The Annals of Probability
1987, Vol. 15, No. 1, 1–39

SPECIAL INVITED PAPER

ANALYSIS OF WIENER FUNCTIONALS (MALLIAVIN CALCULUS) AND ITS APPLICATIONS TO HEAT KERNELS

BY SHINZO WATANABE

Kyoto University

An analysis of Wiener functionals is studied as a kind of Schwartz distribution theory on Wiener space. For this, we introduce, besides ordinary L_p-spaces of Wiener functionals, Sobolev-type spaces of (generalized) Wiener functionals. Any Schwartz distribution on $\mathbf{R}^d$ is pulled back to a generalized Wiener functional by a d-dimensional Wiener map which is smooth and nondegenerate in the sense of Malliavin.

As applications, we construct a heat kernel (i.e., the fundamental solution of a heat equation) by a generalized expectation of the Dirac delta function pulled back by an Itô map, i.e., a Wiener map obtained by solving Itô's stochastic differential equations. Short-time asymptotics of heat kernels are studied through the asymptotics, in terms of Sobolev norms, of the generalized Wiener functional under the expectation.

1. Introduction. The purpose of the present paper is to give an easily accessible exposition of Malliavin's calculus, infinite dimensional differential calculus on Wiener space and its applications. Since Wiener measure was introduced in 1923, many interesting and important probability models have been realized on Wiener space as *Wiener functionals*, typical examples of which are diffusion models corresponding to heat equations on manifolds. These diffusions are realized, as we know well, by solutions of Itô's stochastic differential equations. Thus Itô's calculus produces the most important class of Wiener functionals, sometimes called Itô functionals. However, these Itô functionals, as functionals of paths, are *not* in a class of functionals to which the classical calculus of variations or Fréchet differential calculus on Banach spaces can be applied. It is an important discovery of Malliavin ([15] and [16]) that, under reasonable conditions, these noncontinuous functionals can be differentiated as many times as we want when the differentiation is understood properly. Moreover, he showed that these derivatives can actually be used to produce fruitful results. Examples of such successful applications, due mainly to Malliavin, Kusuoka, Stroock and Bismut, are found, among others, in the problems of regularity, estimates and asymptotics of heat equations, cf., e.g., [5], [10], [14], [16] and [17].

In this paper, we try to give a systematic approach to the Malliavin calculus by using the notion of *generalized Wiener functionals*. In other words, we would

---

Received January 1986; revised June 1986.

AMS 1980 *subject classifications*. Primary 60H10; secondary 28C20, 35K05.

Key words and phrases. Sobolev spaces of Wiener functionals, generalized Wiener functionals, pull-back of Schwartz distributions, asymptotic expansion of Wiener functionals, heat kernels, short-time asymptotics.

treat the Malliavin calculus as a kind of Schwartz distribution theory on Wiener space. A merit of this approach will be that intuitive or heuristic expressions (often used previously in formal computations of Wiener functional expectations) can be given a precise mathematical meaning so that reasoning and computations will be clearer and more direct. For example, if we want to study the density $p(x)$ (with respect to Lebesgue measure) of the law of a d-dimensional Wiener functional $F(w)$, a formal expression for $p(x)$ given by a Wiener functional expectation would be $p(x) = E[\delta_x(F(w))]$, where δ_x is the Dirac delta function at $x \in \mathbf{R}^d$. Of course, $\delta_x(F(w))$ is not a Wiener functional in the sense of a random variable on the Wiener space. Nevertheless, under certain assumptions of regularity and nondegeneracy on the functional $F(w)$ we can realize $\delta_x(F(w))$ as an element in a Sobolev-type space of (generalized) Wiener functionals to which the expectation operator $E(\cdot)$ is naturally extended, so that the above formal expression has a correct mathematical sense. We can study various properties of $p(x)$, in particular regularity properties, by analyzing this expression. In addition, in the case when the functional F depends on a parameter ε (say $0 < \varepsilon \leq 1$) so that $F = F(\varepsilon, w)$, we can also study the asymptotics of $p(\varepsilon, x) = E[\delta_x(F(\varepsilon, x))]$ as $\varepsilon \downarrow 0$ through the asymptotics of $\delta_x(F(\varepsilon, w))$ described in terms of the above Sobolev spaces. This will be our way of studying the short-time asymptotics of heat kernels given below. This approach is actually well suited for the Itô functionals because the Itô calculus can be used quite effectively in evaluating Sobolev norms of these functionals.

Thus, the main application discussed in this paper will be the short-time asymptotics of heat kernels. Needless to say, this problem, closely related to several questions in analysis, geometry and mathematical physics, has been studied by many authors and there exists a huge amount of literature. One standard approach is, of course, through the methods of partial differential equations, cf., e.g., [4] and [18]. A probabilistic approach has been given, in, e.g., Azencott [3] and Molchanov [20], by the use of pinned diffusion processes and pinned Gaussian processes. In this approach, however, the heat kernel is used in an essential way to define a pinned diffusion and so, a knowledge of some analytical properties of the heat kernel is inevitable and the study is usually restricted only to the nondegenerate case, i.e., to the elliptic case. An application of the Malliavin calculus to this problem was first discussed by Bismut [5]. His method is based on the splitting of Wiener space and the use of the implicit function theorem. This approach by Bismut has been much refined and expanded by Kusuoka [12]: Kusuoka introduced the notion of generalized Malliavin calculus and studied the various applications of this powerful method in the asymptotic problems of Wiener functional expectations.

In this paper, we treat this problem by our method of generalized Wiener functionals explained above. Namely, we first construct the heat kernel by a generalized Wiener functional expectation in the form $p(t, x, y) = E[\delta_y(X(t, x, w))]$ and then study the asymptotic expansion of the functional $\delta_y(X(t, x, w))$ in the Sobolev spaces. This method is extremely simple in the case of the heat kernel on the diagonal $p(t, x, x)$. Also, we will obtain the asymptotic expansion of $p(t, x, y)$ off the diagonal, i.e., $x \neq y$, under the assumption (H.2) of

Bismut [5] which is weaker than the ellipticity assumption. We cannot give applications to problems in geometry and mathematical physics here. For some topics, we refer the reader to Bismut [6] and Ikeda and Watanabe [11].

The organization of this paper is as follows. In Section 2, we review the fundamentals of Malliavin calculus, particularly the Sobolev spaces formed from Wiener functionals and the differential calculus defined on them. The pull-back of Schwartz distributions under Wiener maps will be discussed in this context and this notion will play a fundamental role. In Section 3, these ideas will be applied to Itô functionals in order to obtain asymptotic results for heat kernels.

2. The Malliavin calculus in terms of generalized Wiener functionals. In this section, we develop the Malliavin calculus along the lines of [10], [24] and [27].

2.1. *Sobolev spaces of Wiener functionals.* Let (W_0^r, P) be the r-dimensional Wiener space: W_0^r is the space of all continuous paths w: $[0, \infty) \ni t \to w(t) \in R^r$, such that $w(0) = 0$ with the topology of uniform convergence on bounded intervals in $[0, \infty)$ and P is the standard Wiener measure defined on the P-completion of the Borel field over W_0^r. W_0^r (denoted simply by W when there is no confusion) is clearly a real Fréchet space. Let H be the Hilbert subspace of W consisting of $w \in W$ which are absolutely continuous with respect to t and have square-integrable derivatives. Endow H with the norm $\|w\|_H^2 = \int_0^\infty |(dw/dt)(t)|^2\, dt$. This H is often called the *Cameron–Martin subspace of W.* By the Riesz theorem we can identify the dual H' with H and thereby obtain $W' \subsetneq H' = H \subsetneq W$, where $\subsetneq$ denotes the continuous inclusion. Given a real separable Hilbert space E with the norm $|\ |_E$, the L_p-space, $1 \le p < \infty$, of E-valued Wiener functionals is denoted by $L_p(E)$ [i.e., $L_p(E)$ is a real Banach space of all P-measurable E-valued functions on W such that $\|F\|_p^p = \int_W |F(w)|_E^p P(dw) < \infty$ with the usual identification $F = G$ if and only if $F(w) = G(w)$, P-a.a. w]. If $E = \mathbf{R}$, we use L_p in place of $L_p(\mathbf{R})$.

A function F: $W \to E$ is called an *E-valued polynomial* if it is a linear combination of functions $l(w)^m e$, $m \ge 0$, $l \in W'$ and $e \in E$. F is said to be of order at most n if in the preceding, each $m \le n$. The totality of E-valued polynomials is denoted by $\mathscr{P}(E)$ and the totality of E-valued polymials of order at most n by $\mathscr{P}_n(E)$. It is well known that $\mathscr{P}(E) \subset L_p(E)$ for all p and this inclusion is dense. The real Hilbert space $L_2(E)$ is decomposed into a sum of orthogonal subspaces of Wiener's homogeneous chaos:

$$L_2(E) = C_0(E) \oplus C_1(E) \oplus \cdots \oplus C_n(E) \oplus \cdots, \tag{2.1}$$

where $C_0(E) = \{\text{constant } E\text{-valued functions}\}$ and

$$C_n(E) = \overline{\mathscr{P}_n(E)}^{L_2(E)} \cap [C_0(E) \oplus \cdots \oplus C_{n-1}(E)]^\perp, \qquad n = 1, 2, \ldots .$$

The projection onto $C_n(E)$ is denoted by J_n. Thus if $F = \sum J_n F$, $F \in L_2(E)$ and if $F \in \mathscr{P}(E)$, then this decomposition is finite sum and $J_n F \in \mathscr{P}(E)$ for all n. Given a real sequence $\phi = (\phi(n))$, $n = 0, 1, \ldots,$ the linear operator T_ϕ on $\mathscr{P}(E)$

into itself is defined by

$$T_\phi F = \sum_n \phi(n) J_n F, \qquad F \in \mathscr{P}(E). \tag{2.2}$$

If $\phi(n) = -n$, $n = 0, 1, \dots$, the corresponding operator T_ϕ, denoted by L, is called the *Ornstein–Uhlenbeck operator* or the *number operator*. Given $s \in \mathbf{R}$, we have that $(I - L)^s$: $\mathscr{P}(E) \to \mathscr{P}(E)$ is given by T_ϕ with $\phi(n) = (1 + n)^s$, $n = 0, 1, \dots$.

Now, we introduce a family of norms $\| \ \|_{p,s}$, $s \in \mathbf{R}$, $p \in (1, \infty)$, on $\mathscr{P}(E)$ by

$$\|F\|_{p,s} = \left\|(I - L)^{s/2} F\right\|_p, \qquad F \in \mathscr{P}(E), \tag{2.3}$$

where $\| \ \|_p$ is the L_p-norm of $L_p(E)$.

These norms have the following properties, cf. [24] and [27] for proofs:

(2.4) $\|F\|_{p,s} \le \|F\|_{p',s'}$ *for every* $F \in \mathscr{P}(E)$ *if* $p \le p'$ *and* $s \le s'$.

(2.5) (Compatibility). *For every* $p, p' \in (1, \infty)$ *and* $s, s' \in \mathbf{R}$, *if* $F_n \in \mathscr{P}(E)$, $n = 1, 2, \dots$, *satisfy* $\|F_n\|_{p,s} \to 0$ *as* $n \to \infty$ *and* $\|F_n - F_m\|_{p',s'} \to 0$ *as* $n, m \to \infty$, *then* $\|F_n\|_{p',s'} \to 0$ *as* $n \to \infty$.

(Duality). *For every* $p, q \in (1, \infty)$ *such that* $1/p + 1/q = 1$, $s \in \mathbf{R}$ *and* $G \in \mathscr{P}(E)$,

$$\sup \left| \int_W \langle F(w), G(w) \rangle_E P(dw) \right| = \|G\|_{q,-s}, \tag{2.6}$$

where the sup *is taken over* $F \in \mathscr{P}(E)$ *such that* $\|F\|_{p,s} \le 1$ *and* $\langle \ , \ \rangle_E$ *is the inner product of* E.

Let $\mathbf{D}_p^s(E)$ be the Banach space obtained by completing $\mathscr{P}(E)$ with respect to the norm $\| \ \|_{p,s}$ for $p \in (1, \infty)$ and $s \in \mathbf{R}$. By the above properties of the norms, we have

$$\mathbf{D}_{p'}^{s'}(E) \subsetneq \mathbf{D}_p^s(E) \quad \text{if } p \le p' \text{ and } s \le s', \tag{2.7}$$

and, under an obvious identification,

(2.8) $\mathbf{D}_p^s(E)' = \mathbf{D}_q^{-s}(E)$ *if* $p, q \in (1, \infty)$ *such that* $1/p + 1/q = 1$ *and* $s \in \mathbf{R}$.

Set

$$\mathbf{D}^\infty(E) = \bigcap_{s>0} \bigcap_{1<p<\infty} \mathbf{D}_p^s(E) \tag{2.9}$$

and

$$\mathbf{D}^{-\infty}(E) = \bigcup_{s>0} \bigcup_{1<p<\infty} \mathbf{D}_p^{-s}(E). \tag{2.10}$$

Then $\mathbf{D}^\infty(E)$ is a complete countably normed space and $\mathbf{D}^{-\infty}(E)$ is its dual. Since $\mathbf{D}_p^0(E) = L_p(E)$, elements in $\mathbf{D}_p^s(E)$ for $s \ge 0$ are Wiener functionals in the usual sense; but if $s < 0$, some elements of $\mathbf{D}_p^s(E)$ are not classical Wiener functionals and we will call them *generalized Wiener functionals*. It will be convenient to also introduce the following spaces;

$$\tilde{\mathbf{D}}^{\infty}(E) = \bigcap_{s>0} \bigcup_{1<p<\infty} \mathbf{D}_p^s(E) \tag{2.11}$$

and

$$\tilde{\mathbf{D}}^{-\infty}(E) = \bigcup_{s>0} \bigcap_{1<p<\infty} \mathbf{D}_p^{-s}(E). \tag{2.12}$$

The spaces $\mathscr{P}(E), \mathbf{D}_p^s(E), \mathbf{D}^{\infty}(E), \tilde{\mathbf{D}}^{\infty}(E), \dots,$ are denoted simply by $\mathscr{P}, \mathbf{D}_p^s, \mathbf{D}^{\infty}, \tilde{\mathbf{D}}^{\infty}, \dots,$ if $E = \mathbf{R}$.

REMARK 2.1. Sobolev spaces of Wiener functionals have been introduced by, among other, Malliavin [17], Shigekawa [22] and Kusuoka and Stroock [13]. Although their definitions, including ours, are apparently different from each other, Sugita [25] showed that they are, nonetheless, equivalent.

Given real separable Hilbert spaces E_1 and E_2, let $E_1 \otimes E_2$ be their tensor product, i.e., $E_1 \otimes E_2$ is the real Hilbert spaces consisting of all bilinear forms V on $E_1 \times E_2$ with finite Hilbert–Schmidt norm

$$\|V\|_{\mathrm{HS}} = \left\{ \sum_{i,j} V\left(h_i^{(1)}, h_j^{(2)}\right)^2 \right\}^{1/2},$$

where the summation is taken over some (= any) orthonormal bases $h_i^{(1)}$ and $h_j^{(2)}$ of E_1 and E_2, respectively. V can also be identified with a linear operator L: $E_1 \to E_2$ of the Hilbert–Schmidt type by the relation $V(h, h') = \langle Lh, h' \rangle_{E_2}$. For $x \in E_1$ and $y \in E_2$, $x \otimes y \in E_1 \otimes E_2$ is defined by $[x \otimes y](h, h') = \langle x, h \rangle_{E_1} \langle y, h' \rangle_{E_2}$. Meyer [19] obtained the following results (cf. also [24] and [27]):

(2.13) *For every $p, q \in (1, \infty)$ such that $1/p + 1/q = 1/r < 1$ and $k = 0, 1, \dots,$ there exists a positive constant $C_{p,q,k}$ such that*

$$\|F \otimes G\|_{r,k} \le C_{p,q,k} \|F\|_{p,k} \|G\|_{q,k}$$

for all $F \in \mathscr{P}(E_1)$ and $G \in \mathscr{P}(E_2)$.

From this, the map $(F, G) \in \mathscr{P}(E_1) \times \mathscr{P}(E_2) \mapsto F \otimes G \in \mathscr{P}(E_1 \otimes E_2)$ is extended to a continuous map $\mathbf{D}_p^k(E_1) \times \mathbf{D}_q^k(E_2) \mapsto \mathbf{D}_r^k(E_1 \times E_2)$. In particular, $\mathbf{D}^{\infty}$ is an algebra: If $F, G \in \mathbf{D}^{\infty}$, then $F \cdot G \in \mathbf{D}^{\infty}$. More generally, if $F \in \mathbf{D}^{\infty}$ and $G \in \mathbf{D}^{\infty}(E)$, then $F \cdot G \in \mathbf{D}^{\infty}(E)$. Also, if $F \in \mathbf{D}^{\infty}$ and $G \in \tilde{\mathbf{D}}^{\infty}(E)$ or $F \in \tilde{\mathbf{D}}^{\infty}$ and $G \in \mathbf{D}^{\infty}(E)$, then $F \cdot G \in \tilde{\mathbf{D}}^{\infty}(E)$. From (2.13) and (2.6), we can easily deduce the following:

(2.14) *For every $p, q \in (1, \infty)$ such that $1/p + 1/q = 1/r < 1$ and $k = 0, 1, \dots,$ there exists a positive constant $C'_{p,q,k}$ such that*

$$\|F \cdot G\|_{r,-k} \le C'_{p,q,k} \|F\|_{p,k} \|G\|_{q,-k}$$

for all $F, G \in \mathscr{P}$.

Hence $F \cdot \Phi \in \mathbf{D}_r^{-k}$ is well defined if $F \in \mathbf{D}_p^k$ and $\Phi \in \mathbf{D}_q^{-k}$. Thus, $F \cdot \Phi \in \mathbf{D}^{-\infty}$ is well defined if $F \in \mathbf{D}^{\infty}$ and $\Phi \in \mathbf{D}^{-\infty}$ and it is obvious that $F \cdot \Phi$ is the unique element in $\mathbf{D}^{-\infty}$ such that ${}_{\mathbf{D}^{-\infty}}\langle F\cdot\Phi, G\rangle_{\mathbf{D}^{\infty}} = {}_{\mathbf{D}^{-\infty}}\langle \Phi, F\cdot G\rangle_{\mathbf{D}^{\infty}}$ for every $G\in\mathbf{D}^{\infty}$. Similarly $F\cdot\Phi\in\mathbf{D}^{-\infty}$ is defined if $F\in\tilde{\mathbf{D}}^{\infty}$ and $\Phi\in\tilde{\mathbf{D}}^{-\infty}$, and furthermore, $F\cdot\Phi\in\tilde{\mathbf{D}}^{-\infty}$ if $F\in\mathbf{D}^{\infty}$ and $\Phi\in\tilde{\mathbf{D}}^{-\infty}$. More generally, $F\cdot\Phi\in\mathbf{D}^{-\infty}(E)$ is well defined if $F\in\mathbf{D}^{\infty}$ ($\tilde{\mathbf{D}}^{\infty}$) and $\Phi\in\mathbf{D}^{-\infty}(E)$ [resp. $\tilde{\mathbf{D}}^{-\infty}(E)$] or $F\in\mathbf{D}^{\infty}(E)$ [$\tilde{\mathbf{D}}^{\infty}(E)$] and $\Phi\in\mathbf{D}^{-\infty}$ [resp. $\tilde{\mathbf{D}}^{-\infty}$]. Furthermore, $F\cdot\Phi\in\tilde{\mathbf{D}}^{-\infty}(E)$ if $F\in\mathbf{D}^{\infty}$ and $\Phi\in\tilde{\mathbf{D}}^{-\infty}(E)$ or $F\in\mathbf{D}^{\infty}(E)$ and $\Phi\in\tilde{\mathbf{D}}^{-\infty}$.

