# Large Deviations and Stochastic Homogenization (*).

PAOLO BALDI

**Summary.** – *A general theorem is stated providing large deviations estimates for a family of measures on a topological vector space. Applications are given in the second part, where large deviations problems arising in stochastic homogenization are discussed. Another application is given in similar problems connected with Donsker's invariance principle.*

## 0. – Introduction.

In recent papers dealing with problems of stochastic homogenisation (DAL MASO-MODICA [2], FACCHINETTI-RUSSO [12]), the authors prove that certain sequences $\{\mu_n\}_n$ of probability measures on a space of functionals $\mathcal{F}$ do converge to the Dirac mass concentrated on a certain functional $x$.

The aim of this paper is to give a large deviations estimate for this convergence, that is to evaluate the behaviour as $n \to \infty$ of the quantity $\mu_n(A)$, $A$ being a subset not containing $x$. More precisely we shall be able to give in most cases the equivalent, as $n \to \infty$, of $\log \mu_n(A)$.

In § 1 we shall prove an abstract large deviation theorem that will be used in § 2 to derive such estimates for a problem of stochastic homogenisation. In § 3 we deal with a different kind of applications, connected with Donsker invariance principle.

## 1. – An abstract large deviations theorem.

Let $X$ be a topological vector space, $X'$ its dual; for a probability measure $\mu$ on $X$ let us define, for $\alpha \in X'$, its Laplace transform

$$\hat{\mu}(\alpha) = \int_X \exp \langle \alpha, x \rangle \mu(dx)$$

($\hat{\mu}(\alpha) = +\infty$ possibly). If $H: X' \to \overline{\boldsymbol{R}}$ is a convex function, its Legendre transform is defined by

$$x \to \sup_{\alpha \in X'} (\langle \alpha, x \rangle - H(\alpha)) .$$

(*) Entrata in Redazione il 30 ottobre 1985.

Indirizzo dell'A.: Dipartimento di Matematica, Università di Pisa, Via Buonarroti 2, 56100 Pisa, Italia.

It is a positive, convex and lower semicontinuous function, being the supremum of a family of continuous convex functions. Let now be $\{\mu_h\}_{h>0}$ a family of probability laws on $X$.

ASSUMPTION (A). – $\{\mu_h\}_h$ is said to satisfy to assumption (A) if there exists a function $\lambda: \mathbf{R}^+ \to \mathbf{R}^+$ such that, if $H_h = \log \hat{\mu}_h$,

$$\lim_{h\to\infty} \lambda(h) = +\infty \tag{1.1}$$

$$\lim_{h\to+\infty} \frac{1}{\lambda(h)} H_h(\lambda(h)\alpha) = H(\alpha) \tag{1.2}$$

where $H: X' \to \mathbf{R} \cup \{+\infty\}$ is lower semicontinuous and finite in a neighborhood of the origin. By $L$ we shall denote the Legendre transform of $H$.

ASSUMPTION (C). – $\{\mu_h\}_h$ fulfills Assumption (C) if Assumption (A) is verified and, moreover, for every $R > 0$ there exists a compact set $K_R \subset X$ such that

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_h(K_R^c) \leqslant -R \,. \tag{1.3}$$

As it will be made clear later (see proposition 1.5 and the remark following Lemma 1.6), if $X$ is finite-dimensional Assumption (A) implies that $\mu_h \to \delta_x$ in the weak convergence of measures, for some $x \in X$. Assumption (C), which in the finite-dimensional case follows from Assumption (A), is needed to ensure tightness of the family $\{\mu_h\}_h$.

Recall that a convex function $\psi$ is said to be strictly convex at $x_0$ if there exists $\alpha \in X'$ such that

$$\psi(y) > \psi(x_0) + \langle \alpha, y - x_0 \rangle$$

for every $y \neq x_0$.

For every $A \subset X$ let us define

$$\Lambda(A) = \inf_{x\in A} L(x) \,.$$

In this section we shall prove the following

THEOREM 1.1. – Suppose that Assumption (C) holds and that for every $k$ the closure of the set $\{x;\, L(x) \leqslant k,\ L \text{ is strictly convex at } x\}$ contains $\{x;\, L \leqslant k\}$. Then for every Borel subset $A$ of $X$

$$-\Lambda(\mathring{A}) \leqslant \underline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_h(A) \leqslant \overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_h(A) \leqslant -\Lambda(\overline{A}) \,. \tag{1.4}$$

Theorem 1.1 will follow from Corollaries 1.4 and 1.8 below.

THEOREM 1.2. – Under the hypothesis of Theorem 1.1 if moreover there exists a Borel subset $\mathcal{F} \subset X$ such that

i) $\mu_h(\mathcal{F}) = 1$ for every $h > 0$;

ii) $\mathcal{F} \supset \{L < +\infty\}$.

Then for every $A \subset \mathcal{F}$

$$-\Lambda(\mathring{A}_{\mathcal{F}}) \leqslant \varliminf_{h\to+\infty} \frac{1}{\lambda(h)} \log \mu_h(A) \leqslant \varlimsup_{h\to+\infty} \frac{1}{\lambda(h)} \log \mu_h(A) \leqslant -\Lambda(\bar{A}_{\mathcal{F}})$$

$\mathring{A}_{\mathcal{F}}$ and $\bar{A}_{\mathcal{F}}$ being respectively the interior and the closure of $A$ in the topology induced on $\mathcal{F}$ by $X$.

PROOF. – Theorem 1.2 follows easily from Theorem 1.1. Indeed, if $A$ is closed in $\mathcal{F}$, then if $F$ is a closed subset of $X$ such that $A = F \cap \mathcal{F}$, then $\mu_h(A) = \mu_h(F)$ and $\Lambda(A) = \Lambda(F)$. The same argument works if $A$ is open in $\mathcal{F}$.

REMARKS. – 1) Theorems 1.1 and 1.2 are wellknown if $X$ is finite dimensional (GÄRTNER [5], WENTZELL-FREIDLIN [10], § 5.1). For the infinite dimensional case, recall a result by DAWSON-GÄRTNER [11]; their statement holds for a space $X^*$ which is the dual of a Banach space $X$ and is endowed with the weak*-topology and do not require Assumption (C). Their idea is to extend the finite dimensional result by projective limits, whereas our proof actually shows that, under Assumption (C), the finite dimensional proof still works in infinite dimensions.

