Astérisque
327, 2009, p. 93–114

TWO-PARAMETER STOCHASTIC CALCULUS AND MALLIAVIN'S INTEGRATION-BY-PARTS FORMULA ON WIENER SPACE

by

James R. Norris

Dedicated to Jean-Michel Bismut on the occasion of his 60th birthday

Abstract. — The integration-by-parts formula discovered by Malliavin for the Itô map on Wiener space is proved using the two-parameter stochastic calculus. It is also shown that the solution of a one-parameter stochastic differential equation driven by a two-parameter semimartingale is itself a two-parameter semimartingale.

***Résumé* (Calcul stochastique à deux paramètres et formule d'intégration par parties de Malliavin sur l'espace de Wiener)**
La formule d'intégration par parties, qui a été établie par Malliavin pour l'application d'Itô sur l'espace de Wiener, est démontrée en utilisant le calcul stochastique à deux paramètres. On montre aussi que la solution d'une équation différentielle stochastique à un paramètre, guidée par une semimartingale à deux paramètres, est elle-même une semimartingale à deux paramètres.

1. Introduction

The stochastic calculus of variations was conceived by Malliavin [**6, 7, 8**] as follows. Let $(z_t)_{t\geqslant 0}$ denote the Ornstein–Uhlenbeck process on Wiener space $(W, \mathcal{W}, \mu)$ and let $\Phi : W \to \mathbb{R}^d$ denote the (almost-everywhere unique) Itô map obtained by solving a stochastic differential equation in $\mathbb{R}^d$ up to time 1. Then $(z_t)_{t\geqslant 0}$ is stationary and reversible, so, for functions f, g on $\mathbb{R}^d$, setting $F = f \circ \Phi, G = g \circ \Phi$,

$$\mathbb{E}\left[\{F(z_t) - F(z_0)\}\{G(z_t) - G(z_0)\}\right] = -2\mathbb{E}\left[F(z_0)\{G(z_t) - G(z_0)\}\right]. \tag{1}$$

Once certain terms of mean zero are subtracted, a differentiation of this identity with respect to t *inside the expectation* is possible, and leads to the integration-by-parts

2010 Mathematics Subject Classification. — 60H07; 60H15.
Key words and phrases. — Malliavin calculus, two-parameter stochastic calculus, integration by parts formula.

formula on Wiener space

$$\int_W \nabla_i f(\Phi)\Gamma^{ij}\nabla_j g(\Phi)d\mu = -\int_W f(\Phi)LGd\mu, \tag{2}$$

where LG and the *covariance matrix* Γ will be defined below. As is now well known, this formula and its generalizations hold the key to many deep results of stochastic analysis.

Malliavin's proof of the integration-by-parts formula was based on a *transfer principle*, allowing some calculations for two-parameter random processes to be made using classical differential calculus. Stroock [**11, 12, 13**] and Shigekawa [**10**] gave alternative derivations having a a more functional-analytic flavour. Bismut [**1**] gave another derivation based on the Cameron–Martin–Girsanov formula. Elliott and Kohlmann [**3**] and Elworthy and Li [**4**] found further elementary approaches to the formula. The alternative proofs are relatively straightforward. Nevertheless, we have found it interesting to go back to Malliavin's original approach in [**8**] and to review the calculations needed, especially since this can be done now in a more explicit way using the two-parameter stochastic calculus, as formulated in [**9**].

In Section 2 we review in greater detail the various mathematical objects mentioned above. Then, in Section 3, we review some points of two-parameter stochastic calculus from [**9**]. Section 4 contains the main technical result of the paper, which is a regularity property for two-parameter stochastic differential equations. We consider equations in which some components are given by two-parameter integrals and others by one-parameter integrals. It is shown, under suitable hypotheses, that the components which are presented as one-parameter integrals are in fact two-parameter semimartingales. This is useful because one can then compute martingale properties for both parameters by stochastic calculus. The sorts of differential equation to which this theory applies are just one way to realise continuous random processes indexed by the plane. See the survey [**5**] by Léandre for a wider discussion. But this regularity property makes our processes more tractable to analyse than some others. This is illustrated in Section 5, where we do the calculations needed to obtain the integration-by-parts formula.

