THERMODYNAMICAL FORMALISM FOR A MODIFIED SHIFT MAP

STEPHEN MUIR AND MARIUSZ URBAŃSKI

Abstract. We introduce a transfer operator and use it to prove some theorems of a classical flavor from thermodynamical formalism (including existence and uniqueness of appropriately defined Gibbs states and equilibrium states for potential functions satisfying Dini's condition and stochastic laws for Hölder continuous potential and observable functions) in a novel setting: the "alphabet" E is a compact metric space equipped with an *apriori* probability measure ν and an endomorphism T. The "modified shift map" S is defined on the product space $E^{\mathbb{N}}$ by the rule $(x_1x_2x_3\dots) \mapsto (T(x_2)x_3\dots)$. The greatest novelty is found in the variational principle, where a term must be added to the entropy to reflect the transformation of the first coordinate by T after shifting. Our motivation is that this system, in its full generality, cannot be treated by the existing methods of either rigorous statistical mechanics of lattice gases (where only the true shift action is used) or dynamical systems theory (where the apriori measure is always implicitly taken to be the counting measure).

1. Introduction.

1.1. **The Space and Map under consideration.** Let (E, d_0) be a compact metric space equipped "apriori" with a Borel probability measure $\nu : \mathscr{B}(E) \to [0,1]$. Let $T : E \to E$ continuously and surjectively with the additional "not-too-contracting-at-short-range" property that for some constants $\kappa > 0$ and $\delta > 0$

$$d_0(a,b) < \delta \;\Rightarrow\; d_0(Ta,Tb) \geq \kappa\, d_0(a,b) \tag{1}$$

If $\kappa > 1$ then this property is called *distance expanding*, but distance expanding isn't required for any of our present theorems. Assume further that T preserves the Borel sets of E in *both* directions, i.e. $T(\mathscr{B}(E)) \subset \mathscr{B}(E)$ (invariance under T^{-1} follows from continuity), and that the apriori measure is quasi-invariant in both directions, i.e.

$$\nu \circ T^{-1} << \nu \;\text{ and }\; \nu \circ T << \nu.$$

Now, unless T injects the function $\nu \circ T$ is not additive on the whole $B(E)$ (though it is subadditive). Nonetheless the density function $\frac{d\nu\circ T}{d\nu}$ is well defined because Equation (1) implies that T must be a local homeomorphism (see Lemma 1.2), and we've assumed it to be biquasi-invariant. Then integration against the "global measure" $\nu \circ T$ can be defined via the density.

In the language of shifts, E serves as the *alphabet* (physically, the *individual state space*) and the product space $X = E^{\mathbb{N}}$ serves as the (full) *shift space* (physically, the *configuration*

The research of both authors was partially supported by NSF Grant DMS 1001874.

space). For any $0 < q < 1$ the distance function

$$d = \frac{1-q}{q} \sum_{k=1}^{\infty} q^k \, d_0 \circ (\pi_k \times \pi_k),$$

wherein π_k is the kth coordinate projection, makes X into a compact metric space (where the constant $\frac{1-q}{q}$ guarantees the diameter of X to be 1) and the product measure $\nu^{\mathbb{N}}$ makes X into a Borel probability space.

Define the *modified shift map* $S : X \to X$ pointwise by

$$S(x_1 x_2 x_3 \dots) = (T x_2 x_3 \dots). \tag{2}$$

It is continuous under the metric d by the continuity of T under the metric d_0. In the cases where (E, d_0) contains a proper limit point the map S cannot be expansive, much less distance expanding. If, for instance, $a_j \to a$ in E as $j \to \infty$ with $a_j \neq a$ for any $j \geq 1$ then the points $(a_j, a_{j+1}, \dots)$ and $(a_{j+1}, a_{j+2}, \dots)$ in X will be arbitrarily close and remain so under all iterates of S for sufficiently large choices of j.

Whenever $f : (X, d) \to (X', d')$ is a map between metric spaces we use the *modulus of continuity* notation

$$\mathfrak{m}(f, t) = \sup\{d'(f(x), f(y)) : d(x, y) \leq t\}.$$

Lemma 1.1. *For any $t > 0$*

$$\mathfrak{m}(S, t) \leq \frac{1}{q} t + q \, \mathfrak{m}\left(T, \frac{t}{q^2}\right).$$

Proof. By rearranging the definition we see

$$\begin{aligned} d(Sx, Sy) &= \frac{1}{q} d(x, y) - d_0(x_1, y_1) + q \left(d_0(Tx_2, Ty_2) - d_0(x_2, y_2)\right) \\ &\leq \frac{1}{q} d(x, y) + q \left(d_0(Tx_2, Ty_2)\right). \end{aligned}$$

The lemma follows by observing (by definition of d) that $d(x, y) \leq t$ implies $d_0(x_2, y_2) \leq \frac{t}{q^2}$. ∎

1.2. **Statement of Theorems to be Proved.** For a continuous function $\phi : X \to \mathbb{R}$ we will define (Equation (5)) a transfer operator $\mathfrak{L}$ in the spirit of Bowen, Ruelle, Perron-Frobenius, etc., and prove (Lemma 2.3) that

$$p(\phi) = \lim_{n \to \infty} \frac{1}{n} \log \mathfrak{L}^n \mathbb{1}(x)$$

exists independently of x and (Theorem 5.1) is an upper bound on "modified free energies"

$$\int \phi \, d\mu + \mathcal{H}\left(\mu | \nu^{\mathbb{N}}\right) + \int_{a \in E} \log\left(\frac{d\nu \circ T}{d\nu}\right) d\mu \circ \pi_2^{-1},$$

where μ ranges over all S invariant Borel probability measures on X.

For descriptive purposes adopt the convention that a function weighting a transfer operator is a *potential function*, and a function the transfer operator acts on is an *observable function*.

For a potential function $\phi : X \to \mathbb{R}$ satisfying Dini's condition (Equation (8), weaker than Hölder continuity) we use the transfer operator in a typical way to produce an S invariant *Gibbs state* η (Lemma 3.3, Fact 4.5) and prove (Theorem 5.1, Theorem 5.2) that it is totally ergodic (Proposition 4.7) and is the unique *equilibrium state*, i.e. S invariant Borel probability measure on X of maximal modified ϕ-free energy.

For Hölder continuous potential functions ϕ we prove that the central limit theorem (Theorem 6.8) applies to any stochastic process of the form $\{f \circ S^k\}_{k\geq 0}$, where f is a Hölder continuous observable : $X \to \mathbb{R}$ and the underlying probability space is $(X, \mathscr{B}(X), \eta)$. We also show this process has exponential decay of correlations (Theorem 6.7) and, in the cases where $T : E \to E$ is a homeomorphism, it adheres to the law of the iterated logarithm (Theorem 6.9).

Finally we consider the dependence on the apriori measure and show (Theorem 7.2) that for a fixed potential satisfying Dini's condition the pressure and equilibrium state depend continuously on the apriori measure.

1.3. **Inverse Branches in the "local" map.** Before defining the transfer operator, pressure, etc., let us establish an important aspect of the single coordinate map $T : E \to E$ which distinguishes S from a true shift. As a continuous surjection between compact spaces, T is necessarily open. A standard argument involving metric topology shows that if we let $s > 0$,$a \in E$, and

$$r(a, s) = \sup\{r > 0 : \forall\, a \in E \;\; B(Ta, r) \subseteq T(B(a, s))\},$$

then $\inf_{a\in E} r(a, s) > 0$. In particular let $\delta > 0$ be the distance threshold from condition 1 and let

$$r_0 = \inf_{a\in E} r\left(a, \frac{\delta}{2}\right). \tag{3}$$

Lemma 1.2. *On any ball $B(b, r_0)$ where $b \in E$, there are finitely many* inverse branches $T_i^{-1} : B(b, r_0) \to B\left(T_i^{-1}b, \frac{\delta}{2}\right)$ *with the property that $T|_{T_i^{-1}(B(b,r_0))}$ is a homeomorphism from $T_i^{-1}(B(b, r_0))$ to $B(b, r_0)$ (whose inverse is T_i^{-1}). These inverse branches account for all the preimages of any point in the ball $B(b, r_0)$; the number of branches is constant on connected components of E; and*

$$P \;\equiv\; \max\{|T^{-1}(a)| : a \in E\} \;<\; \infty. \tag{4}$$

Proof. First of all notice that even if E has infinitely many connected components, $P < \infty$ still holds. This is simply because E is compact: points with infinitely many preimages would necessarily violate property (1) by having preimages arbitrarily close to each other.

Also by property (1) it follows that T injects on every ball of radius $\frac{\delta}{2}$, and by the definition of r_0 it follows that for every $a \in E$ $B(Ta, r_0) \subseteq\; T\left(B\left(a, \frac{\delta}{2}\right)\right)$. To construct the inverse branches at a point $b \in E$, recall the assumption (surjectivity) that $T^{-1}b \neq \emptyset$ and label $T^{-1}b = \left\{T_1^{-1}b, \dots, T_j^{-1}b\right\}$. The observations in the first sentence of the proof show that for each $c \in B(b, r_0)$ and for each $i : 1 \leq i \leq j$ there is a unique point $T_i^{-1}c \in T^{-1}(c) \cap B\left(T_i^{-1}b, \frac{\delta}{2}\right)$. Because T is continuous and open, the claim of homeomorphism has been proved.

Now, $\cup_{i=1}^{j} T_i^{-1}c$ accounts for all of $T^{-1}(c)$ because if there were another preimage of c then the present reasoning would imply the existence of another preimage of b, which is assumed not to exist. Thus the cardinality of preimages of points under T is constant on balls of radius r_0 and moreover on connected components of E, which can be covered by paths of overlapping balls of radius r_0. □

Said differently, this lemma shows that if $d_0(a,b) < r_0$, then a and b can share the same inverse branches $T_i^{-1}, 1 \le i \le j$, and for each i, $d_0(T_i^{-1}a, T_i^{-1}b) < \delta$.

2. The transfer operator associated to a function $\phi : X \to \mathbb{R}$.

If $\phi : X \to \mathbb{R}$ continuously then the *transfer operator associated to* ϕ is the bounded linear operator mapping $C(X) \to C(X)$ by the rule

$$\mathfrak{L} f(x) = \int_{a \in E} \sum_{b \in T^{-1}(x_1)} f e^{\phi}(abx_2x_3\dots)\, d\nu(a). \tag{5}$$

Fubini's theorem allows a straightforward inductive proof that for all $n \ge 1$

$$\mathfrak{L}^n f(x) = \int_{a \in E^n} \sum_{\substack{b_1 \in T^{-1}(a_2) \\ \vdots \\ b_{n-1} \in T^{-1}(a_n) \\ b_n \in T^{-1}(x_1)}} f e^{s_n\phi}(a_1b_1\dots b_nx_2x_3\dots) d\nu^n(a), \tag{6}$$

wherein appears the *ergodic sum* $s_n\phi = \phi + \phi \circ S + \cdots + \phi \circ S^{n-1}$.

Let

$$\Delta_n(\phi) = \sup\left\{ |s_n\phi(ax) - s_n\phi(ay)| : a \in E^n,\ x, y \in E^{\mathbb{N}} \right\}. \tag{7}$$

We say ϕ *satisfies Dini's condition* if

$$\sum_{n=0}^{\infty} \mathfrak{m}(\phi, 2^{-n}) \;<\; \infty. \tag{8}$$

Notice that ϕ satisfies condition (8) if and only if for every $0 < q < 1$

$$\sum_{n=0}^{\infty} \mathfrak{m}(\phi, q^{n}) \;<\; \infty. \tag{9}$$

Lemma 2.1. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition then*

$$h(t) = \sum_{n=1}^{\infty} \mathfrak{m}\left(\phi, tq^n\right)$$

defines a nondecreasing function $h = h_{\phi,q} : [0,\infty) \to [0,\infty)$ with $h(0) = 0$ and $h(t) \to 0$ as $t \to 0$.

Proof. The claims are all obvious except perhaps the last one. For any $\epsilon > 0$ there is an index n_ϵ so large that $\sum_{n \ge n_\epsilon} \mathfrak{m}(\phi, q^n) < \epsilon$. Then $t < q^{n_\epsilon - 1}$ implies $h(t) < \sum_{n \ge 1} \mathfrak{m}\left(\phi, q^{n+n_\epsilon-1}\right) < \epsilon$. □

The dependence of the function h on the parameters ϕ and q is suppressed when we aren't considering the effect of varying them.

Lemma 2.2. *If $\phi : X \to \mathbb{R}$ continuously then $\Delta_n\phi = o(n)$, and in particular if $\phi : X \to \mathbb{R}$ satisfies Dini's condition then $\{\Delta_n\phi\}_{n\geq 1}$ is bounded.*

Proof. Regardless of any conditions on ϕ the triangle inequality yields

$$\Delta_n\phi \leq \mathfrak{m}\left(\phi, q^n \mathrm{diam} E\right) + \cdots + \mathfrak{m}\left(\phi, q\mathrm{diam} E\right).$$

The second claim now follows from Lemma 2.1: if ϕ satisfies Dini's condition then $\Delta_n\phi \leq h(\mathrm{diam}E) < \infty$ holds for all $n \geq 1$.

For the first claim let $\epsilon > 0$ and observe that mere continuity of ϕ implies that for large enough integers, say for $n \geq n_\epsilon$, it holds that $\mathfrak{m}\left(\phi, q^n\mathrm{diam}E\right) < \epsilon$ and hence

$$\begin{aligned}
\frac{1}{n}\Delta_n\phi &\leq \frac{1}{n}\left(\mathfrak{m}\left(\phi, q^n\mathrm{diam}E\right) + \cdots + \mathfrak{m}\left(\phi, q^{n_\epsilon}\mathrm{diam}E\right) + \cdots + \mathfrak{m}\left(\phi, q\mathrm{diam}E\right)\right) \\
&\leq \frac{n-(n_\epsilon - 1)}{n}\epsilon \;+\; \frac{1}{n}\left(\mathfrak{m}\left(\phi, q^{n_\epsilon - 1}\mathrm{diam}E\right) + \cdots + \mathfrak{m}\left(\phi, q\mathrm{diam}E\right)\right),
\end{aligned}$$

which converges to ϵ as $n \to \infty$. Thus $\overline{\lim}_{n\to\infty}\frac{1}{n}\Delta_n\phi \leq \epsilon$, and the first claim is proved. $\square$

The next lemma introduces a standard formula for a number $p(\phi)$ associated to a continuous function ϕ. Here $p(\phi)$ characterizes the growth rate of the unit function $\mathbb{1} = \mathbb{1}_X : X \to \mathbb{R}$ under iterates of ϕ's transfer operator $\mathfrak{L}$, and it will be seen to have further significance (see Equation (14), Theorem 5.1).

Lemma 2.3. *For any continuous $\phi : X \to \mathbb{R}$*

$$p(\phi) = \lim_{n\to\infty}\frac{1}{n}\log \mathfrak{L}^n\mathbb{1}(x)$$

exists in $\mathbb{R}$ and is independent of x.

