LIDS-P-2077

ONSAGER MACHLUP FUNCTIONALS FOR A CLASS OF NON TRACE CLASS SPDE's

Eddy Mayer Wolf [1]
Department of Mathematics
Technion- Israel Institute of Technology
Haifa 32000, ISRAEL

and
Ofer Zeitouni [2]
Department of Electrical Engineering
Technion- Israel Institute of Technology
Haifa 32000, ISRAEL

6 November 1991

ABSTRACT

An Onsager Machlup functional limit is derived for a class of SPDE's whose principal part is not trace class. The proof uses FKG type inequalities.

1 Introduction

Let u denote the solution to the SPDE (in a bounded domain $D \subset \mathbb{R}^d$, with zero Dirichlet boundary conditions)

$$Pu + F(u) = n, \tag{1.1}$$

where P is an elliptic operator of order $2k$, F is a "nice" operator (for example, a smooth point function of u) and n is a white noise process. An exact definition of what is meant by a solution of (1.1) and the various objects and function spaces involved is given in Section 2.

We are interested in computing limits of the form

$$\lim_{\epsilon \to 0} \frac{\text{Prob}\,(\| u - \phi \|_B < \epsilon)}{\text{Prob}\,(\| u \|_B < \epsilon)} \triangleq \exp J_B(\phi) \tag{1.2}$$

where $\| \quad \|_B$ denotes an appropriate norm and ϕ is a deterministic function satisfying some regularity conditions. The functional $J_B(\phi)$ is called the Onsager-Machlup functional associated with the solution of (1.1).

It is well known that under mild restrictions (which are imposed below), $P : W_0^{2k,2}(D) \to L^2(D)$ possesses a bounded inverse G_0. By Maurin's theorem ([1, Theorem 6.35]) the imbedding

[1]The work of this author was partially supported by the Bernstein Research Fund at the Technion
[2]The work of this author was partially supported by the Center for Intelligent Control Systems at MIT under US Army research office grant DAAL03-86-K0171

of $W^{2k,2}(D)$ into $L^2(D)$ is Hilbert-Schmidt, so that G_0 is actually a Hilbert-Schmidt operator in $L^2(D)$. By pushing Maurin's result the full extent of its validity the same may (and will in the Appendix) be concluded for $G_0 : L^2(D) \to W^{m,2}(D)$ as long as $m < 2k - d/2$.

In [3], it was shown that if $2k > d$ (which now implies that the operator $P^{-1} : L^2(D) \to L^2(D)$ is trace class), if F is a pointwise smooth function and if $\| \quad \|_B$ is taken as the Sobolev norm in $W^{2k-d+\delta}$, some appropriate δ, then the limit in the R.H.S. of (1.2) exists and is given by

$$J_B(\phi) = -\frac{1}{2}\int_D (P\phi_1 + F(\phi_1))^2 dx - \log\det\left((D_{\phi_1}F - D_oF)(P + D_oF)^{-1}\right). \qquad (1.3)$$

On the other hand, in [2] the linear case $F \equiv 0$ was treated. It was shown there that (1.3) holds true even in the domain $d \geq 2k > d/2$ (in which case P^{-1} is only Hilbert-Schmidt but not trace class) and further (1.3) simplifies to

$$\tilde{J}_B(\phi) = -\frac{1}{2}\int_D (P\phi_1 + F(\phi_1))^2 dx. \qquad (1.4)$$

In this paper, we attempt to bridge this gap. Namely, for the case where $d \geq 2k > d/2$, and F is nonlinear, we compute the limit in (1.2). Surprisingly, it turns out (unlike in all other known computations) that the limit (1.2) may become 0 or ∞ even for nice smooth ϕ's and smooth point-functions F. Due to this phenomenon, we turn our attention, instead of (1.2), to the limit

$$\lim_{\epsilon\to 0}\frac{\text{Prob}\,(\|u-\phi_1\|_2 < \epsilon)}{\text{Prob}\,(\|u-\phi_2\|_2 < \epsilon)} \triangleq \exp J(\phi_1, \phi_2)\,, \qquad (1.5)$$

where $\|\cdot\|_2$ denotes throughout the $L^2(D)$ norm. The main result of this paper is roughly as follows:

Let $2k > d/2$. If the operator $K_{\phi_1,\phi_2} \triangleq [D_{\phi_1}F - D_{\phi_2}F](P + D_{\phi_2}F)^{-1}$ is trace class and satisfies some regularity conditions, then the limit in (1.5) exists and is given by (4.36) (the exact statement is given in Theorem 4.1 below).

A few differences between the results of this paper and the results of [3] seem worth emphasizing. First, note that even in the case of trace class operators P^{-1}, the results presented here are under an L^2 norm (and in general, by a similar technique, under appropriate L^2 type Sobolev norms) which differs from the norm used in [3]. Next, we allow here for non trace class operators. Finally, the basic estimates in this paper, which are motivated by [9], make use of FKG type inequalities and thus are different in nature from the estimates in [3]. We believe that these estimates may be of independent interest.

The organization of the paper is as follows: In Section 2, we define, following [3], our basic SPDE and state the appropriate preliminary results (existence, uniqueness, regularity and some Radon-Nykodim derivative computations), and prove a conditional expectation lemma which serves us well in the sequel. In Section 3, we compute the Onsager Machlup functional for a linearized version of the equation and prove the non existence of non trivial limits when the operators involved are not trace class. Finally, in Section 4, we prove the positive result (Theorem 4.1) referred to above, by showing that the linearized equation has the same limiting behaviour as the original equation.

Acknowledgements: We thank Moshe Zakai for a suggestion related to the proof of Theorem 2.2.

2 Preliminaries

We closely follow the notations of [2,3]. Let D denote a bounded domain in $\mathbb{R}^d$ possessing a smooth boundary, with $0 \in D$. $W^{m,2}(D)$ and $W_o^{m,2}(D)$ denote the usual Sobolev spaces equipped with the norm $||\cdot||_{m,2}$. The inner product in $L^2(D)$ is denoted $\langle\cdot,\cdot\rangle$. For any real z, let $\lfloor z \rfloor$ denote the largest integer strictly smaller than z. Let P be a strongly elliptic differential operator of order $2k$ with smooth coefficients and $F : W^{\lfloor 2k-d/2\rfloor,2}(D) \to L^2(D)$ a (possibly) nonlinear transformation. Throughout this paper the following two assumptions will always be implicitly made.

