The Annals of Probability
1978, Vol. 6, No. 5, 751–759

CONVERGENCE RATES OF LARGE DEVIATION PROBABILITIES IN THE MULTIDIMENSIONAL CASE[1]

BY JOSEF STEINEBACH

University of Düsseldorf

Let $\{W_n\}_{n=1,2,\cdots}$ denote a sequence of k-dimensional random vectors on a probability space $(\Omega, \mathscr{A}, P)$. Using moment-generating function techniques sufficient conditions are given for the existence of limits $\rho(A) = \lim_{n\to\infty} [P(W_n \notin k_n A)]^{1/k_n}$ for certain subsets $A \subset R^k$, where $\{k_n\}_{n=1,2,\cdots}$ is a divergent sequence of positive real numbers. The results are multivariate analogs of well-known large deviation theorems on the real line.

1. Introduction. In 1975 G. L. Sievers stated some large deviation results for a sequence $\{P_n\}_{n=1,2,\cdots}$ of probability measures on (R^k, B^k), where R^k is k-dimensional Euclidean space and B^k the system of Borel subsets. Assuming that the limit $e(A) = \lim_{n\to\infty} n^{-1} \log P_n(A)$, $A \in B^k$, exists, Sievers developed methods for determining $e(A)$ from $e(B)$ for "simpler" sets B of the form $B = \{x = (x_1, \cdots, x_k)^T \in R^k : x_1 * a_1, \cdots, x_k * a_k\}$, where the $*$'s are either $\geqq$ or $\leqq$, and $a = (a_1, \cdots, a_k)^T$ is a fixed vector in R^k. The existence of the limits $e(B)$ was also assumed.

In this paper two theorems are proven which give sufficient conditions for the existence of limits such as $e(A)$ or $e(B)$ described above. These conditions correspond to certain properties of the moment-generating functions of the underlying probability measures and are multivariate analogs of well-known assumptions for large deviation theorems on the real line (cf. Sievers (1969), Plachky and Steinebach (1975)).

Theorem 3.1 is concerned with large deviation probabilities for cones in R^k. Included is the case involving Sievers' simple sets B. Theorem 3.2 deals with the existence of the limits $e(A)$ for the complements of certain bounded subsets of R^k.

Both theorems can be applied directly to weighted sums of independent, identically distributed (i.i.d.) random vectors with finite moment-generating functions. For details see [12], Chapter 2.2, where some k-dimensional analogs of the large deviation results of Book (1973, 1975) concerning weighted sums of i.i.d. random variables on the real line are derived.

2. Notations. Certain facts about convex functions. Let $a = (a_1, \cdots, a_k)^T$, $b = (b_1, \cdots, b_k)^T$ be vectors in R^k. Then we write $a \leqq b$ $(a \lhd b)$ if $a_i \leqq b_i$

Received May 17, 1976; revised October 4, 1977.

[1] This research is based on part of the author's doctoral thesis written under Professor Doctor D. Plachky and Professor Doctor H. Klinger at the University of Düsseldorf.

AMS 1970 *subject classification*. Primary 60F10.

Key words and phrases. Large deviations, convergence rates, moment-generating functions.

($a_i < b_i$) for all $i = 1, \cdots, k$, and the set $[a, b] = \{x \in R^k : a \leqq x \leqq b\}$ is referred to as *a k-dimensional closed interval*. Corresponding definitions hold for $(a, b) = \{x \in R^k : a \lhd x \lhd b\}$ (*a k-dimensional open interval*), $[a, b)$, and $(a, b]$ (*a k-dimensional half-open interval*). The *inner product* of $a, b \in R^k$ is denoted as $\langle a, b\rangle$, i.e., $\langle a, b\rangle = \sum_{i=1}^k a_i b_i$. Furthermore, $|a| = \langle a, a\rangle^{\frac{1}{2}}$.

The *closure* (*interior*, *boundary*, *complement*) of $A \subset R^k$ is denoted as $\bar{A}$ (A^0, ∂A, $\complement A$), and, for fixed $\lambda \in R^1$, $a \in R^k$, let $\lambda(a + A) = \{\lambda(a + x) : x \in A\}$. For a positive real number ε the set $U_\varepsilon(x_0) = \{x \in R^k : |x - x_0| < \varepsilon\}$ is called an *ε-neighborhood of x_0*. A subset $C \subset R^k$ is called a *cone* if it is closed under positive scalar multiplication, i.e., $\lambda x \in C$ when $x \in C$ and $\lambda > 0$. The origin itself may or may not be included here. The *polar cone* C^+ is defined as the set $C^+ = \{x \in R^k : \langle x, y\rangle \geqq 0$ for all $y \in C\}$. For the properties of C^+ consult, e.g., Karlin (1962), pages 397–406. A subset $D \subset R^k$ is called *convex* if $\lambda x + (1 - \lambda)y \in D$ when $x, y \in D$, $0 < \lambda < 1$.