We define the H-derivative D: $\mathscr{P}(E)\mapsto\mathscr{P}(H\otimes E)$ by

$$DF(w)[h,e] = \frac{d}{dt}\langle F(w+th), e\rangle_E\Big|_{t=0}, \qquad h\in H,\quad e\in E,\quad F(w)\in\mathscr{P}(E)$$

and

$$D^k\colon \mathscr{P}(E)\mapsto\mathscr{P}\Big(\overbrace{H\otimes H\otimes\cdots\otimes H}^{k}\otimes E\Big)$$

successively by $D^k = D(D^{k-1})$. The following important result is due to Meyer [19] (cf. [24] and [27]): D: $\mathscr{P}(E)\mapsto\mathscr{P}(H\otimes E)$ is uniquely extended to a linear operator $\mathbf{D}^{-\infty}(E)\mapsto\mathbf{D}^{-\infty}(H\otimes E)$ which is continuous in the sense that its restriction: $\mathbf{D}_p^{s+1}(E)\mapsto\mathbf{D}_p^{s}(H\otimes E)$ is continuous for every $p\in(1,\infty)$ and $s\in\mathbf{R}$. The dual operator D^* of D is defined as a linear operator D^*: $\mathbf{D}^{-\infty}(H\otimes E)\mapsto\mathbf{D}^{-\infty}(E)$ which is continuous in the sense that its restriction: $\mathbf{D}_p^{s+1}(H\otimes E)\mapsto\mathbf{D}_p^{s}(E)$ is continuous for every $p\in(1,\infty)$ and $s\in\mathbf{R}$. The operator L is extended uniquely to a linear operator $\mathbf{D}^{-\infty}(E)\mapsto\mathbf{D}^{-\infty}(E)$, continuous in the sense that its restriction: $\mathbf{D}_p^{s+2}(E)\mapsto\mathbf{D}_p^{s}(E)$ is continuous for every $p\in(1,\infty)$ and $s\in\mathbf{R}$ and it holds that

$$L = -D^*D. \tag{2.15}$$

This expression for L shows that it is a second-order differential operator. For chain rules of D, D^* and L, cf. [10], [24], and [27].

2.2. *Pull-back of Schwartz distributions under a Wiener map and the smoothness of conditional expectations.* Let F: $W\to\mathbf{R}^d$ be a d-dimensional Wiener functional. It is said to be *smooth* (in the sense of Malliavin) if $F\in\mathbf{D}^{\infty}(\mathbf{R}^d)$, i.e., $F=(F^1,F^2,\dots,F^d)$ with $F^i\in\mathbf{D}^{\infty}$. In this case

$$\sigma^{ij}(w) = \big\langle DF^i(w), DF^j(w)\big\rangle_H \in\mathbf{D}^{\infty}, \qquad i,j=1,2,\dots,d. \tag{2.16}$$

Note that $DF^i\in\mathbf{D}^{\infty}(H\otimes\mathbf{R})$ and we identify $H\otimes\mathbf{R}=H'$ with H as before. The Wiener functional $\sigma=(\sigma^{ij})$ with values in nonnegative definite symmetric $d\times d$-matrices is called the *Malliavin covariance of F. F* is said to be *nondegenerate* (in the sense of Malliavin) if

(2.17) *For almost all* $w(P)$, $\det\sigma(w)>0$ *and* $[\det\sigma(w)]^{-1}\in\bigcap_{1<p<\infty}L_p$.

In this case, $\gamma=(\gamma^{ij})=\sigma^{-1}$ satisfies $\gamma^{ij}\in\mathbf{D}^{\infty}$, $i,j=1,2,\dots,d$.

Suppose that we are given $F\in\mathbf{D}^{\infty}(\mathbf{R}^d)$ satisfying the nondegeneracy condition (2.17). Then we can show that every Schwartz distribution $T(x)$ on $\mathbf{R}^d$ can be lifted or pulled-back to a generalized Wiener functional $T\circ F$ [denoted also by $T(F)$] in $\mathbf{D}^{-\infty}$ under the Wiener map: $w \in W \to F(w) \in \mathbf{R}^d$. To discuss such notions, we introduce the following family of real Banach spaces of functions and generalized functions on $\mathbf{R}^d$. Let $\mathscr{S}(\mathbf{R}^d)$ be the real Schwartz space of rapidly decreasing C^∞-functions on $\mathbf{R}^d$ and set

$$\|\phi\|_{2k} = \left\|(1 + |x|^2 - \Delta)^k \phi\right\|_\infty, \qquad \phi \in \mathscr{S}(\mathbf{R}^d), \quad k = 0, \pm 1, \pm 2, \dots,$$

where $\| \ \|_\infty$ is the supremum norm and $\Delta = \sum_{i=1}^d (\partial/\partial x^i)^2$. Let $\mathscr{T}_{2k}$ be the completion of $\mathscr{S}(\mathbf{R}^d)$ by the norm $\| \ \|_{2k}$. Then we have

$$\mathscr{S}(\mathbf{R}^d) \subset \cdots \subsetneq \mathscr{T}_2 \subsetneq \mathscr{T}_0 = \hat{C}(\mathbf{R}^d) \subsetneq \mathscr{T}_{-2} \subsetneq \cdots \subset \mathscr{S}'(\mathbf{R}^d),$$

where $\hat{C}(\mathbf{R}^d)$ is the Banach space of all real continuous functions on $\mathbf{R}^d$ tending to 0 at infinity endowed with the supremum norm and $\mathscr{S}'(\mathbf{R}^d)$ is the Schwartz space of real tempered distributions on $\mathbf{R}^d$. Furthermore, $\mathscr{S}(\mathbf{R}^d) = \bigcap_{k=1}^\infty \mathscr{T}_{2k}$ and $\mathscr{S}'(\mathbf{R}^d) = \bigcup_{k=1}^\infty \mathscr{T}_{-2k}$.

THEOREM 2.1 ([10] and [27]). *Let $F \in \mathbf{D}^\infty(\mathbf{R}^d)$ be given and satisfy the nondegeneracy condition* (2.17). *Then for every $p \in (1, \infty)$ and $k = 0, 1, 2, \dots,$ there exists a positive constant $C = C_{p,k}$ such that*

$$\|\phi \circ F\|_{p,-2k} \le C\|\phi\|_{-2k} \quad \text{for all } \phi \in \mathscr{S}(\mathbf{R}^d). \tag{2.18}$$

[*Note that $\phi \circ F \in \mathbf{D}^\infty$.*] *Hence the map $\phi \in \mathscr{S}(\mathbf{R}^d) \mapsto \phi \circ F \in \mathbf{D}^\infty$ can be extended uniquely to a linear map*

$$T \in \mathscr{S}'(\mathbf{R}^d) \mapsto T \circ F \in \mathbf{D}^{-\infty},$$

such that its restriction $T \in \mathscr{T}_{-2k} \mapsto T \circ F \in \mathbf{D}_p^{-2k}$ is continuous for every $p \in (1, \infty)$ and $k = 0, 1, 2, \dots$. In particular, $T \circ F \in \tilde{\mathbf{D}}^{-\infty} = \bigcup_{k=1}^\infty \bigcap_{1<p<\infty} \mathbf{D}_p^{-k}$ for every $T \in S'(\mathbf{R}^d)$.

The proof of this theorem is based on integration by parts on the Wiener space. A typical formula is as follows (cf. [27], page 55 or [10], page 18):

$$\int_W \frac{\partial \phi}{\partial x^i}(F(w))G(w)P(dw) = \int_W \phi(F(w))\Phi_i(w; G)P(dw), \tag{2.19}$$

where

$$\begin{aligned}\Phi_i(w; G) = -\sum_{j=1}^d \Bigg\{ &-\sum_{k,l=1}^d G(w)\gamma^{ik}(w)\gamma^{jl}(w)\left\langle D\sigma^{kl}(w), DF^j(w)\right\rangle_H \\ &+\gamma^{ij}(w)\left\langle DG(w), DF^j(w)\right\rangle_H + \gamma^{ij}(w)G(w)LF^j(w)\Bigg\}\end{aligned} \tag{2.20}$$

if $\phi \in \mathscr{S}(\mathbf{R}^d)$ and $G \in \mathbf{D}^\infty$.

$T \circ F$, denoted also by $T(F)$, is called the *composition of $T \in \mathscr{S}'(\mathbf{R}^d)$ and F, or the lifting or pull-back of $T \in \mathscr{S}'(\mathbf{R}^d)$* under the Wiener map F: $W \mapsto \mathbf{R}^d$. Since $T \circ F \in \tilde{\mathbf{D}}^{-\infty}$, it can act on any test functional in $\tilde{\mathbf{D}}^\infty$, which is much larger than $\mathbf{D}^\infty$. It is now easy to extend the formulas (2.19) and (2.20) to the case

$\phi \in \mathscr{S}'(\mathbf{R}^d)$ and $G \in \tilde{\mathbf{D}}^{\infty}$; only now, integrals are replaced by the coupling between $\tilde{\mathbf{D}}^{-\infty}$ and $\tilde{\mathbf{D}}^{\infty}$.

REMARK 2.2. If $F(w) = F_\alpha(w)$ depends on a parameter $\alpha \in I$, I being an interval in $\mathbf{R}^n$, and if $\sigma^{ij}(\alpha) = \langle DF^i, DF^j \rangle_H$ satisfies that $\{\det \sigma(\alpha)^{-1}, \alpha \in I\}$ is bounded in L_p for every $p \in (1, \infty)$, then it is easy to deduce that the continuity (differentiability) of $\alpha \in I \mapsto F_\alpha \in \mathbf{D}^{\infty}(\mathbf{R}^d)$ implies the continuity (differentiability) of $\alpha \in I \mapsto T(F_\alpha) \in \tilde{\mathbf{D}}^{-\infty}$ in the following sense: $\alpha \mapsto T(F_\alpha) \in \tilde{\mathbf{D}}^{-\infty}$ is continuous at $\alpha = \alpha_0$ if $\delta > 0$ and $s > 0$ exist such that $T(F_\alpha) \in \bigcap_{1<p<\infty} \mathbf{D}_p^{-s}$ for $|\alpha - \alpha_0| < \delta$ and if $T(F_\alpha) \to T(F_{\alpha_0})$ in $\mathbf{D}_p^{-s}$ as $\alpha \to \alpha_0$ for all $p \in (1, \infty)$. It is easy to see, in the differentiable case, that the following holds:

$$\frac{\partial}{\partial \alpha}\{T(F_\alpha)\} = \sum_{i=1}^{d} \frac{\partial T}{\partial x^i}(F_\alpha)\frac{\partial F^i}{\partial \alpha}.$$

Using this notion of the pull-back, the smoothness of the probability laws of Wiener functionals can be discussed as follows. Suppose that $F\colon W \mapsto \mathbf{R}^d$ satisfy the same assumptions as in Theorem 2.1. Noting that δ_x, the Dirac delta function at $x \in \mathbf{R}^d$, is in $\mathscr{T}_{-2m}$ for $m \geq m_0$, where $m_0 = [d/2] + 1$, we see at once from Theorem 2.1 that, for $k = 0, 1, \dots,$ $\mathbf{R}^d \ni x \mapsto \delta_x(F) \in \mathbf{D}_p^{-2m_0-2k}$ is continuously differentiable $2k$-times. Hence, for every $G \in \bigcup_{1<p<\infty} \mathbf{D}_p^{2m_0+2k}$, $R^d \ni x \to \langle \delta_x(F), G \rangle = E[\delta_x(F) \cdot G]$ is C^{2k} and therefore it is C^{∞} if $G \in \tilde{\mathbf{D}}^{\infty}$. In particular, $p_F(x) = \langle \delta_x(F), 1 \rangle = E[\delta_x(F)]$ is a C^{∞}-function on $\mathbf{R}^d$. [Hereafter, we use the following notational convention: The coupling $\langle \Phi, F \rangle = \langle F \cdot \Phi, 1 \rangle$ of, say, $\Phi \in \mathbf{D}_p^{-s}$ and $F \in \mathbf{D}_q^{s}$, where $s > 0$ and $1/p + 1/q \leq 1$, or $\Phi \in \mathbf{D}^{-\infty}$ and $F \in \mathbf{D}^{\infty}$ or $\Phi \in \tilde{\mathbf{D}}^{-\infty}$ and $F \in \tilde{\mathbf{D}}^{\infty}$ will be denoted by $E[\Phi \cdot F] = E[F \cdot \Phi]$. In particular, we denote $\langle \Phi, 1 \rangle$ for $\Phi \in \mathbf{D}^{-\infty}$ by $E(\Phi)$. This notation is compatible with the usual one if Φ or $F \cdot \Phi$ is an integrable random variable.]

Clearly, $p_F(x)$ is the density of probability law of F with respect to the Lebesgue measure on $\mathbf{R}^d$, as is easily seen from

$$\begin{aligned}\int_{\mathbf{R}^d} p_F(x)\phi(x)\,dx &= \int_{\mathbf{R}^d} \phi(x)\langle \delta_x(F), 1 \rangle\,dx \\ &= \left\langle \int_{\mathbf{R}^d} \phi(x)\delta_x(F)\,dx, 1 \right\rangle \\ &= \langle \phi \circ F, 1 \rangle = E[\phi \circ F], \qquad \phi \in \mathscr{S}(\mathbf{R}^d).\end{aligned}$$

In this way we have deduced that the law of F has a C^{∞}-density if $F \in \mathbf{D}^{\infty}(\mathbf{R}^d)$ satisfies the nondegeneracy condition (2.7). Also, it is easy to see that $E[\delta_x(F) \cdot G]$ is a version of $E[G|F = x]p_F(x)$. Thus on a set where p_F is positive, the conditional expectation of $G \in \tilde{\mathbf{D}}^{\infty}$ given F has a smooth version.

REMARK 2.3. In the case $F = F_\alpha$ depending on a parameter as in Remark 2.2, the continuity and differentiability of $p_{F_\alpha}(x)$ in α can be obtained from Remark 2.2.

2.3. *Asymptotic expansions of Wiener functionals.* We consider a family $F(\varepsilon, w)$ of Wiener functionals depending on a parameter $\varepsilon \in (0,1]$. We can speak of its asymptotics as $\varepsilon \downarrow 0$ in terms of Sobolev spaces, e.g., we say $F(\varepsilon, w) = O(\varepsilon^k)$ as $\varepsilon \downarrow 0$ in $\mathbf{D}_p^s$ if $F(\varepsilon, w) \in \mathbf{D}_p^s$ for all $\varepsilon \in (0,1]$ and

$$\limsup_{\varepsilon \downarrow 0} \frac{\|F(\varepsilon, w)\|_{s,p}}{\varepsilon^k} < \infty,$$

where k is some real constant. Based on this notion, we give the following definitions.

Let $F(\varepsilon, w)$, $\varepsilon \in (0,1]$, be a family of elements in $\mathbf{D}^\infty(E)$ and $f_0, f_1, \ldots \in \mathbf{D}^\infty(E)$. We say that $F(\varepsilon, w)$ has the asymptotic expansion

$$(2.21) \qquad F(\varepsilon, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \text{in } \mathbf{D}^\infty(E) \text{ as } \varepsilon \downarrow 0$$

if, for every $p \in (1, \infty)$, $s > 0$ and $k = 1, 2, \ldots$,

$$F(\varepsilon, w) - \left(f_0 + \varepsilon f_1 + \cdots + \varepsilon^{k-1} f_{k-1} \right) = O(\varepsilon^k) \quad \text{in } \mathbf{D}_p^s(E) \text{ as } \varepsilon \downarrow 0.$$

In the case of $F(\varepsilon, w) \in \tilde{\mathbf{D}}^\infty(E)$, $\varepsilon \in (0,1]$, and $f_0, f_1, \ldots \in \tilde{\mathbf{D}}^\infty(E)$, we say that $F(\varepsilon, w)$ has the asymptotic expansion

$$(2.22) \qquad F(\varepsilon, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \text{in } \tilde{\mathbf{D}}^\infty(E) \text{ as } \varepsilon \downarrow 0$$

if, for every $k = 1, 2, \ldots$ and $s > 0$, we can find $p = p_{k,s} \in (1, \infty)$ such that $F(\varepsilon, w) \in \mathbf{D}_p^s(E)$ for all $\varepsilon \in (0,1]$, $f_0, f_1, \ldots, f_{k-1} \in \mathbf{D}_p^s(E)$ and

$$F(\varepsilon, w) - \left(f_0 + \varepsilon f_1 + \cdots + \varepsilon^{k-1} f_{k-1} \right) = O(\varepsilon^k) \quad \text{in } \mathbf{D}_p^s(E) \text{ as } \varepsilon \downarrow 0.$$

Similarly, if $\Phi(\varepsilon, w) \in \mathbf{D}^{-\infty}(E)$, $\varepsilon \in (0,1]$ and $\Phi_0, \Phi_1, \ldots \in \mathbf{D}^{-\infty}(E)$, we say that $\Phi(\varepsilon, w)$ has the asymptotic expansion

$$(2.23) \qquad \Phi(\varepsilon, w) \sim \Phi_0 + \varepsilon \Phi_1 + \cdots \quad \text{in } \mathbf{D}^{-\infty}(E) \text{ as } \varepsilon \downarrow 0$$

if, for every $k = 1, 2, \ldots$, we can find $p = p_k \in (1, \infty)$ and $s = s_k > 0$ such that $\Phi(\varepsilon, w) \in \mathbf{D}_p^{-s}(E)$ for all $\varepsilon \in (0,1]$, $\Phi_0, \ldots, \Phi_{k-1} \in \mathbf{D}_p^{-s}(E)$ and

$$\Phi(\varepsilon, w) - \left(\Phi_0 + \varepsilon \Phi_1 + \cdots + \varepsilon^{k-1} \Phi_{k-1} \right) = O(\varepsilon^k) \quad \text{in } \mathbf{D}_p^{-s}(E) \text{ as } \varepsilon \downarrow 0.$$

Finally, in the case of $\Phi(\varepsilon, w) \in \tilde{\mathbf{D}}^{-\infty}(E)$, $\varepsilon \in (0,1]$ and $\Phi_0, \Phi_1, \ldots \in \tilde{\mathbf{D}}^{-\infty}(E)$, we say that $\Phi(\varepsilon, w)$ has the asymptotic expansion

$$(2.24) \qquad \Phi(\varepsilon, w) \sim \Phi_0 + \varepsilon \Phi_1 + \cdots \quad \text{in } \tilde{\mathbf{D}}^{-\infty}(E) \text{ as } \varepsilon \downarrow 0$$

if, for every $k = 1, 2, \ldots$, we can find $s = s_k > 0$ such that $\Phi(\varepsilon, w) \in \bigcap_{1 < p < \infty} \mathbf{D}_p^{-s}(E)$ for all $\varepsilon \in (0,1]$, $\Phi_0, \ldots, \Phi_{k-1} \in \bigcap_{1 < p < \infty} \mathbf{D}_p^{-s}(E)$ and, for all $p \in (1, \infty)$,

$$\Phi(\varepsilon, w) - \left(\Phi_0 + \varepsilon \Phi_1 + \cdots + \varepsilon^{k-1} \Phi_{k-1} \right) = O(\varepsilon^k) \quad \text{in } \mathbf{D}_p^{-s}(E) \text{ as } \varepsilon \downarrow 0.$$

If (2.23) holds, we obviously have

$$(2.25) \qquad E(\Phi(\varepsilon, w)) \sim E(\Phi_0) + \varepsilon E(\Phi_1) + \cdots \quad \text{as } \varepsilon \downarrow 0$$

in the ordinary numerical sense, i.e., for every $k = 1, 2, \ldots$,

$$E(\Phi(\varepsilon, w)) - \left(E(\Phi_0) + \varepsilon E(\Phi_1) + \cdots + \varepsilon^{k-1} E(\Phi_{k-1}) \right) = O(\varepsilon^k), \quad \text{as } \varepsilon \downarrow 0.$$

The following theorem is easily proved if one uses the continuity of multiplications in Sobolev spaces as expressed in the form of inequalities (2.13) and (2.14).

