Sankhyā : The Indian Journal of Statistics
2010, Volume 72-A, Part 1, pp. 191-207

Gibbs Measures Asymptotics

K. B. Athreya
Iowa State University, Ames, USA
Chii-Ruey Hwang
Academia Sinica, Taipei, Taiwan

Abstract

Let $(\Omega, \mathcal{B}, \nu)$ be a measure space and $H : \Omega \to \mathbb{R}^+$ be $\mathcal{B}$ measurable. Let $\int_\Omega e^{-H} d\nu < \infty$. For $0 < T < 1$ let $\mu_{H,T}(\cdot)$ be the probability measure defined by

$$\mu_{H,T}(A) \equiv \left(\int_A e^{-H/T} d\nu \right) / \left(\int_\Omega e^{-H/T} d\nu \right), \quad A \in \mathcal{B}.$$

In this paper, we study the behavior of $\mu_{H,T}(\cdot)$ as $T \downarrow 0$ and extend the results of Hwang (1980, 1981). When Ω is $\mathbb{R}$ and H achieves its minimum at a single value x_0 (single well case) and $H(\cdot)$ is Hölder continuous at x_0 of order α, it is shown that if X_T is a random variable with probability distribution $\mu_{H,T}(\cdot)$ then as $T \downarrow 0$, *i*) $X_T \to x_0$ in probability; *ii*) $(X_t - x_0)T^{-1/\alpha}$ converges in distribution to an absolutely continuous symmetric distribution with density proportional to $e^{-c_\alpha |x|^\alpha}$ for some $0 < c_\alpha < \infty$. This is extended to the case when H achieves its minimum at a finite number of points (multiple well case). An extension of these results to the case $H : \mathbb{R}^n \to \mathbb{R}^+$ is also outlined.

AMS (2000) *subject classification.* Primary 60B10, 60F99.
Keywords and phrases. Entropy maximization, Gibbs measure, Hamiltonian, Hölder continuous, Laplace's method, simulated annealing, temperature, weak convergence

1 Introduction

Let $(\Omega, \mathcal{B}, \nu)$ be a measure space, $H : \Omega \to \mathbb{R}^+$ be $\mathcal{B}$ measurable and $\int_\Omega e^{-H} d\nu < \infty$. For $0 < T < 1$, let $\mu_{H,T}$ be the probability measure defined by

$$\mu_{H,T}(A) \equiv \left(\int_A e^{-H/T} d\nu \right) / \left(\int_\Omega e^{-H/T} d\nu \right), \quad A \in \mathcal{B}.$$

If $Q(A) \triangleq \mu_{H,1}(A)$, $A \in \mathcal{B}$, then Q is a probability measure and for $0 < T < 1$

$$\mu_{H,T}(A) \equiv \left(\int_A e^{-H/\theta} dQ \right) / \left(\int_\Omega e^{-H/\theta} dQ \right)$$

for $A \in \mathcal{B}$ and $\theta = \frac{T}{T-1}$. As $T \downarrow 0$, $\theta \downarrow 0$ and conversely. Thus without loss of generality we may restrict to the case when $\nu(\cdot)$ is a probability measure. We shall do so in the sequel when needed.

The probability measure $\mu_{H,T}(\cdot)$ is called a Gibbs measure with Hamiltonian H and temperature T. In this paper we extend some of the results of Hwang (1980, 1981) on the behavior of $\mu_{H,T}$ as $T \downarrow 0$. It is intuitively clear that as $T \downarrow 0$, $\mu_{H,T}$ puts more mass on those sets A when e^{-H} is large or equivalently when H is small. Thus it is not surprising that the set $N \triangleq \{\omega : H(\omega) = \text{ess inf } H \text{ w.r.t. } \nu\}$ plays an important role. Here ess inf H w.r.t. ν is a number λ such that for any $\varepsilon > 0$, $\nu\{\omega : H(\omega) < \lambda - \varepsilon\} = 0$ and $\nu\{\omega : H(\omega) < \lambda + \varepsilon\} > 0$. Hwang (1980) considered the cases $i)$ $\nu(N) > 0$, $ii)$ $\nu(N) = 0$ and N a singleton $\{x_0\}$ and $iii)$ $\nu(N) = 0$ and N is a finite set $\{x_1, x_2, \ldots, x_k\}$, $1 < k < \infty$. He showed, when $\Omega = \mathbb{R}^n$ and under some regularity conditions, that as $T \downarrow 0$, in case $i)$

$$\mu_{H,T}(A) \to \frac{\nu(A \cap N)}{\nu(N)} \text{for } A \in \mathcal{B},$$

in case $ii)$ $\mu_{H,T}$ converges weakly to the delta measure at x_0 and in case $iii)$ assuming ν is absolutely continuous with density $f > 0$ on N and H twice differentiable, $\mu_{H,T}$ converges weakly to a probability measure on N with weights proportional to $f(x_i)c_i$ at x_i where c_i^{-2} is the Hessian of H at x_i.

In this paper we study the second order behavior of this convergence. That is, if X_T is a random variable with distribution $\mu_{H,T}$ we study the rate of approach of X_T to its limit as $T \downarrow 0$. We show that in case $ii)$ if $\Omega = \mathbb{R}$ and
if H is Hölder continuous at x_0 of order α then $(X_T - x_0)T^{-1/\alpha}$ converges in distribution to an absolutely continuous continuous symmetric distribution with density proportional to $e^{-c_\alpha |x|^\alpha}$ for some $0 < c_\alpha < \infty$. We prove an appropriate extension of this case to case $iii)$, again assuming $\Omega = \mathbb{R}$. Extensions of these with case $\Omega = \mathbb{R}^n$, $n > 1$ are outlined at the end.

In the next section we give some examples of Gibbs measures from statistical physics, image processing, entropy maximization, perturbations of Hamiltonian systems with white noise. Section 3 is devoted to some preliminary results. Sections 4 and 5 treat the single well and multiple wells respectively. The last section discusses the case $\Omega = \mathbb{R}^n$.

