Probab. Th. Rel. Fields 72, 305-318 (1986)

Probability Theory and Related Fields

© Springer-Verlag 1986

Laplace Approximations for Sums of Independent Random Vectors

E. Bolthausen

Technische Universität Berlin, Fachbereich Mathematik, Straße des 17. Juni 135, 1000 Berlin 12

Summary. Let X_i, $i \in \mathbb{N}$, be i.i.d. B-valued random variables where B is a real separable Banach space, and Φ a mapping $B \to \mathbb{R}$. Under some conditions an asymptotic evaluation of $Z_n = E\left(\exp\left(n\Phi\left(\sum_{i=1}^{n} X_i/n\right)\right)\right)$ is possible, up to a factor $(1+o(1))$. This also leads to a limit theorem for the appropriately normalized sums $\sum_{i=1}^{n} X_i$ under the law transformed by the density $\exp\left(n\Phi\left(\sum_{i=1}^{n} X_i/n\right)\right) \Big/ Z_n$.

§ 1. Introduction

Let B a real separable Banach space with norm $|\ |$ and X_n, $n \in \mathbb{N}$, be a sequence of i.i.d. B-valued random variables with law μ which satisfies:

(1.1) $$\int \exp(t|x|)\,\mu(dx) < \infty \quad \text{for all } t \in \mathbb{R}$$

(1.2) $$\int x\,\mu(dx) = 0.$$

Let Φ be a real-valued Borel-measurable continuous function on B. The aim of this paper is to give an asymptotic evaluation of $E(\exp(n\Phi(S_n/n)))$ as $n \to \infty$ where $S_n = \sum_{i=1}^{n} X_i$.

If there exist real constants $C, D > 0$ with

(1.3) $$\Phi(x) \leqq C + D|x|$$

then it has been proved by Donsker and Varadhan [5] that

(1.4) $$\lim_{n\to\infty} \frac{1}{n} \log E\left(\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right) = \sup_x (\Phi(x) - h(x))$$

where h is the entropy function of μ:

(1.5) $$h(x)=\sup_{\varphi\in B^*}(\varphi(x)-\log M(\varphi)).$$

B^* being the topological dual of B and $M(\varphi)=\int e^{\varphi}\,d\mu$ (see [5] Theorem 5.3 and Sect. 3 of [14] and §2 where some basic facts on the entropy are collected).

Furthermore, under these conditions there is at least one $x^*\in B$ with $\Phi(x^*)-h(x^*)=\sup_{x\in B}(\Phi(x)-h(x))$. This will be proved in §2. We need the stronger condition

(1.6) There is a unique $x^*\in B$ with $\Phi(x^*)-h(x^*)=\sup_{x}(\Phi(x)-h(x))$.

We use x^* exclusively for this point.

We also need that Φ is smooth enough, namely

(1.7) Φ has three continuous Fréchet derivatives on B.

If $b\in B$, we write $D^k\Phi(b)$ for the k-th Fréchet-derivative of Φ at b (when it exists) which is a continuous k-linear form on B.

We write $D^k\Phi(b)[x_1,\dots,x_k]$ for this form at $x_1,\dots,x_k\in B$ and $D^k\Phi(b)[x^k]$ instead of $D^k\Phi(b)[x,\dots,x]$.

Let $d\nu=\exp(D\Phi(x^*))\,d\mu/M(D\Phi(x^*))$. ν has moments of all orders and in §2 we shall prove that

(1.8) $$x^*=\int x\,\nu(dx) \quad \text{holds.}$$

Let ν_0 be ν centered at 0, i.e. $\nu_0=\nu\theta_{x^*}^{-1}$ where $\theta_a\colon B\to B$ is defined by $\theta_a(x)=x-a$.

We need an assumption stating that the maximum in x^* is non-degenerated in some sense. To formulate this we define the mapping $\hat{}\colon B^*\to B$ by $\hat{\varphi}=\int x\varphi(x)\nu_0(dx)$. If $\psi\in B^*$ then $\psi(\hat{\varphi})$ is the covariance of φ and ψ under ν_0: $\Gamma(\varphi,\psi)=\int\varphi(x)\psi(x)\nu_0(dx)$. Then we have

Lemma 1. *For all* $\varphi\in B^*$ $\Gamma(\varphi,\varphi)\geqq D^2\Phi(x^*)[\hat{\varphi}^2]$.

This will be proved in §2. We assume

(1.9) For all $\varphi\in B^*$, with $\hat{\varphi}\neq 0$, $\Gamma(\varphi,\varphi)>D^2\Phi(x^*)[\hat{\varphi}^2]$ holds.

Remark 1. From Lemma 1 it clearly follows that in any case

$$\{\varphi\colon \Gamma(\varphi,\varphi)=D^2\Phi(x^*)[\hat{\varphi}^2]\}$$

is a linear subspace of B^* and it will be shown in §2 that this subspace is finite dimensional (see the remarks following the proof of Lemma 1). So (1.9) just states that this subspace of degenerate directions has dimension 0.

Our main result is the following

Theorem 1. *We assume that* μ *and* Φ *satisfy*

a) (1.1), (1.2), (1.3), (1.6), (1.7), (1.9) *and*

b) ν *satisfies a central limit theorem, i.e.* ν_n *defined by* $\nu_n(A)=\nu_0^{*n}(\sqrt{n}A)$ *converges weakly to a Gaussian measure* γ.

(Here ^{*n} *denotes n-fold convolution.)*

Then

$$\lim_{n\to\infty} \exp(-n(\Phi(x^*)-h(x^*)))\, E\left(\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right)$$
$$=\int \exp(\tfrac{1}{2}D^2\Phi(x^*)[y^2])\,\gamma(dy).$$

Remark 2. If μ has bounded support then (1.3) can be replaced by the weaker condition that Φ is bounded above on bounded sets.