Let f be a real-valued function on a subset D of R^k. Then $(\partial f/\partial t_i)(t_0)$, $i = 1, \cdots, k$, $t = (t_1, \cdots, t_k)^T$, $t_0 = (t_{10}, \cdots, t_{k0})^T$, will denote the *partial derivatives* of f at t_0, $f'(t_0)$ the *gradient* $((\partial f/\partial t_1)(t_0), \cdots, (\partial f/\partial t_k)(t_0))^T$, and $f''(t_0)$ the *Hessian matrix* of second partial derivatives $(\partial^2 f/\partial t_i \partial t_j)(t_0)$. A real-valued function f on a convex subset $D \subset R^k$ is called *convex* if $f(\lambda x + (1 - \lambda)y) \leqq \lambda f(x) + (1 - \lambda)f(y)$ for all $x, y \in D$ and λ in the real open interval $(0, 1)$. Moreover, f is called *strictly convex* if strict inequality holds.

In proving the main results of this paper we shall make frequent use of certain standard facts about convex functions and their conjugates which we collate below.

LEMMA 2.1. *Let f be a convex function on an open convex set $D_0 \subset R^k$. Then*

$$f \quad \textit{is continuous on} \quad D_0\,; \tag{2.1}$$

$$f \quad \textit{is (Fréchet) differentiable almost everywhere in} \quad D_0\,, \tag{2.2}$$

i.e., there is a linear transformation $T: R^k \to R$ such that for sufficiently small $|h|$ $f(t_0 + h) = f(t_0) + T(h) + |h|\varepsilon(t_0, h)$. Here $\varepsilon(t_0, h) \in R$ goes to zero as $|h| \to 0$, and $t_0 \in D \subset D_0$, where $D_0 \backslash D$ has Lebesgue measure zero. Moreover, if t is a real variable, $D_0 \backslash D$ is a countable set. All partial derivatives $(\partial f/\partial t_i)(t_0)$, $i = 1, \cdots, k$, exist in D, and $T(h) = \langle f'(t_0), h\rangle$.

$$f' \quad \textit{is continuous on} \quad D\,; \tag{2.3}$$

$$f' \quad \textit{is isotone on} \quad D\,, \quad \text{i.e.,} \quad \langle f'(t) - f'(s), t - s\rangle \geqq 0 \quad \textit{for all} \quad t \neq s \in D\,, \tag{2.4}$$

and strict inequality holds if f is strictly convex.

Let f be a real-valued function having continuous second partial derivatives on an open convex set $D_0 \subset R^k$. It holds that

$$f \quad \textit{is convex in} \quad D_0 \tag{2.5}$$

iff *the Hessian matrix* $f''(t)$ *is nonnegative definite for each* $t \in D_0$. *Moreover, if the Hessian matrix is positive definite on* D_0, *then* f *is strictly convex, and the set* $f'(D_0) = \{f'(t) : t \in D_0\}$ *is open.*

Let $\{f_n\}_{n=1,2,\cdots}$ *be a sequence of (Fréchet) differentiable convex functions on an open convex set* $D_0 \subset R^k$ *such that* $\lim_{n\to\infty} f_n(t) = f(t)$ *exists for each* $t \in D_0$. *It is apparent that* f *is convex in* D_0. *Moreover, the assertion*

$$\lim_{n\to\infty} f_n'(t) = f'(t) \tag{2.6}$$

holds almost everywhere in D_0.