2 Some examples of Gibbs measures

2.1 Statistical mechanics. Let $S = \{s_1, s_2, \ldots, s_n\}$, $n < \infty$ be a set of n "sites" and $A \equiv \{a_1, a_2, \ldots, a_k\}$, $k < \infty$ be a set of "alphabets" or "spin sizes". Let $\Omega \equiv \{\omega : \omega : S \to A\}$ be the set of all functions from S to A. Each ω in Ω will be referred to as a "configuration". Let $\mathcal{B} = \mathcal{P}(\Omega)$, the power set of Ω and Q be the uniform distributions on Ω. Let $H : \Omega \to \mathbb{R}^+$ and $0 < T < \infty$ be given. The Gibbs distribution $\mu_{H,T}$, in this case, is given by

$$\mu_{H,T}\{\omega\} = \frac{e^{-H(\omega)/T}}{\sum_{\omega' \in \Omega} e^{-H(\omega')/T}}, \quad \text{for all } \omega \in \Omega.$$

The denominator $p_{H,T} \triangleq \sum_{\omega' \in \Omega} e^{-H(\omega')/T}$ is known as the *partition function* with potential or Hamiltonian H and temperature T. In statistical mechanics S is taken to be a finite integer lattice in $\mathbb{R}^3$ of the form $S \triangleq \{(i_1, i_2, i_3) : i_j \text{ an integer } |i_j| \leq m,\ j = 1, 2, 3\}$ and $A = \{-1, +1\}$. Hence, $n = (2m+1)^3$ and $k = 2$. The total number of configurations, the size of Ω, is k^n and here it becomes $2^{(2m+1)^3}$. For $m = 1$, it is 2^{27} a large number. For $m = 2$, it jumps to 2^{125} a very large number indeed. Each configuration ω is of the form $\{\delta_{\underline{i}} : \underline{i} \in S\}$ where $\delta_{\underline{i}}$ is $+1$ or -1 means as "spin up" or "spin down". Typically the Hamiltonian $H(\omega)$ is of the form

$$H(\omega) = \sum_{|\underline{i} - \underline{j}| \leq 1} V(\delta_{\underline{i}}, \delta_{\underline{j}}),$$

where $V : \{-1, 1\}^2 \to \mathbb{R}$ and $|\underline{i} - \underline{j}| = \sum_{s=1}^{3} |i_s - j_s|$. Thus $H(\omega)$ depends on "nearest neighbor interaction". As the temperature T decreases to zero, $\mu_{H,T}(\cdot)$ can be shown to converge to the uniform distribution on the set N of configurations ω of minimal energy, i.e.

$$N = \{\omega : H(\omega) = \inf\{H(\omega') : \omega' \in \Omega\}\}.$$

To find these configurations as well as to obtain a sample from $\mu_{H,T}$ without computing the partition function and to estimate certain averages of the form $\lambda \triangleq \sum_{\omega \in \Omega} g(\omega)\mu_{H,T}(\omega)$, Metropolis et al. (1953) invented a computational technique using Markov chains. This important paper provided the inspiration for the currently popular simulation tool Markov chain Monte Carol (MCMC). See Robert and Casella (2004). The Metropolis-Hastings algorithm in MCMC is due to the adaptation of the Metropolis et al method by Hastings (1970). Similarly, the Gibbs sampler algorithm in MCMC is due to the work of Geman and Geman (1984) who introduced it in their work on image processing. We describe this next.

2.2 Image processing. Here the set S of sites is the matrix of "pixels" (i,j), $1 \le i \le M$, $1 \le j \le M$. At each pixel the color level $\omega(i,j)$ is one of k possible levels $A \equiv \{a_1, a_2, \dots, a_k\}$. A *picture* is a configuration ω, i.e. a map form $S \equiv \{(i,j) : 1 \le i \le M, 1 \le j \le M\}$ to $A \equiv \{a_1, a_2, \dots, a_k\}$. Again, Q is taken to be the uniform distribution on Ω, the set of possible configurations. For $M = 16$, $k = 2$, the size of Ω is 2^{256} a very large number. The Gibbs measure is of the form

$$\mu_{H,T}(\omega) = \frac{e^{-H(\omega)/T}}{\sum_{\omega' \in \Omega} e^{-H(\omega')/T}}.$$

Geman and Geman (1984) have studied this in detail for several special H's.

2.3 Entropy maximization. Let $(\Omega, \mathcal{B}, \nu)$ be a measure space and $h : \Omega \to \mathbb{R}$ be $\mathcal{B}$-measurable and $c \in \mathbb{R}$. Let

$$\mathcal{F}_{h,c} \equiv \left\{ g : g : \Omega \to \mathbb{R}^+ \quad \mathcal{B}\text{-measurable}, \int g d\nu = 1, \int h g d\nu = c \right\}.$$

For any $g : \Omega \to \mathbb{R}^+$, $\mathcal{B}$-measurable with $\int g d\nu = 1$, $E(g,\nu) \equiv -\int g \log g d\nu$ is called the relative entropy of g w.r.t. ν. Consider the problem of maximizing $E(g,\nu)$ w.r.t. g in $\mathcal{F}_{h,c}$ for given h and c. It is shown in Athreya (2009) that if there is a θ in $\mathbb{R}$ such that $\int_\Omega e^{\theta h} d\nu < \infty$, $\int_\Omega |h| e^{\theta h} d\nu < \infty$, and $\int_\Omega h e^{\theta h} d\nu = c \int_\Omega e^{\theta h} d\nu$, then $g_0 \equiv \frac{e^{\theta h}}{\int_\Omega e^{\theta h} d\nu}$ is the unique solution to the above problem, i.e. $g_0 \in \mathcal{F}_{h,c}$ and $E(g_0,\nu) \ge E(g,\nu)$ for all $g \in \mathcal{F}_{h,c}$. Note that

$$\mu(A) \equiv \int_A g_0 d\nu, \quad A \in \mathcal{B}$$

is a Gibbs measure of the form $\mu_{h,\frac{1}{\theta}}$, i.e. with $H = h$ and $T = \frac{1}{\theta}$.