Results of this type have been obtained in the case $B=\mathbb{R}$ by Martin-Löf [11], for Banach spaces with Gaussian μ by Schilder [13], Pincus [12] and Ellis and Rosen [6] and for a Hilbert space in some special situations by Kusuoka and Tamura [10].

Remark 3. The condition b) in Theorem 1 is automatically satisfied in so-called type 2 spaces (as ν has a second moment), especially in L_p spaces for $2\leq p<\infty$ (see Hoffmann-Jørgensen [7]). In other spaces there are useful sufficient conditions for the central limit theorem. (See e.g. [8].) It is desirable to have conditions which only depend on μ and not on Φ. Some of conditions for the validity of the central limit theorem nearly carry over from μ to every possible ν. As an example, we look of the condition of Jain and Marcus for $C(T)$-valued random variables, where T is compact metric space (see [8], Theorem 3.5). If μ on $B=C(T)$ satisfies

$$\mu(\{f\colon |f(s)-f(t)|>V(f)\,\rho(s,t) \text{ for some } s,t\})=0$$

where $V\colon B\to[0,\infty)$ satisfies $\int V^{2+\delta}d\mu<\infty$ for some $\delta>0$ and ρ is a continuous metric on T which satisfies $\int_0^1 H_\rho(\varepsilon)^{\frac{1}{2}}\,d\varepsilon<\infty$ ($H(\ )$ the ε-entropy of T with respect to ρ) then an application of the Hölder-inequality together with Theorem 3.5 of [8] shows that any possible ν of the form $d\nu=e^{\varphi}d\mu/M(\varphi)$, $\varphi\in B^*$, satisfies the central limit theorem (if (1.1) is true).

The Gaussian measure γ is generated as an abstract Wiener measure (as all Gauss measures). We sketch the construction. As Gaussian measures have exponential and therefore second moments, there is a natural mapping $j\colon B^*\to L_2(B,\gamma)$. We denote by H the closure of $j(B^*)$ in $L_2(B,\gamma)$ which is then a Hilbert space. If $\varphi\in B^*$ then the $\hat{\varphi}$ defined above may also be written as $\int x\varphi(x)\,\gamma(dx)$ as it only depends on the covariance form. It is easy to see that $\hat{\varphi}$ depends only on $j(\varphi)$ and the mapping $j(\varphi)\mapsto\hat{\varphi}$ is continuous in the L_2-norm.

So we obtain a continuous mapping $i\colon H\to B$ which can be shown to be one to one. j is one to one if and only if the support of γ is B. i and j are compact linear mappings. (H,B,i) then generates γ as an abstract Wiener measure in the sense of Gross (see [9]).

For proofs of these facts (in a more general setting) see [3].

We write $\langle\ ,\ \rangle_H$ for the inner product in H. We also identify H with the subset $i(H)\subset B$ and shall therefore not distinguish between $x\in H$ and $i(x)\in B$. If $\varphi\in B^*$ then $\hat{\varphi}$ clearly is in H.

Lemma 2. *$Z=\exp(\frac{1}{2}D^2\Phi(x^*)[y^2])\,\gamma(dy)$ is $<\infty$ and γ' defined by*

$$(d\gamma'/d\gamma)(y)=Z^{-1}\exp(\tfrac{1}{2}D^2\Phi(x^*)[y^2])$$

is a centered Gauss measure on B.

The proof will be given in §2.

We can then prove the following central limit theorem.

Theorem 2. *Assume the same conditions as in Theorem 1. Let*

$$dP_n/dP=\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\Big/E\left(\exp\left(n\Phi\left(\frac{S_n}{n}\right)\right)\right).$$

Then the P_n-law of $\sqrt{n}\left(\frac{S_n}{n}-x^*\right)$ converge weakly to γ'.

Similar results have been obtained by Ellis-Rosen for Gaussian laws μ even in degenerate cases, i.e., if (1.9) does not hold and where non Gaussian (finite dimensional) limit laws appear.

The condition b) in Theorem 1 can be reformulated in the following way:

(1.10) For any closed $F\subset B$ $\limsup_{n\to\infty} \nu_n(F)\leqq\gamma(F)$.

The proofs of the Theorems 1 and 2 essentially depend on a Bernstein-type inequality stating roughly that if t is small compared with $\sqrt{n}$, then $\nu_n(tF)$ behaves as is reflected in (1.10). The H-norm governs the large deviations behavior of γ in the following sense ([2], Theorem II 1.6):

$$\limsup_{t\to\infty}\frac{1}{t^2}\,\gamma(tF)\leqq-\inf\{\tfrac{1}{2}|x|_H^2\colon x\in F\} \tag{1.11}$$

(we put $|x|_H=\infty$ if $x\in B\setminus H$). Let $\Gamma(F)=\inf\{\frac{1}{2}|x|_H^2\colon x\in F\}$. Therefore the following result looks quite plausible.

Theorem 3. *If $F\subset B$ is closed then*

$$\limsup_{c\to\infty}\ \sup_{n,t}\left\{\frac{1}{t}\log\nu_n(tF)\colon c\leqq t\leqq\sqrt{n}/c\right\}\leqq-\Gamma(F).$$

In case where B is a Hilbert space and $F=\{x\colon|x|\geqq 1\}$ the proposition follows from results of Yurinskii ([15], Sect. 4). He also obtained results for certain Banach spaces but they are not sharp enough to give the Theorem 3 immediately. But Yurinskii's Theorem 2.1 is basic in the Proof given in §3.

The proofs of the Theorems 1 and 2 are given in §4.

§2. Properties of the Entropy and Proofs of the Lemmas

Let μ be any probability measure on B and define h by (1.5). h obviously is $\geqq 0$ and as a supremum of continuous affine function it is lower semicontinuous with values in $[0, \infty]$ (even in the weak*-topology) and convex. The strong condition (1.1) guarantees that h may be obtained as a contraction of the so-called Kullback-Leibler information (or I-divergence).

