SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

P. BRÉMAUD

La méthode des semi-martingales en filtrage quand l'observation est un processus ponctuel marqué

Séminaire de probabilités (Strasbourg), tome 10 (1976), p. 1-18

<http://www.numdam.org/item?id=SPS_1976__10__1_0>

Université de Strasbourg
Séminaire de Probabilités

Février 1975.

LA METHODE DES SEMI-MARTINGALES EN FILTRAGE QUAND L'OBSERVATION EST UN PROCESSUS PONCTUEL MARQUE

par

P. BREMAUD

Ce qui va suivre est un extrait d'un cours[(1)] sur les processus ponctuels marqués et quelques applications concernant ces processus. Pour ne pas avoir à reproduire tous les chapitres précédents nous commencerons par rappeler brièvement deux ou trois définitions.

Un $(\Omega,\mathcal{F})$ est donné ainsi qu'une histoire $\underline{F} = (F_t, t\in R^+)$ (ou, si on veut, une filtration). Soit $(\widetilde{Y}_n, n\in Z^+)$ une suite de v.a. à valeurs dans $(E',\mathcal{E}')$ et $(T_n, n\in Z^+)$ une suite de $\underline{F}$-temps d'arrêt, $T_o = 0$, $T_n \uparrow\uparrow \infty$. On suppose que $\widetilde{Y}_n$ est F_{T_n}-mesurable, $\forall\, n\in Z^+$. La suite $(T_n,\widetilde{Y}_n, n\in Z^+)$ s'appelle un processus ponctuel marqué ou processus de sauts (cette dernière terminologie s'appliquant surtout au cas où E est vectoriel et où $\widetilde{Y}_n = \Delta Z_{T_n}$ pour un processus c.à.d. $\underline{Z} = (Z_t, t\in R^+)$ à valeurs dans $(E,\mathcal{E})$, étagé et ne sautant qu'aux T_n). Si $(E',\mathcal{E}') = (R,B(R))$ et si $\widetilde{Y}_n = 1$, $\forall\, n\in Z^+-\{0\}$, $\widetilde{Y}_o = 0$, on posera $N_t = \sum_n I(T_n \le t)$. $\underline{N} = (N_t, t\in R^+)$ est alors un processus de comptage. Si on se donne P sur $(\Omega,\mathcal{F})$, $(\underline{N},P)$ est ce qu'on appelle un processus ponctuel stochastique (P.P.S.) sur R^+ . De même, on dira que $(\underline{Y},P)$ est un processus ponctuel marqué stochastique (P.P.M.S.) , si $\underline{Y} = (T_n,\widetilde{Y}_n, n\in Z^+)$. Pour un tel processus, on note $n_t^B = \sum_n I(T_n\le t)I(\widetilde{Y}_n\in B)$, $t\in R^+$, $B\in\mathcal{E}'$, et $\sigma(\underline{Y},t) = \sigma(n_s^B, s\le t, B\in\mathcal{E}')$. En général pour un processus $\underline{Z} = (Z_t, t\in R^+)$ à valeurs dans $(E',\mathcal{E}')$, on posera

(1) U.E.R. Math. de la Décision, Paris IX (Dauphine) : Martingales et processus de sauts, applications, 3e Cycle, 1974-75.

$\sigma(\underline{Z},t) = \sigma(Z_s, s \leq t)$ et $\sigma(\underline{Z}) = (\sigma(\underline{Z},t), t \in R^+)$. Toutes les histoires $\underline{F}$ rencontrées sont supposées satisfaire aux "conditions habituelles" (F_o contient les P-négligeables, $\underline{F}$ continue à droite), sinon on rend $\underline{F}$ telle, et on conserve pour l'histoire ainsi transformée les mêmes notations.

Soit N une mesure de transition de $(\Omega,\mathcal{F})$ dans $]0,\infty[\times E', \mathcal{B}(]0,\infty[)\otimes\mathcal{E}')$, positive, σ-finie et $\underline{\underline{H}}$ une application de $(\Omega\times R^+\times E', \mathcal{F}\otimes\mathcal{B}(R^+)\otimes\mathcal{E}')$ dans $(R,\mathcal{B}(R))$. On dit que $\underline{\underline{H}}$ est $\underline{F}$-prévisible si elle est $\mathcal{P}(\underline{F})\otimes\mathcal{E}'$ mesurable, où $\mathcal{P}(\underline{F})$ est la tribu $\underline{F}$-prévisible sur $R^+\times\Omega$. On dit que N est $\underline{F}$-prévisible si $((N\circ\underline{\underline{H}})_t, t\in R^+)$ est $\underline{F}$-prévisible pour toute $\underline{\underline{H}}$ positive $\underline{F}$-prévisible (où on note $(N\circ\underline{\underline{H}})_t(\omega) = \int_{[0,t]\times E} N(\omega,ds,dx)\, H(s,\omega,x)$). On note $\Delta(N\circ\underline{\underline{H}})_t = (N\circ\underline{\underline{H}})_t - (N\circ\underline{\underline{H}})_{t-}$ et $N(\{t\}\times E') = \Delta(N\circ\underline{\underline{1}})_t$.

Pour tous les résultats sur les P.P.M.S. auxquels on fera allusion, il faut consulter : J. JACOD : Multivariate point processes : Predictable projection, Radon-Nikodym derivatives, representation of martingales (à paraître). Les références n'étant pas données, le lecteur est prié de consulter la bibliographie commentée. Ceci dit, nous nous raccrochons maintenant au cours.

1. LE PROBLEME GENERAL DU FILTRAGE.

1.1. Donnons-nous un espace probabilisé $(\Omega,\mathcal{F},P)$ et deux processus $\underline{X}$ et $\underline{Y}$ sur cet espace. $\underline{X}$, qui prend ses valeurs dans un espace mesurable $(E,\mathcal{E})$, s'appelle l'<u>état</u> d'un "système" $\mathcal{S}$. $\underline{Y}$, qui prend ses valeurs dans un autre espace mesurable $(E',\mathcal{E}')$ s'appelle l'<u>observation</u> du système $\mathcal{S}$. L'observation $\underline{Y}$ contient une information au sujet de l'état $\underline{X}$, en ce sens que $\underline{X}$ et $\underline{Y}$ sont statistiquement corrélés, corrélation qu'on peut représenter (par exemple) par la loi temporelle du couple $(\underline{X},\underline{Y})$ sous la probabilité P .

