The Annals of Probability
1983, Vol. 11, No. 1, 158–167

DOMINATING POINTS AND THE ASYMPTOTICS OF LARGE DEVIATIONS FOR RANDOM WALK ON $\mathbb{R}^d$

BY PETER NEY[1]

University of Wisconsin, Madison

Let $\mu(\cdot)$ be a probability measure on $\mathbb{R}^d$ and B be a convex set with nonempty interior. It is shown that there exists a unique "dominating" point associated with (μ, B). This fact leads (via conjugate distributions) to a representation formula from which sharp asymptotic estimates of the large deviation probabilities $\mu^{*n}(nB)$ can be derived.

1. Introduction and summary. Let $S_n = X_1 + \cdots + X_n$, where $\{X_k; k = 1, 2, \cdots\}$ are i.i.d. random variables with $\mu(B) = P\{X \in B\}$, and $m = EX_1$. The existence and properties of

(1.1) $$\lim_{n\to\infty}[\mu^{*n}(nB)]^{1/n} \equiv \rho(B) = \exp\{\text{entropy of } B\},$$

have been extensively studied. When the $\{X_k\}$ are real valued the subject goes back to Cramér [7] and Chernoff [6]. In the most general setting, Bahadur and Zabell [2] have shown that if $\{X_k\}$ take values in a topological vector space V, then subject to mild regularity conditions, $\rho(\cdot)$ exists for open convex B.

In this paper we consider the case $V = \mathbb{R}^d$.

$$\text{Let: } \mathrm{In}A = \text{interior}, \quad \mathrm{Cl}A = \text{closure}, \quad \partial A = \text{boundary of } A;$$

$$\mathscr{S} = \text{the convex hull of the support of } \mu;$$

$$\varphi(\alpha) = E\exp(\alpha\cdot X), \qquad \alpha \in \mathbb{R}^d;$$

$$\mathscr{D}(\varphi) = \{\alpha : \varphi(\alpha) < \infty\} = \text{"effective domain" of } \varphi;$$

$$H^+(\alpha, v) = \text{the half-space } \{x \in \mathbb{R}^d : x\cdot\alpha \ge v\cdot\alpha\}, \qquad \alpha, v \in \mathbb{R}^d.$$

If $\mathscr{D}$ contains a neighborhood of the origin then for any open set B, $\rho(B)$ exists and can be identified as

(1.2i) $$\rho(B) = \sup\{\tilde\rho(v) : v \in B\},$$

where

(1.2ii) $$\tilde\rho(v) = \inf\{e^{-\alpha\cdot v}\varphi(\alpha) : \alpha \in \mathbb{R}^d\}$$

(Bartfai [5]).

We will see that under a slightly stronger hypoethesis one can go further.

We make the following:

DEFINITION. A point $v_B \in \mathbb{R}^d$ is a *dominating point* of the set B if

(1.3)
(i) $v_B \in \partial B$,
(ii) the equation (in α) $\operatorname{grad}\varphi(\alpha) = v_B\varphi(\alpha)$ has a unique solution $\alpha(v_B) \in \mathscr{D}$; and
(iii) $B \subset H^+(\alpha(v_B), v_B)$.

Received October 1981; revised July 1982.

[1] Research supported by the N.S.F. Part of this work was done while the author was at the Centre de Rech. Math. at the University of Montreal. The hospitality of the Centre is gratefully acknowledged.

AMS 1980 *subject classifications.* Primary 60F10; secondary 60G50.
Key words and phrases. Large deviations, random walk.

Our main result is:

THEOREM. *If $\mathscr{D}$ is an open set, B is convex with* $\text{In}(B \cap \mathscr{S}) \neq \phi$, *and $m \notin B$, then there exists a unique dominating point. This point is in* In $\mathscr{S}$.

The motivation for using the word "dominating" will shortly be apparent. My idea to look for such a point has its germ in a joint paper with F. Spitzer [9] on Martin boundary construction. In various forms, the idea has been under the surface of much of the large deviation literature.

The utility of the above theorem is that it leads to the following

REPRESENTATION FORMULA: *If v_B is a dominating point of B then*

$$\mu^{*n}(nB) = \rho^n(B) \int_{n(B-v_B)} e^{-\alpha(v_B)\cdot x}\, \tilde{\mu}^{*n}(dx), \tag{1.4}$$

where $\tilde{\mu}$ is a probability measure (to be specified) with mean O (the origin), and

$$\alpha(v_B)\cdot x \geq 0 \quad \textit{for} \quad x \in B - v_B. \tag{1.5}$$

The crucial property is (1.5). One can easily get many representations like (1.4); but these are quite useless without (1.5). With this property one can apply *standard central limit estimates* (available in great accuracy, as e.g. in [4]) to estimate the asymptotic behavior of $\mu^{*n}(nB)$. For example, a first consequence is the inequality:

$$c_1 n^{-d/2} \leq \rho^{-n}(B)\mu^{*n}(nB) \leq c_2 n^{-1/2} \tag{1.6}$$

for some $0 < c_1, c_2 < \infty$, which is already a refinement of (1.1). Under somewhat stronger hypotheses one gets asymptotic behavior between these bounds. For any $1/2 \leq \gamma \leq d/2$, there are examples where

$$\mu^{*n}(nB) = \rho^n n^{-\gamma}[c + O(n^{-\delta})], \tag{1.7}$$

and in principle, more refined results can be obtained.

The representation formula (1.4) with (1.5) is valid any time a dominating point exists. The conditions of the theorem are not necessary for this, but at present there do not appear to be other attractive general hypotheses.

From (1.4) one also obtains the corollaries that (i) $\tilde{\rho}(v_B) = \rho(B)$, namely the "point entropy equals the set entropy function"; and (ii) if η is any neighborhood of v_B, then

$$\rho(B \cap \eta) = \rho(B). \tag{1.8}$$

It is this property which suggests calling v_B a "dominating" point.

When the r.v.'s $\{X_k\}$ are $\in \mathbb{R}^1$ and $B = [a, \infty)$, with $EX = m < a$, then the dominating point is trivially $v_B = a$; also $B - v_B = [0, \infty)$, $\alpha > 0$, and $\tilde{\mu}$ is a centering of the 1-dimensional Cramér-Chernoff transform. The representation (1.4) reduces to a form used by Bahadur and Rao [1] to obtain a complete asymptotic expansion of μ^{*n}.

The case when $\mu(\cdot)$ is a multinomial distribution has received special attention in the literature due to its importance in statistical applications. Applying classical asymptotic estimates directly to μ^{*n} it has been shown that in this case (see J. Reeds [10])

$$\log P\{S_n \in nA\} - \log\{n^{-1/2}\rho^n\} = O(1) \tag{1.9}$$

for A a C^∞ manifold, $A \subset d$-simplex. This is consistent with example (i) in Section 5.

The paper is organized as follows. In Section 2 we describe the conjugate (Cramér-Chernoff) transform (a well-known tool in large deviation theory); also we state two lemmas which contain some information we will need about the map $\alpha(\cdot)$ which is defined as the solution of grad $\varphi(\alpha) = v\varphi(\alpha)$, and about the function $\tilde{\rho}(\cdot)$. The representation formula and some first consequences are derived in Section 3; the theorem is proved in

Section 4. This is done by a fixed point argument, which for $d = 2$ involves some simple geometry, and for $d \geq 3$ an elementary homotopy argument. Finally, some further examples of the application of the representation formula are in Section 5.

2. Conjugate distributions and some related functions. For $\alpha \in \mathscr{D}(\varphi)$, the conjugate distribution of $\mu(\cdot)$ is defined by

$$\mu_\alpha(dx) = [\varphi(\alpha)]^{-1} e^{\alpha \cdot x} \mu(dx). \tag{2.1}$$

(When $d = 1$, this is called the Cramér or Chernoff transform.) Let $X^{(\alpha)}$ be the random variable with distribution μ_α and note that

$$Ee^{\tau \cdot X^{(\alpha)}} = \frac{\varphi(\tau + \alpha)}{\varphi(\alpha)}, \qquad \alpha, \tau \in \mathbb{R}^d, \tag{2.2}$$

and

$$EX^{(\alpha)} = \frac{\text{grad } \varphi(\alpha)}{\varphi(\alpha)}. \tag{2.3}$$

The following two lemmas summarize some facts about the solution $\alpha(v)$ of

$$\text{grad } \varphi(\alpha) = v\varphi(\alpha), \qquad v \in \text{In } \mathscr{S}, \tag{2.4}$$

and about the related function

$$\tilde{\rho}(v) = e^{-\alpha(v) \cdot v} \varphi(\alpha(v)), \qquad v \in \text{In } \mathscr{S}. \tag{2.5}$$

These are known results in the theory of exponential families and convexity. As a reference see, for example, the book by O. Barndorff-Nielsen [3].

