MICHEL LEDOUX

# A note on large deviations for Wiener chaos

*Séminaire de probabilités (Strasbourg)*, tome 24 (1990), p. 1-14
<http://www.numdam.org/item?id=SPS_1990__24__1_0>

# A note on large deviations for Wiener chaos

by *Michel Ledoux*

The result of this note is well-known and familiar (it is presented for example, using standard techniques, in the recent work [D-S]). Its purpose is to describe the usefulness and interest of *isoperimetric* methods in large deviation theorems and we present here a simple isoperimetric proof of the large deviation properties for homogeneous Gaussian chaos (even vector valued). The approach suggests some possible further use of isoperimetry in this type of question.

The proof we give is based on, and may be considered as a simple outgrow of, the study by C. Borell [Bo3], [Bo5]. This exposition was actually the opportunity for the author to try to understand and hopefully clarify for the possible readers some aspects of the deep and unfortunately somewhat difficult to read work by C. Borell that develops all the necessary material for the study of this problem.

Let $(E, H, \mu)$ be an abstract Wiener space. That is, let $E$ be a real Banach space with Borel $\sigma$-algebra $\mathcal{B}$ and dual space $E'$. Let further $\mu$ denote a centered Gaussian Radon probability measure on $(E, \mathcal{B})$ in the sense that the law of $\xi \in E'$ under $\mu$ is a real mean zero normal variable with variance $\int \langle \xi, x\rangle^2 d\mu(x)$. By the closed graph theorem, the injection map $E' \to L^2(\mu; \mathbb{R}) = L^2((E,\mu); \mathbb{R})$ is continuous. Since $\mu$ is Radon (i.e. supported by a separable subspace of $E$), it follows further that for each $\xi$ in $E'$, the weak integral $\int x\langle \xi, x\rangle d\mu(x)$ defines an element of $E$. By density, one can map any element $\varphi$ of the closure $E'_2$ of $E'$ in $L^2(\mu; \mathbb{R})$ into an element $\Lambda(\varphi) = \int x\varphi(x)d\mu(x)$ of $E$ and this map is linear and injective. Define then $H$ to be the range of $\Lambda$. Equipped with the natural scalar product $\langle \Lambda(\varphi), \Lambda(\psi)\rangle_H = \langle \varphi, \psi\rangle_{L^2(\mu;\mathbb{R})}$, $H$ is a separable Hilbert space (with norm denoted by $|\cdot|$), dense in the support of $\mu$ and known as the reproducing kernel Hilbert space of the measure $\mu$. Its unit ball $\mathcal{O}$ is a compact subset of $E$. For any orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of $E'_2$, $\mu$ has the same distribution as $\sum_k e_k\Lambda(e_k)$. (This fact puts forward the fundamental Gaussian measurable structure consisting of the canonical Gaussian product measure on $\mathbb{R}^{\mathbb{N}}$ with reproducing kernel Hilbert space $\ell^2$). If $h = \Lambda(\varphi)$ is an element of $H$, we set for simplicity $\widetilde{h} = \Lambda^{-1}(h) = \varphi$; under $\mu$, $\widetilde{h}$ is Gaussian with variance $|h|^2$. Recall finally the Cameron-Martin translation formula [C-M] that indicates that, for any $h$ in $H$, the probability measure $\mu(\cdot + h)$ is absolutely continuous with respect to $\mu$, with density $\exp(\widetilde{h}(\cdot) - |h|^2/2)$.

This classical construction (see e.g. [Ne1], [Ku], [Fe], etc) may be extended to locally convex Hausdorff vector spaces $E$ equipped with a Gaussian Radon probability measure $\mu$ ([Bo2]), but, for the modest purposes of this note, we restrict ourselves to the preceding setting. As an example also, let us mention the classical Wiener space associated with Brownian motion, say on [0,1] for simplicity. Let thus $E$ be the Banach space $C_0([0,1])$ of all real continuous functions $x$ on [0,1] vanishing at the origin and let $\mu$ be the distribution of a standard Brownian motion $(B(t))_{t\in[0,1]}$ starting at the origin . If $m$ is a finitely supported measure on [0,1], $m = \sum_i c_i\delta_{t_i}$, $c_i \in \mathbb{R}$, $t_i \in [0,1]$, clearly $h = \Lambda(m)$ is the element of $E$ given by

$$h(t) = \sum_i c_i(t_i \wedge t), \quad t \in [0,1];$$

it satisfies

$$\int_0^1 h'(t)^2\, dt = \int \langle m, x\rangle^2 d\mu(x).$$

By a standard extension, the reproducing kernel Hilbert space $H$ associated to $\mu$ on $E$ may be identified with the absolutely continuous elements $h$ of $C_0([0,1])$ such that $\int_0^1 h'(t)^2\, dt < \infty$ and $\widetilde{h} = \int_0^1 h'(t) dB(t)$.

