Acta Mathematica Sinica, New Series
1990, Vol.6, No.3, pp. 206 — 219

On Evaluating the Rate Function of Large Deviations for Jump Processes

Chen Mufa (陈木法) Lu Yungang (卢云刚)
Department of Mathematics, Beijing Normal University

Received January 25, 1988

Abstract. This is a sequel to our joint paper[4] in which upper bound estimates for large deviations for Markov chains are studied. The purpose of this paper is to characterize the rate function of large deviations for jump processes. In particular, an explicit expression of the rate function is given in the case of the process being symmetrizable.

1. Introduction

Let $(E, \mathscr{E})$ be an arbitrary measurable space with the property that all the singletons $\{x\}(x \in E)$ belong to $\mathscr{E}$. Denote by ${}_b\mathscr{E}$ the Banach space of all bounded $\mathscr{E}$-measurable real functions with the uniform norm $\|f\| = \sup_{x \in E} |f(x)|$, and denote by $\mathscr{P}(E)$ the set of all probability measures on $(E, \mathscr{E})$. Let $P(t, x, dy)$ $(t \geqslant 0, x \in E)$ be a transition function on $(E, \mathscr{E})$ and $\{T_t\}$ be the semigroup on ${}_b\mathscr{E}$ determined by $P(t, x, dy)$. Let L be the generator of $\{T_t\}$ and denote its domain by $\mathscr{D}$. The rate function (I-function) in Donsker-Varadhan theory[6, 7] is denoted for $\mu \in \mathscr{P}(E)$ by

$$I(\mu) = -\inf_{f \in \mathscr{D}^+} \int_E \frac{Lf}{f}(x)\,\mu(dx),$$

where $\mathscr{D}^+$ consists of the strictly positive functions in $\mathscr{D}$. Since the domain $\mathscr{D}$ is generally not known explicitly, it is interesting to find an explicit expression for the rate functions of some special processes. This has been done for some diffusion processes[8, 10] and is studied in this note for jump processes.

Let us recall some definitions. We say $q(x)$-$q(x, dy)$ is a q-pair if

$$q(\cdot) \text{ and } q(\cdot, A) \text{ are } \mathscr{E}\text{-measurable for each } A \in \mathscr{E};$$
$$q(x, \cdot) \text{ is a nonnegative measure on } \mathscr{E} \text{ for each } x \in E,$$

and

$$q(x, A) \leqslant q(x) \text{ for } x \in E \text{ and } A \in \mathscr{E}.$$

Throughout this paper, we assume that the given q-pair $q(x)$-$q(x, dy)$ is regular. That is, the q-pair is conservative which means that

$$q(x, E) = q(x), \text{ for all } x \in E;$$

and it determines uniquely the so-called jump process (or q-process) $P(t, x, dy)$ such that

$$\frac{d}{dt} P(t, x, A)\Big|_{t=0} = q(x, A) - I_A(x) q(x) \quad \text{for } x \in E, \text{ and } A \in \mathscr{E}$$

and

$$\lim_{t \to 0} P(t, x, \{x\}) = 1 \quad \text{for all } x \in E.$$

For a given q-pair, some general uniqueness criteria were obtained in [6], and some more practical sufficient conditions for uniqueness of q-processes were more recently given in [3].

In the context of diffusion processes[8, 10], the evaluation of the rate function is restricted to the set of probability measures with compact support. In our context, we consider a larger set, i. e.,

$$\mathscr{P}(q) = \left\{ \mu \in \mathscr{P}(E) : \int q(x) \mu(dx) < \infty \right\}.$$

In order to state our results, we need more notations:

$\mathscr{E}$ = the set of all real $\mathscr{E}$-measurable functions,

$\mathscr{E}^+ = \{ f \in \mathscr{E} : f$ has a strictly positive lower bound $\}$,

$\mathscr{E}^0 = \{ f \in \mathscr{E} : 0 < f < \infty \}$,

${}_b\mathscr{E}^+ = {}_b\mathscr{E} \cap \mathscr{E}^+$, ${}_b\mathscr{E}^0 = {}_b\mathscr{E} \cap \mathscr{E}^0$.

For $f \in {}_b\mathscr{E}$, we define

$$\Omega f(x) = \int q(x, dy) f(y) - q(x) f(x), \quad x \in E.$$

Theorem 1. *For every $\mu \in \mathscr{P}(q)$, we have*

$$I(\mu) = -\inf_{f \in \mathscr{D}^+} \int \frac{Lf}{f} d\mu \tag{1}$$

$$= -\inf_{f \in {}_b\mathscr{E}^+} \int \frac{\Omega f}{f} d\mu \tag{2}$$

$$= -\inf_{f \in {}_b\mathscr{E}^+} \int f \, \Omega\left(\frac{1}{f}\right) d\mu \tag{3}$$

$$= -\inf_{f \in {}_b\mathscr{E}^0} \int f \, \Omega\left(\frac{1}{f}\right) d\mu \tag{4}$$

$$= -\inf_{f \in \mathscr{E}^+} \int \frac{\Omega f}{f} d\mu \tag{5}$$

$$= -\inf_{f \in \mathscr{E}^0} \int \frac{\Omega f}{f} d\mu. \tag{6}$$

The theorem says that for $\mu \in \mathscr{P}(q)$, we can use $\mathscr{E}^0$ instead of $\mathscr{D}^+$ to compute the rate function $I(\mu)$. The former one is much easier to handle since the domain $\mathscr{D}$ is quite small, and even the indicator $I_{\{x_0\}}$ is usually not in it. To demonstrate this point, we prove the next result which in special cases is already known[9].