2. Integration-by-parts formula

The Wiener space $(W, \mathcal{W}, \mu)$ over $\mathbb{R}^m$ is a probability space with underlying set $W = C([0,\infty), \mathbb{R}^m)$, the set of continuous paths in $\mathbb{R}^m$. Let $\mathcal{W}^o$ denote the σ-algebra on W generated by the family of coordinate functions $w \mapsto w_s : W \to \mathbb{R}^m$, $s \geqslant 0$, and let μ^o be Wiener measure on $\mathcal{W}^o$, that is to say, the law of a Brownian motion in $\mathbb{R}^m$ starting from 0. Then $(W, \mathcal{W}, \mu)$ is the completion of the probability space $(W, \mathcal{W}^o, \mu^o)$. Write $\mathcal{W}_s$ for the μ-completion of $\sigma(w \mapsto w_r : r \leqslant s)$. Let $X_0, X_1, \dots, X_m$ be vector fields on $\mathbb{R}^d$, with bounded derivatives of all orders. Fix $x_0 \in \mathbb{R}^d$ and consider the stochastic differential equation

$$\partial x_s = X_i(x_s)\partial w_s^i + X_0(x_s)\partial s.$$

Here and below, the index i is summed from 1 to m, and ∂ denotes the Stratonovich differential. There exists a map $x : [0, \infty) \times W \to \mathbb{R}^d$ with the following properties:

- x is a continuous semimartingale on $(W, \mathscr{W}, (\mathscr{W}_s)_{s \geqslant 0}, \mu)$,
- for μ-almost all $w \in W$, for all $s \geqslant 0$ we have

$$x_s(w) = x_0 + \int_0^s X_i(x_r(w))\partial w_r^i + \int_0^s X_0(x_r(w))dr.$$

The first integral in this equation is the Stratonovich stochastic integral. Moreover, for any other such map x', we have $x_s(w) = x'_s(w)$ for all $s \geqslant 0$, for μ-almost all w. We have chosen here a Stratonovich rather than an Itô formulation to be consistent with later sections, where we have made this choice in order to take advantage of the simpler calculations which the Stratonovich calculus allows. The Itô map referred to above is the map $\Phi(w) = x_1(w)$.

We can define on some complete probability space, $(\Omega, \mathscr{F}, \mathbb{P})$ say, a two-parameter, continuous, zero-mean Gaussian field $(z_{st} : s, t \geqslant 0)$ with values in $\mathbb{R}^m$, and with covariances given by

$$\mathbb{E}(z_{st}^i z_{s't'}^j) = \delta^{ij}(s \wedge s')e^{-|t-t'|/2}.$$

Such a field is called an Ornstein–Uhlenbeck sheet. Set $z_t = (z_{st} : s \geqslant 0)$. Then, for $t > 0$, both z_0 and z_t are Brownian motions in $\mathbb{R}^m$ and (z_0, z_t) and (z_t, z_0) have the same distribution. We have now defined all the terms in, and have justified, the identity (1).

Consider the following stochastic differential equation for an unknown process $(U_s : s \geqslant 0)$ in the space of $d \times d$ matrices

$$\partial U_s = \nabla X_i(x_s)U_s\partial w_s^i + \nabla X_0(x_s)U_s\partial s, \quad U_0 = I.$$

This equation may be solved, jointly with the equation for x, in exactly the same sense as the equation for x alone. Thus we obtain a map $U : [0, \infty) \times W \to \mathbb{R}^d \otimes (\mathbb{R}^d)^*$, with properties analogous to those of x. Moreover, by solving an equation for the inverse, we can see that $U_s(w)$ remains invertible for all $s \geqslant 0$, for almost all w. Write U_s^* for the transpose matrix and set $\Gamma_s = U_s C_s U_s^*$, where

$$C_s = \int_0^s U_r^{-1}X_i(x_r) \otimes U_r^{-1}X_i(x_r)dr.$$

Set also

$$L_s = -U_s \int_0^s U_r^{-1}X_i(x_r)\partial w_r^i + U_s \int_0^s U_r^{-1}\{\nabla^2 X_i(x_r)\partial w_r^i + \nabla^2 X_0(x_r)dr\}\Gamma_r,$$
$$+ U_s \int_0^s U_r^{-1}\nabla X_i(x_r)X_i(x_r)dr$$

and define for $G = g \circ \Phi$

$$LG = L_1^i \nabla_i g(x_1) + \Gamma_1^{ij}\nabla_i\nabla_j g(x_1).$$

We have now defined all the terms appearing in the integration-by-parts formula (2). We will give a proof in Section 5.