(A1) $2k > d/2$.

(A2) F possesses a Fréchet derivative D_uF, bounded and continuous in $u \in W^{\lfloor 2k-d/2\rfloor,2}(D)$ and moreover,

$$||DF|| \equiv \sup_{u\in W^{\lfloor 2k-d/2\rfloor,2}(D)} ||D_uF|| < \inf_{0\neq\phi\in W^{2k,2}(D)} \frac{||P\phi||_2}{||\phi||_{2k,2}}.$$

Next, let $W_{x_1,\dots,x_d}$ be a standard Brownian sheet in D with respect to the probability space $(\Omega,\mathcal{F},\mathcal{P}_W)$. Without loss of generality we shall assume that $\Omega = C_0(D) \equiv \{f \in C(D) \big| f(x) = 0 \text{ if } \prod_{i=1}^d x_i = 0\}$. It is known that the linear functional (white noise) $n : \phi \to \int_D \phi(x)dW_x$ defines almost surely an element in $W^{-r,2}(D)$ for any $r > d/2$ (cf. [10, p. 335]), thus in particular in $W^{-2k,2}(D)$. For $f \in C_0(D)$ denote by ∂f the distributional derivative $\frac{\partial^d f}{\partial x_1 \dots \partial x_d}$, and let $H(D) = \{\omega \in C_0(D) : \partial\omega \in L^2(D)\}$. In what follows ∂'s inverse will only be applied to elements of $L^2(D)$ so that we have the representation $\partial^{-1}f(x) = \int_{D\cap R_x} f(\xi)d\xi$, where R_x is the "rectangle" determined by the 2^d vertices $(a_i)_{i=1}^d, a_i = 0$ or x_i.

Given a Banach space X, we say that a mapping $g : C_0(D) \to X$ possesses an $H(D)$ Fréchet derivative at $\omega \in C_0(D)$ if there exists a bounded linear operator $D_\omega g : H(D) \to X$ such that as $||h||_{H(D)} \to 0$, $||g(\omega+h) - g(\omega) - D_\omega g h||_X = o(||h||_{H(D)})$.

Fix $\nu \in W^{-2k,2}(D)$. We shall say that $u \in W^{\lfloor 2k-d/2\rfloor,2}(D)$ is a solution to the equation

$$Pu + F(u) = \nu \tag{2.6}$$

with zero Dirichlet boundary conditions if $\forall\phi \in W_0^{2k,2}(D)$,

$$\int_D [uP^*\phi + F(u)\phi](x)dx = \nu(\phi) \tag{2.7}$$

(with P^* denoting P's adjoint).

The following existence and uniqueness result was proved in [3, Theorem 2.1].

Theorem 2.1 *The equation (2.6) possesses a unique solution in* $W_0^{\lfloor 2k-d/2\rfloor,2}(D)$. *Furthermore,* $u \in C^{\alpha+\lfloor 2k-d/2\rfloor}(D)$ *for some* $\alpha > 0$ *which depends on* k *and* d *but not on* ν.

We shall denote this solution by $G_F(\nu)$. In particular the stochastic PDE

$$Pu + F(u) = n \tag{2.8}$$

should be viewed in this sense ω by ω, and $u(\omega) \equiv G_F(n)$ is its unique solution. Note that by this construction there is also uniqueness in law for (2.8).

We now turn to a representation of the ratio $\frac{\text{Prob}(\|u-\phi_1\|_2<\epsilon)}{\text{Prob}(\|u-\phi_2\|_2<\epsilon)}$. Actually, for $i = 1, 2$, $(u - \phi_i)$ itself satisfies (2.8) if F is replaced by a suitable F_i (which still satisfies **(A2)**). For this reason — and also because it will be later convenient to compare (2.8) to its linearized version — the next result, which is based on Kusuoka's theorem ([7, Theorem 6.4]) and whose proof we defer to the Appendix, is stated in terms of F_1 and F_2 instead of ϕ_1 and ϕ_2.

Theorem 2.2 *For $i = 1, 2$ let $F_i : W^{\lfloor 2k-d/2 \rfloor, 2}(D) \to L^2(D)$ satisfy* **(A2)** *and $u_i = G_{F_i}(n)$ (i.e the unique solution to the SPDE (2.8) for $F = F_i$). Then for any Borel set $B \subset W_0^{2k,2}(D)$,*

$$\frac{\text{Prob}\,(u_1 \in B)}{\text{Prob}\,(u_2 \in B)} = E\left(\Lambda_{1,2} \mid u_2 \in B\right) \tag{2.9}$$

where

$$\begin{aligned}\Lambda_{1,2}(\omega) = \quad & \det{}_2 \quad (I + D_{u_2}(F_1 - F_2) G_{D_{u_2} F_2}) \\ & \exp\left\{ \int_D \left[(F_2 - F_1)(u_2)\right](x) \delta W_x - \frac{1}{2} \int_D \left[(F_2 - F_1)(u_2)\right]^2 (x) dx \right\}\end{aligned} \tag{2.10}$$

with $\det_2(I+Q)$ denoting the Carleman–Fredholm determinant of Q, i.e. if λ_i are the eigenvalues of $(Q^*Q)^{\frac{1}{2}}$ then $\det_2(I+Q) = \prod_1^\infty (1+\lambda_i)e^{-\lambda_i}$, (see [4, XI.9.22]), and the differential δW_x indicating Skorohod integration ([7,8]) with respect to the Brownian sheet W over D.

For our purposes it suffices to recall that given a map $a : C_0(D) \to L^2(D)$ which possesses an $H(D)$–Fréchet derivative $D_\omega a : H(D) \to L^2(D)$ at $\mathcal{P}_W$ almost every $\omega \in C_0(D)$, its Skorohod integral with respect to the Brownian sheet W is defined (in the sense of $L^2(\Omega)$ convergence) by

$$\int a(\omega) \delta W_x = \sum_1^\infty (\langle a,\ e_i \rangle n(e_i) - \langle D_\omega a \hat{e}_i,\ e_i \rangle) \tag{2.11}$$

where $\{e_i\}_{i=1}^\infty$ is a CONS in $L^2(D)$, $n(e_i)$ is the usual Wiener integral of e_i with respect to the Brownian sheet W, and $\hat{e}_i$ is e_i's identification in $H(D)$, namely $\hat{e}_i = \partial^{-1} e_i$.