Theorem 2. *Let $\mu\in\mathscr{P}(q)$. Then $I(\mu)=0$ if and only if μ is a stationary distribution of $\{T_t\}$.*

Now, we would like to give a clearer expression for $I(\mu)$. For this we need a hypothesis:

(H_1) There exist a σ-finite measure*) λ and an $\mathscr{E}\times\mathscr{E}$-measurable function $q(x,y)$ such that

$$q(x,dy)=q(x,y)\lambda(dy), \quad \text{for all } x, y\in E.$$

Theorem 3. *Under (H_1), for each $\mu\in\mathscr{P}(q)$ satisfying $\mu\ll\lambda$, we have*

$$\begin{aligned} I(\mu)=&\frac{1}{2}\iint\left[\sqrt{\frac{d\mu}{d\lambda}(x)q(x,y)}-\sqrt{\frac{d\mu}{d\lambda}(y)q(y,x)}\right]^2\lambda(dx)\lambda(dy)\\ &-\frac{1}{2}\inf_{f\in\mathscr{C}}\iint\left[\sqrt{\frac{d\mu}{d\lambda}(x)\,q(x,y)\frac{f(y)}{f(x)}}\right.\\ &\left.-\sqrt{\frac{d\mu}{d\lambda}(y)q(y,x)\frac{f(x)}{f(y)}}\right]^2\lambda(dx)\lambda(dy), \end{aligned} \tag{7}$$

where $\mathscr{C}$ denotes one of $\mathscr{D}^+$, ${}_b\mathscr{E}^+$, ${}_b\mathscr{E}^0$, $\mathscr{E}^+$ or $\mathscr{E}^0$.

After taking a look at the above expression, the reader may guess that the second term above should vanish in the symmetrizalbe case. This is correct and will be presented in the next theorem.

Recall that a q-pair is called symmetrizable with respect to a σ-finite measure ν on $\mathscr{E}$ if

$$\int fqg\,d\nu=\int gqf\,d\nu$$

for all $0\leqslant f, g\in\mathscr{E}$, where

$$qf(x)=\int q(x,dy)f(y), \quad x\in E.$$

Because of the uniqueness assumption, it is equivalent to say that the q-process (resp. the semigroup $\{T_t\}$) determined by $q(x)$- $q(x,dy)$ is symmetrizable with respect to ν, that is,

$$\int fT_t g\,d\nu=\int gT_t f\,d\nu, \quad t\geqslant 0, \quad f,g\in{}_b\mathscr{E}.$$

We also point out that in the symmetrizable case, the condition

$$\int\nu(dx)\,q(x)<\infty$$

is sufficient for the uniqueness and quite general. For details of these facts, see [2].

Theorem 4. *Let (H_1) hold. Suppose that (H_2): $q(x)$-$q(x,dy)$ is symmetrizable with respect to ν and*

*) All the measures considered in this paper are assumed to be nonnegative and non-trivial.

$$\frac{dv}{d\lambda} > 0, \qquad \text{a. s. } (\lambda).$$

Then for each $\mu \in \mathscr{P}(q)$ *satisfying* $\mu \ll \lambda$ *on* $E^0 = \{x \in E : q(x) > 0\}$, *we have*

$$I(\mu) = \frac{1}{2} \int_{E^0}\int_{E^0}\left[\sqrt{\frac{d\mu}{d\lambda}(x) q(x,y)} - \sqrt{\frac{d\mu}{d\lambda}(y) q(y,x)}\right]^2 \lambda(dx)\lambda(dy). \quad (8)$$

The next result is an alternative of Theorem 4. Instead of $\mathscr{P}(q)$, we consider the set $\mathscr{P}'(q)$ of $\mu \in \mathscr{P}(E)$ satisfying the following conditions:

(i) $\mu \ll \lambda$ on E^0;

(ii) There exist constants $M \geqslant m > 0$ such that

$$0 < m^2 \leqslant \frac{d\mu}{d\lambda}(x) \bigg/ \frac{dv}{d\lambda}(x) \leqslant M^2, \quad \text{a. s. } (\lambda) \text{ on } E^0;$$

(iii) $$\int_{E^0}\int_{E^0}\left[\sqrt{\frac{d\mu}{d\lambda}(x) q(x,y) M/m} - \sqrt{\frac{d\mu}{d\lambda}(y) q(y,x) m/M}\right]^2 \lambda(dx)\lambda(dy) < \infty.$$

Theorem 5. *Under* (H_1) *and* (H_2), (8) *holds for every* $\mu \in \mathscr{P}'(q)$.

In the case E being countable, we use Q-matrix $Q = (q_{ij} : i, j \in E)$ instead of the q-pair, and Theorem 4, for example, can be restated as follows:

Corollary. *Let* $Q = (q_{ij})$ *be a regular* Q*-matrix, symmetrizable with respect to* $(\pi_i > 0 : i \in E)$. *Then for each* $\mu \in \mathscr{P}(Q) = \{ v_i : \sum_i v_i q_i < \infty \}$ *we have*

$$I(\mu) = \frac{1}{2} \sum_i \sum_{j \neq i} [\sqrt{\mu_i q_{ij}} - \sqrt{\mu_j q_{ji}}]^2.$$

Our last result is to show how the rate function determines the Q matrix itself.

