Stochastics, 1987, Vol. 22, pp. 187–199
Photocopying permitted by license only

Printed in Great Britain

A Remark on the Large Deviations of an Ergodic Markov Process

W. H. FLEMING*
Division of Applied Mathematics, Brown University, Providence, RI 02912, USA

S.-J. SHEU†
Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan, Republic of China

and

H. M. SONER*
Department of Mathematics, Carnegie–Mellon University, Pittsburgh, PA 15213, USA

(*Received 12 November 1986; in final form 31 March 1987*)

We give an alternative proof of the Donsker–Varadhan result for the large time behavior of the occupation measure μ_t of a Feller–Markov process $w(t)$, taking values in a compact metric space X. For any weak* continuous functions $\phi(\mu)$ on the space of probability measures on X, we identify $\lim_{t\to\infty} t^{-1} \ln E\exp(\phi(\mu_t))$ through a linearization argument.

KEY WORDS: Ergodic Markov process, Donsker–Varadhan result, linearization argument.

*Partly supported by AFOSR under contract no. 91-0116-0, and by ONR under contract no. N0014-83-K-0542, and by the Institute for Mathematics and its Applications with funds provided by the NSF and ONR.

†Partly supported by National Science Council of Republic of China, and by NSF under Grant no. MCS 8121940.

1. INTRODUCTION

Let $\{w(t), t \geqq 0\}$ be a Feller–Markov process taking values in a compact metric space X. Suppose that there is an invariant measure μ^* corresponding to this process. Then, under some conditions (see (2.1)–(2.3)), the occupation measure $\mu_t(A) = (1/t)\int_0^t X_A(w(s))\,ds$, where X_A is the indicator function of any Borel subset A of X, converges to μ^* exponentially fast. More precisely, one has the well-known result, Theorem 1.1 below, due to M.D. Donsker and S. R. S. Varadhan [2].

We introduce the following notation. Let $C(X)$ denote the space of continuous functions on $X, \mathscr{B}(X)$ the Borel subsets of X, and $P(X)$ the space of probability measures on $(X, \mathscr{B}(X))$. Let L be the infinitesimal generator of the process, $\mathscr{D}^+(L)$ be the space of positive functions which are in the domain of L and $\langle \mu, f\rangle = \int_X f(x)\mu(dx)$. Then the Donsker–Varadhan rate function $I(\mu)$ is given by

$$I(\mu) = \sup_{u \in \mathscr{D}^+(L)} -\left\langle \mu, \frac{Lu}{u}\right\rangle.$$

THEOREM 1.1 *For any weak*-continuous function ϕ on $P(X)$ we have,*

$$\lim_{t\to\infty} \frac{1}{t} \ln E_x \exp(t\phi(\mu_t)) = -\inf_{\mu \in P(X)} [I(\mu) - \phi(\mu)]. \tag{1.1}$$

Moreover, there are extensions of this theorem for the empirical distributions of the process and non-compact X, [2, 12]. In fact, Theorem 1.1 follows from these 'higher level' results via the contaction principle. Also, different approaches to these problems are developed by J. Gärtner [5] and D. Stroock [11]. J. Gärtner made use of the spectral radius function $\lambda(k) = \lim_{t\to\infty}(1/t)\ln E \exp(t\langle k, \mu_t\rangle)$ for $k \in C(X)$. He proved Theorem 1.1 under the assumption that $\lambda(k)$ is Gateaux-differentiable. Notice that, $\lambda(k)$ is the dominant eigenvalue of the operator $L + k$, if there is one. Also, the differentiability of $\lambda(k)$ follows from the uniqueness of the positive eigenfunction. To this end, we note that the existence and the uniqueness of such an eigenfunction is just a simple consequence of the Assumptions (2.1) and (2.2), see Theorem 4.1. Finally, we refer to [1] and [7] for related problems.