Let $f^* : D_0^* \to R$ *denote the conjugate function of* f *defined as* $f^*(t^*) = \sup_{t\in D_0} [\langle t, t^*\rangle - f(t)]$ *with domain* $D_0^* = \{t^* \in R^k : f^*(t^*) < \infty\}$. *Then*

$$f^* \quad \text{is convex on} \quad D_0^* . \tag{2.7}$$

PROOF. See, e.g., Rockafellar (1970) and Roberts and Varberg (1973), though the proofs require slight modification to establish the results (2.4) and (2.6). To obtain (2.6), for instance, note that for $t = (t_1, \cdots, t_k)^T \in D$ and $t_2, \cdots, t_k$ being fixed the convexity of the functions $f_n(\cdot, t_2, \cdots, t_k)$ yields the inequalities

$$\frac{f_n(t_1 - h_1, t_2, \cdots, t_k) - f_n(t_1, \cdots, t_k)}{-h_1} \leqq \frac{\partial}{\partial t_1} f_n(t_1, \cdots, t_k)$$
$$\leqq \frac{f_n(t_1 + h_1, t_2, \cdots, t_k) - f_n(t_1, \cdots, t_k)}{h_1},$$

where $h_1 > 0$ can be chosen arbitrarily small. Letting first n tend to infinity and then h_1 tend to zero, we have by the differentiability of $f(\cdot, t_2, \cdots, t_k)$ at t_1

$$\lim_{n\to\infty} \frac{\partial}{\partial t_1} f_n(t_1, \cdots, t_k) = \frac{\partial}{\partial t_1} f(t_1, \cdots, t_k) .$$

Same arguments for the other components, too, imply the assertion (2.6).

3. Results. As already mentioned, the first theorem states a certain large deviation property for cones C in the Euclidean space R^k.

THEOREM 3.1. *Let* $\{W_n\}_{n=1,2,\cdots}$ *be a sequence of k-dimensional random variables on a probability space* $(\Omega, \mathscr{A}, P)$, *and* D_0 *an open, convex subset of* R^k. *Suppose further that*

(i) $\varphi_n(t) = \int \exp\langle t, W_n\rangle\, dP < \infty$ *for all* $t \in D_0$;

(ii) $\lim_{n\to\infty} k_n^{-1}\psi_n(t) = c(t) \in R$ *for all* $t \in D_0$; *where* $\psi_n(t) = \log \varphi_n(t)$ *and* $\{k_n\}_{n=1,2,\cdots}$ *is a divergent sequence of positive real numbers, remembering here that* $\varphi_n(t)$ *is positive*;

(iii) c *is strictly convex in* D_0.

Now let D *be the subset of* D_0, *where* c' *exists,* C *a cone in* R^k *with* $C^0 \neq \emptyset$, *and* $\{a_n\}_{n=1,2,\cdots} \subset R^k$ *a convergent sequence with*

(iv) $\lim_{n\to\infty} a_n = a \in \{c'(t) : t \in D \cap C^+\}^0$.

Then it follows that

$$\lim_{n\to\infty}[P(W_n \in k_n(a_n + C))]^{1/k_n} = \inf_{t\in D_0}\exp[c(t) - \langle t, a\rangle] \equiv \rho(a)\,, \tag{3.1}$$

and the infimum equals $\exp[c(h) - \langle h, a\rangle]$, *where* h *is the unique solution of* $a = c'(h)$.

The second theorem deals with another sort of multivariate probabilities of large deviations, which appears, for instance, when studying convergence rates in the multidimensional version of the law of large numbers.

THEOREM 3.2. *Let* $\{W_n\}_{n=1,2,\cdots}$ *be a sequence of k-dimensional random variables on a probability space* $(\Omega, \mathscr{A}, P)$ *with* $EW_n = 0$, $n = 1, 2, \cdots$. *Let* $D_0 \subset R^k$ *be an open, bounded convex set including the origin. Suppose also that*

(i′) $\varphi_n(t) = \int \exp\langle t, W_n\rangle\, dP < \infty$ *for all* $t \in D_0$;

(ii′) $\lim_{n\to\infty} k_n^{-1}\psi_n(t) = c(t) \in R$ *for all* $t \in D_0$, *where* $\psi_n(t) = \log\varphi_n(t)$ *and* $\{k_n\}_{n=1,2,\cdots}$ *is a divergent sequence of positive real numbers*;

(iii′) c *is strictly convex in* D_0.

We now let D *be the subset of* D_0, *where* c' *exists, and suppose that* $0 \in D$. *Consider a bounded set* $A \subset R^k$ *with* $0 \in A^0$ *and*

(iv′) $a \in \{c'(t)\colon t \in D\}^0$ *for all* $a \in \partial A$;

(v′) $U_\varepsilon(a^*) \cap (\complement A)^0 \neq \emptyset\ \forall\, \varepsilon > 0$, *where* a^* *is a point in* ∂A *such that the conjugate* c^* *of* c *attains its infimum on* ∂A *at* a^* (cf. (2.7)).

