Probab. Th. Rel. Fields 78, 627–649 (1988)

Probability Theory and Related Fields

The Large Deviation Principle for Hypermixing Processes

T. Chiyonobu[1] and S. Kusuoka[2]

1 Department of Mathematics, Faculty of Science, Nagoya University, Nagoya, 464, Japan
2 Research Institute for Mathematical Science, Kyoto University, Kyoto, 606, Japan

Summary. The large deviation principle of Donsker and Varadhan type is proved under certain hypotheses on the base stationary process. Some examples of processes satisfying those hypotheses are also given.

0. Introduction

In this paper we study large deviations arising from ergodic phenomena for stationary processes. In particular we introduce a new notion of stationary process which implies the large deviation principle, and present some examples including non-Markovian processes.

Let X be a Polish space and Ω be a space of X-valued trajectories $w(\cdot)$ on $(-\infty, \infty)$ with discontinuities of the first kind, normalized to be right continuous. Then Ω is a Polsih space endowed with the Skorokhod topology on bounded intervals. We denote by $\mathscr{F}(I)$ the σ-field in Ω generated by $w(s)$ for $s \in I$ for every closed set in $\mathbb{R}$. (We abbreviate $\mathscr{F}(I)$ to $\mathscr{F}_l$ when $I=[-l, l]$.)

We denote by $\mathscr{M}_s(\Omega)$ the space of stationary measure on Ω. We write $I_1 < I_2$ if $b_1 < a_2$ for any closed intervals $I_1=[a_1, b_1]$ and $I_2=[a_2, b_2]$.

For $w \in \Omega$, and each $t>0$, we define w_t by

$$w_t(s)=w(s), \quad -t \leqq s < t, \quad w_t(s+2t)=w_t(s) \quad \text{for all} \quad s \in (-\infty, \infty).$$

We denote by θ_s the translation map on Ω, i.e. $(\theta_s w)(t)=w(s+t)$. We define $R_{t,w} \in \mathscr{M}_s(\Omega)$ by

$$R_{t,w}=\frac{1}{2t} \int_{-t}^{t} \delta_{\theta_s w_t} ds. \tag{0.1}$$

For each $t>0$, $w \to R_{t,w}$ is $\mathscr{F}([-t, t])$-measurable mapping. For each $P \in \mathscr{M}_s(\Omega)$ we define the probability measure $\Gamma_t=\Gamma_t(P)$ on $\mathscr{M}_s(\Omega)$ by

$$\Gamma_t(B)=P\{w \in \Omega, R_{t,w} \in B\}. \tag{0.2}$$

If P is ergodic, then, by the ergodic theorem,

$$\Gamma_t \to \delta_P \quad \text{as} \quad t \to \infty\,, \tag{0.3}$$

where δ_P is the Dirac measure on $\mathscr{M}_s(\Omega)$ concentrated at P. We are interested in the rate of the convergence of $\Gamma_t(A)$ as $t \to \infty$.

We say that $P \in \mathscr{M}_s(\Omega)$ is hypermixing if it satisfies the following two conditions;

(H.1) There is a decreasing function $\varrho(t) > 1$ defined on (c, ∞) for some $c > 0$ such that

$$\lim_{t \to \infty} t \cdot (\varrho(t) - 1) = 0\,, \tag{0.4}$$

and

$$\|f_1 \cdots f_n\|_1 \leqq \|f_1\|_{\varrho(t)} \cdots \|f_n\|_{\varrho(t)}\,. \tag{0.5}$$

for any bounded $\mathscr{F}(I_i)$-measurable functions f_i, $i = 1, 2, , \dots, n$ where $I_1, \dots, I_n$ are any finite intervals with $I_1 < I_2 < \dots < I_n$ and $\operatorname{dist}(I_i, I_{i+1}) \geqq t$ for every $I = 1, 2, \dots, n-1$. Here, $\|\cdot\|_p$ is the L^p-norm with respect to the measure P.

(H.2) There are decreasing functions $\gamma(t)$ and $c(t)$ defined on (c, ∞) for some $c > 0$ such that

$$\lim_{t \to \infty} t \cdot (\gamma(t) - 1) = 0, \quad \lim_{t \to \infty} c(t) = 0 \tag{0.6}$$

nad

$$\|E_{I_2} E_{I_1} f\|_{\gamma(t)'} \leqq c(t) \|f\|_{\gamma(t)}\,. \tag{0.7}$$

for any two intervals I_1, I_2 with $\operatorname{dist}(I_1, I_2) \geqq t$ and any bounded measurable function f with $E^P[f] = 0$. Here, E_I is the conditional expectation with respect to P given $\mathscr{F}(I)$, and $\gamma(t)'$ is the Hölder conjugate of $\gamma(t)$, i.e. $1/\gamma(t) + 1/\gamma(t)' = 1$.

Our main theorem is the following:

Theorem 1. *Let $P \in \mathscr{M}_s(\Omega)$ and suppose that P is hypermixing. Then*

(1) *For any $Q \in \mathscr{M}_s(\Omega)$,*

$$\mathrm{H(Q)} = \lim_{t \to \infty} \frac{1}{2t} H(t, Q) \quad \textit{exists}.$$

Here $H(t, Q)$ is defined by

$$H(t, Q) = \begin{cases} E^P\left[\left.\dfrac{dQ}{dP}\right|_{F_t} \cdot \log \left.\dfrac{dQ}{dP}\right|_{F_t}\right] \\ \textit{if } Q \textit{ is absolutely continuous w.r.t. } P \textit{ on} \\ F_t \textit{ and the integrand is integrable.} \\ \infty, \textit{ otherwise.} \end{cases}$$

Here, $\left.\dfrac{dQ}{dP}\right|_{F_t}$ *stands for* $\dfrac{dQ|_{F([-t,t])}}{dP|_{F([-t,t])}}$.

(2) $H:\mathcal{M}_{\vartheta}(\Omega)\to[0,\infty]$ *is lower-semicontinuous and affine. Moreover* $\{Q; H(Q)\leqq M\}$ *is compact for every* $M\geqq 0$.

Theorem 2. *Let* $P\in\mathcal{M}_{\vartheta}(\Omega)$ *and suppose that* P *is hypermixing. Then*

$$\overline{\lim_{t\to\infty}}\,\frac{1}{2t}\log\Gamma_t(F)\leqq -\inf_{Q\in F} H(Q), \tag{0.8}$$

for all closed sets F *in* $\mathcal{M}_{\vartheta}(\Omega)$, *and*

$$\underline{\lim_{t\to\infty}}\,\frac{1}{2t}\log\Gamma_t(G)\geqq -\inf_{Q\in G} H(Q). \tag{0.9}$$

for all open sets G *in* $\mathcal{M}_{\vartheta}(\Omega)$.

We will give some examples of stationary probability measures with the hypermixing property in Sects. 4 and 5. In particular, we show that a Gaussian process with some mixing property is hypermixing.

Let us give some remarks. The large deviation principle of the type as in Theorem 2 was first showed for Markov processes with some strong mixing property by Donsker and Varadhan [3]. Since then, several authors have studied the large deviation principle of this type for non-Markov stationary processes (cf. Accardi and Olla [1]; Olla [8], Donsker and Varadhan [4]; Orey [9] and Takahashi [11]). On the other hand, Stroock [12] showed that a symmetric stationary Markov process whose associated semigroup is hypercontractive satisfies the large deviation principle (of a weaker form). In connection with Euclidean field theory, Guerra et al. [4] showed the Gibbs variational equality for Ornstein-Uhlenbeck fields. In their proof, they showed and used the hypermixing property (H.1) for Ornstein-Uhlenbeck fields. A refinement of their work [4] has been given in [6].

1. The Proof of Theorem 1

(1.1) **Lemma.** *Assume that* P *is hypermixing. If* V *is* $\mathcal{F}_l$*-measurable for some* $l>0$ *and*

$$E^P[e^{\varepsilon V}]<\infty$$

for some $\varepsilon>0$, *then for any* $T>0$ *such that* $2(T+l)\varrho(2T)\leqq\varepsilon$,

$$\frac{1}{2t}\log E^P\left[\exp\left(\int_{-t}^{t} V(\theta_s w)\,ds\right)\right] \leqq \frac{1}{2(T+l)\varrho(2T)}\log E^P[\exp(2(T+l)\varrho(2T)V)] \tag{1.2}$$

for any large t.

