μ-presque partout sur T, pour tous les vecteurs x de norme ≤ 1 dans $\mathscr{H}$. S'il en est ainsi, on a*

$$\mathrm{Tr}(v^* v) \leq \|h\|_2^2. \tag{8}$$

Supposons d'abord que v soit un opérateur de Hilbert–Schmidt de $\mathscr{H}$ dans L^2, que l'on représente comme dans (5). Pour tout $n \in I$, la fonction f_n est de norme 1 dans L^2, d'où

$$\sum_{n \in I} \int_T \mu_n^2 |f_n(t)|^2 \, d\mu(t) = \sum_{n \in I} \mu_n^2 < +\infty. \tag{9}$$

Quitte à modifier chacune des fonctions f_n sur un ensemble μ-négligeable, on peut supposer que la fonction h définie par

$$h(t) = \left\{ \sum_{n \in I} \mu_n^2 |f_n(t)|^2 \right\}^{1/2} \tag{10}$$

est finie en tout point de T. Soit alors x un vecteur de norme ≤ 1 dans $\mathscr{H}$. On a donc $\sum_{n \in I} |\langle e_n | x \rangle|^2 \leq 1$; d'après l'inégalité de Cauchy–Schwarz, on a donc

$$\left| \sum_{n \in I} \mu_n \langle e_n | x \rangle \cdot f_n(t) \right| \leq h(t), \tag{11}$$

la série du premier membre convergeant absolument pour tout t dans T. Comme la série $\sum_{n \in I} \mu_n \langle e_n | x \rangle \cdot f_n$ converge en norme dans L^2 vers $v(x)$, un représentant de la classe de fonctions $v(x)$ est fourni par la fonction $t \mapsto \sum_{n \in I} \mu_n \langle e_n | x \rangle \cdot f_n(t)$. D'après (11), on a donc $|v(x)| \leq h$ μ-presque partout sur T, quel que soit le représentant choisi pour $v(x)$.

Réciproquement, supposons qu'il existe une fonction positive h dans L^2 telle que l'on ait $|v(x)| \leq h$ μ-presque partout sur T, pour tout vecteur x de norme ≤ 1 dans $\mathscr{H}$. Soit $(e_\alpha)_{\alpha \in A}$ une base orthonormale dans $\mathscr{H}$, et soit f_α un représentant de la classe de fonctions $v(e_\alpha)$. Soit J une partie finie de A. Si $(\xi_\alpha)_{\alpha \in J}$ est une famille de nombres complexes telle que $\sum_{\alpha \in J} |\xi_\alpha|^2 \leq 1$, on a

$$\left| \sum_{\alpha \in J} \xi_\alpha f_\alpha(t) \right| \leq h(t) \tag{12}$$

μ-presque partout sur T. L'ensemble μ-négligeable exceptionnel dépend de la famille $(\xi_\alpha)_{\alpha \in J}$, mais il existe un ensemble μ-négligeable N tel que l'inégalité

(12) ait lieu chaque fois que t appartient à $T\backslash N$ et que les parties rèelles et imaginaires des ξ_α sont rationnelles. Par continuité, l'inégalité (12) reste valable pour tout t dans $T\backslash N$ quels que soient les nombres complexes ξ_α tels que $\sum_{\alpha\in J}|\xi_\alpha|^2 \leq 1$. On en déduit

$$\sum_{\alpha\in J} |f_\alpha(t)|^2 \leq h(t)^2 \qquad \text{pour } t \text{ dans } T\backslash N, \tag{13}$$

d'où par intégration

$$\sum_{\alpha\in J} \int_T |f_\alpha(t)|^2\, d\mu(t) \leq \int_T h(t)^2\, d\mu(t). \tag{14}$$

En passant à la borne supérieure sur les parties finies J de A, on obtient l'inégalité $\sum_{\alpha\in A}\|f_\alpha\|_2^2 \leq \|h\|_2^2 < +\infty$, donc v est un opérateur de Hilbert–Schmidt, et l'on a $\mathrm{Tr}(v^*v) \leq \|h\|_2^2$ comme annoncé.

IV. Covariances mesurables et séparables

Dans toute cette partie, on note T un ensemble muni d'une tribu $\mathscr{A}$ de parties.

21. Commençons par élucider les relations entre mesurabilité forte et faible. Soient $\mathscr{H}$ un espace de Hilbert séparable, $(e_n)_{n\in I}$ une base orthonormale de $\mathscr{H}$ et φ une application de T dans $\mathscr{H}$. On pose

$$f_n(t) = \langle e_n | \varphi(t)\rangle, \tag{1}$$

$$C(t,t') = \langle \varphi(t) | \varphi(t')\rangle \tag{2}$$

pour t, t' dans T et n dans I.

Lemme. *Les conditions suivantes sont équivalentes:*

(a) *L'application φ de T dans $\mathscr{H}$ est $\mathscr{A}$-mesurable si l'on munit $\mathscr{H}$ de sa tribu borélienne.*

(b) *Pour chaque élément a de $\mathscr{H}$, la fonction scalaire $t \mapsto \langle a|\varphi(t)\rangle$ est $\mathscr{A}$-mesurable.*

(c) *Pour chaque n dans I, la fonction scalaire f_n est $\mathscr{A}$-mesurable.*

(d) *La fonction C est mesurable sur $T \times T$ par rapport* à *la tribu $\mathscr{A} \otimes \mathscr{A}$.*

Pour tout a dans $\mathscr{H}$, l'application $x \mapsto \langle a|x\rangle$ de $\mathscr{H}$ dans $\mathbf{C}$ est continue, donc borélienne; par suite, (a) entraîne (b). Il est clair que (c) est un cas particulier de (b). Enfin, on a

$$C(t,t') = \sum_{n\in I} \overline{f_n(t)} f_n(t') \qquad \text{pour } t, t' \text{ dans } T; \tag{3}$$

comme l'espace de Hilbert $\mathscr{H}$ est séparable, l'ensemble I est dénombrable, donc la mesurabilité des fonctions f_n entraîne celle de C; par suite (c) entraîne (d).

Montrons que (c) *entraîne* (a). Pour tout vecteur $a = \sum_{n \in I} \alpha_n e_n$ de $\mathscr{H}$ et tout nombre réel $r > 0$, notons $B(a, r)$ la boule fermée de centre a et de rayon r dans $\mathscr{H}$; on sait que de telles boules engendrent la tribu borélienne sur l'espace séparable $\mathscr{H}$. Supposons que les fonctions f_n soient $\mathscr{A}$-mesurables. Pour toute partie fine J de I, la fonction $u_J = \sum_{n \in J} |f_n - \alpha_n|^2$ est $\mathscr{A}$-mesurable; de plus, l'ensemble des parties finies de I est dénombrable. Par suite, l'ensemble

$$\varphi^{-1}(B(a, r)) = \bigcap_J u_J^{-1}([0, r^2]) \tag{4}$$

est $\mathscr{A}$-mesurable. Autrement dit, la fonction φ est $\mathscr{A}$-mesurable.

Enfin, *montrons que* (d) *entraîne* (b).† Supposons C mesurable et notons $\mathscr{H}_1$ l'ensemble des vecteurs a de $\mathscr{H}$ tels que la fonction $t \mapsto \langle a | \varphi(t) \rangle$ soit $\mathscr{A}$-mesurable sur T. Il est clair que $\mathscr{H}_1$ est un sous-espace vectoriel fermé de $\mathscr{H}$. Tout vecteur de $\mathscr{H}$ orthogonal à l'image de φ appartient à $\mathscr{H}_1$; de plus, on a $\varphi(T) \subset \mathscr{H}_1$ car les fonctions partielles $t \mapsto C(t_0, t)$ sont $\mathscr{A}$-mesurables. On a donc $\mathscr{H}_1 = \mathscr{H}$, d'où notre assertion.

22. Soit C une covariance sur T. On dit que C est *séparable* si l'espace de Hilbert $\mathscr{E}_C$ est séparable; il revient au même de supposer que C admet un développement de la forme

$$C(t, t') = \sum_{n=1}^{\infty} f_n(t) \overline{f_n(t')}. \tag{5}$$

On dit que C est une *covariance mesurable* sur $(T, \mathscr{A})$ si la fonction C est mesurable sur $T \times T$ par rapport à la tribu $\mathscr{A} \otimes \mathscr{A}$. S'il en est ainsi, toute fonction de $\mathscr{E}_C$ est $\mathscr{A}$-mesurable puisque c'est la limite en norme dans $\mathscr{E}_C$, donc en chaque point de T, d'une suite de fonctions dont chacune est de la forme $t \mapsto \sum_{i=1}^n \xi_i C(t, t_i)$.

Le lemme du n° 21 se reformule comme suit:

LEMME. *Soit C une covariance séparable sur T, avec un développement de la forme* (5). *Les conditions suivantes sont équivalentes*:

(a) *C est une covariance mesurable sur $(T, \mathscr{A})$;*
(b) *$\mathscr{E}_C$ se compose de fonctions $\mathscr{A}$-mesurables sur T;*
(c) *chacune des fonctions f_n est $\mathscr{A}$-mesurable sur T.*

23. Donnons maintenant une application du théorème de Lusin.

† Cette partie de la démonstration reste valable pour des espaces de Hilbert non séparables.

THÉORÈME. *Supposons que T soit un espace polonais*[†] *et μ une mesure positive σ-finie sur la tribu borélienne de T. Soit C une covariance séparable sur T, qui est une fonction borélienne sur $T \times T$. Choisissons un développement de C comme en* (5). *Alors il existe une suite croissante $(K_p)_{p \geq 1}$ de parties compactes de T avec les propriétés suivantes*:

(a) *le complémentaire dans T de la réunion des K_p est μ-négligeable*;

(b) *pour tout entier $p \geq 1$, la restriction de C à $K_p \times K_p$ est continue, et la série* (5) *définissant C converge absolument et uniformément sur $K_p \times K_p$.*

Rappelons que la mesure μ satisfait à la propriété de régularité intérieure: la mesure $\mu(A)$ d'une partie borélienne A de T est la borne supérieure des mesures $\mu(K)$ des parties compactes K de T contenues dans A. De plus, comme on a supposé que la mesure μ est σ-finie, il existe une mesure finie ayant les mêmes ensembles négligeables que μ, et l'on ne restreint pas la généralité en supposant que $\mu(T)$ est fini. Rappelons enfin le théorème de Lusin: si K est une partie compacte de T, f une fonction borélienne sur K et ε un nombre réel strictement positif, il existe une partie compacte L de K telle que $\mu(K \backslash L) < \varepsilon$ et que la restriction de f à L soit une application continue de L dans $\mathbf{C}$.

Construisons par récurrence une suite de parties compactes L_p de T, deux à deux disjointes. Supposons construites les parties $L_1, \ldots, L_p$ et soit A le complémentaire dans T de leur réunion. Choisissons une partie compacte K de A telle que $\mu(A \backslash K) \leq 2^{-p-2}$ et posons $f_0 = \sum_{n=1}^{\infty} |f_n|^2$. Par application du théorème de Lusin, choisissons pour tout entier $n \geq 0$ une partie compacte M_n de K telle que la restriction de f_n à M_n soit continue, et que l'on ait $\mu(K \backslash M_n) < 2^{-p-n-3}$. Posons $L_{p+1} = \bigcap_{n=1}^{\infty} M_n$; par construction, L_{p+1} est disjoint de $L_1, \ldots, L_p$, c'est une partie compacte de T, la restriction de chacune des fonctions f_n à L_{p+1} est continue et l'on a

$$\mu(T \backslash (L_1 \cup \cdots \cup L_{p+1})) = \mu(A \backslash L_{p+1}) \leq \mu(A \backslash K) + \sum_{n=0}^{\infty} \mu(K \backslash M_n)$$

$$\leq 2^{-p-2} + \sum_{n=0}^{\infty} 2^{-p-n-3} = 2^{-p-1}.$$

Posons alors $K_p = L_1 \cup \cdots \cup L_p$; la suite des parties compactes K_p de T est croissante, et comme on a $\mu(T \backslash K_p) \leq 2^{-p}$, la réunion des K_p a un complémentaire μ-négligeable. De plus, la restriction à K_p de chacune des fonctions f_n (pour $n \geq 1$) et $f_0 = \sum_{n=1}^{\infty} |f_n|^2$ est continue.