If λ is a probability measure on B then the Kullback-Leibler information of λ with respect to μ is defined by

$$k(\lambda|\mu)=\begin{cases}\int\left(\log\frac{d\lambda}{d\mu}\right)d\lambda & \text{if } \lambda\ll\mu \text{ and } \log\frac{d\lambda}{d\mu}\in L_1(\lambda)\\ \infty & \text{else.}\end{cases}$$

It is well known (and easy to see) that as a function of λ (with μ fixed) k is convex, strongly convex on $\{\lambda\colon k(\lambda|\mu)<\infty\}$ and lower semicontinuous in the weak topology (see e.g. [4]).

Lemma 3. *Let* (1.1) *be satisfied* (*not necessarily* (1.2)) *then*

a) $h(x)=\inf\{k(\lambda|\mu)\colon \int|y|\,\lambda(dy)<\infty,\ \int y\,\lambda(dy)=x\}$.

b) If $h(x)<\infty$ then there is a unique probability measure λ_x with $k(\lambda_x|\mu)=h(x)$ *and* $\int y\,\lambda_x(dy)=x$.

c) $h(x)=0$ if and only if $x=\int y\,\mu(dy)$.

d) For all $r\in[0,\infty)$ $\{x\colon h(x)\leqq r\}$ is compact in B.

Proof. The facts follow from the considerations in §5 of [5]:

Donsker and Varadhan define $h(x)$ as $\inf\{k(\lambda|\mu)\colon \int y\,\lambda(dy)=x\}$ $(k(\lambda|\mu)=I_\mu(\lambda)$ in their notation, see (5.3) in [5]) a), c), d) of our lemma then follow from (iv), (ii) and (iii) of Theorem 5.2 in [5]. To prove b) we look at a sequence λ_n of probability measures with $k(\lambda_n|\mu)\downarrow h(x)$, $\int y\,\lambda_n(dy)=x$. By Lemma 5.1 of [5] (λ_n) is tight and a straightforward argument shows that any limit point satisfies $\int y\,\lambda_x(dy)=x$ and $h(x)=k(\lambda_x|\mu)$ (see the proof of Lemma 5.1 in [5]). Unicity of λ_x follows from the strong convexity of k. □

Lemma 4. *If* (1.1) *is satisfied, then* $h(x)/|x|\to\infty$ *uniformly as* $|x|\to\infty$.

Proof. If $\lambda>0$, we choose $c>0$ with

$$\int_{|x|\geqq c}\exp((1+\lambda)|x|)\,\mu(dx)\leqq 1.$$

If $x\in B$, let $\varphi\in B^*$ satisfy $|\varphi|_*=1$ and $\varphi(x)=|x|$. Then

$$\begin{aligned}h(x)&\geqq(1+\lambda)\,\varphi(x)-\log M((1+\lambda)\,\varphi)\\&\geqq(1+\lambda)\,|x|-\log\int\exp((1+\lambda)\,|x|)\,\mu(dx)\\&\geqq(1+\lambda)\,|x|-\log\{\exp((1+\lambda)\,c+\int_{|x|\geqq c}e^{(1+\lambda)|x|}\,\mu(dx)\}\\&\geqq(1+\lambda)\,|x|-(1+\lambda)\,c-\log 2.\end{aligned}$$

Therefore, if $|x|\geqq(1+\lambda)\,c+\log 2$ then $h(x)/|x|>\lambda$. □

Using d) of Lemma 3 and Lemma 4, one immediately sees that under condition (1.3) $\sup_x(\Phi(x)-h(x))$ is attained, as is claimed in §1.

In the case where μ has bounded support h is ∞ outside a bounded set and if Φ is bounded on bounded sets, the same conclusion holds true.

Proof of (1.8). By the convexity of h one obtains $h(x^*)+D\Phi(x^*)[x]\leqq h(x^*+x)$ for all $x\in B$ and therefore $\hat{h}$ the entropy function of ν i.e. $\hat{h}=h-D\Phi(x^*)+\log M(D\Phi(x^*))$ is minimal at x^* and so equals zero there. From c) of Lemma 3 one has $\int y\nu(dy)=x^*$. $\square$

Proof of Lemma 1. Let $\varphi\in B^*$ and the probability ν_t on B $(t>0)$ be defined by $d\nu_t=\exp(D\Phi(x^*)+t\varphi)\,d\mu/M(D\Phi(x^*)+t\varphi)$. An easy calculation gives

$$M(D\Phi(x^*)+t\varphi)=e^{t\varphi(x^*)}M(D\Phi(x^*))\left(1+\frac{t^2}{2}\Gamma(\varphi,\varphi)+O(t^3)\right)\quad\text{as } t\to 0. \tag{2.1}$$

Let $a_t=\int x\nu_t(dx)$. Applying (2.1), one obtains

$$a_t=x^*+t\int x(\varphi(x)-\varphi(x^*))\,\nu(dx)+R(t)$$

where $|R(t)|=O(t^2)$ for t near O. $\int x(\varphi(x)-\varphi(x^*))\,\nu(dx)$ equals $\hat{\varphi}$.

