The Annals of Probability
1985, Vol. 13, No. 2, 609–615

MALLIAVIN DERIVATIVES AND DERIVATIVES OF FUNCTIONALS OF THE WIENER PROCESS WITH RESPECT TO A SCALE PARAMETER

BY MOSHE ZAKAI[1]

Technion—Israel Institute of Technology

Let $F(c_0 w)$ be a functional of the Wiener process with variance parameter c_0^2 and let $F(cw)$ be an extension of $F(c_0 w)$ to $F(cw)$, $c \in (0, c_0)$. Relations are derived between the Malliavin derivatives, between the derivatives with respect to the scale parameter $(\partial F(\rho cw)/\partial \rho)_{\rho=1}$ and 'noncoherent derivatives' such as $(dE(F(cw + \sqrt{\varepsilon} c\tilde{w}) \mid w)/d\varepsilon)_{\varepsilon=0}$ where $\tilde{w}$ is another Wiener process independent of w and between the generator of the nontime-homogeneous Ornstein-Uhlenbeck process.

I. Introduction. Let w denote the standard Wiener process on $[0, T]$; for c real, cw will denote the Wiener process with variance parameter c^2 $(E((cw)(t))^2 = c^2 t)$. Let $F(c_0 w)$ be a square integrable functional of $c_0 w$. Then in general $F(c_0 w)$ may not be extendable to a continuous functional on the space of continuous functions on $[0, T]$ starting at zero. Note, however, that $F(c_0 w)$ may always be extended to $F(cw)$, for all c satisfying $0 < c < c_1$ for some c_1, $c_1 \geq c_0$ but such an extension is not unique. Derivatives of functionals of the Wiener process have first been considered for differentiable functions on the space of continuous functions and derivatives of the form $(\partial F(c_0 w + \varepsilon H)/\partial \varepsilon)_{\varepsilon=0}$ where $H(t) = \int_0^t f(\vartheta)\, d\vartheta$; in which case the measure induced by $c_0 w + \varepsilon H$ on the space of continuous functions is absolutely continuous with respect to the measure induced by $c_0 w$ (cf. [4]). Malliavin introduced a different class of derivatives and an associated calculus for functionals $F(c_0 w)$ defined for some fixed c_0 [5]. The Malliavin calculus can be introduced either directly ([2], [7]) or through a time homogeneous infinite dimensional Ornstein-Uhlenbeck process ([5], [6], [8], [9]). For the case where $F(c_0 w)$ is the restriction to $c_0 w$ of a twice differentiable functional on the space of continuous functions starting at zero, the relation between the Malliavin calculus and the calculus of Fréchet derivatives is well known. The purpose of this note is to consider the relation between Malliavin derivatives to certain derivatives related to $F(cw)$ where c is in the interval $(0, c_1)$, and in particular the derivative of F with respect to the scale parameter c.

II. Preliminaries. Let $C_0(T)$ denote the space of real valued continuous functions on $[0, T]$ which vanish at $t = 0$. Let $\mathscr{S}^c$ denote the σ-field induced by Borel measurable subsets of $C_0(T)$ and completed with respect to the measure

Received November 1983; revised April 1984.
[1] The research was supported by the fund for the promotion of research at the Technion.
AMS 1980 *subject classifications.* Primary 60H99; secondary 60J65.
Key words and phrases. Malliavin derivatives, Malliavin calculus, derivatives of Wiener functionals, the infinite dimensional Ornstein-Uhlenbeck process.

induced by the nonstandard Wiener process cw on $[0, T]$. Whenever we consider the functional $F(cw)$ for $0 < c \le c_1$ we will assume that $F(\cdot)$ is defined for all the elements of $C_0(T)$ and for every c, $c \in (0, c_1]$, $F(cw)$ is a Wiener functional, that is, an $\mathscr{F}^c$ adapted random variable. In other words, $F(cw)$ is assumed to be scale invariant measurable in the sense of Johnson and Skoug [3].