In this paper, we give another proof of Theorem 1.1, via a "linearization" argument. The outline of our proof is as follows. A straightforward argument yields an upper bound. Then, using the measure transformation introduced by S. J. Sheu [9] we obtain the following lower bound,

$$\liminf_{t\to\infty}\frac{1}{t}\ln E\exp(t\phi(\mu_t)) \geqq -\inf_{u\in\mathscr{D}^+(L)}[-\phi(\mu^u)+I(\mu^u)],$$

where μ^u is the invariant measure of the Feller–Markov process generated by the operator $L^u(f)=(1/u)[L(uf)-fL(u)]$. When $\phi(u)=\langle\mu,k\rangle$ for some $k\in C(X)$, the map $\mu\to I(\mu)-\phi(\mu)$ is minimized at μ^{u^*} where u^* is the positive eigenfunction corresponding to $L+k$. Consequently, the upper and the lower bounds are the same for linear ϕ. To complete the proof we reduce the problem to the linear case. To do so, we consider functions of the type $\phi(\mu)=F(\langle\mu,f_1\rangle,\dots,\langle\mu,f_k\rangle)$ for some $F\in C^1(R^k)$ and $f_i\in C(X)$. Then, $\inf[I(\mu)-\phi(\mu)]=\inf_{c\in R^k}[\bar{I}(c)-F(c)]$ where $\bar{I}(c)=\inf\{I(\mu):(\langle\mu,f_1\rangle,\dots,\langle\mu,f_k\rangle)=c\}$. Note that $\bar{I}$ is a convex function on R^k and the minimum of a convex plus a C^1-function is the same as the minimum of the same convex function plus a properly chosen linear function, see Lemma 6.3. Thus, the problem reduces to the linear case.

The paper is organized as follows: notation is introduced in Section 2, in Section 3 the measure transformation is defined, the eigenvalue problem is studied in Section 4, Section 5 is devoted to the upper and lower bounds, finally Theorem 1.1 is proved in Section 6.

2. PRELIMINARIES

Let L be the infinitesimal generator of a Markov process $w(t)$, and T_t be the semigroup on $C(X)$, generated by L. Since we assume that $w(t)$ is a Feller–Markov process, L and T_t satisfy the following (see [4])

$$\left.\begin{array}{ll}\text{a)} & T_t: C(X) \to C(X), \\ \text{b)} & \mathscr{D}(L), \text{ the domain of } L, \text{ is dense in } C(X), \\ \text{c)} & L \text{ satisfies the maximum principle, i.e. if } f \in \mathscr{D}(L) \\ & \text{and,} \\ & \qquad f(x_0) = \max f(x), \text{ then } Lf(x_0) \leqq 0, \\ \text{d)} & \text{Range } (\lambda - L) = C(X) \text{ for some } \lambda > 0. \end{array}\right\} \tag{2.1}$$

We also assume that there is $v \in P(X)$ and $p(t, x, y)$ such that

$$\left.\begin{array}{ll}\text{a)} & T_t f(x) = \int_X f(y) p(t, x, y) v(dy), \quad \forall f \in C(X), \\ \text{b)} & \text{there are } a(t), A(t) \text{ such that} \\ & 0 < a(t) \leqq p(t, x, y) \leqq A(t) \quad \forall x, y \in X \text{ and } t > 0, \\ \text{c)} & \lim_{x \to x_0} \| p(t, x, \cdot) - p(t, x_0, \cdot) \|_{L^1(X, v)} = 0, \end{array}\right\} \tag{2.2}$$

where $L^1(X, v)$ is the set of functions on X, which are integrable with respect to v. We note that condition (2.2) is also used in [2.I].

Let Ω_x be the space of right continuous functions $w(t) \in X$ with $w(0) = x$, which also have left limits at every $t \geqq 0$. Note that, the Markov process $w(t)$ induces a measure P_x on Borel subsets of Ω_x. Finally, for $t > 0$, $w \in \Omega_x$ and $A \in \mathscr{B}(X)$, we define the occupation time by,

$$\mu_t(A, w) = \frac{1}{t} \int_0^t X_A(w(s))\, ds,$$

where X_A is the indicator of A. The probability measure $\mu_t(w) = \mu_t(\cdot, w)$ is called the occupation measure.

3. THE MEASURE TRANSFORMATION

Let $\mathscr{D}^+(L) = \{u \in \mathscr{D}(L): u > 0\}$. For each $u \in \mathscr{D}^+(L)$ we define on operator L^u by

$$L^u f = \frac{1}{u}[L(uf) - f Lu], \tag{3.1}$$

with the domain $\mathscr{D}(L^u)=\{f\in C(X): uf\in\mathscr{D}(\mathscr{L})\}$. Notice that, L^u satisfies (2.1) (b)–(c), and consequently generates a Feller–Markov process (see [4]). Let P_x^u be the measure induced by this process on Ω_x. In [9], S.-J. Sheu showed that P_x^u is absolutely continuous with respect to P_x on $\mathscr{F}_t=\sigma(w(s):s\leq t)$, and

$$\frac{dP_x^u}{dP_x}(w)\bigg|_{\mathscr{F}_t}=\frac{u(w(t))}{u(x)}\exp\left[-t\left\langle\mu_t(w),\frac{Lu}{u}\right\rangle\right]. \tag{3.2}$$

Next we shall show that the condition (2.2) is also satisfied by the semigroup $T_t^u f(x)=E_x^u f(w(t))$, where E_x^u denotes the mathematical expectation with respect to P_x^u.