Soit maintenant g une fonction mesurable de $(E,\mathcal{E})$ dans $(R,B(R))$ telle que $E\{|g(X_t)|\} < \infty$, $\forall\, t \in R^+$. On dit qu'on a filtré $g \circ \underline{X}$ par rapport à $\underline{Y}$ si on sait donner une expression explicite de $E\{g(X_t)|\sigma(\underline{Y},t)\}(\omega)$ en fonction de la trajectoire $(Y_s(\omega), s \in [0,t])$ pour presque tout $\omega \in \Omega$. On dit qu'on a filtré $\mathcal{L}(\underline{X})$ par rapport à $\underline{Y}$ si on sait filtrer $g \circ \underline{X}$ par rapport à $\underline{Y}$ pour toute g mesurable bornée de $(E,\mathcal{E})$ dans $(R,\mathcal{B}(R))$.

Il faut bien insister sur le fait qu'on cherche une expression explicite, ou encore (ce que nous considérons comme équivalent) un algorithme qui à partir de $(Y_s(\omega), s \in [0,t])$ fournit $E\{g(X_t)|\sigma(\underline{Y},t)\}(\omega)$ pour tout $t \in R^+$. L'expérience montre qu'il n'est pas facile d'obtenir de tels filtres. En général, la théorie conduit à des préfiltres, c'est-à-dire à des expressions de $E\{g(X_t)|\sigma(\underline{Y},t)\}(\omega)$ en fonction de $(Y_s(\omega), s \in [0,t])$ et de $(E\{f_i(X_s)|\sigma(\underline{Y},s)\}(\omega), s \in [0,t], i \in I)$ où les f_i sont des applications mesurables de $(E,\mathcal{E})$ dans $(R,\mathcal{B}(R))$ tels que $E\{|f_i(X_t)| < \infty$, $\forall\, t \in R^+$, $\forall\, i \in I$. On dira que le préfiltre de $g \circ \underline{X}$ par rapport à $\underline{Y}$ s'"appuie sur les préfiltres de $f_i \circ \underline{X}$ par rapport à $\underline{Y}$, $i \in I$", ou encore pour donner une image qui fera plaisir aux plombiers : "le préfiltre de $g \circ \underline{X}$ fuit par les préfiltres de $f_i \circ \underline{X}$, $i \in I$".

On peut essayer de "colmater les fuites" en calculant les préfiltres de $f_i \circ \underline{X}$ par rapport à $\underline{Y}$ pour tout $i \in I$, ces préfiltres risquent de fuir à leur tour par des préfiltres qui ne sont pas les préfiltres de $f_i \circ \underline{X}$, $i \in I$. Si on arrive à la situation où les préfiltres de $g \circ \underline{X}$ et $f_i \circ \underline{X}$, $i \in I$, s'appuient uniquement sur les préfiltres de $f_i \circ \underline{X}$, $i \in I$, on dira que le préfiltre de $g \circ \underline{X}$ est fermé par les préfiltres de $f_i \circ \underline{X}$, $i \in I$. Si de plus g ne figure pas parmi les $(f_i, i \in I)$ on dira que le préfiltre de $g \circ \underline{X}$ est fermé extérieurement par les préfiltres de $f_i \circ \underline{X}$, $i \in I$. Dans tous les cas, on dira que le préfiltre de $g \circ \underline{X}$ est un préfiltre fermé de dimension card(I). Lorsque le préfiltre fermé est de dimension finie, il est susceptible d'une "algorithmisation" qui en fera un véritable filtre, et cela représente souvent un travail non négligeable. Mais ça n'est encore pas très satisfaisant : on cherche des filtres qui conduisent à un minimum de calculs par exemple des <u>filtres récursifs</u>. Heuristiquement si on a un filtre de $g \circ \underline{X}$ par rapport à $\underline{Y}$, on dira que ce filtre est récursif si, en posant $(\widehat{g \circ \underline{X}})_t = E\{g(X_t)|\sigma(\underline{Y},t)\}$, $(\widehat{g \circ \underline{X}})_{t+dt}$ se calcule à partir de $(\widehat{g \circ \underline{X}})_t$ et de $(Y_s, s \in]t,t+dt])$. Evidemment, la dernière phrase n'a aucun sens mathématique, sauf dans le cas où on remplace l'axe continu des temps R^+, par l'axe discrétisé Z^+ (les processus $\underline{X}$ et $\underline{Y}$ sont alors des suites $(X_n, n \in Z^+)$ et $(Y_n, n \in Z^+)$) : il s'agit, dans ce cas, de calculer $(\widehat{g \circ \underline{X}})_{n+1}$ en fonction de $(\widehat{g \circ \underline{X}})_n$ et de Y_{n+1}.

On peut également "définir" les préfiltres fermés récursifs : supposons que le préfiltre de $g \circ \underline{X}$ soit fermé par les préfiltres de $f_i \circ \underline{X}$, $i \in I$ et que $(\widehat{g \circ \underline{X}})_{t+dt}$ et $(\widehat{f_i \circ \underline{X}})_{t+dt}$ se calculent à partir de $(\widehat{g \circ \underline{X}})_t$, $(\widehat{f_i \circ \underline{X}})_t$, $i \in I$ et $(Y_s, s \in]t,t+dt])$. Un tel préfiltre "fermé récursif" est très satisfaisant, en général, du point de vue des calculs, car il se prête à une "algorithmisation" qui en fera un filtre récursif.

