HIROSHI KUNITA

On the representation of solutions of stochastic differential equations

Séminaire de probabilités (Strasbourg), tome 14 (1980), p. 282-304
<http://www.numdam.org/item?id=SPS_1980__14__282_0>

On the representation of solutions of

stochastic differential equations

Hiroshi Kunita

0. Introduction.

Let us consider the stochastic differential equation

$$(0.1) \qquad d\xi_t = X_0(\xi_t)dt + \sum_{j=1}^{r} X_j(\xi_t)\circ dB_t^j$$

defined on a connected C^∞-manifold M of dimension d. Here X_0, $X_1, \ldots, X_r$ are C^∞-vector fields on M and $B_t = (B_t^1, \ldots, B_t^r)$ is a standard Brownian motion. The symbol $\circ$ denotes the Stratonovich-Fisk integral. Recently a number of authors has expressed the solution directly as a functional of B_t, under some conditions on vector fields $X_0 \ldots, X_r$. In Doss [1] Sussman [7], the solution is expressed in such a way that it is a continuous functional of B_t, if $r=1$ or $X_1, \ldots, X_r$ are commutative. However this is not the case in general if $r \geq 2$ and $X_1, \ldots, X_r$ are not commutative. In fact, Yamato [8] has proved that the solution is a functional of multiple Wiener integrals of B_t, provided that the Lie algebra generated by $X_1, \ldots, X_r$ is nilpotent.

In this paper, we shall consider the similar problem in case that the Lie algebra mentioned above is nilpotent or solvable. In section 2, we will discuss Yamato's result from a different point of view: Applealing Campbell-Hausdorff formula in Lie algebra, we will obtain an explicit expression as a functional of multiple Wiener integrals.

Section 1 is devoted to Campbell-Hausdorff formula. In Section 3, we will discuss the case that the Lie algebra is solvable. We will decompose the equation (0.1) into a chain of equations such that the corresponding Lie algebra of each equation is nilpotent, and then show that the solution of (0.1) is expressed as a composition of solutions of these nilpoitent equations.

1. Campbell-Hausdorff formula.

Given a complete C^∞-vector field X on the manifold M represented as $\sum_{i=1}^{d} X_i(x)\frac{\partial}{\partial x_i}$ with a local coordinate $(x_1, \ldots, x_d)$, we denote by e^{tX} the one parameter group of transformations on M generated by X: This means that $\phi_t(x) \equiv e^{tX}(x)$ satisfies, (i) for each $t \in (-\infty, \infty)$, ϕ_t is a diffeomorphism of M, (ii) $\phi_t \circ \phi_s = \phi_{t+s}$ for any $t, s \in (-\infty, \infty)$, $\lim_{t \downarrow 0} \phi_t(x) = x$ and (iii) it is the solution of the ordinary differential equation $\frac{d\phi_t(x)}{dt} = X(\phi_t(x))$ starting at x, where $X(x) = (X_1(x), \ldots, X_d(x))$. When $t = 1$, we write it as e^X.

Let X and Y be complete C^∞-vector fields. We define the Lie bracket $[X, Y]$ by $XY - YX$. It is often written as $X(\mathrm{ad}Y)$ Campbell-Hausdorff formula is a formula like

$$e^X e^Y = e^{X+Y-\frac{1}{2}[X, Y]+ \ldots}$$

We shall extend the formula to that of n vector fields.

Suppose we are given n C^∞-vector fields $Y_1, \ldots, Y_n$ such that $[\ldots[Y_{i_1}, Y_{i_2}] \ldots,]Y_{i_m}]$, $m = 1, 2, \ldots$ and their linear sums are all complete vector fields. Consider a formal power series

$$(1.1) \qquad Z = \Sigma_{m>0}(-1)^{m-1}m^{-1}\Sigma_{p>0} \frac{1}{p_1^{(1)}!\ldots p_n^{(1)}!p_1^{(2)}!\ldots p_n^{(m)}!} \frac{1}{|p|}$$

$$\times\, Y_1(\mathrm{ad}Y_1)^{p_1^{(1)}-1}(\mathrm{ad}Y_2)^{p_2^{(1)}}\cdots(\mathrm{ad}Y_n)^{p_n^{(1)}}(\mathrm{ad}Y_1)^{p_1^{(2)}}\cdots(\mathrm{ad}Y_n)^{p_n^{(m)}}$$

where $p_1^{(1)}, \ldots, p_n^{(m)}$ are nonnegative integers, $|p| = \Sigma_{1\le i\le n,\ 1\le j\le m} p_i^{(j)}$ and $\Sigma_{p>0}$ means the sum of terms such that $\sum_{i=1}^{n} p_i^{(j)} > 0$ for all $1 \le j \le m$. If $p_1^{(1)} = 0$, we understand the first member as $Y_2(\mathrm{ad}Y_2)^{p_2^{(1)}-1}$ instead of $Y_1(\mathrm{ad}Y_1)^{p_1^{(1)}-1}$. Now the term corresponding $[\ldots[Y_{i_1}, Y_{i_2}]\ldots, Y_{i_m}]$ appears several times in the power series. Summing up all the corresponding term, we denote the coefficient of the above vector field as $c_{i_1\ldots i_m}$. Then the power series is written as

$$(1.2)\qquad Z = \sum_{m=1}^{\infty}\sum_{(i_1,\ldots,i_m)} c_{i_1\ldots i_m} Y^{i_1\ldots i_m},$$

where

$$(1.3)\qquad Y^{i_1\ldots i_m} = [\ldots[Y_{i_1}, Y_{i_2}]\ \ldots]Y_{i_m}]$$

Theorem 1.1. (Campbell-Hausdorff formula) Suppose that (1.2) is absolutely convergent and define a complete vector field. Then it holds

$$(1.4)\qquad e^{Y_n}\cdots e^{Y_1} = e^{Z}.$$

The proof may be found in Jacobson [5] in case $n = 2$. It can be applied to the present case with a simple modification.