By applying the Taylor formula, one obtains for t near O

$$\Phi(a_t)-\Phi(x^*)=D\Phi(x^*)[a_t-x^*]+\frac{t^2}{2}D^2\Phi(x^*)[\hat{\varphi}^2]+o(t^2). \tag{2.2}$$

On the other hand, applying (2.1), one obtains

$$\begin{aligned}h(a_t)\leqq k(\nu_t|\mu)&=D\Phi(x^*)[a_t-x^*]+D\Phi(x^*)[x^*]+t^2\varphi(\hat{\varphi})\\&\quad-\log M(D\Phi(x^*))-\frac{t^2}{2}\Gamma(\varphi,\varphi)+o(t^2)\\&=D\Phi(x^*)[a_t-x^*]+h(x^*)+\frac{t^2}{2}\Gamma(\varphi,\varphi)+o(t^2).\end{aligned}$$

Comparing this with (2.2), one obtains $\Gamma(\varphi,\varphi)\geqq D^2\Phi(x^*)[\hat{\varphi}^2]$ for all $\varphi\in B^*$ as is claimed in Lemma 1.

Remarks. a) We have $\Gamma(\varphi,\varphi)=|\hat{\varphi}|_H^2$ and by continuity, we see that $|y|_H^2\geqq D^2\Phi(x^*)[y^2]$ holds true for all $y\in H$.

b) $S=\{y\in H\colon |y|_H^2=D^2\Phi(x^*)[y^2]\}$ is easily seen to be a closed linear subspace of H. We claim that it is finite dimensional. Indeed, $D^2\Phi(x^*)$ defines a bounded operator $B\to B^*$. Taking compositions with the compact operators $i\colon H\to B$, $j\colon B^*\to H$, we see that $D^2\Phi(x^*)$ defines a compact self-adjoint operator $H\to H$. If $(e_n)_{n\in\mathbb{N}}$ is any countable orthonormal family of vectors in H then $\lim_{n\to\infty} D^2\Phi(x^*)[e_n^2]=0$. This proves the claim.

c) If (1.9) is satisfied, then $\dim(S)=0$.

To prove this, assume $\dim(S)>0$. Then there is a $y\in H$ with $1=|y|_H^2=D^2\Phi(x^*)[y^2]$. Let $z\in H$ satisfy $D^2\Phi(x^*)[y,z]=0$. Then $y_t=(y+tz)/(1+t^2D^2\Phi(x^*)[z^2])^{\frac{1}{2}}$, $t\in\mathbb{R}$, is well defined (at least if $|t|$ is small enough) and satisfies $D^2\Phi(x^*)[y_t^2]=1$. From Lemma 1 one therefore has $|y_t|_H^2\geqq 1$ and so

$|y_t|_H^2$ is minimal at $t=0$. This implies $\langle y, z\rangle_H=0$. If we put $\psi = D^2\Phi(x^*)[y,.]\in B^*$, we therefore have $\hat{\psi}=\alpha y$ for some $\alpha\in\mathbb{R}\setminus\{0\}$.

But this contradicts (1.9).

Proof of Lemma 2. By the consideration in the Remark b) above, $D^2\Phi(x^*)$ defines a compact self-adjoint operator $H\to H$.

We choose an orthonormal base (e_n) of H with $D^2\Phi(x^*)[e_i, e_j]=\lambda_i\delta_{ij}$. $\lambda_i\to 0$ as $i\to\infty$ and from Remark c) we see that $\lambda_i<1$ for all i.

Let ξ_n, $n\in\mathbb{N}$, be a sequence of i.i.d. standard normally distributed random variables, defined on some probability space. Then $\sum_i e_i\xi_i$ converges a.s. and in L_2 to a B-valued random vector with law γ (see [9], p. 157). Then $\sum_i e_i\xi_i/\sqrt{1-\lambda_i}$ converges a.s. and in L_2, too (see Theorem 5.8 of [7]). We denote by γ' the law of this limit which is clearly centered Gaussian. We claim that $\gamma'\ll\gamma$ and $d\gamma'/d\gamma$ has the desired form.

$$D^2\Phi(x^*)[(\sum_i \xi_i e_i)^2]=\sum_i \lambda_i\xi_i^2=\sum_{i:\lambda_i>0}\lambda_i\xi_i^2-\sum_{i:\lambda_i<0}(-\lambda_i)\xi_i^2.$$

As these two summands are independent, and $E|D^2\Phi(x^*)[(\sum_i \xi_i e_i)^2]|<\infty$, we see that $\sum_i|\lambda_i|<\infty$.

We can now apply the Kakutani-criterium (see [9], p. 116) to conclude that the law of $(\xi_1/\sqrt{1-\lambda_1}, \xi_2/\sqrt{1-\lambda_2}, \ldots)$ on $\mathbb{R}^{\mathbb{N}}$ is absolutely continuous with respect to that of $(\xi_1, \xi_2, \ldots)$ with a density

$$\text{const.}\exp\left(\tfrac{1}{2}\sum_i \lambda_i x_i^2\right), \qquad (x_1, x_2, \ldots)\in\mathbb{R}^{\mathbb{N}}.$$

From this one derives in a standard way that

$$(d\gamma'/d\gamma)(y)=\text{const.}\exp\left(\tfrac{1}{2}D^2\Phi(x^*)[y^2]\right).$$

§ 3. Proof of Theorem 3

Let ρ be a probability measure on B which satisfies (1.1), (1.2) and let ρ_n be defined by $\rho_n(A)=\rho^{*n}(\sqrt{n}A)$.

Lemma 5. *If ρ satisfies the central limit theorem then there is linear subset $B_0\subset B$ with a norm $|\ |_0$ on B_0 such that*

(a) *$(B_0, |\ |_0)$ is a Banach space;*

(b) *$\{x\in B_0: |x|_0\leqq 1\}$ is compact in B;*

(c) $\rho(B_0)=1$;

(d) $\int\exp(|x|_0)\,\rho(dx)<\infty$;

(e) $\sup_n\int|x|_0\,\rho_n(dx)<\infty$.

Proof. By the Banach-Mazur theorem, there exists an isometric imbedding of B into $C=C[0,1]$, such that B becomes a closed linear subspace of C.

Let $\{e_j\}$, $j \in \mathbb{N}$, be a normalized Schauder basis of C with associated coordinate functionals $f_j \in C^*$.