Proof. Let $n=\left[\frac{t}{T+l}\right]$, $t'=t-n(T+l)$ and $t_i=-t'+(2i-1)(T+l)$, $i=0,1,2,\dots,n$, $n+1$. We define a function $G(s)$, $-(T+l)\leqq s\leqq T+l$ by

$$G(s)=\begin{cases}\sum_{i=1}^{n+1} V(\theta_{t_i+s}w) & -(T+l)\leqq s< -(T+l)+t' \\ \sum_{i=1}^{n} V(\theta_{t_i+s}w) & -(T+l)+t'\leqq s<(T+l)-t' \\ \sum_{i=0}^{n} V(\theta_{t_i+s}w) & (T+l)-t'\leqq s<T+l.\end{cases}$$

Then, by (0.5), Jensen's inequality and the stationarity of P,

$$\begin{aligned}\text{LHS of (1.2)}&=\frac{1}{2t}\log E^P\left[\exp\left(\int_{-(T+l)}^{T+l} G(s)ds\right)\right]\\ &\leqq\frac{1}{2t}\cdot\frac{1}{2(T+l)}\int_{-(T+l)}^{T+l} ds\log E^P[\exp(2(T+l)G(s))]\\ &=\frac{1}{2t}\cdot\frac{1}{2(T+l)}\left(t'\cdot\log E^P\left[\exp\left(2(T+l)\sum_{i=1}^{n+1} V(\theta_{t_i}w)\right)\right]\right.\\ &\quad+2(T+l-t')\cdot\log E^P\left[\exp\left(2(T+l)\sum_{i=1}^{n} V(\theta_{t_i}w)\right)\right]\\ &\quad\left.+t'\cdot\log E^P\left[\exp\left(2(T+l)\sum_{i=0}^{n} V(\theta_{t_i}w)\right)\right]\right)\\ &\leqq\frac{1}{2t}\cdot\frac{1}{2(T+l)}\left(t'\cdot\log\prod_{i=1}^{n+1} E^P[\exp(2(T+l)\varrho(2T)V(\theta_{t_i}w)]^{1/\varrho(2T)}\right.\\ &\quad+2(T+l-t')\cdot\log\prod_{i=1}^{n} E^P[\exp(2(T+l)\varrho(2T)V(\theta_{t_i}w))]^{1/\varrho(2T)}\\ &\quad\left.+t'\cdot\log\prod_{i=0}^{n} E^P[\exp(2(T+l)\varrho(2T)V(\theta_{t_i}w))]^{1/\varrho(2T)}\right)\\ &=\frac{1}{2t}\cdot\frac{1}{2(T+l)}\left(\frac{2t'(n+1)+2(T+l-t')n}{\varrho(2T)}\right)\\ &\quad\cdot\log E^P[\exp(2(T+l)\varrho(2T)V)].\\ &=\text{RHS of (1.2)}.\end{aligned}$$

Thus (1.2) is proved. Q.E.D.

Let us define a class of the functions on Ω by

$$\Psi=\{f\in\mathbf{C}_b(\Omega)\,;f \text{ is } \mathscr{F}_l\text{-measurable for some } l<\infty\}.$$

(1.3) **Lemma.** *If P is hypermixing, then for any $V \in \Psi$,*

$$\lambda(V)=\lim_{t\to\infty}\frac{1}{2t}\log E^P\left[\exp\int_{-t}^{t} V(\theta_s w_t)ds\right] \tag{1.4}$$

exists.

Proof. Suppose that $V\in\Psi$ be $\mathscr{F}_l$-measurable and $|V|<M$. Let

$$a(t)=\log E^P\left[\exp\left(\int_{-t}^{t} V(\theta_s w)\right)\right].$$

Take an arbitrary $\varepsilon>0$ and fix it. Take $t_0>0$ so that $\varrho(2t_0)-1<\varepsilon/3M$. For this t_0, we take T so that $(t_0+l)/T<\varepsilon/3M$. Then we claim that if $t>3TM/\varepsilon$,

$$a(t)/2t \leqq a(T)/2T+\varepsilon. \tag{1.5}$$

We prove the claim. Let $t=nT+t''$ where $n\in\mathbb{Z}$, $0<t''<T$ and let $t_j=-t+t'' +(2j-1)T, j=1,2,\dots,n$. Then by (0.5) and the stationarity of P and the hypotheses on t_0, T and t,

$$\begin{aligned}
a(t)/2t &\leqq \frac{1}{2t}\log E^P\left[\prod_{j=1}^{n}\exp\left(\int_{-T}^{T} V(\theta_{t_j+s}w)ds\right)\right]+t''M/t\\
&\leqq \frac{1}{2t}\log E^P\left[\prod_{j=1}^{n}\exp\left(\int_{-T+t_0+l}^{T-t_0-l} V(\theta_s(\theta_{t_j}w))ds\right)\right]+[T+n(t_0+l)]M/t\\
&\leqq \frac{1}{2t}\log \prod_{j=1}^{n} E^P\left[\exp\left(\varrho(2t_0)\int_{-T+t_0+l}^{T-t_0-l} V(\theta_s(\theta_{t_j}w))ds\right)\right]^{1/\varrho(2t_0)}\\
&\quad +[T+n(t_0+l)]M/t\\
&\leqq \frac{1}{2t}\log \prod_{j=1}^{n} E^P\left[\exp\left(\varrho(2t_0)\int_{-T}^{T} V(\theta_s w)ds\right)\right]^{1/\varrho(2t_0)}\\
&\quad +[T+2n(t_0+l)]M/t\\
&\leqq \frac{1}{2T}\log E^P\left[\exp\left(\int_{-T}^{T} V(\theta_s w)ds\right)\right]+M(\varrho(2t_0)-1)+[T+2n(t_0+l)]\cdot M/t\\
&\leqq \frac{1}{2T}a(T)+\varepsilon.
\end{aligned}$$

This inequality suggests that for any $\varepsilon>0$,

$$\overline{\lim_{t\to\infty}}\, a(t)/2t \leqq a(T)/2T+\varepsilon$$

for any large T, and thus letting $\varepsilon\to 0$, we get

$$\overline{\lim_{t\to\infty}}\, a(t)/2t \leqq \underline{\lim}_{t\to\infty}\, a(t)/2t.$$

Since $V\in\Psi$, is is easy to show that if

$$\lim_{t\to\infty}\frac{1}{2t}\log E^P\left[\exp\left(\int_{-t}^{t} V(\theta_s w)ds\right)\right]$$

exists, then

$$\lim_{t\to\infty} \frac{1}{2t} \log E^P\left[\exp\left(\int_{-t}^{t} V(\theta_s w_t) ds\right)\right]$$

exists and the two coincides. Thus the lemma is proved. Q.E.D.

Now we define the functions H and $H(t,\cdot)$ on $\mathcal{M}_s(\Omega)$ by

$$H(Q) = \sup\{\int V dQ - \lambda(V),\ V\in\Psi),\tag{1.6}$$

$$H(t,Q) = \sup\{\int V dQ - \log E^P[\exp(V)],\ V\in\Psi,\ V \text{ is } \mathcal{F}_t\text{-measurable}\}\tag{1.7}$$

Then by the fundamental fact (cf. Donsker and Varadhan [2], Theorem 2.1)

$$H(t,Q) = \begin{cases} E^P\left[\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} \log \left.\frac{dQ}{dP}\right|_{\mathcal{F}_t}\right] \\ \text{if } Q \text{ is absolutely continuous w.r.t. } P \text{ on} \\ \mathcal{F}_t \text{ and the integrand is integrable.} \\ \infty, \text{ otherwise}. \end{cases}\tag{1.8}$$

where $\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t}$ stands for $\frac{dQ|_{\mathcal{F}([-t,t])}}{dP|_{\mathcal{F}([-t,t])}}$.

We are now ready to state the main theorem of this section.

(1.9) **Lemma.** *For every*

$$Q\in\mathcal{M}_s(\Omega),\quad \lim_{t\to\infty}\frac{1}{2t}H(t,Q)$$

exists, and

$$H(Q) = \lim_{t\to\infty}\frac{1}{2t}H(t,Q),\tag{1.10}$$

for every $Q\in\mathcal{M}_s(\Omega)$.

Proof. First, we prove $H(Q) \leq \varliminf_{t\to\infty}\frac{1}{2t}H(t,Q)$. For every $V\in\Psi$, we define W by $W(w) = \int_{-t}^{t} V(\theta_s w_t) ds$. Then $W\in\Psi$ and W is $\mathcal{F}_t$-measurable. By stationarity, $\int V dQ = \frac{1}{2t}\int W dQ$ for any $t > l$ if V is $\mathcal{F}_l$-measurable. Thus

$$\begin{aligned} &\int V dQ - \frac{1}{2t}\log E^P\left[\exp\left(\int_{-t}^{t} V(\theta_s w_t) ds\right)\right] \\ &= \frac{1}{2t}(\int W dQ - \log E^P[e^W]) \\ &\leq \frac{1}{2t} H(t,Q). \end{aligned}$$

Letting t tend to infinity, we obtain the inequality.

To prove $\varliminf_{t\to\infty} \frac{1}{2t} H(t,Q) \leqq H(Q)$, we use Lemma (1.1). By the definition of H and (1.2),

$$\frac{1}{2\varrho(2t)(t+l)} \int V dQ - H(Q) \leqq \frac{1}{2\varrho(2t)(t+l)} \log E^P[e^V]$$

which is for $V \in \Psi$, $\mathscr{F}_l$-measurable and for any large $t > c$. And so

$$H(Q) \geqq \frac{1}{\varrho(2t)(t+l)} \cdot \frac{1}{2l} \sup\{\int V dQ - \log E^P[e^V],$$

if V is $\mathscr{F}_l$-measurable$\}$.

Letting l to infinity first, and then letting t to infinity we get the left inequality. Q.E.D.

