

# COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 26 au 30 AVRIL 1977

---

Métriques de Sobolev et fonctions "spline" en traitement du signal

Sobolev metrics and spline functions in time data analysis

Paul KRÉE et Anne-Marie MAZEROLLES  
Paul KRÉE and Anne-Marie MAZEROLLES

Département de Mathématiques-Université de Paris VI  
et de l'Université Paul Sabatier à Toulouse

Department of mathematics University of Paris VI  
and Paul Sabatier University Toulouse

---

## RESUME

L'emploi de l'analyse fonctionnelle, et en particulier des espaces de Sobolev s'est révélée très utile en mécanique stochastique : voir quelques travaux récents concernant l'action du vent, du séisme, de la houle sur les structures ; et aussi la sécurité des constructions.

Les techniques usuelles de traitement du signal utilisent des espaces hilbertiens  $L^2$  et la métrique correspondante : voir par exemple les applications en détection et en économétrie du théorème de Karhunen Loeve. Ceci conduit à des techniques numériques très simples mais ne tient pas compte ni de la régularité des trajectoires, ni de l'ordre chronologique des données. En analyse fonctionnelle en composantes principales et en détection, des métriques de Sobolev peuvent être utilisées. Le travail numérique n'est pas terminé ; un seul exemple est présenté ; mais quelques enseignements pratiques ont été dégagés.

## SUMMARY

Functional analysis, and in particular use of Sobolev spaces, is very useful in some problems of stochastic mechanics : see for example recent works in the domain of structural safety.

Usual techniques of time data analysis are concerned with the Hilbert spaces  $L^2$ , and the corresponding metric : see for example application in detection or in econometric of Karhunen Loeve theorem. This leads to simple techniques but take no account of the regularity of trajectories and of the chronological order of the data.

For principal component functional analysis and detection, metrics induced by Sobolev spaces can be used. Numerical techniques using spline function can be used. Computation work is not achieved ; only a medical problem as been treated.



Les espaces fonctionnels, et en particulier les espaces de Sobolev peuvent être utilisés en statistique des processus stochastiques, donc en traitement du signal (T.S.). Considérons par exemple la mécanique stochastique ; les constructions sont soumises à des actions dynamiques, telles que vent, séisme, houle... Ces signaux aléatoires sont transformés par les structures, en un champ de contraintes aléatoires variables. On cherche alors à évaluer la probabilité de dépassement par ces contraintes, pendant la période d'utilisation prévue, de certains niveaux admissibles, car ceci donne une indication sur la probabilité de ruine de la structure. Dans ce domaine, la considération d'espaces fonctionnels a permis de résoudre les problèmes suivants :

a) Résolution théorique et numérique de l'équation aux dérivées partielles hyperbolique et stochastique régissant la flexion d'une poutre encastrée d'inertie et de raideur variables, soumis à un champ stochastique gaussien de pressions: c'est le schéma pour l'étude des charges dynamiques dans un immeuble de grande hauteur, soumis à une rafale de vent [1].

b) Etude théorique de l'effet dans ce schéma de la non linéarité apparaissant dans la transformation des vitesses en pressions [12].

c) Evaluation de la probabilité de ruine de structures réelles en tenant compte des charges dynamiques, des probabilités de simultanéité des diverses actions. Le but est d'aider à l'amélioration des règlements de construction [13].

- L'étude du filtrage non linéaire, comme dans les plateformes "offshore" est en cours.
- Le but de cet exposé est d'introduire à ces méthodes, et de les illustrer par une application : il s'agit de la description des proximités entre un grand nombre de fonctions, dont on connaît seulement des échantillonnages.

#### 4. Langage de l'analyse fonctionnelle pour l'étude des processus stochastique.

Les recherches évoquées ci-dessus sont effectuées dans un langage mathématique qui est introduit progressivement ci-après, en partant du langage simple de la statistique.

