COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL
ET SES APPLICATIONS

NICE du 26 au 30 AVRIL 1977

FILTRAGE NON LINEAIRE APPROCHE *

Michel FLIESS

Université Paris VIII & Center for Mathematical System Theory, University of Florida at Gainesville

RESUME

Le filtrage stochastique non linéaire reste, en dépit d'efforts tenaces, dans l'enfance. Kailath [7] a, voici peu, exprimé une opinion à laquelle nous souscrivons:

"At the moment, one of the chief benefits of having attacked the nonlinear problem was that it brought to the fore certain difficulties associated with the proper definition of the differential and integral equations used to describe nonlinear operations on white noise....I believe that the situation is somewhat analogous to that of linear filtering in the mid-fifties, when the field was rapidly grinding to a halt amid a welter of numerous attempts at direct extensions of the Wiener filter. The Kalman filter provided a new impulse that moved things out of the doldrums into a new fruitful direction. Similarly in nonlinear filtering it may be that attempts to solve the nonlinear filtering problem along the lines of the successful Kalman linear filter are misdirected. Some new approach needs to be uncovered."

Récemment, l'auteur [3,4] et Sussmann [12] ont prouvé, dans un cadre déterministe, que tout asservissement continu peut être arbitrairement approché par des systèmes réguliers, ou bilinéaires. Ce résultat est ici étendu à des entrées stochastiques très générales. L'intégrale stochastique à laquelle on aboutit, par l'utilisation des séries formelles non commutatives, n'est pas celle d'Itô, mais une généralisation de celle de Stratonovich. Notre théorie est causale alors que celle de Wiener [14] ne l'était pas.

Le filtrage non linéaire sous-optimal ne donnerait ainsi plus lieu à des équations différentielles stochastiques complexes. Il reviendrait à identifier des séries ou polynômes non commutatifs, ou des systèmes réguliers.

SUMMARY

Recently, the author [3,4] and Sussmann [12] have proved, in a deterministic setting, that any continuous system may be approximated by regular (or bilinear) ones. This result is here extended for a wide class of stochastic inputs. The stochastic integral to which we are led by the use of noncommutative formal power series is not Itô's integral but Stratonovich's. Contrary to that of Wiener [14], our theory is causal.

Using the above result, suboptimal nonlinear filtering should no more be related to complicated stochastic differential equations, but should lead to the identification of noncommutative formal power series or polynomials, or of regular systems.

---

* Ce travail a été effectué sous l'auspice de l'U.S. Air Force Grant 76-3034 et de la convention de recherche 76 188 de l'IRIA-SESORI.

I.-Rappels

a)Séries formelles:Soit X^* le monoïde libre engendré par l'alphabet fini,non vide,X.L'élément neutre de X^*,le mot vide,est noté 1.Soient $\underline{R}\langle X\rangle$ et $\underline{R}\langle\langle X\rangle\rangle$ les $\underline{R}$-algèbres des polynômes et des séries formels,à coefficients réels,en les variables associatives $x \in X$ (non commutatives si card $X \geqslant 1$).Un élément $s \in \underline{R}\langle\langle X\rangle\rangle$ est noté:

$$s = \sum \left\{ (s,w)w \mid w \in X^* \right\} \quad [(s,w) \in \underline{R}] \quad .$$

Une série $s \in \underline{R}\langle\langle X\rangle\rangle$ est inversible ssi $(s,1) \neq 0$. Une sous-algèbre R de $\underline{R}\langle\langle X\rangle\rangle$ est dite rationnellement close ssi l'inverse de tout élément inversible de R y appartient encore.L'algèbre $\underline{R}\langle(X)\rangle$ des séries rationnelles est la plus petite sous-algèbre rationnellement close de $\underline{R}\langle\langle X\rangle\rangle$,contenant $\underline{R}\langle X\rangle$.