LEMMA 1. *If $v \in \text{In } \mathscr{S}$ and $\mathscr{D}(\varphi)$ is an open set, then* (2.4) *has a unique solution* $\alpha(v)$. *The function $\alpha : \text{In } \mathscr{S} \to \mathscr{D}(\varphi)$ is $\in C^\infty$, and*

$$EX^{(\alpha(v))} = v. \tag{2.6}$$

REMARKS. (i) It is easily seen that if $v \notin \text{In } \mathscr{S}$, then (2.4) may not have a solution. If $v \in \text{In } \mathscr{S}$ and $\mathscr{D}$ is not open, then it may or may not have a solution. Take for example the mixture of distributions which is (-1) with probability p, and has density $ce^{-y}y^{-4}$, $y \geq 1$, with probability $(1 - p)$. Then $\mathscr{D} = (-\infty, 1]$ (closed on the right). If p is close to 1 then $\varphi'(\alpha) < 0$ for all $\alpha \in \mathscr{D}$ and (2.4) has no solution. If p is small, there will be a solution. Thus the condition "$\mathscr{D}$ open" is sufficient but not necessary.

(ii) "$\mathscr{D}$ Open" also implies that the exponential family generated by μ is "steep" in the language of convexity theory (see [3]). The hypothesis of steepness could replace "$\mathscr{D}$ open" in the lemma.

Now with $\tilde{\rho}(\cdot)$ as defined in (2.5), let

$$R_c = \{v \in \mathbb{R}^\alpha : \tilde{\rho}(v) \geq c\}, \qquad 0 \leq c \leq 1. \tag{2.7}$$

LEMMA 2. *If $\mathscr{D}$ is open, then on* In $\mathscr{S}$

(i) $\tilde{\rho} \in C^{(\infty)}$,
(ii) $\text{grad } \tilde{\rho}(v) = -\alpha(v)\tilde{\rho}(v)$,
(iii) $-\log \tilde{\rho}(v)$ *is strictly convex*,
(iv) $\tilde{\rho}(v) = \inf\{e^{-x \cdot v}\varphi(x) : x \in \mathbb{R}^d\}$ (*as in* (1.2)),
(v) $\tilde{\rho}(m) = 1$,
(vi) $\{R_c\}$ *are convex sets which* $\searrow \{m\}$ *as* $c \nearrow 1$.

3. The representation formula. Let us suppose that the set B has been shown to have a dominating point v_B, and examine some of the consequences. We start with the conjugate transform defined in (2.1), convolve it with itself n times, multiply and divide by $e^{-n(\alpha\cdot v)}$, and make a change of variable in the integral to get

$$\mu^{*n}(nB) = [e^{-\alpha\cdot v}\varphi(\alpha)]^n \int_{n(B-v)} e^{-\alpha\cdot x}\,\mu_\alpha^{*n}(dx + nv), \tag{3.1}$$

for any $\alpha \in \mathscr{D}$, $v \in \mathbb{R}^d$.

We assume throughout this section that $\mathscr{D}$ is open and B is convex with nonempty interior. Hence by the theorem, B has a dominating point v_B. We are free to choose $v = v_B$ and $\alpha = \alpha(v_B)$ (Section 2) in (3.1). This yields

$$\mu^{*n}(nB) = \tilde{\rho}^n(v_B) \int_{n(B-v_B)} e^{-\alpha(v_B)\cdot x}\,\tilde{\mu}^{*n}(dx), \tag{3.2}$$

where $\tilde{\mu}(dx) = \mu_\alpha(dx + v_B)$ and $\tilde{\rho}(v) = e^{-\alpha(v)\cdot v}\varphi(\alpha(v))$. Clearly

$$\int x\, d\tilde{\mu} = O \tag{3.2a}$$

and

$$\alpha(v_B)\cdot x \geq 0 \quad \text{for} \quad x \in B - v_B. \tag{3.2b}$$

This is the formula (1.4).

We observe first that (3.2) implies

$$\rho(B) = \tilde{\rho}(v_B), \tag{3.3}$$

where $\rho(B)$ is defined by (1.1). This follows from the fact that

$$\rho(B) \leq \tilde{\rho}(v_B)$$

since the integral in (3.2) ≤ 1; while from the characterization of $\rho(B)$ in (1.2) and the continuity of $\tilde{\rho}(\cdot)$ on In $\mathscr{S}$ we see that $\rho(B) \geq \tilde{\rho}(v_B)$. Hence (3.3).

Let us derive some first estimates of $\mu^{*n}(nB)$ based on the representation formula.

PROPOSITION. *If the hypotheses of the theorem are satisfied and μ is either lattice or strongly nonlattice, then there exist constants c_1, c_2 such that for sufficiently large n*

$$c_1\tilde{\rho}^n(v_B)n^{-d/2} \leq P\{S_n \in nB\} \leq c_2\tilde{\rho}^n(v_B)n^{-1/2}. \tag{3.4}$$

(Strongly nonlattice means that the modulus of the Fourier transform of μ equals one only at the origin.)