2.4 Random perturbations of Hamiltonian systems. Let $\{X(t) : t \ge 0\}$ be the unique solution of the ordinary differential equation

$$\frac{dX(t)}{dt} = -u'(X(t)), \quad X(0) = x_0,$$

where $u : \mathbb{R} \to \mathbb{R}$ is a C^1 function. For each $t > 0$, consider an Itô process of the form

$$dX^\varepsilon(t) = -u'(X^\varepsilon(t))dt + \sqrt{\varepsilon} dW(t), \quad t \ge 0, \quad X^\varepsilon(0) = x_0,$$

where $\{W(t) : t \ge 0\}$ is a standard Brownian motion. Suppose u is such that

$$c = \int_{\mathbb{R}} e^{-\frac{2u(x)}{\varepsilon}} dx < \infty.$$

Then for all $\varepsilon > 0$, the Itô process $\{X^\varepsilon(t) : t \geq 0\}$ has a unique stationary distribution Π_ε with density

$$f_\varepsilon(x) = \frac{1}{c} e^{-\frac{2u(x)}{\varepsilon}}$$

w.r.t. Lebesgue measure on $\mathbb{R}$. It can be shown that $X(t)$ converges in distribution and also in variation norm to the above stationary distribution Π_ε as $t \to \infty$. Note that the measure Π_ε is a Gibbs measure on $\Omega = \mathbb{R}$ with $H(x) = 2u(x)$ and $T = \varepsilon$. The behavior of Π_ε as $\varepsilon \downarrow 0$ is of interest in the study of small perturbations of Hamiltonian system. See Freidlin and Wentzell (1994).

To find global minima of H, one may consider the following simulated annealing process

$$dX(t) = -u'(X(t))dt + \sqrt{T(t)}dW(t), \quad t \geq 0, \quad X(0) = x_0.$$

With proper annealing rates $T(t)$, $X(t)$ converges to the set of global minima N in probability or weakly to a probability on N. See Hwang and Sheu (1990).

3 Some preliminary results

Let $(\Omega, \mathcal{B}, Q)$ be a probability space. Let $H : \Omega \to \mathbb{R}^+$ be $\mathcal{B}$ measurable. For $0 < T < 1$, let

$$\mu_{H,T}(A) \equiv \left(\int_A e^{-H/T} dQ \right) / \left(\int_\Omega e^{-H/T} dQ \right), \quad A \in \mathcal{B}.$$

Let $\lambda \equiv \operatorname{ess\,inf} H$ w.r.t. Q. That is, for all $\varepsilon > 0$, $Q\{\omega : H(\omega) \leq \lambda - \varepsilon\} = 0$ and $Q\{\omega : H(\omega) \leq \lambda + \varepsilon\} > 0$.

PROPOSITION 3.1. *For all $\varepsilon > 0$,*

$$\mu_{H,T}\{\omega : H(\omega) > \lambda + \varepsilon\} \to 0 \quad \textit{as } T \downarrow 0.$$

PROOF. Fix $\varepsilon > 0$. Then

$$\mu_{H,T}\{\omega : H(\omega) > \lambda + \varepsilon\} = \frac{\int_{\{H > \lambda+\varepsilon\}} e^{-H/T} dQ}{\int_\Omega e^{-H/T} dQ} \leq \frac{\int_{\{H > \lambda+\varepsilon\}} e^{-H/T} dQ}{\int_{\{H \leq \lambda+\varepsilon\}} e^{-H/T} dQ}$$

$$\leq \frac{\int_{\{H > \lambda+\varepsilon\}} e^{-H/T} dQ}{e^{-(\lambda+\varepsilon)/T} Q\{H \leq \lambda + \varepsilon\}} = \frac{\int_{\{H > \lambda+\varepsilon\}} e^{-\frac{(H-(\lambda+\varepsilon))}{T}} dQ}{Q\{H \leq \lambda + \varepsilon\}}.$$

By the bounded convergence theorem, the numerator above goes to zero. By the definition of λ, the denominator is bounded away from zero. $\square$

THEOREM 3.1. *Let* $N \equiv \{\omega : H(\omega) = \lambda \equiv \text{ess inf}\, H \text{ w.r.t. } Q\}$. *Assume* $Q(N) > 0$. *Then for all* $A \in \mathcal{B}$,

$$\mu_{H,T}(A) \to \mu(A) \equiv \frac{Q(A \cap N)}{Q(N)} \text{ as } T \downarrow 0.$$

PROOF.

$$\mu_{H,T}(A) = \frac{\int_{A\cap N} e^{-H/T} dQ + \int_{A\cap N^c} e^{-H/T} dQ}{\int_N e^{-H/T} dQ + \int_{N^c} e^{-H/T} dQ}$$

$$= \frac{e^{-\lambda/T}\left(Q(A\cap N) + \int_{A\cap N^c} e^{-\frac{(H-\lambda)}{T}} dQ\right)}{e^{-\lambda/T}\left(Q(N) + \int_{N^c} e^{-\frac{(H-\lambda)}{T}} dQ\right)}.$$

By the bounded convergence theorem

$$\int_{N^c} e^{-\frac{(H-\lambda)}{T}} dQ \to 0 \quad \text{as } T \downarrow 0,$$

and the given assertion follows. $\square$

REMARK 3.1. By Vitali-Hahn-Saks theorem, Theorem 3.1 can be strengthened to assert that $\|\mu_{H,T} - \mu\|_{TV} \to 0$ as $T \downarrow 0$, where $\|\cdot\|$ denotes the total variation norm.