We shall compute coefficients $c_{i_1\ldots i_m}$. Let us divide the multi-index $I = (i_1, \ldots, i_m)$ to a sequence of shorter ones I_j, $j = 1, \ldots, \ell$ and write it as $\hat{I}$;

(1.5) $$\hat{I} = (I_1, \ldots, I_{k_1})(I_{k_1+1}, \ldots, I_{k_2}) \ldots (I_{k_{\ell-1}+1}, \ldots, I_{k_\ell}),$$

where each index I_k consists of same number $\hat{i}_k$ and these numbers $\hat{i}_k$, $k = 1, \ldots, k_\ell$ satisfies

(1.6) $$\hat{i}_1 > \hat{i}_2 > \ldots > \hat{i}_{k_1} < \hat{i}_{k_1+1} > \ldots > \hat{i}_{k_2} \ldots < \hat{i}_{k_{\ell-1}+1} > \ldots > \hat{i}_{k_\ell}$$

The division $\hat{I}$ is defined uniquely from I. We call this a natural division. We denote the length of I_k (the number of elements in I_k) as n_k. Then $\sum_{k=1}^{k_\ell} n_k = m$. Divide again each index I_k into j_k indices, each of which consists of $n_k^{(i)}$ elements ($i = 1, \ldots, j_k$). Hence it holds $n_k^{(1)} + \ldots + n_k^{(j_k)} = n_k$. Then we have

(1.7) $$c_{i_1 \ldots i_m} = \frac{1}{m} \sum_{s=0}^{\ell-1} \Sigma_* \binom{\ell-1}{s} (-1)^{j_1 + \ldots + j_{k_\ell} - s - 1} (j_1 + \ldots + j_{k_\ell} - s)^{-1}$$
$$\times \frac{1}{n_1^{(1)}! \ldots n_1^{(j_1)}! \ldots n_{k_\ell}^{(1)}! \ldots n_{k_\ell}^{(j_\ell)}!}$$

Here, the sum Σ_* is taken for all subdivisions of I_k, $k = 1, \ldots, k_\ell$, i.e., for all positive integers $n_k^{(i)}$, $i = 1, \ldots, j_k$, $k = 1, \ldots, k_\ell$ such that $\sum_i n_k^{(i)} = n_k$.

Let I' be another multi-index of length m and let

$$\hat{I}' = (I'_1, \ldots, I'_{k'_1}) \ldots (I'_{k'_{\ell-1}+1}, \ldots, I'_{k'_\ell})$$

be its natural division. We say that I and I' are equivalent if for each k I_k and I'_k contain the same number of elements and $k'_1 = k_1, \ldots, k'_\ell = k_\ell$ hold. Note that $c_I = c_{I'}$ holds if I and I' are equivalent.

If each I_k in (1.5) contains a single element, I is divided as

$$\hat{I} = (i_1, \ldots, i_{k_1})(i_{k_1+1}, \ldots, i_{k_2}) \ldots (i_{k_{\ell-1}+1}, \ldots, i_{k_\ell})$$

where $k_\ell = m$. We will call such I as single. In this case, (1.7) becomes

$$(1.8) \qquad c_{i_1 \ldots i_m} = \frac{1}{m} \sum_{s=0}^{\ell-1} \binom{\ell-1}{s} (-1)^{m-s-1} (m-s)^{-1}$$

We shall calculate a few of coefficients

(a) $c_i = 1$

(b) $c_{ij} = -\frac{1}{4}$ if $i > j$, $c_{ij} = \frac{1}{4}$ if $i < j$

(c)
$$c_{ijk} = \begin{cases} \frac{1}{9} & \text{if } i < j < k \text{ or } k < j < i \\ -\frac{1}{18} & \text{if } j < i \ \& \ j < k \text{ or } j > i \ \& \ j > k \\ \frac{1}{36} & \text{if } i \neq j = k \end{cases}$$

2. Representation of solutions (I). Nilpotent case.

Consider the stochastic differential equation on M.

$$(2.1) \qquad d\xi_t = X_0(\xi_t)dt + \sum_{j=1}^{r} X_j(\xi_t) \circ dB_t^j,$$

where $X_0, X_1, \ldots X_r$ are complete C^∞-vector fields. If $X_0, X_1, \ldots X_r$ are commuting, i.e., $[X_i, X_j] = 0$ for each i and j, then the solution of the above equation starting at x is represented as

$$(2.2) \qquad \xi_t(x) = \exp(tX_0 + B_t^1 X_1 + \dots + B_t^r X_r)(x)$$

Here we understand that $tX_0(\omega) + B_t^1(\omega)X_1 + \dots + B_t^r(\omega)X_r$ is a vector field for each t and a. s. ω. This means that $\xi_t(x, \omega)$ equals $\phi_1(x, \omega)$ a.s., where $\phi_s(x, \omega)$ is the solution of the ordinary differential equation

$$\frac{d\phi_s}{dt} = (tX_0(\omega) + \dots + B_t^r(\omega)X_r)(\phi_s),$$

regarding t and ω as parameters. The fact can be proved directly, applying Ito's formula [4] to (2.2). However, if $X_0, \dots, X_r$ are not commuting, the formula (2.2) is not valid. We have to add several terms to the right hand of (2.2). This will be done in Theorem 2.3.