2) Strict convexity of $L$ at $x \in \{L < +\infty\}$ follows from smoothness of $H$ (see, ROCKAFELLAR [7], chap. 2.6).

3) Condition ii) of Theorem 1.2 is a consequence of i) if $\mathcal{F}$ is a closed convex subset of $X$.

4) Under the hypothesis of theorem 1.1 the set $\{L \leqslant \gamma\}$ is compact for every $\gamma \geqslant 0$. Indeed it is closed, $L$ being l.s.c., and if $K_{\gamma+1}$ is as in (1.3) with $R$ replaced by $\gamma+1$, then $K_{\gamma+1}^c$ is an open set and by Theorem 1.1 $\Lambda(K_{\gamma+1}^c) \geqslant \gamma+1$. This implies $\{L \leqslant \gamma\} \subset K_{\gamma+1}$.

PROPOSITION 1.3 (Assumption (C)). – If

$$A_\gamma = \{x;\ L(x) \leqslant \gamma\}$$

then for every open set $U$ containing $A_\gamma$, there exists $\delta > 0$ such that for $h > h_0$

$$\mu_h(U^c) \leqslant \exp[-\lambda(h)(\gamma + \delta)].$$

PROOF. – Let $K_{\gamma+2}$ be a compact set as in (1.3) with $R = \gamma + 2$. Then

$$\mu_h(U^c) \leqslant \mu_h(U^c \cap K_{\gamma+2}) + \mu_h(K^c_{\gamma+2}) . \tag{1.5}$$

Since $L$ is l.s.c., the minimum of $L$ on the compact $U^c \cap K_{\gamma+2}$ is a attained, so that $L(x) \geqslant \gamma + 2\eta$ for every $x \in U^c \cap K_{\gamma+2}$, for some $\eta > 0$.

Thus, for every $x \in U^c \cap K_{2+\gamma}$, there exists $\alpha = \alpha(x) \in X'$ such that

$$\langle \alpha, x \rangle - H(\alpha) > \gamma + \eta .$$

If $S_x = \{y; \langle \alpha(x), y \rangle - H(\alpha(x)) > \gamma + \eta\}$, there exist $x_1, \ldots, x_n$ such that

$$U^c \cap K_{\gamma+2} \subset S_{x_1} \cup \ldots \cup S_{x_n}$$

so that, writing $S_i$ for $S_{x_i}$, $\alpha_i$ for $\alpha(x_i)$,

$$\begin{aligned}\mu_h(U^c \cap K_{\gamma+2}) &\leqslant \sum_{i=1}^n \mu_h(S_i) \leqslant \sum_{i=1}^n \mu_h\{y;\ \exp[\lambda(h)(\langle \alpha_i, y \rangle - H(\alpha_i) - \gamma - \eta)] > 1\} \leqslant \\ &\leqslant \sum_{i=1}^n \int \exp[\lambda(h)(\langle \alpha_i, x \rangle - H(\alpha_i) - \gamma - \eta)]\mu_h(dy) = \\ &= \exp[-\lambda(h)(\gamma + \eta)] \sum_{i=1}^n \exp\left[\lambda(h)\left(\frac{1}{\lambda(h)} H^h(\lambda(h)\alpha_i) - H(\alpha_i)\right)\right].\end{aligned}$$

By (1.2) $(1/\lambda(h))H^h(\lambda(h)\alpha_i) - H(\alpha_i) < \eta/2$ for $h$ large and for every $i$, so that for $h$ large

$$\mu(U^c \cap K_{\gamma+2}) \leqslant n \exp\left[-\lambda(h)\left(\gamma + \frac{\eta}{2}\right)\right].$$

Since for large $h$ $\mu_h(K^c_{\gamma+2}) \leqslant \exp[-\lambda(h)(\gamma+1)]$, from (1.5) the statement is proved with $\delta = \min(\eta/4, 1)$.

COROLLARY 1.4 (Assumption (C)). – For every Borel subset $A$ of $X$

$$\overline{\lim_{h\to\infty}} \frac{1}{\lambda(h)} \log \mu_h(A) \leqslant -\Lambda(\bar{A}) .$$

PROOF. – It suffices to remark that if $\Lambda(\bar{A}) > \gamma$, then $A^c$ is a neighborhood of $\{L \leqslant \gamma\}$ and then apply proposition 1.3.

PROPOSITION 1.5. – If $X = \boldsymbol{R}^m$ then Assumption (A) implies Assumption (C).

PROOF. – Let $(\alpha_1, \ldots, \alpha_m)$ be a basis of $X'$, chosen so that $H(\alpha_i) < +\infty$ for every $i = 1, \ldots, m$.

For fixed $R > 0$ if

$$K_\gamma = \{x;\ |\langle \alpha_i, x\rangle| \leqslant \gamma,\ i = 1, \dots, m\}$$

we shall prove that

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_h(K_\gamma^c) \leqslant -R \tag{1.6}$$

for $\gamma$ large enough. One has

$$\mu_h(K_\gamma^c) \leqslant \sum_{i=1}^{m} \Big(\mu_h(\{x;\ \langle \alpha_i, x\rangle > \gamma\}) + \mu_h(\{x;\ \langle \alpha_i, x\rangle < -\gamma\})\Big)\,.$$

But remembering that $\exp H_h(\alpha) = \int \exp \langle \alpha, x\rangle \mu_h(dx)$

$$\begin{aligned}\mu_h(\{x;\ \langle \alpha_i, x\rangle > \gamma\}) &= \\ &= \mu_h\left\{x;\ \lambda(h)\langle \alpha_i, x\rangle - H_h(\lambda(h)\alpha_i) > \lambda(h)\left(\gamma - \frac{1}{\lambda(h)} H_h(\lambda(h)\alpha_i)\right)\right\} \leqslant \\ &\leqslant \exp\left[-\lambda(h)\left(\gamma - \frac{1}{\lambda(h)} H_h(\lambda(h)\alpha_i)\right)\right] \leqslant \exp\left[-\lambda(h)(\gamma - H(\alpha_i) - 1)\right]\end{aligned}$$

where the last inequality follows from (1.2) for large $h$. Thus if $\gamma \geqslant 2R + H(\alpha_i) + 1$

$$\mu_h(\{x;\ \langle \alpha_i, x\rangle > \gamma\}) \leqslant \exp\{-\lambda(h)2R\}$$

and

$$\mu_h(K_\gamma^c) \leqslant 2m \exp[-2\lambda(h)R] \leqslant \exp[-\lambda(h)R]$$

for $h$ large.