Appliquons alors le lemme de Dini: la série $\sum_{n=1}^{\infty} |f_n|^2$ converge *uniformément* sur l'espace compact K_p. Soit alors $\varepsilon > 0$ et soit $m \geq 1$ un entier

[†] Autrement dit, T est un espace métrique séparable et complet.

tel que l'on ait $\sum_{n>m} |f_n(t)|^2 < \varepsilon$ pour tout t dans K_p. L'inégalité de Cauchy–Schwarz montre alors que l'on a $\sum_{n>m} |f_n(t)|.|f_n(t')| < \varepsilon$ pour t, t' dans K_p. Par suite, la série $\sum_{n=1}^{\infty} f_n(t).\overline{f_n(t')}$ converge absolument et uniformément sur $K_p \times K_p$, et la restriction de C à $K_p \times K_p$ est une fonction continue.

24. Considérons maintenant un espace mesuré σ-fini $(\Omega, \mathscr{F}, \lambda)$ et une fonction complexe X sur $T \times \Omega$. On fait les deux hypothèses suivantes:

(a) *X est mesurable par rapport à la tribu $\mathscr{A} \otimes \mathscr{F}$ sur $T \times \Omega$;*
(b) *pour tout point t de T, l'intégrale $\int_\Omega |X(t,\omega)|^2 \, d\lambda(\omega)$ est finie.*

Pour chaque point t de T, la fonction $\omega \mapsto X(t,\omega)$ sur Ω est mesurable par rapport à $\mathscr{F}$ et de carré intégrable par rapport à λ; elle définit donc un élément X_t de l'espace de Hilbert $L^2 = L^2(\Omega, \mathscr{F}, \lambda)$. La covariance C associée à l'application $t \mapsto X_t$ de T dans L^2 est donnée par la formule

$$C(t,t') = \int_\Omega \overline{X(t,\omega)}.X(t',\omega)\, d\lambda(\omega). \tag{6}$$

LEMME. *La covariance C est séparable et mesurable sur $(T, \mathscr{A})$.*

La formule (6) montre à l'évidence que C est mesurable sur $T \times T$ par rapport à la tribu $\mathscr{A} \otimes \mathscr{A}$ (la mesure λ est supposée σ-finie). Pour montrer que C est séparable, il suffit de montrer que les X_t appartiennent à un sous-espace de Hilbert séparable de L^2. Or, par définition, la tribu $\mathscr{A} \otimes \mathscr{F}$ est engendrée par les ensembles de la forme $A \times B$ avec $A \in \mathscr{A}$ et $B \in \mathscr{F}$, et chaque élément de $\mathscr{A} \otimes \mathscr{F}$ appartient à la sous-tribu engendrée par une suite de tels ensembles $A_n \times B_n$. Appliquons ceci à chacun des ensembles $X^{-1}(I)$ où I est un intervalle à extrémités rationnelles de $\mathbf{R}$. Il existe donc une sous-tribu $\mathscr{A}_0$ de $\mathscr{A}$, et une sous-tribu $\mathscr{F}_0$ de $\mathscr{F}$, toutes deux dénombrablement engendrées, et telles que X soit mesurable par rapport à la tribu $\mathscr{A}_0 \otimes \mathscr{F}_0$. Mais alors $L^2(\Omega, \mathscr{F}_0, \lambda_0)$ (où λ_0 est la restriction de λ à $\mathscr{F}_0$) est un sous-espace séparable de l'espace de Hilbert $L^2 = L^2(\Omega, \mathscr{F}, \lambda)$, contenant tous les X_t. Ceci prouve le lemme.

25. Concluons cette partie en donnant un exemple typique de *covariance mesurable et non séparable*. Prenons pour T l'intervalle $[0,1]$ de $\mathbf{R}$ et pour $\mathscr{A}$ la tribu borélienne sur T.

Définissons une fonction borélienne D sur $T \times T$ par la formule

$$D(t,t') = \begin{cases} 1 & \text{si} \quad t = t', \\ 0 & \text{si} \quad t \neq t'. \end{cases} \tag{7}$$

Autrement dit, D est la fonction caractéristique (ou indicatrice) de la diagonale du carré $[0,1] \times [0,1]$. Si $t_1, \ldots, t_n$ sont des points distincts de

T, et $\xi_1, \ldots, \xi_n$ des nombres complexes, on a

$$\sum_{i,j=1}^{n} D(t_i, t_j)\bar{\xi}_i\xi_j = \sum_{i=1}^{n} |\xi_i|^2, \tag{8}$$

donc D est une covariance sur $[0, 1]$.

Posons $D_t(t') = D(t', t)$. Vu la définition de D, la famille non dénombrable $(D_t)_{t \in T}$ est une base orthonormale de l'espace de Hilbert $\mathscr{E}_D$, donc la covariance D n'est pas séparable.

L'espace $\mathscr{E}_D$ se compose des fonctions f sur T telles que $\sum_{t \in T} |f(t)|^2$ soit fini; une telle fonction est nulle en dehors d'une partie dénombrable de T. Enfin, le produit scalaire dans l'espace $\mathscr{E}_D$ est donné par la formule

$$\langle f \,|\, g \rangle_D = \sum_{t \in T} \overline{f(t)} . g(t). \tag{9}$$

V. Une généralisation du théorème de Mercer

Dans toute cette partie, on note $(T, \mathscr{A}, \mu)$ un espace mesuré σ-fini.

26. Voici d'abord le lemme-clé:

Lemme. *Soit C une covariance mesurable sur $(T, \mathscr{A})$. On suppose que l'intégrale*

$$\gamma = \int_T C(t, t)\, d\mu(t) \tag{1}$$

est finie (ce qui a lieu par exemple si $\mu(T)$ est fini et la fonction C bornée). Alors toute fonction appartenant à $\mathscr{E}_C$ est de carré μ-intégrable. De plus, l'application qui à toute fonction dans $\mathscr{E}_C$ fait correspondre sa classe d'équivalence μ-presque partout est un opérateur de Hilbert–Schmidt j de $\mathscr{E}_C$ dans $L^2 = L^2(T, \mathscr{A}, \mu)$. On a $\mathrm{Tr}(j^*j) \le \gamma$, *avec égalité si C est séparable.*

Posons $h(t) = C(t, t)^{1/2}$ pour t dans T; par hypothèse, la fonction h sur T est de carré μ-intégrable et l'on a $\|h\|_2^2 = \gamma$. D'après la formule (22) du n° 6, on a $|f(t)| \le \|f\|_C . h(t)$ pour $t \in T$ et $f \in \mathscr{E}_C$. Il en résulte que toute fonction $f \in \mathscr{E}_C$ est de carré μ-intégrable, et que j est un opérateur de $\mathscr{E}_C$ dans L^2. Le théorème du n° 20 montre alors que j est un opérateur de Hilbert–Schmidt, et que l'on a $\mathrm{Tr}(j^*j) \le \|h\|_2^2 = \gamma$.

Supposons maintenant C séparable. Si $(f_n)_{n \in I}$ est une base orthonormale de l'espace de Hilbert $\mathscr{E}_C$, on a $\sum_{n \in I} |f_n(t)|^2 = C(t, t)$ pour tout t dans T, et l'ensemble I est dénombrable. Par intégration, on en déduit

$$\mathrm{Tr}(j^*j) = \sum_{n \in I} \|j(f_n)\|_2^2 = \sum_{n \in I} \int_T |f_n(t)|^2\, d\mu(t) = \gamma.$$

27. On conserve les notations et les hypothèses du lemme précédent. Par construction, le noyau $\mathscr{N}$ de l'opérateur $j:\mathscr{E}_C \to L^2$ se compose des fonctions μ-négligeables appartenant à $\mathscr{E}_C$; de plus, $\mathscr{E}_C$ est somme directe de $\mathscr{N}$ et du sous-espace $\mathscr{S}$ orthogonal à $\mathscr{N}$. Pour tout $t \in T$, décomposons l'élément $C_t: t' \mapsto C(t', t)$ de $\mathscr{E}_C$ en $C_t = N_t + S_t$ avec $N_t \in \mathscr{N}$ et $S_t \in \mathscr{S}$. Pour toute fonction f dans $\mathscr{E}_C$, on a

$$f(t) = \langle C_t | f \rangle_C = \langle N_t | f \rangle_C + \langle S_t | f \rangle_C. \tag{2}$$

En particulier, on a $g(t) = \langle N_t | g \rangle$ pour g dans $\mathscr{N}$; si l'on pose $N(t, t') = N_{t'}(t)$, la fonction N est une covariance sur T et l'on a $\mathscr{N} = \mathscr{E}_N$. On introduit de même la covariance S par $S(t, t') = S_{t'}(t)$ et l'on a $\mathscr{S} = \mathscr{E}_S$.

Comme j est un opérateur de Hilbert–Schmidt de $\mathscr{E}_C$ dans L^2, il résulte des résultats rappelés au n° 19 que l'orthogonal du noyau de j est un espace de Hilbert séparable, donc la covariance S est séparable. Comme la covariance C est mesurable sur $(T, \mathscr{A})$, l'espace $\mathscr{E}_C$ se compose de fonctions $\mathscr{A}$-mesurables, et il en donc de même du sous-espace $\mathscr{E}_S$ de $\mathscr{E}_C$; comme la covariance S est séparable, elle est donc aussi mesurable sur $(T, \mathscr{A})$. Comme on a $N = C - S$, la covariance N est mesurable sur $(T, \mathscr{A})$. De plus, par construction, aucune fonction non nulle de $\mathscr{E}_S$ n'est μ-négligeable, et N_t est μ-négligeable pour tout t dans T.

On dira dans la suite que $C = N + S$ exprime la *μ-décomposition* de la covariance C mesurable sur $(T, \mathscr{A})$. Montrons que les propriétés énoncées précédemment caractérisent cette décomposition de C.

LEMME. *Soit C une covariance mesurable sur $(T, \mathscr{A})$ telle que l'intégrale $\int_T C(t, t)\, d\mu(t)$ soit finie. Soient S' et N' deux covariances mesurables sur $(T, \mathscr{A})$, de somme C. On suppose que* 0 *est la seule fonction μ-négligeable appartenant a $\mathscr{E}_{S'}$, et que la fonction $N'_t: t' \mapsto N'(t', t)$ est μ-négligeable pour tout t dans T. Alors, on a $N' = N$ et $S' = S$.*

Posons $\mathscr{N}' = \mathscr{E}_{N'}$ et $\mathscr{S}' = \mathscr{E}_{S'}$. Toute fonction appartenant à $\mathscr{N}'$ est limite en norme dans $\mathscr{N}'$, donc en tout point de T, d'une suite de combinaisons linéaires finies de fonctions N'_t. Donc $\mathscr{N}'$ se compose de fonctions μ-négligeables, et par hypothése, on a $\mathscr{N}' \cap \mathscr{S}' = 0$. Sur l'espace somme $\mathscr{E}' = \mathscr{N}' + \mathscr{S}'$, on définit un produit scalaire par la formule

$$\langle g + h | g' + h' \rangle = \langle g | g' \rangle_{N'} + \langle h | h' \rangle_{S'} \tag{3}$$

pour g, g' dans $\mathscr{N}'$ et h, h' dans $\mathscr{S}'$. Alors $\mathscr{E}'$ est un espace de Hilbert, somme directe des sous-espaces orthogonaux $\mathscr{N}'$ et $\mathscr{S}'$. Pour tout point t de T, introduisons l'élément $S'_t: t' \mapsto S'(t', t)$ de $\mathscr{S}'$. De la relation $C = N' + S'$, on déduit que $C_t = N'_t + S'_t$ appartient à $\mathscr{E}'$; la définition (3) du produit scalaire dans $\mathscr{E}'$ montre aussitôt que l'on a $f(t) = \langle C_t | f \rangle$ pour $t \in T$ et $f \in \mathscr{E}'$.

On a donc $\mathscr{E}' = \mathscr{E}_C$, et comme $\mathscr{N}'$ se compose visiblement des fonctions μ-négligeables de $\mathscr{E}'$, on a $\mathscr{N}' = \mathscr{N}$ avec égalité des produits scalaires associés respectivement à N' et N. On a donc $N' = N$, d'où $S' = C - N' = C - N = S$.

