Diffusion Processes and Related Problems in Analysis, Volume I

Diffusions in Analysis and Geometry

Mark A. Pinsky
Editor

1990

Birkhäuser
Boston • Basel • Berlin

Large Deviations by the Asymptotic Value Method

Wlodzimierz Bryc

0. Introduction. In this paper we present an approach to large deviations which is based on the converse to Varadhan's "asymptotic integral theorem". We call it the asymptotic value approach to the large deviation principle. The method puts less emphasis on rate functions and on the underlying probability theory; in particular, "changes of measure" are not used in our proofs.The asymptotic value approach follows a pattern analogous to the weak convergence of measures: limits over suitable continuous functions replace lower and upper bounds for probabilities. We need to find limits $L(F) = \lim_{V\to\infty} L_V(F)$ for some bounded continuous functions F, knowing the limits $L(f)=\lim_{V\to\infty} L_V(f)$ for much simpler (e. g. linear) functions f. An inspection of the proofs of the theorems shows that few properties of the asymptotic value mapping $f\to L(f)$ are really used: without any reference to probability theory, properties expressed by (3.1)—(3.4) below are responsible for a (non—standard) rate function representation of the asymptotic value; to get the standard rate function representation, we use $L(f\vee g)=L(f)\vee L(g)$ and compactness of the "state space" (or good enough approximations by compact sets, such as the one in the conclusion of lemma L.4.1 below). This might broaden the scope of the large deviation method in applications to those asymptotic problems, where there is no evident probabilistic representation behind the formulas analyzed, compare e. g. van den Berg, Lewis & Pulé [3].

Throughout this paper we assume that the asymptotic value $L(.)$ arises in a probabilistic context, and is given by (1.3) below. We shall show that effective and useful criteria for the large deviation principle follow naturally from the asymptotic value approach. These theorems, when accompanied by theorem 3.1 of de Acosta [8], give short proofs of some non—trivial large deviation principles; the lower—upper bounds pattern of proof is replaced here by the following two steps: verification of "exponential tightness" (see the definition below), and showing that the limit (1.3) exists for a large enough class of functions.

Some aspects of the traditional approach to large deviations were not retained in the paper. We didn't attempt to establish criteria for large deviations "uniform with respect to a starting point", or to separate conditions responsible for upper bounds from those responsible for lower bounds. Both of those aspects of large deviations are very well understood in the context of Markov chains, c. f. de Acosta [10].

Section 1 contains statements of the main results. Theorems T.1.1 and T.1.2 show that if $\mathcal{L}(F)$ exists for all bounded continuous functions F, then it has a rate function representation (1.4) or (1.9) . We also state two related criteria for the large deviation principle. Theorem T.1.3 is our basic asymptotic value criterion; theorem T.1.4 extends results known in the literature, see the commentary preceding corollary C.2.2. Section 2 gives applications of the general theorems in specific situations. Several corollaries of theorems T.1.3 and T.1.4 are stated. Examples (all well known) at the end of section 2 show how the method can be applied and also illustrate the convenience of having several related criteria. In section 3, the non—standard rate function representation T.1.1 is proved. Section 4 contains the proof of T.1.2. In section 5, useful auxiliary results are obtained. In section 6, the main large deviation criteria T.1.3 and T.1.4 are proved. The method of proof is new and is based on theorem T.1.2 .

Other authors used variants of the asymptotic value method less explicitly or less generally, see Baldi [2], Comets [6], Dawson & Gärtner [7], Ellis [15], Gärtner [16], de Acosta [8], de Acosta [9], Kifer [17], Ney & Nummelin [19], Plachky [21], Sievers [24].

1. Notation and the main results. Let $\mathbb{X}$ be a metric space with a metric d(x, y) and the finitely additive Borel field $\mathcal{B}_{f.a.}$. By $\mathcal{P}_{f.a.}(\mathbb{X})$ we denote a complete metric space of all regular <u>finitely</u> additive probability measures on $(\mathbb{X}, \mathcal{B}_{f.a.})$ with the weak topology, $\mathcal{P}(\mathbb{X})$ stands for all <u>countably</u> additive probability measures on $(\mathbb{X}, \mathcal{B})$, where the Borel σ-field $\mathcal{B}$ is generated by $\mathcal{B}_{f.a.}$. By $C_b(\mathbb{X})$ we denote a Banach space of all bounded continuous functions F: $\mathbb{X}\to\mathbf{R}$ with the supremum norm. Through the paper $\{P_v\}_{v\in\mathcal{J}}$ is a family of probability measures, i. e. $P_v\in\mathcal{P}(\mathbb{X})$, $v\in\mathcal{J}$; $\mathcal{J}$ is a fixed unbounded (and not necessarily countable) subset of real numbers $v\geq 1$. To simplify the notation, we write $\{v\geq 1\}$ instead of $v\in\mathcal{J}$.

Following Varadhan [26], we say that $\{P_v\}$ satisfies the large deviation principle with a rate function I: $\mathbb{X}\to[0, \infty]$, if the following two conditions are satisfied

(1.1) $$-\inf\{ I(x): x\in A\}\leq \liminf_{v\to\infty} 1/v \log P_v(A)$$

for each open set $A \subset \mathbb{X}$;

(1.2) $$\limsup_{v\to\infty} 1/v \log P_v(A) \le -\inf\{I(x): x\in A\}$$

for each closed set $A \subset \mathbb{X}$.

We also require $I(.)$ to be lower semicontinuous and with compact level sets $I^{-1}([0, a])$ for each $a \ge 0$. (Then $I(.)$ is determined uniquelly, see e. g. Ellis [15] Theorem II. 3. 2 or Orey [20].)

The following definition is motivated by analogy with weak convergence of measures; the concept occurs explicitly in Deuschel & Stroock [11] and in Lynch & Sethuraman [18] (in the last paper under the name "large deviation tightness"); a number of authors used the same concept less explicitly, see Azencott [1], Baldi [2], de Acosta [8] Theorem 2. 1. (ii), Stroock [25] Theorem 3. 26.

Definition. We say that $\{P_v\}$ is $\mathcal{I}$-exponentially tight, if for each M>0 there exist a compact set $K \subset \mathbb{X}$ such that

$$\sup_{v\in\mathcal{I}} 1/v \log P_v(K^c) \le -M.$$

Note that exponential tightness depends both on the topology of $\mathbb{X}$ and on the index set $\mathcal{I}$. Since through this paper $\mathcal{I}$ is fixed, we shall suppress the $\mathcal{I}$-dependence in our terminology. Exponential tightness of probability measures on $\mathbf{R}^d$ is usually verified using exponential moments and the Chebyshev inequality. For probability measures on infinite dimensional spaces, de Acosta [8] theorem 3. 1 is helpful.

Definition. We say that a family $\{P_v\}_{v\ge1}$ admits an asymptotic value over a class $\mathfrak{F}$ of measurable functions, if

(1.3) $$L(F) = \lim_{v\to\infty} 1/v \log[\int_{\mathbb{X}} \exp\{vF(x)\}\, dP_v(x)]$$

exists and is finite for each function $F \in \mathfrak{F}$.

The following theorem shows that each asymptotic value over $C_b(\mathbb{X})$ has a "rate function representation" (1.4). The proof of (1.4) uses only properties of $L(.)$, listed in Lemma L.3.1 below.

Theorem T.1.1. If a family $\{P_v\}_{v\ge1}$ admits an asymptotic value $L(.)$ over $C_b(\mathbb{X})$, then there exists a lower semicontinuous function J: $\mathcal{P}_{f.a.}(\mathbb{X}) \to [0, \infty]$ such that

(1.4) $\mathcal{L}(F)=\sup\{\mu(F)-\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}(\mathbb{X})\}$,

and the supremum is attained. Furthermore the following variational expression holds

(1.5) $\mathcal{J}(\mu)=\sup\{\mu(F)-\mathcal{L}(F): F\in C_b(\mathbb{X})\}$.

Also

(1.6) $-\inf\{\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}, \mu(A)=1\}\leq \liminf_{v\to\infty} 1/v \log P_v(A)$

for each open set $A\subset \mathbb{X}$;

(1.7) $\limsup_{v\to\infty} 1/v \log P_v(A) \leq -\inf\{\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}, \mu(A)=1\}$

for each closed set $A\subset \mathbb{X}$.

[with the convention $\inf \emptyset=\infty$].

The following variant of T.1.1 gives a converse to Varadhan's [26] theorem 2. 2, and is the basis of our asymptotic value approach to large deviation principles. The proof is simpler than, and independent of, theorem T.1.1; the rate function $I(.)$ need not be convex.

Theorem T.1.2. If a family $\{P_v\}_{v\geq 1}$ admits an asymptotic value $\mathcal{L}(.)$ over $C_b(\mathbb{X})$ and is exponentially tight, then the large deviation principle holds with the rate function $I(x)=\mathcal{J}(\delta_x)$. In particular the following dual variational formulas hold

(1.8) $I(x)=\sup\{F(x)-\mathcal{L}(F): F\in C_b(\mathbb{X})\}$;

(1.9) $\mathcal{L}(F)=\sup\{F(x)-I(x): x\in\mathbb{X}\}$.

The next result is our main criterion for the large deviation principle. Its assumptions are rather technical, but additional flexibility is gained in the choice of the family $\mathcal{G}$. Corollaries in section 2 apply this criterion to a subset $\mathcal{G}$ of the concave functions, with (1.10), (1.11), and (1.13) either trivially satisfied or easily checked.

Recall that a family $\mathcal{G}$ of functions $\mathbb{X}\to\mathbf{R}$ separates points of $\mathbb{X}$, if

$\forall x, y\in\mathbb{X}\ \forall a, b\in \mathbf{R}\ \exists g\in \mathcal{G}$ such that $g(x)= a$ and $g(y)= b$.

Theorem T.1.3. Let $\{P_v\}$ be an exponentially tight family of probability measures. Suppose $\mathcal{G}$ is a subset of the space of all continuous (not necessarily bounded) functions $\mathbb{X}\to\mathbf{R}$, such that the following conditions are satisfied

(1.10) $\mathcal{G}$ separates points of $\mathbb{X}$;

(1.11) $\mathcal{G}$ contains the constant functions and is closed under finite pointwise minima, i. e. if $g_1, ..., g_n \in \mathcal{G}$, then $g_1 \wedge ... \wedge g_n \in \mathcal{G}$;

(1.12) $\{P_V\}$ admits an asymptotic value over $\mathcal{G}$.

