The Annals of Probability
1995, Vol. 23, No. 1, 420–445

MODERATE DEVIATIONS OF DEPENDENT RANDOM VARIABLES RELATED TO CLT[1]

By Wu Liming

Wuhan University

This paper consists of three parts. In the first part, we find a common condition—the C^2-regularity—both for CLT and for moderate deviations. We show that this condition is verified in two important situations: the Lee–Yang theorem case and the FKG system case. In the second part, we apply the previous results to the additive functionals of a Markov process. By means of Feynman–Kac formula and Kato's analytic perturbation theory, we show that the Lee–Yang theorem holds under the assumption that 1 is an isolated, simple and the only eigenvalue with modulus 1 of the operator P_1 acting on an appropriate Banach space $(b\mathscr{E}, C_b(E), L^2 \cdots)$. The last part is devoted to some applications to statistical mechanical systems, where the C^2-regularity becomes a property of the pressure functionals and the two situations presented above become exactly the Lee–Yang theorem case and the FKG system case. We shall discuss in detail the ferromagnetic model and give some general remarks on some other models.

Introduction. The moderate deviation (MD) estimations, like large deviation (LD) estimations, arise from the requirements of statistics. The LD offers a precise estimation associated to LLN (the law of large number), and the MD is often used to give further estimations related to CLT (the central limit theorem) and LIL (the law of iterated logarithm).

Let $(X_k)_{k\geq 1}$ be a sequence of real-valued i.i.d. r.v.'s on $(\Omega, \mathscr{F}, P)$. Set $m = \mathbb{E}X_1$ and $\overline{X}_n = (1/n)\Sigma_1^n X_k$. For many purposes in statistics, one needs to estimate the limit behavior of

$$P\big(\overline{X}_n - m \in A_n\big), \tag{I.1}$$

where A_n, a Borel subset of $\mathbb{R}$, represents the *deviation* of $\overline{X}_n$ from m.

When $A_n = A$, for all n, the term (I.1) is the well known large deviation. If $A_n = (1/\sqrt{n})A$, the CLT tells us the limit behavior of (I.1). If $A_n = (\lambda(n)/\sqrt{n})A$, where $\lambda(n)$ verifies

$$0 < \lambda(n) \to +\infty \quad \text{and} \quad \frac{\lambda(n)}{\sqrt{n}} \to 0, \tag{I.2}$$

Received October 1992.

[1]Partially supported by The Natural Science Foundation of China and the Foundation of Y. D. Fok. This work was completed during the author's visit at Laboratoire de Probabilités, Université Paris VI from January to July 1992.

AMS 1991 *subject classifications.* Primary 60F10, 60F05; secondary 60J25, 60K35.

Key words and phrases. Moderate deviation, CLT, LIL, Markov process, analytic perturbation, Gibbs measure, Lee–Yang theorem, FKG and GHS inequalities.

then the quantity (I.1) becomes the so-called *moderate deviation*. For such A_n, (I.1) can be rewritten as

$$P\left(\frac{\sqrt{n}}{\lambda(n)}\left[\bar{X}_n - m\right] \in A\right). \tag{I.3}$$

So MD is an intermediate estimation between CLT and LD.

Borovgov and Mogulskii (1978, 1980) considered the MD for $(X_k)_{k\geq 1}$ taking values in a Banach space. Under the condition that $\mathbb{E}\exp t\|X_k\| < +\infty$ for all $t > 0$, they obtained that for suitable A,

$$\text{the quantity (I.3)} = \exp\left[-\lambda^2(n)\cdot I(A) + o(\lambda^2(n))\right], \tag{I.4}$$

where $I(A) = \inf_{x\in A}(x^2/2\sigma^2)$ with $\sigma^2 = \mathrm{Var}(X_l)$. In other words, (I.3) verifies a large deviation principle (LDP) with speed $\lambda^{-2}(n)$ and with rate function $I(x) = x^2/2\sigma^2$.

Recently the MD estimations have attracted much attention; see, for example, Baldi (1988), de Acosta (1990), Chen (1990, 1991), Gao (1991) and Ledoux (1992). For $\lambda(n) = n^\alpha$ with $0 < \alpha < \frac{1}{2}$. Chen (1990, 1991) found the necessary and sufficient condition for the MD (I.4) in a Banach space and he obtained the lower bound for general $\lambda(n)$ under a very weak condition. Using the isoperimetric techniques, Ledoux (1992) obtained the necessary and sufficient condition in the general case [i.e., for any sequence $\lambda(n)$ satisfying (I.2)], extending the works of Chen.

However, there are very few works on the MD of dependent random variables until now. The author is only aware of the work of Gao (1991), which discussed the MD of the Doeblin recurrent Markov processes.

The goal of this work is to establish the MD and the CLT for dependent random variables, especially for those arising from Markov processes and from statistical mechanical systems. This paper is composed of three parts. In the first part (Section 1, an abstract), we shall employ the Cramér method of LD theory to establish the MD estimation of a family of probability measures $(\mu_\varepsilon, \beta > 0)$ on $\mathbb{R}$. More precisely, define

$$\Lambda_\varepsilon(t) = \varepsilon\log\int_{\mathbb{R}}\exp(tx/\varepsilon)\,d\mu_\varepsilon(x) := \varepsilon\log Z_\varepsilon(t) \tag{I.5}$$

and assume that

$$\Lambda(t) := \lim_{\varepsilon\to 0}\Lambda_\varepsilon(t) \in (-\infty,+\infty] \quad \text{and } \Lambda < +\infty \text{ on a neighborhood of } 0. \tag{I.6}$$

Define the associated MD

$$\nu_\varepsilon(A) := \mu_\varepsilon\left(x; \frac{x - m(\varepsilon)}{a(\varepsilon)} \in A\right), \tag{I.7}$$

where $m(\varepsilon) = \int_{\mathbb{R}} x\,d\mu_\varepsilon(x) = \Lambda_\varepsilon'(0)$ and $(a(\varepsilon))_{\varepsilon>0}$ satisfies

$$a(\varepsilon) \to 0 \quad \text{and} \quad a(\varepsilon)/\sqrt{\varepsilon} \to +\infty. \tag{I.8}$$

Now introduce the notion of C^2-regularity of $(\mu_\varepsilon)_{\varepsilon\to 0}$:

$$(\mu_\varepsilon)_{\varepsilon\to 0} \text{ is said to be (resp. right, left) } C^2\text{-regular if}$$

$$\text{(I.9)}\quad \Lambda''_\varepsilon(t) \to \Lambda''(t) \quad \text{uniformly for } t\in[-\delta,\delta] \quad (\text{resp. } t\in[0,\delta],\, t\in[-\delta,0]).$$

Our first abstract result (Theorem 1.2) affirms that the right C^2-regularity implies the CLT [i.e., ν_ε with $a(\varepsilon)=\sqrt{\varepsilon}$ tends weakly to $N(0,\Lambda''_+(0))$] and the C^2-regularity implies the MD estimation [i.e., ν_ε defined by (I.7) and (I.8) satisfies LDP with speed $\varepsilon/a^2(\varepsilon)$ with rate function $I(x)=x^2/(2\Lambda''(0))$]. Then the remaining question for us is to check the C^2-regularity. We shall employ two ideas that originated from statistical mechanics. The first is the idea of the Lee–Yang theorem, which consists of locating the zeros of the functions $Z_\varepsilon(t)$ defined by (I.4) (but for t complex), so that the classical theory of analytic functions can be applied. The second is inspired by the FKG systems and the associated GHS inequality, which comprise our second abstract result, Theorem 1.4.

In the second part, we shall consider the applications of the previous abstract results to Markov processes. Using the Feynman–Kac formula, we translate the C^2-regularity in this situation as a property of perturbations of $(P_t)_{t\in T}$ (which is the transition semigroup of the considered process). Under the assumption that 1 is an isolated, simple and the only eigenvalue with modulus 1 in the spectrum of P_t acting on a suitable Banach space (e.g., $b\mathscr{E}, C_b, L^2 \cdots$), we show in Theorem 2.1 that the Lee–Yang theorem holds for bounded additive functionals, by means of Kato's analytic perturbation theory. With the same ideas, we extend Theorem 2.1 to the unbounded case in Sections 2.2 and 2.3 by assuming that the process has some special structures: hyperbounded or reversible w.r.t. an invariant measure m. In particular, we get the CLT and the MD for a wide class of additive functionals, such as sojourn times and local times. As applications, we consider systems of infinite interacting particles and show the CLT and the MD hold under the sufficient condition for the ergodicity given in Liggett (1985). Such examples are provided by the particle systems associated to Gibbs fields at sufficiently high temperature.

We notice also that CLT can be easily obtained by the martingale method [see Jacod and Shiryayev (1987)]. The MD results in this part extend considerably the works of Gao (1991), but the results about CLT are more or less known (except perhaps the uniform convergence for initial laws, established here).

In the last part, we shall apply the results of the first part to statistical mechanics. In the present case, the functions $\Lambda_\varepsilon, \Lambda$ become the pressure functionals and the C^2-regularity becomes a property of the pressure functionals. Thus our results in the first part can be applied successfully to the systems which satisfy the Lee–Yang theorem of FKG and GHS inequalities. We carry out this idea in detail for the ferromagnetic model, and we present many other models to which the results of Section 1 are applicable, such as one-dimensional statistical mechanical systems (including a large class of dynamical systems), the *XY*-model with two or three components, continuous statistical mechanic models and even the $:\phi^4:$ euclidean quantum field. It is worth noticing that the CLT results in this part are known: CLT was obtained by Iogolnizer and Souillard (1979) for Lee–Yang systems and by Newman (1980) for FKG systems (however the methods are different). We are also largely inspired by the works of Ellis [(1985), Chapter 5].

1. An abstract theorem about MD and CLT. The goal of this section is to establish the MD and the CLT under a common condition, the so-called C^2-regularity condition. We shall show that this condition is verified in the following situations: the Lee–Yang theorem case and the FKG and GHS inequalities case.

1.1. *Uniform LDP of Cramér type.* Let $(\mu_\varepsilon^i,\ i \in A,\ \varepsilon > 0)$ be a family of probability measures on $\mathbb{R}$, where A is an index set. For $\varepsilon > 0$ and $i \in A$, define

$$\Lambda_\varepsilon^i(t) = \varepsilon \log \int_{\mathbb{R}} \exp(tx/\varepsilon)\, d\mu_\varepsilon^i(x) := \varepsilon \log Z_\varepsilon^i(t). \tag{1.1}$$

We assume always that there is $\Lambda: \mathbb{R} \to (-\infty, +\infty]$ so that for every $t \in \mathbb{R}$ fixed,

$$\begin{cases} \forall t \in \mathbb{R},\ \Lambda_\varepsilon^i(t) \text{ tends to } \Lambda(t) \text{ uniformly on } i \in A, \text{ as } \varepsilon \to 0 \text{ and} \\ \Lambda \text{ is finite on a neighborhood of zero in } \mathbb{R}. \end{cases} \tag{1.2}$$

The Legendre transformation of Λ is defined as

$$\Lambda^*(s) := \sup_{t \in \mathbb{R}} [ts - \Lambda(t)] \quad \text{for all } s \in \mathbb{R}, \tag{1.3}$$

which is convex, lower semicontinuous and nonnegative on $\mathbb{R}$.

