THERMODYNAMIC FORMALISM OUT OF EQUILIBRIUM, AND GIBBS PROCESSES

S. BEN OVADIA, F. RODRIGUEZ-HERTZ

Abstract. We study the thermodynamic formalism of systems where the potential depends randomly on an exterior system. We define the *pressure out of equilibrium* for such a family of potentials, and prove a corresponding variational principle. We present an application to random dynamical systems. In particular, we study an open condition for random dynamical systems where the randomness is driven by a Gibbs process, and prove hyperbolicity estimates that were previously only known in the i.i.d setting.

Contents

1. Intorduction

Given an ensemble of particles and a finite collection of states, the Helmholtz free energy is the difference between the potential energy of a distribution, and its entropy scaled by the inverse temperature. Minimizing the Helmholtz free energy is equivalent to maximizing the difference between the entropy and the potential energy scaled by the temperature. In Ruelle's thermodynamic formalism ([Rue67]), the variational principle states that given a potential $-\frac{1}{\tau}\phi \in \mathrm{Höl}(\Sigma)$ ($\tau > 0$ is the temperature), where Σ is a one-sided compact and topologically transitive topological Markov shift,

$$\max\{h_\nu(T) + \int \phi d\nu : \nu \text{ is } T\text{-inv. and erg.}\} = P(\phi),$$

where $T:\Sigma\to\Sigma$ is the left-shift, and

$$P(\phi):=\limsup_{n\to\infty}\frac{1}{n}\log\sum_{|\underline{w}|=n,w_{n-1}=a}e^{\sum_{k=0}^{n-1}\phi\circ T^k(\underline{w}\omega_a)},\tag{1}$$

for any $[a]\subseteq\Sigma$ and $\omega_a\in[a]$. Indeed, the value of $P(\phi)$ is independent of the choice of $[a]\subseteq\Sigma$ and $\omega_a\in[a]$, as ϕ is Hölder continuous. That is, Ruelle showed that minimizing the free Helmholtz energy over invariant measures (that is over steady states) can be achieved, and that the minimal value can be expressed as the logarithm of the reciprocal of the spectral radius of an associated operator (called the Ruelle operator).

The variational principle is an optimization problem over a collection of measures, with deep physical motivation. Since the pioneering work of Ruelle, several important extensions have been studied, with different related optimization problems. An important extension is in the study of extensions.

In [LW77] Ledrappier and Walters study the following optimization. Let X and Y be two compact metric spaces, let $T:X\to X$ and $S:Y\to Y$ be two continuous maps, and let $\pi:X\to Y$ be a surjection s.t $\widehat{\pi}\circ T=S\circ\widehat{\pi}$. Then, given an S-invariant ν and a potential $\phi\in C(X)$, Ledrappier and Walters optimize

$$\sup\{h_\mu(T|S)+\int\phi d\mu:\mu\text{ is }T\text{-invariant and }\mu\circ\widehat{\pi}^{-1}=\nu\},$$

where $h_\mu(T|S)$ is the relative entropy. That is, they optimize the contribution of the potential together with the entropy coming from the fibers alone, under the constraint where the invariant measures project to a fixed measure ν on the base.

In [DKS08], Denker, Kifer, and Stadlbauer study a notion of random topological pressure, and prove that it satisfies a relative variational principle as well (see also [Kif01] for a definition of random pressure). In their setting, the fibers are countable Markov shifts, and they study the associated composition of random Ruelle operators w.r.t to a probability preserving transformation on the base dynamics (see [BG95] for results in the compact-shift fibers setting). This is a somewhat opposite approach to the one we are taking, as we study randomness given by a shift space, but the fibers are a general compact metric space; In particular, our fiber dynamics are not given by the random composition of Ruelle operators. In addition, we do not fix a probability measure on the base (nor fibers), but instead seek an optimization over all probabilities. A full description of our setting appears below.

Another extension of Ruelle's variational principle is studied in [BKL23], where the authors study an exponential growth rate as in (1), where $\sum_{k=0}^{n-1}\phi\circ T^k$ is generalized to $n\cdot G(\frac{1}{n}\sum_{k=0}^{n-1}\phi\circ T^k)$ where G is a continuous function.

In this paper we study another extension of Ruelle's variational principle, with applications to random dynamics (in particular to Gibbs processes).

We consider a compact metric space X, and a Hölder continuous skew-product over Σ: $\widehat{F} : \Sigma \times X \to \Sigma \times X$, $\widehat{F}(\omega, t) = (T\omega, F_\omega(t))$, together with a family of Hölder continuous potentials $\phi_t : \Sigma \to \mathbb{R}$, which depend in a Hölder continuous way on $t \in X$. Let $\pi : \Sigma \times X \to \Sigma$ be the projection $(\omega, t) \mapsto \omega$. Then we study the following optimization problem:

$$\sup\{h_{\widehat{\nu}\circ\pi^{-1}}(T) + \int \phi_t(\omega) d\widehat{\nu}(\omega, t) : \widehat{\nu} \text{ is an } \widehat{F}\text{-invariant probability on } \Sigma \times X\}.$$

In other words, we optimize while allowing the potential to be random. We introduce the corresponding *pressure out of equilibrium*:

$$\check{P}(\{\phi_t\}_{t\in X}) := \limsup_{n\to\infty} \sup_{t\in X} \frac{1}{n} \log \sum_{|\underline{w}|=n, w_{n-1}=a} e^{\sum_{k=0}^{n-1} \phi_{F_\omega^k(t)} \circ T^k(\underline{w}\omega_a)},$$

where $F_\omega^k := F_{T^{k-1}\omega} \circ \cdots \circ F_\omega$. Our variational principle assets that (Proposition 3.4 and Theorem 3.8),

$$\check{P}(\{\phi_t\}_{t\in X}) = \max\{h_{\widehat{\nu}\circ\pi^{-1}}(T) + \int \phi_t(\omega) d\widehat{\nu}(\omega, t) : \widehat{\nu} \text{ is an } \widehat{F}\text{-inv.}\}.$$

The notion of pressure out equilibrium rises naturally in the study of skew-products (where we are removing the fiber entropy from the optimization process).

The pressure out of equilibrium also rises naturally when studying random dynamical systems. In the recent years significant progress was made in the study of smooth dynamical systems, where the diffeomorphism is composed randomly. See for example [BQ16, BRH17, DD].

This model is important as every physical system in fact interacts with a larger system, and so the laws of motion themselves may depend randomly on a larger system. In this setting, it is natural to assume that the typical behavior of the larger system is given by a physical measure (e.g SRB) or a measure of maximal entropy- both of which are Gibbs measures. Consequently, we are interested in random dynamical systems where the randomness is driven by a Gibbs process. However, almost all previous results treat the i.i.d case, where the random composition is taken over a finite set of diffeomorphisms.

An important open class of random dynamical systems is systems which satisfy the *uniform expansion on average* condition:

$$\inf_{(x,\xi)\in T^1M} \int_\Sigma \log |d_x f_\omega \xi| d\mu(\omega) > 0,$$

where $f_\omega \in \mathrm{Diff}^{1+\alpha}(M)$ ($\omega \mapsto f_\omega$ may not be constant on partition sets) and M is a closed Riemannian manifold. The prevalence of this condition has been studied for example in [DK07, Pot22, ES23, BEFRH].

In that setting, we prove the following hyperbolicity estimates (Theorem 4.6): $\exists \beta, \gamma, C > 0$ s.t

$$\forall n \geq 0, \quad \sup_{(x,\xi)\in T^1M} \int_\Sigma |d_x f_\omega^n|^{-\beta} d\mu(\omega) \leq Ce^{-\gamma n}, \tag{2}$$

where $f_\omega^n := f_{T^{n-1}\omega} \circ \cdots \circ f_\omega$. One can think of the estimate in (2) as a "non i.i.d version of Azuma's inequality". In this setting, even a Markov process is a non-trivial extension of the i.i.d case (i.e Bernoulli process).

2. Pressure out of equilibrium

Setup:

(1) Let X be a compact metric space.
(2) Let Σ be a one-sided compact topological Markov shift, endowed with the left-shift $T : \Sigma \to \Sigma$.
(3) Let $F : \Sigma \to C_\alpha(X, X)$ be a Hölder continuous map, denoted by $\omega \mapsto F_\omega$, where F_ω is a α-Hölder homeomorphism of X.
(4) Let $\{\phi_t\}_{t\in X} \subseteq \text{Höl}(\Sigma)$ be a equi-ölder family of potentials: $\sup_{t\in X} \|\phi_t\|_{\text{Höl}} < \infty$.
(5) Set $F_\omega^k := F_{T^{k-1}\omega} \circ \cdots \circ F_\omega$.

