COMPARING LENGTH FUNCTIONS ON FREE GROUPS

Richard Sharp

University of Manchester

Abstract. The aim of this note is to study a pair of length functions on a free group, associated to a point in the unprojectivized Outer space. We discuss various characteristics which compare them and describe how these may be unified by a "Manhattan curve" analogous to those used to study pairs of hyperbolic surfaces. Our approach involves the thermodynamic formalism used in the ergodic theory of hyperbolic dynamical systems.

0. Introduction

Let F be a free group on $k \geq 2$ generators and let $\mathcal{C}(F)$ denote the set of its non-trivial conjugacy classes. A natural class of length functions on F, or, more precisely, $\mathcal{C}(F)$, may be obtained from free isometric actions on metric trees (i.e., simplicial $\mathbb{R}$-trees). More precisely, if F acts freely on the metric tree $\mathcal{T}$ then we may define a length function $l : \mathcal{C}(F) \to \mathbb{R}^+$ by

$$l(w) = \inf_{o \in \mathcal{T}} d_{\mathcal{T}}(o, ox),$$

where $x \in w$. It is easy to see that l is well-defined. Furthermore, the image of l is strictly positive. An equivalent formulation is to consider a metric graph Γ with fundamental group identified with F (and some additional non-degeneracy conditions). Then $w \in \mathcal{C}(F)$ corresponds to a unique closed path in Γ (provided backtracking is not allowed) and $l(w)$ is equal to the sum of edge lengths around this path. Such length functions are parametrized by a set $cv(F)$, which is an unprojectivized version of the Culler-Vogtmann Outer space, defined in section 1.

A natural quantity associated to $l \in cv(F)$ is the *critical exponent* $\delta = \delta(l)$. This is defined by

$$\delta = \inf \left\{ s \in \mathbb{R} : \sum_{w \in \mathcal{C}(F)} e^{-sl(w)} < +\infty \right\}.$$

It is a standard result that $\delta > 0$. We also have that

$$\delta = \limsup_{T \to +\infty} \frac{1}{T} \log \#\{w \in \mathcal{C}(F) : l(w) \leq T\}.$$

In fact, there are more precise results for counting conjugacy classes with respect to l; these are described in section 3.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

The aim of this paper is to compare two arbitrary length functions $l_1, l_2 \in cv(F)$. It is convenient to describe this in terms of the so-called Manhattan curve, an object first defined by Burger to compare two convex co-compact representations of a group as isometries of the hyperbolic space $\mathbb{H}^{n+1}$ [6]. A particularly interesting case is given by two points in the Teichmüller space associated to a compact surface [6], [36].

Here we define the Manhattan curve $\mathfrak{M}(l_1, l_2)$ associated to $l_1, l_2 \in cv(F)$ in the following way:

$$\mathfrak{M}(l_1, l_2) = \partial \left\{ (a, b) \in \mathbb{R}^2 \,:\, \sum_{w \in \mathcal{C}(F)} e^{-al_1(w) - bl_2(w)} < +\infty \right\}.$$

(This may be thought of as the natural analogue of the critical exponent for a single length function.) Many numerical characteristics which compare l_1 and l_2 may be read off from $\mathfrak{M}(l_1, l_2)$:

(i) the *intersection*

$$i(l_1, l_2) := \lim_{T \to +\infty} \frac{1}{\#\{w \in \mathcal{C}(F) \,:\, l_1(w) \leq T\}} \sum_{l_1(w) \leq T} \frac{l_2(w)}{l_1(w)};$$

(ii) the *distortion interval*

$$D(l_1, l_2) := \overline{\left\{ \frac{l_2(w)}{l_1(w)} \,:\, w \in \mathcal{C}(F) \right\}};$$

(iii) the *correlation number*

$$\alpha_1(l_1, l_2) := \lim_{\epsilon \to 0} \limsup_{T \to +\infty} \frac{1}{T} \log \# \left\{ w \in \mathcal{C}(F) \,:\, l_1(w) \leq T,\ \frac{l_2(w)}{l_1(w)} \in (1 - \epsilon, 1 + \epsilon) \right\}.$$

We shall study these in section 4 and establish the relation to $\mathfrak{M}(l_1, l_2)$ in section 5.

The earlier parts of the paper are devoted to preliminaries. In section 1, we give some background on length functions and the Outer space. In section 2, we introduce ergodic theoretic ideas in the shape of the shift of finite type Σ^+ formed by infinite reduced words in the generators of F. As we mentioned above, a length function $l \in cv(F)$ is associated (up to equivalence) to a metric graph Γ. It is a crucial point that we always use one fixed shift of finite type associated to a preferred generating set, rather than the shift formed by paths in the oriented line graph of Γ. In section 3, we show that a length function in $cv(F)$ may be encoded in a cohomology class of locally constant functions on Σ^+. So, different points in $cv(F)$ correspond to different cohomology classes of functions on the fixed shift Σ^+. As mentioned in the preceding paragraph, the main results of the paper are in sections 4 and 5. An appendix gives an introduction to the thermodynamic formalism we use.

1. Length Functions and the Outer Space

Let $\mathcal{A} = \{a_1, \ldots, a_k\}$ be a free generating set for F. Any $x \in F$ $(x \neq 1)$ may be written uniquely as a reduced word in these generators:

$$x = x_0 x_1 \ldots x_n,$$

with $x_i \in \mathcal{A} \cup \mathcal{A}^{-1}$ and $x_{i+1} \neq x_i^{-1}$, $i = 0, \ldots, n-1$, and we call $|x| = |x|_{\mathcal{A}} = n$ the word length of x (with respect to $\mathcal{A}$). We may then define a function $\|\cdot\| : \mathcal{C}(F) \to \mathbb{R}$ on conjugacy classes by

$$\|w\| = \min_{x \in w} |x|.$$

The above is only one of many length functions which may be defined and in this section we characterize the class of them that we shall study. (See [14] for a discussion of a more general class of length functions.) These correspond to isometric actions of F on simplicial trees or, equivalently, to equivalence classes of metric graphs with fundamental group isomorphic to F.