LEMMA 1.6. – Suppose that $\{\mu_h\}_h$ satisfies Assumption (C). Let be $\alpha \in X$ such that $H(\alpha) < +\infty$ and denote by $\mu_{h,\alpha}$ the probability law on $X$ defined by

$$d\mu_{h,\alpha}(x) = \exp[\lambda(h)\langle \alpha, x\rangle - H_h(\lambda(h)\alpha)]\, d\mu_h(x)\,.$$

Then $\{\mu_{h,\alpha}\}_h$ also satisfies Assumption (C).

PROOF. – Let $\nu_{h,\alpha}$ be the probability law on $\boldsymbol{R}$ which is the image of $\mu_{h,\alpha}$ through the application $y \to \langle \alpha, y\rangle$. Its Laplace transform is given by

$$\hat{\nu}_{h,\alpha}(\theta) = \hat{\mu}_h(\theta\alpha + \lambda(h)\alpha)\frac{1}{\hat{\mu}_h(\lambda(h)\alpha)}$$

so that

$$\lim_{h\to+\infty} \frac{1}{\lambda(h)} \log \hat{\nu}_{h,\alpha}(\lambda(h)\theta) = H((\theta + 1)\alpha) - H(\alpha)\,.$$

Thus $\{\nu_{h,\alpha}\}_h$ satisfies to Assumption (A) and, by Proposition 1.5, to Assumption (C). Thus, for every $R > 0$ there exists $\gamma > 0$ such that

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \nu_{h,\alpha}([-\gamma, \gamma]^c) \leqslant -R$$

or equivalently if $C_{\alpha,\gamma} = \{x;\ |\langle\alpha, x\rangle| \leqslant \gamma\}$

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_{h,\alpha}(C^c_{\alpha,\gamma}) \leqslant -R\,. \tag{1.7}$$

For $\varrho > 0$ let $K_\varrho$ be a compact set such that

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_h(K^c_\varrho) \leqslant -\varrho\,.$$

Then

$$\mu_{h,\alpha}(K^c_\varrho) \leqslant \mu_{h,\alpha}(K^c_\varrho \cap C_{\alpha,\gamma}) + \mu_{h,\alpha}(C^c_{\alpha,\gamma})\,. \tag{1.8}$$

Since on $C_{\alpha,\gamma}$ $\mu_{h,\alpha}$ has a density with respect to $\mu_h$ which is bounded by $\exp[\lambda(h)\gamma - H_h(\lambda(h)\alpha)]$, if

$$\varrho = R + \gamma - H(\alpha) + 2\,, \quad \text{for large } h$$

$$\begin{aligned}\mu_{h,\alpha}(K^c_\varrho \cap C_{\alpha,\gamma}) &\leqslant \exp[\lambda(h)\gamma - H_h(\lambda(h)\alpha)]\mu_h(K^c_\varrho) \leqslant \\ &\leqslant \exp\Big[\lambda(h)\big(\gamma - H_h(\lambda(h)\alpha) - \varrho + 1\big)\Big] \leqslant \exp[-\lambda(h)R]\,.\end{aligned} \tag{1.9}$$

This combined with (1.7) and (1.8) gives

$$\overline{\lim_{h\to+\infty}} \frac{1}{\lambda(h)} \log \mu_{h,\alpha}(K^c_\varrho) \leqslant -R\,.$$

REMARK. – Under Assumption (C) if there exists $x_0 \in X$ such that $L(x_0) = 0$ whereas $L(x) > 0$ for all $x \neq x_0$, then $\mu_h \to \delta_{x_0}$. Indeed Proposition 1.3 implies that $\mu_h(U^c) \to 0$ for every neighborhood $U$ of $x_0$.

PROPOSITION 1.7 (Assumption (C)). – Suppose that $L$ is strictly convex at $x \in X$. Then for every neighborhood $U$ of $x$ and $\delta > 0$

$$\mu_h(U) \geqslant \exp[-\lambda(h)(L(x) + \delta)]$$

for large $h$.

PROOF. – Let $\alpha \in X'$ be such that

$$L(y) > L(x) + \langle\alpha, y - x\rangle$$

for every $y \neq x$. Thus

$$H(\alpha) = \sup_y [\langle \alpha, y \rangle - L(y)] = \langle \alpha, x \rangle - L(x) \tag{1.10}$$

(since $H$ is itself convex and l.s.c., it coincides with the Legendre transform of $L$) so that for large $h$ $H_h(\lambda(h)\alpha)$ is finite. For such $h$'s let $\mu_{h,\alpha}$ be defined by

$$d\mu_{h,\alpha}(y) = \exp[\lambda(h)\langle \alpha, y \rangle - H_h(\lambda(h)\alpha)]\, d\mu_h(y)$$

so that

$$\begin{aligned} \mu_h(U) = \int_U \exp[-\lambda(h)\langle \alpha, y \rangle + H_h(\lambda(h)\alpha)]\, d\mu_{h,\alpha}(y) \geqslant \\ \geqslant \mu_{h,\alpha}(U) \cdot \min_{y \in U} \exp[-\lambda(h)\langle \alpha, y \rangle + H_h(\lambda(h)\alpha)] . \end{aligned} \tag{1.11}$$

Possibly shrinking $U$, if $y \in U$, $\langle \alpha, y \rangle \leqslant \langle \alpha, x \rangle + \delta/2$ and for large $h$, $y \in U$

$$\langle \alpha, y \rangle - \frac{1}{\lambda(h)} H_h(\lambda(h)\alpha) \leqslant \langle \alpha, x \rangle - H(\alpha) + \frac{\delta}{2} = L(x) + \frac{\delta}{2} .$$

Thus from (1.10)

$$\mu_h(U) \geqslant \mu_{h,\alpha}(U) \exp[-\lambda(h)(L(x) + \delta)]$$

and it is now sufficient to prove that $\mu_{h,\alpha} \to \delta_x$.