28. THÉORÈME. *Toute covariance mesurable sur $(T, \mathscr{A})$ est la somme d'une covariance séparable et mesurable sur $(T, \mathscr{A})$ et d'une covariance qui est une fonction sur $T \times T$ négligeable pour la mesure $\mu \otimes \mu$.*

Par hypothèse, la mesure μ est σ-finie sur $(T, \mathscr{A})$. Il existe donc une mesure ν sur $(T, \mathscr{A})$ ayant les mêmes ensembles négligeables que μ, et telle que $\nu(T)$ soit fini. D'après le théorème de Lebesgue–Fubini, les mesures $\mu \otimes \mu$ et $\nu \otimes \nu$ ont les mêmes ensembles négligeables.

Soit D une covariance mesurable sur $(T, \mathscr{A})$. L'ensemble T_1 des points de T tels que $D(t,t) \neq 0$ appartient à $\mathscr{A}$, et la fonction u définie sur T par

$$u(t) = \begin{cases} D(t,t)^{1/2} & \text{si} \quad t \in T_1, \\ 1 & \text{si} \quad t \in T \backslash T_1 \end{cases} \tag{4}$$

est $\mathscr{A}$-mesurable et positive. Si l'on pose

$$C(t,t') = u(t)^{-1} u(t')^{-1} D(t,t'), \tag{5}$$

on définit une covariance C mesurable sur $(T, \mathscr{A})$. On a $0 \leq C(t,t) \leq 1$ pour tout t dans T, et comme $\nu(T)$ est fini, l'intégrale $\int_T C(t,t)\, d\nu(t)$ est finie.

Introduisons la ν-décomposition $C = N + S$, et posons

$$D_1(t,t') = u(t)u(t')S(t,t'), \tag{6}$$

$$D_2(t,t') = u(t)u(t')N(t,t') \tag{7}$$

pour t, t' dans T. Alors D_1 et D_2 sont deux covariances mesurables sur $(T, \mathscr{A})$, de somme D. Comme S est séparable, il en est de même de D_1. Par hypothèse, l'application $t' \mapsto N(t',t)$ est ν-négligeable pour tout t dans T; d'après le théorème de Lebesgue–Fubini, la fonction N est négligeable pour $\nu \otimes \nu$, donc aussi pour $\mu \otimes \mu$. Enfin, la formule (7) montre que D_2 est elle aussi négligeable pour $\mu \otimes \mu$.

29. Dans la discussion qui suit, nous devrons soigneusement distinguer une fonction f mesurable sur $(T, \mathscr{A})$ de sa classe d'équivalence pour l'égalité presque partout, qu'on notera $\tilde{f}$.

Soit de nouveau C une covariance mesurable sur $(T, \mathscr{A})$ telle que l'intégrale $\int_T C(t,t)\, d\mu(t)$ soit finie. Pour tout point t de T, la fonction $C_t : t' \mapsto \overline{C(t,t')}$ appartient à $\mathscr{E}_C$, donc elle est de carré μ-intégrable. L'intégrale

$$KF(t) = \int_T C(t,t') . f(t')\, d\mu(t') \tag{8}$$

est donc définie si f est une fonction de carré μ-intégrable sur T, quel que soit le point t de T.

LEMME. *L'opérateur j^* de L^2 dans $\mathscr{E}_C$ adjoint de j est défini par la relation $j^*(\tilde{f}) = Kf$ pour toute fonction f de carré μ-intégrable.*

Soit t un point de T. On a

$$\begin{aligned} j^*(\tilde{f})(t) &= \langle C_t | j^*(\tilde{f}) \rangle_C = \langle j(C_t) | \tilde{f} \rangle_2 = \langle \tilde{C}_t | \tilde{f} \rangle_2 \\ &= \int_T \overline{C_t(t')} f(t') \, d\mu(t') = \int_T C(t, t') f(t') \, d\mu(t'). \end{aligned}$$

Ce calcul établit le lemme.

Le lemme précédent montre que, si f est une fonction de carré μ-intégrable sur T, la fonction Kf définie par la formule (8) est aussi de carré μ-intégrable. Par passage aux classes, on définit donc un opérateur $\tilde{K}$ dans L^2, tel que $\tilde{K}\tilde{f} = \widetilde{Kf}$; cet opérateur n'est autre que jj^*.

30. Nous aurons encore besoin d'une remarque sur les vecteurs propres de l'opérateur $\tilde{K}$ dans L^2.

LEMME. *Soient u un élément de L^2 et λ un nombre réel non nul tels que $\tilde{K}u = \lambda u$. Il existe un unique représentant f de u tel que $Kf = \lambda f$.*

La formule de définition de Kf montre aussitôt que la fonction Kf est la même pour tous les représentants f de u; si l'on note Ku cette fonction, on a $\tilde{K}u = \widetilde{Ku}$. L'égalité $\tilde{K}u = \lambda u$ signifie donc que u est la classe de la fonction $f = \lambda^{-1}Ku$; on a $Ku = \lambda f$, et comme f est un représentant de u, on a $Kf = \lambda f$.

Soit f_1 un représentant de u tel que $Kf_1 = \lambda f_1$. Comme f et f_1 diffèrent par une fonction μ-négligeable, on a $Kf_1 = Kf$, d'où $\lambda f_1 = \lambda f$; en simplifiant par la constante $\lambda \neq 0$, on conclut à l'égalité de f_1 et f.

31. Voici maintenant la généralisation annoncée du théorème de Mercer.

THÉORÈME. *Soient $(T, \mathscr{A}, \mu)$ un espace mesuré σ-fini, et C une covariance mesurable sur $(T, \mathscr{A})$ telle que l'intégrale $\gamma = \int_T C(t,t)\, d\mu(t)$ soit finie.*

(a) *Pour toute fonction f de carré μ-intégrable sur T, la fonction Kf définie par*

$$Kf(t) = \int_T C(t, t') . f(t') \, d\mu(t') \tag{9}$$

(pour $t \in T$) *est de carré μ-intégrable sur T.*

(b) *Par passage aux classes, K définit un opérateur $\tilde{K}$ dans $L^2 = L^2(T, \mathscr{A}, \mu)$ qui est hermitien, positif et de trace finie.*

(c) *Soit $\tilde{\mathcal{N}}$ le noyau de $\tilde{K}$. Il existe une suite $(f_n)_{n \in I}$ de fonctions de carré μ-intégrable sur T, et une suite décroissante $(\lambda_n)_{n \in I}$ de nombres réels strictement positifs tels que l'on ait*

$$Kf_n = \lambda_n f_n \qquad \text{pour tout} \quad n \in I, \tag{10}$$

et que la famille des classes $(\tilde{f}_n)_{n \in I}$ soit une base orthonormale du sous-espace de L^2 orthogonal à $\tilde{\mathcal{N}}$. L'ensemble I est fini, ou bien la suite $(\lambda_n)_{n \in I}$ tend vers 0.

(d) *Soit $C = N + S$ la μ-décomposition de la covariance C. Alors S est donnée par le développement absolument convergent*

$$S(t, t') = \sum_{n \in I} \lambda_n f_n(t) \overline{f_n(t')}, \tag{11}$$

et les fonctions S et C sont égales presque partout sur $T \times T$ par rapport à la mesure $\mu \otimes \mu$.

(e) *La trace de $\tilde{K}$ est égale à $\sum_{n \in I} \lambda_n$. On a aussi la relation*

$$\operatorname{Tr}(\tilde{K}) = \int_T S(t, t)\, d\mu(t) \leq \int_T C(t, t)\, d\mu(t), \tag{12}$$

avec égalité si la covariance C est séparable.

D'après le lemme du n° 26, la formule $j(f) = \tilde{f}$ définit un opérateur de Hilbert–Schmidt de $\mathscr{E}_C$ dans L^2. On a vu au n° 29 que $\tilde{K}$ est égal à jj^*, d'où les assertions (a) et (b). L'assertion (c) résulte du théorème de décomposition spectrale rappelé au n° 18, le choix précisé des fonctions propres se faisant au moyen du lemme du n° 30.

Le noyau de j est l'espace $\mathcal{N}$ introduit au n° 27. Comme j est un opérateur de Hilbert–Schmidt, il en est de même de j^*. En appliquant le théorème de structure des opérateurs de Hilbert–Schmidt rappelé au n° 19, on voit que les fonctions $\lambda_n^{-1/2} j^*(\tilde{f}_n)$ forment une base orthonormale du sous-espace $\mathscr{S}$ de $\mathscr{E}_C$ orthogonal à $\mathcal{N}$. Or on a

$$\lambda_n^{-1/2} j^*(\tilde{f}_n) = \lambda_n^{-1/2} K f_n = \lambda_n^{1/2} f_n, \tag{13}$$

et $\mathscr{S}$ est l'espace de Hilbert $\mathscr{E}_S$ associé à la covariance S. La formule (11) résulte aussitôt de là. On sait que, pour tout point t de T, l'application $t' \mapsto N(t', t)$ est μ-négligeable; d'après le théorème de Lebesgue–Fubini, la fonction $N = C - S$ est donc négligeable pour $\mu \otimes \mu$. Ceci établit l'assertion (d).

La formule $\operatorname{Tr}(\tilde{K}) = \sum_{n \in I} \lambda_n$ résulte des rappels du n° 18. De plus, d'après la formule (11), on a $S(t, t) = \sum_{n \in I} \lambda_n |f_n(t)|^2$ pour tout $t \in T$; par intégration de cette série de fonctions positives, on obtient

$$\int_T S(t, t)\, d\mu(t) = \sum_{n \in I} \lambda_n \int_T |f_n(t)|^2\, d\mu(t) = \sum_{n \in I} \lambda_n. \tag{14}$$

Enfin, on a $\tilde{K} = jj^*$, d'où par application du lemme du nº 26 la relation

$$\operatorname{Tr}(\tilde{K}) = \operatorname{Tr}(jj^*) = \operatorname{Tr}(j^*j) \leq \gamma, \tag{15}$$

avec égalité si la covariance C est séparable.

32. Il n'y a pas toujours égalité dans la formule (12). Revenons en effet à l'exemple du nº 25. Nous supposons donc que T est l'intervalle $[0, 1]$ de $\mathbf{R}$, que $\mathscr{A}$ est la tribu borélienne sur T, et que μ est la mesure de Lebesgue. La covariance D est l'indicatrice de la diagonale dans $T \times T$. Pour tout $t \in T$, la fonction D_t est nulle en tout point de T différent de t, donc est μ-négligeable. Par suite, la μ-décomposition de D est donnée par $N = D$ et $S = 0$. L'opérateur intégral associé à D par la formule

$$Kf(t) = \int_T D(t, t')f(t')\,dt' \tag{16}$$

est nul. On a donc $\tilde{K} = 0$, d'où $\operatorname{Tr}(\tilde{K}) = 0$. Cependant, on a $\int_T D(t, t)\,dt = 1$.

33. Pour retrouver le théorème de Mercer usuel, supposons désormais que T soit un espace topologique, $\mathscr{A}$ la tribu borélienne sur T, et que la mesure μ satisfasse à $\mu(U) > 0$ pour toute partie ouverte non vide U de T. On suppose aussi que la covariance C est continue, et que l'intégrale $\int_T C(t, t)\,d\mu(t)$ est finie.

Comme la covariance C est continue, l'espace $\mathscr{E}_C$ se compose de fonctions continues (voir le nº 11). Soit $\mathscr{N}$ l'ensemble des fonctions μ-négligeables appartenant à $\mathscr{E}_C$. Si f appartient à $\mathscr{N}$, l'ensemble des points où f est non nulle est un ouvert U de T tel que $\mu(U) = 0$, d'ou $U = \varnothing$ et $f = 0$. L'orthogonal $\mathscr{S}$ de $\mathscr{N}$ dans $\mathscr{E}_C$ est donc égal à $\mathscr{E}_C$, et dans la μ-décomposition $C = N + S$, on a $N = 0$ et $S = C$; en particulier, la covariance C est séparable.

