The Annals of Probability
1988, Vol. 16, No. 3, 978–990

ON THE UPPER BOUND FOR LARGE DEVIATIONS OF SUMS OF I.I.D. RANDOM VECTORS

BY M. SLABY

Case Western Reserve University

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random vectors with values in $\mathbb{R}^d$, $\mu = \mathscr{L}(X_1)$ and let λ be the convex conjugate of $\log \hat{\mu}$, where $\hat{\mu}$ is the Laplace transform of μ. For every $d \geq 2$, a probability measure μ and an open set A in $\mathbb{R}^d$ are constructed so that

$$\liminf_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) > -\Lambda(A),$$

where $S_n = X_1 + \cdots + X_n$ and $\Lambda(A) = \inf_{x \in A} \Lambda(x)$. It is also shown that if μ satisfies certain regularity conditions, then

$$\limsup_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) \leq -\Lambda(A)$$

holds for all Borel sets in $\mathbb{R}^d$.

1. Introduction. Throughout this paper, unless indicated otherwise, E will denote a finite-dimensional Banach space. For every nonnegative finite Borel measure μ on E, $\hat{\mu}$ will denote its Laplace transform, i.e.,

$$\hat{\mu}(\xi) = \int e^{\xi(x)}\, d\mu(x), \quad \text{for } \xi \in E^*,$$

where E^* is the space of all linear functionals on E.

We shall assume here that

(1.1) $\quad \hat{\mu}(\xi) < \infty, \quad$ for every ξ in some neighborhood of 0 in E^*.

The function $\log \hat{\mu}$: $E^* \to [-\infty, +\infty]$ is convex and lower semicontinuous (l.s.c.).

The Cramér transform λ of μ is defined as the convex conjugate of $\log \hat{\mu}$, i.e.,

$$\lambda(x) = \sup_{\xi \in E^*} [\langle \xi, x \rangle - \log \hat{\mu}(\xi)].$$

λ is then convex and l.s.c.

For every Borel set A in E denote

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

Let now $X_1, X_1, \ldots$ be a sequence of i.i.d. random vectors with values in E, $\mathscr{L}(X_1) = \mu$, and let $S_n = X_1 + X_2 + \cdots + X_n$.

Received January 1987.

AMS 1980 *subject classification*. 60F10.

Key words and phrases. Sums of i.i.d. random vectors, large deviations, upper bound for open sets.

THEOREM 1.1. *Under assumption* (1.1),

(i) $$\liminf_{n\to\infty} \frac{1}{n}\log P\left(\frac{S_n}{n} \in G\right) \geq -\Lambda(G), \text{ for every open set } G,$$

(ii) $$\limsup_{n\to\infty} \frac{1}{n}\log P\left(\frac{S_n}{n} \in F\right) \leq -\Lambda(F), \text{ for every closed set } F.$$

The one-dimensional version of the preceding is due to Cramér and Chernoff, whereas the Banach space-valued case was proved by Donsker and Varadhan in [6]. For proofs we refer to [1] or [5].

For any separable Banach space E, the upper bound, i.e., Theorem 1.1(ii), holds for finite unions of open convex sets (see [1] or [3]); a larger class of sets is considered in [5]. It is easy to show that in the one-dimensional case the upper bound holds for all open sets. One way of seeing it is to observe that the problem can be reduced to the case of open convex sets. On the other hand, Bahadur and Zabell have shown in [3] that Theorem 1.1(ii) does not hold generally for open sets in infinite-dimensional spaces.

Azencott and Ruget [2] and Bártfai [4] have given proofs of the upper bound for any open set in $\mathbb{R}^d$. It has been known for some time that both proofs are incorrect and the question of the validity of the upper bound for open sets in $\mathbb{R}^d$ remains open. (It is our understanding that this question was explicitly raised by S. Zabell in the Workshop on Large Deviations, Institute for Mathematics and Its Applications, University of Minnesota, November 1985.)