Then $\{P_V\}$ satisfies the large deviation principle.

Moreover, if

(1.13) for each $g \in \mathcal{G}$ there is $0<q<1$ and a measurable function $\mathbf{B}$ such that $g \le q\mathbf{B}$ and

(1.14) $$\sup_V [\int_{\mathbb{X}} \exp\{v\mathbf{B}(x)\}\ dP_V(x)]^{1/V} < \infty,$$

then the rate function $I(.)$ is given by

(1.15) $$I(x) = \sup\{g(x) - L(g):\ g \in \mathcal{G}\}.$$

The following large deviation principle criterion is a non-trivial consequence of T.1.3.

Theorem T.1.4. Let $\{P_V\}$ be an exponentially tight family of probability measures. Suppose there is a subset $\mathcal{F}$ of the space of all continuous (not necessarily bounded) functions $\mathbb{X} \to \mathbf{R}$ such that the following conditions hold.

(1.16) $\mathcal{F}$ separates points of $\mathbb{X}$;

(1.17) $\mathcal{F}$ is linear (i. e. closed under finite linear combinations);

(1.18) $\{P_V\}$ admits an asymptotic value over $\mathcal{F}$;

(1.19) $\frac{d}{dt} L((1-t)f_1 + tf_2)\big|_{t=0}$ exists for each $f_1, f_2 \in \mathcal{F}$.

Then $\{P_V\}$ satisfies the large deviation principle with a rate function $I(.)$ given by

(1.20) $$I(x) = \sup\{f(x) - L(f):\ f \in \mathcal{F}\}.$$

Theorem T.1.1 is proved in section 3, theorem T.1.2 is proved in section 4, theorems T.1.3 and T.1.4 are proved in section 6.

2. Corollaries. In this section we list simple but useful consequences of theorems T.1.3 and T.1.4. The following corollary of T.1.3 specifies the family $\mathcal{G}$ and simplifies (1.15) for convex rate functions. The proof is given in section 6.

Corollary C.2.1. Let $\mathbb{V}$ be a locally convex Hausdorff topological linear space with the conjugate space $\mathbb{V}^*$. Suppose $\mathbb{X} \subset \mathbb{V}$ is a metric space in the relative topology, and $\Lambda \subset \mathbb{V}^*$ is a dense linear subspace. Define $\mathcal{G} = \{g:\ g(x) = \min_i\{\lambda_i(x) + c_i\},\ c_i \in \mathbf{R},\ \lambda_i \in \Lambda,\ 1 \le i \le n,\ n \in \mathbb{N}\}$.

Suppose $\{P_V\}$ is exponentially tight and admits an asymptotic value over $\mathcal{G}$.

Then $\{P_V\}$ satisfies the large deviation principle with a rate function $I(x)$ defined by (1.15).

If in addition $\sup_V 1/v \log[\int_{\mathbb{X}} \exp\{v\lambda(x)\}\, dP_V(x)]<\infty$ for each $\lambda\in\mathbb{V}^*$; $\mathbb{X}\subset\mathbb{V}$ is closed, convex and $I(.)$ defined by (1.15) is convex, then $L(\lambda)$ exists for each $\lambda\in\mathbb{V}^*$ and

(2.1) $$I(x)= \sup\{\lambda(x)-L(\lambda): \lambda\in\mathbb{V}^*\}.$$

Remark R.2.1. In a typical application, exponential tightness is verified with the help of de Acosta [8] theorem 3.1. The fact that $\{P_V\}$ admits an asymptotic value over $\mathcal{G}$ is then verified by showing that $v\to\log[\int_{\mathbb{X}} \exp\{v\lambda(x)\}\, dP_V(x)]$ is close to some super-additive function; the limit then exists by e.g. Dunford & Schwartz [13] VIII.1.4 (c. f. our example 2 for this type of argument).

The following corollary of T.1.4 generalizes to the infinite-dimensional setting Gärtner [16] lemmas 1. 1 and 1. 2; see also Dawson & Gärtner [7] Theorem 3. 4, de Acosta [8] Theorem 2. 1, Ellis [15] Theorem II. 6. 1, Plachky [21], Sievers [24] for related results. The result is especially easy to apply to sequences of probability measures obtained from an i. i. d sequence; see our example 1 below, Baldi [2], Dawson & Gärtner [7] (for an application of C.2.4 below). The corollary is proved in section 6.

Corollary C.2.2. Let $\mathbb{V}$ be a locally convex Hausdorff topological linear space with the conjugate space $\mathbb{V}^*$. Suppose $\mathbb{X}\subset\mathbb{V}$ is a metric space in the relative topology and is closed and convex. Let $\{P_V\}$ be an exponentially tight family of probability measures which admits an asymptotic value $L(\lambda)$ for each bounded continuous functional $\lambda\in\mathbb{V}^*$. Furthermore suppose that $L(\lambda)$ is Gateaux differentiable in each direction $\gamma\in\mathbb{V}^*$ and at each point $\lambda\in\mathbb{V}^*$. Then $\{P_V\}$ satisfies the large deviation principle with a rate function $I(.)$ defined by (2.1).

The assumptions of the next corollary eliminate the need to consider concave functions in (1.3) and (1.15) and avoid any explicit differentiability assumption. Similar result in the more specific context of empirical measures and with $\mathbb{V}^*$ replaced by a dense linear subspace of functionals has been recently presented by Kifer [17] (c.f. our C.2.6, which was added to the preliminary draft of this manuscript after seeing Kifer's paper).

Corollary C.2.3. Let $\mathbb{V}$ be a locally convex Hausdorff topological linear space with the conjugate space $\mathbb{V}^*$. Suppose closed and convex $\mathbb{X} \subset \mathbb{V}$ is a metric space in the relative topology. Let $\{P_V\}$ be an exponentially tight family of probability measures which admits an asymptotic value $L(\lambda)$ over $\lambda \in \mathbb{V}^*$. Also suppose that (2.1) defines a strictly convex function $I(.)$, and that $\sup_{\lambda \in \mathbb{O}} L(\lambda) < \infty$ for some open set $\mathbb{O} \subset \mathbb{V}^*$. Then $\{P_V\}$ satisfies the large deviation principle with the rate function $I(.)$.

Proof. The assumptions guarantee that $L(\lambda)$ is continuous, see e. g. Ekeland & Temam [14], p. 12, Proposition 2.5. To use Corollary C.2.2, it is enough to show that $L(\lambda)$ is Gateaux-differentiable at each point $\lambda \in \mathbb{V}^*$. For a finite dimensional vector space $\mathbb{V}$ the result is stated explicitly in Ruelle [22] p. 252. In general, the proof can be sketched as follows. By strict convexity of $I(.)$, the subgradient $\partial L(\lambda)$ is unique. Indeed, $x \in \partial L(\lambda)$ if and only if $L(\lambda)+I(x)=\lambda(x)$, see Ekeland & Temam [14], p. 21, Proposition 5. 1. Therefore $x \in \partial L(\lambda)$ is a unique minimum of the strictly convex function $I(x)-\lambda(x)$. Proposition 5. 2 of Ekeland & Temam [14], p. 23 ends the proof of Gateaux-differentiability of $L(.)$ and the result now follows from Corollary C.2.2.

The following corollary of C.2.2 is essentially a variant of theorem 3. 4 of Dawson & Gärtner [7]. We write it down here to show that it follows from C.2.2, even though the explicit assumption of exponential tightness is absent. Similar variants with no explicit smoothness assumptions can be deduced from C.2.3.

Corollary C.2.4. Let $\mathcal{V}$ be a vector space with a countable Hamel basis. Define $\mathbb{V}$ to be the algebraic dual $\mathcal{V}^a$ with $\sigma(\mathbb{V}, \mathcal{V})$-topology. Let $\mathbb{X}$ be a closed convex subset of $\mathbb{V}$, with the relative topology. Suppose $\{P_V\}$ is a family of probability measures on $(\mathbb{X}, \mathfrak{B})$ which admits an asymptotic value $L(\lambda)$ over all bounded continuous functionals $\lambda \in \mathcal{V}$. Furthermore suppose that $L(\lambda)$ is Gateaux-differentiable at each point $\lambda \in \mathcal{V}$. Then $\{P_V\}$ satisfies the large deviation principle with a rate function $I(x)$ defined by (2.1).

Proof. To apply C.2.2, we need only to verify that $\mathbb{X}$ is a metric space (which is actually an inessential assumption made throughout this paper for the sake of simplification) and that $\{P_V\}$ is exponentially tight. The first claim follows trivially from the fact that under our assumptions the topology of $\mathbb{V}$ is metrizable, see Dunford & Schwartz [13] V. 7. 34.

Exponential tightness follows from Chebyshev's inequality and Dunford & Schwartz [13] V. 4. 1 (use the assumption $|\mathcal{L}(\lambda)|<\infty$ for each fixed $\lambda\in\mathcal{V}$).

Remark R.2.2. As in Dawson & Gärtner [7] theorem 3. 4(iii), the assumption that $\mathbb{X}$ is a closed convex subset of $\mathbb{V}$, needed in C.2.1 and consequently in corollaries C.2.2-C.2.4 to get (2.1), can be replaced by what is actually used in the proof, i. e. by $\{x: I(x)<\infty\}\subset\mathbb{X}$.

The next three corollaries are direct applications of C.2.1 and C.2.2 to empirical distributions of a $\mathbb{Z}^d$-indexed random field. In this application $\mathbb{X}=\mathcal{P}(\mathbb{E})$ is a closed and convex subset of the locally convex Hausdorff topological vector space $\mathbb{V}$ of all signed measures on $\mathbb{E}$, with the topology of weak convergence. The space $\mathbb{E}$ will be either a given metric "state space" $\mathbb{F}$, or its product $\mathbb{F}^{\mathbb{Z}^d}$. The natural normalization corresponds to $\mathcal{J}=\{1, 2^d, 3^d, ...\}$; however we shall index empirical; measures by $n\in\mathbb{Z}$ and substitute n^d for ν in all other places.