Let $(e_k)_{k\in\mathbb{N}} \subset E'$ be any fixed orthonormal basis of $E'_2$ (take any weak-star dense sequence of the unit ball of $E'$ and orthonormalize it with respect to $\mu$ using the Gramm-Schmidt procedure; we choose it in $E'$ for convenience and without any loss in generality). Denote by $(h_k)_{k\in\mathbb{N}}$ the sequence of the Hermite polynomials defined from the generating series

$$\exp(\lambda x - \lambda^2/2) = \sum_{k=0}^{\infty} \lambda^k h_k(x), \quad \lambda, x \in \mathbb{R}.$$

$(\sqrt{k!}\, h_k)$ is an orthonormal basis of $L^2(\gamma; \mathbb{R})$ where $\gamma$ is the canonical Gaussian measure on $\mathbb{R}$. If $\alpha = (\alpha_0, \alpha_1, \ldots) \in \mathbb{N}^{(\mathbb{N})}$, i.e. $|\alpha| = \alpha_0 + \alpha_1 + \cdots < \infty$, set

$$H_\alpha = \sqrt{\alpha!} \prod_k h_{\alpha_k} \circ e_k$$

(where $\alpha! = \alpha_0!\alpha_1! \cdots$). Then the family $(H_\alpha)$ constitutes an orthonormal basis of $L^2(\mu; \mathbb{R})$.

Let now $B$ be a real separable Banach space with norm $\|\cdot\|$. $L^p((E,\mu); B) = L^p(\mu; B)$ $(0 \le p < \infty)$ is the space of all Bochner measurable functions $f$ on $(E, \mu)$

with values in $B$ ($p = 0$) such that $\int \|f\|^p d\mu < \infty$ ($0 < p < \infty$). For each integer $d$, set

$$\mathcal{H}^{(d)}(\mu; B) = \{f \in L^2(\mu; B); \langle f, H_\alpha \rangle = \int f H_\alpha d\mu = 0 \text{ for all } \alpha \text{ such that } |\alpha| \neq d\}.$$

$\mathcal{H}^{(d)}(\mu; B)$ defines the $B$-valued homogeneous Wiener chaos of degree $d$ [Wi]. An element $f$ of $\mathcal{H}^{(d)}(\mu; B)$ can be written as

$$f = \sum_{|\alpha|=d} \langle f, H_\alpha \rangle \, H_\alpha$$

where the multiple sum is convergent (for any finite filtering) $\mu$-almost everywhere and in $L^2(\mu; B)$. (Actually, as a consequence of [Bo3], [Bo4], or the subsequent main result, this convergence also takes place in $L^p(\mu; B)$ for any $p$.) To see it, let, for each $n$, $\mathcal{B}_n$ be the sub-$\sigma$-algebra of $\mathcal{B}$ generated by the functions $e_0, \ldots, e_n$ on $E$ and let $f_n$ be the conditional expectation of $f$ with respect to $\mathcal{B}_n$. Recall that $\mathcal{B}$ may be assumed to be generated by $(e_k)_{k \in \mathbb{N}}$. Then

$$f_n = \sum_{\substack{|\alpha|=d \\ \alpha_i = 0, i > n}} \langle f, H_\alpha \rangle \, H_\alpha \tag{1}$$

as can be checked on linear functionals, and therefore, by the vector valued martingale convergence theorem (cf. [Ne2]), the claim follows.

As a consequence of the Cameron-Martin formula, we may define for any $f$ in $L^0(\mu; B)$ and $h$ in $H$, a new element $f(\cdot + h)$ of $L^0(\mu; B)$. Further, if $f$ is in $L^2(\mu; B)$, for any $h \in H$,

$$\int \|f(x+h)\| \, d\mu(x) \leq \exp(|h|^2/2) \left( \int \|f(x)\|^2 d\mu(x) \right)^{1/2}. \tag{2}$$

Indeed,

$$\int \|f(x+h)\| \, d\mu(x) = \int \exp(\widetilde{h}(x) - |h|^2/2) \, \|f(x)\| \, d\mu(x)$$

from which (2) follows by Cauchy-Schwarz inequality and the fact that $\widetilde{h}(\cdot)$ is Gaussian with variance $|h|^2$.

Let $f$ be in $L^2(\mu; B)$. By (2), for any $h$ in $H$, we can define an element $f^{(d)}(h)$ of $B$ by setting

$$f^{(d)}(h) = \int f(x+h) \, d\mu(x).$$

If $f \in \mathcal{H}^{(d)}(\mu; B)$, $f^{(d)}(h)$ is homogeneous of degree $d$. To see it, we can work by approximation on the $f_n$'s and use then the easy fact (checked on the generating series for example) that, for any real number $\lambda$ and any integer $k$,

$$\int h_k(x+\lambda)\, d\gamma(x) = \frac{1}{k!}\, \lambda^k.$$

Actually, $f^{(d)}(h)$ can be written as the convergent multiple sum

$$f^{(d)}(h) = \sum_{|\alpha|=d} \langle f, H_\alpha \rangle\, h^\alpha$$

where $h^\alpha$ is meant as $e_0(h)^{\alpha_0} e_1(h)^{\alpha_1} \cdots$.