Quelques exemples illustreront ces différentes notions :

1.2. <u>Exemples</u>.

a) Considérons un P.P.S. $(\underline{N},P)$ de $\underline{F}$-intensité $\underline{\lambda}$ [(1)] de la forme $\lambda_t \equiv \Lambda$, Λ étant une variable aléatoire non négative P-intégrable et $\underline{F}$ étant de la forme $F_t = \sigma(\Lambda) \vee \sigma(\underline{N},t)$. $\underline{N}$ est l'observation $(\underline{N} = \underline{Y})$ de l'état $\Lambda (X_t \equiv \Lambda)$ et on veut calculer $E\{\Lambda|\sigma(\underline{N},t)\}$ c'est-à-dire filtrer Λ par rapport à $\underline{N}$. Si H est la distribution de Λ sous P, on sait montrer que :

(1) C'est-à-dire $N_t - \int_0^t \lambda_s \, ds = (P,\underline{F})$ martingale locale, $\underline{\lambda}$ non négatif, $\underline{F}$-progressif.

(1) $$E\{\Lambda|\sigma(\underline{N},t)\} = \frac{\int_0^\infty \lambda^{N_t+1} \exp(-\lambda t)\ H(d\lambda)}{\int_0^\infty \lambda^{N_t} \exp(-\lambda t)\ H(d\lambda)} .$$

Si on pose $\Psi(n,t) = \int_0^\infty \lambda^n \exp(-\lambda t)\ H(d\lambda)$, on a :

(2) $$E\{\Lambda|\sigma(\underline{N},t)\} = \frac{\Psi(N_t+1,t)}{\Psi(N_t,t)} .$$

Comme Ψ est (en principe) précalculable, (2) est une forme trivialement récursive.

b) Le fameux filtre de KALMAN-BUCY est une très bonne illustration de la notion de récursivité. Il nous servira d'introduction à la méthode des semi-martingales.

$\underline{X}$ prend ses valeurs dans $(R^n,\mathcal{B}(R^n))$ et satisfait à une équation différentielle stochastique de ITO :

(3) $$X_t = X_o + \int_0^t A(s)\ X_s\ ds + \oint_0^t B(s)\ dW_s$$

où $\underline{W}$ est un $(P,\underline{F})$ Wiener vectoriel à k-dimensions, X_o est un n-vecteur gaussien F_o-mesurable, $\underline{F}$ une histoire, $(A(t),t\in R^+)$ et $(B(t),t\in R^+)$ des fonctions matricielles de dimensions $n\times n$ et $n\times k$ respectivement. On supposera les dernières fonctions continues (mais l'équation (3) a un sens sous des conditions bien plus larges).

$\underline{Y}$ prend ses valeurs dans $(R^m,\mathcal{B}(R^m))$ et satisfait à :

(4) $$Y_t = \int_0^t F(s)\ X_s\ ds + \oint_0^t G(s)\ dW'_s$$

où $\underline{W}'$ est un $(P,\underline{F})$ Wiener vectoriel à ℓ-dimensions, $(F(t),t\in R^+)$ et $(G(t),t\in R^+)$ des fonctions matricielles continues (ça n'est pas nécessaire dans la théorie générale) de dimensions $m\times n$ et $m\times \ell$ respectivement.

Dans le cas où $X_o = c^{te}$ (par exemple), KALMAN et BUCY fournissent les équations suivantes pour $\hat{X}_t = E\{X_t | \sigma(\underline{Y},t)\}$:

$$(5) \qquad \hat{X}_t = X_o + \int_0^t A(s)\, \hat{X}_s\, ds + \oint_0^t K(s)\, d(Y_s - \int_0^s F(y)\, \hat{X}_u\, du) \;.$$

Le processus $(Y_t - \int_0^t F(u)\, \hat{X}_u\, du, t \in R^+)$ est de la forme $\oint_0^t G(s)\, d\tilde{W}_s$ où $\underline{\tilde{W}}$ est un $(P,\sigma(\underline{Y}))$ Wiener vectoriel à ℓ-dimensions et $(K(t), t \in R^+)$ est une fonction matricielle de dimension $n \times m$: ce sont le processus d'innovations et le gain d'innovations, dans le langage de la théorie du filtrage.

Soit $\Sigma(t) = E\{X_t - \hat{X}_t)(X - \hat{X}_t)^+\}$ la matrice d'erreurs(1), alors $\Sigma(t)$ et $K(t)$ satisfont à

$$(6) \qquad \frac{d\Sigma(t)}{dt} = (B(t) - K(t)\, G(t))\, G(t)\, (B(t) - K(t)\, G(t))^+$$
$$+ (A(t) - K(t)\, F(t))\, F(t)\, \Sigma(t) + \Sigma(t)\, (A(t) - K(t)\, F(t))^+$$

$$(7) \qquad K(t)\, B(t)\, B(t)^+ = B(t)\, G(t)^+ + \Sigma(t)\, F(t)^+ \;.$$

L'équation (5) montre qu'on a un préfiltre fermé à une dimension qui est de plus récursif car, heuristiquement :

$$(8) \qquad \hat{X}_{t+dt} - \hat{X}_t = (A(t) - K(t)\, F(t))\, \hat{X}_t\, dt + K(t)\, dY_t \;.$$

1.3. Le problème du filtrage posé en termes de semi-martingales.

Nous avons dit que la corrélation entre $\underline{X}$ et $\underline{Y}$ peut être donnée par la loi temporelle du couple $(\underline{X},\underline{Y})$ sous la probabilité P. Dans la théorie de KALMAN-BUCY, c'est un couple d'équations stochastiques qui détermine cette loi temporelle. Ces équations sont des équations différentielles stochastiques et décrivent le système $\mathcal{S}$ dans son déroulement dans le temps : elles font clairement apparaître la dynamique du système, alors que la loi temporelle de $(\underline{X},\underline{Y})$ est essentielle-

(1) $A^+ = A$ transposée.