To give some background, we briefly digress to discuss automorphisms of free groups. Let $\mathrm{Aut}(F)$ denote the group of automorphisms of F. An automorphism $\alpha : F \to F$ is called *inner* if it takes the form $\alpha(x) = yxy^{-1}$, for some $y \in F$. Let $\mathrm{Inn}(F)$ denote the the set of all inner automorphisms of F; this is a normal subgroup of $\mathrm{Aut}(F)$. The quotient $\mathrm{Out}(F) = \mathrm{Aut}(F)/\mathrm{Inn}(F)$ is called the group of outer automorphisms of F. In [10], Culler and Vogtmann studied $\mathrm{Out}(F)$ by constructing a space $CV(F)$, called the Outer space, on which it acts in a natural way. (This is analogous to the study of the mapping class group of a surface via its action on Teichmüller space.)

The Outer space is defined in the following way [10], [38]. Let $\mathcal{G}$ be a fixed graph with one vertex $*$ and k edges and identify F with $\pi_1(\mathcal{G})$ so that each a_i, $i = 1, \ldots, k$, corresponds to an (oriented) edge. A metric graph is a graph together with an assignment of a positive length to each edge, making it into a metric space in the obvious way. Let Γ be a metric graph with fundamental group F such that each vertex has valency at least three together with a homotopy equivalence $g : \mathcal{G} \to \Gamma$; we call (Γ, g) a marked metric graph. Consider the set of all marked metric graphs whose edge lengths sum to one. We say that (Γ, g) and (Γ', g') are equivalent if there is an isometry $h : \Gamma \to \Gamma'$ such that $g \circ h$ is homotopic to g'. Then the Outer space $CV(F)$ is defined to be the set of equivalence classes. (An alternative definition is the set of equivalence classes of marked metric graphs under the relation $(\Gamma, g) \sim (\Gamma', g')$ if there is a homothety $h : \Gamma \to \Gamma'$ such that $g \circ h$ is homotopic to g'.)

There is a natural action of $\mathrm{Out}(F)$ on $CV(F)$ as follows. Any $\alpha \in \mathrm{Aut}(F)$ induces a map $A : \mathcal{G} \to \mathcal{G}$. Define $(\Gamma, g) \cdot \alpha = (\Gamma, g \circ A)$. Since the action of inner automorphisms is trivial, this gives a well-defined action of $\mathrm{Out}(F)$.

The universal cover of a marked metric graph Γ (with the lifted metric) is a metric tree (simplicial $\mathbb{R}$-tree) $\mathcal{T}$ on which F acts on the right by isometries. For a choice of $o \in \mathcal{T}$, define a based length function $L : F \to \mathbb{R}$ by

$$L(x) = d_{\mathcal{T}}(o, ox).$$

If we change base point then we obtain a different function but it is easy to see that L only depends the image of o under the covering map $\pi : \mathcal{T} \to \Gamma$. We may eliminate the dependence on the base point by defining a new length function l by

$$l(x) = \inf_{o \in \mathcal{T}} d_{\mathcal{T}}(o, ox).$$

We have $d_{\mathcal{T}}(o, oyxy^{-1}) = d_{\mathcal{T}}(oy, oyx)$, so it is easy to see that $l(x)$ depends only on the conjugacy class of x. Thus we think of l as a function $l : \mathcal{C}(F) \to \mathbb{R}$. Futhermore, l only depends on the point in $CV(F)$ represented by (Γ, g). (Alternatively, identifying $x \in F$ with a homotopy class in $\pi_1(\mathcal{G})$, $l(x)$ may be defined to be the length of the shortest loop in Γ freely homotopic to $g(x)$.)

Lemma 1.1 [9]**.** *The map from $CV(F)$ to the set of functions from $\mathcal{C}(F)$ to $\mathbb{R}^+$ is injective.*

Thus we may identify $CV(F)$ with a subset of $(\mathbb{R}^+)^{\mathcal{C}(F)}$. In this paper, we consider a larger space

$$cv(F) = \{\lambda l : l \in CV(F),\ \lambda > 0\}.$$

Remark. The functions $l : \mathcal{C}(F) \to \mathbb{R}$ are called hyperbolic length functions [1] or translation length functions [9]. The notation $CV(F)$ and $cv(F)$ has been used in [18] and [19], for example. Clearly, $CV(F)$ is the projectivization of $cv(F)$.

The based length function $L : F \to \mathbb{R}$ which arise in the above construction are also completely characterized by the following conditions [7], [24]. Write

$$(x, y)_L = \left(L(x) + L(y) - L(xy^{-1})\right)/2.$$

For $x, y, z \in F$,

(L1) $L(x) = 0$ if and only if $x = 1$;
(L2) $L(x^{-1}) = L(x)$;
(L3) $(x, y)_L \geq 0$;
(L4) $(x, y)_L < (x, z)_L$ implies that $(y, z)_L = (x, y)_L$; and
(L5) $(x, y)_L + (x^{-1}, y^{-1})_L > L(x) = L(y)$ implies that $x = y$.
(L6) $L(x^2) > L(x)$ for every $x \neq 1$.

(Note that $L(x) = (x, x)_L$, so (L3) implies that $L(x) > 0$ for every $x \neq 1$. A function satisfying properties (L1)-(L5) is called a Lyndon length function. (L6) is called the Archimedean property.)

Remark. The functions $l \in cv(F)$ are also characterized by a set of axioms [9], [27].

We shall use the following facts later. For $x, y \in F$ and $o, p, q \in \mathcal{T}$, define Gromov products

$$(x, y) = 1/2(|x| + |y| - |xy^{-1}|)$$

and

$$(p, q)_o = 1/2((d_{\mathcal{T}}(o, p) + d_{\mathcal{T}}(o, q) - d_{\mathcal{T}}(p, q)).$$

As above, write $L(x) = d_{\mathcal{T}}(o, ox)$. Then $(x, y)_L = (ox, oy)_o$. Furthermore, there exist constants $A_1, A_2 > 0$ such that, for all $x \in F$,

$$A_1|x| \leq d_{\mathcal{T}}(o, ox) \leq A_2|x| \tag{1.1}$$

and constants $B_1, B_2, K > 0$ such that, for all $x, y \in F$,

$$B_1(x, y) - K \leq (x, y)_L \leq B_2(x, y) + K. \tag{1.2}$$

These results follow because the map $f : F \to \mathcal{T}$ given by $f(x) = ox$ is a quasi-isometry between F equipped with the word metric $d_{\text{word}}(x, y) = |x^{-1}y|$ and $\mathcal{T}$. Then (1.1) is standard and (1.2) follows from Proposition 15 in Chapitre 5 of [12].