Given thus $f$ in $\mathcal{H}^{(d)}(\mu; B)$, for any $s$ in $B$, set $I_f(s) = |h|^2/2$ whenever $s = f^{(d)}(h)$ for some $h$ in $H$ and $I_f(s) = +\infty$ otherwise. For a subset $S$ of $B$, set $I_f(S) = \inf_{s \in S} I_f(s)$.

We can now state the large deviation properties for the elements $f$ of $\mathcal{H}^{(d)}(\mu; B)$ (see thus [D-S]). We give a new and isoperimetric proof of this result. The case $d = 1$ of course corresponds to the classical large deviation result for Gaussian measures (cf. e.g. [Az], [St1], [D-S], ...). In order to emphasize the interest of isoperimetric methods in this context, we briefly describe below the proof of the upper bound in the case $d = 1$ (and for $\mu$ itself, that is for $f$ the identity map on $E = B$). The proof for higher order chaos will be simply an appropriate extension of this argument.

THEOREM. *Let $\mu_\varepsilon(\cdot) = \mu(\varepsilon^{-1/2}(\cdot))$, $\varepsilon > 0$. Let $d$ be an integer and let $f$ be an element of $\mathcal{H}^{(d)}(\mu; B)$. Then, if $F$ is a closed subset of $B$,*

(i) $$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(x; f(x) \in F) \le -I_f(F).$$

*If $G$ is an open subset of $B$,*

(ii) $$\liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(x; f(x) \in G) \ge -I_f(G).$$

The proof of part (ii) of the theorem follows rather easily from the Cameron-Martin translation formula. Part (i) is rather easy too, but our approach thus rests on the deeper tool of isoperimetric inequalities (first used in the context of large deviations by S. Chevet [Ch]). The *isoperimetric property* of Gaussian measures $\mu$ indicates that if $A$ is a Borel set with measure $\mu(A) = \gamma((-\infty, a])$, $a \in [-\infty, +\infty]$, where $\gamma$ is the canonical Gaussian measure on $\mathbb{R}$, then, for all $t \ge 0$,

(3) $$\mu_*(A + t\mathcal{O}) \ge \gamma((-\infty, a + t])$$

where $\mathcal{O}$ is the unit ball of $H$ and $A + t\mathcal{O} = \{a + th; a \in A, h \in \mathcal{O}\}$ (that is not necessarily measurable justifying therefore the use of the inner measure). In other words, half-spaces are extremal sets for the isoperimetric property on Gauss spaces $(E, \mu)$. The isoperimetric inequality (3) has been established independently in [Bo1] and [S-T] as a consequence of the isoperimetric inequality on the sphere via the Poincaré limit (see [MK]) ; a more intrisic proof was given by A. Ehrhard [Eh]. We will use it in its following simple consequence : if $\mu(A) \geq 1/2$, for all $t \geq 0$,

$$\mu_*(A + t\mathcal{O}) \geq 1 - \exp(-t^2/2) \tag{4}$$

(take $a = 0$ in (3)). In this form, or in a slightly weaker formulation, it may be obtained from rather simple considerations (using for example stochastic calculus) as was shown by B. Maurey and G. Pisier (cf. [Pi], [Le]).

As announced, let us briefly show, using (4), the upper bound (i) for $\mu$, with for simplicity $B = E$ and $f$ the identity map. Let thus $F$ be closed in $E$ and take $0 < r < \inf_{h \in F} |h|^2/2$ so that $(2r)^{1/2}\mathcal{O} \cap F = \emptyset$. Since $\mathcal{O}$ is compact in $E$, there is $\delta > O$ such that

$$((2r)^{1/2}\mathcal{O} + \delta U) \cap F = \emptyset$$

where $U$ is the unit ball of $E$. We can then simply write that, for all $\varepsilon > 0$,

$$\mu_\varepsilon(F) \leq \mu^*(x; x \notin \delta\varepsilon^{-1/2}U + (2r/\varepsilon)^{1/2}\mathcal{O}).$$

Since, for $\varepsilon$ small enough, $\mu(\delta\varepsilon^{-1/2}U) \geq 1/2$, we immediately get that

$$\mu_\varepsilon(F) \leq \exp(-r/\varepsilon)$$

which gives the result since $r < \inf_{h \in F} |h|^2/2$ is arbitrary.