It then follows that

$$\lim_{n\to\infty}[P(W_n \notin k_n A)]^{1/k_n} = \sup_{a\in\partial A}\rho(a) = \rho(a^*)\,, \tag{3.2}$$

where $\rho(a) = \exp[c(h) - \langle h, a\rangle]$ *and* h *is the unique solution of* $a = c'(h)$.

REMARK 3.1. Assumption (iv′), the strict convexity of c, and $c(0) = \lim_{n\to\infty} k_n^{-1}\log\varphi_n(0) = 0$, together yield the inequalities

$$0 < \rho(a) < 1 \qquad \text{for all} \quad a \in \partial A\,.$$

In particular, $\log\rho(a)$ is well defined on ∂A, and even on $B = \{a \in R^k\colon \rho(a) > 0\}$. Accordingly, we have

$$\begin{aligned}-\log\rho(a) &= -\log(\inf_{t\in D_0}\exp[c(t) - \langle t, a\rangle]) \\ &= \sup_{t\in D_0}[\langle t, a\rangle - c(t)] = c^*(a)\,.\end{aligned}$$

Thus, the continuity of ρ on B^0 follows from (2.7) and (2.1). The last equality in (3.2) then holds from the definition of a^* in (v′).

The proofs of both Theorems 3.1 and 3.2 are based on a certain convergence property of the conjugate distributions $P_{t,n}$, $n = 1, 2, \cdots$, of P. Here the $P_{t,n}$ are defined as

$$P_{t,n}(A) = \int_A \frac{\exp\langle t, W_n\rangle}{\varphi_n(t)}\, dP\,, \tag{3.3}$$

for $A \in \mathscr{A}$, and for a fixed $t \in D_0$.

LEMMA 3.1. *Consider the function c in Theorems* 3.1 *or* 3.2. *Further let* (a_1, a_2) *be an open interval in* R^k, *and put* $a = c'(t)$, *where t is a fixed vector in D. Supposing that* $a \in (a_1, a_2)$, *it follows from the assumptions of Theorem* 3.1 *or* 3.2 *that*

$$(3.4) \qquad \lim_{n\to\infty} P_{t,n}(W_n \in k_n(a_1, a_2)) = 1 .$$

REMARK 3.2. Conjugate distributions defined according to (3.3) are known to appear in a certain guise in statistical mechanics when microcanonical measures on a space of possible configurations of a thermodynamical system are defined (cf. Lanford (1973), page 44). The results of Theorems 3.1 and 3.2 thereby could be described as deducing the existence of the thermodynamical limit for the microcanonical ensemble from that of the canonical ensemble (cf. Lanford (1973), page 52, Theorem A5.1 and Corollary A5.2, where the opposite implication is deduced, and also Ruelle (1969), Chapter 3).

REMARK 3.3. The assumptions of Theorems 3.1 and 3.2 may be greatly simplified if $k_n = n$, and W_n is the nth partial sum of a sequence $\{X_i\}_{i=1,2,\cdots}$ of i.i.d. random vectors with finite moment-generating function $\varphi(t) = \int \exp\langle t, X_1\rangle \, dP$ for $t \in D_0$. The distribution of the X_i is not concentrated on a hyperplane of R^k. It follows that $\varphi_n(t) = [\varphi(t)]^n$, and thus c equals the cumulant-generating function $\log \varphi$ of the X_i. From the properties of $\log \varphi$ we know that partial derivatives of any order exist in D_0. In particular, the Hessian matrix $[\log \varphi(t)]''$ exists in D_0, and is positive definite since $[\log \varphi(t)]''$ is the dispersion matrix of X_1 with respect to the conjugate distribution P_t defined as $P_t(A) = \int_A (\exp\langle t, X_1\rangle/\varphi(t)) \, dP$, $A \in \mathscr{A}$, $t \in D_0$. Hence, by (2.5), $\log \varphi$ is strictly convex in D_0, and $\{(\log \varphi(t))' : t \in D_0\}$ is a nonempty open set, if C has a nonempty interior. Assumption (iv) of Theorem 3.1 can thus be fulfilled. Furthermore, relation (v′) in Theorem 3.2 is valid if A is either a closed or convex set.