2. Shifts and Free Groups

In this section we shall consider the shift of finite type $\sigma : \Sigma^+ \to \Sigma^+$ associated to the free group F and free basis $\mathcal{A} = \{a_1, \dots, a_k\}$. (Shifts of finite type and their properties are discussed in greater detail in the appendix.) We define

$$\Sigma^+ = \left\{ x = (x_n)_{n=0}^{\infty} \in \prod_{n=0}^{\infty} (\mathcal{A} \cup \mathcal{A}^{-1}) \, : \, x_{n+1} \neq x_n^{-1} \ \forall n \in \mathbb{Z}^+ \right\},$$

i.e., Σ^+ is the space of infinite reduced words in $\mathcal{A} \cup \mathcal{A}^{-1}$. Of course, Σ^+ be identified with the Gromov boundary of F. The shift map $\sigma : \Sigma^+ \to \Sigma^+$ is defined by $(\sigma x)_n = x_{n+1}$.

In the notation of the appendix, $\Sigma^+ = \Sigma_A^+$, where the index set is $\mathcal{I} = \mathcal{A} \cup \mathcal{A}^{-1}$ and the zero-one $\mathcal{I} \times \mathcal{I}$ matrix A is defined by $A(i,j) = 1$ unless j is the inverse of i. It is easy to see that A is aperiodic. Furthermore, the set $W^{(n)} = W_A^{(n)}$ of allowed words of length n may be identified with the set $\{x \in F \, : \, |x| = n\}$.

A cylinder set is defined by

$$[y_0, \dots y_{n-1}]_m = \{(x_n)_{n=0}^{\infty} \in \Sigma^+ \, : \, x_{m+j} = y_j, \ j = 0, \dots, n-1\}.$$

The cylinder sets generate the topology and Borel σ-algebra on Σ^+.

A simple calculation shows that the topological entropy of $\sigma : \Sigma^+ \to \Sigma^+$ is $h(\sigma) = \log(2k-1)$ and the measure of maximal entropy is the measure μ_0 defined on cylinder sets by

$$\mu_0([x_0, x_1, \dots, x_{n-1}]_m) = \frac{1}{2k(2k-1)^{(n-1)}}$$

and extended to Borel sets by the Kolmogorov extension theorem.

Recall that $\mathcal{C}(F)$ denotes the set of non-trivial conjugacy classes in F. A conjugacy class $w \in \mathcal{C}(F)$ contains a cyclically reduced word in $\mathcal{A} \cup \mathcal{A}^{-1}$, i.e., a reduced word $x_0 x_1 \cdots x_{n-1}$ such that $x_{n-1} \neq x_0^{-1}$. The only other cyclically reduced elements of w are obtained from this by cyclic permutation (and also have word length n) and non-cyclically reduced elements of w have word length greater than n. Therefore it is natural to define the length $\|w\|$ of w (with respect to $\mathcal{A}$) to be

$$\|w\| := n = \min_{x \in w} |x|.$$

It is immediate from the definition that, for $m \geq 1$,

$$\|w^m\| = m\|w\|,$$

where w^m is the conjugacy class $\{x^m \, : \, x \in w\}$. Furthermore, it is clear from the preceding discussion that there is a natural bijection between $\mathcal{C}(F)$ and the set of periodic points of $\sigma : \Sigma^+ \to \Sigma^+$, such that, if $x, \sigma x, \dots, \sigma^{n-1} x$ ($\sigma^n x = x$) corresponds to $w \in \mathcal{C}(F)$ then $\|w\| = n$.

In order to represent elements of F as elements of a shift space (rather than simply as finite words), it is convenient to augment Σ^+ by adding an extra "dummy" symbol 0. Introduce a square matrix A_0, with rows indexed by $\mathcal{A} \cup \mathcal{A}^{-1} \cup \{0\}$, such that

$$A_0(i,j) = \begin{cases} A(i,j) & \text{if } i,j \in \mathcal{A} \cup \mathcal{A}^{-1} \\ 1 & \text{if } i \in \mathcal{A} \cup \mathcal{A}^{-1} \cup \{0\} \text{ and } j = 0 \\ 0 & \text{if } i = 0 \text{ and } j \in \mathcal{A} \cup \mathcal{A}^{-1}. \end{cases}$$

We then define $\Sigma_0^+ = \Sigma_{A_0}^+$. In other words, an element of Σ_0^+ is either an element of Σ^+ or an element of $W^{(n)}$, for some $n \geq 1$, followed by an infinite string of 0s, which we shall denote by $\dot{0}$. Of course, this procedure does not introduce any extra periodic points.

3. Locally Constant Functions

Let $l \in cv(F)$. We wish to encode the lengths $\{l(w) : w \in \mathcal{C}(F)\}$ in terms of the periodic points of $\sigma : \Sigma^+ \to \Sigma^+$ and a function $r : \Sigma^+ \to \mathbb{R}$. In fact, it will turn out that r is locally constant.

The construction of r is somewhat technical. Let $\mathcal{T}$ be a metric tree associated to l. We need to introduce an arbitrarily chosen basepoint $o \in \mathcal{T}$ and write $L(x) = d_{\mathcal{T}}(o, ox)$. The function r that we construct will depend on o only up to the addition of a coboundary.

We begin by relating the value $L(x)$ to the description of $x \in F$ as a word in the generators $\{a_1, \ldots, a_k\}$. For $n \geq 1$ and $m \leq n$, we define the following sets of n-tuples:

$$W_n^{(0)} = \{(0, 0, \ldots, 0)\},$$

$$W_n^{(m)} = \{(x_0, x_1, \ldots, x_{m-1}, 0, \ldots, 0) : (x_0, x_1, \ldots, x_{m-1}) \in W^{(m)}\}$$

and

$$W_n = \bigcup_{m=0}^{n} W_n^{(m)}.$$

The map $\iota : W_n \to \{x \in F : |x| \leq n\}$ defined by $\iota((0, 0, \ldots, 0)) = 1$ and, for $x_0 \neq 0$,

$$\iota((x_0, x_1, \ldots, x_{m-1}, 0, \ldots, 0)) = x_0 x_1 \cdots x_{m-1}$$

is a bijection. We shall abuse notation by using x to denote the sequence in W_n corresponding to $x \in F$. Clearly, ι restricts to a bijection $\iota : W_n^{(n)} = W^{(n)} \to \{x \in F : |x| = n\}$.

