Equilibrium measures of the natural extension of β-shifts

C.-E. PFISTER† and W. G. SULLIVAN‡

† *Section of Mathematics, Faculty of Basic Sciences, EPFL,*
CH-1015 Lausanne, Switzerland
(e-mail: charles.pfister@epfl.ch)
‡ *School of Mathematics and Statistics, UCD, Belfield, Dublin 4, Ireland*
(e-mail: Wayne.Sullivan@ucd.ie)

Abstract. We give a necessary and sufficient condition on β of the natural extension of a
β-shift, so that any equilibrium measure for a function of bounded total oscillations is a
weak Gibbs measure.

1. *Introduction*

We study equilibrium measures of the natural extension of β-shifts. This is an interesting
class of dynamical systems which have been studied in ergodic and number theory since
the fundamental papers [**Re**, **Pa**]. We want to determinate whether an equilibrium measure
for a continuous function φ is a weak Gibbs measure. In [**PS3**] we developed a method
based on a decoupling property [**PS3**, Definition 2.3], which is a slightly weaker condition
than condition (D) in [**Ru**, §4.1]. The results of [**PS3**] are valid only for β such that the
β-shift has the specification property and φ is of bounded total oscillations. This set of β
is the set C_3 in [**Sc**]. Schmeling proved that C_3 has Lebesgue measure 0, but Hausdorff
dimension 1. For the more restricted class of functions satisfying the Bowen condition
[**Bo**], one has a stronger result. Under expansiveness and specification Haydn and Ruelle
[**HR**] proved the equivalence of equilibrium measures and Gibbs measures (in the sense
of Bowen [**Bo**] and Capocaccia [**Ca**]).

However, using basic ideas of [**PS3**, **CTY**], it is possible to obtain a necessary and
sufficient condition on β such that for any function of bounded total oscillations all
equilibrium measures are weak Gibbs measures.

To formulate our main result, Theorem 2.12, precisely we need to recall first some basic
properties of β-shifts. This is done in §2.1. In §2.2 we consider the class of functions

of bounded total oscillations following [**PS3**, §3], and in §2.3 we consider the pressure, establishing two basic estimates for the proof of Theorem 2.12. Our main result is stated in §2.4 and proved in §3. We discuss briefly large deviations for empirical measures in §2.4.

2. *Setting and main result*

2.1. *Beta-shift.*   Let $\beta > 1$ be fixed. The case $\beta \in \mathbb{N}$ is special and corresponds to the full shift. From now on we assume that $\beta \notin \mathbb{N}$. For $t \in \mathbb{R}$, let $\lceil t \rceil := \min\{i \in \mathbb{Z} : i \geq t\}$. We define $\mathsf{b} := \lceil \beta \rceil$. Consider the β-expansion of 1,

$$1 = \sum_{i=1}^{\infty} c_i \beta^{-i},$$

which is given by the algorithm

$$r_0 := 1, \quad c_{i+1} := \lceil \beta\, r_i \rceil - 1, \quad r_{i+1} := \beta\, r_i - c_{i+1}, \quad i \in \mathbb{Z}_+,$$

which ensures that $r_i > 0$ for all $i \in \mathbb{Z}_+$. It follows that $c_1 = \lceil \beta \rceil - 1 > 0$ and $c^\beta := (c_1, c_2, \ldots)$ cannot end with zeros only. For sequences $(a_1, a_2, \ldots)$ and $(b_1, b_2, \ldots)$ the lexicographical order is defined by $(a_1, a_2, \ldots) \prec (b_1, b_2, \ldots)$ if and only if for the smallest index i with $a_i \neq b_i, a_i < b_i$. Let $\mathsf{A} := \{0, \ldots, \mathsf{b} - 1\}$; the (one-sided) β-shift is

$$X^\beta := \{x = (x_1, x_2, \ldots) : x_i \in \mathsf{A},\ T^k x \preceq c^\beta \text{ for all } k \in \mathbb{Z}_+\},$$

where T is the left shift operator. In particular, $T^k c^\beta \preceq c^\beta$ for all $k \in \mathbb{Z}_+$, so that X^β is a shift-invariant closed subset of $\mathsf{A}^{\mathbb{N}}$ (with product topology). The language of the shift X^β is denoted by $\mathcal{L}^\beta$ and the set of the words of length n by $\mathcal{L}_n^\beta$. In this paper the empty word is always denoted by ϵ, $\mathcal{L}_0^\beta = \{\epsilon\}$, while ε is always a positive real number. A word is written $w_1 \cdots w_n$ or simply w. The length of a word w is written $|w|$.

The shift-space X^β can be described by a labeled graph $\mathcal{G}^\beta = (\mathcal{V}, \mathcal{E}^\beta)$ where $\mathcal{V} := \{q_j : j \in \mathbb{Z}_+\}$. The root of the graph is the vertex q_0. There is an edge $q_0 \to q_0$, labeled by k, for each $k = 0, \ldots, \mathsf{b} - 2$, and there is an edge $q_{j-1} \to q_j$ labeled by c_j for each $j \in \mathbb{N}$. Moreover, if the label c_j of $q_{j-1} \to q_j$ is different from 0, then there are c_j edges $q_{j-1} \to q_0$ labeled by $0, \ldots, c_j - 1$. Each word $w_1 \cdots w_n \in \mathcal{L}^\beta$ can always be presented by a path of length n in $\mathcal{G}^\beta$ starting with vertex q_0. For a word $w \in \mathcal{L}^\beta$ we define $q(w)$ as the end vertex of this path starting at q_0 and presenting w. One can concatenate two words w and w' if and only if there is a path η in $\mathcal{G}^\beta$ presenting w and a path η' in $\mathcal{G}^\beta$ presenting w', so that η ends at vertex q and η' starts at vertex q. In particular, one can concatenate w with any words of $\mathcal{L}^\beta$ if $q(w) = q_0$. There is a unique labeled path presenting the infinite sequence c^β, which is the path $(q_0, q_1, q_2, \ldots)$. Let $\mathcal{P}^\beta$ be the set of the prefixes of the sequence c^β, including the empty word ϵ. Let $c_1 \cdots c_n \in \mathcal{P}^\beta$ and suppose that $c_{n+1} = c_{m+2} = \cdots c_{n+m} = 0, c_{n+m+1} \neq 0$. The word $c_1 \cdots c_n$ is presented by a path starting at q_0 and ending at q_n. Since there is only one outgoing edge from each of the vertices $q_n, \ldots, q_{n+m-1}$, the only words w with prefix $c_1 \cdots c_n$ are the words

$$c_1 \cdots c_n, c_1 \cdots c_n 0, c_1 \cdots c_n 00, \ldots, c_1 \cdots c_n \underbrace{0 \cdots 0}_{m}, c_1 \cdots c_n \underbrace{0 \cdots 0}_{m} w', c_1 \cdots c_k,$$

where w' is any word of $\mathcal{L}^\beta$ with first letter $0 \le w'_1 \le c_{n+m+1} - 1$, and $c_1 \cdots c_k \in \mathcal{P}^\beta$ with $k > n + m$. For $u \in \mathcal{P}^\beta$ we set

$$z^\beta(u) := \begin{cases} 0 & \text{if } u = \epsilon, \\ p & \text{if } u = c_1 \cdots c_\ell, \, c_{\ell+1} = \cdots = c_{\ell+p} = 0 \text{ and } c_{\ell+p+1} > 0. \end{cases} \tag{1}$$

$z^\beta(u)$ is a measure of the obstruction to going from vertex $q(u)$ to vertex q_0. We set

$$\bar{z}^\beta(n) := \max\{z^\beta(u) : u \text{ prefix of } c^\beta, \, |u| \le n\}. \tag{2}$$