The following result, taken from [Wu (1991c), Theorem 5.3], is a slight generalization of a well known result due to Cramér, Gärtner and Ellis in the LD theory (they coincide if A is a singleton).

PROPOSITION 1.1. *Assume* (1.1) *and* (1.2).

(a) Λ^* *is a good rate function for the uniform upper bound of LD of* $(\mu_\varepsilon^i,\ \varepsilon \to 0)$ *with speed* ε. *More precisely,* Λ^* *is* inf-*compact* (*i.e.,* $[\Lambda^* \le l]$ *is compact,* $\forall\ l \ge 0$) *and for all closed subsets F in* $\mathbb{R}$,

$$\limsup_{\varepsilon \to 0} \varepsilon \log \sup_{i \in A} \mu_\varepsilon^i(F) \le -\inf_{s \in F} \Lambda^*(s). \tag{1.4}$$

(b) *If Λ is moreover differentiable on the interior of* $[\Lambda < +\infty]$ *and* $|\Lambda'(t)| \to +\infty$ *as t tends to the boundary, then* Λ^* *is also a rate function for the uniform lower bound of LD of* (μ_ε^i) *with speed* ε, *that is, for all open subsets G in* $\mathbb{R}$,

$$\liminf_{\varepsilon \to 0} \varepsilon \log \inf_{i \in A} \mu_\varepsilon^i(G) \ge -\inf_{s \in G} \Lambda^*(s). \tag{1.5}$$

[*If* (a) *and* (b) *hold simultaneously, we say that* (μ_ε^i), *as* $\varepsilon \to 0$, *verifies a uniform* (*for* $i \in A$) *LDP with speed* ε *and with rate function* Λ^*.]

REMARKS. (i) We notice that

$$\Lambda^*(s) = 0 \quad \text{iff } s \in [\Lambda'_-(0), \Lambda'_+(0)]. \tag{1.6}$$

So in Proposition 1.1, if Λ is differentiable at 0, then by the part (a),

$$\forall \delta > 0\colon \limsup_{\varepsilon \to 0} \varepsilon \log \sup_{i \in A} \mu_\varepsilon^i(x\colon |x - \Lambda'(0)| > \delta) < 0, \tag{1.7}$$

that is, μ_ε^i tends to the Dirac measure at the point $\Lambda'(0)$ exponentially. This is a reinforced form of the law of large numbers.

(ii) Let $(X_n)_{n \geq 1}$ be a sequence of i.i.d. r.v.'s with values in $\mathbb{R}$ so that $\mathbb{E}\exp(\delta|X_k|) < +\infty$ for some $\delta > 0$. for $\varepsilon = 1/n$, taking μ_ε as the law of $\overline{X}_n = (1/n)\Sigma_1^n X_k$, we have $\Lambda(t) = \Lambda_\varepsilon(t) = \log \mathbb{E}\exp(tX_1)$. So Proposition 1.1 is applicable and gives us Cramér's theorem. If one takes for $\varepsilon = \lambda^{-2}(n)$, ν_ε as the law of $\sqrt{n}/(\lambda(n))(\overline{X}_n - \mathbb{E}X_1)$, where $\lambda(n)$ verifies (I.2), then $\Lambda(t) = \frac{1}{2}\sigma^2 t^2$. Applying Proposition 1.1 to (ν_ε) one get that $(\nu_\varepsilon)_{\varepsilon \to 0}$ satisfies the LDP with speed $\lambda^{-2}(n)$ and rate function $I(x) = \Lambda^*(x) = x^2/2\sigma^2$. This is just the classical MD.

Thus we can say roughly that LLN is determined by the differentiability of Λ at 0. What we shall show in the next paragraph is that the CLT and the MD are determined, roughly speaking, by the second differentiability of Λ at 0.

1.2. *CLT and MD under C^2-regularity.* We begin with the precise definition of right C^2-regularity.

DEFINITION. $(\mu_\varepsilon^i,\ i \in A)_{\varepsilon \to 0}$ is said to be right (resp. left) C^2-regular uniformly for $i \in A$, if $[\Lambda_\varepsilon^i]''(t) \to \Lambda''(t)$ uniformly for $i \in A$ and $t \in [0, \delta]$ (resp. $t \in [-\delta, 0]$), where $\Lambda''(0)$ is interpreted as the right second derivative $\Lambda''_+(0) := \lim_{t \to 0+}(\Lambda'(t) - \Lambda'_+(0))/t$ [resp. $\Lambda''_+(0)$]. If it is simultaneously right and left C^2-regular uniformly on $i \in A$, we say that it is C^2-regular uniformly on $i \in A$.

The following result illustrates the role of right (left) regularity both in CLT and MD.

THEOREM 1.2. *Let* $(\mu_\varepsilon^i;\ i \in A,\ \varepsilon > 0)$ *be a family of probability measures on* $\mathbb{R}$ *satisfying* (1.2). *If it is right* (*resp. left*) C^2-*regular uniformly for* $i \in A$, *then for any* $(a(\varepsilon))_{\varepsilon > 0}$ *verifying* (I.8) *or* $a(\varepsilon) = \sqrt{\varepsilon}$, *and for all* $t \geq 0$ (*resp.* $t \leq 0$),

$$\limsup_{\varepsilon \to 0} \sup_{i \in A} \left| \frac{\varepsilon}{a^2(\varepsilon)} \log \int_{\mathbb{R}} \exp[t(x - m(i,\varepsilon))a(\varepsilon)/\varepsilon]\, d\mu_\varepsilon^i(x) - \frac{1}{2}\Lambda''(0)t^2 \right| = 0, \tag{1.8}$$

where $m(i,\varepsilon)=\int_{\mathbb{R}} x\,d\mu_\varepsilon^i(x)$. *In particular, let*

$$\nu_\varepsilon^i(\cdot) := \mu_\varepsilon^i\left(x;\frac{x-m(i,\varepsilon)}{a(\varepsilon)}\right), \tag{1.9}$$

then under the right or left C^2*-regularity assumption,* ν_ε^i *with* $a(\varepsilon)=\sqrt{\varepsilon}$ *tends uniformly for* $i\in A$ *to the normal law* $N(0,\Lambda''_+(0)$ *or* $\Lambda''_-(0))$ *as* $\varepsilon\to 0$ *weakly. Additionally, under the* C^2*-regularity assumption,* (ν_ε^i)*, with* $a(\varepsilon)$ *verifying* (I.8), *satisfies the uniform* LDP *for* $i\in A$ *with speed* $\varepsilon/(a^2(\varepsilon))$ *and rate function* $I(x)=x^2/(2\Lambda''(0))$.

We begin by a lemma that collects some elementary facts.

LEMMA 1.3. (a) *Let* $(x_n^i,\ i\in A,\ n\in\mathbb{N})$ *be a family of elements in a separated topological space* X *and* $x\in X$. *For* $x_n^i\to x$ *uniformly for* $i\in A$ *as* $n\to\infty$, *it is necessary and sufficient that* $x_n^{i(n)}\to x$ *for any choice of* $(i(n), n\ge 1)$.

(b) *Let* $(\mu_\varepsilon)_{\varepsilon\ge 0}$ *be a family of probability measures on* $\mathbb{R}$. *If*

$$\int_{\mathbb{R}} e^{tx}\,d\mu_\varepsilon(x)\to\int_{\mathbb{R}} e^{tx}\,d\mu_0(x)\quad \text{for } t\in I \text{ as } \varepsilon\to 0,$$

where I *is an interval so that* $0\in I$ *and* $I^0\neq\varnothing$, *then* μ_ε *tends weakly to* μ_0.

(c) *Let* (f_n) *be a sequence of real convex functions on* (a,b), *converging pointwise to* f. *Then* f *is convex and for any* $t_n\to t$ $(t_n,\ t\in(a,b))$, *and* $x_n\in\partial f_n(t_n)$, *the limit points of* $(x_n)_{n\ge 1}$ *lie in* $\partial f(t)$. *Here* $\partial f(t)=[f'_-(t),f'_+(t)]$ *is the set of subdifferentials of* f *at* t. *In particular, if* f_n *and* f *are moreover continuously differentiable on* (a,b), *then* f'_n *tends to* f' *uniformly on all compact subsets of* (a,b).

PROOF. Part (a) is elementary; (b) is taken from Martin-Löf (1973) and (c) is a well known result in convex analysis. □

We turn now to the proof of Theorem 1.2.

PROOF OF THEOREM 1.2. We begin by showing the crucial estimation (1.8) under the right C^2-regularity assumption. Without loss of generality, we can suppose that $m(i,\varepsilon)=[\Lambda_\varepsilon^i]'(0)=0$. In this case, we can rewrite (1.8) as

$$\limsup_{\varepsilon\to 0}\sup_{i\in A}\left|\frac{1}{a^2(\varepsilon)}\Lambda_\varepsilon^i(a(\varepsilon)t)-\frac{1}{2}\Lambda''_+(0)t^2\right|=0. \tag{1.8'}$$

By the assumption (1.2), there is $\delta>0$ such that Λ_ε^i, $i\in A$, $\varepsilon>0$, are all analytic over $(-2\delta,2\delta)$. Applying twice the Newton–Leibnitz formula, we get

$$\Lambda_\varepsilon^i(u)=\int_0^u[\Lambda_\varepsilon^i]''(s)\cdot(u-s)\,ds\qquad\forall\, u\in[0,\delta).$$

Since Λ is twice continuously differentiable on $[0, \delta)$, we have also

$$\Lambda(u) = \Lambda'_+(0)u + \int_0^u \Lambda''(s)(u-s)\,ds \qquad \forall\, u \in [0, \delta).$$

We show now that $\Lambda'_+(0) = 0$. In fact, by the right C^2-regularity,

$$\Lambda'(u) - \Lambda'_+(0) = \int_0^u \Lambda''(s)\,ds = \lim_{\varepsilon \to 0} \int_0^u [\Lambda^i_\varepsilon]''(s)\,ds = \Lambda'(u),$$

and hence $\Lambda'_+(0) = 0$. Combining these facts, we see that

$$\begin{aligned}
&\sup_{i \in A} \sup_{0 \le u < \delta} |u^{-2}\Lambda^i_\varepsilon(u) - u^{-2}\Lambda(u)| \\
&\quad \le \sup_{i \in A} \sup_{0 \le u < \delta} \int_0^u |[\Lambda^i_\varepsilon]''(s) - \Lambda''(s)|(u-s)\,ds/u^2 \\
&\quad \le \sup_{i \in A} \sup_{0 \le u < \delta} |[\Lambda^i_\varepsilon]''(s) - \Lambda''(s)|/2,
\end{aligned}$$

where the last term tends to zero by the right C^2-regularity condition.

Letting $u = a(\varepsilon)t$, we get thus for every $t \ge 0$ fixed,

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \sup_{i \in A} \left| \frac{1}{a^2(\varepsilon)} \Lambda^i_\varepsilon(a(\varepsilon)t) - \frac{1}{2}\Lambda''_+(0)t^2 \right| \\
&\quad \le \lim_{\varepsilon \to 0} \left| \frac{1}{a^2(\varepsilon)} \Lambda(a(\varepsilon)t) - \frac{1}{2}\Lambda''_+(0)t^2 \right| = 0.
\end{aligned}$$

Finally the uniform CLT follows from (1.8) and Lemma 1.3(a) and (b). MD is a direct consequence of Proposition 1.1. □

1.3. *Two situations for C^2-regularity.* The following result consists of two simple remarks that originated from statistical mechanics.