PROOF. It is necessary only to estimate the integral in (3.2). To do so, we make a change of variable in the integral, $y = Rx$, where R is the rotation that takes $\alpha = (\alpha_1, \cdots, \alpha_d)$ into $R\alpha = (\|\alpha\|, 0, \cdots, 0)$. The resultant measure $\nu(\cdot)$ (say) will still have mean O, and letting $A = \{Rx : x \in B - v_B\}$ we still have

(3.5)
(i) A convex with non-empty interior,
(ii) $O \in \partial A$, and now
(iii) $A \subset$ the right half plane $\{x : x_1 \geq 0\}$.

Thus we must estimate

$$\int_{nA} e^{-\|\alpha\| y}\nu^{*n}(dy) \equiv I_n(nA). \tag{3.6}$$

But for any sphere S about O, clearly

$$I_n(nA) \geq I_n((nA) \cap S) \geq c\nu^{*n}(nA \cap S)$$

for some constant $c > 0$. Since A is convex with nonempty interior and $O \in \partial A$, $nA \cap S$ contains a truncated cone $\mathscr{K}$ with vertex O and In $\mathscr{K} \neq \varnothing$. Thus

$$I_n(nA) \geq c\nu^{*n}(\mathscr{K}), \tag{3.7}$$

and if μ satisfies a local limit condition (e.g. lattice or strongly nonlattice), then (see [4])

$$\nu^{*n}(\mathscr{K}) \geq c'n^{-d/2}. \tag{3.8}$$

The upper bound in (3.4) follows from a comparison with the known 1-dimensional case [1]. □

Recalling (3.3) we also see that

COROLLARY. *Under the hypothesis of the proposition it follows that for any neighborhood $\mathscr{N}$ of v_B*

$$\rho(B \cap \mathscr{N}) = \rho(B).$$

This follows from the proposition and the fact that $v_{B\cap\mathscr{N}} = v_B$.

4. Proof of the theorem. We first give a proof under further restrictions, and then successively remove these.

CASE (A): *B a smooth, bounded, interior set.* Suppose that (i) B is bounded; (ii) ∂B is smooth in the sense that at each point $v \in \partial B$ there is a unique supporting tangent plane $T(v)$ and (exterior) normal unit vector $\mathbb{N}(v)$; (iii) Cl $B \subset$ In $\mathscr{S}$. Consider the normal map:

$$\mathbb{N}: \partial B \to \partial S_1, \tag{4.1}$$

where ∂S_1 is the boundary of the unit ball in $\mathbb{R}^d$; and the map defined in Section 2:

$$\alpha: \partial B \to \mathscr{D}. \tag{4.2}$$

Note that $m \notin B$ implies $O \notin \alpha(B) = \{\alpha(v) : v \in B\}$.

Thus $\mathbb{N}$ and α are both homeomorphisms from ∂B into $\mathbb{R}^d - \{O\}$. If the continuous map

$$F: \partial B \times [0, 1] \to \mathbb{R}^d$$

defined by $F(\lambda, v) = \lambda\alpha(v) + (1 - \lambda)\mathbb{N}(v)$ were also into $\mathbb{R}^d - \{O\}$, then $\mathbb{N}$ and α would be homotopic on $\mathbb{R}^d - \{O\}$. But this is impossible since they have different degrees. Hence $F(\lambda^*, v^*) = O$ for some $0 < \lambda^* < 1$, $v^* \in \partial B$. (Note that $\lambda^* = 0$ or 1 would imply $F(\lambda^*, v^*) = \mathbb{N}(v^*) = O$ or $\alpha(v^*) = O$, which is impossible). Thus

$$\alpha(v^*) = -c\mathbb{N}(v^*) \quad \text{for some} \quad 0 < c < \infty, \tag{4.3}$$

implying that $B \subset H^+(\alpha(v^*), v^*)$.

REMARK. The above argument is particularly simple to visualize when $d = 2$. Fix a point $v_0 \in \partial B$ and let $\Theta_{\mathbb{N}}(v) \in [0, 2\pi]$ be the angle between $\mathbb{N}(v_0)$ and $\mathbb{N}(v)$, $\Theta_\alpha(v) =$ the angle between $\mathbb{N}(v_0)$ and $\alpha(v)$. Now let v run (say in the positive direction) from v_0, around ∂B, and back to v_0. Then $\Theta_{\mathbb{N}}(v)$ increases from 0 to 2π, while $\Theta_\alpha(v)$ returns to the same value $\Theta_\alpha(v_0)$. Thus there must be a point v^* where $\Theta_\alpha(v^*) = \Theta_{\mathbb{N}}(v^*)$.