3. Review of two-parameter stochastic calculus

In [**9**], building on the fundamental works of Cairoli and Walsh [**2**] and Wong and Zakai [**14, 15**], we gave an account of two-parameter stochastic calculus, suitable for the development of a general theory of two-parameter hyperbolic stochastic differential equations. We recall here, for the reader's convenience, the main features of this account.

We take as our probability space $(\Omega, \mathcal{F}, \mathbb{P})$ the canonical complete probability space of an m-dimensional Brownian sheet $(w_{st} : s,t \geqslant 0)$, extended to a process $(w_{st} : s,t \in \mathbb{R})$ by independent copies in the other three quadrants. Thus $w_{st} = (w^1_{st}, \dots, w^m_{st})$ is a continuous, zero-mean Gaussian process, with covariances given by

$$\mathbb{E}(w^i_{st} w^j_{s't'}) = \delta^{ij}(s \wedge s')(t \wedge t'), \quad i,j = 1,\dots,m, \quad s,t \geqslant 0, \quad s',t' \geqslant 0.$$

It will be convenient to define also $w^0_{st} = st$ for all $s,t \in \mathbb{R}$. For $s,t \geqslant 0$, write $\mathcal{F}_{st}$ for the completion with respect to $\mathbb{P}$ of the σ-algebra generated by w_{ru} for $r \in (-\infty, s]$ and $u \in (-\infty, t]$. We say that a two-parameter process $(x_{st} : s,t \geqslant 0)$ is *adapted* if x_{st} is $\mathcal{F}_{st}$-measurable for all $s,t \geqslant 0$, and is *continuous* if $(s,t) \mapsto x_{st}(\omega)$ is continuous on $(\mathbb{R}^+)^2$ for all $\omega \in \Omega$. The *previsible* σ-algebra on $\Omega \times (\mathbb{R}^+)^2$ is that generated by sets of the form $A \times (s,s'] \times (t,t']$ with $A \in \mathcal{F}_{st}$. If we allow $A \in \mathcal{F}_{s\infty}$ in this definition, we get the *s-previsible* σ-algebra.

The classical approach to defining stochastic integrals, by means of an isometry of Hilbert spaces, adapts in a straightforward way from one-dimensional times to two, allowing the construction of stochastic integrals with respect to certain two-parameter processes, in particular with respect to the Brownian sheet. Given an s-previsible process[(1)] $(a_s(t) : s,t \geqslant 0)$, such that

$$\mathbb{E} \int_0^s \int_0^t a_r(u)^2 dr du < \infty$$

for all $s,t \geqslant 0$, we can define, for $i = 1,\dots,m$ and all $t_1, t_2 \geqslant 0$ with $t_1 \leqslant t_2$, one-parameter processes M and A by

$$M_s = \int_0^s \int_{t_1}^{t_2} a_r(t) d_r d_t w^i_{rt}, \quad A_s = \int_0^s \int_{t_1}^{t_2} a_r(t)^2 dr dt. \tag{3}$$

Then M is a continuous $(\mathcal{F}_{s\infty})_{s \geqslant 0}$-martingale, with quadratic variation process $[M] = A$. A localization argument by adapted initial open sets (see below) allows an extension of the integral under weaker integrability conditions. By the Burkholder–Davis–Gundy inequalities, for all $\alpha \in [2,\infty)$, there is a constant $C(\alpha) < \infty$ such that

$$\mathbb{E}\left(\left| \int_{s_1}^{s_2} \int_{t_1}^{t_2} a_s(t) d_s d_t w^i_{st} \right|^\alpha \right) \leqslant C(\alpha) \mathbb{E}\left(\left| \int_{s_1}^{s_2} \int_{t_1}^{t_2} a_s(t)^2 ds dt \right|^{\alpha/2} \right). \tag{4}$$

(1) We write any time parameter with respect to which a process is previsible, here s, as a subscript. Where previsibility is not assumed, here in t, we write the parameter in parentheses.