Dans le théorème du nº 31, on peut donc remplacer les assertions (d) et (e) par les suivantes:

(d′) *Les fonctions f_n sont continues et C admet un développement convergeant absolument et uniformément sur toute partie compacte de T*:

$$C(t, t') = \sum_{n \in I} \lambda_n f_n(t) \cdot \overline{f_n(t')}. \tag{17}$$

(e′) *On a les égalités*

$$\operatorname{Tr}(\tilde{K}) = \sum_{n \in I} \lambda_n = \int_T C(t, t)\,d\mu(t). \tag{18}$$

Voici quelques propriétés supplémentaires:

(f) La famille $(\lambda_n^{1/2} f_n)_{n \in I}$ est une base orthonormale de l'espace de Hilbert $\mathscr{E}_C$. D'aprés les nºˢ Il et 12, toute fonction f de $\mathscr{E}_C$ est continue et elle admet un développement orthogonal $f = \sum_{n \in I} c_n f_n$ avec $\sum_{n \in I} |c_n|^2/\lambda_n$ fini, qui converge en norme dans les espaces de Hilbert $L^2(T, \mathscr{A}, \mu)$ et $\mathscr{E}_C$, et uniformément sur toute partie compacte de T. Les coefficients c_n sont donnés par la formule usuelle

$$c_n = \int_T \overline{f_n(t)} f(t) \, d\mu(t). \tag{19}$$

(g) D'après le lemme du n° 29, la fonction $f = Kg$ appartient à $\mathscr{E}_C$ pour toute fonction g de carré μ-intégrable; on peut donc lui appliquer les résultats de (f).

Pour les résultats classiques sur le théorème de Mercer, on pourra consulter Riesz et Nagy [13].

VI. Covariances mesurables et bornées

Dans toute cette partie, on note $(T, \mathscr{A}, \mu)$ un espace mesuré σ-fini.

34. Commençons par un résultat facile sur les covariances bornées.

Lemme. *Soient C une covariance sur T, et M une constante positive. Les conditions suivantes sont équivalentes:*

(a) *on a $C(t, t) \leq M$ pour tout $t \in T$;*
(b) *on a $|C(t, t')| \leq M$ pour t, t' dans T;*
(c) *on a $|f(t)| \leq M^{1/2}$ pour toute fonction f dans $\mathscr{E}_C$ telle que $\|f\|_C \leq 1$.*

L'équivalence de (a) et (b) résulte de l'inégalité

$$|C(t, t')|^2 \leq C(t, t) . C(t', t'). \tag{1}$$

Par ailleurs, on a $f(t) = \langle C_t | f \rangle_C$, et la condition (c) équivaut à $\|C_t\|_C \leq M^{1/2}$; l'équivalence de (a) et (c) provient alors de la formule $C(t, t) = \|C_t\|_C^2$.

35. On suppose maintenant que C est une covariance mesurable sur $(T, \mathscr{A})$, qui est bornée comme fonction sur $T \times T$. Si u est une fonction μ-intégrable sur T, la fonction Ku sur T définie par

$$Ku(t) = \int_T C(t, t') u(t') \, d\mu(t') \tag{2}$$

est $\mathscr{A}$-mesurable et bornée.

Lemme. *La fonction Ku appartient à $\mathscr{E}_C$ et l'on a*

$$\langle Ku | f \rangle_C = \int_T \overline{u(t)} f(t) \, d\mu(t) \qquad \textit{pour toute fonction} \quad f \in \mathscr{E}_C \tag{3}$$

$$\|Ku\|_C^2 = \int_T \int_T C(t, t') \overline{u(t)} u(t') \, d\mu(t) \, d\mu(t'). \tag{4}$$

Soit M une constante positive telle que l'on ait $|C(t, t')| \leq M$ pour t, t' dans

T. D'après le lemme précédent, on a $|f(t)| \leq M^{1/2}.\|f\|_C$ pour $t \in T$ et $f \in \mathscr{E}_C$; par intégration on en déduit l'inégalité

$$\left|\int_T \overline{u(t)} f(t)\, d\mu(t)\right| \leq M^{1/2}.\|u\|_1\|f\|_C \tag{5}$$

pour tout $f \in \mathscr{E}_C$. Comme l'espace de Hilbert $\mathscr{E}_C$ est son propre dual, il existe un élément g de $\mathscr{E}_C$ tel que l'on ait

$$\int_T \overline{u(t)} f(t)\, d\mu(t) = \langle g | f \rangle_C \qquad \text{pour tout} \quad f \in \mathscr{E}_C. \tag{6}$$

Prenons en particulier $f = C_{t'}$, d'où $f(t) = \overline{C(t',t)}$ et $\langle g | f \rangle_C = \overline{g(t')}$; on en déduit l'égalité de g et de Ku, d'oú la formule (3). Pour obtenir la formule (4), il suffit de faire $f = Ku$ dans la formule (3) et d'appliquer le théorème de Lebesgue–Fubini.

36. *Théorème. Soient C une covariance séparable et mesurable sur $(T, \mathscr{A})$ et M une constante positive telles que l'on ait $|C(t,t')| \leq M$ presque partout sur $T \times T$ par rapport à la mesure $\mu \otimes \mu$. Il existe alors un ensemble μ-négligeable N tel que l'on ait $|C(t,t')| \leq M$ lorsque t et t' appartiennent à $T \backslash N$. En particulier, on a $C(t,t) \leq M$ μ-presque partout sur T.*

Faisons d'abord la démonstration dans le cas ou C est bornée. Vu l'hypothèse faite sur C, la formule (4) entraîne $\|Ku\|_C^2 \leq M.\|u\|_1^2$; d'après la formule (3) et l'inégalité de Cauchy–Schwarz, on a donc

$$\left|\int_T \overline{u(t)} f(t)\, d\mu(t)\right| \leq M^{1/2}.\|f\|_C\|u\|_1 \tag{7}$$

quelles que soient la fonction $f \in \mathscr{E}_C$ et la fonction μ-intégrable u. Comme il est bien connu, on déduit de là l'existence d'un ensemble μ-négligeable N_f (dépendant de f) tel que l'on ait $|f(t)| \leq M^{1/2}\|f\|_C$ pour tout point t de $T \backslash N_f$. Comme la covariance C est séparable, il existe une partie dénombrable dense D dans l'espace de Hilbert $\mathscr{E}_C$. Posons $N = \bigcup_{f \in D} N_f$; c'est un ensemble μ-négligeable. D'après ce qui précède, on a donc

$$|\langle C_t | f \rangle_C| \leq M^{1/2}.\|f\|_C \qquad \text{pour} \quad f \in D \quad \text{et} \quad t \in T \backslash N. \tag{8}$$

Comme D est dense dans $\mathscr{E}_C$, on a donc $\|C_t\| \leq M^{1/2}$ pour t dans $T \backslash N$. En utilisant l'inégalité de Cauchy–Schwarz et la formule $C(t,t') = \langle C_t | C_{t'} \rangle_C$, on en déduit l'inégalité $|C(t,t')| \leq M$ pour t, t' dans $T \backslash N$.

Passons au cas général. Pour tout entier $p \geq 1$, soit T_p l'ensemble des points t de T tels que $C(t,t) \leq p$. On a $|C(t,t')| \leq p$ pour t, t' dans T_p, et la restriction de C à $T_p \times T_p$ est donc une covariance bornée. D'après la première partie de la' démonstration, il existe une partie μ-négligeable N_p de T_p telle que l'on ait $|C(t,t')| \leq M$ pour t, t' dans $T_p \backslash N_p$. L'ensemble $N = \bigcup_{p=1}^{\infty} N_p$ satisfait aux exigences du théorème.

37. On ne peut supprimer l'hypothèse de séparabilité dans le théorème précédent. Reprenons notre exemple familier, où T est l'intervalle $[0,1]$ de $\mathbf{R}$, $\mathscr{A}$ est la tribu borélienne de T et μ la mesure de Lebesgue. Soit D l'indicatrice de la diagonale dans $T \times T$. Alors D est nulle presque partout sur $T \times T$, et l'on a $D(t,t) = 1$ pour tout point t de T.

38. Le théorème suivant renforce le théorème classique de Dunford-Pettis [4] en ce qu'il évite toute hypothèse de séparabilité, mais il ne s'applique qu'aux espaces de Hilbert. Il est essentiel que la mesure μ soit σ-finie, comme le montre l'inclusion de L^1 dans L^2 lorsque T est non dénombrable et que tout point de T est mesurable de mesure 1.

THÉORÈME. *Soient $\mathscr{H}$ un espace de Hilbert, M une constante positive et F une application linéaire continue de $L^1 = L^1(T, \mathscr{A}, \mu)$ dans $\mathscr{H}$ telle que l'on ait $\|F(u)\| \leq M.\|u\|_1$ pour tout $u \in L^1$. Il existe alors un sous-espace de Hilbert séparable $\mathscr{H}_0$ de $\mathscr{H}$ et une application mesurable φ de $(T, \mathscr{A})$ dans $\mathscr{H}_0$ muni de sa tribu borélienne avec les propriétés suivantes:*

(a) *on a $\|\varphi(t)\| \leq M$ pour tout point t de T;*
(b) *l'application F prend ses valeurs dans $\mathscr{H}_0$;*
(c) *on a*

$$\langle F(u) | a \rangle = \int_T \overline{u(t)} \langle \varphi(t) | a \rangle \, d\mu(t) \qquad \textit{pour} \quad u \in L^1 \quad \textit{et} \quad a \in \mathscr{H}. \tag{9}$$

Soit L^∞ l'espace des classes de fonctions mesurables et bornées sur $(T, \mathscr{A}, \mu)$; il s'identifie au dual de L^1 de la manière bien connue. Par suite, il existe une application linéaire Φ de $\mathscr{H}$ dans L^∞ qui satisfait aux relations suivantes:

$$|\Phi a(t)| \leq M.\|a\| \qquad \mu\text{-presque partout sur } T, \tag{10}$$

$$\langle F(u) | a \rangle = \int_T \overline{u(t)} \Phi a(t) \, d\mu(t), \tag{11}$$

pour tout $a \in \mathscr{H}$ et tout $u \in L^1$.

Comme la mesure μ est σ-finie, il existe une fonction mesurable h sur $(T, \mathscr{A})$ telle que $h(t) > 0$ pour tout $t \in T$ et que $\int_T h(t)^2 \, d\mu(t)$ soit fini. Définissons une application linéaire j de L^2 dans $\mathscr{H}$ par $j(f) = F(\tilde{h}f)$. Il est clair qu'on a $j^*(a) = \tilde{h}.\Phi a$ pour tout $a \in \mathscr{H}$. D'après l'inégalité (10), on a $|j^*(a)| \leq M.\tilde{h}$ μ-presque partout pour chaque vecteur a de norme ≤ 1 dans $\mathscr{H}$. D'après le théorème du n° 20, l'opérateur j^* est donc de Hilbert-Schmidt, et il en est par suite de même de j. Par ailleurs, l'ensemble des classes de fonctions $\tilde{h}f$, où f parcourt L^2, est dense dans L^1, et l'image d'un opérateur de Hilbert–Schmidt est séparable d'après les résultats rappelés au n° 19. Autrement dit, il existe un sous-espace de Hilbert séparable $\mathscr{H}_0$ de $\mathscr{H}$ contenant l'image de F.

Le reste de la démonstration est classique. Choisissons une base orthonormale $(e_n)_{n \in I}$ de $\mathscr{H}_0$ et des représentants f_n pour les classes Φe_n. Soient J une partie finie de I et $\xi = (\xi_j)_{j \in J}$ une famille de nombres complexes. On a $\Phi(\sum_{j \in J} \xi_j e_j) = \sum_{j \in J} \xi_j \tilde{f}_j$, et compte tenu de (10), il existe un ensemble μ-négligeable N_ξ tel que l'on ait

$$\left| \sum_{j \in J} \xi_j f_j(t) \right|^2 \leq M^2 \sum_{j \in J} |\xi_j|^2 \tag{12}$$

pour tout t dans $T \backslash N_\xi$. Soit N_J l'ensemble μ-négligeable, réunion des ensembles N_ξ pour toutes les familles ξ où chaque ξ_j ait une partie réelle et une partie imaginaire rationnelles. Par continuité, l'inégalité (12) reste satisfaite pour t dans $T \backslash N_J$ et toute famille ξ; on a donc $\sum_{j \in J} |f_j(t)|^2 \leq M^2$ pour tout t dans $T \backslash N_J$. L'ensemble $N = \bigcup_J N_J$ est μ-négligeable, et l'on a $\sum_{n \in I} |f_n(t)|^2 \leq M^2$ pour tout t dans $T \backslash N$. Soit φ l'application de T dans $\mathscr{H}$ qui est nulle en tout point de N et telle que $\varphi(t) = \sum_{n \in I} \overline{f_n(t)} . e_n$ pour tout t dans $T \backslash N$. On a $\|\varphi(t)\| \leq M$ pour tout point t de T, et l'application φ est mesurable de $(T, \mathscr{A})$ dans $\mathscr{H}_0$ muni de sa tribu borélienne (voir par exemple le n° 21).