4 Single well case

THEOREM 4.1. *Let* $H : \mathbb{R} \to \mathbb{R}^+$ *be Borel measurable. Let* ν *be a measure on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *and let* $\int_{\mathbb{R}} e^{-H} d\nu < \infty$. *Let for* $0 < T \le 1$

$$\mu_{H,T}(A) \equiv \frac{\int_A e^{-H/T} d\nu}{\int_\Omega e^{-H/T} d\nu}, \quad A \in \mathcal{B}(\mathbb{R}).$$

Let $x_0 \in \mathbb{R}$ *be such that*

i) for all $\delta > 0$, *there exists* $\eta > 0$ *such that*

$$a(\delta) \equiv \inf\{H(x) : |x - x_0| \ge \delta\} > b(\eta) \equiv \sup\{H(x) : |x - x_0| \le \eta\}.$$

ii) for all $\eta > 0$, $\nu\{x : |x - x_0| \leq \eta\} > 0$.

Then

$$\mu_{H,T}\{x : |x - x_0| \geq \delta\} \to 0 \quad \text{as } T \downarrow 0.$$

PROOF. For $\delta > 0$, let $\eta > 0$ be such that $i)$ above holds. Then

$$\mu_{H,T}\{x : |x - x_0| \geq \delta\} \leq \frac{\int_{|x-x_0|\geq\delta} e^{-\frac{H(x)}{T}} d\nu}{\int_{|x-x_0|\leq\eta} e^{-\frac{H(x)}{T}} d\nu}$$

$$\leq \frac{\int_{|x-x_0|\geq\delta} e^{-\frac{(H(x)-b(\eta))}{T}} d\nu}{\int_{|x-x_0|\leq\eta} e^{-\frac{(H(x)-b(\eta))}{T}} d\nu} \leq \frac{\int_{|x-x_0|\geq\delta} e^{-\frac{(H(x)-b(\eta))}{T}} d\nu}{\nu\{x : |x - x_0| \leq \eta\}}.$$

On the set $\{x : |x - x_0| \geq \delta\}$, $H(x) - b(\eta) > 0$ and so $\frac{H(x)-b(\eta)}{T} \to \infty$ as $T \downarrow 0$. Also

$$\int_{|x-x_0|\geq\delta} e^{-(H(x)-b(\eta))} d\nu < \infty.$$

So by the dominated convergence theorem

$$\int_{|x-x_0|\geq\delta} e^{-\frac{(H(x)-b(\eta))}{T}} d\nu \to 0 \quad \text{as } T \downarrow 0.$$

Since $\nu\{x : |x - x_0| \leq \eta\} > 0$, the proof is complete. □

REMARK 4.1. It is clear that the same proof works if $\mathbb{R}$ is replaced by a Polish space.

REMARK 4.2. If $H(\cdot)$ is continuous at x_0, then hypothesis $i)$ of Theorem 4.1 can be replaced by the hypothesis that for all $\delta > 0$,

$$a(\delta) \equiv \inf\{H(x) : |x - x_0| \geq \delta\} > H(x_0).$$

Now let X_T be a random variable with probability distribution $\mu_{H,T}$. Theorem 4.1 is the same as saying that under the hypothesis of Theorem 4.1, $X_T \to x_0$ in probability as $T \downarrow 0$. This raises the question of how rapidly does X_T go to x_0 as $T \downarrow 0$, i.e., the rate of convergence. We address the question next when $\Omega = \mathbb{R}$ and $\nu(\cdot)$ is the Lebesgue measure.

THEOREM 4.2. *Let* $H : \mathbb{R} \to \mathbb{R}^+$ *be Borel measurable. Assume:*

i) there exists $x_0 \in \mathbb{R}$, $0 < \alpha, c < \infty$ *such that*

$$\lim_{x \to x_0} \frac{H(x) - H(x_0)}{|x - x_0|^\alpha} = c;$$

ii) for all $\delta > 0$, $a(\delta) \equiv \inf\{H(x) : |x - x_0| \geq \delta\} > H(x_0)$;

iii) $\int_{\mathbb{R}} e^{-H(x)} dx < \infty$.

Let

$$\mu_{H,T}(A) \equiv \frac{\int_A e^{-H/T} dx}{\int_{\mathbb{R}} e^{-H/T} dx}, \quad A \in \mathcal{B}(\mathbb{R}), \quad 0 < T \leq 1.$$

Let X_T *be a random variable with probability distribution* $\mu_{H,T}$*, i.e.* $Pr(X_T \in A) \equiv \mu_{H,T}(A)$, $A \in \mathcal{B}(\mathbb{R})$. *Thus as* $T \downarrow 0$,

$$\frac{(X_T - x_0)c^{1/\alpha}}{T^{1/\alpha}} \xrightarrow{d} X,$$

where X *is a random variable with density* $f_X(x) \equiv e^{-|x|^\alpha} / \int_{\mathbb{R}} e^{-|x|^\alpha} dx$.

Proof. For $T > 0$, $\delta > 0$, $-\infty < a < b < \infty$, let

$$m_1(T,\delta) \equiv \int_{|x-x_0| \leq \delta} e^{-\frac{H(x)}{T}} dx,$$
$$m_2(T,\delta) \equiv \int_{|x-x_0| > \delta} e^{-\frac{H(x)}{T}} dx,$$
$$m_3(T,\delta) \equiv \int_{\frac{aT^{1/\alpha}}{c^{1/\alpha}} \leq (x-x_0) \leq \frac{bT^{1/\alpha}}{c^{1/\alpha}}} e^{-\frac{H(x)}{T}} dx.$$

Let $\lambda \equiv H(x_0)$. We now claim the following

$$\lim_{T \downarrow 0} \frac{e^{\lambda/T} c^{1/\alpha}}{T^{1/\alpha}} m_i(T,\delta) \equiv \begin{cases} \int_{\mathbb{R}} e^{-|u|^\alpha} du, & i = 1, \\ 0, & i = 2, \\ \int_a^b e^{-|u|^\alpha} du, & i = 3. \end{cases}$$

Assuming the validity of this claim here is the proof of Theorem 4.2. For $-\infty < a < b < \infty$,

$$Pr\left(a \leq \frac{(X_T - x_0)}{T^{1/\alpha}} c^{1/\alpha} \leq b\right) = \frac{m_3(T,\delta)}{m_1(T,\delta) + m_2(T,\delta)}$$

$$= \frac{\left(\frac{e^{\lambda/T}c^{1/\alpha}}{T^{1/\alpha}}\right) m_3(T,\delta)}{\left(\frac{e^{\lambda/T}c^{1/\alpha}}{T^{1/\alpha}}\right)(m_1(T,\delta)+m_2(T,\delta))}.$$

By the claim, as $T \downarrow 0$, the above quantity converges to

$$\frac{\int_a^b e^{-|u|^\alpha} du}{\int_{\mathbb{R}} e^{-|u|^\alpha} du}$$

which is the stated assertion.