ment statique. Intuitivement, on sent bien que cette description dynamique du système est adaptée au filtrage récursif. Si on regarde d'un peu loin l'équation de l'état $\underline{X}$, on voit qu'elle appartient au type ci-dessous :

$$(9) \qquad (\mathcal{P}) \begin{cases} X_t = X_o + \int_0^t f_s \, ds + m_t \\ m_t = (P,\underline{F}) \text{ martingale locale} \end{cases}$$

où $\underline{f}$ est adapté à $\underline{F}$ et mesurable. Supposons de plus que $E\{\int_0^t |f_s| ds\} < \infty$, $\forall\ t \in R^+$ et que $\underline{m}$ soit une $(P,\underline{F})$ martingale. Posons :

$$\tilde{m}_t = E\{m_t|\sigma(\underline{Y},t)\} + E\{\int_0^t f_s \, ds|\sigma(\underline{Y},t)\} - \int_0^t E\{f_s|\sigma(\underline{Y},s)\}ds + E\{X_o|\sigma(\underline{Y},t)\} - E\{X_o\} .$$

On a alors :

$$(10) \qquad (\hat{\mathcal{P}}) \begin{cases} \hat{X}_t = E\{X_o\} + \int_0^t \hat{f}_s \, ds + \tilde{m}_t \\ \tilde{m}_t = (P,\sigma(\underline{Y})) \text{ martingale } \quad (\tilde{m}_o = 0) \end{cases}$$

(où on note $\hat{Z}_t = E\{Z_t|\sigma(\underline{Y},t)\}$).

Supposons que $(\underline{Y},P)$ soit un P.P.S. de $\sigma(\underline{Y})$-intensité $\underline{\hat{f}}$. D'après le théorème de réprésentation (voir [12], [13], [14] par exemple), il existe un processus $\underline{K}$ $\sigma(\underline{Y})$-prévisible tel que $\int_0^t |K_s| \hat{f}_s \, ds < \infty$ P-p.s., $\forall\ t \in R^+$ et $\tilde{m}_t = \int_0^t K_s \, d(Y_s - \int_0^s \hat{f}_u \, du)$. Le processus $(Y_t - \int_0^t \hat{f}_s \, ds, t \in R^+)$ s'appelle le processus d'innovations et $\underline{K}$ le gain d'innovations, par analogie avec KALMAN-BUCY. On a donc une expression du même type que (5) :

$$(11) \qquad \hat{X}_t = E\{X_o\} + \int_0^t \hat{f}_s \, ds + \int_0^t K_s \, d(Y_s - \int_0^s \hat{f}_u \, du) .$$

Il faut évidemment déterminer $\underline{\hat{f}}$ et $\underline{K}$. Nous allons le faire, mais dans un cadre plus général.

2. LA METHODE DES SEMI-MARTINGALES LORSQUE L'OBSERVATION EST UN PROCESSUS DE SAUTS.

2.1. Le problème.

On a un espace probabilisé $(\Omega,\mathfrak{F},P)$, une histoire $\underline{F}$, un processus $\underline{X}$ qui satisfait à :

$$(P)\quad \begin{cases} X_t = X_o + \int_0^t f_s \, ds + m_t \\ m_t = (P,\underline{F}) \text{ martingale} \end{cases}$$

où $\underline{f}$ est adapté à $\underline{F}$, mesurable et tel que $E\{\int_0^t |f_s| ds\} < \infty$, $\forall\, t \in R^+$.
On supposera de plus que $\underline{m}$ a P-p.s. toutes ses trajectoires à variation bornée sur tout intervalle borné, et que $\underline{X}$ est borné.

Nous venons de voir que $\hat{\underline{X}}$ $(\hat{X}_t = E\{X_t | \sigma(\underline{Y},t)\}$ satisfait alors à :

$$(\hat{P})\quad \begin{cases} \hat{X}_t = E\{X_o\} + \int_0^t \hat{f}_s \, ds + \tilde{m}_t \\ \tilde{m}_t = (P,\sigma(\underline{Y})) \text{ martingale } (\tilde{m}_o = 0) . \end{cases}$$

$\underline{Y}$ est un processus de sauts à valeurs dans $(E',\mathcal{E}')$ pour son histoire intrinsèque $\sigma(\underline{Y})$ et aussi pour l'histoire $\underline{F}$ plus grosse que $\sigma(\underline{Y})$. On suppose que $(\underline{Y},P)$ admet le $\underline{F}$-noyau prévisible N et le $\sigma(\underline{Y})$-noyau prévisible $\hat{N}$ (1), ce qui signifie que si $\underline{\underline{H}}$ est une application de $R^+ \times \Omega \times E'$ dans R qui est $\underline{F}$-prévisible (resp. $\sigma(\underline{Y})$-prévisible) et positive, on a :

$$(12)\qquad E\{\sum_{n\geq 1} H(T_n,\tilde{Y}_n)\} = E\{\int_{]0,\infty[\times E'} H(s,y)\, N(ds,dy)\}$$

$$(12')\qquad (\text{resp.} = E\{\int_{]0,\infty[\times E'} H(s,y)\, \hat{N}(ds,dy)\}),$$

ou encore, ce qui est équivalent : si $\underline{\underline{H}}$ est une application $\underline{F}$-prévisible

---

(1) ces conditions sont souvent vérifiées (cf. l'article de Jacod cité plus haut).

(resp. $\sigma(\underline{Y})$-prévisible) telle que $\int_{]0,t]\times E'} |H(s,y)|N(ds,dy) < \infty$ P-p.s., $\forall\, t \in R^+$ (resp. $\int_{]0,t]\times E'} |H(s,y)|\hat{N}(ds,dy) < \infty$ P-p.s., $\forall\, t \in R^+$), alors :

(13)
$$(\eta \circ \underline{\underline{H}})_t = \sum_{\substack{n \geq 1 \\ T_n \leq t}} H(T_n, \tilde{Y}_n) - \int_{]0,t]\times E'} H(s,y)\, N(ds,dy)$$

= $(P,\underline{F})$ martingale locale

(13')
(resp. $$(\hat{\eta} \circ \underline{\underline{H}})_t = \sum_{n \geq 1} H(T_n, \tilde{Y}_n) - \int_{]0,t]\times E'} H(s,y)\, \hat{N}(ds,dy)$$

= $(P,\sigma(\underline{Y}))$ martingale locale.

