Probab. Theory Relat. Fields 91, 283–296 (1992)

Probability Theory and Related Fields

Various topologies in the Wiener space and Lévy's stochastic area

Hiroshi Sugita

Department of Information Science, Saga University, Saga 840, Japan

Received February 25, 1991; in revised form August 1, 1991

Summary. In this paper, we observe how Lévy's stochastic area looks when we see it through various topologies in the Wiener space. Our theorem implies that it is quite natural from the viewpoint of topology to define a distinct skeleton of Lévy's stochastic area $S(w)$ for each distinct topology in the Wiener space, or equivalently, for each distinct abstract Wiener space on which the Wiener measure and $S(w)$ are realized. Thus we cannot determine its intrinsic skeleton in the theory of abstract Wiener spaces.

1 Introduction and theorem

To begin with, we introduce the two-dimensional Wiener space. It is a triplet (W, H, P) whose symbols are defined by

$$W := \{w = (w^1, w^2) \in C([0,1] \to \mathbf{R}^2) \mid w(0) = 0 \in \mathbf{R}^2\},$$

$$H := \{h = (h^1, h^2) \in W \mid h(t) \text{ is absolutely continuous and}$$

$$dh/dt \in L^2([0,1] \to \mathbf{R}^2, dt)\},$$

and

$$P := \text{the standard two-dimensional Wiener measure}.$$

Here we endow the spaces W and H with norms $\|\cdot\|_W$ and $\|\cdot\|_H$ respectively which are defined by

$$\|w\|_W := \max\{|w(t)|;\ 0 \leqq t \leqq 1\},$$

$$\|h\|_H := \|dh/dt\|_{L^2([0,1] \to \mathbf{R}^2, dt)},$$

so that W becomes a separable Banach space and H, sometimes called the *Cameron-Martin subspace*, becomes a separable Hilbert space.

Let $l \in W^*$ (the topological dual space of W). Then, since the canonical coupling (l, w), $w \in W$, is a Gaussian random variable with mean zero and variance $\|l\|_H^2$ (l is naturally regarded as an element of H because $W^* \subset H^* \cong H$), the continuous mapping

$$W^* \ni l \mapsto (l, w) \in L^2(W, P)$$

is isometrically continuously extended to the following mapping

$$H \ni h \mapsto (h, w) \in L^2(W, P) .$$

The functional (h, w) is nothing but the so-called Wiener integral

$$(h, w) = \int_0^1 \frac{dh}{dt}(t) \cdot dw(t) := \int_0^1 \left(\frac{dh^1}{dt}(t) dw^1(t) + \frac{dh^2}{dt}(t) dw^2(t) \right) .$$

The Wiener integral (h, w) can be regarded as a natural extension of a continuous linear functional $\langle h, h' \rangle_H$, $h' \in H$, defined on H. Thus, some functions defined on the Cameron-Martin subspace H can be extended naturally to P-measurable functions on W. We call such extension *the stochastic extension.* Since it is well-known that $P(H) = 0$, functions which are defined only on H cannot be observed through the measure P, and we may say that it is the stochastic extension that lifts them up (if possible) to the world of full P-measure, i.e., the Banach space W. At this time, the original functions defined on H are called the *skeletons.*

The theory of Itô's stochastic calculus deals with the most important stochastic extension. It is no longer so definite as the Wiener integral, which we discussed above. In fact it involves not only pure functional analysis but also some probabilistic structures such as martingales.

For example, let us consider the following functional $\varphi(h)$ defined for $h = (h^1, h^2) \in H$.

$$\varphi(h) := \frac{1}{2} \int_0^1 \left(h^1(t) \frac{dh^2}{dt}(t) dt - h^2(t) \frac{dh^1}{dt}(t) dt \right) .$$

If the function h is of C^1-class, the value $\varphi(h)$ is just equal to the (signed) area of the region enclosed by the curve $h(t)$, $0 \leqq t \leqq 1$, and the chord connecting the origin with the point $h(1) = (h^1(1), h^2(1))$. Itô's stochastic extension is, roughly speaking, obtained by replacing the all ordinary differentials (in the above expression, for example, $(dh^1/dt)(t)dt$ and $(dh^2/dt)(t)dt$) by the Stratonovich stochastic differentials ($\circ dw^1(t)$ and $\circ dw^2(t)$, respectively). According to this scheme, the stochastic extension of $\varphi(h)$ will be a double Wiener integral defined on the two-dimensional Wiener space (W, H, P) by

$$S(w) = \frac{1}{2} \int_0^1 (w^1(t) \circ dw^2(t) - w^2(t) \circ dw^1(t)) .$$

The Wiener functional $S(w)$ is known as *P. Lévy's stochastic area* (although it is no longer "area").

Then in what sense is $S(w)$ the stochastic extension of $\varphi(h)$? Or equivalently, in what sense is $\varphi(h)$ the skeleton of $S(w)$? There may be many answers, but we here present only two of them.

(A) The first answer is given by Stroock-Varadhan's polygonal approximation [11]: Let Δ be a finite partition of the interval [0, 1].

$$\Delta : 0 = t_0 < t_1 < \cdots < t_n = 1 ,$$

and define $w_\Delta \in H$, for each $w \in W$ by

$$w_\Delta(t) := \begin{cases} w(t_i) & \text{if } t = t_i \text{ for some } i, \\ \text{linearly interpolated} & \text{if otherwise}. \end{cases}$$

Then for any $\varepsilon > 0$, we have

$$\lim_{|\Delta| \to 0} P(|S(w) - \varphi(w_\Delta)| > \varepsilon) = 0 .$$

(B) The second answer is given by the skew product representation of the two-dimensional Wiener process ([6] [11]): For any $\varepsilon > 0$ and any $h \in W^*$ ($\subset H$), we have

$$P(|S(w) - \varphi(h)| > \varepsilon \,|\, \|w - h\|_W < \delta) \to 0 \quad \text{as } \delta \to 0 .$$

This assertion says that the closer (in the sense of $\|\cdot\|_W$) w is to h, the smaller the fluctuation of $S(w)$ around $\varphi(h)$ becomes.