Since the lower bound holds for all open sets, the positive solution of this problem would prove the convergence of

$$\frac{1}{n}\log P\left(\frac{S_n}{n} \in A\right), \quad \text{for all open sets } A.$$

In Section 2 we construct a probability measure μ on $\mathbb{R}^2$ with bounded support and an open set A such that

$$\liminf_{n\to\infty} \frac{1}{n}\log P\left(\frac{S_n}{n} \in A\right) > -\Lambda(A).$$

In fact, we show that for every separable Banach space E of dimension at least 2, there is a probability measure μ on E and an open set A such that the preceding strict inequality holds; the measure μ has the property that there exists an affine subspace J of E of codimension 2 such that

$$\mu\left(J \cap \overline{\operatorname{conv}\operatorname{supp}\mu}\right) > 0.$$

In Section 3 we show that if μ is a probability measure on a finite-dimensional space E satisfying (1.1) and there is no affine subspace J of codimension 2, such that $\mu(J \cap \overline{\operatorname{conv}\operatorname{supp}\mu}) > 0$, then the upper bound holds for every Borel set A.

In the remainder of this section we recall some known properties of convex functions and especially of the Cramér transform λ. Although the essential definitions and results are given, we refer to [7] for more details on the subject.

Convex functions in this paper are defined on subsets of E and take real values as well as $\pm\infty$. The effective domain of a convex function f is denoted by dom f and is defined as follows:

$$\operatorname{dom} f = \{x | f(x) < \infty\}.$$

It follows immediately from the convexity of f, that dom f is a convex set. By E_f we denote the affine hull of dom f.

A convex function f is said to be proper if

(i) $f(x) < \infty$ for at least one x;
(ii) $f(x) > -\infty$ for every x.

A convex function that is not proper is improper.

We shall quote now three properties of convex functions from [7].

PROPOSITION 1.1 ([7], Theorem 10.1). *Let f be a convex function on E. $f|_A$ is continuous for every relatively open convex set A in the effective domain of f. In particular, $f|_A$ is continuous if A is the relative interior of* dom f, *i.e.,* $A = \operatorname{int}_{E_f}(\operatorname{dom} f)$.

PROPOSITION 1.2 ([7], Theorem 5.3). *If G is a convex subset of $E \times \mathbb{R}$, then*

$$\varphi(u) \overset{\text{df}}{=} \inf\{v | (u, v) \in G\}$$

is a convex function on E.

PROPOSITION 1.3 ([7], Corollary 12.1.2). *Given any proper convex function f on E^*, there exists some $b \in E$ and $\beta \in \mathbb{R}$ such that*

$$f(u) \geq \langle u, b\rangle - \beta, \quad \textit{for every } u \in E^*.$$

Given a measure μ, we denote by C the closed convex hull of the support of μ, i.e.,

$$C = \overline{\operatorname{conv} \operatorname{supp} \mu};$$

and we denote the affine hull of C by E_μ. If $E_\mu = E$ we say that μ is a full measure.

Then, under assumption (1.1), we have the following.

PROPOSITION 1.4 (See [1], [3] or [5]).

$$\overline{\operatorname{dom} \lambda} = C, \tag{1.2}$$

$$\operatorname{int}_{E_\mu}(\operatorname{dom} \lambda) = \operatorname{int}_{E_\mu}(C). \tag{1.3}$$

REMARK 1.1. By Propositions 1.1 and 1.4 $\lambda|_{\operatorname{int}_{E_\mu}(C)}$ is continuous. On the other hand, also by Proposition 1.4, $\lambda(x) = \infty$ for $x \notin C$. Thus $\lambda|_{E_\mu}$ is continuous except perhaps on $\partial_{E_\mu} C$.

REMARK 1.2. By taking $\xi = 0$, one can check that

$$\lambda(x) \geq -\log \mu(E), \quad \text{for every } x \in E.$$

On the other hand, if $\mu \neq 0$ and $b = (1/\mu(E))\int y \, d\mu(y)$, it follows from Jensen's inequality that

$$\lambda(b) = -\log \mu(E).$$

In particular, for a probability measure μ, we have $\lambda(x) \geq 0$ for every x and $\lambda(b) = 0$ if and only if $b = \int y \, d\mu(y)$.

REMARK 1.3. It follows from Remark 1.2 and the convexity of λ that λ is increasing on every ray with vertex b; that is,

$$h(t) \stackrel{\text{df}}{=} \lambda(ta + (1-t)b), \quad \text{where } a \neq b,$$

is an increasing function for $t \geq 0$.

2. A counterexample concerning the upper bound for open sets.

THEOREM 2.1. *There is a probability measure μ on $\mathbb{R}^2$ with bounded support and there is an open set $A \subset \mathbb{R}^2$ such that*

$$\liminf_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) > -\Lambda(A).$$

The proof will follow from three lemmas. We start with the construction of μ and A.

