YAO-ZHONG HU

Calculs formels sur les e.d.s. de Stratonovitch

Séminaire de probabilités (Strasbourg), tome 24 (1990), p. 453-460
<http://www.numdam.org/item?id=SPS_1990__24__453_0>

CALCULS FORMELS SUR LES E.D.S. DE STRATONOVITCH

par Yao-Zhong HU
Institute of Mathematical Physics
Academia Sinica, Wuchang

On trouve des articles difficiles dans les volumes précédents sur les équations différentielles stochastiques au sens de Stratonovitch, mais il y a aussi des choses simples qui manquent. Le premier problème que les mathématiciens se demandent est l'existence et l'unicité de la solution, mais pour les ingénieurs, ils s'intéressent surtout à trouver une forme de calcul explicite. Il y a de nombreux travaux sur la question. Ils reposent sur des résultats de la théorie des équations différentielles qui font partie du "folklore" pour les spécialistes. Ici nous essayons de les présenter avec une technique tout à fait simple. Suivant les conseils de Fliess [5], nous avons regardé les articles de Chen [2][3]. Un travail moderne et plus occupé par les probabilités est l'article de Ben Arous [1], et il y a aussi depuis peu de temps l'article de Strichartz [11] qui est très intéressant. On peut lire notre exposé facilement (par exemple en prenant le TGV de Strasbourg à Paris).

Je remercie P.A. Meyer d'avoir relu cette note et fait plusieurs suggestions pour la clarté.

1. Calculs sur l'exponentielle. Nous commençons par des calculs dans lesquels les grandes lettres $X, Y, Z \ldots$ peuvent avoir (au moins) trois significations différentes : des champs de vecteurs sur une variété (le cas intéressant, pour lequel l'analyse est difficile) des matrices réelles ou complexes (pour lesquelles il y a encore de l'analyse, mais facile) ou simplement des indéterminées non commutatives (pour lesquelles l'analyse n'est presque rien). Les séries formelles non commutatives dépendant d'un paramètre réel t peuvent bien sûr être dérivées ou intégrées par rapport au paramètre, coefficient par coefficient.

L'exponentielle $\xi_t = e^{tX}$ est définie par la série habituelle dans le cas des matrices ou des séries formelles. Dans le troisième cas, celui d'une variété V (supposons la compacte pour simplifier), on peut encore lui donner un sens comme opérateurs en dimension infinie. Le champ de vecteurs X est une *dérivation* de l'algèbre $C^\infty(V)$, et ξ_t est un *automorphisme* de l'algèbre $C^\infty(V)$. On considère la solution $\varphi(t,x)$ de l'équation différentielle

$$(0) \qquad \dot\varphi(t,x) = X(\varphi(t,x)) \quad ; \quad \varphi(0,x) = x$$

et $\xi_t(f)$ pour $f \in C^\infty(V)$ est la fonction $f \circ \varphi(t,\cdot)$. On a alors

$$d\xi_t(f) = \xi_t(Xf)dt \quad ; \quad \xi_0(f) = f \ .$$

L'automorphisme ξ_t est bien donné par la série exponentielle sous la forme

$$\xi_t(f) = \sum_n \frac{t^n}{n!} X^n f \ ,$$

pour t petit, sous des conditions d'analyticité (de la variété, du champ de vecteurs et de la fonction f), auxquelles nous ne nous intéressons pas dans la note. Même sans elles la formule est correcte pour calculer les dérivées n-ièmes.

Un problème important est celui d'avoir une formule de perturbations donnant $e^{t(X+Y)}$ à partir de e^X lorsque X et Y ne commutent pas. Il y a pour cela une formule générale :

$$e^{X+Y} = \sum_n \int_{1>s_1>\ldots>s_n>0} e^{(1-s_1)X} Y e^{(s_1-s_2)X} Y \ldots Y e^{s_n X} \, ds_1 \ldots ds_n \, . \tag{1}$$

La démonstration se fait ainsi : si nous posons $U_t = e^{tX}$, $V_t = e^{t(X+Y)}$, on a

$$\frac{dV_t}{dt} = (X+Y)V_t = XV_t + F(t) \, .$$

Si on fait semblant de connaître le dernier terme $F(t) = YV_t$, la solution de l'équation est alors

$$V_t = U_t + \int_0^t U_{t-s} F(s) \, ds \, .$$

Dans notre cas, on trouve

$$V_t = U_t + \int_0^t U_{t-s} Y V_s \, ds \, ,$$

qui donne alors (1) par itération. La formule (1) est utilisée pour calculer la perturbation d'un générateur X de semi-groupe fortement continu sur un espace de Banach, par un opérateur borné Y : c'est alors de l'analyse un peu sérieuse (Kato [6], p. 497).

