

ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

HARALD CRAMÉR

Décompositions orthogonales de certains processus stochastiques

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 8, série *Mathématiques*, n° 2 (1962), p. 15-21

http://www.numdam.org/item?id=ASCFM_1962__8_2_15_0

© Université de Clermont-Ferrand 2, 1962, tous droits réservés.

L'accès aux archives de la revue « Annales scientifiques de l'Université de Clermont-Ferrand 2 » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

DÉCOMPOSITIONS ORTHOGONALES DE CERTAINS PROCESSUS STOCHASTIQUES

Harald CRAMÉR

Ancien Recteur de l'Université de Stockholm

1 - Soient x, y, \dots des variables complexes aléatoires définies sur un espace de probabilité fixe. Nous supposons que toute variable a une valeur moyenne égale à zéro, et un moment fini de second ordre :

$$E x = 0, \quad E |x|^2 < \infty.$$

Deux variables x et y telles que $E |x - y|^2 = 0$ sont considérées comme identiques.

L'ensemble de toutes les variables aléatoires satisfaisant à ces conditions constitue un espace H de Hilbert, en définissant le produit scalaire et la norme par les relations :

$$(x, y) = E \overline{xy}, \quad \|x\|^2 = E |x|^2.$$

La convergence en norme d'une suite d'éléments x_1, x_2, \dots coïncide alors avec la convergence en moyenne quadratique (m. q.) des variables aléatoires x_1, x_2, \dots

Deux éléments x, y sont orthogonaux, $x \perp y$, si l'on a :

$$(x, y) = E \overline{xy} = 0.$$

Considérons un processus stochastique défini par une famille de variables aléatoires $x(t)$, où t est un paramètre variant d'une manière continue sur tout l'axe réel $(-\infty, \infty)$. Dans un grand nombre d'applications importantes t signifie le temps, et nous emploierons souvent une terminologie qui se réfère à ce cas. Chaque $x(t)$ représentant un point de H , le processus $x(t)$ définit une "courbe" C dans cet espace. Le plus petit sous-espace $H(x) \subset H$ renfermant toute la courbe C sera appelé l'espace du processus $x(t)$. C'est le sous-espace de H engendré par toutes les variables $x(t)$ avec $-\infty < t < +\infty$.

D'autre part, considérons l'arc $C(t)$ de la courbe C formé par tous les points $x(u)$ avec $u \leq t$, et désignons par $H(x, t)$ le plus petit sous-espace de H renfermant $C(t)$. On voit que $H(x, t)$ est le sous-espace engendré par tous les $x(u)$ avec $u \leq t$. Tout élément de $H(x, t)$ est donc, soit une combinaison linéaire d'un nombre fini de variables $x(u_1), \dots, x(u_k)$ avec $u_j \leq t$, soit la limite en m. q. d'une suite convergente de telles combinaisons. Dans un certain sens on peut donc dire que l'espace $H(x, t)$ renferme la totalité de l'information contenue dans le passé et le présent du processus $x(t)$, du point de vue de l'instant t . Evidemment l'espace $H(x, t)$ n'est jamais décroissant quand t croît, et l'union de tous les $H(x, t)$ pour $-\infty < t < +\infty$ coïncide avec $H(x)$. En désignant par $H(x, -\infty)$ l'intersection de tous les $H(x, t)$, on a donc :

$$H(x, -\infty) \subseteq H(x, t) \subseteq H(x).$$

Le sous-espace $H(x, -\infty)$ renferme l'information contenue dans le passé infiniment éloigné du processus $x(t)$. Considérons les deux cas extrêmes :

$$(A) \quad H(x, -\infty) = 0,$$

$$(B) \quad H(x, -\infty) = H(x),$$

où nous avons désigné par O le sous-espace formé par le seul élément zéro de H .

Dans le cas (A), il est évident que le passé infiniment éloigné du processus $x(t)$ ne nous ap-

prend rien sur le développement futur du processus. Toute information contenue dans ce processus doit donc être entrée comme une innovation à un certain moment. Dans ce cas nous dirons que $x(t)$ est un processus parfaitement non-déterministe.