Turning now to the definition of the Malliavin derivative $\mathscr{L}$, let $F(c_0 w)$ be a square integrable functional of $c_0 w$ for some fixed c_0, then $F(c_0 w)$ has the multiple Wiener-Itô integral representation:

$$(1)\quad F(c_0 w) = K + \textstyle\sum_{n=1}^{\infty} \displaystyle\int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(c_0 w_{t_1}) \cdots d(c_0 w_{t_n}).$$

If the series

$$\textstyle\sum_{n=1}^{N} n \displaystyle\int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(c_0 w_{t_1}) \cdots d(c_0 w_{t_n})$$

converges in the mean then we say that $F(c_0 w)$ is in the domain of $\mathscr{L}$ and define ([5], [2], [7], [8])

$$(2)\quad \mathscr{L}F(c_0 w) = \textstyle\sum_{n=1}^{\infty} n \displaystyle\int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(c_0 w_{t_1}) \cdots d(c_0 w_{t_n}).$$

REMARK. Let $\mathscr{L}[n]$ denote the Malliavin operation $\mathscr{L}$ or L as defined in reference $[n]$ then the formal relation between $\mathscr{L}[n]$ and $\mathscr{L}$ as defined by (2) is as follows: $\mathscr{L}[2] = \mathscr{L}[6] = \mathscr{L}[7] = -\mathscr{L}$ and $\mathscr{L}[5] = \mathscr{L}[8] = -\frac{1}{2}\mathscr{L}$. The reason for reversing the sign in the definition of $\mathscr{L}$ is that it appears more natural in the context of derivatives with respect to a scale parameter (cf. equations (8) and (13a)). Let $c\tilde{w}$ denote a Wiener process with variance parameter c^2 which is independent of cw and let $\mathscr{F}^w$ denote the sub σ-field generated by cw. Since $F(cw)$ was assumed to be a Wiener functional of cw for some fixed c and since cw and $(\sqrt{1-\varepsilon}\,cw + \sqrt{\varepsilon}\,c\tilde{w})$ have the same variance parameter, $F(\sqrt{1-\varepsilon}\,cw + \sqrt{\varepsilon}\,c\tilde{w})$ is well defined. Furthermore, since $F(cw)$ was assumed square integrable, it follows immediately that

$$(3)\quad \begin{aligned} &E(F(\sqrt{1-\varepsilon}\,cw + \sqrt{\varepsilon}\,c\tilde{w}) \mid \mathscr{F}^w) \\ &\quad = K + \textstyle\sum_{n=1}^{\infty} (1-\varepsilon)^{n/2} \displaystyle\int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(cw_{t_1}) \cdots d(cw_{t_n}). \end{aligned}$$

Consequently, if $F(cw)$ is in the domain of $\mathscr{L}$ then

$$(4)\quad (1/\varepsilon)F(cw) - E(F(\sqrt{1-\varepsilon}\,cw + \sqrt{\varepsilon}\,c\tilde{w}) \mid \mathscr{F}^w) \xrightarrow[\varepsilon \downarrow 0]{\text{q.m.}} \tfrac{1}{2}\mathscr{L}\, F(cw).$$

The nontime-homogeneous Ornstein-Uhlenbeck process in infinite dimension, can be introduced as follows. Let W_{st}, t, $s \ge 0$, denote the Brownian sheet and consider the process $X_{s,t}$, $s \ge 0$, $0 \le t \le T$ satisfying

$$\partial_s X_{s,t} = -\alpha X_{s,t}\, ds + \partial_s W_{s,t}$$

with $X_{0,t}$, $0 \le t \le T$ as the initial condition. Then

$$X_{s,t} = X_{0,t}e^{-\alpha s} + e^{-\alpha s}\int_0^s e^{\alpha\sigma}\partial_\sigma W_{\sigma,t}. \tag{5}$$

Since $\int_0^s e^{\alpha\sigma}\partial_\sigma W_{\sigma,t} = \int_{R_{s,t}} e^{\alpha\sigma}\, dW_{\sigma,\tau}$ where the r.h.s. of the last equation is a surface integral on $R_{s,t} = \{(\sigma, \tau),\ 0 \le \sigma \le s,\ 0 \le \tau \le t\}$ and since this indefinite surface integral has a.s. a continuous sample version it follows that the solution to (5) is sample continuous iff $X_{0,t}$ is sample continuous in t. It follows from (1) that for any $\vartheta > 0$

$$X_{s+\vartheta,t} = X_{s,t}e^{-\alpha\vartheta} + e^{-\alpha(s+\vartheta)}\int_s^{s+\vartheta} e^{\alpha\sigma}\partial_\sigma W_{\sigma,t}.$$