Soit ${}^N\underline{R}^M$ l'ensemble des matrices à N lignes et M colonnes.Par simplification,${}^N\underline{R}^1$ et ${}^1\underline{R}^M$ sont notés ${}^N\underline{R}$ et $\underline{R}^M$.Une représentation $\rho : X^* \to {}^N\underline{R}^N$ du monoïde libre par des matrices carrées d'ordre N est un homomorphisme de X^* dans le monoïde multiplicatif sous-jacent de ${}^N\underline{R}^N$.

Théorème (Kleene-Schützenberger).-Une série $r \in \underline{R}\langle\langle X\rangle\rangle$ est rationnelle si et seulement s'il existe un entier $N \geqslant 1$,une représentation $\rho : X^* \to {}^N\underline{R}^N$,des matrices ligne $\lambda \in \underline{R}^N$ et colonne $\gamma \in {}^N\underline{R}$ tels que

$$r = \sum \left\{ (\lambda \rho w \gamma) w \mid w \in X^* \right\} \quad .$$

Remarques.-(i)Dans le cas d'une seule variable x,on retrouve les polynômes $\underline{R}[x]$ et les séries $\underline{R}[[x]]$ habituels.Les séries rationnelles sont développements de Taylor à l'origine de fractions rationnelles P/Q, $(P,Q \in \underline{R}[x], Q(0) \neq 0)$.

(ii)Rappelons que les séries rationnelles non commutatives ont été introduites par Schützenberger [9] dans le cadre de la théorie des automates.

b)Application à l'automatique non linéaire(cf. [4]): Soit $g \in \underline{R}\langle\langle X\rangle\rangle$,où $X = \{x_0, x_1, \ldots, x_n\}$ .A $u_1,\ldots,u_n$: $[0,\infty[\to \underline{R}$,associons,via g,à l'instant t,la valeur:

$$(1) \quad y(t) = (g,1) + \sum_{\alpha \geq 0} \sum_{j_\alpha,\ldots,j_0 = 0}^{n} (g, x_{j_\alpha} \ldots x_{j_0}) \int_0^t d\xi_{j_\alpha} \ldots d\xi_{j_0} \, ,$$

où $\xi_0(\tau) = \tau$, $\xi_i(\tau) = \int_0^\tau u_i(\sigma) d\sigma$ $(i=1,\ldots,n)$,et où l'intégrale itérée $\int_0^\tau d\eta_{j_\alpha} \ldots d\eta_{j_0}$ est récursivement définie par:

$$\int_0^\tau d\xi_j = \xi_j(\tau), \int_0^\tau d\xi_{j_\alpha} \ldots d\xi_{j_0} = \int_0^\tau d\xi_{j_\alpha}(\sigma) \int_0^\sigma d\xi_{j_{\alpha-1}} \ldots d\xi_{j_0} .$$

On suppose que la somme infinie (1) est convergente.

g caractérise ainsi une fonctionnelle causale.Deux fonctionnelles causales,définies par des séries non commutatives,ont toujours même valeur ssi les séries sont identiques.Une fonctionnelle causale est dite analytique ssi elle est définie par une série non commutative,appelée série génératrice de la fonctionnelle.Une fonctionnelle causale analytique est dite polynômiale ou rationnelle ssi sa série génératrice l'est.

Soit l'asservissement régulier (ou bilinéaire)

$$(2) \quad \begin{cases} \dot{q}(t) = (A_0 + \sum_{i=1}^{n} u_i(t) A_i) q(t) \\ y(t) = \lambda q(t) \quad , \end{cases}$$

où q(t) appartient à l'espace vectoriel d'état Q de dimension finie [q(0) donné],et où $A_0, A_1, \ldots, A_n : Q \to Q$, $\lambda : Q \to \underline{R}$ sont linéaires.On sait que ce type de systèmes non linéaires généralise les linéaires de la forme

$$\begin{cases} \dot{q}(t) = Fq(t) + Gu(t) \\ y(t) = Hq(t) \quad . \end{cases}$$

En vertu de la formule de Peano-Baker,(2) détermine une fonctionnelle causale analytique,de série génératrice rationnelle donnée par

$$\lambda q(0) + \sum_{\alpha \geq 1} \sum_{j_\alpha,\ldots,j_0 = 0}^{n} \lambda A_{j_\alpha} \ldots A_{j_0} q(0) x_{j_\alpha} \ldots x_{j_0} \quad .$$

Rappelons que les séries génératrices rationnelles jouent quant à la réalisation des asservissements réguliers le même rôle que les fonctions de transfert rationnelles pour les linéaires stationnaires.