THEOREM 2.2. (i) *If* $G(\varepsilon, w) \in \mathbf{D}^\infty$ *and* $F(\varepsilon, w) \in \mathbf{D}^\infty(E)\,[\tilde{\mathbf{D}}^\infty(E)]$, $\varepsilon \in (0,1]$, *have the asymptotic expansions*

(2.26) $$G(\varepsilon, w) \sim g_0 + \varepsilon g_1 + \cdots \quad \textit{in } \mathbf{D}^\infty \textit{ as } \varepsilon \downarrow 0,$$

with $g_i \in \mathbf{D}^\infty$ *and*

(2.27) $$F(\varepsilon, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \textit{in } \mathbf{D}^\infty(E)\ [\textit{resp. } \tilde{\mathbf{D}}^\infty(E)] \textit{ as } \varepsilon \downarrow 0,$$

with $f_i \in \mathbf{D}^\infty(E)\ [\textit{resp. } \tilde{\mathbf{D}}^\infty(E)]$, *then* $G(\varepsilon, w)F(\varepsilon, w)$ *has the asymptotic expansion*

(2.28) $$G(\varepsilon, w)F(\varepsilon, w) \sim h_0 + \varepsilon h_1 + \cdots \quad \textit{in } \mathbf{D}^\infty(E)\ [\textit{resp. } \tilde{\mathbf{D}}^\infty(E)] \textit{ as } \varepsilon \downarrow 0$$

and $h_i \in \mathbf{D}^\infty(E)\ [\textit{resp. } \tilde{\mathbf{D}}^\infty(E)]$ *are determined by the formal multiplication*:

(2.29) $$h_0 = g_0 f_0, \qquad h_1 = g_0 f_1 + g_1 f_0, \qquad h_2 = g_0 f_2 + g_1 f_1 + g_2 f_0, \ldots .$$

(ii) *If* $G(\varepsilon, w) \in \mathbf{D}^\infty\ (\tilde{\mathbf{D}}^\infty)$ *and* $\Phi(\varepsilon, w) \in \mathbf{D}^{-\infty}(E)\,[\textit{resp. } \tilde{\mathbf{D}}^{-\infty}(E)]$, $\varepsilon \in (0,1]$, *have the asymptotic expansions*

(2.30) $$G(\varepsilon, w) \sim g_0 + \varepsilon g_1 + \cdots \quad \textit{in } \mathbf{D}^\infty\ [\textit{resp. } \tilde{\mathbf{D}}^\infty] \textit{ as } \varepsilon \downarrow 0,$$

with $g_i \in \mathbf{D}^\infty$ (*resp.* $\tilde{\mathbf{D}}^\infty$) *and*

(2.31) $$\Phi(\varepsilon, w) \sim \Phi_0 + \varepsilon \Phi_1 + \cdots \quad \textit{in } \mathbf{D}^{-\infty}(E)\ [\textit{resp. } \tilde{\mathbf{D}}^{-\infty}(E)] \textit{ as } \varepsilon \downarrow 0,$$

with $\Phi_i \in \mathbf{D}^{-\infty}(E)\ [\textit{resp. } \tilde{\mathbf{D}}^{-\infty}(E)]$, *then* $G(\varepsilon, w)\Phi(\varepsilon, w)$ *has the asymptotic expansion*

(2.32) $$G(\varepsilon, w)\Phi(\varepsilon, w) \sim \Psi_0 + \varepsilon \Psi_1 + \cdots \quad \textit{in } \mathbf{D}^{-\infty}(E) \textit{ as } \varepsilon \downarrow 0,$$

and $\Psi_i \in \mathbf{D}^{-\infty}(E)$ *are obtained by the formal multiplication*:

(2.33) $$\begin{aligned} &\Psi_0 = g_0 \cdot \Phi_0, \qquad \Psi_1 = g_0 \cdot \Phi_1 + g_1 \cdot \Phi_0, \\ &\Psi_2 = g_0 \cdot \Phi_2 + g_1 \cdot \Phi_1 + g_2 \cdot \Phi_0, \ldots . \end{aligned}$$

Suppose we are given a family $F(\varepsilon, w)$, $\varepsilon \in (0,1]$, of elements in $\mathbf{D}^\infty(\mathbf{R}^d)$. This family is said to be *uniformly nondegenerate* if for each $\varepsilon \in (0,1]$, $F(\varepsilon, w)$ is nondegenerate in the sense of (2.17) and furthermore,

(2.34) $$\limsup_{\varepsilon \downarrow 0} \left\| [\det \sigma(\varepsilon)]^{-1} \right\|_p < \infty \quad \text{for all } p \in (1, \infty),$$

where $\sigma(\varepsilon) = \sigma(\varepsilon, w)$ is the Malliavin covariance of $F(\varepsilon, w)$. If, furthermore, $F(\varepsilon, w)$ has the asymptotic expansion

(2.35) $$F(\varepsilon, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \text{in } \mathbf{D}^\infty(\mathbf{R}^d) \text{ as } \varepsilon \downarrow 0,$$

with $f_i \in \mathbf{D}^\infty$, then, setting $F(0, w) = f_0$, $F(0, w)$ is also nondegenerate with $\|[\det \sigma(0)]^{-1}\|_p = \lim_{\varepsilon \downarrow 0} \|[\det \sigma(\varepsilon)]^{-1}\|_p$.

Finally, if we denote $\sigma(\varepsilon)^{-1} = \gamma(\varepsilon) = (\gamma^{ij}(\varepsilon))$, then $\gamma^{ij}(\varepsilon) \in \mathbf{D}^\infty$, $\varepsilon \in [0,1]$ and it has the asymptotic expansion

$$\gamma^{ij}(\varepsilon) \sim \gamma^{ij}(0) + \varepsilon\gamma_1^{ij} + \varepsilon^2\gamma_2^{ij} + \cdots \quad \text{in } \mathbf{D}^\infty \text{ as } \varepsilon \downarrow 0,$$

with $\gamma_k^{ij} \in \mathbf{D}^\infty$, $i, j = 1,2,\dots,d$, $k = 1,2,\dots$.

THEOREM 2.3. *Let $F(\varepsilon, w) \in \mathbf{D}^\infty(\mathbf{R}^d)$, $\varepsilon \in (0,1]$, satisfy the above assumptions, i.e., it is uniformly nondegenerate and has the asymptotic expansion* (2.35). *Then, for every $T \in \mathscr{S}'(\mathbf{R}^d)$, $T(F(\varepsilon, w)) \in \tilde{\mathbf{D}}^{-\infty}$ (defined for $\varepsilon \in [0,1]$ by Theorem* 2.1) *has the asymptotic expansion in $\tilde{\mathbf{D}}^{-\infty}$ (and a fortiori in $\mathbf{D}^{-\infty}$):*

$$T(F(\varepsilon, w)) \sim \Phi_0 + \varepsilon\Phi_1 + \cdots \quad \textit{in } \tilde{\mathbf{D}}^{-\infty} \textit{ as } \varepsilon \downarrow 0, \tag{2.36}$$

and $\Phi_0, \Phi_1, \ldots \in \tilde{\mathbf{D}}^{-\infty}$ are determined by the formal Taylor expansion

$$\begin{aligned} T\big(f_0 + \left[\varepsilon f_1 + \varepsilon^2 f_2 + \cdots \right]\big) &= \sum_{\mathbf{n}} \frac{1}{\mathbf{n}!} D^{\mathbf{n}} T(f_0)\left[\varepsilon f_1 + \varepsilon^2 f_2 + \cdots\right]^{\mathbf{n}} \\ &= \Phi_0 + \varepsilon\Phi_1 + \cdots, \end{aligned} \tag{2.37}$$

where $\mathbf{n} = (n_1, \ldots, n_d)$ is a multi-index, $\mathbf{n}! = n_1! \cdots n_d!$, $a^{\mathbf{n}} = a_1^{n_1} \cdots a_d^{n_d}$ for $a \in \mathbf{R}^d$ and $D^{\mathbf{n}} = (\partial/\partial x^1)^{n_1} \cdots (\partial/\partial x^d)^{n_d}$. In particular, denoting $\partial^i = \partial/\partial x^i$,

$$\Phi_0 = T(f_0), \qquad \Phi_1 = \sum_{i=1}^d f_1^i \partial^i T(f_0),$$

$$\begin{aligned} \Phi_2 &= \sum_{i=1}^d f_2^i \partial^i T(f_0) + \frac{1}{2!}\sum_{i,j=1}^d f_1^i f_1^j \partial^i \partial^j T(f_0), \\ \Phi_3 &= \sum_{i=1}^d f_3^i \partial^i T(f_0) + \frac{2}{2!}\sum_{i,j=1}^d f_1^i f_2^j \partial^i \partial^j T(f_0) \\ &\quad + \frac{1}{3!}\sum_{i,j,k=1}^d f_1^i f_1^j f_1^k \partial^i \partial^j \partial^k T(f_0), \ldots. \end{aligned} \tag{2.38}$$

PROOF. First, we show that, for every given positive integer k, we can find $s \in \mathbf{R}$ and $\Phi_0, \Phi_1, \ldots, \Phi_{k-1} \in \bigcap_{1<p<\infty} \mathbf{D}_p^s$ such that $T(F(\varepsilon, w)) \in \bigcap_{1<p<\infty} \mathbf{D}_p^s$ for all $\varepsilon \in [0,1]$ [note that $F(0, w) = f_0$] and

$$T(F(\varepsilon, w)) = \Phi_0 + \varepsilon\Phi_1 + \cdots + \varepsilon^{k-1}\Phi_{k-1} + O(\varepsilon^k) \quad \text{in } \mathbf{D}_p^s \text{ as} \downarrow 0 \tag{2.39}$$

for all $1 < p < \infty$. For this we note that we can find a positive integer m and a bounded function $\phi(x)$ on $\mathbf{R}^d$ which is k-times continuously differentiable with bounded derivatives up to kth order such that $T = (1 + |x|^2 - \Delta)^m \phi$. Then, for every $J \in \mathbf{D}^\infty$, we have, by integration by parts [cf. (2.19) and (2.20)], that

$$\langle T(F(\varepsilon, w)), J\rangle = E\left[\phi(F(\varepsilon, w)) l_\varepsilon(J)\right],$$

where $l_\varepsilon(J) \in D^\infty$ is of the form

$$l_\varepsilon(J) = \sum_{i=0}^{2m} \langle P_i(\varepsilon, w), D^i J\rangle_{\overbrace{H\otimes\cdots\otimes H}^{i}},$$

with $P_i(\varepsilon, w) \in \mathbf{D}^\infty(\overbrace{H \otimes \cdots \otimes H}^{i})$, which are polynomials in $F(\varepsilon, w)$, its derivatives and $\gamma(\varepsilon) = \sigma(\varepsilon)^{-1}$. By the assumption on ϕ, we have

$$\phi(F(\varepsilon, w)) = \sum_{|n|\le k-1} \frac{1}{n!} D^n\phi(f_0)[F(\varepsilon, w) - f_0]^n - V_k(\varepsilon, w)$$

and, for every $p' \in (1, \infty)$, we can find $c_1 > 0$ such that

$$\|V_k(\varepsilon, w)\|_{p'} \le c_1 \varepsilon^k \quad \text{for all } \varepsilon \in [0,1].$$

Let $q' \in (1, \infty)$ such that $1/p' + 1/q' = 1$ and choose q such that $q > q' > 1$. Then $c_2 > 0$ exists such that

$$\|l_\varepsilon(J)\|_{q'} \le c_2\|J\|_{q,2m} \quad \text{for all } \varepsilon \in [0,1] \text{ and } J \in \mathbf{D}^\infty.$$

Hence,

$$\begin{aligned}|E(V_k(\varepsilon, w) l_\varepsilon(J))| &\le \|V_k(\varepsilon, w)\|_{p'} \|l_\varepsilon(J)\|_{q'} \\ &\le c_1 c_1 \|J\|_{q,2m} \varepsilon^k\end{aligned}$$

for all $\varepsilon \in [0,1]$ and $J \in \mathbf{D}^\infty$. Also

$$\begin{aligned}&\sum_{|n|\le k-1} \frac{1}{n!} D^n\phi(f_0)[F(\varepsilon, w) - f_0]^n l_\varepsilon(J) \\ (2.40)\quad &= \sum_{i=0}^{2m} \sum_{|n|\le k-1} \left\langle \frac{1}{n!} D^n\phi(f_0)[F(\varepsilon, w) - f_0]^n P_i(\varepsilon, w), D^i J \right\rangle_{\overbrace{H\otimes\cdots\otimes H}^{i}},\end{aligned}$$

and since $[F(\varepsilon, w) - f_0]^n P_i(\varepsilon, w)$ has the asymptotic expansion

$$[F(\varepsilon, w) - f_0]^n P_i(\varepsilon, w) \sim \varepsilon^{|n|} e_{n,i,0}(w) + \varepsilon^{|n|+1} e_{n,i,1}(w) + \cdots \quad \text{in } \mathbf{D}^\infty\left(\overbrace{H \otimes \cdots \otimes H}^{i}\right),$$

we have

$$(2.40) = Z_0(w) + \varepsilon Z_1(w) + \cdots + \varepsilon^{k-1} Z_{k-1}(w) + U_k(\varepsilon, w),$$

where

$$Z_l(w) = \sum_{i=0}^{2m} \sum_{|n|+\nu=l} \left\langle \frac{1}{n!} D^n\phi(f_0) e_{n,i,\nu}, D^i J \right\rangle_{\overbrace{H\otimes\cdots\otimes H}^{i}}, \quad l = 0, 1, \ldots, k-1.$$

Also, we can find $c_3 > 0$ such that

$$|E(U_k(\varepsilon, w))| \leq c_3 \varepsilon^k \|J\|_{q,2m}$$

for all $\varepsilon \in [0,1]$ and $J \in \mathbf{D}^\infty$. Let

$$\Phi_l = \sum_{i=0}^{2m} \sum_{|n|+\nu=l} (D^*)^i \left[\frac{1}{n!} D^n \phi(f_0) e_{n,i,\nu} \right], \qquad l = 0, 1, \ldots, k-1.$$

It is easy to see that $\Phi_l \in \bigcap_{1<p<\infty} \mathbf{D}_p^{-2m}$ and thus, combining the above, we obtain

$$\left| \langle T(F(\varepsilon, w)), J \rangle - \sum_{i=0}^{k-1} \varepsilon^i \langle \Phi_i, J \rangle \right| \leq (c_1 c_2 + c_3) \varepsilon^k \|J\|_{q,2m}$$

for all $\varepsilon \in [0,1]$ and $J \in \mathbf{D}^\infty$. Thus

$$\left\| T(F(\varepsilon, w)) - \sum_{i=0}^{k-1} \varepsilon^i \Phi_i \right\|_{p,-2m} \leq (c_1 c_2 + c_3) \varepsilon^k,$$

where $1/p + 1/q = 1$. It is clear in the above argument that $p \in (1, \infty)$ can be chosen arbitrarily and hence (2.39) is obtained.

It remains only to show that Φ_i can be determined by (2.37). Since

$$\begin{aligned} \Phi_l &= \sum_{i=0}^{2m} \sum_{|n|+\nu=l} (D^*)^i \left[\frac{1}{n!} D^n \phi(f_0) e_{n,i,\nu} \right] \\ &= \sum_{i=0}^{2m} \sum_{|n|+\nu=l} (D^*)^i \left[\frac{1}{n!} D^n \{ (1 + |x|^2 - \Delta)^{-m} T \} (f_0) e_{n,i,\nu} \right], \end{aligned}$$

we see that $T \in \mathscr{S}'(\mathbf{R}^d) \mapsto \Phi_l \in \tilde{\mathbf{D}}^{-\infty}$ is continuous in the sense that, for any positive integer n we can find $s > 0$, $p \in (1, \infty)$ and $K > 0$ such that

$$\|\Phi_l\|_{p,-s} \leq K \|T\|_{-2n}.$$

On the other hand, it is clear that Φ_l is uniquely determined from T and must be given by the expansion (2.37) if $T \in \mathscr{S}(\mathbf{R}^d)$. Thus, it must also be given by (2.37) for general $T \in \mathscr{S}'(\mathbf{R}^d)$. This completes the proof. □

Finally, we give some results for the asymptotic expansions in the space $\tilde{\mathbf{D}}^\infty$.

Theorem 2.4. *Let $G(\varepsilon, w) \in \mathbf{D}^\infty$, $\varepsilon \in (0,1]$, have the asymptotic expansion*

(2.41) $$G(\varepsilon, w) \sim g_0 + \varepsilon g_1 + \cdots \quad \text{in } \mathbf{D}^\infty \text{ as } \varepsilon \downarrow 0,$$

with $g_i \in \mathbf{D}^\infty$. Assume further that there exists p, $1 < p < \infty$, such that

(2.42) $$\sup_\varepsilon E[\exp\{pG(\varepsilon, w)\}] < \infty.$$

Then $\exp\{G(\varepsilon, w)\} \in \tilde{\mathbf{D}}^\infty$ *and it has the asymptotic expansion*

(2.43) $\exp\{G(\varepsilon, w)\} \sim \exp\{g_0\}(1 + \varepsilon\gamma_1 + \varepsilon^2\gamma_2 + \cdots)$ *in* $\tilde{\mathbf{D}}^\infty$ *as* $\varepsilon \downarrow 0$,

where $\exp\{g_0\} \cdot \gamma_i \in \tilde{\mathbf{D}}^\infty$ *and* $\gamma_i \in \mathbf{D}^\infty$ *are determined by the following formal expansion in powers of* ε:

$$1 + \varepsilon\gamma_1 + \varepsilon^2\gamma_2 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}(\varepsilon g_1 + \varepsilon^2 g_2 + \cdots)^n.$$

Thus, $\gamma_1 = g_1$, $\gamma_2 = g_2 + g_1^2$, $\gamma_3 = g_3 + 2g_1g_2 + g_1^3, \ldots$.

PROOF. It is easy to see from (2.42) that $\exp\{G(\varepsilon, w)\} \in \tilde{\mathbf{D}}^\infty$ for $\varepsilon \in [0,1]$; we understand that $G(0, w) = g_0$. Next, we note the following elementary estimates:

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x) \quad \text{and} \quad |R_n(x)| \le e^{x \vee 0}\frac{|x|^n}{n!}.$$

In particular, $R_n(x) \le (e^x + 1)|x|^n/n!$. Then

$$\begin{aligned}\exp\{G(\varepsilon, w)\} &= \exp\{g_0\}\exp[\varepsilon g_1 + \varepsilon^2 g_2 + \cdots] \\ &= \exp\{g_0\}\left(\sum_{k=0}^{n-1} \frac{1}{k!}(\varepsilon g_1 + \varepsilon^2 g_2 + \cdots)^k + r_n(\varepsilon, w)\right)\end{aligned}$$

and $|(\exp g_0) r_n(\varepsilon, w)| \le \exp G(\varepsilon, w) A_n(\varepsilon, w)$ with $\|A_n(\varepsilon, w)\|_p = O(\varepsilon^n)$ as $\varepsilon \downarrow 0$ for every $p > 1$. Clearly, similar estimates can also be obtained for derivatives $D^k(\exp\{G(\varepsilon, w)\})$ and the assertion is obvious from this. □

3. The case of Itô functionals. In this section, we apply the results of the previous section to Itô functionals, i.e., Wiener functionals obtained by solving stochastic differential equations (SDE): This will provide us with probabilistic methods with which to study heat kernels, especially their regularity and asymptotics.