Proof. The first step is to check that for any given $x \in X$ the limit exists. This follows from near-subadditivity of the sequence $\{\log \mathfrak{L}^n\mathbb{1}(x)\}_{n\geq 1}$. Let $1 \leq m < n$, and consider the integrand of $\mathfrak{L}^n\mathbb{1}(x)$ at the point $a \in E^n$. Reverse the order of the summation and factor as

$$\sum_{\substack{b_n\in T^{-1}x_1\\ \vdots \\ b_m\in T^{-1}a_{m+1}}} e^{s_{n-m}\phi}(a_{m+1}b_{m+1}\dots b_n x_2 x_3\dots)\left(\sum_{\substack{b_{m-1}\in T^{-1}a_m\\ \vdots\\ b_1\in T^{-1}a_2}} e^{s_m\phi}(a_1 b_1\dots b_n x_2x_3\dots)\right).$$

Factor further, with an error factor bounded by $e^{\Delta_m(\phi)}$, by replacing $b_m \dots b_n x_2 x_2 \dots$ with an arbitrary tail $y \in E^{\mathbb{N}}$:

$$\leq e^{\Delta_m \phi} \left(\sum_{\substack{b_{m-1} \in T^{-1} a_m \\ \vdots \\ b_1 \in T^{-1} a_2}} e^{s_m \phi}(a_1 b_1 \dots b_{m-1} y) \right) \left(\sum_{\substack{b_n \in T^{-1} x_1 \\ \vdots \\ b_m \in T^{-1} a_{m+1}}} e^{s_{n-m} \phi}(a_{m+1} b_{m+1} \dots b_n x_2 x_3 \dots) \right).$$

Then put the order of summation back to normal and use Fubini's theorem to obtain

$$\mathfrak{L}^n 1\!\!1(x) \leq e^{\Delta_m \phi} \int_{a_1 \dots a_m \in E^m} \sum_{\substack{b_1 \in T^{-1} a_2 \\ \vdots \\ b_{m-1} \in T^{-1} a_m}} e^{s_m \phi}(a_1 b_1 \dots b_{m-1} y)\, d\nu^m(a_1 \dots a_m) \cdot \int_{a_{m+1} \dots a_n \in E^{n-m}} \sum_{\substack{b_m \in T^{-1} a_{m+1} \\ \vdots \\ b_n \in T^{-1} x_1}} e^{s_{n-m} \phi}(a_{m+1} b_{m+1} \dots b_n x_2 x_3 \dots)\, d\nu^{n-m}(a_{m+1} \dots a_n).$$

Finally force the summations to match the integrations as they should in the transfer operator with another error factor bounded by $Pe^{\Delta_m(\phi)}$:

$$\leq e^{2\Delta_m \phi} \int_{a \in E^m} \sum_{\substack{b_1 \in T^{-1} a_2 \\ \vdots \\ b_{m-1} \in T^{-1} a_m}} \sum_{b_m \in T^{-1} y_1} e^{s_m \phi}(a_1 b_1 \dots b_{m-1} b_m y_2 y_3 \dots) d\nu^m(a_1 \dots a_m) \cdot P \int_{a_{m+1} \dots a_n \in E^{n-m}} \sum_{\substack{b_{m+1} \in T^{-1} a_{m+2} \\ \vdots \\ b_n \in T^{-1} x_1}} e^{s_{n-m} \phi}(a_{m+1} b_{m+1} \dots b_n x_2 x_3 \dots) d\nu^{n-m}(a_{m+1} \dots a_n)$$

$$= Pe^{2\Delta_m} \mathfrak{L}^m 1\!\!1(y) \mathfrak{L}^{n-m} 1\!\!1(x).$$

This holds for any y, and letting $y = x$ yields the desired near-submultiplicativity:

$$\mathfrak{L}^n 1\!\!1(x) \leq Pe^{2\Delta_m \phi} \mathfrak{L}^m 1\!\!1(x) \mathfrak{L}^{n-m} 1\!\!1(x). \tag{10}$$

Schematically now there is a sequence $\{L_n = \log \mathfrak{L}^n 1\!\!1(x)\}_{n \geq 1}$ satisfying the near-subadditivity property that for any $1 \leq m < n$

$$L_n \leq C_m + L_m + L_{n-m},$$

where it is known that $C_m = o(m)$ (specifically $C_m = \log(P) + 2\Delta_m \phi$). It is a fairly routine calculation to show that in such a case $\lim_{n\to\infty} \frac{1}{n} L_n$ exists (though it need not equal the infimum of the sequence, as in the truly subadditive case.)

To show that the limit is independent of x examine the integral in the right hand side of Equation (6). Observe that for any two points $x, y \in X$, any point $a \in E^n$ in the domain of integration, and any choice of indices $b_1 \in T^{-1}a_2, \dots, b_{n-1} \in T^{-1}a_n$ in the integrands of $\mathfrak{L}\mathbb{1}(x)$ and $\mathfrak{L}\mathbb{1}(y)$, respectively, the estimate

$$P^{-1}e^{-\Delta_n\phi} \leq \frac{\sum_{b_n^x \in T^{-1}x_1} e^{s_n\phi}(a_1b_1 \dots b_n^x x_2x_3 \dots)}{\sum_{b_n^y \in T^{-1}y_1} e^{s_n\phi}(a_1b_1 \dots b_n^y y_2y_3 \dots)} \leq Pe^{\Delta_n\phi}$$

holds, and therefore in fact

$$P^{-1}e^{-\Delta_n\phi} \leq \frac{\mathfrak{L}^n\mathbb{1}(x)}{\mathfrak{L}^n\mathbb{1}(y)} \leq Pe^{\Delta_n\phi}. \tag{11}$$

Use $\Delta_n\phi = o(n)$ again to conclude $\frac{1}{n}|\log \mathfrak{L}^n\mathbb{1}(x) - \log \mathfrak{L}^n\mathbb{1}(y)| \to 0$ as $n \to \infty$. $\square$

Note that for all continuous $\phi : X \to \mathbb{R}$

$$\lim_{n\to\infty} \frac{1}{n} \log \sup \mathfrak{L}^n\mathbb{1} = \lim_{n\to\infty} \frac{1}{n} \log \inf \mathfrak{L}^n\mathbb{1} = p(\phi)$$

is also true.

Observation 2.4. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition then for every $n \geq 1$ and every $x, y \in X$*

$$P^{-1}e^{-h(diamE)} \leq \frac{\mathfrak{L}^n\mathbb{1}(x)}{\mathfrak{L}^n\mathbb{1}(y)} \leq Pe^{h(diamE)},$$

Proof. Apply the second claim of Lemma 2.2 to Estimate (11). $\square$

Before proceeding to the construction of Gibbs states and the variational principle it is prudent to refine and extend the estimates that established the existence of the pressure.

We consider two points $x, y \in X$ to be *sufficiently close* if $d_0(x_1, y_1) < r_0$. Recall from Lemma 1.2 that if x, y are sufficiently close then they share common inverse branches:

$$T^{-1}x_1 = \{T_i^{-1}x_1\}_{i=1}^j \text{ and } T^{-1}y_1 = \{T_i^{-1}y_1\}_{i=1}^j$$

and for each $i : 1 \leq i \leq j$, $d_0(T_i^{-1}x_1, T_i^{-1}y_1) < \delta$.

Lemma 2.5. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition (8) and $x, y \in X$ are sufficiently close as just described, then for every $n \geq 1$, every point $a \in E^n$, and every inverse branch T_i^{-1} in common between x and y*

$$|s_n\phi(aT_i^{-1}x_1x_2\dots) - s_n\phi(aT_i^{-1}y_1y_2\dots)| \leq h\left(\max\left(1, \frac{1}{\kappa}\right)d(x,y)\right).$$

Proof. By Property (1) $d_0(T_i^{-1}x_1, T_i^{-1}y_1) \leq \frac{1}{\kappa}d_0(x_1, y_1)$. Thus for any $n \geq 1$ and any $k : 1 \leq k \leq n$

$$\begin{aligned} d(a_k \dots a_nT_i^{-1}x_1x_2\dots, a_k\dots a_nT_i^{-1}y_1, y_2\dots) &= q^{n-k+1}d(T_i^{-1}x_1x_2\dots, T_i^{-1}y_1, y_2\dots) \\ &\leq q^{n-k+1}\max\left(1, \frac{1}{\kappa}\right)d(x,y). \end{aligned}$$

Now by the triangle inequality the difference in question $|s_n\phi(aT_i^{-1}x_1x_2\dots) - s_n\phi(aT_i^{-1}y_1y_2\dots)|$ is bounded above by the first n terms of the series $h\left(\max\left(1, \frac{1}{\kappa}\right)d(x,y)\right)$. $\square$

From now on denote

$$\tilde{\kappa} = \max\left(1, \frac{1}{\kappa}\right). \tag{12}$$

Lemma 2.6. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition, $f : X \to \mathbb{R}$ is continuous, and $x, y \in X$ are sufficiently close as just described, then for every $n \geq 1$*

$$|\mathfrak{L}^n f(x) - \mathfrak{L}^n f(y)| \leq \mathfrak{L}^n \mathbb{1}(x) \mathfrak{m}\left(f, q^n \tilde{\kappa} d(x,y)\right) + \|f\|_\infty \mathfrak{L}^n \mathbb{1}(y) |1 - e^{h(\tilde{\kappa} d(x,y))}|. \tag{13}$$

Proof. Combine the difference of the two integrals $\mathfrak{L}^n f(x)$ and $\mathfrak{L}^n f(y)$ into one integral $d\nu^n$. In the integrand, for each $a \in E^n, b_1 \in T^{-1}a_2, \ldots, b_{n-1} \in T^{-1}a_n$, and each i indexing a common inverse branch of T at x_1 and y_1, add and subtract a term of the form

$$f(a_1 b_1 \ldots b_{n-1} T_i^{-1} y_1\, y_2 \ldots) e^{s_n \phi}(a_1 b_1 \ldots b_{n-1} T_i^{-1} x_1\, x_2 \ldots).$$

With the triangle inequality this leads to

$$\begin{aligned} |\mathfrak{L}^n f(x) - \mathfrak{L}^n f(y)| \leq & \int_{a \in E^n} \sum_{\substack{b_1 \in T^{-1} a_2 \\ \vdots \\ b_{n-1} \in T^{-1} a_n}} \sum_{i=1}^{j} \Bigg(\\ & e^{s_n \phi}(a_1 b_1 \ldots T_i^{-1} x_1 x_2 \ldots) \Big| f(a_1 b_1 \ldots T_i^{-1} x_1 x_2 \ldots) - f(a_1 b_1 \ldots T_i^{-1} y_1 y_2 \ldots) \Big| + \\ & + \Big| f(a_1 b_1 \ldots T_i^{-1} y_1 y_2 \ldots) \Big| \Big| e^{s_n \phi}(a_1 b_1 \ldots T_i^{-1} x_1 x_2 \ldots) - e^{s_n \phi}(a_1 b_1 \ldots T_i^{-1} y_1 y_2 \ldots) \Big| \Bigg) d\nu^n(a). \end{aligned}$$

Separate this integral into the integral of the sum of the left hand terms of each summand and the integral of the sum of the right hand terms of each summand. In the left hand terms

$$|f(a_1 b_1 \ldots T_i^{-1} x_1 x_2 \ldots) - f(a_1 b_1 \ldots T_i^{-1} y_1 y_2 \ldots)| \leq \mathfrak{m}\left(f, q^n \tilde{\kappa} d(x,y)\right).$$

In the right hand terms factor out $e^{s_n \phi}(a_1 b_1 \ldots T_i^{-1} y_1 y_2 \ldots)$ and use Lemma 2.5 to complete the estimate. □

Lemma 2.6 shows a sort of local equicontinuity in the sequence $\{\mathfrak{L}^n f\}_{n \geq 1}$; with the right normalization of the operator it will become truly equicontinuous. The normalization uses the dual transfer operator and is a key point in demonstrating the existence of an equilibrium state (with respect to S) for the potential function ϕ satisfying Dini's condition.

3. The Dual Transfer Operator.

The vector space C of continuous complex valued functions on $X = E^{\mathbb{N}}$ is a Banach space with $\| \, \|_\infty$, and for any $\phi \in C$ the associated transfer operator $\mathfrak{L}$ is a bounded linear operator on C. Let C^* be the dual space of bounded linear functionals on C, and let $\mathfrak{L}^*$ be the *adjoint* or *dual* bounded linear operator on C^*, defined by $\mathfrak{L}^* \mu(f) = \mu(\mathfrak{L} f)$. The transformation

$$C^* \ni \ \mu \mapsto \frac{1}{\mathfrak{L}^* \mu(\mathbb{1})} \mathfrak{L}^* \mu \ \in C^*$$

is not linear, but it is clearly continuous (with respect to the weak* topology on C^*) and it preserves the (weak*) compact subset $C^*_{+,1}$ of positive linear functionals with norm 1. By Schauder and Tichonov's fixed point theorem there must be a functional $\gamma \in C^*_{+,1}$ fixed by this transformation. In terms of $\mathfrak{L}$ this means that for all $f \in C$

$$\gamma(\mathfrak{L}f) = \gamma(\mathfrak{L}\mathbb{1})\,\gamma(f). \tag{14}$$

We see in Corollary 5.3 that γ is unique. Naively for now adopt the convention that, assuming a potential function ϕ has been fixed, γ always refers to an arbitrary eigenvector of $\mathfrak{L}^*$ as defined in Equation (14).

Observation 3.1. *If $\phi \in C$ then $\log\gamma(\mathfrak{L}\mathbb{1}) = p(\phi)$.*

Proof. Iterating Equation (14) and evaluating both sides at $\mathbb{1}$ yields

$$\gamma(\mathfrak{L}^n\mathbb{1}) = (\gamma(\mathfrak{L}\mathbb{1}))^n\,\gamma(\mathbb{1}) = (\gamma(\mathfrak{L}\mathbb{1}))^n\,.$$

So

$$\log\gamma(\mathfrak{L}\mathbb{1}) = \frac{1}{n}\log\gamma(\mathfrak{L}^n\mathbb{1}) \left\{ \begin{array}{l} \le \frac{1}{n}\log\sup\mathfrak{L}^n\mathbb{1} \\ \ge \frac{1}{n}\log\inf\mathfrak{L}^n\mathbb{1} \end{array} \right. \to p(\phi) \quad \text{as} \quad n\to\infty.$$

$\square$

Next comes the normalization, promised in the previous section, which makes the sequence of iterates of the unit function under the transfer operator equicontinuous. The *normalized transfer operator* weighted by a potential function $\phi \in C$ is defined by

$$\mathfrak{L}_0 \equiv e^{-p(\phi)}\mathfrak{L} \tag{15}$$

Notice $\mathfrak{L}_0^*\gamma = \gamma$. The claimed equicontinuity and its first consequence are elaborated in the next lemma.