If $H_h^\alpha(\beta) = \log \hat{\mu}_{h,\alpha}(\beta)$, $H_h^\alpha(\beta) = H_h(\beta + \lambda(h)\alpha) - H_h(\lambda(h)\alpha)$ so that

$$\lim_{h \to +\infty} \frac{1}{\lambda(h)} H_h^\alpha(\lambda(h)\beta) = H(\alpha + \beta) - H(\alpha) \overset{\text{def}}{=} H^\alpha(\beta)$$

and the Legendre transform $L^\alpha$ of $H^\alpha$ is given by

$$L^\alpha(y) = L(y) + \langle \alpha, y \rangle - H(\alpha) .$$

Thus, recalling (1.10), $L^\alpha(x) = 0$ whereas $L^\alpha(y) > 0$, for $y \neq x$, since $L$, and then $L^\alpha$, are strictly convex at $x$. We may then apply Lemma 1.6 and the remark preceding this proposition.

COROLLARY 1.8. – Under the Assumptions of Theorem 1.1 for every Borel subset $A$ of $X$

$$-\Lambda(\mathring{A}) \leqslant \varliminf_{h \to +\infty} \frac{1}{\lambda(h)} \log \mu_h(A) .$$

PROOF. – We may suppose that $A$ is open. If $\Lambda(A) = +\infty$ there is nothing to prove. If not let $x \in A$ be such that $L$ is strictly convex at $x$ and

$$L(x) \leqslant \Lambda(A) + \eta\,. \tag{1.12}$$

Since $A$ is a neighborhood of $x$, by Proposition 1.7

$$\mu_h(A) \geqslant \exp\left[-\lambda(h)\big(L(x) - \eta\big)\right]$$

which combined with (1.12), $\eta$ being arbitrary, concludes the proof.

## 2. – Application: a problem in stochastic homogenization.

DEFINITION 2.1. – Let $X$ be a topological space, $\{f_n\}_n$ a sequence of real extended functions; we shall say that $\{f_n\}_n$ $\Gamma(X^-)$-converges to $f$ if for every $x \in X$

$$\sup_{\mathcal{U}\in\mathfrak{I}_x} \overline{\lim_{n\to\infty}} \inf_{\mathcal{U}} f_n = \sup_{\mathcal{U}\in\mathfrak{I}_x} \underline{\lim_{n\to\infty}} \inf_{\mathcal{U}} f_n = f(x)$$

$\mathfrak{I}_x$ denoting the class of all neighborhoods of $x$.

The notion of $\Gamma(X^-)$-convergence ($\Gamma$-convergence from now on) is natural in the study of variational problems as Theorem 2.2 will make clear. For more details s see DE GIORGI-FRANZONI [3] or § 1 of DAL MASO-MODICA [2].

A family of real extended functions on $X$ is said to be *equicoercive* if for every $t \in \boldsymbol{R}$ there exists a compact set $K_t \subset X$ such that $\{f_n \leqslant t\} \subset K_t$ for every $n$.

THEOREM 2.2. – Let $\{f_n\}_n$ be a sequence of real extended functions on the topological space $X$. Suppose that

i) $f_n$ $\Gamma$-converges to $f$;

ii) the sequence $\{f_n\}_n$ is equicoercive.

Suppose that for every $n$ $f_n$ attains its minimum at $x_n \in X$ and set $m_n = f_n(x_n)$. Then $f$ attains its minimum in $X$ and

$$\min_{x\in X} f(x) = \lim_{n\to\infty} m_n\,.$$

Moreover any converging subsequence of $\{x_n\}_n$ does converge to a point of minimum for $f$.

We shall consider the following case:

$$X = \{u \in W^{1,2}(\,0,1]),\ u(0) = 0,\ u(1) = 1\} \quad \text{and for } 0 < c < C$$

let $\mathcal{F}$ be the set of all functionals on $X$ of the form

$$F(u) = F_a(u) = \tfrac{1}{2}\int_0^1 a(t)u'(t)^2\,dt \tag{2.1}$$

where $a$ is measurable and $c \leqslant a(t) \leqslant C$ for every $t \in [0, 1]$. On $X$ we consider the $L^2$ topology.

THEOREM 2.3 (SPAGNOLO [9]). – Let $\mathcal{K} \subset L^\infty([0, 1])$ be the subset of all functions $x$ such that $1/C \leqslant x(t) \leqslant 1/c$ for every $t \in [0, 1]$ and let $d$ be a distance on $\mathcal{K}$ inducing the $\sigma(L^\infty, L^1)$ topology. Then the distance $\tilde{d}$ on $\mathcal{F}$ defined by $\tilde{d}(F_a, F_b) = d(1/a, 1/b)$ is such that $F_{a_n} \xrightarrow{\Gamma} F_a$ if and only if $\tilde{d}(F_{a_n}, F_a) \to 0$.

Thus the set of all functionals of the form (2.1) endowed with the topology induced by $\Gamma$-convergence is compact and metrizable.