En dérivant, on obtient (D désignant $\frac{d}{dt}|_{t=0}$)

$$D e^{X+tY} = \int_0^1 e^{(1-u)X} Y e^{uX} \, du = e^X \int_0^1 e^{-uX} Y e^{uX} \, du \, . \tag{2}$$

En particulier, on trouve une formule utile pour plus tard : si $Z(t)$ dépend différentiablement du paramètre t

$$\frac{d}{dt} e^{Z(t)} = e^{Z(t)} \int_0^1 e^{-uZ(t)} \dot{Z}(t) e^{uZ(t)} \, du \, . \tag{3}$$

Maintenant on voit arriver l'algèbre de Lie : on a la formule

$$e^{-uX} Y e^{uX} = \exp(-u \, \mathrm{ad}(X))Y \tag{4}$$

où $\mathrm{ad}(X)Y = XY - YX$, et nous définissons $\mathrm{ad}^p(X)$ par récurrence, et $\exp(\mathrm{ad}(X))$ est la série exponentielle. Dans le cas des matrices c'est presque évident, car pour chaque X on a deux groupes à un paramètre d'applications linéaires de l'espace des matrices Y dans lui même, et il suffit de vérifier qu'ils ont la même tangente pour $u = 0$. Pour les autres cas il faut peut être réfléchir plus. Dans la formule (2) à droite nous faisons le changement (4) dans l'intégrale et nous intégrons la série (dans le cas des matrices ce n'est pas trop difficile à justifier)

$$D e^{X+tY} = e^X \sum_p \frac{(-1)^p}{(p+1)!} \mathrm{ad}^p(X)Y = e^X f(\mathrm{ad}_X)Y \tag{5}$$

où $f(z)$ est la fonction entière $(1-e^{-z})/z$. C'est une formule très connue dans la théorie des groupes de Lie, elle donne l'application linéaire tangente à l'exponentielle avec les crochets de Lie itérés. On peut d'ailleurs faire la même chose avec l'exponentielle toute entière en sortant juste e^X à gauche des intégrales dans (1).

Maintenant nous faisons encore une remarque : on peut faire l'inversion de la formule

$$X = \int_0^1 e^{-uZ} Y e^{uZ} \, du \, .$$

En effet, cette relation est de la forme

$$X = Y + \sum_{m+n>0} a_{mn} Z^m Y Z^n \, , \tag{6}$$

avec des coefficients $a_{mn} = (-1)^m / m!n!(m+n+1)$. Ecrivons cela sous la forme

$$Y = X - \sum \ldots$$

et itérons :

$$Y = X - \sum a_{mn} Z^m X Z^n + \sum a_{mn} a_{pq} Z^{m+p} X Z^{n+q} - \ldots$$

Comme il n'y a qu'un nombre fini de termes d'un type $Z^j X Z^k$ donné, cela définit une série formelle non commutative (dans le cas des matrices on voit facilement que la série converge pour $\|Z\|$ assez petit). Une meilleure façon de faire l'inversion est de partir de (5) et de poser $g(z) = z/(1-e^{-z})$, analytique au voisinage de 0 (ses coefficients g_p sont à peu près les nombres de Bernoulli). On a alors

$$Y = g(\mathrm{ad}_Z) X \, . \tag{7}$$

2. Equations dépendant du temps (et de Stratonovitch). Maintenant nous regardons une équation différentielle comme (0), mais le champ de vecteurs dépend du temps