For each prefix $u = c_1 \cdots c_n$ of c^β we define a new word $\widehat{u}$ as follows. Let c_j be the last letter in $c_1 \cdots c_n$ which is different from 0. We set

$$\widehat{c}_\ell := \begin{cases} c_\ell & \text{if } \ell \ne j, \\ c_j - 1 & \text{if } \ell = j. \end{cases} \tag{3}$$

The word $\widehat{u} := \widehat{c}_1 \cdots \widehat{c}_n$ differs from $u = c_1 \cdots c_n$ by a single letter and $q(\widehat{u}) = q_0$. For any word $w \in \mathcal{L}^\beta$ there is a unique decomposition of w into

$$w = vs(w) \quad \text{where } s(w) \text{ is the largest suffix of } w \text{ belonging to } \mathcal{P}^\beta. \tag{4}$$

We extend the definition (1) to any word w by setting

$$z^\beta(w) := z^\beta(s(w)),$$

and extend the transformation $u \mapsto \widehat{u}$ to any word by setting

$$\widehat{w} := \begin{cases} w & \text{if } s(w) = \epsilon, \\ v\widehat{u} & \text{if } s(w) = u. \end{cases} \tag{5}$$

By convention we set $\widehat{\epsilon} = \epsilon$. The words $\widehat{w}$ can be freely concatenated since $q(\widehat{w}) = q_0$ (see Lemma 2.2).

LEMMA 2.1. *Let $a = a_1 \cdots a_k$ and $b = b_1 \cdots b_\ell$ be two prefixes of c^β. If $ab \in \mathcal{L}^\beta$, then ab is a prefix of c^β.*

Proof. By hypothesis $a_i = c_i$, $i = 1, \ldots, k$, and $b_j = c_j$, $j = 1, \ldots, \ell$. Let $w = ab \in \mathcal{L}^\beta$. Suppose that w is not a prefix of c^β. Then there exists j, $1 \le j \le \ell$, so that

$$c_m = w_{k+m} = c_{k+m}, \quad \text{for } 1 \le m < j, \quad \text{and} \quad c_j = w_{k+j} < c_{k+j}.$$

Hence, $w_{k+1} \cdots w_{k+\ell} = c_1 \cdots c_\ell \prec c_{k+1} \cdots c_{k+\ell}$, a contradiction with $T^k c^\beta \preceq c^\beta$. $\square$

LEMMA 2.2.
(a) *Let $w = vu$, $s(w) = u$. Then $q(v) = q_0$ and $s(\widehat{w}) = s(v\widehat{u}) = \epsilon$.*
(b) *Let $p_1 := z^\beta(c_1)$, where c_1 the first character of c^β. Then the mapping on $\mathcal{L}^\beta$, $w \mapsto \widehat{w}$, is at most $(p_1 + 2)$-to-one, and $s(\widehat{w}) = \epsilon$.*

Proof. (a) Let $w = vu$, $u = s(w)$ and $\widehat{w} = v\widehat{u}$. If $q(v) \ne q_0$, then $s(v) \ne \epsilon$, so that, by Lemma 2.1, u is not maximal, a contradiction. Since $s(v) = \epsilon$ and $q(\widehat{u}) = q_0$, $s(\widehat{w}) = s(\widehat{u}) = \epsilon$.

(b) Let $\widehat{\mathcal{P}}^\beta := \{\widehat{u} : u \in \mathcal{P}^\beta\}$. The list is ordered according to increasing length. The first $p_1 + 2$ words in the list $\widehat{\mathcal{P}}^\beta$ are

$$\epsilon, \ \widehat{c}_1, \ \widehat{c}_1 0, \ \widehat{c}_1 00, \dots, \widehat{c}_1 \underbrace{0 \cdots 0}_{p_1}. \tag{6}$$

On the other hand, if $|\widehat{u}| > p_1 + 1$, then the first character of $\widehat{u}$ is c_1.

Let $\widehat{w} = w'$. Let $\widehat{s}(w')$ be the largest suffix of w' among the first $|w'| + 1$ elements of the list $\widehat{\mathcal{P}}^\beta$. We write $w' = v'u'$ with $u' = \widehat{s}(w')$. Let $w = vu$, $u = s(w)$, such that $\widehat{w} = w'$. We have $|u| \le |u'|$, otherwise $\widehat{w} = w'$ would imply that $\widehat{s}(w')$ is not maximal. In particular, if $\widehat{s}(w') = \epsilon$, then $s(w) = \epsilon$ and $w' = w$.

Suppose that $\widehat{w} = w'$ and $p_1 + 2 \le |u| < |u'|$. Then the first character of u' is c_1 and also the first character of $\widehat{u}$ is c_1 since $|u| \ge p_1 + 2$. By hypothesis $|u| < |u'|$. This implies that $v = v'a$ with a a prefix of u'. Hence the first letter of a is the first letter of u', which is c_1, and the letter following a is the first letter of $\widehat{u}$, which is c_1. We have $u' = a\widehat{u} = \widehat{s}(w') \in \widehat{\mathcal{P}}^\beta$. By definition of the map $w \mapsto \widehat{w}$ (see (5) and (3)) we conclude that a is a prefix of c^β. By Lemma 2.1, au is a prefix of c^β, contradicting the maximality of u. Therefore $|u| = |u'|$, and in this case the mapping $w \mapsto \widehat{w}$ is two-to-one. In the remaining cases $|s(w)| \le p_1 + 1$. Therefore the mapping $w \mapsto \widehat{w}$ is at most $(p_1 + 2)$-to-one. $\qquad\square$

Definition 2.3. The natural extension Σ^β of X^β is

$$\Sigma^\beta = \{x \in \mathtt{A}^{\mathbb{Z}} : \text{ for all } k \in \mathbb{Z}, \ (x_k, x_{k+1}, \dots) \in X^\beta\}.$$

It is henceforth simply called the *β-shift*.

The language of Σ^β is also $\mathcal{L}^\beta$. Let $k < \ell$, $[k, \ell] = \{k, k+1, \dots, \ell-1, \ell\}$, and $x \in \Sigma^\beta$. The projection $J_{[k,\ell]} : \Sigma^\beta \to \mathcal{L}^\beta$ is defined as

$$x \mapsto J_{[k,\ell]}(x) := x_{[k,\ell]} \equiv x_k x_{k+1} \cdots x_\ell.$$

Let $w = x_k x_{k+1} \cdots x_{k+m-1} \in \mathcal{L}_m^\beta$. We can always extend w to the left by 0, that is, there exist $y \in \Sigma^\beta$, $y_j = 0$, $j < k$, and $y_j = x_j$, $j = k, \dots, k+m-1$. We can also extend w to the right by 0. If $q(w) = q_0$, this is clear. If $w = vu$, $s(w) = u \ne \epsilon$, then $u = c_1 \cdots c_p$ for some $p \ge 1$. When $c_{p+1} \ne 0$, we may change c_{p+1} into 0. When $c_{p+1} = \cdots = c_{p+r} = 0$, but $c_{p+r+1} \ne 0$, we may change c_{p+r+1} into 0. Hence there exist $y \in \Sigma^\beta$, $y_j = x_j$, for all $j = k, \dots, k+m-1$, and $y_j = 0$ for all $j < k$ and $j \ge m$.

