A short proof of Schilder's large deviation theorem

Abstract

The note presents the short proof by S. Chevet [3] of Schilder's large deviation theorem for the Wiener measure [10]. The argument covers more generally the case of arbitrary Gaussian measures on a Banach space.

On some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, let $W = (W(t))_{t \in [0,1]}$ be a standard Brownian motion or Wiener process with law the Wiener measure μ on the Banach space $E = C([0,1])$ of real continuous functions on $[0,1]$ and reproducing kernel Hilbert space $\mathcal{H}$ identified as the subspace of $E = C([0,1])$ consisting of the absolutely continuous functions $h : [0,1] \to \mathbb{R}$, with almost everywhere derivative h' in $\mathrm{L}^2([0,1])$ (for the Lebesgue measure).

Setting for $h \in \mathcal{H}$,

$$|h|_{\mathcal{H}} = \left(\int_0^1 h'(t)^2 \, dt \right)^{1/2},$$

the rate function $\mathcal{I} : E \to [0, +\infty]$ which will govern the large deviation properties of εW as $\varepsilon \to 0$ is defined as

$$\mathcal{I}(x) = \begin{cases} \frac{1}{2}|x|_{\mathcal{H}}^2 & \text{if } x \in \mathcal{H}, \\ +\infty & \text{if } x \notin \mathcal{H}. \end{cases} \tag{1}$$

In the large deviation language ([11, 5, 4]), this rate function is a good rate function in the sense that its level sets $\{\mathcal{I} \leq a\}$, $a \geq 0$, are compact in E (due to the compactness of the $\mathcal{H}$-balls in E).

The following theorem, going back to the work of M. Schilder [10], presents the large deviation behavior of the law of εW as $\varepsilon \to 0$. For a subset A of E, let

$$\mathcal{I}(A) = \inf_{x \in A} \mathcal{I}(x).$$

Theorem 1 (Schilder's large deviation theorem)**.** *For any closed set F in E,*

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon W \in F) \,\leq\, -\mathcal{I}(F).$$

For any open set O in E,

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon W \in O) \,\geq\, -\mathcal{I}(O).$$

The purpose of this note is to present a short proof of this result in the general context of a Gaussian measure on a Banach space as emphasized by S. Chevet [3], relying on the Gaussian isoperimetric inequality, actually milder Gaussian concentration properties. General references on (Gaussian) large deviations include [11, 5, 4, 9, 2] etc.

1 The Gaussian large deviation principle

Let X be a centered Gaussian random vector on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a real separable Banach space E equipped with its Borel σ-algebra $\mathcal{B}$, and with norm $\|\cdot\|$. The law μ of X on the Borel sets of E gives rise to an abstract Wiener space structure $(E, \mathcal{H}, \mu)$, in which the Hilbert space $\mathcal{H} \subset E$, with scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, is the reproducing kernel Hilbert space associated to the covariance structure of μ (cf. [7, 5, 8, 2]...). In case of the Wiener measure on $E = C([0,1])$, $\mathcal{H}$ is identified with the subspace of E consisting of the absolutely continuous functions with almost everywhere derivative in $\mathrm{L}^2([0,1])$ as described above.

Large deviations for general Gaussian measures go back to M. Donsker and S. Varadhan [6]. The study of [6] actually addresses the large deviation principle for sums of independent Banach space valued random variables, the Gaussian case being a particular case.

Consider as before the rate function $\mathcal{I}$ as defined in (1), and for a subset A of E, set similarly $\mathcal{I}(A) = \inf_{x \in A} \mathcal{I}(x)$.

Theorem 2 (The Gaussian large deviation principle)**.** *For any closed set F in E,*

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in F) \,\leq\, -\mathcal{I}(F). \tag{2}$$

For any open set O in E,

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in O) \,\geq\, -\mathcal{I}(O). \tag{3}$$

Applied to complements of balls, this theorem easily produces the limit

$$\lim_{t \to \infty} t^2 \log \mathbb{P}\big(\|X\| \geq t\big) \,=\, -\frac{1}{2\sigma^2}$$

where

$$\sigma^2 = \sup_{\xi \in E^*, \|\xi\| \le 1} E\big(\langle \xi, X \rangle^2\big)$$

together with the observation that $\sigma = \sup_{|h|_{\mathcal{H}} \le 1} \|h\|$.