Let now be $(X_n)_n$ a sequence of i.i.d. random variables on a probability space $(\Omega, \mathcal{A}, P)$ taking values in $[c, C]$ and define

$$a_n(t) = X_i \quad \text{if} \quad \frac{i-1}{n} \leqslant t < \frac{i}{n}$$

so that (by theorem 2.3) $\omega \to F_{a_n(\omega)}$ defines a r.v. taking values in $\mathcal{F}$. It can be shown that $F_{a_n} \to F_{a_\infty}$ a.s. where

$$a_\infty \equiv \frac{1}{E[1/X_1]} \tag{2.2}$$

(see DAL MASO-MODICA [2] for details, however we shall not make use of this fact which will even follow from the results of this section) so that if $\mu_n$ denotes the law of $F_{a_n}$

$$\mu_n \to \delta_{F_{a_\infty}}.$$

In the following $\gamma_n$ will denote the law on $L^\infty([0, 1])$ which is the image of $\mu_n$ through the application $F_a \to 1/a$. Of course the support of $\gamma_n$ is contained in $\mathcal{K}$ for every $n$. The results of section 1 will now be used to prove a large deviations result for the sequence $(\mu_n)_n$. Let us denote by $\nu$ the law of $1/X_1$, $\hat{\nu}$ its Laplace transform and

$$\lambda(x) = \sup_\theta\,[\theta x - \log \hat{\nu}(\theta)] \tag{2.3}$$

so that $\lambda$ is the Cramer transform of $\nu$; $\lambda$ is strictly convex, since $\log \hat{\nu}$ is differentiable and $\lambda(x) = +\infty$ if $x \notin [1/C, 1/c]$ (see ROCKAFELLAR [7], chap. 2.6). Define

$$J(f) = \int_0^1 \lambda(f(s))\,ds\,.$$

Obviously, from the properties of $\lambda$, $J$ is strictly convex and $\{J < +\infty\} \subset \mathcal{K}$. For $A \subset L^\infty([0,1])$ we define

$$I(A) = \inf_{f \in A} J(f)$$

and also

$$\hat{J}(F_a) = J\left(\frac{1}{a}\right), \qquad \hat{I}(B) = \inf_{F_a \in B} \hat{J}(F_a)$$

for every $F_a \in \mathcal{F}$ and $B \subset \mathcal{F}$.

THEOREM 2.4. – For every Borel subset $A \subset L^\infty([0,1])$

$$-I(\mathring{A}) \leqslant \varliminf_{n\to\infty} \frac{1}{n} \log \gamma_n(A) \leqslant \varlimsup_{n\to\infty} \frac{1}{n} \log \gamma_n(A) \leqslant -I(\bar{A})$$

and for every Borel subset $B \subset \mathcal{F}$.

$$-\hat{I}(\mathring{B}) \leqslant \varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(B) \leqslant \varlimsup_{u} \frac{1}{n} \log \mu_n(B) \leqslant -\hat{I}(\bar{B}) .$$

The theorem will follow from theorems 1.1, 1.2, 2.3 and the lemmas below.

First remark that, since $\mathcal{F}$ and $\mathcal{K}$ are compact spaces (1.3) is automatically satisfied so that only Assumption (A) is to be proved. Since $L^1([0,1])$ is obviously the topological dual of the space $L^\infty([0,1])$ endowed with the $\sigma(L^\infty, L^1)$ topology we have first to compute

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{\gamma}_n(ng)$$

for $g \in L^1(\,0,1])$ where

$$\hat{\gamma}(g) = \int \exp \langle f, g\rangle \gamma_n(df) = E\left[\exp\left\langle\frac{1}{a_n}, g\right\rangle\right]$$

(here $\langle\ \rangle$ denotes the duality between $L^1$ and $L^\infty$). An explicit computation gives easily, the r.v. being independent,

$$\text{(2.4)} \qquad \begin{cases} \hat{\gamma}_n(ng) = E\left[\exp\left(\sum_{i=1}^n \frac{n}{X_i} \int_{(i-1)/n}^{i/n} g(s)\, ds\right)\right] = \prod_{i=1}^n E\left[\exp\left(\frac{n}{X_i} \int_{(i-1)/n}^{i/n} g(s)\, ds\right)\right] \\ \frac{1}{n} \log \hat{\gamma}_n(ng) = \frac{1}{n} \sum_{i=1}^n \log \hat{\nu}\left(n \int_{(i-1)/n}^{i/n} g(s)\, ds\right) \end{cases}$$

(remember that $\nu$ is the law of $1/X_1$). In the following we shall write $\Phi = \log \hat{\nu}$; $\Phi$ is convex and continuous.

LEMMA 2.5. – For every $g \in L^1([0,1])$ define

$$g_n(t) = n \int_{(i-1)/n}^{i/n} g(s)\, ds \quad \text{if} \quad \frac{i-1}{n} \leqslant t < \frac{i}{n}.$$

Then $g_n \xrightarrow{L^1} g$.

PROOF. – If $g$ is continuous the statement follows from a uniform continuity argument. If $g \in L^1([0,1])$, then for every $\varepsilon > 0$ there exist $g^1$, $g^2$, with $g^1$ continuous and $\|g^2\|_1 \leqslant \varepsilon$. Moreover, since $\|g_n\|_1 \leqslant \|g\|_1$ for every $n$,

$$\overline{\lim_{n\to\infty}} \|g_n - g\|_1 \leqslant \overline{\lim_{n\to\infty}} (\|g^1 - g^1_n\|_1 + \|g^2 - g^2_n\|) \leqslant \overline{\lim_{n\to\infty}} (\|g^2\|_1 + \|g^2_n\|_1) \leqslant 2\varepsilon$$

which, $\varepsilon$ being arbitrary, completes the proof.

LEMMA 2.6. – For every $g \in L^1([0,1])$

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{\gamma}_n(ng) = \int_0^1 \Phi(g(s))\, ds.$$

PROOF. – Starting from (2.4) and with the notations of Lemma 2.4

$$\frac{1}{n} \sum_{i=1}^n \Phi\left(n \int_{(i-1)/n}^{i/n} g(s)\, ds\right) = \int_0^1 \Phi(g_n(s))\, ds \tag{2.5}$$

and by Jensen inequality

$$\frac{1}{n} \sum_{i=1}^n \Phi\left(n \int_{(i-1)/n}^{i/n} g(s)\, ds\right) \leqslant \int_0^1 \Phi(g(s))\, ds. \tag{2.6}$$

Moreover, by Lemma 2.4, every subsequence of $(g_n)_n$ has a subsequence $(g_{n_k})_k$ converging to $g$ a.e. Thus $\Phi(g_{n_k}) \to \Phi(g)$ a.e. Since $\Phi$ is bounded from below by an affine function and $g_{n_k} \to g$ in $L^1$, by a variant of Fatou's lemma

$$\underline{\lim_{k\to+\infty}} \int_0^1 \Phi(g_{n_k}(s))\, ds \geqslant \int_0^1 \Phi(g(s))\, ds$$

so that, the subsequence being arbitrary,

$$\underline{\lim_{n\to+\infty}} \int_0^1 \Phi(g_n(s))\, ds \geqslant \int_0^1 \Phi(g(s))\, ds$$

that, combined with (2.6) and (2.5), completes the proof.

