LARGE DEVIATIONS PRINCIPLES FOR STATIONARY PROCESSES

BY

STEVEN OREY

AND

STEPHEN PELIKAN

IMA Preprint Series # 227
March 1986
LARGE DEVIATIONS PRINCIPLES FOR STATIONARY PROCESSES

by

Steven Orey

and

Stephen Pelikan

Consider a discrete time stationary process \((\omega_n)\) with \(n\) varying over the integers \(\mathbb{Z}\) and \(\omega_n\) taking values in a complete separable metric space \(M\). Then \(\omega = (\omega_n) \in M^\mathbb{Z} := \Omega\), and using the product topology \(\Omega\) is again metrizable as a complete separable metric space. Let \(\mathcal{M}(\Omega)\) or \(\mathcal{M}\) denote the class of probability measures on the Borel sets at \(\Omega\). Let \(\theta: \Omega \to \Omega\) be the shift \((\theta\omega)_n = \omega_{n+1}\). Let \(\mathcal{M}_\theta\) be the class of \(\mu \in \mathcal{M}\) such that \(\mu = \mu \circ \theta^{-1}\), that is the class of stationary measures, determining stationary stochastic processes.

For \(\omega \in \Omega\) let \(\delta_\omega\) be the element in \(\mathcal{M}\) assigning mass one to \(\{\omega\}\), and let \(\epsilon_n^\omega\) be the \(\mathcal{M}\)-valued random variable whose value at \(\omega\) is given by \(\epsilon_n^\omega = n^{-1} \sum_{k=0}^{n-1} \delta_{\theta^k\omega}\). Note that for a Borel subset \(\Lambda\) of \(\Omega\), \(\epsilon_n^\Lambda = \sum_{k=0}^{n-1} \theta^k \in \Lambda\) represents \(n^{-1}\) times the number of \(k\) between 0 and \(n-1\) such that \(\theta^k \omega \in \Lambda\).

Given \(\mu \in \mathcal{M}_\theta\), the random variable \(\epsilon_n\) has a distribution function \(\mu \circ \epsilon_n^{-1}\), and our concern is with the validity of a large deviations principle and the identification of the corresponding rate function. Here we use the familiar terminology of [10]. Precise definitions are given in the body of the paper.

The pioneering work on this question is Donsker and Varadhan [2], where this problem was discussed (actually for continuous time) in case the underlying shift is Markovian with transition probabilities having good continuity and mixing properties. Extensions to some non-Markovian situations are included in [4].

In the present paper we always assume \(M\) to be compact. In Section 1 a certain mixing condition (RM) is introduced, and it is shown that the condition alone suffices for the validity of a uniform large deviation principle. Actually we work with a class of processes slightly wider than the stationary ones. The condition (RM) does not appear to be comparable with other familiar
mixing conditions in ergodic theory. Section 2 briefly recapitulates some results from [4]. If (RM) is supplemented by a condition (CD) the rate function in the large deviation principle can be identified as a relative entropy as in [2]. In Section 3 a certain class of Gibbs measures is considered (the same class studied in Bowen [1]), it is shown that (RM) and (CD) hold, and an explicit identification for the rate function is given. Finally in Section 4 some examples and counterexamples are given, and connections with the work of Takahashi [8], [9] are given.

In Bowen [1] the theory of Gibbs states is developed with the motivation of applying it to the symbolic dynamics of Anosov diffeomorphisms. We also have such applications in mind; these will be given in a separate paper.

For stationary Gaussian sequences the large deviations principle has been beautifully worked out in Donsker and Varadhan [3].

1. A large deviations principle. For every integer \( n \) let 
\[
(M_n, B_n) = (M, B),
\]
where \( M \) is a compact separable metric space and \( B \) the corresponding Borel field. Let
\[
(\Omega, \mathcal{F}) = \prod_{n=-\infty}^{\infty} (M_n, B_n)
\]
be the product space, endowed with the product topology. Let \( \mathcal{M}(\Omega) \) denote the class of probability measures on \( (M, \mathcal{F}) \), with the topology of weak convergence. Usually we write \( \mathcal{M} \) for \( \mathcal{M}(\Omega) \). Then \( \Omega \) and \( \mathcal{M} \) are also compact separable metric spaces. The shift operator \( \theta: \Omega \to \Omega \) is defined by \( (\theta \omega)_i = \omega_{i+1} \) for every integer \( i \). If \( \omega \in \Omega \) \( \omega^- \) denotes the one sided sequence \( (\ldots, \omega_{-1}, \omega_0) \), and let \( \Omega^- = \{ \omega^-: \omega \in \Omega \} \). For any \( \mu \in \mathcal{M} \) the system \( (\Omega, \mathcal{F}, \theta, \mu) \) is called a shift. Let \( \mathcal{M}_0 \) be the class of \( \mu \in \mathcal{M} \) such that \( \mu = \mu \circ \theta^{-1} \). A \( \mu \in \mathcal{M}_0 \) is called a stationary measure and the corresponding shift is a stationary shift. For \( -\infty < m < n < \infty \) let \( \mathcal{F}_{m,n} = \prod_{i=m}^{n} B_i \) and put \( \mathcal{F}_m = \mathcal{F}_{m,m} \). For \( \mu \in \mathcal{M} \), \( \mu_{m,n} \) will denote the restriction of \( \mu \) to \( \mathcal{F}_{m,n} \), and \( \mu_m = \mu_{m,m} \). The
measure \( \mu \in \mathcal{M} \) is **homogeneous** if for every \( A \in \mathcal{F} \) and \( m < n \)

\[
\mu[A | \mathcal{F}_{m,n}](\omega) = \mu[\theta^{-1}A | \mathcal{F}_{m+1,n+1}](\theta \omega)
\]

holds for \( \mu \)-a.e. \( \omega \). Let \( \mathcal{M}_{\theta}^1 \) be the class of homogeneous probability measures. Note \( \mathcal{M}_{\theta}^1 \subseteq \mathcal{M}_{\theta} \). For example, Markov processes with stationary transition probabilities belong to \( \mathcal{M}_{\theta}^1 \).

For each \( A \in \mathcal{F} \) the random variable \( \mu(A | \mathcal{F}_{-\infty,0}) \) is defined only up to \( \mu \)-null sets. Under our assumptions there exist **regular conditional probabilities**, that is the choice of random variables can be made so that \( \mu( A | \mathcal{F}_{-\infty,0})(\omega) \) is a probability measure on \( (\Omega, \mathcal{F}) \) for each \( \omega \). We shall denote a choice of such a regular conditional probability by \( (\mu^*_\omega) \) or simply \( \mu^* \), so that \( \mu^*_\omega(A) \) is a version of \( \mu(A | \mathcal{F}_{-\infty,0})(\cdot) \). If \( \mu \in \mathcal{M}_{\theta}^1 \) we will always require that for every \( \omega \in \Omega, B \in \mathcal{F}, \) and every non-negative integer \( m \),

\[
\mu^*_\omega(\theta^{-m}B | \mathcal{F}_{-\infty,m})(\eta) = \mu^*_{\theta^m \eta}(B)
\]

holds for \( \eta^*_\omega \)-a.e. \( \eta \in \Omega \).

