SEPTIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 28 MAI au 2 JUIN 1979

QUELQUES REMARQUES SUR LE "PARADOXE" DE WONG ET ZAKAI ; APPLICATION A L'IDENTIFICATION NON LINEAIRE.

Michel FLIESS

Université Paris VIII et Laboratoire des Signaux et Systèmes, CNRS-ESE, Plateau du Moulon, 91190 GIF/YVETTE

RESUME

La modélisation par équations différentielles stochastiques d'Itô

$$dq = A_o(q) + A_1(q)db_1$$

de systèmes dynamiques bruités se heurte au fait que les trajectoires physiquement réalisables ont, mathématiquement, une probabilité nulle. Wong et Zakai ont jeté les premières lumières sur cette situation embarrassante en montrant la convergence d'approximations régulières vers l'équivalent de Stratonovich.

Des phénomènes intéressants, encore incompris, apparaissent avec deux bruits (ou plus), c'est-à-dire quand A_1db_1 est remplacé par $A_1db_1+A_2db_2$. Il est important de noter que ce manque de robustesse du calcul stochastique ne tient pas à la théorie des probabilités, mais a une nature géométrique. Il est relié aux propriétés de l'algèbre de Lie engendrée par les champs de vecteurs A_1 et A_2. Les variables non commutatives permettent de résoudre le problème dans le cas simple des systèmes réguliers (ou bilinéaires).

SUMMARY

In using the Itô stochastic differential equations

$$dq = A_o(q)\ dt + A_1(q)\ db_1$$

as a model for a dynamical system perturbed by white noise, one has to face the fact that realistic, i.e. smooth trajectories, occur, mathematically, with probability zero. The pionnering work in clarifying this apparently embarrassing situation was done by Wong and Zakai, who demonstrated convergence of sequences of smooth approximations to the Stratonovich version.

Interesting things, which are not yet understood, happen when there are two (or more) noises, i.e. when A_1db_1 is replaced by $A_1db_1+A_2db_2$. It is important to notice that this failure of robustness in stochastic calculus has not to do with probability theory, but has a geometrical nature. It is related to the properties of the Lie algebra generated by the vector fields A_1 and A_2. By using non-commutative variables, we solve the problem for a restricted class of non-linear systems, i.e. the regular (or bilinear) ones.

QUELQUES REMARQUES SUR LE "PARADOXE" DE WONG ET ZAKAI ; APPLICATION A L'IDENTIFICATION NON LINEAIRE.

Michel FLIESS

I.- INTRODUCTION

Le bruit blanc étant une idéalité mathématique, il faut s'interroger sur le comportement d'une équation différentielle stochastique lorsqu'on prend pour entrées des approximations. Ce problème de robustesse conditionne, à l'évidence, la validité pratique de la théorie.

Or ces questions n'ont reçu, à la suite de Wong et Zakai (14,15,16), de réponse satisfaisante que pour une bruit scalaire. Le cas vectoriel semble dépasser les possibilités actuelles des techniques disponibles. C'est à une rapide revue de ces difficultés qu' est consacrée cette communication. A la suite d'autres auteurs, on insiste sur leur nature avant tout d'ordre géométrique.

Cependant, le problème reçoit une solution complète dans le cas simple des systèmes réguliers (ou bilinéaires). L'outil privilégié est constitué par les séries formelles en plusieurs variables non commutatives.

II.- UN EXEMPLE DE DIFFICULTES EN IDENTIFICATION NON LINEAIRE (Lobry (9)).

Soit l'équation différentielle stochastique dans $\underline{R}^2$:

$$(1)\quad \begin{cases} dq_1 = \Theta dt + q_2 db_1 \\ dq_2 = db_2, \end{cases}$$

où $q_1(0)=q_2(0)=0$. b_1 et b_2 sont deux browniens standard indépendants. On suppose savoir a priori que le paramètre Θ ne peut prendre que deux valeurs 0 et $\Theta_o>0$. Il n'est pas nécessaire de préciser si (1) est prise au sens d'Itô ou de Stratonovich, puisqu'ici ils coïncident. La variable aléatoire $q_1(1)$ a pour espérance Θ et pour variance 1/2.