4. Proofs.

PROOF OF LEMMA 3.1. Put $W_n = (W_{1n}, \cdots, W_{kn})^T$, $a_j = (a_{1j}, \cdots, a_{kj})^T$, $j = 1, 2$, $a = (a_1, \cdots, a_k)^T$. Then,

$$(4.1) \qquad P_{t,n}(W_n \notin k_n(a_1, a_2)) \leqq \sum_{i=1}^k P_{t,n}(W_{in} \leqq k_n a_{i1}) + \sum_{i=1}^k P_{t,n}(W_{in} \geqq k_n a_{i2}) .$$

Now, for negative τ_{i1}, it follows that

$$(4.2) \qquad P_{t,n}(W_{in} \leqq k_n a_{i1}) \leqq \int \frac{\exp\{\langle t, W_n\rangle + \tau_{i1}(W_{in} - k_n a_{i1})\}}{\varphi_n(t)} \, dP = \frac{\varphi_n(t + \tau_{i1} e_i) \exp(-k_n \tau_{i1} a_{i1})}{\varphi_n(t)} ,$$

where e_i denotes the ith column of a $k \times k$ identity matrix. Remembering that the cumulant-generating functions $\psi_n = \log \varphi_n$ are convex and differentiable in D_0, we have

$$(4.3) \qquad \log \varphi_n(t) \geqq \log \varphi_n(t + \tau_{i1} e_i) - \langle \psi_n'(t + \tau_{i1} e_i), \tau_{i1} e_i\rangle .$$

Relations (4.2) and (4.3) yield the upper bound

$$P_{t,n}(W_{in} \leqq k_n a_{i1}) \leqq \exp\left\{k_n \tau_{i1}\left[k_n^{-1}\frac{\partial\psi_n}{\partial t_i}(t + \tau_{i1}e_i) - a_{i1}\right]\right\}. \tag{4.4}$$

For fixed t_j, $j \neq i$, where $t = (t_1, \cdots, t_k)^T$, it holds that ψ_n and c are strictly convex functions of t_i, $i = 1, \cdots, k$. Using (2.2), (2.3), and (2.4) and recalling that $a = c'(t) \in (a_1, a_2)$, we can choose τ_{i1} such that $(\partial c/\partial t_i)(t + \tau_{i1}e_i)$ both exists and lies inside the open interval (a_{i1}, a_{i2}). In particular,

$$\frac{\partial c}{\partial t_i}(t + \tau_{i1}e_i) > a_{i1}, \qquad i = 1, \cdots, k. \tag{4.5}$$

Finally, we may note that in view of (2.6)

$$\lim_{n\to\infty} k_n^{-1}\frac{\partial\psi_n}{\partial t_i}(t + \tau_{i1}e_i) = \frac{\partial c}{\partial t_i}(t + \tau_{i1}e_i), \qquad i = 1, \cdots, k. \tag{4.6}$$

Using (4.4), (4.5), and (4.6) we thus have

$$\lim_{n\to\infty} P_{t,n}(W_{in} \leqq k_n a_{i1}) = 0, \qquad i = 1, \cdots, k. \tag{4.7}$$

By similar arguments one also obtains

$$\lim_{n\to\infty} P_{t,n}(W_{in} \geqq k_n a_{i2}) = 0, \qquad i = 1, \cdots, k. \tag{4.8}$$

Use of (4.1) renders the proof complete.

Large deviation results on the real line analogous to (3.1) or (3.2) are usually proven by deriving upper and lower bounds for the probabilities to be studied, and then showing that both bounds are equal. The upper bound is thereby often easy to obtain whereas the hard work is in the lower bound (see, e.g., Bahadur (1971), page 5, and Landford (1973), page 40). In the same vein the proofs of Theorems 3.1 and 3.2 are divided into two parts. The easier upper bound is first derived and then an estimate of the lower bound is obtained.

Proof of Theorem 3.1: (a) *Upper bound*: The relation $W_n \in k_n(a_n + C)$ is equivalent to $k_n^{-1}(W_n - k_n a_n) \in C$. Then, from the definition of the polar cone C^+, $\langle t, W_n - k_n a_n\rangle \geqq 0$ for $t \in C^+$. Hence, it follows that $P(W_n \in k_n(a_n + C)) \leqq \varphi_n(t)\exp(-k_n\langle t, a_n\rangle)$, for $t \in D \cap C^+$ and using assumption (ii),

$$\limsup_{n\to\infty}\,[P(W_n \in k_n(a_n + C))]^{1/k_n} \leqq \exp[c(h) - \langle h, a\rangle] \tag{4.9}$$

for $h \in D \cap C^+$. Since $a = c'(h)$ and $h \in D_0$ the equality $\exp[c(h) - \langle h, a\rangle] = \inf_{t\in D_0}\exp[c(t) - \langle t, a\rangle]$ is obvious.