Let $\mathbb{F}$ be a Polish space. Suppose $\{X_z\}_{z\in\mathbb{Z}^d}$ is an $\mathbb{F}$-valued random field. Define $\mathcal{P}(\mathbb{F})$-valued empirical distributions

$$\nu_n=n^{-d}\sum_{k\in\mathcal{C}_n}\delta_{\{X_k\}},$$

where $\mathcal{C}_n=\{z: 1\le z_j\le n\}$.

Define also $\mathcal{P}(\mathbb{F}^{\mathbb{Z}^d})$-valued empirical fields

$$\mu_n=n^{-d}\sum_{k\in\mathcal{C}_n}\delta_{\{X_{z+k}\}_{z\in\mathbb{Z}^d}}.$$

We shall say that empirical distributions/fields are exponentially tight, if the induced probability measures on $\mathcal{P}(\mathbb{E})$ are exponentially tight (after re-indexation $\nu=n^d$).

Corollary C.2.5. Let C_o be a dense linear subset of $C_b(\mathbb{F})$. Suppose for each $k\ge1$ and every $F_1, F_2, ..., F_k\in C_o$ there exists

$$\lim_{n\to\infty} n^{-d}\log(E\{\exp(\min_{1\le i\le k}\sum_{z\in\mathcal{C}_n}F_i(X_z))\})=\mathcal{L}(F_1, F_2, ..., F_k).$$

If the empirical measures $\{\nu_n\}$ are exponentially tight, then $\{\nu_n\}$ satisfies the large deviation principle with a rate function $I_o(.)$: $\mathcal{P}(\mathbb{F})\to[0,\infty]$ given by

$$I_o(p)=\sup\{p(F_1)\wedge p(F_2)\wedge...\wedge p(F_k)-\mathcal{L}(F_1, ..., F_k): F_1, F_2, ..., F_k\in C_o, k\in\mathbb{N}\}.$$

If in addition $I_o(.)$ is convex, then it can be identified by using k=1 only, i.e.

(2.2) $$I(p)= \sup\{p(F)-\mathcal{L}(F): F\in C_o\}.$$

Proof. This is essentially C.2.1 in the application-adjusted notation. (The integrability condition for (2.1) is here superfluous, since $X=\mathcal{P}(\mathbb{F})$ is bounded; by continuity, the supremum in (2.2) can be taken over $F\in C_0$ rather than over $F\in C_b(\mathbb{F})$, as would follow from (2.1).)

Remark R.2.3. Convexity of $I_0(.)$ can be verified whenever for each $p, q\in\mathcal{P}(\mathbb{F})$ and $F_1, F_2, ..., F_k\in C_0$ one can find $G_1, G_2, ..., G_k, H_1, H_2, ..., H_k\in C_0$ such that simultaneously we have $2L(F_1, ..., F_k)\geq L(G_1, ..., G_k)+L(H_1, ..., H_k)$ and $(p+q)(F_1)\wedge(p+q)(F_2)\wedge...\wedge(p+q)(F_k)=2p(G_1)\wedge p(G_2)\wedge...\wedge p(G_k)=2q(H_1)\wedge q(H_2)\wedge...\wedge p(H_k)$. In practical instances such G_i and H_i are obtained by adding suitable constants to functions F_i.

Corollary C.2.6. Let C_0 be a dense linear subset of $C_b(F)$. Suppose for each $k\geq 1$ and every $F\in C_0$ there exists

$$\lim_{n\to\infty} n^{-d}\log(E\{\exp(\Sigma_{z\in \mathcal{C}_n} F(X_z))\})=L(F).$$

Suppose furthermore that $\frac{d}{dt}L(F+tG)$ exists for each $F, G\in C_0$.

If the empirical measures $\{\nu_n\}$ are exponentially tight on $\mathcal{P}(\mathbb{F})$, then $\{\nu_n\}$ satisfies the large deviation principle with a rate function $I(.)$: $\mathcal{P}(\mathbb{F})\to[0,\infty]$ given by (2.2).

Proof. If $C_0=C_b(\mathbb{F})$, this is C.2.2 in the application adjusted notation. In the general case the large deviation principle is proved exactly as in C.2.2; then (2.2) is established by an argument used in the proof of (2.1), using the additional fact that our X is bounded in $\mathbb{V}$-norm. The detailed proof is omitted.

Corollary C.2.7 Let C_0 be any dense subset of $C_b(\mathbb{F}^{\mathbb{Z}^d})$. Put $Y_z=(X_{z+r})_{r\in\mathbb{Z}^d}\in\mathbb{F}^{\mathbb{Z}^d}$. Suppose for each $k\geq 1$ and every $F_1, F_2, ..., F_k\in C_0$ there exists

$$\lim_{n\to\infty} n^{-d}\log(E\{\exp(\min_{1\leq i\leq k}\Sigma_{z\in\mathcal{C}_n} F_i(Y_z))\})=L(F_1, F_2, ..., F_k).$$

If the empirical fields $\{\mu_n\}$ are exponentially tight, then $\{\mu_n\}$ satisfies the large deviation principle with a rate function $I_0(.)$: $\mathcal{P}(\mathbb{F}^{\mathbb{Z}^d})\to[0,\infty]$ given by

$$I_0(p)=\sup\{p(F_1)\wedge p(F_2)\wedge...\wedge p(F_k)-L(F_1, ..., F_k): F_1, F_2, ..., F_k\in C_0, k\in\mathbb{N}\}.$$

If in addition $I_0(.)$ is convex, then it can be identified by (2.2).

Proof. This is essentially C.2.1 in the application adjusted notation.

Remark R.2.4. In a typical application, C_0 consists of all those bounded continuous functions $F: \mathbb{F}^{\mathbb{Z}^d} \to \mathbf{R}$, which depend on finite number of coordinates only.

The following examples illustrate both the use of the asymptotic value method and the convenience of having several different criteria. We consider sums of random vectors only; similar examples with empirical measures, or "empirical processes" could have been presented equally easily. The first example is the large deviation principle for i. i. d. random vectors, see Donsker & Varadhan [12] theorem 5.3; the second example gives a result that seems to follow via the contraction principle from known large deviation results for the empirical measure of a Markov chain; the third example is synthetic and shows how non-convex rate functions can be handled by the asymptotic value method. The upper bounds in a related to our example 3, but more general setting, have been obtained in de Acosta [8] theorem 5. 1.

Example 1. Suppose $(\mathbb{V}, \|\,.\,\|)$ is a separable Banach space with the conjugate space $\mathbb{V}^*$. Let $\{X_i\}$ be a sequence of $\mathbb{V}$-valued i. i. d. random variables such that $E\{\exp(\alpha\|X_1\|)\}<\infty$ for each $\alpha\in\mathbf{R}$. Then $\{(X_1+\ldots+X_n)/n\}_{n=1, 2, \ldots}$ satisfies the large deviation principle with a rate function $I(v)=\sup_{\lambda\in\mathbb{V}^*}\{\lambda(v)-\log E\{\exp(\lambda(X))\}$.

This follows from C.2.2: By the independence assumption $L(\lambda)=\log E\{\exp(\lambda(X))\}$; hence $L(\lambda)$ exists for each $\lambda\in\mathbb{V}^*$ and is Gateaux-differentiable. Exponential tightness follows from theorem 3. 1 of de Acosta [8].

Example 2. Suppose $(\mathbb{V}, \|.\|)$ is a separable Banach space with the conjugate space $\mathbb{V}^*$. Let $\{X_n\}$ be a $\mathbb{V}$-valued stationary Markov chain such that $E\{\exp(\alpha\|X\|)\}<\infty$ for each $\alpha\in\mathbf{R}$. Let $\pi(dx)$ be a distribution of X_1. Suppose that there is $C<\infty$ such that 1-step transition probabilities $\Pi(\mathbf{x}, dy)$ satisfy

(2.3) $$\Pi(\mathbf{x}, A)\le C\Pi(\mathbf{y}, A) \text{ for } \pi\otimes\pi\text{-almost all } \mathbf{x}, \mathbf{y}\in\mathbb{V} \text{ and all Borel sets } A.$$

Then $\{(X_1+\ldots+X_n)/n\}_{n=1, 2, \ldots}$ satisfies the large deviation principle (with a convex rate function; below we omit the convexity argument).

This follows from C.2.1 applied to $\Lambda=\mathbb{V}^*$. Inequality (2.3) implies $E\{\exp(\sum_{i=1}^{n} q(X_i))\}\leq[CE\{\exp(q(X_1))\}]^n$ for any semi-norm q(.), hence theorem 3. 1 of de Acosta [8] guarantees exponential tightness. We shall verify that $\{(X_1+\ldots+X_n)/n\}$ admits an asymptotic value over the family $\mathcal{G}$, defined in C.2.1. Fix $g(.)\in\mathcal{G}$ and put $M_n=\pi\text{-ess inf } E_x\{\exp(g(X_1+\ldots+X_n))\}$. By the integrability assumption each M_n is finite. Also $M_{n+m}\geq M_nM_n$, since g(.) is concave. This shows that $n^{-1}\log M_n$ has a finite limit, see e. g. Dunford & Schwartz [13] VIII. 1. 4. It remains only to notice that by (2.3) $M_n\leq E\{\exp(g(X_1+\ldots+X_n))\}\leq CM_n$ which shows that $\mathcal{L}(g)$ exists.

Example 3. Suppose $\{X_k\}$ is an infinite [0, 1]-valued exchangeable sequence. If a tail σ-field $\bigcap_{n\geq 1}\sigma(X_n, X_{n+1}, \ldots)$ is finitely generated, then $\{(X_1+\ldots+X_n)/n\}_{n\geq 1}$ satisfies the large deviation principle (with, in general, a non-convex rate function).

This follows from C.2.1 applied to $\mathbb{V}=\Lambda=\mathbf{R}$ and the compact set $\mathbb{X}=[0, 1]$. By the de Finetti theorem, the distribution of $\{X_k\}$ is a mixture of product measures $\otimes_k\pi_\theta(dx_k)$ with a discrete mixing measure $\alpha(d\theta)$. Take a concave continuous function $g\in\mathcal{G}$. Then

$$n^{-1}\log E\{\exp(g(X_1+\ldots+X_n))\}=$$
$$n^{-1}\log\int\exp(g(x_1+\ldots+x_n)\pi_\theta(dx_1)\ldots.\ \pi_\theta(dx_n)\alpha(d\theta)\to\max_\theta L_\theta(g),$$

where the maximum is taken over a finite number of values $\theta\in$ supp α only, and $L_\theta(.)$ is an asymptotic value corresponding to the product measure $\otimes_k\pi_\theta(dx_k)$, $\theta\in$ supp α.