On supposera de plus que les noyaux N et $\hat{N}$ satisfont à $E\{N(]0,t]\times E')\} = E\{\hat{N}(]0,t]\times E')\} < \infty$, $\forall\, t \in R^+$, ce qui implique que $\eta \circ \underline{\underline{H}}$ (resp. $\hat{\eta} \circ \underline{\underline{H}}$) sont de vraies martingales si $E\{\int_{]0,t]\times E'} H(s,y)\, N(ds,dy)\} < \infty$, $\forall\, t \in R^+$ (resp. $E\{\int_{]0,t]\times E'} H(s,y)\, \hat{N}(ds,dy)\} < \infty$, $\forall\, t \in R^+$).

Voici un des résultats fondamentaux de la théorie des P.P.M.S. (voir JACOD ; voir aussi [13]) :

<u>Représentation des</u> $(P,\sigma(\underline{Y}))$ <u>martingales</u> [14] .

<u>Soit</u> $\underline{M}$ <u>une</u> $(P,\sigma(\underline{Y}))$ <u>martingale locale telle que</u> $M_0 = 0$. <u>Alors il existe une application</u> $\sigma(\underline{Y})$-<u>prévisible</u> $\underline{\underline{H}}$ <u>telle que</u> $\int_{]0,t]\times E'} |H(s,y)|\hat{N}(ds,dy) < \infty$ P-<u>p.s.</u>, $\forall\, t \in R^+$ <u>et</u> $M_t = (\hat{\eta} \circ \underline{\underline{H}})_t$ P-<u>p.s.</u>, $\forall\, t \in R^+$.

De ce résultat, il découle que $\hat{\underline{X}}$ a la forme :

(14)
$$\hat{X}_t = E\{X_0\} + \int_0^t \hat{f}_s\, ds + (\hat{\eta} \circ \underline{\underline{K}})_t$$

où $\underline{\underline{K}}$ est une application $\sigma(\underline{Y})$-prévisible telle que $\int_{]0,t]\times E'} |K(s,y)|\hat{N}(ds,dy) < \infty$ P-p.s., $\forall\, t \in R^+$. $\underline{\underline{K}}$ s'appelle le <u>gain</u> du filtre de $\underline{X}$ par rapport à $\underline{Y}$.

Nous allons maintenant calculer $\underline{\underline{K}}$, ou plutôt, en donner une "expression".

2.2. Calcul du gain.

a) Nous allons commencer par un préliminaire : il existe 3 applications $\sigma(\underline{Y})$-prévisibles $\underline{\underline{X}}^{(1)}$, $\underline{\underline{X}}^{(2)}$ et Ψ telles que, pour toute application $\underline{\underline{H}}$ $\sigma(\underline{Y})$-prévisible positive, on ait, pour tout $t \in R^+$:

$$(15)\quad E\{\int_{]0,t]\times E'} H(s,y)\, X^{(1)}(s,y)\hat{N}(ds,dy)\} = E\{\int_{]0,t]\times E'} H(s,y)\, X_{s-}N(ds,dy)\}$$

$$(16)\quad E\{\int_{]0,t]\times E'} H(s,y)\, X^{(2)}(s,y)\hat{N}(ds,dy)\} = E\{\int_{]0,t]\times E'} H(s,y)\, X_{s-}\hat{N}(ds,dy)\}$$

$$(17)\quad E\{\int_{]0,t]} H(s,y)\, \Psi(s,y)\, \hat{N}(ds,dy)\} = E\{\sum_{n\geq 1} H(T_n,\tilde{Y}_n)\, \Delta m_{T_n} I(T_n \leq t)\} .$$

Nous allons prouver l'existence de telles applications seulement dans le cas où (15), (16) et (17) sont vérifiées sur $[0,T]$, $T \in R^+$. Le cas général s'obtient facilement par recollement des processus obtenus lorsque T décrit R^+ .

Soit μ_1 , μ_2 , μ_3 et ν les mesures définies sur $]0,\infty[\times\Omega\times E', P(\sigma(\underline{Y}))\otimes\mathcal{E}')$ par leurs actions sur les applications $\sigma(\underline{Y})$-prévisibles positives $\underline{\underline{H}}$:

$$(18)\quad \mu_1(\underline{\underline{H}}) = E\{\int_{]0,T]\times E'} H(s,y)\, X_{s-}\, N(ds,dy)\} = E\{\sum_{\substack{n\\ T_n \leq T}} H(T_n,\tilde{Y}_n)\, X_{T_n-}\}$$

$$(19)\quad \mu_2(\underline{\underline{H}}) = E\{\int_{]0,T]\times E'} H(s,y)\, X_{s-}\, \hat{N}(ds,dy)\}$$

$$(20)\quad \mu_3(\underline{\underline{H}}) = E\{\sum_{n\geq 1} H(T_n,\tilde{Y}_n)\, \Delta m_{T_n} I(T_n \leq t)\}$$

$$(21)\quad \nu(\underline{\underline{H}}) = E\{\int_{]0,\infty[\times E'} H(s,y)\, \hat{N}(ds,dy)\} = E\{\sum_{n\geq 1} H(T_n,\tilde{Y}_n)\} .$$

Comme $\underline{X}$ est borné, $\underline{\Delta m}$ est borné. Ce qui entraîne, puisque $\hat{N}(]0,t]\times E')$ et $N(]0,t]\times E')$ sont finis P-p.s., $\forall\, t\in R^+$, que les mesures μ_1, μ_2, μ_3 et ν sont σ-finies. D'autre part, il est facile de voir que μ_1, μ_2 et μ_3 sont absolument continues par rapport à ν, ce qui entraîne l'existence de $\underline{\underline{X}}^{(1)}$, $\underline{\underline{X}}^{(2)}$ et $\underline{\Psi}$ (les dérivées de Radon-Nikodym $\frac{d\mu_1}{d\nu}$, $\frac{d\mu_2}{d\nu}$, $\frac{d\mu_3}{d\nu}$).

b) Soit $\underline{\underline{H}}$ un processus $\sigma(\underline{Y})$-prévisible positif tel que $\hat{\eta}\circ\underline{\underline{H}}$ soit une $(P,\sigma(\underline{Y}))$ martingale bornée. Nous allons calculer $E\{X_t(\hat{\eta}\circ\underline{\underline{H}})_t\}$ et $E\{\hat{X}_t(\hat{\eta}\circ\underline{\underline{H}})_t\}$ ce qui nous permettra, en écrivant l'égalité entre ces deux quantités, de trouver une forme nécessaire de $\underline{\underline{K}}$.