THEOREM 1.4. (a) *Let $Z^i_\varepsilon(z) := \int_{\mathbb{R}} e^{zx}\,d\mu^i_\varepsilon(x)$, $i \in A$, $\varepsilon > 0$ (which are all well defined and holomorphic on a neighborhood of 0 in $\mathbb{C}$). If there is a ball $B(0, \delta)$ in $\mathbb{C}$ such that $|Z^i_\varepsilon(z)| > c > 0$ for all $z \in B(0, \delta)$, $i \in A$, $\varepsilon > 0$, then $\Lambda^i_\varepsilon(z) = \varepsilon \log Z^i_\varepsilon(z)$ converge to a function $\tilde{\Lambda}(z)$ uniformly for $i \in A$ and z belonging to any compact subset of $B(0, \delta)$. $\tilde{\Lambda}$ is holmorphic on $B(0, \delta)$, coincides with $\Lambda(t)$ for $z = t$ real and $[\Lambda^i_\varepsilon]^{(k)}$ tends to $\tilde{\Lambda}^{(k)}$ in the same way. In particular, $(\mu^i_\varepsilon, i \in A)$ is uniformly C^2-regular.*

(b) *If there is $\delta > 0$ so that $[\Lambda^i_\varepsilon]'$ are all concave (or convex) on $[0, \delta)$ and Λ is twice continuously differentiable on $[0, \delta)$, and if*

$$(1.10) \qquad [\Lambda^i_\varepsilon]''(0) \to \Lambda''(0) \quad \text{uniformly for } i \in A, \text{ as } \varepsilon \to 0,$$

then $(\mu^i_\varepsilon, i \in A)$ is uniformly right C^2-regular.

PROOF. Part (a) follows from the classic Montel and Hurwitz theorems in complex analysis and (b) is elementary by Lemma 1.3(c). □

REMARKS. (i) Part (a) is inspired by the Lee–Yang theorem in statistical mechanics, and we shall see in Section 3 that it becomes the Lee–Yang theorem for statistical mechanical systems.

(ii) If one applies part (b) to statistical mechanical systems (as we shall do in Section 3), the condition (1.10) will be guaranteed by FKG inequality and the concave property is a consequence of GHS inequality.

As the readers should see clearly, the results of this section are not difficult (even very easy). However, the new point of view adopted here avoids fastidious estimations in many applications and it often gives a unified way to understand the limit behavior of dependent r.v.'s. These will be justified in the next sections.

2. Markov processes with exponential convergence. In this section, we discuss MD, CLT for additive functionals of a Markov process $(X_t)_{t\in\mathbb{T}}$, where $\mathbb{T}=\mathbb{N}$ or $\mathbb{R}^+$. Using the Feynman–Kac formula, we can translate the C^2-regularity as a property of the perturbations of $(P_t)_{t\in\mathbb{T}}$. Applying Kato's analytic perturbation theory, we show that the Lee–Yang theorem holds for bounded additive functionals under the assumption that 1 is an isolated, simple and the only eigenvalue with modulus 1, of P_t acting on an appropriate Banach space (e.g., $b\mathscr{E}, C_b, L^2$, etc.). This is the content of Section 2.1. In Sections 2.2 and 2.3, we shall extend them to unbounded additive functionals and, in particular, we shall discuss sojourn times and local times.

2.1. *Main results: bounded case.* We describe first the framework of this section. Let E be a Polish space and $\mathscr{E}$ its Borel σ-field. We denote by $b\mathscr{E}$ (resp. C_b) the space of real bounded measurable (resp. continuous) functions on E and by $M_b(E)$ [resp. $M_1(E)$] the space of all measures of finite variation (probability) on $(E,\mathscr{E})$. For a measurable function f and a measure μ, we write $\|f\|=\sup_{x\in E}|f(x)|$ and $\mu(f)=\int f\,d\mu$.

Let $(P_t)_{t\in\mathbb{T}}$ be a semigroup of Markov kernels on E with $P_0(x,\cdot)=\delta_x$, where $\mathbb{T}=\mathbb{N}$ or $\mathbb{R}^+$. Let $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\in\mathbb{T}},(X_t)_{t\in\mathbb{T}},(\theta_t)_{t\in\mathbb{T}},(\mathbb{P}_x)_{x\in E})$ be a Markov process (P_t) as semigroup of transition. We suppose always that (P_t) possesses an invariant probability measure m, and in the continuous time case, (X_t) is càdlàg, that is, $X.\in\mathbb{D}(\mathbb{R}^+,E)$. We write $\mathbb{P}_\mu=\int\mathbb{P}_x\,\mu(dx)$ for an initial measure $\mu\in M_1(E)$ and we denote P_1 by P for simplicity.

Define the empirical measures (L_t) by

$$(2.1)\qquad L_t(\omega)=\begin{cases}\dfrac{1}{t}\displaystyle\sum_{s=0}^{t}\delta_{X_s(\omega)}, & \text{if } \mathbb{T}=\mathbb{N},\\[2ex] \dfrac{1}{t}\displaystyle\int_0^t\delta_{X_s(\omega)}\,ds, & \text{if } \mathbb{T}=\mathbb{R}^+\end{cases}$$

and set for $f \in b\mathscr{E}$,

$$(2.2)\quad V(f) := \begin{cases} \langle (f - m(f))^2 \rangle_m + 2 \sum_{k=1}^{\infty} \langle P_k f, f - m(f) \rangle_m, & \text{if } \mathbb{T} = \mathbb{N}, \\ 2\int_0^{\infty} \langle P_t f, f - m(f) \rangle_m \, dt, & \text{if } \mathbb{T} = \mathbb{R}^+. \end{cases}$$

The main result of this section is the following.

THEOREM 2.1. (a) *If* 1 *is an isolated, simple and the only eigenvalue with modulus* 1 *of* P_{t_0} *acting on* $(b\mathscr{E}, \|\cdot\|)$ *for one* $t_0 \in \mathbb{T}$, *then for every* $f \in b\mathscr{E}$:

(2.3) $\mathbb{P}_x(\sqrt{t}\,[L_t(f) - m(f)] \in \cdot)$ *converge weakly to* $N(0, V(f))$ *as* $T \to \infty$, *and this convergence is uniform for initial states* $x \in E$;

(2.4) *as* $t \to +\infty$, $(\mathbb{P}_x(\sqrt{t}/(\lambda(t))[L_t(f) - m(f)] \in \cdot))_{x \in E}$ *satisfies a uniform* LDP *for* $x \in E$ *with speed* $\lambda^{-2}(t)$ *and rate function* $I_f(s) = s^2/(2V(f))$,

where $\lambda(t)$ *satisfies* (I.2).

(b) *Assume* (P_t) *is Feller. Part* (a) *holds if one changes* $b\mathscr{E}$ *by* $C_b(E)$.

(c) *If* 1 *is an isolated, simple and the only eigenvalue with modulus* 1 *of* P_{t_0} *acting on* $L^2(m)$ *for some* $t_0 \in \mathbb{T}$, *then for* $f \in b\mathscr{E}$, *the CLT and the MD still hold for* $\{L_t(f) - m(f), t \to \infty\}$ *under the laws* $\mathbb{P}_{h\cdot m}$, $h \in L^2(m)$, *and they are uniform for the initial distributions* $\mu \in U_m$. *Here* $U_m = \{\mu = h \cdot m \in M_1(E);\ \langle h^2 \rangle \le l\}$ *where* l *is an arbitrary constant.*

REMARKS. (i) For a Markov kernel P, we have the equivalence between the following properties [see Revuz (1976)]:

(a) $P = P_{t_0}$ satisfies the assumption (a) of Theorem 2.1.

(b) P is quasicompact on $b\mathscr{E}$, Harris recurrent and aperiodic.

(c) P is Doeblin recurrent and aperiodic.

(d) P is Harris recurrent, aperiodic and E is a small set [i.e., $\exists$ ($c > 0$, $k \in \mathbb{N}$ and $\nu \in M_1(E)$) so that $P_k(x, dy) \ge c \cdot \nu(dy)$].

(e) P is Harris positively recurrent and for any $f \in b\mathscr{E}$,

$$P^n f \text{ tends to } m(f) \text{ uniformly on } E.$$

(ii) In the discrete time case, the Doeblin recurrence is a necessary condition for the uniform LDP of $(L_t)_{t\to\infty}$ on $(M_1(E), \tau)$, where τ is the weak topology $\sigma(M_b(E), b\mathscr{E})$. This is indicated in Wu (1991b).

(iii) The CLT is well known under the assumption of Doeblin recurrence (except perhaps the uniform convergence stated above).

(iv) The assumption in (c) is weaker than that of (a) or of (b). In fact, suppose that the assumption in (b) holds, which is obviously weaker than that of (a). We have

$$\exists\, c > 0 \text{ so that } \|P_{t_0} f - m(f)\| \le e^{-c}\|f\|, \qquad \forall f \in C_b(E).$$

On the other hand, $P_{t_0} - m$ is contractive on all L^p, $p \geq 1$. Thus by the theory of interpolation of operators, the norm of P_{t_0} as an operator on $L_0^2(m)$ is strictly less than 1 [see Stein and Weiss (1975)]. So the desired result follows.

PROOF OF THEOREM 2.1 IN THE DISCRETE TIME CASE. For $f \in b\mathscr{E}$, define

$$P_f(x, dy) := e^{f(x)} P(x, dy). \tag{2.5}$$

We have the following simplified Feynman–Kac formula in the present case:

$$(P_f)^n g(x) = \mathbb{E}^x \exp\left[\sum_{k=0}^{n-1} f(X_k)\right] g(X_n). \tag{2.6}$$

Thus taking μ_ε^x as the law of $L_n(f)$ under $\mathbb{P}_x$ for $\varepsilon = 1/n$, then the function Λ_ε^x defined in Section 1 equals $(P_f)^n 1(x)$. Consequently, the C^2-regularity becomes a property of the perturbation operator P_f of P. Our basic tool will be the following result in analytic perturbation theory, taken from Kato (1984).

THEOREM [Kato (1984), [K], Chapter 7, Theorems 1.7 and 1.8]. *Let B be a Banach space and $(A(z))_{z \in D}$ be a holomorphic family of bounded operators on B, where D is an open domain in $\mathbb{C}$ [i.e., $z \to A(z)x$ is differentiable on D for every $x \in B$]. Suppose that $A(z_0)$ has an isolated point λ_0 in its spectrum and the corresponding eigenspace E_0 is one dimensional. Let E_1 be the subspace corresponding to the remaining part of the spectrum of $A(z_0)$. Then there is an open neighborhood U of z_0 such that $\forall\, z \in U$, the spectrum $\Sigma(z)$ of $A(z)$ is separated into two parts: $\Sigma(z) = \Sigma_0(z) \cup \Sigma_1(z)$, which satisfies:*

(i) *$\Sigma_0(z) = \{\lambda(z)\}$, where $\lambda(z)$ is an isolated and simple eigenvalue of $A(z)$ and $\lambda(z)$ is holomorphic on U with $\lambda(z_0) = \lambda_0$.*

(ii) *If one denotes by $E_0(z)$ [resp. $E_1(z)$] the subspace corresponding to $\Sigma_0(z)$ and $\Sigma_1(z)$, then the family of projections $(J_0(z))_{z \in U}$ on $E_0(z)$ along $E_1(z)$ is also holomorphic.*

We now introduce some functions:

(2.7) $G^1(z)$ = the complex number with the largest modulus in the spectrum of P_{zf} regarded as an operator on $(b\mathscr{E}, \|\cdot\|)$;

(2.8) $G^2(z)$ = the complex number with the largest modulus in the spectrum of P_{zf} regarded as an operator on $(C_b(E), \|\cdot\|)$;

(2.9) $G^3(z)$ = the complex number with the largest modulus in the spectrum of P_{zf} regarded as an operator on $L^2(m)$.