We begin with the following lemma, which should be compared with Proposition 4 of [30].

Lemma 3.1. *There exists an integer $N \geq 1$ such that if $n \geq N$ and $x_0 x_1 \cdots x_{n-1}$ is a reduced word then*

$$L(x_0 x_1 \cdots x_{n-1}) - L(x_1 \cdots x_{n-1}) = L(x_0 x_1 \cdots x_{N-1}) - L(x_1 \cdots x_{N-1}).$$

Proof. Note that

$$(x_0 x_1 \cdots x_{N-1}, x_0 x_1 \cdots x_{n-1}) = 1/2(N + n - (n - N)) = N.$$

Write $x = x_0x_1\cdots x_{n-1}$ and $y = x_0x_1\cdots x_{N-1}$. We need to show that $L(y) - L(x_0^{-1}y) = L(x) - L(x_0^{-1}x)$ or, equivalently

$$L(x_0) + L(y) - L(x_0^{-1}y) = L(x_0) + L(x) - L(x_0^{-1}x).$$

Using property (L2), this may be rewritten as

$$L(x_0^{-1}) + L(y^{-1}) - L(y^{-1}(x_0^{-1})^{-1}) = L(x_0^{-1}) + L(x^{-1}) - L(x^{-1}(x_0^{-1})^{-1})$$

or as

$$(y^{-1}, x_0^{-1})_L = (x^{-1}, x_0^{-1})_L.$$

By property (L4), this will hold provided

$$(x^{-1}, x_0^{-1})_L < (x^{-1}, y^{-1})_L.$$

By inequality (1.2), this will be satisfied if $N = (x^{-1}, y^{-1}) \geq B_1^{-1}(x^{-1}, x_0^{-1})_L + B_1^{-1}K$. Since $B_1^{-1}(x^{-1}, x_0^{-1})_L + B_1^{-1}K \leq B_1^{-1}B_2(x^{-1}, x_0^{-1}) + 2B_1^{-1}K$ and $(x^{-1}, x_0^{-1}) = 1$, this will be satisfied provided $N \geq B_1^{-1}(B_2 + 2K)$. $\square$

We now introduce a function on Σ_0^+, which "encodes" the values $\{L(x) : x \in F\}$.

Definition. Define a locally constant function $r : \Sigma_0^+ \to \mathbb{R}$ by

$$r((x_n)_{n=0}^{\infty}) = L(x_0 \cdots x_{N-1}) - L(x_1 \cdots x_{N-1}), \tag{3.1}$$

where N is chosen so that the conclusion of Lemma 3.1 holds.

The following result is immediate from the definition.

Lemma 3.2. *Suppose that $|x| = n$ and $x = x_0x_1\cdots x_{n-1}$. Then*

$$\begin{aligned} L(x) = r((&x_0, x_1, \ldots, x_{N-1}), (x_1, x_2, \ldots, x_N)) + \cdots \\ &+ r((x_{n-N}, x_{n-N+1}, \ldots, x_{n-1}), (x_{n-N+1}, \ldots, x_{n-1}, 0)) + \cdots \\ &+ r((x_{n-2}, x_{n-1}, 0, \ldots, 0), (x_{n-1}, 0, \ldots, 0)) + r((x_{n-1}, 0, \ldots, 0), (0, 0, \ldots, 0)). \end{aligned}$$

Proof. We have

$$\begin{aligned} L(x) = &(L(x_0x_1\cdots x_{n-1}) - L(x_1x_2\cdots x_{n-1})) \\ &+ (L((x_1x_2\cdots x_{n-1}) - L(x_2x_3\cdots x_{n-1})) + \cdots \\ &+ (L(x_{n-2}x_{n-1}) - L(x_{n-1})) + (L(x_{n-1}) - L(1)). \end{aligned}$$

The result then follows from the definition of r. $\square$

Remark. This construction has been used by Lalley [23], Bourdon [5] and Pollicott and Sharp [31],[32] in the context of convex co-compact group actions on hyperbolic space or, more generally, CAT(-1) spaces to obtain Hölder continuous functions.

We now describe how we may recover the l-lengths of conjugacy classes. For a point $x \in \Sigma_0^+$, $r^n(x) = r(x) + r(\sigma x) + \cdots + r(\sigma^{n-1}x)$.

Lemma 3.3. *Let $\sigma^n x = x$ correspond to the conjugacy class w containing the cyclically reduced word $x_0 x_1 \cdots x_{n-1}$. Then*

$$l(w) = r^n(x).$$

In particular, the cohomology class of r only depends on l.

Proof. Let $x^{(m)}$ denote the reduced word obtained from the m-fold concatenation of $x_0 x_1 \cdots x_{n-1}$. Since r is locally constant, there exists $N \geq 1$ such that

$$|r^{mn}(x) - r^{mn}(x^{(m)}, 0, 0, \dots)| \leq 2N \|r\|_\infty.$$

Noting that $r^{mn}(x) = m r^n(x)$, the above estimate gives us that

$$r^n(x) = \lim_{m \to +\infty} \frac{1}{m} r^{mn}(x^{(m)}, 0, 0, \dots) = \lim_{m \to +\infty} \frac{1}{m} L(x^{(m)}).$$

Thus it remains to show that this last quantity is equal to $l(w)$. We shall do this by proving inequalities in both directions.

First observe that $x^{(m)}$ is a cyclically reduced word in w^m. Therefore

$$m l(w) = l(w^m) \leq L(x^{(m)})$$

and so $l(w) \leq \lim_{m \to +\infty} m^{-1} L(x^{(m)}))$.