If $x \in C$, let $\pi_k(x) = \sum_{j=k+1}^{\infty} f_j(x) e_j$ and let F_k be the subspace of C spanned by $e_1, \ldots, e_k$.

It is easy to see that there is a constant $M > 0$ such that

(3.1) $|\pi_k(x)| \leqq M d(x, F_k)$ for all k and $x \in C$
where $d(x, F_k) = \min\{|x-y|: y \in F_k\}$.

Indeed, by the Banach-Steinhaus theorem

$$b = \sup\left\{\left|x - \sum_{j=1}^{k} f_j(x) e_j\right| : k \in \mathbb{N},\ x \text{ with } |x| \leqq 1\right\} < \infty.$$

To $x \in B$ we choose $x^{(k)} \in F_k$ with $|x - x^{(k)}| \leqq 2 d(x, F_k)$. Then $\sum_{j=1}^{k} f_j(x^{(k)}) e_j = x^{(k)}$ and therefore $\left|\sum_{j=k+1}^{\infty} f_j(x) e_j\right| \leqq 2 b\, d(x, F_k)$.

As ρ_n is tight (in B and therefore in C), it follows that for all $\varepsilon > 0$ $\lim_{k\to\infty} \sup_n \rho_n(\{x: |\pi_k(x)| > \varepsilon\}) = 0$. Therefore

$$\begin{aligned}\sup_n \int |\pi_k(x)|\, \rho_n(dx) &\leqq \varepsilon + \sup_n \int_{\{x:\, \pi_k(x) > \varepsilon\}} |\pi_k(x)|\, \rho_n(dx) \\ &\leqq \varepsilon + M \sup_n \left(\int |x|^2 \rho_n(dx)\right)^{\frac{1}{2}} \rho_n(\{x: |\pi_k(x)| > \varepsilon\})^{\frac{1}{2}}.\end{aligned}$$

By a result of the Acosta and Giné (see [1]) $\int |x|^2 \rho_n(dx)$ converges, so it follows that

$$\lim_{k\to\infty} \sup_n \int |\pi_k(x)|\, \rho_n(dx) = 0.$$

If $k \in \mathbb{N}$, let $n_k^0 \in \mathbb{N}$ satisfy

(3.2) $$\sup_n \int |\pi_m(x)|\, \rho_n(dx) \leqq k^{-3} \quad \text{for } m \geqq n_k^0.$$

From (1.1) and (3.1) one obtains

$$\lim_{k\to\infty} \int \exp(t\,|\pi_k(x)|)\, \rho(dx) = 1 \quad \text{for all } t \in \mathbb{R}.$$

We choose $n_k \geqq n_k^0$ increasing in k with

(3.3) $$\int \exp(2k 2^k |\pi_m(x)|)\, \rho(dx) \leqq 2 \quad \text{for } m \geqq n_{k-1}.$$

If $x \in C$, let $x_1 = x - \pi_{n_1}(x)$ and $x_k = \pi_{n_{k-1}}(x) - \pi_{n_k}(x)$ for $k \geqq 2$. We put $|x|_0 = \sum_{k=1}^{\infty} k |x_k| \in [0, \infty]$.

Let $C_0 = \{x \in C: |x|_0 < \infty\}$. $(C_0, |\ |_0)$ clearly is a (separable) Banach space.

We put $B_0 = B \cap C_0$. Then $(B_0, |\ |_0)$ is a Banach space, too. So (a) is proved.

To prove (b) we remark that $\{x \in C: |x|_0 \leqq 1\}$ is compact in C, so its intersection with B is also compact.

(c) clearly follows from (d). This follows by estimating as follows:

$$\int e^{|x|_0} \rho(dx) = \int \exp\left(\sum_{k=1}^{\infty} k|\pi_{n_{k-1}}(x) - \pi_{n_k}(x)|\right) \rho(dx)$$
$$\leqq \prod_{k=1}^{\infty} \{\int \exp(k 2^k (|\pi_{n_k}(x)| + |\pi_{n_{k-1}}(x)|)) \rho(dx)\}^{2^{-k}}$$
$$\leqq \prod_{k=1}^{\infty} \{\int \exp(2k 2^k |\pi_{n_k}(x)|) \rho(dx) \int \exp(2k 2^k |\pi_{n_{k-1}}(x)|) \rho(dx)\}^{2^{-k-1}}$$

which by (3.3) is $< \infty$.

To see (e) we use (3.2) to estimate

$$\int |x|_0 \rho_n(dx) = \sum_{k=1}^{\infty} \int k|\pi_{n_k}(x) - \pi_{n_{k-1}}(x)| \rho_n(dx)$$
$$\leqq \sum_{k=1}^{\infty} \int k(|\pi_{n_k}(x)| + |\pi_{n_{k-1}}(x)|) \rho_n(dx),$$

which is bounded uniformly in n. □

The following lemma is a corollary of a result of Yurinskii:

Lemma 6. *Under the conditions of Lemma* 5

$$\limsup_{c \to \infty} \sup_{n,t} \left\{\frac{1}{t^2} \log \rho_n(\{x: |x|_0 > t\}): c \leqq t \leqq \sqrt{n}/c\right\} < 0.$$

Proof. If $m \in \mathbb{N}$, then

$$\int |x|_0^m \rho(dx) = m \int_0^{\infty} \lambda^{m-1} \rho(\{x: |x|_0 > \lambda\}) d\lambda$$
$$\leqq m \int e^{|x|_0} \rho(dx) \int_0^{\infty} \lambda^{m-1} e^{-\lambda} d\lambda \leqq m! \int e^{|x|_0} \rho(dx).$$

Therefore ρ satisfies the condition (2.1) in [15] with $H=1$ and $b_j^2 = 2\int e^{|x|_0} \rho(dx)$. By Lemma 5(e) one has $\int |x|_0 \rho^{*n}(dx) = O(\sqrt{n})$.