Il reste à vérifier la formule (9). Elle est vraie par construction lorsque a est l'un des vecteurs e_n, et elle est trivialement satisfaite lorsque a est orthogonal à $\mathscr{H}_0$. Pour u fixée dans L^1, les deux membres de (9) sont des formes linéaires continues du vecteur a de $\mathscr{H}$; le cas général se déduit aussitô de là.

39. D'après un théorème connu de Grothendieck [7], toute forme sequilinéaire continue sur L^1 admet la représentation

$$\Gamma(u, v) = \int_T \int_T C(t, t') \overline{u(t)} v(t') \, d\mu(t) \, d\mu(t'), \tag{13}$$

où C est une fonction mesurable et bornée sur $(T \times T, \mathscr{A} \otimes \mathscr{A})$. Si M est une constante positive, les relations

$$|\Gamma(u, v)| \leq M . \|u\|_1 \|v\|_1 \qquad \text{pour } u, v \text{ dans } L^1 \tag{14}$$

et

$$|C(t, t')| \leq M \qquad \text{presque partout sur } T \times T \text{ pour } \mu \otimes \mu \tag{15}$$

sont équivalentes. Si la forme sesquilinéaire Γ est positive, elle satisfait à l'inégalité de Cauchy–Schwarz $|\Gamma(u, v)|^2 \leq \Gamma(u, u) . \Gamma(v, v)$ et la relation (14) équivaut encore à la suivante

$$\Gamma(u, u) \leq M . \|u\|_1^2 \qquad \text{pour } u \text{ dans } L^1. \tag{16}$$

THÉORÈME. *Soit Γ une forme sesquilinéaire positive et continue sur L^1. Il existe une covariance C séparable, mesurable et bornée sur $(T, \mathscr{A})$ satisfaisant à la relation* (13) *ci-dessus.*

D'après le lemme du n° 3, il existe un espace de Hilbert $\mathscr{H}$ et une application linéaire continue F de L^1 dans $\mathscr{H}$ tels que l'on ait

$$\Gamma(u, v) = \langle F(u) | F(v) \rangle \qquad \text{pour } u, v \text{ dans } L^1. \tag{17}$$

D'après le théorème du n° 38, on peut supposer que l'espace $\mathscr{H}$ est séparable et qu'il existe une application mesurable et bornée φ de $(T, \mathscr{A})$ dans $\mathscr{H}$ satisfaisant à la relation

$$\langle F(u) | a \rangle = \int_T \overline{u(t)} \langle \varphi(t) | a \rangle \, d\mu(t) \tag{18}$$

pour $u \in L^1$ et $a \in \mathscr{H}$.

Notons C la covariance associée à $\mathscr{H}$ et φ; elle est séparable, bornée et mesurable sur $(T \times T, \mathscr{A} \otimes \mathscr{A})$. Soit $\mathscr{F}$ le sous-espace vectoriel fermé de $\mathscr{H}$ engendré par l'image de φ. Comme au n° 5, définissons un isomorphisme H de $\mathscr{F}$ sur $\mathscr{E}_C$ par la relation $Ha(t) = \langle \varphi(t) | a \rangle$ pour $a \in \mathscr{F}$ et $t \in T$. La formule (18) équivaut donc à la suivante

$$\langle HF(u) | f \rangle_C = \int_T \overline{u(t)} f(t) \, d\mu(t) \qquad \text{pour} \quad u \in L^1 \quad \text{et} \quad f \in \mathscr{E}_C. \tag{19}$$

Si l'on prend en particulier pour f une fonction de la forme $C_{t'} : t \mapsto C(t, t')$, on voit que $HF(u)$ n'est autre que la fonction notée Ku au n° 35. On a donc

$$\Gamma(u, v) = \langle F(u) | F(v) \rangle = \langle HF(u) | HF(v) \rangle_C = \langle Ku | Kv \rangle_C \tag{20}$$

pour u, v dans L^1. Enfin si l'on fait $f = Kv$ dans la formule (3) du n° 35, on obtient

$$\langle Ku | Kv \rangle_C = \int_T \int_T C(t, t') \overline{u(t)} v(t') \, d\mu(t) \, d\mu(t') \tag{21}$$

d'après le théorème de Lebesgue–Fubini. Le théorème résulte des formules (20) et (21).

40. Terminons cette partie par un corollaire important du théorème précédent. Soit C une fonction mesurable et bornée sur $(T \times T, \mathscr{A} \otimes \mathscr{A})$. On définit la forme sesquilinéaire Γ sur L^1 par la formule (13). Si A et B sont deux éléments de $\mathscr{A}$, on pose

$$\Gamma[A, B] = \int_A \int_B C(t, t') \, d\mu(t) \, d\mu(t'). \tag{22}$$

COROLLAIRE. *Les conditions suivantes sont équivalentes:*

(a) *La fonction C est égale presque partout* (pour $\mu \otimes \mu$) *à une covariance mesurable et bornée.*

(b) *La fonction C est égale presque partout à une covariance séparable, mesurable et bornée.*

(c) *On a $\Gamma(u, u) \geq 0$ pour toute fonction μ-intégrable u sur T.*

(d) *La matrice d'éléments $\Gamma[A_i, A_j]$ est hermitienne positive pour toute suite $(A_i)_{1 \leq i \leq n}$ d'éléments de $\mathscr{A}$.*

Si C est une covariance, on a $\Gamma(u, u) = \|Ku\|_C^2$ avec les notations du n° 35, donc (a) implique (c). Le théorème du n° 39 exprime que (c) implique (b), et il est clair que (b) implique (a).

Avec les notations de (d), on a

$$\sum_{i,j=1}^{n} \Gamma[A_i, A_j]\bar{\xi}_i\xi_j = \Gamma(u, u) \tag{23}$$

où la fonction u prend la valeur constante ξ_j sur A_j, et est nulle hors de $A_1 \cup \cdots \cup A_n$. Comme les classes de telles fonctions sont denses dans L^1, les conditions (c) et (d) sont équivalentes.

D'après le corollaire précédent, toute covariance mesurable et bornée est égale presque partout à une covariance séparable, mesurable et bornée. On a donc retrouvé un cas particulier du théorème du n° 28.

VII. Une inégalité intégrale

Dans toute cette partie, on note $(T, \mathscr{A}, \mu)$ un espace mesuré σ-fini.

41. Soit C une covariance mesurable sur $(T, \mathscr{A})$. On pose

$$C[u] = \int_T \int_T C(t, t')\overline{u(t)}u(t')\, d\mu(t)\, d\mu(t') \tag{1}$$

pour toute fonction mesurable u sur $(T, \mathscr{A})$ pour laquelle l'intégrale précédente converge absolument. Nous nous proposons de montrer que $C[u]$ *est toujours positif*.

Nous connaissons déjà plusieurs cas particuliers:

(a) Supposons que la mesure μ soit portée par un ensemble fini $A = \{t_1, \ldots, t_n\}$ au sens que $\mu(T \backslash A)$ est nul, et que toute partie de A appartient à $\mathscr{A}$. Posons $\xi_i = u(t_i)\mu(\{t_i\})$ pour $1 \leq i \leq n$. On a alors $C[u] = \sum_{i,j=1}^{n} C(t_i, t_j)\bar{\xi}_i\xi_j$ et l'inégalité $C[u] \geq 0$ exprime la définition d'une covariance.

(b) Au n° 15, nous avons considéré le cas où T est un espace localement compact, muni de sa tribu borélienne $\mathscr{A}$, où la mesure μ donne une masse finie à toute partie compacte de T, où C est continue et u continue à support compact.

(c) Supposons que les intégrales $\int_T C(t,t)\,d\mu(t)$ et $\int_T |u(t)|^2\,d\mu(t)$ soient finies. D'après l'inégalité de Cauchy–Schwarz, l'intégrale $\int_T C(t,t)^{1/2}|u(t)|\,d\mu(t)$ est finie et l'inégalité $|C(t,t')| \leq C(t,t)^{1/2}C(t',t')^{1/2}$ montre que l'intégrale définissant $C[u]$ converge absolument. Utilisons les notations du n° 31; d'après le théorème de Lebesgue–Fubini, on a $C[u] = \langle \tilde{u} | \tilde{K}\tilde{u} \rangle$, et comme l'opérateur $\tilde{K}$ est hermitien positif dans $L^2(T, \mathscr{A}, \mu)$, on a bien $C[u] \geq 0$.

(d) D'après le corollaire du n° 40, on a encore $C[u] \geq 0$ lorsque la covariance C est bornée et que la fonction u est μ-intégrable.

42. Pour démontrer l'inégalité $C[u] \geq 0$, nous ferons une réduction préliminaire.

Comme la mesure μ est σ-finie, il existe une fonction $\mathscr{A}$-mesurable v sur T, à valeurs strictement positives et finies, et une mesure ν sur $\mathscr{A}$ telles que $\nu(T)$ soit fini et que l'on ait $\mu(A) = \int_A v(t)\,d\nu(t)$ pour tout A dans $\mathscr{A}$. Posons

$$D(t,t') = C(t,t')\overline{u(t)v(t)}u(t')v(t') \tag{2}$$

pour t, t' dans T. Il est immédiat que D est une covariance mesurable sur $(T, \mathscr{A})$ et que l'on a

$$C[u] = \int_T \int_T D(t,t')\,d\nu(t)\,d\nu(t'). \tag{3}$$

Par ailleurs, soit T_p l'ensemble des points t de T tels que $D(t,t) \leq p$. Alors T est la réunion de la suite croissante $(T_p)_{p \geq 1}$, et l'on a donc

$$C[u] = \lim_{p \to \infty} \int_{T_p} \int_{T_p} D(t,t')\,d\nu(t)\,d\nu(t'). \tag{4}$$

On est donc ramené à prouver l'inégalité

$$\int_S \int_S D(t,t')\,d\nu(t)\,d\nu(t') \geq 0 \tag{5}$$

où S est l'un des ensembles T_p.

D'après l'inégalité $|D(t,t')|^2 \leq D(t,t) \cdot D(t',t')$, la fonction D est bornée sur $S \times S$; le nombre $c = \nu(S)$ est positif et fini, et l'inégalité (5) est clairement vérifiée lorsque $c = 0$. Supposons donc $c > 0$ et fixons un entier $n \geq 1$. Soient $t_1, \ldots, t_n$ des points de S; comme D est une covariance sur S, on a

$$\sum_{i,j=1}^{n} D(t_i, t_j) \geq 0. \tag{6}$$

Posons

$$a_{ij} = \int_S \cdots \int_S D(t_i, t_j)\,d\nu(t_1) \cdots d\nu(t_n). \tag{7}$$

Le calcul de cette intégrale est immédiat:

$$a_{ii} = c^{n-1} \int_S D(t,t)\,d\nu(t), \tag{8}$$

$$a_{ij} = c^{n-2} \int_S \int_S D(t,t')\,d\nu(t)\,d\nu(t') \qquad \text{si} \quad i \neq j. \tag{9}$$

Intégrons l'inégalité (6) par rapport à $t_1, \ldots, t_n$. Comme il y a n couples (i,j) avec $i = j$ et $n^2 - n$ couples avec $i \neq j$, on trouve aprés division par $c^{n-2}n^2$ l'inégalité

$$\frac{c}{n} \int_S D(t,t)\,d\nu(t) + \frac{n-1}{n} \int_S \int_S D(t,t')\,d\nu(t)\,d\nu(t') \geq 0. \tag{10}$$

L'inégalité (5) résulte de là en faisant tendre n vers l'infini. On remarquera que la réduction au cas où D est bornée est faite pour assurer que l'intégrale $\int_S D(t,t)\,d\nu(t)$ est finie.

43. *Remarques.* (a) La démonstration de l'inégalité (5) est inspirée d'un artifice de Riesz [12], repris par Weil [18] dans l'étude des fonctions de type positif sur un groupe localement compact. On en donnera une "explication" au n° suivant.

(b) La démonstration précédente utilise une propriété moins forte que la définition d'une covariance, à savoir:

La matrice $(C(t_i, t_j))_{1 \leq i,\, j \leq n}$ est hermitienne positive pour presque tout système $(t_1, \ldots, t_n)$ de points de T, par rapport à la mesure $\mu \otimes \cdots \otimes \mu$ (PC)
(n facteurs).