Lemma 3.2. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition then $e^{p(\phi)}$ is an eigenvalue of $\mathfrak{L}$ with a positive eigenvector $\rho \in C$ that is bounded away from 0.*

Proof. Let $n \ge 1$. Because γ is positive and $\gamma(\mathfrak{L}^n\mathbb{1}) = e^{np(\phi)}$ there must be points $y, z \in X$ satisfying

$$\mathfrak{L}^n\mathbb{1}(y) \le e^{np(\phi)} \le \mathfrak{L}^n\mathbb{1}(z).$$

By Observation 2.4,

$$P^{-1}e^{-h(\mathrm{diam}E)}\mathfrak{L}^n\mathbb{1}(z) \le \mathfrak{L}^n\mathbb{1}(x) \le Pe^{h(\mathrm{diam}E)}\mathfrak{L}^n\mathbb{1}(y),$$

and combining these yields

$$P^{-1}e^{np(\phi)-h(\mathrm{diam}E)} \le \mathfrak{L}^n\mathbb{1}(x) \le Pe^{np(\phi)+h(\mathrm{diam}E)},$$

i.e.

$$P^{-1}e^{-h(\mathrm{diam}E)} \le \mathfrak{L}_0^n\mathbb{1}(x) \le Pe^{h(\mathrm{diam}E)} \tag{16}$$

for every $x \in X$. Now take $f = \mathbb{1}$ in Lemma 2.6 and divide by $e^{np(\phi)}$ to obtain

$$\begin{aligned} \left|\mathfrak{L}_0^n\mathbb{1}(x) - \mathfrak{L}_0^n\mathbb{1}(y)\right| &\le \mathfrak{L}_0^n\mathbb{1}(y)\left|1 - e^{h(\tilde{\kappa}d(x,y))}\right| \\ &\le Pe^{\mathrm{diam}E}\left|1 - e^{h(\tilde{\kappa}d(x,y))}\right|, \end{aligned}$$

which holds for all $x, y \in X$ with $d_0(x_1, y_1) < 2r_0$. Thus the sequence of functions $\{\mathfrak{L}_0^n 1\!\!1\}_{n\geq 1}$ is equicontinuous. Because

$$\mathfrak{m}\left(\frac{1}{n}\mathfrak{L}_0^n 1\!\!1, \delta\right) = \frac{1}{n}\mathfrak{m}\left(\mathfrak{L}_0^n 1\!\!1, \delta\right)$$

the sequence $\left\{\frac{1}{n}\mathfrak{L}_0^n 1\!\!1\right\}_{n\geq 1}$ is equicontinuous too. By Arzela and Ascoli's theorem there is a $\| \, \|_\infty$ convergent subsequence $\left\{\frac{1}{n_j}\mathfrak{L}_0^{n_j} 1\!\!1\right\}_{j\geq 1}$ converging to a uniformly continuous limit ρ. Because $\mathfrak{L}_0$ is continuous

$$\begin{aligned}|\mathfrak{L}_0\rho - \rho| &= \left|\lim_{j\to\infty} \mathfrak{L}_0 \frac{1}{n_j}\mathfrak{L}_0^{n_j} 1\!\!1 - \frac{1}{n_j}\mathfrak{L}_0^{n_j} 1\!\!1\right| \\ &= \lim_{j\to\infty} \left|\mathfrak{L}_0^{n_j}\frac{1}{n_j}\left(\mathfrak{L}_0 1\!\!1 - 1\!\!1\right)\right| = 0\end{aligned}$$

because $\{\mathfrak{L}_0^{n_j}\}_{j\geq 1}$ is bounded in operator norm. Thus $\mathfrak{L}_0\rho = \rho$.

Because all the functions $\mathfrak{L}_0^{n_j} 1\!\!1$, $j \geq 1$ are bounded below $P^{-1}e^{-h(\mathrm{diam}E)}$, their limit ρ is too. Note for use in Theorem 7.2 that Estimate (16) proves that ρ is in fact bounded from 0 and ∞ by constants that are independent of the apriori measure ν. $\square$

Lemma 3.3. *There is an S-invariant functional $\eta \in C^*_{+,1}$ that is boundedly and continuously equivalent to γ.*

Proof. We have only to check that $\eta = \frac{1}{\rho\gamma(1\!\!1)}\rho\gamma$ is S-invariant, because the boundedness and continuity of the density function ρ was proved in Lemma 3.2. Both facts $\mathfrak{L}_0^*\gamma = \gamma$ and $\mathfrak{L}_0\rho = \rho$ are used:

$$\rho\gamma(f \circ S) = \gamma(\rho\, f \circ S) = \gamma(\mathfrak{L}_0(\rho\, f \circ S)) = \gamma(f\, \mathfrak{L}_0\rho) = \gamma(f\rho) = \rho\gamma(f).$$

Scaling by $\frac{1}{\rho\gamma(1\!\!1)}$ won't change the invariance. $\square$

4. Gibbs states associated to a potential function $\phi : X \to \mathbb{R}$

There is a natural Borel probability measure associated to each functional in $C^*_{+,1}$; we let the two go by the same name. If $n \geq 1$ is specified and $A \in \mathscr{B}(E^n)$ then the *cylinder set* notation

$$[A] = \pi^{-1}_{1\ldots n}(A)$$

can be used without ambiguity.

Definition 4.1. *We define a* Gibbs state *for $\phi \in C$ as a measure $\mu \in C^*_{+,1}$ for which there exist constants $c_1, c_2 > 0$ such that for every $n \geq 1$, every set $A \in \mathscr{B}(E^n)$ with $\nu^n(A) > 0$, and every point $x \in X$*

$$\frac{1}{c_1 c_2^n} \leq \frac{\mu([A])}{\mathfrak{L}^n 1\!\!1_{[A]}(x)} \leq \frac{c_1}{c_2^n}. \tag{17}$$

Note that the proof of the following observation contains the fact that $\nu^n(A) = 0$ if and only if $\mathfrak{L}^n 1\!\!1_{[A]}$ is identically equal to 0, which prevents division by 0 in the preceding definition.

Observation 4.2. *If $\phi \in C$ and μ is a Gibbs state for ϕ then the finite projections $\mu \circ \pi_{1\dots n}^{-1}$, $n \geq 1$ are equivalent to the apriori measures ν^n on $(E^n, \mathscr{B}(E^n))$, respectively.*

Proof. Let $n \geq 1$ and a consider any rectangle $A = A_1 \times \cdots \times A_n \in \mathscr{B}(E^n)$.

$$\mathfrak{L}^n \mathbb{1}_{[A]}(x) = \int_{a \in E^n} \sum_{\substack{b_1 \in T^{-1} a_2 \\ \vdots \\ b_{n-1} \in T^{-1} a_n \\ b_n \in T^{-1} x_1}} \mathbb{1}_{A_1}(a_1) \mathbb{1}_{A_2}(b_1) \dots \mathbb{1}_{A_n}(b_{n-1}) \cdot$$

$$\cdot\, e^{s_n \phi}(a_1 b_1 \dots b_n x_2 \dots)\, d\nu^n(a)$$

$$= \int_{a \in E \times T A_2 \times \cdots \times T A_n} \sum_{\substack{b_1 \in T^{-1} a_2 \cap A_2 \\ \vdots \\ b_{n-1} \in T^{-1} a_n \cap A_n \\ b_n \in T^{-1} x_1}} e^{s_n \phi}(a_1 b_1 \dots b_n x_2 \dots)\, d\nu^n(a). \tag{18}$$

Assuming $\nu \circ T << \nu$ as we do, this shows $\mathfrak{L}^n \mathbb{1}_{[A]}(x) = 0$ for all $x \in X$ if $\nu^n(A) = 0$. Thus, by the right hand side of Definition 4.1, any Gibbs state γ has $\gamma \circ \pi_{1\dots n}^{-1} << \nu^n$.

Conversely, because the integrand is positive and bounded away from 0, and because of the left hand side of Definition 4.1 it follows that if $\gamma \circ \pi_{1\dots n}^{-1}(A) = 0$ then $\mathfrak{L}^n \mathbb{1}_{[A]}$ is identically equal to 0 and therefore either $\nu(A_1)$ or one of $\nu(TA_2), \dots, \nu(TA_n)$ must be 0. By quasi-invariance ($\nu \circ T^{-1} << \nu$) it holds that if $\nu(TA_k) = 0$ then $\nu(A_k) = 0$, and so we conclude that $\nu^n << \gamma \circ \pi_{1\dots n}^{-1}$. □

Observation 4.3. *A functional $\mu \in C_{+,1}^*$ is a Gibbs state for $\phi \in C$ if and only if there are constants $c_1, c_2 > 0$ such that for every $n \geq 1$ and every $\mathscr{B}_n$ measurable $f \in C$*

$$\frac{1}{c_1 c_2^n} \leq \frac{\mu(f)}{\mathfrak{L}^n f(x)} \leq \frac{c_1}{c_2^n}. \tag{19}$$

Proof. It is a standard limiting argument in measure theory. If the measure μ is known to be a Gibbs state then Estimate (4.1) extends to $\mathscr{B}_n$ measurable simple functions by linearity of μ and $\mathfrak{L}$; then it extends to all $\mathscr{B}_n$ measurable functions by the Lebesgue dominated convergence theorem. Conversely, if the functional μ is known to be a Gibbs state then Estimate (19) applies to all continuous "bump" functions f that could be used to approximate indicator functions $\mathbb{1}_{[A]}$, and so the measure constructed to represent μ as integration (in the Riesz representation theorem) is bounded in the same way. □

Observation 4.4. *If $\phi : X \to \mathbb{R}$ is continuous and has a Gibbs state then the constant c_2 in the definition of a Gibbs state must be $e^{p(\phi)}$.*

Proof. For every $n \geq 1$, taking the set A in Definition 4.1 to be E^n yields

$$\frac{1}{c_1 c_2^n} \leq \frac{1}{\mathfrak{L}^n \mathbb{1}(x)} \leq \frac{c_1}{c_2^n}.$$

Taking logs, etc. yields

$$\frac{\log \mathfrak{L}^n \mathbb{1}(x) - \log c_1}{n} \leq \log c_2 \leq \frac{\log \mathfrak{L}^n \mathbb{1}(x) + \log c_1}{n},$$

wherein $\log c_2$ is seen to be bounded between two sequences converging to same limit $p(\phi)$. $\square$

In light of Observation 4.4, an equivalent but more streamlined definition of a Gibbs state would be any $\mu \in C^*_{+,1}$ for which there exists a constant c_1 such that for all $n \geq 1$, all $A \in \mathscr{B}(E^n)$, and all $x \in X$

$$\frac{1}{c_1} \leq \frac{\mu(\mathbb{1}_{[A]})}{\mathfrak{L}_0^n \mathbb{1}_{[A]}(x)} \leq c_1. \tag{20}$$

Fact 4.5. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition then any solution γ of Equation (14) is a Gibbs state for ϕ, and $\eta = \frac{1}{\rho\gamma(\mathbb{1})}\rho\gamma$ is an S-invariant Gibbs state for ϕ.*

Proof. Fix $n \geq 1$ and $A \in \mathscr{B}(E^n)$, and consider the integrands of $\mathfrak{L}^n \mathbb{1}_{[A]}(x)$ and $\mathfrak{L}^n \mathbb{1}_{[A]}(y)$ for arbitrary $x, y \in X$. For each choice of $a \in E^n, b_1 \in T^{-1}a_2, \ldots, b_{n-1} \in T^{-1}a_n, b_n^x \in T^{-1}x_1$, and $b_n^y \in T^{-1}y_1$, the number

$$\{0,1\} \ni \mathbb{1}_{[A]}(a_1 b_1 \ldots b_{n-1} b_n^x x_2 \ldots) = \mathbb{1}_{[A]}(a_1 b_1 \ldots b_{n-1} b_n^y y_2 \ldots)$$

and in fact only depends on $a, b_1, \ldots, b_{n-1}$ (not b_n^x or b_n^y). This means Estimate (11) and Observation 2.4 can be generalized to say that

$$P^{-1} e^{-h(\mathrm{diam} E)} \leq \frac{\mathfrak{L}^n \mathbb{1}_{[A]}(x)}{\mathfrak{L}^n \mathbb{1}_{[A]}(y)} \leq P e^{h(\mathrm{diam} E)},$$

holds for any $n \geq 1$, $A \in \mathscr{B}(E^n)$, and $x, y \in X$. Changing both instances of $\mathfrak{L}$ to $\mathfrak{L}_0$ has the non-effect of multiplying by 1 in the center expression. Therefore, integrating $d\gamma(x)$ (or, equivalently, treating each term as a function of x and applying γ)

$$P^{-1} e^{-h(\mathrm{diam} E) - np(\phi)} \leq \frac{\gamma(\mathfrak{L}_0^n \mathbb{1}_{[A]})}{\mathfrak{L}_0^n \mathbb{1}_{[A]}(y)} \leq P e^{h(\mathrm{diam} E) - np(\phi)},$$

which, bearing in mind the $\mathfrak{L}_0^*$ invariance of γ, proves the first claim. The second claim is a trivial consequence of Lemma 3.3 because it is clear by definition that any functional boundedly equivalent to a Gibbs state is also a Gibbs state. $\square$

Lemma 4.6. *Let $\phi : X \to \mathbb{R}$ satisfy Dini's condition and let η be an S invariant Gibbs state for ϕ. Then there is a constant $c_2 > 0$ such that for any $n \geq 1$, $A \in \mathscr{B}(E^n)$, $B' \in \mathscr{B}(E)$, and $B'' \in \mathscr{B}(X)$*

$$\eta\left(\pi_{1\ldots n}^{-1}(A) \cap \pi_{n+1}^{-1}(B') \cap \pi_{n+2\ldots}^{-1}(B'')\right) \geq c_2 \eta\left(\pi_{1\ldots n}^{-1} A\right) \eta\left(\pi_1^{-1}(TB') \cap \pi_{2\ldots}^{-1} B''\right) \tag{21}$$

holds.