$$dx(t) = X(t, x(t))\, dt \quad ; \quad x(0) = x \tag{8}$$

(nous écrirons quelquefois $X_t(x)$, quelquefois $X(t)$). Un cas intéressant est celui des champs de vecteurs constants par morceaux (donc non continus). Par exemple Y pour $0 \le t \le 1$ et X pour $1 < t \le 2$, et alors on calcule e^{X+Y}. En probabilités, on s'intéresse à des équations à "plusieurs temps"

$$dx(t) = \sum_{i=0,\ldots,d} X_i(t, x(t))\, dt^i \tag{9}$$

parce que le premier temps sera $dt^0 = dt$, et les autres seront par exemple $dt^i = g^i(t)\,dt$ (équations avec contrôle) ou bien $dt^i = dB^i(t)$ (équations différentielles stochastiques de Stratonovitch pour des mouvements browniens, ou même pour des semimartingales continues). Les calculs formels sont les mêmes pour les mouvements browniens que pour les équations avec contrôle, en remplaçant $g_i(t)$ par le "bruit blanc" $\dot{B}^i(t)$, c'est le "principe de transfert" des équations de Stratonovitch, bien connu des ingénieurs (McShane [9]).

Donc il suffit d'obtenir les formules explicites dans le cas des termes de contrôle, et alors on est ramené à l'équation (8) plus simple en posant

$$(10) \qquad X_t(x) = X_0(t,x) + \sum_{i>0} X_i(t,x) g^i(t) \,.$$

Donc il suffit d'étudier (8).

Comme pour (0) on peut regarder l'équation (8) comme la définition d'un automorphisme $\xi_t(f) = f \circ \varphi(t)$ de $C^\infty(V)$

$$(11) \qquad d\xi_t(f) = \xi_t(X_t f)\, dt \quad ; \quad \xi_0 = I \,.$$

Dans le cas des matrices ou des séries formelles, il faut faire attention car il y a deux équations différentielles qui s'écrivent $\dot{\xi}_t = \xi_t X_t$ ou bien $X_t \xi_t$: c'est la première qu'on étudie. La série exponentielle est remplacée par des intégrales multiples

$$(12) \qquad \xi_t = \sum_n \int_{0<s_1<\ldots<s_n<t} X_{s_1} \ldots X_{s_n}\, ds_1 \ldots ds_n \,.$$

Bien sûr pour traiter le cas des équations différentielles stochastiques il faut remplacer X_t par sa valeur (10) et développer toutes les intégrales, ce qui donne un grand nombre d'intégrales multiples de Stratonovich. Ceci est étudié chez Ben Arous [1]. Maintenant, ce qu'on essaye de faire, c'est d'écrire pour chaque t

$$(13) \qquad \xi_t = e^{Z(t)} \,,$$

où $Z(t)$ est un champ de vecteurs (dans le cas des é. d. s. un champ de vecteurs aléatoire). Le résultat important, c'est qu'on peut en principe calculer $Z(t)$ par une formule universelle d'algèbre de Lie sur les champs $X(t)$. Cela s'appelle la *formule de Campbell-Hausdorff*, ou pour ne pas faire de la peine, Baker-Campbell-Hausdorff-Dynkin, mais on a quand même fait de la peine à Chen qui a généralisé la formule BCHD des groupes de Lie, et peut être à Strichartz qui a donné une nouvelle formule. On trouve la formule de BCHD classique quand le champ $X(t)$ est constant par morceaux. Plutôt, on va avoir deux formules universelles, une avec les développements de Volterra et une avec les crochets de Lie. Nous les prenons dans Strichartz [11] qui est le travail le plus récent.

Peut être on peut remarquer la relation entre ces calculs sur les équations différentielles et les intégrales itérées de Chen [2] [3] : une seule courbe $u(t)$ dans $\mathbb{R}^n$ détermine un champ de vecteurs dépendant du temps $X(t) = \sum_i u^i(t) D_i$ et on a $\xi_t(f) = f(\cdot + (u(t) - u(0))$. Si on résout l'équation différentielle par la méthode ci-dessous (en oubliant que les D_i commutent : il ne faut pas regrouper les termes) on voit apparaître la série de Chen du chemin $u(t)$.