To prove uniqueness, suppose that there were two dominating points v_1^* and v_2^*. Then by (3.3) $\tilde{\rho}(v_1^*) = \rho(B) = \tilde{\rho}(v_2^*)$. Hence $v_i^* \in \partial R_{\rho(B)}$, $i = 1, 2$ (see (2.7)). Also of course $v_i^* \in \partial B$, $i = 1, 2$. By (4.3) and Lemma 2 (ii) we see that $R_{\rho(B)}$ and B are in opposite half spaces, with a separating hyperplane containing v_i^*, $i = 1, 2$. But this is impossible unless $v_1^* = v_2^*$.

CASE (B). *B a bounded interior set.* For any set A, let $A_\varepsilon = \cup_{v\in A}(v + S_\varepsilon)$ where $S_\varepsilon = \{x : \|X\| \le \varepsilon\}$, i.e. A_ε is the ε-neighborhood of A. Suppose that B is a bounded set with Cl $B \subset$ In $\mathscr{S}$. Then B_ε is bounded, convex, and smooth (in the sense of case (A)). If ε is small enough, then Cl $B_\varepsilon \subset$ In $\mathscr{S}$. Hence by Case (A), for each $k \ge 1$, there exists a unique dominating point v_k or $B_{(1/k)}$. By compactness there is a subsequence $v_{k_i} \to v_0 \in \partial B$. If $x \in B \subset B_{1/k_i}$, then by the dominating property of v_{k_i}, $(x - v_{k_i}) \cdot \alpha(v_{k_i}) \ge 0$, and by continuity $(x - v_0) \cdot \alpha(v_0) \ge 0$; i.e. v_0 is dominating for B. Uniqueness follows exactly as in part (a).

CASE (C) *General case.* Finally we remove the boundedness and "interior" restrictions, and consider any convex B with nonempty interior.

To end this we must define yet another class of sets. Let $\mathscr{S}_\varepsilon^- = \mathscr{S} - (\partial\mathscr{S})_\varepsilon$ = the interior ε-approximation of $\mathscr{S}$, and S_r = the ball of radius r, center O. Let $\{\delta_k; k = 1, 2, \cdots\}$, $\{r_k; k = 1, 2, \cdots\}$ be sequences such that $0 \le \delta_k \searrow 0$, $\infty > r_k \nearrow \infty$ as $k \to \infty$. We define

$$B^{(k)} = \mathscr{S}_{\delta_k}^- \cap S_{r_k} \cap B. \tag{4.4}$$

Since In $B \ne \phi$, one can choose δ_1 small enough and r_1 large enough so that In $B^{(1)} \ne \phi$. Hence

$$\mu^{*n_0}(n_0 B^{(1)}) > 0 \text{ for some integer } n_0. \tag{4.5}$$

Of course $B^{(1)}, B^{(2)}, \cdots$, are all bounded interior sets with In $B^{(k)} \ne \phi$, and $B^{(k)} \nearrow B$ as $k \to \infty$. Thus by case (b), there exists a unique dominating point v_k for $B^{(k)}$, $k = 1, 2, \cdots$. There are two possibilities:

CASE (i). There exists a $k_0 \ge 1$, such that $v_{k_0} = v_{k_0+1}$ or

CASE (ii). $v_k \ne v_{k+1}$ for all $k \ge 1$.

To prove the theorem in Case (i) we use

LEMMA 3. *Suppose B and $C_1, C_2, \cdots$ are convex set such that*

(1) Cl $C_i \subset$ In C_{i+1}, $i = 1, 2, \cdots$
(2) $B \subset (\cup C_i)$
(3) In$(B \cap C_1) \ne \phi$.

If H is a hyperplane which supports both $(B \cap C_i)$ at v and $(B \cap C_j)$ at v for some $i < j$, then H supports B at v.

This follows from the following fact: Consider two convex sets A and E, with $A \subset E$, and suppose that there is a point v in the "interior" of $\partial A \cap \partial E$ in the sense that

$$(\mathscr{N}_v \cap \partial A) \subset (\partial A \cap \partial E) \tag{4.6}$$

for some neighborhood $\mathscr{N}_v$ of v. If there is a supporting hyperplane H of A at v, then H also supports E (at v). Now apply this to the setup in Lemma 3, with $A = B \cap C_j$ and $E = B$. The fact that H supports both $B \cap C_i$ and $B \cap C_j$ at v implies that $v \in$ interior of $\partial(B \cap C_j)$, hence of $\partial(B \cap C_j) \cap \partial B$. Hence H also supports B. □

Returning to the proof of the theorem, let $\{B^{(k)}\}$ as defined in (4.4) play the role of $\{C_k\}$ in Lemma 3. In case (i) $B^{(k_0)}$ and $B^{(k_0+1)}$ have a common dominating point, say v^*, and supporting hyperplane $H(\alpha(v^*), v^*) = \{x : x \cdot \alpha(v^*) \ge v \cdot \alpha(v^*)\}$. By the lemma, then, we also have $v^* \in \partial B$ and $B \subset H^+(\alpha(v^*), v^*)$, i.e. v^* is a dominating point for B. Clearly also $v^* \in$ In $\mathscr{S}$. Uniqueness follows as in part (a). Thus, in Case (i), the theorem is proved.