Under the assumption of Theorem 2.1(a) [resp. (b) and (c)], G^1 (resp. G^2 and G^3) is well defined and holomorphic over a neighborhood U of 0 in $\mathbb{C}$. We turn now to the proof of Theorem 2.1. in the discrete time case.

We begin with the case $t_0 = 1$.

First it is easy to show that $(z \to P_{zf})_{z \in \mathbb{C}}$ is a holomorphic family of bounded operators on $B = b\mathscr{E}$ or $C_b(E)$ or $L^2(m)$. In the following, we deal only with the case where $B = b\mathscr{E}$. The other cases can be treated in the same way.

Identifying $A(z)$ with P_{zf}, λ_0 with 1 and using the notations in Kato's theorem above, we have $\lambda(z) = G^1(z)$. Since $I - J_0(z)$ is the projection on $E_1(z)$ along $E_0(z)$, we have

$$P_{zf}^n - \left[G^1(z)\right]^n J_0(z) = P_{zf}^n(I - J(z)).$$

On the other hand, because the largest modulus in the spectrum of P restricting to E_1 is strictly smaller than 1 by the hypothesis, we can choose a neighborhood U of 0 sufficiently small in $\mathbb{C}$, such that

$$\begin{aligned} &\|P_{zf}^n - \left[G^1(z)\right]^n J_0(z)\| \le e^{-\gamma n} \quad \text{and} \\ &|G^1(z)| > e^{-\gamma} \qquad \text{where } \gamma > 0, \\ &\inf_{x \in E} |[J_0(z)1](x)| > 1/2 \quad \text{for any } z \text{ in } U. \end{aligned} \tag{2.10}$$

Therefore we have by (2.10),

$$\sup_{z \in U} \sup_{x \in E} \left| \frac{1}{n} \log P_{zf}^n 1(x) - \log G^1(z) \right| \to 0, \quad \text{as } n \to \infty. \tag{2.11}$$

Thus with the identifications given at the beginning of this paragraph, we see that the condition in Theorem 1.4(a) is verified. Hence the CLT (2.3) and the MD (2.4) follow from the Theorem 1.2 if we note that

$$\frac{1}{\sqrt{n}} \|\mathbb{E} \cdot L_n(f) - m(f)\| = \frac{1}{\sqrt{n}} \left\| \frac{1}{n} \sum_0^{n-1} P_k f - m(f) \right\| \to 0, \tag{2.12}$$

which is obvious by the assumption.

For the general case where $t_0 \neq 1$, we shall consider the perturbation $(P_{zf})^{t_0}$ of P_{t_0}. Following the same proof as above, we can get the CLT (2.3) and MD (2.4) for t tending to infinity along the subsequence $\{nt_0, n = 1, 2, \dots\}$. The passage to the whole sequence is easy. □

Proof of Theorem 2.1 in the continuous time case. The proof of Theorem 2.1 becomes more difficult in the continuous time case, because $(P_t)_{t \ge 0}$ is not strongly continuous, the semigroup theory cannot be applied directly. Fortunately the Feynman–Kac formula holds always and it will help us to overcome this difficulty.

In the continuous time case, in place of (2.5) and (2.6), we define the Feynman–Kac operators $(P_t^f)_{t \ge 0}$ as follows:

$$P_t^f g(x) := \mathbb{E}^x g(X_t) \exp \int_0^t f(X_s)\, ds. \tag{2.5'}$$

By the Markov property, we see that (P_t^f) is a semigroup. Instead of regarding $(P_t^f)_{t \ge 0}$ as a perturbation of the semigroup (P_t), we shall consider P_t^f for t fixed.

LEMMA 2.2. *For every $f \in b\mathscr{E}$ and $t > 0$ fixed, $(z \to P_t^{zf})_{z\in\mathbb{C}}$ is a holomorphic family of bounded operators on $B = b\mathscr{E}$ or $L^2(m)$. Furthermore, it is holomorphic on $C_b(E)$ if $f \in C_b(E)$ and if (P_t) is Feller.*

Its proof is elementary, so it is omitted.

Now let us modify the definition of G^i $(i = 1, 2, 3)$ given in (2.7), (2.8) and (2.9) as follows:

$$G^1(z)\ (\text{resp. } G^2; G^3) := \text{the spectral radius of } P_{t_0}^{zf} \text{ considered as an operator on } B = b\mathscr{E}\ [\text{resp. } C_b(E), L^2(E)],$$

where t_0 is the index given in the assumption of Theorem 2.1.

We prove now Theorem 2.1(c) in the continuous time case. By Lemma 2.2 and Kato's theorem, we get with the same arguments as in the discrete time case that

$$(2.11') \quad \sup_{z\in U}\sup_{\mu\in U_m}\left|\frac{1}{n}\log\int\left[P_{nt_0}^{zf}1\right](x)\,d\mu - \log G^3(z)\right| \to 0 \quad \text{as } n\to\infty.$$

where U is a neighborhood of 0 in $\mathbb{C}$ and U_m is defined in Theorem 2.1. Now applying the results of Section 1 with $A = U_m$, $\varepsilon = 1/n$ and $i = \nu \in U_m$, μ_ε^i = the law of $(1/nt_0)\int_0^{nt_0} f(X_s)\,ds$ under P_ν, we get them the CLT (2.3) and the MD (2.4) for t tending to infinity along $t = nt_0$.

For the control of the limits as $T\to\infty$ along the whole $\mathbb{R}^+$, it is enough to notice that

$$\left|\int_0^t f(X_s)\,ds - \int_0^{[t/t_0]t_0} f(X_s)\,ds\right| \le \|f\|\cdot t_0.$$

Finally (a) and (b) can be shown in the same way. □

2.2. *Some complements to Theorem* 2.1.

2.2.1. *Discrete time case: MD for empirical measures* (L_t).

THEOREM 2.3. *In the context of Theorem* 2.1(c), $(P_\mu[\sqrt{n}/(\lambda(n))\,(L_n - m) \in \cdot],\ \mu\in U_m)$ *satisfies a uniform LDP on* $(M_b(E), \tau)$ *with speed* $\lambda^{-2}(n)$ *and with rate function* $I\colon M_b(E)\to[0,+\infty]$ *given by*

$$(2.13) \qquad I(\nu) = \begin{cases} \frac{1}{2}\int f(Q+Q^*-I)^{-1} f\,dm, & \text{if } \nu \ll m \text{ and } f := d\nu/dm \in L_0^2(m),\\ +\infty, & \text{otherwise},\end{cases}$$

where $Q = I + P + P^2 + \cdots$ *is the potential operator* [*acting on* $L_0^2(m)$] *and* Q^* *is its adjoint on* $L_0^2(m)$.

NOTE. Since $Q + Q^* - I \ge I$, it is invertible on $L_0^2(m)$.

PROOF. By Theorem 2.1, we have for every $f\in b\mathscr{E}$ that

$$\frac{1}{\lambda^2(n)}\log\mathbb{E}^\mu\exp\lambda(n)\sqrt{n}\,(L_n(f) - m(f))$$

converges uniformly to $\Lambda(f) := \frac{1}{2}V(f)$ for $\mu \in U_m$, as $n \to \infty$. The so-defined function $\Lambda\colon b\mathscr{E} \to \mathbb{R}$ is convex, Gateaux differentiable, and its Legendre transformation $\Lambda^*\colon M_g(E) \to [0, +\infty]$ defined by

$$\Lambda^*(\nu) := \sup_{f \in b\mathscr{E}} [\nu(f) - \Lambda(f)]$$

can be calculated by means of variational calculus and it equals $I(\cdot)$ given by (2.13). Now by Theorem 4.1 in Wu (1991c) [see also Dawson-Gärtner (1989)], it remains to show that:

if $\nu \in b\mathscr{E}^*$ (the algebric dual) verifies
$\overline{\Lambda}^*(\nu) = \sup_{f \in b\mathscr{E}}[\nu(f) - \Lambda(f)] < +\infty$, then ν is a measure of finite variation, that is $\nu \in M_b(E)$.

For proving such ν is a measure, we have only to show that $\lim_{n \to \infty}\nu(f_n) = 0$ for every sequence (f_n) in $b\mathscr{E}$ verifying $\sup_n\|f_n\| < \infty$ and $\lim f_n(x) = 0$ $\forall x \in E$. To do this, since $V(f) \le 2\langle f - m(f), Q(f) - m(f)\rangle$ by definition (2.2), and Q is bounded on $L_0^2(m)$, we have

$$\Lambda(cf_n) \to 0 \quad \text{for any such sequence } (f_n) \text{ and for all } c \in \mathbb{R}.$$

Hence

$$+\infty > \sup_n [c \cdot \nu(f_n) - \Lambda(cf_n)] \ge c \cdot \limsup_{n \to \infty}\Big(\text{or } \liminf_{n \to \infty}\Big)\nu(f_n).$$

Because c is arbitrary, we obtain the desired result. □

REMARKS. In the same way, the MD in this theory holds uniformly for $\mu \in M_1(E)$ under the assumption of part (a) or (b) of Theorem 2.1. This result extends the previous works of Gao (1991, 1993) on the Doeblin recurrent Markov chains, because there are many Markov processes with noncompact state space, which satisfies the condition of Theorem 1.2(c), but not that of the part (a).

Recently Gao (1993) discovered that the Doeblin recurrence is also necessary for the uniform MD w.r.t. $\mu \in M_1(E)$.

Unbounded case. The following result extends Theorem 2.1 to the unbounded case.

THEOREM 2.4. *In the context of Theorem* 2.1(c), *suppose moreover that*

(H) $\exists\, t > 0, \exists\, q > p > 1$ *so that* $P_t\colon L^p(m) \to L^q(m)$ *is bounded*

(say, P_t is hyperbounded), then for every real measurable function f on E satisfying

$$\int_E \exp(\delta|f|)\, dm < +\infty \quad \textit{for some } \delta > 0 \tag{2.14}$$

$[L_t(f) - m(f), t \to \infty]$ *satisfies CLT and MD uniformly for the family of laws* $\{\mathbb{P}_\mu, \ \mu \in U_m^p\}$ *for every* $p > 1$, *where* $U_m^p := \{\mu \in M_1(E);\ \mu \ll m$ *and* $\|d\mu/dm\|_p \le l\}$ *where l is an arbitrary constant.*

PROOF. We begin with the situation where $t_0 = 1$ in the assumption of Theorem 2.1 and $t = 1$ in (H) above; in other words, P is hyperbounded and has a gap near 1 in its spectrum.