We will consider \( \mu \in \mathcal{M}_{\theta}^1 \) satisfying the following **ratio-mixing condition**:

(RM) There exists a non-decreasing function \( m(n) \) such that \( 0 < m(n) < n \), \( m(n)/n \) approaches zero as \( n \) tends to infinity, and

\[
\lim_{n \to \infty} \frac{1}{n} \sup \{ \log \frac{\mu^*_\omega(A)}{\mu^*_n(A)} : \eta \in \Omega^-, \omega \in \Omega^-, A \in \mathcal{F}_{m(n),n} \} = 0.
\]

Denote by \( \varepsilon_n \) the \( \mathcal{M} \)-valued random variable whose value at \( \omega \) is

\[
\varepsilon_{n,\omega} = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\theta^k \omega}.
\]
For a fixed \( \mu \in \mathcal{M}_\theta^1 \) and choice of \( \mu^*_\omega = (\mu^*_\omega) \), let \( Q_{\omega} \) denote the distribution of \( \varepsilon_n \) under \( \mu^*_\omega \), that is

\[
Q_{\omega}(A) = \mu^*_\omega [\varepsilon_n, \varepsilon_n \in A]
\]

for \( A \) a Borel subset of \( \mathcal{M} \). Let \( Q_{n, \mu} \) be the distribution of \( \varepsilon_n \) under \( \mu \). Also define

\[
\mu_{\inf}(A) = \inf_{\omega \in \Omega} \mu^*_\omega (A), \ A \in \mathcal{B}
\]

and

\[
Q_n(A) = \mu_{\inf} [\varepsilon_n, \varepsilon_n \in A].
\]

The purpose of this section is to prove the following theorem.

**1.1 Theorem** Let \( \mu \in \mathcal{M}_\theta^1 \) with \( \mu^* \) satisfying (RM). Then there exists a lower semi-continuous function \( K: \mathcal{M} \rightarrow [0, \infty] \) such that

(i) \( \lim_{n \to \infty} \frac{1}{n} \log \inf_{\omega \in \Omega} Q_{n, \omega} (A) > -\inf \{K(\nu): \nu \in A\}, \ A \ open; \)

(ii) \( \lim_{n \to \infty} \frac{1}{n} \log \sup_{\omega \in \Omega} Q_{n, \omega} (A) < -\inf \{K(\nu): \nu \in A\}, \ A \ closed. \)

**Remark 1.** The \( K \) in Theorem 1 is easily shown to be unique, see [4].

**Remark 2.** Our proof exploits the sub-additivity of \( Q_n(A) \) for suitable \( A \). It is quite similar to the proof given by Stroock of Theorem 6.2 of [7].

We proceed by a sequence of lemmas. First we introduce the class \( I \) of all subsets \( A \) of \( \mathcal{M} \) representable in the form

\[
A = \{\nu \in \mathcal{M}: \ \int Y_i \ d\gamma - c_i < \varepsilon, \ i = 1, 2, \ldots, p\}
\]

where \( p \) is some positive integer, \( c_i \) are real numbers, \( \varepsilon > 0 \), and \( Y_i \) are real valued, continuous (hence bounded) random variables, and there exist non-negative
integers \( N_1 \) and \( N_2 \) so that \( Y_i \) is \( \mathcal{F}_{N_1, N_2} \)-measurable for \( 1 < i < p \). Then \( \mathcal{R} \) is a basis of convex open sets for the topology at \( M \). Denote by \( \mathcal{R}_0 \) the subclass of \( \mathcal{R} \) consisting of all \( \mathcal{A} \in \mathcal{R} \) for which \( N_2 = 0 \).

1.2 Lemma For non-negative integers \( n \) and \( m \) and \( \mathcal{A} \in \mathcal{R}_0 \),

\[
q_{n+m}(\mathcal{A}) > q_n(\mathcal{A})q_m(\mathcal{A})
\]

Proof. Observe that

\[
u^*_\omega \left[ \varepsilon_{n+m} \in \mathcal{A} \right] = \mu^*_\omega \left[ \left( \frac{m}{m+n} \varepsilon_m + \frac{n}{m+n} \varepsilon_n \circ \theta_m \right) \in \mathcal{A} \right]
\]

\[
> \mu^*_\omega \left[ \varepsilon_m \in \mathcal{A}, \varepsilon_n \circ \theta_m \in \mathcal{A} \right]
\]

\[
= \int \left[ \varepsilon_m \in \mathcal{A} \right] \mu^*_\omega \left[ \theta_m^{-1} \left[ \varepsilon_n \in \mathcal{A} \right] \right](\mathcal{N}) \mu^* \left( d\mathcal{N} \right)
\]

\[
= \int \left[ \varepsilon_m \in \mathcal{A} \right] \left( \theta_m^{-1} \right) \left( \varepsilon_n \in \mathcal{A} \right)(\mathcal{N}) \mu^* \left( d\mathcal{N} \right)
\]

\[
> q_n(\mathcal{A})q_m(\mathcal{A}).
\]

where the first inequality follows from the convexity of \( \mathcal{A} \), the second equality holds because \( \left[ \varepsilon_m \in \mathcal{A} \right] \in \mathcal{F}_{-\infty, m} \) since \( \mathcal{A} \in \mathcal{R}_0 \), and the third inequality uses (1.2). Taking the infimum over \( \Omega^- \) gives the lemma. \( \square \)

1.3 Lemma For \( \mathcal{A} \in \mathcal{R}_0 \), \( q_n(\mathcal{A}) = 0 \) or \( q_n(\mathcal{A}) > 0 \) for all sufficiently large \( n \).

Proof. Let \( \mathcal{A} \in \mathcal{R}_0 \), and assume \( q_m(\mathcal{A}) > 0 \). The parameter \( \varepsilon \) enters in the definition of \( \mathcal{A} \); let us write \( \mathcal{A} = \mathcal{A}^\varepsilon \), and for \( 0 < \varepsilon' \) let \( \mathcal{A}^{\varepsilon'} \) be defined like \( \mathcal{A}^\varepsilon \) but with \( \varepsilon' \) in place of \( \varepsilon \). Assume \( \mathcal{A}^\varepsilon \neq \Omega \). For \( \varepsilon' < \varepsilon, \mathcal{A}^{\varepsilon'} \subseteq \mathcal{A}^\varepsilon \) and indeed there exists \( c > 0 \) such that dist \((\nu', \nu) > c\) whenever \( \nu' \in \mathcal{A}^{\varepsilon'} \) and \( \nu \in \mathcal{A}^\varepsilon \). By taking \( \varepsilon' < \varepsilon \) but sufficiently big we can
ensure \( Q_m(\varphi^{e'}) > 0 \). By Lemma 1.2, \( Q_{km}(\varphi^{e'}) > 0 \) for \( k = 1, 2, \ldots \). For \( n < m \) write \( n = qm + r, 0 < r < m \). Then \( Q_{qm}(\varphi^{e'}) > 0 \). For \( n \) sufficiently big \( \text{dist} (\varepsilon_{qm}, \varepsilon) < \varepsilon \), and so \([\varepsilon_{qm} \in \varphi^{e'}] \subseteq [\varepsilon_n \in \varphi^{e}]\) and hence \( Q_n(\varphi^{e'}) > Q_{qm}(\varphi^{e'}) > 0 \). \( \square \)

1.4 Lemma \[ \lim_{n \to \infty} \frac{1}{n} \log Q_n(\varphi) = \sup \frac{1}{n} \log Q_n(\varphi) = : \Lambda(\varphi), \varphi \in \Gamma. \]

Proof By Lemma 1.2, \(- \frac{1}{n} \log Q_n(\varphi)\) is a sub-additive function of \( n \) for \( \varphi \in \Gamma \); by Lemma 1.3 this function is strictly positive for all big \( n \), or else identically zero. This suffices. \( \square \)

Consider \( \varphi \in \Gamma \) and write \( \varphi = \varphi^{e} \) as above. Then

\[ \lim_{n \to \infty} \frac{1}{n} \log Q_n(\varphi^{e}) = \Lambda(\varphi^{e}) \]

is an increasing function of \( e \), and hence it has at most denumerably many points of discontinuity.