We remark that in the case $2k > d$ which was treated in [3], the Skorohod integral in (2.10) could be decomposed into an Ogawa integral and a correction term of the trace form, each existing separately. Here, since the trace is not finite in general, one has to use the Skorohod integral.

We conclude this section with the following lemma, which will turn out to be crucial in the evaluation of exponential estimates. Here and throughout, a non trace class operator $T : K \to K$ (where K is a separable Hilbert space) will be said to have infinite trace (denoted: $\operatorname{tr} T = \infty$) if $\sum_i \langle Te_i, e_i \rangle = \infty$ for any CONS $\{e_i\}$ in K. A similar definition holds for $\operatorname{tr} T = -\infty$. Note that if $\operatorname{tr} T = \infty$ then $\sum_{i \in I} \langle Te_i, e_i \rangle > -c$ for some positive c and any $I \subset \mathcal{N}$.

Lemma 2.3 *a) Let $T : \ell^2 \to \ell^2$ be a deterministic trace class operator, and let T_{ij} denote its canonical (i, j) element. Let η_i be a deterministic sequence, with $\sum_i \frac{1}{\eta_i^2} < \infty$. Let a_i be a sequence of unit variance, independent Gaussian random variables with $\sum_i E(a_i^2) < \infty$. Then*

$$E(\exp(\sum_{i,j} a_i a_j T_{ij}) | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \to_{\epsilon \to 0} 1, \tag{2.12}$$

and, for any deterministic sequence c_i with $\sum_i c_i^2 < \infty$,

$$E(\exp(\sum_{i,j}(a_i + c_i)a_j T_{ij})| \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \to_{\epsilon\to 0} 1. \tag{2.13}$$

b) Let $T : \ell^2 \to \ell^2$ be a deterministic, Hilbert-Schmidt operator, and let $T_{i,j}$ denote its canonical (i,j) element. Assume that $\operatorname{tr} T = \infty$. Let η_i be a deterministic sequence, with $\sum_i \frac{1}{\eta_i^2} < \infty$. Let $\{a_i\}$ be a sequence of independent Gaussian random variables as above. Then

$$\lim_{\epsilon\to 0} E(\exp(\sum_{i\neq j} a_i a_j T_{ij} + \sum_i (a_i^2 - 1)T_{ii})| \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) = 0. \tag{2.14}$$

Similarly, under the same assumptions but $\operatorname{tr} T = -\infty$,

$$\lim_{\epsilon\to 0} E(\exp(\sum_{i\neq j} a_i a_j T_{ij} + \sum_i (a_i^2 - 1)T_{ii})| \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) = \infty. \tag{2.15}$$

Remark: All infinite sums involving a_i above are to be interpreted in the sense of L^2 convergence, which is ensured since T is Hilbert–Schmidt.

Proof of Lemma 2.3: a) Note first that, since under the conditioning, $\sum_{i,j=1}^N a_i a_j T_{ij} \to_{\epsilon\to 0} 0$ uniformly in ω for each fixed deterministic N, the proof of part a) follows once the following estimates are proved for each deterministic (not necessarily positive) constant c:

$$E(\exp c \sum_{i=1}^{\infty} T_{ij} a_i | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \underset{\epsilon \to 0}{\longrightarrow} 1 \tag{2.16}$$

for each $1 \le j \le N$, and

$$E(\exp c \sum_{i,j=N}^{\infty} a_i a_j T_{ij} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \le \exp C(c,N) \tag{2.17}$$

where $C(c,N) \underset{N\to\infty}{\longrightarrow} 0$. Note that since $\sum_i T_{ij}^2 < \infty$, (2.16) is proved like Theorem 1 in [9]. To see (2.17), let $\tilde{T}_{ij}$ denote the operator with $\tilde{T}_{ij} = T_{ij}$ if $i \neq j$ and $\tilde{T}_{ii} = |T_{ii}|$. Since T_{ij} is trace class, so is $\tilde{T}_{ij}$, and it is clearly enough to prove (2.17) with $\tilde{T}$ replacing T. Denote by $\mathcal{F}_{|a|}$ the sigma algebra generated by the sequence $\{|a_i|, i = 1, \ldots\}$, and define

$$A = E(\exp |c| \sum_{ij=N}^{\infty} a_i a_j \tilde{T}_{ij} | \mathcal{F}_{|a|}). \tag{2.18}$$

We claim that A is a nondecreasing function of $|a_i|$ for each i. Indeed, by symmetry, it is enough to check that A increases in $|a_N|$, and that follows from the equality

$$A = E(B \exp(|c| a_N^2 \tilde{T}_{NN}) \cosh(|c| a_N \sum_{j=N+1}^{\infty} a_j T_{Nj}) | \mathcal{F}_{|a|}^{N+1}) \tag{2.19}$$

where B depends only on $a_i, i \geq N+1$ and $\mathcal{F}_{|a|}^{N+1}$ denotes the sigma field generated by $\{|a_i|, i \geq N+1\}$. One concludes that A is a nondecreasing function of $|a_N|$, and by symmetry this implies that A is nondecreasing as a function of each of the random variables $|a_i|$. On the other hand, the function $\mathbf{1}_{\sum_i \frac{a_i^2}{\eta_i^2} > \epsilon}$ is also clearly nondecreasing in each of the $|a_i|$. Since $|a_i|$ are independent random variables, they are associated [5], and hence, by the FKG inequality,

$$E(A\mathbf{1}_{\sum_i \frac{a_i^2}{\eta_i^2} > \epsilon}) \geq E(A)E(\mathbf{1}_{\sum_i \frac{a_i^2}{\eta_i^2} > \epsilon}), \tag{2.20}$$

which implies that

$$E(A|\sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \leq E(A) = E(\exp |c| \sum_{ij=N}^{\infty} a_i a_j \tilde{T}_{ij}) \leq \exp C(c,N) \tag{2.21}$$

where $C(c,N)$ is a constant which depends on the trace of $c\tilde{T}_N$, where $\tilde{T}_N$ denotes the truncation of the operator $\tilde{T}$ such that $\tilde{T}_N(i,j) = 0$ if either i or j are smaller than N, and $C(c,N) \to 0$ as $N \to \infty$. The proof of (2.12) follows by noting that

$$E(\exp c \sum_{i,j=N}^{\infty} a_i a_j T_{ij} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) = E(A|\sum_i \frac{a_i^2}{\eta_i^2} < \epsilon). \tag{2.22}$$