We conclude the proof by showing that Case (ii) is impossible. In this case, necessarily

$v_{k+1} \in (B^{(k)})^c$ and hence either

or (ii1.) $\| v_k \| \to \infty$

(ii2.) $v_{k'} \to v_0 \in \partial\mathscr{S}$ on a subsequence $\{k'\}$.

We will use the fact, whose proof we defer to the end, that

$$\lim_{\| v_k \| \to \infty} \tilde{\rho}(v_k) = 0. \tag{4.7}$$

But

$$\rho(B_1) = \lim_{n\to\infty} P\{S_n \in nB^{(1)}\}^{1/n} \le \lim P\{S_n \in nB^{(k)}\}^{1/n} = \tilde{\rho}(v_k),$$

while (δ_1, r_1) in (3.4) can be chosen so that $\rho(B_1) > 0$. Hence if $\| v_k \| \to \infty$ we are led to a contradiction.

It remains only to consider Case (ii2). To this end we use the conjugate transform (2.1), $\alpha = \alpha(v_k)$, plus a little manipulation, to write

$$\mu(B^{(1)}) = \tilde{\rho}(v_k) \int_{B^{(1)}} e^{-\alpha(v_k)\cdot(x-v_k)} \mu_{\alpha(v_k)}\,(dx).$$

Since $\tilde{\rho}(v_k) \le 1$ and $B^{(1)} \subset B^{(k)} \subset H^+(\alpha(v_k), v_k)$, this implies that

$$\mu(B^{(1)}) \le \mu_{\alpha_{v_k}}(B^{(1)}). \tag{4.8}$$

But if $v_{k'} \to v_0$ then $EX_{\alpha(v'_k)} \to v_0 \in \mathscr{S}$ and $E(u \cdot X_{\alpha(v'_k)}) \to u \cdot v_0$ for any unit vector u orthogonal to a supporting hyperplane T of $\mathscr{S}$ at v_0. Since $u \cdot v_0 \ge \text{ess sup}(u \cdot X_{\alpha(v'_k)})$, we must have $u \cdot X_{\alpha(v'_k)} \to_{\text{a.s.}} u \cdot v_0$, i.e. the limit distribution of $X_{\alpha(v'_k)}$ is singular and concentrated on T. But this means that $\mu_{\alpha(v_k)}(B^{(1)}) \to 0$, contradicting (4.8), since $B^{(1)}$ can be chosen so that $\mu(B^{(1)}) > 0$.

Thus, to complete the proof of existence of a dominating point we need only prove (4.7). If this is false, there must exist a subsequence $\{k'\}$ and $\delta > 0$ such that $\liminf \tilde{\rho}(v_{k'}) \ge \delta$, and a further subsequence $\{k''\}$ such that $v_{k''}/| v_{k''} | \to w$ (say). Also, for each $\{k''\}$ there will exist a small sphere $\mathscr{S}_{k''}$ interior to $\mathscr{S}$ such that $v_{k''}$ is a dominating point of $\mathscr{S}_{k''}$. Let u be a unit vector with $u \cdot w > 0$. There exists a sequence $j(k)$ such that $\mathscr{S}_{k''} \subset H^+(u, v_{j(k'')})$ and $u \cdot v_{j(k'')} \to \infty$ as $k'' \to \infty$. Now

$$\tilde{\rho}(v_{k''}) = \lim_{n\to\infty} P\{S_n \in n\mathscr{S}_{k''}\}^{1/n} \le \lim_{n\to\infty} P\{u \cdot S_n \ge n(u \cdot v_{j(k'')})\}^{1/n}. \tag{4.9}$$

But $u \cdot S_n = u \cdot X_1 + \cdots + u \cdot X_n$ is a sum of r.v.'s on $\mathbb{R}^1$ with entropy function $\tilde{\rho}_1(\cdot)$ (say), and in this one-dimensional case it is easily shown that $\tilde{\rho}_1(a) \to 0$ as $a \to \infty$. Hence the right sight of (4.9) can be made arbitrarily small for large n and k''. This proves (4.7) and the existence of a dominating point in In $\mathscr{S}$, and completes the proof of the theorem. □

5. Examples of asymptotic calculations. We will now illustrate how, with additional assumptions on the set B, one can use the formula (3.2) to get more precise information on the asymptotic behavior $P\{S_n \in nB\}$. It would be nice to have a systematic treatment of this matter in terms of the geometry of B, but here we limit ourselves to some examples. (See also J. Reeds [10]).