Now let \( \varphi \in \Gamma \), with \( e, n, n_1, n_2 \) as in the definition of \( \Gamma \). Let \( \varphi = \varphi^{n_2} \) and note \( \varphi \in \Gamma \).

Define \( \Gamma^1 \) to be the subclass of \( \Gamma \) consisting of all \( \varphi \in \Gamma \) such that the \( e \) corresponding to \( \varphi \) is a continuity point of the function \( \Lambda(\varphi^{e}) \). Then \( \Gamma^1 \) is a basis of convex sets for the topology of \( \mathcal{H} \).

1.5 Lemma For \( \varphi \in \Gamma^1 \), \[ \lim_{n \to \infty} \frac{1}{n} \log Q_n(\varphi) = \Lambda(\varphi) \] exists.

Proof. Let \( \varphi \in \Gamma^1 \), then \( \varphi \in \Gamma \) and if \( e \) corresponds to \( \varphi \), it is a continuity point of \( \Lambda(\varphi^{e}) \). For \( e' < e < e'' \) and \( n \) sufficiently big

\[ [\varepsilon_n \in \varphi^{e'}] \subseteq [\varepsilon_n \in \varphi^{e}] \subseteq [\varepsilon_n \in \varphi^{e''}] \]

Applying \( Q_n \), taking logarithms and dividing by \( n \), and then letting \( n \) tend to infinity the first and last terms go to \( \Lambda(\varphi^{e'}) \) and \( \Lambda(\varphi^{e''}) \) respectively. Since \( e \) is a continuity point of \( \Lambda(\varphi^{e}) \) the result follows. \( \square \)
Now we define
\[ K(v) = - \inf \{ \Lambda(\lambda): v \in \lambda, \lambda \in \mathbb{A} \}. \]

It is easily verified that \( K \) is lower semicontinuous and convex, and if \( \mathbb{A} \) is closed and \( \mathbb{A}(\delta) \) is a \( \delta \)-neighborhood of \( \mathbb{A} \) then
\[ \lim_{\delta \to 0} \inf \{ K(v): v \in \mathbb{A}(\delta) \} = \inf \{ K(v): v \in \mathbb{A} \}. \]

(Complete details for the analogous results in [7] are given in Section 6 of that monograph).

**Proof of Theorem 1.1.** To prove (i) let \( v \in \mathbb{A} \), where \( \mathbb{A} \) is open. There exists \( a^1 \in \mathbb{A} \) such that \( v \in a^1 \) and \( a^1 \subseteq a \). Then \( q_n(a) \geq q_n(a^1) \) and so
\[ \lim_{n \to \infty} \frac{1}{n} \log q_n(a) \geq \Lambda(a^1) \geq -K(a). \]

Taking the supremum over all \( v \in \mathbb{A} \) in the last relation gives (i).

For the proof of (ii) choose \( \mathbb{A} \) to be closed (hence compact). Let \( c > \inf \{ K(v): v \in \mathbb{A} \} \). For \( \alpha > 0 \), choose for each \( v \in \mathbb{A} \) a neighborhood \( \mathcal{N} \) of \( v \) of diameter less than \( \alpha \) with \( \mathcal{N} \in \mathbb{A} \) and \( \Lambda(\mathcal{N}) < c \). Recall the constants \( N_1 \) and \( N_2 \) entering the definition of \( \mathcal{N} \). With \( m(n) \) the function introduced in (RM), define
\[ k(n) = m(N_2 + 2n) + N_1 \]

Then \( k(n) < n \) for \( n > n_0 \), say. This guarantees
\[ (1.3) \quad -N_1 + k(n) \geq m(k(n)) + N_2 + n. \]

There exist neighborhoods \( \mathcal{N}' \) and \( \mathcal{N}'' \) of \( v \) and a positive integer \( n' \) such that \( \mathcal{N}' \in \mathbb{A} \), \( \mathcal{N}'' \in \mathbb{A} \), \( \mathcal{N}'' \subseteq \mathcal{N}' \subseteq \mathcal{N} \), and for \( n > n' \),
\[ \{ e_n \in \mathcal{N} \} \supseteq \theta^{k(n)} \{ e_n \in \mathcal{N}' \} \supseteq \{ e_n \in \mathcal{N}'' \} \]
The event in the middle is $\mathcal{F}_{-N_1 + k(n), k(n) + N_2 + n}$ - measurable. Now choose a finite set $v_1, \ldots, v_q$ such that the corresponding $\mathcal{N}_i^1, \ldots, \mathcal{N}_i^q$ cover $A$.

For each $1 \leq i \leq q$ there will be a corresponding $\mathcal{N}_i^1$ depending on integers $N_1^{(i)}, N_2^{(i)}$. Now

\begin{equation}
\mu^* \left[ \varepsilon_n \in A \right] \leq \sum_{i=1}^{q} \mu^* \left[ \varepsilon_n \in \mathcal{N}_i^q \right] < \sum_{i=1}^{q} \mu^* \left[ \vartheta^{k(n)} \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] \right].
\end{equation}

Let $k_i^{(n)}$ be defined as in (2) with $N_1^{(i)}$ and $N_2^{(i)}$ for $N_1$ and $N_2$. Put

\[ \delta_i(n) = \sup_{\omega \in \Omega} \log \mu^* \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] - \log \mu^* \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] \]

and observe that (1.3) allows us to apply (RM) to conclude that $\delta_i(n)/n$ tends to zero as $n$ approaches infinity. Then from (1.4)

\begin{equation}
\lim_{n \to \infty} \frac{1}{n} \log \sup_{\omega \in \Omega} \mu^* \left[ \varepsilon_n \in A \right] < \lim_{n \to \infty} \frac{1}{n} \log \sum_{i=1}^{q} \mu^* \left[ \vartheta^{k(n)} \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] \right]
\end{equation}

\begin{equation}
= \max_{1 \leq i \leq q} \lim_{n \to \infty} \frac{1}{n} \log \sup_{\omega \in \Omega} \mu^* \left[ \vartheta^{k(n)} \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] \right]
\end{equation}

\begin{equation}
= \max_{1 \leq i \leq q} \lim_{n \to \infty} \left( \log \mu^* \left[ \vartheta^{k(n)} \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] \right] + \delta_i(n) \right)
\end{equation}

\begin{equation}
< \max_{1 \leq i \leq q} \lim_{n \to \infty} \left( \log \mu^* \left[ \varepsilon_n \in \mathcal{N}_i^1 \right] + \delta_i(n) \right)
\end{equation}

\begin{equation}
= \max_{1 \leq i \leq q} \Lambda(\mathcal{N}_i) < c
\end{equation}

and since $c$ was an arbitrary number greater than $-\inf \{K(v) : v \in A\}$ (ii) follows.
2. Entropy. In their work [2] Donsker and Varadhan not only prove the existence of a rate function $K$, but they identify it as a relative entropy. As was shown in [4] their proof can be adapted to more general situations. In addition to the hypothesis (RM) of Section 1, one needs an assumption of continuous dependence.

(CD) For every real valued continuous random variable $Y$ which is $\mathcal{F}_{-\infty,1}^\omega$-measurable, $\omega - \mu^*\omega (Y)$ is a continuous function.