Proof. The inequality will follow from a kind of near-supermultiplicativity in the sequence of iterates under $\mathfrak{L}$ of indicator functions for the appropriate cylinder sets. Separating the integrals and making substitutions as in the proof of Lemma 2.3 leading up to Estimate (10), one ascertains that

$$\mathfrak{L}^{m+n}\mathbb{1}_{[A\times B'\times B'']}(x) \geq \frac{1}{Pe^{2\Delta_m\phi}}\mathfrak{L}^m\mathbb{1}_{[A]}(x)\mathfrak{L}^n\mathbb{1}_{[TB'\times B'']}(x) \tag{22}$$

holds for every $m, n \geq 1$, $A \in \mathscr{B}(E^m)$, $B' \in \mathscr{B}(E)$, $B'' \in \mathscr{B}(E^{n-1})$, and $x \in X$. With this and the bounds from the definition of η as a Gibbs state, (using first the lower bound on the measure of the longer cylinder, then the upper bound on the measure of the two shorter cylinders)

$$\begin{aligned}
\eta\left([A\times B'\times B'']\right) &\geq \frac{1}{c_1e^{(m+n)p(\phi)}}\mathfrak{L}^{m+n}\mathbb{1}_{[A\times B'\times B'']}(x) \\
&\geq \frac{1}{c_1Pe^{(m+n)p(\phi)+2\Delta_m\phi}}\mathfrak{L}^m\mathbb{1}_{[A]}(x)\mathfrak{L}^n\mathbb{1}_{[TB'\times B'']}(x) \\
&\geq \frac{1}{c_1^3Pe^{2\Delta_m\phi}}\eta([A])\eta([TB'\times B'']) \\
&\geq \frac{1}{c_1^3Pe^{2h(\operatorname{diam}E)}}\eta([A])\eta([TB'\times B'']).
\end{aligned}$$

The constant

$$c_2 = \frac{1}{c_1^3Pe^{2h(\operatorname{diam}E)}}$$

satisfies the claims of the lemma. To prove it we'll first extend this last estimate to allow any $B'' \in \mathscr{B}(X)$. Use S-invariance of η and surjectivity of T to see that for any $l \geq 1$

$$\begin{aligned}
\eta\left([A\times B'\times B'']\right) &\geq c_2\eta([A])\eta([T(B')\times B'']) \\
&= c_2\eta([A])\eta([E^l\times T^{-1}(T(B'))\times B'']) \\
&\geq c_2\eta([A])\eta([E^l\times B'\times B'']).
\end{aligned} \tag{23}$$

Now consider two measures on $\mathscr{B}(X)$:

$$m_1(B) = \eta\left([A\times B']\cap\left(E^{n+1}\times B\right)\right) \tag{24}$$

$$m_2(B) = \eta\left([T(B')]\cap(E\times B)\right) \tag{25}$$

Observe that $m_1 << m_2$, because of the set containment

$$\begin{aligned}
[A\times B']\cap(E^{n+1}\times B) &\subseteq [E^n\times B']\cap(E^{n+1}\times B) \\
&\subseteq S^{-n}\left([T(B')]\cap(E\times B)\right)
\end{aligned}$$

and S-invariance of η. In terms of these two auxiliary measures Estimate (23) shows that for any $k \geq 1$ and any set $B \in \mathscr{B}_k$

$$m_1(B) \geq c_2\eta([A])m_2(B)$$

holds; this suffices to conclude $\frac{dm_1|_{\mathscr{B}_k}}{dm_2|_{\mathscr{B}_k}} \geq c_2\mu([A])\ m_2$ -a.e.. Now Theorem 35.7 of [2] (application of the martingale convergence theorem to Radon-Nikodym derivatives) tells us that $\frac{dm_1|_{\mathscr{B}_k}}{dm_2|_{\mathscr{B}_k}} \to \frac{dm_1}{dm_2}$ m_2-a.e. as $k \to \infty$ and therefore $\frac{dm_1}{dm_2} \geq c_2\mu([A])$ m_2-a.e., i.e.

$$m_1(B) \geq c_2\eta([A])m_2(B)$$

holds for every $B \in \mathscr{B}(X)$, which is this lemma's statement. □

Proposition 4.7. *If η is an S-invariant Gibbs state for a continuous potential $\phi : X \to \mathbb{R}$, then it is totally ergodic, i.e. $(X, \mathscr{B}, \eta, S^n)$ is an ergodic measure preserving dynamical system for every iterate $n \geq 1$.*

Proof. Suppose to the contrary that there was an iterate $n \geq 1$ for which S^n was not η ergodic. There would be a set $B_S \in \mathscr{B}$ of η non-trivial, non-full measure for which

$$\begin{aligned} B_S = S^{-n}B_S &= \{(a_1 \dots a_n bx) : a_1, \dots, a_n, b \in E,\ x \in X,\ \text{and } (Tb, x) \in B_S\} \\ &= E^n \times \{(b, x) : b \in E,\ x \in X,\ \text{and } (Tb, x) \in B_S\} \\ &=: E^n \times \hat{B}_S,\ \mathrm{mod}\,\eta. \end{aligned}$$

Although we won't need it, we can give an explicit description $\hat{B}_S = \sigma(S^{-1}(B))$, where σ is the usual left shift. More generally, iterating S^n $j \geq 1$ times leaves

$$B_S = S^{-jn}B_S = E^{jn} \times \hat{B}_S,\ \mathrm{mod}\,\eta.$$

The function

$$m := A \mapsto \eta(A \cap B_S)$$

defines a positive Borel measure on X, and clearly $m << \eta$. For any $j \geq 1$ and any $A = A' \times E^{\mathbb{N}} \in \mathscr{B}_{jn-1}$

$$\begin{aligned} m(A) &= \eta\left([A'] \cap \left(E^{jn} \times \hat{B}_S\right)\right) \\ &= \eta([A']) - \eta\left([A'] \cap \left(E^{jn} \times \hat{B}_S{}^{\complement}\right)\right) \\ &= \eta(A) - \eta\left(\pi^{-1}_{1\dots jn-1}(A') \cap \pi^{-1}_{jn}(E) \cap \pi^{-1}_{jn+1\dots}\left(\hat{B}_S{}^{\complement}\right)\right) \\ \left(\begin{array}{c}\text{by lemma}\\ 4.6\end{array}\right) &\leq \eta(A)\left(1 - c_2\eta\left(\pi^{-1}_1(TE) \cap \pi^{-1}_{2\dots}\left(\hat{B}_S{}^{\complement}\right)\right)\right) \\ \left(\begin{array}{c}\text{by surjectivity}\\ \text{of } T\end{array}\right) &= \eta(A)\left(1 - c_2\eta\left(E \times \hat{B}_S{}^{\complement}\right)\right) \\ &= \eta(A)\left(1 - c_2\eta\left((B_S)^{\complement}\right)\right). \end{aligned}$$

Since this holds for every $j \geq 1$ and every $\mathscr{B}_{jn-1}$-measurable set A, this bound extends to the ($\mathscr{B}_{jn-1}$-measurable) density between the restrictions of the respective measures:

$$\frac{dm|_{\mathscr{B}_{jn-1}}}{d\eta|_{\mathscr{B}_{jn-1}}} \leq 1 - c_2\eta\left((B_S)^{\complement}\right),\ \eta - a.e.$$

Again invoke [2] Theorem 35.7 to conclude that the densities between the restricted measures converge $\eta - a.e.$ to $\frac{dm}{d\eta}$, and therefore that $m(A) \leq \eta(A)\left(1 - c\eta\left((B_S)^{\complement}\right)\right)$ holds for every $A \in \mathscr{B}$. We have arrived at the contradiction

$$\eta(B_S) = m(B_S) \leq \eta(B_S)\left(1 - c_2\eta\left((B_S)^{\complement}\right)\right) < \eta(B_S),$$

which finishes the proof. □

5. Variational Principle and Equilibrium States.

For $\mu \in C^*_{+,1,S}$ let

$$H(\mu \circ \pi^{-1}_{1\ldots n}|\nu^n) = \mu\left(-\log\left(\frac{d\mu \circ \pi^{-1}_{1\ldots n}}{d\nu^n}\right) \circ \pi_{1\ldots n}\right)$$

if the n^{th} projection of μ is absolutely continuous with respect to ν^n, and $-\infty$, otherwise. Basic entropy theory (see [7], section III.4) states that

$$\mathcal{H}(\mu|\nu^{\mathbb{N}}) \equiv \lim_{n\to\infty} \frac{H(\mu \circ \pi^{-1}_{1\ldots n}|\nu^n)}{n} = \inf_{n\geq 1} \frac{H(\mu \circ \pi^{-1}_{1\ldots n}|\nu^n)}{n} \in [-\infty, 0]. \tag{26}$$

For reasons that should become clear, we add a term involving the transfer of ν under T and call it the *modified entropy*:

$$\mathcal{H}_S(\mu|\nu^{\mathbb{N}}) \equiv \mathcal{H}(\mu|\nu^{\mathbb{N}}) + \int_E \log\frac{d\nu \circ T}{d\nu}\, d\mu \circ \pi_2^{-1}. \tag{27}$$

Finally define the modified Gibbs free energy for ϕ as

$$C^*_{+,1,S} \ni \mu \mapsto G_\phi(\mu) \equiv \mu(\phi) + \mathcal{H}_S(\mu|\nu^{\mathbb{N}}). \tag{28}$$

Although this definition doesn't really require μ to be S-invariant, the relationship between $p(\phi)$ and $G(\mu)$ stated in the variational principle depends on it. And if G_ϕ attains its supremum at some invariant state μ, then μ is called an *equilibrium state* for ϕ.

Theorem 5.1. *Variational Principle. For all continuous $\phi : X \to \mathbb{R}$ the number $p(\phi)$ is an upper bound on G_ϕ, and if $\phi : X \to \mathbb{R}$ satisfies Dini's condition then*

$$p(\phi) = \sup G_\phi$$

and every S-invariant Gibbs state for ϕ is an equilibrium state for ϕ.

Proof. (That $p(\phi)$ is an upper bound.) Partition the alphabet E into finitely many measurable sets $W \in \mathcal{W}$ of diameter $< r_0$. For each $W \in \mathcal{W}$ label a full set of inverse branches $\{T^{-1}_{W,i}\}_{i=1}^{j_W}$; they are defined on some ball of radius r_0 containing W. The branches of the preimage of a given one of these balls are disjoint. So, for each $W \in \mathcal{W}$, $T^{-1}W$ is the disjoint union of $\{T^{-1}_{W,i}W\}_{i=1}^{j_W}$ and therefore $\left\{T^{-1}_{W,i}(W) : W \in \mathcal{W},\, 1 \leq i \leq j_W\right\}$ is a refinement of the partition $T^{-1}\mathcal{W}$.

Recall from Lemma 1.2 that each $W \in \mathcal{W}$ is homeomorphic to $T(W)$ via $T|_W$, and therefore $\nu|_W \circ T|_W$ is a well defined finite measure on $(W, \mathscr{B}(W))$. Then "globalize" the measure by defining

$$\nu \circ T(A) = \sum_{W\in\mathcal{W}} \nu|_W \circ T|_W(W \cap A), \tag{29}$$

a probability measure on $(E, \mathscr{B}(E))$.

Observe that for any $n \geq 1$

$$\mathcal{W}_n = \{E \times W_2 \times \cdots \times W_n \,:\, W_k \in \mathcal{W},\, 2 \leq k \leq n\}$$

is a partition of E^n and calculate

$$\begin{aligned}
\mathfrak{L}^n \mathbb{1}(x) &= \sum_{\mathcal{W}_n} \int_{a \in E \times W_2 \times \ldots W_n} \sum_{(i_2,\ldots,i_n)=(1,\ldots,1)}^{(j_{W_2},\ldots,j_{W_n})} \sum_{b \in T^{-1}x_1} \cdot \\
&\quad \cdot\, e^{s_n \phi}(a_1 T^{-1}_{W_2,i_2}(a_2) \ldots T^{-1}_{W_n,i_n}(a_n) b y_2 \ldots)\, d\nu^n(a) \\
&= \sum_{b \in T^{-1}x_1} \sum_{\mathcal{W}_n} \sum_{(i_2,\ldots,i_n)=(1,\ldots,1)}^{(j_{W_2},\ldots,j_{W_n})} \int_{a \in E \times T^{-1}_{W_2,i_2}(W_2) \times \cdots \times T^{-1}_{W_n,i_n}(W_n)} \cdot \\
&\quad \cdot\, e^{s_n \phi}(a_1 \ldots a_n b x_2 \ldots)\, d\nu \otimes (\nu \circ T)^{n-1}(a) \\
&= \sum_{b \in T^{-1}x_1} \int_{a \in E^n} e^{s_n \phi}(a b x_2 \ldots) d\nu \otimes (\nu \circ T)^{n-1}(a)
\end{aligned}$$

$$\begin{pmatrix} \text{for} \\ \text{any} \\ y \in X \end{pmatrix} \quad \geq e^{-\Delta_n \phi} \int_{a \in E^n} e^{s_n \phi}(a y) d\nu \otimes (\nu \circ T)^{n-1}(a). \tag{30}$$

Notice that the S invariance of μ, along with the identities

$$\pi_1 \circ S = T \circ \pi_2,$$

$$\pi_k \circ S = \pi_{k+1} \;\forall\; k \geq 2,$$

now imply

$$\mu \circ \pi_1^{-1} = \mu \circ \pi_2^{-1} \circ T^{-1}$$

and for all $k \geq 2$

$$\mu \circ \pi_2^{-1} = \mu \circ \pi_k^{-1}. \tag{31}$$

Using Fubini's theorem and the fact that the density between product measures is the product of the densities, this goes to show that

$$\int_{E^n} \log \frac{d\nu \otimes (\nu \circ T)^{n-1}}{d\nu^n}\, d\mu \circ \pi_{1\ldots n}^{-1} = \sum_{k=2}^{n} \int_E \log \frac{d\nu \circ T}{d\nu}\, d\mu \circ \pi_k^{-1} =$$

$$= (n-1) \int_E \log \frac{d\nu \circ T}{d\nu}\, d\mu \circ \pi_2^{-1}.$$

For $n \geq 1$ and $y \in X$ define the continuous map

$$\pi_{1\ldots n,y}^{-1} : E^n \to X \quad \text{by} \quad a \mapsto (a, y).$$

$$\phi_k^y := X \ni x \mapsto \phi(x_1 \ldots x_k y),$$

By change of variables and S-invariance of μ it holds for any indices $0 \le j \le n-1$ that

$$\begin{aligned}
\int_{E^n} \phi \circ S^j \circ \pi^{-1}_{1\ldots n,y}\, d\mu \circ \pi^{-1}_{1\ldots n} &= \int_X \phi \circ S^j \circ \pi^{-1}_{1\ldots n,y} \circ \pi_{1\ldots n}\, d\mu \\
&= \int_X \phi \circ \pi^{-1}_{1\ldots n,y} \circ \pi_{1\ldots n-j} \circ S^j\, d\mu \\
&= \int_{E^{n-j}} \phi \circ \pi^{-1}_{1\ldots n-j,y}\, d\mu \circ \pi^{-1}_{1\ldots n-j} \\
&\to \int \phi\, d\mu \quad \text{as} \quad n-j \to \infty.
\end{aligned}$$

The argument for the convergence is that of the convergence of Riemann sums and holds for continuous ϕ.

Since we have just described a convergent sequence of numbers, the sequence of averages must converge to the same limit:

$$\begin{aligned}
\frac{1}{n}\int_{a\in E^n} s_n\phi(ay)\, d\mu \circ \pi^{-1}_{1\ldots n}(a) &= \frac{1}{n}\sum_{j=0}^{n-1}\int_{a\in E^n} \phi \circ S^j(ay)\, d\mu \circ \pi^{-1}_{1\ldots n}(a) \\
&= \frac{1}{n}\sum_{j=0}^{n-1}\int_{a\in E^{n-j}} \phi(ay)\, d\mu \circ \pi^{-1}_{1\ldots n-j}(a) \\
&\to \int \phi\, d\mu \text{ as } n \to \infty.
\end{aligned}$$

It is now clear that we can calculate the modified Gibbs free energy as

$$G(\nu,\phi,\mu) = \lim_{n\to\infty} \frac{1}{n}\int_{E^n} s_n\phi \circ \pi^{-1}_{1\ldots n,y} - \log\frac{d\mu\circ\pi^{-1}_{1\ldots n}}{d\nu^n} + \log\frac{d\nu\otimes(\nu\circ T)^{n-1}}{d\nu^n}\ \ d\mu\circ\pi^{-1}_{1\ldots n}. \tag{32}$$

Using $\mu \circ \pi^{-1}_{1\ldots n} << \nu^n$ and consolidating the logs we rewrite the integral as

$$\begin{aligned}
&\int_{E^n} \log\left(\frac{e^{s_n\phi}(\centerdot, y)\frac{d\nu\otimes(\nu\circ T)^{n-1}}{d\nu^n}}{\frac{d\mu\circ\pi^{-1}_{1\ldots n}}{d\nu^n}}\right)\frac{d\mu\circ\pi^{-1}_{1\ldots n}}{d\nu^n}\ d\nu^n \\
\left(\begin{array}{c}\text{by}\\ \text{Jensen's}\\ \text{inequality}\end{array}\right) &\le \log\int_{E^n} e^{s_n\phi}(\centerdot, y)\frac{d\nu\otimes(\nu\circ T)^{n-1}}{d\nu^n}\, d\nu^n \\
\left(\begin{array}{c}\text{by}\\ \text{estimate}\\ 30\end{array}\right) &\le \Delta_n\phi + \log\mathfrak{L}^n\mathbb{1}(x)
\end{aligned}$$

for any $x \in X$. Dividing by n and taking the limit, the term $\Delta_n\phi$ vanishes by first claim in Lemma 2.2. This half of the variational principle is proved.