LEMMA 3.1 *T_t^u satisfies* (2.2).

Proof Using (3.1) we obtain the following

$$T_t^u f(x)=\int_X f(y)p^u(t,x,y)\nu(dy),$$

where

$$p^u(t,x,y)=p(t,x,y)\frac{u(y)}{u(x)}E_x\left[\exp\left(-t\left\langle\mu_t,\frac{Lu}{u}\right\rangle\right)\bigg|w(t)=y\right].$$

Due to the above representation, (2.2) (a) and (b) hold. Now, approximate p^u by

$$p_\varepsilon^u(t,x,y)=p(t,x,y)\frac{u(y)}{u(x)}g_\varepsilon(t,x,y),$$

where

$$g_\varepsilon(t,x,y)=E_x\left[\exp\left(-\int_\varepsilon^t\frac{Lu}{u}(\omega(s))\right)ds\bigg|\omega(t)=y\right].$$

Then, it is easy to show that

$$g_\varepsilon(t,x,y)=\frac{1}{p(t,x,y)}\int p(\varepsilon,x,z)g_0(t-\varepsilon,z,y)p(t-\varepsilon,z,y)\nu(dz).$$

Therefore, $g_\varepsilon(t,x,\cdot)\in L^1(\nu)$ is continuous in x. One completes the proof, after observing that $p^u_\varepsilon(t,x,\cdot)$ converges to $p^u(t,x,\cdot)$ in L^1, uniformly in x, as $\varepsilon\to 0$. □

Now observe that (2.2) implies that the corresponding Markov process is ergodic (see [3]). This observation together with the above lemma yield that for each $u\in\mathscr{D}^+(L)$ there is a unique invariant measure μ^u corresponding to $\{P^u_x : x\in X\}$, i.e. $\langle\mu^u, L^u f\rangle=0$ for all $f\in\mathscr{D}(L^u)$. The following result indicates the relation between this transformation and the rate function.

LEMMA 3.2 *$I(\mu)=-\langle u, Lu/u\rangle$ for some $u\in\mathscr{D}^+(L)$ if and only if $\mu=\mu^u$.*

Proof First, suppose that $I(\mu)=-\langle\mu, Lu/u\rangle$. Consider the map $f\to J(f)=-\langle\mu, Lf/f\rangle$ for $f\in\mathscr{D}^+(L)$. Since J is maximized at u, we have $(\partial/\partial\varepsilon)J(u+\varepsilon h)|_{\varepsilon=0}=0$ for all $h\in\mathscr{D}(L)$. But, $(\partial/\partial\varepsilon)J(u+\varepsilon h)|_{\varepsilon=0}=-\langle\mu, L^u(h/u)\rangle$. Therefore, $\mu=\mu^u$.

To prove the converse, observe that $T^u_t f(x)\leqq\log T^u_t e^f(x)$ for all $f\in C(X)$. Therefore,

$$\frac{1}{t}(T^u_t-I)f(x)\leqq\frac{1}{t}\log(T^u_t e^f(x)/e^f(x))$$

and

$$-\left\langle\mu^u,\frac{1}{t}\log\left(\frac{T^u_t e^f}{e^f}\right)\right\rangle\leqq 0,\quad\forall f\in C(X)\text{ and }t>0. \tag{3.3}$$

Also,

$$\frac{1}{t}\log\frac{T_t v}{v}=\frac{1}{t}\int_0^t\frac{L^u T^u_s v}{T^u_s v}\,ds$$

and as t tends to zero, it converges to $L^u v/v$ for all $v\in\mathscr{D}^+(L^u)$. Thus, $-\langle\mu^u, L^u v/v\rangle\leqq 0$. The following calculation together with this observation complete the proof of the lemma.

$$\begin{aligned} I(\mu^u)&=\sup_{v\in\mathscr{D}^+(L)}-\left\langle\mu^u,\frac{Lv}{v}\right\rangle\\ &=\sup_{v\in\mathscr{D}^+(L)}\left[-\left\langle\mu^u,\frac{L^u(v/u)}{v/u}\right\rangle\right]-\left\langle\mu^u,\frac{Lu}{u}\right\rangle\end{aligned}$$

□

4. THE EIGENVALUE PROBLEM

The following result follows from [6]. But under the Assumptions (2.1) and (2.2) a simpler proof is available.