Let $\mu \in \mathcal{M}_\Theta$, $\nu \in \mathcal{M}$, and let $(\mu^*\omega)$, $(\nu\omega)$ denote regular conditional probabilities given $\mathcal{F}_{-\infty,0}^\omega$ as in Section 1; by $\mu^*\omega |_1$ and $\nu\omega |_1$ we denote the restrictions to $\mathcal{F}_1$. The entropy $H(\nu; \mu^*)$ will now be defined. Set $H(\nu; \mu^*) = \infty$ if $\nu \notin \mathcal{M}_\Theta$ or if $\nu\omega |_1 \ll \mu^*\omega |_1$ fails on a set of $\omega$ having positive $\nu$-measure. In the remaining cases

$$H(\nu, \mu^*) = \int_\Omega \left[ \int_M \log \frac{d\nu\omega |_1}{d\mu^*\omega |_1} (y) \nu\omega |_1 (dy) \right] \nu(d\omega)$$

Note that $H(\nu; \mu^*)$, depends on $\mu^*$, and not just $\mu$, because one is integrating with respect to $\nu$ and $\mu$-null sets will not usually be $\nu$-null sets.

Throughout this paper we are taking $M$ to be compact.

The next result, Theorem 2, is included in [4]. To make this paper self-contained we include a brief proof.

2. Theorem. Let $\mu \in \mathcal{M}_\Theta$ with $\mu^*$ satisfying (RM) and (CD). Then Theorem 1.1 holds with $K(\nu) = H(\nu; \mu^*)$.

Proof. It will be shown how to adapt the ideas of Donsker and Varadhan [2].

Actually in [2] it is not the random variable $\varepsilon_n$ but a closely related random variable $\varepsilon'$ which is studied. For each positive integer $n$ define $\pi_n : \Omega \to \Omega$ by $(\pi_n \omega)_i = \omega_i$, $0 \leq i < n$ and $(\pi_n \omega)_{j+n} = (\pi_n \omega)_j$ for all $j$. Then
\[ \epsilon_n^i, \omega := \epsilon_n, \pi_n(\omega). \] Note \( \epsilon_n^i \in M_\omega. \) As \( n \to \infty, \) \( \text{dist}(\epsilon_n, \omega, \epsilon_n^i, \omega) \to 0 \) uniformly in \( \omega, \) and the distributions of the random variables \( (\epsilon_n) \) have the same rate function as those of the random variables \( (\epsilon_n^i). \) Hence \( K \) is infinite off \( M_\omega. \)

In [2] the underlying measure is assumed to be Markovian. We can easily reduce our situation to a Markovian one by keeping track of the past. That is, instead of the shift on sequences \( \omega = (\omega_n) \) consider the shift on sequences \( \omega^* = (\omega_n^\ast) \) where \( \omega_n^\ast = (\omega_n, \omega_{n-1}, \ldots). \) Now the condition (CD) is just the Feller property for the Markov process we have obtained. The Feller property was assumed in [2]. Now all the results on entropy developed in [2] are available to us. Since we are here working only with the compact state space, the results of [2] immediately imply the upper bound (ii) of Theorem 1.1 with \( H(\nu, \mu^\ast) \) in place of \( K. \)

Finally we must argue that the lower bound (i) of Theorem 1.1 holds with \( H(\nu, \mu^\ast) \) in place of \( K. \) Our Markov process \( (\omega^\ast) \) will never satisfy the assumptions of [2], (there is not even a reference measure in our case). However, going back to the original shift on \( \Omega, \) the condition (RM) is exactly what is needed to imitate [2]. We sketch this briefly.

Start first with \( \nu \in M_\theta \) which is ergodic and satisfies \( H(\nu, \mu^\ast) \to \infty. \) Using the properties of \( H \) one obtains as in [2] that \( (\nu, \omega_0) \to \nu, \) \( (\mu, \omega) \to \nu \) with corresponding Radon-Nikodym derivative \( \psi_n(\omega) = \psi_n(\omega^\ast, \omega_1, \ldots, \omega_n). \) Then \( \psi_{n+m} = \psi_n \cdot \psi_m \circ \theta_n \) \( \nu \) a.e. so that the ergodic theorem immediately implies the following "conditional Shannon-McMillan Theorem,"

\begin{equation}
\lim_{n \to \infty} \frac{1}{n} \log \psi_n = H(\nu; \mu^\ast), \quad \nu \text{ a.e.}
\end{equation}

Now for \( \alpha > 0, \)
\[
\mu^*_\omega^{-\left[\varepsilon'_n \in \mathcal{N}\right]} > \int_{\varepsilon'_n \in \mathcal{N}} e^{-\frac{1}{n} \log \psi_n} \, d\nu^*_\omega^{-}\left[\varepsilon'_n \in \mathcal{N}\right] - n(H(\nu; \mu^*) + \alpha) \right]
\]

Let \( \mathcal{N} \) be a neighborhood of \( \nu \). Then as \( n \) approaches infinity the second factor in the last member tend to 1 \( \nu \)-a.e., so that

\[
(2.2) \quad \lim_{n \to \infty} \frac{1}{n} \log \mu^*_\omega^{-\left[\varepsilon'_n \in \mathcal{N}\right]} > -H(\nu; \mu^*), \nu \text{- a.e.}
\]

To obtain (i) of Theorem 1.1 we need (2.2) to hold for all \( \omega^* \), and indeed the inequality is to hold uniformly in \( \omega^* \). Let \( m(n) \) be as in condition (RM) of Section 1. Since \( m(n)/n \to 0 \), there exist neighborhoods \( \mathcal{N}' \) and \( \mathcal{N}'' \) of \( \nu \) and a positive integer \( n' \) such that \( \mathcal{N} \supseteq \mathcal{N}' \supseteq \mathcal{N}'' \) and

\[
(2.3) \quad \left[\varepsilon'_n \in \mathcal{N}\right] \supseteq \theta^{m(n)}\left[\varepsilon'_n \in \mathcal{N}'\right] \supseteq \left[\varepsilon'_n \in \mathcal{N}''\right], \quad n > n'
\]

According to (2.2) there exists an \( \omega^* \in \Omega^* \) and for every \( \alpha > 0 \) an \( n_\alpha \) such that

\[
\frac{1}{n} \log \mu^*_\omega^{-\left[\varepsilon'_n \in \mathcal{N}''\right]} > -(H(\nu; \mu^*) + \frac{\alpha}{2}), \quad n > n_\alpha.
\]

Now it follows from (2.3) and condition (RM) that there exists \( \mathcal{N}_\alpha \) such that

\[
\frac{1}{n} \log \mu^*_n^{-\left[\varepsilon'_n \in \mathcal{N}\right]} > -(H(\nu; \mu^*) + \alpha), \quad n > n'_\alpha, \quad n^* \in \Omega^*.
\]

Since \( \text{dist}(\varepsilon'_n, \omega^*, \varepsilon_n, \omega) \to 0 \) uniformly, the corresponding relation with \( \mathcal{N}_\alpha \) in place of \( \mathcal{N}_\alpha' \) is also valid.

It remains to consider the case that \( \nu \in \mathcal{M}_\theta \) is not ergodic. This can be handled exactly as in [2]. \( \square \)
Later we consider the situation in which \( M = \{0, 1, \ldots, N-1\} \). In that case (CD) reduces to the condition

\[
(2.4) \quad \omega^* + \mu^*_{\omega} [\eta: \eta_1 = i] \text{ is a continuous function of } \omega^- \text{ for } 0 < i < N.
\]

3. Gibbs measures. Our compact space will now be a finite set \( M = \{0, 1, \ldots, N-1\} \). For any finite sequence \((x_0, x_1, \ldots, x_{n-1})\) let \([x_0, x_1, \ldots, x_{n-1}] = \{\omega \in \Omega(M): \omega_i = x_i, 0 < i < n\}\). Also set \( B_i = \{\omega \in \Omega(M): \omega_i = i\}\).