Let $V_\alpha(x) = (V_\alpha^i(x))$, $\alpha = 0, 1, \ldots, r$, be a system of $\mathbf{R}^d$-valued functions defined on $\mathbf{R}^d$: $x \in \mathbf{R}^d \to V_\alpha(x) \in \mathbf{R}^d$. We suppose that $V_\alpha^i(x)$ are C^∞-functions with bounded derivatives of all orders. Let $\hat{V}_\alpha$ be the vector field defined by $\hat{V}_\alpha = \sum_{i=1}^d V_\alpha^i(x)\,\partial/\partial x^i$, $\alpha = 0, 1, \ldots, r$. Set

$$A^{ij}(x) = \sum_{\alpha=1}^{r} V_\alpha^i(x) V_\alpha^j(x) \tag{3.1}$$

and define the second-order differential operator A by

$$A = \tfrac{1}{2}\sum_{\alpha=1}^{r} \hat{V}_\alpha^2 + \hat{V}_0. \tag{3.2}$$

Note that the main (second-order) term of A coincides with

$$\frac{1}{2}\sum_{i,j=1}^{d} A^{ij}(x)\frac{\partial^2}{\partial x^i\,\partial x^j}.$$

Let $(W = W_0^r, P)$ be, as before, the r-dimensional Wiener space and consider the following SDE on $\mathbf{R}^d$

$$(3.3)\qquad \begin{aligned} dX_t &= \sum_{\alpha=1}^{r} V_\alpha(X_t)\circ dw^\alpha(t) + V_0(X_t)\,dt,\\ X_0 &= x \in \mathbf{R}^d, \end{aligned}$$

or, in the component form of $X_t = (X_t^1, \dots, X_t^d)$,

$$(3.3')\qquad \begin{aligned} dX_t^i &= \sum_{\alpha=1}^{r} V_\alpha^i(X_t)\circ dw^\alpha(t) + V_0^i(X_t)\,dt\\ X_0^i &= x^i, \qquad i = 1,2,\dots,d. \end{aligned}$$

Here $w = (w^\alpha(t))$ is a generic element of W, which is clearly a realization of an r-dimensional Wiener process under the measure P and $\circ$ indicates the stochastic differential in the Stratonovich sense, cf. [9]. It is well known ([9]) that the unique solution $X_t = X(t,x,w)$ to the above equation (3.3) exists for every $x \in \mathbf{R}^d$ such that

(i) $t \to X(t,x,w)$ is a sample path of A-diffusion process starting at x,
(ii) with probability one, $(t,x) \to X(t,x,w)$ is continuous and $x \to X(t,x,w)$ is a diffeomorphism of $\mathbf{R}^d$.

Let $Y_t = (Y_j^i(t,x,w))$ be defined by $Y_j^i(t,x,w) = (\partial X^i/\partial x^j)(t,x,w)$. Then Y_t is determined as the unique solution of the following equation in the matrix notation: denoting $\partial V_\alpha(x) = ((\partial/\partial x^j)V_\alpha^i(x))$, $\alpha = 0,1,\dots,r$,

$$(3.4)\qquad \begin{aligned} dY_t &= \sum_{\alpha=1}^{r} \partial V_\alpha(X_t)\,Y_t\circ dw^\alpha(t) + \partial V_0(X_t)\,Y_t\,dt,\\ Y_0 &= I\text{: the identity matrix.} \end{aligned}$$

(3.3) and (3.4), combined together, define the SDE for the process $r_t = (X_t, Y_t)$ on $\mathbf{R}^d \times GL(d,\mathbf{R})$. The following results are well known (cf. [9], [10] and [27]).

THEOREM 3.1. *Let $t > 0$ and $x \in \mathbf{R}^d$ be fixed. Then $X(t,x,w)$ is smooth in the sense that*

$$(3.5)\qquad X(t,x,w) \in \mathbf{D}^\infty(\mathbf{R}^d), \quad \text{i.e., } X^i(t,x,w) \in \mathbf{D}^\infty,\ i = 1,2,\dots,d.$$

Furthermore, the Malliavin covariance $\sigma_t^{ij}(w) = \langle DX^i(t,x,w), DX^j(t,x,w)\rangle_H$ is given by

$$(3.6)\qquad \sigma_t^{ij} = \sum_{\alpha=1}^{r}\int_0^t \left(Y_tY_s^{-1}V_\alpha(X_s)\right)^i\left(Y_tY_s^{-1}V_\alpha(X_s)\right)^j ds.$$

Based on the formula (3.6), we can study the nondegeneracy, in the sense of (2.17), of $X(t, x, w)$. Consider the following condition:

(H.1) $\dim \mathscr{L}\{[V_{\alpha_k}, [V_{\alpha_{k-1}}, [\ldots, [V_{\alpha_1}, V_{\alpha_0}]\ldots](x)$: $0 \le k \le k_0$, where $\alpha_0 \in \{1, 2, \ldots, r\}$ and $\alpha_i \in \{0, 1, \ldots, r\}$ if $1 \le i \le k_0\} = d$ for some k_0.

Here,

$$[L_1, L_2]^i(x) = \sum_{j=1}^{d} \left\{ L_1^j(x) \frac{\partial}{\partial x^j}\left(L_2^i(x)\right) - L_2^j(x) \frac{\partial}{\partial x^j}\left(L_1^i(x)\right) \right\}$$

if $L_1 = (L_1^i(x))$ and $L_2 = (L_2^i(x))$ and $\mathscr{L}$ denotes the linear hull in $\mathbf{R}^d$.

THEOREM 3.2 (Kusuoka and Stroock [14]). *If* (H.1) *is satisfied at* $x \in \mathbf{R}^d$, *then for every* $t > 0$, $X(t, x, w) \in \mathbf{D}^\infty(\mathbf{R}^d)$ *satisfies the nondegeneracy condition* (2.17). *More precisely, there exists a positive integer n depending only on* k_0 *in* (H.1) *and, for each* $1 < p < \infty$, *a positive constant* $c = c(p, x)$ *such that*

$$\left\|(\det \sigma_t)^{-1}\right\|_p \le c t^{-n} \quad \textit{for all } t > 0. \tag{3.7}$$

If (H.1) *is satisfied everywhere in a domain D of* $\mathbf{R}^d$, *then the estimate* (3.7) *holds uniformly in* $x \in K$ *for any bounded set* $K \Subset D$.

This theorem is a consequence of precise estimates of Kusuoka and Stroock [14]. Also this follows from a key-lemma as stated in [10] and [27] combined with a scaling property of the solution $X(t, x, w)$: For $\varepsilon > 0$, $\{X(\varepsilon^2 t, x, w),\ t > 0\}$ is equivalent in law to $\{X^\varepsilon(t, x, w),\ t > 0\}$, where X_t^ε is the solution to SDE (3.3) in which V_α is replaced by εV_α for $\alpha = 1, 2, \ldots, r$ and V_0 by $\varepsilon^2 V_0$.

Thus, if (H.1) is satisfied at $x \in \mathbf{R}^d$ then for any Schwartz distribution T, $T(X(t, x, w)) \in \mathbf{D}^{-\infty}$ is defined for every $t > 0$. In particular, $\delta_y(X(t, x, w))$ is defined for every y and $p(t, x, y) = E[\delta_y(X(t, x, w))]$, which is a C^∞-function in y, coincides with the fundamental solution of the heat equation $\partial/\partial t = A$. More generally, if $c(x)$ is a real C^∞-function on $\mathbf{R}^d$ such that $\exp \int_0^t c(X(s, x, w))\, ds \in \tilde{\mathbf{D}}^\infty$ [it is sufficient to assume, for example, that $c(x)$ and its derivatives are of polynomial growth and $c(x) \le K$ for some constant $K > 0$], then

$$p^c(t, x, y) = E\left[\exp\left(\int_0^t c(X(s, x, w))\, ds\right) \delta_y(X(t, x, w))\right]$$

is the fundamental solution of the heat equation $\partial/\partial t = A + c$.

Now assume that V_α, $\alpha = 0, 1, \ldots, r$, are bounded on $\mathbf{R}^d$. Let $\phi(\xi)$ be a C^∞-function on R^d such that $\phi(\xi) = 1$, $|\xi| \le \frac{1}{3}$ and $\phi(\xi) = 0$, $|\xi| \ge \frac{2}{3}$. Let $y \in \mathbf{R}^d$ and define $\psi(z) = \phi((z - y)/|x - y|)$ if $x \ne y$ and $\phi(z) = 1$ if $x = y$. Then $\psi \cdot \delta_y = \delta_y$ and hence

$$p(t, x, y) = E\left[\psi(X(t, x, w))\delta_y(X(t, x, w))\right].$$

By an integration by parts, this is easily seen to be a finite sum $\sum_i E(a_i(X(t, x, w))F_i(w))$, where $a_i(x)$ is the form of a finite sum $a_i(x) =$

$\sum_j b_j(x) D^{\beta_j}\psi(x)$ with bounded continuous functions $b_j(x)$, D^{β_j} is the differential operator in the usual multi-index notation and finally, $F_i(w) \in \mathbf{D}^\infty$ is a polynomial in the components of $X(t,x,w)$, their derivatives and $\det \sigma_t(w)^{-1}$ [cf. (2.19) and (2.20)]. The following estimate is well known, cf. [9], page 342: There exist positive constants $a_1 \geq 1$, a_2, a_3 such that

$$P\left(|X(t,x,w) - x| \geq \frac{|x-y|}{3}\right) \leq a_1 \exp\left[-\frac{a_2|x-y|^2}{t}\right]$$

provided $0 < t < a_3|x-y|$. Hence

$$P\left(|X(t,x,w) - x| \geq \frac{|x-y|}{3}\right)$$
$$\leq a_1 \exp\left[\frac{a_2}{a_3|x-y|}\right]\exp\left[-\frac{a_2|x-y|^2}{t}\right],$$

and combining this with (3.7), we can obtain the following estimate:

Suppose that V_α, $\alpha = 0,1,\dots,r$, are bounded and that (H.1) *is satisfied everywhere in a domain D of $\mathbf{R}^d$. Then, for every compact set $K \Subset D$ and $T > 0$, we can determine positive constants c_1 and c_2 such that*

$$p(t,x,y) \leq \frac{c_1}{t^\nu}\exp\left[-\frac{c_2|x-y|^2}{t}\right] \tag{3.8}$$

for all $t \in (0,T]$, $x \in K$ and $y \in \mathbf{R}^d$.

Here $\nu > 0$ is a constant depending only on d and $\max k(x)$ in K, where $k(x)$ is k_0 in (H.1) at x. Similar estimates can also be obtained for derivatives $D_y^\beta p(t,x,y) = E[(-1)^{|\beta|}(D^\beta \delta_y)(X(t,x,w))]$ with different ν and c_1, c_2, cf. [14] for more precise statements.

Now we shall study the short-time asymptotics of the fundamental solution $p(t,x,y)$. For this, we introduce a parameter $\varepsilon \in (0,1]$ in SDE (3.3) as follows:

$$dX_t = \varepsilon \sum_{\alpha=1}^{r} V_\alpha(X_t)\circ dw^\alpha(t) + \varepsilon^2 V_0(X_t)\,dt, \tag{3.9}$$
$$X_0 = x.$$

We denote the solution of this equation by $X^\varepsilon(t,x,w)$. It is easy to see that $\{X^1(\varepsilon^2 t,x,w),\, t>0\}$ is equivalent in law to $\{X^\varepsilon(t,x,w),\, t>0\}$. Hence, if (H.1) is satisfied at x,

$$p(\varepsilon^2,x,y) = E\left[\delta_y\left(X^1(\varepsilon^2,x,w)\right)\right] = E\left[\delta_y\left(X^\varepsilon(1,x,w)\right)\right],$$

and, for this reason, we study the family $X^\varepsilon(1,x,w) \in \mathbf{D}^\infty(\mathbf{R}^d)$, $\varepsilon \in (0,1]$. First of all, we introduce the following notation: for $\mathbf{i} = (i_1,\dots,i_m) \in \{0,1,2,\dots,r\}^m$, we set $\alpha(\mathbf{i}) = \#\{\nu\colon i_\nu = 0\}$ and $\|\mathbf{i}\| = m + \alpha(\mathbf{i})$. Also, let

$$S^{\mathbf{i}}(t,w) = \int_0^t \circ\, dw^{i_1}(t_1)\int_0^{t_1}\circ\, dw^{i_2}(t_2)\cdots\int_0^{t_{m-1}}\circ\, dw^{i_m}(t_m)$$

be a multiple stochastic integral in the Stratonovich sense for $\mathbf{i} = (i_1, \dots, i_m)$, where we set $w^0(t) = t$.

THEOREM 3.3. *Let $x \in \mathbf{R}^d$ be fixed. Then $X^\varepsilon(1, x, w) \in \mathbf{D}^\infty(\mathbf{R}^d)$ has the asymptotic expansion*

$$X^\varepsilon(1, x, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \text{in } \mathbf{D}^\infty(\mathbf{R}^d) \text{ as } \varepsilon \downarrow 0, \tag{3.10}$$

and $f_n \in \mathbf{D}^\infty(\mathbf{R}^d)$, $n = 0, 1, \dots$, are given by

$$f_0 = x \tag{3.11}$$

and

$$f_n = \sum_{\mathbf{i},\ \|\mathbf{i}\| = n} \left(\hat{V}_{i_m} \circ \cdots \circ \hat{V}_{i_2}\right)(V_{i_1})(x) S^{\mathbf{i}}(1, w), \qquad n = 1, 2, \dots. \tag{3.12}$$

In particular,

$$f_1(w) = \sum_{\alpha=1}^{r} V_\alpha(x) w^\alpha(1). \tag{3.13}$$

The expansion (3.10) *is uniform in $x \in K$ for any bounded set K in $\mathbf{R}^d$.*

The proof is easily provided by successive applications of the Itô formula:

$$\begin{aligned}
&X^\varepsilon(1, x, w) - x \\
&= \varepsilon \sum_{\alpha=1}^{r} \int_0^1 V_\alpha(X_s^\varepsilon) \circ dw_s + \varepsilon^2 \int_0^1 V_0(X_s^\varepsilon)\, ds \\
&= \varepsilon \sum_{\alpha=1}^{r} V_\alpha(x) w^\alpha(1) + \varepsilon \sum_{\alpha=1}^{r} \int_0^1 [V_\alpha(X_s^\varepsilon) - V_\alpha(x)] \circ dw_s^\alpha + \varepsilon^2 \int_0^1 V_0(X_s^\varepsilon)\, ds \\
&= \varepsilon f_1 + \varepsilon^2 \Bigg(\sum_{i_1=1}^{r} \sum_{i_2=1}^{r} \int_0^1 \left\{ \int_0^s \hat{V}_{i_2}(V_{i_1})(X_u^\varepsilon) \circ dw^{i_2}(u) \right\} \circ dw^{i_1}(s) \\
&\qquad + \int_0^1 V_0(X_s^\varepsilon)\, ds + \varepsilon \sum_{i_1=1}^{r} \int_0^1 \left\{ \int_0^s \hat{V}_0(V_{i_1})(X_u^\varepsilon)\, du \right\} \circ dw^{i_1}(u) \Bigg) \\
&= \varepsilon f_1 + \varepsilon^2 \left(V_0(x) + \sum_{i_1=1}^{r} \sum_{i_2=1}^{r} \hat{V}_{i_2}(V_{i_1})(x) S^{(i_1, i_2)}(1) \right) \\
&\quad + \varepsilon^2 \Bigg(\sum_{i_1=1}^{r} \sum_{i_2=1}^{r} \int_0^1 \left\{ \int_0^s [\hat{V}_{i_2}(V_{i_1})(X_u^\varepsilon) - \hat{V}_{i_2}(V_{i_1})(x)] \circ dw^{i_2}(u) \right\} \circ dw^{i_1}(s) \\
&\quad + \int_0^1 [V_0(X_s^\varepsilon) - V_0(x)]\, ds \\
&\quad + \varepsilon \sum_{i=1}^{r} \int_0^1 \left\{ \int_0^s [\hat{V}_0(V_i)(X_u^\varepsilon) - \hat{V}_0(V_i)(x)]\, du \right\} \circ dw^i(u) \Bigg) \cdots.
\end{aligned}$$

Continuing this, it is easy to see that

$$X^{\varepsilon}(1,x,w)=f_0+\varepsilon f_1+\cdots+\varepsilon^n f_n+O(\varepsilon^{n+1})$$

in L^p for every $p\in(1,\infty)$, and since $D^kX^{\varepsilon}(i,x,w)[h_1,\ldots,h_k]$ are determined successively by SDE ([9] and [27]), it is easy to show that

$$D^kX^{\varepsilon}(1,x,w)=f_0^{(k)}+\varepsilon f_1^{(k)}+\cdots+\varepsilon^n f_n^{(k)}+O(\varepsilon^{n+1})$$

in $L_p(\overbrace{H\otimes\cdots\otimes H}^{k})$ for every $p\in(1,\infty)$. □

$X^{\varepsilon}(1,x,w)$ is not uniformly nondegenerate in the sense of (2.34) because $f_0=x$ which is completely degenerate. However, if we consider $F(\varepsilon,w)=(X^{\varepsilon}(1,x,w)-x)/\varepsilon$, $\varepsilon>0$, then we have the following:

THEOREM 3.4. *The family* $F(\varepsilon,w)=(X^{\varepsilon}(1,x,w)-x)/\varepsilon$, $\varepsilon>0$, *is uniformly nondegenerate in the sense of* (2.34) *if and only if* $(A^{ij}(x))$ *defined by* (3.1) *is nondegenerate, i.e.*, $\det(A^{ij}(x))>0$.

PROOF. Since $F(\varepsilon,w)\sim f_1+\varepsilon f_2+\cdots$, where the f_i are those in (3.12) and f_1 that in (3.13) whose Malliavin covariance coincides with $A^{ij}(x)$, $A^{ij}(x)$ must be nondegenerate in order that $F(\varepsilon,w)$ is uniformly nondegenerate.