THEOREM 4.1 *For any* $k \in C(X)$ *there is a solution* $(u^k, \lambda(k)) \in \mathscr{D}^+(L) \times R$ *of the following equation*

$$Lu^k(x) + k(x)u^k(x) = \lambda(k)u^k(x). \tag{4.1}$$

Moreover, the eigenfunction u^k *is unique up to a multiplicative constant.*

Proof Let $\hat{T}_t$ be the semigroup generated by $L+k$. Then,

$$\hat{T}_t f(x) = \int_X f(y)G(t,x,y)\nu(dy) \tag{4.2}$$

where

$$G(t,x,y) = p(t,x,y)E_x\left[\exp \int_0^t k(w(s))\,ds \,\middle|\, w(t) = y\right].$$

Proceeding as in the proof of Lemma 3.1, one can show that $G(t,x,y)$ satisfies (2.2). This implies the existence (Krasnoselski [8], Theorem 2.8, p. 72), and the uniqueness ([8], Theorem 2.10, p. 76) of $(u_t^k, \lambda_t(k)) \in C(X) \times R$ satisfying,

$$\left.\begin{array}{ll} \text{i)} & \hat{T}_t u_t^k(x) = \lambda_t(k)u_t^k(x), \quad \forall\, t > 0 \\ \text{ii)} & u_t^k > 0, \quad \lambda_t(k) > 0 \\ \text{iii)} & \max[u_t^k(x) : x \in X] = 1. \end{array}\right\} \tag{4.3}$$

Set $u_1 = u_1^k, \lambda_1 = \lambda_1(k)$. Then, $\hat{T}_1\hat{T}_s u_1 = \hat{T}_s\hat{T}_1 u_1 = \lambda_1 \hat{T}_s u_1$. Due to the uniqueness of $u_t^k, \hat{T}_s u_1 = c_s u_1$ for some constant $c_s > 0$. Also, $c_{s+t}u_1 = \hat{T}_{s+t}u_1 = \hat{T}_s\hat{T}_t u_1 = c_s c_t u_1$. Therefore, by strong continuity $c_s = e^{\lambda s}$ for some $\lambda \in R$. Moreover,

$$\lim_{t \to \infty} t^{-1}(\hat{T}_t - I)u_1 = \lambda u_1.$$

From the Feynman–Kac representation of $\hat{T}_t$, (4.1), and the right-continuity of sample paths at $t=0$, one concludes that $\mathscr{D}(L)=\mathscr{D}(L+k)$. Then $u_1 \in \mathscr{D}^+(L)$ and the pair (u_1, λ) is a solution of (4.1).

The last assertion of the Theorem follows from the uniqueness of u_t^k. □

5. UPPER AND LOWER BOUNDS

The following theorem is proved in [2], but for completeness we give the proof.

THEOREM 5.1 *For all weak* continuous functions ϕ on $P(X)$ we have*

$$\limsup_{t\to\infty} \frac{1}{t} \ln E_x \exp(t\phi(\mu_t)) \leqq -\inf_{\mu\in P(X)} [I(\mu)-\phi(\mu)] \tag{5.1}$$

$$\liminf_{t\to\infty} \frac{1}{t} \ln E_x \exp(t\phi(\mu_t)) \geqq -\inf_{u\in\mathscr{D}^+(L)} [I(\mu^u)-\phi(\mu^u)] \tag{5.2}$$

where $I(\mu)$ is as in Theorem 1.1.