On désire rejeter avec un risque inférieur à $\varepsilon>0$ l'hypothèse $\Theta=0$ sur la base de N observations $y_1, \ldots, y_N$ de $q_1(1)$. La théorie élémentaire des tests d'hypothèses permet d'établir le résultat suivant :

Il existe une fonction $\phi(\varepsilon,N)$ à valeurs réelles positives telle que la règle de décision

si $\dfrac{y_1+\ldots+y_N}{N} \geq \phi(\varepsilon,N)$, *on rejette l'hypothèse $\Theta=0$, sinon on l'accepte,*

définisse un test le plus puissant de rejet de l'hypothèse $\Theta=0$ contre l'hypothèse $\Theta=\Theta_o>0$. Pour tout $\varepsilon>0$ donné, la fonction $N \to \phi(\varepsilon,N)$ tend vers zéro.

Supposons maintenant que la variable aléatoire observée ne soit pas $q_1(1)$, mais une variable aléatoire $\tilde{q}_1(1)$ définie à partir de l'équation

$$(2)\quad \begin{cases} d\tilde{q}_1 = \Theta dt + \tilde{q}_2 d\tilde{b}_1 \\ d\tilde{q}_2 = d\tilde{b}_2, \end{cases}$$

où $\tilde{b}_1$ et $\tilde{b}_2$ sont des processus dont on sait qu'ils sont, en un certain sens, "proches" de b_1 et b_2. En l'absence d'information supplémentaire sur $\tilde{b}_1$ et $\tilde{b}_2$, nous allons montrer, par un contre-exemple, que l'on ne peut continuer à appliquer le précédent test d'hypothèse.

Définissons des processus $(\tilde{b}_1^{(n)}, \tilde{b}_2^{(n)})$ qui convergent, presque sûrement, uniformément vers (b_1,b_2). Pour cela, introduisons une application $T^{(n)}$ de l'ensemble $C^o([0,1],\underline{R}^2)$ des applications continues de $[0,1]$ dans $\underline{R}^2$, dans lui-même. L'application $T^{(n)}$ fait correspondre au couple de fonctions continues (u_1,u_2) le couple de fonctions polygonales, c'est-à-dire affines par morceaux, $(\tilde{u}_1^{(n)},\tilde{u}_2^{(n)})$ définies aux points $\frac{k}{n}$, $\frac{k+1/2}{n}$:

$$\begin{cases} \tilde{u}_1^{(n)}(\frac{k}{n}) = u_1(\frac{k}{n}) \ ; \ \tilde{u}_1^{(n)}(\frac{k+1/2}{n}) = \tilde{u}_1^{(n)}(\frac{k+1}{n}) = u_1(\frac{k+1}{n}) \\ \tilde{u}_2^{(n)}(\frac{k}{n}) = \tilde{u}_2^{(n)}(\frac{k+1/2}{n}) = u_2(\frac{k}{n}) \ ; \ \tilde{u}_2^{(n)}(\frac{k+1}{n}) = u_2(\frac{k+1}{n}) \end{cases}$$

si $\left(u_1(\frac{k+1}{n})-u_1(\frac{k}{n})\right)\left(u_2(\frac{k+1}{n})-u_2(\frac{k}{n})\right) \leq 0$;

$$\begin{cases} \tilde{u}_1^{(n)}(\frac{k}{n}) = \tilde{u}_2^{(n)}(\frac{k+1/2}{n}) = u_1(\frac{k}{n}) \ ; \ \tilde{u}_1^{(n)}(\frac{k+1}{n}) = u_1(\frac{k+1}{n}) \\ \tilde{u}_2^{(n)}(\frac{k}{n}) = u_2(\frac{k}{n}) \ ; \ \tilde{u}_2^{(n)}(\frac{k+1/2}{n}) = \tilde{u}_2^{(n)}(\frac{k+1}{n}) = u_2(\frac{k+1}{n}) \end{cases}$$

sinon.