Given \( \mu \in \mathcal{M}_\theta \), the condition (CD) of the previous section reduces to

\[
(3.1) \quad \omega^- + \mu^*_{\omega^-} (B_i) \text{ is a continuous function, } 0 < i < N.
\]

For given a conditional probability distribution \( (\mu^*)_\omega^- \) defined only for \( B_i: 0 < i < N \) one naturally extends it to \( \mathcal{F}_\omega^- \). For example, if \( \omega^- = (\ldots, \omega_{-1}, \omega_0) \), \( B = \Theta[i, j, k] \), \( \mu^*_{\omega^-} (B) \) is defined as follows: let \( \omega^* = (\ldots, \omega_{-1}, \omega_0, j) \) and set

\[
\mu^*_{\omega^-} (B) = \delta_{\omega_0, i} \mu^*_{\omega_j} (B_j) \mu^*_{\omega^-} (B_k).
\]

If (3.1) holds \( \mu^* \) will satisfy (CD).

Instead of \( \Omega(M) \) we will also consider a subshift \( \Omega_A \) determined by an \( N \times N \) matrix \( A \) of 0's and 1's. That is, \( \Omega_A = \{\omega \in \Omega: A \omega_i \omega_{i+1} = 1 \text{ for all } i\} \). All concepts can be relativized to \( \Omega_A \) and this will be indicated by an \( A \) in the notation, e.g. \( \mathcal{F}_A = \{B \cap \Omega_A: B \in \mathcal{F}\} \), \( \mathcal{M}_\Theta^A \) is the class of probability measures on \( (\Omega_A, \mathcal{F}_A) \), and \( \mathcal{M}_\Theta^A \) is the class of \( \Theta \)-invariant members of \( \mathcal{M}_A \). Any \( \mu \in \mathcal{M}_\Theta^A \) may be extended in exactly one way to a \( \mu \in \mathcal{M}_\Theta \); evidently \( \mu(B) = 0 \) for \( B \in \mathcal{F} \setminus \mathcal{F}_A \). This will allow us to start with \( \mu \in \mathcal{M}_\Theta^A \) and apply the results of the previous section. In working with \( \Omega_A \) we will assume that there exists a positive integer \( m \) such that

\[
(3.2) \quad (A^m)_{ij} > 0, 0 < i < N, \ 0 < i < N
\]
A finite or infinite sequence \((x_i)\) of non-negative integers less than \(N\) is \(A\)-admissible if \(A_{x_i x_{i+1}} = 1\) for every pair of consecutive elements \(x_i\) and \(x_{i+1}\) of \(x\).

3.1. Proposition. Assume \(\mu \in \mathcal{M}_A\) satisfies (3.1) with \(\omega\) varying over \(\Omega_A\) and also

\[
\inf \{ \mu^*_\omega^{-i}(B_i) : \omega^{-i} \text{ is } A \text{-admissible} \} > 0.
\]

Then (RM) holds with \(m(n) = m\). Considering \(\mu\) as an element of \(\mathcal{M}_\theta\), one can define \((\mu^*_\omega)\) so that (CD) and (RM) continue to hold for the extended family.

Proof. If \(x^-\) and \(y^-\) belong to \(\Omega_A^-\) and \(x^-i\) and \(y^-i\) also belong to \(\Omega_A^-\), we can write

\[
\mu^*_x^{-i}(B_i) = \mu^*_y^{-i}(B_i) (1 + \alpha(x^-y^-,i))
\]

and define

\[
\alpha_s = \sup \{ \alpha(x^-y^-,i) : x_k = y_k, -s < k < 0 \}.
\]

Our assumptions imply

\[
(3.4) \quad \lim_{s \to \infty} \alpha_s = 0
\]

In verifying (RM) it will suffice to consider the supremum over all sets \(B \in \mathcal{F}_{m,n}^A\) having the form \(B = \theta^{-m} \left[ i_m, i_{m+1}, \ldots, i_n \right] \). Then

\[
(3.5) \quad \mu^*_\omega^{-i}(B) = \sum \mu^*_\omega^{-i}(B_k) \mu^*_\omega^{-i}(B_{k_1}) \ldots \mu^*_\omega^{-i}(B_{i_m}) \ldots \mu^*_\omega^{-i}(B_{i_n})
\]

where the sum extends over all \(k = (k_1, \ldots, k_{m-1})\) such that \((\omega_0, k_1, \ldots, k_{m-1}, i_m)\) is admissible; by our assumption the sum contains at least one term, and at most
$N^{m-1}$ terms and each factor in the sum is bounded below by a positive constant (by
3.1)), and bounded above by a constant times $\pi_{t=0}^{n-m}(1 + a_t)$ and by (3.4)

$$\lim_{n \to \infty} n^{-1} \log \pi_{t=0}^{n-m}(1 + a_t) = 0.$$ 

The final assertion of the Proposition is easily seen to be true. \qed

Following the exposition of Bowen [1] a certain class of $\mu \in \mathcal{M}_0^A$ will be
introduced and called Gibbs measures. (According to traditional terminology this
class ought to be called "Gibbs measures with translation invariant exponentially
decreasing interactions", see Ruelle [5], Section 5.18).

For $\beta \in (0,1)$ define the metric $d_{\beta}(\omega, \eta)$ on $\Omega_A$ by

$$d_{\beta}(\omega, \eta) = \beta^n$$

where $n$ is the least non-negative integer such that $\omega_n \neq \eta_n$ or $\omega_{-n} \neq \eta_{-n}$. Let $H_A$
denote the class of real-valued functions $\phi$ on $\Omega_A$ which are Holder con-
tinuous. $H_A$ does not depend on $\beta$. Note $\phi \in H_A$ if and only if there exists a
positive constant $b$ and $\gamma \in (0,1)$ such that

$$\sup \{ |\phi(x) - \phi(y)| : x \in \Omega_A, y \in \Omega_A, x_i = y_i \text{ for } |i| < n \} < b \gamma^n, n = 0,1, \ldots$$

We rely on the following theorem: For $\phi \in H_A$ there exists a unique $\mu \in \mathcal{M}_0^A$
and unique real number $P$ for which one can find positive constants $c_1$ and $c_2$
such that

$$(3.6) \quad c_1 < \frac{\mu[\omega_0, \omega_1, \ldots, \omega_{n-1}]}{\exp(-Pn + \sum_{k=0}^{n-1} \phi(\theta^k \omega))} < c_2$$

for every $\omega \in \Omega_A$ and $n > 0$. For a proof of the theorem see [1]. The measure $\mu$
is called Gibbs measure corresponding to $\phi$. (The constant $P$ is equal to the
pressure of $\phi$, see [1], Theorem 1.22. If $\phi$ and $\phi'$ belong to $H_A$ they deter-
mine the same Gibbs measure if and only if there exists a constant $c$ and a
$\psi \in H_A$ such that $\phi - \phi' = c + \psi \theta - \psi$, see Ruelle [5], Theorem 5.21). For
\( \phi \in H_A \) there exists \( \phi' \in H_A \) such that \( \phi'(x) = \phi'(y) \) for all \( x \in \Omega_A, \; y \in \Omega_A \) with \( x_i = y_i \) for \( i > 0 \), and \( \phi' \) determines the same Gibbs measure as \( \phi \); see [1] Lemma 1.6.