In order to achieve the proof of Theorem 2.3 we only need to prove that $J$ is the Legendre transform of $g \to \int_0^1 \Phi(g(s))\, ds$.

LEMMA 2.7.

$$J(f) = \int_0^1 \lambda(f(s)) = \sup_{g \in L^1([0,1])} \langle f, g \rangle - \int_0^1 \Phi(g(s))\, ds\,.$$

PROOF. – Remembering (2.3) one has immediately

$$J(f) \leqslant \int_0^1 \sup_g \left(g(s)f(s) - \Phi(g(s)\right) ds = \int_0^1 \lambda(f(s))\, ds\,.$$

Moreover the function $\theta \to \theta x - \Phi(\theta)$ is concave and vanishes at $\theta = 0$, so that it is non negative on an interval $I_x$ (finite, infinite or reduced to $\{0\}$) having 0 as an endpoint. For every $n$, let $\theta_n(x) \in I_x$ be a point such that

$$\begin{aligned} &\theta_n(x)x - \Phi(\theta_n(x)) \geqslant \lambda(x) - \frac{1}{n}, \quad &&\text{if } \lambda(x) < +\infty\,, \\ &\theta_n(x)x - \Phi(\theta_n(x)) \geqslant n &&\text{if } \lambda(x) = +\infty\,. \end{aligned}$$

This can be done in such a way that $x \to \theta_n(x)$ is a measurable function (use a section theorem e.g.). If now $\theta_{n,m}(x) = \theta_n(x) \wedge m \vee - m$, $\theta_{n,m}(x)$ is still in $I_x$, so that

$$\theta_{n,m}(x)x - \Phi(\theta_{n,m}(x)) \geqslant 0\,. \tag{2.7}$$

If we set $g_{n,m}(s) = \theta_{n,m}(f(s))$, $g_{n,m}$ is bounded, and thus in $L^1$; moreover

$$J(f) \geqslant \int_0^1 [f(s)g_{n,m}(s) - \Phi(g_{n,m}(s))]\, ds\,.$$

Since the integrand is non negative by (2.7), letting $m \to +\infty$, by Fatou's lemma

$$J(f) \geqslant \int_0^1 \Big[f(s)\theta_n(f(s)) - \Phi\big(\theta_n(f(s))\big)\Big]\, ds\,. \tag{2.8}$$

Now let $A_f = \{s;\, \lambda(f(s)) = +\infty\}$. If mis $(A_f) = 0$ then (2.8) gives

$$J(f) \geqslant \int_0^1 \lambda(f(s))\, ds - \frac{1}{n}$$

otherwise

$$J(f) \geqslant n \operatorname{mis}(A_f)\,.$$

In either case, $n$ being arbitrary, the statement is proved.

REMARK. – Since $J$ is strictly convex on $\{x;\, J(x) < +\infty\}$ and $J(x) = 0$ if and only if $x \equiv 1/(E[1/X_1])$, Theorem 2.4 implies that $F_{a_n} \to F_{a_\infty}$ in distribution, $a_\infty$ being defined by (2.2) (see the remark preceding Proposition 1.7).

Let now be $m\colon \mathcal{F} \to \boldsymbol{R}$ the application associating to every functional $F_a$ its minimum

$$m(F_a) = \min_{u \in X} \int_0^1 a(s) u'(s)^2\, ds\,. \tag{2.9}$$

By Theorem 2.2 $m$ is continuous on $\mathcal{F}$. Setting $M_n(\omega) = m(F_{a_n(\omega)})$, $M_n$ is a r.v. and $M_n \to m(F_{a_\infty})$ a.s. (and then in distribution) where $a_\infty$ is given by (2.2). We shall now derive from Theorem 2.4 a large deviation estimate for this convergence. Indeed if we define

$$\begin{cases} l(x) \;=\; \inf\limits_{m(F)=x} \hat{J}(F)\,, & x \in \boldsymbol{R} \\ L(A) = \inf\limits_{x \in A} l(x)\,, & A \subset \boldsymbol{R}\,. \end{cases} \tag{2.10}$$

It is then easy to see that, $m$ being continuous, for every Borel subset $A$ of $\boldsymbol{R}$

$$-L(\mathring{A}) \leqslant \varliminf_{n\to\infty} \frac{1}{n} \log P\{M_n \in A\} \leqslant \varlimsup_{n\to\infty} \frac{1}{n} \log P\{M_n \in A\} \leqslant -L(\bar{A})\,. \tag{2.11}$$

Moreover the function $l$ may be explicitly calculated

LEMMA 2.8.

$$l(x) = \lambda\left(\frac{1}{x}\right).$$

PROOF. – The minimum in (2.9) equals $\left(\int_0^1 (1/a(s))\, ds\right)^{-1}$ by elementary Calculus of Variations so that from (2.10) in order to compute $l(x)$ one has to estimate the infimum of $a \to \int_0^1 \lambda(1/a(s))\, ds$ on the set of all $a$ such that $\int_0^1 (1/a(s))\, ds = 1/x$. Choosing $a \equiv x$ gives $l(x) \geqslant \lambda(1/x)$, while $l(x) \leqslant \lambda(1/x)$ follows by Jensen inequality.

REMARK. – A direct computation shows that the sequence $\{M_n\}_n$ is a subadditive process (see KINGMAN [8] for a precise definition).

(2.11) and lemma 2.8 are then an example of large deviations estimate for the ergodic subadditive convergence theorem.