(1.11) **Lemma.** *H is lower semi-continuous and affine. Moreover, $\{Q \in \mathscr{M}_s(\Omega); H(Q) \leqq c\}$ is compact in $\mathscr{M}_s(\Omega)$ for every $c \geqq 0$.*

Proof. Since H is the supremum of continuous linear functions, H is lower semi-continous and also convex. By convexity, if Q is expressed as $Q = \int R m(dR)$ where m is a probability measure on $\mathscr{M}_s(\Omega)$, then

$$H\left(\int R m(dR)\right) \leqq \int H(R) m(dR). \tag{1.12}$$

To prove the converse inequality, note that

$$\log(\lambda a + (1-\lambda)b) \geqq \log(\lambda a) + \log((1-\lambda)b) \geqq \log a + \log b + \log \lambda(1-\lambda)$$

for $a, b > 0$ and $0 \leqq \lambda \leqq 1$. Thus if we set

$$a = \frac{dQ}{dP}\bigg|_t \text{ and } b = \frac{dR}{dP}\bigg|_t,$$

and if we let t tend to infinity, by (1.10) we get

$$H(\lambda Q + (1-\lambda)R) \geqq \lambda H(Q) + (1-\lambda) H(R),$$

and so by induction,

$$H\left(\sum_{i=1}^n a_i Q_i\right) \geqq \sum_{i=1}^n a_i H(Q_i) \tag{1.13}$$

for any $a_i \geqq 0$ with $\sum_{i=1}^n a_i = 1$ and $Q_i \in \mathscr{M}_s(\Omega)$, $i = 1, 2, \ldots, n$.

For each $n \geqq 1$, there is a compact set K_n in $\mathscr{M}_s(\Omega)$ and finite points $R_1^{(n)}, \ldots, R_{N(n)}^{(n)}$ in $\mathscr{M}_s(\Omega)$ such that

$$m(\mathscr{M}_s(\Omega)/K_n) \leqq 1/2^n$$

and

$$\bigcup_{i=1}^{N(n)} U\left(\frac{1}{n}, R_i^{(n)}\right) \supset K_n$$

where $U(r, R) = \{P \in \mathscr{M}_s(\Omega); d(P, R) < r\}$, $d(\,,)$ is a metric on $\mathscr{M}_s(\Omega)$ compatible with the weak topology in $\mathcal{M}_s(\Omega)$. Let

$$A_1^{(n)}=U\left(\frac{1}{n},\ R_1^{(n)}\right),$$

and

$$A_i^{(n)}=U\left(\frac{1}{n},\ R_i^{(n)}\right)\setminus\bigcup_{j=1}^{i-1} U\left(\frac{1}{n},\ R_j^{(n)}\right),\quad i=1,2,\dots,N_n.$$

Let $R^{(n)}$ be the transformation on $\mathcal{M}_s(\Omega)$ defined by

$$R^{(n)}(R)=\sum_{i=1}^{N(n)}\chi_{A_i^{(n)}}(R)\cdot\frac{1}{m(A_i^{(n)})}\int_{A_i^{(n)}} Rm(dR)$$
$$+\chi_{M_s(\Omega)\setminus\bigcup_{i=1}^{N(n)}A_i^{(n)}}(R)\cdot\frac{1}{m(M_s\setminus\cup A_c^{(n)})}\int_{M_s\setminus\cup A_i^{(n)}} Rm(dR),$$

where $\chi_A(R)=1$ if $R\in A$ and $=0$ if $R\notin A$, then it is easily shown that

$$\int R^{(n)}(R)m(dR)=Q$$

and

$$R^{(n)}(R)\to R,\quad\text{in weak sense,}\quad m\text{-a.e. } R.$$

Thus, by Fatou's lemma, lower semi-continuity of H and (1.13), we get

$$H(Q)\geqq\int\varliminf_{n\to\infty} H(R^{(n)}(R))m(dR)\geqq\int H(R)m(dR)$$

and affinity is proved.

Finally, by (1.6) and Lemma (1.1), for every $Q\in\mathcal{M}_s(\Omega)$,

$$H(Q)\geqq E^Q[V]-\frac{1}{2(T+l)\varrho(2T)}\log E^P[\exp(2(T+l)\varrho(2T)V)],$$

where V is any $\mathcal{F}_l$-measurable function in Ψ and T is some positive number. Therefore, if $H(Q)\leqq L$, then

$$E^Q[V]\leqq L+\frac{1}{\alpha}\log E^P[e^{\alpha V}],\quad\text{where}\quad\alpha=2(T+l)\varrho(2T).$$

Since P is Radon, we can take a compact set $K(M)$ in $\mathbf{D}_X([-l,l])$ such that $P(K(M)^c)<e^{-\alpha M}$ for each $M>0$. If V is such that $V=0$ on $K(M)$ and $\|V\|_{\mathbf{C}_b}<M$, then

$$E^Q[V]\leqq L+\frac{1}{\alpha}\log(1+e^{\alpha M}/e^{\alpha M})\leqq L+\frac{1}{\alpha}\log 2$$

and so $Q(K(M)^c)\leqq\left(L+\frac{1}{\alpha}\log 2\right)/M$. This means that $\{Q\,;H(Q)\leqq L\}$ is a tight family of measures on $\mathbf{D}_X([-l,l])$, but for stationary measures, tightness on $\Omega([-l,l])$ is equivalent to tightness on Ω, (in fact, if we let $\tilde{K}(M)=\bigcap_{i=-\infty}^{\infty}\theta_{(2i-1)l}K(M)$, then $\tilde{K}(M)$ is compact in Ω and $Q(\tilde{K}(M)^c)$

$\leq \left(L + \frac{1}{\alpha} \log 2 \right) / M. \Big)$ and so it is tight in $\mathcal{M}_s(\Omega)$. Since H is lower semicontinuous, $\{Q; H(Q) \leqq L\}$ is compact in $\mathcal{M}_s(\Omega)$. Q.E.D.

By Lemma (1.9) and Lemma (1.11), we get Theorem 1.

2. The Proof of Theorem 2: Upper Estimate

(2.1) **Lemma.** *Let F be a compact set in* $\mathcal{M}_s(\Omega)$. *Then*

$$\overline{\lim_{t \to \infty}} \frac{1}{2t} \log \Gamma_t(F) \leqq - \inf_{Q \in F} H(Q). \tag{2.2}$$

Proof. Let $l = \inf_{Q \in F} H(Q)$. Fix an arbitrary $\varepsilon > 0$. For each $Q \in F$, we can choose a $V_Q \in \Psi$ so that $\int V_Q dQ - \lambda(V_Q) \geqq l - \varepsilon$. Next, for each $Q \in F$, choose an open neighborhood B_Q of Q so that

$$\sup_{R \in B_Q} |\int V_Q dQ - \int V_Q dR| < \varepsilon.$$

Because F is compact, we can choose $Q_1, Q_2, \ldots, Q_N \in F$ so that $F \subset \bigcup_{i=1}^{N} B_{Q_i}$. Clearly,

$$\overline{\lim_{t \to \infty}} \frac{1}{2t} \log \Gamma_t(F) \leqq \overline{\lim_{t \to \infty}} \frac{1}{2t} \log \sup_i \Gamma_t(B_{Q_i}).$$

But for $Q \in F$,

$$\Gamma_t(B_Q) \leqq E^P \left[exp \left(\int_{-t}^{t} V_Q(\theta_s w_t) ds, \; R_{t,w} \in B_Q \right) \right] \cdot \sup_{R \in B_Q} \exp(-2t \int V_Q dR)$$

$$\leqq E^P \left[\exp \left(\int_{-t}^{t} V_Q(\theta_s w_t) ds \right) \right] \cdot \exp(-2t(\lambda(V_Q) + l - 2\varepsilon)),$$

and thus

$$\overline{\lim_{t \to \infty}} \frac{1}{2t} \log \Gamma_t(B_Q) \leqq l - 2\varepsilon.$$

As ε is arbitrary, we obtain (2.2). Q.E.D.

(2.3) **Lemma.** *For each* $M > 0$, *there exists a compact set* $C(M)$ *in* $\mathcal{M}_s(\Omega)$ *such that*

$$\overline{\lim_{t \to \infty}} \frac{1}{2t} \log \Gamma_t(C(M)^c) \leqq -M. \tag{2.4}$$

Proof. Since P is Radon, we can take a positive $\mathcal{F}_1$-measurable function V such that $\{V \leqq L\}$ is compact in $\mathbf{D}_x[-1, 1]$ and

$$E^P[\exp(2\varrho(2T)(T+1)V)]^{1/\{2\varrho(2T)(T+1)\}} \leqq e$$

for some T. Then by the estimate (1.2), we have

$$\frac{1}{2t}\log E^P[\exp(\int V(\theta_s w)ds)] \leqq e^{2t}$$

for large t. Given $M\in(0,\infty)$ set $K(M)=\{w\in\Omega, V(w)\leqq M^2\}$. Then $K(M)$ is compact in $\mathbf{D}_X([-1,1])$, and

$$\begin{aligned}\Gamma_t\{Q, Q(K(M)^c)\geqq M^{-1}\} &\leqq P\left\{w, \exp\left(\int_{-t}^{t} V(\theta_s w)ds\right)\geqq e^{2tM}\right\}\\ &\leqq E\left[\exp\left(\int_{-t}^{t} V(\theta_s w)ds\right)\right]/e^{2tM}\leqq e^{-2t(M-1)}.\end{aligned}$$

Hence if we let $\tilde{C}(M)=\bigcap_{l=1}^{\infty}\left\{Q, Q(K(M+l)^c)\leqq\frac{1}{M+l}\right\}$, then $\tilde{C}(M)$ is tight in $\mathscr{M}_1(\mathbf{D}[-1,1])$, and thus, by stationarity, tight in $\mathscr{M}_s(\Omega)$. Let $C(M)$ be the closure of $\tilde{C}(M)$. Then $C(M)$ is compact in $\mathscr{M}_s(\Omega)$, and

$$\begin{aligned}\Gamma_t\{Q, Q\in C(M)^c\} &\leqq \sum_{l=1}^{\infty}\Gamma_t\left\{Q, Q(K(M+l)^c)\geqq\frac{1}{M+l}\right\}\\ &\leqq \sum_{l=1}^{\infty} e^{-2t(M+l-1)}=e^{-2Mt}/(1-e^{-2Mt})\end{aligned}$$

for large t. Thus the Lemma is proved. Q.E.D.