(2.13) follows from (2.12) by combining the proof in [9] with the fact (see [6], pg. 536-537) that, if for all constants c, $\limsup_{\epsilon \to 0} E(\exp(cA_k)|\sum_i a_i^2/\eta_i^2 < \epsilon) \leq 1$, $k = 1,2$, then $\lim_{\epsilon \to 0} E(\exp(A_1 + A_2)| \sum_i a_i^2/\eta_i^2 < \epsilon) = 1$.

b) We prove (2.14), the proof of (2.15), being similar, is ommitted. Towards this end, note first that, for each $N < \infty$, $\sum_{i \neq j, i,j \leq N} a_i a_j T_{ij} + \sum_{i=1}^N (a_i^2 - 1) T_{ii} \to_{\epsilon \to 0} -\sum_{i=1}^N T_{ii}$, uniformly in a_i in the conditioning set. Further note that, by Theorem 1 in [9], for each constant $0 < l \leq N$, and any constant k,

$$\lim_{\epsilon \to 0} E(\exp k c_l \sum_{i=N}^{\infty} a_i T_{ij} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) = 1$$

because $\sum_{i=N}^{\infty} T_{ij}^2 < \infty$. It therefore suffices to show that there exists a constant k, independent of N and ϵ, such that

$$E(\exp \sum_{i \neq j, i,j \geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty} (a_i^2 - 1) T_{ii} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) < k.$$

By taking N large enough, one may assume that all eigenvalues of $(T^*T)^{1/2}$ are bounded by $1/2$. Repeating now the argument used in the proof of part a) of this lemma, one concludes that

$$\begin{aligned} E(\exp \sum_{i \neq j, i,j \geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty} (a_i^2 - 1) T_{ii} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) &\leq \\ E(\exp \sum_{i \neq j, i,j \geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty} (a_i^2 |T_{ii}| - T_{ii}) | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) &\leq \\ e^c E(\exp \sum_{i \neq j, i,j \geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty} (a_i^2 |T_{ii}| - |T_{ii}|)), \end{aligned} \tag{2.23}$$

where the last inequality follows from the assumption on the negative part of T_{ii}. Let now $\bar{T}$ denote the operator with $\bar{T}_{ij} = T_{ij}$ if $i \neq j$ and $\bar{T}_{ii} = |T_{ii}|$. Clearly, $\bar{T}$ is also a Hilbert–Schmidt operator, and

$$E(\exp \sum_{i\neq j, i,j\geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty}(a_i^2-1)T_{ii} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) \leq e^c E(\exp \sum_{i\neq j, i,j\geq N} a_i a_j \bar{T}_{ij} + \sum_{i=N}^{\infty}(a_i^2-1)\bar{T}_{ii}). \tag{2.24}$$

Let λ_i denote the eigenvalues of $(\bar{T}^* + \bar{T})/2$, and note that $\sum_{i=1}^{\infty}\lambda_i^2 < \infty$, and that, by the remark above, $\lambda_i < 1/2$. Therefore, using (2.24), it follows that for some constants k_1, k_2, k_3,

$$\begin{aligned} E(\exp \sum_{i\neq j, i,j\geq N} a_i a_j T_{ij} + \sum_{i=N}^{\infty}(a_i^2-1)T_{ii} | \sum_i \frac{a_i^2}{\eta_i^2} < \epsilon) &\leq k_1 \exp(k_2 \sum_i (Ea_i)^2) \prod_{i=1}^{\infty}(1-2\lambda_i)^{-1/2}e^{-\lambda_i} \\ &< k_3, \end{aligned} \tag{2.25}$$

which completes the proof of the lemma. □

3 Probability ratios for linearized equations

Let u_{ϕ_i}, $i = 1, 2$ denote the solutions to the equations

$$Pu_{\phi_i} + F(\phi_i) + D_{\phi_i}F(u_{\phi_i} - \phi_i) = n \tag{3.26}$$

where $\phi_i \in W_o^{4k,2}$. In this section, we derive a modified Onsager Machlup functional for solutions of (3.26), namely

Theorem 3.1 *a) Assume that $(D_{\phi_1}F - D_{\phi_2}F)P^{-1} : L^2(D) \to L^2(D)$ is a trace class operator. Then,*

$$\begin{aligned} \hat{J}(\phi_1, \phi_2) &\equiv \log \lim_{\epsilon\to 0} \frac{\text{Prob}\,(||u_{\phi_1} - \phi_1||_2 < \epsilon)}{\text{Prob}\,(||u_{\phi_2} - \phi_2||_2 < \epsilon)} \\ &= -\frac{1}{2}\int_D ((P\phi_1 + F(\phi_1))^2 - (P\phi_2 + F(\phi_2))^2)dx \\ &\quad + \log\det[I + (D_{\phi_1}F - D_{\phi_2}F)(P + D_{\phi_2}F)^{-1}] \end{aligned} \tag{3.27}$$

b) Assume that $\text{tr}[(D_{\phi_1}F - D_{\phi_2}F)P^{-1}] = \infty$ *(respectively, $= -\infty$). Then $J(\phi_1, \phi_2) = \infty$ (respectively, $J(\phi_1, \phi_2) = -\infty$).*

Proof: The theorem follows from theorem 3.2 below by the substitution $A_i = D_{\phi_i}F$, $\Psi_i = F(\phi_i) + P\phi_i$. □