3. Résolution formelle de l'équation différentielle. Pour essayer de calculer $Z(t)$, la première idée vient d'écrire

$$Z_t = \log \Big(I + \sum_{n>0} \int_{0<s_1<\ldots<s_n<t} X_{s_1} \ldots X_{s_n}\, ds_1 \ldots ds_n \Big)$$

c'est à dire

$$Z(t)=\sum_m \frac{(-1)^{m+1}}{m}\Big(\sum_{n>0}\int_{0<s_1<\ldots<s_n<t} X_{s_1}\ldots X_{s_n}\,ds_1\ldots ds_n\Big)^m .$$

Cela a l'air très compliqué, mais on peut quand même s'en servir. Appelons $J(n)$ l'intégrale multiple d'ordre n et considérons les $X(s)$ comme des indéterminées (de degré 1). Alors le terme $Z_r(t)$ de degré r dans $Z(t)$ est une somme sur toutes les décompositions en entiers $p_1+\ldots+p_m=r$ (m est maintenant le nombre d'entiers) et on a

$$Z_r(t)=\sum_{(p_j)} \frac{(-1)^{m+1}}{m} J(p_1)\ldots J(p_r)\,. \tag{14}$$

Ce produit d'intégrales est aussi l'intégrale de $X(s_1)\ldots X(s_r)$ sur le domaine $0<s_1<\ldots<s_{q_1}<t\,;\ 0<s_{q_1+1}<\ldots<s_{q_2}<t\,;\ldots$ en posant $q_j=p_1+\ldots+p_j$.

Nous allons maintenant arranger la formule (14) en recopiant Strichartz [11]. On ne voit peut être pas clairement en lisant Strichartz que ce calcul est fait *avant* de passer aux crochets de Lie. Nous considérons d'abord chaque facteur $X(s_1)\ldots X(s_r)$, et cherchons son coefficient dans la somme (14). On appelle e le nombre d'erreurs d'ordre (inversions) dans la suite s_i, c'est à dire le nombre d'entiers i tels que $s_i>s_{i+1}$. On compte d'abord le nombre $\nu(m)$ de fois qu'il apparaît dans la somme avec le même entier m. Il s'agit de placer les $m-1$ "barrières" $0<q_1<\ldots<q_{m-1}<r$ de manière que les s_i croissent entre 0 et q_1, entre q_1+1 et q_2, ... q_{m-1} et r. On doit obligatoirement placer une barrière aux endroits i où il y a une erreur. Cela fixe e barrières, et il en reste à placer $m-1-e$ (bien sûr on doit avoir $m\geq e+1$) ce qui laisse $\binom{r-e-1}{m-1-e}$ possibilités. Maintenant le coefficient total est (en posant $m=e+1+j$, $0\leq j\leq r-e-1$)

$$\sum_{m=e+1}^{r-1}\frac{(-1)^m\nu(m)}{m+1}=(-1)^e\sum_{j=0}^{r-e-1}\frac{(-1)^j}{e+1+j}\binom{r-e-1}{j}$$

Le coefficient binômial est le coefficient de x^j dans $(1-x)^{r-e-1}$ et on fait apparaître le dénominateur en multipliant par x^e et en intégrant de 0 à 1. Il reste donc

$$(-1)^e\int_0^1(1-x)^{r-e-1}x^e\,dx=(-1)^e\frac{(r-e-1)!\,e!}{r!}\,.$$

Maintenant on revient à (14) et on veut intégrer juste sur le simplexe croissant $0<t_1<\ldots<t_r<t$ en faisant opérer les permutations σ de $1,\ldots,r$. On désigne par $e=e(\sigma)$ le nombre d'erreurs dans l'ordre des entiers $\sigma(i)$. Alors $Z_r(t)$ est égal à

$$\sum_{\sigma\in S_r}(-1)^e\frac{e!(r-1-e)!}{r!}\int_{0<t_1<\ldots<t_r<t} X(t_{\sigma(1)})\ldots X(t_{\sigma(r)})\,dt_1\ldots dt_r\,. \tag{15}$$