Theorem 6. *Denote by* $\mathscr{P}_0$ *the set of all probability measures on E with finite supports. Let* $Q = (q_{ij})$ *and* $\overline{Q} = (\overline{q}_{ij})$ *be two regular* Q*-matrices with symmetric measures* $(\pi_i > 0)$ *and* $(\overline{\pi}_i > 0)$ *respectively. Then the following assertions are equivalent.*

(i) $I(\mu) = \overline{I}(\mu)$ *for all* $\mu \in \mathscr{P}_0$.

(ii) $I(\mu) = \overline{I}(\mu)$ *for all* $\mu \in \mathscr{P}(Q) \cap \mathscr{P}(\overline{Q})$,

(iii) $q_i = \overline{q}_i$, $q_{ij} q_{ji} = \overline{q}_{ij} \overline{q}_{ji}$, *for all* $i, j \in E$.

Remark 1. For symmetrizable Markov processes, some characterization of the rate function was obtained by Donsker and Varadhan[7] and Stroock[9]. In the former paper, the square root of L is used, in the second the spectrum representation is used. Hence their expressions for the rate function are not as explicit as ours. To get a more concrete impression, let us consider the case of E being countable. From [9 ; Throem (7.44)and Lemma (7.38)], we have

$$I(\mu) = \lim_{t \downarrow 0} \frac{1}{t} (\varphi - P_t \varphi, \varphi)_{L^2(\pi)}, \qquad \mu \in \mathscr{P}(E).$$

where $\varphi_i = \sqrt{\mu_i / \pi_i}$ $(i \in E)$. Hence, for each $\mu \in \mathscr{P}(E)$, we have

$$I(\mu) = \lim_{t \downarrow 0} \frac{1}{t} \left[\sum_i \pi_i \varphi_i^2 - \sum_i \pi_i \varphi_i \sum_j P_{ij}(t) \varphi_j\right]$$

$$=\lim_{t\downarrow 0}\frac{1}{t}\left[\sum_i \mu_i-\sum_{i,j}\pi_i\sqrt{\frac{\mu_i\mu_j}{\pi_i\pi_j}}\,P_{ij}(t)\right]$$

$$=\lim_{t\downarrow 0}\frac{1}{t}\left[1-\sum_{i,j}\sqrt{\mu_i\mu_j P_{ij}(t)}\cdot\sqrt{\pi_i P_{ij}(t)/\pi_j}\right]$$

$$=\lim_{t\downarrow 0}\frac{1}{t}\left[1-\sum_{i,j}\sqrt{\mu_i\mu_j P_{ij}(t)P_{ji}(t)}\right]$$

$$=\lim_{t\downarrow 0}\frac{1}{2t}\sum_{i,j}[\mu_i P_{ij}(t)-2\sqrt{\mu_i\mu_j P_{ij}(t)P_{ji}(t)}+\mu_j P_{ji}(t)]$$

$$=\lim_{t\downarrow 0}\frac{1}{2t}\sum_{i,j}[\sqrt{\mu_i P_{ij}(t)}-\sqrt{\mu_j P_{ji}(t)}]^2$$

$$\geqslant\frac{1}{2}\sum_{i,j}[\sqrt{\mu_i q_{ij}}-\sqrt{\mu_j q_{ji}}]^2.$$

A further assertion is that if φ is nice enough (for example, it belongs to the domain of the generator L, or its extension), then the above inequality becomes an equality. But our corollary simply says that this happens whenever $\mu\in\mathscr{P}(Q)$.

Remark 2. As we have said above, if $\mu\in\mathscr{P}(E)$, we have

$$I(\mu)\geqslant\frac{1}{2}\sum_i\sum_{j\neq i}(\sqrt{\mu_i q_{ij}}-\sqrt{\mu_j q_{ji}})^2\equiv J(\mu)$$

and hence $J(\mu)=\infty$ implies $I(\mu)=\infty$. A natural question is what happens when $\mu\in\mathscr{P}(E)\setminus\mathscr{P}(Q)$ but $J(\mu)<\infty$. Is it still true that $I(\mu)=J(\mu)$? Unfortunately, we have nothing to say about it at this moment.

Remark 3. As we have said before, the second term of (7) vanishes in the symmetrizable case. However, even not in this case, it can happen also. Here is a trivial example. Take $E=\{0,1,2,\cdots\}$, $-q_{ii}=q_i=q_{i,i+1}=\lambda i$, $i\geqslant 0$ and $q_{ij}=0$ if $j\neq i, i+1$. For $\varepsilon>0$, set $f_0=1, f_{i+1}=\varepsilon f_i/q_{i,i+1}$, $i\geqslant 1$. Then $f\in\mathscr{E}^0$ and

$$\sum_i\sum_{j\neq i}\left[\sqrt{\mu_i q_{ij}f_j/f_i}-\sqrt{\mu_j q_{ji}f_i/f_j}\right]^2=\sum_i\mu_i q_{i,i+1}f_{i+1}/f_i=\varepsilon,$$

hence

$$I(\mu)\leqslant J(\mu)-\varepsilon.$$

Since ε is arbitrary, we obtain $I(\mu)=J(\mu)$.