2.2. *Functions of bounded total oscillations.* We recall the definition of a function of bounded total oscillations. For details we refer to [**PS3**], §3. Let $\bar{f} \in C(\mathtt{A}^{\mathbb{Z}})$ be a continuous function defined on the *full shift* $\mathtt{A}^{\mathbb{Z}}$. On $\mathtt{A}^{\mathbb{Z}}$ we define for each $i \in \mathbb{Z}$,

$$\delta_i(\bar{f}) := \sup\{|\bar{f}(x) - \bar{f}(y)| : x, y \in \mathtt{A}^{\mathbb{Z}}, \text{ where } x_k = y_k \text{ for all } k \ne i\}$$

and

$$\|\bar{f}\|_\delta := \sum_{i \in \mathbb{Z}} \delta_i(\bar{f}).$$

A function has *bounded total oscillations* if $\|\bar{f}\|_\delta < \infty$. On a subshift $X \subset \mathtt{A}^\mathbb{Z}$, $\delta_i(f)$ may not make sense. If $f \in C(X)$ is a continuous function on X and has a continuous extension $\bar{f} \in C(\mathtt{A}^\mathbb{Z})$ on $\mathtt{A}^\mathbb{Z}$, we write

$$f \approx \bar{f} \iff f(x) = \bar{f}(x) \quad \text{for all } x \in X.$$

An extension $\bar{f}$ of f exists [**PS3**, Proposition 3.2]. For $f \in C(X)$ we define

$$\|f\|_\delta := \inf\{\|\bar{f}\|_\delta : \bar{f} \in C(\mathtt{A}^\mathbb{Z}), \ \bar{f} \approx f\}.$$

A function $f \in C(X)$ has *bounded total oscillations* if $\|f\|_\delta < \infty$. Examples of functions of bounded total oscillations are given in [**PS3**]. The set of bounded total oscillations is a Banach space $\mathcal{B}(\Sigma^\beta)$ with the norm [**PS3**, Proposition 3.1]

$$\|f\|_\delta + \sup_{x \in X} |f(x)|.$$

We prove two basic estimates for functions with bounded total oscillations. For convenience, from now on we write $\bar{f}$ for a continuous extension of f to $\mathtt{A}^\mathbb{Z}$. The arguments do not require that $\bar{f}$ satisfies $\|\bar{f}\|_\delta = \|f\|_\delta$ but just that $f \approx \bar{f}$ and $\|\bar{f}\|_\delta < \infty$. Fundamental to many of the arguments is the following lemma.

LEMMA 2.4. *Let $x, y \in X$ and $\Gamma := \{j : x_j \neq y_j\}$. Then for $\Lambda \subset \mathbb{Z}$ and $f \in C(X)$,*

$$\sum_{i \in \Lambda} |f(T^i x) - f(T^i y)| \leq \sum_{i \in \Lambda} \sum_{j \in \Gamma} \delta_{j-i}(\bar{f}) \leq \infty.$$

Proof. Since Γ is at most countable, we can list the elements of Γ, so that $\Gamma = \{j_1, j_2, \ldots\}$. We define a sequence of elements of $\mathtt{A}^\mathbb{Z}$ as follows. Let $z^{j_0} := x$. For $j_\ell \in \Gamma$, set

$$z_k^{j_\ell} := \begin{cases} y_k & \text{if } k = j_1, \ldots, j_\ell, \\ x_k & \text{if } k \in \mathbb{Z} \setminus \{j_1, \ldots, j_\ell\}. \end{cases}$$

Then

$$|f(T^i x) - f(T^i y)| \leq \sum_{j_k \in \Gamma} |\bar{f}(T^i z^{j_{k-1}}) - \bar{f}(T^i z^{j_k})| \leq \sum_{j_k \in \Gamma} \delta_{j_k}(\bar{f} \circ T^i).$$

The lemma follows from the identity

$$\delta_j(\bar{f} \circ T^i) = \delta_{j-i}(\bar{f}).$$

Indeed,

$$\begin{aligned} \delta_j(\bar{f} \circ T^i) &= \sup\{|(\bar{f} \circ T^i)(x) - (\bar{f} \circ T^i)(y)| : x_k = y_k \text{ for all } k \neq j\} \\ &= \sup\{|\bar{f}(T^i x) - \bar{f}(T^i y)| : (T^{-i} T^i x)_k = (T^{-i} T^i y)_k \text{ for all } k \neq j\} \\ &= \sup\{|\bar{f}(x') - \bar{f}(y')| : (T^{-i} x')_k = (T^{-i} y')_k \text{ for all } k \neq j\} = \delta_{j-i}(\bar{f}). \quad \square \end{aligned}$$

LEMMA 2.5. *Let f be a function of bounded total oscillations on X. Given $\varepsilon > 0$, there exists N_ε such that for $m \geq N_\varepsilon$,*

$$\sup \left\{ \sum_{1 \leq i \leq m} |f(T^i x) - f(T^i y)| : x, y, x_k = y_k \text{ for all } k \in \{1, \ldots, m\} \right\} \leq m\varepsilon \quad (7)$$

and

$$\sup \left\{ \sum_{j \notin \{1,\ldots,m\}} |f(T^j x) - f(T^j y)| : x, y, x_k = y_k \text{ for all } k \notin \{1, \ldots, m\} \right\} \leq m\varepsilon. \quad (8)$$

Proof. Let $\varepsilon > 0$ be given. There exists r_ε so that $\sum_{k:|k|>r_\varepsilon} \delta_k(\bar{f}) \leq \varepsilon/2$. If $m > 2r_\varepsilon$, then for $x_{[1,m]} = y_{[1,m]}$ the sum over $[1, m]$ of $|T^i f(x) - T^i f(y)|$ can be written as over

$$[1, r_\varepsilon] \cup [r_\varepsilon + 1, m - r_\varepsilon] \cup [m + 1 - r_\varepsilon, m].$$

For i in the middle interval and $j \notin [1, m]$ we have $|i - j| > r_\varepsilon$, so that by Lemma 2.4,

$$|f(T^i x) - f(T^i y)| \leq \sum_{j \notin \{1,\ldots,m\}} \delta_j(\bar{f} \circ T^i) \leq \sum_{k:|k|>r_\varepsilon} \delta_k(\bar{f}) \leq \varepsilon/2.$$

For the i in the outside intervals we use $|f(T^i x) - f(T^i y)| \leq 2\|f\|$ to yield

$$\frac{1}{m} \sum_{1 \leq i \leq m} |f(T^i x) - f(T^i y)| \leq \frac{4r_\varepsilon \|f\|}{m} + \frac{\varepsilon(m - 2r_\varepsilon)}{2m} \leq \varepsilon \quad \text{for } m \text{ large enough}.$$

The proof of the second statement is similar. □

2.3. *Equilibrium measure and pressure.* In our setting a shift-invariant (Borel) probability measure ν is an equilibrium measure for a continuous function φ if and only if ν is a tangent functional to the pressure p at φ (see [Wa, Theorems 8.2 and 9.5]).

Definition 2.6. An invariant probability measure ν is a *tangent functional to the pressure* p *at* φ if

$$p(\varphi + f) \geq p(\varphi) + \int f \, d\nu \quad \text{for all continuous functions } f.$$

The set of tangent functionals to the pressure at φ is denoted $\partial p(\varphi)$.