Dans le cas (B), au contraire, toute information relativement au processus est déjà contenue dans son passé infiniment éloigné, et nous dirons alors que $x(t)$ est un processus déterministe.

Dans le cas général, on aura $O \subset H(x, -\infty) \subset H(x)$, et nous dirons que $x(t)$ est non-déterministe, sans être parfaitement non-déterministe.

Deux processus $x(t)$ et $y(t)$ sont dits orthogonaux l'un à l'autre, si nous avons $x(s) \perp y(t)$ pour tous s et t réels.

2 - Soit $x(t)$ un processus donné tel que pour tout t :

$$E x(t) = 0, \quad E |x(t)|^2 < \infty. \quad (1)$$

Nous dirons que la relation :

$$x(t) = u(t) + v(t) \quad (2)$$

définit une décomposition orthogonale du processus $x(t)$ si les processus $u(t)$ et $v(t)$ sont orthogonaux l'un à l'autre, et si l'on a pour tout t :

$$u(t) \in H(x, t), \quad v(t) \in H(x, t).$$

THEOREME 1 - Pour tout processus $x(t)$ satisfaisant à (1), il y a une décomposition orthogonale (2) uniquement déterminée telle que $u(t)$ est parfaitement non-déterministe, tandis que $v(t)$ est déterministe.

Dans mes travaux antérieurs sur ce sujet ⁽¹⁾, j'ai énoncé ce théorème sans donner une démonstration détaillée. Voici comment on peut le démontrer.

Désignons par $P_t x$ la projection orthogonale d'un élément arbitraire de H sur le sous-espace $H(x, t)$, et posons :

$$v(t) = P_{-\infty} x(t), \quad u(t) = x(t) - v(t).$$

D'abord il est évident que les deux composantes $u(t)$ et $v(t)$ appartiennent à $H(x, t)$. En effet, $v(t)$ appartient à $H(x, -\infty)$, qui est un sous-espace de $H(x, t)$, tandis que $u(t) = x(t) - v(t) \in H(x, t)$.

D'autre part, $v(t)$ étant la projection de $x(t)$ sur $H(x, -\infty)$, et $x(t) = u(t) + v(t)$, on voit que $u(t)$ doit être orthogonal à $H(x, -\infty)$ pour tout t , et il s'ensuit que $u(s) \perp v(t)$ pour tout s et t . Les processus $u(t)$ et $v(t)$ sont donc orthogonaux l'un à l'autre.

En désignant par $H(u, t)$ et $H(v, t)$ les sous-espaces de H qui, pour les processus $u(t)$ et $v(t)$, correspondent à $H(x, t)$, il suit des remarques précédentes qu'on a pour tout t :

$$\begin{aligned} H(u, t) &\subseteq H(x, t), \\ H(u, t) &\perp H(x, -\infty), \\ H(v, t) &\subseteq H(x, -\infty). \end{aligned} \quad (3)$$

En faisant tendre t vers $-\infty$ dans les deux premières de ces relations, on voit tout de suite qu'on a :

$$H(u, -\infty) = 0 ;$$

$u(t)$ est donc un processus parfaitement non-déterministe.

Or, la relation $x(t) = u(t) + v(t)$ montre que le sous-espace $H(x, t)$ est contenu dans la somme vectorielle des sous-espaces mutuellement orthogonaux $H(u, t)$ et $H(v, t)$, qui sont eux-mêmes des sous-espaces de $H(x, t)$; donc :

$$H(x, t) = H(u, t) + H(v, t). \quad (4)$$

 (1) Voir : Colloque sur le Calcul des Probabilités, Paris 1959, p.103-112 ; Proceedings of the Fourth Berkeley Symposium 1961, Vol. II, p.57-78 ; Arkiv för Matematik, Stockholm 1961, Vol. 4, Nr 19.