2. Proof of the Results

In this part, we will prove the results successively.

Proof of Theorem 1.

By the definitions, we immediately get the equalities (1), (3) and (5). We now prove (2).

It is known that the strong generator and the weak generator of $\{T_t\}$ both coincide with the operator Ω defined on ${}_b\mathscr{E}$. Let $\tilde{\mathscr{D}}$ be the domain of the weak generator. From [11; p. 162, Lemma 3] and [2 ; Lemma 6.1.1], it follows that

$$\begin{aligned}\tilde{\mathscr{D}} \subset \mathscr{B}_0 &= \{f \in {}_b\mathscr{E} : \|T_t f - f\| \to 0 \text{ as } t \to 0\} \\ \subset \tilde{\mathscr{B}}_0 &= \{f \in {}_b\mathscr{E} : T_t f(x) \to f(x) \text{ for } x \in E \text{ as } t \to 0\} \\ &= {}_b\mathscr{E}.\end{aligned} \tag{9}$$

Take an $f \in {}_b\mathscr{E}^+$ and put

$$f_n(x) = n\int_0^{1/n} T_t f(x)dt \qquad \text{for } x \in E \text{ and } n \geqslant 1.$$

It follows that $f_n \in \tilde{\mathscr{D}} \cap {}_b\mathscr{E}^+$, $0 < \delta \equiv \inf_{x \in E} f(x) \leqslant f_n \leqslant \|f\|$, $n \geqslant 1$ and $f_n(x) \to f(x)$ for each $x \in E$ as $n \to \infty$. Since

$$\begin{aligned}\left|\frac{\Omega f_n}{f_n}(x)\right| &= \left|q(x) - \frac{1}{f_n(x)}\int q(x, dy) f_n(y)\right| \\ &\leqslant \left(1 + \frac{\|f\|}{\delta}\right) q(x), \qquad x \in E,\end{aligned}$$

$\mu \in \mathscr{P}(q)$, and applying the dominated convergence theorem we get

$$\lim_{n \to \infty} \int \frac{\Omega f_n}{f_n} d\mu = \int \frac{\Omega f}{f} d\mu. \tag{10}$$

Similarly, for each $f \in \mathscr{B}_0 \cap {}_b\mathscr{E}^+$, we can choose a sequence $\{f_n\}_1^\infty \subset \mathscr{D}^+$ such that $\delta \leqslant f_n \leqslant \|f\|$ and $\|f_n - f\| \to 0$ as $n \to \infty$. Again by using the fact that $\mu \in \mathscr{P}(q)$ and the dominated convergence theorem, we have the same conclusion as given in (10). Hence we have proved that

$$\inf_{f \in \mathscr{D}^+} \int \frac{\Omega f}{f} d\mu = \inf_{f \in \mathscr{B}_0 \cap {}_b\mathscr{E}^+} \int \frac{\Omega f}{f} d\mu$$

and

$$\inf_{f \in \tilde{\mathscr{D}} \cap {}_b\mathscr{E}^+} \int \frac{\Omega f}{f} d\mu = \inf_{f \in {}_b\mathscr{E}^+} \int \frac{\Omega f}{f} d\mu.$$

Combining this with

$$\mathscr{D}^+ \subset \tilde{\mathscr{D}} \cap {}_b\mathscr{E}^+ \subset \mathscr{B}_0 \cap {}_b\mathscr{E}^+ \subset {}_b\mathscr{E}^+,$$

we get (2).

In order to prove (4), we fix an $f \in {}_b\mathscr{E}^0$. Without loss of generality we may and will assume that

$$\iint \left[\frac{f(y)}{f(x)} - 1\right] q(x, dy)\, \mu(dx) = \int f\, \Omega\left(\frac{1}{f}\right) d\mu < \infty.$$

This is equivalent to

$$\iint\left[\frac{f(y)}{f(x)}+1\right]q(x,dy)\mu(dx)<\infty . \tag{11}$$

Let $f_n = f + 1/n$, $n \geqslant 1$. Then $f_n \in {}_b\mathscr{E}^+$ and

$$\int f_n \Omega\left(\frac{1}{f_n}\right)d\mu$$

$$=\iint\left[\frac{f(x)}{f(y)}-1\right]f(y)/[f(y)+1/n]\, q(x,dy)\mu(dx).$$

Hence the condition (11) and the dominated convergence theorem give us

$$\lim_{n\to\infty}\int f_n \Omega\left(\frac{1}{f_n}\right)d\mu=\int f\Omega\left(\frac{1}{f}\right)d\mu,$$

which proves the equality (4).

Similarly, one can prove the last equality (6).

Proof of Throrem 2.

This subsection is independent of the remainder of the paper and can be safely skipped.