For each $n \in \mathbb{N}$ we choose a set $E^n \subset \Sigma^\beta$ with the following properties:

$$J_{[-n,n]}(E^n) = \mathcal{L}_{2n+1}^\beta \quad \text{and} \quad (x, x' \in E^n, \ J_{[-n,n]}(x) = J_{[-n,n]}(x')) \implies x = x'. \quad (9)$$

Let φ be a continuous function and set

$$\Xi^n(\varphi) := \sum_{x \in E^n} \exp \left(\sum_{j \in [-n,n]} \varphi(T^j x) \right) \quad \text{and} \quad P_{E^n}(\varphi) := \frac{1}{2n+1} \ln \Xi^n(\varphi). \quad (10)$$

The pressure $p(\varphi)$ is defined as

$$p(\varphi) = \lim_{n \to \infty} P_{E^n}(\varphi). \quad (11)$$

The result in (11) is independent of the choice of the sets E^n. From now on we choose E^n so that if $x \in E^n$, then $x_j = 0$, for all $|j| > n$.

Let $[k, \ell] \subset [-n, n]$, $x \in E^n$. Set

$$x_k^- := x_{(-\infty, k-1]} \quad \text{and} \quad x_\ell^+ := x_{[\ell+1, \infty)}.$$

By our choice of E^n we can extend (4) and (5) to the infinite sequence x_k^- since $x_j = 0$ for $j < -n$. Let $s(x_k^-)$ be the largest suffix $\in \mathcal{P}^\beta$ of x_k^-, and

$$\widehat{x_k^-} := y \widehat{s(x_k^-)} \quad \text{where } x_k^- = y s(x_k^-).$$

Let

$$E_{[k,\ell]}^n(v) := \{x \in E^n : x_{[k,\ell]} = v\} \quad \text{and} \quad E_{[k,\ell]}^{*,n}(v) := \{x \in E_{[k,\ell]}^n(v) : s(x_k^-) = \epsilon\},$$

and set

$$\Xi_{[k,\ell]}^n(v) := \sum_{x \in E_{[k,\ell]}^n(v)} \exp\left(\sum_{j \in [-n,n]} \varphi(T^j x)\right), \tag{12}$$

$$\Xi_{[k,\ell]}^{*,n}(v) := \sum_{x \in E_{[k,\ell]}^{*,n}(v)} \exp\left(\sum_{j \in [-n,n]} \varphi(T^j x)\right).$$

We have

$$\Xi^n(\varphi) = \sum_v \Xi_{[k,\ell]}^n(v). \tag{13}$$

Lemmas 2.7 and 2.8 give basic estimates used in the proof of Theorem 2.12.

LEMMA 2.7. *Let $[k, \ell] \subset [-n, n]$ and $\|\varphi\|_\delta < \infty$. Then*

$$\Xi_{[k,\ell]}^{*,n}(v) \leq \Xi_{[k,\ell]}^n(v) \leq (z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta} \Xi_{[k,\ell]}^{*,n}(\widehat{v}).$$

Proof. This first inequality follows from $E_{[k,\ell]}^{*,n}(v) \subset E_{[k,\ell]}^n(v)$. We define a map $\widehat{f}$ from $E_{[k,\ell]}^n(v)$ to $E_{[k,\ell]}^{*,n}(\widehat{v})$ by setting

$$\widehat{f} : E_{[k,\ell]}^n(v) \to E_{[k,\ell]}^{*,n}(\widehat{v}), \quad x = x_k^- v x_\ell^+ \mapsto \widehat{f}(x) := \widehat{x_k^-} \widehat{v} x_\ell^+. \tag{14}$$

Since $\widehat{x_k^-}$ and $\widehat{v}$ are presented by paths in $\mathcal{G}^\beta$ with end-point q_0, $\widehat{f}(x)$ is well defined and $\widehat{f}(x) \in E_{[k,\ell]}^{*,n}(\widehat{v})$. By Lemma 2.2 the map $x_k^- \mapsto \widehat{x_k^-}$ is at most $(z^\beta(c_1) + 2)$-to-one. Hence the map $\widehat{f}$ is at most $(z^\beta(c_1) + 2)$-to-one (v is fixed, hence $\widehat{v}$ is also fixed). The sequences x and $x' = \widehat{f}(x)$ differ at most at two coordinates, so that by Lemma 2.4,

$$\left| \sum_{j \in [-n,n]} (\varphi(T^j x) - \varphi(T^j x')) \right| \leq 2\|\varphi\|_\delta.$$

Hence

$$\sum_{x \in E^n_{[k,\ell]}(v)} \exp\left(\sum_{j \in [-n,n]} \varphi(T^j x)\right) \le e^{2\|\varphi\|_\delta} \sum_{x \in E^n_{[k,\ell]}(v)} \exp\left(\sum_{j \in [-n,n]} \varphi(T^j x')\right)$$

$$\le (z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta} \sum_{\substack{y \in E^n_{[k,\ell]}(\widehat{v}): \\ s(y_k^-) = \epsilon}} \exp\left(\sum_{j \in [-n,n]} \varphi(T^j y)\right)$$

$$= (z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta} \; \Xi^{*,n}_{[k,\ell]}(\widehat{v}). \qquad \square$$

LEMMA 2.8. *Let $[k, \ell] \subset [-n, n]$ and $\|\varphi\|_\delta < \infty$. Then*

$$\Xi^{*,n}_{[k,\ell]}(v) \ge |\mathbb{A}|^{-(z^\beta(v)+1)} e^{-(z^\beta(v)+2)\|\varphi\|_\delta} \; \Xi^{*,n}_{[k,\ell]}(\widehat{v}).$$

Proof. If $s(v) = \epsilon$, then $v = \widehat{v}$ and the inequality is trivial. Let v with $s(v) \ne \epsilon$; we define a map $f : E^{*,n}_{[k,\ell]}(\widehat{v}) \to E^{*,n}_{[k,\ell]}(v)$,

$$f(x_k^- \widehat{v} x_\ell^+) := x_k^- v x'^+_\ell,$$

with

$$x'^+_{\ell+1} = \cdots = x'^+_{\ell+z^\beta(v)+1} = 0 \quad \text{and} \quad x'^+_j = x_j \quad \text{if } j > \ell + z^\beta(v) + 1.$$

This map is at most $|\mathbb{A}|^{z^\beta(v)+1}$-to-one. We have

$$\sum_{j \in [-n,n]} \varphi(T^j(x_k^- \widehat{v} x_\ell^+)) = \sum_{j \in [-n,n]} (\varphi(T^j(x_k^- \widehat{v} x_\ell^+)) - \varphi(T^j(x_k^- \widehat{v} x'^+_\ell)))$$

$$+ \sum_{j \in [-n,n]} (\varphi(T^j(x_k^- \widehat{v} x'^+_\ell)) - \varphi(T^j(x_k^- v x'^+_\ell)))$$

$$+ \sum_{j \in [-n,n]} \varphi(T^j(x_k^- v x'^+_\ell)).$$

The configurations $x_k^- \widehat{v} x_\ell^+$ and $x_k^- \widehat{v} x'^+_\ell$ differ at most at $i = \ell + 1, \ldots, \ell + z^\beta(v) + 1$, so that

$$\left| \sum_{j \in [-n,n]} (\varphi(T^j(x_k^- \widehat{v} x_\ell^+)) - \varphi(T^j(x_k^- \widehat{v} x'^+_\ell))) \right| \le (z^\beta(v) + 1)\|\varphi\|_\delta. \qquad (15)$$

The configurations $x_k^- \widehat{v} x'^+_\ell$ and $x_k^- v x'^+_\ell$ differ at one coordinate, so that

$$\left| \sum_{j \in [-n,n]} (\varphi(T^j(x_k^- \widehat{v} x'^+_\ell)) - \varphi(T^j(x_k^- v x'^+_\ell))) \right| \le \|\varphi\|_\delta.$$