1.A. La statistique usuelle se traduit mathématiquement par des problèmes d'analyse et de théorie de

la mesure sur des espaces vectoriels (e.v.) de dimension finie. En effet si l'on a  $p$  variables aléatoires  $\xi_1, \dots, \xi_p$  de loi conjointe  $P$ , on ne perd rien en prenant comme espace probabilisé  $\mathbb{R}^p$  muni de la probabilité  $P$ . Pour  $i=1 \dots p$ , la v.a.  $\xi_i$  est représentée par la  $i^{\text{e}}$  projection canonique

$$(1) \quad \mathbb{R}^p \quad x = (x_1, \dots, x_p) \longmapsto x_i \in \mathbb{R} .$$

Pratiquement cela signifie que si l'on fait  $N$  expériences indépendantes, on tire au hasard  $N$  points de  $\mathbb{R}^p$ , en suivant la loi  $P$ .

- la moyenne est le vecteur  $\bar{\xi}$  de composantes

$$\bar{\xi}_i = \mathcal{E}(\xi_i).$$

- On peut centrer les variables  $\xi_i$ , c'est-à-dire placer l'origine de  $\mathbb{R}^p$  au point moyen  $\bar{\xi}$ .

- Si  $\bar{\xi} = (\bar{\xi}_1, \dots, \bar{\xi}_p)$  est centré, la matrice de corrélation  $[C_{\bar{\xi}}]$  a pour terme générique

$$(2) \quad [C_{\bar{\xi}}]_{ij} = \mathcal{E}(\xi_i \xi_j).$$

On sait comment construire à partir d'un échantillon de taille  $N$  relatif à  $\bar{\xi}$ , des approximations de la moyenne et de la corrélation

(3) Pour bien comprendre ce qui suit, il est commode de donner des définitions géométriques de moyenne et corrélation, c'est-à-dire sans faire intervenir de base particulière dans l'e.v.  $E = \mathbb{R}^p$ . On introduit le dual  $E'$  de  $E$ . Tout élément  $\ell$  de  $E'$ , autrement dit toute forme linéaire  $\ell$  sur  $E$  est interprété comme une v.a. sur  $E$ . On rappelle que le dual de  $E'$  est  $E$ .

(3') Alors  $\bar{\xi}$  est tel que  $\ell(\bar{\xi}) = \mathcal{E}(\ell)$  pour toute  $\ell \in E'$ . Sur  $E'$  on a une forme bilinéaire symétrique  $C_{\bar{\xi}}$  :

$$(4) \quad (\ell_1, \ell_2) \longmapsto C_{\bar{\xi}}(\ell_1, \ell_2) = \mathcal{E}(\ell_1 \ell_2).$$

(5) Alors  $\ell \longmapsto C(\ell, \ell)$  est une forme quadratique positive sur  $E'$ . Et il existe un seul opérateur linéaire  $L$  de  $E'$  dans  $E$  tel que

$$\forall \ell_1, \ell_2 \in E' \quad \langle L \ell_1, \ell_2 \rangle = C_{\bar{\xi}}(\ell_1, \ell_2)$$

(6) Cet opérateur est encore noté  $C_{\bar{\xi}}$ . Autrement dit on identifie les trois ensembles suivants : les formes bilinéaires symétriques sur  $E'$ , les formes quadratiques sur  $E'$ , et les opérateurs linéaires de  $E'$  dans  $E$ .

(7) L'analyse en composantes principales (ACP) popularisée en France par certains hebdomadaires, a pour but de décrire un ensemble  $(\xi^1, \dots, \xi^N)$  de points de  $E$ , que l'on considère comme le résultat d'un échantillon de taille  $N$  relatif à une même v.a. vectoriel  $\xi$  du second ordre, à valeurs dans  $E$ . Dans cette analyse,  $E$  est muni d'une distance euclidienne, ce qui permet d'identifier  $E'$  et  $E$ . On centre d'abord  $\xi$ . La théorie spectrale montre qu'il existe une base orthonormée  $(V_1, \dots, V_p)$  formée par des vecteurs propres de l'opérateur  $C_\xi$ ; les valeurs propres correspondantes sont notées  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ . On note  $P_i$  le projecteur orthogonal sur  $V_i$ . Alors les variables  $P_i \xi$  sont telles que

$$(8) \mathbb{E}((P_i \xi) \cdot (P_j \xi)) = \langle C_\xi P_i \xi, P_j \xi \rangle = \lambda_i \lambda_j \langle V_i, V_j \rangle = \lambda_i \delta_{ij}$$