Supposons l'espace fonctionnel $\mathfrak{F}$ des entrées $\{u_i\}$ localement L^α $(1 \leq \alpha \leq \infty)$ et muni de la topologie correspondante.Une fonctionnelle causale est dite continue si c'est une application continue $[0,\infty[\times \mathfrak{F} \to \underline{R}$.

Théorème (cf. [3,4]).-Toute fonctionnelle causale continue peut,dans un compact de $[0,\infty[\times\mathfrak{S}$,être arbitrairement approchée par une fonctionnelle causale analytique,que l'on peut choisir polynômiale ou rationnelle.

Corollaire.-Toute fonctionnelle causale continue peut, dans un compact de $[0,\infty[\times\mathfrak{S}$,être arbitrairement approchée par un asservissement régulier de dimension finie.

Sous cette deuxième forme,le résultat a été trouvé indépendamment par Sussmann [12].

A la suite des travaux de Volterra et Wiener,on a souvent,en ingéniérie,tenté de modéliser les systèmes non linéaires par des séries de Volterra,ainsi définies dans le cas d'une entrée scalaire (cf. Barrett [1]):

$$y(t) = h_0(t) + \int_0^t h_1(t,\tau_1)u(\tau_1)d\tau_1 + \int_0^t\int_0^t h_2(t,\tau_2,\tau_1)u(\tau_2)u(\tau_1)d\tau_2 d\tau_1 + \dots .$$

Le lien entre cet outil classique et nos séries non commutatives est frappant.

Théorème (cf. [3,4]).-Une série de Volterra définit une fonctionnelle causale analytique si et seulement si,pour tout $k \geq 0$,le noyau $h_k(t,\tau_k,\dots,\tau_1)$ est une fonction analytique de $t,\tau_k,\dots,\tau_1$.

II.-Séries de Chen et processus stochastiques

a)Définition des séries de Chen:Soit $\{\xi_0(\tau),\xi_1(\tau),\dots,\xi_n(\tau) \mid 0\leq\tau\leq t\}$ un chemin continu,à variations bornées, d'un espace affine réel de dimension n+1.Chen [2] lui a associé la série de $\underline{R}\langle\langle X\rangle\rangle$ ($X=\{x_0,x_1,\dots,x_n\}$):

$$(3)\qquad 1 + \sum_{\alpha\geq 1}\sum_{j_\alpha,\dots,j_0=0}^{n} x_{j_\alpha}\dots x_{j_0}\int_0^t d\xi_{j_\alpha}\dots d\xi_{j_0} \quad ,$$

où l'intégrale itérée $\int_0^t d\xi_{j_\alpha}\dots d\xi_{j_0}$ est,comme précédemment,définie récursivement:

$$\int_0^\tau d\xi_j = \xi_j(\tau) - \xi_j(0) \quad ,$$

$$\int_0^t d\xi_{j_\alpha}\dots d\xi_{j_0} = \int_0^t d\xi_{j_\alpha}(\tau)\int_0^\tau d\xi_{j_{\alpha-1}}\dots d\xi_{j_0} \quad .$$

(3) est la série de Chen du chemin,qu'elle détermine à une translation près.La série de Chen de deux chemins mis bout à bout est le produit de leurs séries de Chen.Le logarithme d'une série de Chen est un élément de Lie.

A l'entrée $\{u_1(\tau),\dots,u_n(\tau) \mid 0\leq\tau\leq t\}$,associons le chemin

$$\{\xi_0(\tau) = \tau,\ \xi_1(\tau) = \int_0^\tau u_1(\sigma)d\sigma,\dots,\ \xi_n(\tau) = \int_0^\tau u_n(\sigma)d\sigma \mid 0\leq\tau\leq t\} \quad ,$$

et la série de Chen correspondante.La valeur (1) ou, si l'on préfère,la sortie sont obtenues par dualité canonique entre cette série de Chen et la série génératrice.

b)Topologies pour les séries de Chen et entrées généralisées (cf. [6]):Comme Sussmann [12] l'a souligné, il est nécessaire,avec les systèmes non linéaires,de sortir du cadre de la théorie des distributions de L. Schwartz pour représenter des entrées généralisées, comme l'impulsion de Dirac ou le bruit blanc.Les séries de Chen n'ont été introduites que pour cela.