On commence par appliquer la formule d'intégration par parties pour les intégrales de Stieltjes à $\hat{X}_t(\hat{\eta}\circ\underline{\underline{H}})_t$:

$$(22)\quad \hat{X}_t(\hat{\eta}\circ\underline{\underline{H}})_t = \int_0^t \hat{X}_{s-}d(\hat{\eta}\circ\underline{\underline{H}})_s + \int_0^t (\hat{\eta}\circ\underline{\underline{H}})_s d\hat{X}_s$$

$$= \int_0^t \hat{X}_{s-}d(\hat{\eta}\circ\underline{\underline{H}})_s + \int_0^t (\hat{\eta}\circ\underline{\underline{H}})_s\, \hat{f}_s\, ds + \int_0^t ((\hat{\eta}\circ\underline{\underline{H}})_{s-} + \Delta(\hat{N}\circ\underline{\underline{H}})_s)d(\hat{\eta}\circ\underline{\underline{K}})_s$$

$$+ \sum_{n\geq 1} H(T_n,Y_n)(K(T_n,Y_n) + \Delta(\hat{N}\circ\underline{\underline{K}})_{T_n})\, I(T_n\leq t).$$

D'où :

$$(23)\qquad \hat{X}_t(\hat{\eta}\circ\underline{\underline{H}})_t = \int_0^t \hat{X}_{s-}\, d(\hat{\eta}\circ\underline{\underline{H}})_s + \int_0^t ((\hat{\eta}\circ H)_{s-} + \Delta(\hat{N}\circ\underline{\underline{H}})_s)\, d(\hat{\eta}\circ\underline{\underline{K}})_s$$

$$+ (\hat{\eta}\circ\underline{\underline{Z}})_t - (\hat{N}\circ\underline{\underline{Z}})_t + \int_0^t (\hat{\eta}\circ\underline{\underline{H}})_s\, \hat{f}_s\, ds$$

où $Z(t,y) = H(t,y)(K(t,y) + \Delta(\hat{N}\circ\underline{\underline{K}})_t)$.

Les trois premiers termes du 2me membre de (23) sont des martingales locales car ils s'écrivent sous la forme $\hat{\eta}\circ\underline{\underline{Y}}$ où $\underline{\underline{Y}}$ est $\sigma(\underline{Y})$-prévisible tel que $(\hat{N}\circ|\underline{\underline{Y}}|)_t<\infty$ P-p.s., $\forall\, t\in R^+$ (utiliser le fait que $\underline{X}$, $\underline{\underline{H}}$ et $\hat{\eta}\circ\underline{\underline{H}}$ sont bornées et (13)). Soit $(S_n, n \in Z^+)$ la suite de $\sigma(\underline{Y})$-temps d'arrêt qui localise ces martingales. Comme $\underline{\hat{X}}$ et $\hat{\eta} \circ \underline{\underline{H}}$ sont bornés, on a $\lim E\{\hat{X}_{t \wedge S_n} (\hat{\eta} \circ \underline{\underline{H}})_{t \wedge S_n}\} = E\{\hat{X}_t (\hat{\eta} \circ \underline{\underline{H}})_t\}$ d'où :

$$E\{\hat{X}_t (\hat{\eta} \circ \underline{\underline{H}})_t\} = E\{(\hat{N} \circ \underline{\underline{Z}})_t\} + E\{\int_0^t (\hat{\eta} \circ \underline{\underline{H}})_s \hat{f}_s \, ds\} . \tag{24}$$

Calculons maintenant $X_t (\hat{\eta} \circ \underline{\underline{H}})_t$:

$$X_t (\hat{\eta} \circ \underline{\underline{H}})_t = \int_0^t X_{s-} \, d(\hat{\eta} \circ \underline{\underline{H}})_s + \int_0^t (\eta \circ \underline{\underline{H}})_s \, dX_s . \tag{25}$$

Soit

$$(26)\quad X_t (\hat{\eta} \circ \underline{\underline{H}})_t = \int_0^t X_{s-} \, d(\eta \circ \underline{\underline{H}})_s + \int_0^t X_{s-} \, d(N \circ \underline{\underline{H}})_s - \int_0^t X_{s-} \, d(\hat{N} \circ \underline{\underline{H}})_s + \int_0^t (\hat{\eta} \circ \underline{\underline{H}})_s f_s \, ds$$

$$+ \int_0^t ((\hat{\eta} \circ \underline{\underline{H}})_{s-} + \Delta(\hat{N} \circ \underline{\underline{H}})_s) dm_s + \sum_{n \geq 1} H(T_n, \tilde{Y}_n) \, \Delta m_{T_n} \, I(T_n \leq t)$$