Theorem 3.2 *For $i = 1, 2$, let $A_i : W^{\lfloor 2k-d/2\rfloor,2}(D) \to L^2(D)$ be a bounded linear operator, let $\Psi_i \in W_0^{2k,2}(D)$ such that the operator $F(u) = A_iu + \Psi_i$ satisfies* **(A2)**. *Let v_i, $i = 1, 2$, denote the unique solutions to*

$$Pv_i + A_iv_i + \Psi_i = n. \tag{3.28}$$

Let $Q = (A_1 - A_2)(P + A_2)^{-1}$.

a) If Q is trace class then

$$\lim_{\epsilon\to 0}\frac{\text{Prob}\,(||v_1||_2<\epsilon)}{\text{Prob}\,(||v_2||_2<\epsilon)}=\det(I+Q)\exp\left(\frac{1}{2}(||\Psi_2||_2^2-||\Psi_1||_2^2)\right). \tag{3.29}$$

b) If $\operatorname{tr} Q=\infty$ (respectively, $-\infty$) in the sense that $\sum_{i=1}^{\infty}\langle Qe_i,e_i\rangle=\infty$ (respectively, $-\infty$) for any CONS $\{e_i\}\in L^2(D)$, then the limit above is ∞ (respectively, 0).

Remark: Note that Q is trace class if and only if the operator $\tilde{Q}=(A_1-A_2)P^{-1}$ is trace class.

Proof: By Theorem 2.2,

$$\begin{aligned}\frac{\text{Prob}\,(||v_1||_2<\epsilon)}{\text{Prob}\,(||v_2||_2<\epsilon)} &= E_\epsilon\left(\det{}_2(I+Q)\exp(\int_D(A_2-A_1)v_2\delta W_x\right. \\ &\quad \left.+\int_D(\Psi_2-\Psi_1)\delta W_x-\frac{1}{2}||(A_2-A_1)v_2+\Psi_2-\Psi_1||_2^2)\right),\end{aligned} \tag{3.30}$$

where here and henceforth, E_ϵ denotes the expectation $E(\cdot|\,||v_2||_2<\epsilon)$. Since the integrand in $\int_D(\Psi_2-\Psi_1)\delta W_x$ is deterministic, the latter may be integrated by parts, resulting in

$$\int_D(\Psi_2-\Psi_1)\delta W_x=\int_D(\Psi_1-\Psi_2)\Psi_2dx+\int_D((\Psi_2-\Psi_1)A_2v_2+v_2P^*(\Psi_2-\Psi_1))\,dx.$$

Therefore, the expression in the exponent of (3.30) differs from $\int_D(A_2-A_1)v_2\delta W_x+\frac{1}{2}(||\Psi_2||_2^2-||\Psi_1||_2^2)$ by a term which converges to zero with $||v_2||_2$, uniformly in ω. It remains therefore only to show that

$$\lim_{\epsilon\to 0}E_\epsilon\exp\int_D(A_2-A_1)v_2\delta W_x=\exp(-\operatorname{tr}Q)\,\in[0,\infty].$$

Let (e_i,λ_i) denote the eigenfunctions and eigenvalues associated with the Karhunen–Loeve expansion of v_2, namely $v_2=\sum_i\xi_ie_i/\lambda_i$ where ξ_i are independent Gaussian random variables with means $b_i=-\langle\lambda_i(P+A_2)^{-1}\Psi_2,e_i\rangle$ and unit variance, and $\sum_i 1/\lambda_i^2<\infty$, which implies that $\sum_i b_i^2<\infty$. Note that, by (2.11),

$$\begin{aligned}\int_D(A_2-A_1)v_2\delta W_x &= \sum_{ij,i\neq j}(\xi_j-b_j)\xi_i\langle(A_2-A_1)e_i/\lambda_i,(P+A_2)e_j/\lambda_j\rangle \\ &\quad -\sum_i(\xi_i^2-1-\xi_ib_i)\langle(A_2-A_1)e_i/\lambda_i,(P+A_2)e_i/\lambda_i\rangle.\end{aligned} \tag{3.31}$$

To see part a), it therefore suffices to prove that

$$E_\epsilon(\exp\sum_{ij}\xi_i(\xi_j-b_j)T_{ij}\to_{\epsilon\to 0}1. \tag{3.32}$$

where $T_{ij}=\langle(A_2-A_1)e_i/\lambda_i,(P+A_2)e_j/\lambda_j\rangle)$. Since $\sum_i E(\xi_i)^2<\infty$, (3.32) follows from (2.13) once we prove that $T_{ij}=\langle Te_i,e_j\rangle$ for some trace class operator T. Towards this end, define the operators $\Lambda:W^{2k,2}(D)\to L^2(D)$ and $U:L^2(D)\to L^2(D)$ by $\Lambda\phi=\sum_i\lambda_i\langle\phi,e_i\rangle e_i$ and $U=(P+A_2)\Lambda^{-1}$. It is easy to verify that $T_{ij}=\langle U^*QUe_i,e_j\rangle$, and moreover $T=U^*QU$ is trace class since U is bounded and Q is trace class by assumption. The proof of part a) is completed.

To see part b) of the theorem, note that by the same proof as above, the claim follows once we show that if $\mathrm{tr}\,(T) = \infty$ then

$$\lim_{\epsilon \to 0} E\left(\exp(\sum_{i \neq j} \xi_i \xi_j T_{ij} + \sum_i (\xi^2 - 1) T_{ii}) | \sum_i \frac{\xi_i^2}{\lambda_i^2} < \epsilon\right) = 0.$$

This however follows from the assumptions of the theorem and part b) of Lemma 2.3 once one notes that T is Hilbert-Schmidt and $\sum_i b_i^2 < \infty$. □

4 Onsager-Machlup functional for comparable functions

We prove here the following theorem:

Theorem 4.1 *Let P and F satisfy conditions* (**A1**) *and* (**A2**)*. Let $\phi_1, \phi_2 \in C_o^\infty(D)$ and assume that the operator*

$$K_{\phi_1,\phi_2} \triangleq [D_{\phi_1} F - D_{\phi_2} F] P^{-1} : \ L^2(D) \to L^2(D) \tag{4.33}$$

is trace class. Further, assume that

$$\lim_{||u - \phi_i||_2 \to 0} || \, D_u F P^{-1} - D_{\phi_i} F P^{-1} \, ||_{HS} = 0 \tag{4.34}$$

where $|| \cdot ||_{HS}$ denotes the Hilbert–Schmidt norm of an operator. Finally, assume that there exists a deterministic trace class operator $\bar{T} : L^2(D) \to L^2(D)$ such that for any $u \in L^2$ such that $|| \, u - \phi_i \, ||_2$ is small enough,

$$|\langle (D_u F P^{-1} - D_{\phi_i} F P^{-1}) \psi, \psi \rangle| \le \langle \bar{T} \psi, \psi \rangle \tag{4.35}$$

for any $\psi \in L^2(D)$.