Our basic assumption in this section is that the Lie algebra $L = L(X_0, X_1, \dots, X_r)$ generated by $X_0, X_1, \dots, X_r$ is nilpotent of step p, i.e.,

$$[\dots[X_{i_1}, X_{i_2}] \dots]X_{i_m}] = 0$$

holds whenever $i_1, \dots, i_m \in \{0, 1, \dots, r\}$ and $m > p$. The algebra L is then a finite dimensional vector space, obviously. Then any element of L is a complete (or proper) vector field (See Palais [6], p.95). Under the same condition, Yamato [8] showed that the solution ξ_t of equation (2.1) is a functional of multiple Wiener integrals of B_t of degrees less than or equal to p. We will obtain the functional in a more explicit manner, making use of Campbell-Hausdorff formula.

We begin with notations on multi-index. We shall divide a multi-index $I = (i_1, \ldots, i_m)$ to shorter ones; $I = I_1 \ldots I_q$ $(q \leq m)$, where each I_k consists of the same element $\hat{i}_j$. Given positive integers $k_1 < k_2 < \ldots < k_\ell = q$, we define a divided index of I as

$$\Delta I = (I_1, \ldots, I_{k_1})(I_{k_1+1}, \ldots, I_{k_2}) \ldots (I_{k_{\ell-1}+1}, \ldots, I_{k_\ell}) \tag{2.3}$$

(This time we do not assume relation (1.6)). If each I_k contains a single element (or at most two), we say that ΔI is single (or double). The equivalence of two indices ΔI and $\Delta I'$ is defined similarly as in Section 1. Suppose now we are given an index I and a divided one ΔI. ΔI is not equal to the natural division of I. But if there is an index I' such that its natural division $\hat{I}'$ is equivalent to ΔI, then we set $c_{\Delta I} = c_{I'}$ for convention.

Let $B_t = (B_t^1, \ldots, B_t^r)$ be a standard r-dimensional Brownian motion. We set $B_t^0 = t$ for convention. Given a single divided index ΔI, we define the multiple Wiener integral $B_t^{\Delta I}$ as

$$B_t^{\Delta I} = \int \cdots \int_A dB_{t_1}^{i_1} \ldots dB_{t_m}^{i_m} \tag{2.4}$$

where

$$A = \{t_{k_1} < \ldots < t_1 < t, \ \ldots, \ t_{k_\ell} < \ldots < t_{k_{\ell-1}+1} < t, \ t_{k_i} < t_{k_i+1}, \quad i = 1, \ldots, \ell\} \tag{2.5}$$

If ΔI is a double index, we define

$$(2.6)\qquad B_t^{\Delta I} = \int \cdots \int_A dB_{t_1}^{I_1} \cdots dB_{t_\ell}^{I_{k_\ell}} ,$$

where

$$(2.7)\qquad B_t^{I_k} = \begin{cases} B_t^{i_k} & \text{if } I_k = \{i_k\} \quad \text{(single)} \\ t & \text{if } I_k = \{i_k, i_k\} \quad \text{(double)} \end{cases}$$

Lemma 2.1. Let $\xi_t(x)$ be the solution of (2.1) with $\xi_0 = x$. Then it is represented as

$$(2.8)\qquad \xi_t(x) = (\exp W_t)(x),$$

where

$$(2.9)\qquad W_t = tX_0 + \cdots + B_t^r X_r + \sum_{J:1<|J|\leq p} \{\Sigma^*_{\Delta J}\, c_{\Delta J} B_t^{\Delta J}\} X^J$$

$$X^J = [\ldots[X_{j_1}, X_{j_2}] \ldots]X_{j_m}] \qquad (J = (j_1, \ldots, j_m))$$

Here $\Sigma^*_{\Delta J}$ is the sum for all single and double divided indices of J.

Proof. For a fixed positive integer n and positive time t, set $\delta B_k^j = B^j_{\frac{k}{n}t} - B^j_{\frac{k-1}{n}t}$ and define

$$Y_1 = \frac{t}{n}X_0 + \delta B_1^1 X_1 + \cdots + \delta B_1^r X_r$$

$$\vdots$$

$$Y_n = \frac{t}{n}X_0 + \delta B_n^1 X_1 + \cdots + \delta B_n^r X_r$$

Set

$$(2.10) \qquad \xi_t^{(n)}(x) = (\exp Y_n \cdots \exp Y_1)(x)$$

Then, clearly it is the value at 1 of the solution of equation

$$\frac{d\xi'_s}{ds} = (X_0 + \frac{n}{t}\sum_{j=1}^{r} \delta B_k^j X^j)(\xi'_s) \quad \text{if} \quad (\frac{k-1}{n})t \le s \le \frac{k}{n}t$$

There is a subsequence of $\xi_t^{(n)}(x)$ converging to $\xi_t(x)$ a.s.