(b) *Lower bound*: Assumption (iv) implies the existence of a k-dimensional open interval $(a_1, a_2) \subset U_\varepsilon(a) \cap (a + C)^0$ with $(a_1, a_2) \cap \{c'(t): t \in D\} \neq \emptyset$ for an arbitrary $\varepsilon > 0$. Let $t = t(\varepsilon)$ be such that $c'(t) \in (a_1, a_2)$. Then for n sufficiently large

$$\begin{aligned} &P(W_n \in k_n(a_n + \dot{C})) \\ &\quad \geqq P(W_n \in k_n(a_1, a_2)) \\ &\quad = \varphi_n(t)\int_{\{W_n \in k_n(a_1,a_2)\}} \exp(-\langle t, W_n\rangle)\frac{\exp\langle t, W_n\rangle}{\varphi_n(t)}\,dP \\ &\quad \geqq \varphi_n(t)\exp(-k_n\langle t, a_0\rangle)P_{t,n}(W_n \in k_n(a_1, a_2)), \end{aligned} \tag{4.10}$$

where $a_0 = (a_{10}, \cdots, a_{k0})^T$ is defined by $a_{i0} = a_{i1}$ (if $t_i \leqq 0$) and $a_{i0} = a_{i2}$ (if $t_i > 0$). In view of Lemma 3.1 and assumption (i) relation (4.10) yields

$$\begin{aligned} &\liminf_{n\to\infty} [P(W_n \in k_n(a_n + C))]^{1/k_n} \\ (4.11) \qquad &\geqq \exp[c(t) - \langle t, a_0\rangle] \\ &\geqq \inf_{t\in D_0} \exp[c(t) - \langle t, a_0\rangle] = \rho(a_0)\,. \end{aligned}$$

Now $\lim_{\varepsilon\to 0} a_0 = a \in \{c'(t) : t \in D \cap C^+\}^0$, and $\rho(a) > 0$. By the continuity of ρ in a neighborhood of a (cf. Remark 3.1) relation (4.11) then implies that

$$(4.12) \qquad \liminf_{n\to\infty} [P(W_n \in k_n(a_n + C))]^{1/k_n} \geqq \rho(a)\,.$$

From (4.9) and (4.12) the proof is now complete.

Before the proof of Theorem 3.2 we note some useful facts:

REMARK 4.1. Using (2.4) the strict convexity of c implies that the solution h of $a = c'(h)$ in Theorem 3.2 is unique. Hence $h = h(a)$, and $h \neq 0$ if $a \neq 0$, because $\rho(0) = \inf_{t\in D_0} \exp[c(t)] = \exp[c(0)]$ in view of the equalities $c'(0) = \lim_{n\to\infty} k_n^{-1}\psi_n'(0) = \lim_{n\to\infty} k_n^{-1}EW_n = 0$.

REMARK 4.2. For $h(a) \neq 0$ we define the hyperplane

$$H(a) = \{x \in R^k : \langle h(a), x - a\rangle = 0\}$$

with the open half-spaces

$$\begin{aligned} H^+(a) &= \{x \in R^k : \langle h(a), x - a\rangle > 0\}\,, \\ H^-(a) &= \{x \in R^k : \langle h(a), x - a\rangle < 0\}\,. \end{aligned}$$

As mentioned in Remark 3.1 it follows that

$$\rho(a) = \exp[c(h(a)) - \langle h(a), a\rangle] < 1$$

for all $a \in \partial A$, which in turn implies that

$$(4.13) \qquad \langle h(a), a\rangle > c(h(a))\,,$$

and $0 \in H^-(a)$, since $c(h(a)) > c(0) = 0$, for $a \in \partial A$. Moreover, from $c'(0) = 0 \in A^0$, $c'(h(a)) = a \in \partial A$, and the continuity of c' on D (cf. (2.3)), it follows that the vectors $h(a)$ are outside a neighborhood of the origin. Hence a real number c_0 exists such that

$$(4.14) \qquad c(h(a)) \geqq c_0 > 0 = c(0) = \inf_{t\in D_0} c(t)$$

for all $a \in \partial A$.