2 Proof of the upper-bound (2)

The proof by S. Chevet [3] of the upper-bound (2) in Theorem 2 relies on isoperimetric and concentration inequalities which provide a very convenient tool.

Let F be closed in E, and take r such that $0 < r < \mathcal{I}(F)$. By the very definition of $\mathcal{I}(F)$,

$$F \cap \sqrt{2r}\, \mathcal{K} = \emptyset,$$

where $\mathcal{K}$ is the (closed) unit ball in $\mathcal{H}$. Since F is closed and $\mathcal{K}$ is compact in E, there exists $\eta > 0$ such that it still holds true that

$$F \cap \big[\sqrt{2r}\, \mathcal{K} + B_E(0, \eta)\big] = \emptyset$$

where $B_E(0, \eta)$ is the ball with center the origin and radius η for the norm $\| \cdot \|$ in E. Clearly

$$\lim_{\varepsilon \to 0} \mathbb{P}\big(\varepsilon X \in B_E(0, \eta)\big) = \lim_{\varepsilon \to 0} \mathbb{P}\big(X \in B_E(0, \tfrac{\eta}{\varepsilon})\big) = 1.$$

The Gaussian isoperimetric inequality for the law of X (cf. [9, 8, 2]) expresses that, whenever $\mathbb{P}(X \in A) \ge \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{1}{2}x^2} dx$ for some $a \in \mathbb{R}$,

$$\mathbb{P}(X \in A + s\, \mathcal{K}) \ge \Phi(a + s)$$

for every $s \ge 0$. For $\varepsilon > 0$ small enough, $\mathbb{P}(X \in B_E(0, \frac{\eta}{\varepsilon})) \ge \frac{1}{2} = \Phi(0)$. Hence,

$$\mathbb{P}(\varepsilon X \in F) \le \mathbb{P}\big(\varepsilon X \notin \sqrt{2r}\, \mathcal{K} + B_E(0, \eta)\big) \le 1 - \Phi\bigg(\frac{\sqrt{2r}}{\varepsilon}\bigg) \le e^{-r/\varepsilon^2}.$$

Therefore

$$\limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in F) \le -r,$$

which is the result since $r < \mathcal{I}(F)$ is arbitrary.

As mentioned above, the full strength of the Gaussian isoperimetric inequality is not really needed, and weaker concentration inequalities are enough to achieve the conclusion. For example, as emphasized in [8], the latter easily produce that

$$\mathbb{P}(X \in A + s\, \mathcal{K}) \ge 1 - e^{-\frac{1}{2}s^2 + \delta(\mu(A))s}$$

for every $s \ge 0$, where $\delta(\mu(A)) \to 0$ as $\mu(A) \to 1$, so that the proof may be developed similarly.

3 Proof of the lower-bound (3)

The lower-bound (3) in Theorem 2 classically relies on the Cameron-Martin translation formula which formulates, in the present general setting, that, for any h in $\mathcal{H}$, the shifted probability measure $\mu(\cdot + h)$ is absolutely continuous with respect to μ, with density given by the formula

$$\mu(B+h) \,=\, e^{-\frac{1}{2}|h|^2_{\mathcal{H}}} \int_B e^{-\widetilde{h}} d\mu \tag{4}$$

for every Borel set B in E, where $\widetilde{h} : E \to \mathbb{R}$ is (centered) Gaussian under μ with variance $|h|^2_{\mathcal{H}}$ [9, 8, 2]. In case of Brownian motion $W = (W(t))_{t\in[0,1]}$ on $E = C([0,1])$ $\widetilde{h} = \int_0^1 h'(t)dW(t)$, so that the shifted measure $\mu(\cdot + h)$ has density

$$\exp\bigg(-\frac{1}{2}\int_0^1 h'(t)^2\, dt - \int_0^1 h'(t)dW(t)\bigg)$$

with respect to μ.