Conversely, suppose that $\det(A^{ij}(x))>0$. Denoting $A(x)=(A^{ij}(x))$ and $\sigma(\varepsilon)=(\sigma^{ij}(\varepsilon))$, where $\sigma(\varepsilon)$ is the Malliavin covariance of $F(\varepsilon,w)$, we have

$$\sigma(\varepsilon)=\int_0^1 Y_1^{\varepsilon}(Y_s^{\varepsilon})^{-1}A(X_s^{\varepsilon})^t\left[Y_1^{\varepsilon}(Y_s^{\varepsilon})^{-1}\right]ds,$$

and, since εV_α, $\alpha=1,2,\ldots,r$ and ε^2V_0 are obviously bounded in $\varepsilon\in(0,1]$, it is easy to see that $P[\tau<1/n]\le c_1\exp(-c_2n^{c_3})$, $n=1,2,\ldots$, where c_i, $i=1,2,3$, are positive constants independent of ε and n, $\tau=\min\{s\colon (Y_s^{\varepsilon})^{-1}A(X_s^{\varepsilon})^t(Y_s^{\varepsilon})^{-1}\le\frac{1}{2}A(x)\}$ ($A_1\le A_2$ if and only if A_2-A_1 is nonnegative definite), and that $\sup_\varepsilon\|\det(Y_1^{\varepsilon})^{-1}\|_p<\infty$ for all $p>1$. Now

$$\det\sigma(\varepsilon)\ge(\det Y_1^{\varepsilon})^2\det\int_0^{1\wedge\tau}(Y_s^{\varepsilon})^{-1}A(X_s^{\varepsilon})^t(Y_s^{\varepsilon})\,ds$$
$$\ge\frac{1}{2^d}(\det(Y_1^{\varepsilon}))^2\det A(x)(1\wedge\tau)^d$$

and it is easy to conclude from this that

$$\sup_{\varepsilon\in(0,1]}\left\|\det\sigma(\varepsilon)^{-1}\right\|_p<\infty\quad\text{for all } p\in(1,\infty).$$ □

Suppose that $\det(A^{ij}(x))>0$. Then by Theorem 2.3, $T(F(\varepsilon,w))$ has the asymptotic expansion in $\tilde{\mathbf{D}}^{-\infty}$. Since

$$\delta_x(X^{\varepsilon}(1,x,w))=\delta_x(x+\varepsilon F(\varepsilon,w))=\varepsilon^{-d}\delta_0(F(\varepsilon,w)),$$

we have the following:

THEOREM 3.5. *Suppose that* $\det(A^{ij}(x))>0$. *Then* $\delta_x(X^{\varepsilon}(1,x,w))$ *has the following asymptotic expansion in* $\tilde{\mathbf{D}}^{-\infty}$ *as* $\varepsilon\downarrow0$:

$$\delta_x(X^{\varepsilon}(1,x,w))\sim\varepsilon^{-d}(\Phi_0+\varepsilon\Phi_1+\varepsilon^2\Phi_2+\cdots)\tag{3.14}$$

and $\Phi_k \in \tilde{\mathbf{D}}^{-\infty}$ *can be obtained explicitly by Theorem* 2.3:

$$\Phi_k(w) = \sum_{\mathbf{j},\mathbf{n}} \frac{1}{|\mathbf{j}|!} \partial^{\mathbf{j}}\delta_0(f_1) f_{n_1}^{j_1} f_{n_2}^{j_2} \cdots f_{n_l}^{j_l}, \tag{3.15}$$

where, $f_n = (f_n^i)$ *is given by* (3.12), *and the summation extends over all* $\mathbf{j} = (j_1, \dots, j_l) \in \{1, 2, \dots, d\}^l$ *and* $\mathbf{n} = (n_1, \dots, n_l)$, $n_i \geq 2$, $l = 0, 1, \dots,$ *such that* $n_1 + n_2 + \cdots + n_l - l = k$. *Also,* $\partial^{\mathbf{j}} = (\partial/\partial x^{j_1})(\partial/\partial x^{j_2}) \cdots (\partial/\partial x^{j_l})$ *and* $|\mathbf{j}| = l$ *if* $\mathbf{j} = (j_1, \dots, j_l)$. *In particular,* $\Phi_0 = \delta_0(f_1)$.

From this, we see that $E(\Phi_k(w)) = 0$ if k is odd since, then, $\Phi_k(-w) = -\Phi_k(w)$ and the map $w \to -w$ preserves the measure P.

COROLLARY. *Suppose that* $\det(A^{ij}(x)) > 0$. *Then* $p(t, x, x)$ *has the asymptotic expansion as* $t \downarrow 0$:

$$p(t, x, x) \sim t^{-d/2}(c_0 + c_1 t + \cdots) \tag{3.16}$$

and c_i *is given by*

$$c_i = E(\Phi_{2i}), \qquad i = 0, 1, \dots, \tag{3.17}$$

where Φ_i *is given by* (3.15). *In particular,* $c_0 = [(2\pi)^{d/2}\det A(x)]^{-1}$.

If $\det(A^{ij}(x)) > 0$, then the inverse $(A_{ij}(x))$ of $(A^{ij}(x))$ induces a Riemannian structure and coefficients c_i in (3.16) can be described in terms of the Riemann curvature tensor and its covariant derivatives. The problem of computing these coefficients was discussed by McKean and Singer [18]. We can also apply our result to this problem: We introduce the normal coordinate system around x and compute c_i by the above method. Then the computation is reduced to that of

$$E\left(\prod_{\nu=1}^{l} S^{i_\nu}(1, w)\, \partial^k \delta_0(w(1))\right)$$
$$= (-1)^{|k|} \partial_x^k \left\{ E\left(\prod_{\nu=1}^{l} S^{i_\nu}(1, w) | w(1) = x \right) (2\pi)^{-d/2} e^{-|x|^2/2} \right\} \bigg|_{x=0}.$$

We can carry out the computation for small k and l and thereby obtain c_0, c_1 and c_2 explicitly. The computation is quite complicated for larger k and l but by analyzing the expectation, we can show that this gives a linear combination, with universal coefficients, of certain products of Kronecker's δ's, cf. Uemura [26]. This fact is essentially a consequence of a well-known fact that a moment of a system of centered Gaussian random variables is a sum of products of pairwise second moments. In this way, we can also rediscover a result of [18] that c_i is a linear combination with universal coefficients of certain monomials in components of curvature tensor and its covariant derivatives. The type of these monomials can be specified and the universal coefficients can be determined by explicit computation for simple manifolds. In [18], this result is obtained by Weyl's invariant theory together with some geometric considerations. Cf. also [4].

If $\det(A^{ij}(x)) = 0$, then the asymptotic expansion of $\delta_x(X^\varepsilon(1, x, w))$ is generally not easy to obtain. However, there are some cases in which we can obtain such expansions. Suppose, for example, that *there exists a diffeomorphism* $\xi \in \mathbf{R}^d \to \theta(\xi) \in \mathbf{R}^d$ *such that* $\theta(x) = 0$, *derivatives are all polynomial growth and, for some* $k_i \in N$, $i = 1, 2, \ldots, d$,

$$\begin{pmatrix} \varepsilon^{-k_1} & & & 0 \\ & \varepsilon^{-k_2} & & \\ & & \ddots & \\ 0 & & & \varepsilon^{-k_d} \end{pmatrix} \theta(X^\varepsilon(1, x, w)) = F(\varepsilon, w) \tag{3.18}$$

has the asymptotic expansion

$$F(\varepsilon, w) \sim f_0 + \varepsilon f_1 + \cdots \quad \text{in } \mathbf{D}^\infty(\mathbf{R}^d) \text{ as } \varepsilon \downarrow 0,$$

with $f_i \in \mathbf{D}^\infty(\mathbf{R}^d)$, *and furthermore*, $F(\varepsilon, w)$ *is uniformly nondegenerate*. Then, setting $\hat{k} = k_1 + k_2 + \cdots + k_d$,

$$\delta_x(X^\varepsilon(1, x, w)) = \varepsilon^{-\hat{k}} \left| \det\left(\frac{\partial \theta}{\partial \xi} \right)(x) \right| \delta_0(F(\varepsilon, w))$$

and by Theorem 2.3,

$$\delta_0(F(\varepsilon, w)) \sim \Phi_0 + \varepsilon \Phi_1 + \cdots \quad \text{in } \tilde{\mathbf{D}}^{-\infty} \text{ as } \varepsilon \downarrow 0.$$

Hence, $\delta_x(X^\varepsilon(1, x, w))$ has the following expansion:

$$\delta_x(X^\varepsilon(1, x, w)) \sim \varepsilon^{-\hat{k}} \left| \det\left(\frac{\partial \theta}{\partial \xi} \right)(x) \right| (\Phi_0 + \varepsilon \Phi_1 + \cdots) \quad \text{in } \tilde{\mathbf{D}}^{-\infty} \text{ as } \varepsilon \downarrow 0. \tag{3.19}$$

In the case that $\det A(x) > 0$, the above assumption is satisfied by taking $\theta(\xi) = \xi - x$ and $k_i = 1$ for all i.

Another typical case that the above assumption is satisfied is when $V_0 = 0$ and vector fields $\hat{V}_1, \ldots, \hat{V}_r$ are free of order s at x (cf. [8] and [21]) and furthermore, they, together with their commutators of length s, span $\mathbf{R}^d$ at x. In this case, θ: $\mathbf{R}^d \to \mathbf{R}^d$ is given, in a neighborhood of x, by

$$\theta\left(\left[\exp \sum_{I \in B} u_I L_{[I]} \right]_x \right) = (u_I)_{I \in B},$$

where $L_{[I]}$, $I \in B$, denotes the family of commutators of length $\leq s$ such that $L_{[I]}$, $I \in B$ forms a basis for $T_x(\mathbf{R}^d)$. Here we follow the notation of [8]: $I = (\alpha_1, \ldots, \alpha_k)$ is a sequence of indices $\alpha \in \{1, 2, \ldots, r\}$ and $L_{[I]} = (\operatorname{ad} \hat{V}_{\alpha_1} \circ \operatorname{ad} \hat{V}_{\alpha_2} \cdots \circ \operatorname{ad} \hat{V}_{\alpha_{k-1}}) \hat{V}_{\alpha_k}$ $(\operatorname{ad} \hat{V}_1(\hat{V}_2) = \hat{V}_1 \hat{V}_2 - \hat{V}_2 \hat{V}_1)$.

In particular, $\#B = d$. $k_I = k$, $I \in B$, if $I = (\alpha_1, \ldots, \alpha_k)$. In this case, a free nilpotent Lie group G of step s exists with the Lie algebra generated by r left-invariant vector fields $Y_1, Y_2, \ldots, Y_r$ such that if g_t is the solution of the following SDE on G:

$$\begin{aligned} dg_t &= \sum_{\alpha=1}^r Y_\alpha(g_t) \circ dw^\alpha(t), \\ g_0 &= e \quad \text{(identity)} \end{aligned} \tag{3.20}$$

and if the $\mathbf{R}^d$-valued process $(g_t^I)_{I\in B}$ is defined by

$$\exp\left(\sum_{I\in B} g_t^I Y_{[I]}\right) = g_t,$$

where $Y_{[I]}$ is the commutator of $Y_1,\dots,Y_r$ corresponding to I, then the first term f_0 in the expansion of $F(\varepsilon, w)$ coincides with $(g_1^I)_{I\in B}$.

Finally, we shall discuss the short-time asymptotics of $p(t,x,y)$ off the diagonal, i.e., $x \neq y$. This problem was discussed by Bismut [5] as an application of Wiener functional analysis; the splitting of the Wiener space and the use of an implicit function theorem. Here we will treat this problem by our method introduced above.

Let $V_\alpha(x) = (V_\alpha^i(x))$, $\alpha = 0,1,\dots,r$, be as above. In the following, we assume, for simplicity,

(H.1′) the condition (H.1) strengthened as follows:
$\dim \mathscr{L}\{[V_{\alpha_k},[V_{\alpha_{k-1}},[\dots,[V_{\alpha_1},V_{\alpha_0}]\dots](x);\ 0\le k\le k_0,$
$\alpha_i \in \{1,\dots,r\}\} = d$ for some k_0.

Let $h = (h^\alpha(t))_{\alpha=1}^r \in H$, an element in the Cameron–Martin subspace of $W = W_0^r$ and consider the following SDE on $\mathbf{R}^d$ for each $\varepsilon \in (0,1]$:

$$dX_t = \varepsilon\sum_{\alpha=1}^r V_\alpha(X_t)\circ dw^\alpha(t) + \varepsilon^2 V_0(X_t)\,dt + \sum_{\alpha=1}^r V_\alpha(X_t)\dot h^\alpha(t)\,dt, \tag{3.21}$$

where $\dot h^\alpha(t) = (dh^\alpha/dt)(t)$. We denote this solution by $X^{\varepsilon,h}(t,x,w)$. Thus $X^{\varepsilon,h}(t,x,w) = X^\varepsilon(t,x,w+h/\varepsilon)$, where $X^\varepsilon(t,x,w)$ is the solution to (3.9). If $V_0 = 0$, we may also consider, at least formally, that $X^{\varepsilon,h}(t,x,w) = X(t,x,\varepsilon w + h)$, where $X(t,x,w)$ is the solution to (3.3). It is easy to see that $X^{\varepsilon,h}(1,x,w) \in \mathbf{D}^\infty(\mathbf{R}^d)$ for every $\varepsilon \in (0,1]$, $h \in H$ and $x \in \mathbf{R}^d$. Furthermore, we can easily obtain its asymptotic expansion. In the following, the time interval of $w \in W$ and $h \in H$ are restricted to $[0,1]$: In particular, $\|h\|_H^2 = \int_0^1 |dh/dt|^2(s)\,ds$.

First of all, we shall introduce some notation. For $h \in H$, consider the following dynamical system on $\mathbf{R}^d$:

$$\frac{d\xi_t}{dt} = \sum_{\alpha=1}^r V_\alpha(\xi_t)\dot h^\alpha(t), \qquad \xi_0 = x, \tag{3.22}$$

and denote the solution ξ_t by $\xi^{x,h}(t)$. The corresponding variational equation is given, in $d\times d$-matrix notation, by

$$\frac{d\Xi_t}{dt} = \sum_{\alpha=1}^r \partial V_\alpha(\xi^{x,h}(t))\,\Xi_t \dot h^\alpha(t), \qquad \Xi(0) = I, \tag{3.23}$$

where $\partial V_\alpha(x) = ((\partial V^i/\partial x^j)(x))$. We denote the solution Ξ_t by $\Xi^{x,h}(t)$.

Theorem 3.6. *For fixed $x \in \mathbf{R}^d$ and $h \in H$, $X^{\varepsilon, h}(1, x, w)$ has the asymptotic expansion*

$$X^{\varepsilon, h}(1, x, w) \sim f_0^{(h)} + \varepsilon f_1^{(h)} + \cdots \quad \text{in } \mathbf{D}^\infty(\mathbf{R}^d) \text{ as } \varepsilon \downarrow 0, \tag{3.24}$$

and $f_i^{(h)} \in \mathbf{D}^\infty(\mathbf{R}^d)$ is obtained successively by

$$f_0^{(h)} = \xi^{x, h}(1), \tag{3.25}$$

$$f_1^{(h)} = \Xi^{x, h}(1) \int_0^1 \Xi^{x, h}(s)^{-1} V_\alpha(\xi^{x, h}(s))\, dw^\alpha(s), \tag{3.26}$$

$$\begin{aligned} f_2^{(h)} = \Xi^{x, h}(1) \int_0^1 \Xi^{x, h}(s)^{-1} \big[& \tfrac{1}{2} \partial^2 V_\alpha(\xi^{x, h}(s)) \eta^{x, h}(s) \otimes \eta^{x, h}(s) \dot{h}^\alpha(s)\, ds \\ & + \partial V_\alpha(\xi^{x, h}(s)) \eta^{x, h}(s) \circ dw^\alpha(s) \\ & + V_0(\xi^{x, h}(s))\, ds \big] \cdots, \end{aligned} \tag{3.27}$$

where

$$\eta^{x, h}(t) = \Xi^{x, h}(t) \int_0^t \Xi^{x, h}(s)^{-1} V_\alpha(\xi^{x, h}(s))\, dw^\alpha(s). \tag{3.28}$$

[*Note that $\eta^{x, h}(t)$ is a $\mathbf{R}^d$-valued continuous Gaussian process.*] *The asymptotic expansion* (3.24) *is uniform in (x, h) on any bounded set in $\mathbf{R}^d \times H$.*

In the above, we omitted the summation sign Σ_α and used the notation

$$\partial^2 V_\alpha(x) = \left(\frac{\partial^2 V_\alpha^i}{\partial x^j\, \partial x^k}(x) \right),$$

$$\left[\partial^2 V_\alpha \cdot \eta \otimes \eta \right]^i = \sum_{j, k} \left(\frac{\partial^2 V_\alpha^i}{\partial x^j\, \partial x^k} \right) \eta^j \eta^k.$$

We omit the proof of this theorem because it is rather routine. See the proof of Lemma 3.4 below in which some of necessary computations will be made.

Next, we study the conditions for uniform nondegeneracy of $\sigma^h(\varepsilon) = (\sigma^{ij;\, h}(\varepsilon))$ defined by

$$\begin{aligned} \sigma^h(\varepsilon) &= \left\langle D \frac{\left(X^{\varepsilon, h}(1, x, w) - f_0^{(h)} \right)}{\varepsilon}, D \frac{\left(X^{\varepsilon, h}(1, x, w) - f_0^{(h)} \right)}{\varepsilon} \right\rangle_H \\ &= \varepsilon^{-2} \left\langle D X^{\varepsilon, h}(1, x, w), D X^{\varepsilon, h}(1, x, w) \right\rangle_H. \end{aligned}$$

Lemma 3.1. *If $x \in \mathbf{R}^d$ and $h \in H \cap C^2([0,1])$ satisfies*

$$\inf_{l \in S^{d-1}} \sum_{\alpha=1}^r \left\{ \langle l, V_\alpha(x) \rangle^2 + \left\langle l, \sum_{\beta=1}^r \left[\dot{h}^\beta(0) V_\beta(x), V_\alpha(x) \right] \right\rangle^2 \right\} > 0, \tag{3.29}$$

then $\sigma^h(\varepsilon)$ is uniformly nondegenerate in the sense that

$$\limsup_{\varepsilon \downarrow 0} \left\| \det \sigma^h(\varepsilon)^{-1} \right\|_p < \infty \quad \text{for all } p \in (1, \infty).$$

Proof. First we note that $\varepsilon_0 > 0$ can be chosen so that for all $0 \le \varepsilon \le \varepsilon_0$,

$$\inf_{l \in S^{d-1}} \sum_{\alpha=1}^{r} \left\{ \langle l, V_\alpha(x) \rangle^2 + \left\langle l, \sum_{\beta=1}^{r} \left[\dot{h}^\beta(0) V_\beta + \varepsilon^2 V_0, V_\alpha \right](x) \right\rangle^2 \right\} > 0. \tag{3.30}$$

Let $Y_t^{\varepsilon, h}$ be the solution to

$$\begin{aligned} dY_t &= \varepsilon\, \partial V_\alpha\left(X_t^{\varepsilon, h}\right) Y_t \circ dw^\alpha(t) + \varepsilon^2\, \partial V_0\left(X_t^{\varepsilon, h}\right) Y_t\, dt \\ &\quad + \partial V_\alpha\left(X_t^{\varepsilon, h}\right) Y_t \dot{h}^\alpha(t)\, dt, \\ Y_0 &= I. \end{aligned}$$

Then

$$\sigma^{ij;h}(\varepsilon) = \sum_{\alpha=1}^{r} \int_0^1 \left[Y_1^{\varepsilon, h} \left(Y_s^{\varepsilon, h}\right)^{-1} V_\alpha\left(X_s^{\varepsilon, h}\right) \right]^i \left[Y_1^{\varepsilon, h} \left(Y_s^{\varepsilon, h}\right)^{-1} V_\alpha\left(X_s^{\varepsilon, h}\right) \right]^j ds.$$