Although the first answer seems to be a convincing one, it is known that approximations of w by elements of H other than Stroock-Varadhan's do not always imply the approximation of $S(w)$ by $\varphi(h)$ (see [8]).

Then how about the second answer? If we replace the norm $\|\cdot\|_W$ by other metrics, will it be still valid? As a matter of fact, from the viewpoint of Gross's theory of abstract Wiener spaces, the Banach space W (accordingly, as well as its norm $\|\cdot\|_W$) is not essential. Namely, any other Banach space will also do, only if it densely contains H and its norm is *measurable* [2, 3]. Thus it makes sense to ask if (B) will be still valid when we change the norm $\|\cdot\|_W$ by other measurable norms.

The answer to this question is unfortunately "*No*". To the contrary, (and strange to say,) we will show the following theorem in the forthcoming sections.

Theorem *There exists an everywhere dense set $A \subset \mathbf{R}$ such that for each $a \in A$, we can find a measurable norm $\|\cdot\|^{(a)}$ so that*

$$P(|S(w) - \varphi(h) - a| > \varepsilon \,|\, \|w - h\|^{(a)} < \delta) \to 0, \quad \text{as } \delta \to 0 ,$$

for each $\varepsilon > 0$ and each $h \in H$.

Remark. Possibly, we may take $A = \mathbf{R}$ in this theorem. So far, however, we have shown it only for a countable everywhere dense set A. Nevertheless, it is enough for our purpose.

For each $a \in A$, let $W^{(a)}$ be the completion of H with respect to the norm $\|\cdot\|^{(a)}$. Then this theorem claims that if we take the Banach space $W^{(a)}$, instead of the space W of continuous functions, it is more natural to consider $S(w) - a$ to be the stochastic extension of $\varphi(h)$, and equivalently, $\varphi(h) + a$ to be the skeleton of $S(w)$. Thus we can conclude that from the viewpoint of topology, Itô's stochastic extension is not an intrinsic concept in the theory of abstract Wiener spaces.

2 Some lemmas

In this section, we will prepare some lemmas for the proof of the theorem.

Let $Q: H \to H$ be an orthogonal projection with $\dim QH < \infty$. Then it has an expression like

$$Qh = \sum_{j=0}^{n} \langle h_j, h \rangle_H h_j, \quad h \in H ,$$

where $\{h_0, \ldots, h_n\}$ is an orthonormal base of the range QH. Hence we can well-define an H-valued Wiener functional Qw by

$$Qw = \sum_{j=0}^{n} (h_j, w)h_j, \ w \in W .$$

Namely, $Qw \in H$ is defined for P-a.e.w and it is independent of the particular choice of the base $\{h_0, \ldots, h_n\}$.

A sequence of orthogonal projections $\{Q_n\}$ is called *an approximating sequence of projections*, if $\dim Q_n H < \infty$ and $Q_n \nearrow I_H$ (the identity operator of H), strongly.

Under these preparations, we present the definition of measurable semi-norm, which is equivalent to that of Gross [2, 3].

Definition 1 A semi-norm $\|\cdot\|_1$ on H is said to be *measurable*, if there exists a Wiener functional $q(w) < \infty$, P-a.e. such that for any approximating sequence of projections $\{Q_n\}$, a sequence of Wiener functionals $\{\|Q_n w\|_1\}$ converges to $q(w)$ in probability with respect to P. If, in addition, $\|\cdot\|_1$ is a norm, it is called *a measurable norm.*

The limit functional $q(w)$ will be denoted by $\|w\|_1$ and also called a measurable semi-norm or a measurable norm if $\|\cdot\|_1$ is a norm.

First we introduce a well-known theorem obtained by K. Itô and M. Nisio.

Lemma 1 [7] *Let $\|\cdot\|_1$ be a measurable semi-norm. Then for any approximating sequence of projections $\{Q_n\}$, we have $\|w - Q_n w\|_1 \to 0$, as $n \to \infty$, P-a.e.$w \in W$.*

Lemma 2 *Let $h \in H$ and $\|\cdot\|_1$ be a measurable semi-norm. Then for arbitrary $\varepsilon, \delta > 0$, we have*

$$P(|(h, w)| > \varepsilon, \|w\|_1 < \delta) \leqq P(|(h, w)| > \varepsilon)P(\|w\|_1 < \delta) .$$

If, in addition, $\|\cdot\|_1$ is a measurable norm, for an arbitrary $\varepsilon > 0$, we have

$$P(|(h, w)| > \varepsilon \mid \|w\|_1 < \delta) \to 0, \quad as\ \delta \to 0 .$$

Proof. (i) Let $\{Q_n\}$ be an approximating sequence of projections. Then since $Q_n w \in H$ is naturally regarded as a finite-dimensional Gaussian random vector, Lemma A1 (see Appendix) implies that

$$P(|\langle h, Q_n w\rangle_H| > \varepsilon, \|Q_n w\|_1 < \delta) \leqq P(|\langle h, Q_n w\rangle_H| > \varepsilon)P(\|Q_n w\|_1 < \delta) .$$