Proof For any $u\in\mathscr{D}^+(L)$, (3.1) yields

$$E_x \exp[t\phi(\mu_t)] = E_x^u \frac{u(x)}{u(w(t))} \exp\left[t\left(\phi(\mu_t) + \left\langle \mu_t, \frac{Lu}{u}\right\rangle\right)\right]. \tag{5.3}$$

Choose a neighborhood $U_\varepsilon(\mu^u)$ of μ^u such that $|\phi(\mu)-\phi(\mu^u)|<\varepsilon$ and $|\langle \mu, Lu/u\rangle - \langle \mu^u, Lu/u\rangle| \leqq \varepsilon$, for all $\mu\in U_\varepsilon(\mu^u)$. Then, (5.3) yields that

$$E_x \exp[t\phi(\mu_t)] \geqq E_x^u\left[\frac{u(x)}{u(w(t))} \exp\left[t\left(\phi(\mu_t) + \mu_t, \frac{Lu}{u}\right)\right]; \mu_t\in U_\varepsilon(\mu^u)\right]$$

$$\geqq \exp\left[t\left(\phi(\mu^u) + \left\langle \mu^u, \frac{Lu}{u}\right\rangle - 2\varepsilon\right)\right] E_x^u\left[\frac{u(x)}{u(w(t))}; \mu_t \in U_\varepsilon(\mu^u)\right], \tag{5.4}$$

where $E_x^u[\ldots.; A]$ denotes the integral over the set A. Recall that $-\langle \mu^u, Lu/u\rangle = I(\mu^u)$ and $P_x^u(\mu_t\in U_\varepsilon(\mu^u))$ converges to one. Therefore, (5.4) yields the lower bound.

Let $\mathscr{C}=\{C_j: j=1,\ldots,K\}$ be a finite open covering of $P(X)$ and $\mathscr{D}_\varepsilon^+(L)=\{u\in\mathscr{D}^+(L): \varepsilon\leqq u(x)\leqq 1/\varepsilon\}$. For each $(u_1,\ldots,u_K)\in\mathscr{D}_\varepsilon^+(L)$, we have

$$\frac{1}{t}\ln E_x\exp[t\phi(\mu_t)]\leqq\frac{1}{t}\ln\sum_{j=1}^{K}E_x[\exp(t\phi(\mu_t));\mu_t\in C_j]$$

$$\leqq\frac{1}{t}\ln\left[K\sup_j E_x^{u_j}\left[\frac{u_j(x)}{u_j(w(t))}\exp t\left(\phi(\mu_t)\right.\right.\right.$$

$$\left.\left.\left.+\left\langle\mu_t,\frac{Lu_j}{u_j}\right\rangle\right);\mu_t\in C_j\right]\right]$$

Therefore,

$$\frac{1}{t}\ln E_x\exp[t\phi(\mu_t)]\leqq$$

$$\frac{1}{t}\ln\left(\frac{K}{\varepsilon^2}\right)+\sup_j\inf_{u\in\mathscr{D}_\varepsilon^+(L)}\sup_{\mu\in C_j}\left[\phi(\mu)+\left\langle\mu,\frac{Lu}{u}\right\rangle\right]$$

Since the above inequality holds for every $t>0$, $\varepsilon>0$ and open covering $\mathscr{C}$, we obtain

$$\limsup_{t\to\infty}\frac{1}{t}E_x\exp[t\phi(\mu_t)]\leqq\inf_{\{C_j\}}\sup_j\inf_{u\in\mathscr{D}^+(L)}\sup_{\mu\in C_j}\left[\phi(\mu)+\left\langle\mu,\frac{Lu}{u}\right\rangle\right]\tag{5.5}$$

In the rest of the proof, we will show that the right-hand side of (5.5) is in fact equal to $\sup_{\mu\in P(X)}[\phi(\mu)-I(\mu)]=l$. Given $\varepsilon>0$ and $\mu\in P(X)$ there is $u_\mu\in C(X)$ such that $\phi(\mu)+\langle\mu,Lu_\mu/u_\mu\rangle)\leqq l+\varepsilon$. Since the map $\mu\to\langle\mu,Lu/u\rangle$ is continuous for each $u\in\mathscr{D}^+(L)$, there is an open neighborhood N_μ of every $\mu\in P(X)$ such that $\phi(\eta)+\langle\eta,Lu_\mu/u_\mu\rangle\leqq l+2\varepsilon$ for all $\eta\in N_\mu$. Pick a finite covering $\{N_{\mu_i}:i=1,\ldots,K\}$.