First of all we assume that $I(\mu)=0$. Since $\mu\in\mathscr{P}(q)$, from Theorem 1, it follows that

$$\int \frac{\Omega f}{f}\, d\mu \geqslant 0 \quad \text{for each } f\in {}_b\mathscr{E}^+ .$$

We now fix an $f\in {}_b\mathscr{E}^+$; then there exists a $\delta>0$ such that $1+\varepsilon f\in {}_b\mathscr{E}^+$ for every $\varepsilon\in[-\delta,\infty)$, and so

$$F_f(\varepsilon)=\int \frac{\Omega(1+\varepsilon f)}{1+\varepsilon f}\, d\mu \geqslant 0$$

for every $\varepsilon\in[-\delta,\infty)$. As a function of ε, F_f has a minimum 0 at $\varepsilon=0$, therefore

$$\frac{d}{d\varepsilon}F_f(0)=\int \Omega f\, d\mu=0 .$$

This implies that

$$\int q(x)f(x)\mu(dx)=\iint q(x,dy)f(y)\,\mu(dx)$$

for each $f\in {}_b\mathscr{E}^+$, and hence for each $0\leqslant f\in\mathscr{E}$. This conclusion is equivalent to the claim that μ is a stationary distribution of $\{T_t\}$(see [2 ; Chapter 12]).

Conversely, we assume that μ is a stationary distribution of $\{T_t\}$. From the last part of the proof of [7 ; Lemma 2.5] it follows that

$$I_t(\mu)\equiv -\inf_{f\in {}_b\mathscr{E}^+}\int \log\frac{T_t f}{f}\, d\mu=0 .$$

On the other hand, the proof of [7 ; Lemma 3. 1] also works in our case, hence we have

$$I(\mu)=\lim_{t\to 0}\frac{1}{t}I_t(\mu).$$

Combining the above two facts, we get the assertion of the theorem.

Proof of Theorem 3.

For convenience, if a σ-finite measure α on $(E,\mathscr{E})$ (resp., $(E^0, E^0\cap\mathscr{E})$) is absolutely continuous with respect to λ, we write $\alpha(x)=\frac{d\alpha}{d\lambda}(x)$ for $x\in E$ (resp. E^0).

Now, let $\mu\ll\lambda$ and define

$$h(\mu;f;x,y)=\left(\sqrt{\mu(x)q(x,y)f(y)/f(x)}-\sqrt{\mu(y)q(y,x)f(x)/f(y)}\right)^2,$$

$$h(\mu;x,y)=h(\mu;1;x,y),\quad x,y\in E.$$

Denote by $H(\mu)$ and $H(\mu,f)$ the integrals of $h(\mu;x,y)$ and $h(\mu;f;x,y)$ with respect to $\lambda\times\lambda$ respectively.

Since $\mu\in\mathscr{P}(q)$, we have

$$\frac{1}{2}H(\mu)\leqslant\frac{1}{2}\iint[q(x,dy)\mu(dx)+q(y,dx)\mu(dy)]$$

$$=\int q(x)\mu(dx)<\infty,$$

and so by Theorem 1, we obtain

$$\begin{aligned}
I(\mu)&=-\inf_{f\in\mathscr{F}}\int\frac{\Omega f}{f}\,d\mu\\
&=\sup_{f\in\mathscr{F}}\iint\mu(x)\left[1-\frac{f(y)}{f(x)}\right]q(x,y)\lambda(dx)\lambda(dy)\\
&=\sup_{f\in\mathscr{F}}\iint\mu(y)\left[1-\frac{f(x)}{f(y)}\right]q(y,x)\lambda(dx)\lambda(dy)\\
&=\frac{1}{2}\sup_{f\in\mathscr{F}}\iint[(1-f(y)/f(x))\mu(x)q(x,y)\\
&\qquad+(1-f(x)/f(y))\mu(y)q(y,x)]\lambda(dx)\lambda(dy)\\
&=\frac{1}{2}\sup_{f\in\mathscr{F}}[H(\mu)-H(\mu,f)]\\
&=\frac{1}{2}H(\mu)-\frac{1}{2}\inf_{f\in\mathscr{F}}H(\mu;f).
\end{aligned}$$

□

To prove Theorem 4, we need some preparations.

Lemma 1. *Under* (H_1) *and* (H_2), *we have*

$$\nu(x)q(x,y)=\nu(y)q(y,x),\ \text{a.s.}\ (\lambda\times\lambda)\ \text{on}\ E^0\times E^0.$$

Proof. We first show that $\nu\ll\lambda$ on E^0 so that $\nu(x)=\frac{d\nu}{d\lambda}(x)$ has a meaning.

Let $\lambda(A)=0$. Since ν is a symmetric measure of $q(x,dy)$, it follows that

$$0=\int\nu(dx)\int_A q(x,y)\,\lambda(dy)=\int\nu(dx)q(x,A)$$
$$=\int_A\nu(dx)q(x),$$

and so $\nu\ll\lambda$ on E^0. Now, by the symmetrizability, we have

$$\int_A q(x,B)\nu(x)\lambda(dx)=\int_B q(x,A)\nu(x)\lambda(dx),\quad A,B\in E^0\bigcap\mathscr{E},$$

hence

$$\int_A\left[\int_B\nu(x)q(x,y)\lambda(dx)\right]\lambda(dy)$$
$$=\int_B\left[\int_A\nu(x)q(x,y)\lambda(dy)\right]\lambda(dx)$$
$$=\int_B\nu(x)q(x,A)\,\lambda(dx)$$
$$=\int_A\nu(x)q(x,B)\lambda(dx)$$
$$=\int_A\int_B\nu(y)q(y,x)\lambda(dx)\lambda(dy),\quad A,B\in E^0\bigcap\mathscr{E}.$$

Starting from this and using the monotone class theorem, the conclusion follows immediately.