3. Proof of T.1.1. The following properties of $\mathcal{L}(.)$ follow immediately from (1.3) and the proof is omitted[1].

Lemma L.3.1. If a family $\{P_v\}_{v\geq 1}$ admits an asymptotic value $\mathcal{L}(.)$ over $C_b(\mathbb{X})$, then

(3. 1) $\mathcal{L}((F+G)/2)\leq\mathcal{L}(F)/2+\mathcal{L}(G)/2$

for each F, G $\in C_b(\mathbb{X})$;

(3. 2) $\inf_{x\in\mathbb{X}}[F(x)-G(x)]\leq\mathcal{L}(F)-\mathcal{L}(G)\leq\sup_{x\in\mathbb{X}}[F(x)-G(x)]$

for each F, G $\in C_b(\mathbb{X})$;

(3.3) $\mathcal{L}(F+\text{cons})=\mathcal{L}(F)+\text{const}$;

(3.4) $\mathcal{L}(0)=0$.

[1] (3.1) follows from Hölder's inequality.

Denote

(3.5) $\mathcal{L}_\nu(F) = 1/\nu \log[\int_{\mathbb{X}} \exp\{\nu F(x)\}\, dP_\nu(x)],\ \nu=1, 2, \ldots$

Lemma L.3.2. If $\mathcal{L}(g_1), \ldots, \mathcal{L}(g_n)$ exist for some measurable functions $g_i(x)$, $1\le i\le n$, then $\mathcal{L}(\max\{g_1, \ldots, g_n\})$ exists and

$$\mathcal{L}(\max\{g_1, \ldots, g_n\})= \max\{ \mathcal{L}(g_1), \ldots, \mathcal{L}(g_n)\}.$$

Proof. Since $\mathcal{L}_\nu(\max\{g_1, \ldots, g_n\})\ge \max\{ \mathcal{L}_\nu(g_1), \ldots, \mathcal{L}_\nu(g_n)\}$, we have

$$\liminf \mathcal{L}_\nu(\max\{g_1, \ldots, g_n\})\ge \max\{ \mathcal{L}(g_1), \ldots, \mathcal{L}(g_n)\}.$$

It remains to show that

(3.6) $\limsup \mathcal{L}_\nu(\max\{g_1, \ldots, g_n\})\le\max\{ \mathcal{L}(g_1), \ldots, \mathcal{L}(g_n)\}.$

Without losing generality we may assume that

$$\max\{ \mathcal{L}(g_1), \ldots, \mathcal{L}(g_n)\}=\mathcal{L}(g_1).$$

Fix $\varepsilon>0$ and let ν_o be such that for each $\nu>\nu_o$ and every $1\le i\le n$

$$1/\nu\log\int \exp\{\nu g_i(x)\}dP_\nu(x)\le \mathcal{L}(g_i)+\varepsilon.$$

Then $\int \exp\{\nu g_i(x)\}dP_\nu(x)\le e^{\nu\mathcal{L}(g_i)+\nu\varepsilon}\le e^{\nu\mathcal{L}(g_1)+\nu\varepsilon}$.

Therefore $\mathcal{L}_\nu(\max\{g_1, \ldots, g_n\})\le 1/\nu \log\sum_{i=1}^{n} \int \exp\{\nu g_i(x)\}dP_\nu(x)\le$

$1/\nu \log[ne^{\nu\mathcal{L}(g_1)+\nu\varepsilon}]\to\mathcal{L}(g_1)+\varepsilon$ as $\nu\to\infty$, which proves (3.6).

Note (to be used in the proof of T.1.3). The proof actually shows that

$$\limsup_{\nu\to\infty} \mathcal{L}_\nu(\max\{g_1, \ldots, g_n\})=$$

$$\max\{ \limsup_{\nu\to\infty} \mathcal{L}_\nu(g_1), \ldots, \limsup_{\nu\to\infty} \mathcal{L}_\nu(g_n)\}.$$

Proof of (1.4). Let $\mathcal{J}(.)$ be defined by (1.5) and fix $F_o\in C_b(\mathbb{X})$. By the definition of $\mathcal{J}(.)$, we need to show that

(3.7) $\mathcal{L}(F_o)= \sup_\mu \inf_F \{\mu(F_o) -\mu(F)+\mathcal{L}(F)\},$

where the supremum is taken over all $\mu\in \mathcal{P}_{f.a.}(\mathbb{X})$ and the infimum is taken over all $F\in C_b(\mathbb{X})$.

Moreover, since by (1.5) $\mathcal{J}(\mu)\ge\mu(F_o)-\mathcal{L}(F_o)$ for each $\mu\in \mathcal{P}_{f.a.}(\mathbb{X})$, therefore

$$\mathcal{L}(F_o)\ge\sup_\mu \inf_F \{\mu(F_o) -\mu(F)+\mathcal{L}(F)\}.$$

Hence to prove (3.7), it remains to show that there is $\sigma\in\mathcal{P}_{f.a.}(\mathbb{X})$ such that

(3.8) $\mathcal{L}(F_o)\le\sigma(F_o) -\sigma(F)+\mathcal{L}(F)$ for each $F\in C_b(\mathbb{X})$.

Also, for this σ, the supremum in (1.4) will be attained.

To find σ, define the following sets of functions. Let

$\mathfrak{M}=\{F\in C_b(\mathbb{X}): \inf_x [F(x)-F_0(x)]>0\}$ and let $\mathfrak{N}$ be a set of all finite convex combinations of functions g(x) of the form $g(x)= F(x)+ L(F_0)- L(F)$, where $F\in C_b(\mathbb{X})$. It is easily seen that $\mathfrak{M}$ and $\mathfrak{N}$ are convex; also $\mathfrak{M}$ is open and non-empty (e. g. $1+F_0\in\mathfrak{M}$). Furthermore $\mathfrak{M}$ and $\mathfrak{N}$ are disjoint. Indeed, take an arbitrary $\mathfrak{N}\ni g=\sum_{k=1}^{n} \alpha_k F_k+L(F_0)-\sum_{k=1}^{n} \alpha_k L(F_k)$.

Then $\inf_x\{g(x)-F_0(x)\}=\inf_x\{\sum_{k=1}^{n} \alpha_k F_k(x) -F_0(x)\}- \sum_{k=1}^{n} \alpha_k L(F_k)+ L(F_0)\leq$

$$\inf_x\{\sum_{k=1}^{n} \alpha_k F_k(x) -F_0(x)\} -L(\sum_{k=1}^{n} \alpha_k F_k)+ L(F_0)\leq 0,$$

where the first inequality follows from (3. 1) and the second follows from (3. 2) applied to $F=\sum_{k=1}^{n} \alpha_k F_k(x)$ and $G= F_0$. Therefore $g\notin \mathfrak{M}$.

Convex and open $\mathfrak{M}$ can be separated from disjoint convex $\mathfrak{N}$ by a linear functional, i. e. there is $0\neq f^*\in C_b{}^*(\mathbb{X})$ such that for some $\alpha\in \mathbf{R}$

(3.9) $$f^*(\mathfrak{N})\leq \alpha< f^*(\mathfrak{M}),$$

see e. g. Ekeland & Temam [14] Ch. 1, section 1. 2, or Dunford & Schwartz [13] V. 2. 8.

Claim: f^* is non-negative.

Indeed, it is easily seen that $F_0(.)$ belongs to $\mathfrak{N}$, and, as a limit of $\varepsilon+ F_0(x)$ as $\varepsilon\searrow 0$, F_0 belongs also to the closure of $\mathfrak{M}$. Therefore by (3.9) we have $\alpha=f^*(F_0)$. To end the proof, take a function F with $\inf_x F(x)>0$. Then $F_1= F+ F_0\in \mathfrak{M}$ and by (3.9)

$$f^*(F)= f^*(F_1- F_0)= f^*(F_1)- f^*(F_0)> \alpha- f^*(F_0)= 0.$$

This ends the proof of the claim.

Without loosing generality, we may assume $f^*(1)= 1$; then it is well known, see e. g. Bergström [5] Ch. 2 Section 4 theorem 1, that $f^*(F)= \sigma(F)$ for some $\sigma\in\mathcal{P}_{f.a.}(\mathbb{X})$ (Dunford & Schwartz [13] IV. 6. 2 implies that σ is regular). It remains to check that σ satisfies (3.8). To this end observe that since $F+L(F_0)- L(F)\in \mathfrak{N}$, by (3.9) we have

$\sigma(F)+ L(F_0)- L(F)\leq \alpha= \sigma(F_0)$ for every $F\in C_b(\mathbb{X})$. This ends the proof of (1.4).

Proof of the lower and upper bounds.

For $A\subset \mathbb{X}$ and $\varepsilon>0$ denote $[A]^\varepsilon=\{x: d(x, A)\leq\varepsilon\}:=\{x: \inf_{y\in A} d(x, y)\leq\varepsilon\}$. Clearly $[A]^\varepsilon$ is closed and $[A]^0$ is the closure of A.

Proposition P.3.1. If a family $\{P_\nu\}$ admits an asymptotic value $\mathcal{L}(.)$ over $C_b(\mathbb{X})$ and $\mathcal{J}(.)$ is defined by (1.5), then for any measurable set $A\subset\mathbb{X}$ we have

(3.10)) $$-\inf\{\mathcal{J}(\mu):\ \mu\in\mathcal{P}_{f.a.},\ \mu(\mathrm{int}(A))=1\}\le\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A);$$

(3.11) $$\limsup_{\nu\to\infty} 1/\nu\log P_\nu(A)\le-\lim_{\varepsilon\searrow 0}\inf\{\mathcal{J}(\mu):\ \mu\in\mathcal{P}_{f.a.},\ \mu([A]^\varepsilon)\ge 1-\varepsilon\}.$$

[with the convention $\inf\emptyset=\infty$].