Letting $n \to \infty$, since $|(h, \cdot)|$ is obviously a measurable semi-norm, Definition 1 implies

$$P(|(h, w)| > \varepsilon, \|w\|_1 < \delta) \leqq P(|(h, w)| > \varepsilon)P(\|w\|_1 < \delta) .$$

(ii) Let W_1 be the completion of H with respect to $\|\cdot\|_1$, and $\{h_n\}_{n=0}^{\infty}$ be a complete orthonormal system (CONS) of H such that each h_n is taken from W_1^*. Define an approximating sequence of projections $\{Q_n\}$ by

$$Q_n := \sum_{j=0}^{n} \langle h_j, \cdot\rangle_H h_j, \quad n = 0, 1, 2, \ldots$$

Since $\langle h, \cdot - Q_n \cdot \rangle_H = \langle (I_H - Q_n)h, \cdot \rangle_H$, we see by Lemma 1 that for any $\varepsilon, \eta > 0$, there exists an $n \in N$ such that

$$P\left(|(h, w - Q_n w)| > \frac{\varepsilon}{2} \right) < \eta .$$

And hence the above assertion (i) implies

$$P\left(|(h, w - Q_n w)| > \frac{\varepsilon}{2} \middle| \, \| w \|_1 < \delta \right) < \eta \tag{1}$$

On the other hand, since all the norms are equivalent in each finite dimensional space, there exists a $\delta_0 > 0$ such that

$$P\left(|\langle h, Q_n w \rangle_H| > \frac{\varepsilon}{2} \middle| \, \| Q_n w \|_1 < \delta \right) = 0, \quad 0 < \delta < \delta_0 .$$

Now noting that Q_n is a bounded linear operator when we regard it as a mapping $W_1 \to W_1^*$, we see that for a certain $c > 0$,

$$P\left(|\langle h, Q_n w \rangle_H| > \frac{\varepsilon}{2} \middle| \, \| w \|_1 < \delta \right) = 0, \quad 0 < \delta < \delta_0 / c . \tag{2}$$

Combining (1) and (2), for $0 < \delta < \delta_0/c$, we have

$$\begin{aligned} P(|(h, w)| > \varepsilon | \, \| w \|_1 < \delta) &\leqq P(|\langle h, Q_n w \rangle_H| > \varepsilon/2 \text{ or } |(h, w - Q_n w)| \\ &\qquad > \varepsilon/2 | \, \| w \|_1 < \delta) \\ &= P(|(h, w - Q_n w)| > \varepsilon/2 | \, \| w |_1 < \delta) \\ &< \eta , \end{aligned}$$

which completes the proof. □

The norm $\| \cdot \|_W$ of W is of course measurabe and it is well-known that the Hölder norm with index less than 1/2 is also measurable. Among others, we will fully make use of *Hilbertian* measurable semi-norms, which have the following expression

$$\| h \|_\lambda = \left(\sum_{n=0}^{\infty} \lambda_n \langle h_n, h \rangle_H^2 \right)^{1/2}, \quad h \in H ,$$

where $\lambda_n \geqq 0$, $\sum \lambda_n < \infty$ and $\{h_n\}$ is a CONS of H. At this time, we have

$$\| w \|_\lambda = \left(\sum_{n=0}^{\infty} \lambda_n (h_n, w)^2 \right)^{1/2}, \; w \in W .$$

Lemma 3 *Let $\{h_n\}_{n=0}^{\infty}$ be a CONS and let $\| \cdot \|_\lambda$ be a Hilbertian measurable semi-norm defined as above and $\| \cdot \|_\nu$ be another Hilbertian measurable semi-norm defined by*

$$\| w \|_\nu := \left(\sum_{n=0}^{\infty} \nu_n (h_n, w)^2 \right)^{1/2},$$

where $v_n \geqq 0$, *and* $\sum v_n < \infty$. *Then for arbitrary* $\varepsilon, \delta > 0$, *we have*

$$P(\|w\|_\lambda > \varepsilon, \|w\|_\nu < \delta) \leqq P(\|w\|_\lambda > \varepsilon) P(\|w\|_\nu < \delta) .$$

If, in addition, $\|\cdot\|_\nu$ *is a norm (i.e.,* $v_n > 0$*), for an arbitrary* $\varepsilon > 0$, *we have*

$$P(\|w\|_\lambda > \varepsilon \mid \|w\|_\nu < \delta) \to 0, \quad as\ \delta \to 0 .$$

Making use of Lemma A2 in Appendix instead of Lemma A1, we are able to prove Lemma 3 in the same way as Lemma 2.

Now we introduce the eigen-function expansion formulas for the functional $\varphi(h)$ and Lévy's stochastic area $S(w)$. The former one is nothing but a quadratic form $\langle Kh, h\rangle_H$, $h \in H$, where $K: H \to H$ is a Hilbert-Schmidt operator given by

$$(3) \qquad K = \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\langle f_{2n}, \cdot\rangle_H f_{2n} + \langle f_{2n+1}, \cdot\rangle_H f_{2n+1}) - \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\langle g_{2n}, \cdot\rangle_H g_{2n} + \langle g_{2n+1}, \cdot\rangle_H g_{2n+1}) .$$

Here $\{f_n, g_n\}_{n=0}^{\infty}$ is, as a whole, a CONS of H, to which we can give explicit expressions using sine and cosine functions. We therefore have

$$\varphi(h) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\langle f_{2n}, h\rangle_H^2 + \langle f_{2n+1}, h\rangle_H^2) - \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\langle g_{2n}, h\rangle_H^2 + \langle g_{2n+1}, h\rangle_H^2) .$$

This K is also the kernel of Lévy's stochastic area as a double Wiener integral. Namely, we have

$$(4) \qquad S(w) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\{(f_{2n}, w)^2 - 1\} + \{(f_{2n+1}, w)^2 - 1\}) - \sum_{n=0}^{\infty} \frac{1}{(2n+1)2\pi} (\{(g_{2n}, w)^2 - 1\} + \{(g_{2n+1}, w)^2 - 1\}) .$$

The above convergence is in the sense of $L^2(W, P)$. (See [5] for details.) Since the operator K is of Hilbert-Schmidt class but not of trace class, $S(w)$ is by no means continuous in any measurable norm [12].