Pour assurer la continuité d'une fonctionnelle causale analytique,la topologie la moins fine sur les séries de Chen est,selon la nature de la série génératrice g:

- T_P définie par l'ensemble des semi-normes d'indice $w\in X^*$: $\|s\|_w = |(s,w)|$,si $g\in\underline{R}\langle X\rangle$;
- T_E définie par l'ensemble des normes d'indice $M = 1, 2,\dots$:

$$|||s|||_M = \sum\{|(s,w)|\,M^{|w|} \mid w\in X^*\} \quad ,$$

où $|w|$ désigne la longueur,si les coefficients de g croissent au plus exponentiellement,ce qui est le cas des séries rationnelles.

Soient $\overline{C}_P$ et $\overline{C}_E$ les complétés,pour ces topologies, de l'ensemble des séries de Chen obtenues à partir d'entrées localement Lebesgue-intégrables.

Application.-Pour k infini,la série de Chen de l'entrée $\{d_i^{(k)}(\tau) = a_i k \mid 0\leq\tau\leq 1/k\}$ ($i=1,\dots,n$; $a_i\in\underline{R}$) a pour limite,relativement à T_P et T_E,$\exp(\sum_{i=1}^{n} a_i x_i)$.Cette représentation de l'impulsion de Dirac donne, si on l'applique à une série génératrice g, comme sortie à l'instant zéro:

$$(g,1) + \sum_{\alpha\geq 0} \sum_{i_\alpha,\ldots,i_0=1}^{n} (g, x_{i_\alpha}\ldots x_{i_0}) \frac{a_{i_\alpha}\ldots a_{i_0}}{(\alpha+1)!} \quad .$$

Si on l'applique au système régulier (3), il vient:

$$\lambda \exp\left(\sum_{i=1}^{n} a_i A_i\right) q(0) \quad .$$

c) Processus stochastiques: Les hypothèses, fort larges, sont inspirées de McShane [8]. Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité, $[0,T]$ un intervalle fermé ($T>0$) avec:

(i) $\{\mathcal{B}_\tau \mid 0 \leq \tau \leq T\}$ est une famille de σ-algèbres telle que, si $0 \leq \tau \leq \tau' \leq T$, $\mathcal{B}_\tau \subset \mathcal{B}_{\tau'}$;

(ii) le processus vectoriel considéré $\{\eta_1(\tau,\omega), \ldots, \eta_n(\tau,\omega)\}$ est adapté à $\mathcal{B}_\tau$ et satisfait à :

$$E[(\eta_{i_\alpha}(\tau') - \eta_{i_\alpha}(\tau))\ldots(\eta_{i_0}(\tau') - \eta_{i_0}(\tau)) \mid \mathcal{B}_\tau] < K^{\alpha+1}(\tau'-\tau)$$

$(0 \leq \tau \leq \tau' \leq T ; \alpha = 0,1,2,\ldots ; K \text{ cte} > 0)$.

Parmi bien d'autres, les processus de Wiener-Lévy et de Poisson vérifient ces conditions.

Considérons une subdivision $0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_k \leq t$ de $[0,t]$, $0 \leq t \leq T$, et la série de Chen:

$$(4) \quad \exp[(t-\tau_k)x_0 + \sum_{i=1}^{n} (\eta_i(t) - \eta_i(\tau_k))x_i]\ldots \exp[\tau_1 x_0 + \sum_{i=1}^{n} (\eta_i(\tau_1) - \eta_i(0))x_i] \quad .$$

<u>Théorème</u>.-Lorsque le pas de la subdivision tend vers zéro, (4) a, relativement à T_P et T_E, une limite en moyenne quadratique.