The statement of our lemma now follows easily from Yurinskii's Theorem 2.1. □

Proof of Theorem 3. We first assume that F is compact in B and satisfies $\Gamma(F) < \infty$. For any $r > 0$, $\{x \in B: |x|_H \leqq r\}$ is compact in B and therefore closed. Given $\varepsilon > 0$ $\{x \in B: \frac{1}{2}|x|_H^2 > \Gamma(F) - \varepsilon\}$ is open in B and contains F. Therefore, there exists a finite covering of F with open balls $U_1, \ldots, U_m$ which satisfy $\frac{1}{2}|x|_H^2 > \Gamma(F) - \varepsilon$, whenever $x \in \bigcup_{j=1}^m U_j$. (We assume $\Gamma(F) - \varepsilon > 0$). Let $C_j = \{x: \frac{1}{2}|x|_H^2 \leqq \Gamma(U_j)\}$, then C_j is compact, convex and disjoint from U_j. So there exists a hyperplane separating C_j and U_j, i.e. there is a $\varphi_j \in B^*$ with

(3.4) $$U_j \subset \{x\colon \varphi_j(x) > 1\} \subset \{x\colon \tfrac{1}{2}|x|_H^2 > \Gamma(U_j)\}.$$

So we obtain

(3.5) $$\nu_n(tF) \leqq \sum_{j=1}^{m} \nu_n \varphi_j^{-1}((t, \infty)).$$

By using a standard argument in one dimensional large deviation theory, this is easily seen to be $\leqq \sum_{j=1}^{m} \exp(-n h_j(t/\sqrt{n}))$ where h_j is the entropy function of $\nu_0 \varphi_j^{-1}$, i.e. $h_j(x) = \sup_{\lambda \in \mathbb{R}} (\lambda x - \log \int e^{\lambda y} \nu_0 \varphi_j^{-1}(dy))$.

As ν_0 is centered, h_j is smooth near 0 except when $\nu_0 \varphi_j^{-1}(\{0\}) = 1$ in which case $\nu_n \varphi_j^{-1}((t, \infty)) = 0$ for all $t > 0$. So this case doesn't bother us.

In all other cases, one has $h_j(0) = h_j'(0) = 0$ and $h_j''(0) = \sigma_j^{-2}$ where $\sigma_j^2 = \int x^2 \nu_0 \varphi_j^{-1}(dx) = \Gamma(\varphi_j, \varphi_j)$.

Therefore there exists a $\delta > 0$ such that for all j and $t/\sqrt{n} \leqq \delta$ one has $h_j(t/\sqrt{n}) \geqq (1-\varepsilon) t^2 / 2\sigma_j^2 n$.

From the second inclusion in (3.4) one obtains $\Gamma(U_j) \leqq 1/2\sigma_j^2$. Using this, together with (3.5), one obtains

(3.6) $$\nu_n(tF) \leqq m \exp(-t^2(1-\varepsilon)(\Gamma(F) - \varepsilon)) \quad \text{for } t/\sqrt{n} \leqq \delta.$$

This proves Theorem 3 in this case. If $\Gamma(F) = \infty$ then the proof is similar.

Let now F be closed.

Let $|\ |_0$ be constructed according to Lemma 5 (for ν). For $D_t = \{x\colon |x|_0 > t\}$ we have by Lemma 6 $a = \limsup_{c \to \infty} \sup \frac{1}{t^2} \log \nu_n(D_t) < 0$, the sup being over those t, n with $c \leqq t \leqq \sqrt{n}/c$.

For any $r > 0$ one has

$$\begin{aligned}\limsup_{c\to\infty} \sup \frac{1}{t^2} \log \nu_n(tF) &\leqq \max\Big(\limsup_{c\to\infty} \sup \frac{1}{t^2} \nu_n(t(F \cap D_r^c)),\\ &\qquad \limsup_{c\to\infty} \sup \frac{1}{t^2} \nu_n(D_{rt})\Big)\\ &\leqq \max(-\Gamma(F \cap D_r^c), r^2 a) \leqq \max(-\Gamma(F), r^2 a).\end{aligned}$$

Letting $r \to \infty$ the theorem is proved. □

§ 4. Proof of the Theorem 1 and 2

If $\underline{x} = (x_1, \ldots, x_n) \in B^n$, we write $s_n(\underline{x})$ or for short just $s_n = \sum_{i=1}^{n} x_i$. Then

(4.1) $$\begin{aligned} E(\exp(n\Phi(S_n/n))) &= \int \exp(n\Phi(s_n(\underline{x})/n)) \mu^n(d\underline{x})\\ &= \exp(n(\log M(D\Phi(x^*)) - D\Phi(x^*)[x^*] + \Phi(x^*)))\\ &\quad \cdot \int \exp(n(\Phi(s_n/n) - \Phi(x^*) - D\Phi(x^*)[s_n/n - x^*])) \nu^n(d\underline{x}).\end{aligned}$$

The factor before the integral is just $\exp(n(\Phi(x^*)-h(x^*)))$. We split the integral into three parts:

Let

$$A_1(c_1, n)=\{\underline{x}\in B^n\colon |s_n/n-x^*|\leqq c_1/\sqrt{n}\}$$
$$A_2(c_1, c_2, n)=\{\underline{x}\in B^n\colon c_1/\sqrt{n}<|s_n/n-x^*|\leqq c_2\}$$
$$A_3(c_2, n)=\{\underline{x}\in B^n\colon c_2<|s_n/n-x^*|\}$$

and we write the integral in (4.1) as

$$=\int_{A_1}+\int_{A_2}+\int_{A_3}=I_1(c_1, n)+I_2(c_1, c_2, n)+I_3(c_2, n).$$

We shall prove:

(4.2) $\hat{I}_1(c_1)=\lim_{n\to\infty} I_1(c_1, n)$ exists for all but countably many $c_1>0$.