Proof of (3.10). Fix an open set $A\subset\mathbb{X}$ and μ_o, such that $\mu_o(A)=1$. Since μ_o is regular, therefore for each $M\ge 1$ we may choose a closed set $C_M\subset A$ such that $\mu_o(C_M)\ge 1-1/M^2$. Let F_M: $\mathbb{X}\to[-M, 0]$ be a continuous function such that $F_M(x)=-M$ for $x\notin A$, $F_M(x)=0$ for $x\in C_M$. Then

$$1/\nu \log[\int_{\mathbb{X}} \exp\{\nu F_M(x)\}\ dP_\nu(x)]\le 1/\nu \log[e^{-\nu M}P_\nu(A^c)+P_\nu(A)].$$

Therefore

$$\mathcal{L}_\nu(F_M)\le 1/\nu \log 2+\max\{-M; 1/\nu \log P_\nu(A)\}.$$

Considering separately two cases: $\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)=-\infty$ and $\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)>-\infty$, we obtain

$$\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)\ge\liminf_{M\to\infty}\mathcal{L}(F_M).$$

Indeed, if $\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)=-\infty$, then $\mathcal{L}_\nu(F_M)\le 1/\nu \log 2-M$ for all large enough ν, so that $\liminf_{M\to\infty}\mathcal{L}(F_M)=-\infty$. And if $q=\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)>-\infty$, then $\mathcal{L}_\nu(F_M)\le 1/\nu \log(2)+1/\nu \log P_\nu(A)$ for all large enough ν and for each $M>-q$.

Therefore by (1.4)

$$\liminf_{\nu\to\infty} 1/\nu \log P_\nu(A)\ge\liminf_{M\to\infty}\sup_\mu\{\mu(F_M)-\mathcal{J}(\mu):\ \mu\in\mathcal{P}_{f.a.}(\mathbb{X})\}\ge$$
$$\liminf_{M\to\infty}\{\mu_o(F_M)-\mathcal{J}(\mu_o)\}.$$

It remains to notice that since $\mu_o(F_M)\ge-M\mu_o(C_M{}^c)\ge-1/M\to 0$, therefore $\liminf_{M\to\infty}\{\mu_o(F_M)-\mathcal{J}(\mu_o)\}\ge-\mathcal{J}(\mu_o)$. Since μ_o is an arbitrary element of $\mathcal{P}_{f.a.}(\mathbb{X})$ such that $\mu_o(A)=1$, this ends the proof of (1.6).

Proof of (3.11). Fix a closed set $A\subset\mathbb{X}$ and $\varepsilon>0$. Let F_M: $\mathbb{X}\to[-M, 0]$ be a continuous function such that $F_M(x)=-M$ for $x\notin[A]^\varepsilon$ and $F_M(x)=0$ for $x\in A$. Then

$$1/\nu \log[\int_{\mathbb{X}} \exp\{\nu F_M(x)\}\ dP_\nu(x)]\ge 1/\nu \log P_\nu(A).$$

Hence $\limsup_{V\to\infty} 1/V \log P_V(A) \leq$

$\inf_M \limsup_{V\to\infty} 1/V \log[\int_{\mathbb{X}} \exp\{VF_M(x)\}\, dP_V(x)]=\inf_M \mathcal{L}(F_M)$.

By (1.4) we obtain $\limsup_{V\to\infty} 1/V \log P_V(A) \leq$

$\inf_M\{\sup_\mu \{\mu(F_M)-\mathcal{J}(\mu): M>0, \mu\in \mathcal{P}_{f.a.}(\mathbb{X}), \mu([A]^\varepsilon)\geq 1-\varepsilon\} \vee$

$\sup_\mu \{\mu(F_M)-\mathcal{J}(\mu): M>0, \mu\in \mathcal{P}_{f.a.}(\mathbb{X}), \mu([A]^\varepsilon)<1-\varepsilon\}\} \leq$

$\sup_\mu \{-\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}(\mathbb{X}), \mu([A]^\varepsilon)\geq 1-\varepsilon\} \vee \inf_M \sup_\mu \{\mu(F_M): M>0, \mu\in \mathcal{P}_{f.a.}(\mathbb{X}),$

$\mu([A]^\varepsilon)<1-\varepsilon\}$.

Notice that if $\mu([A]^\varepsilon)<1-\varepsilon$, then $\mu(F_M)\leq -M\varepsilon\to -\infty$. Hence

$\limsup_{V\to\infty} 1/V \log P_V(A) \leq \sup_\mu \{-\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}(\mathbb{X}), \mu([A]^\varepsilon)\geq 1-\varepsilon\}=$

$-\inf_\mu \{-\mathcal{J}(\mu): \mu\in \mathcal{P}_{f.a.}(\mathbb{X}), \mu([A]^\varepsilon)\geq 1-\varepsilon\}$,

which ends the proof of (3.11).

Proof of (1.7). Since $\mathcal{P}_{f.a.}(\mathbb{X})$ is compact and $\mathcal{J}(.)$ is lower semicontinuous, by the standard subsequence argument we can pass in (3.11) to the limit as $\varepsilon\to 0$.

4 Proof of T.1.2.

Lemma L.4.1. If $\{P_V\}$ is an exponentially tight family of probability measures, then for each $M>0$ there is a compact set $K\subset\mathbb{X}$ with the following property.

If **h** is a measurable function such that $h(x)\leq M$ and $\limsup_{V\to\infty} \mathcal{L}_V(h)\geq 0$, then

$$\lim_{V\to\infty}[\mathcal{L}_V(h)- 1/V \log\int_K \exp\{Vh(x)\}\, dP_V(x)] = 0.$$

(Recall that $\mathcal{L}_V$ is defined by (3.5).)

Proof. Let K be a compact set such that $\sup_V 1/V \log P_V(K^c)\leq -2M$.

Since $\mathcal{L}_V(h)=$

$$1/V \log[\int_K \exp\{Vh(x)\}\, dP_V(x)+ \int_{K^c} \exp\{Vh(x)\}\, dP_V(x)]\leq$$

$$1/V \log 2 +1/V\log([\int_K \exp\{Vh(x)\}\, dP_V(x)]\vee \exp(VM)P_V(K^c)),$$

therefore

(4.1) $\mathcal{L}_V(h)\leq$

$$1/V \log 2+ 1/V\log[\int_K \exp\{Vh(x)\}\, dP_V(x)]\vee[M+\sup_V 1/V\log P_V(K^c)]\leq$$

$$1/V \log 2+ (-M)\vee 1/V\log[\int_K \exp\{Vh(x)\}\, dP_V(x)].$$

Since $M>0$ and $\limsup_{V\to\infty} \mathcal{L}_V(h)\geq 0$, from (4.1) we obtain

$$\limsup_{v\to\infty} L_v(h) \le \limsup_{v\to\infty} 1/v \log[\int_K \exp\{vh(x)\}\, dP_v(x).$$

This, together with the trivial inequality $1/v \log[\int_K \exp\{vh(x)\} dP_v(x)] \le L_v(h)$, ends the proof.

Proposition P.4.1. If $\{P_v\}$ admits an asymptotic value $L(.)$ over $C_b(\mathbb{X})$ and $I(.)$ satisfies (1.9), then (1.1) holds and

(4.2) $$\limsup_{v\to\infty} 1/v \log P_v(A) \le -\lim_{\varepsilon\searrow 0} \inf\{I(x): x \in [A]^\varepsilon\}.$$

[with the usual convention $\inf \emptyset = \infty$].

Proof. (4.2) is proved in the same way, as (3.11), except that (1.9) should be used in place of (1.4).

Proof of T.1.2.

As a trivial consequence of (1.8) we get

(4.3) $$L(F) \ge \sup\{F(x) - I(x): x \in \mathbb{X}\}.$$

To prove (1.9) pick a bounded continuous function F. By lemma L.4.1 there is a compact set K such that both $L(F)$ and $L(F\chi_K)$ exist and are equal; here χ_K denotes the indicator function of K. By (4.3) we need only to show that

(4.4) $$\limsup_{v\to\infty} L_v(F\chi_K) \le \sup\{F(x) - I(x): x \in K\}$$

for each compact set $K \subset \mathbb{X}$.

To prove (4.4) fix $F_o \in C_b(\mathbb{X})$ and $\varepsilon > 0$. Let $s = \sup\{F_o(x) - I(x): x \in K\}$. By (1.8) and (3.3), for each $x \in K$, there is $F_x \in C_b(\mathbb{X})$ such that $F_o(x) - F_x(x) < s + \varepsilon$ and $L(F_x) = 0$.

This means that open sets $U_x = \{y \in X: F_o(y) - F_x(y) < s + \varepsilon\}$ cover K, and we may choose a finite covering $U_{x(1)}, \ldots, U_{x(k)}$. Then

$$F(x) < \max_{1\le i\le k} F_{x(i)}(x) + s + \varepsilon,\ x \in K.$$

By (3. 2) and (3.3)

$$\limsup_{v\to\infty} L_v(F\chi_K) \le s + \varepsilon + \limsup_{v\to\infty} L_v(\max_{1\le i\le k} F_{x_i}(x)\chi_K) \le$$

$$s + \varepsilon + L(\max_{1\le i\le k} F_{x_i}(x)) = s + \varepsilon,$$

because Lemma L.3.2 gives $L(\max_{1\le i\le k} F_{x_i}(x)) = 0$. This concludes the proof of (1.9).

Since (1.9) has been proved, by P.4.1 we need only to show that $I(.)$ has compact level sets and

(4.5) $\lim_{\varepsilon\searrow 0}\inf\{I(x): x\in [A]^{\varepsilon}\}=\inf\{I(x): x\in A\}$,

where $I(.)$ is defined by (1.8) and A is a closed set. This is proved by the standard subsequence argument, since, by P.4.2, sets $A\cap I^{-1}[0, \lambda]$ are compact for each $\lambda\geq 0$.

Proposition P.4.2. If $\{P_v\}$ admits an asymptotic value $L(.)$ over $C_b(X)$ and $I(.)$ is defined by (1.8), then the following two conditions are equivalent:

(i) $\{P_v\}$ is exponentially tight;

(ii) $I^{-1}[0, \lambda]$ is a compact set for each $\lambda>0$.

Proof. Since (1.9) was proved, (1.1) follows by P.4.1 and the proof of Stroock [25] Theorem 3. 26 gives the implication (i)⇒(ii). For the implication (ii)⇒(i) see Lynch & Sethuraman [18] lemma 2. 6.

The following proposition complements T.1.1 and P.3.1.