That $p(\phi)$ is the least upper bound, given Dini's condition on ϕ.

By definition (see Formula (4.1)), if γ is a Gibbs state for ϕ and $A = A_1 \times \cdots \times A_n \in \mathscr{B}(E^n)$ then

$$\frac{1}{c_1 e^{np(\phi)}} \leq \frac{\frac{\gamma\circ\pi_{1\dots n}^{-1}(A)}{\nu^n(A)}}{\frac{\nu^n(A_1\times TA_2\times\cdots\times TA_n)}{\nu^n(A)}\frac{\mathfrak{L}^n\mathbb{1}_{[A]}(x)}{\nu^n(A_1\times TA_2\times\cdots\times TA_n)}} \leq \frac{c_1}{e^{np(\phi)}} \tag{33}$$

holds for any $x \in X$.

To see why we've made a simple expression complicated, recall Equation (18). If $\mathrm{diam} A_k < \delta$ for each $2 \leq k \leq n$ in that formula then T injects on each A_k i.e. $T^{-1}a_k \cap A_k$ is a singleton for each a_k in the domain of integration TA_k. Calling $T^{-1}a_k \cap A_k \equiv \{b_{k-1}\}$ for $2 \leq k \leq n$ reduces equation 18 to

$$\mathfrak{L}^n\mathbb{1}_{[A]}(x) = \int_{a\in E\times TA_2\times\cdots\times TA_n} \sum_{b_n\in T^{-1}x_1} e^{s_n\phi}(a_1b_1\dots b_nx_2\dots)\,d\nu^n(a). \tag{34}$$

Taking each set A_k, $1 \leq k \leq n$ to be a ball of small radius around a singleton a'_k, we have

$$\frac{\mathfrak{L}^n\mathbb{1}_{[A]}(x)}{\nu^n(E\times TA_2\times\cdots\times TA_n)} \to \sum_{b_n\in T^{-1}x_1} e^{s_n\phi}(a_1\dots a_nb_nx_2\dots).$$

as all the radii shrink to 0.

Regarding the other two factors in the middle of Estimate (33), it is known that $\nu\circ T << \nu$ and $\gamma\circ\pi_{1\dots n}^{-1} << \nu^n$. Therefore

$$\frac{1}{c_1 e^{np(\phi)}} \leq \frac{\frac{d\gamma\circ\pi_{1\dots n}^{-1}}{d\nu^n}(a')}{\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}(a'_2\dots a'_n)\sum\limits_{b_n\in T^{-1}x_1} e^{s_n\phi}(a'b_nx_2\dots)} \leq \frac{c_1}{e^{np(\phi)}} \tag{35}$$

holds for ν^n-a.e. $a' \in E^n$.

By the right hand side

$$\begin{aligned} np(\phi) - \log c_1 &\leq -\log\frac{d\gamma\circ\pi_{1\dots n}^{-1}}{d\nu^n}(a') + \log\sum_{b_n\in T^{-1}x_1} e^{s_n\phi}(a'b_nx_2\dots) + \\ &\quad + \log\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}(a'_2\dots a'_n) \\ \begin{pmatrix}\text{for}\\ \text{any}\\ y\in X\end{pmatrix} &\leq -\log\frac{d\gamma\circ\pi_{1\dots n}^{-1}}{d\nu^n}(a') + s_n\phi(a'y) + \log P + \Delta_n\phi + \\ &\quad + \log\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}(a'_2\dots a'_n). \end{aligned}$$

Integrating the last line $d\gamma\circ\pi_{1\dots n}^{-1}(a')$ and dividing by n results in a quantity that differs by only $\frac{1}{n}\Delta_n\phi$ from the nth term of the right hand side of Equation (32) (taking $\mu = \gamma$). It was already shown that those terms converge to

$$G(\gamma) = \gamma(\phi) + \mathcal{H}_S(\gamma|\nu^{\mathbb{N}}),$$

as $n \to \infty$, and by Lemma 2.2 the difference $\frac{1}{n}\Delta_n\phi$ vanishes in the limit.

This finishes the proof that every S-invariant Gibbs state for a continuous $\phi : X \to \mathbb{R}$ is also an equilibrium state for ϕ. In light of Fact 4.5, this finishes the theorem. □

If ϕ is continuous but does not satisfy Dini's condition (8) then the upper bound on the free energy still holds, but we are not guaranteed the existence of an S-invariant Gibbs state to prove that it is the least upper bound. In fact, Dini's condition was not used in the proof of Theorem 5.1 in any other way than to ensure the existence of an invariant Gibbs state. It follows that if ϕ fails to satisfy Dini's condition but is continuous and happens to have an invariant Gibbs state, then that state is still an equilibrium state.

We turn to the question of uniqueness. As a matter of fact uniqueness of S-invariant Gibbs states for Dini potential functions has already been proven. It follows from Theorem 4.7 and the facts that all Gibbs measures must be boundedly equivalent while distinct ergodic measures must be mutually singular. However it still seems possible that there could be another type of equilibrium state which is not a Gibbs state. The following theorem rules this out.

Theorem 5.2. *If ϕ does satisfy Dini's condition, then the S-invariant Gibbs state $\eta = \frac{1}{\rho\gamma(\mathbb{1})}\rho\gamma$ constructed in Lemma 3.3 is in fact the only equilibrium state for ϕ.*

Proof. It suffices to demonstrate that there is only one ergodic equilibrium state, because the following three facts show that for a continuous potential function ϕ the set of all equilibrium states for ϕ is the weakly* closed convex hull of the set of ergodic equilibrium states for ϕ.

(1) The set of all equilibrium states for a continuous potential ϕ is convex and w* closed (therefore w* compact). This follows from w* upper semi-continuity and affinity of the entropy $\mathcal{H}(\mu|\nu^{\mathbb{N}})$ and w* continuity and linearity of $\mu \mapsto \mu(\phi)$ and $\mu \mapsto \int_E \frac{d\nu\circ T}{d\nu} d\mu \circ \pi_2^{-1} = \mu\left(\frac{d\nu\circ T}{d\nu} \circ \pi_2\right)$.
(2) By (1) and the Krein-Milman theorem the set of all equilibrium states for $\phi \in C$ is the w* closed convex hull of its extreme points.
(3) Also owing to affinity of the entropy, the extreme points of the set of equilibrium states for $\phi \in C$ must all be ergodic.

Since distinct ergodic measures are mutually singular and since we already know that the measure η of Fact 4.5 is an ergodic equilibrium state (by Theorem 4.7 and the latter half of Theorem 5.1), it suffices to prove that no equilibrium state can be mutually singular against $\eta = \rho\gamma$, where ρ is normalized to make η a probability measure.

In the first place observe that, because the entropy $\mathcal{H}(\mu|\nu^{\mathbb{N}})$ is the infimum of its finite stage approximations, for any equilibrium state μ for ϕ

$$\begin{aligned} p(\phi) &\leq \mu(\phi) + \frac{1}{n}\int_{E^n} -\log\left(\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\right) d\mu\circ\pi_{1\ldots n}^{-1} + \int_E \log\frac{d\nu\circ T}{d\nu}\, d\mu\circ\pi_2^{-1} \\ &= \mu\left(\frac{1}{n}s_n\phi - \frac{1}{n}\log\left(\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\right)\circ\pi_{1\ldots n} + \frac{1}{n-1}\log\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\circ\pi_{2\ldots n}\right). \end{aligned}$$

Multiplying through by n and rearranging yields

$$0 \leq \mu\left(\log\left(e^{s_n\phi - np(\phi)}\,\frac{\left(\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\right)^{\frac{n}{n-1}}\circ\pi_{2\ldots n}}{\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\circ\pi_{1\ldots n}}\right)\right). \tag{36}$$

To introduce $\eta = \rho\gamma$ use the left hand side of Estimate (35), replacing γ with η, to obtain that for every $x \in X$ and $a \in E^n$ and $b \in T^{-1}(x_1)$

$$e^{s_n\phi}(abx_2\ldots) \leq \sum_{b'\in T^{-1}(x_1)} e^{s_n\phi}(ab'x_2\ldots) \leq c_1 e^{np(\phi)}\frac{\frac{d\eta\circ\pi_{1\ldots n}^{-1}}{d\nu^n}(a)}{\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}(a_2\ldots a_n)}.$$

Using this in Estimate (36) yields

$$0 \leq \mu\left(\log c_1 + \log\left(\frac{\frac{d\eta\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\circ\pi_{1\ldots n}}{\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\circ\pi_{2\ldots n}}\;\frac{\left(\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\right)^{\frac{n}{n-1}}\circ\pi_{2\ldots n}}{\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\circ\pi_{1\ldots n}}\right)\right),$$

wherein the big logarithm can be split

$$-\log\frac{\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}}{\frac{d\eta\circ\pi_{1\ldots n}^{-1}}{d\nu^n}}\circ\pi_{1\ldots n} + \frac{1}{n-1}\log\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\circ\pi_{2\ldots n}.$$

Recall by Formula (31) that

$$\mu\left(\log\frac{d(\nu\circ T)^{n-1}}{d\nu^{n-1}}\circ\pi_{2\ldots n}\right) = (n-1)\mu\left(\log\frac{d\nu\circ T}{d\nu}\circ\pi_2\right)$$

and because $\eta\circ\pi_{1\ldots n}^{-1} << \nu^n$

$$\frac{\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}}{\frac{d\eta\circ\pi_{1\ldots n}^{-1}}{d\nu^n}} = \frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\nu^n}\,\frac{d\nu^n}{d\eta\circ\pi_{1\ldots n}^{-1}} = \frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\eta\circ\pi_{1\ldots n}^{-1}}.$$

Thus for any equilibrium state μ for ϕ and all $n \geq 1$ we've proved

$$\mu\left(\log\frac{d\mu\circ\pi_{1\ldots n}^{-1}}{d\eta\circ\pi_{1\ldots n}^{-1}}\circ\pi_{1\ldots n}\right) \leq \log c_1 + \mu\left(\log\frac{d\nu\circ T}{d\nu}\circ\pi_2\right). \tag{37}$$

This is a statement of the boundedness of density between an arbitrary equilibrium measure μ and an S-invariant Gibbs measure η.

The remainder of the proof is to show that if there were an equilibrium state $\mu \perp \eta$, then it would have to satisfy

$$\frac{d\mu|_{\mathscr{B}_n}}{d\eta|_{\mathscr{B}_n}} \to \infty\ \mu-\text{a.e.}\ \text{ as }\ n\to\infty \tag{38}$$

and hence, invoking Fatou's lemma,

$$\mu\left(\log\frac{d\mu|_{\mathscr{B}_n}}{d\eta|_{\mathscr{B}_n}}\right) \to \infty\ \text{ as }\ n\to\infty.$$

The contradiction of Divergence (38) with Estimate (37) will prove the theorem.

To prove (38) fix a number $t > 0$ and let, for every $n \geq 1$,

$$\Omega_n(t) = \left\{ \frac{d\mu|_{\mathscr{B}_n}}{d\eta|_{\mathscr{B}_n}} \leq t \right\}.$$

Then let

$$\Omega(t) = \cap_{n\geq 1} \cup_{m\geq n} \Omega_m(t),$$

i.e. the set where $\frac{d\mu|_{\mathscr{B}_n}}{d\eta|_{\mathscr{B}_n}}$ is infinitely often $\leq t$. We will show that

$$1 = \mu\left(\Omega(t)^{\complement}\right), \tag{39}$$

i.e. the set where $\frac{d\mu|_{\mathscr{B}_n}}{d\eta|_{\mathscr{B}_n}}$ is eventually $> t$ has full μ measure. This will imply convergence (38).

Let $n_0 \geq 1$ and consider any set $A \in \mathscr{B}_{n_0}$ with $A \subseteq \Omega(t)$. Apply the typical disjointing to the n_0th intersectand of $\Omega(t)$:

$$\cup_{m\geq n_0}\Omega_m(t) = \cup_{m\geq n_0}\left(\Omega_m(t) \setminus \cup_{n_0\leq l<m}\Omega_l(t)\right) \equiv \cup_{m\geq n_0}\Omega^{\bullet}_m(t).$$

Use the fact that for all $m \geq n_0$

$$A \cap \Omega^{\bullet}_m(t) \in \mathscr{B}_m$$

to estimate

$$\begin{aligned}
\mu(A) &= \sum_{m\geq n_0} \mu\left(A \cap \Omega^{\bullet}_m(t)\right) \\
&= \sum_{m\geq n_0} \mu|_{\mathscr{B}_m}\left(A \cap \Omega^{\bullet}_m(t)\right) \\
&= \sum_{m\geq n_0} \eta|_{\mathscr{B}_m}\left(\frac{d\mu|_{\mathscr{B}_m}}{d\eta|_{\mathscr{B}_m}} \mathbb{1}_{A\cap\Omega^{\bullet}_m(t)}\right) \\
&\leq \sum_{m\geq n_0} t\eta\left(A \cap \Omega^{\bullet}_m(t)\right) \\
&= t\eta(A).
\end{aligned}$$

Let us see for how large a class of sets this estimate

$$\mu(A) \leq t\eta(A)$$

remains valid. For each $n \geq 1$ let

$$\mathscr{B}_n \cap \Omega(t) \equiv \{A \in \mathscr{B}_n : A \subseteq \Omega(t)\};$$

it is a σ-algebra of subsets of $\Omega(t)$. The finite unions of sets from

$$\cup_{n\geq 1}\mathscr{B}_n \cap \Omega(t)$$

therefore form an algebra of subsets of $\Omega(t)$, call it $\mathscr{A}(t)$. Notice $\sigma(\mathscr{A}(t)) = \mathscr{B}(X) \cap \Omega(t)$. This is because $\mathscr{A}(t)$ contains all the basic open sets in the restriction of the product topology on $X = E^{\mathbb{N}}$ to $\Omega(t)$, the product topology coincides with our (compact and hence separable) metric topology and is therefore 2nd countable, and hence any σ-algebra containing the basic open sets contains all the open sets and hence the Borel σ-algebra.

Now, the class

$$\mathscr{C} \equiv \{A \in \mathscr{B}(X) \cap \Omega(t) : \mu(A) \le t\eta(A)\}$$

is a monotone class and we have just finished showing that it contains the algebra $\mathscr{A}(t)$. The monotone class theorem then states that the estimate $\mu(A) \le t\eta(A)$ holds for all $A \in \sigma(\mathscr{A}(t)) = \mathscr{B}(X) \cap \Omega(t)$.