Definition 2.1. *We define the* pressure out of equilibrium *as*

$$\check{P}(\{\phi_t\}_{t\in X}) := \limsup_{n\to\infty} \frac{1}{n} \log \sup_{t\in X} \check{Z}_n(\phi_t, a),$$

where

$$\check{Z}_n(\phi_t, a) := \sum_{|\underline{w}|=n, w_{n-1}=a} e^{\sum_{k=0}^{n-1} \phi_{F^k_{\underline{w}\omega_a}(t)} \circ T^k(\underline{w}\omega_a)}$$

and $[a] \subset \Sigma$ *and* $\omega_a \in [a]$.

Remark:

(1) Note, since Σ is compact and topologically transitive, the definition of the POE does not depend on a. Since the family $\{\phi_t\}_{t\in X}$ is equi-Hölder, the definition of the POE also does not depend on the choice of $\omega_a \in [a]$.
(2) Note, $\check{Z}_n(\phi_t, a)$ is not of the composit form $L_{\phi_n} \circ \cdots \circ L_{\phi_1} \mathbb{1}_{[a]}$! For each word of length n, the sequence of potentials in its corresponding weight depends on the word.
(3) The notation of Definition 2.1 is abused, as $\check{Z}_n(\phi_t, a)$ should in fact be $\check{Z}_n(\{\phi_s\}_{s\in X}, t, a)$. However, in order to not have too heavy of a notation, we keep the abused, yet clear notation of Definition 2.1.

Lemma 2.2. *The following limit exists:*

$$\check{P}(\{\phi_t\}_{t\in X}) = \lim_{n\to\infty} \frac{1}{n} \log \sup_{t\in X} \check{Z}_n(\phi_t, a).$$

Proof. First we note that since Σ is compact, as in the remark following Definition 2.1, there exists a constant $C_{\Sigma,\phi} > 1$ so we can write

$$\begin{aligned}\check{P}(\{\phi_t\}_{t\in X}) &= \limsup_{n\to\infty}\sup_t \frac{1}{n}\log\Big(C_{\Sigma,\phi}^{\pm 1}\sum_{[a]\subseteq\Sigma}\sum_{|\underline{w}|=n, w_{n-1}=a} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\underline{w}\omega_a}(t)}\circ T^k(\underline{w}\omega_a)}\Big)\\ &= \limsup_{n\to\infty}\sup_t \frac{1}{n}\log\Big(\sum_{[a]\subseteq\Sigma}\sum_{|\underline{w}|=n, w_{n-1}=a} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\underline{w}\omega_a}(t)}\circ T^k(\underline{w}\omega_a)}\Big)\\ &= \limsup_{n\to\infty}\frac{1}{n}\log Z_n(\{\phi_t\}_t),\end{aligned}$$

where $Z_n(\{\phi_t\}_t) := \sup_t \sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t)}\circ T^k(\theta_{\underline{w}})}$ and $\theta_{\underline{w}} \in [\underline{w}]$ maximizes $\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t)}\circ T^k(\theta_{\underline{w}})$.

Then we notice that for all $n,m\in\mathbb{N}$,

$$\log Z_{n+m}(\{\phi_t\}_t) \le \log Z_n(\{\phi_t\}_t) + \log Z_m(\{\phi_t\}_t),$$

since the collection of admissible words of length $n+m$ is smaller or equal to the concatenation of all words of length n and all words of length m, and $\sup_t\{A_t\cdot B_t\} \le \sup_t A_t \cdot \sup_t B_t$.Then by Fekete's lemma, we are done. □

3. Variational principle

The proof of the variational principle for Hölder continuous potentials on TMSs goes through Ruelle operators. We cannot use the techniques of Ruelle operators in this setting, as the POE is not the spectral radius of a Ruelle operator, nor of a random composition of Ruelle operators!

3.1. Upper bound.

Definition 3.1 (Maximal asymptotic instability). *Let $\widehat{F} : \Sigma\times T^1M \to \Sigma\times X$, $\widehat{F}((\omega,t)) := (T\omega, F_\omega(t))$, and let $\widehat{\phi} := \Sigma\times X\to\mathbb{R}$, $\widehat{\phi}((\omega,t)) := \phi_t(\omega)$. Then $\widetilde{\phi}:\Sigma\to\Sigma$,*

$$\widetilde{\phi}(\omega) := \limsup_{n\to\infty}\sup_t \frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}^k(\omega,t),$$

is called the maximal asymptotic instability potential *of $\{\phi_t\}_{t\in X}$.*

Remark: The terminology stems from the comparison between $\widetilde{\phi}$ and ϕ (i.e the case where $\{\phi_t\}_{t\in X}$ is a singleton).

Definition 3.2. *If substituting the* $\limsup\sup_t$ *by a* $\liminf\inf_t$ *in Definition 3.1 does not change $\widetilde{\phi}$ for ν-a.e ω, then we say that $\{\phi_t\}_{t\in X}$ is* asymptotically stable w.r.t ν.

Theorem 3.3.

$$\check{P}(\{\phi_t\}_t) \le \sup_{\nu\text{ erg. inv. prob.}}\Big\{h_\nu(T) + \int\widetilde{\phi}(\omega)d\nu(\omega)\Big\}.$$

This proof follows the ideas of Walters in [Wal75].

Proof. Let t_n s.t $\sup_t \sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t)}\circ T^k(\theta_{\underline{w}})} = \sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t_n)}\circ T^k(\theta_{\underline{w}})}$ for some $\theta_{\underline{w}}\in[\underline{w}]$. Set

$$\widetilde{\nu}_n := \frac{\sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t_n)}\circ T^k(\theta_{\underline{w}})}\delta_{\theta_{\underline{w}}}}{\sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t_n)}\circ T^k(\theta_{\underline{w}})}}, \text{ and } \nu_n := \frac{1}{n}\sum_{j=0}^{n-1}\widetilde{\nu}_n\circ T^{-j}.$$

Choose $n_j\uparrow\infty$ s.t $\nu_{n_j}\to\nu$, and note $\frac{1}{n_j}\log\sum_{|\underline{w}|=n_j} e^{\sum_{k=0}^{n_j-1}\phi_{F^k_{\theta_{\underline{w}}}(t_{n_j})}\circ T^k(\theta_{\underline{w}})}\to$ $\check{P}(\{\phi_t\}_t)$ (recall Lemma 2.2).

It is enough to show that $h_\nu(T)+\int\widetilde{\phi}(\omega)d\nu(\omega)\geq\check{P}(\{\phi_t\}_t)$ since the l.h.s respects the ergodic decomposition, and it is clear that ν is invariant. Let $\mathcal{C}$ be the partition into cylinders of length one. Then,

$$\begin{aligned}
&H_{\widetilde{\nu}_n}\Big(\bigvee_{i=0}^{n-1}T^{-j}[\mathcal{C}]\Big)+\int\sum_{j=0}^{n-1}\phi_{F^j_\omega(t_n)}\circ T^j(\omega)d\widetilde{\nu}_n\\
&=\sum_{|\underline{w}|=n}\widetilde{\nu}_n(\{\theta_{\underline{w}}\})\cdot\Big(\sum_{j=0}^{n-1}\phi_{F^j_{\theta_{\underline{w}}}(t_n)}\circ T^j(\theta_{\underline{w}})-\log\widetilde{\nu}(\{\theta_{\underline{w}}\})\Big)\\
&=\log\sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1}\phi_{F^k_{\theta_{\underline{w}}}(t_n)}\circ T^k(\theta_{\underline{w}})}\equiv\log\sum_{|\underline{w}|=n}e^{S_n(\theta_{\underline{w}})}.
\end{aligned}\tag{3}$$