Après calculs sur (2), on obtient :

$$\tilde{q}_1^{(n)}(1) = \Theta + \sum_{k=}^{n-1} b_2(\tfrac{k}{n})\left(b_1(\tfrac{k+1}{n}) - b_1(\tfrac{k}{n})\right) + \sum_{k=0}^{n-1} \psi\left((b_1(\tfrac{k+1}{n}) - b_1(\tfrac{k}{n}))(b_2(\tfrac{k+1}{n}) - b_2(\tfrac{k}{n}))\right),$$

où $(\alpha,\beta) \to \psi(\alpha,\beta)$ est la fonction qui vaut $\alpha\beta$ si le produit $\alpha\beta$ est positif, zéro sinon.

Les propriétés des intégrales gaussiennes donnent:

$$E(\tilde{q}_1^{(n)}(1)) = \Theta + 1/\pi.$$

Ceci montre que l'utilisation du test précédent conduit à rejeter l'hypothèse $\Theta=0$ avec un risque qui tend vers 1 quand N tend vers l'infini, ce qui est contraire au résultat souhaité.

Remarque - Il faut convenir que les processus $\tilde{b}_1^{(n)}, \tilde{b}_2^{(n)}$ sont peu naturels du fait de la dépendance des accroissements aux instants $\frac{k}{n}$, $\frac{k+1/2}{n}$, $\frac{k+1}{n}$.

III.- ILLUSTRATION DE LA NATURE DU PROBLEME

Considérons une équation différentielle stochastique sous la forme générale

$$(3) \quad dq = A_o(q)dt + \sum_{i=1}^{n} A_i(q)db_i,$$

où les $A_o, A_1, \ldots, A_n$ sont des champs de vecteurs. Les $b_1, \ldots, b_n$ sont des browniens indépendants. Il faut entendre (3) au sens de Stratonovich.

Généralement au cas vectoriel le résultat fameux de Wong et Zakai (15), McShane (11,12), dans ses intéressantes publications sur les équations différentielles stochastiques, a montré, sous certaines conditions techniques, que si l'on approchait les browniens par des fonctions polygonales, la solution approchée tendait vers la solution exacte. Si (3) était prise au sens d'Itô, il faut, pour la limite des solutions approchées, introduire les termes correctifs bien connus, d'où parfois le mot "paradoxe".

Si l'approximation est quelconque, il n'y a de résultats que dans des cas particuliers :

- le bruit est scalaire (n=1) (cf. Wong et Zakai (15, 16), Wong (14), McShane (11,12), Sussmann (13), Freedman et Willems (7)).
- les champs de vecteurs $A_1, \ldots, A_n$ commutent (McShane (11), Sussmann (13), Freedman et Willems (7)). Cette situation, qui contient la précédente, est, comme le constate Lobry (9,10), difféomorphe au cas linéaire. On ne peut donc s'attendre à des phénomènes non linéaires spécifiques.

Dans le cas général, on se heurte à des difficultés considérables, qui, de l'avis de McShane (11), p. 226, n'ont rien à voir avec la théorie des probabilités. Il s'agit de phénomènes non linéaires qui ont une explication géométrique, voire combinatoire. Paraphrasons Sussmann (13):

De façon heuristique, il faut considérer chaque entrée $\dot{b}_i = db_i/dt$ comme une suite infinie de sauts séparés par un intervalle de temps infinitésimal. L'effet de ces sauts sur l'état q est déterminé par les champs de vecteurs A_i. Les difficultés proviennent de l'ordre d'arrivée de ces sauts dans un intervalle infinitésimal. Il est ainsi possible de prendre deux suites de sauts, dont l'effet sur tout intervalle fini est le même, mais dont l'ordre sur tout intervalle infinitésimal diffère, de sorte qu'un effet "macroscopique" apparaît.