(4.3) $\lim_{c_1\to\infty} \hat{I}_1(c_1)=\int \exp((1/2)D^2\Phi(x^*)[y^2])\,\gamma(dy)$.

(4.4) $\lim_{c_1\to\infty} \sup_n I_2(c_1, c_2, n)=0$ for small enough c_2.

(4.5) $\lim_{n\to\infty} I_3(c_2, n)=0$ for all $c_2>0$.

From (4.2)–(4.5) the Theorem 1 clearly follows.

Proof of (4.2) and (4.3). If $x\in B$ then

$$\Phi(x+x^*)-\Phi(x^*)=D\Phi(x^*)[x]+(1/2)D^2\Phi(x^*)[x^2] + (1/6)D^3\Phi(x^*+\theta x)[x^3] \tag{4.6}$$

where $0\leqq\theta\leqq 1$. Therefore

$$n(\Phi(x+x^*)-\Phi(x^*)-D\Phi(x^*)[x])=(1/2)D^2\Phi(x^*)[(\sqrt{n}x)^2]+O(1/\sqrt{n})$$

uniformly in x as long as $|x|\leqq c_1/\sqrt{n}$.

By the central limit theorem

$$\lim_{n\to\infty} I_1(c_1, n)=\int_{|y|\leqq c_1} \exp((1/2)D^2\Phi(x^*)[y^2])\,\gamma(dy)$$

except for countably many c_1. From this (4.2) and (4.3) follow.

Proof of (4.4). If $\varepsilon>0$, let $A_\varepsilon=\{x\in B\colon (1/2)D^2\Phi(x^*)[x^2]+\varepsilon|x|^2\geqq 1\}$. We claim that for sufficiently small $\varepsilon>0$ one has $\Gamma(A_\varepsilon)>1$. To prove this, let $B_r=\{x\in B\colon |x|^2\geqq r\}$ and $\bar{A}_\delta=\{x\in B\colon (1/2)\mathrm{D}^2\Phi(x^*)[x^2]\geqq 1-\delta\}$. Then $A_\varepsilon\subset B_r\cup\bar{A}_{r\varepsilon}$ and so

$$\Gamma(A_\varepsilon)\geqq\min(\Gamma(B_r), \Gamma(\bar{A}_{r\varepsilon})). \tag{4.7}$$

As $\{x\in B\colon |x|_H^2\leqq c\}$ is compact in B, it follows from Remark c) following the proof of Lemma 1 that $\Gamma(\bar{A}_0)>1$. If $\delta>0$ then

$$\Gamma(\bar{A}_\delta)=\inf\{(1/2)|x|_H^2\colon (1/2)D^2\Phi(x^*)[x^2]\geqq 1-\delta\}=(1-\delta)\Gamma(\bar{A}_0).$$

Therefore, $\Gamma(\bar{A}_\delta)>1$ for small enough $\delta>0$.

As $\{x \in B\colon |x|_H^2 \leqq c\}$ is bounded in B we have $\Gamma(B_r) \to \infty$ for $r \to \infty$. So we conclude from (4.7) that $\Gamma(A_\varepsilon) > 1$ for small enough $\varepsilon > 0$.

Using (4.6) one obtains for small enough c_2

$$I_2(c_1, c_2, n) = \int_{c_1 < |y(x)| \leqq c_2 \sqrt{n}} \exp\{(1/2) D^2 \Phi(x^*)[y(x)^2] + (1/6\sqrt{n}) D^3 \Phi(x^* + \theta y(\underline{x})/\sqrt{n})[y(\underline{x})^3]\} \, v^n(d\underline{x})$$

where $y(\underline{x}) = \sqrt{n}(s_n(\underline{x})/n - x^*)$.

For given $\varepsilon > 0$ we can choose c_2 small enough in order that

$$|D^3 \Phi(x^* + \theta y(\underline{x})/\sqrt{n})[y(\underline{x})^3]/6\sqrt{n}| \leqq \varepsilon |y(\underline{x})|^2$$

in the domain of integration. Therefore

$$\begin{aligned} I_2 &\leqq \int_{c_1 < |y| \leqq c_2 \sqrt{n}} \exp((1/2) D^2 \Phi(x^*)[y^2] + \varepsilon |y|^2) \, v^n(d\underline{x}) \\ &= \int_{-\infty}^{\infty} e^t v^n(\{\underline{x}\colon (1/2) D^2 \Phi(x^*)[y(\underline{x})^2] + \varepsilon |y(\underline{x})|^2 \geqq t, \\ &\qquad c_1 < |y(\underline{x})| \leqq c_2 \sqrt{n}\}) \, dt \\ &= \int_{-\infty}^{\infty} e^t v^n(\{\underline{x}\colon y(\underline{x}) \in \sqrt{t} A_\varepsilon, c_1 < |y(\underline{x})| \leqq c_2 \sqrt{n}\}) \, dt. \end{aligned}$$

According to Theorem 3, we find $c < \infty$ and $q > 1$ such that

$$v^n(\{\underline{x}\colon y(\underline{x}) \in \sqrt{t} A_\varepsilon\}) \leqq \exp(-qt)$$

whenever $c \leqq t \leqq \sqrt{n}/c$. We choose c_2 small enough such that

$$v^n(\{y \in \sqrt{t} A_\varepsilon, c_1 < y \leqq c_2 \sqrt{n}\}) = 0$$

if $t > \sqrt{n}/c$. Therefore, if $d > c$ then

$$I_2 \leqq \int_{-\infty}^{d} e^t v^n(y > c_1) \, dt + \int_d^{\infty} e^t e^{-qt} \, dt.$$

So $\limsup_{c_1 \to \infty} \sup_n I_2(c_1, c_2, n) \leqq \int_d^{\infty} e^{-(q-1)t} \, dt$.