Fix $1\leq q<n$, and for any $0\leq j\leq q-1$ set $a(j):=[\frac{n-j}{q}]$. Then $\bigvee_{i=0}^{n-1}T^{-j}[\mathcal{C}]=\bigvee_{r=0}^{a(j)-1}T^{-rq+j}\bigvee_{i=0}^{q-1}T^{-i}[\mathcal{C}]\vee\bigvee_{l\in R}T^{-l}[\mathcal{C}]$, where $\#R\leq 2q$. Then,

$$\begin{aligned}
\log\sum_{|\underline{w}|=n}e^{S_n(\theta_{\underline{w}})}=&H_{\widetilde{\nu}_n}\Big(\bigvee_{i=0}^{n-1}T^{-j}[\mathcal{C}]\Big)+\int S_n(\omega)d\widetilde{\nu}_n(\omega)\\
\leq&\sum_{r=0}^{a(j)-1}H_{\widetilde{\nu}_n}\Big(T^{-rq+j}\bigvee_{i=0}^{q-1}T^{-i}[\mathcal{C}]\Big)\\
&+H_{\widetilde{\nu}_n}\Big(\bigvee_{l\in R}T^{-l}[\mathcal{C}]\Big)+\int S_n d\widetilde{\nu}_n\\
\leq&\sum_{r=0}^{a(j)-1}H_{\widetilde{\nu}_n\circ T^{-(rq+j)}}\Big(\bigvee_{i=0}^{q-1}T^{-i}[\mathcal{C}]\Big)+2q\log\#\mathcal{C}+\int S_n d\widetilde{\nu}_n.
\end{aligned}\tag{4}$$

Summing (4) over $0 \le j \le q-1$ and dividing by n, we get

$$\begin{aligned}\frac{q}{n}\log\sum_{|\underline{w}|=n} e^{S_n(\theta_{\underline{w}})} \le& \frac{1}{n}\sum_{p=0}^{n-1} H_{\widetilde{\nu}_n\circ T^{-p}}\Big(\bigvee_{i=0}^{q-1} T^{-i}[\mathcal{C}]\Big) + \frac{2q^2}{n}\log\#\mathcal{C} + q\int \frac{1}{n}S_n d\widetilde{\nu}_n \\ \le& H_{\nu_n}\Big(\bigvee_{i=0}^{q-1} T^{-i}[\mathcal{C}]\Big) + \frac{2q^2}{n}\log\#\mathcal{C} + q\int \frac{1}{n}S_n d\widetilde{\nu}_n \\ =& q\cdot\Big(\frac{1}{q}H_{\nu_n}\Big(\bigvee_{i=0}^{q-1} T^{-i}[\mathcal{C}]\Big) + \frac{2q}{n}\log\#\mathcal{C} \\ &+ \int_{\Sigma\times T^1M} \widehat{\phi}\, d\big(\frac{1}{n}\sum_{i=0}^{n-1}\widehat{\nu}_n\circ \widehat{F}^{-k}\big)\Big), \end{aligned} \tag{5}$$

where $\widehat{\nu}_n := \widetilde{\nu}_n \times \delta_{t_n}$.

Let $\pi : \Sigma\times X \to \Sigma$ be the projection onto the first coordinate, then π is continuous and $\widehat{\nu}_n\circ\pi^{-1} = \widetilde{\nu}_n$. Moreover, $\pi\circ\widehat{F} = T\circ\pi$, and hence

$$\big(\frac{1}{n}\sum_{i=0}^{n-1}\widehat{\nu}_n\circ\widehat{F}^{-k}\big)\circ\pi^{-1} = \nu_n. \tag{6}$$

We may assume w.l.o.g that $\frac{1}{n_j}\sum_{i=0}^{n_j-1}\widehat{\nu}_n\circ\widehat{F}^{-k} \xrightarrow[j\to\infty]{} \widehat{\nu}$, and so

$$\widehat{\nu}\circ\pi^{-1} = \nu. \tag{7}$$

Dividing by q, sending $n_j\to\infty$, and then sending $q\to\infty$, on both sides of (5), we conclude that for all $n\in\mathbb{N}$,

$$\begin{aligned}\check{P}(\{\phi_t\}_t) \le& h_\nu(T) + \int\widehat{\phi}\,d\widehat{\nu} = h_\nu(T) + \int\frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}\,d\widehat{\nu}\ (\because \widehat{\nu} = \widehat{\nu}\circ\widehat{F}^{-1}) \\ \le& h_\nu(T) + \int\sup_t\Big\{\frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}(\omega,t)\Big\}d\widehat{\nu},\end{aligned} \tag{8}$$

and in fact since n is arbitrary,

$$\begin{aligned}\check{P}(\{\phi_t\}_t) \le& h_\nu(T) + \limsup_n\int\sup_t\Big\{\frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}(\omega,t)\Big\}d\widehat{\nu} \qquad (9)\\ \le& h_\nu(T) + \int\limsup\sup_t\Big\{\frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}(\omega,t)\Big\}d\widehat{\nu}\ (\because \text{Fatou lemma}) \\ =& h_\nu(T) + \int\widetilde{\phi}\,d\widehat{\nu} = h_\nu(T) + \int\widetilde{\phi}\circ\pi\,d\widehat{\nu} \\ =& h_\nu(T) + \int\widetilde{\phi}\,d(\widehat{\nu}\circ\pi^{-1}) = h_\nu(T) + \int\widetilde{\phi}\,d\nu.\end{aligned}$$

□

Proposition 3.4.

$$\max\left\{h_{\widehat{\nu}\circ\pi^{-1}}(T)+\int\widehat{\phi}d\widehat{\nu}:\widehat{\nu}\ \mathit{erg.}\ \widehat{F}\text{-}\mathit{inv.}\right\}=\max\left\{h_\nu(T)+\int\widetilde{\phi}d\nu:\nu\ \mathit{erg.}\ T\text{-}\mathit{inv.}\right\},$$

where $\pi:\Sigma\times X\to\Sigma$ is the projection onto Σ. In particular the maximum is attained.

Proof. Given an $\widehat{F}$-invariant probability $\widehat{\nu}$, write $\nu:=\widehat{\nu}\circ\pi^{-1}$. Then,

$$\begin{aligned}\int\widehat{\phi}d\widehat{\nu}&=\limsup_n\int\widehat{\phi}d\widehat{\nu}=\limsup_n\int\frac{1}{n}\sum_{j=0}^{n-1}\widehat{\phi}\circ\widehat{F}^j(\omega,t)d\widehat{\nu}(\omega,t) \qquad (10)\\&\le\limsup_n\int\sup_t\frac{1}{n}\sum_{j=0}^{n-1}\widehat{\phi}\circ\widehat{F}^j(\omega,t)d\widehat{\nu}(\omega,t)\\&=\limsup_n\int\sup_t\frac{1}{n}\sum_{j=0}^{n-1}\widehat{\phi}\circ\widehat{F}^j(\omega,t)d\widehat{\nu}\circ\pi^{-1}(\omega)\\&\le\int\limsup_n\sup_t\frac{1}{n}\sum_{j=0}^{n-1}\widehat{\phi}\circ\widehat{F}^j(\omega,t)d\widehat{\nu}\circ\pi^{-1}(\omega)\\&=\int\widetilde{\phi}d\nu\ (\because\text{ Fatou's lemma}).\end{aligned}$$

Hence it is clear that,

$$\max\left\{h_{\widehat{\nu}\circ\pi^{-1}}(T)+\int\widehat{\phi}d\widehat{\nu}:\widehat{\nu}\text{ erg. }\widehat{F}\text{-inv.}\right\}\le\sup\left\{h_\nu(T)+\int\widetilde{\phi}d\nu:\nu\text{ erg. }T\text{-inv.}\right\},$$

where the maximum on l.h.s exists since π is continuous and Σ is entropy expansive.

We continue to show the other inequality. Let ν_n s.t $P_{\nu_n}(\widetilde{\phi})\to\sup_{\nu'}\left\{h_{\nu'}(T)+\int\widetilde{\phi}d\nu'\right\}$. Given $\omega\in\Sigma$, set $n_j(\omega)$ to be the first integer larger than $n_{j-1}(\omega)$ s.t $\frac{1}{n_j(\omega)}\sum_{k=0}^{n_j(\omega)-1}\widehat{\phi}\circ\widehat{F}^k(\omega,t_{n_j(\omega)}(\omega))\ge\limsup\sup_t\frac{1}{n}\sum_{k=0}^{n-1}\widehat{\phi}\circ\widehat{F}^k(\omega,t)-\frac{1}{j}$.