The following is the key lemma of this paper.

Lemma 4 *Let* $\{h_n, k_n\}_{n=0}^{\infty}$ *be as a whole an orthonormal system of* H *and* $\{\lambda_n\}_{n=0}^{\infty}$ *be a sequence of real numbers such that* $\lambda_n > 0$ *and* $\sum \lambda_n^2 < \infty$. *Define a double Wiener integral* $F(w)$ *and a measurable semi-norm* $\|\cdot\|_0$, *respectively, as follows.*

$$(5) \qquad F(w) := \sum_{n=0}^{\infty} \lambda_n \{(h_n, w)^2 - 1\} - \sum_{n=0}^{\infty} \lambda_n \{(k_n, w)^2 - 1\} ,$$

$$\|w\|_0 := \left(\sum_{n=0}^{\infty} \lambda_n^2 (h_n, w)^2 + \sum_{n=0}^{\infty} \lambda_n^2 (k_n, w)^2 + \|w\|_\perp^2 \right)^{1/2} ,$$

where $\|\cdot\|_\perp$ *is an arbitrary measurable semi-norm which degenerates on the subspace spanned by* $\{h_n, k_n\}$.

Then for any $\varepsilon > 0$, it holds that

$$P(|F(w)| > \varepsilon | \, \|w\|_0 < \delta) = O\left(\exp\left(-\frac{c\varepsilon}{\delta^2}\right)\right), \quad as\ \delta \to 0\,,$$

where $c > 0$ is a constant independent of ε and δ.

Proof. Since $F(w)$ and $\|w\|_\perp$ are mutually independent random variables, it is enough to prove the lemma in case $\|w\|_\perp = 0$. In particular, we may assume that $\{h_n, k_n\}$ spans the whole space H.
First define $\tilde{h}_n, \tilde{k}_n \in H$, $n = 0, 1, \ldots$, by

$$\tilde{h}_n := \frac{h_n + k_n}{\sqrt{2}}, \quad \tilde{k}_n := \frac{h_n - k_n}{\sqrt{2}}\,.$$

Then $\{\tilde{h}_n, \tilde{k}_n\}$ is again a CONS of H. By using the base $\{\tilde{h}_n, \tilde{k}_n\}$, $F(w)$ is rewritten as

$$\begin{aligned} F(w) &= \sum_{n=0}^{\infty} \lambda_n((h_n, w)^2 - (k_n, w)^2) \\ &= 2 \sum_{n=0}^{\infty} \lambda_n(\tilde{h}_n, w)(\tilde{k}_n, w)\,. \end{aligned}$$

Let H_1, H_2 be closed subspaces of H spanned by $\{\tilde{h}_n\}$ and $\{\tilde{k}_n\}$, respectively. Accordingly, we define induced measures $P_i := P \circ Q_i^{-1}$, $i = 1, 2$, where $Q_i: H \to H_i$ is the orthogonal projection onto H_i, and measurable semi-norms $\|\cdot\|_i$, $i = 1, 2$, by

$$\|\cdot\|_1 := \left(\sum_{n=0}^{\infty} \lambda_n^2 \langle \tilde{h}_n, \cdot \rangle_H^2\right)^{1/2}, \quad \|\cdot\|_2 := \left(\sum_{n=0}^{\infty} \lambda_n^2 \langle \tilde{k}_n, \cdot \rangle_H^2\right)^{1/2}.$$

Finally, we define Hilbert spaces W_i as the completion of H with respect to $\|\cdot\|_i$, $i = 1, 2$, respectively.

Then a triplet $(W_1 \oplus W_2, H_1 \oplus H_2, P_1 \otimes P_2)$ forms an abstract Wiener space. Since $H = H_1 \oplus H_2$, $P = P_1 \otimes P_2$, we may consider everything on this abstract Wiener space instead of the original Wiener space. Hence we once again rewrite $F(w)$ as a two-variable function $F(w_1, w_2)$. That is,

$$F(w_1, w_2) = 2 \sum_{n=0}^{\infty} \lambda_n(\tilde{h}_n, w_1)(\tilde{k}_n, w_2), \quad w_1 \in W_1,\ w_2 \in W_2\,.$$

Now if we fix a $w_1 \in W_1$, then $F(w_1, \cdot)$ becomes a linear functional on W_2 and hence its distribution under P_2 is centered Gaussian. Since $\{(\tilde{k}_n, w_2)\}$ is a sequence of independent random variables with variance 1, the variance of $F(w_1, \cdot)$ is just equal to

$$4 \sum_{n=0}^{\infty} (\lambda_n(\tilde{h}_n, w_1))^2 = 4\|w_1\|_1^2\,.$$

Consequently, for sufficiently small $\delta_0 > 0$, there exists constants $c, c_1 > 0$ such that for any w_1 with $\|w_1\|_1 \leqq \delta \leqq \delta_0$,

$$P_2(|F(w_1, w_2)| > \varepsilon) \leqq c_1 \exp\left(-\frac{c\varepsilon}{\delta^2}\right).$$