Now to the proof of the claim, by hypothesis $i)$ we can write

$$H(x) = \lambda + |x - x_0|^\alpha (1 + h(x)),$$

where $\lim_{x\to x_0} h(x) = 0$ and have, $\sup\{|h(x)| : |x - x_0| \le \delta\} < \infty$, for $\delta > 0$ small. Thus

$$m_1(T,\delta) = \frac{e^{-\lambda/T}T^{1/\alpha}}{c^{1/\alpha}} \int_{|u| \le \frac{\delta c^{1/\alpha}}{T^{1/\alpha}}} e^{-|u|^\alpha\left(1+h\left(x_0+\frac{uT^{1/\alpha}}{c^{1/\alpha}}\right)\right)} du.$$

By the dominated convergence theorem, for $\delta > 0$ small the integral above converges as $T \downarrow 0$ to

$$\int_{\mathbb{R}} e^{-|u|^\alpha} du$$

proving the claim for $i = 1$. Next

$$m_2(T,\delta) = e^{-\frac{a(\delta)}{T}} \int_{|x-x_0|\ge\delta} e^{-\frac{(H(x)-a(\delta))}{T}} dx$$

where $a(\delta)$ is as in hypothesis $ii)$. Then

$$\frac{e^{\lambda/T}c^{1/\alpha}}{T^{1/\alpha}} m_2(T,\delta) = \frac{e^{-\frac{(a(\delta)-\lambda)}{T}}c^{1/\alpha}}{T^{1/\alpha}} \int_{|x-x_0|\ge\delta} e^{-\frac{(H(x)-a(\delta))}{T}} dx.$$

Since $a(\delta) > H(x_0) = \lambda$ and $H(x) \ge a(\delta)$ for $|x - x_0| \ge \delta$ and for $0 < T < 1$,

$$\int_{|x-x_0|\ge\delta} e^{-\frac{(H(x)-a(\delta))}{T}} dx \le \int_{|x-x_0|\ge\delta} e^{-(H(x)-a(\delta))} dx$$
$$\le \left(\int_{\mathbb{R}} e^{-H(x)} dx\right) e^{a(\delta)} < \infty.$$

Also it follows that

$$\frac{e^{-\left(\frac{a(\delta)-\lambda}{\lambda}\right)}}{T^{1/\alpha}} c^{1/\alpha} \to 0 \quad \text{as } T \downarrow 0.$$

Thus the claim is proved for the case $i = 2$. Finally

$$\frac{e^{\frac{\lambda}{T}} c^{1/\alpha}}{T^{1/\alpha}} m_3(T,\delta) = \int_a^b e^{-|u|^\alpha \left(1+h\left(x_0 + \frac{uT^{1/\alpha}}{c^{1/\alpha}}\right)\right)} du.$$

Now arguing as for the first case of the claim, the above integral converges to $\int_a^b e^{-|u|^\alpha} du$ establishing the claim for case $i = 3$. $\square$

Corollary 4.1. *Suppose (i) and (ii) of Theorem 4.2. hold and H is twice differentiable at x_0 with $H''(x_0) = \sigma^2$, $0 < \sigma^2 < \infty$. Then as $T \downarrow 0$*

$$\frac{X_T - x_0}{T^{1/2}} \xrightarrow{d} N(0, 1/\sigma^2)$$

Remark 4.3. A simple change of variables yields

$$\int_{\mathbb{R}} e^{-|u|^\alpha} du = \frac{2}{\alpha}\Gamma\left(\frac{1}{\alpha}\right)$$

where $\Gamma(p) \equiv \int_0^\infty e^{-u} u^{p-1} du$, $0 < p < \infty$.

Some examples:

i) Let

$$H(x) = \begin{cases} -\log\cos x, & |x| \le \pi/2, \\ 0, & |x| > \pi/2. \end{cases}$$

Then $x_0 = 0$, $\lambda = 0$, $H''(0) = 1$. Let $0 < T < 1$. So if X_T is a random variable such that

$$Pr(X_T \in A) = \frac{\int_{A \cap [-\frac{\pi}{2}, \frac{\pi}{2}]} (\cos x)^{1/T} dx}{\int_{[-\frac{\pi}{2}, \frac{\pi}{2}]} (\cos x)^{1/T} dx}, \quad A \in \mathcal{B}(\mathbb{R}),$$

then as $T \downarrow 0$, $\frac{X_T}{\sqrt{T}} \xrightarrow{d} N(0,1)$.

ii) Let

$$H(x) = \begin{cases} -\log(\cos |x|^\beta), & |x| \le \pi/2, \\ 0, & |x| > \pi/2, \end{cases} \qquad 0 < \beta < \infty.$$

Then $x_0 = 0$, $\lambda = 0$, $\alpha = 2\beta$ and $c = 1/2$. So if X_T is a random variable such that

$$Pr(X_T \in A) = \frac{\int_{A\cap[-\frac{\pi}{2},\frac{\pi}{2}]}(\cos|x|^\beta)^{1/T}dx}{\int_{[-\frac{\pi}{2},\frac{\pi}{2}]}(\cos|x|^\beta)^{1/T}dx}, \quad A \in \mathcal{B}(\mathbb{R}),$$

then $\frac{X_T}{T^{1/\alpha}} \xrightarrow{d} X$ where X has pdf $f_X(x) = \frac{2}{\alpha}\frac{1}{\Gamma(\frac{1}{\alpha})}e^{-|x|^\alpha}$, $-\infty < u < \infty$ and $\alpha = 2\beta$.