The measure μ. Let

$$F(x) = \begin{cases} 0, & \text{if} \quad x \leq 0, \\ 1/\log(1/x), & \text{if} \quad 0 < x < 1/e, \\ 1, & \text{if} \quad 1/e \leq x. \end{cases}$$

Then F is a distribution function. Let ν_F be the associated probability measure. For $x \in [0, 1/e]$, define

$$\gamma(x) = (x, x^2)$$

and

$$\nu = \nu_F \circ \gamma^{-1}.$$

Finally, we define

$$\mu = p\delta_0 + q\nu, \quad \text{where } p, q > 0, \; p + q = 1.$$

The set A. For any $\alpha > 0$, define

$$h(x) = \alpha^{-1} \log\log(1/x)$$

and

$$f(x) = x^2 h(x).$$

Since $\lim_{x\to 0+} h(x) = +\infty$, we have that $x^2 < f(x)$ for small x.

For every $t > 0$, define

$$A_t = \{(x, y) | 0 < y < f(x), 0 < x < t\}.$$

A_t is open for every $t > 0$.

The set A will be chosen eventually to be A_t for small enough t.

LEMMA 2.1. *For every t positive,*

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A_t\right) \geq \log p - \alpha.$$

PROOF. Since

$$\begin{aligned} P\left(\frac{S_n}{n} \in A_t\right) &= (p\delta_0 + q\nu)^{*n}(nA_t) \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \nu^{*(n-k)}(nA_t) \geq np^{n-1} q\nu(nA_t), \end{aligned}$$

we have

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A_t\right) \geq \log p + \liminf_{n\to\infty} \frac{1}{n} \log \nu(nA_t).$$

Now

$$\begin{aligned} \nu(nA_t) &= \nu\left(\left\{(x, y) | 0 < \frac{y}{n} < f\left(\frac{x}{n}\right), 0 < \frac{x}{n} < t\right\}\right) \\ &= \nu_F\left(\left\{x \in [0, 1/e] | n < h\left(\frac{x}{n}\right), x < tn\right\}\right) \\ &= \nu_F\left(\left\{x | x < ne^{-e^{\alpha n}}\right\}\right), \end{aligned}$$

for large enough n. Thus

$$\nu(nA_t) = F(ne^{-e^{\alpha n}}) = \frac{1}{e^{\alpha n} - \log n} \geq e^{-\alpha n}$$

and so

$$\liminf_{n\to\infty} \frac{1}{n} \log \nu(nA_t) \geq -\alpha. \qquad \square$$

LEMMA 2.2. *Let λ be the Cramér transform of μ. Then*

$$\liminf_{x\to 0+} \lambda(x, f(x)) \geq -\log p + 2\alpha.$$

PROOF. From the definition,

$$\lambda(x, f(x)) = \sup_{u, v} [ux + vf(x) - \log(p + q\hat{\nu}(u, v))].$$

It is hard to solve this extremal problem exactly, since $\hat{\nu}(u,v)$ cannot be computed explicitly. In order to get a suitable lower bound, we try to choose values of u and v (depending on x) which, roughly speaking, make $ux + vf(x)$ constant and $\hat{\nu}(u,v)$ negligible for small x. Let $u = u(x) = 2\beta/x$ and $v = v(x) = -(\beta/f(x))$. Then $ux + vf(x) = \beta$, and $\lambda(x, f(x)) \geq \beta - \log(p + q\hat{\nu}(u,v))$ for any x. Now to conclude the proof it is enough to show for every $\beta < 2\alpha$

$$\lim_{x \to 0+} \hat{\nu}(u(x), v(x)) = 0,$$

where

$$\hat{\nu}(u(x), v(x)) = \int_0^{1/e} e^{u(x)t + v(x)t^2}\, dF(t).$$

Let $\varphi(t) = u(x)t + v(x)t^2$. Then φ is increasing on $[0, -u(x)/2v(x)]$ and decreasing on $[-u(x)/2v(x), 1/e]$;

$$\varphi_{\max} = \varphi\left(-\frac{u(x)}{2v(x)}\right) = \varphi(xh(x)) = \beta h(x)$$

and

$$\varphi\left(-\frac{u(x)}{v(x)}\right) = \varphi(2xh(x)) = 0.$$

For small x, $0 < x < 2xh(x) < 1/e$.