(2.5) **Proposition.** *For any closed set F in $\mathscr{M}_s(\Omega)$,*

$$\overline{\lim_{t\to\infty}}\frac{1}{2t}\log\Gamma_t(F)\leqq-\inf_{Q\in F}H(Q). \tag{2.6}$$

Proof. Since $F\subset(F\cap C(M))\cup C(M)^c$ and $F\cap C(M)$ is compact for each M, by the preceeding two lemmas,

$$\overline{\lim_{t\to\infty}}\frac{1}{2t}\log\Gamma_t(F)\leqq\max\left\{-\inf_{Q\in F\cap C(M)}H(Q), -M\right\}.$$

Since M is arbitrary, by letting $M\to\infty$, we get (2.6). Q.E.D.

3. The Proof of Theorem 2; Lower Estimate

(3.1) **Proposition.** *Assume P is hypermixing. Let $Q\in\mathscr{M}_s(\Omega)$ be such that $H(Q)<\infty$ and N be any neighborhood of Q, then*

$$\underline{\lim_{t\to\infty}}\frac{1}{2t}\log\Gamma_t(N)\geqq-H(Q). \tag{3.2}$$

Thus if G is open in $\mathscr{M}_s(\Omega)$, we have

$$\varliminf_{t\to\infty} \frac{1}{2t} \log \Gamma_t(G) \geqq -\inf_{Q\in G} H(Q). \tag{3.3}$$

Proof.

Step 1. By virtue of the affinity of H, it is sufficient to prove (3.2) for any Q such that $Q=\sum_{i=1}^{n} a_i Q_i$, where $a_i>0$, $\sum_{i=1}^{n} a_i=1$ and Q_i, $i=1,2,\dots,n$ are ergodic stationary measures. Take a neighborhood N_i of Q_i, $i=1,2,\dots,n$, $\varepsilon>0$ and $l>0$ such that if $R_i\in N_i$ for each i and $\left\|R-\sum_{i=1}^{n} a_i R_i\right\|_l<\varepsilon$, where $\|\cdot\|_l$ is the the variation norm on $(\mathbf{D}[-l,l],\mathscr{F}_l)$, then $R\in N$.

For each $t>0$ and $\delta>0$, let $T=t+(n-1)\delta t$,

$$t_m=-T+\left((m-1)\delta+2\sum_{i=1}^{m-1} a_i+a_m\right)t, \quad \text{and} \quad I_m=[t_m-a_m t, t_m+a_m t].$$

Then, I_m's are subintervals in $[-T,T]$ such that $\operatorname{dist}(I_m,I_{m+1})=\delta t$, $m=1 2,\dots,n-1$. Then

$$\begin{aligned}
&\left\|\sum_{i=1}^{n} a_i R_{a_i t,\theta_{t_i}w}-R_{T,w}\right\|_l\\
&\quad=\left\|\frac{1}{2T}\int_{-T}^{T}\delta_{\theta_s w_T}ds-\frac{1}{2t}\sum_{i=1}^{n}\int_{-a_i t}^{a_i t}\delta_{\theta_s(\theta_{t_i}w)_{a_i t}}ds\right\|_l\\
&\quad\leqq\frac{\delta t}{2T}(n-1)+\left(\frac{1}{2T}-\frac{1}{2t}\right)\sum_{i=1}^{n}2(a_i t-l)+2nl\left(\frac{1}{2T}-\frac{1}{2t}\right)\\
&\quad\leqq 2(n-1)\delta+2nl/t.
\end{aligned}$$

Thus we can take δ so that

$$\left\|\sum_{i=1}^{n} a_i R_{a_i t,\theta_{t_i}w}-R_{T,w}\right\|_l<\varepsilon,$$

for large t. We fix the $\delta>0$ and re-define $T=T(t)=\{1+(n-1)\delta\}t$, for each t. Let $A_i=A_i(t)=\{w; R_{a_i t,\theta_{t_i}w}\in N_i\}$, then by the preceeding argument,

$$\{w; R_{T,w}\in N\}\supset\bigcap_{i=1}^{n} A_i. \tag{3.4}$$

Step 2. We take an increasing function $\varphi(t)$ on (c,∞) for some positive c satisfying

$$\lim_{t\to\infty}\varphi(t)(\varrho(t)-1)=0, \quad \lim_{t\to\infty}\varphi(t)/t=\infty, \quad \text{and} \tag{3.5}$$
$$\varphi(t)(\gamma(t)-1)<M, \quad \text{for some } M.$$

We introduce for simplicity, the notations

$$F_i(t)=\frac{dQ_i}{dP}\bigg|_{\mathscr{F}(I_i)},$$

$$G_i(t)=\min(\log F_i(t),\varphi(\delta t)),\quad i=1.2,\dots,n,$$

$$Z(t)=E^P\left[\prod_{i=1}^n \exp(G_i(t))\right],$$

and we define $Q_t\in\mathscr{M}_1(\mathbf{D}[-T,T])$ by

$$Q_t(A)=E^P\left[\frac{1}{Z(t)}\prod_{i=1}^n \exp(G_i(t)),A\right].$$

Then by (3.4) and Jensen's inequality,

$$\begin{aligned}P\{w;R_{T,w}\in N\}&\geqq P\left\{\bigcap_{i=1}^n A_i\right\}\\ &=Z(t)\cdot E^{Q_t}\left[\prod_{i=1}^n \exp(-G_i(t)),\bigcap_{i=1}^n A_i\right]\\ &=Z(t)\cdot E^{Q_t}\left[\prod_{i=1}^n \exp(-G_i(t)|\bigcap_{i=1}^n A_i\right]\cdot Q_t\left(\bigcap_{i=1}^n A_i\right)\\ &\geqq Z(t)\cdot\exp\left(E^{Q_t}\left[-\sum_{i=1}^n G_i|\bigcap_{i=1}^n A_i\right]\right)\cdot Q_t\left(\bigcap_{i=1}^n A_i\right)\\ &=Z(t)\cdot\exp\left\{E^{Q_t}\left[-\sum_{i=1}^n G_i,\bigcap_{i=1}^n A_i\right]\Big/Q_t\left(\bigcap_{i=1}^n A_i\right)\right\}\cdot Q_t\left(\bigcap_{i=1}^n A_i\right).\end{aligned}$$

In the next step, we shall show that

$$\lim_{t\to\infty} Z(t)=1\quad\text{and}\quad \lim_{t\to\infty} Q_t\left(\bigcap_{i=1}^n A_i\right)=1.$$

Thus, using these facts, we get

$$\varliminf_{t\to\infty}\frac{1}{2t}\log\Gamma_t(N)\geqq-\sum_{i=1}^n a_i\cdot\varlimsup_{t\to\infty}\frac{1}{2a_iT}E^{Q_t}\left[G_i,\bigcap_{i=1}^n A_i\right].\tag{3.6}$$

Step 3. We claim that

$$\lim_{t\to\infty} Z(t)=1,\tag{3.7}$$

and

$$\lim_{t\to\infty} Q_t\left(\bigcap_{i=1}^n A_i\right)=1.\tag{3.8}$$

Proof of the claim. First we prove (3.7). Since

$$\begin{aligned}1-Z(t)&=1-E\left[\prod_{i=1}^n e^{G_i(t)}\right]\\ &=E[(1-e^{G_1})+e^{G_1}(1-e^{G_2})+\dots+e^{G_1}e^{G_2}\dots e^{G_{n-1}}(1-e^{G_n})],\end{aligned}$$

it is sufficient to prove that

(3.9) $$\lim_{t\to\infty} E^P[\exp(G_m(t))]=1,$$

and

(3.10) $$\lim_{t\to\infty} E^P\left[\prod_{i=1}^{m-1} \exp(G_i(t))\cdot(1-\exp(G_m(t)))\right]=0.$$

By definition of F_m and G_m,

$$E[\exp(G_m(t))]\geqq E[F_m(t),\ \log F_m(t)<\varphi(\delta t)]$$
$$=Q_m(\log F_m(t)<\varphi(\delta t)),$$

and so

$$\begin{aligned}0&\leqq 1-E[\exp(G_m(t))]\\ &\leqq Q_m(\log F_m(t)\geqq\varphi(\delta t))\\ &\leqq E^{Q_m}[|\log F_m(t)|]/\varphi(\delta t)\\ &=\frac{2a_m}{\delta}\cdot\frac{\delta t}{\varphi(\delta t)}\cdot\frac{1}{2a_m t}\,E^P[F_m(t)|\log F_m(t)|].\end{aligned}$$

Since

$$\lim_{t\to\infty}\frac{1}{2a_m t}\,E^P[F_m(t)|\log F_m(t)|]=H(Q_m)<\infty,$$

and

$$\lim_{t\to\infty}\frac{\delta t}{\phi(\delta t)}=0,$$

(3.9) is proved.