5 Multiple well case

THEOREM 5.1. *Let H, $\mu_{H,T}$ be as in Theorem 4.2. Suppose there exists a finite set $N = \{x_1, x_2, \ldots, x_m\}$, $2 \le m < \infty$ such that*

i) $H(x_i) = \lambda$, $0 \le \lambda < \infty$, $1 \le i \le m$,

ii) for all $\delta > 0$, $a(\delta) = \inf\{H(x) : |x - x_i| \ge \delta, i = 1, 2, \ldots, m\} > \lambda$,

iii) there exists $0 < \alpha < \infty, 0 < c_i < \infty, i = 1, 2, \ldots, m$ *such that*

$$\lim_{x\to x_0} \frac{H(x) - H(x_i)}{|x - x_i|^\alpha} = c_i, \quad \text{for } i = 1, 2, \ldots, m.$$

(α does not change with x_i)

Let X_T be a random variable with distribution $\mu_{H,T}$. Then as $T \downarrow 0$ $X_T \xrightarrow{d} Y$, where Y is a random variable with distribution

$$P(Y = x_i) = \frac{\frac{1}{c_i^{1/\alpha}}}{\sum_{j=1}^m \frac{1}{c_j^{1/\alpha}}}, \quad 1 \le i \le m.$$

PROOF. For $0 < T < 1$, $\delta > 0$, let

$$m_i(T, \delta) = \int_{|x-x_i|<\delta} e^{-\frac{H(x)}{T}}dx, \quad i = 1, 2, \ldots, m,$$
$$m(T, \delta) = \int_{|x-x_i|\ge\delta, i=1,2,\ldots,m} e^{-\frac{H(x)}{T}}dx.$$

The argument to establish claim in the proof of Theorem 4.2 yields that as $T \downarrow 0$ for each $\delta > 0$ sufficiently small such that $A_i \equiv \{|x - x_i| < \delta\}$ $i = 1, 2, \ldots, m$ are disjoint

$$\frac{e^{\lambda/T} c_i^{1/\alpha}}{T^{1/\alpha}} m_i(T, \delta) \to \int_{\mathbb{R}} e^{-|u|^\alpha} du$$

for $1 \le i \le m$ and

$$\frac{e^{\lambda/T} m(T, \delta)}{T^{1/\alpha}} \to 0.$$

Since for $\delta > 0$ small

$$P(|X_T - x_i| < \delta) = \frac{m_i(T, \delta)}{\sum_{j=1}^m m_j(T, \delta) + m(T, \delta)},$$

it follows that $X_T \xrightarrow{d} Y$. □

REMARK 5.1. It is important to note that the Hölder constant α for all x_i is the same but c_i can change with i.

The following extension of Theorem 5.1 is straightforward. It allows for the Hölder constant to vary with x_i.

THEOREM 5.2. *Let* $H, \mu_{H,T}, N$ *be as in Theorem* 5.1. *Assume also i) and ii) of Theorem 5.1 hold. Now assume*

iii)′ for all i, *there exists* $0 < \alpha_i < \infty$, $0 < c_i < \infty$ *such that*

$$\lim_{x \to x_i} \frac{H(x) - H(x_i)}{|x - x_i|^{\alpha_i}} = c_i.$$

Let $J \equiv \{i : 1 \le i \le m, \alpha_i = \max_{1 \le j \le m} \alpha_j\}$. *Then* $X_T \xrightarrow{d} Y'$ *where*

$$P(Y' = i) = \frac{\frac{1}{c_i^{1/\alpha}}}{\sum_{j \in J} \frac{1}{c_j^{1/\alpha}}}, \quad i \in J.$$

An analog of Theorem 4.2 is the following:

THEOREM 5.3. *Under the hypothesis of Theorem* 5.1, *for* $\delta > 0$ *sufficiently small (such that* $A_i \equiv \{|x - x_i| < \delta\}, i = 1, 2, \ldots, m$ *are disjoint)*

$$Pr\left(a \le \frac{(X_T - x_i) c_i^{1/\alpha}}{T^{1/\alpha}} \middle| |X_T - x_i| < \delta\right) \to \frac{\int_a^b e^{-|u|^\alpha} du}{\int_{\mathbb{R}} e^{-|u|^\alpha} du}.$$

PROOF. Same as Theorem 4.2. □

REMARK 5.2. One way to combine Theorems 5.1 and 5.3 is the following: For $T > 0$ and small, X_T has the same distribution as

$$Y + T^{1/\alpha} \frac{1}{(C(Y))^{1/\alpha}} (Z + \eta(T)),$$

where Y, Z and $\eta(T)$ are independent random variables such that Y is as in Theorem 5.1, Z has a symmetric distribution in $\mathbb{R}$ with density proportional to $e^{-|x|^\alpha}$, $\eta(T) \xrightarrow{P} 0$ as $T \downarrow 0$ and $c(x_i) = c_i, 1 \le i \le m$.

6 Gibbs measures on $\mathbb{R}^n$

6.1 Single well case.

THEOREM 6.1. *Let $H : \mathbb{R}^n \to \mathbb{R}^+$ be Borel measurable. Assume*

i) Assume $\int_{\mathbb{R}^n} e^{-H(x)} dx < \infty$. Let for $0 < T < 1$

$$\mu_{H,T}(A) \equiv \left(\int_A e^{-H(x)/T} dx \right) / \left(\int_{\mathbb{R}^n} e^{-H(x)/T} dx \right)$$

for $A \in \mathcal{B}(\mathbb{R}^n)$.

ii) Let $x_0 \in \mathbb{R}^n$ be such that for all $\delta > 0$, there exists $\eta > 0$ such that

$$a(\delta) \equiv \inf \{H(x) : |x - x_0| \ge \delta\} > b(\eta) \equiv \sup \{H(x) : |x - x_0| \le \eta\}.$$

Then for all $\delta > 0$, $\mu_{H,T} \{x : |x - x_0| > \delta\} \to 0$ as $T \downarrow 0$.