Il s'ensuit en particulier :

$$H(x, -\infty) \subseteq H(u, t) + H(v, t)$$

En combinant ceci avec les deux dernières relations (3), on obtient :

$$H(v, t) = H(x, -\infty)$$

pour tout t , ce qui donne :

$$H(v, -\infty) = H(v) = H(x, -\infty). \quad (5)$$

$v(t)$ est donc un processus déterministe. De (4) et (5) on déduit aussi :

$$H(x, t) = H(u, t) + H(x, -\infty),$$

ce qui montre que $H(u, t)$ est le complément orthogonal de $H(x, -\infty)$ par rapport à $H(x, t)$.

Il reste à démontrer que la décomposition est unique. Supposons donc que :

$$x(t) = u_1(t) + v_1(t)$$

est une décomposition orthogonale satisfaisant aux conditions du théorème, mais différente de celle qu'on vient de considérer. Posons :

$$u_1(t) = u(t) + z(t), \quad (6)$$

$$v_1(t) = v(t) - z(t),$$

$u(t)$ et $v(t)$ étant définis comme ci-dessus. Alors on a, puisque les composantes $u_1(t)$ et $v_1(t)$ doivent appartenir à $H(x, t)$,

$$H(x, t) \subseteq H(u_1, t) + H(v_1, t) \subseteq H(x, t),$$

$$H(x, t) = H(u_1, t) + H(v_1, t) = H(u_1, t) + H(v_1),$$

$$H(v_1) = H(x, -\infty) = H(v).$$

On voit maintenant que $H(u_1, t)$ est le complément orthogonal de $H(x, -\infty)$ par rapport à $H(x, t)$, c'est-à-dire identique à $H(u, t)$. Il suit donc de (6) que nous avons :

$$z(t) \in H(u, t),$$

$$z(t) \in H(v).$$

Or, les sous-espaces $H(u, t)$ et $H(v)$ sont orthogonaux l'un à l'autre. Donc $z(t)$ doit se réduire à zéro pour tout t , et il s'ensuit que :

$$u_1(t) = u(t),$$

$$v_1(t) = v(t).$$

La démonstration du théorème est donc achevée.

3 - Considérons maintenant un processus $x(t)$ qui est parfaitement non-déterministe, de manière que $x(t)$ coïncide avec sa composante non-déterministe $u(t)$, tandis que la composante déterministe $v(t)$ se réduit à zéro. Supposons encore, pour éliminer des cas trop irréguliers, que les limites en m. q.

$$\begin{aligned} x(t+0) &= \text{l. i. m. } x(t+h), \\ x(t-0) &= \text{l. i. m. } x(t-h) \end{aligned} \quad (7)$$

existent pour tout t , quand h tend vers zéro par des valeurs positives.

Dans mon travail dans l'Arkiv f. Matematik cité plus haut j'ai montré que, sous ces conditions, $x(t)$ admet une décomposition orthogonale dans un sens généralisé, que je vais maintenant rappeler.

Le sous-espace $H(x,t)$ défini ci-dessus n'est jamais décroissant quand t croît et l'on a, $x(t)$ étant parfaitement non-déterministe, $H(x, -\infty) = 0$. Les compléments orthogonaux :

$$\Delta H(x,t) = H(x,t + \Delta t) - H(x,t),$$

$$\Delta H(x,u) = H(x,u + \Delta u) - H(x,u)$$

sont orthogonaux l'un à l'autre si les intervalles $(t, t + \Delta t]$ et $(u, u + \Delta u]$ sont disjoints. On voit que le sous-espace $\Delta H(x,t)$ est engendré par les innovations reçues par le processus $x(t)$ pendant l'intervalle de temps $(t, t + \Delta t]$. Si, pour tout $h > 0$, on a :

$$H(x,t+h) - H(x,t-h) \supset 0, \quad (8)$$

le processus reçoit donc effectivement des innovations dans tout intervalle $(t-h, t+h]$, et l'ensemble de toutes les valeurs de t telles que (8) est satisfaite s'appelle le spectre d'innovations du processus $x(t)$.