Then

$$\begin{aligned} J(\phi_1, \phi_2) \ &= \ \frac{1}{2} \int_D \left((P\phi_2 + F(\phi_2))^2 - (P\phi_1 + F(\phi_1))^2 \right) dx \\ &\quad + \log \det [I + \ \ (D_{\phi_1} F - D_{\phi_2} F)(P + D_{\phi_2} F)^{-1}] \end{aligned} \tag{4.36}$$

where $\det(A)$ *denotes the Fredholm determinant of A.*

Remark Note that (4.35) implies that, for any sequence a_i, $i = 1, \ldots,$

$$| \sum_{j,k} \langle a_k a_j (D_u F P^{-1} - D_{\phi_i} F P^{-1}) e_k, e_j \rangle | \le \sum_{j,k} a_k a_j \langle \bar{T} e_k, e_j \rangle \tag{4.37}$$

in the sense that the inequality holds whenever the RHS is smaller than ∞.

Proof of Theorem 4.1

In view of Theorem 3.1, Theorem 2.2 and an analysis similar to the one done in Section 3, it suffices to check that, for $i = 1, 2$,

$$\lim_{\epsilon \to 0} \frac{\mathrm{Prob}\,(||u - \phi_i||_2 < \epsilon)}{\mathrm{Prob}\,(||u_{\phi_i} - \phi_i||_2 < \epsilon)} = \lim_{\epsilon \to 0} E(\tilde{\Lambda} | \, ||u_{\phi_i} - \phi_i||_2 < \epsilon) = 1 \tag{4.38}$$

where

$$\tilde{\Lambda} = \exp\left(\int_D (F(u_{\phi_i}) - F(\phi_i)) - D_{\phi_i}F(u_{\phi_i} - \phi_i))\delta W_x - \frac{1}{2}\|F(u_{\phi_i}) - F(\phi_i) - D_{\phi_i}F(u_{\phi_i} - \phi_i)\|_2^2 dx\right)$$
$$\cdot \quad \det{}_2(I + D_{u_{\phi_i}}F - D_{\phi_i}F)(P + D_{\phi_i}F)^{-1} \tag{4.39}$$

is the Radon-Nykodim derivative between the measures defined by $u - \phi_i$ and $u_{\phi_i} - \phi_i$. Note however that, denoting by ϕ either ϕ_1 or ϕ_2,

$$\int_D (F(u_\phi) - F(\phi) - D_\phi F(u_\phi - \phi))^2 dx \to_{\|u_\phi - \phi\|_2 \to 0} 0$$

uniformly. Moreover, by our assumptions, the Hilbert-Schmidt norm of $(D_{u_\phi}F - D_\phi F)(P + D_\phi F)^{-1}$ converges to zero uniformly with $\|u_\phi - \phi\|_2$, and hence

$$\det{}_2(I + D_{u_\phi}F - D_\phi F)(P + D_\phi F)^{-1} \to_{\|u_\phi - \phi\|_2 \to 0} 1$$

Note next that by Taylor's generalized theorem (see [11], pg. 148) $F(u_\phi) - F(\phi) - D_\phi F(u_\phi - \phi) = \int_0^1 (D_{tu_\phi + (1-t)\phi} - D_\phi F)(u_\phi - \phi)dt$. Combining the above, it follows that it suffices to prove that

$$E(\exp \int_D \int_0^1 (D_{tu_\phi + (1-\phi)\phi}F - D_\phi F)(u_\phi - \phi)dt\delta W_x | \, \|u_\phi - \phi\|_2 < \epsilon) \to_{\epsilon \to 0} 1,$$

where the expectation is with respect to the Gaussian measure associated with u_ϕ. By using the same Karhunen-Loeve expansion as in Section 3, it follows that it is enough to prove that

$$E\left(\exp \int_0^1 (\sum_{ij}(T_{ij}^t - T_{ij}^0)(a_i + b_i)(a_j + b_j) - \sum_i (T_{ii}^t - T_{ii}^0)) | \sum_i \frac{(a_i + b_i)^2}{\lambda_i^2} < \epsilon\right) \to_{\epsilon \to 0} 1 \tag{4.40}$$

where T_{ij}^t denotes the i, j element of $D_{tu_\phi + (1-t)\phi}F(P + D_\phi F)^{-1}$. Since $\sum_i |T_{ii}^t - T_{ii}^0| \to_{\epsilon \to 0} 0$ uniformly, it is enough to show that for any constant c,

$$E(\exp c \sum_{ij} \int_0^1 (T_{ij}^t - T_{ij}^0)(a_i + b_i)(a_j + b_j)dt | \sum_i \frac{(a_i + b_i)^2}{\lambda_i^2} < \epsilon) \to_{\epsilon \to 0} 1.$$

By our assumption,

$$E(\exp c \int_0^1 \sum_{ij}(T_{ij}^t - T_{ij}^0)(a_i + b_i)(a_j + b_j)dt | \sum_i \frac{(a_i + b_i)^2}{\lambda_i^2} < \epsilon) \leq$$
$$E(\exp |c| \sum_{ij} \bar{T}_{ij}(a_i + b_i)(a_j + b_j) | \sum_i \frac{(a_i + b_i)^2}{\lambda_i^2} < \epsilon). \tag{4.41}$$