**3. – Application: large deviations connected with Donsker invariance principle.**

Let $(X_n)_n$ be a sequence of i.i.d. random vectors taking values in $\mathbf{R}^d$, of law $\mu$. Let us set $S_0 = 0$, $S_n = X_1 + \ldots + X_n$ and, for every $t \in [0, 1]$

$$\eta_n(t) = \begin{cases} S_k & \text{if} \quad t = \dfrac{k}{n} \\ \text{interpolated linearly} & \text{if} \quad \dfrac{k}{n} < t < \dfrac{k+1}{n} \end{cases}$$

$\eta_n$ then defines a r.v. taking values in $\mathcal{C} = \mathcal{C}([0, 1], \mathbf{R}^d)$, the set of all continuous paths from $[0, 1]$ to $\mathbf{R}^d$, endowed with the topology of uniform convergence; $\mathcal{C}_0$ will denote the set of all paths $u \in \mathcal{C}$ such that $u(0) = 0$. $\mathcal{C}_0$ is a closed convex set of $\mathcal{C}$ and $\eta_n \in \mathcal{C}_0$ a.s. for every $n$.

Let $\hat{\mu}$ be the Laplace transform of $\mu$ and $\lambda$ the corresponding Cramer transform, namely

$$\lambda(x) = \sup_\theta \left[(\theta, x) - \log \hat{\mu}(\theta)\right]. \tag{3.1}$$

Let $\Gamma$ be the covariance matrix of $\mu$ and $Q^*$ its conjugate quadratic form, defined by

$$\tfrac{1}{2} Q^*(x) = \sup_\theta \left((\theta, x) - \tfrac{1}{2}(\Gamma\theta, \theta)\right)$$

(if $\Gamma$ is invertible then $Q^*(x) = (\Gamma^{-1}x, x)$) and define for $u \in \mathcal{C}$

$$l(u) = \begin{cases} \tfrac{1}{2}\displaystyle\int_0^1 Q^*(u'(s))\, ds & \text{if } u \text{ is absolutely continuous and } u(0) = 0 \\ +\infty & \text{otherwise} \end{cases} \tag{3.2}$$

and for every subset $A \subset \mathcal{C}$

$$L(A) = \inf_{u \in A} l(u).$$

The main theorem of this section is the following

THEOREM 3.1. – If $\hat{\mu}$ is finite in a neighborhood of 0 and $(b_n)_n$ is a sequence such that

$$\lim_{n\to\infty} \frac{b_n}{\sqrt{n}} = +\infty, \quad \lim_{n\to\infty} \frac{b_n}{n} = 0$$

then for every $A \subset \mathcal{C}_0$

$$-L(\mathring{A}) \leqslant \varliminf_{n\to\infty} \frac{n}{b_n^2} \log P\left\{\frac{1}{b_n}\eta_n \in A\right\} \leqslant \varlimsup_{n\to\infty} \frac{n}{b_n^2} \log P\left\{\frac{1}{b_n}\eta_n \in A\right\} \leqslant -L(\bar{A})$$

the operations of closure and of interior being considered in the topology of $\mathcal{C}_0$.

This theorem was first proved by A. A. MOGUL'SKII [6] (see also BOROVKOV [1]). It is however interesting to see how Theorem 3.1 may be derived by Theorems 1.1 and 1.2. In the rest of this section we will sketch this program.

By Riesz theorem the topological dual of $\mathcal{C}$ is given by the set of all $\nu = (\nu_1, \dots, \nu_d)$, each $\nu_j$ being a signed measure on $([0,1], \mathcal{B}(0,1]))$ of bounded variation. Let be

$$H_n(\nu) = \log E\left[\exp\left\langle\frac{1}{b_n}\eta_n, \nu\right\rangle\right]$$

where $\langle u, \nu\rangle = \int_0^1 u(s)\nu(ds)$, then easily, if $\Phi = \log \hat{\mu}$

$$\frac{n}{b_n^2} H_n\left(\frac{b_n^2}{n}\nu\right) = \frac{n}{b_n^2}\sum_{i=1}^n \Phi\left(\frac{b_n}{n}\left[\int_{(i-1)/n}^{i/n}\left(s - \frac{i-1}{n}\right)d\nu(s) + \int_{i/n}^1 d\nu(s)\right]\right). \tag{3.3}$$

Let us denote $\nu_1(s) = (\nu_1([s,1]), \dots, \nu^d([s,1]))$.

LEMMA 3.2. – Under the hypothesis of Theorem 3.1

*a*) $\lim\limits_n \frac{n}{b_n^2} H_n\left(\frac{b_n^2}{n}\nu\right) = \frac{1}{2}\int_0^1 (\Gamma\nu_1(s), \nu_1(s))\, ds \overset{\text{def}}{=} H(\nu)$;

*b*) $l(u) = \sup\limits_\nu (\langle u, \nu\rangle - H(\nu))$

$l$ being defined in (3.2).

Lemma 3.2 *a*) follows directly by (3.3), using the fact that at the origin

$$\Phi(x) = \tfrac{1}{2}(\Gamma x, x) + o(|x|^2)$$

whereas lemma 3.2 *b*) is proved by arguments similar to those of lemma 2.7.

Since $l$ is certainly strictly convex at every $u \in \mathcal{C}_0$ such that $l(u) < +\infty$, in order to achieve the proof of Theorem 3.1, we only need to prove (1.3).

Let $\delta = (\delta_k)_k$ be a sequence of positive real numbers decreasing to 0 and define

$$A_{\delta,k} = \left\{x \in \mathcal{C};\ w_x(\delta_k) \leqslant \frac{1}{k}\right\}$$

$$A_\delta = \bigcap_{k=1}^\infty A_{\delta,k}$$

$w_x$ being the modulus of continuity of $x$ defined by

$$w_x(\delta) = \sup_{\substack{0\leqslant t_1\leqslant t_2\leqslant 1\\ t_2-t_1\leqslant\delta}} |x(t_2) - x(t_1)|\,.$$

Then by Ascoli-Arzela's theorem $A_\delta$ is a compact subset of $\mathcal{C}$. We prove now that for every $R > 0$ there exist $\delta = (\delta_k)_k$ such that

$$\overline{\lim_{n\to\infty}}\, \frac{n}{b_n^2} \log P\left\{\frac{1}{b_n}\eta_n \notin A_\delta\right\} < -R\,.$$

We shall suppose from now on that $d = 1$. The computations for the case $d > 1$ will be only more cumbersome. From now on $\sigma^2 > 0$ will denote the variance of $X_1$, so that $\Gamma = \sigma^2$ and $Q^*(x) = (1/\sigma^2)x^2$.