Lemma 2. *Let λ be a σ-finite measure on $(E,\mathscr{E})$ and g be a nonnegative $\mathscr{E}$-measurable function. Then there exists a $\mu\in\mathscr{P}(E)$ such that*

(i) $\mu\ll\lambda$;

(ii) $\dfrac{d\mu}{d\lambda}>0$, a.s. (λ),

(iii) $\int g(x)\mu(dx)<\infty$.

Proof. Choose a sequence of disjoint sets $\{B_n\}_1^\infty\subset\mathscr{E}$ such that $0<\lambda(B_n)<\infty$ and $\sum_{n=1}^\infty B_n=E$. Set

$$f(x)=\sum_{n=1}^\infty(\lambda(B_n)+n^2)^{-2}I_{B_n}(x),\qquad x\in E.$$

Then $f \in {}_b\mathscr{E}^0$ and

$$\int f(x)\lambda(dx) = \sum_{n=1}^{\infty} (\lambda(B_n)+n^2)^{-2}\lambda(B_n)$$
$$< \sum_{n=1}^{\infty} (\lambda(B_n)+n^2)^{-1} < \infty .$$

Now if we take

$$\overline{\mu}(x) = (g(x)+f(x)^{-1})^{-2}, \quad x \in E,$$

then the measure μ defined by

$$\mu(A) = \int_A \overline{\mu}(x)\lambda(dx) \Big/ \int \overline{\mu}(x)\lambda(dx), \qquad A \in \mathscr{E}$$

will have the required properties. Indeed, it is clear that $\overline{\mu} \in \mathscr{E}^+$ and $0 < \int \overline{\mu}(x)\lambda(dx) < \infty$, and so $\mu \in \mathscr{P}(E)$, $\mu \ll \lambda$ and $\frac{d\mu}{d\lambda} = \overline{\mu} / \int \overline{\mu}\, d\lambda > 0$, a.s.$(\lambda)$. Finally,

$$\int g(x)\mu(dx) = \left[\int \overline{\mu}\, d\lambda\right]^{-1} \int g(x)\overline{\mu}(x)\lambda(dx)$$
$$\leqslant \left[\int \overline{\mu}\, d\lambda\right]^{-1} \int f(x)\lambda(dx) < \infty .$$

Proof of Theorem 4.

The proof is split into five steps.

Step 1. In the proof of Lemma 1, we have seen that $\nu \ll \lambda$ on E^0. Hence the symmetrizability implies that

$$0 = \int_{E\setminus E^0} \nu(dx)q(x) = \int_E \nu(dx)q(x, E\setminus E^0)$$
$$= \int_{E^0} \nu(dx)q(x, E\setminus E^0) = \int_{E^0} \nu(x)q(x, E\setminus E^0)\lambda(dx),$$

where $\nu(x) = \frac{d\nu}{d\lambda}(x)$, $x \in E^0$. By the assumption (H_2), we have

$$\nu(\cdot) > 0, \text{ a.s.}(\lambda) \text{ on } E^0$$

and so

$$q(\cdot, E\setminus E^0) = 0, \text{ a.s.}(\lambda) \text{ on } E^0 .$$

Furthermore, for each nonnegative $\mathscr{E}$-measurable function f, we have

$$\int_{E\setminus E^0} q(\cdot, y)f(y)\lambda(dy) = 0, \text{ a.s.}(\lambda) \text{ on } E^0 .$$

At last we get

$$\iint \mu(dx)\{1 - f(y)/f(x)\}q(x, dy)$$

$$= \int_{E^0} \mu(dx) \int_E \{1 - f(y)/f(x)\} q(x, y) \lambda(dy)$$

$$= \int_{E^0} \mu(dx) \int_{E^0} \{1 - f(y)/f(x)\} q(x, y) \lambda(dy) . \tag{12}$$

Step 2. Starting from (12) and using the approach used in the proof of Theorem 3, we obtain

$$I(\mu) = \frac{1}{2} \int_{E^0} \int_{E^0} h(\mu; x, y) \lambda(dx) \lambda(dy)$$
$$- \frac{1}{2} \inf_{f \in \mathscr{F}} \int_{E^0} \int_{E^0} h(\mu; f; x, y) \lambda(dx) \lambda(dy) .$$

Thus the proof of Theorem 4 is reduced to the particular case $E = E^0$. From now on, we will assume that $E = E^0$.

Step 3. Because of $0 < v < \infty$, a . s . (λ), without loss of generality we may and will assume that $v \in \mathscr{E}^0$. At the moment, we also assume that $0 < \mu(x) < \infty$ for any $x \in E$. Then $\mu / v \in \mathscr{E}^0$, and so

$$\inf_{f \in \mathscr{E}^0} H(\mu; f) \leqslant H(\mu; \sqrt{\mu/v}).$$

But by Lemma 1, we have

$$\mu(x) q(x, y) \sqrt{\mu(y) v(x) / \mu(x) v(y)}$$
$$= \sqrt{\mu(x)}\, q(y, x) v(y) \sqrt{\mu(y) / v(x) v(y)}$$
$$= \mu(y) q(y, x) \sqrt{\mu(x) v(y) / \mu(y) v(x)} , \quad \text{a. s. } (\lambda \times \lambda) .$$

This implies that $H(\mu; \sqrt{\mu/v}) = 0$, and hence

$$I(\mu) = \frac{1}{2} H(\mu)$$

which is exactly what we want.