Then,

$$\begin{aligned}
\sup_{\nu'}\left\{h_{\nu'}(T)+\int\widetilde{\phi}d\nu'\right\} &= \lim_n\left(h_{\nu_n}(T)+\int\widetilde{\phi}d\nu_n\right)\\
&= \lim_n\Bigg(h_{\nu_n}(T)\\
&\qquad+\int\lim_j\frac{1}{n_j(\omega)}\sum_{k=0}^{n_j(\omega)-1}\widehat{\phi}\circ\widehat{F}^k(\omega,t_{n_j(\omega)}(\omega))d\nu_n\Bigg)\\
(\because\text{ dominated convergence}) &= \lim_n\Bigg(h_{\nu_n}(T)\\
&\qquad+\lim_j\int\frac{1}{n_j(\omega)}\sum_{k=0}^{n_j(\omega)-1}\widehat{\phi}\circ\widehat{F}^k(\omega,t_{n_j(\omega)}(\omega))d\nu_n\Bigg)\\
&= \lim_n\left(h_{\nu_n}(T)+\lim_j\int\widehat{\phi}d\widehat{\nu}_{n,j}\right), \qquad (11)
\end{aligned}$$

where $\widehat{\nu}_{n,j} := \int\frac{1}{n_j(\omega)}\sum_{k=0}^{n_j(\omega)-1}(\delta_\omega\times\delta_{t_{n_j(\omega)}})\circ\widehat{F}^{-k}d\nu_n(\omega)$. Assume w.l.o.g that $\widehat{\nu}_{n,j}\to\widehat{\nu}_n\to\widehat{\nu}$, and that $\nu_n\to\nu$.

Claim: $\widehat{\nu}_n\circ\pi^{-1}=\nu_n$ and $\widehat{\nu}\circ\pi^{-1}=\nu$.

Proof: Let $g\in C(\Sigma)$, then by the point-wise ergodic theorem and dominated convergence,

$$\int g\circ\pi d\widehat{\nu}_{n,j}=\int\frac{1}{n_j(\omega)}\sum_{k=0}^{n_j(\omega)-1}g\circ T^k d\nu_n\xrightarrow[j\to\infty]{}\int gd\nu_n.$$

Hence $\widehat{\nu}_n\circ\pi^{-1}=\nu_n$. Therefore, by the continuity of π, $\widehat{\nu}\circ\pi^{-1}=\nu$. QED

Since (Σ,T) is entropy expansive, $\limsup h_{\nu_n}(T)\geq h_\nu(T)$. Plugging this back in (11), we get

$$\sup_{\nu'}\left\{h_{\nu'}(T)+\int\widetilde{\phi}d\nu'\right\}\leq h_{\widehat{\nu}\circ\pi^{-1}}(T)+\int\widehat{\phi}d\widehat{\nu}. \qquad (12)$$

Then $\sup_{\nu'}\left\{h_{\nu'}(T)+\int\widetilde{\phi}d\nu'\right\}=\max_{\widehat{\nu}'}\left\{h_{\widehat{\nu}'\circ\pi^{-1}}(T)+\int\widehat{\phi}d\widehat{\nu}'\right\}$. We now continue to prove that the supremum is attained over ν. By (12) and (10), since $\widehat{\nu}$ is $\widehat{F}$-invariant, and since $\widehat{\nu}\circ\pi^{-1}=\nu$,

$$\sup_{\nu'}\left\{h_{\nu'}(T)+\int\widetilde{\phi}d\nu'\right\}\leq h_\nu(T)+\int\widetilde{\phi}d\nu.$$

□

Remark: Proposition 3.4 implies that the maximum of $\max_\nu\left\{h_\nu(T)+\int\widetilde{\phi}d\nu\right\}$ is indeed attained, even though $\widetilde{\phi}$ is not continuous. Even when $\{\phi_t\}_t=\{\phi\}$, $\widetilde{\phi}(\omega)$ still depends on the empirical measure of ω. See also Lemma 3.5 and the remark following it.

3.2. **Lower bound.** The following proposition is our main statement, as it effectively allows us to carry out the non-trivial operation of taking the supremum of $t \in X$ to be after taking the limit on n in almost all estimates.

Lemma 3.5. $\widetilde{\phi} \circ T = \widetilde{\phi}$.

Proof. For all $\omega \in \Sigma$, $F_\omega : X \to X$ is a homeomorphism, hence

$$\begin{aligned}\widetilde{\phi}(\omega) &= \limsup_n \sup_{t \in X} \frac{1}{n} \sum_{k=0}^{n-1} \widehat{\phi}(T^k\omega, F_\omega^k(t)) \\ &= \limsup_n \Big(\sup_{F_\omega(t) \in X} \frac{1}{n} \sum_{k=0}^{n-2} \widehat{\phi}(T^k T\omega, F_{T\omega}^k(F_\omega(t))) + \frac{\widehat{\phi}(\omega, F_\omega^{-1}(F_\omega(t)))}{n} \Big) \\ &= \limsup_n \sup_{F_\omega(t) \in X} \frac{1}{n} \sum_{k=0}^{n-1} \widehat{\phi}(T^k T\omega, F_{T\omega}^k(F_\omega(t))) = \widetilde{\phi} \circ T(\omega). \end{aligned} \tag{13}$$

□

Proposition 3.6 (Main Proposition)**.** *For all $\omega \in \Sigma$ there exists a closed $\mathcal{L}_\omega \subseteq \mathcal{P}(X)$ s.t $\forall \sigma \in \mathcal{L}_\omega$,*

$$\widetilde{\phi}(\omega) = \int \widehat{\phi}(\omega, t) d\sigma,$$

and $\mathcal{L}_{T\omega} = \mathcal{L}_\omega$. In particular, given an ergodic T-invariant probability ν on Σ, there exists $\sigma_\nu \in \mathcal{P}(X)$ s.t

$$\sigma_\nu \in \mathcal{L}_\omega \text{ for } \nu\text{-a.e } \omega \in \Sigma.$$

Proof. Given $\omega \in \Sigma$, let $n_j(\omega) \uparrow \infty$ and $t_j(\omega) \in X$ s.t

(I) $\lim_{j\to\infty} \sup_{t\in X} \frac{1}{n_j} \sum_{k=0}^{n_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) = \widetilde{\phi}(\omega)$,

(II) $\frac{1}{n_j} \sum_{k=0}^{n_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t_j(\omega))) - \sup_{t\in X} \frac{1}{n_j} \sum_{k=0}^{n_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) \xrightarrow[j\to\infty]{} 0$,

(III) $\frac{1}{n_j(\omega)} \sum_{k=0}^{n_j(\omega)-1} \delta_{F_\omega^k(t_j(\omega))}$ has a weak-* limit in $\mathcal{P}(X)$.

In total, there exists $\sigma_\omega \in \mathcal{P}(X)$ s.t

$$\begin{aligned} \widetilde{\phi}(\omega) &= \lim_j \frac{1}{n_j(\omega)} \sum_{k=0}^{n_j(\omega)-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t_j(\omega))), \\ \text{and } &\lim_j \frac{1}{n_j(\omega)} \sum_{k=0}^{n_j(\omega)-1} \delta_{F_\omega^k(t_j(\omega))} = \sigma_\omega, \end{aligned} \tag{14}$$

where σ_ω is a probability measure on X. Then,

$$\begin{aligned}\widetilde{\phi}(\omega) &= \lim_j \frac{1}{n_j(\omega)} \sum_{k=0}^{n_j(\omega)-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t_j(\omega))) \\ &= \lim_j \int \widehat{\phi} d(\frac{1}{n_j(\omega)} \sum_{k=0}^{n_j(\omega)-1} \delta_{F^k_\omega(t_j(\omega))}) = \int \widehat{\phi}(\omega, t) d\sigma_\omega(t). \qquad (15)\end{aligned}$$

Denote by $\mathcal{L}_\omega = \{\sigma_\omega :$ achieved by n_j and t_j satisfying (I),(II),(III) above$\}$. By (13), for any admissible subsequence $n_j(\omega)$, we can choose $n_j(T\omega) = n_j(\omega) - 1$. Moreover, for all $t \in X$,

$$\frac{1}{n_j} \sum_{k=0}^{n_j-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) = \frac{1}{n_j - 1} \sum_{k=0}^{n_j-1-1} \widehat{\phi}(T\omega, F_\omega(t)) + o_j(1).$$

Hence, we can choose $t_j(T\omega) = F_\omega(t_j(\omega))$. By the definition of σ_ω in (14),

$$\begin{aligned}\sigma_{T_\omega} &= \lim_j \frac{1}{n_j - 1} \sum_{k=0}^{n_j-1-1} \delta_{F^k_{T\omega}(t_j(T\omega))} = \lim_j \frac{1}{n_j - 1} \sum_{k=0}^{n_j-1-1} \delta_{F^k_{T\omega}(F_\omega(t_j(\omega)))} \\ &= \lim_j \frac{1}{n_j - 1} \sum_{k=0}^{n_j-1-1} \delta_{F^k_{T\omega}(F_\omega(t_j(\omega)))} = \lim_j \frac{1}{n_j - 1} \sum_{k=0}^{n_j-1-1} \delta_{F^{k+1}_\omega(t_j(\omega))} = \sigma_\omega.\end{aligned}$$

This provides us with a one-to-one matching between all elements of $\mathcal{L}_\omega$ and those of $\mathcal{L}_{T\omega}$.