Given \( v \in \mathcal{M} \), the notations \( (v, \psi) \) and \( \int \psi dv \) will be used with the same significance. If \( v \in \mathcal{M}_0 \), \( h_v \) will denote the Kolmogorov-Sinai entropy of \( v \) with respect to the shift \( \theta \). The principal result of the section can now be stated.

### 3.2 Theorem

Let \( \phi \in H_A, \mu \) the corresponding Gibbs measure extended to \( \Omega, P \) the pressure, as in (3.6). Then there exists \( \mu^* \) satisfying (CD) and (RM).

The large deviation principle as given in Theorem 1.1 holds and

\[
K(v) = H(v; \mu^*) = -h_v - (v, \phi) + P \quad v \in \mathcal{M}_0.
\]

where \( (v, \phi) = \infty \) if \( v \in \mathcal{M}_0 \setminus \mathcal{M}_A \).

The conditional probabilities \( \mu^*_x(\theta^{-1}[i]) \), with \( x^- \in \Omega^- \) and \( 0 < i < N \), will be obtained as limits. Assume \( \mu \) is a Gibbs measure. Let

\[
\mu_{(n)}(x^-)(i) = \frac{\mu[x_{-n}, \ldots, x_{0}, i]}{\mu[x_{-n}, \ldots, x_0]}
\]

which is well defined if \( (x_{-n}, \ldots, x_{0}, i) \) is \( A \)-admissible; otherwise make the following convention: \( \mu_{(n)}(x^-)(i) = 0 \) if \( A_{x_0 i} = 0 \), and if \( A_{x_0 i} = 1 \) \( \mu_{(n)}(x^-)(i) = \mu(s)(i) \) where \( s \) is the largest integer such that \( (x_{-s}, \ldots, x_{0}, i) \) is \( A \)-admissible.

### 3.3 Lemma

For a Gibbs measure \( \mu \), \( \mu_{(n)}(i) \) converges uniformly in \( x^- \) as \( n \to \infty, \; 0 < i < N \).

**Proof.** It will suffice to investigate the ratios on the right side at (3.8) for \( x \in \Omega_A \). To avoid negative subscripts define now

\[
\mu_{n}(x) = \frac{\mu[x_0, \ldots, x_n, x_{n+1}]}{\mu[x_0, \ldots, x_n]}.
\]
It has to be shown that

\[(3.9) \limsup_{s \to \infty} \{ |\mu_{s+k}(x) - \mu_{s+k}(y)| : x \in \Omega_A, y \in \Omega_A, x_i = y_i \text{ for } k \leq i \leq k+s+1, k \geq 0 \} = 0 \]

It follows from the definition of Gibbs measure that \( \mu_n(x) \) is bounded away from zero for \( x \in \Omega_A \). For (3.9) more information about \( \mu \) is needed. As shown in [1] there is a useful formula for \( \mu[x_0, \ldots, x_n] \) involving an auxiliary measure \( \mu' \in \mathcal{M}_\theta \), a strictly positive function \( h \in \mathcal{H}_A \) such that \( h(x) = h(y) \) if \( x, y \in \Omega_A, x_i = y_i \text{ for } i > 0 \). Since \( h \) does not depend on the past write \( h(x_0 x_1 \ldots) \) for \( h(x) \). If \( z = (z_0, z_1, \ldots) \) then \( (x_0 \ldots x_n z) \) denotes the sequence \( (x_0, \ldots, x_n, z_0, z_1, \ldots) \). With \( \lambda = e^p \) the formula for \( \mu[x_0, \ldots, x_n] \) reads

\[(3.10) \mu[x_0, \ldots, x_n] = \lambda^{-(n+1)} \int \exp \left\{ \sum_{j=0}^m (\phi(\theta^j(x_0 \ldots x_n z)) h(x_0 \ldots x_n z)) \right\} \mu'(dz) \]

the integral extending over all \( z = (z_0, z_1, \ldots) \) such that \( (x_0 \ldots x_n z) \) is \( A \)-admissible. Assume \( (x_0, \ldots, x_n) \) is admissible and let \( w = (w_0, w_1, \ldots) \) be any sequence such that \( (x_0, \ldots, x_n w) \) is admissible. Let \( n = k + s, s > 1 \). On the right side of (3.10) replace the argument \( (x_0 \ldots x_n z) \) by \( (x_0 \ldots x_{n-1} w) \) in those terms of the exponential sum corresponding to \( 0 < j < k \), and make the same replacement in the argument of \( h \). One obtains

\[\lambda^{-(n+1)} \exp \left\{ \sum_{j=0}^k \phi(\theta^j(x_0 \ldots x_{n-1} w)) h(x_0 \ldots x_{n-1} w) \right\} \int \exp \left\{ \sum_{j=k+1}^{k+s} \phi(\theta^j(x_0 \ldots x_n z)) \right\} \mu'(dz) \]

and the ratio of \( \mu[x_0 \ldots x_n] \) to this quantity differs from 1 by a quantity going to zero exponentially as \( s \to \infty \). In like manner \( \mu[x_0 \ldots x_{n-1}] \) is approximated by

\[\lambda^{-n} \exp \left\{ \sum_{j=0}^k \phi(\theta^j(x_0 \ldots x_{n-1})) h(x_0 \ldots x_{n-1} w) \right\} \int \exp \left\{ \sum_{j=k+1}^{k+s-1} \phi(\theta^j(x_0 \ldots x_{n-1} z)) \right\} \mu'(dz) \]
Hence defining for \( n = k+s \)

\[
\alpha_{n,s}(x) = \lambda^{-1} \int \exp\left\{ \sum_{j=k+1}^{k+s} \phi(\theta^j(x_0 \cdots x_n z)z) \right\} u'(dz) / \int \exp\left\{ \sum_{j=k+1}^{k+s+1} \phi(\theta^j(x_{1} \cdots x_{n-1} z)) \right\} u'(dz)
\]

one finds

\[
\lim_{s \to \infty} \sup_{n > s} \left\{ \frac{u[x_0 \cdots x_n]}{u[x_0 \cdots x_{n-1}]} - \alpha_{n,s}(x) \right\} = 0
\]

Note that if \( y_i = x_i \) for \( k < i < k - s \) then \( \alpha_{n,s}(x) = \alpha_{n,s}(y) \). So (3.9) is proved, and the lemma follows. \( \square \)

**Proof of Theorem 3.2.** Let \( \mu \) be a Gibbs measure. Using Lemma 3.3 define

\[
\mu^*(B_i) = \lim_{n \to \infty} \mu^{(n)}_i
\]

for \( \omega \in \Omega_A \). Then the conditions of Proposition 3.1 hold and so the conclusion of this proposition applies. Now Theorem 1.1 and Theorem 2 apply. This justifies all parts of Theorem 3.2 except the second equality in (3.7).