Notice that P: $L^p \to L^p$ also has a gap near 1 in its spectrum for every $p > 1$ under the assumption of Theorem 2.1(c) (by Riesz–Thorin's interpolation theorem). Consider the operator

$$P^f(x, dy) := e^{f(y)}P(x, dy). \tag{2.15}$$

We have another version of the Feynman–Kac formula given by

$$(P^f)^n g(x) = \mathbb{E}^x g(X_n) \exp \sum_{k=1}^{n} f(X_k). \tag{2.16}$$

We show now that for any $q > 1$, there is a constant $\delta > 0$ so that $P^{\delta|f|}$ is bounded on $L^q(m)$ under the hypothesis (2.14). In fact, for $q > 1$ fixed, by Riesz–Thorin's interpolation theorem, there is $p \in (1, q)$ such that P is bounded from L^p to L^q. Let a, b be two conjugated numbers. For a nonnegative $g \in L^q$, we have

$$\begin{aligned}
\int_E (P^f g)^q \, dm &\le \int_E [\mathbb{E}^x \exp af(X_1)]^{q/a} [\mathbb{E}^x g^b(X_1)]^{q/b} \, dm \\
&\le \left[\int_E \exp aqf(s) \, dm\right]^{1/a} \left\{\int_E [Pg^b(x)]^q \, dm\right\}^{1/b} \\
&\le \left[\int_E \exp aqf(x) \, dm\right]^{1/a} \|g^b\|_p^{q/b}.
\end{aligned}$$

Choosing b so that $bp = q$, we get that P^f is bounded on L^q once f verifies

$$\int_E \exp aqf(x) \, dm < +\infty,$$

and so the affirmation above follows.

Now consider the family of operators $(z \to P^{zf})$, which are bounded for z in a neighborhood of 0 in $\mathbb{C}$, by what has been proved. To show it is holomorphic, we need only to show that $z \to \langle u, P^{zf} v \rangle_m$ is analytic for every $u, v \in b\mathscr{E}$. This is easy. Now we can apply Kato's perturbation theorem and the remaining proof is the same as that of Theorem 2.1 [what will be obtained is the limit behavior of $L_n \circ \theta(f)$, in place of $L_n(f)$. The remaining passage is easy]. □

REMARKS. (i) This theorem extends the results of Borovkov and Mogulskii (1980) for the i.i.d. case. Condition (2.14) is the same as their condition. Unfortunately, we have not found the counterpart of the necessary and sufficient condition for MD in the i.i.d. case, obtained previously by Chen and Ledoux.

(ii) In Wu (1991–1993, 1991b), it is shown that the hyperboundedness (H) is also a sufficient condition for the LDP of L_n $(n \to \infty)$ on $(M_1(E), \tau)$; see these papers for a detailed discussion of hyperboundedness.

2.2.2. *Continuous time case.* With the same proof as that of Theorem 2.4, we can establish the following theorem.

THEOREM 2.3′. *In the context of Theorem* 2.1(*c*), $\{P_\mu[\sqrt{t}/(\lambda(t))(L_t - m) \in \cdot],\ \mu \in U_m\}$ *satisfies as* $t \to \infty$ *a uniform LDP on* $(M_b(E), \tau)$ *with speed* $\lambda^{-2}(t)$ *and with rate function* I: $M_b(E) \to [0, +\infty]$ *given by*

$$I(\nu) = \sup\{\nu(f) - \tfrac{1}{2}\langle f, (Q + Q^*) f\rangle;\ f \in L_0^2(m)\}$$

$$(2.17)\qquad = \begin{cases} \tfrac{1}{2}\langle f, (Q + Q^*)^{-1} f\rangle, & \\ \qquad \text{if } \nu \ll m \text{ and} & \\ \qquad f := d\nu/dm \in L_0^2(m) \cap \operatorname{Dom}\left[(Q + Q^*)^{-1/2}\right], & \\ +\infty, \qquad \text{otherwise}, & \end{cases}$$

where $Q = \int_0^\infty (P_t - m)\,dt$ *is the potential operator* [*acting on* $L_0^2(m)$] *and* Q^* *is its adjoint on* $L_0^2(m)$.

REMARKS. The last term in (2.17) is well defined: since $Q + Q^*$ is self-adjoint, injective on $L_0^2(m)$ and its range is dense in $L_0^2(m)$, $(Q + Q^*)^{-1}$ is self-adjoint too on $L_0^2(m)$ and $\langle f, (Q + Q^*)^{-1} f\rangle$ is then its quadratic form.

If (P_t) is symmetric w.r.t. m, then $I(\nu)$ can be explicitly calculated out by means of its Dirichlet form as follows:

$$(2.17')\qquad I(\nu) = \begin{cases} \tfrac{1}{4}\mathscr{E}(f, f), & \text{if } \nu \ll m \text{ and } f := d\nu/dm \in L_0^2(m) \cap \operatorname{Dom}(\mathscr{E}), \\ +\infty, & \text{otherwise}, \end{cases}$$

where $\mathscr{E}(f, g) := \langle \sqrt{-\mathscr{L}} f, \sqrt{-\mathscr{L}} g\rangle_m$ for $f, g \in \operatorname{Dom}(\mathscr{E}) = \operatorname{Dom}(\sqrt{-\mathscr{L}})$ and $\mathscr{L}$ is the generator of the semigroup (P_t) on $L^2(m)$.

In the continuous time case, there are many other important additive functionals besides those studied in Theorem 2.1, such as sojourn times and local times. This is our goal in this paragraph and the next.

Let $A = (A_t)_{t \ge 0}$ be an additive functional (i.e., $A_{t+s} = A_t + A_s \circ \theta_t$, $\forall\, t, s \ge 0$) such that A_t is $\mathscr{F}_t$-measurable, $\forall\, t \ge 0$. We associate A with the Feynman–Kac operator [as in (2.5′)]

$$(2.18)\qquad P_t^A g(x) := \mathbb{E}^x g(x_t) e^{A_t}.$$

THEOREM 2.5 (in the continuous time case $\mathbb{T} = \mathbb{R}^+$). (a) *If* (P_t) *verifies the assumption of Theorem* 2.1(a) *and if* A_t *is bounded for every* $t \ge 0$, *then the limit*

$$(2.19)\qquad V(A) := \lim_{t \to \infty} \frac{1}{t}\left[\mathbb{E}^m A_t^2 - (\mathbb{E}^m A_t)^2\right]$$

exists and we have CLT and MD about the limit behavior of $A_t - \mathbb{E}^m A_t$ *uniformly for the laws* $(\mathbb{P}_x,\ x \in E)$, *as in Theorem* 2.1(a).

(b) *Under the assumption of Theorem* 2.1(c), *if* A *verifies* $\exists\, \delta > 0$, $\exists\, t^0 > 0$, *so that* $P_{t^0}^{\pm\delta A}$ *are bounded operators on* $L^2(m)$ *and*

$$\sup_{0\le s\le t^0} \mathbb{E}^m \exp(\delta|A_s|) < \infty, \tag{2.20}$$

then the limit in (2.19) *exists and the corresponding CLT and MD hold uniformly for the laws* $(\mathbb{P}_\mu,\ \mu \in U_m)$.

PROOF. Considering $z \to P_{t_0}^{zA}$ as a analytical perturbation of P_{t_0}, we can show (a) in the same way as in Theorem 2.1. For (b), we assume without loss of generality that $t_0 = 1$ in the assumption of Theorem 2.1(c) and consider $P_{t^0}^{zf}$ as a perturbation of P_{t^0} (in the general case, consider $P_{t_0 t^0}^{zf}$). We get as in Theorem 2.1 the corresponding CLT and MD, but only for $t \to \infty$ along $t = nt^0$, $n \in \mathbb{N}$. For the limit along the whole $\mathbb{R}^+$, there is no question for CLT since

$$\sup\left\{\mathbb{E}^\mu\left[A_t - \mathbb{E}^m A_t - (A_{nt^0} - \mathbb{E}^m A_{nt^0})\right]^2,\ \mu \in U_m,\ t \in \left[nt^0, (n+1)t^0\right]\right\} \to 0,$$
$$\text{as } n \to \infty.$$

For MD, by a comparison lemma in the LD theory [see Wu (1991c), Section 1], we have only to show that

$$\lim_{L\to\infty} \limsup_{t\to\infty} \lambda^{-2}(t) \sup_{\mu\in U_m} \mathbb{P}_\mu\left\{\frac{1}{\sqrt{t}\,\lambda(t)}|A_t - A_{[t/t^0]t^0}| > L\right\} = -\infty,$$

which can be established easily by Markov's inequality and by assumption (2.20). □

REMARKS. For unbounded additive functionals, what is difficult to check in the assumption of Theorem 2.6(b) is the boundedness of the operator P_t^f on $L^2(m)$. As in the discrete time case (Theorem 2.4), this boundedness is guaranteed by condition (2.20) under the hyperboundedness of (P_t). The following result is then counterpart of Theorem 2.4 in continuous time.

THEOREM 2.4′. *In the context of Theorem* 2.1(c), *if* (P_t) *is hyperbounded in the sense of* (H) *and if* (2.20) *is satisfied, then* $A = (A_t)_{t\ge 0}$ *satisfies* CLT *and* MD *uniformly for* $\mathbb{P}_\mu$, $\mu \in U_m^p$, *where* $p > 1$ *is arbitrary and* U_m^p *is given in Theorem* 2.4.

The theorem has many consequences, shown by the following corollary.

COROLLARY 2.6. *In the context of Theorem* 2.4′, *assume moreover that* $(X_t)_{t\ge 0}$ *is continuous. Then the affirmations of Theorem* 2.4′ *are valid for* $A = (A_t)$ *given by*:

(i) $A_t = \int_0^t f(X_s)\,ds$ *with f satisfying*

$$\exists\, \delta > 0 \text{ such that } \int_E \exp(\delta|f|)\,dm < \infty. \tag{2.21}$$

(ii) $(A_t = M_t)$ *is a continuous martingale with* $\langle M\rangle_t = \int_0^t f(X_s)\,ds$, *where f satisfies* (2.21).

(iii) (A_t) *is the local time* $L_t^0(f)$ *of the semimartingale* $(f(X_t))_{t\ge 0}$ *at* 0, *where* f *and* f^2 *belong to the domain of* $\mathscr{L}$ [*in* $L^2(m)$] *and satisfy the condition*

$$(2.22) \qquad |f|, \mathscr{L}f, \Gamma(f,f) := \mathscr{L}f^2 - 2f\mathscr{L}f \quad \textit{satisfy all } (2.21).$$

PROOF. By Theorem 2.4′, we have only to justify condition (2.20). Part (i) follows from Theorem 2.4′.

(ii) Consider $\{\exp(\lambda M_t - \lambda^2/2\langle M\rangle_t)\}_{t\ge 0}$, which is local martingale for any $\lambda \in \mathbb{R}$. It is a martingale for all λ small enough by (i) above and by Novikov's criterion. Thus we have for any couple of conjugated numbers (p,q) and $a > 0$,

$$\begin{aligned}\mathbb{E}^m \exp \lambda M_t &= \mathbb{E}^m \exp(\lambda M_t - a\langle M\rangle_t + a\langle M\rangle_t)\\ &\le \left[\mathbb{E}^m \exp(\lambda p M_t - ap\langle M\rangle_t)\right]^{1/p}\left[\mathbb{E}^m \exp(aq\langle M\rangle_t)\right]^{1/q}.\end{aligned}$$

Choosing $p = q = 2$ and $a = \lambda^2$, then by this estimate we see that for $|\lambda|$ sufficiently small, $\mathbb{E}^m \exp \lambda M_t$ is bounded on finite intervals of time. Thus $A = M$ satisfies condition (2.20).