Si, pour une certaine valeur de t appartenant au spectre d'innovations, le premier membre de (8) ne tend pas vers zéro avec h , nous dirons que t est un point de discontinuité du spectre ; dans le cas contraire, c'est un point de continuité. J'ai démontré (l.c.) que, si les limites (7) existent pour tout t , l'espace $H(x)$ est séparable, et l'ensemble des points de discontinuité du spectre d'innovations est au plus dénombrable.

Soit maintenant z une variable aléatoire quelconque dans $H(x)$, et posons :

$$z(t) = P_t z, \quad (9)$$

P_t désignant, comme plus haut, la projection sur le sous-espace $H(x,t)$. Il suit alors des propriétés de la famille $H(x,t)$ que $z(t)$ est un processus stochastique à accroissements orthogonaux. Dans mon travail cité, j'ai démontré le théorème suivant.

THEOREME 2 - Soit $x(t)$ un processus parfaitement non-déterministe tel que les limites (7) existent pour tout t . Alors il existe un nombre N uniquement déterminé, qui s'appelle la multiplicité spectrale du processus $x(t)$, et qui peut être un entier positif fini, ou bien égal à $+\infty$, ayant les propriétés suivantes. On peut trouver N éléments z_1, \dots, z_N de $H(x)$ tels que les processus correspondants $z_1(t), \dots, z_N(t)$ définis par (9) sont orthogonaux l'un à l'autre, et que l'on a pour tout t :

$$x(t) = \sum_{n=1}^N \int_{-\infty}^t g_n(t,u) dz_n(u), \quad (10)$$

$$H(x,t) = \sum_{n=1}^N H(z_n,t), \quad (11)$$

où les g_n sont des fonctions non aléatoires telles que :

$$\sum_{n=1}^N \int_{-\infty}^t |g_n(t,u)|^2 E |dz_n(u)|^2 < \infty.$$

Enfin, N est le plus petit nombre ayant ces propriétés.

On voit que (10) donne une décomposition du processus donné dans une infinité d'éléments :

$$g_n(t,u) dz_n(u), \quad (12)$$

qui sont tous orthogonaux l'un à l'autre, et qui appartiennent tous à $H(x,t)$. C'est donc une généralisation de la simple décomposition orthogonale $x(t) = u(t) + v(t)$ considérée auparavant.

Chaque élément (12) peut être considéré comme l'effet à l'instant t d'une innovation $dz_n(u)$ entrée dans l'intervalle $(u, u + du)$. Donc (10) donne une représentation linéaire de $x(t)$ au moyen de toutes les innovations appartenant au passé et au présent du processus, du point de vue de l'instant t . Le processus vectoriel :

$$\{z_1(t), \dots, z_N(t)\}$$

peut être regardé comme le processus d'innovations lié au processus donné $x(t)$. On voit que, dans le cas général, le processus d'innovations sera multi-dimensionnel, sa dimensionnalité étant égale à la multiplicité spectrale N de $x(t)$.

4 - Le théorème 2 donne une généralisation de certaines propriétés bien connues des processus stationnaires à une classe plus générale de processus stochastiques. Pour les processus stationnaires, on sait que la multiplicité est toujours égale à l'unité : $N = 1$. Il est donc bien naturel de poser la question s'il existait des processus $x(t)$ satisfaisant aux conditions du théorème 2, ayant une multiplicité N donnée, finie ou infinie. Dans mon travail cité, j'ai répondu à cette question par l'affirmative, en construisant un exemple d'un tel processus.

Cependant, dans cet exemple le spectre d'innovations est purement discontinu, et ne contient en effet qu'un seul point. Il serait d'un intérêt bien plus grand de connaître un exemple d'un processus $x(t)$ à spectre d'innovations partout continu, ayant une multiplicité N donnée.