D'où 
$$\xi = \sum_{\lambda_i \neq 0} \sqrt{\lambda_i} \cdot \frac{P_i \xi}{\sqrt{\lambda_i}}$$

les variables  $\lambda_i^{-1/2} P_i \xi$  formant un système orthogonal. Si l'on remplace  $\xi$  par les deux premiers termes de la somme, on commet une erreur de variance  $\sum_{i \geq 2} \lambda_i$ . Par conséquent, parmi les projecteurs orthogonaux de rang deux de  $E$ , le meilleur est celui qui projette  $E$  sur le sous espace engendré par  $V_1$  et  $V_2$ .

1.B. Souvent les ingénieurs spécialistes de T.S. peuvent éviter dans leurs raisonnements une étude fine relative à l'espace probabilisé, espace qui est maintenant de dimension infinie. Malheureusement, ceci n'est pas toujours possible; c'est par exemple le cas en théorie de la détection. Il faut alors considérer de tels problèmes comme des problèmes d'analyse et de théorie de la mesure relatifs à des e.v. de dimension infinie; c'est le point de vue de la thèse d'Edith Mourier, et qui fait actuellement l'objet de nombreux travaux mathématiques.

Pratiquement, cela signifie que l'on prend pour espace probabilisé un espace localement convexe séparé complet  $E$  (de Banach en général), dont les éléments sont des fonctions; on dit que  $E$  est un espace fonctionnel. La tribu borélienne de cet espace est muni d'une loi de probabilité  $P$  (ou de plusieurs, comme par exemple en théorie de la détection). Les trajectoires du processus étudié s'obtient en tirant au hasard un point  $w$  de  $E$ . Exemple: Si  $E$  est l'espace  $C(\mathbb{R})$  des fonctions

continues sur  $\mathbb{R}$ , et si  $P$  est la mesure de Wiener, on obtient ainsi le mouvement brownien. Ainsi on peut reprendre sans modification les points (3) et (5) qui précèdent, en introduisant l'ensemble  $E'$  des formes linéaires continues sur  $E$ , et en le munissant de la topologie faible.

- Soit  $L^2(E)$  l'espace de Hilbert des variables aléatoires du second ordre sur  $E$ . On suppose que toute  $\ell \in E'$  est dans  $L^2(E)$  et que l'application  $\ell \in E' \rightarrow \ell \in L^2(E)$  est faiblement continue.
- La trajectoire moyenne  $\xi$  du processus est encore définie par (3). On peut centrer le processus en prenant pour origine de  $E$  le point  $\xi$ .
- Sur  $E'$ , on peut encore définir par (4) la forme bilinéaire symétrique  $C_\xi$ . Les identifications (6) sont encore possibles; d'où un opérateur de corrélation  $C_\xi$  de  $E'$  dans  $E$ , et une forme quadratique positive  $\ell \rightarrow \langle C_\xi \ell, \ell \rangle$  sur  $E'$ .
- Par la suite, on suppose  $C_\xi$  injectif, ce qui signifie que cette forme quadratique ne s'annule que pour  $\ell = 0$ . Elle définit un produit scalaire sur  $E'$ . On a donc une injection  $j$  de  $E'$  dans l'espace de Hilbert  $H$  obtenu en complétant  $E'$  pour le produit scalaire défini par  $C_\xi$ . En identifiant  $H$  à son dual, et en transposant  $j$ , on obtient le schéma.

$$(9) \quad E' \xleftarrow{j} H \cong H' \xleftarrow{j'} E$$

L'espace  $H$  est l'espace des v.a. linéaires du second ordre sur  $E$ ; car son produit scalaire coïncide avec celui de  $L^2(E)$ . L'espace  $H'$  est appelé l'espace reproduisant associé au processus étudié. Il coïncide avec l'espace introduit par E. Parzen [10] en utilisant d'autres méthodes.