The proof of (i) also sheds some light on the structure of Gaussian polynomials as developed by C. Borell, and in particular the homogeneous structures. As is clear indeed from [Bo3], [Bo5] (and the proof below), the theorem may be shown to hold for all Gaussian polynomials, i.e. elements of the closure in $L^0(\mu; B)$ of all continuous polynomials from $E$ into $B$ of degree less than or equal to $d$. As we will see, $\mathcal{H}^{(d)}(\mu; B)$ may be considered as a subspace of all *homogeneous* Gaussian polynomials of degree $d$ (at least if $\mu$ is infinite dimensional), and hence, the elements of $\mathcal{H}^{(d)}(\mu; B)$ are $\mu$-almost everywhere $d$-homogeneous. In particular, (i) and (ii) of the theorem are equivalent to say that (changing moreover $\varepsilon$ into $t^{-2}$)

$$\limsup_{t \to \infty} \frac{1}{t^2} \log \mu(x; f(x) \in t^d F) \leq -I_f(F) \tag{i'}$$

and

$$\liminf_{t \to \infty} \frac{1}{t^2} \log \mu(x; f(x) \in t^d G) \geq -I_f(G), \tag{ii'}$$

and these are the properties we will actually establish.

Before turning to the proof of the theorem, let us mention a few applications and illustrations. If we take $F$ and $G$ in the theorem to be the complement $U^c$ of the (open or closed) unit ball $U$ of $B$, one checks immediately that

$$I_f(U^c) = \frac{1}{2}\left(\sup_{|h|\leq 1} \|f^{(d)}(h)\|\right)^{-2/d}$$

so that

$$\lim_{t\to\infty} \frac{1}{t^{2/d}} \log \mu(x; \|f(x)\| > t) = -\frac{1}{2}\left(\sup_{|h|\leq 1} \|f^{(d)}(h)\|\right)^{-2/d}.$$

In particular, when $d = 1$,

$$\sup_{|h|\leq 1} \|f^{(1)}(h)\| = \sup_{\|\xi\|\leq 1} \left(\int \langle \xi, f(x)\rangle^2 d\mu(x)\right)^{1/2}.$$

In the setting of the classical Wiener space $E = C_0([0,1])$ equipped with the Wiener measure $\mu$, *and when* $B = E$, K. Itô [It] (see also [Ne1] and the recent approach [St2]) identified the elements $f$ of $\mathcal{H}^{(d)}(\mu; E)$ with the multiple stochastic integrals

$$f = \left(\int_0^t \int_0^{t_1} \cdots \int_0^{t_{d-1}} g(t_1, \dots, t_d)\, dB(t_1) \cdots dB(t_d)\right)_{t\in[0,1]}$$

where $g$ deterministic is such that

$$\int_0^t \int_0^{t_1} \cdots \int_0^{t_{d-1}} g(t_1, \dots, t_d)^2 dt_1 \cdots dt_d < \infty.$$

If $h$ belongs to the reproducing kernel Hilbert of the Wiener measure, then

$$f^{(d)}(h) = \left(\int_0^t \int_0^{t_1} \cdots \int_0^{t_{d-1}} g(t_1, \dots, t_d)\, h'(t_1) \cdots h'(t_d)\, dt_1 \cdots dt_d\right)_{t\in[0,1]}.$$

**Proof of the theorem.** Let us start with the simpler property (ii). Recall $f_n$ from (1). We can write (explicitely on the Hermite polynomials), for all $x$ in $E$, $h$ in $H$ and $t$ real number,

$$f_n(x + th) = \sum_{k=0}^{d} t^k f_n^{(k)}(x, h).$$

If $P(t) = a_0 + a_1 t + \cdots + a_d t^d$ is a polynomial of degree $d$ in $t \in \mathbb{R}$ with vector coefficients $a_0, a_1, \ldots, a_d$, there exist real constants $c(i,k,d)$, $0 \le i,k \le d$, independent of $P$, such that, for all $k = 0, \ldots, d$,

$$a_k = c(0,k,d)P(0) + \sum_{i=1}^{d} c(i,k,d)P(2^{i-1}).$$

Hence, for all $h$,

$$f_n^{(k)}(\cdot, h) = c(0,k,d)f_n(\cdot) + \sum_{i=1}^{d} c(i,k,d)f_n(\cdot + 2^{i-1}h)$$

from which we deduce together with (2) that, for all $k = 0, \ldots, d$,

$$\int \|f_n^{(k)}(x,h)\| \, d\mu(x) \le C(k,d;h) \left( \int \|f_n(x)\|^2 d\mu(x) \right)^{1/2}$$

for some constants $C(k,d;h)$ thus only depending on $k$, $d$ and $h \in H$. In the limit, we conclude that there exist, for every $h$ in $H$ and $k = 0, \ldots, d$, elements $f^{(k)}(\cdot, h)$ of $L^1(\mu; B)$ such that

$$f(\cdot + th) = \sum_{k=0}^{d} t^k f^{(k)}(\cdot, h)$$

for all $t \in \mathbb{R}$, with

$$\int \|f^{(k)}(x,h)\| \, d\mu(x) \le C(d,k;h) \left( \int \|f(x)\|^2 d\mu(x) \right)^{1/2}$$

and $f^{(0)}(\cdot, h) = f(\cdot)$, $f^{(d)}(\cdot, h) = f^{(d)}(h)$ (since $\int f(x+th) \, d\mu(x) = t^d f^{(d)}(h)$). As a main consequence, we get that, for all $h$ in $H$,

$$\lim_{t \to \infty} \frac{1}{t^d} \int \|f(x+th) - t^d f^{(d)}(h)\| \, d\mu(x) = 0. \tag{5}$$

This limit can be made uniform in $h \in \mathcal{O}$ but we will not use this observation in this form later (that is in the proof of (i); we use instead a stronger property, (7) below).