The construction of N_{μ_i}'s implies that

$$\sup_i \inf_{u\in\mathscr{D}^+(L)} \sup_{\mu\in N_{\mu_i}} \left[\phi(\mu)+\left\langle\mu,\frac{Lu}{u}\right\rangle\right]$$

$$\leqq \sup_i \sup_{\mu\in N_{\mu_i}} \left[\phi(\mu)+\left\langle\mu,\frac{Lu_{\mu_i}}{u_{\mu_i}}\right\rangle\right] < l+2\varepsilon$$

Therefore, the left-hand side of (5.5) is less than l. □

6. PROOF OF THEOREM 1.1

We will first prove the theorem when $\phi(\mu)=\langle\mu,k\rangle$.

LEMMA 6.1 *Let $k\in C(X)$ and $\mu^*\in P(X)$ be such that*

$$-I(\mu^*)+\langle\mu^*,k\rangle=\sup_{\mu\in P(X)}\left[-I(\mu)+\langle\mu,k\rangle\right]$$

Then, $\mu^*=\mu^{u^k}$ where u^k is the positive eigenfunction of the operator $L+k$.

Proof The choice of μ^*, Lemma 3.2 and Eq. 4.1 yield

$$\lambda(k)=\left\langle\mu^*,\frac{Lu^k}{u^k}+k\right\rangle \geqq -I(\mu^*)+\langle\mu^*,k\rangle$$

$$\geqq -I(\mu^{u^k})+\langle\mu^{u^k},k\rangle=\left\langle\mu^{u^k},\frac{Lu^k}{u^k}+k\right\rangle=\lambda(k)$$

Therefore, $I(\mu^*)=-\langle\mu^*,Lu^k/u^k\rangle$ and Lemma 3.2 implies that $u^*=\mu^{u^k}$. □

Remark 6.2 In view of Theorem 5.1, the above result proves (1.1) for linear ϕ. But, in the linear case a sharper asymptotic formula is available. That is,

$$\lim_{t\to\infty} E_x \exp\left[\int_0^t k(w(s))\,ds\right]e^{-\lambda(k)t}=u^k(x)\left\langle\mu^{u^k},\frac{1}{u^k}\right\rangle. \tag{6.1}$$

One proves (6.1) by choosing $u=u^k$ in (5.3).

We need one final result in the proof of Theorem 1.1.

LEMMA 6.3 *Let $\bar{I}$ be a real-valued, convex, lower-semicontinuous function on R^k and F be a real valued, bounded, continuously differentiable function on R^k. Then,*

$$\sup_{c\in R^k}[F(c)-\bar{I}(c)]=\sup_{c\in R^k}[F(c^*)+\nabla F(c^*)\cdot(c-c^*)-\bar{I}(c)],$$

where ∇ denotes the gradient and $c^*\in R^k$ is an extremal point of $F(c)-\bar{I}(c)$, i.e.

$$F(c^*)-\bar{I}(c^*)=\sup_{c\in R^*}[F(c)-\bar{I}(c)].$$

Proof Suppose to the contrary. Then, there are $c_0\in R^k$ and $\delta>0$ such that

$$F(c^*)+\nabla F(c^*)\cdot(c_0-c^*)-\bar{I}(c_0)=F(c^*)-\bar{I}(c^*)+\delta.$$

For each $\tau\in[0,1]$ we have,

$$\begin{aligned}
&F(c^*)+\nabla F(c^*)\cdot(c_0-c^*)\tau-I(c^*+\tau(c_0-c^*))\\
&\geqq(\tau-1)\bar{I}(c^*)-\tau\bar{I}(c_0)+F(c^*)+\nabla F(c^*)\cdot(c_0-c^*)\tau\\
&=F(c^*)-\bar{I}(c^*)+\tau\delta\\
&\geqq F(c^*+\tau(c_0-c^*))-\bar{I}(c^*+\tau(c_0-c^*))+\tau\delta
\end{aligned}$$

The above inequality yields,

$$\nabla F(c^*)\cdot(c_0-c^*)\geqq\frac{1}{\tau}[F(c^*+\tau(c_0-c^*))-F(c^*)]+\delta.$$

But this inequality contradicts the differentiability of F at c^*. □

Proof of Theorem 1.1 Since the class of functions ϕ of the form $\phi(\mu)=F(\langle\mu,f_1\rangle,\ldots,\langle\mu,f_K\rangle)$ for some $f_i\in C(X)$ and F as in Lemma 6.3 is dense in $C(P(X))$, it suffices to prove Theorem 1.1 for this class of functions. Pick $\mu^* \in P(X)$ such that $\phi(\mu^*) - I(\mu^*) = \sup\{\phi(\mu) - I(\mu): \mu \in P(X)\}$. We will show that $\mu^* = \mu^u$ for some $u \in C(X)$ and in view of Theorem 5.1 this implies (1.1). For $c \in R^K$ define $\bar{I}(c)$ by

$$\bar{I}(c) = \inf\{I(\mu): \mu \in P(X) \quad \text{and} \quad (\langle \mu, f_1 \rangle, \ldots, \langle \mu, f_K \rangle) = c\},$$