On the other hand,

$$\begin{aligned} &P_1 \otimes P_2(|F(w_1, w_2)| > \varepsilon,\ \|w_1\|_1^2 + \|w_2\|_2^2 < \delta^2) \\ &= \int_{\|w_1\|_1^2 < \delta^2} P_1(\mathrm{d}w_1) P_2(|F(w_1, w_2)| > \varepsilon,\ \|w_2\|_2^2 < \delta^2 - \|w_1\|_1^2) . \end{aligned}$$

For each fixed w_1, $F(w_1, w_2)$ is linear in w_2 and hence Lemma 2 implies that the above value is

$$\begin{aligned} &\leqq \int_{\|w_1\|_1^2 < \delta^2} P_1(\mathrm{d}w_1) P_2(|F(w_1, w_2)| > \varepsilon) P_2(\|w_2\|_2^2 < \delta^2 - \|w_1\|_1^2) \\ &\leqq c_1 \exp\left(-\frac{c\varepsilon}{\delta^2}\right) \int_{\|w_1\|_1^2 < \delta^2} P_1(\mathrm{d}w_1) P_2(\|w_2\|_2^2 < \delta^2 - \|w_1\|_1^2) \\ &= c_1 \exp\left(-\frac{c\varepsilon}{\delta^2}\right) P_1 \otimes P_2(\|w_2\|_2^2 + \|w_1\|_1^2 < \delta^2) . \end{aligned}$$

Note that for any $h = h_1 + h_2 \in H = H_1 \oplus H_2$, we have

$$\begin{aligned} \|h_2\|_2^2 + \|h_1\|_1^2 &= \sum_{n=0}^{\infty} \lambda_n^2 (\langle \tilde{h}_n, h\rangle_H^2 + \langle \tilde{k}_n, h\rangle_H^2) \\ &= \sum_{n=0}^{\infty} \lambda_n^2 (\langle h_n, h\rangle_H^2 + \langle k_n, h\rangle_H^2) \\ &= \|h\|_0^2 . \end{aligned}$$

Now we therefore have

$$P(|F(w)| > \varepsilon,\ \|w\|_0 < \delta) \leqq c_1 \exp\left(-\frac{c\varepsilon}{\delta^2}\right) P(\|w\|_0 < \delta) ,$$

which completes the proof. □

Since Lévy's stochastic area $S(w)$ has the expression (4), Lemma 4 implies

$$P(|S(w)| > \varepsilon \,|\, \|w\|^{(0)} < \delta) \to 0, \quad \text{as } \delta \to 0,$$

for a measurable norm

$$\|w\|^{(0)} := \left(\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \{(f_{2n}, w)^2 + (f_{2n+1}, w)^2 + (g_{2n}, w)^2 + (g_{2n+1}, w)^2\} \right)^{1/2} .$$

Thus Lemma 4 does not seem to support our theorem, but as a matter of fact, it does imply the theorem. The trick consists in the elementary fact that a conditionally convergent series can have various different sums by rearranging the sequence. The readers will find it in the Eqs. (8), (10) and so on in the next section.

3 Proof of theorem

We will prove the theorem in three steps. Throughout this section, ε denotes an arbitrary positive number.

Step 1. First we will prove that

(6) $$P(|S(w)-a|>\varepsilon \mid \|w\|^{(a)}<\delta)\to 0, \quad \text{as } \delta\to 0,$$

for $a=(1/2\pi)\log 2$ and

$$\|w\|^{(a)}:=\left(\sum_{n=1}^{\infty}\frac{1}{n^2}\{(f_n,w)^2+(g_{2n-1},w)^2+(g_{2n},w)^2\}+(f_0,w)^2+(g_0,w)^2\right)^{1/2}.$$

Put

$$F_0(w):=\sum_{n=1}^{\infty}\frac{1}{n}((f_n,w)^2-1)-\sum_{n=1}^{\infty}\frac{1}{n}((g_n,w)^2-1) \quad \text{and}$$

$$G(w):=S(w)-\frac{1}{2\pi}F_0(w).$$

Then we have

$$\begin{aligned} G(w) &= \left\{\frac{1}{2\pi}(f_0,w)^2+\sum_{n=1}^{\infty}\frac{1}{2\pi}\left(\frac{1}{2n}-\frac{1}{2n+1}\right)(g_{2n},w)^2\right\} \\ &\quad -\left\{\frac{1}{2\pi}(g_0,w)^2+\sum_{n=1}^{\infty}\frac{1}{2\pi}\left(\frac{1}{2n}-\frac{1}{2n+1}\right)(f_{2n},w)^2\right\} \\ &=: G_1(w)-G_2(w). \end{aligned}$$

Since $\frac{1}{2\pi}\left(\frac{1}{2n}-\frac{1}{2n+1}\right)>0$ and $\sum_{n=1}^{\infty}\frac{1}{2\pi}\left(\frac{1}{2n}-\frac{1}{2n+1}\right)<\infty$, Lemma 3 implies that

(7) $$P(|G_i(w)|>\varepsilon \mid \|w\|^{(a)}<\delta)\to 0, \quad \text{as } \delta\to 0, \quad i=1,2.$$

On the other hand, we can rewrite $F_0(w)$ as follows.