REMARK 4.3. For $\varepsilon > 0$ and $a \in \partial A$ let $H_\varepsilon(a)$ denote *the* supporting hyperplane of the ε-neighborhood $U_\varepsilon(a)$ which parallels $H(a)$ and is tangential to $U_\varepsilon(a)$ at a unique point $a_\varepsilon \in H^-(\mathrm{a})$. We then have

$$H_\varepsilon(a) = \{x \in R^k : \langle h(a), x - a_\varepsilon\rangle = 0\}\,,$$

and $H_\varepsilon(a)$ also determines the open half-spaces

$$\begin{aligned} H_\varepsilon^+(a) &= \{x \in R^k : \langle h(a), x - a_\varepsilon\rangle > 0\}\,, \\ H_\varepsilon^-(a) &= \{x \in R^k : \langle h(a), x - a_\varepsilon\rangle < 0\}\,. \end{aligned}$$

Using (4.13), (4.14) and the boundedness of $\{h(a) : a \in \partial A\}$, a positive real number ε_0 can be chosen such that

$$0 \in H_\varepsilon^-(a) \quad \text{for all} \quad a \in \partial A \quad \text{and all} \quad \varepsilon \in (0, \varepsilon_0) . \tag{4.15}$$

REMARK 4.4. For every $\varepsilon \in (0, \varepsilon_0)$ there exists a finite number of neighborhoods $U_\varepsilon(a)$, $a \in \partial A$ (say) $U_\varepsilon(a_1), \cdots, U_\varepsilon(a_K)$, where $K = K(\varepsilon)$ and with

$$\partial A \subset \bigcup_{i=1}^K U_\varepsilon(a_i) . \tag{4.16}$$

We now show that

$$B_K \equiv \bigcap_{i=1}^K H_\varepsilon^-(a_i) \subset A^0 . \tag{4.17}$$

First we note from (4.15) that B_K is a convex set containing the origin. Also,

$$a \notin B_K \quad \text{for all} \quad a \in \partial A \tag{4.18}$$

from (4.16) and the definition of $H_\varepsilon^-(a_i)$. But then $a \notin B_K$ for all $a \in \complement\bar{A}$, as otherwise $a \in B_K$ for some $a \in \complement\bar{A}$ and $0 \in B_K$ would imply that $\{\lambda a : \lambda \in [0, 1]\}$ is a subset of B_K. This would contradict (4.18).

PROOF OF THEOREM 3.2: (a) *Upper bound*: From (4.17) we have for $\varepsilon \in (0, \varepsilon_0)$

$$\begin{aligned} P(W_n \notin k_n A) &\leqq P(W_n \notin k_n(\textstyle\bigcap_{i=1}^K H_\varepsilon^-(a_i))) \\ &= P(W_n \in k_n(\textstyle\bigcup_{i=1}^K (H_\varepsilon(a_i) \cup H_\varepsilon^+(a_i)))) \\ &\leqq \textstyle\sum_{i=1}^K P(W_n \in k_n(H_\varepsilon(a_i) \cup H_\varepsilon^+(a_i))) \\ &= \textstyle\sum_{i=1}^K P(\langle h(a_i), W_n - k_n a_{i,\varepsilon}\rangle \geqq 0) . \end{aligned} \tag{4.19}$$

Here $a_{i,\varepsilon}$ is the unique point in $H^-(a_i)$ at which the supporting hyperplane $H_\varepsilon(a_i)$ of $U_\varepsilon(a_i)$ that parallels $H(a_i)$ is tangential to $U_\varepsilon(a_i)$ (cf. Remark 4.3).