(10) Exemple

Pour le mouvement brownien, on trouve pour  $H'$  l'espace (de Sobolev!) des fonctions  $\varphi$  continues sur  $\mathbb{R}$ , nulles à l'origine, à dérivée de carré sommable: le produit scalaire de  $H'$  étant

$$\langle \varphi, \psi \rangle = \int_{-\infty}^{+\infty} \varphi'(x) \psi'(x) dx$$

On pourrait penser à priori qu'il ne fallait pas commencer le raisonnement avec l'espace  $E$ , mais plutôt avec l'espace  $H'$  muni d'une probabilité  $P'$  telle que  $j'(P') = P$ . Ceci est faux car  $P'$  n'existe pas en général:  $P'$  est seulement une fonction additive sur l'algèbre de Boole  $A$  des



cylindres de  $H'$ , qui ne se prolonge pas en une fonction  $\sigma$ -additive (i.e. une mesure) sur la tribu engendrée par  $A$ .

1.C. Ceci conduit à considérer une situation encore plus générale.

Soit  $E$  un espace fonctionnel et soit  $E'$  son dual topologique comme ci-dessus. Un processus linéaire  $R$  (de type faible deux) basé sur  $E'$  est une application linéaire faiblement continue de  $E'$  dans un espace de v.a. du second ordre noté  $L^2(\Omega)$ . La moyenne de  $R$  est le vecteur  $\underline{R}$  de  $\mathcal{L}$  tel que

$$(11) \quad \forall e \in \mathcal{L} \quad e(\underline{R}) = \mathcal{E}(R e).$$

On peut centre  $R$

Si  $R$  est centré la corrélation de  $R$  est la forme bilinéaire suivante sur  $E'$

$$(12) \quad e_1, e_2 \longrightarrow \mathcal{E}(R e_1 \cdot R e_2).$$

On dit que  $R$  est décomposé par une v.a.  $\xi$  s'il existe une v.a.  $\xi$  à valeurs dans  $E$  telle que  $R(e) = e \circ \xi$  pour tout  $e \in E'$ . Il en est toujours ainsi en dimension finie.

Soit  $F-F'$  un autre couple d'e.v. analogue à  $E-E'$ . Si  $A$  est un opérateur linéaire de  $F'$  dans  $E'$  à domaine dense, le composé de  $R$  et de  $A$  est défini si  $RA$  admet un prolongement linéaire faiblement continu de  $F'$  dans  $L^2(\mathcal{L})$ .

La théorie permet alors, pour un processus donné, de trouver un espace de Hilbert  $H$  contenant presque sûrement les trajectoires : voir par exemple [7].

§.2. Une analyse en composantes principales

Par un simple exercice d'analyse mathématique, les considérations qui précèdent permettent de reformuler la théorie usuelle de l'ACP :

(13) Théorème : (voir [2] dans le cas particulier où  $H = L^2(\mathcal{L})$  et  $R$  décomposé). Soient  $H$  et  $X$  deux espaces de Hilbert réels et soit  $R$  un opérateur linéaire continu de  $H$  dans  $X$ , d'opérateur de corrélation  $C_R$ . Soit  $M$  la mesure spectrale de  $C_R$  définie sur le spectre  $S$  de  $C_R$ .

- a) Pour tout  $x \in H$ ,  $R(dM(\lambda)x)$  est une mesure spectrale  $S \longrightarrow X$  telle que

$$(14) \quad R x = \int_S R(dP(\lambda)x), \text{ pour tout } x \text{ de } H$$

- b) Quelles que soient les fonctions boréliennes bornées  $\varphi$  et  $\psi$  sur  $S$

$$(15) \quad \int_S \varphi(u) R(dM(u)x), \int_S \psi(v) R(dM(v)x) \overline{x} = \int_S \lambda \varphi(\lambda) \psi(\lambda) \langle dM(\lambda)x, x \rangle$$

Par exemple si  $C_R$  est compact de spectre  $S = \{\lambda_1, \lambda_2, \dots\}$ .

$$R x = \sum R(P_{\lambda_i} x)$$

$$\langle R(P_{\lambda_i} x), R(P_{\lambda_j} x) \rangle = \lambda_i \delta_{ij}$$

On retrouve donc l'ACP usuelle.