Now suppose that $v \in w$. It is clear that

$$\lim_{m \to +\infty} m^{-1} L(v^m) = \lim_{m \to +\infty} m^{-1} L(x^{(m)}).$$

On the other hand, since $\{L(v^m)\}_{m \geq 1}$ is a subadditive sequence,

$$\lim_{m \to +\infty} m^{-1} L(v^m) = \inf_{m \geq 1} m^{-1} L(v^m)$$

and so $\lim_{m \to +\infty} m^{-1} L(x^{(m)}) \leq L(v)$. Hence

$$\lim_{m \to +\infty} m^{-1} L(v^m) \leq \inf_{v \in w} L(v) = l(w).$$

The final statement follows from Proposition A.1 in the appendix, i.e., the cohomology class of r only depends on the data $r^n(x)$, where $\sigma^n x = x$. □

Lemma 3.4. *$r : \Sigma^+ \to \mathbb{R}$ is cohomologous to a strictly positive locally constant function.*

Proof. It is clear from Lemma 3.2 that that r^N is strictly positive. Futhermore, r is cohomologous to $N^{-1} r^N$. □

Let $LC(\Sigma^+)$ denote the linear space of cohomology classes $[f]$ locally constant functions $f : \Sigma^+ \to \mathbb{R}$. We say that a class $[f] \in LC(\Sigma^+)$ is positive if $f^n(x) > 0$, whenever $\sigma^n x = x$, for any (and hence all) $f \in [f]$. Let $LC^+(\Sigma^+)$ denote the set of positive classes in $LC(\Sigma^+)$. By Lemmas 3.3 and 3.4, the above construction gives us a well-defined map

$$\mathfrak{S} : cv(F) \to LC^+(\Sigma^+).$$

Lemma 3.5. *The map $\mathfrak{S} : cv(F) \to LC^+(\Sigma^+)$ is injective and, for $l \in cv(F)$, $c > 0$, $\mathfrak{S}(cl) = c\mathfrak{S}(l)$.*

Proof. If $l_1, l_2 \in cv(F)$ and $[r_1] = \mathfrak{S}(l_1) = \mathfrak{S}(l_2) = [r_2]$ then $r_1^n(x) = r_2^n(x)$, whenever $\sigma^n x = x$. Hence, by Lemma 3.3, $l_1(w) = l_2(w)$, for all $w \in \mathcal{C}(F)$, and so, by Lemma 1.1, $l_1 = l_2$.

Remark. $\mathfrak{S} : cv(F) \to LC^+(\Sigma^+)$ is not a surjection. It is shown in [9] that a function $l : F \to \mathbb{R}$ is a hyperbolic length function for an isometric action of F on an $\mathbb{R}$-tree only if it is well-defined as a function $l : \mathcal{C}(F) \to \mathbb{R}$ and

(i) for $x, y \in F$, either

$$l(xy) = l(xy^{-1}) \quad \text{or} \quad \max\{l(xy), l(xy^{-1})\} \le l(x) + l(y);$$

(ii) for $x, y \in F$ with $l(x) > 0$ and $l(y) > 0$, either

$$l(xy) = l(xy^{-1}) < l(x) + l(y) \quad \text{or} \quad \max\{l(xy), l(xy^{-1})\} = l(x) + l(y).$$

(Here the action is not assumed to be free. The conditions were shown to be sufficient in [27]) It is easy to construct elements of $LC^+(\Sigma^+)$ for which these are violated.

It follows from the above construction that there are rather precise results for counting conjugacy classes with respect to l. We say that $l \in cv(F)$ is mixing if $\{l(w) : w \in \mathcal{C}(F)\}$ is not contained in a discrete subgroup of $\mathbb{R}$. Let $\pi_l(T) = \#\{w \in \mathcal{C}(F) : l(w) \le T\}$. If l is mixing then

$$\pi_l(T) \sim \frac{e^{\delta T}}{\delta T}, \quad \text{as} \quad T \to +\infty.$$

If l is not mixing then

$$\pi_l(T) \sim \frac{e^{2\pi\delta/a}}{e^{2\pi\delta/a} - 1} \, \frac{2\pi e^{2\pi\delta[aT/2\pi]/a}}{aT}, \quad \text{as} \quad T \to +\infty,$$

where $a > 0$ is the smallest number such that $\{l(w) : w \in \mathcal{C}(F)\} \subset a\mathbb{Z}$. This may be derived from the results in [28], [29] and an alternative direct proof may be found in [13].

Remark. There is an alternative way of relating a length $l \in cv(F)$ to a shift of finite type and locally constant function (and of thus obtaining the asymptotics for $\pi_l(T)$). To describe this, let Γ be a choice of metric graph corresponding to l. Let E denote the set of *oriented* edges in Γ and let l_e denote the length of an edge $e \in E$. (Of course, $l_e = l_{\bar{e}}$, where $\bar{e}$ is the same geometric edge as e but with the reversed orientation.) The *oriented line graph* associated to Γ is the directed graph with vertices E and a directed edge (e, e') if and only if the final vertex of e coincides with the initial vertex of e', and $e' \ne \bar{e}$. We may then define a shift of finite type $\Sigma^+(\Gamma)$ to be the subset of $E^{\mathbb{Z}^+}$ formed by all (one-sided) infinite paths in the oriented line graph. A locally constant function $\omega : \Sigma^+(\Gamma) \to \mathbb{R}^+$ is then defined by $\omega((e_n)_{n=0}^{\infty}) = l_{e_0}$. Despite it simplicity, this construction has disadvantages when one wishes to compare two length functions. Then we would have two different shifts and, in particular, the periodic orbits corresponding to a given conjugacy class would generally have different periods.

4. Intersection, Distortion and Correlation

In this section we discuss a number of quantities which may be used to compare two length functions $l_1, l_2 \in cv(F)$. First we define the intersection $i(l_1, l_2)$ to be the quantity

$$i(l_1, l_2) = \lim_{T \to +\infty} \frac{1}{\pi_{l_1}(T)} \sum_{l_1(w) \le T} \frac{l_2(w)}{l_1(w)}. \tag{4.1}$$

(Note that this is not necessarily symmetric.) This is a direct analogy of the intersection of two Riemannian metris on a smooth manifold M considered in [3], [4], [6], [8], [11], [21], [41]. If M is a compact and the metrics have negative sectional curvature then the intersection is given by an expression of the form (4.1), involving the lengths of the closed geodesic in each free homotopy class with respect to the two metrics.

The intersection may be characterized directly in terms of functions $r_i \in \mathfrak{S}(l_i)$, $i = 1, 2$, or, more precisely their integrals. We write δ_1, δ_2 for the critical exponents of $l_, l_2$, respectively, and $\mu^1 = \mu_{-\delta_1 r_1}$ for the equilibrium state of $-\delta_1 r_1$. (Of course, this measure only depends on l_i or, equivalently, the cohomology class $[r_i] = \mathfrak{S}(l_i)$.)