Therefore $X_{s,\cdot}$ is a Markov process in the parameter s with the state space being functions in the t parameter on $[0, T]$. Let $Y(\tau) = \int_s^{s+\vartheta} e^{\alpha\sigma}\partial_\sigma W_{\sigma,\tau}$ then $Y(\tau)$ is a one parameter Brownian motion in the τ parameter with the variance $EY^2(\tau) = \tau\int_s^{s+\vartheta} e^{2\alpha\sigma}\, d\sigma = \tau e^{2\alpha s}(e^{2\alpha\vartheta} - 1)/2\alpha$. Therefore if $X_{s,\cdot}$ is a Brownian motion with a variance parameter c^2 then the pair $(X_{s,\cdot}; X_{s+\vartheta,\cdot})$ has the same probability law as the pair

$$(cw,\ e^{-\alpha\vartheta}cw + ((1/2\alpha)(1 - e^{-2\alpha\vartheta}))^{1/2}\tilde{w}). \tag{6}$$

Let $X_{0,\cdot} = cw$ and set $\alpha = (2c^2)^{-1}$ then the process $X_{s,\cdot}$ is time homogeneous in the time parameters. It follows immediately from (4) and (6) that if $F(cw)$ is in the Domain of $\mathscr{L}$ then it is in the domain of the Generator of the Markov process $X_{s,\cdot}$ and

$$\mathscr{G}F(cw) = -(\tfrac{1}{2}c^2)\mathscr{L}F(cw). \tag{7}$$

III. Relations between the different derivatives. The relation between the Malliavin derivative $\mathscr{L}$ and derivatives of $F(cw)$ with respect to the scale parameter c is given by the following result.

PROPOSITION 1. *Let $F(cw)$, $c < c_0$ be a scale invariant measurable functional. Assume that for $0 < c < c_0$ $(F(cw) - F((1-\varepsilon)cw))/\varepsilon$ converges in quadratic mean as $\varepsilon \to 0$ and denote the limit as $\partial F((\rho cw)/\partial\rho)_{\rho=1}$* or *$D^\rho F(cw)$. Also assume that for $0 < c < c_0$ $(E\{F(cw + \sqrt{\varepsilon}c\tilde{w}) \mid \mathscr{F}^w\} - F(cw))/\varepsilon$ converges in the mean as $\varepsilon \to 0$ and denote the limit as $\tilde{D}F(cw)$. Furthermore, assume that $D^\rho F(cw)$ is continuous in the mean with respect to the scale parameter c, $0 < c < c_0$. Then*

$$\mathscr{L}F(cw) = D^\rho F(cw) - 2\tilde{D}F(cw) \tag{8}$$

and for $EX_{s,t}^2 < c_0 t$

$$\mathscr{G}F(X_{s,\cdot}) = -(1/2c^2)\mathscr{L}F(X_{s,\cdot}) + ((1/2c^2) - \alpha)D^\rho(X_{s,\cdot}). \tag{9}$$

REMARK. Given $F(c, w)$ for some fixed c_1, each of the derivatives $D^\rho F(c, w)$ and $\tilde{D}F(c, w)$ depends on how $F(c, w)$ was extended to $c \ne c_1$ while the left hand side of equation (8) is independent of such an extension.

PROOF. For $c < c_0$, arbitrary c_1 and ε sufficiently small

$$
\begin{aligned}
&F((1-2\varepsilon\alpha)^{1/2}cw+\sqrt{\varepsilon}c_1\tilde{w}) \\
&\qquad = F(cw+(\varepsilon/(1-2\varepsilon\alpha))^{1/2}c_1\tilde{w}) \\
&\qquad\qquad - D^\rho F(cw+((\varepsilon/(1-2\varepsilon\alpha)))^{1/2}c_1\tilde{w})(1-(1-2\varepsilon\alpha)^{1/2})+o(\varepsilon)
\end{aligned}
\tag{10}
$$