En effet, cette hypothèse suffit pour assurer que l'inégalité (6) a lieu presque partout sur $T \times \cdots \times T$(n facteurs) par rapport à la mesure $\nu \otimes \cdots \otimes \nu$. De plus, l'inégalité $|D(t,t')|^2 \leq D(t,t)\cdot D(t',t')$ a lieu presque partout sur $S \times S$ par rapport à $\nu \otimes \nu$, donc D est égale presque partout sur $S \times S$ à une fonction bornée. Ceci suffit à assurer que les intégrales a_{ij} sont définies. Le reste de la démonstration est sans changement.

La condition (PC) est malheureusement assez peu utilisable. Si C est presque partout (pour $\mu \otimes \mu$) égale à une covariance, on aura encore $C[u] \geq 0$ chaque fois que cette intégrale est définie; mais C ne satisfait pas nécessairement à la condition (PC). Par exemple, avec les notations du n° 25, la fonction $-D$ est égale presque partout (pour la mesure de Lebesgue) à la covariance nulle; elle ne satisfait pas à (PC), puisque l'on a $-D(t,t) = -1$, alors que la condition (PC) exige que l'on ait $C(t,t) \geq 0$ μ-presque partout sur T.

44. Donnons maintenant une *interprétation probabiliste* des résultats précédents. Pour simplifier, on suppose que T est un espace localement compact à base dénombrable et $\mathscr{A}$ sa tribu borélienne, que C est une co-

variance continue et bornée sur T, et u une fonction continue et bornée sur T; on suppose aussi que l'on a $\mu(T) = 1$.

Soit Ω l'espace topologique produit d'une suite de copies T_n de T (pour $n \geq 1$); on munit Ω de sa tribu borélienne $\mathscr{F}$ et de la mesure de probabilité $P = \otimes_{n=1}^{\infty} \mu_n$, avec $\mu_n = \mu$ pour tout entier $n \geq 1$. Etant donné un élément $\omega = (t_n)_{n \geq 1}$ de Ω et un entier $n \geq 1$, on note $\gamma_{n,\omega}$ la mesure de probabilité sur T qui alloue la masse $1/n$ à chacun des points $t_1, \ldots, t_n$ (répétitions permises). On considére la famille $(\gamma_{n,\omega})_{\omega \in \Omega}$ comme une mesure aléatoire γ_n sur T, définie par rapport à l'espace probabilisé $(\Omega, \mathscr{F}, P)$.

Rappelons qu'une suite de mesures de probabilité μ_n sur T tend vaguement vers une mesure de probabilité μ sur T si l'on a

$$\lim_{n \to \infty} \int_T f(t)\, d\mu_n(t) = \int_T f(t)\, d\mu(t) \tag{11}$$

pour toute fonction f continue et bornée sur T.

Une des formes de la *loi forte des grands nombres* affirme que la suite des mesures aléatoires γ_n tend vaguement vers μ avec probabilité 1; autrement dit, il existe un ensemble Ω_1 dans $\mathscr{F}$ tel que $P(\Omega_1) = 1$ et que la suite des mesures $\gamma_{n,\omega}$ tende vaguement vers μ pour tout ω dans Ω_1.

Définissons une suite de variables aléatoires X_n par la formule

$$X_n = \int_T \int_T C(t, t')\overline{u(t)}u(t')\, d\gamma_n(t)\, d\gamma_n(t'). \tag{12}$$

Vu la définition d'une covariance, on a $X_n \geq 0$ [voir le cas (a) du n° 41]. La loi forte des grands nombres montre alors que la suite des variables aléatoires X_n tend avec probabilité 1 vers la constante

$$C[u] = \int_T \int_T C(t, t')\overline{u(t)}u(t')\, d\mu(t)\, d\mu(t').$$

Autrement dit, si l'on pose

$$D(t, t') = C(t, t')\overline{u(t)}u(t') \tag{13}$$

pour t, t' dans T, on a

$$X_n(\omega) = n^{-2} \sum_{i,j=1}^{n} D(t_i, t_j) \qquad \text{pour} \quad \omega = (t_1, t_2, \ldots) \tag{14}$$

et

$$C[u] = \lim_{n \to \infty} X_n(\omega) \qquad \text{pour } \omega \text{ dans } \Omega_1. \tag{15}$$

Comme on a $X_n(\omega) \geq 0$ pour tout ω dans Ω, on a bien $C[u] \geq 0$.

Remarquons que la suite des variables aléatoires X_n est majorée par une constante, et la formule (15) entraine $C[u] = \lim_{n \to \infty} E[X_n]$. Or d'après la formule (14), on a

$$E[X_n] = n^{-2} \int_T \cdots \int_T \sum_{i,j=1}^{n} D(t_i, t_j)\, d\mu(t_1) \cdots d\mu(t_n); \tag{16}$$

on retrouve donc le calcul fait au n° 42.

VIII. Applications et exemples

45. *Théorie des groupes*

Montrons comment l'on retrouve les propriétés élémentaires des fonctions de type positif sur un groupe au moyen des résultats généraux sur les covariances.

Soit G un groupe topologique, d'élément unité e, opérant continuement sur un espace topologique T. On dira qu'une covariance C sur T est *invariante* par G si l'on a $C(g \cdot t, g \cdot t') = C(t, t')$ pour t, t' dans T et g dans G; supposons qu'il en soit ainsi. Si g est un élément de G et f une fonction dans $\mathscr{E}_C$, le lemme du n°5 montre que la fonction $U_g f : t \mapsto f(g^{-1} \cdot t)$ appartient à $\mathscr{E}_C$ et qu'elle a même norme que f dans $\mathscr{E}_C$. Supposons de plus que C soit continue sur $T \times T$; le produit scalaire $\langle C_t | U_g C_{t'} \rangle$ est égal à $C(g^{-1} \cdot t, t') = C(t, g \cdot t')$, donc est fonction continue de g pour t, t' fixés dans T. Rappelons que les combinaisons linéaires finies des fonctions C_t sont denses dans l'espace de Hilbert $\mathscr{E}_C$. On conclut que *l'application $g \mapsto U_g$ est une représentation unitaire continue du groupe G dans l'espace de Hilbert $\mathscr{E}_C$.*

Supposons en particulier que l'on ait $T = G$, le groupe G opérant sur lui-même par translations à gauche. La formule

$$C(g, g') = c(g^{-1} g') \qquad \text{pour } g, g' \text{ dans } G \tag{1}$$

définit une bijection de l'ensemble des covariances invariantes (à gauche) sur G sur l'ensemble des fonctions de type positif sur G. Soit c une fonction continue de type positif sur G; alors $\bar{c} = C_e$ appartient à $\mathscr{E}_C$ et l'on a

$$c(g) = \langle \bar{c} | U_g \bar{c} \rangle_C \qquad \text{pour tout } g \text{ dans } G; \tag{2}$$

ceci précise le résultat connu selon lequel *toute fonction continue de type positif sur G est un coefficient d'une représentation unitaire continue de G.*

Un autre cas intéressant est celui où T est un espace homogène pour G. Choisissons un point t_0 de T, de stabilisateur H; l'application $gH \mapsto g \cdot t_0$ est alors un homéomorphisme de G/H sur T. La formule

$$C(g.t_0, g'.t_0) = c(g^{-1} g') \qquad \text{pour } g, g' \text{ dans } G \tag{3}$$

définit alors une bijection de l'ensemble des covariances invariantes continues sur T sur l'ensemble des fonctions continues de type positif sur G, telles que $c(hgh') = c(g)$ pour $g \in G$, et h, h' dans H. L'élément $a = C_{t_0}$ de $\mathscr{E}_C$ satisfait à $U_h a = a$ pour tout $h \in H$, et l'on a $c(g) = \langle a | U_g a \rangle_C$. Ces remarques sont le point de départ de la théorie des *fonctions sphériques*, qui s'occupe du cas où G est localement compact et H compact.

Le théorème du n° 15 implique le résultat classique suivant: supposons le groupe G localement compact; une fonction continue c sur G est de type positif si et seulement si l'on a l'inégalité

$$\int_G \int_G c(g^{-1}g')\overline{u(g)}u(g')\,dg\,dg' \geq 0 \tag{4}$$

pour toute fonction u sur G, qui est continue à support compact. De même, le théorème du n° 39 est à rapprocher du théorème selon lequel toute forme linéaire continue Φ sur $L^1(G)$ telle que l'on ait $\Phi(u^* * u) \geq 0$ pour tout $u \in L^1(G)$, est de la forme

$$\Phi(u) = \int_G c(g)u(g)\,dg \tag{5}$$

où c est une fonction de type positif sur G. Mais ledit théorème du n° 39 ne suffit pas à assurer que toute fonction de type positif qui est mesurable et bornée est égale presque partout à une fonction continue de type positif.

46. *Un théorème de finitude*

Il s'agit d'un corollaire facile du théorème du n° 20, dont voici l'énoncé:

Soit $(T, \mathscr{A}, \mu)$ un espace mesuré tel que $\mu(T)$ soit fini, de sorte que L^∞ est un sous-espace de L^2. Tout sous-espace vectoriel E de L^∞ qui est fermé dans L^2 est de dimension finie.

La démonstration est la suivante. Si l'on pose $c = \mu(T)^{1/2}$, on a

$$\|f\|_2 \leq c.\|f\|_\infty \qquad \text{pour toute fonction} \quad f \in L^\infty. \tag{6}$$

Si une suite de fonctions $f_n \in E$ converge dans L^∞ vers une fonction f, elle converge donc aussi vers f dans L^2, et l'on a $f \in E$ par hypothése. Autrement dit, l'espace vectoriel E est fermé à la fois dans L^∞ et dans L^2, et il est complet pour les deux normes $\|.\|_2$ et $\|.\|_\infty$. Il résulte alors de l'inégalité (6) et du théorème du graphe fermé que ces deux normes sont équivalentes sur E; il existe donc une constante positive M telle que l'on ait

$$\|f\|_\infty \leq M.\|f\|_2 \qquad \text{pour toute fonction} \quad f \in E. \tag{7}$$

On peut appliquer le théorème du n° 20 en prenant pour h la constante M; par suite, l'inclusion j de E dans L^2 est un opérateur de Hilbert–Schmidt et l'on a $\mathrm{Tr}(j^*j) \leq M^2$. Soit $(f_\alpha)_{\alpha \in A}$ une base orthonormale de E; on a $\|j(f_\alpha)\|_2 = 1$ pour tout $\alpha \in A$, d'où $\sum_{\alpha \in A} 1 = \mathrm{Tr}(j^*j) \leq M^2$. On a prouvé que E est de dimension finie au plus égale à M^2.

Le théorème de finitude précédent a été utilisé dans la théorie des fonctions automorphes. Montrons comment, dans la théorie des équations aux dérivées partielles elliptiques, il permet de *déduire le théorème de finitude du théorème de régularité*. Supposons que T soit une variété compacte de classe C^∞, que $\mathscr{A}$ soit la tribu borélienne sur T, et que la mesure μ ait localement une densité de classe C^∞ (dans un système de coordonnées locales). Soit D un opérateur différentiel elliptique défini sur T, et soit D^* son adjoint; notons E l'espace vectoriel des fonctions f de classe C^∞ telles que $Df = 0$. Le théorème de régularité affirme que E est aussi l'espace des solutions faibles, c'est-à-dire des fonctions $f \in L^2$ telles que $\int_T \overline{f(t)} D^*g(t)\, d\mu(t) = 0$ pour toute fonction g de classe C^∞ sur T. Par suite, E est un sous-espace vectoriel fermé de L^2; toute fonction $f \in E$ est continue sur l'espace compact T, donc bornée, et le théorème de finitude ci-dessus montre que E est de dimension finie.

47. *Processus aléatoires du second ordre*

Soit $(\Omega, \mathscr{F}, P)$ un espace probabilisé, et soit $X = (X_t)_{t \in T}$ une famille de variables aléatoires relative à $(\Omega, \mathscr{F}, P)$. Supposons pour simplifier que l'espérance de chaque X_t soit nulle. On dit que X est un *processus du second ordre* si $E[|X_t|^2]$ est fini pour tout $t \in T$. La *covariance* C du processus X est alors définie par

$$C(t, t') = E[\bar{X}_t X_{t'}]; \tag{8}$$

autrement dit, c'est la covariance associée à l'application $t \mapsto X_t$ de T dans l'espace de Hilbert $L^2(\Omega, \mathscr{F}, P)$.