(8) $$\begin{aligned} F_0(w) &= \left\{\sum_{n=1}^{\infty}\frac{1}{n}((f_n,w)^2-1)-\sum_{n=1}^{\infty}\frac{1}{n}((g_{2n},w)^2-1)\right\} \\ &\quad -\frac{1}{2}\left\{\sum_{n=1}^{\infty}\frac{1}{n}((g_{2n-1},w)^2-1)-\sum_{n=1}^{\infty}\frac{1}{n}((g_{2n},w)^2-1)\right\} \\ &\quad -\sum_{n=1}^{\infty}\left(\frac{1}{2n-1}-\frac{1}{2n}\right)(g_{2n-1},w)^2+\sum_{n=1}^{\infty}\left(\frac{1}{2n-1}-\frac{1}{2n}\right) \\ &=: F_1(w)-\tfrac{1}{2}F_2(w)-F_3(w)+\log 2. \end{aligned}$$

Owing to Lemma 4, we have

(9) $$P(|F_i(w)|>\varepsilon \mid \|w\|^{(a)}<\delta)\to 0, \quad \text{as } \delta\to 0, \quad i=1,2,$$

and owing to Lemma 3, we have (9) also for $i=3$. Consequently, we see that

$$P(|F_0(w)-\log 2|>\varepsilon \mid \|w\|^{(a)}<\delta)\to 0, \quad \text{as } \delta\to 0.$$

Since $S(w)=G(w)+(1/2\pi)F_0(w)$, (7) and the above fact imply (6) for $a=(1/2\pi)\log 2$.

Step 2. Now putting

$$A := \left\{ a = k2^{-m}\left(\frac{1}{2\pi}\right)\log 2 \,|\, k = 0, \pm 1, \ \pm 2, \ldots, \quad m = 0, 1, 2, \ldots \right\},$$

which is everywhere dense in $\mathbf{R}$, we will show (6) for each $a \in A$.
First we note that F_0 can also be rewritten as follows.

$$F_0(w) = \left\{ \sum_{n=1}^{\infty} \frac{1}{n}((f_{2n}, w)^2 - 1) - \sum_{n=1}^{\infty} \frac{1}{n}((g_n, w)^2 - 1) \right\} \tag{10}$$

$$+ \frac{1}{2}\left\{ \sum_{n=1}^{\infty} \frac{1}{n}((f_{2n-1}, w)^2 - 1) - \sum_{n=1}^{\infty} \frac{1}{n}((f_{2n}, w)^2 - 1) \right\}$$

$$+ \sum_{n=1}^{\infty} \left(\frac{1}{2n-1} - \frac{1}{2n}\right)(f_{2n-1}, w)^2 - \sum_{n=1}^{\infty} \left(\frac{1}{2n-1} - \frac{1}{2n}\right)$$

$$=: F_1'(w) + \tfrac{1}{2}F_2'(w) + F_3'(w) - \log 2.$$

Following the same procedure of Step 1, (10) implies (6) for $a = -(1/2\pi)\log 2$ if we choose a suitable measurable norm $\|\cdot\|^{(a)}$. Now rewrite the term $F_2(w)$ in (8) in the same manner as above, say,

$$F_2 = F_{21}' + \tfrac{1}{2}F_{22}' + F_{23}' - \log 2 .$$

Then we have

$$F_0 = F_1 - \tfrac{1}{2}F_{21}' - \tfrac{1}{4}F_{22}' - \tfrac{1}{2}F_{23}' - F_3 + \tfrac{1}{2}\log 2 ,$$

which implies (6) for $a = 2^{-1}(1/2\pi)\log 2$ (and a suitable measurable norm $\|\cdot\|^{(a)}$). Continuing this procedure, it is easy to see that (6) holds for each $a = 2^{-m}(1/2\pi)\log 2$, $m = 0, 1, \ldots$. It is easy to see that it also holds for $a = -2^{-m}(1/2\pi)\log 2$, only if we exchange the roles of (8) and (10).
Thus we have obtained the following expression for each $m = 0, 1, 2, \ldots$,

$$F_0(w) = F_1(w) + G_m(w) + 2^{-m}\log 2 ,$$

where G_m as well as F_1 has the following property.

$$P(|G_m(w)| > \varepsilon \,|\, \|w\|^{(a)} < \delta) \to 0, \quad \text{as } \delta \to 0, \quad (a = 2^{-m}(1/2\pi)\log 2) .$$

Now rewrite the term $F_1(w)$ here in the same manner as above recursively, say,

$$F_1(w) = F_1''(w) + G_m''(w) + 2^{-m}\log 2 ,$$

then we have

$$F_0(w) = F_1''(w) + G_m''(w) + G_m(w) + 2\cdot 2^{-m}\log 2 .$$

Applying this procedure again and again, we will obtain (6) for every $a = k2^{-m}(1/2\pi)\log 2$, $k = 1, 2, \ldots$. It is clear that it also holds for all negative integers k.