Next, if we define an operator E_t by the conditional expectation given $\mathscr{F}\left(\bigcup_{i=1}^{m-1} I_i\right)$, by Hölder's inequality and (0.7),

$$\begin{aligned}&E^P\left[\prod_{i=1}^{m-1}\exp(G_i(t))(1-\exp(G_m(t)))\right]\\ &=E^P\left[\prod_{i=1}^{m-1}\exp(G_i(t))\cdot E_t(1-\exp(G_m(t)))\right]\\ &\leqq\left\|\prod_{i=1}^{m-1}\exp(G_i(t))\right\|_{\gamma(\delta t)}\cdot\|E_t(1-\exp(G_m(t)))\|_{\gamma(\delta t)'}\\ &\leqq\prod_{i=1}^{m-1}\|\exp(G_i(t))\|_{\varrho(\delta t)\gamma(\delta t)}\cdot\|E_t(1-\exp(G_m(t)))\|_{\gamma(\delta t)'}.\end{aligned}$$

Since

$$\|\exp(G_i(t))\|_{\varrho(\delta t)\gamma(\delta t)}\leqq\exp(\varphi(\delta t)(\varrho(\delta t)\gamma(\delta t)-1)$$

and

$$\varphi(\delta t)(\varrho(\delta t)\gamma(\delta t)-1)\leqq 2\varphi(\delta t)(\varrho(\delta t)-1)+\varphi(\delta t)(\gamma(\delta t)-1)$$

as t is large, $\|\exp(G_i(t))\|_{\varrho(\delta t)\gamma(\delta t)}$ is bounded above by (3.5). As for $\|E_t(1-e^{G_m})\|_{\gamma(\delta t)'}$, if we set

$$\alpha(t)=E^P[1-\exp(G_m(t))]$$

and

$$\beta(t)=1-\exp(G_m(t))-\alpha(t),$$

then $E^P[\beta(t)]=0$ and by (3.7) $\lim_{t\to\infty}\alpha(t)=0$. Since

$$\begin{aligned}\|\beta(t)\|_{\gamma(\delta t)} &\leqq |1-\alpha(t)|+\|\exp(G_m(t))\|_{\gamma(\delta t)}\\ &\leqq |1-\alpha(t)|+\exp((\gamma(\delta t)-1)\cdot\varphi(\delta t))\end{aligned}$$

and $(\gamma(\delta t)-1\varphi(\delta t)$ is bounded in t, $\|\beta(t)\|_{\gamma(\delta t)}$ is bounded in t. By (H.2) of the hypermixing property,

$$\begin{aligned}\|E_t(1-\exp(G_m(t))\|_{\gamma(\delta t)'} &\leqq \alpha(t)+\|E_t\beta(t)\|_{\gamma(\delta t)'}\\ &\leqq \alpha(t)+c(\delta t)\|\beta(t)\|_{\gamma(\delta t)},\end{aligned}$$

and so

$$\lim_{t\to\infty}\|E_t(1-G_m(t))\|_{\gamma(\delta t)'}=0.$$

Therefore with the preceeding consideration, we get (3.10).

Finally, we show (3.8). Since $1-Q_t(\cap A_i)=Q_t(\cup A_i^c)\leqq\Sigma Q_t(A_i^c)$, it suffices to show that $\lim_{t\to\infty}Q_t(A_i^c)=0$. Following the similar argument as in Step 2, we see that

$$\begin{aligned}Q_t(A_i^c) &= E^P\left[\frac{dQ_t}{dP}, A_i^c\right]\\ &\leqq Z_t^{-1}E[\exp(\varrho(\delta t)G_i(t)), A_i^c]^{1/\varrho(\delta t)}\cdot\prod_{\substack{j=1\\ j\neq i}}^n E^P[\exp(\varrho(\delta t)G_j(t))]\\ &\leqq Z_t^{-1}Q_i(A_i^c)\cdot e^{n\varphi(\delta t)(\varrho(\delta t)-1)}.\end{aligned}$$

By the ergodic theorem, $\lim_{t\to\infty}Q_i(A_i^c)=0$. Thus with (3.5) and (3.7), we get (3.8).

Step 4. By (0.5),

$$\begin{aligned}E^{Q_t}\left[G_i, \bigcap_{j=1}^n A_j\right] &= Z^{-1}\cdot E^P\left[G_i\prod_{i=1}^n\exp(G_j), \bigcap_{i=1}^n A_j\right]\\ &\leqq Z^{-1}E^P[|G_ie^{G_i}\chi_{A_i}|^{\varrho(\delta t)}]^{\varrho(\delta t)-1}\cdot\\ &\quad\cdot\prod_{j\neq i}E^P[|e^{G_j}\cdot\chi_{A_j}|^{\varrho(\delta t)}]^{\varrho(\delta t)^{-1}}\\ &\leqq Z^{-1}\cdot\{E^P[|G_ie^{G_i}|, A_i]\cdot\prod_{j\neq i}E^P[e^{G_j}, A_j]\cdot\\ &\quad\cdot\varphi(\delta t)^{(\varrho(\delta t)-1)}\cdot e^{\varphi(\delta t)(\varrho(\delta t)-1)n}\}^{\varrho(\delta t)^{-1}}\\ &\leqq Z^{-1}\cdot E^P[|F_i(t)\log F_i(t)|]\cdot\prod_{j\neq i}Q_j(A_j)\\ &\quad\cdot\varphi(\delta t)^{(\varrho(\delta t)-1)}\cdot e^{\varphi(\delta t)(\varrho(\delta t)-1)n}.\end{aligned}$$

By the ergodic theorem, $\lim_{t\to\infty} Q_j(A_j)=1$, and by (3.5),

$$\lim_{t\to\infty} \varphi(\delta t)^{(\varrho(\delta t)-1)} \cdot e^{\varphi(\delta t)(\varrho(\delta t)-1)n}=1,$$

and se we get

$$\begin{aligned}&\varliminf_{t\to\infty} \frac{1}{2a_i T} E^{Q_t}\left[G_i(t), \bigcap_{i=1}^{n} A_j\right]\\ &\leqq \varliminf_{t\to\infty} \frac{1}{2a_i T} E^P[|F_i(t)\cdot\log F_i(t)|]\\ &\leqq \varliminf_{t\to\infty} \frac{1}{2a_i T} E^P[F_i(t)\cdot\log F_i(t)+e^{-1}]=H(Q_i).\end{aligned}$$

Thus, with (3.6) and the affinity of H, we get

$$\varliminf_{t\to\infty} \frac{1}{2t} \log\Gamma_t(N) \geqq -\sum_{i=1}^{n} a_i H(Q_i) = -H(Q),$$

and the theorem is proved. Q.E.D.

By Propositions (2.5) and (3.1), we get Theorem 2 stated in the Introduction.

4. ε-Markov Case

(4.1) **Definition.** Let ε be a nonnegative number. We say that $P\in\mathcal{M}_s(\Omega)$ is *ε-Markov* if $\mathcal{F}(I_1)$ and $\mathcal{F}(I_2)$ are independent under conditional probability of P with respect to $\mathcal{F}(I_1\cap I_2)$ for any pair of intervals I_1, I_2 in $\mathbb{R}$ such that $I_1\cup I_2=\mathbb{R}$ and $\operatorname{dis}(I_1^c, I_2^c)>\varepsilon$.

The following lemma is clear from the above definition.

(4.2) **Lemma.** *Assume P is ε-Markov. Suppose that I_1, $I_2, \ldots, I_n$ are the closed intervals of length longer that ε and $I_1<I_2<\ldots<I_n$, then,*

$$E_{I_1}E_{I_2}\ldots E_{I_n}=E_{I_1}E_{I_n}. \tag{4.3}$$

In this section, we consider an ε-Markov process with the following property;

($\mathcal{K}$); There is a $T>0$ such that

$$\|E_{I_2}E_{I_1}\|_{2,4}\leqq 1 \tag{4.4}$$

for any two intervals I_1 and I_2 with $\operatorname{dist}(I_1, I_2)\geqq T$, where $\|\cdot\|_{p,q}$ is the operator norm of the operator from $L^p(\Omega, F, P)$ to $L^q(\Omega, F, P)$.

Then, by Lemma (4.2) and interpolation theory, it is easy to prove the following

(4.5) **Lemma.** *Assume an ε-Markov process satisfies ($\mathcal{K}$). Then there is an $\alpha>0$ such that*

$$\|E_{I_2}E_{I_1}\|_{p,q}\leqq 1 \tag{4.6}$$

whenever $(p-1)/(q-1)\leqq\exp(-\alpha\cdot\operatorname{dist}(I_1, I_2))$.

The next lemma is the direct modification to the ε-Markov case of Theorem (III-7), Guerra et al. [4].