Let

$$I_1(x) = \int_0^x e^{\varphi(t)}\, dF(t),$$

$$I_2(x) = \int_x^{2xh(x)} e^{\varphi(t)}\, dF(t),$$

$$I_3(x) = \int_{2xh(x)}^{1/e} e^{\varphi(t)}\, dF(t).$$

Then $\hat{\nu}(u(x), v(x)) = I_1(x) + I_2(x) + I_3(x)$.

$$I_1(x) \leq e^{u(x)x} F(x) = e^{2\beta} F(x),$$

hence $\lim_{x \to 0+} I_1(x) = 0$.

$$I_3(x) = \int_{2xh(x)}^{1/e} \exp\left(-\frac{\beta t}{f(x)}(t - 2xh(x))\right) dF(t),$$

hence by Lebesgue's dominated convergence theorem $\lim_{x \to 0+} I_3(x) = 0$.

$$I_2(x) \leq e^{\varphi_{\max}}[F(2xh(x)) - F(x)],$$

but

$$F(2xh(x)) - F(x) = \frac{1}{\log 1/2xh(x)} - \frac{1}{\log 1/x} = \frac{\log 2h(x)}{(\log 1/x)^2\left[1 - \dfrac{\log 2h(x)}{\log 1/x}\right]}.$$

Thus

$$e^{\beta h(x)}[F(2xh(x)) - F(x)] = \frac{\log 2h(x)}{(\log 1/x)^{2-\beta/\alpha}} \frac{1}{1 - \dfrac{\log 2h(x)}{\log 1/x}}.$$

Let $w = \log 1/x$. Then, for every $k > 0$ and $s > 0$,

$$\lim_{w\to\infty} \frac{\log(k\log w)}{w^s} = \lim_{w\to\infty} \frac{\log(k\log w)}{k\log w} \lim_{w\to\infty} \frac{k\log w}{w^s} = 0.$$

Hence

$$\lim_{x\to 0+} \frac{\log 2h(x)}{\log 1/x} = \log_{x\to\infty +} \frac{\log 2h(x)}{(\log 1/x)^{2-\beta/\alpha}} = 0, \quad \text{for } \beta < 2\alpha,$$

and therefore also $\lim_{x\to 0+} I_2(x) = 0$. Summarizing, $\lim_{x\to 0+}\hat{\nu}(u(x), v(x)) = 0$ and so

$$\liminf_{x\to 0+} \lambda(x, f(x)) \geq -\log p + \beta, \quad \text{for every } \beta < 2\alpha. \qquad \square$$

LEMMA 2.3. *For small enough t,*

$$\Lambda(A_t) = \inf_{x<t} \lambda(x, f(x)).$$

PROOF. Let $b = (b_1, b_2)$ be the barycenter of μ, i.e., $b = \int x\, d\mu(x)$. Then $b_1 > 0_1$, $b_2 > 0$ and $\lambda(b) = 0$, by Remark 1.2.

$$\lambda(0) = \sup[-\log(p + q\hat{\nu}(u, v))] \leq -\log p.$$

Thus, by Remark 1.3, for every $0 < x < b_1$,

$$\lambda\left(x, \frac{b_2}{b_1}x\right) \leq \lambda(0) \leq -\log p.$$

By Lemma 2.2, for small enough x,

$$\lambda(x, f(x)) \geq -\log p + \alpha.$$

For t small enough, $f(x) < (b_2/b_1)x$ and $\lambda(x, f(x)) > \lambda(x, (b_2/b_1)x)$ for every $0 < x < t$. Now suppose there is $y < f(x)$ such that $\lambda(x, f(x)) > \lambda(x, y)$, where $0 < x < t$. Then, for every $0 < s < 1$,

$$\lambda(x, f(x)) > (1-s)\lambda\left(x, \frac{b_2}{b_1}x\right) + s\lambda(x, y),$$

which contradicts the convexity of λ because

$$y < f(x) < \frac{b_2}{b_1}x.$$

This proves that $\Lambda(A_t) \geq \inf_{x<t}\lambda(x, f(x))$.