Now we finish the proof: if $\mu \perp \eta$ then there would be a set $F \in \mathscr{B}(X)$ with μ concentrated on F and η concentrated on $F^{\complement}$. It would follow that

$$\mu(\Omega(t)) = \mu(\Omega(t) \cap F) \le t\eta(\Omega(t) \cap F) = 0,$$

which proves Statement (39) and thereby proves Statement (38) and finishes the proof. □

Corollary 5.3. *If ϕ satisfies Dini's condition then the eigenvalue $e^{p(\phi)}$ has one dimensional eigenspaces for both the transfer operator $\mathfrak{L}$ and the dual operator $\mathfrak{L}^*$. This means there is a unique normalized eigenvector $\gamma \in C^*_{+,1}$ for which $\mathfrak{L}^*_0\gamma = \gamma$ and a unique $\rho \in C$ for which $\gamma(\rho) = 1$ and $\mathfrak{L}_0\rho = \rho$.*

In other words there are well defined maps

$$\phi \mapsto \gamma_\phi \in C^*_{+,1} \tag{40}$$

$$\phi \mapsto \rho_\phi \in C \tag{41}$$

for the components of the unique equilibrium state $\eta_\phi = \rho_\phi\gamma_\phi$ for ϕ.

Proof. Suppose there are $\gamma \ne \gamma' \in C^*_{+,1}$ such that $\mathfrak{L}^*\gamma = e^{p(\phi)}\gamma$ and $\mathfrak{L}^*\gamma' = e^{p(\phi)}\gamma'$. Take any positive function ρ as constructed in Lemma 3.2 that satisfies $\mathfrak{L}\rho = e^{p(\phi)}\rho$. Without loss of generality let $0 < \rho\gamma(\mathbb{1}) \le \rho\gamma'(\mathbb{1})$. There must be a Borel set A such that $\gamma(A) \ne \gamma'(A)$, and by replacing A with $A^{\complement}$ if necessary we may assume $\gamma(A) > \gamma'(A)$. It follows that $\rho\gamma(A) > \rho\gamma'(A)$ and so too

$$\frac{\rho\gamma(A)}{\rho\gamma(\mathbb{1})} > \frac{\rho\gamma'(A)}{\rho\gamma'(\mathbb{1})}.$$

But now we have two distinct positive normalized functionals $\frac{\rho\gamma}{\rho\gamma(\mathbb{1})}$, $\frac{\rho\gamma'}{\rho\gamma'(\mathbb{1})}$ which have both been constructed to be equilibrium states for ϕ, according to Fact 4.5 and the latter half of Theorem 5.1. This contradicts Theorem 5.2.

Now that we know γ is unique let ρ be constructed as in Lemma 3.2 and suppose there is a continuous function $\rho' \ne \rho$ satisfying $\mathfrak{L}\rho' = e^{p(\phi)}\rho'$. If ρ' is a scaling of ρ then we are done, and otherwise we can conclude that scaling the functions won't make them equal:

$$\frac{\rho}{\rho\gamma(\mathbb{1})} \ne \frac{\rho'}{\rho'\gamma(\mathbb{1})}.$$

By a standard argument of measure theory it follows that the respective weightings of γ are also distinct:

$$\frac{\rho\gamma}{\rho\gamma(\mathbb{1})} \ne \frac{\rho'\gamma}{\rho'\gamma(\mathbb{1})}.$$

Again this contradicts the uniqueness of equilibria stated in Theorem 5.2. □

6. Stochastic Laws for Hölder continuous ϕ and Hölder continuous observable

For $0 < \alpha \leq 1$ let $\mathcal{H}_\alpha$ be the space of locally Hölder continuous functions

$$f : X \to \mathbb{R} \text{ with } \sup_{d(x,y)<2r_0} \frac{|f(x) - f(y)|}{d(x,y)^\alpha} < \infty.$$

Assume in this section that $\operatorname{diam}(E) < 1$, which causes no loss of generality and leads to the nice property that $\mathcal{H}_\alpha$ increases (in the sense of set containment) as $0 < \alpha < 1$ decreases.

In this section we'll see that if the potential function ϕ is Hölder continuous then for any Hölder continuous observable function f. the stochastic process $\{f \circ S^n\}_{n\geq 1}$ exhibits exponential decay of correlations, satisfies the central limit theorem, and, in the case $T : E \to E$ is continuous and bijective, i.e. is a homeomorphism, the law of the iterated logarithm.

6.1. **The Theorem of Ionescu Tulcea and Marinescu.**

All the probabilistic laws stated above depend on a decomposition of the transfer operator according to a famous theorem published in [5]. The conclusions of the theorem are stated at the beginning of the proof of Proposition 6.6, where they are used. There are four hypotheses, which we address next. As a matter of terminology, we say that an operator O on a Banach space $(B, \| \|_B)$ *acts with the Romanian inequality* (or two-norm inequality) on a Banach space $(C, \| \|_C)$, where $C \subseteq B$ if constants $R > 0$ and $0 < r < 1$ can be found so that

$$\|Oc\|_C \leq R\|c\|_B + r\|c\|_C$$

holds for all vectors $c \in C$. Actually this theorem applies to any Banach spaces B, C for which C is dense as a subset of $(B, \| \|_B)$, but we are here only concerned with the specific case $B = C(X),\ C = H_\alpha$.

Ionescu Tulcea and Marinescu's theorem requires

1. If $0 < k < \infty$, $f_n \in H_\alpha$ with $\|f_n\|_\alpha \leq k$ for every $n \geq 1$, and $f_n \to f$ in $C(X), \| \|_\infty$, then $f \in H_\alpha$ with $\|f\|_\alpha \leq k$, too.
2. Considered as an operator on H_α, the set of norms of iterates $\{\|\mathfrak{L}_0^n\|\}_{n\geq 1}$ is bounded.
3. Some iterate of $\mathfrak{L}_0$ acts with the Romanian two-norm inequality on H_α.
4. $\mathfrak{L}_0$ is a compact transformation from H_α to $C(X)$, i.e. if $A \subset H_\alpha$ is bounded in $\| \|_\alpha$ then $\mathfrak{L}_0(A)$ is conditionally compact in $\| \|_\infty$.

Condition (1) is simply a property of metric topology. Because $\mathfrak{L}_0$ is a positive operator, Condition (2) follows from Estimate (16) in the proof of Lemma 3.2 that $\{\|\mathfrak{L}_0^n \mathbb{1}\|_\infty\}_{n\geq 1}$ is bounded. Condition (3) is the subject of the following lemma, and with the Arzela-Ascoli theorem Condition (4) follows from Estimate (44) below.

Lemma 6.1. *If $\phi \in \mathcal{H}_\beta$ for some $0 < \beta < 1$ then for every $0 < \alpha \leq \beta$ there is an iterate $\mathfrak{L}_0^n$ of ϕ's normalized transfer operator which acts with the Romanian inequality (or two norm inequality) on the Banach subspace $\mathcal{H}_\alpha \subset (C(X), \| \|_\infty)$.*

Proof. First of all we claim that $\alpha \leq \beta$ guarantees H_α to be invariant under $\mathfrak{L}_0$. This will be justified by Estimate (44) below. For this proof we recall Lemma 2.6 and sharpen its statement with the added strength of Hölder continuity. First of all divide Estimate (13) (in the statement of Lemma 2.6) by λ^n to obtain expressions for $\mathfrak{L}_0^n$ from those for $\mathfrak{L}^n$. Use the fact that $\mathfrak{L}_0^n 1\!\!1(x) \leq Pe^{h(\operatorname{diam}E)}$ (see Estimate (16)) independently of n and x to obtain

$$|\mathfrak{L}_0^n f(x) - \mathfrak{L}_0^n f(y)| \leq Pe^{h(\operatorname{diam}E)} \left(\mathfrak{m}(f, q^n \tilde{\kappa} d(x,y)) + \|f\|_\infty |1 - e^{h(\tilde{\kappa} d(x,y))}|\right). \tag{42}$$

Recall the norm $\| \ \|_\alpha = \mathscr{V}_\alpha + \| \ \|_\infty$ which makes H_α into a Banach space. Of course the Hölder property yields immediately that $\mathfrak{m}(f,s) \leq \mathscr{V}_\alpha(f) s^\alpha$ and $\mathfrak{m}(\phi, s) \leq \mathscr{V}_\beta(\phi) s^\beta$. This latter estimate implies that

$$\begin{aligned} h(t) &= \sum_{n=1}^\infty \mathfrak{m}(\phi, tq^n) \\ &\leq \mathscr{V}_\beta(\phi) \sum_{n=1}^\infty t^\beta q^{\beta n}. \end{aligned}$$

Use all of this and the fact $|1 - e^x| \leq |x| e^{|x|}$ to update Estimate (42) to read

$$|\mathfrak{L}_0^n f(x) - \mathfrak{L}_0^n f(y)| \leq Pe^{h(\operatorname{diam}E)} \left(\mathscr{V}_\alpha(f) q^{\alpha n} (\tilde{\kappa} d(x,y))^\alpha + \|f\|_\infty \mathscr{V}_\beta(\phi) (\tilde{\kappa} d(x,y))^\beta \sum_{n=1}^\infty q^{\beta n} \ e^{h(\tilde{\kappa} d(x,y))} \right). \tag{43}$$

Take the supremum of the right hand side over all $x, y : \ d(x,y) \leq t$ for a fixed $t : \ 0 < t \leq 2r_0$, divide both sides by t^α, and then take an upper bound on the leftover factor $t^{\beta-\alpha} e^{h(\tilde{\kappa} t)}$ from the facts that $t \leq 2r_0$, $\beta - \alpha \geq 0$, and $h(t)$ is nondecreasing. This yields

$$t^{-\alpha} \mathfrak{m}\left(\mathfrak{L}_0^n f, t\right) \leq Pe^{h(\operatorname{diam}E)} \left(\mathscr{V}_\alpha(f) \tilde{\kappa}^\alpha q^{\alpha n} + \|f\|_\infty \mathscr{V}_\beta(\phi) \tilde{\kappa}^\beta (2r_0)^{\beta - \alpha} \sum_{n=1}^\infty q^{\beta n} \ e^{h(2r_0 \tilde{\kappa})} \right). \tag{44}$$

By definition the right hand side is an upper bound on $\mathscr{V}_\alpha\left(\mathfrak{L}_0^n f\right)$; this shows that $\mathfrak{L}_0$ does indeed preserve the space H_α. Now we glean the Romanian inequality by associating appropriate terms:

$$\begin{aligned} &\|\mathfrak{L}_0^n f\|_\alpha = \|\mathfrak{L}_0^n f\|_\infty + \mathscr{V}_\alpha(\mathfrak{L}_0^n f) \leq \\ &\leq Pe^{h(\operatorname{diam}E)} \left(\|f\|_\infty + \mathscr{V}_\alpha(f) \tilde{\kappa}^\alpha q^{\alpha n} + \|f\|_\infty \mathscr{V}_\beta(\phi) \tilde{\kappa}^\beta (2r_0)^{\beta-\alpha} \sum_{n=1}^\infty q^{\beta n} \ e^{h(2r_0 \tilde{\kappa})} \right) \\ &= Pe^{h(\operatorname{diam}E)} \left(\tilde{\kappa}^\alpha q^{\alpha n} \|f\|_\alpha + \left(1 - \tilde{\kappa}^\alpha q^{\alpha n} + \mathscr{V}_\beta(\phi) \tilde{\kappa}^\beta (2r_0)^{\beta-\alpha} \sum_{n=1}^\infty \ q^{\beta n} e^{h(2r_0 \tilde{\kappa})} \right) \|f\|_\infty \right). \end{aligned} \tag{45}$$

The key to the Romanian inequality is that $r < 1$; this is accomplished by taking n so large that the coefficient of $\|f\|_\alpha$ in this last upper bound is < 1. □

6.2. **The fully normalized transfer operator.** There is an "even more normalized" transfer operator than $\mathfrak{L}_0 = e^{-p(\phi)} \mathfrak{L}$. Let

$$\mathfrak{L}_1 f = \frac{1}{\rho} \mathfrak{L}_0(\rho f); \tag{46}$$

notice that, because of the fact that ρ is bounded away from 0 and ∞, $\mathfrak{L}_1$ is conjugate to $\mathfrak{L}_0$ by a bounded invertible linear operator with a bounded inverse. It retains the characteristic interaction with iterates of S that is shared by $\mathfrak{L}$ and $\mathfrak{L}_0$, namely that for all $n \geq 1$ and all continuous functions f, g

$$\mathfrak{L}_1^n(f\, g \circ S^n) = g\, \mathfrak{L}_1^n(f).$$

Notice too that

$$\mathfrak{L}_1(\mathbb{1}) = \frac{\mathfrak{L}_0(\rho)}{\rho} = \frac{\rho}{\rho} = \mathbb{1},$$

and

$$\mathfrak{L}_1^*\eta = \eta, \text{ i.e. } \eta(\mathfrak{L}_1 f) = \gamma(\mathfrak{L}_0(\rho f)) = \eta(f)\ \forall f \in C(X).$$

The following observation facilitates checking Gordin's condition for the central limit theorem, see Theorem 6.8.

Observation 6.2. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition and $\eta = \eta_\phi$ is the S-invariant Gibbs state introduced in Fact 4.5 then for every $f \in C(X)$ and every $n \geq 1$ a version of $E_\eta[f|S^{-n}(\mathscr{B})]$ is given by $\mathfrak{L}_1^n f \circ S^n$.*

Proof. As a general rule, a the composition of a $\mathscr{B}$-measurable function on the left with S^n on the right is $S^{-n}(\mathscr{B})$ measurable. A typical event (set) in $S^{-n}(\mathscr{B})$ can be represented without loss of generality as $S^{-n}(B)$ for some $B \in \mathscr{B}$. Using only general principles of integration and the established properties of the objects in this paper,

$$\begin{aligned}\eta\left(\mathbb{1}_{S^{-n}(B)}\,(\mathfrak{L}_1^n f)\circ S^n\right) &= \eta\left((\mathbb{1}_B\,\mathfrak{L}_1^n f)\circ S^n\right) = \eta\left(\mathbb{1}_B\,\mathfrak{L}_1^n f\right)\\ &= \eta\left(\mathfrak{L}_1^n(\mathbb{1}_B\circ S^n\, f)\right) = \eta\left(\mathbb{1}_B\circ S^n\, f\right) = \eta\left(\mathbb{1}_{S^{-n}(B)}\, f\right).\end{aligned}$$

Thus the conditional expectation is achieved. □

The convergence introduced in the following theorem is strengthened to exponentially fast in Proposition 6.6. The proof is standard, as in [8], for instance.