To see that $\mathcal{L}_\omega$ is closed, let $\sigma_i \in \mathcal{L}_\omega$ s.t $\sigma_i \to \sigma$. For each σ_n, let $\{n^i_j\}$ s.t $\widehat{\phi}(\omega) = \lim_j \sup_t \frac{1}{n^i_j} \sum_{k=0}^{n^i_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t))$ and t^i_j s.t

$$\sup_t \frac{1}{n^i_j} \sum_{k=0}^{n^i_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) - \frac{1}{n^i_j} \sum_{k=0}^{n^i_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t^i_j)) \xrightarrow[j\to\infty]{} 0.$$

Then consider $\frac{1}{n^i_{j_i}} \sum_{k=0}^{n^i_{j_i}-1} \delta_{F^k_\omega(t^i_{j_i})}$, where $j_i \uparrow \infty$ is a sequence of j's so

(1) $d(\frac{1}{n^i_{j_i}} \sum_{k=0}^{n^i_{j_i}-1} \delta_{F^k_\omega(t^i_{j_i})}, \sigma_i) \leq \frac{1}{i}$,

(2) $\left| \sup\limits_t \frac{1}{n^i_{j_i}} \sum\limits_{k=0}^{n^i_{j_i}-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) - \limsup \sup\limits_t \frac{1}{n} \sum\limits_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) \right| \leq \frac{1}{i}$,

(3) $\left| \sup_t \frac{1}{n^i_j} \sum_{k=0}^{n^i_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t)) - \frac{1}{n^i_j} \sum_{k=0}^{n^i_j-1} \widehat{\phi} \circ \widehat{F}^k((\omega, t^i_j)) \right| \leq \frac{1}{i}$,

(4) $\frac{1}{n^i_{j_i}} \sum_{k=0}^{n^i_{j_i}-1} \delta_{F^k_\omega(t^i_{j_i})} \xrightarrow[i\to\infty]{} \sigma'$.

Then σ' must coincide with σ, and $\sigma' \in \mathcal{L}_\omega$.

We continue to prove that given an erg. T-inv. ν, $\exists \sigma_\nu \in \mathcal{P}(X)$ s.t $\sigma_\nu \in \mathcal{L}_\omega$ for ν-a.e ω.

Let $i \in \mathbb{N}$, and since $\mathcal{P}(X)$ is a metric separable space (hence Lindelöf), we have a countable cover $\{B(\sigma_j^i, \frac{1}{2^i})\}_{j\in\mathbb{N}}$. Define $E_{\sigma_j^i} := \{\omega : B(\sigma_j^i, \frac{1}{2^i}) \cap \mathcal{L}_\omega \neq \varnothing\}$, and note that $E_{\sigma_j^i} = T^{-1}[E_{\sigma_j^i}]$ since $\mathcal{L}_{T\omega} = \mathcal{L}_\omega$ for all ω. Then there exists $\sigma^i_{j_i(\nu)}$ s.t $\nu(E_{\sigma^i_{j_i(\nu)}}) = 1$.

We cover $B(\sigma^i_{j_i(\nu)}, \frac{1}{2^i})$ by balls of radius $\frac{1}{2^{i+1}}$ intersected with $B(\sigma^i_{j_i(\nu)}, \frac{1}{2^i})$, and continue this way, and we get that $\exists\{\sigma^i_{j_i(\nu)}\}_{i\geq 0}$ s.t $\exists!\sigma_\nu \bigcap_{i\geq 0} \overline{B(\sigma^i_{j_i(\nu)}, \frac{1}{2^i})}$ while for any $\omega \in \bigcap_{i\geq 0} E_{\sigma^i_{j_i(\nu)}}$ (which is a full ν-measure set), and for all $i \geq 0$, $d(\sigma^i_{j_i(\nu)}, \mathcal{L}_\omega) \leq \frac{1}{2^i}$. Since $\mathcal{L}_\omega$ is closed, $\sigma_\nu \in \mathcal{L}_\omega$, and we are done. $\square$

Definition 3.7. *The invariant family of probability measures $\mathcal{L}_\omega$ given by Proposition 3.6 is called the* instability kernel *of ω.*

Theorem 3.8 (Variational Principle)**.**

$$\check{P}(\{\phi_t\}_t) = \max_{\nu \text{ erg. inv. prob.}} \Big\{ h_\nu(T) + \int \widetilde{\phi}(\omega) d\nu(\omega) \Big\}.$$

Proof. We already showed the upper bound in Theorem 3.3, and by Proposition 3.4 we know that the maximum is attained. We are only left to show the lower bound. Let ν be an ergodic T-invariant probability measure.

Let σ_ν be the probability measure given by Proposition 3.6. By (15), for all k, $\int \widetilde{\phi}(\omega) d\nu(\omega) = \int \widehat{\phi}(T^k\omega, F^k_\omega(t)) d\sigma_\nu(t) d\nu(\omega)$. Then,

$$\begin{aligned}
&\exp\Big(n(h_\nu(T) + \int \widetilde{\phi} d\nu)\Big) \\
&= \exp\Big(n(h_\nu(T) + \int\int \frac{1}{n}\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) d\sigma_\nu(t) d\nu(\omega))\Big) \\
(\because \text{Fubini}) &= \exp\Big(n(h_\nu(T) + \int\int \frac{1}{n}\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) d\nu(\omega) d\sigma_\nu(t))\Big) \\
(\because \text{Jensen}) &\leq \int \exp\Big(n(h_\nu(T) + \int \frac{1}{n}\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) d\nu(\omega))\Big) d\sigma_\nu(t) \\
&\leq \sup_t \exp\Big(n(h_\nu(T) + \int \frac{1}{n}\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) d\nu(\omega))\Big).
\end{aligned}$$

Let $\epsilon > 0$ and let $n_\epsilon \in \mathbb{N}$ s.t $\nu(K_\epsilon) \geq 1 - \epsilon$, where

$$K_\epsilon := \Big\{\omega \in \Sigma : \forall n \geq n_\epsilon, \nu([\omega_0, \dots, \omega_{n-1}]) = e^{nh_\nu(T) \pm \epsilon n}\Big\}.$$

Then,for all $n \geq n_\epsilon$,

$$\begin{aligned}
&\exp\Big(n(h_\nu(T) + \int \widetilde{\phi} d\nu)\Big) \\
&\leq \sup_t \exp\Big(n(h_\nu(T) + \int_{K_\epsilon} \frac{1}{n}\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\omega, t) d\nu(\omega))\Big) e^{\epsilon\|\widehat{\phi}\|_\infty n} \\
&\leq \sup_t \frac{e^{\epsilon\|\widehat{\phi}\|_\infty n}}{\nu(K_\epsilon)} \int\limits_{K_\epsilon} \exp\Big(nh_\nu(T) + \sum_{k=0}^{n-1} \nu(K_\epsilon)\widehat{\phi} \circ \widehat{F}^k(\omega, t)\Big) d\nu(\omega) \\
&\quad (\because \text{Jensen}) \\
&\leq e^{\epsilon\|\widehat{\phi}\|_\infty n} \sup_t \sum_{\substack{|\underline{w}|=n: \\ [\underline{w}]\cap K_\epsilon \neq \varnothing}} \nu([\underline{w}]) e^{nh_\nu(T)} e^{\sum_{k=0}^{n-1} \nu(K_\epsilon)\widehat{\phi}\circ\widehat{F}^k(\theta_{\underline{w}}\omega, t) + C_{\{\phi_t\}_{t\in X}}} \\
&\leq e^{2\epsilon\|\widehat{\phi}\|_\infty n + C_{\{\phi_t\}_{t\in X}}} \sup_t \sum_{\substack{|\underline{w}|=n: \\ [\underline{w}]\cap K_\epsilon \neq \varnothing}} e^{-nh_\nu(T)+\epsilon n} e^{nh_\nu(T)} e^{\sum_{k=0}^{n-1} \widehat{\phi}\circ\widehat{F}^k(\theta_{\underline{w}}\omega, t)} \\
&\leq e^{2\epsilon(\|\widehat{\phi}\|_\infty+1)n + C_{\{\phi_t\}_{t\in X}}} \check{Z}_n(\{\phi_t\}_{t\in X}),
\end{aligned}$$

where $Z_n(\{\phi_t\}_t) := \sup_t \sum_{|\underline{w}|=n} e^{\sum_{k=0}^{n-1} \phi_{F^k_{\theta_{\underline{w}}}(t)} \circ T^k(\theta_{\underline{w}})}$ and $\theta_{\underline{w}} \in [\underline{w}]$ maximizes $\sum_{k=0}^{n-1} \widehat{\phi} \circ \widehat{F}^k(\theta_{\underline{w}}, t)$.