Step 3. Let $a \in A$ and $\|\cdot\|^{(a)}$ be such that (6) holds. Take an arbitrary $h \in H$ and put

$$M(w) := \exp(-(h, w) - \tfrac{1}{2}\|h\|_H^2) .$$

Define a probability measure $\tilde{P}$ on W by

$$\tilde{P}(\mathrm{d}w) := M(w)P(\mathrm{d}w) .$$

Then it is well-known that the distribution law of $w - h$ under $\tilde{P}$ is just equal to P(Cameron-Martin's theorem). Then we have

$$P(|S(w-h) - a| > \varepsilon \mid \|w - h\|^{(a)} < \delta)$$

$$= \frac{\tilde{P}(|S(w-h) - a| > \varepsilon,\ \|w - h\|^{(a)} < \delta)}{\tilde{P}(\|w - h\|^{(a)} < \delta)} \times \frac{P(\|w\|^{(a)} < \delta)}{P(\|w\|^{(a)} < \delta)}$$

$$= \frac{E[M(w)\mathbf{1}_B(w) \mid \|w\|^{(a)} < \delta]}{E[M(w) \mid \|w\|^{(a)} < \delta]}$$

$$= \frac{E[\exp(-(h, w))\mathbf{1}_B(w) \mid \|w\|^{(a)} < \delta]}{E[\exp(-(h, w)) \mid \|w\|^{(a)} < \delta]} ,$$

where $B := \{w \mid |S(w) - a| > \varepsilon\}$. Using Schwarz inequality the numerator of the above expression can be bounded by

$$E[\exp(-2(h, w)) \mid \|w\|^{(a)} < \delta]^{1/2} \times P(|S(w) - a| > \varepsilon \mid \|w\|^{(a)} < \delta)^{1/2} .$$

Although M is not necessarily continuous, it is not difficult to see, using Lemma 2 that the conditional expectation of $\exp(-(h, w))$ given $\{\|w\|^{(a)} < \delta\}$, converges to one as δ tends to zero. This is because the conditional law of $|(h, w)|$ is bounded by the law of this variable (Lemma 2).—The limit of the conditional expectation of exponentials of Gaussian random variables has been studied by Shepp and Zeitouni in a recent work [10], using some techniques which are similar to those of this paper.
Consequently, we have

$$(11) \qquad P(|S(w-h) - a| > \varepsilon \mid \|w - h\|^{(a)} < \delta) \to 0, \quad \text{as } \delta \to 0 .$$

Noting that

$$S(w - h) = S(w) - 2(Kh, w - h) - \varphi(h) ,$$

where K is the operator given by (3), and that

$$P(|(Kh, w - h)| > \varepsilon \mid \|w - h\|^{(a)} < \delta) \to 0, \quad \text{as } \delta \to 0 ,$$

which is proved by Lemma 2 and a similar argument used in (11), we finally see that

$$P(|S(w) - \varphi(h) - a| > \varepsilon \mid \|w - h\|^{(a)} < \delta) \to 0, \quad \text{as } \delta \to 0 .$$

This completes the proof. □

4 Appendix

The aim of this Appendix is to prove the following two lemmas, from which we derived Lemma 2 and Lemma 3 in Sect. 2.

Lemma A1 (cf. [4]) *Let μ be the standard Gaussian measure on $\mathbf{R}^n$.*

$$\mu(dx) := (2\pi)^{-n/2}\exp(-|x|^2/2)dx, \quad x \in \mathbf{R}^n$$

Then for an arbitrary semi-norm $\|\cdot\|$ in $\mathbf{R}^n$ and for an arbitrary $\xi \in \mathbf{R}^n$, it holds that

$$\mu(|\langle \xi, x\rangle| < a, \|x\| < b) \geqq \mu(|\langle \xi, x\rangle| < a)\mu(\|x\| < b), \quad a, b > 0\,,$$

where $\langle\cdot,\cdot\rangle$ stands for the inner product of $\mathbf{R}^n$.

Lemma A2 *Let μ be the same as Lemma A1. For arbitrary $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\nu = (\nu_1, \ldots, \nu_n)$ with $\lambda_k, \nu_k \geqq 0$, define semi-norms $\|\cdot\|_\lambda$ and $\|\cdot\|_\nu$ respectively by*

$$\|x\|_\lambda := \left(\sum_{k=1}^{n} \lambda_k x_k^2\right)^{1/2}, \quad \|x\|_\nu := \left(\sum_{k=1}^{n} \nu_k x_k^2\right)^{1/2}, \quad x = (x_1, \ldots, x_n) \in \mathbf{R}^n\,.$$

Then it holds that

$$\mu(\|x\|_\lambda < a, \|x\|_\nu < b) \geqq \mu(\|x\|_\lambda < a)\mu(\|x\|_\nu < b), \quad a, b > 0\,.$$

The assertion of these two lemmas seems to be valid for any pair of semi-norms. In fact, when $n \leqq 2$, L.D. Pitt [9] proved it for any pair of semi-norms. But so far, it seems to be still an open problem for $n \geqq 3$.

In this paper, we will approach to the problem by means of the following inequality, which is an elementary case of the so-called FKG-inequality.

Lemma A3 *Let $f, g\colon [0, \infty) \to \mathbf{R}$ be non-increasing functions and $m(dx)$ be a probability measure without atoms on $[0, \infty)$. If f and g are square integrable with respect to m, then we have*

$$\int_0^\infty f(x)g(x)m(\mathrm{d}x) \geqq \int_0^\infty f(x)m(\mathrm{d}x) \int_0^\infty g(x)m(\mathrm{d}x)\,.$$

Proof. Put $\bar{f} := \int_0^\infty f(x)m(\mathrm{d}x)$. Since f is non-increasing, there exists an $x_0 \in [0, \infty)$ such that

$$\begin{cases} f(x) - \bar{f} \geqq 0, & 0 \leqq x < x_0, \\ f(x) - \bar{f} \leqq 0, & x_0 < x < \infty\,. \end{cases}$$

On the other hand, g is non-increasing too, and hence putting $\beta := g(x_0)$, we have

$$\begin{cases} g(x) - \beta \geqq 0, & 0 \leqq x < x_0, \\ g(x) - \beta \leqq 0, & x_0 < x < \infty\,. \end{cases}$$

We therefore have

$$(f(x) - \bar{f})(g(x) - \beta) \geqq 0, \quad x \in [0, x_0) \cup (x_0, \infty)\,.$$

Then we see that

$$0 \leqq \int_0^\infty (f(x) - \bar{f})(g(x) - \beta)m(\mathrm{d}x)$$

$$= \int_0^\infty f(x)g(x)m(\mathrm{d}x) - \int_0^\infty f(x)m(\mathrm{d}x) \int_0^\infty g(x)m(\mathrm{d}x). \qquad \square$$

Now let us prove Lemma A1. Since the Gaussian measure μ is invariant under any rotation around the origin, it is enough to show

$$\mu(|x_1| < a, \|x\| < b) \geqq \mu(|x_1| < a)\mu(\|x\| < b)\,,$$

where x_1 is the first coordinate of $x = (x_1, \ldots, x_n)$.