It is easy to see that $\sup_\varepsilon \|(Y_1^{\varepsilon, h})^{-1}\|_p < \infty$ for all $p \in (1, \infty)$. By the same standard reasoning as in [10] or [27], it is sufficient to find, for each $n = 1, 2, \dots$, a stopping time $\tau \le n^{-1}$ such that

$$P\left[\tau < \frac{1}{n} \right] \le c_1 \exp(-c_2 n^{c_3}) \tag{3.31}$$

[here and in the following, $c_1, c_2, \dots$ and $a_1, a_2, \dots$ are positive constants independent of n, $\varepsilon \in (0, \varepsilon_0]$ and $l \in S^{d-1}$] and, for all $l \in S^{d-1}$ and $\varepsilon \in (0, \varepsilon_0]$,

$$P\left[\int_0^\tau \left\{ \sum_{\alpha=1}^{r} \left\langle \left(Y_s^{\varepsilon, h}\right)^{-1} V_\alpha\left(X_s^{\varepsilon, h}\right), l \right\rangle^2 ds \le n^{-c_5} \right] \le c_6 \exp(-c_7 n^{c_8}), \tag{3.32}$$

for some c_5, c_6, c_7, c_8. But by (3.30), we can find $\tau \le n^{-1}$ satisfying (3.31) such that

$$\begin{aligned} &\int_0^\tau \left(\sum_{\alpha=1}^{r} \left\langle \left(Y_s^{\varepsilon, h}\right)^{-1} V_\alpha\left(X_s^{\varepsilon, h}\right), l \right\rangle^2 \right. \\ &\left. \quad + \left\langle \left(Y_s^{\varepsilon, h}\right)^{-1} \left[\sum_{\beta=1}^{r} \dot{h}^\beta(s) V_\beta + \varepsilon^2 V_0, V_\alpha \right] \left(X_s^{\varepsilon, h}\right), l \right\rangle^2 \right) ds \ge c_9 n^{-1} \end{aligned} \tag{3.33}$$

on $\{\tau = n^{-1}\}$,

$\sum_{\alpha=0}^{r} |V_\alpha(X_t^{\varepsilon, h}) - V_\alpha(x)| \le 1$ and $\|I - Y^\varepsilon(t)\| < \frac{1}{4}$ for $t \in [0, \tau]$. Since

$$\begin{aligned} \left(Y_t^{\varepsilon, h}\right)^{-1} V_\alpha\left(X_t^{\varepsilon, h}\right) &= V_\alpha(x) + \sum_\beta \int_0^t \left(Y_s^{\varepsilon, h}\right)^{-1} \left[V_\beta, V_\alpha\right]\left(X_s^{\varepsilon, h}\right) \dot{h}^\beta(s)\, ds \\ &\quad + \varepsilon \sum_\beta \int_0^t \left(Y_s^{\varepsilon, h}\right)^{-1} \left[V_\beta, V_\alpha\right]\left(X_s^{\varepsilon, h}\right) \circ dw^\beta(s) \\ &\quad + \varepsilon^2 \int_0^t \left(Y_s^{\varepsilon, h}\right)^{-1} \left[V_0, V_\alpha\right]\left(X_s^{\varepsilon, h}\right) ds, \end{aligned}$$

we can apply the key lemma of [10] or [27] to conclude that, if $\varepsilon < \varepsilon_0$,

$$P\Bigg[\int_0^T \Big\langle \sum_{\alpha=1}^r (Y_s^{\varepsilon,h})^{-1} V_\alpha(X_s^{\varepsilon,h}), l \Big\rangle^2 ds < n^{-a_1},$$

$$\int_0^T \sum_{\alpha=1}^r \Big\langle (Y_s^{\varepsilon,h})^{-1}\Big[\sum_{\beta=1}^r \dot h^\beta(s) V_\beta + \varepsilon^2 V_0, V_\alpha\Big](X_s^{\varepsilon,h}), l \Big\rangle^2 ds \ge \frac{c_9}{(2n)}\Bigg]$$

$$\le a_2 \exp(-a_3 n^{a_4}) \quad \text{for some } a_1, a_2, a_3, a_4.$$

Combining this with (3.33), we can conclude that, except on a set of probability $\le c_1\exp(-c_2 n^{c_3}) + a_2\exp(-a_3 n^{a_4})$,

$$\int_0^T \Big(\sum_{\alpha=1}^r \big\langle (Y_s^{\varepsilon,h})^{-1} V_\alpha(X_s^{\varepsilon,h}), l \big\rangle^2 \Big) ds$$

is no less than $1/n^{a_1}$ or $c_9/(2n)$. Hence (3.32) is established. □

Bismut introduced in [5] the following assumption:

(H.2) for every $\lambda \in \mathbf{R}^r$, $\lambda \neq 0$,
$\dim \mathscr{L}\{V_1(x), \dots, V_r(x), [V_1, Y](x), \dots, [V_r, Y](x)\} = d$,
where $Y = \sum_{i=1}^r \lambda^i V_i$.

Corollary. *If* (H.2) *is satisfied at* x, *then* $\sigma^h(\varepsilon)$ *is uniformly nondegenerate provided that* $h \in C^2([0,1])$ *and* $\dot h(0) \neq 0$.

Let $x \neq y \in \mathbf{R}^d$ be given and assume that (H.2) is satisfied at x. Set

$$K_x^y = \{h \in H;\ \xi^{x,h}(1) = y\}. \tag{3.34}$$

Under the assumption (H.1′) everywhere, K_x^y is not empty. Let $\bar h$ be an element in K_x^y which minimizes the norm $\|h\|_H$: $\|\bar h\|_H = \min h \in K_x^y \|h\|_H$. As is easily seen, such an $\bar h$ always exists. By the Lagrange multiplier method, we can show that $\bar\lambda \in \mathbf{R}^d$ exists uniquely such that

$$\dot{\bar h}^\alpha(s) = \big\langle \bar\lambda, \Xi^{x,\bar h}(1)\Xi^{x,\bar h}(s)^{-1} V_\alpha(\xi^{x,\bar h}(s))\big\rangle, \qquad \alpha = 1, \dots, r, \quad \text{a.a. } s \in [0,1]. \tag{3.35}$$

From this it is easy to conclude that $t \to \bar h(t)$ is smooth and (3.35) holds for all $s \in [0,1]$.

Lemma 3.2. $\dot{\bar h}(0) \neq 0$.

Proof. We have

$$\Xi^{x,\bar h}(s)^{-1} V_\alpha(\xi^{x,\bar h}(s)) = V_\alpha(x) + \sum_{\beta=1}^r \int_0^s \Xi^{x,\bar h}(u)^{-1}[V_\beta, V_\alpha](\xi^{x,\bar h}(u))\dot{\bar h}^\beta(u)\, du.$$

Hence, if $\dot{\bar{h}}^\alpha(0) = \langle \lambda, \Xi^{x,\bar{h}}(1)V_\alpha(x)\rangle = 0$ for every $\alpha = 1,2,\ldots,r$, then

$$\dot{\bar{h}}^\alpha(s) = \sum_{\beta=1}^{r} \int_0^s \left\langle \lambda, \Xi^{x,\bar{h}}(1)\Xi^{x,\bar{h}}(u)^{-1}\left[V_\beta, V_\alpha\right]\left(\xi^{x,\bar{h}}(u)\right)\dot{\bar{h}}^\beta(u)\right\rangle du,$$

and hence it is easy to deduce from this that

$$\left|\dot{\bar{h}}(s)\right| \le \text{const}\int_0^s \left|\dot{\bar{h}}(u)\right| du.$$

This implies that $\dot{\bar{h}}(s) = 0$ and hence $\bar{h}(t) = 0$ but, since $x \neq y$, this is an obvious contradiction. □

Thus, from the above corollary, we have

THEOREM 3.7. *If* (H.2) *is satisfied at* x, *then*

$$F(\varepsilon, w) = \left(X^{\varepsilon,\bar{h}}(1, x, w) - y\right)/\varepsilon \in \mathbf{D}^\infty(\mathbf{R}^d)$$

is uniformly nondegenerate in the sense of (2.34).

Next, we introduce the following assumption.

(H.3) $\bar{h} \in K_x^y$ *which minimizes the H-norm in* K_x^y *is unique*.

Choose a C^∞-function $\phi(x)$ on $\mathbf{R}^1$ such that $0 \le \phi(x) \le 1$, $\phi(x) = 0$ if $|x| > 1$ and $\phi(x) = 1$ if $|x| \le \frac{1}{2}$. Set, for $\delta > 0$ and $\varepsilon > 0$,

$$\chi_\delta(\varepsilon, w) = \phi\left(\frac{1}{\delta}\int_0^1 \left|X^\varepsilon(s, x, w) - \xi^{x,\bar{h}}(s)\right|^2 ds\right), \tag{3.36}$$

where $X^\varepsilon(s, x, w)$ is the solution to SDE (3.9). It is easy to see that $\chi_\delta(\varepsilon, w) \in \mathbf{D}^\infty$ and

$$\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right) = \phi\left(\frac{1}{\delta}\int_0^1 \left|X^{\varepsilon,\bar{h}}(s, x, w) - \xi^{x,\bar{h}}(s)\right|^2 ds\right)$$

is also in $\mathbf{D}^\infty$ for every $\delta > 0$ and $\varepsilon > 0$. We know that

$$\int_0^1 \left|X^{\varepsilon,\bar{h}}(s, x, w) - \xi^{x,\bar{h}}(s)\right|^2 ds = O(\varepsilon^2)$$

in $\mathbf{D}^\infty$ as $\varepsilon \downarrow 0$ and it is easy to deduce from this that, for each $\delta > 0$,

$$\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right) = 1 + O(\varepsilon^{-n}) \tag{3.37}$$

in $\mathbf{D}^\infty$ as $\varepsilon \downarrow 0$ for every $n = 1,2,\ldots$.

LEMMA 3.3. *Suppose that* (H.3) *is satisfied. Then, for every* $\delta > 0$, $c > 0$ *exists such that the following estimate holds*:

$$\begin{aligned} 0 &\le E\left(\left(1 - \chi_\delta(\varepsilon, w)\right)\delta_y\left(X^\varepsilon(1, x, w)\right)\right) \\ &= O\left(\exp\left\{-\frac{1}{(2\varepsilon^2)}\left(\|\bar{h}\|_H^2 + c\right)\right\}\right) \quad \text{as } \varepsilon \downarrow 0. \end{aligned} \tag{3.38}$$

Proof. We have, for every $\delta' > 0$,

$$0 \le E\big((1-\chi_\delta(\varepsilon,w))\delta_y(X^\varepsilon(1,x,w))\big)$$
$$= E\left((1-\chi_\delta(\varepsilon,w))\phi\left(\frac{|X^\varepsilon(1,x,w)-y|^2}{\delta'^2}\right)\delta_y(X^\varepsilon(1,x,w))\right)$$

and, by an integration by parts on the Wiener space, it is easy to see that this is a finite sum of the form

$$\sum E\left(F_k(\varepsilon,w)\phi^{(l)}\left(\frac{|X^\varepsilon(1,x,w)-y|^2}{\delta'^2}\right)\right.$$
$$\left.\times(1-\phi)^{(m)}\left(\frac{1}{\delta}\int_0^1|X^\varepsilon(s,x,w)-\xi^{x,\bar h}(s)|^2\,ds\right)C_j(X^\varepsilon(1,x,w))\right),$$

where $F_k(\varepsilon,w)$ is a polynomial in components of $X^\varepsilon(1,x,w)$,

$$\frac{1}{\delta}\int_0^1|X^\varepsilon(s,x,w)-\xi^{x,\bar h}(s)|^2\,ds,$$

their derivatives and $\gamma(\varepsilon)=\sigma(\varepsilon)^{-1}$, $\sigma(\varepsilon)$ being the Malliavin covariance of $X^\varepsilon(1,x,w)$, and $C_j(x)$ is a bounded continuous function on R^d. By (3.7), we have $E(|F_k(s,w)|^p)^{1/p}=O(\varepsilon^{-l})$ for all $p>1$ with some $l>0$ and hence

$$E\big((1-\chi_\delta(\varepsilon,w))d_y(X^\varepsilon(1,x,w))\big)$$
$$\le K/\varepsilon^l P\left(\int_0^1|X^\varepsilon(s,x,w)-\xi^{x,\bar h}(s)|^2\,ds\ge\frac{\delta}{2},\ |X^\varepsilon(1,x,w)-y|\le\delta'\right)^{1/q},$$

where $1/p+1/q=1$. Therefore,

$$\limsup_{\varepsilon\downarrow 0}(2\varepsilon^2)\log E\big((1-\chi_\delta(\varepsilon,w))\delta_y(X^\varepsilon(1,x,w))\big)$$
$$\le \limsup_{\varepsilon\downarrow 0}(2\varepsilon^2)\log P\left(\int_0^1|X^\varepsilon(s,x,w)-\xi^{x,\bar h}(s)|^2\,dx \ge\frac{\delta}{2},\ |X^\varepsilon(1,x,w)-y|\le\delta'\right), \tag{3.39}$$

since q may be chosen arbitrarily close to 1.

We assume for a moment that $V_0=0$ to make the following argument simpler. Then, as remarked above, we may consider roughly $X^\varepsilon(s,x,w)=X(s,x,\varepsilon w)$, where $X(s,x,w)$ is the solution to SDE (3.3) and $X(s,x,h)=\xi^{x,h}(s)$ if $h\in H$. If it were true that $(s,w)\in[0,1]\times W\to X(s,x,w)\in R^d$ is continuous with respect to the uniform topology on W, a standard large deviation result for Wiener measure could be applied to yield that

(3.40) *The right-hand side of* (3.39) $\le -\inf\|h\|_H^2$,

where the infimum is taken over $h\in H$ such that

$$\int_0^1|\xi^{x,h}(s)-\xi^{x,\bar h}(s)|^2\,ds\ge\frac{\delta}{2},\qquad |\xi^{x,h}(1)-y|\le\delta'. \tag{3.41}$$

Of course, this is not true. However, we can fortunately justify (3.40), including the general case of $V_0 \neq 0$, by virtue of the approximation theorem for solutions of SDE by continuous functionals due to Azencott and Stroock [cf. Stroock [23], Lemma (4.8)]. Finally we note that $\delta' > 0$ exists such that the above $\inf\|h\|_H$ satisfies $\inf\|h\|_H > \|\bar{h}\|_H$. For, if otherwise, we may choose, for $\delta' = 1/n$, $h_n \in H$ satisfying (3.41) and $\limsup\|h_n\|_H \leq \|\bar{h}\|_H$. We can choose a subsequence of $\{h_n\}$ converging to $\tilde{h} \in H$ weakly. Then $\|\tilde{h}\|_H \leq \|\bar{h}\|_H$ and $\xi^{x,\tilde{h}}(1) = y$ because, as is easily shown in general, if $h_k \to h_\infty$ weakly in H then $\xi^{x,h_k}(s) \to \xi^{x,h_\infty}(s)$ uniformly in $s \in [0,1]$. This implies that $\tilde{h} \in K_x^y$ and, by (H.3), $\tilde{h} = \bar{h}$. But this contradicts $\int_0^1 |\xi^{x,\tilde{h}}(s) - \xi^{x,\bar{h}}(s)|^2 ds \geq \delta/2$. Now the proof of (3.38) is completed. □

Set

$$S^{\bar{h}}(\varepsilon, w) = f_2^{(\bar{h})} + \varepsilon f_3^{(\bar{h})} + \cdots = \frac{1}{\varepsilon^2}\left(X^{\varepsilon,\bar{h}}(1, x, w) - y - \varepsilon f_1^{(\bar{h})}\right). \tag{3.42}$$

LEMMA 3.4. *For every $M > 0$, $\delta > 0$ can be chosen such that*

$$\sup_{\varepsilon \in (0,1]} E\left(\exp\left\{M\left\langle \bar{\lambda}, S^{\bar{h}}(\varepsilon, w) - f_2^{(\bar{h})}\right\rangle\right\} I_{\{\int_0^1 |X^{\varepsilon,\bar{h}}(s,x,w) - \xi^{x,\bar{h}}(s)|^2 ds \leq \delta\}}\right) < \infty. \tag{3.43}$$

PROOF. In this proof, we write X_s^ε, ξ_s and h for $X^{\varepsilon,\bar{h}}(s,x,w)$, $\xi^{x,\bar{h}}(s)$ and $\bar{h}$, for simplicity. Also, we note that the following estimates can be obtained uniformly in $\varepsilon \in (0,1]$. Setting $\tilde{V}_0(x) = \frac{1}{2}\sum_{\alpha=1}^r \hat{V}_\alpha(V_\alpha)(x)$,

$$X_t^\varepsilon - \xi_t = \varepsilon\int_0^t V_\alpha(X_s^\varepsilon)\, dw_s^\alpha + \varepsilon^2\int_0^t \tilde{V}_0(X_s^\varepsilon)\, ds + \int_0^t [V_\alpha(X_s^\varepsilon) - V_\alpha(\xi_s)]\, \dot{h}^\alpha(s)\, ds$$

(we omit the summation sign). Hence, setting $\eta_t^\varepsilon = (X_t^\varepsilon - \xi_t)/\varepsilon$, we have

$$\eta_t^\varepsilon = \int_0^t V_\alpha(X_s^\varepsilon)\, dw_s^\alpha + \varepsilon\int_0^t \tilde{V}_0(X_s^\varepsilon)\, ds + \int_0^t \partial V_\alpha^\varepsilon(s)\, \eta_s^\varepsilon \dot{h}^\alpha(s)\, ds,$$

where $\partial V_\alpha^\varepsilon(s) = \int_0^1 \partial V_\alpha(\xi_s + u(X_s^\varepsilon - \xi_s))\, du$. Let η_t be determined by

$$\eta_t = \int_0^t V_\alpha(\xi_s)\, dw_s^\alpha + \int_0^t \partial V_\alpha(\xi_s)\, \eta_s \dot{h}^\alpha(s)\, ds.$$

Then this η_t is the same given by (3.28) with $h = \bar{h}$. Define a $d \times d$-matrix valued process $\Xi^\varepsilon(t)$ by

$$d\Xi_t^\varepsilon = \partial V_\alpha^\varepsilon(t)\Xi_t^\varepsilon \dot{h}^\alpha(t)\, dt, \qquad \Xi_0^\varepsilon = I.$$

Then η_t^ε is given by

$$\eta_t^\varepsilon = \Xi_t^\varepsilon \int_0^t (\Xi_s^\varepsilon)^{-1}\left[V_\alpha(X_s^\varepsilon)\, dw_s^\alpha + \varepsilon V_0(X_s^\varepsilon)\, du\right].$$

Note that Ξ_t^ε is bounded since $\partial V_\alpha^\varepsilon(t)$ is bounded. From this, it is easy to deduce that, for every $\delta > 0$, there exists $p_0 > 0$ such that

$$E\left(\exp\left\{p_0 \operatorname*{Max}_{0 \leq t \leq 1} |\eta_t^\varepsilon|^2\right\} I_{[\int_0^1 |X_u^\varepsilon - \xi_u|^2 du \leq \delta]}\right) < \infty. \tag{3.44}$$

Indeed, $\int_0^1 |X_u^\varepsilon - \xi_u|^2\,du \le \delta$ implies that $\int_0^1 |X_s^\varepsilon|^2\,du \le C$ for some $C > 0$. Noting that every component of $\int_0^t (\Xi_s^\varepsilon)^{-1} V_\alpha(X_s^\varepsilon)\,dw_s^\alpha$ is a Brownian motion time changed: $B(\phi_t)$ with $\phi_t \le C' \int_0^t |X_s^\varepsilon|^2\,ds \le C''$ for $0 \le t \le 1$, (3.44) is easily obtained.