(4.7) **Lemma** (Sandwich estimate). *Set* $\beta(t)=1+2\cdot(e^{\alpha(t-\varepsilon)}-1)^{-1}$, $t>\varepsilon$, *where* α *is the same as in Lemma (4.5). Then,*

$$\|E_{I_1}fE_{I_2}\|_{2,2}\leqq\|f\|_{\beta(t)} \tag{4.8}$$

for any bounded $\mathscr{F}(I)$*-measurable function* f, *where* I_1, I_2, I *are closed intervals such that* $I_1<I<I_2$, *the length of* I_1 *and* I_2 *is longer than* ε *and* $\operatorname{dist}(I_1,I)\geqq t$ *and* $\operatorname{dist}(I_2,I)\geqq t$.

Proof. Let J_1, J_2 be closed intervals of length ε such that $I_1<J_1\leqq I\leqq J_2<I_2$, $\operatorname{dist}(J_i,I)=0$ and $\operatorname{dist}(I_i,J_i)=t-\varepsilon$, $i=1,2$. Then, by (4.3) and (4.5),

$$\begin{aligned}\|E_{I_1}uE_{I_2}\|_{2,2}&=\|E_{I_1}E_{J_1}^2uE_{J_2}^2E_{I_2}\|_{2,2}\\&\leqq\|E_{I_1}E_{J_1}\|_{2,p}\|E_{J_1}uE_{J_2}\|_{p,q}\|E_{J_2}E_{I_2}\|_{q,2}\\&\leqq\|E_{J_1}uE_{J_2}\|_{p,q}\end{aligned}$$

if $p-1=e^{\alpha(t-\varepsilon)}$ and $q-1=e^{-\alpha(t-\varepsilon)}$. Thus by Hölder's ineqaulity,

$$\begin{aligned}\|E_{J_1}uE_{J_2}\|_{p,q}&\leqq\sup_{v_2}\{\|E_{J_1}(uv_2)\|_q/\|v_2\|_p,\ v_2\text{ is }\mathscr{F}(J_2)\text{-measurable}\}\\&\leqq\sup\{\|v_1uv_2\|_1/\|v_1\|_{q'}\|v_2\|_p,\ v_i\text{ is }\mathscr{F}(J_i)\text{-measurable},\ i=1,2\}.\\&\leqq\|u\|_{\beta(t)}\sup\{\|v_1v_2\|_{\beta(t)'}/\|v_1\|_{q'}\|v_2\|_p,\ v_i\text{ is }\mathscr{F}(J_i)\text{-measurable},\ i=1,2\}\\&\leqq\|u\|_{\beta(t)}.\end{aligned}$$

Therefore, the lemma is proved. Q.E.D.

(4.9) **Lemma.** *Assume that* P *is* ε*-Markov. Then* P *satisfies* (H.1) *of the hypermixing property if* P *satisfies* $(\mathscr{K})$.

Proof. For any closed intervals $I_1,I_2,\dots,I_n$ such that $I_1<I_2<\dots<I_n$ and $\operatorname{dist}(I_l,I_{l+1})\geqq t$, we take intervals $J_1,\dots,J_{n+1}$ of length ε satisfying that $J_1<I_1<J_2<I_2<J_3<\dots<I_n<J_{n+1}$, $\operatorname{dist}(J_1,I_1)=\operatorname{dist}(I_1,J_2)$, $\operatorname{dist}(I_l,J_{l+1})=\operatorname{dist}(J_{l+1},I_{l+1})$, $l=2,\dots,n$, and $\operatorname{dist}(J_n,I_n)=\operatorname{dist}(I_n,J_{n+1})$. Then, by (4.3) and (4.8), we have

$$\begin{aligned}&\|f_1\cdot f_2\cdot\ \ f_n\|_1\\&=(1,E_{J_1}f_1E_{J_2}\cdot E_{J_2}f_2E_{J_3}\cdot\dots E_{J_n}f_nE_{J_{n+1}}1)_{L^2}\\&\leqq\|E_{J_1}f_1E_{J_2}\|_{2,2}\cdot\dots\|E_{J_n}f_nE_{J_{n+1}}\|_{2,2}\\&\leqq|f_1\|_{\beta(s)}\cdot\dots\|f_2\|_{\beta(s)},\end{aligned}$$

where $s=(t-\varepsilon)/2$, and so P satisfies (H.1) of the hypermixing property. Q.E.D.

Next, we show that $(\mathscr{K})$ derives (H.2) of the hypermixing property in the ε-Markov case. The idea of the proof of the following lemma is due to B. Simon.

(4.10) **Lemma.** *Let $p>3$. If $T; L^2(m)\to L^p(m)$ is a symmetric contraction where m is a probability measure on a measurable space such that $T1=1$. Then*

(4.11) $$\|Tf-(1,f)\|_2 \leqq (1/p-3)^{1/2}\|f\|_2$$

for any $f\in L^2(m)$.

Proof. For any $f\in L^2(m)$, we write $f=a1+g$, where $(1,g)_2=0$. Then, $\|f\|_2^2=a^2+\|g\|_2^2$ and so $\|f\|_2^{2l}=(a^2+\|g\|_2^2)^l=a^{2l}+la^{2(l-1)}\|g\|_2^2+0(a^{2(l-2)})$. On the other hand, since $(1,Tg)=(T1,g)=(1,g)=0$,

$$\|Tf\|_{2l}^{2l}=a^{2l}+l(2l-1)a^{2(l-1)}\|Tg\|_2^2+0(a^{2l-3}).$$

So if

$$\|Tf\|_{2l}\leqq\|f\|_2, \text{ for every } f\in L^2(m),$$

then

$$l(2l-1)a^{2(l-1)}\|Tg\|_2^2\leqq la^{2(l-1)}\|g\|_2^2+0(a^{2l-3}) \quad \text{for every} \quad a\in R.$$

Letting $a\to\infty$, we get

$\|Tg\|_2\leqq(1/2l-1)^{1/2}\|g\|_2$ for every positive integer l. Thus we obtain the desired estimate. Q.E.D.

(4.12) **Lemma.** *Assume $P\in M_s(\Omega)$ is* ε-Markov. Then P satisfies (H.2) of the hypermixing property if P satisfies $(\mathscr{K})$.

Proof. If K_1 and K_2 are the domains such that $\operatorname{dist}(K_1,K_2)=t$, by (4,5), $\|E_{K_1}E_{K_2}\|_{p,p'}\leqq 1$ if $p\geqq e^{-\alpha t/2}+1$, where p' is a conjugate of p. Especially, if u_2 is $\mathscr{F}(K_2)$-measurable, then

(4.13) $$\|E_{K_1}u_2\|_{p(t)'}\leqq\|u_2\|_{p(t)}, \quad p(t)=e^{-\alpha t/2}+1.$$

By Lemma (4.10), if $E[u_2]=0$, then

(4.14) $$\|E_{K_1}u_2\|_2\leqq e^{-\alpha t/5}\|u_2\|_2$$

for any large t. Now let I_1, I_2 be the domains with $\operatorname{dist}(I_1,I_2)=t$ and let J_1, J_2 be the intervals of length ε such that $I_1<J_1<J_2<I_2$ and $\operatorname{dist}(I_1,J_1)=\operatorname{dist}(J_1,J_2)=\operatorname{dist}(J_2,I_2)$. We set $q(t)=e^{-\alpha t/4}$, then by Lemma (4.2), Lemma (4.5) and (4.14), for any $\mathscr{F}(I_2)$-measurable function f_2 with $E[f_2]=0$,

$$\begin{aligned}\|E_{I_1}f_2\|_{q(t)'}&=\|E_{I_1}E_{J_1}E_{J_2}f_2\|_{q(t)'}\\&\leqq\|E_{J_1}E_{J_2}f_2\|_2\leqq e^{-\alpha t/20}\|E_{J_2}f_2\|_2\\&\leqq e^{-\alpha t/20}\|f_2\|_{q(t)}.\end{aligned}$$

Therefore P satisfies (H.2). Q.E.D.

By Lemma (4.9) and (4.12), we get the main theorem of this section.

(4.15) **Theorem.** *Assume P is ε-Markov. Then P is hypermixing if P satisfies the condition $(\mathscr{K})$.*

(4.16) **Collorary.** *Assume P is a stationary Markov provess whose associated semigroup is hypercontractive. Then P is hypermixing.*

Remark. By Theorem 1 and Corollary (4.16), the stationary Markov process whose associated semigroup is hypercontractive satisfies the large deviation principle. This result has been shown, though only on the empirical measure level, by Stroock [12].

5. Gaussian Case

Let $\mathbf{X}=(X(t), t\in\mathbb{R})$, $X(t)=(X_k(t))_{k=1,\dots d}$ be the $\mathbb{R}^d$-valued Gaussian stationary process on (Ω, F, P) with mean zero and covariance matrix

(5.1) $$R_{k,l}(t)=E^P(X_k(t+s)X_l(s)),\quad t,s\in\mathbb{R},\quad 1\leqq k,l\leqq d.$$

We assume that the spectral measure of the process has a density, i.e.

(5.2) $$R_{k,l}(t)=\int e^{i\xi t} f_{k,l}(\xi)d\xi.$$

We may assume that $(f_{k,l}(\xi))$ is hermitian and non-negative definite. In this section, we give certain criteria on $(f_{k,l}(\xi))$ for the associated Gaussian process to be hypermixing. For this purpose, we take $\mathbf{X}$ as a Gaussian random field indexed by a Hilbert space. Throughout the section, we follow the notation of Simon [11].