Mais l'on retrouve aussi la décomposition de Karhunen-Loeve. En effet soit  $(X_t)$  un processus à valeurs dans  $\mathbb{R}^k$  paramétré dans un ensemble  $T$  muni d'une mesure  $dt$ , tel que

$$\iint_{\mathcal{L} \times T} \|X_t\|^2 dt \quad dP(\omega) < \infty$$

On peut représenter  $(X_t)$  par le processus linéaire

$$H=L^2(T) \longrightarrow X=L^2(\mathcal{L}, \mathbb{R}^k)$$

$$\varphi \longmapsto R\varphi = \int_T X_t \varphi(t) dt$$

L'opérateur  $C_R$  a pour noyau  $U(t, t') = \mathcal{E}(\langle X_t, X_{t'} \rangle)$ .

Il admet la décomposition spectrale suivante

$$C_R = \sum_{j=1}^{\infty} \lambda_j \psi_j \otimes \psi_j$$

où les  $\psi_j$  forment un système orthonormal de  $L^2(T)$ .

$$D'où \quad R\varphi = \sum \frac{R\psi_j}{\sqrt{\lambda_j}} (\varphi, \psi_j) \sqrt{\lambda_j}$$

$$(16) \text{ soit pour presque tout } t \quad X_t = \sum_j \frac{R\psi_j}{\sqrt{\lambda_j}} \psi_j(t) \sqrt{\lambda_j}$$

§.3. Pratique de l'ACP fonctionnelle

Ce paragraphe est la contribution de A.M. Mazerolles [9].

Fonctions spline et espaces de Sobolev.

Soit  $D$  un domaine régulier de  $\mathbb{R}^d$  et soit  $s$  un entier positif. L'espace de Banach  $C^s(D)$  des fonctions  $s$  fois continuellement dérivables sur  $D$  n'est pas pratique en calcul des variations, car des fonctionnelles très naturelles n'y atteignent pas leur minimum. De plus  $C^s(D)$  n'est pas hilbertien. L'espace de Sobolev  $H^s(D)$  est l'espace des  $\varphi \in L^2(D)$  dont les dérivées distribution, jusqu'à l'ordre  $s$ , appartiennent à  $L^2(D)$ . C'est un espace de Hilbert pour la norme

Métriques de Sobolev et fonctions "spline" en traitement du signal

Sobolev metrics and spline functions in time date analysis

$$(17) \|\varphi\|_s = \left( \sum_{|\alpha| \leq s} a_\alpha \int_D \left| \left( \frac{\partial}{\partial x} \right)^\alpha \varphi(x) \right|^2 dx \right)^{1/2}$$

où les  $a_\alpha$  sont des constantes positives fixées.

(18) Les espaces de Sobolev sont excellents en calcul des variations. Par exemple ([5] [8]) pour un intervalle  $I=D$  de  $\mathbb{R}^d = \mathbb{R}$ , parmi toutes les  $\varphi$  de  $H^s(I)$ , celle qui minimise  $\int_I \varphi^{(s)}(t)^2 dt$  en prenant des valeurs données  $a_i$  en des points donnés  $t_i$  de  $T$ , est une fonction polynomiale de degré au plus  $2s-1$  dans chaque intervalle  $[t_i, t_{i+1}]$ ; c'est une fonction spline.

(19) Remarque.

On définit aussi l'espace de Sobolev vectoriel  $H^s(I, \mathbb{R}^q)$ . C'est l'espace des  $\vec{\varphi} = (\varphi_1, \dots, \varphi_q)$  avec  $\varphi_j \in H^s(I)$ , muni de la norme

$$\|\vec{\varphi}\| = \left( \sum_{j=1}^q b_j \|\varphi_j\|_s^2 \right)^{1/2}$$

où les  $b_j$  sont des constantes fixées. On définit de même des fonctions "spline" vectorielles, chacune de leurs composantes se calculant indépendamment des autres, à l'aide des algorithmes de [5] relatifs aux fonctions "spline" scalaires".