Now

$$\eta_t^\varepsilon - \eta_t = \int_0^t [V_\alpha(X_s^\varepsilon) - V_\alpha(\xi_s)]\,dw_s^\alpha + \varepsilon \int_0^t \tilde{V}_0(X_s^\varepsilon)\,ds + \int_0^t \partial V_\alpha(\xi_s)[\eta_s^\varepsilon - \eta_s]\,\dot{h}^\alpha(s)\,ds + \tfrac{1}{2}\int_0^t \partial^2 V_\alpha^\varepsilon(s)\eta_s^\varepsilon \otimes (X_s^\varepsilon - \xi_s)\dot{h}^\alpha(s)\,ds,$$

where

$$\partial^2 V_\alpha^\varepsilon(s) = 2\int_0^1 du \int_0^u \partial^2 V_\alpha(\xi_s + v(X_s^\varepsilon - \xi_s))\,dv.$$

Hence, if Ξ_t is defined by (3.23) with $h = \bar{h}$, then

$$\eta_t^\varepsilon - \eta_t = \Xi_t \int_0^t \Xi_s^{-1}\{(V_\alpha(X_s^\varepsilon) - V_\alpha(\xi_s))\,dw_s^\alpha + \varepsilon[\tilde{V}_0(X_s^\varepsilon) - \tilde{V}_0(\xi_s)]\,ds + \varepsilon\tilde{V}_0(\xi_s)\,ds + \tfrac{1}{2}\partial^2 V_\alpha^\varepsilon(s)\,\eta_s^\varepsilon \otimes (X_s^\varepsilon - \xi_s)\dot{h}^\alpha(s)\,ds\}. \tag{3.45}$$

From this, it is easy to conclude [from (3.44) and the fact that a component of $\int_0^t \Xi_s^{-1}(V_\alpha(X_s^\varepsilon) - V_\alpha(\xi_s))\,dw_s^\alpha$ is a Brownian motion time changed: $B(\psi_t)$ with $\psi_t \le C\int_0^t |X_s^\varepsilon - \xi_s|^2\,ds$] that, for every $p > 0$, we can find $\delta_0 > 0$ such that

$$E\left[\exp\left\{p \operatorname*{Max}_{0\le t\le 1} |\eta_t^\varepsilon - \eta_t|^2\right\} I_{[\int_0^1 |X_u^\varepsilon - \xi_u|^2\,du \le \delta_0]}\right] < \infty. \tag{3.46}$$

Let $\theta_t^\varepsilon = (\eta_t^\varepsilon - \eta_t)/\varepsilon$ and θ_t be determined by

$$\theta_t = \int_0^t \partial V_\alpha(\xi_s)\eta_s\,dw_s^\alpha + \int_0^t \tilde{V}_0(\xi_s)\,ds + \int_0^t \partial V_\alpha(\xi_s)\theta_s \dot{h}_s^\alpha\,ds + \tfrac{1}{2}\int_0^t \partial^2 V_\alpha(\xi_s)\eta_s \otimes \eta_s\,ds.$$

Then,

$$\theta_t^\varepsilon - \theta_t = \int_0^t \partial V_\alpha(\xi_s)(\eta_s^\varepsilon - \eta_s)\,dw_s^\alpha + \tfrac{1}{2}\int_0^t \partial^2 V_\alpha^\varepsilon(s)\eta_s^\varepsilon \otimes (X_s^\varepsilon - \xi_s)\,dw_s^\alpha + \int_0^t (\tilde{V}_0(X_s^\varepsilon) - \tilde{V}_0(\xi_s))\,ds + \int_0^t \partial V_\alpha(\xi_s)(\theta_s^\varepsilon - \theta_s)\dot{h}_s^\alpha\,ds + \tfrac{1}{2}\int_0^t [\partial^2 V_\alpha^\varepsilon(s)\eta_s^\varepsilon \otimes \eta_s^\varepsilon - \partial^2 V_\alpha(\xi_s)\eta_s \otimes \eta_s]\dot{h}_s^\alpha\,ds.$$

Hence

$$\begin{aligned}\langle \theta_1^\varepsilon - \theta_1, \bar{\lambda}\rangle = \Big\langle {}^t\Xi_1\bar{\lambda}, \int_0^1 \Xi_s^{-1}\{&\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon - \eta_s)\,dw_s^\alpha \\ &+ \tfrac{1}{2}\partial^2 V_\alpha^\varepsilon(s)\,\eta_s^\varepsilon \otimes (X_s^\varepsilon - \xi_s)\,dw_s^\alpha \\ &+ [\tilde{V}_0(X_s^\varepsilon) - \tilde{V}_0(\xi_s)]\,ds \\ &+ \tfrac{1}{2}[\partial^2 V_\alpha^\varepsilon(s)\,\eta_s^\varepsilon \otimes \eta_s^\varepsilon - \partial^2 V_\alpha(\xi_s)\,\eta_s \otimes \eta_s]\dot{h}_s^\alpha\,ds\}\Big\rangle \\ = I_1 + I_2 + I_3 + I_4.\end{aligned}$$

If we can show that, for every $p > 0$, $\delta_1 > 0$ exists such that

$$(3.47)\qquad E\Big(\exp(pI_i)I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_1]}\Big) < \infty, \qquad i = 1,2,3,4,$$

this implies that, for every $p > 0$, $\delta_2 > 0$ exists such that

$$(3.48)\qquad E\Big(\exp(p\langle\theta_1^\varepsilon-\theta_1,\bar\lambda\rangle)I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_2]}\Big) < \infty,$$

and then the proof of (3.43) will be complete since $\theta_1^\varepsilon - \theta_1 = S^h(\varepsilon, w) - f_2^{(h)}$. Therefore, it remains only to show (3.47). As for I_1, setting $\mu = {}^t\Xi_1\bar\lambda$,

$$E\left(\exp\left\{p\left\langle\mu,\int_0^1\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\,dw_s^\alpha\right\rangle\right\}I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_1]}\right)$$

$$= E\left(\exp\left\{p\left\langle\mu,\int_0^1\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\,dw_s^\alpha\right\rangle\right.\right.$$

$$\left.-p^2\sum_\alpha\int_0^1\left\langle\mu,\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\right\rangle^2 ds\right\}$$

$$\left.\times\exp\left\{p^2\sum_\alpha\int_0^1\left\langle\mu,\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\right\rangle^2 ds\right\}I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_1]}\right)$$

$$\le E\left(\exp\left\{2p\left\langle\mu,\int_0^1\Xi_s^{-1}\,\partial v_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\,dw_s^\alpha\right\rangle\right.\right.$$

$$\left.\left.-\frac{(2p)^2}{2}\sum_\alpha\int_0^1\left\langle\mu,\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\right\rangle^2 ds\right\}\right)^{1/2}$$

$$\times E\left(\exp\left\{2p^2\sum_\alpha\int_0^1\left\langle\mu,\Xi_s^{-1}\,\partial V_\alpha(\xi_s)\,(\eta_s^\varepsilon-\eta_s)\right\rangle^2 ds\right\}I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_1]}\right)^{1/2}$$

$$\le E\left(\exp\left\{\text{const.}\,p^2\int_0^1|\eta_s^\varepsilon-\eta_s|^2\,ds\right\}I_{[\int_0^1|X_u^\varepsilon-\xi_u|^2\,du\le\delta_1]}\right)^{1/2} < \infty$$

for some δ_1 by (3.46).

The estimates for $i = 2,3,4$ are more easily obtained and details are omitted. □

Finally, we introduce the following condition:

(H.4)

$$(3.49)\qquad E\Big(\exp\langle\bar\lambda, f_2^{(\bar h)}\rangle\,\Big|\,f_1^{(\bar h)} = 0\Big) < \infty.$$

REMARK 3.1. Since $f_1^{(\bar h)} = \eta_1$ is a nondegenerate Gaussian random variable and $f_2^{(\bar h)} = \theta_1$ is a nice quadratic functional of a Gaussian process (η_t, w_t), (3.49) is equivalent to

$$\text{(3.49}')\quad \textit{there exists } p,\ 1 < p < \infty, \textit{ such that } \quad E\Big(\exp\big[p\langle\bar{\lambda}, f_2^{(\bar{h})}\rangle\big]\Big| f_1^{(\bar{h})} = 0\Big) < \infty$$

or

$$\text{(3.49}'')\quad \textit{there exists } p,\ 1 < p < \infty, \textit{ such that } \quad E\Big(\exp\big[p\langle\bar{\lambda}, f_2^{(\bar{h})}\rangle\big] I_{[|f_1^{(\bar{h})}|<\delta]}\Big) < \infty \quad \textit{for any } \delta > 0,\ \textit{cf.}\ [4].$$

LEMMA 3.5. *Suppose that* (H.3) *and* (H.4) *are satisfied. Then there exist* p_1, $1 < p_1 < \infty$, *and* $\delta > 0$ *such that*

$$\text{(3.50)}\quad \sup_{0<\varepsilon\le 1} E\Big[\exp\{p_1\langle\bar{\lambda}, S^{\bar{h}}(\varepsilon, w)\rangle\} I_{[\int_0^1 |X^{\varepsilon,\bar{h}}(s,x,w) - \xi^{x,\bar{h}}(s)|^2\, ds < \delta]} \times I_{[|f_1^{(h)} + \varepsilon S^{\bar{h}}(\varepsilon, w)| < \delta_1]}\Big] < \infty \quad \textit{for any } \delta_1 > 0.$$

PROOF. We use the same notation as in the proof of Lemma 3.4. First, we remark that the conditional expectation in (3.49′) coincides with the unconditional expectation

$$\text{(3.51)}\qquad E(\exp[p\langle\bar{\lambda}, \tilde{f}_2\rangle]) < \infty,$$

where $\tilde{f}_2$ is obtained from $f_2^{(\bar{h})}$ by replacing η_t and w_t by $\eta_t - A_1(t)\eta_1$ and $w_t - A_2(t)\eta_1$, respectively. Here $A_1(t)$ and $A_2(t)$ are given by $A_1(t) = E[\eta_t \otimes \eta_1]E[\eta_1 \otimes \eta_1]^{-1}$ and $A_2(t) = E[w_t \otimes \eta_1]E[\eta_1 \otimes \eta_1]^{-1}$. Thus,

$$\tilde{f}_2 = f_2^{(\bar{h})} + R$$

and R is given in the form $R = R_1(\eta, w)\eta_1 + R_2(\eta_1)\eta_1$, where $R_1(\eta, w) \in \mathbf{R}^d \otimes \mathbf{R}^d$, every component of which is a linear combination of components of $\Xi_1\int_0^1 \Xi_s^{-1}\, \partial^2 V_\alpha(\xi_s)\, \eta_s \otimes A_1(s)\dot{h}_s^\alpha\, ds$, $\Xi_1\int_0^1 \Xi_s^{-1}\, \partial V_\alpha(\xi_s)\, \eta_s A_2(s)\, ds$ and $\Xi_1\int_0^1 \Xi_s^{-1}\, \partial V_\alpha(\xi_s) A_1(s)\, dw_s^\alpha$; $R_2(\eta_1) \in \mathbf{R}^d \otimes \mathbf{R}^d$, every component of which is a linear combination of components of η_1. Next, we note that it is easy to deduce from (3.46) that, for every $M > 0$, there exists $\delta' > 0$ such that

$$\text{(3.52)}\quad \sup_{0<\varepsilon\le 1} E\Big(\exp\Big[M|\eta_1 - \eta_1^\varepsilon| \operatorname*{Max}_{0\le s\le 1}(|\eta_s| + |w_s|)\Big] I_{[\int_0^1 |X_u^\varepsilon - \xi_u|^2\, du < \delta']}\Big) < \infty.$$

Indeed, choosing $\rho > 0$ such that

$$E\Big(\exp\Big[\rho \operatorname*{Max}_{0\le s\le 1}(|\eta_s| + |w_s|)^2\Big]\Big) < \infty$$

and noting that

$$M|\eta_1 - \eta_1^\varepsilon| \operatorname*{Max}_{0\le s\le 1}(|\eta_s| + |w_s|) \le \frac{\rho}{2}\operatorname*{Max}_{0\le s\le 1}(|\eta_s| + |w_s|)^2 + \frac{M^2}{2\rho}|\eta_1 - \eta_1^\varepsilon|^2,$$

(3.52) follows from (3.46).

By (3.52), it is clear that, for given $M > 0$, $\delta'' > 0$ can be chosen so that

$$\text{(3.53)}\qquad \sup_\varepsilon E\Big(\exp[M\langle\lambda, R - R_\varepsilon\rangle] I_{[\int_0^1 |X_u^\varepsilon - \xi_u|^2\, du < \delta'']}\Big) < \infty,$$

where

$$R_\varepsilon = R_1(\eta, w)\eta_1^\varepsilon + R_2(\eta_1)\eta_1^\varepsilon.$$

Now (3.51) combined with (3.43) and (3.53) implies that, for any p', $1 < p' < p$, we can choose $\delta > 0$ such that

$$\sup_\varepsilon E\Big(\exp\big[p'\langle\bar\lambda, S^{\bar h}(\varepsilon, w) + R_\varepsilon\rangle\big] I_{[\int_0^1 |X_u^\varepsilon - \xi_u|^2 du < \delta]}\Big) < \infty.$$

The assertion of the lemma follows immediately from this since

$$E\Big(\exp\big[N\langle\bar\lambda, R_\varepsilon\rangle\big] I_{[|\eta_1^\varepsilon| < \delta_1]}\Big) < \infty$$

for every $N > 0$ and $\delta_1 > 0$. □

LEMMA 3.6. *Suppose that* (H.3) *and* (H.4) *are satisfied. Then there exists* $\delta > 0$ *such that, for every* $\delta_1 > 0$,

$$F(\varepsilon, w) := \exp\big(\langle\bar\lambda, S^{\bar h}(\varepsilon, w)\rangle\big)\chi_\delta\Big(\varepsilon, w + \frac{\bar h}{\varepsilon}\Big)\phi\Big(\frac{1}{\delta_1^2}\big| f_1^{(\bar h)} + \varepsilon S^{\bar h}(\varepsilon, w)\big|^2\Big) \in \tilde{\mathbf{D}}^\infty$$

for all $\varepsilon \in (0,1]$ *and, for every* $k = 1, 2, \ldots,$

$$(3.54)\quad F(\varepsilon, w) = \exp\big(\langle\bar\lambda, f_2^{(\bar h)}\rangle\big)\phi\Big(\frac{1}{\delta_1^2}|f_1^{(\bar h)}|^2\Big)\big(1 + \varepsilon\gamma_1 + \cdots + \varepsilon^k\gamma_k\big) + F_k(\varepsilon, w),$$

where $F_k(\varepsilon, w)$ *satisfies*

$$(3.55)\qquad F_k(\varepsilon, w)T\big(f_1^{(\bar h)}\big) = O(\varepsilon^{k+1}) \quad \textit{in } \mathbf{D}^{-\infty} \textit{ as } \varepsilon \downarrow 0$$

for any $T \in \mathscr{S}'(\mathbf{R}^d)$ *having its support in* $\{|x| < \delta_1/2\}$. *Furthermore,* $\gamma_i \in \mathbf{D}^\infty$, $i = 1, 2, \ldots,$ *are obtained by the formal expansion*

$$(3.56)\qquad \begin{aligned}&\sum_{n=0}^\infty \frac{1}{n!}\Big(\varepsilon\langle\lambda, f_3^{(\bar h)}\rangle + \varepsilon^2\langle\lambda, f_4^{(\bar h)}\rangle + \cdots\Big)^n\\ &\quad = 1 + \varepsilon\gamma_1 + \varepsilon^2\gamma_2 + \cdots.\end{aligned}$$

PROOF. The proof is similar to that of Theorem 2.4 once we take into account Lemma 3.5, (3.37) and the following fact:

$$\begin{aligned}&\phi\Big(\delta_1^{-2}\big| f_1^{(h)} + \varepsilon S^h(\varepsilon, w)\big|^2\Big)\\ &\quad = \phi\big(\delta_1^{-2}|f_1^{(h)}|^2\big) + \sum_{m=1}^k \phi^{(m)}\big(\delta_1^{-2}|f_1^{(h)}|^2\big)G_m(\varepsilon, w) + O(\varepsilon^{k+1}) \quad \text{in } \mathbf{D}^\infty\end{aligned}$$

for some $G^{(m)} \in \mathbf{D}^\infty$ and $\phi^{(m)}(\delta_1^{-2}|f_1^{(h)}|^2)T(f_1^{(h)}) = 0$ for $m \geq 1$ if T is supported in $\{x < \delta_1/2\}$. □

All the above preparations completed, we are now going to discuss the short-time asymptotic expansion of the fundamental solution $p(t, x, y)$. We assume (H.1′) everywhere (this is to guarantee that $K_x^y \neq \phi$) and (H.2), (H.3) and (H.4) for x, $y \in \mathbf{R}^d$, $x \neq y$. Let $\bar h$ be the unique H-norm minimizing element in K_x^y. Choose $\delta > 0$ as in Lemma 3.6. Then, we have

$$
\begin{aligned}
p(\varepsilon^2, x, y) &= E\left[\delta_y(X^\varepsilon(1,x,w))\right] \\
&= E\left[\delta_y(X^\varepsilon(1,x,w))\chi_\delta(\varepsilon,w)\right] + E\left[\delta_y(X^\varepsilon(1,x,w))(1-\chi_\delta(\varepsilon,w))\right] \qquad (3.57)\\
&= I_1 + I_2
\end{aligned}
$$

and, by (3.38),

$$I_2 = O\left(\exp\left(-\frac{1}{2\varepsilon^2}\left(\|\bar{h}\|_H^2 + c\right)\right)\right) \quad \text{as } \varepsilon \downarrow 0 \text{ for some } c > 0. \tag{3.58}$$

By the Cameron–Martin theorem,

$$I_1 = E\left[\exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2} - \frac{1}{\varepsilon}(\bar{h}, w)_H\right)\delta_y(X^{\varepsilon,\bar{h}}(1,x,w))\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\right], \tag{3.59}$$

where

$$(\bar{h}, w)_H = \sum_{\alpha=1}^r \int_0^1 \frac{d\bar{h}^\alpha}{dt}(s)\, dw^\alpha(s).$$

Noting (3.26) and (3.35), we have

$$(\bar{h}, w)_H = \langle \bar{\lambda}, f_1^{(\bar{h})}\rangle, \tag{3.60}$$

$\langle\ ,\ \rangle$ being the inner product in $\mathbf{R}^d$. Hence,

$$
\begin{aligned}
I_1 &= \exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2}\right)E\left[\exp\left(-\frac{1}{\varepsilon}\langle\bar{\lambda}, f_1^{(\bar{h})}\rangle\right)\right.\\
&\qquad\left.\times\delta_y\left(y + \varepsilon f_1^{(\bar{h})} + \varepsilon^2 S^{\bar{h}}(\varepsilon, w)\right)\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\right]\\
&= \exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2}\right)E\left[\exp\left(-\frac{1}{\varepsilon}\langle\bar{\lambda}, f_1^{(\bar{h})}\rangle\right)\delta_0\left(\varepsilon f_1^{(\bar{h})} + \varepsilon^2 S^{\bar{h}}(\varepsilon,w)\right)\right.\\
&\qquad\left.\times\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\right]\\
&= \varepsilon^{-d}\exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2}\right)E\left[\exp\left(-\frac{1}{\varepsilon}\langle\bar{\lambda}, f_1^{(\bar{h})}\rangle\right)\delta_0\left(f_1^{(\bar{h})} + \varepsilon S^{\bar{h}}(\varepsilon, w)\right)\right. \qquad (3.61)\\
&\qquad\left.\times\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\right]\\
&= \varepsilon^{-d}\exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2}\right)E\left[\exp\left(\langle\bar{\lambda}, S^{\bar{h}}(\varepsilon, w)\rangle\right)\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\right.\\
&\qquad\left.\times\delta_0\left(f_1^{(\bar{h})} + \varepsilon S^{\bar{h}}(\varepsilon, w)\right)\right]\\
&= \varepsilon^{-d}\exp\left(-\frac{\|\bar{h}\|_H^2}{2\varepsilon^2}\right)E\left(F(\varepsilon, w)\delta_0\left(f_1^{(h)} + \varepsilon S^h(\varepsilon, w)\right)\right).
\end{aligned}
$$

By Lemma 3.6, we know that $F(\varepsilon, w)$ satisfies (3.54) and (3.55) for every k. Also, we know by Theorem 3.7 that $f_1^{(\bar{h})} + \varepsilon S^{\bar{h}}(\varepsilon, w) = (X^{\varepsilon, \bar{h}}(1, x, w) - y)/\varepsilon$ is uniformly nondegenerate and hence, by Theorem 2.3,

$$\delta_0\big(f_1^{(\bar{h})} + \varepsilon S^{\bar{h}}(\varepsilon, w)\big) \sim \Phi_0 + \varepsilon\Phi_1 + \cdots \quad \text{in } \tilde{\mathbf{D}}^{-\infty} \text{ as } \varepsilon \downarrow 0 \tag{3.62}$$

and $\Phi_i \in \tilde{\mathbf{D}}^{-\infty}$ are determined explicitly in terms of $f_i^{(\bar{h})}$ (see the proof of Lemma 3.4 for how these $f_l^{(\bar{h})}$ are obtained successively). Hence

$$\begin{aligned} &\exp\big(\big\langle\bar{\lambda}, S^{\bar{h}}(\varepsilon, w)\big\rangle\big)\chi_\delta\left(\varepsilon, w + \frac{\bar{h}}{\varepsilon}\right)\delta_0\big(f_1^{(\bar{h})} + \varepsilon S^{\bar{h}}(\varepsilon, w)\big) \\ &\qquad \sim \Psi_0 + \varepsilon\Psi_1 + \cdots \quad \text{in } \mathbf{D}^{-\infty} \text{ as } \varepsilon \downarrow 0, \end{aligned} \tag{3.63}$$

and this expansion is obtained explicitly by formally multiplying the following (3.64) and (3.62):

$$\exp\big(\big\langle\bar{\lambda}, f_1^{(\bar{h})}\big\rangle\big)\big(1 + \varepsilon\gamma_1 + \varepsilon^2\gamma_2 + \cdots\big) \quad \textit{with } \gamma_i \textit{ determined by } (3.56). \tag{3.64}$$

Therefore, we have obtained the following theorem:

THEOREM 3.8. *Under the assumptions* (H.1′) *everywhere and* (H.2), (H.3), (H.4) *for* $x, y \in \mathbf{R}^d$, $x \neq y$, *the fundamental solution* $p(t, x, y)$ *of* $\partial/\partial t = A$ *has the asymptotic expansion*

$$p(t, x, y) \sim \exp\left(-\frac{\|\bar{h}\|_H^2}{2t}\right)t^{-d/2}\big(c_0 + c_1 t + \cdots\big) \quad \textit{as } t \downarrow 0 \tag{3.65}$$

with

$$c_i = E(\Psi_{2i}). \tag{3.66}$$

[*We can show as in Theorem* 3.5 *that* $E(\Psi_i) = 0$ *for odd* i.] *In particular,*

$$c_0 = E\Big[\exp\big(\big\langle\bar{\lambda}, f_2^{(\bar{h})}\big\rangle\big)\delta_0\big(f_1^{(\bar{h})}\big)\Big]. \tag{3.67}$$

REMARK 3.2. It is easy to obtain the asymptotic expansion $E[J(\varepsilon, w)\,\delta_y(X^\varepsilon(1, x, w)]$ if $J(\varepsilon, w) \in \mathbf{D}^\infty$ and $J(\varepsilon, w + \bar{h}/\varepsilon)$ has the asymptotic expansion in $\mathbf{D}^\infty$ as $\varepsilon \downarrow 0$. In this case, (3.63) is multiplied by $J(\varepsilon, w + \bar{h}/\varepsilon)$ which has the asymptotic expansion in $\mathbf{D}^{-\infty}$ by Theorem 2.2.