(we use the convention that inf over an empty set is $+\infty$). Therefore,

$$\phi(\mu^*) - I(\mu^*) = F(c^*) - \bar{I}(c^*) = \sup_{c \in R^K} [F(c) - \bar{I}(c)]. \tag{6.2}$$

where $c^* = (\langle \mu^*, f_1 \rangle, \ldots, \langle \mu^*, f_K \rangle)$. Now, Lemma 6.3 and the definition of $\bar{I}$ yield

$$\begin{aligned} \sup_{c \in R^K} [F(c) - \bar{I}(c)] &= \sup_{c \in R^K} [F(c^*) + \nabla F(c^*) \cdot (c - c^*) - \bar{I}(c)] \\ &= \sup_{\mu \in P(X)} [\langle \mu, k \rangle - I(\mu)] \\ &= \langle \mu^*, k \rangle - I(\mu^*). \end{aligned}$$

where

$$k(x) = F(c^*) + \nabla F(c^*) \cdot [(f_1(x), \ldots, f_K(x)) - c^*].$$

Therefore, Lemma 6.1 implies that $\mu^* = \mu^{u^k}$. $\square$

Remark 6.4 The technique, developed in this paper, also applies to discrete-time Markov processes. In this case, the Donsker–Varadhan rate function $I(\mu)$ is given by

$$I(\mu) = \sup_{u \in C^+(X)} - \left\langle \mu, \ln \frac{\pi u}{u} \right\rangle$$

where $C^+(X)$ is the set of strictly positive continuous functions on X, and $\pi u(x) = E_x u(w(1))$. Then, the analogue of the transformation (3.1) is,

$$\pi^u f(x) = \frac{\pi(uf)(x)}{\pi u(x)}, \quad u \in C^+(X).$$

The only other change is the form of the eigenvalue problem (4.1), that is: for $k \in C(X)$ one needs to study the following equation:

$$\pi u^k(x) = e^{\lambda(k)k(x)}u(x).$$

Acknowledgements

The authors wish to thank L. Stettner for providing a simpler proof of Theorem 4.1, and D. Stroock for valuable comments and in particular for pointing out to us Gärtner's paper [5].

References

[1] R. Azencott, *Grandes deviations et applications*, Springer Lecture Notes in Math. No. 774, 1978.

[2] M. D. Donsker and S. R. S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I, II, III, *Comm. Pure Appl. Math.* **27** (1975), 1–47, **28** (1975), 279–301, **29** (1976), 389–461.

[3] J. Doob, *Stochastic Processes*, John Wiley & Sons, New York, 1952.

[4] S. N. Either and T. G. Kurtz, *Markov Processes: Existence and Approximation*, John Wiley & Sons, New York, 1986.

[5] J. Gärtner, On large deviations from the invariant measure, *Theory of Probability and its Applications* **22**(1) (1977), 24–39.

[6] G. Greiner, Spectrum and asymptotic stark stetiger Halbgruppen positiven operatoren, *Sitzungsber, Heidelberg Akad. d. Wiss.* (math.-naturw. Klasse), **3**, Abhandlung (1982).

[7] W. H. Fleming, Logarithmic transformations and stochastic control, *Advances in Filtering and Optimal Stochastic Control* (W. H. Fleming and L. G. Gorostiza, eds), Springer Lecture Notes in Control and Information Science, 1982, pp. 131–141.

[8] M. A. Krasnoselskii, *Positive Solutions of Operator Equations*, P. Noordhoff, Gronigen, The Netherlands, 1964.

[9] S.-J. Sheu, Stochastic control and principal eigenvalue, *Stochastics* **11** (1984), 191–211.

[10] S.-J. Sheu, On the solution of the dynamic programming equation of a discounted control problem associated with a recurrent Markov process, preprint.

[11] D. Stroock, *An Introduction to the Theory of Large Deviations*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[12] S. R. S. Varadhan, *Large Deviations and Applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, 1984.