<u>Exemple</u> (cf. (11,4,5)) - Considérons l'équation

$$(4) \quad dq = (A_1 d\xi_1 + A_2 d\xi_2)q,$$

où A_1 et A_2 sont des matrices carrées. Posons

$$\xi_1^{(k)}(t) = \frac{\sin k^2 t}{k}, \quad \xi_2^{(k)}(t) = \frac{\cos k^2 t}{k}$$

Un calcul élémentaire montre que, pour k infini, la limite de la solution est

$$q(t) = (\exp \tfrac{t}{2}(A_1 A_2 - A_1 A_2))\, q(0)$$

On constate le crochet de Lie. Cet exemple sera repris plus tard.

<u>Commentaire</u> - Au sens de la théorie des distributions les limites de $\xi_1^{(k)}$, $\xi_2^{(k)}$, ou des entrées associées $u_1^{(k)}(t) = \dot{\xi}_1^{(k)}(t) = k \cos k^2 t$, $u_2^{(k)}(t) = \dot{\xi}_2^{(k)}(t) = -k \sin k^2 t$ sont nulles. Avec un système linéaire, la contribution limite, pour k infini, serait nulle. Il en irait de même si, en (4), on ne laissait qu'une entrée pour supprimer l'autre. C'est donc bien à l'intéraction non linéaire des deux composantes qu'est dû l'effet calculé.

IV.- <u>RESOLUTION DANS LE CAS DES SYSTEMES REGULIERS (OU BILINEAIRES)</u> (cf. (4)).

a) <u>Généralités</u>.

Rappelons qu'un système régulier (ou bilinéaire) est de la forme

$$(5) \quad \begin{cases} \dot{q}(t) = (A_o + \sum_{i=1}^{n} u_i(t)\, A_i)q(t) \\ y(t) = \lambda q(t) \end{cases}$$

Le vecteur d'état q appartient à un <u>R</u>-espace vectoriel Q de dimension finie (q(0) est donné). Les applications $A_o, A_1, \ldots, A_n : Q \to Q$, $\lambda : Q \to \underline{R}$ sont R-linéaires. Les entrées $u_1, \ldots, u_n : [0, \infty[\to \underline{R}$ sont, pour simplifier, supposées continues.

Afin d'introduire de façon heuristique les variables non commutatives, écrivons la sortie y à l'aide du développement de Peano-Baker (cf. (8)) :

$$(6) \quad y(t) = \lambda q(0) + \sum_{\nu \geq o} \sum_{j_o, \ldots, j_\nu = o}^{n} \lambda A_{j_\nu} \ldots A_{j_o} q(0) \int_o^t d\xi_{j_\nu} \ldots d\xi_{j_o}$$

Définissons l'intégrale itérée $\int_o^t d\xi_{j_\nu} \ldots d\xi_{j_o}$ par récurrence sur la longueur :

$$\xi_o(\tau) = \tau, \quad \xi_i(\tau) = \int_o^\tau u_i(\sigma)d\sigma \; (i=1, \ldots, n)$$

$$\int_o^t d\xi_j = \xi_j(t) \quad (j=0,1,\ldots,n)$$

$$\int_o^t d\xi_{j_\nu} \ldots d\xi_{j_\nu} = \int_o^t d\xi_{j_\nu}(\tau) \int_o^\tau d\xi_{j_{\nu-1}} \ldots d\xi_{j_o}$$

(l'intégrale de gauche est prise au sens de Stieltjes). Introduisons l'alphabet $X = \{x_o, x_1, \ldots, x_n\}$ et faisons correspondre à (6) la série formelle en les variables non commutatives $x \in X$:

$$g = \lambda q(0) + \sum_{\nu \geq o} \sum_{j_o, \ldots, j_\nu = o}^{n} \lambda A_{j_\nu} \ldots A_{j_o} q(0)\, x_{j_\nu} \ldots x_{j_o}$$

g, appelée série <u>génératrice</u> de (5), caractérise entièrement son comportement entrée-sortie (cf. (3)). Sa rationalité équivaut à la finitude de la dimension de l'espace d'état Q. Elle joue quant à la réalisation un rôle analogue aux fonctions de transfert dans le cas linéaire.