In the elliptic case, i.e., $\dim \mathscr{L}\{V_1, V_2, \ldots, V_r\} = d$ everywhere, $(A^{ij}(x))$ in (3.1) is strictly positive definite and its inverse $(A_{ij}(x))$ induces a Riemannian structure. In this case, $\xi^{x, \bar{h}}$ is the minimal geodesic connecting x and y and (H.4) just corresponds to the condition that x and y are not conjugate along $\xi^{x, \bar{h}}$. Cf. [5] and [20].

REMARK 3.3. The above method can be extended to the case of $E(\exp\{G(\varepsilon, w)/\varepsilon^2\}J(\varepsilon, w)\delta_y(X^\varepsilon(1, x, w)))$ if $G(\varepsilon, w)$ is a nice Itô functional such that

$$G\left(\varepsilon, w + \frac{h}{\varepsilon}\right) \sim \Gamma_0(h) + \varepsilon\Gamma_1(h, w) + \cdots, \qquad h \in H, \quad \text{in } \mathbf{D}^\infty \text{ as } \varepsilon \downarrow 0.$$

In this case, $\bar{h}$ is replaced by the minimizing element in K_x^y of $\|h\|_H^2/2 - \Gamma_0(h)$.

EXAMPLE 3.1 (The Heisenberg group). Let $d = 3$, $r = 2$ and V_i, $i = 0, 1, 2$, be given by ($x = (x_1, x_2, x_3) \in \mathbf{R}^3$)

$$V_0 = 0, \qquad V_1 = \frac{\partial}{\partial x_1} + 2x_2 \frac{\partial}{\partial x_3}, \qquad V_2 = \frac{\partial}{\partial x_2} - 2x_1 \frac{\partial}{\partial x_3}.$$

Since $[V_1, V_2] = -4\,\partial/\partial x_3$, (H.1′) and (H.2) are satisfied everywhere. Let $0 = (0,0,0)$ be the origin of $\mathbf{R}^3$ and we consider the asymptotic properties of $p(t, 0, x)$, $x \neq 0$, as $t \downarrow 0$. First of all, we introduce the following notation: For $w(t) = (w_1(t), w_2(t)) \in W_0^2$ and $h(t) = (h_1(t), h_2(t)) \in H$,

$$(h, w)_H = \sum_{i=1}^{2} \int_0^1 \dot{h}_i(t)\, dw_i(t), \tag{3.68}$$

$$\begin{aligned} h^{\perp}(t) &= (h_2(t), -h_1(t)), \\ I(h)(t) &= \left(\int_0^t h_1(s)\, ds, \int_0^t h_2(s)\, ds \right), \end{aligned} \tag{3.69}$$

$$S(t, w) = \int_0^t w_2(s)\, dw_1(s) - w_1(s)\, dw_2(s) \tag{3.70}$$

(Lévy's stochastic area integral).

If $x_1^2 + x_2^2 \neq 0$, $\bar{h} \in K_0^x$ minimizing the H-norm is unique and given by

$$\begin{aligned} \bar{h}_1(t) &= \alpha(1 - \cos(2\sigma t)) + \beta \sin(2\sigma t), \\ \bar{h}_2(t) &= \alpha \sin(2\sigma t) - \beta(1 - \cos(2\sigma t)) \quad \text{if } \sigma \neq 0 \end{aligned} \tag{3.71}$$

and

$$\bar{h}_1(t) = \alpha t, \qquad \bar{h}_2(t) = \beta t \quad \text{if } \sigma = 0. \tag{3.72}$$

α, $\beta \in R$ and $|\sigma| < \pi$ are uniquely determined by the condition

$$\bar{h}_1(1) = x_1, \qquad \bar{h}_2(1) = x_2 \quad \text{and} \quad 2\langle \bar{h}, I(\bar{h}^{\perp}) \rangle_H = x_3. \tag{3.73}$$

In particular, $\sigma = 0$ if and only if $x_3 = 0$.

Now $X^{\varepsilon, \bar{h}}(1, w) = (X_i^{\varepsilon, \bar{h}}(1, w))_{i=1}^3$ are given by

$$\begin{aligned} X_1^{\varepsilon, \bar{h}}(1, w) &= x_1 + \varepsilon w_1(1), \\ X_2^{\varepsilon, \bar{h}}(1, w) &= x_2 + \varepsilon w_2(1), \\ X_3^{\varepsilon, \bar{h}}(1, w) &= x_3 + \varepsilon\left[4(I(\bar{h}^{\perp}), w)_H + 2(x_1 w_2(1) - x_2 w_1(1))\right] \\ &\quad + 2\varepsilon^2 S(1, w). \end{aligned} \tag{3.74}$$

Hence

$$f_1^{(\bar{h})} = \left(w_1(1), w_2(1), 4(I(\bar{h}^{\perp}), w)_H + 2(x_1 w_2(1) - x_2 w_1(1))\right) \tag{3.75}$$

and

$$f_2^{(\bar{h})} = (0, 0, 2S(1, w)). \tag{3.76}$$

It is easy to see

$$\langle \bar{\lambda}, f_2^{(\bar{h})} \rangle = \sigma \cdot S(1, w). \tag{3.77}$$

It is well known that

$$E\Big(\exp\big\{\big\langle\bar\lambda,\, f_2^{(\bar h)}\big\rangle\big\}\,\Big|\,w_1(1)=0,\ w_2(1)=0\Big)=\sigma/\sin\sigma<\infty,$$

since $|\sigma|<\pi$ and, a fortiori,

$$E\Big(\exp\big\{\big\langle\lambda,\, f_2^{(\bar h)}\big\rangle\big\}\,\Big|\,f_1^{(h)}=0\Big)<\infty.$$

Hence (H.4) is also satisfied and

$$\begin{aligned}E\Big(\exp\big\{\big\langle\bar\lambda,\, f_2^{(\bar h)}\big\rangle\big\}\delta_0\big(f_1^{(\bar h)}\big)\Big) &= (2\pi)^{-3/2}(\det C)^{-1/2}\\ &\quad\times E\Big(\exp\big\{\big\langle\lambda,\, f_2^{(\bar h)}\big\rangle\big\}\,\Big|\,f_1^{(\bar h)}=0\Big),\end{aligned}$$

where C is the covariance matrix (= the Malliavin covariance) of $f_1^{(\bar h)}$. This expectation can be computed by a standard technique to obtain

$$\begin{aligned}p(\varepsilon^2,0,x) &\sim (2\pi\varepsilon^2)^{-3/2}\exp\left\{-(2\varepsilon^2)^{-1}\big(x_1^2+x_2^2\big)\left(\frac{\sigma}{\sin\sigma}\right)^2\right\}\\ &\quad\times\frac{\sigma}{2}\left(\frac{\sin\sigma}{\sin\sigma-\sigma\cos\sigma}\right)^{1/2}\big(x_1^2+x_2^2\big)^{-1/2}.\end{aligned}\tag{3.78}$$

Cf. Bismut [5], Azencott [2] and Gaveau [7].

Next, consider the case $x=(0,0,x_3)$ with $x_3\neq 0$ and assume $x_3>0$ for simplicity. In this case, (H.3) is no longer satisfied and the set of $h\in K_0^x$ minimizing the H-norm is a one-dimensional submanifold $\{h^\theta,\theta\in[0,2\pi)\}$ of H given by

$$\begin{aligned}h_1^\theta(t) &= r\sin\theta(1-\cos 2\pi t)+r\cos\theta\sin 2\pi t,\\ h_2^\theta(t) &= r\sin\theta\sin 2\pi t-r\cos\theta(1-\cos 2\pi t),\end{aligned}\tag{3.79}$$

where

$$r=\big\{(4\pi)^{-1}x_3\big\}^{1/2}.\tag{3.80}$$

Hence, we cannot apply our result obtained above. We can, however, proceed as follows. First, we note the following general formula:

$$\int_0^\pi \delta_0(\xi\sin\theta-\eta\cos\theta)|\xi\cos\theta+\eta\sin\theta|\,d\theta=1$$

if $\xi,\eta\in R$, $(\xi,\eta)\neq(0,0)$ and δ_0 is the Dirac delta function on R^1. From this, we have

$$\begin{aligned}p\big(\varepsilon^2,0,(0,0,x_3)\big) &= E\Big[\delta_{(0,0,x_3)}\big(\varepsilon w_1(1),\varepsilon w_2(1),2\varepsilon^2 S(1,w)\big)\Big]\\ &= \int_0^\pi E\Big[\delta_{(0,0,x_3)}\big(\varepsilon w_1(1),\varepsilon w_2(1),2\varepsilon^2 S(1,w)\big)\\ &\qquad\times\delta_0\big(\varepsilon(h^{\theta-\pi/2},w)_H\big)\big|\varepsilon(h^\theta,w)_H\big|\Big]\,d\theta.\end{aligned}$$

If we apply the Cameron–Martin theorem to the translation $w \to w + h^\theta/\varepsilon$, the above is equal to

$$\int_0^\pi \exp\left\{-\frac{\|h^\theta\|_H^2}{2\varepsilon^2}\right\} E\Bigg[\exp\left\{-\frac{(h^\theta, w)_H}{\varepsilon}\right\} \times \delta_{(0,0,x_3)}\left(\varepsilon w_1(1), \varepsilon w_2(1), x_3 + \frac{2\varepsilon}{\pi}(h^\theta, w)_H + 2\varepsilon^2 S(1,w)\right) \times \delta_0\left(\langle h^{\theta-\pi/2}, h^\theta\rangle_H + \varepsilon(h^{\theta-\pi/2}, w)_H\right) \times \left| \|h^\theta\|_H^2 + \varepsilon(h^\theta, w)_H\right| \Bigg] d\theta.$$

Noting that $\langle h^\theta, h^{\theta'}\rangle_H = 4\pi^2 r^2 \cos(\theta - \theta')$, the above is equal to

$$\begin{aligned}
&\exp\left\{-\frac{4\pi^2 r^2}{(2\varepsilon^2)}\right\} \int_0^\pi E\Bigg[\exp\left\{-\frac{(h^\theta, w)_H}{\varepsilon}\right\} \\
&\qquad \times \delta_{(0,0,0)}\left(\varepsilon w_1(1), \varepsilon w_2(1), \frac{2\varepsilon}{\pi}(h^\theta, w)_H + 2\varepsilon^2 S(1,w)\right) \\
&\qquad \times \delta_0\left(\varepsilon(h^{\theta-\pi/2}, w)_H\right)\left|4\pi^2 r^2 + \varepsilon(h^\theta, w)_H\right| \Bigg] d\theta \\
&= \varepsilon^{-4} 2\pi^3 r^2 \exp\left\{-\frac{\pi x_3}{(2\varepsilon^2)}\right\} \\
&\quad \times \int_0^\pi E\Big[\exp\{\pi S(1,w)\}\delta_{(0,0,0,0)}\big(w_1(1), w_2(1), (h^\theta, w)_H \\
&\qquad\qquad + \varepsilon\pi S(1,w), (h^{\theta-\pi/2}, w)_H\big) \\
&\qquad \times \left|1 + \varepsilon(4\pi^2 r^2)^{-1}(h^\theta, w)_H\right|\Big]\, d\theta \\
&\sim \varepsilon^{-4} 2\pi^3 r^2 \exp\left\{-\frac{\pi x_3}{(2\varepsilon^2)}\right\} \\
&\quad \times \int_0^\pi E\left[\exp\{\pi S(1,w)\}\delta_{(0,0,0,0)}\big(w_1(1), w_2(1), (h^\theta, w)_H, (h^{\theta-\pi/2}, w)_H\big)\right] d\theta.
\end{aligned}$$

This (generalized) expectation can be easily computed to obtain

$$p(t, 0, x) \sim (8t^2)^{-1} \exp\left\{-\frac{\pi x_3}{(2t)}\right\};$$

cf. Gaveau [7] and Azencott [2].

We note that the above generalized expectations are well defined: Generally, the coupling of a positive generalized Wiener functional and a Wiener functional, which has a continuous version on W and satisfies a certain growth condition, is well defined.

REMARK 3.4. Our probabilistic methods for short-time asymptotics of heat kernels can also be applied to heat equations for geometrical objects. In particular, it provides a probabilistic approach to discuss heat equations on differential forms which are related to several important problems in geometry and mathematical physics. For such topics, cf. Bismut [6] and Ikeda and Watanabe [11].

REFERENCES

[1] AZENCOTT, R. (1980). Grandes deviations et applications. *École d'Été de Saint-Flour VIII, 1978. Lecture Notes in Math.* **774** 1–176. Springer, Berlin.

[2] AZENCOTT, R. (1981). Diffusions invariantes sur le groupe d'Heisenberg; une étude de cas d'après B. Gaveau. *Geodesiques et Diffusions en Temps Petit. Astérisque* **84–85** 227–235.

[3] AZENCOTT, R. (1984). Densités des diffusions en temps petit: developements asymptotiques. *Seminaire de Prob. XVIII, 1982 / 1983. Lecture Notes in Math.* **1059** 402–498. Springer, Berlin.

[4] BEALS, M., FEFFERMAN, C. and GROSSMAN, R. (1983). Strictly pseudoconvex domains in C^n. *Bull. Amer. Math. Soc.* (*N.S.*) **8** 125–322.

[5] BISMUT, J.-M. (1984). Large deviations and the Malliavin calculus. *Progress in Math.* **45**. Birkhäuser, Boston.

[6] BISMUT, J.-M. (1984). The Atiyah–Singer theorems: A probabilistic approach. 1. The index theorem, 2. The Lefschetz fixed point formulas. *J. Funct. Anal.* **57** 56–99, 329–348.

[7] GAVEAU, B. (1977). Principe de moindre action, propagation de la chaleur, estimées sous elliptiques sur certains groupes nilpotents. *Acta Math.* **139** 95–153.

[8] HÖRMANDER, L. and MELIN, A. (1978). Free systems of vector fields. *Ark. Mat.* **16** 83–88.

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

[10] IKEDA, N. and WATANABE, S. (1984). An introduction to Malliavin's calculus. *Stochastic Analysis. Proc. Taniguchi Symp. Katata 1982*, Kinokuniya 1–52.

[11] IKEDA, N. and WATANABE, S. Malliavin calculus of Wiener functionals and its applications. In *Proc. Warwick Symp. on Stochastic Analysis.* To appear.

[12] KUSUOKA, S. (1984). The generalized Malliavin calculus based on Brownian sheet and Bismut's expansion for large deviation. *Stochastic Processes—Mathematics and Physics. Proc. First Bibos-Symp. Lecture Notes in Math.* **1158** 141–157. Springer, Berlin.

[13] KUSUOKA, S. and STROOCK, D. (1984). Applications of the Malliavin calculus, Part I. *Stochastic Analysis. Proc. Taniguchi Symp. Katata 1982*, Kinokuniya 271–306.

[14] KUSUOKA, S. and STROOCK, D. (1985). Applications of the Malliavin calculus, Part II. *J. Fac. Sci. Univ. Tokyo Sect. 1A Math.* **32** 1–76.

[15] MALLIAVIN, P. (1978). Stochastic calculus of variation and hypoelliptic operators. *Proc. Intern. Symp. SDE Kyoto 1976*, Kinokuniya 195–263.

[16] MALLIAVIN, P. (1978). C^k-hypoellipticity with degeneracy. *Stochastic Analysis* 199–214, 327–340. Academic, New York.

[17] MALLIAVIN, P. (1984). Analyse differentielle sur l'espace de Wiener. *Proc. ICM Warszawa 1983* **2** (PWN); secondary 1089–1096.

[18] MCKEAN, H. P. and SINGER, I. M. (1967). Curvature and eigenvalues of the Laplacian. *J. Differential Geom.* **1** 43–69.

[19] MEYER, P. A. (1983). Quelques resultats analytiques sur semigroupe d'Ornstein-Uhlenbeck en dimension infinie. *Theory and Applications of Random Fields. Lecture Notes in Contr. and Inform. Sci.* **49** 201–214. Springer, Berlin.

[20] MOLCHANOV, S. A. (1975). Diffusion processes and Riemannian geometry. *Russian Math. Surveys* **30** 1–63.

[21] ROTHSCHILD, L. P. and STEIN, E. M. (1976). Hypoelliptic differential operators and nilpotent groups. *Acta Math.* **137** 247–320.

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

[23] STROOCK, D. W. (1984). *An Introduction to the Theory of Large Deviations*. Springer, New York.

[24] SUGITA, H. (1985). Sobolev spaces of Wiener functionals and Malliavin's calculus. *J. Math. Kyoto Univ.* **25** 31–48.

[25] SUGITA, H. (1985). On a characterization of Sobolev spaces over an abstract Wiener space. *J. Math. Kyoto Univ.* **25** 717–725.

[26] UEMURA, H. On the short time expansion of the fundamental solution of heat equations by the method of Wiener functionals. *J. Math. Kyoto Univ.* To appear.

[27] WATANABE, S. (1984). *Lectures on Stochastic Differential Equations and Malliavin Calculus*. Springer, Berlin.

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCES
KYOTO UNIVERSITY
KYOTO 606, JAPAN