Proposition P.4.3. If a family $\{P_v\}$ of probability measures admits an asymptotic value $L(.)$ over $C_b(X)$ and $I(x)=J(\delta_x)$ is defined by (1.8), then (1.1) holds and

(4.6) $\limsup_{v\to\infty} 1/v\log P_v(K)\leq -\inf\{I(x): x\in K\}$

for each compact set $K\subset X$.

Proof. Inequality (1.1) is a trivial consequence of (1.6). To prove (4.6) put F=0 in (4.4) to obtain

$$\limsup_{v\to\infty} 1/v\log P_v(K)=\limsup_{v\to\infty} L_v(F\chi_K)\leq$$
$$\sup\{F(x)-I(x): x\in K\}= -\inf\{I(x): x\in K\}.$$

5. Extension lemmas. The results of this section will be used to identify the rate function.

Lemma L.5.1. Suppose $\{P_v\}$ is tight. If **f** is a continuous function, then there exist a constant M>0 such that

(5.1) $\lim_{v\to\infty}[L_v(f)-L_v(f\vee(-M))]=0.$

(Recall that L_v is defined by (3.5).)

Proof. From the tightness assumption and continuity of **f** it follows that there are $0<\rho<1/2$ and M>0 such that $P_v(x: f(x)\geq -M)>\rho$. This implies that

$$\frac{P_v(x: f(x)\geq -M)}{P_v(x: f(x)<-M)}\geq \rho/(1-\rho).$$

Therefore for each $v\geq 1$

(5.2) $1/\nu \log P_\nu(x: f(x)\geq -M) \geq 1/\nu \log P_\nu(x: f(x)<-M)+1/\nu \log(\rho/(1-\rho))$.

Define $g(x)=f(x)\vee(-M)$. Since $f\leq g$, therefore $L_\nu(f)\leq L_\nu(g)$, which shows that $\limsup_{\nu\to\infty}[L_\nu(f)-L_\nu(g)]\leq 0$. To analyze the lower limit, observe that

(5.3) $L_\nu(g)=1/\nu \log[e^{-\nu M}P_\nu(x: f(x)<-M)+\int_{f\geq -M} \exp\{\nu f(x)\}\, dP_\nu(x)]\leq$

$1/\nu \log 2+\max\{-M+1/\nu \log P_\nu(f(x)<-M);\ 1/\nu \log\int_{f\geq -M} \exp\{\nu f(x)\}\, dP_\nu(x)\}$.

However, by (5.2) we have

$1/\nu \log\int_{f\geq -M}\exp\{\nu f(x)\}dP_\nu(x)\geq -M+1/\nu \log P_\nu(f(x)\geq -M)\geq$

$-M+1/\nu \log P_\nu(f(x)<-M)+1/\nu \log(\rho/(1-\rho))$,

and by our choice of ρ we have $\log(\rho/(1-\rho))\leq 0$. Therefore (5.3) implies

$L_\nu(g)\leq 1/\nu \log 2-1/\nu \log(\rho/(1-\rho))+1/\nu \log\int_{f\geq -M}\exp\{\nu f(x)\}dP_\nu(x)\leq$

$1/\nu \log 2-1/\nu \log(\rho/(1-\rho))+L_\nu(f)$.

This shows that $\liminf_{\nu\to\infty}[L_\nu(f)-L_\nu(g)]\geq 0$ and the lemma is proved.

Lemma L.5.2. Suppose a family $\{P_\nu\}$ of probability measures admits an asymptotic value $L(.)$ over $C_b(X)$ and is tight. Let $\mathbf{B}$: $X\to\mathbf{R}$ be a measurable function such that (1.14) holds. If f: $X\to\mathbf{R}$ is a continuous function such that for some $0<q<1$

(5.4) $f(x)\leq q\mathbf{B}(x)$,

then $L(f)$ exists. Moreover $L(f)$ is finite and there are constants M_0, N_0 such that $L(f)=L((M\vee f)\wedge N)$ for each $M\leq M_0$ and each $N\geq N_0$.

Proof. From Lemma L.5.1 it follows that for each continuous function f there is a constant $M>0$ such that if $g(x)=f(x)\vee(-M)$, then both $L_\nu(f)$ and $L_\nu(g)$ have the same lim sup and lim inf. Therefore it is enough to show that

$\liminf_{\nu\to\infty}L_\nu(g)=\limsup_{\nu\to\infty}L_\nu(g)$.

Define $H(x)=g(x)\wedge N$, where $N>0$ will be chosen later. Since $g\geq H$ and H is bounded and continuous, therefore $L(H)$ exists and $\liminf_{\nu\to\infty}L_\nu(g)\geq L(H)$.

We shall show that $\limsup_{\nu\to\infty}L_\nu(g)\leq L(H)$ for all large enough N. To this end observe that

(5.5) $L_\nu(g)\leq 1/\nu \log 2+$

$\max\{1/\nu\log\int_{g\geq N}\exp\{\nu g(x)\}dP_\nu(x);\ 1/\nu\log\int_{g<N}\exp\{\nu g(x)\}dP_\nu(x)\}$.

By the tightness assumption, there are $0<\rho<1/2$ and $N_0>0$ such that

$P_\nu(x\colon g(x)<N_0)>\rho$. Since $g\geq -M$, this implies that for $N>N_0$

$1/\nu \log\int_{g<N} \exp\{\nu g(x)\}dP_\nu(x)\geq 1/\nu \log[e^{-\nu M}P_\nu(g<N)]\geq -M+1/\nu\log \rho\geq -M+\log \rho$.

By Hölder's inequality we have $\int_{g\geq N} \exp\{\nu g(x)\}dP_\nu(x)\leq$

$(\int \exp\{\nu g(x)/q\}dP_\nu(x))^q(P_\nu(g\geq N))^{1-q}$. Hence using Chebyshev's inequality, (5.4) and (1.14) we obtain

$$\int_{g\geq N} \exp\{\nu g(x)\}\ dP_\nu(x)\leq$$
$$[\int_{\mathbb{X}} \exp\{\nu B(x)\}\ dP_\nu(x)]^q\ [e^{-\nu N/q}\int_{\mathbb{X}} \exp\{\nu g(x)/q\}\ dP_\nu(x)]^{1-q}\leq$$
$$[\int_{\mathbb{X}} \exp\{\nu B(x)\}\ dP_\nu(x)]\ e^{-\nu N(1-q)/q}.$$

Therefore

$$1/\nu \log\int_{g\geq N} \exp\{\nu g(x)\}dP_\nu(x)\leq C-N(1-q)/q,$$

where $C= \log \sup_{\nu\geq 1}[\int_{\mathbb{X}} \exp\{\nu B(x)\}\ dP_\nu(x)]^{1/\nu}$.

In particular $1/\nu \log\int_{g\geq N} \exp\{\nu g(x)\}dP_\nu(x)\leq -M+\log \rho$ for all N large enough. This shows that (5.5) amounts to

$$L_\nu(g)\leq 1/\nu \log 2+ 1/\nu \log\int_{g\leq N} \exp\{\nu g(x)\}\ dP_\nu(x)\leq 1/\nu \log 2+ L_\nu(H),$$

provided that $N>N_0$ and $N>(C+M-\log \rho)q(1-q)^{-1}$. Hence $\limsup_{\nu\to\infty} L_\nu(g) \leq L(H)$.

Proposition P.5.1. Suppose $\mathbb{X}$ is a metric subset of a linear space $\mathbb{V}$ with the dual $\mathbb{V}^*$. If $\{P_\nu\}$ admits an asymptotic value $L(.)$ over $C_b(\mathbb{X})$, $\{P_\nu\}$ is tight and

$\sup_\nu[\int_{\mathbb{X}} \exp(\nu\lambda(x))dP_\nu(x)]^{1/\nu}< \infty$ for each $\lambda\in \mathbb{V}^*$, then $L(\lambda)$ exists and is finite, $\lambda\in\mathbb{V}^*$.

Proof. Since $\exp\{|\lambda|\}\leq \exp(\lambda)+\exp(-\lambda)$, therefore $\sup_\nu[\int_{\mathbb{X}}\exp\{\nu 2|\lambda(x)|\}\ dP_\nu(x)]^{1/\nu}<\infty$ for each functional $\lambda\in\mathbb{V}^*$, and the result follows from L.5.2, applied to $f(.)=\lambda(.)$, $q=1/2$ and $B(x)=2|\lambda(x)|$.

The following result extends (1.9) to non-bounded continuous functions, compare de Acosta [9] Lemma 6. 1.

Proposition P.5.2. If for some continuous function **f** the assumptions of L.5.2 are satisfied and (1.9) holds for all $F\in C_b(\mathbb{X})$, then $L(f)=\sup_x\{f(x)-I(x)\}$.

Indeed, L.5.2 shows that one can find constants M, N such that $L(f)=L((M\vee f)\wedge N)$, and the equality is not affected by decreasing M, or increasing N. Take $M< L(f)$. By (1.9) applied to $L((M\vee f)\wedge N)$ we get

$L(f)= \sup_{x:f(x)>M}\{f(x)\wedge N-I(x)\}\vee \sup_{x:f(x)\leq M}\{M-I(x)\}$. Since $M< L(f)$ and $I(.)\geq 0$, this means that $L(f)= \sup_x\{N\wedge f(x)- I(x)\}$, for each large enough N. Therefore $L(f)= \sup_N \sup_x\{N\wedge f(x)-I(x)\}= \sup_x\{f(x)-I(x)\}$.

6. Proofs of the Large Deviation Principle Criteria. We begin with the two part proof of Theorem T.1.3: in Part A we show that the large deviation principle holds; in Part B we identify the rate function.

Proof of Theorem T.1.3. Part A (large deviation principle). Fix a bounded continuous function F. By T.1.2, it is enough to prove that $L(F)$ exists. By adding a constant, see (3.3), without loosing generality we may assume $0\leq F\leq M$ for some $M\geq 0$. Let $\varepsilon>0$ be fixed. We shall show that

(6.1) $$\lim\sup_{v\to\infty} L_v(F)\leq \lim\inf_{v\to\infty} L_v(F)+2\varepsilon.$$

Since $F\geq 0$, we have $\lim\inf_{v\to\infty} L_v(F)\geq 0$. Therefore from Lemma L.4.1 follows that we need only to show that

(6.2) $$\lim\sup_{v\to\infty} 1/v\log[\int_K \exp\{vF(x)\}dP_v(x)]\leq$$
$$\lim\inf_{v\to\infty} 1/v \log\int_K \exp\{vF(x)\}dP_v(x) + 2\varepsilon,$$

where K is a compact set from the conclusion of Lemma L.4.1.