(20) Le problème pratique de l'ACP fonctionnelle.

On se donne  $N$  courbes  $t \rightarrow \xi^i(t)$  de  $\mathbb{R}^q$ , paramétrées dans  $I=[0,1]$ . On se propose de décrire les proximités entre ces courbes. On interprète ces données comme un échantillon de taille  $N$  relatif à une fonction aléatoire  $(\xi_t)_t$ . On suppose usuellement en détection que  $H = L^2(I, \mathbb{R}^q)$ . C'est aussi ce que suppose J.C. Deville [3] lorsqu'il cherche à décrire des proximités entre données économétriques. Pratiquement, on prend un pas de discrétisation de  $1/n$ , et l'on représente  $(\xi_t)_t$  par les variables aléatoires  $\xi(j) = \xi_{j/n}$  et l'on est ramené à l'ACP usuelle. Noter qu'il faut supposer les trajectoires de  $(\xi_t)_t$  p.s. continues pour pouvoir définir les  $\xi(j)$ . Comme on cherche une proximité entre courbes paramétrées et non entre familles de v.a., il faut se donner un procédé (appelé ici après relèvement) associant une courbe paramétrée à toute suite de nombres  $\xi(j)$ .

(21) J.C. Deville propose

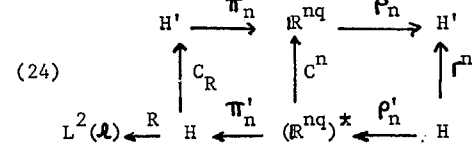
- le relèvement constant par intervalle.
- le relèvement affine par intervalle.
- le relèvement par polynôme d'interpolation de

Lagrange.

(22) A.M. Mazerolles introduit les concepts suivants pour réaliser pratiquement l'ACP pour les données de (13).

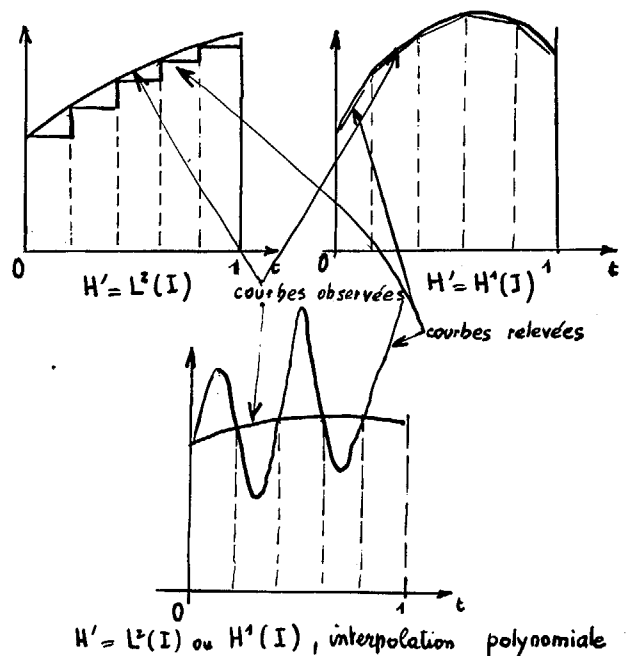
Soit  $\pi_n$  un opérateur linéaire continu surjectif de  $H$  sur  $\mathbb{R}^{nq}$  appelé opérateur d'échantillonnage. Soit  $\rho_n$  un inverse à droite de  $\pi_n$ . C'est un opérateur linéaire continu de  $\mathbb{R}^{nq}$  dans  $H$  tel que  $\pi_n \rho_n y = y$  pour tout  $y \in \mathbb{R}^{nq}$ . Un tel opérateur étant forcément injectif,  $\mathbb{R}^{nq}$  se trouve identifié à un sous espace de  $H$ ;  $\mathbb{R}^{nq}$  et donc muni d'une métrique.