where $o(\varepsilon)$ denotes a random variable with a second moment which is $o(\varepsilon^2)$. Therefore

$$
\begin{aligned}
&(1/\varepsilon)E\{F((1-2\varepsilon\alpha)^{1/2}cw+\sqrt{\varepsilon}c_1\tilde{w})\mid\mathscr{F}^w\}-F(cw) \\
&\qquad \underset{\varepsilon\to 0}{\overset{\text{q.m.}}{\longrightarrow}} \left(\frac{c_1}{c}\right)^2\tilde{D}F(cw)-\alpha D^\rho F(cw).
\end{aligned}
\tag{11}
$$

In the stationary case $c_1 = 1$, $c^2 = (2\alpha)^{-1}$ therefore the left hand side of (11) converges to $-\alpha\mathscr{L}F(cw)$ and (8) follows. Replacing equation (10) by

$$
\begin{aligned}
&F(e^{-\alpha\varepsilon}cw+((1/2\alpha)(1-e^{-2\alpha\varepsilon}))^{1/2}\tilde{w}) \\
&\qquad = F(cw+((1/2\alpha)(e^{2\alpha\varepsilon}-1))^{1/2}\tilde{w}) \\
&\qquad\qquad - D^\rho(cw+((1/2\alpha)(e^{2\alpha\varepsilon}-1))^{1/2}\tilde{w})(1-e^{-\alpha\varepsilon})+o(\varepsilon),
\end{aligned}
$$

it follows from (6) that

$$\mathscr{G}F(X_{s,\cdot}) = (1/c^2)\tilde{D}F(cw)-\alpha D^\rho F(cw)$$

and (9) follows by substituting $\tilde{D}$ from (8) into the last equation.

REMARK. Equation (8) can be used to yield an expression for $\tilde{D}$ for solutions of stochastic differential equations in terms of known results for $\mathscr{L}$ and an easily derivable equation for D^ρ.

Two special cases of (8) and (9) will be considered now.

Let X denote the Banach space of continuous functions on $[0, T]$ starting at zero under the sup norm. Assume that $F(c_0w)$ can be extended to all $x \in X$ and is a continuous functional on X for this extension which will be denoted $F(x)$. It is easy to show that such an extension, if it exists, is unique. Let $F(x)$ be twice continuously Fréchet differentiable on X:

$$F(x+\varepsilon y) = F(x)+\varepsilon F_x(y)+\tfrac{1}{2}\varepsilon^2F_{x,x}(y,y)+\eta(\varepsilon;x,y)$$

where, for every x in X, $F_x(y)$ is a linear form in y; $F_{x,x}(y, z)$ is a bilinear form in $y, z \in X$ and $\eta(\varepsilon; x, y)$ is $o(\varepsilon^2)$ for every $x, y \in X$. Further assume that $F_{cw}(w)$, $F_{cw}(\tilde{w})$, $F_{cw,cw}(\tilde{w}, \tilde{w})$ are square integrable and $E\eta^2(\varepsilon; w, w)$ and $E\eta^2(\varepsilon; w, \tilde{w})$ are upper bounded by $K\varepsilon^4$. In this case $E(F_{cw}(\tilde{w})) = 0$, because of the linearity of $F_w(\tilde{w})$ in $\tilde{w}$;

$$D^\rho(cw) = F_{cw}(cw) \tag{12a}$$

and

$$\tilde{D}F(cw) = \tfrac{1}{2}E\{F_{cw,cw}(c\tilde{w}, c\tilde{w})\mid\mathscr{F}^w\}.$$

Let $h^i(\cdot)$, $i = 1, 2, \cdots$ be any complete orthonormal sequence on $[0, T]$ and γ_i

independent Gaussian normalized random variables. Then $\sum \gamma_i \int_0^s h^i(\vartheta)\, d\vartheta$ converges in the mean to the standard Wiener process on $[0, T]$, further assume that h^i are such that the convergence is uniform. Then for $x \in X$ fixed and since γ_i, γ_j, $i \neq j$ are independent

$$E\{F_{xx}(w, w)\} = \lim_{N\to\infty} E \sum_1^N \gamma_i^2 F_{xx}\left(\int_0 h^i(\vartheta)\, d\vartheta, \int_0 h^i_\vartheta\, d\vartheta\right)$$

$$= \sum_1^\infty F_{xx}\left(\int_0 h^i(\vartheta)\, d\vartheta, \int_0 h^i(\vartheta)\, d\vartheta\right).$$