Therefore, $h_\nu(T) + \int \widetilde{\phi} d\nu \leq \check{P}(\{\phi_t\}_{t\in X}) + \epsilon(1 + 2\|\widehat{\phi}\|_\infty)$ for all $\epsilon > 0$. Then by sending $\epsilon \to 0$, we are done. □

4. Hyperbolicity estimates for a uniformly expanding on average Gibbs process

In this section we apply the results of §3 in the setting of Gibbs processes.

Setup:

(1) Let M be a closed Riemannian manifold of dimension $d \geq 2$.
(2) Let $f : \Sigma \to \mathrm{Diff}^{1+\alpha}(M)$ be a Hölder continuous map, denoted by $\omega \mapsto f_\omega$.
(3) Let μ be a T-invariant probability measure on Σ, and assume further that μ is a Gibbs measure for the potential $\psi \in \mathrm{Höl}(\Sigma)$.
(4) Assume w.l.o.g that $P(\psi) = 0$, where $P(\cdot)$ denotes the Gurevich pressure (otherwise replaces ψ by $\psi - P(\psi)$ which does not change μ).
(5) Given $t = (x, \xi) \in T^1M$, write $\varphi_t(\omega) := \log|d_x f_\omega \xi|$, and note that $\sup_{t\in T^1M} \|\varphi_t\|_{\mathrm{Höl}} < \infty$.
(6) Assume that (M, f, μ) admits *uniform expansion on average*: $\exists \chi > 0$ s.t

$$\inf_{t\in T^1M} \int_\Sigma \varphi_t(\omega) d\mu(\omega) \geq \chi.$$

Remark:

(1) Note that items (4) and (5) above hold when μ is an SRB measure on an Anosov system. This requires reducing the case of a two-sided shift to a one-sided shift, see Lemma A.1.
(2) We do not require $\omega \mapsto f_\omega$ to be constant on partition sets.
(3) The random dynamical system $f_\omega \sim \mu$ is called a *Gibbs process*. It naturally extends Bernoulli processes and Markov processes.

Gibbs processes translate naturally to the setting of thermodynamic formalism out of equilibrium (recall §2):

(1) $X := T^1M$,
(2) $F_\omega(x,\xi) := (f_\omega(x), \frac{d_x f_\omega \xi}{|d_x f_\omega \xi|})$.

We wish to prove that there exist $\beta, \gamma > 0$ s.t for all n sufficiently large, for all $(x,\xi) \in T^1M$,

$$\int_\Sigma |d_x f_\omega^n \xi|^{-\beta} d\mu(\omega) \leq e^{-\gamma n}. \tag{16}$$

Remark: By the Markov inequality, (16) implies that for all $(x,\xi) \in T^1M$, every word $\omega \in \Sigma$ expands uniformly, aside for an exponentially small exceptional set of words. Such estimates have proved themselves to be the fundamental estimates which are used in the study of exponential mixing of random dynamical system (see for example [DD]). Previously, such estimates were shown for i.i.d dynamical systems (i.e μ is a Bernoulli measure, and $\omega \mapsto f_\omega$ is constant on partition sets). However, the physical motivation suggests that the randomness which drives our dynamics should be given by a Gibbs measure, hence motivating our extension.

4.1. **Reduction to thermodynamic formalism out of equilibrium.**

Theorem 4.1. *Assume that (Σ, T) is topologically transitive. Then, for all $\beta \in (0,1)$, for all $n \in \mathbb{N}$ and all $t = (x,\xi) \in T^1M$,*

$$\int |d_x f_\omega^n \xi|^{-\beta} d\mu = \widehat{C}_{\psi,\varphi}^{\pm 1} \cdot \check{Z}_n(\phi_t, a),$$

where, $\phi_t := \psi + \beta\varphi_t$, $\widehat{C}_{\psi,\varphi} \geq 1$ is a global constant and $[a] \subset \Sigma$.

Proof. Let $n \in \mathbb{N}$ and $t = (x,\xi) \in T^1M$. For all $[a] \subset \Sigma$ fix some $\omega_a \in [a]$, and let $\theta_{\underline{w}} \in [\underline{w}]$ be some element of the cylinder. Set $\varphi_t^{(n)}(\omega) := \sum_{k=0}^{n-1} \varphi_{F_\omega^k(t)}(T^k\omega)$, $\psi^{(n)} = \sum_{k=0}^{n-1} \psi \circ T^k$, and $\phi_t^{(n)}(\omega) := \sum_{k=0}^{n-1} \phi_{F_\omega^k(t)}(T^k\omega)$. Then,

$$\begin{aligned}\int |d_x f_\omega^n \xi|^{-\beta} d\mu = \int e^{-\beta\varphi_t^{(n)}(\omega)} d\mu = C_\varphi^{\pm 1} \sum_{|\underline{w}|=n} \mu([\underline{w}]) e^{-\beta\varphi_t^{(n)}(\theta_{\underline{w}})} \\ =(C_\psi C_\varphi)^{\pm 1} \sum_{|\underline{w}|=n} e^{\psi^{(n)}(\theta_{\underline{w}}) - \beta\varphi_t^{(n)}(\theta_{\underline{w}})} = (C_\psi C_\varphi)^{\pm 1} \sum_{|\underline{w}|=n} e^{\phi_t^{(n)}(\theta_{\underline{w}})} \\ =(C_\psi C_\varphi)^{\pm 1} \sum_{[a]\subset\Sigma} C_a^{\pm 1} \sum_{|\underline{w}|=n, w_{n-1}=a} e^{\phi_t^{(n)}(\underline{w}\cdot\omega_a)},\end{aligned} \tag{17}$$

where $\omega_a \in [a]$ is any element of $[a]$, and $C_a := N_a \cdot \sup_t \|\phi_t\|_\infty$ where $\forall[b] \subset \Sigma$, $a \xrightarrow{n_{ab}} b \xrightarrow{n_{ba}} a$ with $n_{ab}, n_{ba} \leq N_a$.

$\square$

Remark:

(1) The proof Theorem 4.1 in fact implies the following more general statement:
$$\forall t \in X, \int e^{\beta \sum_{k=0}^{n-1} \varphi_{F_\omega^k(t)} \circ T^k(\omega)} d\mu_\psi(\omega) = \widehat{C}_{\psi,\varphi}^{\pm 1} \cdot \check{Z}_n(\psi + \beta\varphi_t, a).$$

(2) Theorem 4.1 demonstrates that our definition of the pressure out of equilibrium (recall §2) is in fact optimal. If one wishes to gain hyperbolicity estimates such as in (16), then the POE is optimal quantity to study.

Corollary 4.2.
$$\limsup \frac{1}{n} \log \sup_t \int |d_x f_\omega^n \xi|^{-\beta} d\mu \leq \check{P}(\{\phi_t\}_t).$$

Remark: By Corollary 4.2, proving that the POE is negative for all $\beta > 0$ sufficiently small, would prove (16).

Our plan to prove (16) is composed of three steps:

(1) Reducing the expression of (16) to an estimate of the POE (§4.1).
(2) Providing a bound to the POE in terms of topological pressure (§3, §4.3).
(3) Perturbative theory of Ruelle operators to bound the topological pressure (§4.2).

We start by recalling the perturbative theory of Ruelle operators, and gaining the desired estimates for the spectral radius.