We shall next apply Campbell-Hausdorff formula to the right hand side of (2.10). It holds

$$(2.11) \qquad \sum_{(i_1,\ldots,i_m)} c_{i_1\ldots i_m} Y^{i_1\ldots i_m}$$

$$= \sum_{j_1,\ldots,j_m} \{ \sum_{(i_1,\ldots,i_m)} c_{i_1\ldots i_m} \delta B_{i_1}^{j_1} \cdots \delta B_{i_m}^{j_m} \} X^{j_1\ldots j_m}$$

$$= \sum_{J:|J|=m} \{ \sum_{\Delta J} c_{\Delta J} \sum_{I:\hat{I}\sim\Delta J} \delta B_{i_1}^{j_1} \cdots \delta B_{i_m}^{j_m} \} X^J$$

Here $\Sigma_{I:\hat{I}\sim\Delta J}\delta B_{i_1}^{j_1} \cdots \delta B_{i_m}^{j_m}$ means the sum for all indices I such that $\hat{I}$ is equivalent to ΔJ. The sum converges to $B_t^{\Delta J}$ if ΔJ is a single or double index. If ΔJ is more than double (i.e., ΔJ contains a subindex I_k with more than two elements), then the sum converges to 0. Therefore, (2.11) converges a.s. to

$$\sum_J \{ \sum_{\Delta J}{}^{*} c_{\Delta J} B_t^{\Delta J} \} X^J .$$

This proves that the sum of (2.11) for m = 1, 2, ... converges to W_t a.s.

Then the exponential map converges a.s. to e^{W_t}. The proof is complete.

We shall next calculate multiple Wiener integrals in (2.9) in cases that $|J|$ are 2 and 3. We introduce notations.

$$[B^i, B^j]_t = \int_0^t B_s^i dB_s^j - \int_0^t B_s^j dB_s^i$$

This indicates the stochastic area enclosed by the Brownian curve (B_s^i, B_s^j), $0 \leq s \leq t$ and its chord. Similarly, we set

$$[[B^i, B^j], B^k]_t = \int_0^t [B^i, B^j]_s dB_s^k - \int_0^t B_s^k d[B^i, B^j]_s$$

Lemma 2.2 (i) Coefficient of X^{ij} in (2.9) equals $\frac{1}{2}[B^i, B^j]_t$ if $i \neq j$. (ii) Coefficient of X^{ijk} equals $\frac{1}{18}[[B^i, B^j], B^k]_t$ if i, j, k are different or $0 = j = k \neq i$. If $0 < j = k \neq i$, it equals $\frac{1}{36} tB_t^i$ plus the above quantity.

Proof. The coefficient of X^{ij} equals

$$c_{(i)(j)} B_t^{(i)(j)} + c_{(ij)} B_t^{(ij)}$$

$$= \frac{1}{4} \iint_{0<s<u<t} dB_s^i dB_u^j - \frac{1}{4} \iint_{0<u<s<t} dB_s^i dB_u^j = \frac{1}{4}[B^i, B^j]_t$$

The coefficient of X^{ji} is then equal to $\frac{1}{4}[B^j, B^i]_t$. Since $X^{ij} = -X^{ji}$, joining these two terms, we see that coefficient of X^{ij} is $\frac{1}{2}[B^i, B^j]_t$.

We shall next consider coefficient of X^{ijk}. If i, j, k are different or if $0 = j = k \neq i$, terms corresponding to double indices are 0 and what we have is

$$c_{(i)(j)(k)} B_t^{(i)(j)(k)} + c_{(i)(jk)} B_t^{(i)(jk)} + c_{(ij)(k)} B_t^{(ij)(k)} + c_{(ijk)} B_t^{(ijk)}$$

$$= \frac{1}{9} \iiint_{0<t_i<t_j<t_k<t} dB^i_{t_i} dB^j_{t_j} dB^k_{t_k} - \frac{1}{18} \iiint_{(0<t_i<t_j<t)\cap(0<t_k<t_j<t)} dB^i_{t_i} dB^j_{t_j} dB^k_{t_k}$$

$$- \frac{1}{18} \iiint_{(0<t_j<t_i<t)\cap(0<t_j<t_k<t)} dB^i_{t_i} dB^j_{t_j} dB^k_{t_k} + \frac{1}{9} \iiint_{0<t_k<t_j<t_i<t} dB^i_{t_i} dB^j_{t_j} dB^k_{t_k}$$

$$= \frac{1}{18} \{[[B^i, B^j], B^k]_t + [[B^j, B^k], B^i]_t\}$$

Similarly, the coefficient of X^{jik} is

$$\frac{1}{18} \{[[B^j, B^i], B^k]_t + [[B^i, B^k], B^j]_t\}$$

Since $X^{jik} = -X^{ijk}$, we join these two and see that the coefficient of X^{ijk} is

$$\text{(2.12)} \quad \frac{1}{18} \{2[[B^i, B^j], B^k]_t + [[B^j, B^k], B^i]_t - [[B^i, B^k], B^j]_t\}$$

We have on the other hand Jacobi identity

$$[[B^i, B^j], B^k]_t + [[B^k, B^i], B^j]_t + [[B^j, B^k], B^i]_t = 0.$$

Substitute the above to (2.12), then we see that the coefficient of X^{ijk} is $\frac{1}{18} [[B^i, B^j], B^k]_t$.

If $i \neq j = k \neq 0$ the coefficient of X^{ijk} contains terms with double indices. These are

$$c_{(ijj)} \int_0^t s dB^i_s + c_{(i)(jj)} \int_0^t B^i_s ds = \frac{1}{36} t B^i_t$$

which should be added to the quantity obtained above.

Summarizing these two lemmas, we establish the following theorem.