b) Séries de Chen

Soit $\{\xi_o(\tau),\xi_1(\tau),\ldots,\xi_n(\tau)|0\leq\tau\leq t\}$ un chemin de $\underline{R}^{n+1}$, où les ξ_j sont supposées continues, à variations bornées. On lui attache la série non commutative

$$(7) \quad 1+\sum_{\nu\geq o}\ \sum_{j_o,\ldots,j_\nu=o}^{n} x_{j_\nu}\ldots x_{j_o}\int_o^t d\xi_{j_\nu}\ldots d\xi_{j_o}.$$

Chen (1,2) montre que (7) caractérise le chemin à une translation près. Nous appellerons (7) série de Chen du chemin.

Aux entrées $\{u_1(\tau),\ldots,u_n(\tau)|0\leq\tau\leq t\}$, associons le chemin

$$\{\xi_o(\tau)=\tau,\xi_1(\tau)=\int_o^\sigma u_1(\sigma)d\sigma,\ldots,\xi_2(\tau)=\int_o^\sigma u_n(\sigma)d\sigma|0\leq\tau\leq t\}$$

et sa série de Chen.

Il est clair, par la formule (6), que la valeur numérique $y(t)$ de la sortie s'obtient par dualité canonique entre la série génératrice g et la série de Chen. En un certain sens, on peut donc éliminer les fonctions entrées et les remplacer par des séries formelles.

Comme on le vérifie aisément, les coefficients de g croissent au plus exponentiellement en fonction de la longueur des mots. Considérons la sous-algèbre des séries formelles non commutatives telles que, pour tout entier $M>0$,

$$\sum_{\nu\geq o}\ \sum_{j_o,\ldots,j_\nu=o}^{n} |(s,x_{j_\nu}\ldots x_{j_o})|M^\nu<\infty$$

($(s,x_{j_\nu}\ldots x_{j_o})$ est le coefficient du mot $x_{j_\nu}\ldots x_{j_o}$). Cette famille de normes définit une topologie T_E qui est la topologie la plus faible rendant continues les fonctionnelles de séries génératrices rationnelles, ou, plus généralement, dont les coefficients croissent au plus exponentiellement. Soit $\bar{C}_E$ le complété des séries de Chen pour ces topologies.

Exemples - La série de Chen limite pour k infini du chemin $\{\xi_1^{(k)}(\tau)=\frac{\sin k^2\tau}{k},\xi_2^{(k)}(\tau)=\frac{\cos k^2\tau}{k}|0\leq\tau\leq t\}$ est $\exp\frac{t}{2}(x_1x_2-x_2x_1)$. Cela explique l'exemple du § III.

c) Intégrales stochastiques itérées.

Posons

$$\eta_o(\tau)=\tau,\eta_i(\tau)=b_i(\tau)=\int_o^\tau db_i \ (i=1),\ldots n).$$

L'intégrale itérée stochastique $\int_o^t d\eta_{j_\nu}\ldots d\eta_{j_o}$ est définie par récurrence sur la longueur comme précédemment, mais ici les intégrales sont prises au sens de Stratonovich.

On peut montrer que, pour tout entier $M>0$, la série

$$\sum_{\nu\geq o}\ \sum_{j_o,\ldots,j_\nu=o}^{n} |\int_o^t d\eta_{j_\nu}\ldots d\eta_{j_o}|M^\nu$$

est, p.s., convergente. La série de Chen

$$1+\sum_{\nu\geq o}\ \sum_{j_o,\ldots,j_\nu=o}^{n} x_{j_\nu}\ldots x_{j_o}\int_o^t d\eta_{j_\nu}\ldots d\eta_{j_o}$$

est donc, p.s., un élément de $\bar{C}_E$.