Since $A_t \cap \{(x, y) | y > x^2\} \subset \operatorname{int} C$ for small enough t, by Remark 1.1,

$$\Lambda(A_t) = \Lambda(A_t \cup \{(x, f(x)) | 0 < x < t\}) \leq \inf_{x<t} \lambda(x, f(x)). \qquad \square$$

PROOF OF THEOREM 2.1. It follows from Lemma 2.2 that

$$\lim_{t \to 0+} \inf_{0<x<t} \lambda(x, f(x)) \geq -\log p + 2\alpha.$$

By Lemma 2.3, then

$$\lim_{t \to 0+} \Lambda(A_t) \geq -\log p + 2\alpha.$$

Thus, for small enough t, $\Lambda(A_t) \geq -\log p + 3/2\alpha$. On the other hand, by Lemma 2.1,

$$\liminf_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A_t\right) \geq \log p - \alpha > \log p - 3/2\alpha \geq -\Lambda(A_t). \qquad \square$$

REMARK 2.1. The curve $\gamma(x) = (x, x^2)$ is not essential in the construction of Theorem 2.1. One could use other curves $\gamma(x) = (x, g(x))$, where $g(x)$ is convex, increasing and $g(0) = 0$, and then find an appropriate function f so that $\lim_{n \to \infty}(\nu(nA_t))^{1/n}$ exists and is positive. The essential part of the construction is the distribution function F; the key point in its use is the proof of $\lim_{x \to 0+} I_2(x) = 0$ in Lemma 2.2.

REMARK 2.2. For every separable Banach space E with $\dim E > 1$, there are a probability measure μ on E and an open set A such that

$$\liminf_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) > -\Lambda(A);$$

μ is such that $\mu(J \cap \partial C) > 0$ for some affine subspace J of codimension 2.

PROOF OF REMARK 2.2. If $\dim E = 2$, then we have μ and A from Theorem 2.1.

If $\dim E > 2$, then we may assume without loss of generality that $E = \mathbb{R}^2 \oplus F$, where F is a Banach space (any subspace of dimension 2 has a closed complement). Let μ_1 and A_1 be as in Theorem 2.1 and let μ_2 be a probability measure on F such that $\int \|x\| \, d\mu_2(x) < \infty$. Take $\mu = \mu_1 \times \mu_2$ and $A = A_1 \times A_2$, where A_2 is an open set in F such that $\operatorname{conv} \operatorname{supp} \mu_2 \subset A_2$. Then $\hat{\mu}(u, v) = \hat{\mu}_1(u) \cdot \hat{\mu}_2(v)$, where $u \in \mathbb{R}^2$, $v \in F^*$. For any $x \in \mathbb{R}^2$ and $y \in F$,

$$\begin{aligned} \lambda_\mu(x, y) &= \sup_{u, v} \left[\langle u, x\rangle + \langle v, y\rangle - \log(\hat{\mu}_1(u) \cdot \hat{\mu}_2(u))\right] \\ &= \sup_{u, v} \left[\langle u, x\rangle - \log \hat{\mu}_1(u) + \langle v, y\rangle - \log \hat{\mu}_2(v)\right] \\ &= \sup_{u} \left[\langle u, x\rangle - \log \hat{\mu}_1(u)\right] + \sup_{v} \left[\langle v, y\rangle - \log \hat{\mu}_2(v)\right]. \end{aligned}$$

Thus

$$\lambda_\mu(x, y) = \lambda_{\mu_1}(x) + \lambda_{\mu_2}(y).$$

Let $X_n = (Y_n, Z_n)$, $\mathscr{L}(Y_n) = \mu_1$, $\mathscr{L}(Z_n) = \mu_2$ so that $\mathscr{L}(X_n) = \mu$. Let $S_n =$

$\sum_{i=1}^n X_i$. Then

$$P\left(\frac{S_n}{n} \in A\right) = P\left(\left(\frac{\sum_{i=1}^n Y_i}{n}, \frac{\sum_{i=1}^n Z_i}{n}\right) \in A_1 \times A_2\right)$$
$$= P\left(\frac{\sum_{i=1}^n Y_i}{n} \in A_1\right) P\left(\frac{\sum_{i=1}^n Z_i}{n} \in A_2\right)$$
$$= P\left(\frac{\sum_{i=1}^n Y_i}{n} \in A_1\right).$$

Thus

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) = \liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{\sum_{i=1}^n Y_i}{n} \in A_1\right) > -\Lambda_{\mu_1}(A_1).$$

But

$$\Lambda(A) = \inf_{(x, y) \in A} \lambda(x, y) = \inf_{x \in A_1} \inf_{y \in A_2} \left[\lambda_{\mu_1}(x) + \lambda_{\mu_2}(y)\right]$$
$$= \Lambda_{\mu_1}(A_1) + \Lambda_{\mu_2}(A_2).$$

By Remark 1.2, $\Lambda_{\mu_2}(A_2) = 0$ because A_2 contains the barycenter of μ_2. Thus $\Lambda(A) = \Lambda_{\mu_1}(A)$ and so

$$\liminf_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) > -\Lambda(A). \qquad \square$$

3. Upper bound for probability measures satisfying certain regularity conditions. As it was pointed out in Section 1, there are known upper bound results for a restricted class of sets. The following theorem gives the upper bound for all Borel sets but with a restriction on the measure.