Theorem 2.3. Suppose that the Lie algebra generated by X_0, ..., X_r is nilpotent of step p. Then the solution of equation (2.1) with $\xi_0 = x$ is represented as $\xi_t(x) = (\exp W_t)(x)$, where

$$(2.13) \quad W_t = \sum_{i=1}^{r} B_t^i X^i + \frac{1}{2} \sum_{i<j} [B^i, B^j]_t [X_i, X_j] + \frac{1}{18} \sum_{i<j,k} [[B^i, B^j], B^k]_t [[X^i, X^j], X^k] + \frac{1}{36} \sum_{i=0}^{r} \sum_{j=1}^{r} tB_t^i [[X_i, X_j], X_j] + \sum_{J;3<|J|\leq p} \{ \sum_{\Delta J}^{*} c_{\Delta J} B_t^{\Delta J} \} X^J$$

Example (Yamato [8]) Consider the equation in R^3 where $X_0 = 0$, $X_1 = \frac{\partial}{\partial x_1} + 2x_2 \frac{\partial}{\partial x_3}$ and $X_2 = \frac{\partial}{\partial x_2} - 2x_1 \frac{\partial}{\partial x_3}$. Then $[X_1, X_2] = -4 \frac{\partial}{\partial x_3}$ and $[[X_1, X_2], X_1] = [[X_1, X_2], X_2] = 0$. Hence the corresponding Lie algebra is nilpotent of step 2. The solution is then written as $\xi_t(x) = \exp W_t(x)$, where

$$W_t = B_t^1 X_1 + B_t^2 X_2 + \frac{1}{2}[B^1, B^2]_t [X_1, X_2]$$

$$= B_t^1 \frac{\partial}{\partial x_1} + B_t^2 \frac{\partial}{\partial x_2} + 2\{B_t^1 x_2 - B_t^2 x_1 - [B^1, B^2]_t\} \frac{\partial}{\partial x_3}$$

Therefore

$$\exp W_t(x) = \begin{cases} x_1 + B_t^1 \\ x_2 + B_t^2 \\ x_3 + 2\{B_t^1 x_2 - B_t^2 x_1 - [B^1, B^2]_t\} \end{cases}$$

where $x = (x_1, x_2, x_3)$.

We shall mention that similar representation is valid for a more general class of stochastic differential equation. Let us consider a vector field valued stochastic process. Let $X(t, x, \omega) = X(t, \omega)$ be a stochastic process such that for each $t > 0$, it is a C^∞-vector field for almost all ω. We assume that it is continuous in t for almost all ω and F_t-adapted, where F_t, $t \geq 0$ is a given family of increasing σ-fields. Suppose we are given $r + 1$ vector field valued stochastic processes $X_0, X_1, \ldots, X_r$. We will call that $\{X_0, X_1, \ldots, X_r\}$ is nilpotent of step p, if

$$[\ \ldots\ [X_{i_1}, X_{i_2}]\ \ldots,]X_{i_m}](t_1, \ldots, t_m\omega) = 0 \qquad \text{a.s. p}$$

holds for any $i_1, \ldots, i_m \in \{0, 1, \ldots, r\}$ and $t_1, \ldots, t_m \geq 0$ if $m > p$.

Let $B_t = (B_t^1, \ldots, B_t^r)$ be a F_t-Brownian motion. Consider the stochastic differential equation

$$(2.14) \qquad \xi_t = x + \int_0^t X_0(s, \xi_s, \omega)ds + \sum_{j=1}^r \int_0^t X_j(s, \xi_s, \omega)\circ dB_s^j$$

Theorem 2.2. Suppose that for each i, $X_i(t, \omega)$ is a complete vector field for any t and a.s. ω. Suppose further that the Lie algebra generated by $\{X_0, X_1, \ldots, X_r\}$ is finite dimensional and nilpotent. Then the solution of (2.14) is represented as $\exp W(t)$, where

$$W(t, x) = \int_0^t X_0(s, x)ds + \sum_{j=1}^r \int_0^t X_j(s, x)\circ dB_s^j$$

$$+ \sum_{J:1<|J|\leq p} \sum_{\Delta J}^{*} c_{\Delta J} \int_A X^J(\hat{s}_1, \ldots, \hat{s}_\ell) \circ dB_{s_1}^{J_1} \ldots \circ dB_s^{J_\ell},$$

where A is the set of (2.5), $B_s^{J_k}$ is defined by (2.7) and $\hat{s}_k = s_k$ if $|J_k| = 1$ and $\hat{s}_k = (s_k, s_k)$ if $|J_k| = 2$. The sum Σ^* is taken for all single or double divided indices ΔJ of J.

The proof is similar to that of Lemma 2.1.

3. Representation of solutions (II). Solvable case

Let $L = L(X_0, X_1, \ldots, X_r)$ be the Lie algebra of vector fields generated by $X_0, X_1, \ldots, X_r$. Define a chain of Lie algebras as $L_1 = [L, L]$, $L_2 = [L_1, L_1]$, $\ldots$, $L_n = [L_{n-1}, L_{n-1}]$. Then $L \supset L_1 \supset L_2 \supset \ldots$ and L_i is an ideal in L_{i-1}. The Lie algebra L is called solvable if there exists p such that $L_p = \{0\}$. By the definition, nilpotent Lie algebra is solvable.

Consider the stochastic differential equation

$$(3.1) \qquad d\xi_t = X_0(\xi_t)dt + \sum_{j=1}^{r} X_j(\xi_t) \circ dB_t^j.$$

The purpose of this section is to show that the above equation is decomposed to a chain of equations whose coefficients are nilpotent vector fields if L is a finite dimensional solvable algebra. We will then prove that the solution of (3.1) is expressed as a composition of solutions of these equations.