To establish (ii), let $s = f^{(d)}(h)$, $h \in H$, belong to $G$ (if no such $s$ exists, then $I_f(G) = +\infty$ and (ii) then holds trivially). Since $G$ is open, there is $\delta > 0$ such that the ball $B(s,\delta)$ of center $s$ and radius $\delta$ is contained in $G$. Therefore, if $A = A(t) = \{x; f(x) \in t^d B(s,\delta)\}$, by Cameron-Martin,

$$\mu(x; f(x) \in t^d G) \ge \mu(A) = \int_{A-th} \exp(t\widetilde{h}(x) - t^2|h|^2/2) d\mu(x).$$

Further, by Jensen's inequality,

$$\mu(A) \geq \exp(-t^2|h|^2/2)\mu(A-th)\exp\left(\frac{t}{\mu(A-th)}\int_{A-th}\widetilde{h}(x)d\mu(x)\right).$$

By (5),

$$\mu(A-th) = \mu(x; \|f(x+th) - t^d f^{(d)}(h)\| \leq \delta t^d) \geq \frac{1}{2}$$

for all $t \geq t_0$ large enough. By centering and Cauchy-Schwarz,

$$\int_{A-th}\widetilde{h}(x)d\mu(x) = -\int_{(A-th)^c}\widetilde{h}(x)d\mu(x) \geq -|h|\mu((A-th)^c)^{1/2} \geq -\frac{|h|}{\sqrt{2}}.$$

Thus, for all $t \geq t_0$,

$$\frac{t}{\mu(A-th)}\int_{A-th}\widetilde{h}(x)d\mu(x) \geq -\sqrt{2}\,t|h|,$$

and hence, summarizing,

$$\mu(x; f(x) \in t^d G) \geq \frac{1}{2}\exp\left(-\frac{1}{2}t^2|h|^2 - \sqrt{2}\,t|h|\right).$$

It follows that

$$\liminf_{t\to\infty}\frac{1}{t^2}\log\mu(x; f(x) \in t^d G) \geq -\frac{|h|^2}{2} = -I_f(s)$$

and since $s$ is arbitrary in $G$, property (ii′) is satisfied. As a consequence of what we will develop now, (ii) is satisfied as well.

We now turn to (i) and in the first part of this investigation, we closely follow C. Borell [Bo3], [Bo5]. We start by showing that every element $f$ of $\mathcal{H}^{(d)}(\mu; B)$ is limit (at least if the dimension of the support of $\mu$ is infinite), $\mu$-almost everywhere and in $L^2(\mu; B)$, of a sequence of $d$-homogeneous polynomials. In particular, $f$ is $\mu$-almost everywhere $d$-homogeneous justifying therefore the equivalences between (i) and (ii) and respectively (i′) and (ii′). Assume thus in the following that $\mu$ is infinite dimensional. We can actually always reduce to this case by appropriately tensorizing $\mu$, for example with the canonical Gaussian measure on $\mathbb{R}^{\mathbb{N}}$. Recall that $f$ is limit almost surely and in $L^2(\mu; B)$ of the $f_n$'s of (1). The finite sums $f_n$ can be decomposed in their homogeneous components as

$$f_n = f_n^{(d)} + f_n^{(d-2)} + \cdots,$$

where, for any $x$ in $E$,

$$f_n^{(k)}(x) = \sum_{i_1,\ldots,i_k=0}^{\infty} b_{i_1,\ldots,i_k}\, e_{i_1}(x)e_{i_2}(x)\cdots e_{i_k}(x) \tag{6}$$

with only finitely many $b_{i_1,\dots,i_k}$ in $B$ non zero. The main observation is that the constant 1 is limit of homogeneous polynomials of degree 2; indeed, simply take

$$p_n(x) = \frac{1}{n+1}\sum_{k=0}^{n} e_k(x)^2.$$

Since $p_n$ and $f_n^{(k)}$ belong to $L^p(\mu;\mathbb{R})$ and $L^p(\mu;B)$ respectively for all $p$, and since $p_n - 1$ tends there to 0, it is easily seen that there exists a subsequence $m_n$ of the integers such that $(p_{m_n}-1)(f_n^{(d-2)}+f_n^{(d-4)}+\cdots)$ converges to 0 in $L^2(\mu;B)$. This means that $f$ is limit in $L^2(\mu;B)$ of $f_n^{(d)}+p_{m_n}(f_n^{(d-2)}+f_n^{(d-4)}+\cdots)$, that is limit of a sequence of polynomials $f'_n$ whose decomposition in homogeneous polynomials