Theorem 4.1. *For $l_1, l_2 \in cv(F)$, let $r_1, r_2 : \Sigma^+ \to \mathbb{R}$ be locally constant functions with $r_i \in \mathfrak{S}(l_i)$, $i = 1, 2$. Then the limit defining $i(l_1, l_2)$ exists and is equal to*

$$\frac{\int r_2 \, d\mu^1}{\int r_1 \, d\mu^1}.$$

Furthermore, if L_1 and L_2 are based length functions corresponding to l_1 and l_2, then

$$i(l_1, l_2) = \lim_{n \to +\infty} \frac{L_2(x_0 x_1 \cdots x_{n-1})}{L_1(x_0 x_1 \cdots x_{n-1})} \quad \textit{for } \mu^1\textit{-a.e. } (x_n)_{n=0}^{\infty} \in \Sigma^+.$$

Proof. Consider the r_1-suspended semi-flow $\sigma_t^{r_1} : \Sigma^{r_1} \to \Sigma^{r_1}$ over Σ^+. Choose $R : \Sigma^{r_1} \to \mathbb{R}$ so that $\mathfrak{I}(R) = r_2$. If the periodic $\sigma_t^{r_1}$-orbit τ corresponds to the σ-orbit $x, \sigma x, \ldots, \sigma^{n-1}x$ (and the conjugacy class w) then τ has period $\lambda(\tau) = r_1^{|w|}(x) = l_1(w)$ and we have

$$\pi_{l_1}(T) = \#\{\tau : \lambda(\tau) \le T\} \quad \text{and} \quad \int_0^{l_1(w)} R(\sigma_t(x, 0))dt = r_2^{|w|}(x) = l_2(w).$$

It is a standard result for suspended flows that

$$\lim_{T \to +\infty} \frac{1}{\#\{\tau : \lambda(\tau) \le T\}} \sum_{\lambda(\tau) \le T} \int R \, dm_\tau = \int R \, dm_0 = \frac{\int r_2 \, d\mu^1}{\int r_1 \, d\mu^1}$$

[29], where m_0 is the measure of maximal entropy for $\sigma_t^{r_1}$, giving the first part of the result.

Now consider $x = (x_n)_{n=0}^{\infty} \in \Sigma^+$. Choose the r_i, so that they are related by equation (3.1) to the based length functions L_i, $i = 1, 2$. Then

$$L_i(x_0 x_1 \cdots x_{n-1}) = r_i^n(x_0 x_1 \cdots x_{n-1}, \dot{0}).$$

It is easy to see that

$$\lim_{n\to+\infty} \frac{r_i^n(x_0x_1\cdots x_{n-1},\dot{0})}{n} = \lim_{n\to+\infty} \frac{r^n(x)}{n}$$

and, by the ergodic theorem, for μ^1-a.e. x, the limit is equal to $\int r_i d\mu^1$. Combinining these observations gives the second equality. $\square$

Remark. Kapovich [18] has defined an intersection pairing on $cv(F) \times \mathrm{Curr}(F)$, where $\mathrm{Curr}(F)$ denotes the space of geodesic currents on F. There is a natural embedding of $cv(F)$ in $\mathrm{Curr}(F)$, using the Patterson-Sullivan currents described by Kapovich and Nagnibeda in [19]. However, unlike the case of surfaces, the intersection does not extend to $\mathrm{Curr}(F) \times \mathrm{Curr}(F)$. There is a natural identication between $\mathrm{Curr}(F)$ and $\mathcal{M}_\sigma$ and it is natural to conjecture that, for $l \in cv(F)$ and $\mu \in \mathcal{M}_\sigma$,

$$I(l,\mu) = \int r \, d\mu,$$

where $r \in \mathfrak{S}(l)$, but we have been unable to prove it. If ϕ is an automorphism of F then $i(\|\cdot\|, \|\phi(\cdot)\|)$ is equal to the generic stretch of ϕ [16], [37].

Theorem 4.2. *If $\{w_m\}_{m=1}^\infty$ is a sequence in $\mathcal{C}(F)$ such that the associated periodic orbit measures μ_{w_m} converge weak* to μ_1 then*

$$i(l_1,l_2) = \lim_{m\to+\infty} \frac{l_2(w_m)}{l_1(w_m)}.$$

Proof. By Lemma 3.3, $l_i(w_m)/|w_m| = \int r_i \, d\mu_{w_m}$. Suppose that μ_{w_m} converge to μ^1. Then

$$\lim_{m\to+\infty} \int r_i \, d\mu_{w_m} = \int r_i \, d\mu^1,$$

$i = 1,2$. The result now follows from Theorem 4.1. $\square$

Theorem 4.3. *For $l_1, l_2 \in cv(F)$, we have*

$$i(l_1,l_2) \geq \frac{\delta_1}{\delta_2}$$

with equality if and only if l_2 is a constant multiple of l_1 (i.e., l_1 and l_2 define the same point in $CV(F)$.)

Proof. The inequality may be rewritten as

$$\frac{\int r_2 \, d\mu^1}{\int r_1 \, d\mu^1} \geq \frac{h(\mu^1)}{\int r_1 \, d\mu^1} \frac{\int r_2 \, d\mu^2}{h(\mu^2)}.$$

Rearranging, this becomes

$$\frac{h(\mu^2)}{\int r_2 \, d\mu^2} \geq \frac{h(\mu^1)}{\int r_2 \, d\mu^1},$$

which is true by the variational principle. Furthermore, we have equality if and only if $\mu^1 = \mu^2$, i.e., if and only if $-\delta_2 r_2$ is cohomologous to $-\delta_1 r_1 + c$, $c \in \mathbb{R}$. If this holds then, since $P(-\delta_1 r_1) = P(-\delta_2 r_2) = 0$, we must have $c = 0$, so $-\delta_2 r_2$ is cohomologous to $-\delta_1 r_1$. By Lemma 3.5, this is equivalent to $l_2 = (\delta_1/\delta_2) l_1$. $\square$

The intersection defined above quantifies the typical distortion when l_2 is compared to l_1, where "typical" is with respect to the measure $\mu_{-\delta_1 r_1}$. It is natural to generalize this notion to consider all possible distortions. As for intersection, we give a definition based to conjugacy classes and then show that is has other equivalent characterizations.

Set

$$\mathcal{D}(l_1, l_2) = \left\{ \frac{l_2(w)}{l_1(w)} \, : \, w \in \mathcal{C}(F) \right\} \quad \text{and} \quad D(l_1, l_2) = \overline{\mathcal{D}(l_1, l_2)}.$$

Proposition 4.1.

$$D(l_1, l_2) = I(r_1, r_2) := \left\{ \frac{\int r_2 \, d\mu}{\int r_1 \, d\mu} \, : \, \mu \in \mathcal{M}_\sigma \right\}.$$

In particular, if l_1 is not a constant multiple of l_2 then the closure $D(l_1, l_2)$ is a non-trivial interval.