(iii) Recall Tanaka's formula for local time,

$$|f(X_t)| - |f(X_0)| = \int_0^t \operatorname{sign}(t(X_s))\, df(X_s) + L_0^f,$$

and notice

$$f(X_t) - f(X_0) = M_t + \int_0^t \mathscr{L}f(X_s)\, ds,$$

where (M_t) is a (continuous) martingale with $\langle M\rangle_t = \int_0^t \Gamma(f,f)(X_s)\, ds$ (see [D-M]). Now it is easy to deduce (2.20) from (i) and (ii) by means of Hölder's inequality. □

2.3. *Some further studies.*

2.3.1. *Symmetric case.* The hyperboundedness is often difficult to check. In the symmetric case, it is possible to handle the boundedness of (P_t^A) assumed in Theorem 2.5 in another way. The following result is taken from Wu (1994).

THEOREM 2.7. *Suppose* $(P_t)_{t\in\mathbb{R}^+}$ *is symmetric w.r.t. m. Let* $A_t = \int_0^t f(X_s)\, ds$. *Then* $(P_t^f)_{t\ge 0}$ *is a strongly continuous semigroup on* $L^2(m)$ *iff*

$$\Lambda(f) := \sup\left\{\int u^2 f\, dm - \mathscr{E}(u,u);\ u \in \mathrm{Dom}(\mathscr{E}) \textit{ and } \langle u^2\rangle_m \le 1\right\} < +\infty.$$

In particular, all assumptions in Theorem 2.5(b) *for* (A_t) *will be verified once*

$$(2.23) \qquad \Lambda(\pm\delta f) < +\infty \quad \textit{for some } \delta > 0.$$

2.3.2. *Multivariate functionals.* It is interesting to extend the previous results to multivariate additive functionals, such as $A_n = \sum_{k=0}^{n-1} f(X_k, \dots, X_{k+d})$, where $f\colon E^d \to \mathbb{R}$ in the discrete time case, or $A_t = \int_0^t f(\theta_s X.)\, ds$, where f: $\mathbb{D}([0,\infty), E) \to \mathbb{R}$ is $\mathscr{F}_{t_0}$-measurable. For this purpose, consider the process $(X_t^{(s)})_{t\in\mathbb{T}}$ defined by

$$X_t^{(s)} := \begin{cases} (X_t, X_{t+1}, \dots, X_{t+s-1}) \in E^s, & \text{if } \mathbb{T} = \mathbb{N} \text{ and } t, s \in \mathbb{N}, \\ X_{[t,t+s]} \in \mathbb{D}([0,s], E), & \text{if } \mathbb{T} = \mathbb{R}^+ \text{ and } t, s \in \mathbb{R}^+. \end{cases}$$

Then the multivariate functionals of (X_t) above become the adapted additive functionals of $(X_t^{(s)})$. The following result allows us to apply the previous results to such functionals.

PROPOSITION 2.8. *If $(X_t)_{t\in\mathbb{T}}$ satisfies the assumption of the part* (a) [*resp.* (b); (c)] *of Theorem* 2.1, *then $(X_t^{(s)})_{t\in\mathbb{T}}$ satisfies the same property both in the discrete or continuous time case.*

PROOF. We deal only with the discrete time case and only with the situation of the part (c) of Theorem 2.1. The other cases can be treated similarly.

Assume without loss of generality that $t_0 = 1$. Denote by $m^{(d)}$ the law of $X_0^{(d)}$ under $\mathbb{P}_m$, and by $P^{(d)}$ the one-step transition kernel of $(X_n^{(d)})_{n\in\mathbb{N}}$. For every function $f \in L^2(E^d, m^{(d)})$, we denote by $g(x)$ the function $\mathbb{E}^x f(X_0^{(d)})$, and we have

$$[P^{(d)}]^n f(x_0, \dots, x_{d-1}) = \mathbb{E}^{x_{d-1}}\mathbb{E}\big(f(X_n^{(d)})|X_n\big) = P^{n-d}g(x_{d-1}) \qquad \forall\, n \geq d,$$

which implies

$$\|[P^{(d)}]^n f - m^{(d)}(f)\|_{L^2(E^d, m^{(d)})} = \|P^{n-d}g - m(g)\|_{L^2(E,m)}.$$

As n increases to infinity, the last term above tends exponentially to $m(g)$ by the assumption of Theorem 2.1(c). This implies the desired result. □

REMARKS. From this result, we can get CLT and MD for those functionals presented above. However, it does not allow us to obtain CLT and MD for additive functionals such as $A_n = \Sigma_0^{n-1}(f(X_k, X_{k+1}, \dots)$, where f depends on an infinite number of variables (i.e., f depends on the future). However, this can be done in a more restricted framework; see Section 3.3.1.

More generally one can consider for a two-sided Markov process $(X_n)_{n\in\mathbb{Z}}$ the additive functionals depending both on the past and on the future. In this case, one has no results even for i.i.d. sequences.

2.3.3. *Some examples.*

1. *Ornstein–Uhlenbeck process.* For an Ornstein–Uhlenbeck process (even infinite dimensional), its semigroup of transition is hypercontractive and then all assumptions in this section are satisfied. An interesting example is the following. Let (X_t) be the canonical Ornstein–Uhlenbeck process in $\mathbb{R}^d$, that is, it satisfies

$$dX_t = dB_t + \tfrac{1}{2}X_t\,dt.$$

Its unique invariant measure is $m = N(0,1)$. Take $f(x) = |x|^2$ and consider the local time $L_t^1(f)$ of $(f(X_t))$ at 1. It is easy to see that all the conditions given in (2.22) are verified. Moreover, since $\Gamma(f,f)(x) = 1$ on $\{f = 1\} = S(0,1)$ (the unit sphere), $L_t^1(f)$ is also the local time of (X_t) at $S(0,1)$. Thus we get the CLT and MD for this local time with

$$V(L_t^1(f)) = \int\int_{S(0,1)^2} G(x,y)\,dx\,dy, \tag{2.23}$$

where $G(x,y)$ is given by $\int_0^\infty (P_t - m)\,dt(x,dy) = G(x,y)\,dy$ (i.e., G is the Green function) and dx in (2.23) is the usual Riemann measure on $S(0,1)$.

2. *Diffusions on compact Riemann manifolds.* Let E be a C^∞ compact connected Riemann manifold and dx its Riemann measure. Let $\mathscr{L}$ be a C^∞ second order differential operator which is the generator of a diffusion on E. Denote by (P_t) the corresponding semigroup of transition. Assume $\mathscr{L}$ is hyperelliptic. Thus by Hörmander's well known theorem, $P_t(x,dy) \ll m$ and its density $p_t(x,y)$ is C^∞ and strictly positive. Then it is hyperbounded by Theorem 4.12 in Wu (1991b).

Let f be a C^∞ real function on E such that $\Gamma(f,f) > 0$ on $S = \{f = 0\}$, where Γ is the associated square field operator. The local time $L_t(S)$ of the diffusion (X_t) can be written by

$$dL_t(S) = \Gamma(f,f)^{-1}(X_2)\,dL_t^0(f).$$

It is clear that all the assumptions in Corollary 2.7(c) are satisfied. Thus we get CLT and MD for $L_t(S)$.

3. *Systems of infinite interacting particles.* In the book by Liggett (1985), there are several examples of systems of infinite interacting particles which have an exponential convergence in L^2 to their equilibrium measures, such as the stochastic Ising model (due to Holley). Liggett (1985), Chapter 1, Theorem 4.1] also gives a very general condition which ensures the exponential convergence, but only for a class of functions (not for all $f \in L^2$ as we require in Theorem 2.1). Nevertheless if the systems is reversible, we have the following proposition.

PROPOSITION 2.9. *Let m be a reversible probability measure of (P_t). Assume that there is a dense subset $\mathscr{D}$ of $L^2(m)$ such that $\exists\,\gamma > 0$, $\forall\, f \in \mathscr{D}$, $\exists$ a constant, $C(f)$: $\|P_t f - m(f)\|_{L^2(m)} \le C(f)\cdot e^{-\gamma t}$. Then 1 is an isolated and simple eigenvalue of P_1 and the distance of 1 from the rest of the spectrum is larger than $e^{-\gamma}$.*

PROOF. Let $\mathscr{L}$ be the generator of (P_t) on $L^2(m)$. It is self-adjoint and negative definite. So it admits a spectral decomposition $\mathscr{L} = -\int_0^{+\infty}\lambda\,dE_\lambda$. We show now $E_\lambda f = E_0 f = m(f)$ for $\lambda \in (0,\gamma)$. In fact for any such λ fixed, we have $\forall\, f \in \mathscr{D}$,

$$\|P_t(E_\lambda f - m(f))\|_2 \ge e^{-\lambda t}\cdot\|E_\lambda f - m(f)\|_2,$$

which implies $\|E_\lambda f - m(f)\|_2 = 0$ for $f \in \mathscr{D}$ and then $E_\lambda f = m(f)$ by our assumption. □

REMARKS. (i) By Proposition 2.9 all reversible systems verifying the conditions of the Theorem 4.1 of Liggett [(1985), Chapter 1] satisfy the condition of our Theorem 2.1(c). In particular, our results in this section are applicable to the systems associated to Gibbs fields at sufficiently high temperature.

(ii) For nonreversible systems, we do not know whether Proposition 2.12 still holds. Nevertheless we can do something in the spirit of analytic perturbation. To be more precise, let us use the framework of Liggett (1985). The space of configurations $E = W^T$ (W is compact and T is countable). For $f \in C(E)$, define

$$\Delta_f(i) = \sup\{|f(\eta) - f(\xi)|; \eta(j) = \xi(j) \text{ for all } j \neq i \text{ in } T\},$$

$$\mathbb{D}(E) = \left\{ f \in C(E); \|f\| := \sum_{i \in T} \Delta_f^{(i)} < +\infty \right\}.$$

The generator Ω of a system of infinite interacting particles is given in Proposition 3.2 in Liggett [(1985), Chapter 1]. We have the following proposition.

PROPOSITION 2.10. *Under the condition of Theorem* 4.1 *in Liggett* (1985), *let* m *be the unique invariant measure. Then for* $f \in \mathbb{D}(E)$, $\{L_t(f) - m(f), t \to \infty\}$ *satisfies the CLT and the MD as in Theorem* 2.1.

In fact, by Lemma 3.7 in Liggett (1985), we can show easily that the family of closed operators $\{\Omega + zf;\ z \in \mathbb{C}\}$ acting on $(\mathbb{D}(E), \|\cdot\|)$ is analytic. By Theorem 3.9 in Liggett (1985), the semigroup (P_t) (associated to Ω), acting on $\mathbb{D}(E)$ has a gap near 1 in its spectrum and 1 is a simple eigenvalue. Thus we can apply the analytic perturbation theory as in the proof of Theorem 2.1.