PROOF. Same as in Theorem 3.1. □

Let X_T be a random vector in $\mathbb{R}^n$ with distribution $\mu_{H,T}(\cdot)$. A natural question suggested by Theorem 6.1 is what is the rate at which $X_T \to x_0$ as $T \downarrow 0$. To answer this let us start with an example.

EXAMPLE 6.1. Let $n = 2$, $H(x_1, x_2) = x_1^2 + x_2^4$. Then the hypotheses of Theorem 6.1 hold with $x_0 = (0, 0)$. Let $X_T = (X_{T_1}, X_{T_2})$ be distributed as $\mu_{H,T}(\cdot)$. Then for $(a_i, b_i) \in \mathbb{R}^2$, $i = 1, 2$,

$$P \left(a_1 \le \frac{X_{T_1}}{\sqrt{T_1}} \le b_1, a_2 \le \frac{X_{T_2}}{T^{1/4}} \le b_2 \right)$$

$$= \frac{\int_{a_1\sqrt{T_1}\le x_1\le b_1\sqrt{T}, a_2T^{1/4}\le x_2\le b_2T^{1/4}} e^{-\frac{x_1^2+x_2^4}{T}} dx_1dx_2}{\int_{\mathbb{R}^2} e^{-\frac{(x_1^2+x_2^4)}{T}} dx_1dx_2}.$$

By the change of variable $x_1 = \sqrt{T}u_1$, $x_2 = T^{1/4}u_2$, the right side above becomes

$$\frac{T^{3/4}\int_{a_1\le u\le b_1, a_2\le u_2\le b_2} e^{-(u_1^2+u_2^4)} du_1du_2}{T^{3/4}\int_{\mathbb{R}^2} e^{-(u_1^2+u_2^4)} du_1du_2}.$$

Canceling $T^{3/4}$ we see that for $0 < T < 1$, $(X_{T_1}, X_{T_2}) \sim (X_1, X_2)$ where (X_1, X_2) is a random vector with an absolutely continuous distribution with density proportional to $e^{-(u_1^2+u_2^4)}$ in $\mathbb{R}^2$. This suggests the following:

Theorem 6.2. *Let $H : \mathbb{R}^n \to \mathbb{R}^+$ be Borel measurable. Assume:*

i) $\int_{\mathbb{R}^n} e^{-H(x)}dx < \infty$,

ii) for all $\delta > 0$, $a(\delta) \equiv \inf\{H(x) : |x| > \delta\} > H(0)$,

iii) there exists $\alpha_1, \alpha_2, \ldots, \alpha_n \in (0, \infty)$ *such that for all* $(u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$,

$$\frac{H(T^{\alpha_1}u_1, T^{\alpha_2}u_2, \ldots, T^{\alpha_n}u_n)}{T} \to g(u_1, u_2, \ldots, u_n) \in \mathbb{R}$$

as $T \downarrow 0$.

iv) $\int \sup_{0<T<1} e^{-\frac{H(T^{\alpha_1}u_1, T^{\alpha_2}u_2, \ldots, T^{\alpha_n}u_n)}{T}} du_1du_2\ldots du_n < \infty$.

For $0 < T < 1$, *let* $X_T = (X_{T_1}, X_{T_2}, \ldots, X_{T_n})$ *be a random vector with distribution* $\mu_{H,T}$ *as in Theorem* 6.1. *Then*

$$\left(\frac{X_{T_1}}{T^{\alpha_1}}, \frac{X_{T_2}}{T^{\alpha_2}}, \ldots, \frac{X_{T_n}}{T^{\alpha_n}}\right) \xrightarrow{d} X \equiv (X_1, X_2, \ldots, X_n)$$

and X has an absolutely continuous distribution in $\mathbb{R}^n$ with density proportional to $e^{-g(x_1,x_2,\ldots,x_n)}$.

Proof. Easy verification using the dominated convergence theorem and hypotheses *iii*) and *iv*). □

Thus different coordinates of X_T could go to zero at different rates.

Example 6.2. Let $n = 2$, $H(x_1, x_2) = x_1^2 + x_2^4 + x_1x_2^2 + x_1^2x_2^2$. Then $x_0 = (0, 0)$ and

$$H(\sqrt{T}u_1, T^{1/4}u_2) = Tu_1^2 + Tu_2^4 + Tu_1u_2^2 + T^{3/2}u_1^2u_2^2.$$

Thus the hypotheses of Theorem 6.2 are satisfied with

$$g(u) = u_1^2 + u_2^4 + u_1 u_2^2.$$

So

$$\left(\frac{X_{T_1}}{\sqrt{T}}, \frac{X_{T_2}}{T^{1/4}} \right) \xrightarrow{d} (X_1, X_2),$$

which has a probability density proportional to

$$\exp\left(-(u_1^2 + u_2^4 + u_1 u_2^2)\right).$$

Another set of sufficient conditions for studying the second order properties of X_T as $T \downarrow 0$ is given below. The proof is similar to that of Theorem 4.2.

THEOREM 6.3. *Let $H(\cdot)$ satisfy the hypotheses of Theorem* 6.1. *Let $\phi : \mathbb{R}^n \to \mathbb{R}^+ = [0, \infty)$ be such that*

i) $\frac{H(x)-H(x_0)}{(\phi(x-x_0))^\alpha} \to 1$ *as* $x \to x_0$,

ii) $\phi(\beta x) = |\beta| \phi(x)$ *for* $\beta \in \mathbb{R}$, $x \in \mathbb{R}^n$,

iii) $0 < m \equiv \inf \{\phi(x) : \|x\| = 1\} \leq M \equiv \sup \{\phi(x) : \|x\| = 1\} < \infty$,

iv) $\int_{\mathbb{R}^n} e^{-\phi(u)^\alpha} du < \infty$.