La preuve est une adaptation de celles des théorèmes 2.3 de McShane [8, p. 63] et 1 de [5]. La limite est notée:

$$(5) \quad 1 + \sum_{\alpha\geq 0} \sum_{j_\alpha,\ldots,j_0=0}^{n} x_{j_\alpha}\ldots x_{j_0} \int_0^t d\eta_{j_\alpha}\ldots d\eta_{j_0} \quad ,$$

où $\eta_0(\tau) = \tau$. L'intégrale itérée stochastique $\int_0^t d\eta_{j_\alpha}\ldots d\eta_{j_0}$ est définie, en moyenne quadratique, comme le coefficient du mot $x_{j_\alpha}\ldots x_{j_0}$ dans la limite (5). Lorsque l'on a un processus de Wiener-Lévy, on retrouve l'intégrale symétrique de Stratonovich [11] et non celle d'Itô (cf. [5]).

En raison de (1), (5) est appelée <u>trajectoire</u> du "bruit" associé à $\{\eta_i(\tau)\}$, c'est-à-dire de la "dérivée formelle" $\{\dot\eta_i(\tau)\}$. Dans le cas d'un processus de Wiener-Lévy, on obtient ainsi une définition nouvelle du bruit blanc gaussien (cf. [5]). Il en va de même pour processus de Poisson et bruit blanc poissonien.

<u>Remarque</u>.-Pour un processus de Wiener-Lévy, la limite (5) est obtenue presque sûrement (cf. [5]), ce qui conduit, à la suite de Sussmann [12,13], à une interprétation "non probabiliste" des fonctionnelles causales rationnelles et des asservissements réguliers bruités.

<u>III.-Filtrage non linéaire approché</u>

a) <u>Théorème d'approximation</u>: Soit $E(F_t)$ l'espérance à l'instant t, si elle existe, d'une fonctionnelle causale stochastique du processus $\{\eta_i(\tau)\}$. Soient I un intervalle de la forme $[0,T]$, $T<\infty$, ou $[0,\infty[$, m une mesure positive complète sur I, de masse finie, telle que:

- les polynômes soient denses dans $L^2(I,m)$;
- pour toute fonctionnelle causale analytique polynômiale P, $E(P_t)$ soit m-intégrable sur I.

Soit $H_{I,m}$ l'ensemble des fonctionnelles causales F, telles que $\int_I E(F_t^2)dm < \infty$, modulo l'équivalence $\int_I E(F'_t - F''_t)^2 dm = 0$.

<u>Proposition</u>.-$H_{I,m}$, qui contient les fonctionnelles analytiques polynômiales, est un hilbert relativement au produit scalaire $\int_I E(F'_t F''_t)dm$.

<u>Théorème</u>.-Les fonctionnelles causales analytiques polynomiales sont denses dans $H_{I,m}$.

La démonstration reprend celle du théorème 3 de [5] en utilisant, ici, le lemme 2.1 d'approximation de Segal [10].

<u>Remarques</u>.-(i) En [5], le théorème d'approximation est énoncé pour les seules fonctionnelles du bruit blanc gaussien.

(ii) Dans de nombreux cas, les fonctionnelles causales rationnelles appartiennent à $H_{I,m}$. Elles y sont alors denses, de même que les asservissements réguliers de dimension finie.

b) Filtrage non linéaire sous-optimal: Nous ne faisons qu'esquisser les calculs nécessaires pour obtenir les filtres sous-optimaux. En particulier, nous n'abordons pas la détermination des systèmes sous forme régulière.

Soit $z(t)$ un signal supposé, pour simplifier, scalaire. Il s'agit de trouver, dans $H_{I,m}$, la fonctionnelle F minimisant:

$$\int_I E(z(t) - F_t)^2 dm \quad . \tag{6}$$

Remarque.- La validité de (6) et le choix de la mesure m ne sont pas étudiées ici.