By the Stone-Weierstrass theorem, see e. g. Schaefer [23] p. 243, there is a finite collection $\{g_i\}_{1\leq i\leq n}$ of functions in $\mathcal{G}$, such that

(6.3) $$\sup_{x\in K}|F(x)-\max_i g_i(x)|\leq\varepsilon.$$

Moreover, since $0\leq F(x)\leq M$, passing to $g_j\wedge M$ if necessary, we may assume $g_j\leq M$ for all j and we can also take $g_0=0$.

From (6.3) we obtain

$$\lim\sup_{v\to\infty} 1/v\log[\int_K \exp\{vF(x)\}dP_v(x)]\leq$$
$$\varepsilon+ \lim\sup_{v\to\infty} 1/v\log[\int_K \exp\{v[\max_i g_i(x)]\}dP_v(x)].$$

Since by Lemma L.3.2 (see the note following its proof)

$$\limsup_{v\to\infty} 1/v\log[\int_X \exp(v\max_i g_i(x))dP_v(x)] = \max_i \mathcal{L}(g_i(x)),$$

therefore

(6.4) $$\limsup_{v\to\infty} 1/v\log[\int_K \exp\{vF(x)\}dP_v(x)] \leq \varepsilon + \max_i \mathcal{L}(g_i(x)).$$

Using (6.3) once more we get

$$\liminf_{v\to\infty} 1/v\log[\int_K \exp\{vF(x)\}dP_v(x)] \geq$$

$$\liminf_{v\to\infty} 1/v\log[\int_K \exp\{v\max_i g_i(x)\}dP_v(x)] - \varepsilon.$$

Notice that since $\max_i \mathcal{L}(g_i(x)) \geq \mathcal{L}(g_0) = 0$, by Lemma L.4.1

$$\liminf_{v\to\infty} 1/v\log[\int_K \exp\{v\max_i g_i(x)\}\, dP_v(x)] = \max_i \mathcal{L}(g_i(x))$$

and hence

(6.5) $$\liminf_{v\to\infty} 1/v\log[\int_K \exp\{vF(x)\}dP_v(x)] \geq \max_i \mathcal{L}(g_i(x)) - \varepsilon.$$

Together (6.5) and (6.4) imply (6.2) and the large deviation principle is proved.

Proof of Theorem T.1.3. Part B (rate function identification).

Let $\vee\mathcal{G} = \{\max\{g_i(.)\}: g_i \in \mathcal{G},\ 1\leq i\leq n,\ n\geq 1\}$.

T.1.2 says that the rate function $I_0(.)$ is defined by (1.8). Therefore

(6.6) $I_0(x) \geq \sup\{h(x) - \mathcal{L}(h): h \in \vee\mathcal{G} \cap C_b(X)\}$.

Claim 1. $I_0(x) \leq \sup\{h(x) - \mathcal{L}(h): h \in \vee\mathcal{G} \cap C_b(X)\}$.

Indeed, let F be a continuous function, $0\leq F\leq M$ and let $\varepsilon>0$, $x_0 \in X$ be fixed. Let $K=K_M$ be a compact set from the conclusion of Lemma L.4.1, enlarged to ensure $x_0 \in K$. From part I of the proof we see that there is $h \in \vee\mathcal{G} \cap C_b(X)$, $h\geq 0$ such that

$|F(x) - h(x)| < \varepsilon$ at each $x \in K$;

$|\mathcal{L}(F) - \mathcal{L}(h)| < \varepsilon$.

Therefore

$F(x) - \mathcal{L}(F) \leq h(x) - \mathcal{L}(h) + 2\varepsilon$.

Since $\varepsilon>0$ is arbitrary, this concludes the proof of Claim 1.

Claim 2. $\sup\{h(x) - \mathcal{L}(h): h \in \vee\mathcal{G} \cap C_b(X)\} = \sup\{h(x) - \mathcal{L}(h): h \in \vee\mathcal{G}\}$

Indeed, inequality "$\leq$" is trivial. To prove "$\geq$" fix x_0 and $h = g_1 \vee \ldots \vee g_n \in \vee\mathcal{G}$. From L.5.2. applied to $f := h$, $B(x) := \max B_i(x)$ and $q = \max_i q_i$ [here B_i and q_i come from (1.13), applied to each g_i, $1\leq i\leq n$], one can find numbers M, N such that

(6.7) $\mathcal{L}(h) = \mathcal{L}((h\vee M)\wedge N)$.

(Lemma L.5.2 can be applied because the large deviation principle holds)

We can also decrease M and increase N so that (6.7) holds and $h(x_0)=(h(x_0)\vee M)\wedge N$. Then

$$h(x_0)-\mathcal{L}(h)=(h(x_0)\vee M)\wedge N-\mathcal{L}((h\vee M)\wedge N)\leq\sup\{h(x_0)-\mathcal{L}(h): h\in\vee\mathcal{G}\cap C_b(\mathbb{X})\}.$$

Since h and x_0 were arbitrary, this ends the proof of Claim 2.

Claim 3. $\sup\{h(x)-\mathcal{L}(h): h\in\vee\mathcal{G}\}=\sup\{g(x)-\mathcal{L}(g): g\in\mathcal{G}\}(=I(x))$.

Indeed, inequality "≥" is trivial. To prove "≤" observe that by (1.15) $\mathcal{L}(g)\geq \sup_x\{g(x)-I(x)\}$. Lemma L.3.2 implies that $\mathcal{L}(\max_i g_i)\geq \max_i \sup_x\{g_i(x)-I(x)\}$. Therefore $\sup\{h(x)-\mathcal{L}(h): h\in\vee\mathcal{G}\}=\sup_{g_i}\{\max_i g_i(x)-\mathcal{L}(\max_i g_i)\}\leq$

$\sup_{g_i}\inf_y\{\max_i g_i(x)-\max_i g_i(y)+I(y)\}\leq I(x)$, where the last inequality was obtained by taking $y=x$. This proves Claim 3.

Claims 1, 2 and 3 together with (6.6) end the proof of (1.15).

Proof of C.2.1. By the Hahn-Banach theorem, $\mathbb{V}^*$ separates points of $\mathbb{X}$, see e. g. Ekeland & Temam [14] Ch. 1. page 5 Corollary 1. 2. Since Λ is dense and linear, it separates points of $\mathbb{X}$, too, and all the assumptions of Theorem T.1.3 hold, see the proof of Proposition P.5.1 for T.1.3 (1.13). Therefore the large deviation principle holds with a rate function $I_0(.)$ given by (1.15). It remains to show that if $\mathbb{X}$ is a convex closed subset of V and $I_0(.)$ is convex, then $I_0(.)$ is defined by (2.1).

Since for each $\lambda\in\mathbb{V}^*$ the assumptions of L.5.2 are satisfied (take e.g. $q=1/2$, $B=2|\lambda|$), by Proposition P.5.2 the large deviation principle implies $\mathcal{L}(\lambda)=\sup\{\lambda(x)-I_0(x): x\in\mathbb{X}\}$. However, since $\mathbb{X}$ is closed and convex, (1.15) gives $I_0(v)=\infty$ at each $v\notin\mathbb{X}$. Therefore $\mathcal{L}(\lambda)=\sup_{v\in\mathbb{V}}\{\lambda(v)-I_0(v)\}$ for every $\lambda\in\mathbb{V}^*$ and the well known result on the bi-conjugate of a convex lower semicontinuous function, see e. g. Ekeland & Temam [14] Ch. 1, gives (2.1).

Proof of T.1.4.

The following lemma is the main step in the proof of T.1.4.

Lemma L.6.1. Fix $n\geq1$. Let measurable functions $f_i(x)$, $1\leq i\leq n$, be such that

(i) $\mathcal{L}(f)$ exists and is a finite number for each linear combination $f=\sum_{i=1}^n \alpha_i f_i$;

(ii) $\frac{d}{dt}\mathcal{L}((1-t)f+tg)_{|t=0}$ exists for each pair of convex combinations $f=\sum_{i=1}^n \alpha_i f_i$, $g=\sum_{i=1}^n \beta_i f_i$.

Then $L(\min\{f_1, ..., f_n\})$ exists and

(6.8) $L(\min\{f_1, ..., f_n\}) = \inf\{L(\alpha_1 f_1 + ... + \alpha_n f_n) : \alpha_i \geq 0, \alpha_1 + ... + \alpha_n = 1\}$.

Proof. Since $(\alpha_1, ..., \alpha_n) \to L(\alpha_1 f_1 + ... + \alpha_n f_n)$ is continuous (as a convex and bounded function), therefore the infimum on the right hand side of (6.8) is attained. Suppose the infimum is attained at the point $f_0 = \sum_{i=1}^{n} \alpha_i f_i$, where $\alpha_i \geq 0$, $i = 1, 2, ..., n$, and $\sum_{i=1}^{n} \alpha_i = 1$.

Then $f_1 \wedge ... \wedge f_n = f_0 \wedge f_1 \wedge ... \wedge f_n$ and we need to show that

(6.9) $$L(f_0) = \inf\{L(\alpha_0 f_0 + \alpha_1 f_1 + ... + \alpha_n f_n): \alpha_i \geq 0, \forall\, i, \alpha_0 + ... + \alpha_n = 1\},$$

implies $L(f_0 \wedge f_1 \wedge ... \wedge f_n) = L(f_0)$.

