COMMENTARII MATHEMATICI
UNIVERSITATIS SANCTI PAULI
Vol. 47, No. 1 1998

ed. RIKKYO UNIV/MATH
IKEBUKURO TOKYO
171 JAPAN

# On Variational Principle and the Metrics Associated with a Potential of Bounded Variation

by

Masamichi YOSHIDA

(Received August 26, 1997)

**Abstract.** We introduce a "family of metrics" associated with a potential and give another proof of the variational characterization of Gibbs measures without the use of "local specification property".

## 1. Introduction

In [2], F. Ledrappier showed using the Ruelle's nonlinear ergodic theorem ([4]) for a $C^2$ map on an interval that the absolutely continuous invariant measures are characterized by a variational principle. When one considers the behavior of a one-sided shift under a given potential, one can define the Gibbs measure associated with the potential and show that this measure is also characterized by a variational principle ([1]). Indeed, R. Bowen showed this characterization using the "local specification property" of Gibbs measures ([1] Theorem 1.2). In this paper using Ledrappier's technique and a "family of metrics associated with a potential" one can give another proof of the variational characterization of Gibbs measures without the use of "local specification property". This "family of metrics" inform us of the unstable nature of a one-sided shift under a given potential. Following Ledrappier's technique one can find the density of the associated Gibbs measure.

## 2. Family on metrics associated with potential

The set-up is the following. Let $B=\{0, 1\}^{\boldsymbol{Z}_+}$ and $T$ be the one-sided shift on $B$, that is, $(Tb)_n=b_{n+1}$ for each $b\in B$ and $n\in\boldsymbol{Z}_+$. In order to consider the natural extension of $(B, T)$, prepare "source space" $X=\{0, 1\}^{\boldsymbol{N}}$ and the one-sided shift $S$ on $X$. So the natural extension of $(B, T)$ is $\tilde{B}=B\times X$. Here we shall write an element of $X$ as $x=x_1x_2\cdots$, of $B$ as $b=\cdots b_1b_0$, and of $\tilde{B}$ as $[b, x]=\cdots b_1b_0x_1x_2\cdots$. Let $\sigma$ be the two-sided shift on $\tilde{B}$ i.e. $\sigma: [b, x]\mapsto[Tb, b_0x]$ where $b_0x=b_0x_1x_2\cdots\in X$. Define the holonomy mapping from $b\in B$ to $c\in B$ by

$$h_b(c): [c, X]\to[b, X]: [c, x]\mapsto[b, x]\,.$$

When a potential on the original space $B$ is given, we can define the "Jacobian" of the holonomy with respect to this potential as follows.

Let $\boldsymbol{D}(B)$ be the space of functions on $B$ with bounded logarithmic variations. Thus a positive-valued function $f\colon B\to(0,\infty)$ belongs to $\boldsymbol{D}(B)$ if there are constants $C>0$ and $0<\alpha<1$ so that $\mathrm{var}_k(f)=\sup\{|\log f(b)-\log f(c)|: b_n=c_n \text{ for all } n\le k\}\le C\alpha^k$ for each $k\in\boldsymbol{Z}_+$. In this paper we shall regard $\boldsymbol{D}(B)$ as the space of potentials (Usually $\{\log f\,|\,f\in\boldsymbol{D}(B)\}$ is the space of potentials [1]). Associated with $f\in\boldsymbol{D}(B)$ we can define its cocycle by the formula $f(0,b)=1$ and $f(n,b)=\prod_{i=0}^{n-1}f(T^ib)$ and put a metric on each $[c,X]$ in the following way. Fix $\beta\in(0,1)$ and set $\hat{f}=\beta\dfrac{f}{\|f\|}$ where $\|f\|$ is the sup-norm on $B$. For two distinct points $x,y$ in $X$, one can define their initial common word $w(x,y)$ and its length $l(x,y)$, that is, $l(x,y)=\min\{n\in\boldsymbol{N}: x_{n+1}\ne y_{n+1}\}$ and $w(x,y)=x_1x_2\cdots x_{l(x,y)}$ where it is empty word if $l(x,y)=0$. Put a metric on $[c,X]$ such that for $x,y\in X$,

$$d_c(x,y)=\begin{cases}\hat{f}(l(x,y),cw(x,y))\,, & \text{if } x\ne y\\ 0\,, & \text{if } x=y\end{cases}$$

where $cw(x,y)=\cdots c_1c_0x_1x_2\cdots x_{l(x,y)}\in B$. Notice that if $f$ is constant, the associated metrics are usual one. Indeed we can show that these are metrics.