Step 4. We now remove the extra assumption in the last step that $\mu > 0$. By Lemma 2, we can choose an $\alpha \in \mathscr{P}(E)$ such that $\alpha \ll \lambda$, $\alpha(x) > 0$ on E and $\int q(x) \alpha(dx) < \infty$. Hence

$$\inf_{f \in \mathscr{E}} H(\alpha; f) = 0$$

and

$$I(\alpha) = \frac{1}{2} H(\alpha) \leqslant \int q(x) \alpha(dx) < \infty .$$

Define

$$\mu_n = \frac{1}{n} \alpha + \frac{n-1}{n} \mu , \quad n \geqslant 1 .$$

Then $\mu_n \in \mathscr{P}(E)$, $\mu_n \ll \lambda$, $\mu_n(x) > 0$ on E and $\int q(x) \mu_n(dx) < \infty$ for each $n \geqslant 1$. Also, as we have proved in the last step

$$I(\mu_n) = \frac{1}{2} H(\mu_n), \qquad n \geqslant 1 .$$

We now show that the assertion

$$\lim_{n\to\infty} I(\mu_n)=I(\mu) \tag{13}$$

implies very quickly the conclusion of Theorem 4. Indeed, noticing that

$$h(\mu_n;x,y)\leqslant(\alpha(x)+\mu(x))q(x,y)+(\alpha(y)+\mu(y))q(y,x),$$

$$\iint(\alpha(x)+\mu(x))q(x,y)\lambda(dx)\lambda(dy)$$

$$=\int(\alpha(x)+\mu(x))q(x)\lambda(dx)$$

$$=\int q(x)\alpha(dx)+\int q(x)\mu(dx)<\infty\ , \tag{14}$$

$$\iint(\alpha(y)+\mu(y))q(y,x)\lambda(dx)\lambda(dy)<\infty \tag{15}$$

and using the dominated convergence theorem, we get

$$\lim_{n\to\infty} I(\mu_n)=\frac{1}{2}\lim_{n\to\infty} H(\mu_n)=\frac{1}{2}H(\mu)$$

which plus (13) gives us the conclusion of Theorem 4.

Step 5. Let us return to proving (13). By the convexity property of I, we have

$$I(\mu_n)\leqslant\frac{1}{n}I(\alpha)+\frac{n-1}{n}I(\mu)$$

and so

$$\overline{\lim_{n\to\infty}}\, I(\mu_n)\leqslant I(\mu).$$

On the other hand, noticing (14) and (15), it follows that

$$\underline{\lim_{n\to\infty}}\, I(\mu_n)=\underline{\lim_{n\to\infty}}\sup_{f\in\mathscr{B}^+}\left[-\int\frac{Lf}{f}\,d\mu\right]$$

$$=\underline{\lim_{n\to\infty}}\left[\int q(x)\mu_n(x)\lambda(dx)\right.$$

$$\left.-\inf_{f\in\mathscr{B}^+}\iint q(x,y)\mu_n(x)f(y)/f(x)\lambda(dx)\lambda(dy)\right]$$

$$\geqslant\underline{\lim_{n\to\infty}}\int q(x)\mu_n(x)\lambda(dx)$$

$$-\overline{\lim_{n\to\infty}}\inf_{f\in\mathscr{B}^+}\iint q(x,y)\mu_n(x)f(y)/f(x)\lambda(dx)\lambda(dy)$$

$$\geqslant\underline{\lim_{n\to\infty}}\int q(x)\mu_n(x)\lambda(dx)$$

$$-\inf_{f\in\mathscr{B}^+}\overline{\lim_{n\to\infty}}\iint q(x,y)\mu_n(x)f(y)/f(x)\lambda(dx)\lambda(dy)$$

$$= \int q(x)\mu(x)\lambda(dx)$$
$$- \inf_{f\in \mathscr{D}^+} \iint q(x,y)\,\mu(x)f(y)/f(x)\lambda(dx)\lambda(dy)$$
$$= I(\mu).$$

This finishes the proof of (13).

Proof of Theorem 5.

Put

$$\tilde{h}(\mu;x,y) = [\sqrt{\mu(x)q(x,y)m/M} - \sqrt{\mu(y)q(y,x)M/m}\,]^2$$
$$+ [\sqrt{\mu(x)q(x,y)M/m} - \sqrt{\mu(y)q(y,x)m/M}\,]^2$$

and denote by $\tilde{H}(\mu)$ the integral of $\tilde{h}(\mu;x,y)$ with respect to $\lambda\times\lambda$. Because $\mu\in\mathscr{P}'(q)$, we have

$$H(\mu) \leqslant \frac{1}{2}\tilde{H}(\mu) \leqslant \infty.$$

Hence the proof of Theorem 4 also gives us

$$I(\mu) = \frac{1}{2}\int_{E^0}\int_{E^0} h(\mu;x,y)\lambda(dx)\lambda(dy)$$
$$- \frac{1}{2}\inf_{f\in\mathscr{D}^+}\int_{E^0}\int_{E^0} h(\mu;f;x,y)\lambda(dx)\lambda(dy).$$