We start with (3.2), and again change variables by the rotation R as in Section 3. We assume through this section that this rotation has been carried out. Thus the integral

$$I_n = \int_{n\tilde{B}} e^{-\|\alpha\| x_1}\, \nu^{*n}\,(dx), \quad \tilde{B} = R(B - v_B) \tag{5.1}$$

must be estimated.

The finer estimates will always be slightly different in the lattice and nonlattice cases. We will here consider the nonlattice case, and will assume that μ (and hence ν) satisfies the

"Cramer condition"

$$\limsup_{|\alpha|\to\infty} |\varphi(i\alpha)| < 1.$$

The lattice case can be treated similarly. We classify the results according to the curvature of ∂B at v_B.

CASE (i): ∂B has "small curvature" near v_B in the sense that for some $\varepsilon > 0$, $0 < r < \infty$, and $0 \le \gamma < ½$,

$$[\tilde{B} \cap H(0, \varepsilon)] \supset [\Gamma(r, \gamma) \cap H(0, \varepsilon)],$$

where

$$H(0, \varepsilon) = \{x \in \mathbb{R}^d : 0 \le x_1 \le \varepsilon\}$$

and

$$\Gamma(r, \gamma) = \{x \in \mathbb{R}^d : (x_2^2 + \cdots x_d^2)^{1/2} \le r x_1^\gamma\}.$$

In this case $\mu^{*n}(nB)$ will be seen to behave similar to the 1-dimensional case, namely

$$\mu^{*n}(nB) = c\rho^n n^{-1/2}[1 + o(n^{-\delta})], \qquad \delta > 0. \tag{5.2}$$

Thus the behavior of μ^{*n} is in this case not very sensitive to the shape of B.

CASE (ii). B has "γ-order contact" with its tangent plane at v_B, in the sense that

$$\tilde{B} = \{x \in \mathbb{R}^d : x_2^2 + \cdots + x_d^2 \le r x_1^{2\gamma} + o(x_1^{2\gamma})\}$$

for some $½ < \gamma \le 1$, $0 < r < \infty$. In this case

$$\mu^{*n}(nB) = c\rho^n n^{(d-1)(1/2-\gamma)-1/2}[1 + O(n^{-\delta})], \ \delta > 0. \tag{5.3}$$

Thus in this case the behavior of μ^{*n} is sensitive to shape and dimension.

In the following calculations $\underset{\sim}{c}$, $\underset{\sim}{c}_i$, will always denote constants; not necessarily the same ones each time.

CASE (iii). If $\gamma = ½$ in either (i) or (ii) above then we get the weaker result

$$\mu^{*n}(nB) \overset{\cup}{\cap} n^{-1/2} \rho^n, \tag{5.4}$$

where $a_n \overset{\cup}{\cap} b_n$ means $\log a_n - \log b_n = O(1)$. To get a sharper estimate here, it seems that the actual shape of B would have to be more carefully taken into account.

To prove (5.2) observe that

$$0 \le \int_0^\infty e^{-\|\alpha\| x_1} \nu_1^{*n}(dx) - \int_{n\tilde{B}} e^{-\|\alpha\| x_1} \nu^{*n}(dx),$$

where ν_1^{*n} = the first marginal of ν^*,

$$\le \int_{n[\Gamma(r,\gamma)\cap H(0,\varepsilon)]^c} e^{-\|\alpha\| x_1} \nu^{*n}(dx)$$

$$\le \nu^{*n}\{n[\Gamma(r, \gamma) \cap H(0, \varepsilon)]^c\}$$

$$\le \nu^{*n}(n\Gamma^c(r, \gamma)) + \nu^{*n}(H(n\varepsilon, \infty)).$$

Now by the 1-dimensional theory [1],

$$\int_0^\infty e^{-\|\alpha\| x_1} \nu^{*n}(dx) = cn^{-1/2}[1 + o(n^{-\delta})],$$

while $\nu^{*n}(H(n\varepsilon, \infty))$ is exponentially small. Thus to prove (5.2) it is sufficient to show that

(5.5) $$\nu^{*n}(n\Gamma^c) = O(n^{-1/2-\delta}) \quad \text{for some } \delta > 0.$$

But $n\Gamma^c = \{x: (x_2^2 + \cdots + x_d^2)^{1/2} > rn^{1-\gamma}x_1^\gamma\}$, and hence for any $\Theta > 0$

$$\nu^{*n}(n\Gamma^c) \le \nu_1^{*n}(0, n^{-\Theta}) + \nu^{*n}\{x: (x_2^2 + \cdots + x_d^2)^{1/2} \ge rn^{1-\gamma-\gamma\Theta}\}.$$