Notice that since $f_0 \wedge f_1 \wedge ... \wedge f_n \leq f_0$, therefore $\limsup_{V\to\infty} L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) \leq L(f_0)$, and to prove $L(f_0 \wedge f_1 \wedge ... \wedge f_n) = L(f_0)$ we need only to show that

(6.10) $$\liminf_{V\to\infty} L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) \geq L(f_0).$$

To end the proof fix $\varepsilon > 0$ and let $\theta = \varepsilon/(1+\varepsilon) \in (0, 1)$. Obviously we have

(6.11) $$f_0 = \theta(f_0 \wedge f_1 \wedge ... \wedge f_n) + (1-\theta)\max_{0\leq k\leq n}\{(1+\varepsilon)f_0 - \varepsilon f_k\}.$$

Using the fact that each $L_V(.)$ is non-decreasing and convex, compare L.3.1, from (6.11) we obtain

(6.12) $$L_V(f_0) \leq \theta L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) + (1-\theta)L_V(\max_{0\leq k\leq n}\{(1+\varepsilon)f_0 - \varepsilon f_k\}).$$

This by Lemma L.3.2 implies

(6.13) $$L(f_0) \leq \theta \liminf_{V\to\infty} L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) + (1-\theta)\max_{0\leq k\leq n} L((1+\varepsilon)f_0 - \varepsilon f_k).$$

We shall deduce (6.10) from (6.13) by considering separately the following two cases.

Case 1. There is $\varepsilon > 0$ such that $\max_{1\leq k\leq n} L((1+\varepsilon)f_0 - \varepsilon f_k) \leq L(f_0)$.

In this case (6.13) implies $L(f_0) \leq \theta \liminf_V L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) + (1-\theta)L(f_0)$, and (6.10) follows, since $\theta > 0$.

Case 2. $\max_{1\leq k\leq n} L((1+\varepsilon)f_0 - \varepsilon f_k) > L(f_0)$ for each $\varepsilon > 0$.

In this case there is an index k ($1 \leq k \leq n$) such that

$$L((1+\varepsilon_r)f_0 - \varepsilon_r f_k) = \max_{1\leq i\leq n} L((1+\varepsilon_r)f_0 - \varepsilon_r f_i) = \max_{0\leq i\leq n} L((1+\varepsilon_r)f_0 - \varepsilon_r f_i)$$

for the infinite number of values $\varepsilon_r > 0$, $r = 1, 2, ...$ and without losing generality we may assume $\varepsilon_r \to 0$. Then (6.13) implies

$$L(f_0) \leq \theta \liminf_V L_V(f_0 \wedge f_1 \wedge ... \wedge f_n) + (1-\theta)L((1+\varepsilon_r)f_0 - \varepsilon_r f_k).$$

Since $\theta = \varepsilon/(1+\varepsilon)$, this in turn implies that

$$\liminf_{V\to\infty} L_V(f_0\wedge f_1\wedge\dots\wedge f_n)-L(f_0)\geq[L(f_0)-L((1+\varepsilon_r)f_0-\varepsilon_r f_k)]/\varepsilon_r.$$

Passing to the limit as $r\to\infty$, and after taking into account the differentiability assumption, we obtain

$$\liminf_{V\to\infty} L_V(f_0\wedge f_1\wedge\dots\wedge f_n)-L(f_0)\geq$$

$$\lim_{r\to\infty}[L(f_0)-L((1+\varepsilon_r)f_0-\varepsilon_r f_k)]/\varepsilon_r=\partial/\partial\varepsilon\, L((1-\varepsilon)f_0+\varepsilon f_k)_{|\varepsilon=0}\geq 0,$$

where the last inequality follows from the assumption that $L(f_0)\leq L((1-\varepsilon)f_0+\varepsilon f_k)$ for each $\varepsilon>0$, see (6.9).

Proof of Theorem T.1.4. Without loosing generality we may assume that $\mathfrak{F}$ contains the constant functions, see (3.3). Consider

$$\mathcal{G}=\wedge\mathfrak{F}=\{g: g(x)=f_1(x)\wedge\dots\wedge f_n(x),\ n\geq 1,\ f_i\in\mathfrak{F}\}.$$

It is easy to check that $\wedge f$ satisfies the assumptions of Theorem T.1.3. Indeed, T.1.3. (1.13) holds with q=2, $\mathbf{B}=2(f\vee 0)$ [condition (1.14) for **B** is checked similarly as in the proof of Proposition P.5.1]; T.1.3. (1.10) is assumed for $\mathfrak{F}$ so it holds for $\wedge\mathfrak{F}$, too; T.1.3. (1.12) holds by Lemma L.6.1 ; T.1.3. (1.11) holds by the definition of $\wedge\mathfrak{F}$.

Therefore by Theorem T.1.3. the large deviation principle holds with a rate function

(6.14). $$I_0(x)=\sup\{g(x)-L(g): g\in\wedge\mathfrak{F}\}.$$

Let $I(.)$ be defined by (1.20). Since the supremum in (1.20) is taken over the smaller set, therefore $I(.)\leq I_0(.)$. To prove the converse inequality, take $g=\min_{1\leq j\leq n} f_j\in\wedge\mathfrak{F}$. By (6.8), there are numbers $\alpha_j\geq 0$, j=1, 2, ..., n, $\sum_{j=1}^n\alpha_j=1$, such that $g(x)-L(g)=g(x)-L(\sum_{j=1}^n\alpha_j f_j)$. Since $\min_j f_j\leq\sum_{j=1}^n\alpha_j f_j$, therefore this implies $g(x)-L(g)\leq\sum_{j=1}^n\alpha_j f_j(x)-L(\sum_{j=1}^n\alpha_j f_j)\leq I(x)$.

Since $g\in\wedge\mathfrak{F}$ was arbitrary, this ends the proof.

Proof of C.2.2. We shall show that C.2.1. can be applied. By Lemma L.6.1 $L(g)$ is defined for each $g=f_1\wedge\dots\wedge f_n$, where $f_i(x)=\lambda_i(x)+c_i$, $c_i\in\mathbf{R}$, $\lambda_i\in\mathbb{V}^*$, $1\leq i\leq n$, $n\in\mathbb{N}$. To check that (1.15) defines a convex function, we use Lemma L.6.1 again. Indeed, by (6.8) $L(g)=L(\sum_{k=1}^n\alpha_k f_k)$ for some $\alpha_k\geq 0$, $\sum_{k=1}^n\alpha_k=1$. Since trivially $g(x)\leq\sum_{k=1}^n\alpha_k f_k(x)$, therefore, since each $f(.)$ is linear, we get

$$g(\mu(x))-L(g)\leq[\sum_{k=1}^n\alpha_k f_k](\mu(x))-L(\sum_{k=1}^n\alpha_k f_k)=$$

$$\mu([\sum_{k=1}^{n} \alpha_k f_k](x)) - L(\sum_{k=1}^{n} \alpha_k f_k) \leq \sup_f\{\mu(f) - L(f)\} \leq \mu(I(x))$$

for each (say discrete) measure μ. Since **g** was arbitrary, this shows $I(\mu(x)) \leq \mu(I(x))$, i. e. $I(.)$ is convex.

Acknowledgement. The subject of this research was suggested by conversations with A. Korzeniowski and W. Smolenski. The author also would like to thank A. de Acosta, Y. Kifer, S. Pelikan and W. Smolenski for references and/or preprints.

REFERENCES

[1] R. Azencott, *Grandes deviations et applications*, Lecture Notes in Math. 774 (1980), Springer, Berlin.

[2] P. Baldi, Large Deviations and Stochastic Homogenization, Annali di Matematica Pura et Applicata 151 (1988), pp. 161-177.

[3] M. van den Berg, J. T. Lewis & J. V. Pulé, The Large Deviation Principle and Some Models of an Interacting Boson Gas, Comm. Math. Physics 118 (1988), pp. 61-85.

[5] H. Bergström, *Weak convergence of measures*, Academic Press, New York, 1982.

[6] F. Comets, Large Deviation Estimates for a Conditional Probability Distribution. Application to Random Interaction Gibbs Measures, Probab. Th. Rel. Fields 80 (1989), pp. 407-432.

[7] D. A. Dawson, J. Gärtner, Large Deviations from the McKean-Vlasov Limit for Weakly Interacting Diffusions, Stochastics 20 (1987), pp. 247-308.

[8] A. de Acosta, Upper bounds for Large Deviations of Dependent Random Vectors, Zeitsch. Wahrsch. verw. Gebiete 69 (1985), pp. 551-565.

[9] A. de Acosta, Large Deviations for vector-valued functionals of a Markov chain: lower bounds, Ann. Probab. 16 (1988), pp. 925-960.

[10] A. de Acosta, Large deviations for empirical measures of Markov chains, Case Western Reserve Univ. preprint #88-101 (1988).

[11] J. Deuschel & D. W. Stroock, *Large Deviations*, Academic Press, Boston, 1989.

[12] M. D. Donsker, S. R. S. Varadhan, Asymptotic Evaluation of Certain Markov Process Expectations for Large Time -III, Comm. Pure Appl. Math. 29 (1976), pp. 389-461.

[13] N. Dunford, J. T. Schwartz, *Linear Operators I*, Interscience, New York, 1958.

[14] I. Ekeland, R. Temam, *Convex analysis and variational problems*, North Holland, Amsterdam, 1976.

[15] R. Ellis, Entropy, *Large Deviations and Statistical Mechanics*, Springer-Verlag, New York, 1985.

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

[17] Y. Kifer, Large Deviations in Dynamical Systems and Stochastic Processes, preprint (1989).

[18] J. Lynch, J. Sethuraman, Large deviations for processes with independent increments, Ann. Probab. 15 (1987), pp. 610-627.

[19] P. Ney, Nummelin, Markov additive process I. Eigenvalue properties and limit theorems, Ann. Probab. 15 (1987), pp. 561-592, Markov additive process II pp. 593-609.

[20] S. Orey, Large Deviations in ergodic theory, In: Seminar in Stochastic Processes 1984, E. Cinclar, K. L. Chung, R. K. Getoor ed., Birhhäuser, Boston, 1985.

[21] D. Plachky, On a theorem by G. L. Sievers, Ann. Math. Stat. 42 (1971), pp. 1442-1443.

[22] D. Ruelle, *Thermodynamic Formalism*, Adison-Wesley, London, 1978.

[23] H. Schaefer, *Topological Vector Spaces*, Springer-Verlag, New York, 1971.

[24] G. L. Sievers, On the probability of large deviations and exact slopes, Ann. Math. Stat. 40 (1969), pp. 1908-1921.

[25] D. W. Stroock, *An introduction to the theory of large deviations*, Springer, New York, 1984.

[26] S. Varadhan, *Large Deviations and Applications*, SIAM, Philadelphia, 1984.

Wlodzimierz Bryc
Department of Mathematics
University of Cincinnati
Cincinnati, OH 45221-0025