THEOREM 3.1. *Let* $\mu = \mathscr{L}(X_1)$ *satisfy the following conditions*:

(i) $\hat{\mu}(\xi) < \infty$ *for all* ξ *in a neighborhood of* 0 *in* E^*.
(ii) *There is no affine subspace* J *of* E_μ *of codimension* 2 *such that*

$$\mu\left(J \cap \partial_{E_\mu} C\right) > 0.$$

Then, for every Borel set A *in* E,

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) \le -\Lambda(A).$$

Assumption (ii) is not necessary; however, in view of Remark 2.2, it is not superfluous and provides what appears to be the simplest hypothesis, which rules out the cases described in the remark. The proof will follow from the upper bound for closed sets [Theorem 1.1(ii)] and three lemmas.

REMARK 3.1. (a) Let $a \in E$ and T_a: $E \to E$ be defined by $T_a(x) = x + a$. If $\mu_a = \mu \circ T_a^{-1}$, $C_a = \overline{\operatorname{conv} \operatorname{supp} \mu_a}$ and λ_a is the Cramér transform of μ_a, then, for every x in E,

$$\lambda_a(x+a) = \lambda(x)$$

and

$$C_a = C + a.$$

(b) Let f be a linear 1–1 function from E to another finite-dimensional space F. If $\mu_f = \mu \circ f^{-1}$, $C_f = \overline{\text{conv}}\,\text{supp}\,\mu_f$ and λ_f is the Cramér transform of μ_f, then, for every $x \in E$,

$$\lambda_f(f(x)) = \lambda(x)$$

and

$$C_f = f(C).$$

LEMMA 3.1. *Let μ be a full probability measure on E such that $\hat{\mu}$ is finite in a neighborhood of 0 and there is no affine subspace J of codimension 2 such that $\mu(J \cap \partial C) > 0$. Then $\lambda|_H$ is continuous for every supporting hyperplane H of C.*

PROOF. Let H be a supporting hyperplane of C and let $a \in H$. By the l.s.c. of λ,

$$\lambda(a) \le \liminf_{x \to a} \lambda(x) \le \liminf_{x \to a,\, x \in H} \lambda|_H(x).$$

Thus if a is a point of discontinuity of $\lambda|_H$, then $\lambda(a) < \infty$.

Now

$$H = \text{int}_H(\text{dom}\,\lambda \cap H) \cup (H \setminus \text{dom}\,\lambda) \cup \partial_H(\text{dom}\,\lambda \cap H).$$

By Proposition 1.1, $\lambda|_H$ is continuous in $\text{int}_H(\text{dom}\,\lambda \cap H)$. For $a \in H \setminus \text{dom}\,\lambda$, $\lambda(a) = \infty$, hence $\lambda|_H$ is also trivially continuous in $H \setminus \text{dom}\,\lambda$. It is enough then to show that

$$\lambda(a) = \infty, \quad \text{for every } a \in \partial_H(\text{dom}\,\lambda \cap H).$$

Let then $a \in \partial_H(\text{dom}\,\lambda \cap H)$ and let J be a supporting hyperplane of $\text{dom}\,\lambda \cap H$ at a in H. By Remark 3.1, we can assume without loss of generality that $E = X \oplus \mathbb{R}^2$, where X is a finite-dimensional Banach space such that

$$\dim X = \dim E - 2.$$

We can also assume that $a = 0 = (0,0,0)$, $H = X \times \mathbb{R} \times \{0\}$, $J = X \times \{(0,0)\}$, $C \subseteq \{(x,y,z) | x \in X,\ y \in \mathbb{R},\ z \ge 0\}$ and $H \cap \text{dom}\,\lambda \subset \{(x,y,z) | x \in X,\ y \ge 0,\ z = 0\}$. Then $E^* = X^* \oplus \mathbb{R}^2$. For any $x \in X$, let $h = (x, -1, 0)$. Then $h \in H \setminus \text{dom}\,\lambda$; hence $\lambda(h) = \infty$ and so

$$\lambda(h) = \sup_{(u,v,w) \in \text{dom}\,\hat{\mu}} [\langle u, x\rangle - v - \log \hat{\mu}(u,v,w)] = \infty.$$