Let then $h \in O \cap \mathcal{H}$. Since O is open, there exists $\eta > 0$ such that $h + B_E(0,\eta) \subset O$, and thus

$$\mathbb{P}(\varepsilon X \in O) \,\geq\, \mathbb{P}\big(\varepsilon X \in h + B_E(0,\eta)\big).$$

The Cameron-Martin translation formula (4) therefore yields that

$$\begin{aligned} \mathbb{P}\big(\varepsilon X \in h + B_E(0,\eta)\big) \,&=\, \mu\big(\tfrac{h}{\varepsilon} + B_E(0,\tfrac{\eta}{\varepsilon})\big) \\ &=\, \exp\bigg(-\frac{|h|^2_{\mathcal{H}}}{2\varepsilon^2}\bigg) \int_{B_E(0,\frac{\eta}{\varepsilon})} \exp\bigg(-\frac{\widetilde{h}}{\varepsilon}\bigg) d\mu. \end{aligned} \tag{5}$$

By Jensen's inequality,

$$\int_{B_E(0,\frac{\eta}{\varepsilon})} \exp\bigg(-\frac{\widetilde{h}}{\varepsilon}\bigg) d\mu \,\geq\, \mu\big(B_E(0,\tfrac{\eta}{\varepsilon})\big) \exp\bigg(-\int_{B_E(0,\frac{\eta}{\varepsilon})} \frac{\widetilde{h}}{\varepsilon} \cdot \frac{d\mu}{\mu(B_E(0,\frac{\eta}{\varepsilon}))}\bigg).$$

Now

$$\int_{B_E(0,\frac{\eta}{\varepsilon})} \widetilde{h}\, d\mu \,\leq\, \int_E |\widetilde{h}| d\mu \,\leq\, \bigg(\int_E \widetilde{h}^2 d\mu\bigg)^{1/2} \,=\, |h|_{\mathcal{H}}.$$

For every $\varepsilon > 0$ small enough, $\mu(B_E(0,\frac{\eta}{\varepsilon})) \geq \frac{1}{2}$ (for example). As a consequence of the preceding,

$$\int_{B_E(0,\frac{\eta}{\varepsilon})} \exp\bigg(-\frac{\widetilde{h}}{\varepsilon}\bigg) d\mu \,\geq\, \frac{1}{2} \exp\bigg(-\frac{2|h|_{\mathcal{H}}}{\varepsilon}\bigg).$$

Implementing into (5), for $\varepsilon > 0$ small enough,

$$\mathbb{P}(\varepsilon X \in O) \,\geq\, \frac{1}{2} \exp\bigg(-\frac{|h|^2_{\mathcal{H}}}{2\varepsilon^2} - \frac{2|h|_{\mathcal{H}}}{\varepsilon}\bigg)$$

from which it follows that

$$\liminf_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}(\varepsilon X \in O) \;\geq\; -\frac{1}{2}\,|h|^2_{\mathcal{H}} \;=\; -\mathcal{I}(h).$$

This result for any $h \in O \cap \mathcal{H}$ yields the announced lower-bound (3).

The preceding combined upper and lower arguments actually produce a measurable version of the large deviation principle, without referring to any topology associated to the underlying abstract Wiener space (cf. [1, 8]).

References

[1] G. Ben Arous, M. Ledoux. Schilder's large deviation principle without topology. *Pitman Res. Notes Math. Ser.* 284, 107–121 (1993).

[2] V. Bogachev. *Gaussian measures.* Math. Surveys Monogr. 62. American Mathematical Society (1998).

[3] S. Chevet. Gaussian measures and large deviations. Probability in Banach spaces IV. Lecture Notes in Math. 990, 30-46 (1983). Springer.

[4] A. Dembo, O. Zeitouni. *Large Deviations Techniques and Applications.* Springer (1998).

[5] J.-D. Deuschel, D. Stroock. *Large deviations.* Academic Press (1989).

[6] M. D. Donsker, S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time III. *Comm. Pure Appl. Math.* 29, 389–461 (1976).

[7] L. Gross. Abstract Wiener spaces. *Proc. 5th Berkeley Symp. Math. Stat. Prob.* 2, 31–42 (1965).

[8] M. Ledoux. Isoperimetry and Gaussian Analysis. *École d'Été de Probabilités de St-Flour 1994. Lecture Notes in Math.* 1648, 165–294. Springer (1996).

[9] M. Lifshits. *Gaussian random functions.* Kluwer (1994).

[10] M. Schilder. Asymptotic formulas for Wiener integrals. *Trans. Amer. Math. Soc.* 125, 63–85 (1966).

[11] S. R. S. Varadhan. *Large Deviations and Applications.* SIAM Publications, Philadelphia (1984).