It is well-known (cf. Rozanov [10]) that there exist random measures $\{Z_k(d\xi), k=1,\dots,d\}$ on $\mathbb{R}$ such that

(5.3) $$Z_k(A)\perp Z_l(B)\quad\text{if}\quad A\cap B=\phi,\quad k,l=1,2,\dots,d,$$

(5.4) $$X_k(t)=\int e^{i\xi t} Z_k(d\xi),\quad t\in\mathbb{R},$$

(5.5) $$E^P(Z_k(d\xi)Z_l(d\xi))=f_{k,l}(\xi)d\xi.$$

We define separable Hilbert spaces by

(5.6) $$\mathbf{H}=\{\varphi=(\varphi_k)_{k=1,\dots d};\ \|\varphi\|_{\mathbf{H}}^2=\sum_{k,l=1}^{d}\int \varphi_k(\xi)\overline{\varphi_l(\xi)}f_{k,l}(\xi)d\xi<\infty\}$$

with an inner product

(5.7) $$(\varphi,\psi)=\sum_{k,l=1}^{d}\int \varphi_k(\xi)\psi_l(\xi)f_{k,l}(\xi)d\xi,\quad \varphi,\psi\in\mathbf{H}$$

and

(5.8) $$\mathbf{H}_I=\text{the closed linear hull of } \psi_k^t(\xi),\ k=1,\dots,d,\ t\in I \text{ in } \mathbf{H},$$

where $\psi_k^t(\xi)=(0,\dots,0,e^{it\xi},0,\dots,0)$, $e^{it\xi}$ is the k-th cordinate, for every closed set I in $\mathbb{R}$ with the same inner product as (5.7). We denote by e_I the orthogonal projection of $\mathbf{H}$ onto $\mathbf{H}_I$. We define a linear map $\Phi;\mathbf{H}\to L^2(\Omega,F,P)$ by

(5.9) $$\Phi(\varphi)=\sum_{k=1}^{d}\int \varphi_k(\xi)Z_k(d\xi).$$

(5.10) **Lemma.** $\mathbf{X}$ *is the Gaussian random process indexed by* $\mathbf{H}$ (cf. Simon [11], p. 15) *under the map* Φ.

Proof. We only have to show that $E^P(\Phi(\varphi)\Phi(\psi)) = (\varphi, \psi)_{\mathbf{H}}$ for $\varphi, \psi \in \mathbf{H}$. But by (5.3) and (5.5),

$$LHS = E^P\left(\sum_{k,l=1}^{d} \int \varphi_k(\xi)\psi_l(\xi) Z_k(d\xi)\overline{Z_l(d\xi)}\right)$$

$$= \sum_{k,l=1}^{d} \int \varphi_k(\xi)\psi_l(\xi) f_{k,l}(\xi)d\xi = \text{RHS}. \quad \text{Q.E.D.}$$

Let Γ be the "second quantized operator" (Simon [11], p. 25) on the space of contraction operators on $\mathbf{H}$. Let e_I be the orthogonal projection operator of $\mathbf{H}$ onto the subspace $\mathbf{H}_I$. Then

$$\Gamma(e_I) = E_I (= E[\cdot \mid \mathscr{F}(I)]),$$

since

$$\mathscr{F}(I) = \sigma(X_k(t), t \in I, k = 1, 2, \ldots, d)$$
$$= \sigma(\Phi(\psi_k(t)), t \in I, k = 1, 2, \ldots, d).$$

The following is essential in the proof of the main theorem of this section.

(5.11) **Theorem** (Nelson [7]). *Let A be a contraction from* $\mathbf{H}$ *to* $\mathbf{H}$. *Let* $1 \leqq p \leqq q \leqq \infty$. *Then a necessary and sufficient condition for $\Gamma(A)$ to be a contraction from $L^p(\Omega, F, P)$ to $L^q(\Omega, F, P)$ is that* $\|A\|^2 \leqq (p-1)/(q-1)$.

(5.12) **Lemma.** *Let I_1 and I_2 be disjoint open sets in* $\mathbb{R}$. *Then*

(5.13) $$\|E_{I_1}E_{I_2}\|_{p,q} \leqq 1 \quad \textit{if and only if} \quad (p-1)/(q-1) \geqq \|e_{I_1}e_{I_2}\|^2$$

where e_{I_i} is the projection onto $\mathbf{H}_{I_i}$, $i = 1, 2$. Especially, if $p \geqq \|e_{I_1}e_{I_2}\| + 1$, *then*

(5.14) $$\|f_1 \cdot f_2\|_1 \leqq \|f_1\|_p \cdot \|f_2\|_p$$

for any two bounded $\mathscr{F}(I_i)$-measurable functions f_i, $i = 1, 2$.

Proof. Since $E_{I_1}E_{I_2} = \Gamma(e_{I_1})\Gamma(e_{I_2}) = \Gamma(e_{I_1}e_{I_2})$, (5.13) is a direct consequence of Lemma (5.11). Now let q be the conjugate of p in (5.13), then if f_2 is $\mathscr{F}(I_2)$-measurable,

$$\|E_{I_1}E_{I_2}f_2\|_q = \|E_{I_1}f_2\|_q$$
$$= \sup\{\|f_1 \cdot f_2\|_1 / \|f_1\|_p, f_1 \text{ is } \mathscr{F}(I_1)\text{-measurable.}\}$$
$$\leqq \|f_2\|_p,$$

if $\|e_{I_1}e_{I_2}\|_{\text{op}}^2 \leqq (p-1)^2$. Thus,

$$\|f_1 \cdot f_2\|_1 \leqq \|f_1\|_p \cdot \|f_2\|_p \quad \text{if} \quad p \geqq \|e_{I_1}e_{I_2}\| + 1. \quad \text{Q.E.D.}$$

We define a decreasing function $\tau(t)$ on $(0, \infty)$ by

$$\tau(t) = \sup\{\|e_{I_1}e_{I_2}\|, I_1 \text{ and } I_2 \text{ are closed sets in } \mathbb{R} \text{ such that } \operatorname{dist}(I_1, I_2) \geqq t\}.$$

(5.15) **Lemma.** *The Gaussian stationary process P is hypermixing if* $\lim_{t\to\infty} t \cdot \tau(t) = 0$.

Proof.

Step 1. We first show (H.1) of the hypermixing property. Let I_i, $i=1,2,\dots,n$ be intervals such that $\operatorname{dist}(I_i, I_{i+1}) \geqq t$, $i=1,2,\dots,n-1$. We divide the intervals into two groups; $\{I_1, I_3, \dots)$ and $\{I_2, I_4, \dots\}$. Then since $\operatorname{dist}(I_1 \cup I_3 \cup \dots, I_2 \cup I_4 \cup \dots) \geqq t$, by (5.15) we get

$$\|f_1 \cdot f_2 \cdots f_n\|_1 \leqq \|f_1 \cdot f_3 \cdots\|_{1+\tau(t)} \cdot \|f_2 \cdot f_4 \cdots\|_{1+\tau(t)}.$$

Next, we split each group of intervals into two by the same way. Then, since $\operatorname{dist}(I_1 \cup I_5 \cup I_9 \cup \dots, I_3 \cup I_7 \cup I_{11} \cup \dots) \geqq 2t$,

$$\begin{aligned} &\|f_1 \cdot f_3 \cdots\|_{1+\tau(t)} \\ &\quad \leqq \|f_1 \cdot f_5 \cdots\|_{(1+\tau(t))(1+\tau(2t))} \cdot \|f_3 \cdot f_7 \cdots\|_{(1+\tau(t))(1+\tau(2t))}. \end{aligned}$$

Successive use of the hypermixing property as in the above way leads to

$$\tag{5.16} \begin{aligned} &\|f_1 \cdot f_2 \cdots f_n\|_1 \\ &\leqq \prod_{i=1}^{n} \|f_i\|_{(1+\tau(t))(1+\tau(2t))\dots(1+\tau(2^n t))} \end{aligned}$$

Let $\varrho(t) = \prod_{k=1}^{\infty} (\tau(kt)+1)$. The RHS converges for each t since

$$\text{RHS} \leqq \exp\left(\sum_{k=1}^{\infty} \tau(2^k t)\right).$$

Moreover, since

$$\lim_{t\to\infty} t\left(\sum_{k=1}^{\infty} \tau(2^k t)\right) = 0,$$

$$\lim_{t\to\infty} t \cdot (\varrho(t)-1) \leqq \lim_{t\to\infty} t\left(\exp\left(\sum_{k=1}^{\infty} \tau(2^k t)\right) - 1\right) = 0.$$

By (5.16), we get

$$\tag{5.17} \|f_1 \cdot f_n\|_1 \leqq \|f_1\|_{\varrho(t)} \cdots \|f_n\|_{\varrho(t)}.$$

Thus P satisfies (H.1) of hypermixing property.