(23) Du point de vue probabiliste, le composé de  $\pi_n'$  avec  $R$  est un processus linéaire basé sur  $(\mathbb{R}^{nq})^*$ , décomposé par une v.a. à valeurs dans  $\mathbb{R}^{nq}$  notée  $(\xi(0), \dots, \xi(n-1))$ . Il suffit de faire l'ACP de cette variable aléatoire. Dont l'opérateur de corrélation est noté  $C^n$ . Comme  $\rho_n$  injecte  $\mathbb{R}^{nq}$  dans  $H$ , on peut considérer que cette ACP est relative à une v.a. de corrélation  $\Gamma^n$  à valeurs dans  $H$ . Le diagramme suivant est commutatif.



On peut alors étudier la convergence de cette dernière ACP vers l'ACP de  $R$ , en supposant par exemple  $C_R$  compact,  $\sup \|\rho_n \pi_n\| < \infty$ , et  $(\rho_n \pi_n x)_n \rightarrow x$  pour tout  $x \in H'$ .

(25) Voici trois exemples généraux avec  $q=1$  et  $n=5$





(26) Exemple pratique.

a) Un premier exemple a été traité, les données ayant été fournies par M. Pradères du laboratoire de médecine nucléaire de l'Université P. Sabatier à Toulouse. Une substance iodée ( $I^{131}$ ) a été injectée par voie intravéneuse à 66 individus à l'instant  $t=0$ . L'expérience durant 12 minutes, on prend cette durée comme unité de temps. Aux instants  $t_j = \epsilon$  ;  $0,4'$  ;  $0,9'$  ;  $2'$  ;  $3,4'$  ;  $5'$  ;  $6,7'$  ;  $8,8'$  et  $12'$ , on mesure la radioactivité  $Z^D(j)$  et  $Z^G(j)$  dans le rein droit et dans le rein gauche de chaque sujet. Le problème pratique est de comparer les courbes relatives aux reins droits, les courbes relatives aux reins gauches, et les couples de courbes relatives aux deux reins.: dans ce dernier cas  $q=2$ . L'ACP fonctionnelle a montré que pour chaque courbe plane, les deux facteurs de forme principaux sont la rapidité du taux de montée de radioactivité, et le taux résiduel de radioactivité, à la fin de l'expérience.

b) D'autres exemples vont être traités ; ils concernent des données de l'INSEE et ont été exposés oralement au colloque.

(27) Conclusions de l'étude pratique.

Bien que les essais numériques ne soient pas terminés, les résultats suivants sont déjà acquis.

a) L'ACP fonctionnelle par interpolation polynomiale n'a aucun intérêt statistique : ceci est dû aux grandes oscillations des polynômes d'interpolation et à la non convergence pour  $n \rightarrow \infty$ .

b) En général l'ACP dans  $L^2(I)$  avec relèvement trivial, est améliorée par l'ACP utilisant des fonctions "spline" d'ordre un.

En ce qui concerne l'utilisation de fonctions "spline" d'ordre supérieur, il faut bien noter que la norme (17) s'écrit pour  $D=I$  et  $s=2$

$$\| \varphi \|_2^2 = \left( \int_I \varphi(x)^2 dx + a_1 \int_I \varphi'(x)^2 dx + a_2 \int_I \varphi''(x)^2 dx \right)^{1/2}$$

Plus  $a_1$  et  $a_2$  sont grands, plus on donne de l'importance aux variations de pente et de courbure. Il y a donc de nombreux essais numériques à faire pour se rendre compte des valeurs optimales : pour un type donné de courbes, il y a des valeurs optimales de  $a_1$  et  $a_2$  à trouver.

§.4. Remarques terminales.

Le domaine de l'ACP fonctionnel est a peine

effleuré ci-dessus même du point de vue théorique ; voici quelques prolongements possibles.