Choose a small Θ, $0 < \Theta < ½$ so that $1 - \gamma - \gamma\Theta \equiv \eta > ½$. Then

(5.6i) $$\nu_1^{*n}(0, n^{-\Theta}) = P\{0 \le U_1 + \cdots + U_n \le n^{-\Theta}\}, \qquad \Theta > 0,$$

and

(5.6ii) $$\nu^{*n}((x_2^2 + \cdots + x_d^2)^{1/2} \ge rn^\eta) = P\{\| V_1 + \cdots + V_n \| \ge rn^\eta\}, \qquad \eta > ½,$$

where $\{U_i\}$ are i.i.d. 1-dimensional r.v.'s with $EU_i = 0$, $\{V_i\}$ are i.i.d. $(d-1)$-dimensional r.v.'s with $EV_i = 0$, and all moments of U_i and V_i exist. Both probabilities are of at most the order of magnitude required in (5.5), and hence (5.2) follows.

To prove (5.3), we make a change of variable $y_i = x_i/\sqrt{n}$ in (5.1), let $\lambda^{(n)}(dx) = \nu^{*n}(dx/\sqrt{n})$, and choose δ so that $(1-\gamma)/2\gamma < \delta < ½$, which is possible since $\gamma > ½$. We can then write

(5.7) $$I_n = \int_0^{n^{-\delta}} e^{-\|\alpha\|\sqrt{n}y_1} \int_{D_n} \lambda^{(n)}(dy) + O(e^{-\|\alpha\|n^{(1/2-\delta)}}),$$

where $D_n = \{y \in \mathbb{R}^d: y_2^2 + \cdots + y_d^2 \le [r + o(1)]n^{1-\gamma}y_1^{2\gamma}\}$. Applying Corollary 20.3 of [4], the main term in the estimate of (5.7) is

(5.8) $$\int_0^{n^{-\delta}} e^{-\|\alpha\|\sqrt{n}y_1} \int_{D_n} \phi_{o,\Sigma}(y)\, dy = \int_0^{n^{-\delta}} e^{-\|\alpha\|\sqrt{n}y_0} \int_{D_n} (2\pi|\Sigma|)^{-d/2}[1 + O(n^{1-\gamma-2\delta\gamma})]\, dy,$$

where $\phi_{o,\Sigma}$ is the normal density in $\mathbb{R}^d$ with O mean and covariance matrix Σ, and $\Sigma\,| =$ det Σ; and (5.7) =

$$= c[1 + O(n^{1-\gamma-2\gamma\delta})] \int_0^{n^{-\delta}} e^{-\|\alpha\|\sqrt{n}y_1}[n^{1-\gamma}y_1^{2\gamma}]^{(d-1)/2}\, dy_1.$$

Changing variables again ($\sqrt{n}\, y_1 = z$), we conclude that this

(5.9) $$= cn^{(d-1)(1/2-\gamma)-1/2}[1 + O(n^{-\delta_0})] \quad \text{for some} \quad \delta_0 > 0.$$

The remaining terms in the central limit expansion of (5.7) are like (5.8) times $O(n^{-1/2})$. This implies (5.3).

The proof of (5.4) goes like that of (5.2), but now we only get (5.5) with $\delta = 0$, and this weakens the result.

Acknowledgment. I would like to thank J. Cannon for some very helpful discussions on the proof of the theorem, S. Zabell for pointing out some relevant literature, and the referee for several suggestions.

REFERENCES

[1] Bahadur, R. R., and Rao, R. R. (1960). On deviations of the sample mean. *Ann. Math. Statist.* **31** 43–54.

[2] Bahadur, R. R., and Zabell, S. L. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probability* **7** 587–621.

[3] Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory.* Wiley, New York.

[4] BHATTACHARYA, R. N., and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

[5] BARTFAI, P. (1978). Large deviation of the sample mean in Euclidean spaces. *Purdue Univ. Statist. Dept. Mimeo Series No.* 78–13.

[6] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[7] CRAMÉR, H. (1938). Sur un nouveau théorème-limite de la théorie de probabilités. *Act. Sci. Ind.* **736** 5–23.

[8] FELLER, W. (1971). *An Introduction to Probability Theory and its Applications,* Vol. II 2nd Ed. Wiley, New York.

[9] NEY, P., and SPITZER, F. (1966). The Martin boundary for random walk. *Trans. Amer. Math. Soc.* **121** 116–132.

[10] REEDS, JAMES (1980). Correction terms for multinomial large deviations. *Asymptotic Theory of Statistical Tests and Estimation.* 287–305. Academic, New York.

MATHEMATICS DEPARTMENT
VAN VLECK HALL
480 LINCOLN DR.
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706