The differential of smooth map is needed for our discussion. Let Φ be a diffeomorphism of the manifold M. The differential Φ_* is an automorphism of the space of vector fields defined by

$$(3.2) \qquad \Phi_* X(f)(x) = X(f \circ \Phi)(\Phi^{-1}(x)), \qquad \forall f \in C^\infty(M),$$

where $C^{\infty}(M)$ is the space of all real C^{∞}-functions on M.

Let $\{Y_0, \ldots, Y_r\}$ be a nilpotent subset of L. Consider the stochastic differential equation

$$(3.3) \qquad d\zeta_t = Y_0(\zeta_t)dt + \sum_{j=1}^{r} Y_j(\zeta_t)\circ dB_t^j .$$

The solution $\zeta_t(x)$ starting at x is represented as (2.6), so that ζ_t may be regarded as a diffeomorphism for each $t > 0$ and almost all ω. Set

$$(3.4) \qquad Z_j = (\zeta_t^{-1})_*(X_j - Y_j), \qquad j = 0, \ldots, r.$$

These are vector field valued stochastic processes. Consider

$$(3.5) \qquad d\eta_t = Z_0(\eta_t)dt + \sum_{j=1}^{r} Z_j(\eta_t)\circ dB_t^j$$

Proposition 3.1. Solutions of equations (3.1), (3.3) and (3.5) are linked by the relation $\xi_t = \zeta_t \circ \eta_t$.

Proof. Using a local coordinate, we shall write $\eta_t = (\eta_t^1, \ldots, \eta_t^d)$ etc. We put t as B_t^0 for convention. Let f be a C^{∞}-function. By Ito's formula we have

$$(3.6) \qquad f(\zeta_t \circ \eta_t) - f(x) = \sum_k \int_0^k \frac{\partial f}{\partial x_k}(\zeta_t(x))\circ d\zeta_t^k(x)\Big|_{x=\eta_t}$$
$$+ \sum_{\ell} \int_0^t \frac{\partial}{\partial x_{\ell}}(f\circ\zeta_t)(\eta_t)\circ d\eta_t^{\ell}$$

Since $d\zeta_t^k = \sum Y_j^k(\zeta_t)\circ dB_t^j$, the first term of the right hand side equals

$$\sum_k \int_0^t \sum_j Y_j^k(\zeta_t(x))\frac{\partial f}{\partial x_k}(\zeta_t(x))\circ dB_t^j \Big|_{x=\eta_t}$$

$$= \sum_j \int_0^t Y_j f(\zeta_t \circ \eta_t)\circ dB_t^j.$$

The second term is

$$\sum_\ell \int_0^t \sum_j Z_j^\ell(\eta_t)\frac{\partial}{\partial x_\ell}(f\circ\zeta_t)(\eta_t)\circ dB_t^j$$

$$= \sum_j \int_0^t Z_j(f\circ\zeta_t)(\eta_t)\circ dB_t^j$$

$$= \sum_j \int_0^t (X_j - Y_j)(f\circ\zeta_t\circ\zeta_t^{-1})(\zeta_t\circ\eta_t)\circ dB_t^j$$

$$= \sum_j \int_0^t (X_j - Y_j)f(\zeta_t\circ\eta_t)\circ dB_t^j.$$

Therefore we have

$$f(\zeta_t\circ\eta_t) - f(x) = \sum_j \int_0^t X_j f(\zeta_t\circ\eta_t)\circ dB_t^j$$

Since this holds for any C^∞-function, we see that $\zeta_t\circ\eta_t(x)$ is the solution of (3.1) starting at x. The proof is complete.

Now we shall decompose vector fields $X_0, \ldots, X_r$ into sums of vector fields

$$X_0 = X_0^{(1)} + \ldots + X_0^{(n)}, \quad \ldots, \quad X_r = X_r^{(1)} + \ldots + X_r^{(n)}$$

such that

$$L^{(1)} = L(X_0^{(1)}, \ldots, X_r^{(1)}) , \ldots, L^{(n)} = L(X_0^{(n)}, \ldots, X_r^{(n)})$$

are all nilpotent Lie algebra. Such a decomposition exists always, although it is not unique. For example, let us choose a basis of $L = L(X_0, \ldots, X_r)$ and denote it as $Y_1, \ldots, Y_n$. Then each X_i is written as $X_i = \sum_{j=1}^{n} a_{ij} Y_j$. Setting $X_i^{(j)} = a_{ij} Y_j$, for example, we have a decomposition mentioned above.

Let us now consider a chain of stochastic differential equations.

$$(3.7) \qquad d\zeta_t^{\ell} = \sum_{j=0}^{r} X_j^{\ell} \circ dB_t^j, \qquad \ell = 1, \ldots, n$$

$$(3.8) \qquad d\xi_t^{\ell} = \sum_{j=0}^{r} (\xi_t^{\ell-1})_*^{-1} \cdots (\xi_t^1)_*^{-1} X_j^{\ell} \circ dB_t^j , \qquad \ell = 2, 3, \ldots n$$

where $\xi_t^1 = \zeta_t^1$ and

$$(3.9) \qquad d\eta_t^{\ell} = \sum_{j=0}^{r} (\zeta_t^{\ell})_*^{-1} \{ (\xi_t^{\ell-1})_*^{-1} \cdots (\xi_t^1)_*^{-1} X_j^{\ell} - X_j^{\ell} \} \circ dB_t^j, \qquad \ell = 2, 3, \ldots, n$$

Since $L^{(\ell)}$ is nilpotent, the solution $\zeta_t^{\ell}(x)$ is a diffeomorphism of M for each t and a.s. ω. Hence the differential $(\zeta_t^{\ell})_*^{-1}$ is well defined. In order to show the analogous fact for $\xi_t^{\ell}(x)$ and $\eta_t^{\ell}(x)$, we require

Lemma 3.2. Coefficients of equations on η_t^{ℓ} are nilpotent.