L'exemple suivant d'un processus $x(t)$ avec ces propriétés a été construit et discuté au cours d'une correspondance entre M. A. N. Kolmogorov et moi. Soit :

$$z_1(t), z_2(t), \dots$$

une suite infinie de processus orthogonaux l'un à l'autre et avec des accroissements orthogonaux, et supposons que l'on a pour tous n et t :

$$E z_n(t) = 0, \quad E |z_n(t)|^2 = \int_{-\infty}^t f(u) du,$$

où $f(u)$ est partout continue, bornée et positive, avec un moment fini de second ordre.

Soit d'autre part A_1, A_2, \dots une suite d'ensembles disjoints de nombres réels, tout A_n étant de mesure positive dans chaque intervalle Δ :

$$\text{Mes} (\Delta \cap A_n) > 0;$$

désignons par $b_n(s)$ la fonction caractéristique de l'ensemble A_n , et posons pour $u \leq t$

$$g_n(t, u) = \int_u^t b_n(s) ds.$$

On trouve alors sans difficulté que la série :

$$x(t) = \sum_1^{\infty} \int_{-\infty}^t g_n(t, u) dz_n(u) = \sum_1^{\infty} \int_{-\infty}^t b_n(u) z_n(u) du \quad (13)$$

converge en m. q. pour tout t , et représente un processus $x(t)$ satisfaisant aux conditions du théorème 2. On a évidemment :

$$H(x, t) \subseteq \sum_1^{\infty} H(z_n, t). \quad (14)$$

D'autre part, on a :

$$\int_{-\infty}^t b_n(u) z_n(u) du = \int_{-\infty}^t b_n(u) dx(u),$$

$$z_n(t) = \lim_{h \rightarrow 0} \frac{\int_{t-h}^t b_n(u) z_n(u) du}{\text{Mes} ([t-h, t] \cap A_n)},$$

la dernière relation ayant lieu dans chaque point de continuité de la fonction $z_n(t)$, c'est-à-dire avec une probabilité égale à l'unité pour chaque t . On tire de là :

$$\begin{aligned} z_n(t) &\in H(x, t), \\ H(z_n, t) &\subseteq H(x, t), \end{aligned}$$

et d'après (14), pour tout t ,

$$H(x, t) = \sum_1^{\infty} H(z_n, t)$$

Pour les compléments orthogonaux correspondant à un intervalle $(t, t + \Delta t]$, on obtient :

$$\Delta H(x, t) = \sum_1^{\infty} \Delta H(z_n, t),$$

et on voit que le spectre d'innovations de $x(t)$ est partout continu, et contient tout l'axe réel.

Nous allons montrer explicitement que $x(t)$ a une multiplicité $N = \infty$. Supposons, en effet, que N a une valeur finie, et que :

$$x(t) = \sum_{n=1}^N \int_{-\infty}^t h_n(t, u) dy_n(u) \quad (15)$$

est pour $x(t)$ l'analogie de la représentation (10) du théorème 2. Alors les $y_n(u)$ sont des processus orthogonaux l'un à l'autre, et à accroissements orthogonaux, tels que

$$H(x, u) = \sum_{n=1}^N H(y_n, u) = \sum_{n=1}^{\infty} H(z_n, u),$$

et pour un intervalle $(u, u + \Delta u)$:

$$\Delta H(x, u) = \sum_{n=1}^N \Delta H(y_n, u) = \sum_{n=1}^{\infty} \Delta H(z_n, u).$$

On peut donc écrire :

$$\begin{aligned} E |dy_n(u)|^2 &= dJ_n(u), \\ E dy_m(u) d\overline{z_n(v)} &= \begin{cases} 0 & \text{pour } u \neq v, \\ dK_{mn}(u) & \text{pour } u = v, \end{cases} \end{aligned}$$

J_n étant non décroissante, tandis que K_{mn} est à variation bornée dans tout intervalle fini. On a alors pour $u < u + \Delta u < t < t + \Delta t$:

$$\sum_{m=1}^N \int_u^{u+\Delta u} \Delta_t h_m(t, u) dy_m(u) = \sum_{m=1}^{\infty} \int_t^{t+\Delta t} b_m(s) ds \cdot \int_u^{u+\Delta u} dz_m(u), \quad (16)$$

où nous avons écrit :

$$\Delta_t h_m(t, u) = h_m(t + \Delta t, u) - h_m(t, u).$$

Il s'ensuit :

$$\sum_{m=1}^N \int_u^{u+\Delta u} |\Delta_t h_m(t, u)|^2 dJ_m(u) = \int_u^{u+\Delta u} f(u) du \cdot \sum_{m=1}^{\infty} \left(\int_t^{t+\Delta t} b_m(s) ds \right)^2.$$