Let μ_1 and μ_y be the marginal distributions of $x_1 \in \mathbf{R}$ and $y = (x_2, \ldots, x_n) \in \mathbf{R}^{n-1}$, respectively. It is clear that the both two are again Gaussian and that $\mu = \mu_1 \otimes \mu_y$. Then we have

$$\mu(|x_1| < a, \|x\| < b) = 2 \int_0^\infty \mathbf{1}_{[0,a]}(x_1)\mu_y(\{y \in \mathbf{R}^{n-1} \,|\, \|(x_1, y)\| < b\})\mu_1(\mathrm{d}x_1)\,.$$

The indicator function $\mathbf{1}_{[0,a]}(x_1)$ of the interval $[0, a]$ is clearly non-increasing in $x_1 \geqq 0$ and it is known that the function

$$\mu_y(\{y \in \mathbf{R}^{n-1} \,|\, \|(x_1, y)\| < b\})$$

is also non-increasing in $x_1 \geqq 0$. (See [1], the proof of Theorem 6.1.) Applying Lemma A3, we complete the proof of Lemma A1.

We will next prove Lemma A2 by induction. For $n = 1$, the assertion is clear. Suppose $n \geqq 2$, then we have

$$\mu(\|x\|_\lambda < a, \|x\|_\nu < b) = 2 \int_0^\infty \mu_1(\mathrm{d}x_1)\mu_y(\|(x_1, y)\|_\lambda < a, \|(x_1, y)\|_\nu < b)\,.$$

Now it follows from the hypothesis of the induction that

$$\mu_y(\|(x_1, y)\|_\lambda < a, \|(x_1, y)\|_\nu < b) \geqq \mu_y(\|(x_1, y)\|_\lambda < a)\mu_y(\|(x_1, y)\|_\nu < b), \quad x_1 \geqq 0\,.$$

Furthermore, functions $\mu_y(\|(x_1, y)\|_\lambda < a)$ and $\mu_y(\|(x_1, y)\|_\nu < b)$ are both non-increasing in $x_1 \geqq 0$, and hence we have

$$\mu(\|x\|_\lambda < a, \|x\|_\nu < b) \geqq 2 \int_0^\infty \mu_1(\mathrm{d}x_1)\mu_y(\|(x_1, y)\|_\lambda < a)\mu_y(\|(x_1, y)\|_\nu < b)$$

$$\geqq 2 \int_0^\infty \mu_1(\mathrm{d}x_1)\mu_y(\|(x_1, y)\|_\lambda < a) \cdot 2 \int_0^\infty \mu_1(\mathrm{d}x_1)\mu_y(\|(x_1, y)\|_\nu < b)$$

$$= \mu(\|x\|_\lambda < a)\mu(\|x\|_\nu < b)\,. \qquad \square$$

Acknowledgement. The author would like to thank the referee for presenting him an improved proof of Step 3 of the main Theorem.

References

1. Borell, C.: Convex measures on locally convex spaces. Ark. Mat. **12**, (2) 239–252 (1974)
2. Gross, L.: Measurable functions on Hilbert space. Trans. Am. Math. Soc. **105**, 372–390 (1962)
3. Gross, L.: Abstract Wiener spaces. Proceedings of the Fifth Berkeley Symposium on Math. Statist. and Prob. II, Part I. University of California pp. 31–42, 1967
4. Das Gupta, S., Eaton, M.L., Olkin, I., Perlman, M., Savage, L.J., Sobell, M.: Inequalities on the probability content of convex regions for elliptically contoured distributions, Proc. of Sixth Berkeley Symp. on Math. Statist. and Prob. II. University of California, pp. 241–264, 1972
5. Hida, T.: Brownian motion. New York Heidelberg Berlin: Springer 1980
6. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes, 2nd edn. Amsterdam Tokyo: North-Holland Kodansha 1989
7. Itô, K., Nisio, M.: On the convergence of sums of independent Banach space valued random variables. Osaka J. Math. **5**, 35–48 (1968)
8. Kallianpur, G.: Some remarks on Hu and Meyer's paper and finitely additive canonical Hilbert space. (preprint)
9. Pitt, L.D.: A Gaussian correlation inequality for symmetric convex sets. Ann. Probab. **5**, 470–474 (1977)
10. Shepp, L.A., Zeitouni, O.: A note on conditional exponential moments and Onsager Machlup functionals. (preprint)
11. Stroock, D.W., Varadhan, S.R.S.: On the support of diffusion processes with applications to the strong maximum principle, Proc. of the Sixth Berkeley Symp. on Math. Statist. and Prob. III. University of California 333–359, 1972
12. Sugita, H.: Hu-Meyer's multiple Stratonovich integral and essential continuity of multiple Wiener integral. Bull. Sci. Math., II. Sér. **113**, 463–474 (1989)