Since $A_\delta^c \subset \bigcup_{k=1}^{\infty} A_{\delta,k}^c$, it will be sufficient to prove that, for a suitable sequence $\delta$,

$$P\left\{\frac{1}{b_n}\eta_n \notin A_{\delta,k}\right\} \leqslant \varepsilon_k \exp\left[-R\frac{b_n^2}{n}\right[ \tag{3.4}$$

for every $n$, where $\{\varepsilon_k\}_k$ is positive and sommable.

If

$$B_t^k = \left\{x;\ \sup_{t\leqslant s\leqslant t+\delta_k} |x(s) - x(t)| > \frac{1}{3k}\right\}$$

then by the triangle inequality

$$A_{\delta,k}^c \subset \bigcup_{i\leqslant\delta_k^{-1}} B_{i\delta_k}^k \tag{3.5}$$

and ($[t]$ denoting the integral part of $t$ from now on)

$$\begin{aligned}
&P\left\{\frac{1}{b_n}\eta_n \notin B_{i\delta_k}^k\right\} = P\left\{\frac{1}{b_n}\eta_n \notin B_0^k\right\} = \qquad (3.6)\\
&= P\left\{\sup_{m\leqslant[\delta_k n]+1} \frac{1}{b_n}|S_m| > \frac{1}{3k}\right\} \leqslant cP\left\{\frac{1}{b_n}|S_{[\delta_k n]+1}| > \frac{1}{3k}\right\} = \\
&= cP\left\{\frac{1}{[\delta_k n]+1}|S_{[\delta_k n]+1}| > \frac{b_n}{3k[\delta_k h]+1}\right\} \leqslant \\
&\leqslant c\left\{\exp\left(-([\delta_k n]+1)\lambda\left(\frac{b_n}{3k([\delta_k n])+1}\right)\right) + \exp\left(-([\delta_k n]+1)\lambda\left(-\frac{b_n}{3k([\delta_k n]+1)}\right)\right)\right\}
\end{aligned}$$

where we used a classical maximal lemma (see FELLER [4], Lemma 2, p. 192) and Cramers's results on large deviations for the sums of i.i.d. r.v. Since $\lambda(x) \sim (1/2\sigma^2)x^2$

at the origin and $b_n/n \to 0$, for $n$ large (depending on $k$)

$$\exp\left\{-([\delta_k n]+1)\,\lambda\left(\frac{b_n}{3k([\delta_k n]+1)}\right)\right\}\leqslant \leqslant \exp\left\{-\frac{b_n^2}{27\sigma^2k^2([\delta_k n]+1)}\right\}\leqslant \exp\left\{-\frac{b_n^2}{n}\frac{1}{54\sigma^2k^2\delta_k}\right\}$$

treating similarly the other term in the last member of (3.6), from (3.5), (3.6) one gets

$$P(A^c_{\delta,k})\leqslant \frac{2c}{\delta_k}\exp\left[-\frac{b_n^2}{n}\frac{1}{54\sigma^2k^2\delta_k}\right]$$

for every $n\geqslant n(k)$. Choosing $\delta_k = 1/k^3$

$$P\left\{\frac{1}{b_n}\eta_n\in A^c_{\delta,k}\right\}\leqslant 2ck^3\exp\left[-\frac{b_n}{n}\frac{k}{54\sigma^2}\right]\leqslant \frac{1}{2^k}\exp\left[-\frac{b_n^2}{n}\frac{k}{108\sigma^2}\right] \tag{3.7}$$

for every $n\geqslant n'(k)$. Now just remark that the quantity $P\{(1/b_n)\eta_n\in A^c_{\delta,k}\}$ is decreasing in $\delta_k$ and tends to 0 as $\delta_k\to 0$ for every fixed $n$. Thus (3.7) is true for every $n$, if we take $\delta_k$ small enough. This proves (3.5) and concludes the proof of theorem 3.1.

## BIBLIOGRAPHY


[1] A. A. BOROVKOV, *Boundary value problems for random walks and large deviations in function spaces*, Theory Probab. Appl., **12** (1967), pp. 635-654.

[2] G. DAL MASO - L. MODICA, *Nonlinear stochastic homogenization*, to appear.

[3] E. DE GIORGI - T. FRANZONI, *Su un tipo di convergenza variazionale*, Atti Accad. Naz. Lincei, **58** (1975), pp. 842-850.

[4] W. FELLER, *An introduction to probability theory and its applications*, vol. I, second edition J. Wiley, 1957.

[5] J. GARTNER, *On large deviations from the invariant measure*, Theory Probab. Appl., **22** (1977), pp. 24-39.

[6] A. A. MOGUL'SKII, *Large deviations for trajectories of multidimensional random walks*, Theory Probab. Appl., **21** (1976), pp. 309-323.

[7] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton Math. Series 28, Princeton University Press, 1970.

[8] J. F. C. KINGMAN, *Subadditive processes*, in Ecole d'été de Probabilités de St. Flour V-1975, Springer, 1976.

[9] S. SPAGNOLO, *Sulla convergenza delle soluzioni delle equazioni paraboliche e ellittiche*, Ann. Scuola Norm. Sup. Pisa, **21** (1967), pp. 657-699.

[10] A. D. WENTZELL - M. I. FREIDLIN, *Radom perturbations of dynamical systems*, Springer, 1984.

[11] D. A. DAWSON - J. GARTNER, *Long time fluctuations of weakly interacting diffusions*, to appear.

[12] G. FACCHINETTI - L. RUSSO, *Un caso unidimensionale di convergenza stocastica*, Boll. Un. Mat. Ital., sect. C, **2** (1983), pp. 159-170.