Therefore

$$
\begin{aligned}
\Xi_{[k,\ell]}^{*,n}(\widehat{v}) &= \sum_{x \in E_{[k,\ell]}^{*,n}(\widehat{v})} \exp \sum_{j \in [-n,n]} \varphi(T^j x) \\
&= \sum_{y \in E_{[k,\ell]}^{*,n}(v)} \sum_{\substack{x \in E_{[k,\ell]}^{*,n}(\widehat{v}): \\ f(x)=y}} \exp \sum_{j \in [-n,n]} \varphi(T^j x) \\
&\leq |\mathbb{A}|^{z^\beta(v)+1} e^{(z^\beta(v)+2)\|\varphi\|_\delta} \sum_{y \in E_{[k,\ell]}^{*,n}(v)} \exp \sum_{j \in [-n,n]} \varphi(T^j y) \\
&= |\mathbb{A}|^{z^\beta(v)+1} e^{(z^\beta(v)+2)\|\varphi\|_\delta} \; \Xi_{[k,\ell]}^{*,n}(v).
\end{aligned}
$$

$\square$

LEMMA 2.9. *Let $w \in \mathcal{L}_m^\beta$ and $w^\sharp \in \Sigma^\beta$,*

$$
w_j^\sharp := \begin{cases} w_j & \text{if } 1 \leq j \leq m, \\ 0 & \text{otherwise.} \end{cases} \tag{16}
$$

The pressure $p(\varphi)$ is equal to

$$
\lim_{m \to \infty} \frac{1}{m} \ln \sum_{w \in \mathcal{L}_m^\beta : s(w)=\epsilon} \exp \sum_{j=1}^m \varphi(T^j w^\sharp).
$$

Proof. The configurations $w^\sharp$ and $\widehat{w}^\sharp$ differ at most at one coordinate, so that

$$
\left| \sum_{j=1}^m (\varphi(T^j \widehat{w}^\sharp) - \varphi(T^j w^\sharp)) \right| \leq \|\varphi\|_\delta.
$$

The map $w \mapsto \widehat{w}$ is at most $(z^\beta(c_1)+2)$-to-one (Lemma 2.2). Therefore

$$
\begin{aligned}
\sum_{w \in \mathcal{L}_m^\beta} \exp \sum_{j=1}^m \varphi(T^j w^\sharp) &\leq e^{\|\varphi\|_\delta} \sum_{w \in \mathcal{L}_m^\beta} \exp \sum_{j=1}^m \varphi(T^j \widehat{w}^\sharp) \\
&\leq e^{\|\varphi\|_\delta} (z^\beta(c_1)+2) \sum_{\widehat{w} \in \mathcal{L}_m^\beta} \exp \sum_{j=1}^m \varphi(T^j \widehat{w}^\sharp) \\
&= e^{\|\varphi\|_\delta} (z^\beta(c_1)+2) \sum_{w \in \mathcal{L}_m^\beta : s(w)=\epsilon} \exp \sum_{j=1}^m \varphi(T^j w^\sharp).
\end{aligned}
$$

Hence,

$$
p(\varphi) \leq \lim_{m \to \infty} \frac{1}{m} \ln \sum_{w \in \mathcal{L}_m^\beta : s(w)=\epsilon} \exp \sum_{j=1}^m \varphi(T^j w^\sharp).
$$

On the other hand,

$$\sum_{w \in \mathcal{L}_m^\beta} \exp \sum_{j=1}^m \varphi(T^j w^\sharp) \geq e^{-\|\varphi\|_\delta} \sum_{w \in \mathcal{L}_m^\beta} \exp \sum_{j=1}^m \varphi(T^j \widehat{w}^\sharp)$$

$$\geq e^{-\|\varphi\|_\delta} \sum_{w \in \mathcal{L}_m^\beta : s(w)=\epsilon} \exp \sum_{j=1}^m \varphi(T^j w^\sharp).$$

Hence,

$$p(\varphi) \geq \lim_{m \to \infty} \frac{1}{m} \ln \sum_{w \in \mathcal{L}_m^\beta : s(w)=\epsilon} \exp \sum_{j=1}^m \varphi(T^j w^\sharp). \qquad \square$$

2.4. *Main result.*

Definition 2.10. An invariant probability measure ν is a *weak Gibbs measure for a continuous function ψ*, if for any $\delta > 0$ there exists N_δ such that for $m \geq N_\delta$,

$$\sup_{x \in X} \left| \frac{1}{m} \ln \nu([x_0 \cdots x_{m-1}]) - \frac{1}{m} \sum_{\ell=0}^{m-1} \psi(T^\ell(x)) \right| \leq \delta,$$

where $[x_0 \cdots x_{m-1}] = \{y \in \Sigma^\beta : y_0 \cdots y_{m-1} = x_0 \cdots x_{m-1}\}$.

Remark 2.11. Our definition of weak Gibbs measure is stated so that if ν is weak Gibbs for ψ, then $p(\psi) = 0$ [**PS2**]. If $\psi = \varphi - p(\varphi)$, then it equivalent to

$$e^{-\delta m} \leq \frac{\nu([x_0 \cdots x_{m-1}])}{\exp(-mp(\varphi) + \sum_{\ell=0}^{m-1} \varphi(T^\ell x))} \leq e^{\delta m}.$$

THEOREM 2.12. *Let $\beta > 1$ and φ be a function of bounded total oscillations on Σ^β.*
(1) *If ν is an equilibrium measure for φ and if*

$$\lim_{n \to \infty} \frac{\overline{z}^\beta(n)}{n} = 0,$$

   then ν is a weak Gibbs measure for $\psi = \varphi - p(\varphi)$.
(2) *If ν is an equilibrium measure for φ and if*

$$\limsup_{n \to \infty} \frac{\overline{z}^\beta(n)}{n} > 0,$$

   then ν is not a weak Gibbs measure for $\psi = \varphi - p(\varphi)$.

   If ν is a weak Gibbs measure, then the empirical measures satisfy a large-deviations principle [**PS2**]. Large deviations for (one-sided) β-shifts, for any $\beta > 1$, and equilibrium measures have been proved by Climenhaga, Thompson and Yamamoto [**CTY**] for the class of functions satisfying the Bowen condition. From the estimates of Lemma 3.2 [**PS1**, Proposition 4.3 and Theorem 3.1], the result of [**CTY**] is also valid for all equilibrium measures for functions φ of bounded total oscillations. This is important since the Bowen

condition implies uniqueness of the equilibrium measure for β-shifts, while this is not necessarily the case for bounded total oscillations functions.

3. *Proof of Theorem 2.12*

Let φ be a function of bounded total oscillations on Σ^β. In §3.1 we prove upper and lower bounds for $\nu([y_0 \cdots y_{m-1}])$ for any equilibrium measure ν of φ. There is no restriction on $\beta > 1$. In §3.2 we prove Theorem 2.12.

3.1. *Upper and lower bounds.* We first assume that there is a unique tangent functional ν to the pressure at φ. The result is then extended to any φ of bounded total oscillations using a theorem of Mazur and a theorem of Lanford and Robinson (see, for example, [**Ru**, Appendix A.3.7]).