For \( \nu \in \mathcal{M}_0^A \) define

\[
R^n(\omega) = \frac{1}{n} \log \frac{\nu[\omega_0, \cdots \omega_{n-1}]}{\mu[\omega_0, \cdots \omega_{n-1}]}
\]

By (3.6)

\[
R^n(\omega) = \frac{1}{n} \log \nu[\omega_0, \cdots \omega_{n-1}] - \frac{1}{n} \sum_{k=0}^{n-1} \phi(\theta^k \omega) + \frac{1}{n} \sum_{k=0}^{n-1} \phi^k(\omega) + \alpha_n(\omega)
\]

where \( \alpha_n(\omega) \to 0 \) uniformly in \( \omega \). One obtains immediately that as \( n \to \infty \)

\[
(v, R^n) \to -h_v - (v, \phi) + \mathcal{P}
\]
We wish to identify the right side of (3.11) as $H(v; \mu^*)$. For $\omega \in \Omega_A$ put

$$f_{kn}(\omega) = \frac{v[\omega_k, \ldots, \omega_n]}{\mu[\omega_k, \ldots, \omega_n]}, \quad r_n(\omega) = \frac{f_{-n,1}(\omega)}{f_{-n,0}(\omega)}$$

and note that if the denominator in $r_n$ vanishes so does the numerator; in that case interpret the ratio arbitrarily (e.g. set it equal to 1). Now

$$R_n^v = n^{-1} \log f_{0,n-1}. \quad \text{Note}$$

(3.12) \hspace{1cm} \frac{1}{n} \log f_{0,n} = \frac{1}{n} \log f_{0,0} + \frac{1}{n} \sum_{j=0}^{n-1} \log r_j \circ \theta^j$$

It will be shown that

(3.13) \hspace{1cm} (v, \log r_n) + H(v; \mu^*)$$

and then (3.12) will imply $(v, n^{-1} \log f_{0,n}) + H(v; \mu^*)$, as desired.

Recall the notation introduced in (3.8); $v^{(n)}(i)$ will have the corresponding definition, where here and in the subsequent discussion $0/0$ may be interpreted as having the value one. Also $(v_{\omega^-})$ will denote regular conditional probabilities $v[\cdot | \mathcal{F}_{\omega,0}](\omega)$. Write

$$r_n(\omega) = \frac{v[\omega_n, \ldots, \omega_0, \omega_1]}{\mu[\omega_n, \ldots, \omega_0, \omega_1]} \cdot \frac{\mu[\omega_n, \ldots, \omega_0]}{\mu[\omega_n, \ldots, \omega_0]} = \frac{v^{(n)}(\omega_1)}{\mu^{(n)}(\omega_1)}$$

and so

(3.14) \hspace{1cm} (v, \log r_n) = \int (v_{\omega^-}, \log r_n) v(d\omega^-)$$

$$= \int \left( \sum_{i=0}^{N-1} v^{(n)}(i) \log v^{(n)}(i) \right) v(d\omega^-) - \int \left( \sum_{i=0}^{N-1} v^{(n)}(i) \log \mu^{(n)}(i) \right) v(d\omega^-)$$

By the martingale convergence theorem $v^{(n)}(i) + (v_{\omega^-})_1(i) v$ - a.e., and since $\mu$ is a Gibbs measure $\log \mu^{(n)}(i) + \log (\mu_{\omega^-})_1(i)$ as $n \to \infty$. It follows then from (3.14) that $(v, \log r_n) + H(v; \mu^*)$. This justifies (3.7) for $v \in \mathcal{M}_0^A$. If
then for some \( k \) \( \nu_{0,k} \) is not absolutely continuous with respect to \( \nu_{0,k}^\infty \). By the basic properties of entropy (see [2]) this implies \( H(\nu;\mu^\infty) = \infty \) and again (3.7) holds.

For the class of Gibbs measures treated here one can actually verify a stronger condition than (RM). Namely, there exists a constant \( c \) such that
\[
\log \frac{\mu^\infty}{\nu_{\omega^-}(A)} < c, \text{ for } n \in \Omega^-, \omega^- \in \Omega^-, A \in \mathcal{F}_{m,n}, n \geq m.
\]

4. Examples and Remarks. The following examples and remarks give some more insight into the applicability of large deviation principles to stationary shifts.

4.1 Example (Adapted from Sokal [6]). This example will present a strongly mixing shift on \( M = \{-1, +1\} \) for which the large deviations principle fails. According to a basic result of Donsker and Varadhan (cf [2], Corollary 1.7) if \( \mu \in \mathcal{M}_\theta \) is a stationary shift obeying a large deviation principle with rate function \( K \), then for any bounded and continuous random variable \( Y \)
\begin{equation}
(4.1) \quad \lim_{u \to \infty} \frac{1}{n} \log \mu \left( \exp \left\{ \sum_{k=0}^{n-1} Y \circ \theta^k \right\} \right) = \sup_{\nu \in \mathcal{M}} \{ <Y, \nu> - K(\nu) \}.
\end{equation}

In the present example the limit in (4.1) will fail to exist for \( Y(\omega) = \omega_0 \). It remains to specify \( \mu \in \mathcal{M}_\theta \). The measure \( \mu \) is concentrated on those \( \omega \) consisting of a sequence of blocks of +1's followed by an equal number of -1's. The length of each block is the same as the number of +1's and -1's, (i.e. twice the number of +1's). It is assumed that the lengths of successive blocks forms a sequence of independent identically distributed random variables. Thus any time coordinate, say 0, belongs to a block of random length 2L. Let \( a_n = \mu[L=n] \). Of course, since \( \mu \) is to be stationary, given that \( L = n \), the coordinate 0 is equally likely to occupy the first, second, ... 2n th position of the block. Let \( N \) be an integer, \( N \geq 2 \), and choose \( 0 < \alpha < (N-1)/N \). Let \( n_1 = N^i \) \( i = 1,2, \ldots \) and
let \( a(n_i) = c_\alpha e^{-\alpha n_i} \), \( a(n) = 0 \) if \( n \not\in N_i \) for every \( i \), with \( c_\alpha \) so that 
\[ \sum a(n) = 1. \]
This determines \( \mu \in M_\theta \). Now let \( S_k = \sum_{i=0}^{k-1} \omega_i \), \( g(k) = \mu(\exp(S_k)) \).

Our claim is that as \( n \) tends to infinity \( n^{-1} \log g(n) \) oscillates. To see this note \( \mu[ S_{n_i} = n_i ] = P[ 0 \text{ occupies first position of block of length } 2n_i ] \)
\( = a(n_i)/(2n_i) \), and this implies

\[ g(n_i) > (2n_i)^{-1} \exp \{ n_i(1-\alpha) \}. \]

Next consider \( k_i = \beta n_i \) where \( 1 < \beta < N \). If, for given \( \omega \), \( k_i \) is in a block of length \( 2n \), with \( n < n_i \), then \( S_{k_i} < n < n_i \); if \( k_i \) is in a block of length \( 2n \) with \( n < n_i+1 \) use \( S_{k_i} < n_i \). Note there exists \( c'_\alpha \) such that
\[ \mu[ k_i \text{ in block of length } > 2n_i + 1 ] < c'_\alpha \exp\{-\alpha n_i + 1\}. \]
Hence \( \mu(\exp(S_{k_i})) < \exp(n_i) + c'_\alpha \exp\{-\alpha n_i + 1 + \beta n_i\} \). Now choose \( \beta = N\alpha + 1 \) and find

\[ g(k_i) < (1+c'_\alpha)e^{n_i} = (1+c'_\alpha) \exp \{ k_i/\beta \}. \]

Our assumptions insure \( \beta^{-1} < (1-\alpha) \), and oscillation is established.

\[ 4.2 \text{. Example} \]
Consider a shift \( (\Omega(M), \mathcal{F}, \theta, \mu) \), with \( \mu \in M_\theta \) such that the \( (\omega_i) \)
are \( m \)-dependent (i.e. \( \ldots, \omega_{n-1}, \omega_n \)) and \( (\omega_{n+m+1}, \omega_{n+m+2}, \ldots) \) are independent.