Lemma 6.3. *If $\phi : X \to \mathbb{R}$ satisfies Dini's condition (8) and $\eta = \rho\gamma$ is the S invariant Gibbs state introduced in Fact 4.5 then for every $f \in C(X)$ the sequence $\mathfrak{L}_1^n f$, $n \geq 1$ converges uniformly to the constant function $\eta(f)\mathbb{1}$.*

Proof. First of all check that $\mathfrak{L}_1$ is a monotone operator on $C(X)$ with respect to its natural pointwise ordering. It follows that, pointwise,

$$\mathfrak{L}_1 f \leq \mathfrak{L}_1((\sup f)\mathbb{1}) = \sup f\, \mathbb{1}$$

and therefore

$$\sup \mathfrak{L}_1 f \leq \sup f.$$

For any $f \in C(X)$ the sequence $\{\mathfrak{L}_0^n f\}_{n\geq 1}$ is bounded by a slight extension of Estimate (13) and $\|\ \|_\infty$ equicontinuous by Estimate (42). By the Arzela-Ascoli theorem this shows that the operator $\mathfrak{L}_0$ is almost periodic. $\mathfrak{L}_1$ is almost periodic by its conjugacy to $\mathfrak{L}_0$. By compactness of $C(X)$ there is a uniform limit

$$f^* = \lim_{j\to\infty} \mathfrak{L}_1^{n_j} f \tag{47}$$

of some subsequence $\{n_j\}_{j\geq 1}$ of iterates for every continuous observable function f. The above monotonicity argument says

$$\sup f^* \geq \sup \mathfrak{L}_1 f^* \geq \sup \mathfrak{L}_1^2 f^* \geq \dots .$$

We claim that in fact all the suprema are equal. To see it, let $\epsilon > 0$. Taking only one half of the compound inequality (which holds for large enough j)

$$\|f^* - \mathfrak{L}_1^{n_j}\|_\infty < \epsilon$$

leaves the pointwise inequality

$$f^* \geq \mathfrak{L}_1^{n_j} f - \epsilon 1\!\!1.$$

Monotonicity of $\mathfrak{L}_1$ implies that for every $q \geq 1$

$$\mathfrak{L}_1^q f^* \geq \mathfrak{L}_1^{n_j+q} f - \epsilon 1\!\!1 \;\text{ and hence }\; \sup \mathfrak{L}_1^q f^* \geq \sup \mathfrak{L}_1^{n_j+q} f - \epsilon.$$

Choose $k_j \geq 1$ large enough that $n_{j+k_j} \geq n_j + q$ for each $j \geq 1$, which implies

$$\sup \mathfrak{L}_1^{n_j+q} f \geq \sup \mathfrak{L}_1^{n_{j+k_j}} f.$$

So we have

$$\begin{aligned} \sup f^* &\geq \sup \mathfrak{L}_1^q f^* \\ &\geq \overline{\lim}_{j\to\infty} \sup(\mathfrak{L}_1^{n_j+q} f) - \epsilon \\ &\geq \overline{\lim}_{j\to\infty} \sup(\mathfrak{L}_1^{n_{j+k_j}} f) - \epsilon \\ &= \sup f^* - \epsilon. \end{aligned}$$

The last equality follows from the defining equation (47) for f^*. This proves the claim that $\sup f^* = \sup \mathfrak{L}_1^q f^*$ for every $q \geq 1$.

By compactness of X there exists, for each $q \geq 1$, a point $x^q \in X$ for which $\mathfrak{L}_1^q f^*(x^q) = \sup \mathfrak{L}_1^q f^* = \sup f^*$. The positivity of the operator $\mathfrak{L}_1$ and the fact $\mathfrak{L}_1(1\!\!1) = 1\!\!1$ show that for any $x \in X$ the expression

$$\mathfrak{L}_1^q f^*(x)$$

is an integral against a probability measure of a convex combination of the values

$$f^*(a_1 b_1 \dots b_q x_2 \dots),\; (a_1, \dots, a_q) \in E^q,\; b_1 \in T^{-1}a_2,\; \dots\; , b_q \in T^{-1}x_1,$$

which are all bounded above by $\sup f^*$. Taking $x = x^q$ we have an integral of a convex combination of the values $f^*(a_1 b_1 \dots b_q x_2^q \dots)$ which is equal to the upper bound $\sup f^*$ on the values. Thus $f^*(a_1 b_1 \dots b_q x_2^q \dots) = \sup f^*$ for ν^q-a.e. $a \in E^q$ and every $b_1 \in T^{-1}a_2, \; \dots \; , b_q \in T^{-1}x_1^q$. We have proved this for every $q \geq 1$. Assuming that the topological support *supp*(ν) is not less than the full alphabet E, continuity of f^* leads to the conclusion that f^* is constant. In fact, f^* is the constant function $\eta(f)1\!\!1$, because $\mathfrak{L}_1^*$ invariance of η implies $\eta(f) = \lim_{j\to\infty} \eta(\mathfrak{L}_1^{n_j} f) = \eta(f^*)$.

Since $n_j,\; j \geq 1$, was an arbitrary subsequence along which the iterates of f under $\mathfrak{L}_1$ converged and we have shown that relative compactness of the full sequence $\mathfrak{L}_1^n f,\; n \geq 1$, follows from the Arzela-Ascoli theorem, we can conclude that the full sequence converges uniformly to the constant function $\eta(f)1\!\!1$. □

Lemma 6.4. *The number* 1 *is the only unitary eigenvalue of* $\mathfrak{L}_1$ *and* $ker(\mathfrak{L}_1 - I) = \mathbb{C}1\!\!1$.

Proof. It has already been shown that $\mathfrak{L}_1(\mathbb{1}) = \mathbb{1}$, thus $\mathbb{C}\mathbb{1} \subseteq ker(\mathfrak{L}_1 - I)$. For the rest that has been claimed suppose $z \in \mathbb{C}$, $|z| = 1$, and $\mathfrak{L}_1 f = zf$ for some nonzero $f \in C(X)$. Then by lemma 6.3 $\mathfrak{L}_1^n f = z^n f \to \eta(f)\mathbb{1}$ uniformly as $n \to \infty$, and even pointwise convergence here implies $z = 1$ and therefore f is constant. □

Lemma 6.5. *The number* 1 *is the only unitary eigenvalue of* $\mathfrak{L}_0$ *and* $ker(\mathfrak{L}_0 - 1) = \mathbb{C}\rho$.

Proof. By definition $\rho \in ker(\mathfrak{L}_0 - 1)$. For the rest, let $z \in \mathbb{C}$, $|z| = 1$, and $\mathfrak{L}_0 f = zf$ for some nonzero $f \in C(X)$. Then $\mathfrak{L}_1\left(\frac{f}{\rho}\right) = \frac{zf}{\rho}$ and hence (as in Lemma 6.4) $z = 1$ and $\frac{f}{\rho}$ is constant. □

These corollaries now permit the final bit of analysis to be done before being able to conclude the probability laws. Specifically we apply the theorem of Ionescu Tulcea and Marinescu to provide conditions on the potential and observable functions ϕ, f that guarantee the convergence in Proposition 6.3 to be exponentially fast.

Proposition 6.6. *If* ϕ *and* f *are Hölder continuous then*

$$\|\mathfrak{L}_1^n f - \eta(f)\mathbb{1}\|_\infty \to 0$$

exponentially fast as $n \to \infty$.

Proof. If ϕ and f are Hölder with different exponents then use the larger class (the one with the smaller exponent) as the subspace of $C(X)$ in Ionescu Tulcea and Marinescu's theorem. By virtue of Section 6.1, the theorem applies to $\mathfrak{L}_1$ and states that $\mathfrak{L}_1$ is of the form $\mathfrak{Q} + \sum_{i=1}^m z_i \mathfrak{P}_i$, where $\{z_i\}_{i=1}^m$ is contained in the unit circle of $\mathbb{C}$ and comprises all the eigenvalues of $\mathfrak{L}_0$, the operators $\mathfrak{P}_i$ are projections orthogonal to each other and to $\mathfrak{Q}$, and the spectrum of $\mathfrak{Q}$, which is the essential spectrum of $\mathfrak{L}_1$, is contained in some closed disk about the origin of radius $r_Q < 1$. By Lemma 6.5 we know in fact that $\{z_i\}_{i=1}^m = \{1\}$, and by the spectral radius formula we can conclude that for every constant $r : r_Q < r < 1$ there exists a constant $c_Q > 0$ such that for every $n \geq 1$

$$\|\mathfrak{Q}^n\| \leq c_Q r^n. \tag{48}$$

So

$$\mathfrak{L}_1 = \mathfrak{Q} + \mathfrak{P},$$

and for every $n \geq 1$

$$\mathfrak{L}_1^n f = \mathfrak{Q}^n f + \mathfrak{P} f.$$

Since

$$\mathfrak{L}_1^n \to \eta(f)\mathbb{1}$$

and

$$\mathfrak{Q}^n f \to 0\mathbb{1}$$

uniformly as $n \to \infty$ it must be that

$$\eta(f)\mathbb{1} = \mathfrak{P} f$$

and hence

$$\|\mathfrak{L}_1^n f - \eta(f)\mathbb{1}\|_\infty = \|\mathfrak{Q}^n f\|_\infty \leq c_Q r^n \|f\|_\infty.$$

This proves the proposition. □

6.3. **Statement of the laws of probability for the dynamical system.**

Theorem 6.7. *If the potential ϕ and the observable f are Hölder continuous then there is an* exponential decay of correlations *in the stochastic process $\{f \circ S^n\}$ on the probability space $(X, \mathscr{B}(X), \eta)$.*

Proof. The theorem claims that $corr_\eta(f\circ S^m, f\circ S^n) \to 0$ exponentially fast as $|n-m| \to \infty$. This provides a memorable statement for the theorem, but we are really in a position to prove a stronger result. Let $1 \leq m \leq n$ and $g \in \mathscr{L}_1(\eta)$ and expand

$$corr_\eta(f \circ S^m, g \circ S^n) = \frac{\eta\left((f \circ S^m - \eta(f \circ S^m))(g \circ S^n - \eta(g \circ S^n))\right)}{(\eta((f \circ S^m - \eta(f \circ S^m))^2)\eta((g \circ S^n - \eta(g \circ S^n))^2))^{\frac{1}{2}}}.$$

By S-invariance of η the denominator reduces to

$$\left(\left(\eta(f^2) - (\eta(f))^2\right)\left(\eta(g^2) - (\eta(g))^2\right)\right)^{\frac{1}{2}},$$

so exponential decay of the correlation will follow from exponential decay of the numerator as $n - m \to \infty$.

To streamline the formulas let $f_0 = f - \eta(f)\mathbb{1}$ and $g_0 = g - \eta(g)\mathbb{1}$. Then the numerator of the correlation (the covariance) becomes

$$\begin{aligned}
\eta(f_0 \circ S^m \, g_0 \circ S^n) &= \eta(f_0 \, g_0 \circ S^{n-m}) \\
&= \eta\left(\mathfrak{L}_1^{n-m}(f_0 \, g_0 \circ S^{n-m})\right) \\
&= \eta\left(g_0 \, \mathfrak{L}_1^{n-m} f_0\right) \\
&\leq \|g_0\|_1 \|\mathfrak{L}_1^{n-m} f_0\|_\infty.
\end{aligned}$$

Because $\eta(f_0) = 0$, Proposition 6.6 shows that this last norm decays exponentially fast as $n - m \to \infty$. □

Theorem 6.8. *If the potential ϕ and the observable f are Hölder continuous with $\eta(f) = 0$, then the central limit theorem holds for the stochastic process $\{f \circ S^k\}_{k\geq 1}$ on the probability space $(X, \mathscr{B}(X), \eta)$, i.e.*

$$\sigma = \lim_{k\to\infty} \left(\frac{\eta((s_k f)^2)}{k}\right)^{\frac{1}{2}}$$

exists and if $\sigma > 0$ then for every $z \in \mathbb{R}$

$$\eta\left(\left\{\frac{s_k f}{\sqrt{k}} < z\right\}\right) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2\sigma^2}} \, du \text{ as } k \to \infty.$$

Proof. According to a famous criterion published by Gordin ([4], Theorem 2), it suffices to show that

$$\sum_{k=0}^{\infty} \eta\left((E_\eta[f|S^{-k}(\mathscr{B})])^2\right) < \infty.$$

By Observation 6.2 the kth term of this series is

$$0 < \eta\left((E_\eta[f|S^{-k}(\mathscr{B})])^2\right) = \eta(f \, E_\eta[f|S^{-k}(\mathscr{B})])$$
$$\eta\left(f\,(\mathfrak{L}_1^k f) \circ S^k\right) \leq \|f\|_\infty \eta\left(|\mathfrak{L}_1^k f| \circ S^k\right)$$

$$= \|f\|_\infty \eta\left(|\mathfrak{L}_1^k f|\right) \le \|f\|_\infty \|\mathfrak{L}_1^k f\|_\infty.$$

Because $\eta(f) = 0$, Proposition 6.6 finishes the proof. □

It is no real restriction in Theorem 6.8 to require the observable function f to have expected value $\eta(f) = 0$ because for any Hölder function f we can apply the preceding theorem to $f_0 = f - \eta(f)\mathbb{1}$. Then

$$\eta\left(\left\{\frac{s_k f_0}{\sqrt{k}} < z\right\}\right) = \eta\left(\left\{\frac{s_k f - k\eta(f)\mathbb{1}}{\sqrt{k}} < z\right\}\right),$$

converges to the normal probability $N_\sigma((-\infty, z))$ and we have recovered a statement of the central limit theorem for uncentered (i.e having nonzero expectation) random variables.

Theorem 6.9. *If the local map $T : E \to E$ is a homeomorphism and the potential ϕ and observable f are Hölder continuous then the law of the iterated logarithm holds for the stochastic process $\{f \circ S^k\}_{k\ge 1}$ on the probability space $(X, \mathscr{B}(X), \eta)$, i.e.*

$$\sigma = \lim_{k\to\infty} \left(\frac{\eta((s_k f)^2)}{k}\right)^{\frac{1}{2}}$$

exists and

$$\sqrt{2}\sigma = \overline{\lim}_{k\to\infty} \frac{s_k f - k\eta(f)}{\sqrt{k \log(\log(k))}}$$

holds η-a.e.

Proof. The extra requirement on T ensures that $T^{-1}(\mathscr{B}(E)) = \mathscr{B}(E)$ and not a smaller σ-algebra. Thus we have

$$\vee_{j=1}^n S^{-1}(\mathscr{B}_1) = \mathscr{B}(E) \otimes T^{-1}(\mathscr{B}(E)) \otimes \cdots \otimes T^{-1}(\mathscr{B}(E)) \otimes \{\emptyset, E^{\mathbb{N}}\} = \mathscr{B}_n.$$

We rely now on the result of Phillips and Stout ([1], c.f. [6]) that the law of the iterated logarithm applies to this process if the following two conditions are met:

(1) the dynamical system is φ-mixing and
(2) $\|f - \eta(f|\mathscr{B}_n)\|_2 \to 0$ on the order of $\frac{1}{n^2}$ or faster.