$$f'_n = f'^{(d)}_n + f'^{(d-2)}_n + \cdots$$

is such that $f'^{(1)}_n$, or $f'^{(0)}_n$ and $f'^{(2)}_n$, according as $d$ is odd or even, can be taken to be 0. Repeating this procedure, $f$ is indeed seen to be limit in $L^2(\mu;B)$ of a sequence $(g_n)$ of $d$-homogeneous polynomials (i.e. polynomials of the type of (6)).

The important property in order to establish (i′) is the following. It improves upon (5) and claims that, in the preceding notations, i.e. if $f$ is limit of the sequence $(g_n)$ of $d$-homogeneous polynomials,

$$\lim_{t\to\infty}\frac{1}{t^d}\sup_n\int\sup_{|h|\le 1}\|g_n(x+th)-t^d g_n(h)\|^2 d\mu(x) = 0. \tag{7}$$

To establish this property, given

$$g_n(x) = \sum_{i_1,\dots,i_d=0}^{\infty} b^n_{i_1,\dots,i_d}\, e_{i_1}(x)e_{i_2}(x)\cdots e_{i_d}(x)$$

(with only finitely many $b^n_{i_1,\dots,i_d}$ non zero), let us consider the (unique) multilinear symmetric polynomial $\widehat{g}_n$ on $E^d$ such that $\widehat{g}_n(x,\dots,x) = g_n(x)$; $\widehat{g}_n$ is given by

$$\widehat{g}_n(x_1,\dots,x_d) = \sum_{i_1,\dots,i_d=0}^{\infty} \widehat{b}^n_{i_1,\dots,i_d}\, e_{i_1}(x_1)\cdots e_{i_d}(x_d),\quad x_1,\dots,x_d\in E,$$

where

$$\widehat{b}^n_{i_1,\dots,i_d} = \frac{1}{d!}\sum_{\sigma} b^n_{\sigma(i_1),\dots,\sigma(i_d)},$$

the sum being running over all permutations $\sigma$ of $\{1,\dots,d\}$. As is well-known [M-O], [B-S], we have the following polarization formula : letting $\varepsilon_1,\dots,\varepsilon_d$ be independent symmetric Bernoulli random variables and denoting by $\mathbb{E}$ expectation with respect to them,

$$\widehat{g}_n(x_1,\dots,x_d) = \frac{1}{d!}\,\mathbb{E}(g_n(\varepsilon_1 x_1+\cdots+\varepsilon_d x_d)\,\varepsilon_1\cdots\varepsilon_d). \tag{8}$$

We adopt the notation $x^{d-k}y^k$ for $(x,\dots,x,y,\dots,y)$ in $E^d$ where $x$ is repeated $(d-k)$-times and $y$ $k$-times. Then, for any $x,y$ in $E$, we have

$$g_n(x+y)=\sum_{k=0}^{d}\binom{d}{k}\widehat{g}_n(x^{d-k}y^k). \tag{9}$$

To establish (7), we see from (9) that it suffices to show that for all $k=1,\dots,d-1$,

$$\sup_n\int\sup_{|h|\le 1}\|\widehat{g}_n(x^{d-k}h^k)\|^2d\mu(x)<\infty. \tag{10}$$

Let $k$ be fixed. By orthogonality,

$$\begin{aligned}
&\sup_{|h|\le 1}\|\widehat{g}_n(x^{d-k}h^k)\|^2\\
&\le\sup_{\|\xi\|\le 1}\ \sup_{|h_1|\le 1,\dots,|h_k|\le 1}\langle\xi,\widehat{g}_n(x,\dots,x,h_1,\dots,h_k)\rangle^2\\
&\le\sup_{\|\xi\|\le 1}\sum_{i_{d-k+1},\dots,i_d=0}^{\infty}\left|\sum_{i_1,\dots,i_{d-k}=0}^{\infty}\langle\xi,\widehat{b}^n_{i_1,\dots,i_d}\rangle e_{i_1}(x)\cdots e_{i_{d-k}}(x)\right|^2\\
&=\sup_{\|\xi\|\le 1}\int\cdots\int\langle\xi,\widehat{g}_n(x,\dots,x,y_1,\dots,y_k)\rangle^2d\mu(y_1)\cdots d\mu(y_k)\\
&\le\int\cdots\int\|\widehat{g}_n(x,\dots,x,y_1,\dots,y_k)\|^2d\mu(y_1)\cdots d\mu(y_k).
\end{aligned}$$