3. Applications to statistical mechanical systems. In this section, we begin by recalling some notions of statistical mechanics on $\mathbb{Z}^d$, such as interaction potentials, pressure functionals, Gibbs states (or measures) and so forth. We shall see that the Cramér functional Λ introduced in Section 1 is just the pressure functional and then the C^2-regularity becomes a property of thermodynamical limit procedure. We shall explain why part (a) of Theorem 1.6. becomes the Lee–Yang theorem and indicate how to apply Theorem 1.6(b) to monotone functionals of FKG systems. We shall realize in detail for the ferromagnetic model, based on the previous works of Newman (1980) and Ellis (1985). Finally we shall give some discussions on several other models to which the results of Section 1 are applicable.

3.1. *Notations and ideas.* Let E (state space) be a Polish space, $\alpha(dx)$ be a probability measure on E and $\Omega = E^{\mathbb{Z}^d}$ be the space of configurations. A statistical mechanical system with configurations in Ω is determined by its potentials of interaction, denoted by $\Phi = (\phi_V)_{V_{\text{finite in } Z^d}}$. Φ is always assumed to satisfy:

1. ϕ_V is $\mathscr{F}_V$-measurable, where $\mathscr{F}_V := \sigma\{\omega \to \omega_V\}$.
2. $\phi_V \cdot \theta_i = \phi_{V+i}$, where θ_i is the shift on Ω, that is, $\theta_i \omega(j) = \omega(i+j)$.
3. $\sum_{0 \in V} \|\phi_V\|_\infty < \infty$.

For $V \subset \mathbb{Z}^d$ finite, the Hamiltonian on V with boundary condition η_{V^c} is given by

$$H_V(\omega_V | \eta_{V^c}) := \sum_{w:\, w \cap v \neq \varnothing} \phi_w(\omega_v, \eta_{v^c})$$

and the formal sum $H := \sum_V \phi_V$ is called Hamiltonian. A probability measure μ on Ω is called a *Gibbs state* (*or measure*) *associated to* Φ if its conditional distribution on E^V subject to the boundary condition η_{V^c}, $\mu(d\omega_V | \mathscr{F}_{V^c})(\eta_{V^c})$, is given by

$$\pi_V(\omega_V | \eta_{V^c}) := \frac{\exp(-H_V(\omega_V | \eta_{V^c})) \prod_{i \in V} \alpha(d\omega_i)}{Z_V(\eta_{V^c})}, \tag{3.1}$$

where $Z_{V^c}(\eta_{V^c})$, the normalization constant, is called a *partition functional in* V. The set of all Gibbs states of Φ will be denoted by $G(\Phi)$.

Let $F \in C_b(\Omega)$. It is known that the limit of

$$\frac{1}{|V|} \log \int_{E^V} \exp\left[\sum_{i \in V} F \cdot \theta_i\right] \pi_V(d\omega_V | \eta_{V^c}) := P_V(\Phi, F; \eta_{V^c})$$

as V increases to $\mathbb{Z}^d$ in Von Hove's sense [see Ruelle (1969)], exists uniformly on η (in this section, when we say $V \to \mathbb{Z}^d$, it is always taken in the Von Hove sense). This limit is the so-called *pressure functional*, which is denoted by $P(\Phi, F)$. Consequently for $\mu \in G(\phi)$, we have

$$\begin{aligned} P(\Phi, F) &= \lim_{V \to \mathbb{Z}^d} P_V(\Phi, F; \mu) \\ &:= \lim_{V \to \mathbb{Z}^d} \frac{1}{|V|} \log \int_{E^V} \exp\left[\sum_{i \in V} F \circ \theta_i\right] \pi_V(d\omega_V | \eta_{V^c}). \end{aligned} \tag{3.2}$$

Let $L_V := (1/|V|) \sum_{i \in V} \delta_{\theta_i \omega}$ be the empirical field. The LD of $\mu(L_V \in \cdot)$ on $M_1(\Omega)$ as $V = V_n = \{i \in \mathbb{Z}^d;\ |i| = \max(i_1, \dots, i_d) \le n\} \to \mathbb{Z}^d$, has been obtained by Comets (1986), Föllmer and Orey (1988) and Olla (1988). Here we shall discuss the corresponding CLT and MD estimations.

Considering a (formal) correspondence $|V|^{-1} \to \varepsilon$ and taking μ_ε as the law of $L_V(F)$ under μ for such ε, we have the following identifications:

$$\Lambda_\varepsilon(t) = P_V(\Phi, tF; \mu) \quad \text{and} \quad \Lambda(t) = P(\Phi, tF). \tag{3.3}$$

Having these identifications, we see that Theorem 1.4(a) becomes the Lee–Yang theorem (well known in statistical mechanics). For Theorem 1.6(b), notice that the convexity of $t \to P_V(\Phi, F; \mu)$ [or of $-P_V(\Phi, F; \mu)$] is implied by

$$\frac{d^3}{dt^3} P_V(\Phi, tF; \mu) \ge 0 \quad (\text{or } \le 0),$$

which is the so-called GHS inequality for monotone functions of attractive systems (i.e., FKG systems). Then to apply Theorem 1.6(b) to this class of models, the main question that remains to be checked is

$$\lim_{V\to\mathbb{Z}^d} P_V''(\Phi, tF;\mu)|_{t=0} = P_{0+}''(\Phi,\cdot F) < +\infty.$$

Finally notice that in the present case $\Lambda''(0)$ (if it exists) is the *susceptibility of the observable F of the system.*

3.2. *The ferromagnetic model.* Let $\Omega = \{+1, -1\}^{\mathbb{Z}^d}$ and define the formal Hamiltonian by

$$H^h(\omega) := \Sigma - J(i-j)\,\omega_i\omega_j + \Sigma h\,\omega_i,$$

where $J\colon \mathbb{Z}^d \to \mathbb{R}^+$ verifies $\Sigma_{i\in\mathbb{Z}^d} J(i) < +\infty$. This is the so-called ferromagnetic model. We denote by $G(\beta, h)$ the set of Gibbs states associated to βH, where $\beta = 1/kT$ is the inverse temperature.

Considering the spin variable $F(\omega) = \omega_0$, we have $\Sigma_{i\in V} F\circ\theta_i = \Sigma_{i\in V}\omega_i := S_V$. In this paragraph we shall discuss the limit behavior of S_V as V increases to $\mathbb{Z}^d$ in Von Hove's sense. We begin by recalling the well known structure of $G(\beta, h)$.

Let $\mu_{\beta,h}^+$ (resp. $\mu_{\beta,h}^-$) be the high (resp. low) Gibbs state. There is a critical inverse temperature β_c such that

$G(\beta, h)$ is a singleton for $(\beta,h)\in\mathscr{U} := \{(\beta,h); h\neq 0\}\cup\{(\beta,0);\beta<\beta_c\}$,

$$\mu_{\beta,h}^+ \neq \mu_{\beta,h}^- \quad \text{for } \beta>\beta_c \text{ and } h=0.$$

THEOREM 3.1. *Set $m(\beta,h,+) := \int\omega_0\,d\mu_{\beta,h}^+ := \langle\omega_0\rangle_{\beta,h,+}$. The limit*

$$\sigma^2(\beta,h,+) := \lim_{V\to\mathbb{Z}^d}\frac{1}{|V|}\langle[S_V - |V|\cdot m(\beta,h,+)]^2\rangle \tag{3.4}$$

exists always as an extended real number. Under the assumption that $\sigma^2(\beta,h,+)<\infty$, we have

$$\frac{1}{\sqrt{|V|}}[S_V - |V|\cdot m(\beta,h,+)] \xrightarrow{\mathscr{D}} N(0,\sigma^2(\beta,h,+)) \quad \text{as } V\to\mathbb{Z}^d. \tag{3.5}$$

If moreover $(\beta,h)\in\mathscr{U}$, then as $V\to\mathbb{Z}^d$, the family of measures

$$\mu_{\beta,h}\left(\frac{1}{\lambda(|V|)\sqrt{|V|}}[S_V - |V|\cdot m(\beta,h,+)]\in\cdot\right)$$

satisfies LDP *with speed $\lambda^{-2}(|V|)$ and with rate function $I(t) = t^2/(2\,\sigma^2(\beta, h,+))$.*

PROOF. We begin with the case $h\neq 0$. In this situation the Lee–Yang theorem holds [see Glimm and Jaffe (1981), Theorem 4.6.2, page 68] and it affirms that $P_V(\beta H^h, zF)$ tends to $P(\beta H^h, zF)$ uniformly for $z\in B(0,\delta)$ in $\mathbb{C}$, as V increases to $\mathbb{Z}^d$ in Von Hove's sense. Then we have as $V\to\mathbb{Z}^d$,

$$\frac{1}{|V|}\langle[S_V - |V|\cdot m(\beta,h,+)]^2\rangle = \frac{d^2}{dt^2}P_V(\beta H^h, tF)\bigg|_{t=0} \to \frac{d^2}{dt^2}P(\beta H^h, tF)\bigg|_{t=0}.$$

So (3.4) is valid and the CLT and MD follows from Theorem 1.6(a).

We treat now the situation where $h = 0$. Recall first some known results.

LEMMA 3.2. *Set*

$$\Lambda_V^{+,\beta}(t) = \frac{1}{|V|}\log\int\exp\left[\sum_{i\in V} tF(\omega_i)\right] d\mu_\beta^+.$$

Then we have:

(i) $[\Lambda_V^{+,\beta}(t)]'_t$ *is concave on* $t \in [0, +\infty)$.

(ii) (3.4) *is valid and* $\sigma^2(\beta, 0, +) = \sigma^2(\beta, +) = \sum_{i\in\mathbb{Z}^d}\langle\omega_0, \omega_i\rangle_{\beta,+}$.

(iii) *Note* $\Lambda^\beta(t) = \lim \Lambda_V^{+,\beta}(t)$. *Then* $[\Lambda^\beta(t)]''_{t=0+} = \sigma^2(\beta, +)$.

PROOF. Part (i) is a direct consequence of GHS inequality [see Ellis (1985), pages 167–168]. Part (ii) is deduced from FKG inequality [it is valid for any monotone function of FKG systems; see Newman (1980)].

(iii) Ellis [(1985), Lemma V.7.4] has shown it for $(\beta, h) \in \mathscr{U}$. Now we show it in the general case.

For $t > 0$, by the Lee–Yang theorem, we have $\sigma^2(\beta, t, +) = \sigma^2(\beta, t) = [\Lambda^\beta(t)]''$. So for proving (iii), it is enough to show that as t decreases to 0,

$$\sigma^2(\beta, t) = \sum_{i\in\mathbb{Z}^d}\langle\omega_0, \omega_i\rangle_{\beta,t} \to \sum_{i\in\mathbb{Z}^d}\langle\omega_0, \omega_i\rangle_{\beta,+}$$

by the result in (ii). To show this, recall that as t decreases to 0,

$$\langle\omega_0, \omega_k\rangle_{\beta,t} \text{ decreases to } \langle\omega_0, \omega_k\rangle_{\beta,+}.$$

Thus the desired result follows from Fatou's lemma. □

By this lemma, under the assumption $\sigma^2(\beta, +) < +\infty$, the family of measures $\{\mu_\beta^+((1/|V|)S_V \in \cdot), V \subset \mathbb{Z}^d\}$ will be right C^2-regular by Theorem 1.6 as $V \to \mathbb{Z}^d$ in Von Hove's sense. So the CLT (3.5) follows. For the MD, notice that $\mu^{+,\beta} = \mu^{-,\beta} = \mu^\beta$ for $\beta < \beta_c$. Then by symmetry, $\{\mu_\beta((1/|V|)S_V \in \cdot), V \subset \mathbb{Z}^d\}$ will be also left C^2-regular. Thus Theorem 3.1 is a direct consequence of the results in Section 1. □

REMARKS. (i) The CLT has been obtained by Newman (1980). The MD in Theorem 3.1 is new, but its proof, as the reader should see clearly, lies still on the results and the techniques already developed by Newman, Martin-Löf and Ellis, among others.