Then (RM) certainly holds. However easy examples show (even for \( M = \{0,1,2\} \))
that (CD) may fail.

\[ 4.3 \text{ Example} \]
Piecewise monotonic maps of the unit interval onto itself have been extensively studied. In [4] it was shown that under familiar hypotheses these maps have an invariant measure and the corresponding symbolic dynamics give rise to a stationary shift satisfying (RM). Thus Theorem 1.1 gives the existence of a rate function. From [8] an interesting expression for the rate function becomes available.
4.4 Example. Let \( \mu \) be the Bernoulli shift on \( \Omega(M), M = \{0,1\} \), with 
\[ \mu[\omega_0 = 0] = \mu[\omega_0 = 1] = 1/2. \]
Define \((\eta_i)\) by \( \eta_i = \sum_{k=1}^{\infty} 2^{-k} \omega_{i+k} \). This gives rise to a new shift on \([0,1]\), for which the large deviation principle will hold; this follows from the Donsker-Varadhan contraction principle, or see the discussion in [4]. This shift is deterministic (\( \eta_0 \) determines all \( \eta_i \) with \( i > 0 \)), and (RM) does not hold. This indicates that it may be difficult to find necessary conditions for the large deviation principle.

4.5 Remark. Instead of the condition (CD), which gives a continuous dependence on the past, one can condition on the future and ask for continuous dependence. This is the approach of Takahashi [8]. In the case of a Gibbs measure \( \mu \) for \( \phi \in H_A \), it follows from the formula (3.10) that

\[
\frac{\mu[x_0, x_1, \ldots, x_n]}{\mu[x_1, \ldots, x_n]} \xrightarrow{n \to \infty} j_\mu(x_0 x_1 \ldots) = \frac{h(x_0 x_1 \ldots)}{h(x_1 \ldots)} \lambda^{-1} \exp \{\phi(x_0 x_1 \ldots)\}
\]

as \( n \to \infty \) uniformly for \( x \in \Omega_A \). So \( j_\mu \) is continuous and it follows from [9] that

\[
K(\nu) = -h_\nu + (\nu, -\log j_\mu)
\]

which agrees with (3.7) and (3.15). Indeed one can extend Theorem 3.2 to a wider class of Gibbs measures if one uses the results of [5], exercise 2, p. 97.
Bibliography


Recent IMA Preprints (continued)

241 Harry Kesten, Scaling Relations for 2D-Percolation
242 A. Leizarowitz, Infinite Horizon Optimization for Markov Process with Finite States Spaces
243 Louis H.Y. Chen, The Rate of Convergence In A Central Limit Theorem for Dependent Random Variables with Arbitrary Index Set
244 G. Kallianpur, Stochastic Differential Equations in Duals of Nuclear Spaces with some Applications
245 Tzu-Shuh Chiang, Yunshong Chow and Yuh-Jia Loe, Evaluation of Certain Functional Integrals
246 L. Karp and M. Pinsky, The First Eigenvalue of a Small Geodesic Ball in a Riemannian Manifold
247 Chi-Sing Man, Towards An Acoustoelastic Theory for Measurement of Residual Stress
248 Andreas Stoll, Invariance Principles for Brownian Intersection Local Time and Polymer Measures
249 R.W.R. Darling, Rate of Growth of the Coalescent Set in a Coalescing Stochastic Flow
250 R. Cohen, F. Hardt, D. Kinderlehrer, S.Y. Lin, M. Luskin, Minimum Energy for Liquid Crystals: Computational Results
251 Suzanne M. Lanher, Viscosity Solutions for Weakly Coupled Systems of Properties Order PDE's
252 M. Cranston, E. Fabes, Z. Zhao, Condition Gauge and Potential Theory for the Schr?dinger Operator
253 H. Brezis, J.M. Coron, E.H. Lieb, Harmonic Maps with Defects
254 A. Carverhill, Flows of Stochastic Dynamical Systems: Nontriviality of the Lyapunov Spectrum
255 A. Carverhill, Constructing a 'Lifted' Stochastic System in a Product Case
256 R.J. Williams, Local Time and Excursions of Reflected Brownian Motion
257 H. Follmer, S. Orey, Large Deviations for the Empirical Field of a Gibbs Measure
258 A. Leizarowitz, Characterization of Optimal Trajectories On an Infinite Horizon
259 Y.Giga, T. Miyakawa, H. Osada, Two Dimensional Navier Stokes Flow with Measures As Initial Vorticity
261 V. Perez-Abreu, Decompositions of Semimartingales On Duals of Countably Nuclear Spaces
262 J.M. Ball, Does Rank-One Convexity Imply Quasiconvexity?
264 K.A. Pericak-Spector, On Radially Symmetric Simple Waves in Elasticity
265 P.N. Shivakumar, Chi-Sing Man, Simon W. Robkin, Modelling of the Heart and Pericardium at End-Diastole
266 Jose-Luis Menaldi, Probabilistic View of Estimates for Finite Difference Methods
267 Robert Hardt, Harold Rosenberg, Open Book Structures and Unicity of Minimal Submanifolds
268 Bernardo Cockburn, The Quasi-Monotone Schemes for Scalar Conservation Laws
269 H.R. Jauslin, W. Zimmermann, Jr., Dynamics of a Model for an Ac Josephson Effect in Superfluid 4He
270 A.K. Kapila, Introductory Lecture on Reacting Flows
271 J.C. Taylor, Do Minimal Solutions of Heat Equations Characterize Diffusions?
272 J.C. Taylor, The Minimal Eigenfunctions Characterize the Ornstein-Uhlenbeck Process
273 Chi-Sing Man, Quan-Xin Sun, On the Significance of Normal Stress Effects in the Flow of Glaciers
274 Omar Hijab, On Partially Observed Control of Markov Processes
275 Lawrence Gray, The Behavior of Processes with Statistical Mechanical
276 R. Hardt, D. Kinderlehrer, M. Luskin, Remarks About the Mathematical Theory of Liquid Crystals

278 H. Chipot, T. Sideris, On the Abelian Higgs Model
279 C. Foles, B. Nicolaenko, G.R. Sell, R. Temam, Inertial Manifolds for the Kuramoto-Sivashinsky Equation and an Estimate of their Lowest Dimension
280 R. Duran, On the Approximation of Miscible Displacement in Porous Media by a Method of Characteristics Combined with a Mixed Method
282 V. Tversky, Dispersion Bulks Parameters for Coherent Propagation in Correlated Random Distributions
283 W.Th.F. den Hollander, Mixing Properties for Random Walk in Random Scenery
284 H.R. Jauslin, Non-differentiable Potentials for Nonequilibrium Steady State
285 K. Hoyer, G.R. Sell, Homoclinic Orbits and Bernoulli Bundles in Almost Periodic Systems
286 J. Douglas, Jr., Y. Yuan, Finite Difference Methods for the Transient Behavior of a Semiconductor Device
287 Li Kaitai, Yan Ningning, The Extrapolation for Boundary Finite Elements
288 R. Durrett, R.H. Schonmann, Stochastic Growth Models
289 David Kinderlehrer, Remarks about Equilibrium Configurations of Crystals
290 D.G. Aronson, J.L. Vazquez, Eventual C-R Regularity and Convexity for Flow In One-Dimensional Porous Media
291 L.R. Scott, J.M. Boyle, B. Bagheri, Distributed Data Structures for Sclen Computation
293 L.A. Caffarelli, Optimal Regularity for One-Dimensional Por Medium Flow
294 Haim Brezis, Liquid Crystals and Energy Estimates for S2-Valued Maps