The probabilities for the sum can be bounded by

$$\begin{aligned} &P(\langle h(a_i), W_n - k_n a_{i,\varepsilon}\rangle \geqq 0) \\ &\quad \leqq E \exp\langle h(a_i), W_n - k_n a_{i,\varepsilon}\rangle \\ &\quad = \varphi_n(h(a_i)) \exp(-\langle h(a_i), k_n a_i\rangle) \exp\langle h(a_i), k_n(a_i - a_{i,\varepsilon})\rangle . \end{aligned} \tag{4.20}$$

As already mentioned, the set $\{h(a) : a \in \partial A\}$ is bounded, i.e., we may say $|h(a)| \leqq h_0$ for all $a \in \partial A$. Now

$$|\langle h(a_i), k_n(a_i - a_{i,\varepsilon})\rangle| \leqq |h(a_i)|\,|k_n(a_i - a_{i,\varepsilon})| \leqq k_n h_0 \varepsilon . \tag{4.21}$$

Finally, assumption (ii′) yields

$$\begin{aligned} \lim_{n\to\infty} &[\varphi_n(h(a_i)) \exp(-k_n\langle h(a_i), a_i\rangle)]^{1/k_n} \\ &= \exp[c(h(a_i)) - \langle h(a_i), a_i\rangle] = \rho(a_i) . \end{aligned} \tag{4.22}$$

Using relations (4.19)—(4.22) it follows that

$$\begin{aligned} \limsup_{n\to\infty} [P(W_n \notin k_n A)]^{1/k_n} &\leqq e^{h_0\varepsilon} \max_{i=1,\cdots,K} \rho(a_i) \\ &\leqq e^{h_0\varepsilon} \sup_{a\in\partial A} \rho(a) . \end{aligned} \tag{4.23}$$

Upon letting ε tend to zero in (4.23) the upper bound is derived, i.e.,

$$\limsup_{n\to\infty} [P(W_n \notin k_n A)]^{1/k_n} \leqq \sup_{a\in\partial A} \rho(a) = \rho(a^*) . \tag{4.24}$$

(b) *Lower bound*: For every $\varepsilon > 0$ assumptions (iv′) and (v′) imply the existence of a k-dimensional open interval

$$(a_1, a_2) \subset U_\varepsilon(a^*) \cap (\complement A)^0 \qquad \text{with} \quad (a_1, a_2) \cap \{c'(t) : t \in D\} \neq \emptyset \, .$$

By the same arguments as those used in part (b) of the proof of Theorem 3.1 it then follows that

$$\liminf_{n\to\infty} [P(W_n \notin k_n A)]^{1/k_n} \geqq \rho(a^*) \, . \tag{4.25}$$

This completes the proof.

Acknowledgments. I wish to express my thanks to Professor D. Plachky for introducing me to the problems of large deviations and to Professor H. Klinger for many helpful discussions. Moreover, I should like to thank a referee for his very helpful remarks on an earlier version of the paper which led to a more generalized form of Theorem 3.1 in addition to many technical improvements. I am also thankful for the referee's comments on certain related problems in statistical mechanics as mentioned in Remark 3.2.

REFERENCES

[1] BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics.* Regional Conference Series in Applied Mathematics, No. 4, Philadelphia.

[2] BOOK, S. A. (1973). A large deviation theorem for weighted sums. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **26** 43–49.

[3] BOOK, S. A. (1975). Convergence rates for a class of large deviation probabilities. *Ann. Probability* **3** 516–525.

[4] KARLIN, S. (1959). *Mathematical Methods and Theory in Games, Programming, and Economics* **1**. Addison-Wesley, London.

[5] LANFORD, O. E. (1973). Entropy and equilibrium states in classical statistical mechanics. In *Statistical Mechanics and Mathematical Problems.* Battelle Seattle 1971 Rencontres (A. Lenard, ed.). *Lecture Notes in Physics*, No. 20. Springer-Verlag, Berlin.

[6] PLACHKY, D. and STEINEBACH, J. (1975). A theorem about probabilies of large deviations with an application to queueing theory. *Period. Math. Hungar.* **6** 343–345.

[7] ROBERTS, A. W. and VARBERG, D. E. (1973). *Convex Functions.* Academic Press, New York.

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

[9] RUELLE, D. (1969). *Statistical Mechanics.* Benjamin, New York.

[10] SIEVERS, G. L. (1969). On the probability of large deviations and exact slopes. *Ann. Math. Statist.* **40** 1908–1921.

[11] SIEVERS, G. L. (1975). Multivariate probabilities of large deviations. *Ann. Statist.* **3** 897–905.

[12] STEINEBACH, J. (1976). Exponentielles Konvergenzverhalten von Wahrscheinlichkeiten großer Abweichungen. Thesis, Univ. of Düsseldorf.

INSTITUT FÜR STATISTIK UND DOKUMENTATION
UNIVERSITÄT DÜSSELDORF
4000 DÜSSELDORF
WEST GERMANY