Step 2. Secondly, we show (H.2). Since

$$E_{I_1} E_{I_2} = \Gamma(A)\Gamma(\|e_{I_1} e_{I_2}\|), \quad A = e_{I_1} e_{I_2} / \|e_{I_1} e_{I_2}\|$$

for closed intervals I_1 and I_2, and since $\Gamma(A)$ is a contraction on any $L^p(\Omega, F, P)$ (cf. Simon [11]), it suffices to show that

$$\|\Gamma(\tau(t))f\|_{q(t)'} \leqq c(t) \|f\|_{q(t)}$$

for every bounded measurable function f with $E[f]=0$, where $c(t)$ and $q(t)$ are decreasing functions such that

$$\lim_{t\to\infty} t(q(t)-1) = 0 \quad \text{and} \quad \lim_{t\to\infty} c(t) = 0.$$

Since $\lim_{t\to\infty} t\cdot\tau(t)=0$, it is possible to take functions $\tau'(t)$ and $\tau''(t)$ on $(0,\infty)$ such that $\tau(t)=\tau'(t)^2\tau''(t)$ and

$$\lim_{t\to\infty}\tau'(t)=0,\quad \lim_{t\to\infty}\tau''(t)=0 \quad\text{and}\quad \lim_{t\to\infty} t^{1/2}\cdot\tau'(t)=0.$$

Thus taking $q(t)=\tau'(t)^{1/2}+1$ and $c(t)=\tau''(t)^{1/2}$, we have, by Theorem (5.11) and Lemma (4.10),

$$\begin{aligned}\|\Gamma(\tau(t))f\|_{q(t)'}&=\|\Gamma(\tau'(t))\Gamma(\tau''(t))\Gamma(\tau'(t))f\|_{q(t)'}\\ &\leqq\|\Gamma(\tau''(t))\Gamma(\tau'(t))f\|_2\\ &\leqq c(t)\|\Gamma(\tau'(t))f\|_2\leqq c(t)\|f\|_{q(t)}\end{aligned}$$

for all f with $E[f]=0$. Thus the lemma is proved. Q.E.D.

(5.18) **Theorem.** *Assume the spectral density matrix $f(\xi)=(f_{k,l}(\xi))$ satisfies for some $c>0$, $\sigma>0$, $\gamma\geqq 0$ and $\zeta\geqq 0$,*

$$\left|\frac{d}{d\xi}f_{k,l}(\xi+x)-\frac{d}{d\xi}f_{k,l}(\xi+x+h)\right|\leqq c|h|^\sigma(1+|h|)^\gamma(1+|x|)^\zeta\varrho(\xi) \tag{5.19}$$

for all $\xi\in\mathbb{R}$, $x\in\mathbb{R}$ and $h\in\mathbb{R}$, where $\varrho(\xi)$ is the smallest eigenvalue of $f(\xi)$. Then P is hypermixing.

Remark. The idea of the following proof is essencialy due to Kesten and Papanicolau [5], Theorem 4, and Kolmogorov and Rozanov (cf. Rozanov [10], Lemma 10.6, p. 189), and so we omit the detailed calculation in the proof.

Proof. From Lemma (5.15), it is sufficient to prove that $\lim_{t\to\infty} t\cdot\tau(t)=0$ under the above condition.

Now we take for any $T>0$ and $m\in\mathbb{N}$

$$g_m^T(x)=T\cdot c_m^{-1}\left(\frac{\sin Tx}{Tx}\right)^{2m} \tag{5.20}$$

where

$$c_m=\int_{\mathbb{R}}\left(\frac{\sin x}{x}\right)^{2m}dx$$

and

$$f_{k,l,m}^T(\xi)=\int g_m^T(x)\{2f_{k,l}(\xi+x)-f_{k,l}(\xi+2x)\}dx \tag{5.21}$$

Then since $\tilde g_m^T(t)=0$ if $|t|>T$,

$$\tilde f_{m,k,l}^T(t)=\tilde g_m^T(t)\tilde f_{\kappa,l}(t)=0 \quad\text{if}\quad |t|>T. \tag{5.22}$$

And also, since $\int g_m^T(\xi)d\xi=1$,

$$f_{k,l}(\xi)-f_{k,l,m}^T(\xi)=\int g_m^T(x)(f_{k,l}(\xi)-2f_{k,l}(\xi+x)+f_{k,l}(\xi+2x))dx.$$

Let $F_\xi(x)=f_{k,l}(\xi)-2f_{k,l}(\xi+x)+f_{k,l}(\xi+2x)$, then $F_\xi(0)=0$ and

$$\left|\frac{d}{dx}F_\xi(x)\right|=2\left|\frac{d}{dx}f_{k,l}(\xi+x)-\frac{d}{dx}f_{k,l}(\xi+2x)\right|$$
$$\leqq 2c|x|^\sigma(1+|x|)^{\gamma+\zeta}\varrho(\xi),$$

and so

$$|F_\xi(x)|\leqq 2c^{\gamma+\zeta+1}|x|^{1+\sigma}(1+|x|)^{\gamma+\zeta}\varrho(\xi).$$

Thus if we define $f_{k,l}^T(\xi)$ by $f_{k,l,m}^T(\xi)$ for sufficiently large m, we can show that

$$|f_{k,l}(\xi)-f_{k,l}^T(\xi)|\leqq A\cdot T^{-(\sigma+1)}\varrho(\xi). \tag{5.23}$$

Let I_1 and I_2 be intervals such that $\operatorname{dist}(I_1,I_2)\geqq T$, and let $f=(f_k)\in\mathbf{H}_{I_1}$ and $g=(g_1)\in\mathbf{H}_{I_2}$, $(k,l=1,2,\dots,N)$ be

$$f_k(\xi)=\sum_n a_{k,n}e^{i\xi t_n}\quad\text{(finite sum)},\quad t_n\in I_1$$

$$g_l(\xi)=\sum_m b_{l,m}e^{i\xi t_m}\quad\text{(finite sum)},\quad t_m\in I_2.$$

Then since $|t_n-t_m|>T$ for every $t_n\in I_1$ and $t_m\in I_2$

$$(f,g)_H=\int_{\mathbb{R}}\sum_{k,l}f_k(\xi)g_l(\xi)f_{k,l}(\xi)d\xi$$
$$=\int\sum_{k,l}f_k(\xi)g_l(\xi)(f_{k,l}(\xi)-f_{k,l}^T(\xi))d\xi.$$

Thus by (5.23) and the Schwarz inequality

$$|(f,g)_H|\leqq AT^{-(\sigma+1)}\sum_{k,l}\int|f_k(\xi)|g_l(\xi)|\varrho(\xi)d\xi$$
$$\leqq NAT^{-(\sigma+1)}\left(\int\sum_{k,l}f_k(\xi)f_l(\xi)f_{k,l}(\xi)d\xi\right)^{1/2} \tag{5.24}$$
$$\cdot\left(\int\sum_{k,l}g_k(\xi)g_l(\xi)f_{k,l}(\xi)d\xi\right)^{1/2}$$
$$\leqq NAT^{-(\sigma+1)}\|f\|_H\cdot\|g\|_H.$$

Now we recall that

$$\tau(t)=\sup\{\|e_{I_1}e_{I_2}\|_{\mathbf{H}},\operatorname{dist}(I_1,I_2)\geqq t\}$$
$$=\sup\{(f,g)_{\mathbf{H}},F\in\mathbf{H}_1,g\in\mathbf{H}_2,\|f\|_{\mathbf{H}}=\|h\|_{\mathbf{H}}=1,\operatorname{dist}(I_1,I_2)\geqq t\}.$$

Thus (5.24) shows that $\lim_{t\to\infty}t\cdot\tau(t)=0$, and the theorem is proved. Q.E.D.

References

1. Accardi, L., Olla, S.: Donsker and Varadhan's theory for stationary processes. (Preprint)
2. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, IV. Commun. Pure Appl. Math. **36**, 183–212 (1983)

3. Donsker, M.D., Varadhan, S.R.S.: Large deviations for stationary Gaussian processes. Commun. Math. Phys. **97**, 187–210 (1985)
4. Guerra, F., Rosen, L., Simon, B.: The $P(\phi)_2$ Euclidean quantum field theory as classical statistical mechanics, Ann. Math. **101**, 111–259 (1975)
5. Kesten, H., Papanicolaou, G.C.: A limit theorem for turbulent diffusion. Commun. Math. Phys. **65**, 97–138 (1979)
6. Kusuoka, S.: The variational principle for stationary Guassian Markov fields. Theory and Application of Random Fields, Proc. of the IFIP-WG 7/1 Working Conference, Bangalore, India, 1982 (Lecture Notes Control Inf. Sci., vol. 49, pp. 179–187) Berlin Heidelberg New York: Springer 1983
7. Nelson, E.: The free Markov field, J. Funct. Anal. **12**, 211–227 (1973)
8. Olla, S.: Large deviation for almost Markovian processes. Probab. Th. Rel. Fields **76**, 395–409 (1987)
9. Orey, S.: On the Shannon-Perez-Moy theorem. Proc. on Particle systems, random media, and large deviations (Brunswick, Maine, 1984). Contemp. Math. **41**, 319–327 (1985)
10. Rozanov, Yu.A.: Stationary random processes. San Francisco: Holden-Day 1967
11. Simon, B.: The $P(\phi)_2$ Euclidean (quantum) field theory. Princeton: Princeton University Press 1974
12. Stroock, D.: An introduction to the theory of large deviations. Berlin Heidelberg New York: Springer 1984
13. Takahashi, Y.: Entropy functional (free energy) for dynamical systems and their random perturbations, Proc. Taniguchi Symp. on Stoc. Anal. at Katata and Kyoto. Tokyo: Kinokuniya North Holland 1982

Received September 9, 1987