Soit $\mathscr{H}_X$ le plus petit sous-espace vectoriel fermé de $L^2(\Omega, \mathscr{F}, P)$ contenant les X_t. Les propriétés du second ordre du processus X sont celles qui ne dépendent que de l'espace de Hilbert $\mathscr{H}_X$ et de la famille des éléments X_t de $\mathscr{H}_X$. Par exemple, la continuité du second ordre du processus X signifie la continuité de l'application $t \mapsto X_t$ de T dans $\mathscr{H}_X$; comme on l'a vu au n° 9, ceci équivaut à la continuité de C. Soit $\mathscr{E}_C$ l'espace de Hilbert de fonctions sur T associé à la covariance C. Comme au n° 5, on définit un isomorphisme H de l'espace de Hilbert $\mathscr{H}_X$ sur l'espace de Hilbert $\mathscr{E}_C$ qui associe à une variable aléatoire $Y \in \mathscr{H}_X$ la fonction $t \mapsto E[\bar{X}_t Y]$. Cet isomorphisme transforme X_t en C_t, de sorte que la famille des éléments C_t de l'espace de Hilbert $\mathscr{E}_C$ fournit un "modèle analytique" pour l'étude des propriétés du second ordre du processus X.

Par exemple, supposons que $(T, \mathscr{A}, \mu)$ soit un espace mesuré σ-fini et que la covariance C soit mesurable sur $(T, \mathscr{A})$. Soit u une fonction $\mathscr{A}$-mesurable sur T telle que l'intégrale $\int_T C(t,t)^{1/2} \cdot |u(t)|\, d\mu(t)$ soit finie. En raisonnant comme au n° 35, on montre que la formule

$$Ku(t) = \int_T C(t, t') u(t')\, d\mu(t') \tag{9}$$

définit une fonction $Ku \in \mathscr{E}_C$, et qu'on a

$$\langle Ku|f\rangle_C = \int_T \overline{u(t)} f(t)\, d\mu(t) \qquad \text{pour tout} \quad f \in \mathscr{E}_C. \tag{10}$$

Par l'isomorphisme H de $\mathscr{H}_X$ sur $\mathscr{E}_C$, la fonction Ku correspond à une variable aléatoire $X[u] \in \mathscr{H}_X$; elle est caractérisée par la formule suivante, qui ne fait que traduire la formule (10):

$$E[X[u] \cdot Y] = \int_T u(t) E[X_t \cdot Y]\, d\mu(t) \qquad \text{pour tout} \quad Y \in \mathscr{H}_X. \tag{11}$$

Ce résultat justifie la notation intégrale $\int_T u(t) \cdot X_t\, d\mu(t)$ pour $X[u]$. Prenant $Y = \overline{X[u]}$, on obtient le cas particulier

$$E\left[\left|\int_T u(t) \cdot X_t\, d\mu(t)\right|^2\right] = \int_T \int_T C(t, t') \overline{u(t)} u(t')\, d\mu(t)\, d\mu(t'). \tag{12}$$

L' intégrale au second membre a été notée $C[u]$ précédemment; la formule (12) rend évident le fait qu'elle soit positive. Voici expliqués les mystères des intégrales aléatoires du second ordre!

48. *Covariances séparables et développement de Karhunen* [9]

Avec les notations précédentes, supposons que la covariance C du processus du second ordre X soit *séparable et mesurable* sur $(T, \mathscr{A})$. Interprétons d'abord le théorème du n° 36: supposons qu'il existe une constante positive M telle que l'on ait $|C(t, t')| \leq M$ presque partout sur $T \times T$ (par rapport à $\mu \otimes \mu$); il existe alors un processus du second ordre $X' = (X'_t)_{t \in T}$ tel que l'on ait $E[|X'_t|^2] \leq M$ pour tout $t \in T$, et $X'_t = X_t$ pour presque tout $t \in T$ (par rapport à μ).

Soit $(f_n)_{n \in I}$ une suite de fonctions sur T, dont les conjuguées $\overline{f}_n$ forment une base orthonormale de l'espace de Hilbert $\mathscr{E}_C$. Il lui correspond donc par H une base orthonormale $(Z_n)_{n \in I}$ de $\mathscr{H}_X$ telle que l'on ait

$$f_n(t) = E[\overline{Z}_n X_t] \qquad \text{pour} \quad n \in I \quad \text{et} \quad t \in T. \tag{13}$$

On a donc dans l'espace de Hilbert $\mathscr{H}_X$ le développement orthogonal

$$X_t = \sum_{n \in I} f_n(t) . Z_n \tag{14}$$

généralement attribué à Karhunen [9]. La covariance C se développe comme suit:

$$C(t, t') = \sum_{n \in I} \overline{f_n(t)} . f_n(t'). \tag{15}$$

Les intégrales aléatoires du second ordre admettent le développement en série

$$\int_T u(t).X_t\,d\mu(t) = \sum_{n\in I} c_n Z_n, \tag{16}$$

dont les coefficients sont donnés par les intégrales absolument convergentes

$$c_n = \int_T u(t) f_n(t)\,d\mu(t). \tag{17}$$

Examinons le problème des *versions* du processus X. Par définition, une version de X est une fonction $\hat{X}$ sur $T \times \Omega$ telle que, pour tout $t \in T$, la fonction $\omega \mapsto X(t,\omega)$ soit un représentant de la classe $X_t \in L^2(\Omega,\mathscr{F},P)$. Une version $\hat{X}$ de X est *mesurable* si c'est une fonction mesurable par rapport à la tribu $\mathscr{A} \otimes \mathscr{F}$. Le lemme du n° 24 admet l'interprétation suivante: *si le processus X admet une version mesurable, sa covariance C est séparable et mesurable.* Réciproquement, supposons que la covariance C soit séparable et mesurable et introduisons le développement de Karhunen. Si l'espace $\mathscr{E}_C$ est de dimension finie, on définit une version mesurable du processus X par la formule:

$$\hat{X}(t,\omega) = \sum_{n\in I} f_n(t).\hat{Z}_n(\omega); \tag{18}$$

l'ensemble I est fini, les fonctions f_n sont mesurables sur $(T,\mathscr{A})$, et $\hat{Z}_n$ est un représentant de la classe $Z_n \in L^2(\Omega,\mathscr{F},P)$. Lorsque $\mathscr{E}_C$ est de dimension infinie, la série (18) ne converge pas telle quelle. Supposons pour simplifier que I soit l'ensemble des entiers $n \geq 1$; il existe alors une suite d'entiers $n(1) < n(2) < n(3) < \cdots$ telle que la suite des fonctions

$$X_p(t,\omega) = \sum_{n=1}^{n(p)} f_n(t).\hat{Z}_n(\omega) \tag{19}$$

converge presque partout sur $T \times \Omega$ par rapport à la mesure $\mu \otimes P$. Utilisant le théorème de Lebesgue–Fubini, on conclut qu'il existe un processus du second ordre $X' = (X'_t)_{t\in T}$ admettant une version mesurable, et tel que l'on ait $X'_t = X_t$ pour presque tout $t \in T$ (par rapport à μ).

Supposons le processus X *gaussien.* Alors $(Z_n)_{n\in I}$ est une suite indépendante de variables aléatoires gaussiennes de moyenne nulle et variance 1. Si l'ensemble I est infini, on sait que l'on a $\sum_{n\in I} |\hat{Z}_n(\omega)|^2 = +\infty$ avec probabilité 1, quelles que soient les versions $\hat{Z}_n$ des variables aléatoires Z_n. Par suite, la probabilité qu'une trajectoire appartienne à l'espace $\mathscr{E}_C$ est nulle; il faut choisir pour les trajectoires du processus X un espace fonctionnel plus grand que $\mathscr{E}_C$, dont les éléments seront en général plus "irréguliers" que ceux de $\mathscr{E}_C$. Par exemple, dans le cas du mouvement brownien, toute fonction de

$\mathscr{E}_C$ satisfait à une condition de Hölder d'ordre $\frac{1}{2}$, soit

$$|f(t) - f(t')| \leq M . |t - t'|^{1/2}. \tag{20}$$

Or Wiener [19] a prouvé que le processus du mouvement brownien satisfait à une condition de Hölder d'ordre $\frac{1}{2}$ avec probabilité 0, mais à une condition de Hölder d'ordre $c < \frac{1}{2}$ arbitraire avec probabilité 1.

49. *La covariance du mouvement brownien*

Le mouvement brownien est un processus $X = (X_t)_{t \in \mathbf{R}}$ satisfaisant aux conditions formulées par Einstein en 1905:

(a) *si $t < t'$, la variable aléatoire réelle $X_{t'} - X_t$ est gaussienne, de moyenne 0 et de variance $t' - t$;*

(b) *si $t_0 < t_1 < \cdots < t_n$, les variables aléatoires $Y_i = X_{t_i} - X_{t_{i-1}}$* (pour $1 \leq i \leq n$) *sont indépendantes.*

Les hypothéses précédentes ne concernant que les différences $X_{t'} - X_t$, il faut ajouter une condition de normalisation; nous conviendrons qu'on a $X_0 = 0$.

D'aprés l'hypothèse (a), on a

$$E[(X_t - X_{t'})^2] = |t - t'|; \tag{21}$$

la fonction d_C associée comme au n° 9 à la covariance C de X est donc la distance sur $\mathbf{R}$ définie par $d_C(t, t') = |t - t'|^{1/2}$. La covariance C s'obtient par la relation

$$2C(t, t') = E[X_t^2] + E[X_{t'}^2] - E[(X_t - X_{t'})^2], \tag{22}$$

et l'on a donc

$$C(t, t') = \tfrac{1}{2}[|t| + |t'| - |t - t'|]. \tag{23}$$

On a en particulier $C(t, t') = 0$ si t et t' sont de signes contraires, et l'on a

$$C(t, t') = C(-t, -t') = \inf(t, t') \qquad \text{si} \quad t \geq 0, \quad t' \geq 0. \tag{24}$$

Du point de vue "analytique", pour montrer que C est effectivement une covariance, il suffit de construire un espace de Hilbert $\mathscr{H}$ et une application φ de $\mathbf{R}$ dans $\mathscr{H}$ satisfaisant aux conditions suivantes:

(a′) *le vecteur $\varphi(t) - \varphi(t')$ est de norme $|t - t'|^{1/2}$;*

(b′) *si $t_1 < t_1' < t_2 < t_2'$, les vecteurs $\varphi(t_1') - \varphi(t_1)$ et $\varphi(t_2') - \varphi(t_2)$ sont orthogonaux;*

(c′) *on a $\varphi(0) = 0$.*

Voici une solution: on prend pour $\mathscr{H}$ l'espace $L^2(\mathbf{R}, \lambda)$ où λ est la mesure de Lebesgue; si $t \leq t'$, on note $I_{t,t'}$ l'indicatrice de l'intervalle $]t, t']$, et l'on pose $I_{t',t} = -I_{t,t'}$; enfin, on pose $\varphi(t) = I_{0,t}$. On a alors $I_{t,t'} = \varphi(t') - \varphi(t)$ et la vérification des conditions (a′) à (c′) est immédiate.

L'espace $\mathscr{E}_C$ correspondant à la covariance du mouvement brownien sera noté W_1. D'après la construction générale du n° 5, on définit un isomorphisme H de $L^2(\mathbf{R}, \lambda)$ sur W_1 par la formule $Hu(t) = \langle \varphi(t) | u \rangle$, c'est-à-dire

$$Hu(t) = \begin{cases} \int_0^t u(x)\, dx & \text{si} \quad t \geq 0, \\ -\int_t^0 u(x)\, dx & \text{si} \quad t \leq 0. \end{cases} \tag{25}$$

Par suite, W_1 se compose des fonctions absolument continues f sur $\mathbf{R}$, telles que $f(0) = 0$, et dont la dérivée f' (définie presque partout) est de carré intégrable pour λ; la norme dans W_1 est donnée par

$$\|f\|_1 = \left\{ \int_{\mathbf{R}} |f'(t)|^2\, dt \right\}^{1/2}. \tag{26}$$

On peut définir toute une suite d'espaces de Hilbert W_m analogues. On pose $W_0 = L^2(\mathbf{R}, \lambda)$, et pour tout entier $m \geq 1$, on note W_m l'espace des fonctions f sur $\mathbf{R}$ satisfaisant aux conditions suivantes:

(α) *la fonction f est $m - 1$ fois continument différentiable, et l'on a $f^{(j)}(0) = 0$ pour $0 \leq j \leq m - 1$;*

(β) *la dérivée $f^{(m-1)}$ d'ordre $m - 1$ est absolument continue, et sa dérivée $f^{(m)}$* (définie presque partout) *est de carré intégrable pour λ.*

La norme dans W_m est donnée par

$$\|f\|_m = \left\{ \int_{\mathbf{R}} |f^{(m)}(t)|^2\, dt \right\}^{1/2}. \tag{27}$$

On définit un isomorphisme H_m de $L^2(\mathbf{R}, \lambda)$ sur W_m en généralisant la formule (25):

$$(m-1)!.H_m u(t) = \begin{cases} \int_0^t (t-x)^{m-1} u(x)\, dx & \text{si} \quad t \geq 0, \\ -\int_t^0 (t-x)^{m-1} u(x)\, dx & \text{si} \quad t \leq 0. \end{cases} \tag{28}$$

Grâce à cette formule, on montre facilement que W_m est l'espace $\mathscr{E}_{C_m}$ associé à une covariance continue C_m caractérisée par les propriétés suivantes: on a $C_m(t, t') = 0$ si t et t' sont de signes contraires, on a $C_m(t, t') = C_m(t', t) =$

$C_m(-t,-t') = C_m(-t',-t)$, et enfin, on a

$$C_m(t,t') = \frac{1}{(2m-1)!}\sum_{i=0}^{m-1}(-1)^{m-i+1}\binom{m}{i}t^{2m-i-1}t'^i \tag{29}$$

lorsque $0 \leq t \leq t'$.