The last expression is known as trace $D^2F(x)$. Therefore

$$\tilde{D}F(cw) = \tfrac{1}{2} \text{ trace } D^2F(cw), \tag{12b}$$

which yields the relation $\mathscr{L}F = F_{cw}(cw) - \text{trace } D^2F(cw)$, which was used by Shigekawa [2], [7] to introduce $\mathscr{L}F$.

A second special case is the following. Let $F(c_0w)$ be a square integrable functional for some $c_0 > 0$. A natural extension of $F(c_0w)$ to $c < c_0$ is to set

$$\begin{aligned} F(cw) &= K + \sum_{n=1}^\infty \left(\frac{c}{c_0}\right)^n \int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(c_0w_{t_1}) \cdots d(c_0w_{t_n}) \\ &= K + \sum_{n=1}^\infty \int_0^T \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n)\, d(cw_{t_1}) \cdots d(cw_{t_n}) \end{aligned} \tag{13}$$

where $f_n(\cdots)$ and K are as in (1). In order to point out the difference between this special case and the previous one, consider the functional $F(w) = w^2(T) - T$. This functional can be extended as a functional on the Banach space of continuous functions on $[0, T]$ and for this extension $F_1(cw) = (cw(T))^2 - T$; on the other hand $F(w) = 2 \int_0^T \int_0^{t_2} dw_{t_2}\, dw_{t_1}$ therefore the extension defined by equation (13) yields $F_2(cw) = 2c^2 \int_0^T w_\vartheta\, dw_\vartheta$. Consequently, (since for $c \neq 1$, $F_1(cw) \neq F_2(cw)$) $F_1(cw)$ and $F_2(cw)$ are two different extensions of $F(w)$. Incidently, a case in which the extension defined by (3) yields a solution to the problem of representing a given process as "signal plus independent white noise" is given in [10]. For the extension defined by (13) we obviously have

$$\mathscr{L}F(cw) = D^\rho F(cw) \tag{14a}$$

therefore (9) and by (11)

$$\mathscr{G}F(X_{s,\cdot}) = -\alpha LF(X_{s,\cdot}) \tag{14b}$$

and

$$\tilde{D}F(X_{s,\cdot}) = 0. \tag{14c}$$

IV. The bilinear form associated with the Malliavin calculus. We turn now to the bilinear map or the "opérateur carré du champ" $\langle DF, DG \rangle$. Let $h^i(s)$ be a complete orthonormal sequence on $[0, T]$ and assume that $(F(cw + \varepsilon \int_0 h^i_\vartheta\, d\vartheta) - F(cw))/\varepsilon$ converges in quadratic mean and denote the limit by $D_{h^i}F(cw)$. Assume that $\sum_{i=1}^{N} (D_{h^i}F(cw))^2$ converges in L_1 and denote the limit by

$$\langle DF, DF \rangle = \sum_{i=1}^{\infty} D_{h^i}(F(cw))^2. \tag{15}$$

If $\langle DF, DF \rangle$ and $\langle DG, DG \rangle$ exist then $\langle DF, DG \rangle = \langle DG, DF \rangle$ is defined by polarization.

A different way to introduce $\langle DF, DF \rangle$ is the following. Let Q denote the class of process $\{u_t, t \in [0, T]\}$ such that, a.s., $u.$ is measurable and square integrable on $[0, T]$, for every t, u_t is adapted to the σ-field generated by cw (but not necessarily to the σ-field generated by $\{cw_\vartheta, 0 \le \vartheta \le t\}$), and the measure induced by $cw. + \varepsilon \int_0 u_\vartheta \, d\vartheta$ on $C_0(T)$ is equivalent to the measure induced by cw for every $0 < \varepsilon \le \varepsilon_0$. It can then be shown [11] that

$$\langle DF, DF \rangle = \sup_{u \in Q} \frac{(D_u(F(cw)))^2}{\int_0^T u_\vartheta^2 \, d\vartheta} \tag{16}$$