To discuss the φ-mixing let us refine our notation somewhat and define

$$\mathscr{B}_{m\ldots n} = \{E^{m-1} \times B \times E_n^{\mathbb{N}} \ : \ B \in \mathscr{B}(E^{n-m})\},$$

a sub σ-algebra of $\mathscr{B}(X)$. Let $1 \le m \le n-2$, $A \in \mathscr{B}_{1\ldots m}$ and $B \in \mathscr{B}_{n\ldots\infty}$. We can represent these sets as $A = A' \times E^{\mathbb{N}_m}$ and $B = E^n \times B'$ for some sets $A' \in \mathscr{B}(E^m)$ and $B' \in \mathscr{B}(E^{\mathbb{N}_n})$. We'll soon use the fact that

$$\mathbb{1}_B = \mathbb{1}_{E\times B'} \circ S^{n-1}.$$

For φ-mixing we are supposed to analyze

$$\begin{aligned}
\eta(A \cap B) - \eta(A)\eta(B) &= \eta\left((\mathbb{1}_A - \eta(A)\mathbb{1})(\mathbb{1}_B - \eta(B)\mathbb{1})\right)\\
&= \eta\left((\mathbb{1}_A - \eta(A)\mathbb{1})(\mathbb{1}_{E\times B'} - \eta(B)\mathbb{1}) \circ S^{n-1}\right)\\
&= \eta\left(\mathfrak{L}_1^{n-1}\left((\mathbb{1}_A - \eta(A)\mathbb{1})(\mathbb{1}_{E\times B'} - \eta(B)\mathbb{1}) \circ S^{n-1}\right)\right)\\
&= \eta\left((\mathbb{1}_{E\times B'} - \eta(B)\mathbb{1})\mathfrak{L}_1^{n-1}(\mathbb{1}_A - \eta(A)\mathbb{1})\right)
\end{aligned}$$

$$= \eta(A)\eta\left((\mathbb{1}_{E\times B'} - \eta(B)\mathbb{1})\mathfrak{L}_1^{n-1}\left(\frac{\mathbb{1}_A}{\eta(A)} - \mathbb{1}\right)\right).$$

Because

$$\|\mathbb{1}_{E\times B'} - \eta(B)\mathbb{1}\|_\infty \leq 1,$$

we conclude

$$|\eta(A\cap B) - \eta(A)\eta(B)| \leq \eta(A)\left|\eta\left(\mathfrak{L}_1^{n-1}\left(\frac{\mathbb{1}_A}{\eta(A)} - \mathbb{1}\right)\right)\right|$$

$$= \eta(A)|\eta\left(\mathfrak{L}_1^{n-1-m}\left(\frac{\mathfrak{L}_1^m(\mathbb{1}_A)}{\eta(A)} - \mathbb{1}\right)\right)$$

$$\leq \eta(A)\|\mathfrak{L}_1^{n-1-m}\|\left\|\frac{\mathfrak{L}_1^m(\mathbb{1}_A)}{\eta(A)} - \mathbb{1}\right\|_\infty$$

$$\leq \eta(A)c_Q r^{n-1-m}c_1,$$

wherein $c_Q > 0$ and $0 < r < 1$ come from the decomposition of the operator $\mathfrak{L}_1$ in corollary 6.6 and $c_1 > 1$ is the constant guaranteed by the definition of η as a Gibbs measure to satisfy

$$\frac{1}{c_1} \leq \frac{\eta([A'])}{\mathfrak{L}_1^m \mathbb{1}_{[A']}(x)} \leq c_1$$

for all $m \geq 1$, $A' \in \mathscr{B}(E^m)$, and $x \in X$.

We have proven that $(X, \mathscr{B}, \eta, S)$ is a φ-mixing system with

$$\varphi(n-m) = c_Q c_1 r^{n-1-m}.$$

To arrive at Condition (2) we fix an $n \geq 1$ and observe that for any measurable rectangle $A_1 \times \cdots \times A_n \subset E^n$ and any $x \in [A_1 \times \cdots \times A_n]$

$$\frac{\int_{[A_1\times\cdots\times A_n]} \eta(f|\mathscr{B}_{1\ldots n})\,d\eta}{\eta([A_1\times\cdots\times A_n])} = \frac{\int_{[A_1\times\cdots\times A_n]} f\,d\eta}{\eta([A_1\times\cdots\times A_n])} \in$$

$$\in \left(f(x) - \mathscr{V}_\alpha(f)\mathrm{diam}\left([A_1\times\cdots\times A_n]\right)^\alpha, f(x) + \mathscr{V}_\alpha(f)\mathrm{diam}\left([A_1\times\cdots\times A_n]\right)^\alpha\right).$$

It thus follows by the martingale convergence theorem that

$$E_\eta(f|\mathscr{B}_{1\ldots n})(ax) \in \left(f(at) - \mathscr{V}_\alpha(f)\frac{q^{\alpha n}}{(1-q)^\alpha}, f(at) + \mathscr{V}_\alpha(f)\frac{q^{\alpha n}}{(1-q)^\alpha}\right).$$

Thus

$$|f - \eta(f|\mathscr{B}_{1\ldots n})| \leq \mathscr{V}_\alpha(f)\frac{q^{\alpha n}}{(1-q)^\alpha}$$

holds η-a.e., which clearly gives exponential decay of the associated 2-norms, too. This is even faster than the required power-law decay. □

7. Weak Continuous Dependence of the Equilibrium State on the apriori measure

We prove that for a fixed potential function $\phi : X \to \mathbb{R}$ satisfying Dini's condition the pressure and equilibrium state depend continuously on the apriori measure η, where apriori measures are treated with the weak topology and functionals with the weak* topology. In this section dependence on ϕ is suppressed in all notation and dependence on ν, which has heretofore been suppressed, is notated with subscripts.

Lemma 7.1. *If ν_n is a sequence of Borel probability measures on E converging weakly to ν and $\gamma_n \subset C^*$ is a sequence of functionals converging weakly* to γ then $\mathfrak{L}^*_{\nu_n}\gamma_n$ converges weakly* to $\mathfrak{L}^*_\nu\gamma$.*

Proof. Let $f \in C$. Applying the convergence $\gamma_n \to \gamma$ to the estimate

$$|\gamma_n(\mathfrak{L}_{\nu_n} f) - \gamma(\mathfrak{L}_\nu f)| \le |\gamma_n(\mathfrak{L}_\nu f) - \gamma(\mathfrak{L}_\nu f)| + |\gamma_n(\mathfrak{L}_{\nu_n} f) - \gamma_n(\mathfrak{L}_\nu f)|$$

yields

$$\overline{\lim}_{n\to\infty}|\gamma_n(\mathfrak{L}_{\nu_n} f) - \gamma(\mathfrak{L}_\nu f)| \le \overline{\lim}_{n\to\infty}|\gamma_n(\mathfrak{L}_{\nu_n} f - \mathfrak{L}_\nu f)|. \tag{49}$$

Now let $\epsilon > 0$ and $0 < \delta_1 < \mathfrak{m}(fe^\phi, \frac{\epsilon}{3P})$. For any a priori measure ν' and any two points $x, y \in X$ with $d(x,y) < \delta_2 \equiv \min\left(r_0, \frac{\delta_1}{q\tilde{\kappa}}\right)$ we have

$$|\mathfrak{L}_{\nu'} f(x) - \mathfrak{L}_{\nu'} f(y)| = \left|\nu'\left(\sum_{i=1}^{|T^{-1}(x_1)|} fe^\phi(\centerdot\, T_i^{-1}(x_1)x_2\dots) - fe^\phi(\centerdot\, T_i^{-1}(y_1)y_2\dots)\right)\right|, \tag{50}$$

wherein Lemma 1.2 has shown that, because $d_0(x_1,y_1) < r_0$ they have common inverse branches with paired points satisfying

$$d_0(T_i^{-1}x_1, T_i^{-1}y_1) \le \frac{1}{\kappa} d_0(x_1,y_1)$$

and therefore

$$d(\centerdot\, T_i^{-1}(x_1)x_2\dots, \centerdot\, T_i^{-1}(y_1)y_2\dots) \le q\tilde{\kappa} d(x,y) \le \delta_1.$$

Now by definition of δ_1 and the triangle inequality estimate 50 yields

$$|\mathfrak{L}_{\nu'} f(x) - \mathfrak{L}_{\nu'} f(y)| \le P\left(\frac{\epsilon}{3P}\right) = \frac{\epsilon}{3}. \tag{51}$$

Let X_ϵ be a finite δ_2 net in X endowed with an arbitrary order. Define a map $X \ni x \mapsto \overline{x} \in X_\epsilon$ by the rule that $\overline{x}$ is the first element of X_ϵ within δ_2 of x. By Estimate (51) it follows that for any apriori measure ν'

$$\|\mathfrak{L}_{\nu'} f - \mathfrak{L}_{\nu'} f \circ \overline{}\|_\infty \le \frac{\epsilon}{3}.$$

By the weak convergence $\nu_n \to \nu$ it follows that for all $\overline{x} \in X_\epsilon$ there is an index $n_{\overline{x},\epsilon}$ beyond which

$$|\mathfrak{L}_{\nu_n} f(\overline{x}) - \mathfrak{L}_\nu f(\overline{x})| < \frac{\epsilon}{3}.$$

Let

$$n_\epsilon = \max_{\overline{x}\in X_\epsilon} n_{\overline{x},\epsilon}$$

and observe that for $n \geq n_\epsilon$ we have, in the right hand side of Estimate (49),

$$|\gamma_n(\mathfrak{L}_{\nu_n} f - \mathfrak{L}_\nu f)| \leq$$
$$\|\mathfrak{L}_{\nu_n} f - \mathfrak{L}_{\nu_n} f \circ \text{-}\|_\infty + \|\mathfrak{L}_{\nu_n} f \circ \text{-} - \mathfrak{L}_\nu f \circ \text{-}\|_\infty + \|\mathfrak{L}_\nu f \circ \text{-} - \mathfrak{L}_\nu f\|_\infty$$
$$< \epsilon.$$

This finishes the lemma. □

Theorem 7.2. *Treat the space of Borel probability measures ν on E with the weak topology. Then for any $\phi : X \to \mathbb{R}$ satisfying Dini's condition the maps*

$$\nu \mapsto p_\nu(\phi) \in \mathbb{R},$$
$$\nu \mapsto \gamma_{\phi,\nu} \in C^*_{+,1},$$

and

$$\nu \mapsto \eta_{\phi,\nu} \in C^*_{+,1,S}$$

are continuous.

Proof. Let $\nu_j \to \nu$, and for each $j \geq 1$ find a functional $\gamma_j \in C^*_{+,1}$ for which

$$\mathfrak{L}^*_{\nu_j} \gamma_j = e^{p(\nu_j)} \gamma_j.$$

Because the sequence of pressures is bounded, every subsequence j_k, $k \geq 1$, has a convergent subsequence among the pressures

$$p(\nu_{j_{k_l}}) \to p_1 \in \mathbb{R} \text{ as } l \to \infty.$$

By Banach-Alaoglu's compactness theorem there is a further weakly* convergent subsequence

$$\gamma_{j_{k_{l_m}}} \to \gamma_1 \in C^*_{+,1} \text{ as } m \to \infty.$$

By Lemma 7.1 we have the weak* convergence

$$\mathfrak{L}^*_{\nu_{j_{k_{l_m}}}} \gamma_{j_{k_{l_m}}} \to \mathfrak{L}^*_\nu \gamma_1 \text{ as } m \to \infty.$$

Then again it is easy to see that

$$e^{p(\nu_{j_{k_{l_m}}})} \gamma_{j_{k_{l_m}}} \to e^{p_1} \gamma_1 \text{ as } m \to \infty$$

also holds weakly*, and therefore

$$\mathfrak{L}^*_\nu \gamma_1 = e^{p_1} \gamma_1.$$

By Observation 3.1 p_1 must be equal to $p_\nu = p_\nu(\phi)$ and by Fact 4.5 γ_1 is a Gibbs state for ϕ with a priori measure ν. Thus by Corollary 5.3 $\gamma_1 = \gamma_{\phi,\nu}$. Repeat the argument first choosing a convergent subsequence of γs and then refining it to a convergent subsubsequence of pressures. Together we've proven that $p_\nu(\phi)$ and $\gamma_\nu = \gamma_{\phi,\nu}$ are the only possible limits of the sequences $\{p_{\nu_j}\}_{j\geq 1}$ and $\{\gamma_{\nu_j}\}_{j\geq 1}$, respectively, as $j \to \infty$. With compactness this proves both sequences truly converge to their respective limits.

Now finally suppose there is a convergent subsequence of the equilibrium measures:

$$\eta_{\nu_{j_k}} \to \eta_1 \text{ as } k \to \infty.$$

We've just proven that $\gamma_{\nu_{j_k}} \to \gamma_\nu$. According to Estimate (16) in Lemma 3.2

$$\frac{1}{Pe^{h(\mathrm{diam}E)}} \le \rho_{\phi,\nu_{j_k}} \le Pe^{h(\mathrm{diam}E)}$$

holds independently of k. It follows that

$$\left(Pe^{h(\mathrm{diam}E)}\right)^{-2} \le \frac{\rho_{\nu_{j_k}}}{\gamma_{\nu_{j_k}}(\rho_{\nu_{j_k}})} \le \left(Pe^{h(\mathrm{diam}E)}\right)^{2}$$

also holds independently of k and hence, by definition of η, that for every $k \ge 1$ and every $B \in \mathscr{B}(X)$ with $\gamma_{\nu_{j_k}}(B) > 0$,

$$\left(Pe^{h(\mathrm{diam}E)}\right)^{-2} \le \frac{\eta_{\nu_{j_k}}(B)}{\gamma_{\nu_{j_k}}(B)} \le \left(Pe^{h(\mathrm{diam}E)}\right)^{2}.$$

Then for all sets $B \in \mathscr{B}(X)$ with $\gamma_\nu(B) > 0$ and $\gamma_\nu(\partial B) = 0$ it follows from the weak convergence that (and for all continuous functions $B : X \to \mathbb{R}$)

$$\left(Pe^{h(\mathrm{diam}E)}\right)^{-2} \le \frac{\eta_1(B)}{\gamma_\nu(B)} \le \left(Pe^{h(\mathrm{diam}E)}\right)^{2}.$$

Thus η_1 is a Gibbs measure for ν, and moreover it is invariant because it is the weak* limit of invariant measures. By the uniqueness theorem (Theorem 5.2) this shows $\eta_1 = \eta_{\phi,\nu}$. $\square$

References

[1] Phillips, Stout: Almost sure invariance principles for partial sums of weakly dependent random variables. Mem. AMS 161 (1975)
[2] Billingsley, Patrick.: Probability and Measure. Wiley Interscience, (1995)
[3] Georgii, H-O.: Gibbs Measures and Phase Transitions. DeGruyter, (1981)
[4] Gordin, M.I.: The central limit theorem for stationary processes. Soviet Math Dokl. Vol. 10 (1969) No.5 p.1174-1176
[5] Ionescu Tulcea, C.T., Marinescu, G.: Théorie ergodique pour des classes d'opérations non complètement continues. Ann. of Math. (2) 52, 140-147 (1950)
[6] Przytycki, F., Urbański, M. Conformal Fractals - Ergodic Theory Methods. Cambridge University Press, (2009)
[7] Simon, B.: The Statistical Mechanics of Lattice Gases, Vol. I. Princeton University Press, Princeton (1993).
[8] Urbański, M., Zdunik, A.: Ergodic Theory for Holomorphic Endomorphisms of Complex Projective Spaces. Preprint (2009).

Stephen Muir
Department of Mathematics
University of California, Riverside
900 Big Springs Rd., Surge Bldg. 239
Riverside, CA 92521-0135

Mariusz Urbański
University of North Texas
Department of Mathematics
1155 Union Circle #311430
Denton, TX 76203-5017

E-mail address: muirst@ucr.edu and urbanski@unt.edu