(ii) If $h \neq 0$, then $\sigma^2(\beta, h) < +\infty$ by Lee–Yang theorem. In the case where $h = 0$, whether $\sigma^2(\beta, 0, +) < \infty$ for all $\beta \neq \beta_c$ is a very interesting open question.

Assuming β_c as a normal critical point, Simon proved that for any $\beta < \beta_c$, $\langle\omega_0, \omega_k\rangle$ has an exponential decay as $|k| \to \infty$ and then $\sigma^2(\beta, +) < \infty$.

Even in the situation of phase transition, Martin-Löf (1973) showed for β sufficiently large (i.e., at low temperature), $(\Lambda^{\beta}(t))^{(n)}_{t=0+}$ exists for all $n \in \mathbb{N}$. In this way, we rediscover the CLT in Martin-Löf (1973).

Notice however that for the *Ising model* in dimension 2, the question above has a positive response by the well known Onsager explicit formula.

Of course the most striking situation is at the critical point $\beta = \beta_c$. In this critical case, only very few results are known [see the discussions of Ellis (1985), Chapter 5]. It is widely believed that the CLT would be broken. With the same analysis as in Chapter 5 of Ellis (1985), the MD estimation in Theorem 3.1 should change its form (the speed and rate function).

By the method of explicit calculus (but difficult!), Ellis and Newman (1978a) found the weak convergence theorem of the Curie–Weiss model at $\beta = \beta_c$. Their results have been extended in many respects by many authors, but always restricted to the class of mean field models.

In our previous work [Wu (1991a)], the LDP of μ_β^+ as $\beta \to +\infty$ is established.

3.3. *Other statistical mechanical systems.*

3.3.1. *One-dimensional situation.* For one-dimensional statistical mechanical systems, the pressure functionals are usually very regular. We present now some typical situations.

Assume

$$\exists r > 1 \text{ such that } \sum_{0 \in V} r^{|V|} \cdot \|\phi_V\|_\infty < +\infty. \tag{3.6}$$

Then for $F \in C_b(E)$, which depends only on a finite number of variables, it is shown in Ruelle (1978) that the Lee–Yang theorem holds, that is, $P_V(\phi, zF)$ tends to $P(\Phi, zF)$ uniformly for z in a neighborhood of 0 in $\mathbb{C}$. We have then CLT and MD of $\sum_{i=-n}^{n}(F \cdot \theta_i - \langle F \rangle_\mu)$ under μ, where μ is the unique Gibbs measure corresponding to Φ.

In the Ruelle (1978) book, he presented many dynamical systems which are isomorphic to the models above, such as expanding mappings on a compact metrisable space, Anosov diffeomorphisms or, more generally, Smale mapping. However, there is one thing worth notice. The CLT and the MD above are valid only for the functions F depending on a finite number of variables, and we do not know which functions of the dynamic system could be transformed as such functions under the isomorphism above. Fortunately the Lee–Yang theorem is known to hold for a large class of functions of dynamical systems, such as the Hölder-continuous functions of the Smale mappings; see Ruelle (1978).

3.3.2. *High temperature case.* In the framework of Section 3.1, let $F \in C_b(E)$, which is $\mathscr{F}_V$-measurable for some finite subset V in $\mathbb{Z}^d$. Then it is well known that the Lee–Yang theorem holds for β sufficiently large, that is, $P_V(\beta\Phi, zF) \to P(\beta\Phi, zF)$ uniformly for z in a neighborhood of 0 in $\mathbb{C}$, and $G(\beta\Phi) = \{\mu\}$. Thus by Theorem 1.6(a), we have CLT and MD for the limit behavior of $\sum_{i \in V}[F \circ \theta_i - \langle F \rangle_\mu]$, as V increases to $\mathbb{Z}^d$.

3.3.3. *Discrete Euclidean field model.* Let $\Omega = \mathbb{R}^{\mathbb{Z}^d}$ and $\alpha(dx) := Z^{-1}e^{-P(x)}\,dx$, where $P(x) = \lambda x^4 - \mu x^2$ with $\lambda > 0$, $\mu \in \mathbb{R}$, Z is the normalization constant. The formal Hamiltonian of the discrete Euclidean quantum field model is given by

$$H^h(\omega) := \sum_{|i-j|=1} -\omega_i \omega_j.$$

Consider $S_V := \sum_{i \in V} \omega_i$. Since the FKG and GHS inequalities hold and the Lee–Yang theorem holds under the same condition as for the ferromagnetic model, we have a similar result as that of Theorem 3.1.

3.3.4. *Continuous statistical mechanical systems.* In the continuous statistical model (i.e., on $\mathbb{R}^d$), the space of configurations will be $\Omega := \{\sum_i \delta_{x_i}$ (denumerable sum); $x_i \in \mathbb{R}^d\}$. The readers can consult Ruelle's (1978) book for the language below.

Let μ^β be a Gibbs measure associated to a potential $\Phi = \{\phi;\ V \subset \mathbb{R}^d$ bounded$\}$ and to the inverse temperature β. Let N_V be the number of particles contained in V [i.e., $N_V(\omega) = \omega(V)$]. The LD of N_V can be obtained under the assumption that the pressure functional is differentiable (in fact, this is a direct consequence of Proposition 1.1).

For the superstable and temporal potentials, it is shown that the Lee–Yang theorem holds at sufficiently high temperature. Therefore by Theorem 1.6, the CLT and MD hold in this case. This phenomena occurs also in the hard rod models.

3.3.5. *Other models.* The Lee–Yang theorem is a general phenomenon at high temperature. However, the FKG and GHS inequalities are the special properties of systems. It is known that these two inequalities hold for many other models besides those discussed above. For example, it is valid for *XY* model with two or three components, even for $:\phi^4:$ Euclidean quantum field model and so forth; see Glimm and Jaffe [(1981), Chapters 5, 18 and 20] for such discussions and also Ellis and Newman (1978b).

Acknowledgments. I am grateful to Dr. Gao for useful conversations, and to Professor J. Jacod for his kind invitation to Paris VI, where this work were accomplished.

REFERENCES

Baldi, P. (1988). Large deviations and stochastic homogenisation *Ann. Mat. Pura Appl.* **151** 161–177.

Borovkov, A. A. and Mogulskii, A. A. (1978). Probabilities of large deviations in topological vector space I. *Siberian Math. J.* **19** 697–709.

Borovkov, A. A. and Mogulskii. A. A. (1980). Probabilities of large deviations in topological vector space II. *Siberian Math. J.* **21** 12–26.

Chen, X. (1990). On the lower bounds of i.i.d. random variables in a Banach space. *Chinese J. Appl. Probab. Statist.* To appear.

Chen, X. (1991). The moderate deviations of independent random vectors in a Banach space. *Chinese J. Appl. Probab. Statist.* **7** 24–32.

Comets, F. (1986). Grandes déviations pour champs de Gibbs sur $\mathbb{Z}^d$. *C. R. Acad. Sci. Paris Sér. I Math.* **303** 511.

DE ACOSTA, A. (1990). Moderate deviations and associated Laplace approximations for sums of independent random variables. *Trans. Amer. Math. Soc.* To appear.

DAWSON, D. A. and GAERTNER, J. (1987). Long time fluctuation of weakly interacting diffusions. *Stochastics* **20** 247–308.

ELLIS, R. S. (1985). *Entropy, Large Deviations and Statistical Mechanics*. Springer, Berlin.

ELLIS, R. S. and NEWMAN, C. M. (1978a). Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **44** 117–139.

ELLIS, R. S. and NEWMAN, C. M. (1978b). Necessary and sufficient conditions for the GHS inequality with applications to probability and analysis. *Trans. Amer. Math. Soc.* **237** 83–99.

FÖLLMER, H. and OREY, S. (1988). Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16** 961–977.

GAO, F. Q. (1991). Moderate deviations of Markov processes. Preprint.

GAO, F. Q. (1993). Characterization of moderate deviations for Markov process. In *Advances in Mathematics*. Pekin University.

GLIMM, J. and JAFFE, A. (1981). *Quantum Physics: A Functional Integral Point of View*. Springer, Berlin.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.

IAGOLNIZER, D. and SOUILLARD, B. (1979). Lee–Yang theory and normal fluctuations. *Phys. Rev. B.* **19** 1515–1518.

JACOD, J. and SHIRYAYEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, Berlin.

KATO, T. (1984). *Perturbation Theory for Linear Operators*, 2nd ed. (2nd corrected printing). Springer, Berlin.

LEDOUX, M. (1992). Sur les deviations modérées des sommes de variables aléatoires vectorielles independantes de même loi. *Ann. Inst. H. Poincaré* **28** 267–280.

LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.

MARTIN-LÖF, A. (1973). Mixing properties, differentiability of the free energy, and the central limit theorem for a pure phase in the Ising model at low temperature. *Comm. Math. Phys.* **32** 75–92.

NEWMAN, C. M. (1980). Normal fluctuations and FKG inequalities. *Comm. Math. Phys.* **74** 119–128.

OLLA, S. (1988). Large deviations for Gibbs random fields *Probab. Theory Related Fields* **77** 343–357.

REVUZ, D. (1976). *Markov Chains*. North-Holland, Amsterdam.

RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*. Benjamin, New York.

RUELLE, D. (1978). *Thermodynamic Formalism*. Addison-Wesley, Reading, MA.

STEIN, E. M. and WEISS, G. (1975). *Introduction to Fourier Analysis on Euclidean Space*. Princeton Univ. Press.

WU, L. M. (1991–1993). Grandes déviations pour les processus essentiellement irréductibles I. En temps discret. *C.R. Acad. Sci. Paris Ser. 1 Math.* **312** 609–614 (1991); II. En temps continu. **314** 941–946 (1992); III. Applications. **316** 853–858 (1993).

WU, L. M. (1991a). Grandes déviations pour les mesures de Gibbs lorsque la température tend vers zéro. *Ann. Inst. H. Poincaré* **27** 273–290.

WU, L. M. (1991b). Grandes déviations pour processus de Markov I: en temps discret. Preprint included in Habilitation à diriger des recherches, Université Paris VI.

WU, L. M. (1991c). Some general methods of large deviations and applications Preprint.

WU, L. M. (1994). Perturbations of Dirichlet forms, ground state diffusions and large deviations *J. Funct. Anal.* **119**. To appear.

DÉPARTEMENT DE MATHÉMATIQUES APPLIQUÉES
UNIVERSITÉ CLERMONT-FERRAND II
63177 AUBIERE
FRANCE