By the polarization formula (8),

$$\widehat{g}_n(x,\dots,x,y_1,\dots,y_k)=\frac{1}{d!}\,\mathbb{E}(g_n((\varepsilon_{k+1}+\cdots+\varepsilon_d)x+\varepsilon_1y_1+\cdots+\varepsilon_ky_k)\,\varepsilon_1\cdots\varepsilon_k)).$$

Therefore, we obtain from the rotational invariance of Gaussian distributions and homogeneity that

$$\begin{aligned}
&(d!)^2\int\sup_{|h|\le 1}\|\widehat{g}_n(x^{d-k}h^k)\|^2d\mu(x)\\
\le&\mathbb{E}\int\int\cdot\int\|g_n((\varepsilon_{k+1}+\cdot+\varepsilon_d)x+\varepsilon_1y_1+\cdot+\varepsilon_ky_k)\|^2d\mu(x)d\mu(y_1)\cdot d\mu(y_k)\\
=&\mathbb{E}\int\|g_n(((\varepsilon_{k+1}+\cdots+\varepsilon_d)^2+k)^{1/2}x)\|^2d\mu(x)\\
=&\mathbb{E}(((\varepsilon_{k+1}+\cdots+\varepsilon_d)^2+k)^{d/2})\int\|g_n(x)\|^2d\mu(x).
\end{aligned}$$

Hence (10) and therefore (7) are established.

We can now conclude the proof of (i′) and thus of the theorem. It is intuitively clear that

$$\lim_{n\to\infty} \sup_{|h|\le 1} \|g_n(h) - f^{(d)}(h)\| = 0. \tag{11}$$

This property is an easy consequence of (7). Indeed, for all $n$ and $t > 0$,

$$\begin{aligned}
&\sup_{|h|\le 1} \|g_n(h) - f^{(d)}(h)\| \\
&\le \sup_m \sup_{|h|\le 1} \|g_m(h) - \frac{1}{t^d}\int g_m(x+th)\,d\mu(x)\| \\
&\qquad\qquad + \sup_{|h|\le 1} \frac{1}{t^d}\|\int g_n(x+th) - f(x+th)\|\,d\mu(x) \\
&\le \sup_m \int \sup_{|h|\le 1} \|g_m(h) - \frac{1}{t^d} g_m(x+th)\|\,d\mu(x) \\
&\qquad\qquad + \sup_{|h|\le 1} \frac{1}{t^d}\int \|g_n(x+th) - f(x+th)\|\,d\mu(x)
\end{aligned}$$

and, using (2) and (7), the limit in $n$ and then in $t$ yields (11). Let now $F$ be closed in $B$ and take $0 < r < I_f(F)$. The definition of $I_f(F)$ indicates that $(2r)^{d/2} f^{(d)}(\mathcal{O}) \cap F = \emptyset$ where we recall that $\mathcal{O}$ is the unit ball of $H$, compact in $E$. Therefore, since $f^{(d)}(\mathcal{O})$ is clearly seen to be compact in $B$ by (11), and since $F$ is closed, one can find $\delta > 0$ such that

$$((2r)^{d/2} f^{(d)}(\mathcal{O}) + 2\delta U) \cap (F + \delta U) = \emptyset \tag{12}$$

where $U$ is the (closed) unit ball of $B$. By (11), there exists $n_0 = n_0(\delta)$ large enough such that for all $n \ge n_0$,

$$(2r)^{d/2} g_n(\mathcal{O}) \subset (2r)^{d/2} f^{(d)}(\mathcal{O}) + \delta U. \tag{13}$$

Let thus $n \ge n_0$. For any $t > 0$, we can write

$$\begin{aligned}
&\mu(x; f(x) \in t^d F) \\
&\le \mu(x; \|f(x) - g_n(x)\| > \delta t^d) + \mu(x; g_n(x) \in t^d(F + \delta U)) \\
&\le \mu(x; \|f(x) - g_n(x)\| > \delta t^d) + \mu^*(x; x \notin A + t\sqrt{2r}\mathcal{O})
\end{aligned} \tag{14}$$

where

$$A = A(t,n) = \{a; \sup_{|h|\le 1} t^{-d}\|g_n(a + t\sqrt{2r}h) - t^d (2r)^{d/2} g_n(h)\| \le \delta\}.$$

To justify the second inequality in (14), observe that if $x = a + t\sqrt{2r}h$ with $a \in A$ and $|h| \le 1$, then

$$\frac{1}{t^d}\, g_n(x) = \frac{1}{t^d}\,[g_n(a + t\sqrt{2r}h) - t^d (2r)^{d/2} g_n(h)] + (2r)^{d/2} g_n(h),$$

so that the claim follows by (12), (13) and the definition of $A$. By (7), let now $t_0 = t_0(\delta)$ be large enough so that, for all $t \geq t_0$,

$$\sup_n \frac{1}{t^d} \int \sup_{|h|\leq 1} \|g_n(x + t\sqrt{2r}h) - t^d(2r)^{d/2} g_n(h)\|^2 d\mu(x) \leq \frac{\delta^2}{2}\,.$$