where $D_uF(cw) = \partial(F(cw + \varepsilon \int_0 u_\vartheta \, d\vartheta)/\partial\varepsilon)_{\varepsilon=0}$. We may therefore denote $\langle DF, DG \rangle$ as the $(gradient)^2$ operation. As introduced above, the $(gradient)^2$ operation on $F(c_0w)$ does not depend on any possible extension of $F(c_0w)$ to other values of c. In terms of the Ornstein-Uhlenbeck process this operation may be introduced as follows [5], [9]. Let C_s^F denote the continuous martingale

$$C_s^F = F(X_{s,\cdot}) - F(X_{0,\cdot}) - \int_0^s \mathscr{G}F(X_{\vartheta,\cdot}) \, d\vartheta$$

and set

$$\Gamma(F, G) = 2c_s^2(d/ds)[C^F, C^G]_s \tag{17}$$

where $[C^F, C^G]_s$ denotes the cross quadratic variation (or cross increasing) process associated with the martingales C^F and C^G and c_s^2 is defined by $c_s^2 t = EX_{s,t}^2$.

In the stationary case it is well known that $\Gamma(F, G) = \langle DF, DG \rangle$ (cf. Theorem 7.1 of [2]). It seems interesting to note that the same result holds in the nontime homogeneous case.

PROPOSITION 2. *If F, G and FG satisfy the assumptions of Proposition* 1, *then*

$$\Gamma(F(X_{s,\cdot}), G(X_{s,\cdot})) = \langle DF(X_{s,\cdot}), DG(X_{s,\cdot}) \rangle. \tag{18}$$

PROOF. $\langle DF, DG \rangle$ as defined by (15) depends only on the variance parameter c^2 (but is invariant of α) and, as is well known,

$$\langle DF, DG \rangle = \mathscr{L}(FG) - F\mathscr{L}G - G\mathscr{L}F. \tag{19}$$

On the other hand, a direct calculation as in [9] yields that as in the time homogeneous case

$$\Gamma(F, G) = 2c^2(\mathscr{G}(FG) - F\cdot\mathscr{G}G - G\cdot\mathscr{G}F). \tag{20}$$

Substituting (9) in (20), noting that $D^\rho(FG) - FD^\rho G - GD^\rho F = 0$ and comparing with (19) yields (18).

REFERENCES

[1] CAMERON, R. H. (1954). The translation pathology in Wiener space. *Duke Math. J.* **21** 623–628.

[2] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North Holland/Kodansha. Amsterdam.

[3] JOHNSON, G. W. and SKOUG, D. L. (1979). Scale invariant measurability in Wiener space. *Pacific J. Math.* **83** 157–176.

[4] KUO, H. H. (1975). Gaussian measures in Banach spaces. *Lecture Notes in Math.* **463** Springer, New York.

[5] MALLIAVIN, P. (1978). Stochastic Calculus of Variations and Hypoelliptic Operators. *Proc. of Intern. Symp. Stoch. D. Eqs., Kyoto* 1976. K. Ito, Ed. 195–263, Kinokuniya-Wiley 195–263.

[6] MEYER, P. A. (1982). Note Sur Les Processus d'Ornstein-Uhlenbeck. *Seminaire de Probabilités* XVI, 95–133. *Lecture Notes in Math.* **920** Springer, New York.

[7] SHIGEKAWA, I. (1980). Derivatives of Wiener functionals and absolute continuity of induced measures. *J. Math. Kyoto Univ.* **20** 263–289.

[8] STROOCK, D. (1981). The Malliavin calculus, a functional analytical approach. *J. Funct. Anal.* **44** 212–257.

[9] WILLIAMS, D. (1981). To begin at the beginning. *Lecture Notes in Math.* **851** 1–55. Springer, New York.

[10] WONG, E. and ZAKAI, M. (1982). A characterization of kernels associated with the multiple integral representation of some functionals of the Wiener process. *Systems and Control Letters* **2** 94–98.

[11] ZAKAI, M. (1985). The Malliavin calculus. *Acta Applicandae Math.* To Appear.

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION
HAIFA, 32000, ISRAEL