By an (s,t)-*semimartingale*, s-*semimartingale*, t-*semimartingale*, we mean, respectively, previsible processes $(x_{st}: s,t \geqslant 0)$, $(p_{st}: s,t \geqslant 0)$, $(q_{st}: s,t \geqslant 0)$ for which we may write

$$\begin{aligned}
&x_{st} - x_{s0} - x_{0t} + x_{00} \\
&\quad = \sum_{i=0}^{m} \int_0^s \int_0^t (x''_{ru})_i d_r d_u w^i_{ru} + \sum_{i,j=0}^{m} \int_0^s \int_{-1}^t \left(\int_{-1}^s \int_0^t (x''_{ru}(r',u'))_{ij} d_{r'} d_u w^j_{r'u} \right) d_r d_{u'} w^i_{ru'}
\end{aligned}$$

and

$$p_{st} - p_{0t} = \sum_{i=0}^{m} \int_0^s \int_{-1}^t (p'_{rt}(u'))_i d_r d_{u'} w^i_{ru'}, \quad q_{st} - q_{s0} = \sum_{i=0}^{m} \int_{-1}^s \int_0^t (q'_{su}(r'))_i d_{r'} d_u w^i_{r'u}.$$

Here, $(x''_{st}: s,t \geqslant 0)$ is a previsible process, having components $(x''_{st})_i$, subject to certain local integrability conditions, which are implied, in particular, by almost sure local boundedness. The process $(x''_{st}(r,u): s,t \geqslant 0, r,u \in \mathbb{R})$ is required to be previsible in (ω,s,t) and (Borel) measurable in (r,u), with $x''_{st}(r,u) = 0$ for $r > s$ or $u > t$, and is subject to similar local integrability conditions. The inner and outer parts of the second integral are both cases of the stochastic integral at (3), or its t-analogue, or of the usual Lebesgue integral, and the value of the iterated integral is unchanged if we reverse the order in which the integrals are taken. The integrals appearing in the expression for x_{st} are called stochastic integrals of the *first* and *second kind*. The processes $(p'_{st}(u): s,t \geqslant 0, u \in \mathbb{R})$ and $(q'_{st}(r): s,t \geqslant 0, r \in \mathbb{R})$ are required to be previsible in (ω,s,t) and measurable in u and r, respectively, with $p'_{st}(u) = 0$ for $u > t$ and $q'_{st}(r) = 0$ for $r > s$, and are subject to similar local integrability conditions. For fixed $t \geqslant 0$, if $(x_{s0}: s \geqslant 0)$ is a continuous $(\mathcal{F}_{s0})_{s\geqslant 0}$-semimartingale, then $(x_{st}: s \geqslant 0)$ is a continuous $(\mathcal{F}_{st})_{s\geqslant 0}$-semimartingale, in the usual one-parameter sense. Also $(p_{st}: s \geqslant 0)$ is a continuous $(\mathcal{F}_{st})_{s\geqslant 0}$-semimartingale, for all $t \geqslant 0$.

The heuristic formulae

$$\begin{aligned}
d_s d_t x_{st} &= \sum_{i=0}^{m} (x''_{st})_i d_s d_t w^i_{st} + \sum_{i,j=0}^{m} \int_{-1}^s \int_{-1}^t (x''_{st}(r,u))_{ij} d_s d_u w^i_{su} d_r d_t w^j_{rt}, \\
d_s p_{st} &= \sum_{i=0}^{m} \int_{-1}^t (p'_{st}(u))_i d_s d_u w^i_{su}, \\
d_t q_{st} &= \sum_{i=0}^{m} \int_{-1}^s (q'_{st}(r))_i d_r d_t w^i_{rt}
\end{aligned}$$

provide a good intuition in representing the two-parameter increment

$$d_s d_t x_{st} = x_{s+ds,t+dt} - x_{s,t+dt} - x_{s+ds,t} + x_{st}$$

and the one-parameter increments $d_s p_{st} = p_{s+ds,t} - p_{st}$ and $d_t q_{st} = q_{s,t+dt} - q_{st}$ in terms of a linear combinations of increments, and of products of increments of the Brownian sheet.