Proof. We will prove that coefficients of the equation are in L_1, since L_1 is nilpotent ([5], p. 51). We first consider the case $\ell = 2$. Since $\xi_t^1(x) = \zeta_t^1(x) = \exp W_t$ where W_t is the vector field valued stochastic process of (2.9), it is a diffeomorphism of M for each t and a.s. ω. Hence the differential $(\xi_t^1)_*^{-1}$ is well defined. We shall show that $(\xi_t^1)_*^{-1}X_j^\ell - X_j^\ell$ belongs to L_1 a. s. P for each j and ℓ, following the argument of Ichihara-Kunita [3]. Let us choose $Y_1, \ldots, Y_n$ as a basis of L, such that $Y_1, \ldots, Y_k$ $(k < n)$ is a basis of L_1. Set $Y_k(s) = (e^{sW_t})_* Y_k$, the parameter t being fixed. Then it is known that

$$\frac{dY_k(s)}{ds} = (e^{sW_t})_*[W_t, Y_k]$$

Since $[W_t, Y_k]$ in L_1, it is written as

$$[W_t, Y_k] = \sum_{i=1}^{n} a_{ki}(t)Y_i, \qquad a_{ki}(t) = 0 \text{ if } k < i \leq n.$$

Then the above equation derives a system of linear differential equations

$$\frac{dY(s)}{ds} = AY(s), \quad Y(s) = [Y_1(s), \ldots, Y_n(s)], \quad A = (a_{ki}(s))$$

The solution is then written as

$$Y(s) = e^{As}Y(0) = \sum_{p=0}^{\infty} \frac{s^p}{p!} A^p Y(0)$$

Note that $a_{mi}^{(p)} = 0$ if $k < i \leq n$, where $(a_{mi}^{(p)}) = A^p$. Then $Y(s) - Y(0)$ is a linear sum of $Y_1, \ldots, Y_k$. Since X_j^ℓ is written as a linear sum of $Y_1, \ldots, Y_n$, $(e^{sW_t})_* X_j^\ell - X_j^\ell$ also a linear sum of $Y_1, \ldots, Y_k$, so that it is in L_1. Since $(e^{-W_t})_* = (e^{W_t})_*^{-1}$, we see that $(e^{W_t})_*^{-1} X_j^\ell - X_j^\ell$ is in L_1 a.s. P for any $t > 0$, $0 \leq j \leq r$, $1 \leq \ell \leq n$.

Now noting that L_1 is an ideal in L, we can show similarly as the above argument that $(\zeta_t^2)_*^{-1}$ maps L_1 into itself. Therefore $(\zeta_t^2)_*^{-1}\{(\xi_t^2)_*^{-1}X_j^\ell - X_j^\ell\}$ is in L_1 a.s. P for any $t > 0$, $0 \leq j \leq r$ and $1 \leq \ell \leq n$. We have thus shown that coefficients of equation on η_t^2 are nilpotent.

The solution ξ_t^2 has the decompotion $\zeta_t^2 \circ \eta_t^2$ by Lemma 3.1. Hence ξ_t^2 is a diffeomorphism of M for each t and a.s. ω. Thus equations (3.8) and (3.9) are well defined for $\ell = 3$. We then see that coefficients of equation on η_t^3 are nilpotent as before. Repeating this argument, it turns out that coefficients of equations on η_t^ℓ are nilpotent for all $\ell = 2, 3, \ldots, n$. The proof is complete.

We can now show the following theorem.

Theorem 3.3. Suppose that the Lie algebra generated by X_0, $X_1, \ldots, X_r$ is finite dimensional and solvable. Then the solution of the equation (3.1) is represented as

$$(3.10) \qquad \xi_t = \zeta_t^1 \circ \zeta_t^2 \circ \eta_t^2 \circ \cdots \circ \zeta_t^n \circ \eta_t^n ,$$

where ζ_t^ℓ and η_t^ℓ are solutions of equations (3.7) and (3.9) with nilpotent coefficients.

Proof. We have $\xi_t = \xi_t^1 \circ \cdots \circ \xi_t^n$ by Lemma 3.1. Furthermore it holds $\xi_t^\ell = \zeta_t^\ell \circ \eta_t^\ell$ for $\ell = 2, \ldots, n$. Hence we get the representation (3.10).

If coefficients of equations on ξ_t^ℓ in (3.8) are already nilpotent, it is not necessary to decompose it to ζ_t^ℓ and η_t^ℓ, and we may obtain a shorter decomposition of ξ_t. This occurs if X_j^ℓ, $j = 0, \ldots, r$ are in the derived ideal L_1, since $(\xi_t^{\ell-1})_*^{-1} \cdots (\xi_t^1)_*^{-1} X_j^\ell$ are in L_1 for any $j = 0, \ldots, r$. We shall discuss two examples.