By letting $d \to \infty$, (4.4) follows.

Proof of (4.5). If $c_3 > c_2$ then

$$\int_{c_2 < |s_n/n - x^*| \leqq c_3} \exp(n(\Phi(s_n/n) - \Phi(x^*) - D\Phi(x^*)[s_n/n - x^*])) \, v^n(d\underline{x})$$

converges to 0 exponentially fast as $n \to \infty$ by standard large deviation results (see [2] and [5]). (Remark that the sum in the square brackets remains bounded in the domain of integration.)

So it remains to estimate $I_3(c_3, n)$ for arbitrary large c_3. There is an $M>0$ such that

$$I_3(c_3, n) \leqq \int_{|s_n/n - x^*| \geqq c_3} \exp(M|s_n|)\, v^n(d\underline{x})$$
$$\leqq \left(\int \exp(2M|s_n|)\, v^n(d\underline{x})\, v^n(|s_n/n - x^*| \geqq c_3)\right)^{\frac{1}{2}}$$
$$\leqq \left(\int \exp(2M|x|)\, v(dx)\right)^{n/2} v^n(|s_n/n - x^*| \geqq c_3)^{\frac{1}{2}}.$$

Using Lemma 4 and standard large deviation results, one sees that the second factor goes to 0 exponentially fast with an arbitrary large exponential rate if c_3 is chosen sufficiently large. So $\lim_{n\to\infty} I_3(c_3, n) = 0$ follows and so (4.5).

Therefore, Theorem 1 is proved. □

Proof of Theorem 2. We only sketch the proof as it is straightforward from the method in the proof of Theorem 1.

First of all, it follows from the proof of (4.4) that

$$\lim_{t\to\infty} \sup_n E\left(e^{n(h(x^*) - \Phi(x^*))}\, e^{n\Phi\left(\frac{S_n}{n}\right)}\, 1_{\{|\sqrt{n}(S_n/n - x^*)|_0 \geqq t\}}\right) = 0.$$

Therefore the sequence of laws of $\sqrt{n}\left(\frac{S_n}{n} - x^*\right)$ under $\hat{P}_n$ are tight. If $\varphi \in B^*$, then using the same method as in the proof of Theorem 1, one obtains for $\lambda \in \mathbb{R}$

$$\lim_n E\left(e^{i\lambda\varphi(\sqrt{n}(S_n/n - x^*))}\, e^{n\Phi(S_n/n)}\right) / E\left(e^{n\Phi\left(\frac{S_n}{n}\right)}\right)$$
$$= Z^{-1} \int e^{i\lambda\varphi(x) + D^2\Phi(x^*)[x^2]}\, \gamma(dx)$$

with $Z = \int e^{\frac{1}{2}D^2\Phi(x^*)[x^2]}\, \gamma(dx)$ and this, by Lemma 2, equals $\int e^{i\lambda\varphi(x)}\, \gamma'(dx)$.

So Theorem 2 is proved. □

Acknowledgements. The author thanks the refree for suggesting a number of improvements and pointing at an error in the original manuscript.

References

1. de Acosta, A., Giné, E.: Convergence of moments and related functionals in the general central limit theorem in Banach spaces. Z. Wahrscheinlichkeitstheor. Verw. Geb. **48**, 213–231 (1979)
2. Azencott, R.: Grandes déviations et applications. Saint-Flour VIII 1978. Lecture Notes in Math. **774**. Berlin-Heidelberg-New York: Springer 1980
3. Borell, C.: Gaussian Radon measures on locally convex spaces. Math. Scand. **38**, 265–284 (1976)
4. Csziszar, I.: *I*-divergence geometry of probability distributions and minimization problems. Ann. Probab. **3**, 146–158 (1975)
5. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time III. Comm. Pure Appl. Math. **29**, 389–461 (1976)
6. Ellis, R.S., Rosen, J.S.: Asymptotic analysis of Gaussian integrals I, II. Trans. Am. Math. Soc. **273**, 447–481 and Comm. Math. Phys. **82**, 153–181 (1982)
7. Hoffmann-Jørgensen, J.: Probability in Banach spaces, Saint-Flour VI-1976, Lecture Notes in Math. **598**. Berlin-Heidelberg-New York: Springer 1977

8. Jain, N.C.: Central limit theorem in Banach spaces, Proc. First Conf. on Prob. in Banach Spaces Oberwolfach 1975. Lecture Notes Math. **526**. Berlin-Heidelberg-New York: Springer 1976
9. Kuo, H.-H.: Gaussian measures in Banach spaces. Lecture Notes Math. **463**. Berlin-Heidelberg-New York: Springer 1975
10. Kusuoka, Sh., Tamura, Y.: The convergence of Gibbs measures associated with mean field potentials. J. Fac. Sci., Univ. Tokyo, Sect. 1A **31**, 223–245 (1984)
11. Martin-Löf, A.: Laplace approximation for sums of independent random variables. Z. Wahrscheinlichkeitstheor. Verw. Geb. **59**, 101–115 (1982)
12. Pincus, M.: Gaussian processes and Hammerstein integral equations. Trans. Am. Math. Soc. **134**, 193–216 (1968)
13. Schilder, M.: Some asymptotic formulae for Wiener integrals. Trans. Am. Math. Soc. **125**, 63–85 (1966)
14. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Comm. Pure Appl. Math. **19**, 261–286 (1966)
15. Yurinskii, V.V.: Exponential inequalities for sums of random vectors. J. Multivariate Anal. **6**, 473–499 (1966)

Received July 16, 1984; in revised form September 2, 1985