Proof. For $w \in \mathcal{C}(F)$, let $\tau = \tau_w$ be the corresponding periodic $\sigma_t^{r_1}$. The characterization of $D(l_1, l_2)$ then follows from the identity

$$\frac{l_2(w)}{l_1(w)} = \int R \, dm_\tau = \frac{\int r_2 \, d\mu_w}{\int r_1 \, d\mu_w}$$

and the fact that the periodic orbit measures m_τ are weak* dense in $\mathcal{M}_{\sigma^{r_1}}$. If l_1 is not a constant multiple of l_2 then r_1 is not cohomologous to a constant multiple of r_2. Hence, by Proposition A.2, $I(r_1, r_2)$ is a non-trivial interval. $\square$

In [37], we considered the case where $l_1(w) = \|w\|$ and $l_2(w) = \|\phi(w)\|$, where ϕ is an automorphism of F. In this setting, Kapovich [17] showed that $\|\phi(w)\|/\|w\|$ may take any rational value in $D(l_1, l_2))$ and we studied the growth rate of the number of $w \in \mathcal{C}(F)$ (counted by $\|w\|$) for which $\|\phi(w)\|/\|w\| = \rho$, for any rational $\rho \in \text{int}(D(l_1, l_2))$. For arbitrary (non-discrete) l_1 and l_2, $\mathcal{D}(l_1, l_2)$ is hard to characterize. In particular, for any $\rho \in \mathcal{D}(l_1, l_2)$, it is not clear that $\#\{w \in \mathcal{C}(F) : l_1(w) \leq T, \ l_2(w)/l_1(w) = \rho\}$ has good asymptotic properties. Thus, in this general setting, it is more natural to study those $w \in \mathcal{C}(F)$ for which $l_2(w)/l_1(w)$ are close to some prescribed value. With this in mind, for two length functions $l_1, l_2 \in cv(F)$ and $\rho \in \text{int}(D(l_1, l_2))$, we define the ρ-correlation number $\alpha_\rho(l_1, l_2)$ by

$$\alpha_\rho(l_1, l_2) = \lim_{\epsilon \to 0} \limsup_{T \to +\infty} \frac{1}{T} \log \# \left\{ w \in \mathcal{C}(F) \, : \, l_1(w) \leq T, \ \frac{l_2(w)}{l_1(w)} \in (\rho - \epsilon, \rho + \epsilon) \right\}.$$

Note that if $\delta_1 = \delta_2$ then $1 \in \text{int}(D(l_1, l_2))$. The next result, from the thesis of T. White, shows that this also holds for lengths in $CV(F)$.

Lemma 4.2 [40]**.** *Suppose that $l_1, l_2 \in CV(F)$ (i.e., they are associated to graphs with total side length 1). There exist $w', w'' \in \mathcal{C}(F)$ such that $l_1(w') < l_2(w')$ and $l_1(w'') > l_2(w'')$. In particular $1 \in \mathrm{int}(D(l_1, l_2))$.*

In the following theorem, we shall characterize the ρ-correlation number $\alpha_\rho(l_1, l_2)$ in terms of a variational principle for entropy over a set of measures satisfying a constraint. A more precise asymptotic holds if we impose an extra condition on the lengths. We say that $l_1, l_2 \in cv(F)$ are *independent* if, for $a, b \in \mathbb{R}$,

$$\{al_1(w) + bl_2(w) : w \in \mathcal{C}(F)\} \subset \mathbb{Z} \quad \Longrightarrow \quad a = b = 0.$$

(By setting $a = 0$ or $b = 0$, this implies that l_1 and l_2 are each non-discrete.)

Theorem 4.4.

(i) *Suppose that $l_1, l_2 \in cv(F)$. Then, for every $\rho \in \mathrm{int}(D(l_1, l_2))$, $\alpha_\rho(l_1, l_2) > 0$ and satisfies $\alpha_\rho(l_1, l_2) = \mathfrak{h}_{r_1,r_2}(\rho)$, where*

$$\mathfrak{h}_{r_1,r_2}(\rho) = \sup\left\{\frac{h(\mu)}{\int r_1 \, d\mu} : \mu \in \mathcal{M}_\sigma \text{ and } \frac{\int r_2 \, d\mu}{\int r_1 \, d\mu} = \rho\right\}.$$

(ii) *Suppose that $l_1, l_2 \in cv(F)$ are independent. Then, for any $\rho \in \mathrm{int}(D(l_1, l_2))$,*

$$\#\{w \in \mathcal{C}(F) : l_1(w) \in (T, T+\epsilon), \; l_2(w) \in (\rho T, \rho T + \epsilon)\} \sim C(\rho, \epsilon) \frac{e^{\alpha_\rho(l_1,l_2)}}{T^{3/2}},$$

as $T \to +\infty$.

Proof. We apply results from the periodic theory of hyperbolic flows, which immediately carry over to suspended semi-flows over subshifts of finite type. For part (i), observe that

$$\begin{aligned}
&\#\left\{w \in \mathcal{C}(F) : l_1(w) \leq T, \; \frac{l_2(w)}{l_1(w)} \in (\rho - \epsilon, \rho + \epsilon)\right\} \\
&= \#\left\{\tau : \lambda(\tau) \leq T, \; \int R \, dm_\tau \in (\rho - \epsilon, \rho + \epsilon)\right\}.
\end{aligned}$$

By large deviations results for periodic orbits [20], this has growth rate

$$\begin{aligned}
&\lim_{T \to +\infty} \frac{1}{T} \log \# \left\{\tau : \lambda(\tau) \leq T, \; \int R \, dm_\tau \in (\rho - \epsilon, \rho + \epsilon)\right\} \\
&= \sup\left\{h(m) : m \in \mathcal{M}_{\sigma^{r_1}} \text{ and } \int R \, dm \in (\rho - \epsilon, \rho + \epsilon)\right\} \\
&= \sup\left\{\frac{h(\mu)}{\int r_1 \, d\mu} : \mu \in \mathcal{M}_\sigma \text{ and } \frac{\int r_2 \, d\mu}{\int r_1 \, d\mu} \in (\rho - \epsilon, \rho + \epsilon)\right\} \\
&= \sup\{\mathfrak{h}_{r_1,r_2}(\rho') : \rho' \in (\rho - \epsilon, \rho + \epsilon)\}.
\end{aligned}$$

Since $\mathfrak{h}_{r_1,r_2}$ is analytic, we obtain $\alpha_\rho(l_1, l_2) = \mathfrak{h}_{r_1,r_2}(\rho)$, as required. Part (ii) follows from local limit results for suspended semi-flows [2], [22], [34]. $\square$

Corollary 4.4.1. $\alpha_1(l_1, l_2) = \alpha_1(l_2, l_1)$.