LEMMA 1. *For each $c\in B$, $d_c$ is a metric on $X$.*

*Proof.* We show that $d_c$ is a non-archimedean metric i.e. $d_c(x,z)\le\max\{d_c(x,y),d_c(y,z)\}$. If $l(x,z)>l(x,y)$, then $T^{l(x,z)-l(x,y)}(cw(x,z))=cw(x,y)$ and hence

$$d_c(x,z)=\left(\frac{\beta}{\|f\|}\right)^{l(x,z)}\left(\prod_{i=0}^{l(x,z)-l(x,y)-1}f(T^i(cw(x,z)))\right)\prod_{j=l(x,z)-l(x,y)}^{l(x,z)-1}f(T^j(cw(x,z)))$$

$$\le\beta^{l(x,z)-l(x,y)}\left(\frac{\beta}{\|f\|}\right)^{l(x,y)}\left(\prod_{k=0}^{l(x,y)-1}f(T^k(cw(x,y)))\right)<d_c(x,y)\,.$$

If $l(x,z)=l(x,y)$, then $w(x,z)=w(x,y)$ and $d_c(x,z)=d_c(x,y)$. If $l(x,z)<l(x,y)$, then $l(x,z)=l(y,z)$ and hence $w(x,z)=w(y,z)$ and $d_c(x,z)=d_c(y,z)$. ∎

Next we define the "Jacobian" $D_{[b,x]}(c)$ of the mapping $h_b(c)$ at $[c,x]$ by the formula $D_{[b,x]}(c)=\lim_{y\to x}\dfrac{d_b(x,y)}{d_c(x,y)}$ where "$y\to x$" means the convergence with respect to the product topology of $X$. In fact the above limit exists.

LEMMA 2. *For each $b,c\in B$, $D_{[b,x]}(c)$ exists and equals to $\prod_{n\ge1}\dfrac{f(bx_1\cdots x_n)}{f(cx_1\cdots x_n)}$ where $bx_1\cdots x_n=\cdots b_1b_0x_1\cdots x_n\in B$.*

*Proof.* Let $x\ne y\in X$ be arbitrarily given.

$$\frac{d_b(x,y)}{d_c(x,y)}=\prod_{i=0}^{l(x,y)-1}\frac{f(bx_1\cdots x_{l(x,y)-i})}{f(cx_1\cdots x_{l(x,y)-i})}=\prod_{j=1}^{l(x,y)}\frac{f(bx_1\cdots x_j)}{f(cx_1\cdots x_j)}$$

by letting $j=l(x,y)-i$. Since $l(x,y)\to\infty$ as $y\to x$ and $f\in \boldsymbol{D}(B)$, we get the desired result. ∎

LEMMA 3. *For each $b,c\in B$ and $x\in X$,*

$$D_{[b,x]}(cb_{n-1}\cdots b_0)=D_{\sigma^n[b,x]}(c)\frac{f(n,cb_{n-1}\cdots b_0)}{f(n,b)}.$$

*Proof.* Indeed

$$D_{[b,x]}(cb_{n-1}\cdots b_0)=\prod_{k\ge 1}\frac{f((T^nb)b_{n-1}\cdots b_0x_1\cdots x_k)}{f(cb_{n-1}\cdots b_0x_1\cdots x_k)}$$

$$=\left(\prod_{k\ge 1}\frac{f((T^nb)b_{n-1}\cdots b_0x_1\cdots x_k)}{f(cb_{n-1}\cdots b_0x_1\cdots x_k)}\right)\left(\prod_{i=0}^{n-1}\frac{f((T^nb)b_{n-1}\cdots b_i)}{f(cb_{n-1}\cdots b_i)}\right)\frac{f(n,cb_{n-1}\cdots b_0)}{f(n,b)}$$

$$=D_{\sigma^n[b,x]}(c)\frac{f(n,cb_{n-1}\cdots b_0)}{f(n,b)}.$$

∎

## 3. Proof of the variational principle

One can define the exponential pressure $\lambda_f\ge 0$ and $f$-conformal measure $\nu_f$ on $B$ of a potential $f$ ([1]). They satisfies the following condition. For each $i\in\{0,1\}$,

$$\frac{d\nu_f\circ T_{[i]_0}}{d\nu_f}(b)=\lambda_f f(b)\quad \nu_f\text{-}a.e.\ b\in[i]_0$$