Comme $\underline{X}$ et $\underline{\underline{H}}$ sont bornés et que $E\{\int_0^t |dm_s|\} < \infty$ ($\underline{\Delta m}$ borné), on voit que le 1er et le 5me terme de (26) sont des martingales (proposition 2 de [1]). D'où en utilisant (15), (16) et (17) :

$$(27)\quad E\{X_t (\hat{\eta} \circ \underline{\underline{H}})_t\} = E\{\int_{]0,t] \times E'} H(s,y)(X^{(1)}(s,y) - X^{(2)}(s,y) + \Psi(s,y)) \, \hat{N}(ds,dy)\}$$

$$+ E\{\int_0^t (\hat{\eta} \circ \underline{\underline{H}})_s f_s \, ds .$$

c) En égalant $E\{X_t (\hat{\eta} \circ \underline{\underline{H}})_t\}$ et $E\{\hat{X}_t (\hat{\eta} \circ \underline{\underline{H}})_t\}$ et comme $E\{\int_0^t (\hat{\eta} \circ \underline{\underline{H}}) f_s \, ds\} = = E\{\int_0^t (\hat{\eta} \circ \underline{\underline{H}})_s \hat{f}_s \, ds\}$, on obtient

$$(28)\quad E\{\int_{]0,t] \times E'} H(s,y)(K(s,y) + \Delta(\hat{N} \circ \underline{\underline{K}})_s) \, \hat{N}(ds,dy)\}$$

$$= E\{\int_{]0,t] \times E'} H(s,y)(X^{(1)}(s,y) - X^{(2)}(s,y) + \Psi(s,y)) \, \hat{N}(ds,dy)\}$$

où $\underline{\underline{H}}$ est, rappelons-le, $\sigma(\underline{Y})$-prévisible, positif, borné, tel que $\hat{\eta}\circ\underline{\underline{H}}$ soit une martingale bornée. Les $\underline{\underline{H}}$ de la forme $H(t,\omega,y) = 1_{[\![0,T]\!]}(t,\omega)\, 1_B(y)$, T $\sigma(\underline{Y})$-temps d'arrêt, $B\in\mathcal{E}'$ satisfont à ces conditions et engendrent $P(\sigma(\underline{Y}))\otimes\mathcal{E}'$. On a donc :

$$(29)\qquad K(t,y) + \Delta(\hat{N}\circ\underline{\underline{K}})_t = X^{(1)}(t,y) - X^{(2)}(t,y) + \Psi(t,y) = U(t,y)$$

sur N^c où N est un ensemble ν-négligeable de $\mathcal{P}(\sigma(\underline{Y}))\otimes\mathcal{E}'$. D'après (29), on a :

$$(30)\qquad (\hat{N}\circ\underline{\underline{K}})_t + \sum_{s\leq t}\Delta(\hat{N}\circ\underline{\underline{K}})_s\,\hat{N}(\{s\}\times E') = (\hat{N}\circ\underline{\underline{U}})_t\ ,$$

donc, sur N^c :

$$(31)\qquad K(t,y) = U(t,y) - \Delta(\hat{N}\circ\underline{\underline{U}})_t/(1+\hat{N}(\{t\}\times E'))\ .$$

D'où finalement :

THEOREME . <u>En dehors d'un évènement</u> N <u>de</u> $(R^+\times\Omega\times E', \mathcal{P}(\sigma(\underline{Y}))\otimes\mathcal{E}')$ <u>de</u> ν-<u>mesure nulle</u> :

$$(32)\qquad K(t,y) = U(t,y) - \Delta(\hat{N}\circ\underline{\underline{U}})_t/(1+\hat{N}(\{t\}\times E'))$$

<u>où</u> $\underline{\underline{U}} = X^{(1)} - X^{(2)} + \underline{\underline{\Psi}}$, $\underline{\underline{X}}^{(1)}$, $\underline{\underline{X}}^{(2)}$, $\underline{\underline{\Psi}}$ <u>étant définis à une</u> ν-<u>équivalence près par</u> (15), (16), (17) .

d) Le gain (32) n'est défini pour l'instant qu'en dehors de N . Nous allons voir que ça n'a pas d'importance. En effet, la coupe $N(\omega)$ de N par ω a P-p.s. une mesure nulle pour les mesures sur $(R^+\times E', \mathcal{B}(R^+)\otimes\mathcal{E}')$ dans $(R,\mathcal{B}(R))$: $h\to\sum_{0<T_n} h(T_n,\tilde{Y}_n)$ et $h\to\int_{]0,\infty[\times E'} h(s,y)\,\hat{N}(ds,dy)$. Or $\underline{\underline{K}}$ n'intervient dans le préfiltre $(\hat{\mathcal{P}})$ que par $\sum_{\substack{n\geq 1\\ T_n\leq t}} K(T_n,\tilde{Y}_n)$ et $\int_{]0,t]\times E'} K(s,y)\,\hat{N}(ds,dy)$. On peut

donc prendre pour $K(t,\omega,y)$ une valeur quelconque sur N sans changer le pré-filtre qui s'écrit donc :

$$\hat{X}_t = E\{X_o\} + \int_0^t \hat{f}_s \, ds + \sum_{T_n \le t} \Big(U(T_n, \tilde{Y}_n) - \Delta(\hat{N} \circ \underline{\underline{U}})_{T_n} / (1 + \hat{N}(\{T_n\} \times E')) \Big) \tag{33}$$

$$- \int_{[0,t] \times E'} [U(s,y) - \Delta(\hat{N} \circ \underline{\underline{U}})_s / (1 + \hat{N}(\{s\} \times E'))] \hat{N}(ds,dy)$$

où $\underline{\underline{U}} = \underline{\underline{X}}^{(1)} - \underline{\underline{X}}^{(2)} + \underline{\underline{Y}}$, $X^{(1)}$, $X^{(2)}$ et $\underline{\underline{Y}}$ sont définis par (15), (16) et (17) (à un ensemble de ν-mesure nulle près).

Remarque.