That is, for all $n$ and all $t \geq t_0$, $\mu(A(t,n)) \geq 1/2$. By (4), it follows that

$$\mu^*(x; x \notin A + t\sqrt{2r}\mathcal{O}) \leq \exp(-rt^2). \tag{15}$$

Fix now $t \geq t_0 = t_0(\delta)$. Choose $n = n(t) \geq n_0 = n_0(\delta)$ large enough in order that

$$\mu(x; \|f(x) - g_n(x)\| > \delta t^d) \leq \exp(-rt^2).$$

Together with (14) and (15), it follows that for all $t \geq t_0$,

$$\mu(x; f(x) \in t^d F) \leq 2\exp(-rt^2).$$

$r < I_f(F)$ being arbitrary, the proof of (i′) and therefore of the theorem is complete.

Note that it would of course have been possible to work directly on $f$ rather than on the approximating sequence $(g_n)$ in the preceding proof; this approach however avoids several measurability questions and makes everything more explicit.

## References


[Az] R. Azencott. Grandes déviations et applications. Ecole d'Eté de Probabilités de St-Flour 1978. Lecture Notes in Math. 774, 1–176 (1978). Springer-Verlag.

[B-S] J. Bochnak, J. Siciak. Polynomials and multilinear mappings in topological vector spaces. Studia Math. 39, 59–76 (1971).

[Bo1] C. Borell. The Brunn-Minskowski inequality in Gauss space. Invent. Math. 30, 207–216 (1975).

[Bo2] C. Borell. Gaussian Radon measures on locally convex spaces. Math. Scand. 38, 265–284 (1976).

[Bo3] C. Borell. Tail probabilities in Gauss space. Vector Space Measures and Applications, Dublin 1977. Lecture Notes in Math. 644, 71–82 (1978). Springer-Verlag.

[Bo4] C. Borell. On polynomials chaos and integrability. Prob. Math. Statist. 3, 191–203 (1984).

[Bo5] C. Borell. On the Taylor series of a Wiener polynomial. Seminar Notes on multiple stochastic integration, polynomial chaos and their integration. Case Western Reserve University, Cleveland (1984).

[C-M] R. H. Cameron, W. T. Martin. Transformations of Wiener integrals under translations. Ann. Math. 45, 386–396 (1944).

[Ch] S. Chevet. Gaussian measures and large deviations. Probability in Banach spaces IV, Oberwolfach 1982. Lecture Notes in Math. 990, 30–46 (1983).

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

[Eh] A. Ehrhard. Symétrisation dans l'espace de Gauss. Math. Scand. 53, 281–301 (1983).

[Fe] X. Fernique. Gaussian random vectors and their reproducing kernel Hilbert spaces. Technical report, University of Ottawa (1985).

[It] K. Itô. Multiple Wiener integrals. J. Math. Soc. Japan 3, 157–164 (1951).

[Ku] H.-H. Kuo. Gaussian measures in Banach spaces. Lecture Notes in Math. 436 (1975). Springer-Verlag.

[Le] M. Ledoux. Inégalités isopérimétriques et calcul stochastique. Séminaire de Probabilités XXII. Lecture Notes in Math. 1321, 249–259 (1988). Springer-Verlag.

[MK] H. P. McKean. Geometry of differential space. Ann. Probability 1, 197–206 (1973).

[M-O] S. Mazur, W. Orlicz. Grundlegende Eigenschaften der polynomischen Operationen. Studia Math. 5, 50–68, 179–189 (1935).

[Ne1] J. Neveu. Processus aléatoires gaussiens. Presses de l'Université de Montréal (1968).

[Ne2] J. Neveu. Martingales à temps discret. Masson (1972).

[Pi] G. Pisier. Probabilistic methods in the geometry of Banach spaces. Probability and Analysis, Varenna (Italy) 1985. Lecture Notes in Math. 1206, 167–241 (1986). Springer-Verlag.

[St1] D. Stroock. An introduction to the theory of large deviations. Springer-Verlag (1984).

[St2] D. Stroock. Homogeneous chaos revisited. Séminaire de Probabilités XXI. Lecture Notes in Math. 1247, 1–7 (1987). Springer-Verlag.

[S-T] V. N. Sudakov, B. S. Tsirel'son. Extremal properties of half-spaces for spherically invariant measures. J. Soviet. Math. 9, 9–18 (1978); translated from Zap. Nauch. Sem. L.O.M.I. 41, 14–24 (1974).

[Wi] N. Wiener. The homogeneous chaos. Amer. Math. J. 60, 897–936 (1930).

*Institut de Recherche Mathématique Avancée,*
*Laboratoire associé au C.N.R.S.,*
*Université Louis-Pasteur,*
*7, rue René-Descartes, F-67084 Strasbourg.*