En vertu du théorème d'approximation, restreignons-nous à trouver la fonctionnelle polynomiale P, de polynôme générateur p de degré inférieur ou égal à d, minimisant:

$$\int_I E(z(t) - P_t)^2 dm = \int_I E[z(t)^2 - 2z(t)P_t + P_t^2] dm \quad . \tag{7}$$

La fonctionnelle P^2 est aussi polynomiale, de polynôme générateur $p ш p$, où ш est le produit de Hurwitz, ou "shuffle product" (cf. [4]). En différentiant (7) successivement par rapport à chacun des coefficients inconnus de p, on obtient un système d'équations linéaires pour les calculer.

Un inconvénient grave de la méthode précédente est le suivant: il faut entièrement recommencer les calculs si l'on veut un filtre plus précis donné par un polynôme de degré $d' > d$. Ce n'est plus le cas une fois déterminée une base orthonormée de $H_{I,m}$. Pour cela, on peut appliquer la méthode de Gram-Schmidt à partir d'un ordre total sur le monoïde libre X^*.

Remarques.- (i) A notre connaissance, la théorie de Volterra-Wiener ne conduit pas à un théorème d'approximation stochastique causal. Elle n'en a pas moins été utilisée pour le filtrage sous-optimal.

(ii) Barrett [1], par exemple, aboutit, pour construire le filtre, à des équations intégrales fort complexes.

Conclusion.- Les résultats d'approximations déterministe et stochastique laissent entrevoir de vastes applications pour les séries non commutatives et les asservissements réguliers. D'autant plus que les variables non commutatives conduisent à des manipulations combinatoires aisées à programmer.

Bibliographie

[1] Barrett (J.F.).-The use of functionals in the analysis of non-linear physical systems. J. Electron. Control, 15, 1963, p. 567-615.

[2] Chen (K.T.).-Integration of paths, a faithful representation of paths by non-commutative formal power series. Trans. Amer. Math. Soc., 89, 1958, p. 395-407.

[3] Fliess (M.).-Séries de Volterra et séries formelles non commutatives. C. R. Acad. Sc. Paris, A-280, 1975, p. 965-967.

[4] Fliess (M.).-Un outil algébrique: les séries formelles non commutatives. In Mathematical Systems Theory (G. Marchesini et S.K. Mitter, éd.), p. 122-148, Lect. Notes Econom. Math. Syst., vol. 131, Springer-Verlag, Berlin, 1976.

[5] Fliess (M.).-Intégrales itérées de K.T. Chen, bruit blanc gaussien et filtrage non linéaire. C. R. Acad. Sc. Paris, A-284, 1977.

[6] Fliess (M.) et Jacob (G.).-Topologies pour certaines fonctions de lignes non linéaires; application aux asservissements. C. R. Acad. Sc. Paris, A-282, 1976, p. 321-324.

[7] Kailath (T.).-A view of three decades of linear filtering theory. IEEE Trans. Inform. Th., 20, 1974, p. 146-181.

[8] McShane (E.J.).-Stochastic calculus and stochastic models. Academic Press, New York, 1974.

[9] Schützenberger (M.P.).-On the definition of a family of automata. Inform. Control, 4, 1961, p. 245-270.

[10] Segal (I.E.).-Tensor algebras over Hilbert spaces. I. Trans. Amer. Math. Soc., 81, 1956, p. 106-134.

[11] Stratonovich (R.L.).-Conditional Markov processes and their applications to the theory of optimal control. Moscou, 1966. Trad. anglaise: Elsevier, New York, 1968.

[12] Sussmann (H.J.).-Semigroup representations, bilinear approximation of input-output maps, and generalized inputs. In Mathematical Systems Theory (G.

Marchesini et S.K. Mitter,éd.),p. 172-191,Lect. Notes Econom. Math. Syst.,vol. 131,Springer-Verlag,Berlin, 1976.

[13] Sussmann (H.J.).-On generalized inputs and white noise.Proc. 1976 IEEE Conf. Decision Control, p. 809-814,Clearwater (FL),Dec. 1976.

[14] Wiener (N.).-Nonlinear problems in random theory. Wiley,New York,1958.

Adresse postale: 38,rue Godefroy-Cavaignac,
75011 Paris.