Cette identité ayant lieu pour tout intervalle $(u, u + \Delta u)$ tel que $u < u + \Delta u < t$, on voit que la contribution de la composante non absolument continue de $J_m(u)$ - s'il y en a une - à la valeur du premier membre doit se réduire à zéro, de manière qu'on aura, presque partout en u :

$$\sum_{m=1}^N |\Delta_t h_m(t, u)|^2 J'_m(u) = f(u) \sum_{m=1}^{\infty} \left(\int_t^{t+\Delta t} b_m(s) ds \right)^2.$$

En multipliant (16) par $\int_u^{u+\Delta u} \overline{dz_n(u)}$, on obtient :

$$\sum_{m=1}^N \int_u^{u+\Delta u} \Delta_t h_m(t, u) dK_{mn}(u) = \int_t^{t+\Delta t} b_n(s) ds \cdot \int_u^{u+\Delta u} f(u) du.$$

et par un raisonnement analogue à celui de tout à l'heure :

$$\sum_{m=1}^N \Delta_t h_m(t, u) K'_{mn}(u) = f(u) \int_t^{t+\Delta t} b_n(s) ds \quad (17)$$

presque partout en u. En donnant N + 1 intervalles disjoints $(t_j, t_j + \Delta t_j)$, on peut toujours supposer les ensembles A_1, A_2, \dots choisis tels que la matrice :

$$\left\{ \int_{t_j}^{t_j + \Delta t_j} b_n(s) ds \right\} \quad (n = 1, 2, \dots; j = 1, 2, \dots, N + 1)$$

aura le rang N + 1. Or, la matrice dont les éléments sont donnés par le premier membre de l'identité (17), en remplaçant successivement t par t_1, \dots, t_{N+1} et en prenant $n = 1, 2, \dots$, sera au plus du rang N. Il y a donc ici une contradiction, et la représentation (15) ne peut pas avoir lieu. Le processus $x(t)$ a donc une multiplicité infinie.

En modifiant cet exemple d'une manière évidente, on peut construire un processus ayant une multiplicité finie N donnée d'avance.

DISCUSSION

M. FORTET - A quelles décompositions corrélatives de la fonction de covariance correspondent les décompositions indiquées pour la fonction aléatoire ?

M. CRAMER - Dans le cas stationnaire les composantes non déterministes et déterministes correspondent à des composantes déterminées du spectre de la fonction de covariance.

Pour les classes de processus harmonisable il y a correspondance incomplète entre les composantes du spectre de la fonction de covariance et cette décomposition.

M. FORTET - Quel est le mode de convergence utilisé lorsque $N \rightarrow +\infty$?

M. CRAMER - Si $N \rightarrow +\infty$ la convergence est définie en moyenne quadratique, la condition de convergence est :

$$\sum_1^{\infty} \int_{-\infty}^t |g_n(t, u)|^2 dF_n(u) < \infty$$

Lorsque F_n est définie par :

$$E |dz_n(u)|^2 = dF_n(u)$$

M. ITO - Can you use the same method to construct an example which needs a countable number of orthogonal processes ?

M. CRAMER - Yes, by the same method. I treated a general example in my paper.

M. BASS - Dans la représentation :

$$X(t) = \sum_{n=1}^N \int_{-\infty}^t g_n(t, u) dZ_n(u)$$

Les propriétés d'orthogonalité de $Z_n(u)$ relativement à u et à n présentent de l'analogie. Y-a-t-il de l'intérêt à remplacer n par une variable continue, et $Z_n(u)$ par un accroissement correspondant à cette variable ?