- La propriété variationnelle (18) s'étend aux domaines  $D$  de dimension quelconque du moins pour  $s=1$ . Autrement dit, étant une décomposition simplifiée de  $D = \bigcup_i \omega_i$  de  $D$  et étant donnée une fonction  $\varphi$  définie sur la réunion des frontières  $\partial \omega_i$  des simplexes, affine sur chaque face, parmi toutes les  $f$  de  $H^1(D)$  coïncidant avec  $\varphi$  sur  $\bigcup_i \partial \omega_i$ , celle qui minimise  $\sum_k \int_D \frac{\partial f}{\partial x_k}(x)^2 dx$  est celle qui est affine sur chaque simplexe. On peut donc étendre les méthodes qui précèdent aux fonctions de plusieurs variables, en utilisant des splines d'ordre un.

2. Soit  $Z$  une v.a. de carré intégrable à valeurs dans un espace de Hilbert réel. Si  $(z_1, z_2, \dots)$  sont les composantes de  $Z$  par rapport à une base orthonormée, Y. Escoufier [4] associe à  $Z$  l'opérateur linéaire  $U_Z = \sum_i z_i \otimes z_i$  de  $L^2(\mathcal{A})$ . Cet opérateur appartient à l'espace hilbertien  $K$  des opérateurs de Hilbert Schmidt de  $L^2(\mathcal{A})$ . Si donc on a  $q$  processus  $(z_t^1), \dots, (z_t^q)$  représentés par des v.a. du second ordre à valeurs dans des espaces de Hilbert  $H_1, \dots, H_q$ , on peut leur associer  $q$  vecteurs  $U_1, \dots, U_q$  de  $K$ . Pour analyser les proximités entre les  $q$  processus, il suffit d'analyser les proximités entre les  $q$  vecteurs  $U_j$  de l'hilbertien  $K$ , en faisant une ACP usuelle sur ces  $q$  vecteurs. On peut faire ainsi une ACP sur des v.a. à valeurs dans des espaces de Hilbert donnés  $H_1, \dots, H_q$ , sans connaître de structure hilbertienne sur la somme des  $H_q$ . Pour  $q=2$ , ceci peut être appliqué aux couples de courbes.

3. Les techniques évoquées dans ce travail n'ont aucun intérêt pour les processus stationnaires. Dans ce cas, c'est une ACP toute différente qui doit être effectuée [1].

Il est probable que les techniques ci-dessus ont un intérêt pratique pour l'étude de certains problèmes de détection.

## BIBLIOGRAPHIE

- [1] D.R. Brillinger : Time series data analysis and theory, Holt Rinehart and Winston Inc 1975.  
 [2] J. Dauxois et A. Pousse : Thèse d'état. Univer-

---

sité Paul Sabatier Toulouse 1976.

- 3 J.C. Deville : Méthodes statistiques et numériques de l'analyse harmonique. Annales de l'INSEE n°15, Paris 1975.
- 4 Y. Escoufier : Le traitement des variables vectorielles Biometrics 29, p.751-760. Décembre 1973.
- 5 T.N.E. Greville : Theory and applications of spline functions. Académie Press 1969.
- 6 P. Krée : Introduction à la statistique des processus (livre à paraître).
- 7 P. Krée : Séminaire à l'Institut H. Poincaré. Publié par le secrétariat de mathématique, 1ère année 1974-75.
- 8 P.J. Laurent : Approximation et optimisation Hermann 1972.
- 9 A.M. Mazerolles : Problèmes numériques et statistiques relatifs à l'ACP des variables aléatoires à valeurs dans un Hilbert. Thèse de 3ème cycle, à soutenir en mai-Juin 1977 à l'Université P. Sabatier de Toulouse.
- 10 E. Parzen : Time series analysis Papers. Holden days.
- 11 C. Soize : Thèse de 3ème cycle de l'Université de Paris VI juin 1975.
- 12 C. Soize : Transformation quadratique d'un processus gaussien et application à la dynamique stochastique. Journal de Mécanique décembre 1976.
- 13 C. Trézos : Méthodes probabilistes en sécurité des constructions. Article dans la revue "Constructions métalliques" juin 1977.