Il est alors licite d'appeler tout élément de $\bar{C}_E$ trajectoire du bruit blanc gaussien. Il est aisé de munir $\bar{C}_E$ d'une mesure de probabilité à savoir celle correspondant à la mesure de Wiener. Rappelons que jusqu'à présent les seules définitions rigoureuses du bruit blanc étaient données dans le cadre de la théorie des distributions de L. Schwartz.

Par l'intermédiaire de g, on voit donc que la sortie y de (5) dépend des trajectoires du bruit blanc continûment dans la topologie T_E. Le problème de l'approximation est résolu.

Remarques - (i) - On retrouve les crochets de Lie en rappelant que le logarithme d'une série de Chen (7) est toujours un élément de Lie.

(ii) - Dans les topologies fonctionnelles classiques, il n'est guère possible de lier les composantes d'une fonction $\underline{R}\to\underline{R}^N$ $(N\geq 2)$. C'est l'avantage des séries de Chen d'y parvenir et de résoudre ainsi le problème.

(iii) - La méthode n'est, malheureusement, pas généralisable à des équations plus générales. En effet, il n'y a pas, p.s., convergence de la série génératrice associée, quand on y remplace les mots par des intégrales itérées stochastiques. C'est ce type de divergence qui est aussi à la base de bien des difficultés mathématiques rencontrées en mécaniques statistique et quantique.

(iv) - Un récent article de Yamato (17) conduit à penser que le problème d'approximation peut encore être résolu, grâce aux intégrales itérées, pour les équations (3), où l'algèbre de Lie engendrée par les champs de vecteurs $A_o,A_1,\ldots,A_n$ est nilpotente.

BIBLIOGRAPHIE

1.- CHEN (K.-T.) - Intégration of paths - a faithful representation of paths by non-commutative formal power series, Trans. Amer. Math. Soc., 89, 1958, 395-407.

2.- CHEN (K.-T.) - Iterated path integrals, Bull. Amer. Math. Soc., 83, 1977, 831-879.

3.- 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 Ecomm. Math. Syst. 131, Springer-Verlag, Berlin, 1976.

4.- FLIESS (M.) - Intégrales itérées de K.-T. Chen, bruit blanc gaussien et filtrage non linéaire, C.R. Acad. Sc. Paris, A-284, 1977, 459-462.

5.- FLIESS (M.) - Mise en évidence d'effets non linéaires nouveaux grâce aux variables non commutatives, J. Physique, Coll. C. 5, 8, 1978, 21-22.

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, 321-324.

7.- FREEDMAN (M.I.) et WILLEMS (J.C.) - Smooth representation of systems with differentiated inputs, IEEE Trans. Autom. Control, 23, 1978, 16-21.

8.- GANTMACHER (F.R.) - Théorie des matrices (traduit du russe), t.2, Dunod, Paris, 1966.

9.- LOBRY (C.) - Sur l'identification de processus non linéaires par des méthodes stochastiques, C.R. Acad. Sc. Paris, A-284, 1977, 1473-1476.

10.- LOBRY (C.) - Article en préparation.

11.- McSHANE (E.J.) - Stochastic calculus and stochastic models, Academic Press, New York, 1974.

12.- McSHANE (E.J.) - Stochastic differential equations J. Multivar. Anal., 5, 1975, 121-177.

13.- SUSSMANN (H.J.) - On the gap between deterministic and stochastic ordinary differential equations, Ann. Proba., 6, 1978, 19-41.

14.- WONG (E.) - Stochastic processes in information and dynamical systems, McGraw-Hill, New York, 1971

15.- WONG (E.) et ZAKAI (M.) - On the relation between ordinary and stochastic differential equations, Int. J. Engin. Sci., 3, 1965, 213-229.

16.- WONG (E.) et ZAKAI (M.) - Riemann-Stieltjes approximations of stochastic integrals, Z. Wahrschein verw. Geb., 12, 1969, 87-97.

17.- YAMATO (Y.) - Stochastic differential equations and nilpotent Lie algebras, Z. Wahrschein verw. Geb. 47, 1979, 213-229.