4. La formule de B-C-H-D .Maintenant on revient au calcul avant la formule (14) et on utilise à partir de (12) une méthode moins brutale. On dérive (12) par rapport à t et on compare les formules (11) et (3). On obtient

$$e^{Z(t)} \int_0^1 e^{-uZ(t)} \dot{Z}(t) e^{uZ(t)} \, du = e^{Z(t)} X(t) .$$

Mais on sait inverser cette équation par la formule (6), qui est de la forme

$$\dot{Z}(t) = X_t + \sum_{m+n>0} b_{mn} Z_t^m X_t Z_t^n . \tag{16}$$

Pour faire apparaître l'algèbre de Lie on utilise plutôt la formule (7) : on a $Z_0 = 0$ et

$$\dot{Z}_t = g(\mathrm{ad}(Z_t)) X_t = \sum_p \mathrm{ad}^p(Z_t) X_t , \tag{17}$$

donc la série (16) est en fait d'une forme très spéciale. Si on résout l'équation (17) par approximations successives, on a d'abord

$$Z_t^1 = \int_0^t X(s) ds$$

ensuite

$$Z_t^2 = \sum_p g_p \int_0^t \mathrm{ad}^p(X_s) X_t \, ds$$

et ainsi de suite. Il est clair par récurrence que $Z(t)$ est un élément de Lie, c'est à dire que les coefficients $Z_r(t)$ du développement de $Z(t)$ en série formelle s'expriment complètement au moyen des crochets de Lie d'ordre r formés avec les champs $X(s_i)$. Bien sûr ici ce n'est pas tout à fait de l'algèbre car on a une infinité d'indéterminées et des intégrales, mais on peut prendre des champs constants par morceaux et passer à la limite après.

On arrive au point amusant d'algèbre : quand on sait d'avance qu'un polynôme homogène de degré r est un élément de Lie, on peut arranger automatiquement ce polynôme pour l'écrire avec les crochets de Lie. Voilà comment on fait (cette idée a été inventée par Dynkin avant de faire des probabilités). Un polynôme homogène P de degré r en des indéterminées X_i est une combinaison linéaire unique de monômes $X_{i_1} \dots X_{i_r}$ (on n'écrit pas de puissances : on répète les indéterminées s'il le faut). On définit le *monôme de Lie* $[X_{i_1} \dots X_{i_r}]$ par récurrence

$$[X_{i_1} \dots X_{i_r}] = [\, [X_{i_1} \dots X_{i_{r-1}}], X_{i_r} \,] ,$$

et on définit un polynôme $[P]$ en remplaçant dans P les monômes par les monômes de Lie. *Alors si P est déjà un élément de Lie, on a* $[P] = rP$. Ceci est un peu d'algèbre mais pas difficile, voir Postnikov [10] p. 104. On peut alors transformer directement la formule (15) en une formule d'algèbre de Lie. Voilà comment Strichartz écrit sa formule (on a juste mis des crochets dans (15) et divisé par r)

$$Z(t) = \sum_r \sum_{\sigma \in S_r} \frac{(-1)^{e(\sigma)}}{r^2 \binom{r-1}{e(\sigma)}} \int_{0<t_1<\ldots<t_r<t} [\, X(t_{\sigma(1)}) \ldots X(t_{\sigma(r)}) \,] \, dt_1 \ldots dt_r \tag{18}$$

Strichartz explique bien au début de son article que cette formule a moins de termes inutiles que la formule BCHD qu'on trouve dans les livres. Tout de même, quand on se rend compte qu'on n'a pas du tout utilisé la symétrie des crochets de Lie on se dit qu'on doit pouvoir la réduire encore.

On ne s'est pas du tout occupé de la convergence de la série. Le problème est bien sûr tout à fait différent dans le cas déterministe et probabiliste. Pour ce dernier cas il faut surtout consulter Ben Arous [1].

REFERENCES

[1] BEN AROUS (G.). Flots et séries de Taylor stochastiques. *Prob. Th. Rel. Fields,* 81, 1989, p.29–77.

[2] CHEN (K.T.). Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. Math.*, 65, 1957, p.163–178.

[3] CHEN (K.T.). Expansion of solutions of differential systems. *Arch. Rat. Mech. Anal.*, 13, 1963, p.348–363.