Suppose now that $\lambda(a) = \lambda(0) < \infty$, i.e.,

$$\lambda(0) = \sup_{u,v,w} [-\log \hat{\mu}(u,v,w)] < \infty.$$

Then, for every $x \in X$, there is a sequence $\xi_n = (u_n, v_n, w_n)$ such that

$$(*) \qquad (\xi_n) \subset \text{dom}\,\hat{\mu} \quad \text{and} \quad \lim_{n \to \infty} (\langle u_n, x\rangle - v_n) = +\infty.$$

Let

$$G = \{(u, v) \in X^* \times \mathbb{R} | \hat{\mu}(u, v, w) < \infty \text{ for some } w\}.$$

Then: (1) G is convex. (2) There is a u_0 such that $(u_0, v) \in G$ for arbitrarily large negative v.

(1) follows from the convexity of $\hat{\mu}$.

To prove (2) we define

$$\varphi(u) = \inf\{v | (u, v) \in G\}.$$

By Proposition 1.2, φ is a convex function. Now it is enough to show that there is u_0 such that $\varphi(u_0) = -\infty$. Since G contains a neighborhood of 0, $\varphi(u) < +\infty$ in a neighborhood of 0. Therefore, it is enough to show that φ is not a proper convex function.

Suppose that φ is a proper convex function. By Proposition 1.3, there is $x \in X$ and $\beta \in \mathbb{R}$ such that, for every $u \in X^*$, $\varphi(u) \geq \langle u, x \rangle - \beta$. Hence, for every $(u, v) \in G$,

$$\langle u, x \rangle - v \leq \langle u, x \rangle - \varphi(u) \leq \beta,$$

which contradicts (∗).

Therefore, φ must be improper and so (2) is proved. Now

$$\begin{aligned} e^{-\lambda(0)} &= \inf_{u, v, w} \hat{\mu}(u, v, w) \\ &= \inf_{u, v, w} \left[\int_J e^{\langle u, x \rangle} \, d\mu(x, y, z) + \int_{H \setminus J} e^{\langle u, x \rangle + vy} \, d\mu(x, y, z) \right. \\ &\qquad \left. + \int_{E \setminus H} e^{\langle u, x \rangle + vy + wz} \, d\mu(x, y, z) \right]. \end{aligned}$$

Take $u = u_0$. Then $(u_0, v) \in G$ for arbitrarily large negative v and

$$\lim_{v \to -\infty} \int_{H \setminus J} e^{\langle u_0, x \rangle + vy} \, d\mu(x, y, z) = 0,$$

by Lebesgue's dominated convergence theorem.

Let $\varepsilon > 0$ and $v_0 \in \mathbb{R}$ be such that $(u_0, v_0) \in G$ and

$$\int_{H \setminus J} e^{\langle u_0, x \rangle + v_0 y} \, d\mu(x, y, z) < \varepsilon.$$

By the definition of G, there is $w_0 \in \mathbb{R}$ such that $\hat{\mu}(u_0, v_0, w_0) < \infty$. Again, by Lebesgue's dominated convergence theorem,

$$\lim_{w \to -\infty} \int_{E \setminus H} e^{\langle u_0, x \rangle + v_0 y + wz} \, d\mu(x, y, z) = 0.$$

Finally, by the assumption,

$$\mu(J) = \mu(J \cap C) = \mu(J \cap H \cap C) = \mu(J \cap \partial C) = 0,$$

hence

$$\int_J e^{ux} \, d\mu(x, y, z) = 0.$$

Thus, for every $\varepsilon > 0$, $e^{-\lambda(0)} \leq \varepsilon$, and, therefore,

$$\lambda(a) = \lambda(0) = \infty.$$

This contradicts our initial assumption that $\lambda(a) < \infty$. Thus $\lambda(a) = \infty$ for every $a \in \partial_H(\operatorname{dom} \lambda \cap H)$, as claimed. □

LEMMA 3.2. *If there is no affine subspace J of codimension 2 in E_μ such that $\mu(J \cap \partial_{E_\mu} C) > 0$, then $\lambda|_C$ is continuous.*

PROOF. By Remark 3.1, we can assume that $E = E_\mu$. By Proposition 1.1, $\lambda|_C$ is continuous on $\operatorname{Int} C$. Let $a \in \partial C$ and let H be a supporting hyperplane of C at a. Let $b = \int x \, d\mu$. Then $b \notin H$. Let P be the stereographic projection from the barycenter b onto H. P is defined on the open halfspace L such that $\partial L = H + b - a$ and $H \subset L$.