where $[i]_0=\{b\in B: b_0=i\}$ and $\nu_f\circ T|_{[i]_0}(E)=\nu_f(TE)$ for each Borel subset $E\subset[i]_0$ (Notice that $TE$ is a Borel subset of $B$ since $T|_{[i]_0}:[i]_0\to B$ is continuous injective). So the $f$-conformality implies that $T$ is strongly non-singular i.e. $\nu_f(TE)=0$ for each Borel subset $E\subset[i]_0$ with $\nu_f(E)=0$ and that the distortion of $T$ with respect to $\nu_f$ is $f$ up to multiplication by a constant. For the completeness we shall prove the existence of $\lambda_f$ and $\nu_f$ (See [1] theorem 1.7, p. 14). First define $L_f$ on the space of continuous functions on $B$ in the following way: $(L_f\phi)(b)=\sum_{c\in T^{-1}(b)}\frac{\phi(c)}{f(c)}=\sum_{i=0}^{1}\frac{\phi(bi)}{f(bi)}$. So $L_f$ is a positive linear operator and one can define the dual operator $L_f^*$ on the space of finite Borel measures on $B$ and its normalization $\mathscr{L}_f$ on the space of probability measures on $B$, that is, $\mathscr{L}_f\nu=L_f^*\nu/L_f^*\nu(B)$. Since $\mathscr{L}_f$ is continuous on $P(B)$, $\mathscr{L}_f$ has a fixed point $\nu_f$ by the Schauder-Tychonoff's theorem. Let $\lambda_f=L_f^*\nu_f(B)$. Then $\lambda_f$ and $\nu_f$ satisfies the above condition. Indeed for each subcylinder $C\subset[i]_0$,

$$\int_C\lambda_f f\,d\nu_f=\int_B\chi_C f\,dL_f^*\nu_f=\int_B L_f(\chi_C f)d\nu_f=\int_B\chi_C(bi)d\nu_f(b)=\int_C d\nu_f\circ T_{[i]_0}$$

where $\chi_C$ is the characteristic function of $C$ (this is continuous). Here we remark that $\lambda_f = \exp P(T, -\log f)$ where $P(T, -\log f)$ is the "pressure" of $-\log f$ (See [1], p. 30). To any $T$-invariant probability measure $\mu$ on $B$ we assign a $\sigma$-invariant measure $\tilde{\mu}$ on $\tilde{B}$ in a homogeneous way i.e. for each cylinder $\tilde{C} = \{[b, x] : b \in [i_{l-1} \cdots i_0]_0^{l-1}$ and $x_1 \cdots x_n = j_1 \cdots j_n\}$ (where $i_{l-1}, \cdots, i_0, j_1, \cdots, j_n \in \{0, 1\}$ and $[i_{l-1} \cdots i_0]_0^{l-1} = \{b \in B \mid b_{l-1} \cdots b_0 = i_{l-1} \cdots i_0\}$), $\tilde{\mu}(\tilde{C}) = \mu(\pi(\sigma^{-n}\tilde{C}))$ where $\pi : \tilde{B} \to B$ is the natural projection. Consider the refining and generating sequence of partitions of $\tilde{B}$ given by

$$L_0 = \{[B, x] : x \in X\},$$
$$L_n = \sigma^{-n} L_0 = \{[B x_1 \cdots x_n, S^n x] : x \in X\}.$$

We write the element of $L_n$ containing $[b, x]$ as $L_n[b, x]$, that is, $L_n[b, x] = \{[c, x] : c \in [b_{n-1} \cdots b_0]_0^{n-1}\}$. By lemma 3, we can show the following.

LEMMA 4. *For each $T$-invariant probability measure $\mu$ on $B$ and $n \geq 1$,*

$$n \int_B \log(\lambda_f f) d\mu = -\int_{\tilde{B}} \log \int_{\pi L_n[b,x]} \frac{D_{[b,x]}}{\nu_f(D_{[b,x]})} d\nu_f d\tilde{\mu}([b, x])$$

$$\textit{where} \quad \nu_f(D_{[b,x]}) = \int_B D_{[b,x]} d\nu_f .$$

*Proof.*

$$-\int_{\tilde{B}} \log \int_{\pi L_n[b,x]} \frac{D_{[b,x]}}{\nu_f(D_{[b,x]})} d\nu_f d\tilde{\mu}([b, x])$$

$$= -\int_{\tilde{B}} \log \int_{[b_{n-1} \cdots b_0]_0^{n-1}} \frac{1}{\lambda_f^n \nu_f(D_{[b,x]})} \frac{D_{[b,x]}(d)}{f(n, d)} d\nu_f \circ T^n|_{[b_{n-1} \cdots b_0]_0^{n-1}}(d) d\tilde{\mu}([b, x])$$

$$= n \log \lambda_f + \int_{\tilde{B}} \log \nu_f(D_{[b,x]}) d\tilde{\mu}([b, x]) - \int_{\tilde{B}} \log \int_B \frac{D_{\sigma^n[b,x]}(c)}{f(n, b)} d\nu_f(c) d\tilde{\mu}([b, x])$$

(by letting $d = c b_{n-1} \cdots b_0$ and lemma 3)

$$= n \int_B \log(\lambda_f f) d\mu \text{ (by the invariance of } \mu \text{ and } \tilde{\mu}).$$

∎

Now we can give another proof of the variational characterization of Gibbs measures. We write the conditional probability of $E \subset \tilde{B}$ at $[b, x]$ about the Borel $\sigma$-algebra $\boldsymbol{B}(L_n)$ generated by $L_n$ with respect to $\tilde{\mu}$ as $\tilde{\mu}[E \mid L_n]([b, x])$, the conditional entropy of $L_n$ under $L_0$ with respect to $\tilde{\mu}$ as $H_{\tilde{\mu}}(L_n \mid L_0)$, the entropy of $T$ with respect to $\mu$ as $h_\mu(T)$ and the entropy of $\sigma$ with respect to $\tilde{\mu}$ as $h_{\tilde{\mu}}(\sigma)$.