When there is a unique tangent functional to the pressure at φ we can estimate $\nu([y_0 \cdots y_{m-1}])$ using a classical result about differentiability of a convex function, here the pressure, which is a pointwise limit of convex functions [**Ro**, Theorem 25.7]. Let $\bar{u} \in \mathcal{L}_m^\beta$ be fixed and set

$$I_{\bar{u}}(y) := \begin{cases} 1 & \text{if } y_0 \cdots y_{m-1} = \bar{u}, \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\nu(I_{\bar{u}}) = \frac{d}{dt} \lim_{n \to \infty} P_{E^n}(\varphi + t\, I_{\bar{u}}) \Big|_{t=0} = \lim_{n \to \infty} \frac{d}{dt} P_{E^n}(\varphi + t\, I_{\bar{u}}) \Big|_{t=0}$$

$$= \lim_{n \to \infty} \frac{1}{2n+1} \sum_{j=-n}^{n} \frac{\sum_{x \in E^n} I_{\bar{u}}(T^j x) \exp \sum_{i=-n}^{n} \varphi(T^i x)}{\sum_{x \in E^n} \exp \sum_{i=-n}^{n} \varphi(T^i x)}. \tag{17}$$

Let $j = k$ and $\ell := k + m - 1$. Then a term in (17) is written as a ratio of partition functions (see (12) and (13))

$$\frac{\sum_{x \in E^n} I_{\bar{u}}(T^j x) \exp \sum_{i=-n}^{n} \varphi(T^i x)}{\sum_{x \in E^n} \exp \sum_{i=-n}^{n} \varphi(T^i x)} = \frac{\Xi_{[k,\ell]}^n(\bar{u})}{\Xi^n(\varphi)} = \frac{\Xi_{[k,\ell]}^n(\bar{u})}{\sum_v \Xi_{[k,\ell]}^n(v)}. \tag{18}$$

The core of the proof involves estimating the ratio of partition functions $\Xi_{[k,\ell]}^n(v)/\Xi_{[k,\ell]}^n(\bar{u})$, uniformly in $[k,\ell] \subset [-n, n]$ using Lemmas 2.7 and 2.8. Since in (17) we take the limit $n \to \infty$, it is sufficient to consider the cases where $[k,\ell] \subset [-n, n]$.

LEMMA 3.1. *Let $\bar{u} \in \mathcal{L}_m^\beta$. For any $\varepsilon > 0$ and continuous φ such that $\|\varphi\|_\delta < \infty$ there exists $N_{\varepsilon,\varphi}$ so that if y is such that $J_{[1,m]}(y) = \bar{u}$ and $m \geq N_{\varepsilon,\varphi}$, then*

$$\nu(I_{\bar{u}}) \leq K_{\varphi,\varepsilon}^+(m, \beta) \exp\left(\sum_{j=1}^{m} \varphi(T^j y) - m p(\varphi) \right),$$

where

$$K_{\varphi,\varepsilon}^+(m, \beta) = (z^\beta(c_1) + 2) e^{3\|\varphi\|_\delta} e^{5m\varepsilon}.$$

$N_{\varepsilon,\varphi}$ is chosen so that all of the inequalities (7), (8) and (23) are satisfied for $m \geq N_{\varepsilon,\varphi}$.

Proof. Let $\varepsilon > 0$ and φ be given and $N_{\varepsilon,\varphi}$ defined as above. We consider the term in the sum (17) with $j = k$ and $[k, \ell] \subset [-n, n]$ (see (18)),

$$\frac{\Xi^n_{[k,\ell]}(\bar{u})}{\Xi^n(\varphi)} = \frac{\Xi^n_{[k,\ell]}(\bar{u})}{\sum_v \Xi^n_{[k,\ell]}(v)}.$$

We have

$$\sum_v \Xi^n_{[k,\ell]}(v) \geq \sum_{v:s(v)=\epsilon} \Xi^n_{[k,\ell]}(v) \geq \sum_{v:s(v)=\epsilon} \Xi^{*,n}_{[k,\ell]}(v).$$

From now on $s(v) = \epsilon$. By Lemma 2.7,

$$\Xi^n_{[k,\ell]}(\bar{u}) \leq (z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta} \Xi^{*,n}_{[k,\ell]}(\widehat{\bar{u}}).$$

Hence,

$$\frac{\Xi^{*,n}_{[k,\ell]}(v)}{\Xi^n_{[k,\ell]}(\bar{u})} \geq \frac{1}{(z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta}} \frac{\Xi^{*,n}_{[k,\ell]}(v)}{\Xi^{*,n}_{[k,\ell]}(\widehat{\bar{u}})}.$$

Since $s(v) = \epsilon$ and $s(\widehat{\bar{u}}) = \epsilon$, if $s(x_k^-) = \epsilon$, then

$$x_k^- v x_\ell^+ \in E^{*,n}_{[k,\ell]}(v) \iff x_k^- \widehat{\bar{u}} x_\ell^+ \in E^{*,n}_{[k,\ell]}(\widehat{\bar{u}}). \tag{19}$$

We write

$$\Xi^{*,n}_{[k,\ell]}(v) = \sum_{x \in E^{*,n}_{[k,\ell]}(v)} \frac{\exp(\sum_{j=-n}^n \varphi(T^j(x_k^- v x_\ell^+)))}{\exp(\sum_{j=-n}^n \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+)))}$$

$$\times \exp\left(\sum_{j=-n}^n \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+)) \right). \tag{20}$$

Let $v^\sharp$ and $\widehat{\bar{u}}^\sharp$ be defined as in (16) with $1 \leq j \leq m$ replaced by $k \leq j \leq \ell$. Then

$$\sum_{j=-n}^n (\varphi(T^j(x_k^- v x_\ell^+)) - \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+)))$$

$$= \sum_{j \notin [k,\ell]} (\varphi(T^j(x_k^- v x_\ell^+)) - \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+)))$$

$$+ \sum_{j \in [k,\ell]} (\varphi(T^j(x_k^- v x_\ell^+)) - \varphi(T^j v^\sharp)) + \sum_{j \in [k,\ell]} (\varphi(T^j \widehat{\bar{u}}^\sharp) - \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+)))$$

$$+ \sum_{j \in [k,\ell]} (\varphi(T^j v^\sharp) - \varphi(T^j \widehat{\bar{u}}^\sharp)). \tag{21}$$

By Lemma 2.5, if $m \geq N_{\varepsilon,\varphi}$, then

$$\left| \sum_{j=-n}^n (\varphi(T^j(x_k^- v x_\ell^+)) - \varphi(T^j(x_k^- \widehat{\bar{u}} x_\ell^+))) - \sum_{j \in [k,\ell]} (\varphi(T^j v^\sharp) - \varphi(T^j \widehat{\bar{u}}^\sharp)) \right| \leq 3\varepsilon m.$$

Let $y \in \Sigma^\beta$ be such that $J_{[k,\ell]}(y) = \bar{u}$. By definition of $\bar{u}^\sharp$,

$$J_{[k,\ell]}(\bar{u}^\sharp) = \bar{u} \quad \text{and} \quad \bar{u}_i^\sharp = 0 \quad \text{for all } i \notin [k, \ell].$$

By Lemmas 2.4 and 2.5, if $m \geq N_{\varepsilon,\varphi}$, then

$$
\begin{aligned}
& \left| \sum_{j \in [k,\ell]} (\varphi(T^j \widehat{\bar{u}}^\sharp) - \varphi(T^j y)) \right| \\
& \leq \left| \sum_{j \in [k,\ell]} (\varphi(T^j \widehat{\bar{u}}^\sharp) - \varphi(T^j \bar{u}^\sharp)) \right| + \left| \sum_{j \in [k,\ell]} (\varphi(T^j \bar{u}^\sharp) - \varphi(T^j y)) \right| \\
& \leq \|\varphi\|_\delta + m\varepsilon.
\end{aligned}
\tag{22}
$$

$N_{\varepsilon,\varphi}$ has been chosen so that (see Lemma 2.9)

$$
e^{-m\varepsilon} \leq \frac{e^{mp(\varphi)}}{\sum_{\substack{v \in \mathcal{L}_m^\beta: \\ s(v)=\epsilon}} \exp \sum_{j=1}^m \varphi(T^j v^\sharp)} \leq e^{m\varepsilon}.
\tag{23}
$$