4.2. **Perturbative theory of Ruelle operators.**

Setup:

(1) For $\beta > 0$, $\phi_t := \psi + \beta\varphi_t$.
(2) $L_{\phi_t} : C_c(\Sigma) \to C_c(\Sigma)$ is the associated *Ruelle operator* defined by
$$(L_{\phi_t} h)(\omega) := \sum_{T\omega'=\omega} e^{\phi_t(\omega')} h(\omega').$$
(3) By assumption μ is a Gibbs measure of ψ, hence $\mu = h_\psi \cdot p_\psi$ where $L_\psi h_\psi = e^{P(\psi)} h_\psi$ and $L_\psi^* p_\psi = e^{P(\psi)} p_\psi$. Recall that $P(\psi) = 0$. Then one can check that $\exists C_\psi > 1$ s.t for every word $|\underline{w}| = n$,
$$\mu([\underline{w}]) = C_\psi^{\pm 1} e^{\psi^{(n)}(\theta_{\underline{w}})}, \text{ for any } \theta_{\underline{w}} \in [\underline{w}].$$

Theorem 4.3 (Aaronson & Denker)**.** *Ruelle operators of Hölder continuous potentials on compact TMSs are quasi-compact.*

Corollary 4.4. *$L_{\phi_t} = P_t + N_t$ where*
1. *$P_t h := e^{P(\phi_t)} h_{\phi_t} \cdot \int h d\nu_{\phi_t}$,*
2. *$N_t P_t = P_t N_t = 0$,*
3. *$\|N_t^n\| = O(\rho^n)$, with $\rho \in (0, e^{P(\phi_t)})$.*

See [Sar09, § 5.2] for details.

Theorem 4.5. *For all $\beta > 0$ sufficiently small, for all $t \in T^1 M$, $P(\phi_t) \leq -\frac{\chi}{2}\beta$.*

Proof. Note, $\beta \mapsto P(\psi - \beta\varphi_t)$ is analytic in a neighborhood of 0.

By the mean value theorem, for all $\beta_0 > 0$ sufficiently small, there exist $\xi_0 \in (0, \beta_0)$ and $\xi_1 \in (0, \xi_0)$ s.t

$$\begin{aligned}
P(\psi - \beta_0\varphi_t) =& P(\psi) - \beta_0 \frac{d}{d\beta}\Big|_{\beta=\xi_0} P(\psi - \beta\varphi_t) = -\beta_0 \frac{d}{d\beta}\Big|_{\beta=\xi_0} P(\psi - \beta\varphi_t) \\
=& -\beta_0 \frac{d}{d\beta}\Big|_{\beta=0} P(\psi - \beta\varphi_t) \\
& -\beta_0 \cdot \Big(\frac{d}{d\beta}\Big|_{\beta=\xi_0} P(\psi - \beta\varphi_t) - \frac{d}{d\beta}\Big|_{\beta=0} P(\psi - \beta\varphi_t) \Big) \\
=& -\beta_0 \frac{d}{d\beta}\Big|_{\beta=0} P(\psi - \beta\varphi_t) - \beta_0 \cdot \xi_0 \cdot \frac{d^2}{d\beta^2}\Big|_{\beta=\xi_1} P(\psi - \beta\varphi_t).
\end{aligned}$$

By the linear response formula,

$$\frac{d}{d\beta}\Big|_{\beta=0} P(\psi - \beta\varphi_t) = -\int \varphi_t d\mu \leq -\chi,$$

so

$$\begin{aligned}
P(\psi - \beta_0\varphi_t) \leq& -\chi\beta_0 - \beta_0 \cdot \xi_0 \cdot \frac{d^2}{d\beta^2}\Big|_{\beta=\xi_1} P(\psi - \beta\varphi_t) \\
\leq& -\chi\beta_0 + \beta_0^2 \cdot \Big| \frac{d^2}{d\beta^2}\Big|_{\beta=\xi_1} P(\psi - \beta\varphi_t) \Big|. \qquad (18)
\end{aligned}$$

By the Green-Kubo formula, and since $\frac{d^2}{dx^2}\Big|_{x=a} f(x) = \frac{d^2}{dx^2}\Big|_{x=0} f(x+a)$,

$$\begin{aligned}
\frac{d^2}{d\beta^2}\Big|_{\beta=\xi_1} P(\psi - \beta\varphi_t) =& \frac{d^2}{d\beta^2}\Big|_{\beta=0} P((\psi - \xi_1\varphi_t) - \beta\varphi_t) \qquad (19) \\
=& \mathrm{Var}_{\mu_{\psi-\xi_1\varphi_t}}(\overline{\varphi_t}) + 2\sum_{k\geq 1} \mathrm{Cov}_{\mu_{\psi-\xi_1\varphi_t}}(\overline{\varphi_t}, \overline{\varphi_t} \circ T^k),
\end{aligned}$$

where $\overline{\varphi_t} := \varphi_t - \int \varphi_t d\mu_{\psi-\xi_1\varphi_t}$.

Since

$$\sup_t \|\overline{\varphi_t}\|_{\text{Höl}} < \infty,$$

we get that for all $\beta_0 > 0$ sufficiently small the r.h.s of (19) is bounded uniformly in t, as it depends only on the rate of mixing of the operator $e^{-P(\psi-\xi_1\varphi_t)} L_{\psi-\xi_1\varphi_t}$ which is a perturbation of L_ψ of multiplicative size

$e^{2\xi_1 \cdot \sup_t \|\varphi_t\|_\infty}$, where $\xi_1 \le \beta_0$. Therefore, there exists $K_\psi > 0$ s.t

$$\sup_t \left| \frac{d^2}{d\beta^2}\Big|_{\beta=\xi_1} P(\psi - \beta\varphi_t) \right| \le K_\psi,$$

whenever β_0 is smaller than half the spectral gap of L_ψ divided by $\sup_t \|\varphi_t\|_\infty$. Plugging this back in (18), we get that for all $\beta_0 > 0$ sufficiently small,

$$P(\psi - \beta_0\varphi_t) \le -\chi\beta_0 + \beta_0^2 K_\psi \le -\frac{\chi}{2}\beta_0.$$

□

Remark: Theorem 4.5 is essential, as linear algebra (or functional analysis) can estimate the difference between L_ψ and L_{ϕ_t}, and consequently $|P(\psi) - P(\phi_t)|$, but much finer analysis which relies on additional information is needed in order to know the sign of $P(\psi) - P(\phi_t)$, including a bound from below. This makes Theorem 4.5 independent from the rest of the techniques we use, and crucial. Its proof is where we use the assumption of uniform expansion on average.

4.3. **Bounding the POE by the topological pressure.** Theorem 4.6 below relies on the variational principle proved in §3.

Theorem 4.6.

$$\limsup \frac{1}{n} \log \sup_{(x,\xi)\in T^1M} \int |d_x f_\omega^n \xi|^{-\beta} d\mu \le -\frac{\chi\beta}{2}.$$

Proof. By Corollary 4.2, it is enough to show that the POE is bounded by $-\frac{\chi\beta}{2}$. Then,

$$\begin{aligned}
\check{P}(\{\phi_t\}_t) &= \sup_{\nu \text{ erg. inv. prob.}} \left\{ h_\nu(T) + \int \widetilde{\phi}(\omega) d\nu(\omega) \right\} \ (\because \text{Theorem } 3.8) \\
&= \sup_{\nu \text{ erg. inv. prob.}} \left\{ h_\nu(T) + \int\int \widehat{\phi}(\omega, t) d\sigma_\nu(t) d\nu(\omega) \right\} \ (\because \text{Proposition } 3.6) \\
&= \sup_{\nu \text{ erg. inv. prob.}} \left\{ \int \left[h_\nu(T) + \int \widehat{\phi}(\omega, t) d\nu \right] d\sigma_\nu \right\} \\
&= \sup_{\nu \text{ erg. inv. prob.}} \left\{ \int \left[h_\nu(T) + \int \phi_t d\nu \right] d\sigma_\nu \right\} \\
&\le \sup_{\nu,\eta \text{ erg. inv. prob.}} \left\{ \int \left[h_\eta(T) + \int \phi_t d\eta \right] d\sigma_\nu \right\} \\
&= \sup_\nu \left\{ \sup_\eta \int \left[h_\eta(T) + \int \phi_t d\eta \right] d\sigma_\nu \right\} \\
&\le \sup_\nu \left\{ \int \sup_\eta \left[h_\eta(T) + \int \phi_t d\eta \right] d\sigma_\nu \right\} \\
&= \sup_\nu \left\{ \int P(\phi_t) d\sigma_\nu \right\} \le \sup_\nu \left\{ \int -\frac{\beta\chi}{2} d\sigma_\nu \right\} = -\frac{\beta\chi}{2} \ (\because \text{Theorem } 4.5).
\end{aligned}$$

□

Appendix A. Reduction to a one-sided shift

Let $\Sigma^{\pm}$ be a two-sided shift.