For every $x \in L$, $P(x) = h$ if $h \in H$ and $h - b \in \{t(x-b) | t \in \mathbb{R}\}$. P is continuous on L. For every $x \in C \cap L$, $Px = t(x - b) + b$ with $t \geq 1$. By Remark 1.3,

$$\lambda(P(x)) \geq \lambda(x), \quad \text{for every } x \in C \cap L.$$

Let $(x_n) \subset C$ be such that $\lim_{n \to \infty} x_n = a$. We can assume that $(x_n) \subset C \cap L$. Then

$$\limsup_{n \to \infty} \lambda(x_n) \leq \lim_{n \to \infty} \lambda(P(x_n)) = \lambda(a),$$

by the continuity of P and Lemma 3.1.

Since also $\lambda(a) \leq \liminf_{n \to \infty} \lambda(x_n)$ by the l.s.c. of λ, it follows that

$$\lim_{n \to \infty} \lambda(x_n) = \lambda(a). \qquad \square$$

COROLLARY 3.1. *Under the assumption of Lemma 3.3, for every $A \subset C$,*

$$\Lambda(A) = \Lambda(\bar{A}).$$

PROOF. Let $a \in \bar{A}$ and let (x_n) be a sequence in A such that $\lim_{n \to \infty} x_n = a$. By the continuity of $\lambda|_C$, $\lim_{n \to \infty} \lambda(x_n) = \lambda(a)$ and hence

$$\lambda(a) \geq \inf_{x \in A} \lambda(x) = \Lambda(A). \qquad \square$$

PROOF OF THEOREM 3.1. Let A be any Borel set in E. Then $A = (A \cap C) \cup (A \setminus C)$.

$$\limsup_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) = \limsup_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A \cap C\right) \leq -\Lambda(\overline{A \cap C}),$$

by Theorem 1.1(ii).

By Corollary 3.1, $\Lambda(\overline{A \cap C}) = \Lambda(A \cap C) = \Lambda(A)$. Therefore

$$\limsup_{n \to \infty} \frac{1}{n} \log P\left(\frac{S_n}{n} \in A\right) \leq -\Lambda(A). \qquad \square$$

COROLLARY 3.2. *If $X_1, X_2, \dots$ is a sequence of i.i.d. random vectors in $\mathbb{R}^2$ and $\mu = \mathscr{L}(X_1)$ has no atoms on the boundary of $C = \overline{\text{conv supp}\,\mu}$, then*

$$\limsup_{n\to\infty} \frac{1}{n}\log P\left(\frac{S_n}{n} \in A\right) \le -\Lambda(A), \quad \textit{for any Borel set } A.$$

Acknowledgment. The results in this paper form a part of the author's Ph.D. dissertation at Case Western Reserve University in Cleveland, under the supervision of Professor A. de Acosta. The author would like to thank him for suggesting the problem and many valuable comments as well as constructive criticism during preparation of this paper.

REFERENCES

[1] AZENCOTT, R. (1980). Grandes deviations et applications. *Ecole d'Été de Probabilités de Saint-Flour VIII, 1978. Lecture Notes in Math.* **774** 1–176. Springer, Berlin.

[2] AZENCOTT, R. and RUGET, G. (1977). Mélanges d'équations différentielles et grands écarts a la loi de grands nombres. *Z. Wahrsch. verw. Gebiete* **38** 1–54.

[3] BAHADUR, R. R. and ZABELL, S. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probab.* **7** 587–621.

[4] BÁRTFAI, P. (1978). Large deviations of the sample mean in Euclidean spaces. Mimeograph Series 78-13, Dept. Statistics, Purdue Univ.

[5] DE ACOSTA, A. (1985). On large deviations of sums of independent random vectors. *Lecture Notes in Math.* **1153** 1–14. Springer, Berlin.

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

[7] ROCKAFELLAR, R. T. (1970). *Convex Analysis.* Princeton Univ. Press, Princeton, N.J.

DEPARTMENT OF MATHEMATICS AND STATISTICS
CASE WESTERN RESERVE UNIVERSITY
CLEVELAND, OHIO 44106