Le lecteur familier avec la théorie des martingales constatera que les équations du filtre sont valables même lorsque $\underline{m}$ n'est pas à variations bornées. En effet, d'après DOLEANS-MEYER [1] :

$$X_t(\hat{\eta} \circ \underline{\underline{H}})_t = \int_0^t X_{s-} \, d(\hat{\eta} \circ \underline{\underline{H}})_s + \int_0^t (\hat{\eta} \circ \underline{\underline{H}})_s f_s \, ds$$

$$+ \int_0^t (\hat{\eta} \circ \underline{\underline{H}})_{s-} \, dm_s + [\hat{\eta} \circ \underline{\underline{H}}, \underline{m}]_t$$

et comme $\hat{\eta} \circ \underline{\underline{H}}$ est une somme compensée de sauts $[\hat{\eta} \circ \underline{\underline{H}}, \underline{m}]_t = \sum_{s \le t} \Delta m_s \, \Delta(\hat{\eta} \circ \underline{\underline{H}})_s$.
D'où $\int_0^t (\hat{\eta} \circ \underline{\underline{H}})_{s-} \, dm_s + [\hat{\eta} \circ \underline{\underline{H}}, \underline{m}]_t = \int_0^t ((\hat{\eta} \circ \underline{\underline{H}})_{s-} + \Delta(\hat{N} \circ \underline{\underline{H}})_s) \, dm_s + \sum_{T_n \le t} \Delta m_{T_n} \, H(T_n, \tilde{Y}_n)$.
Le reste suit exactement comme dans le cas " $\underline{m}$ à variation bornée".

En ce qui concerne les applications concrètes de la théorie ci-dessus, nous renvoyons à [10] où on trouvera d'ailleurs une présentation un peu plus simple en ce sens que les exemples traités (estimation de l'état d'une file d'attente par rapport à sa sortie ; estimation du temps de "désordre" d'une machine par rapport au "processus des plaintes des usagers") ne font intervenir que des observations du type processus de comptage. Les équations obtenues sont les analogues des équations de FUJISAKI-KALLIANPUR-KUNITA [5] en des termes très généraux. Dans [12], on trouvera des exemples conduisant à des équations qui ressemblent plus à celles de [5], en prenant $\underline{X} = (X_t) = (f \circ x_t)$ où $(\underline{x}, P^{\bullet}, \underline{\mathcal{X}})$ est un processus de Markov homogène, fellérien, à valeurs dans E métrique compact, et où f est un élément du domaine $\mathcal{D}(A)$ du générateur infinitésimal A, i.e. est tel que le processus

$$m_t = f(x_t) - f(x_o) - \int_0^t Af(x_s)ds$$

soit une $(P^{\mu}, \underline{\mathcal{X}})$-martingale pour toute loi initiale μ.

R E F E R E N C E S

[1] C. DOLEANS, P.A. MEYER — Intégrales stochastiques par rapport aux martingales locales.
In Sém. Prob. Strasbourg IV, Lect. Notes in Math. 124, Springer-Verlag, Berlin (1970).

L'origine historique du filtrage récursif est le fameux papier :

[2] R.E. KALMAN, R.S. BUCY — New results in linear filtering and prediction theory.
Trans. Amer. Soc. Mech. Eng. Séries D, J. Basic Eng., 83 (1961).

On trouvera un exposé des méthodes modernes (utilisant les martingales) du filtrage récursif pour le cas où le bruit d'observation est un Wiener, dans le chapitre 4 de :

[3] E. WONG — Stochastic Processes in information theroy and dynamical Systems.
Mc Graw-Hill, séries in Syst. Sciences (1972).

La méthode des semi-martingales a été utilisée dans les 2 articles suivants qui ne traitent cependant que des martingales qui sont des intégrales stochastiques de processus de Wiener :

[4] A.V. BALAKRISHNAN — A martingale approach to linear recursive state estimation.
SIAM J. of Control, 10 (4) (1972).

[5] FUJISAKI, KALLIANPUR, KUNITA — Stochastic Différential equation for the non linear filtering problem.
Osaka Math. J., 9 (1) (1972).

La portée plus générale de la méthode des semi-martingales a été montrée dans les 3 thèses et rapports suivants :

[6] M.H.A. DAVIS — Non-linear filtering with point process observation.
Research Report 73/8 (1973), Dept of Comp. and Control, Imperial College, London.

[7] A. SEGALL — A martingale approach to modeling, estimation and detection of jump processes.
Ph. D. Dissertation, Dept of El. Eng., U. of Stanford (1973) ; Tech. Report 7050-21, Center for Systems research, U. of Stanford.

[8] J. VAN SCHUPPEN — Estimation theory for continuous time process, a martingale approach.
Ph. D. Dissertation, Dept of E.E.C.S., U. of Cal., Berkeley (1973) ; Memo. ERL M-405, El. Res. Lab., Dept of E.E.C.S., U. of Cal., Berkeley.

L'application de la méthode des semi-martingales au cas où l'observation est un processus de sauts a été faite, dans le cas des noyaux continus par

[9] K. BOEL, P. VARAIYA, E. WONG — Martingales and jump processes.
Part II : Applications, Memo ERL M-409, Elect. Res. Lab.,Dept. of E.E.C.S., U. of Cal. Berkeley (1973)

où des équations différentes de celles qu'on trouve dans cet exposé (et sans doute moins directement utilisables) sont dérivées. Deux exemples explicites menés "jusqu'au bout" sont traités dans :

[10] P. BREMAUD — Estimation de l'état d'une file d'attente et du temps de panne d'une machine par la méthode des semi-martingales.
A paraître dans Advances in Applied Probability (1975).

On trouvera aussi des exemples dans :

[11] M.H.A. DAVIS, T. KAILATH, A. SEGALL — Non-linear Filtering with Counting Observations.
A paraître dans IEEE Transactions on Information Theory (1975).

D'autres développements sont indiqués dans :

[12] P. BREMAUD, J.JACOD — Revue des résultats récents sur les systèmes dynamiques où interviennent les processus ponctuels marqués : modélisation par les martingales, estimation, contrôle, indentification et information.
En préparation.

En ce qui concerne la représentation des martingales, les deux références utiles dans le cas où les sauts T_n sont rangeables en ordre croissant, et les noyaux "quelconques" :

[13] C.S. CHOU, P.A. MEYER — Sur la représentation des martingales comme intégrales stochastiques dans les processus ponctuels.
A paraître in Sém. Prob. VIII, Springer-Verlag, Berlin (1974)

et pour le cas général des processus ponctuels marqués, par une méthode différente :

[14] J. JACOD — Multivariate point processes : predictable projection, Radon-Nikodym derivatives, representation of martingales.
A paraître dans Z. für Wahrscheinlichkeitstheorie.