The proof is concluded by an application of part a) of Lemma 2.3. □

Examples

a) The following particular case is of interest: let P^{-1} be a positive, self adjoint Hilbert-Schmidt operator, let λ_1 denote its maximal eigenvalue. Let $F(u) = f(P^{-1}u)$, where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a C^{2k+1} function with $|f'| < 1/\lambda_1$. We check that $F(\cdot)$ satisfies the assumptions of Theorem 4.1. Indeed, note that $D_\phi F = f'(P^{-1}\phi)P^{-1}$. It follows that $K_{\phi_1,\phi_2} = (f'(P^{-1}\phi_1) - f'(P^{-1}\phi_2))P^{-2}$ is trace class being the product of a bounded (in $L^2(D)$) operator and a trace class operator there. By exactly the same argument, (4.34) holds true. Finally, to see (4.37) for an appropriate operator $\bar{T}$, note that

$$\begin{aligned} \langle P(f'(P^{-1}u) - f'(P^{-1}\phi))P^{-1}\phi, \phi\rangle &\leq \|P\|_{2k\to 0}\|(f'(P^{-1}u) - f'(P^{-1}\phi))\|_{2k\to 2k}\|P^{-1}\|_{0\to 2k}\|\phi\|_2 \\ &\leq c(\|u-\phi\|_2)\|\phi\|_2 \end{aligned} \tag{4.42}$$

where $\|\cdot\|_{a\to b}$ denotes the operator norm from $W^{a,2}(D)$ to $W^{b,2}(D)$ and $c(\|u-\phi\|_2)$ denotes a constant which depends only on $\|u-\phi\|_2$. In particular, for smooth ϕ such that $\psi = P\phi \in L^2(D)$, it follows that

$$\langle (f'(P^{-1}u) - f'(P^{-1}\phi))P^{-2}\psi, \psi\rangle \leq c(\|u-\phi\|_2)\langle P^{-1}\psi, P^{-1}\psi\rangle = c(\|u-\phi\|_2)\langle \psi, P^{-2}\psi\rangle. \tag{4.43}$$

It follows that by taking $\bar{T} = c(\|u-\phi\|_2)P^{-2}$, (4.37) holds.

b) To see a truly nonlinear situation in which $J(\phi_1, \phi_2)$ is degenerate, let P^{-1} be Hilbert–Schmidt but not trace class on $L^2(D)$, and let $F(\cdot)$ be a smooth function such that $F(u) = c|u|$ for $|u| > 1/2$, where $|c| < 1/\lambda_1$. Let $\phi_1 = 1$, $\phi_2 = -1$. It is easy to check that, due to the local nature of the conditions in Theorem 4.1, (4.33,4.34,4.37) are satisfied, and hence by following the argument in the proof of the theorem, it follows that (4.38) holds true. On the other hand, part b) of Theorem 3.1 applies in this situation, and combining the two one concludes that the Onsager-Machlup functional is trivial.

5 Appendix

Lemma 5.1 *In addition to* **(A2)** *assume that F is linear. Then $G_F : L^2(D) \longrightarrow W^{\lfloor 2k-d/2\rfloor,2}(D)$ is Hilbert–Schmidt.*

Proof: Assumption **(A2)** implies that $FG_0 : L^2(D) \rightarrow L^2(D)$ is bounded with $\|FG_0\| < 1$, from which it may be concluded that $I + FG_0$ has a bounded inverse in $L^2(D)$. Since $G_F = G_0(I + FG_0)^{-1}$, we may assume with no loss of generality that $F \equiv 0$, in which case the result has already been shown to hold as a result of Maurin's theorem (cf. the introduction). □

Proof of Theorem 2.2: Recall that $\mathcal{P}_W$ is the probability measure associated with the Brownian sheet. Denote $E_0 = \{\omega \in C_0(D) : \partial\omega \in W^{-2k,2}(D)\}$ and recall that $\mathcal{P}_W(E_0) = 1$. Next, define $T, S : C_0(D) \rightarrow C_0(D)$ by

$$T\omega = \begin{cases} \omega + \partial^{-1}(F_1 - F_2)(G_{F_2}\partial\omega) = \partial^{-1}(P + F_1)(G_{F_2}\partial\omega) & \omega \in E_0 \\ \omega & \omega \in C_0(D)\backslash E_0 \end{cases} \tag{5.44}$$

$$S\omega = \begin{cases} \omega + \partial^{-1}(F_2 - F_1)(G_{F_1}\partial\omega) = \partial^{-1}(P + F_2)(G_{F_1}\partial\omega) & \omega \in E_0 \\ \omega & \omega \in C_0(D)\backslash E_0 \end{cases} \tag{5.45}$$

It follows by inspection (use the second definition of S and T) that $TS\omega = ST\omega = \omega$, $\forall\omega \in C_0(D)$, i.e. T is a bijection. We now define a new probability measure on $(\Omega, \mathcal{F})$, $\mathcal{P} = \mathcal{P}_W \circ T$. An expression for the Radon Nikodym derivative $\frac{d\mathcal{P}}{d\mathcal{P}_W}$ will be provided by Kusuoka's theorem ([7, Theorem 6.4]) as long as the following conditions hold:

(a) $T_0 \equiv T - I : C_0(D) \to H(D)$ and possesses a Hilbert-Schmidt $H(D)$–Fréchet derivative $D_\omega T_0 : H(D) \to H(D)$ at every $\omega \in C_0(D)$

(b) $h \in H(D) \to D_{\omega+h}T_0$ is continuous in the Hilbert–Schmidt norm $\forall\omega \in C_0(D)$.

(c) $I_{H(D)} + D_\omega T_0 : H(D) \to H(D)$ is invertible $\forall\omega \in C_0(D)$.

In verifying (a)–(c) note first that both E_0 and $C_0(D)\backslash E_0$ are closed under perturbations by elements of $H(D)$ so that each row in (5.44) and (5.45) may be considered separately. Moreover when $\omega \in C_0(D)\backslash E_0$ everything becomes trivial so it will be assumed that $\omega \in E_0$.