Example 1. (Linear System) Consider

$$(3.11)\qquad d\xi_t = A\xi_t dt + CB_t,$$

where A is a d×d-matrix and C is a d×r-matrix. Corresponding vector fields are

$$X_0 = \sum_i(\sum_j a_{ij}x_j)\frac{\partial}{\partial x_i}\ ,\quad X_j = \sum_i c_{ij}\frac{\partial}{\partial x_i}$$

It holds

$$(adX_0)^n X_j = (-1)^n \sum_i (A^n C)_{ij}\frac{\partial}{\partial x_i}$$

The Lie algebra L generated by X_0 and X_j is the linear span of X_0 and $(ad^n X_0)X_j$, n = 0, 1, 2, ... The derived ideal $L_1 = [L, L]$ is the linear span of $(adX_0)^n X_j$, n = 1, 2, ... Since latters are commuting each other, it holds $L_2 = [L_1, L_1] = \{0\}$. Hence L is solvable.

Consider two equations

$$d\xi_t^1 = X_0 dt$$

$$d\xi_t^2 = \sum_{j=1}^{r} (\xi_t^1)_*^{-1} X_j \circ dB_t^j$$

Then it holds $\xi_t^1(x) = e^{tA}x$ and

$$(\xi_t^1)_*^{-1} X_j = \sum_i (e^{-tA}C)_{ij}\frac{\partial}{\partial x_i}\ ,$$

so that

$$d\xi_t^2 = e^{-tA}C\circ dB_t$$

The solution is $\xi_t^2(x) = x + \int_0^t e^{-sA}CdB_s$. Therefore we have

$$\xi_t = \xi_t^1\circ\xi_t^2 = e^{tA}(x + \int_0^t e^{-sA}CdB_s).$$

This is a well known formula for the solution of the linear system.

Example 2. (Bilinear system) Consider the equation

(3.12) $$d\xi_t = A_0\xi_t dt + \sum_{j=1}^{r} A_j\xi_t\circ dB_t^j$$

where $A_j = (a_{k,\ell}^{(j)})$, $j = 0, \ldots, r$ are d×d-triangular matrices such that $a_{k_\ell}^{(j)} = 0$ if $k > \ell$. Corresponding vector fields are

$$X_j = \sum_k(\sum_\ell a_{k\ell}^{(j)}x_\ell)\frac{\partial}{\partial x_k}, \quad j = 0, \ldots, r.$$

It holds $[X_i, X_j] = -\sum_k(\sum_\ell [A_i, A_j]_{k\ell}x_\ell)\frac{\partial}{\partial x_k}$, where $[A_i, A_j] = A_iA_j - A_jA_i$. Hence $L(X_0, X_1, \ldots, X_r)$ is isomorphic to the matrix Lie algebra $L(A_0, \ldots, A_r)$. The derived ideal L_1 of $L(A_0, \ldots, A_r)$ consists of nilpotent matrices as is easily seen. Thus $L(A_0, \ldots, A_r)$, or equivalently, $L(X_0, \ldots, X_r)$ is a solvable Lie algebra.

We shall decompose matrices A_j to sums of diagonal matrices $D_j = (\delta_{k\ell}a_{k\ell}^{(j)})$ and nilpotent ones $N_j = ((1 - \delta_{k\ell})a_{k\ell}^{(j)})$, where $\delta_{k\ell}$ is Kronecker's delta. Consider

$$d\xi_t^1 = \sum_{j=0}^{r} D_j \xi_t^j \circ dB_t^j$$

where $B_t^0 = t$. The solution is then written as

$$(3.13) \qquad \xi_t^1(x) = e^{W_t}x, \quad W_t = \sum_{j=0}^{r} B_t^j D_j$$

It holds $(\xi_t^1)_*^{-1} N_j = e^{-W_t} N_j e^{W_t}$. Consider

$$d\xi_t^2 = \sum_{j=0}^{r} e^{-W_t} N_j e^{W_t} \xi_t^2 \circ dB_t^j$$

Since $e^{-W_t} N_j e^{W_t}$, $j = 0, \ldots, r$ are nilpotent matrices, the solution is represented as $\xi_t^2(x) = e^{V_t}x$, where

$$(3.14) \qquad V_t = \sum_{j=0}^{r} \int_0^t e^{-W_s} N_j e^{W_s} \circ dB_s^j$$

$$+ \frac{1}{2} \sum_{i<j} \iint_{0<s<u<t} [e^{-W_s} N_i e^{W_s}, e^{-W_u} N_j e^{W_u}] \circ (dB_s^i dB_u^j - dB_s^j dB_u^i) + \ldots$$

The solution of (3.12) is then written as,

$$\xi_t(x) = e^{W_t} e^{V_t} x.$$

Department of Applied Science
Faculty of Engineering
Kyushu University
Fukuoka, 812 Japan

References

[1] Doss, H: Liens entre équations différentielles stochastiques et ordinaires, Ann. Inst. H. Poincaré 13, 99-125(1977).

[2] Ichihara, K - Kunita, H: A classification of the second order degenerate elliptic operators and its probabilistic characterization, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 30, 235-254(1974).

[3] Ichihara, K - Kunita, H: Supplements and corrections to the above paper, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 39, 81-84(1977).

[4] Itô, K: Stochastic differentials, Appl. Math. Optimization 1, 374-384(1975).

[5] Jacobson, N: Lie algebras, New York-London. Wiley 1962.

[6] Palais, R. S : A global formulation of the Lie theory on transformation groups, Memoirs of the American Mathematical Society, No. 22 (1957).

[7] Sussman, H. J : On the gap between deterministic and stochastic ordinary differential equations, The Annals of Probability 6, 19-41(1978).

[8] Yamato, Y: Stochastic differential equations and nilpotent Lie algebras, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 47, 213-229(1979).