Let X_T be a random vector with distribution $\mu_{H,T}$. Then

$$\frac{(X_T - x_0)}{T^{1/\alpha}} \xrightarrow{d} X,$$

where X is a random vector with density

$$f_X(x) = c e^{-(\phi(x))^\alpha} \quad \text{for some } 0 < c < \infty.$$

EXAMPLE 6.3. 1. Let $H(\cdot)$ satisfy *i*), *ii*) of Theorem 6.1 and in addition be C^2 at x_0 such that there is a positive definite nonsingular matrix D such that

$$H(x) - H(x_0) = \frac{1}{2} \langle D(x - x_0), (x - x_0) \rangle + o(|x - x_0|^2).$$

Here set $\phi(x) = (\frac{1}{2} \langle Dx, x \rangle)^{1/2}$.

2. Let $H(\cdot)$ satisfy *i*) and *ii*) and be C^4 at x_0 with all second derivatives at x_0 vanishing but with $H(x) - H(x_0) = p(x - x_0) +$ higher order terms where $p(\cdot)$ is a homogeneous polynomial of order 4. Here set $\phi(x) = (p(x))^{1/4}$.

3. $H(x) - H(x_0) = \|A(x - x_0)\|^\alpha +$ higher order terms, where A is a nonsingular matrix. Here set $\phi(x) = \|Ax\|$.

6.2 Multiple well case.

THEOREM 6.4. *Let $H : \mathbb{R}^n \to \mathbb{R}^+$ $(n \geq 2)$ be Borel measurable. Assume*

i) $\int_{\mathbb{R}^n} e^{-H(x)}dx < \infty$,

ii) there exists $N \equiv \{x_i : 1 \leq i \leq m\} \subset \mathbb{R}^n$, $m \geq 2$ *such that for all* $\delta > 0$,

$$a(\delta) = \inf\{H(x) : |x - x_0| \geq \delta\} > \lambda$$

and for all i, $H(x_i) = \lambda, 0 < \lambda < \infty$,

iii) $\lim_{x \to x_0} H(x) = \lambda$ *for all* $1 \leq i \leq m$,

iv) there exists $\{\alpha_{ij} : 1 \leq j \leq n, 1 \leq i \leq m\}$, $0 < \alpha < \infty$ *such that* $\alpha_{ij} \geq 0$, *and* $\sum_{j=1}^{n} \alpha_{ij} = 1$, *for all* i *such that, as* $T \downarrow 0$,

$$\frac{1}{T}[H(x_{i1} + T^{\alpha\alpha_{i1}}u_1, x_{i2} + T^{\alpha\alpha_{i2}}u_2, \ldots, x_{in} + T^{\alpha\alpha_{in}}u_n) \\ - H(x_{i1}, x_{i2}, \ldots, x_{in})] \to g_i(u_1, u_2, \ldots, u_n),$$

v) for all i,

$$\sup_{0<T<1} e^{-\frac{H(x_{i1}+T^{\alpha\alpha_{i1}}u_1, x_{i2}+T^{\alpha\alpha_{i2}}u_2, \ldots, x_{in}+T^{\alpha\alpha_{in}}u_n)}{T}} \in L^1(\mathbb{R}^n).$$

Let X_T *be a* $\mathbb{R}^n$ *valued random vector with probability distribution*

$$\mu_{H,T}(A) = \frac{\int_A e^{-H/T}dx}{\int_{\mathbb{R}^n} e^{-H/T}dx}, \quad A \in \mathcal{B}(\mathbb{R}^n).$$

Then as $T \downarrow 0$, $X_T \xrightarrow{d} Y$ *where* Y *is a random vector with distribution*

$$P(Y = x_i) = \frac{\int_{\mathbb{R}^n} e^{-g_i(x)}dx}{\sum_{j=1}^{m} \int_{\mathbb{R}^n} e^{-g_j(x)}dx}, \quad 1 \leq i \leq m.$$

Further, there exists $\delta_0 > 0$ *such that for all* $0 < \delta < \delta_0$, *as* $T \downarrow 0$,

$$P\left(a_j \leq \frac{(X_T)_j - x_{ij}}{T^{1/\alpha}} \leq b_j, 1 \leq j \leq n \,\middle|\, |X_T - x_i| < \delta\right) \\ \to \frac{\int_{a_j \leq u_j \leq b_j, j=1,2,\ldots,n} e^{-g_i(u_1,u_2,\ldots,u_n)}du_1du_2\ldots du_n}{\int_{\mathbb{R}^n} e^{-g_i(u_1,\ldots,u_n)}du_1du_2\ldots du_n}.$$

PROOF. The proof is similar to that of Theorem 5.2. □

A similar extension of Theorem 6.3 to the multiple well case can be formulated and proved.

References

ATHREYA, K.B. (2009). Entropy maximization. *Proc. Indian Acad. Sci. (Math. Sci.)*, **119**, 531–539.

FREIDLIN, M. AND WENTZELL, A. (1994). *Random Perturbations of Hamiltonian Systems.* Springer.

GEMAN, S. and GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 721–741.

HASTINGS, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.

HWANG, C.R. (1980). Laplace's method revised: weak convergence of probability measures. *Ann. Probab.*, **8**, 1177–1182.

HWANG, C.R. (1981). A generalization of Laplace's method. *Proc. Amer. Math. Soc.*, **82**, 446–451.

HWANG, C.R. and SHEU, S.J. (1990). Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker-Planck operators and simulated annealing. *Acta Appl. Math.*, **19**, 253–295.

METROPOLIS, N., ROSENBLUTH, A., ROSENBLUTH, M., TELLER, A. and TELLER, E. (1953). Equations of state of calculation by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1092.

ROBERT, C.P. and CASELLA, G. (2004). *Monte Carlo Statistical Methods.* Second Edition. Springer Texts in Statistics. Springer-Verlag, New York.

K. B. ATHREYA
DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IA 50011-1210 USA
E-mail: kbathreya@gmail.com

CHII-RUEY HWANG
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
TAIPEI, TAIWAN, 11529
E-mail: crhwang@sinica.edu.tw

Paper received January 2009; revised September 2009.