THEOREM 5. *Let $\mu$ be a $T$-invariant probability measure on $B$.*
*If $\mu$ satisfies the formula $h_\mu(T) = \int_B \log(\lambda_f f) d\mu$, then $\mu$ is equivalent to $\nu_f$.*

*Proof.* By the assumption and lemma 4,

$$-\int_{\tilde{B}}\log\int_{\pi L_n[b,x]}\frac{D_{[b,x]}}{\nu_f(D_{[b,x]})}d\nu_f d\tilde{\mu}([b,x])=nh_\mu(T)=nh_{\tilde{\mu}}(\sigma)=h_{\tilde{\mu}}(\sigma^n)$$
$$=H_{\tilde{\mu}}(L_n\,|\,L_0)\quad(\text{since }\{L_n\}_{n\geq 0}\text{ is generating})$$
$$=-\int_{\tilde{B}}\log\tilde{\mu}[L_n[b,x]\,|\,L_0([b,x)d\tilde{\mu}([b,x]).$$

Thus

$$\int_{\tilde{B}}\log\frac{\displaystyle\int_{\pi L_n[b,x]}\frac{D_{[b,x]}}{\nu_f(D_{[b,x]})}d\nu_f}{\tilde{\mu}[L_n[b,x]\,|\,L_0]([b,x])}d\tilde{\mu}([b,x])=0.$$

Accordingly if we define a probability measure $\tilde{\nu}$ on $\boldsymbol{B}(L_n)$ so that

$$\tilde{\nu}(\tilde{E})=\int_{\tilde{B}}\int_{\pi\tilde{E}}\frac{D_{[b,x]}}{\nu_f(D_{[b,x]})}d\nu_f d\tilde{\mu}([b,x])\qquad\text{for each}\quad E\in\boldsymbol{B}(L_n),$$

then $\int_{\tilde{B}}\log(d\tilde{\nu}/d\tilde{\mu})d\tilde{\mu}=0$. By the concavity of the logarithm, $\tilde{\nu}=\tilde{\mu}$ on $\boldsymbol{B}(L_n)$. Since $\{L_n\}_{n\geq 0}$ is generating, we can extend $\tilde{\nu}$ to a probability measure on the Borel $\sigma$-algebra of $\tilde{B}$ and $\tilde{\nu}=\tilde{\mu}$. Here for each Borel subset $E\subset B$, $\mu(E)=\tilde{\mu}(\pi^{-1}E)=\tilde{\nu}(\pi^{-1}E)$. Therefore $\mu\ll\nu_f$ by the definition of $\tilde{\nu}$. Conversely if $\mu(E)=0$, then $\int_E(D_{[b,x]}/\nu_f(D_{[b,x]}))d\nu_f=0$ $\tilde{\mu}$-a.e. $[b,x]\in\tilde{B}$ and hence $\nu_f(E)=0$. ∎

Finally one can prove that if $\nu_f$ is non-atomic and ergodic under $T$ and $\mu$ is a $T$-invariant probability measure equivalent to $\nu_f$ (such $\mu$ is unique by the ergodicity), then $\mu$ satisfies the formula $h_\mu(T)=\int_B\log(\lambda_f f)d\mu$ ([3] corollary 3).

So we get the variational characterization of Gibbs measures.

## References

[1] R. Bowen; *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Springer Lecture Notes in Mathematics 470. Berlin-Heidelberg-New York: Springer 1975.
[2] F. Ledrappier; Some properties of absolutely continuous invariant measures on an interval. *Ergod. Th. Dynam. Sys.* **1** (1981), 77–93.
[3] W. Parry; On Rohlin's formula for entropy. *Acta math. Acad. Sci. Hungar.* **15** (1964), 107–113.
[4] D. Ruelle; Ergodic theory of differentiable dynamical systems. *Publ. Math. IHES* **50** (1979), 27–58.

Department of Mathematics
Osaka City University, 3–3–138
Sugimoto Sumiyoshi-ku, Osaka 558, Japan

*E-mail address*: yoshida@sci.osaka-cu.ac.jp