[4] DAVIES (E.B.). *One Parameter Semi-groups,* Academic Press, 1980.

[5] FLIESS (M.) et NORMAN-CYROT (D.). Algèbres de Lie nilpotentes, formule de Campbell-Baker-Hausdorff et intégrales itérées de Chen. *Sém. Prob. XVI,* LN 920, p.257-267, Springer 1982.

[6] KATO (T.). *Perturbation theory for linear operators,* 2nd edition, Springer 1976.

[7] KUNITA (H.). On the representation of solutions of stochastic differential equations. *Sém. Prob. XIV,* LN 784, p.282–304, Springer 1980.

[8] MARCUS (S.I.). Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics,* 4, 1981, p.223–245.

[9] McSHANE (E.J.). Stochastic differential equations *J. Multiv. Anal.*, 6, 1975, p.121–177.

[10] POSTNIKOV (M.M.). *Leçons de géométrie : Groupes et algèbres de Lie.* Editions MIR, Moscou 1982.

[11] STRICHARTZ (R.S.). The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *J. Funct. Anal.*, 72, 1987, p.320–345.

Appendice. Pour éviter aux lecteurs de consulter un livre, voici le schéma de la démonstration du lemme de Dynkin utilisé à la fin de l'exposé. Chaque étape exige un calcul très simple, que nous ne détaillons pas. La lettre X désigne n'importe laquelle des indéterminées X^i.

Pour la clarté, nous réservons le crochet aux commutateurs; nous notons $X^{[\alpha]}$ le monôme de Lie $[X^\alpha]$, et si $P \in \mathcal{P}$ (l'algèbre des polynômes non commutatifs) nous écrirons comme Postnikov $\sigma(P)$ au lieu de $[P]$.

1) On commence par démontrer que le crochet $[X^{[\alpha]}, X^{[\beta]}]$ de deux monômes de Lie est une combinaison linéaire de monômes de Lie. C'est trivial pour tout α si $X^{[\beta]} = X$, et on raisonne par récurrence sur la longueur de β : il n'y a là dedans que l'identité de Jacobi. Par conséquent les combinaisons linéaires de monômes de Lie forment une sous-algèbre de Lie $\mathcal{L}$ de $\mathcal{P}$.

2) On définit une application bilinéaire $M, P \longmapsto \theta_M(P)$ de $\mathcal{P} \times \mathcal{P}$ dans $\mathcal{P}$ par les règles suivantes

$$\theta_1(P) = P \ , \ \theta_X(P) = [P, X] \ , \ \theta_{RS} = \theta_S \theta_R \ .$$

On a alors trois lemmes faciles

— *Si* $M \in \mathcal{L}, P \in \mathcal{P}$ *on a* $\theta_M(P) = [P, M]$.

— *Si* $M \in \mathcal{P}, P \in \mathcal{P}$ *on a* $\theta_M(\sigma(P)) = \sigma(PM)$.

— *Si* $M \in \mathcal{L}, P \in \mathcal{L}$ *on a* $\sigma([M,P]) = [\sigma(P),M] + [P,\sigma(M)]$.

Le premier : on peut supposer que M est un monôme de Lie. On raisonne par récurrence en montrant que si la propriété est vraie pour M (quel que soit P) elle est aussi vraie pour $[M,X]$. Le second : on se ramène à vérifier que si la propriété est vraie pour M (quel que soit P) elle l'est pour MX , et c'est facile. Pour le troisième, il suffit de remplacer $[M,P]$ par $MP - PM$ et d'appliquer le lemme précédent.

Alors le théorème de Dynkin :

— *Si* $P \in \mathcal{L}$ *est homogène de degré* n *on a* $\sigma(P_n) = nP_n$,

se démontre par récurrence sur n : il suffit de démontrer que s'il est vrai pour P il l'est aussi pour $[P,X]$. Cela découle très facilement du dernier lemme. Ce théorème a été démontré indépendamment de Dynkin par W. Specht et par F. Wever, dans les mêmes années 1948–49.

Référence supplémentaire : P. Cartier, Démonstration algébrique de la formule de Hausdorff, *Bull. SMF,* 84, 1956, p. 241–249.