50. *Problème de Sturm-Liouville*

Nous considérons un intervalle compact $I = [a,b]$ de $\mathbf{R}$, et deux fonctions mesurables positives p et q sur I; on suppose qu'il existe une constante $c > 0$ telle que l'on ait $p(x) \geq c$ pour tout x dans I. Notons $\mathscr{E}$ l'espace vectoriel des fonctions (complexes) f sur I, nulles en a et b, absolument continues de dérivée f' (définie presque partout), pour lesquelles l'intégrale $\int_I \{p(x)|f'(x)|^2 + q(x)|f(x)|^2\}\,dx$ soit finie. C'est un espace de Hilbert dans lequel le produit scalaire s'exprime ainsi:

$$\langle f\,|\,g\rangle = \int_I \{p(x)\overline{f'(x)}g'(x) + q(x)\overline{f(x)}g(x)\}\,dx. \tag{30}$$

Si une fonction $f \in \mathscr{E}$ satisfait à $\langle f\,|\,f\rangle \leq 1$, on a les inégalités

$$|f(x)| \leq (b-a)^{1/2}/c^{1/2}, \qquad |f(x)-f(y)| \leq |x-y|^{1/2}/c^{1/2} \tag{31}$$

pour x, y dans I. Autrement dit, la boule unité fermée dans l'espace de Hilbert $\mathscr{E}$ est un ensemble équicontinu uniformément borné. D'après le critère général des nos 6 et 11, il existe donc une covariance continue G sur l'intervalle $I = [a,b]$ telle que $\mathscr{E} = \mathscr{E}_G$.

Soit $L^2(I)$ l'espace de Hilbert formé des fonctions sur I de carré intégrable pour la mesure de Lebesgue; comme on a $\int_I G(x,x)\,dx < +\infty$, on peut appliquer les résultats généraux de la partie V. En particulier, l'espace de Hilbert $\mathscr{E}$ est contenu dans $L^2(I)$ et l'application j qui associe à chaque fonction $f \in \mathscr{E}$ sa classe $\tilde{f} \in L^2(I)$ est un opérateur de Hilbert-Schmidt de $\mathscr{E}$ dans L^2. L'opérateur $G = jj^*$ est hermitien, positif, de trace finie dans $L^2(I)$; c'est l'opérateur intégral de noyau G:

$$Gu(x) = \int_I G(x,y)u(y)\,dy. \tag{32}$$

Comme l'espace $\mathscr{E}$ contient les fonctions continument dérivables dans I, nulles en a et b, il est dense dans $L^2(I)$, et l'opérateur G est donc injectif et d'image dense dans $L^2(I)$. Notons D l'inverse de G; le domaine de D se compose des fonctions $f \in \mathscr{E}$ pour lesquelles il existe une fonction $Df \in L^2(I)$ telle que l'on ait

$$\langle f\,|\,g\rangle = \int_I \overline{Df(x)}\cdot g(x)\,dx \qquad \text{pour tout} \quad g \in \mathscr{E}. \tag{33}$$

Compte tenu de la définition (30) du produit scalaire dans $\mathscr{E}$, on voit que le domaine $\mathscr{D}$ de D se compose des fonctions f satisfaisant aux hypothèses suivantes:

(a) on a $f(a) = f(b) = 0$;
(b) la fonction f est absolument continue, de dérivée f' (définie presque partout);
(c) la fonction pf' est absolument continue, et si $(pf')'$ est sa dérivée (définie presque partout), la fonction $-(pf')' + qf$ appartient à $L^2(I)$.

L'opérateur D est donné par la formule:

$$Df = -(pf')' + qf; \tag{34}$$

ce n'est autre que l'opérateur auto-adjoint non borné associé par la méthode de Friedrichs à l' "intégrale d'énergie" (30); la fonction G n'est autre que la *fonction de Green* de l'opérateur différentiel D du second ordre.

En appliquant de manière classique le théorème de Mercer, on obtient le développement de la fonction de Green sous la forme

$$G(x, y) = \sum_{n=1}^{\infty} \lambda_n^{-1} f_n(x) \cdot \overline{f_n(y)}, \tag{35}$$

où les λ_n sont les valeurs propres de D et f_n ses fonctions propres. La suite $(\lambda_n)_{n \geq 1}$ est strictement croissante et tend vers $+\infty$, les fonctions f_n appartiennent à $\mathscr{D}$ et l'on a $Df_n = \lambda_n f_n$ pour tout entier $n \geq 1$, avec la normalisation $\int_I |f_n(x)|^2 \, dx = 1$.

Prenons par exemple $a = 0$, $b = 1$, $p(x) = 1$, $q(x) = 0$ pour tout $x \in [0, 1]$. L'espace $\mathscr{E}$ se compose des fonctions absolument continues f sur $[0, 1]$ telles que $f(0) = f(1) = 0$, et dont la dérivée f' est de carré intégrable sur $[0, 1]$; le produit scalaire dans $\mathscr{E}$ est défini par

$$\langle f | g \rangle = \int_0^1 \overline{f'(x)} g'(x) \, dx. \tag{36}$$

La covariance G se déduit facilement des résultats du n° 49, d'où

$$G(x, y) = \inf(x, y) - xy. \tag{37}$$

L'opérateur D a pour domaine l'espace $\mathscr{D}$ des fonctions f sur $[0, 1]$ telles que $f(0) = f(1) = 0$, qui admettent une dérivée absolument continue f', et dont la dérivée seconde $f'' = (f')'$ soit de carré intégrable sur $[0, 1]$; on a $Df = -f''$. Les valeurs propres et fonctions propres sont données par

$$\lambda_n = \pi^2 n^2, \qquad f_n(x) = 2^{1/2} \sin(\pi n x). \tag{38}$$

Par application de la formule (35), on retrouve le développement connu

$$\inf(x, y) = xy + \sum_{n=1}^{\infty} \frac{2}{(\pi n)^2} \sin(\pi n x)\cdot\sin(\pi n y) \tag{39}$$

pour x, y dans l'intervalle $[0, 1]$.

Compte tenu de la formule (24), on a donc un développement en série de la covariance de la restriction du processus du mouvement brownien à l'intervalle $[0, 1]$. Le développement de Karhunen correspondant s'écrit:

$$X_t = tZ_0 + 2^{1/2} \sum_{n=1}^{\infty} Z_n \cdot \sin(\pi n t)/(\pi n) \qquad \text{pour} \quad 0 \leq t \leq 1. \tag{40}$$

Cette représentation du mouvement brownien comme série de Fourier aléatoire a été introduite par Wiener [19] et souvent utilisée depuis.

51. *Autres Applications*

Nous contenterons d'indiquer quelques thèmes. Tout d'abord, la méthode indiquée au n° 50 pour construire la fonction de Green d'un opérateur différentiel ordinaire se généralise à des intervalles infinis, des conditions aux limites plus générales, et des opérateurs d'ordre plus grand que 2.

Les espaces de Hilbert associés à une covariance se prêtent bien aux problémes d'interpolation et d'approximation. Par exemple, l'espace W_1 est adapté à l'interpolation linéaire, alors que les espaces W_m sont liés aux fonctions-splines. En traduisant en termes de processus du second ordre, on rencontre la théorie de la prédiction linéaire de Wiener–Kolmogoroff . . . .

Remerciements

Ce travail est dédié à Laurent Schwartz, à qui je dois une part importante de ma formation comme mathématicien et comme homme. Les thèmes étudiés ici ont de nombreux points de contact avec les préoccupations mathématiques de L. Schwartz; en particulier, ses séminaires sur la thèse de Grothendieck en 1954, et sur les applications radonifiantes vers 1970, et ses travaux sur les sous-espaces hilbertiens [15, 16] m'ont été l'occasion de m'intéresser à ce sujet, ou d'y revenir.

Le théorème de Mercer généralisé a été mis au point en collaboration avec Xavier Fernique, qui l'a utilisé dans un travail [5] publié en 1970. Je le remercie des fructueuses discussions poursuivies dans le cadre du Séminaire de Probabilités de Strasbourg.

Enfin, Nicolas Bourbaki a fait sans fard de nombreuses remarques pertinentes sur une version préliminaire de ce travail, et m'a encouragé à trouver le "bon cadre" dans lequel exposer la théorie. J'espère que la présentation actuelle est conforme à ses canons mathématiques.

RÉFÉRENCES

1. N. ARONSZAJN, Theory of reproducing kernels, *Trans. Amer. Math. Soc.* **68** (1950), 337–404.
2. N. BOURBAKI, "Espaces vectoriels topologiques," nouvelle éd., Masson, Paris, 1981.
3. P. CARTIER, Classes de formes bilinéaires sur les espaces de Banach [d'après Grothendieck], Séminaire Bourbaki, 13e année 1960, 1961, exp. 211, 14 pp., Benjamin, New York, 1966.
4. N. DUNFORD ET B. J. PETTIS, Linear operations on summable functions, *Trans. Amer. Math. Soc.* **47** (1940), 323–392.
5. X. FERNIQUE, Régularité des processus gaussiens, *Invent. Math.* **12** (1971), 304–320.
6. F. GANTMACHER, "Théorie des matrices," Tome 1, Dunod, Paris, 1966.
7. A. GROTHENDIECK, "Produits tensoriels topologiques et espaces nucléaires", Memoirs Amer. Math. Soc. Vol. 16, 1955.
8. A. GROTHENDIECK, Résumé de la théorie métrique des produits tensoriels topologiques, *Bol. Soc. Mat. Sao Paulo* **8** (1956), 1–79.
9. K. KARHUNEN, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, *Ann. Acad. Sci. Fenn.* no 37 (1947).
10. J.-L. KRIVINE, Constante de Grothendieck et fonctions de type positif sur les sphères, *Adv. in Math.* **31** (1979), 16–30.
11. M. LOEVE, "Probability theory," 2nd ed., Van Nostrand Reinhold, New York, 1960 (voir la section 34 plus particulièrement).
12. F. RIESZ, Über Sätze von Stone und Bochner, *Acta Sci. Math.* (*Szeged*) **6** (1933), 184–198.
13. F. RIESZ ET B. S. NAGY, "Leçons d'analyse fonctionnelle", Akad. Kiado, Budapest, 1953.
14. E. SCHMIDT, Auflösung der allgemeinen linearen Integralgleichung, *Math. Ann.* **64** (1907), 161–174.
15. L. SCHWARTZ, Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants), *J. Analyse Math.* **13** (1964), 115–256.
16. L. SCHWARTZ, Sous-espaces hilbertiens et antinoyaux associés, Séminaire Bourbaki, 14e année 1961, 1962, exp. 238, 18 pages, Benjamin, New York, 1966.
17. G. SZEGÖ, "Orthogonal Polynomials," Colloquium Publications 4th ed., Amer. Math. Soc., Providence, Rhode Island, 1975.
18. A. WEIL, "L'intégration dans les groupes topologiques et ses applications", Actualités Scientifiques et Industrielles, Hermann, Paris, 1940.
19. N. WIENER, Differential space, *in* "Selected Papers," MIT Press, Cambridge, Massachusetts, 1964.