For simplicity, from now on we assume that $E=E^0$. Set

$$\mathscr{E}_{m,M} = \{f\in\mathscr{E}: m\leqslant f\leqslant M\}.$$

For each $f\in\mathscr{E}_{m,M}$, we can choose a sequence $\{f_n\}_1^\infty \subset \mathscr{D}^+$ such that $m\leqslant f_n\leqslant M$ and $f_n(x)\to f(x)$ for each $x\in E$ as $n\to\infty$ as we did in the proof of Theorem 1. On the other hand, for each $g\in\mathscr{E}_{m,M}$, we have

$$h(\mu;g;x,y) \leqslant \tilde{h}(\mu;x,y)/2, \qquad x,y\in E.$$

Therefore for each $\mu\in\mathscr{P}'(q)$, by the dominated convergence theorem, we get

$$\lim_{n\to\infty} H(\mu,f_n) = H(\mu,f),$$

and hence

$$\inf_{f\in\mathscr{D}^+} H(\mu,f) = \inf_{f\in\mathscr{D}^+\cup\mathscr{E}_{m,M}} H(\mu,f). \tag{16}$$

But the assumptions of Theorem 5 give us

$$\sqrt{\mu/v} \in \mathscr{E}_{m,M}, \quad \text{a.s. } (\lambda),$$

and Lemma 1 gives us

$$h(\mu;\sqrt{\mu/v};x,y) = 0, \qquad \text{a.s. } (\lambda\times\lambda);$$

thus by (16) we have

$$\inf_{f \in \mathscr{D}^+} H(\mu, f) \leqslant H(\mu, \sqrt{\mu/\nu}) = 0.$$

This completes the proof of Theorem 5.

Proof of Theorem 6.

Clearly, (ii)$\Rightarrow$(i). From Corollary, it follows that (iii)$\Rightarrow$(ii). Now we need only to show that (i)$\Rightarrow$(iii).

Without loss of generality, we may and will assume that $E=\{0, 1, 2, \cdots\}$. Denote by δ_k the pointmass at k. Then we have

$$I(\delta_k)=\left\{\sum_{i<k} + \sum_{i\geqslant k}\right\}\sum_{j<i}[\sqrt{\delta_{ki}q_{ij}} - \sqrt{\delta_{kj}q_{ji}}]^2$$

$$=\sum_{i\geqslant k}\sum_{j<i}(\sqrt{\delta_{ki}q_{ij}} - \sqrt{\delta_{kj}q_{ji}})^2$$

$$=\sum_{j<k} q_{kj} + \sum_{i>k} q_{ki} = q_k.$$

The same proof shows that $\overline{I}(\delta_k)=\overline{q}_k$. Hence $q_k=\overline{q}_k$ for every $k\in E$.

Next , assume that $n<m$ and take $\mu=\frac{1}{2}(\delta_n+\delta_m)$. Then we have

$$I(\mu)=\sum_i\sum_{j>i}(\sqrt{\mu_i q_{ij}} - \sqrt{\mu_j q_{ji}})^2 = \frac{1}{2}(q_m+q_n)-\sqrt{q_{nm}q_{mn}},$$

$$\overline{I}(\mu)=\frac{1}{2}(\overline{q}_m+\overline{q}_n)-\sqrt{\overline{q}_{nm}\overline{q}_{mn}}.$$

Combining the above facts with the condition (i), we get (iii) immediately.

Additional Note. The answer to the question mentioned in (9) has been obtained and will be presented elsewhere.

Acknowledgement.

The authors wish to express their thanks to Professor S. J. Yan for his comments.

References

[1] Chen Mufa, Reversible Markov processes in abstract state space (in Chinese), *Chin. Ann. of Math.*, **1** (1980), 437— 451.

[2] Chen Mufa, Jump Processes and Particle Systems (in Chinese), Press of Beijing Normal Univ., 1986.

[3] Chen Mufa, Coupling for jump processes, *Acta Mathematica Sinica, New Series*, **2** (1986), 123— 136.

[4] Chen Mufa and Lu Yungang , Large deviations for Markov chains, *Acta Math. Scientai*, 1 (1990).

[5] Chen Mufa and Zheng Xiaogu, Uniqueness criterion for q-processes, *Sci. Sinica*, **XXVI** (1983), 11 — 24.

[6] Donsker, M. D. and Varadhan, S. R. S., On the principal eigenvalue of second order elliptic differential operators, SDE Conference in Japan, 1976, Edited by K. Itô, 41 — 48.

[7] Donsker, M. D. and Varadhan, S. R. S., Asymptotic evaluation of certain Markov process expectation for large time (I), *Comm. Pure Appl. Math.*, **28** (1975), 1 — 47.

[8] Pinskey, R., On evaluating the Donsker-Varadhan I-function, *Ann. of Prob.*, **13** (1985), 342 — 362.

[9] Stroock, D. W., An Introduction to the Theory of Large Deviations, Springer-Verlag, 1984.

[10] Stroock, D. W., On the rate at which a homogeneous diffusion approaches a limit, an application of large deviation theory to certain stochastic integrals, **14** (1986), 840 — 859 .

[11] Wang , Z. K., Stochastic Processes (in Chinese), Sci. Press, Beijing , 1965.