From (18), (20), $s(v) = \epsilon$ and the above estimates, taking into account (19), which allows the use of the elementary inequalities for positive real numbers a_i and b_i,

$$
\inf_i \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} = \frac{\sum_{i=1}^n (a_i/b_i) \, b_i}{\sum_{i=1}^n b_i} \leq \sup_i \frac{a_i}{b_i},
\tag{24}
$$

we get

$$
\begin{aligned}
\frac{\sum_v \Xi_{[k,\ell]}^n(v)}{\Xi_{[k,\ell]}^n(\bar{u})} & \geq \frac{\sum_{v:s(v)=\epsilon} \Xi_{[k,\ell]}^{*,n}(v)}{\Xi_{[k,\ell]}^n(\bar{u})} \geq \frac{1}{(z^\beta(c_1)+2)e^{2\|\varphi\|_\delta}} \frac{\sum_{v:s(v)=\epsilon} \Xi_{[k,\ell]}^{*,n}(v)}{\Xi_{[k,\ell]}^n(\widehat{\bar{u}})} \\
& \geq e^{-4m\varepsilon - \|\varphi\|_\delta} \frac{e^{-\sum_{j \in [k,\ell]} \varphi(T^j y)}}{(z^\beta(c_1)+2)e^{2\|\varphi\|_\delta}} \sum_{v:s(v)=\epsilon} \exp \sum_{j \in [k,\ell]} \varphi(T^j v^\sharp) \\
& \geq \frac{e^{-5m\varepsilon}}{(z^\beta(c_1)+2)e^{3\|\varphi\|_\delta}} \exp\left(-\sum_{j \in [k,\ell]} \varphi(T^j y) + mp(\varphi) \right).
\end{aligned}
$$

The result follows from (17) by taking the limit $n \to \infty$. $\qquad \square$

LEMMA 3.2. *Let* $\bar{u} \in \mathcal{L}_m^\beta$. *For any* $\varepsilon > 0$ *and continuous* φ *such that* $\|\varphi\|_\delta < \infty$ *there exists* $N_{\varepsilon,\varphi}$ *so that if* y *is such that* $J_{[1,m]}(y) = \bar{u}$ *and* $m \geq N_{\varepsilon,\varphi}$, *then*

$$
\nu(I_{\bar{u}}) \geq K_{\varphi,\varepsilon}^-(m, \beta, \bar{u}) \exp\left(\sum_{j=1}^m \varphi(T^j y) - mp(\varphi) \right),
$$

where

$$
K_{\varphi,\varepsilon}^-(m, \beta, \bar{u}) = \frac{|\mathbb{A}|^{-(z^\beta(\bar{u})+1)} e^{-(z^\beta(\bar{u})+2)\|\varphi\|_\delta}}{(z^\beta(c_1)+2)^2 e^{5m\varepsilon+3\|\varphi\|_\delta}}.
$$

$N_{\varepsilon,\varphi}$ *is chosen so that all of the inequalities* (7), (8) *and* (23) *are satisfied for* $m \geq N_{\varepsilon,\varphi}$.

Proof. Let $\varepsilon > 0$ and φ be given and $N_{\varepsilon,\varphi}$ defined as above. We consider the term in the sum (17) with $j = k$ and $[k, \ell] \subset [-n, n]$ (see (18)),

$$\frac{\Xi^n_{[k,\ell]}(\bar{u})}{\Xi^n(\varphi)} = \frac{\Xi^n_{[k,\ell]}(\bar{u})}{\sum_v \Xi^n_{[k,\ell]}(v)}.$$

As in the proof of Lemma 3.1, if $s(v) = \epsilon$, then we can estimate the ratio $\Xi^{*,n}_{[k,\ell]}(v)/\Xi^{*,n}_{[k,\ell]}(\widehat{u})$ using (24) since (19) holds. If $y \in \Sigma^\beta$ is such that $J_{[k,\ell]}(y) = \bar{u}$ and $s(v) = \epsilon$, then (see (21) and (22))

$$\frac{\Xi^{*,n}_{[k,\ell]}(v)}{\Xi^{*,n}_{[k,\ell]}(\widehat{u})} \leq e^{4m\varepsilon + \|\varphi\|_\delta} \exp\left(\sum_{j \in [k,\ell]} \varphi(T^j v^\sharp) - \varphi(T^j y) \right).$$

Let $m \geq N_{\varepsilon,\varphi}$. Then

$$
\begin{aligned}
\frac{\sum_v \Xi^n_{[k,\ell]}(v)}{\Xi^n_{[k,\ell]}(\bar{u})} &\leq \sum_v \frac{(z^\beta(c_1) + 2)e^{2\|\varphi\|_\delta}}{|\mathbb{A}|^{-(z^\beta(\bar{u})+1)}e^{-(z^\beta(\bar{u})+2)\|\varphi\|_\delta}} \frac{\Xi^{*,n}_{[k,\ell]}(\widehat{v})}{\Xi^{*,n}_{[k,\ell]}(\widehat{u})} \\
&\leq \frac{(z^\beta(c_1) + 2)^2 e^{2\|\varphi\|_\delta}}{|\mathbb{A}|^{-(z^\beta(\bar{u})+1)}e^{-(z^\beta(\bar{u})+2)\|\varphi\|_\delta}} \sum_{v:s(v)=\epsilon} \frac{\Xi^{*,n}_{[k,\ell]}(v)}{\Xi^{*,n}_{[k,\ell]}(\widehat{u})} \\
&\leq \frac{(z^\beta(c_1) + 2)^2 e^{5m\varepsilon+3\|\varphi\|_\delta}}{|\mathbb{A}|^{-(z^\beta(\bar{u})+1)}e^{-(z^\beta(\bar{u})+2)\|\varphi\|_\delta}} \exp\left(\sum_{j \in [k,\ell]} -\varphi(T^j y) + mp(\varphi) \right).
\end{aligned}
$$

For the first inequality we use Lemmas 2.7 and 2.8, and for the second inequality, where we replace $\widehat{v}$ by v with $s(v) = \epsilon$, we use Lemma 2.2 and $s(\widehat{v}) = s(v) = \epsilon$. The result follows from (17) by taking the limit $n \to \infty$. $\qquad\square$

We now remove the restriction that ν is the unique equilibrium measure for φ. The pressure is convex and continuous on the Banach space $\mathcal{B}(\Sigma^\beta)$ of bounded total oscillations functions. The set $\mathcal{R} \subset \mathcal{B}$ of φ, such that $\partial p(\varphi) = \{\nu\}$ has a unique element ν, is residual (theorem of Mazur).