Concerning (a), the range of $T_0 = \partial^{-1}(F_1 - F_2)G_{F_2}\partial$ is clearly in $H(D)$ since the range of $(F_1 - F_2)$ is in $L^2(D)$. We obtain the $H(D)$–Fréchet derivative by the standard chain and inverse differentiation rules:

$$D_\omega T_0 h = \partial^{-1}D_{(G_{F_2}\partial\omega)}(F_1 - F_2)G_{D_{(G_{F_2}\partial\omega)}F_2}\partial h \qquad h \in H(D). \tag{5.46}$$

This operator may be described by the graph

$$H(D) \xrightarrow{\partial} L^2(D) \xrightarrow{B_\omega \equiv G_{A^{(2)}_\omega}} W^{\lfloor 2k-d/2 \rfloor,2}(D) \xrightarrow{A_\omega \equiv A^{(1)}_\omega - A^{(2)}_\omega} L^2(D) \xrightarrow{\partial^{-1}} H(D)$$

where

$$A^{(i)}_\omega = D_{(G_{F_2}\partial\omega)}F_i \qquad i = 1, 2.$$

Clearly ∂ and ∂^{-1} are isometries, so we may actually restrict our attention to the operator $A_\omega B_\omega$. Now, A_ω is bounded by assumption and B_ω is Hilbert–Schmidt by lemma 5.1, from which it may be concluded that the composition is Hilbert–Schmidt. Moreover

$$||D_\omega T_0||_{HS} \leq ||A_\omega||\ ||B_\omega||_{HS}. \tag{5.47}$$

As for (b),

$$||A_{\omega+h}B_{\omega+h} - A_\omega B_\omega||_{HS} \leq ||A_{\omega+h}||\ ||B_{\omega+h} - B_\omega||_{HS} + ||A_{\omega+h} - A_\omega||\ ||B_\omega||_{HS}. \tag{5.48}$$

When $h \to 0$, $G_{F_2}\partial(\omega + h) \to (G_{F_2}\partial\omega)$ in $W^{\lfloor 2k-d/2 \rfloor,2}(D)$ so that

$$||A^{(i)}_{\omega+h} - A^{(i)}_\omega|| \to 0 \tag{5.49}$$

by the continuity property of the Fréchet derivative assumed in **(A2)**; this takes care of the second term in the right hand side of (5.48), while the first term can be seen to converge to 0 as $h \to 0$ by writing $B_{\omega+h} - B_\omega = B_{\omega+h}(A^{(2)}_\omega - A^{(2)}_{\omega+h})B_\omega$, by noting that $||A_{\omega+h}||$ and $||B_{\omega+h}||$ are uniformly bounded in h (which follows from assumption **(A2)**) and by applying (5.49) once again.

Finally, property (c) follows by applying the chain rule to the identity $ST\omega = \omega$. Namely one has $D_{T\omega}S\ D_\omega T = I_{HD}$ (where D always stands for the $H(D)$–Fréchet derivative operator), which shows that $D_\omega T = I_{H(D)} + D_\omega T_0$ is invertible. Here we have made use of the fact S is $H(D)$–Fréchet differentiable just as T was previously shown to be, since both transformations are of the same form differing only in that the roles of F_1 and F_2 are interchanged.

We may now conclude from [7, Theorem 6.4] that $\mathcal{P} \ll \mathcal{P}_W$ and $\frac{d\mathcal{P}}{d\mathcal{P}_W} = \Lambda_{1,2}$ (defined in (2.10)). By definition,

$$\tilde{W}(\omega) \equiv W(T\omega) = \left(W + \partial^{-1}(F_1 - F_2)(u_2)\right)(\omega)$$

is a Brownian sheet with respect to $\tilde{P}$. Since $\forall\phi \in W_0^{2k,2}(D)$

$$\int_D [u_2 P^*\phi + F_2(u_2)\phi]\, dx = \int_D \phi dW_x = \int_D \phi \left[d\tilde{W}_x - (F_1 - F_2)(u_2)dx\right]$$

we obtain $\forall\phi \in W_0^{2k,2}(D)$

$$\int_D [u_2 P^*\phi + F_1(u_2)]\, dx = \int_D \phi d\tilde{W}_x.$$

By the uniqueness in law, u_2 solves (2.8) with respect to $\tilde{P}$ for $F = F_1$. Thus, for any Borel set $B \subset W_0^{2k,2}(D)$

$$\begin{aligned} \frac{P(u_1 \in B)}{P(u_2 \in B)} &= \frac{\tilde{P}(u_2 \in B)}{P(u_2 \in B)} = \frac{\int_\Omega \Lambda_{1,2}(\omega)\mathbf{1}_{\{u_2(\omega)\in B\}}(\omega)dP(\omega)}{P(u_2 \in B)} \\ &= E\left(\Lambda_{1,2} | \, u_2 \in B\right). \end{aligned}$$

□

References

[1] Adams, R., Sobolev Spaces, Academic Press (1975).

[2] Dembo, A. and Zeitouni, O., "Maximum a-posteriori estimation of elliptic Gaussian fields observed via a nonlinear channel", J. Multivariate Analysis, 35 (1990), pp. 411-425.

[3] Dembo, A. and Zeitouni, O., "Onsager Machlup functionals and maximum a posteriori estimation for a class of non Gaussian random fields", J. Multivariate Analysis, 36 (1991), pp. 243-262.

[4] Dunford, E. and Shwartz, J.T., Linear Operators, Interscience publishers, 1957.

[5] Esary, J.D., Proschan, F. and Walkup, D.W., " Association of random variables, with applications", Ann. Math. Stat. 38 (1967), pp. 1466-1474.

[6] Ikeda, N. and Watanabe, S., Stochastic Diferential Equations and Diffusion Processes, 2nd edition, North-Holland, 1989.

[7] Kusuoka, S., " The nonlinear transformation of Gaussian measure on Banach space and its absolute continuity", J. Fac. Sci. Tokyo Univ., Sec I.A. (1985), pp. 567-597.

[8] Nualart, D. and Zakai, M., "Generalized stochastic integrals and the Malliavin calculus", Prob. Theory and Related Fields, 73 (1986), pp. 255-280.

[9] Shepp, L.A. and Zeitouni, O., "A note on conditional exponential moments and the Onsager Machlup functional", to appear, Annals of Probability.

[10] Walsh, J.B., "An introduction to stochastic partial differential equations", in Lecture Notes in Mathematics # 1180, (1986), pp. 266-437. Springer-Verlag, Berlin– New-York.

[11] Zeidler, E., Nonlinear Functional Analysis and its Applications, Vol. I, Springer, 1986.