Lemma A.1. *Assume that for every $\omega, \omega' \in \Sigma^{\pm}$ which lie on the same stable leaf, the limit $\lim_{n\to\infty} f_\omega^{-n} \circ f_{\omega'}^n$ converges in $C^{1+\alpha}$ norm with a uniform exponential rate. Then there exists a Hölder continuous mapping $\omega \mapsto f_\omega^+$ s.t $f^+_{(\omega_i)_{i\in\mathbb{Z}}} = f^+_{(\omega_i)_{i\geq 0}}$ and f_ω^+ is $C^{1+\alpha}$-cohomologous to f_ω, where the coboundary depends in a Hölder manner on ω.*

Proof. For every $[a] \subset \Sigma^{\pm}$, fix $\omega_a \in [a]$. Denote by $[\omega_a, \omega]$ the Smale bracket. Define $C_\omega := \lim_{n\to\infty} f_\omega^{-n} \circ f^n_{[\omega_{\omega_0},\omega]}$. Note that C_ω is a $C^{1+\alpha}$ diffeomorphism, as it admits an inverse $C_\omega^{-1} = \lim_{n\to\infty} f^{-n}_{[\omega_{\omega_0},\omega]} \circ f_\omega^n$. Then set $f_\omega^+ := C_{T\omega}^{-1} \circ f_\omega \circ C_\omega$, which is a $C^{1+\alpha}$ diffeomorphism.

We first show that $\omega \mapsto f_\omega^+$ is Hölder continuous. Let $\omega, \omega' \in \Sigma^{\pm}$ with $d(\omega, \omega) = e^{-N} < 1$, and write $\omega_0 = \omega_0' = a$. It is enough to show that $\omega \mapsto C_\omega$ is Hölder continuous. By the uniform exponential convergence, let $K > 0$ and and $\theta \in (0,1)$ s.t

$$d_{C^{1+\alpha}}(C_\omega, C_{\omega'}) \leq 2Ke^{-\theta n} + d_{C^{1+\alpha}}(f_\omega^n \circ f^{-n}_{[\omega_a,\omega]}, f^n_{\omega'} \circ f^{-n}_{[\omega_a,\omega']}). \tag{20}$$

Choose $n = r \cdot N$, where $r \in (0,1)$ will be specified later. Let $K, \tau > 0$ s.t $\omega \mapsto f_\omega$ is (K', τ)-Hölder continuous. Then one can check by induction that $d(f_\omega^n \circ f_{\omega'}^{-n}, Id) \leq B^n \cdot e^{-\tau(N-n)}$ for all N large enough so $B^n \cdot e^{-\tau(N-n)} \leq 1$, where B is a constant depending on $\max_\omega\{\|f_\omega\|_{C^{1+\alpha}}, \|f_\omega^{-1}\|_{C^{1+\alpha}}\}$, K, and K'. By choosing $r > 0$ sufficiently small we can guarantee that

$$d(f_\omega^n \circ f_{\omega'}^{-n}, Id) \leq K'' e^{-\frac{\tau}{2}N}.$$

Then, for all N large enough (and $r > 0$ sufficiently small),

$$d_{C^{1+\alpha}}(f_\omega^n \circ f^{-n}_{[\omega_a,\omega]}, f^n_{\omega'} \circ f^{-n}_{[\omega_a,\omega']}) \leq K''' e^{-\frac{\tau}{3}N}.$$

Finally, putting this together with (20), we conclude that $\omega \mapsto C_\omega$ is Hölder continuous.

We continue to show that f_ω^+ depends only the non-negative coordinates of ω:

$$\begin{aligned} f_\omega^+ =& C_{T\omega}^{-1} \circ f_\omega \circ C_\omega = \lim_n f^{-n}_{[\omega_{\omega_1}, T\omega]} \circ f^n_{T\omega} \circ f_\omega \circ f_\omega^{-n-1} \circ f^{n+1}_{[\omega_{\omega_0},\omega]} \\ =& \lim_n f^{-n}_{[\omega_{\omega_1}, T\omega]} \circ f^{n+1}_{[\omega_{\omega_0},\omega]}, \end{aligned}$$

which depends only on $(\omega_i)_{i\geq 0}$. □

Remark:

$$\int \log|d_x f_\omega^+ \xi| d\mu = \int \log\left|d_{C_\omega(x)} f_\omega \frac{d_x C_\omega \xi}{|d_x C_\omega \xi|}\right| + \log|d_x C_\omega \xi| + \log\left|d_{f_\omega \circ C_\omega} C_{T\omega}^{-1} \frac{d_x f_\omega \circ C_\omega \xi}{|d_x f_\omega \circ C_\omega \xi|}\right| d\mu.$$

Then if $(\{f_\omega\}, \mu)$ satisfies the uniform expansion on average property with a constant χ, and $\max_\omega d(C_\omega, Id)$ is sufficiently small w.r.t χ and $\min_\omega \|d.f_\omega\|_{\text{co}}$, then $(\{f_\omega^+\}, \mu)$ also satisfies the uniform expansion on average condition. This can always be arranged by refining the alphabet of $\Sigma^\pm$ into all admissible words of length m for some sufficiently large m.

References

[BEFRH] Aaron Brown, Alex Eskin, Simion Filip, and Federico Rodriguez-Hertz. Measure rigidity for generalized u-gibbs states and stationary measures via the factorization method. arXiv:2502.14042.

[BG95] Thomas Bogenschütz and Volker Mathias Gundlach. Ruelle's transfer operator for random subshifts of finite type. *Ergodic Theory Dynam. Systems*, 15(3):413–447, 1995.

[BKL23] Jérôme Buzzi, Benoît Kloeckner, and Renaud Leplaideur. Nonlinear thermodynamical formalism. *Ann. H. Lebesgue*, 6:1429–1477, 2023.

[BQ16] Yves Benoist and Jean-Fran¸cois Quint. *Random walks on reductive groups*, volume 62 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Cham, 2016.

[BRH17] Aaron Brown and Federico Rodriguez Hertz. Measure rigidity for random dynamics on surfaces and related skew products. *J. Amer. Math. Soc.*, 30(4):1055–1132, 2017.

[DD] Jon DeWitt and Dima Dolgopyat. Expanding on average diffeomorphisms of surfaces: exponential mixing. arXiv:2410.08445.

[DK07] Dmitry Dolgopyat and Raphaël Krikorian. On simultaneous linearization of diffeomorphisms of the sphere. *Duke Math. J.*, 136(3):475–505, 2007.

[DKS08] Manfred Denker, Yuri Kifer, and Manuel Stadlbauer. Thermodynamic formalism for random countable Markov shifts. *Discrete Contin. Dyn. Syst.*, 22(1-2):131–164, 2008.

[ES23] Rosemary Elliott Smith. Uniformly expanding random walks on manifolds. *Nonlinearity*, 36(11):5955–5972, 2023.

[Kif01] Yuri Kifer. On the topological pressure for random bundle transformations. In *Topology, ergodic theory, real algebraic geometry*, volume 202 of *Amer. Math. Soc. Transl. Ser. 2*, pages 197–214. Amer. Math. Soc., Providence, RI, 2001.

[LW77] Fran¸cois Ledrappier and Peter Walters. A relativised variational principle for continuous transformations. *J. London Math. Soc. (2)*, 16(3):568–576, 1977.

[Pot22] Rafael Potrie. A remark on uniform expansion. *Rev. Un. Mat. Argentina*, 64(1):11–21, 2022.

[Rue67] D. Ruelle. A variational formulation of equilibrium statistical mechanics and the Gibbs phase rule. *Comm. Math. Phys.*, 5:324–329, 1967.

[Sar09] Omri M. Sarig. Lecture notes on thermodynamic formalism for topological markov shifts, May 2009.

[Wal75] Peter Walters. A variational principle for the pressure of continuous transformations. *Amer. J. Math.*, 97(4):937–971, 1975.

S. Ben Ovadia, Einstein Institute of Mathematics, Hebrew University, 91904 Jerusalem, Israel. *E-mail address*: `snir.benovadia@mail.huji.ac.il`

F. Rodriguez-Hertz, The Pennsylvania State University, State College, PA 16801, United States. *E-mail address*: `fjr11@psu.edu`