Let $\varphi \in \mathcal{B}$ be an arbitrary function of bounded total oscillations and ν be an equilibrium measure for φ, such that there exists a sequence $\varphi_k \in \mathcal{R}$ with the properties that $\lim_k \varphi_k = \varphi$ and $\lim_k \nu_k = \nu$, $\{\nu_k\} = \partial p(\varphi_k)$ (weak convergence). For such φ let $N_{\varepsilon,\varphi}$ be defined as in Lemmas 3.1 and 3.2. Let $\varepsilon' > \varepsilon$. By our choice of r_ε (see the proof of Lemma 2.5),

$$\sum_{k:|k|>r_\varepsilon} \delta_k(\varphi_k) \leq \sum_{k:|k|>r_\varepsilon} \delta_k(\varphi) + \|\varphi_k - \varphi\|_\delta \leq \frac{\varepsilon}{2} + \|\varphi_k - \varphi\|_\delta$$

and $\|\varphi_k\| \leq \|\varphi\| + \|\varphi_k - \varphi\|$. Since $|p(\varphi) - p(\varphi_k)| \leq \|\varphi - \varphi_k\|$, for $m \geq N_{\varepsilon,\varphi}$, the upper and lower bounds of Lemmas 3.1 and 3.2 are true for ν_k and φ_k, with constants $K^+_{\varphi_k,\varepsilon'}(m, \beta)$ and $K^-_{\varphi_k,\varepsilon'}(m, \beta, \bar{u})$, provided that k is large enough. Since $\varepsilon' > \varepsilon$ is arbitrary, Lemmas 3.1 and 3.2 are true for ν. This is also the case for any μ in the weak-closed convex hull of such ν. By the theorem of Lanford and Robinson this set coincides with the set of equilibrium measures for φ.

3.2. *Proof of Theorem 2.12.* (1) Suppose that $\lim_{n\to\infty} \bar{z}^\beta(n)/n = 0$. Then

$$\lim_{m\to 0} \sup_{\bar{u}\in\mathcal{L}_m^\beta} \frac{z^\beta(\bar{u})}{m} = 0.$$

The estimates of Lemmas 3.1 and 3.2 prove that the equilibrium measure ν for φ is a weak Gibbs measure for $\psi = \varphi - p(\varphi)$.

(2) Suppose that $\limsup_{n\to\infty} \bar{z}^\beta(n)/n > 0$. There exist an increasing diverging sequence $\{m_k\}_k$ and $w^k \in \mathcal{L}_{m_k}^\beta$ so that $\lim_k z^\beta(w^k)/m_k = a > 0$. Let

$$\widetilde{w^k} := w^k \underbrace{0\cdots 0}_{z^\beta(w^k)}.$$

By definition $|\widetilde{w_k}| = m_k + z^\beta(w^k)$ and $\nu(I_{w^k}) = \nu(I_{\widetilde{w^k}})$. If $y \in [w^k]$, then by Lemma 3.1 (and m_k large enough),

$$\frac{\nu(I_{w^k})}{\exp(-m_k p(\varphi) + \sum_{\ell=1}^{m_k} \varphi(T^\ell y))} = \frac{\nu(I_{\widetilde{w^k}})}{\exp(-m_k p(\varphi) + \sum_{\ell=1}^{m_k} \varphi(T^\ell y))}$$

$$\le K_{\varphi,\varepsilon}^+(m_k + z^\beta(w^k), \beta) \exp\left(\sum_{j=1}^{z^\beta(w^k)} (\varphi(T^{m_k+j} y) - p(\varphi)) \right).$$

We can compare $\sum_{j=1}^{z^\beta(w^k)} \varphi(T^{m_k+j} y)$ with $z^\beta(w^k)\varphi(0)$, where 0 is the configuration with all coordinates equal to 0. For $z^\beta(w^k) \ge N_{\varepsilon,\varphi}$,

$$\left| \sum_{j=1}^{z^\beta(w^k)} (\varphi(T^{m_k+j} y)) - z^\beta(w^k)\varphi(0) \right| \le \varepsilon z^\beta(w^k),$$

so that

$$\frac{\nu(I_{w^k})}{\exp(-m_k p(\varphi) + \sum_{\ell=1}^{m_k} \varphi(T^\ell y))} \le K_{\varphi,\varepsilon}^+(m_k + z^\beta(w^k), \beta) e^{z^\beta(w^k)(\varphi(0)-p(\varphi)+\varepsilon)}.$$

For any T-invariant probability measure μ and any continuous function φ,

$$p(\varphi) \ge h_T(\mu) + \int \varphi \, d\mu,$$

where $h_T(\mu)$ is the (metric) entropy of μ. A T-invariant probability measure ν is an equilibrium measure for φ if and only if

$$p(\varphi) = h_T(\nu) + \int \varphi \, d\nu.$$

The support of a measure is the complement of the union of the open sets of measure 0. For the β-shift any equilibrium measure has support Σ^β, since by Lemma 3.2 all cylinder sets have positive measure and the cylinder sets generate the topology. Therefore the Dirac measure δ_0 cannot be an equilibrium measure. Hence

$$p(\varphi) > h_T(\delta_0) + \int \varphi \, \delta_0 = \varphi(0) \tag{25}$$

because the entropy of δ_0 is zero. Inequality (25) implies that ν is not a weak Gibbs measure for $\psi = \varphi - p(\varphi)$. Indeed, for x such that $J_{[1,m_k]}(x) = w_k$,

$$\frac{1}{m_k}\left(\ln \nu(I_{w^k}) - \sum_{j=1}^{m_k} \psi(T^j x) \right) \leq \frac{\ln K_{\varphi,\varepsilon}^+(m_k + z^\beta(w^k), \beta)}{m_k}$$
$$+ \frac{z^\beta(w^k)(\varphi(0) - p(\varphi) + \varepsilon)}{m_k}.$$

Taking the limit $k \to \infty$ and observing that ε is as small as we wish in that limit,

$$\lim_{k \to \infty} \frac{1}{m_k}\left(\ln \nu(I_{w^k}) - \sum_{j=1}^{m_k} \psi(T^j x) \right) \leq \lim_{k \to \infty} \frac{z^\beta(w^k)}{m_k} \underbrace{(6\varepsilon + \varphi(0) - p(\varphi))}_{< 0 \text{ if } \varepsilon \text{ is small enough}} < 0.$$

REFERENCES

[Bo] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)*, 2nd revised edn. Ed. J.-R. Chazottes. Springer, Berlin, 2008.

[Ca] D. Capocaccia. A definition of Gibbs state for a compact set with a $\mathbb{Z}^\nu$ action. *Comm. Math. Phys.* **48** (1976), 85–88.

[CTY] V. Climenhaga, D. J. Thompson and K. Yamamoto. Large deviations for systems with non-uniform structure. *Trans. Amer. Math. Soc.* **369** (2017), 4167–4192.

[HR] N. T. A. Haydn and D. Ruelle. Equivalence of Gibbs and equilibrium states for homeomorphisms satisfying expansiveness and specification. *Comm. Math. Phys.* **148** (1992), 155–167.

[Pa] W. Parry. On the β-expansions of real numbers. *Acta Math. Hungar.* **11** (1960), 401–416.

[PS1] C.-E. Pfister and W. G. Sullivan. Large deviations estimates for dynamical systems without the specification property. Application to the β-shifts. *Nonlinearity* **18** (2005), 237–261.

[PS2] C.-E. Pfister and W. G. Sullivan. Weak Gibbs measures and large deviations. *Nonlinearity* **31** (2018), 49–53.

[PS3] C.-E. Pfister and W. G. Sullivan. Asymptotic decoupling and weak Gibbs measures for finite alphabet shift spaces. *Nonlinearity* **33** (2020), 4799–4817.

[Re] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Hungar.* **8** (1957), 477–493.

[Ro] R. T. Rockafellar. *Convex Analysis (Princeton Mathematical Series, 28)*. Princeton University Press, Princeton, NJ, 1970.

[Ru] D. Ruelle. *Thermodynamic Formalism (Encyclopedia of Mathematics and its Applications, 5)*. Addison-Wesley, Reading, MA, 1978.

[Sc] J. Schmeling. Symbolic dynamics for β-shifts and self-normal numbers. *Ergod. Th. & Dynam. Sys.* **17** (1997), 675–694.

[Wa] P. Walters. *An Introduction to Ergodic Theory*. Springer, New York, 1982.