Proof. From Theorem 4.4, $\alpha_1(l_1, l_2) = \mathfrak{h}_{r_1,r_2}(1)$ and $\alpha_1(l_2, l_1) = \mathfrak{h}_{r_2,r_1}(1)$. Clearly, we have

$$\begin{aligned}\mathfrak{h}_{r_1,r_2}(1) &= \sup\left\{\frac{h(\mu)}{\int r_1\, d\mu} : \mu \in \mathcal{M}_\sigma \text{ and } \int r_1\, d\mu = \int r_2\, d\mu\right\} \\ &= \sup\left\{\frac{h(\mu)}{\int r_2\, d\mu} : \mu \in \mathcal{M}_\sigma \text{ and } \int r_1\, d\mu = \int r_2\, d\mu\right\} = \mathfrak{h}_{r_2,r_1}(1),\end{aligned}$$

as required. □

If ϕ is an automorphism of F and $l_1 = \|\cdot\|$ and $l_2 = \|\phi(\cdot)\|$ then the critical exponents are both $\log(2k-1)$ and $\alpha_1(l_1, l_2)$ is related to the quantity $\mathrm{Curl}(\phi)$ [26], [37].

Remark. As noted in the appendix, $\mathfrak{h}_{r_1,r_2}$ is a continuous function on the closed interval $D(l_1, l_2) = [\rho_{\min}, \rho_{\max}]$. Typically, the value of $\mathfrak{h}_{r_1,r_2}$ is zero at the endpoints but it is easy to construct examples where $\mathfrak{h}_{r_1,r_2}(\rho_{\min}) > 0$ or $\mathfrak{h}_{r_1,r_2}(\rho_{\max}) > 0$.

Consider the graph $\mathcal{G}$, with one vertex and k edges. Make this into a metric graph by assigning lengths $0 < l_1^{(1)} < l_1^{(2)} < \cdots < l_1^{(k)}$ to the edges corresponding to the generators $a_1^{\pm 1}, a_2^{\pm 1}, \ldots, a_k^{\pm 1}$, respectively. Repeat the procedure with a diffferent set of lengths $0 < l_2^{(1)} < l_2^{(2)} < \cdots < l_2^{(k)}$ to obtain another metric graph. These give rise to two length functions $l_1, l_2 \in cv(F)$.

Suppose that we choose the edge lengths so that

$$\frac{l_2^{(1)}}{l_1^{(1)}} < \min_{2\le j\le k} \frac{l_2^{(j)}}{l_1^{(j)}}.$$

Let $r_1, r_2 : \Sigma^+ \to \mathbb{R}$ be functions in $\mathfrak{S}(l_1), \mathfrak{S}(l_2)$, respectively. It is easy to see that $\min D(l_1, l_2) = \min I_{r_1,r_2}$ is only attained for convex combinations of the periodic point measures supported on $(a_1, a_1, 0\ldots)$ and $(a_1^{-1}, a_1^{-1}, 0\ldots)$. In paticular, $\mathfrak{h}_{r_1,r_2}(\rho_{\min}) = 0$.

Now modify the lengths so that

$$\frac{l_2^{(1)}}{l_1^{(1)}} = \frac{l_2^{(2)}}{l_1^{(2)}} < \min_{3\le j\le k} \frac{l_2^{(j)}}{l_1^{(j)}}.$$

Then $\min D(l_1, l_2) = \min I_{r_1,r_2}$ is attained, in particular, for the measure of maximal entropy for the subshift of finite type obtained by restricting to the symbols $\{a_1, a_1^{-1}, a_2, a_2^{-1}\}$. A simple caluclation then shows that

$$\mathfrak{h}_{r_1,r_2}(\rho_{\min}) \ge \frac{2\log 3}{l_1^{(1)} + l_1^{(2)}} > 0.$$

Of course, a small perturbation returns us to the situation of zero entropy at the endpoints. For analogous results in the context of Hölder continuous function on subshifts of finite type, see [25].

5. The Manhattan Curve

In this section, we discuss the Manhattan curve $\mathfrak{M}(l_1, l_2)$ associated to a pair of lengths $l_1, l_2 \in cv(F)$. This is defined by

$$\mathfrak{M}(l_1, l_2) = \partial \left\{ (a, b) \in \mathbb{R}^2 : \sum_{w \in \mathcal{C}(F)} e^{-al_1(w) - bl_2(w)} < +\infty \right\}.$$

As noted in the introduction, this may be thought of as the natural analogue of the critical exponent for a single length function.

The following theorem describes how various quantities may be read off from $\mathfrak{M}(l_1, l_2)$.

Theorem 5.1. *Let $l_1, l_2 \in cv(F)$.*

- (i) *$\mathfrak{M}(l_1, l_2)$ is a straight line if and only if l_1 is a constant multiple of l_2.*
- (ii) *$\mathfrak{M}(l_1, l_2)$ is real analytic.*
- (iii) *The set of slopes of normals to $\mathfrak{M}(l_1, l_2)$ is equal to $\mathrm{int}(D(l_1, l_2))$. Furthermore, $\mathfrak{M}(l_1, l_2)$ has asymptotes whose normals have slopes equal to $\max D(l_1, l_2)$ and $\min D(l_1, l_2)$.*
- (iv) *$\mathfrak{M}(l_1, l_2)$ passes through $(\delta_1, 0)$, where its normal has slope equal to $i(l_1, l_2)$ (and through $(0, \delta_2)$, where its normal has slope equal to $1/i(l_2, l_1)$).*

If $l_1, l_2 \in CV(F)$ then

- (v) *there is a unique point $(a, b) \in \mathfrak{M}(l_1, l_2)$ where the normal has slope 1 and $a + b = \alpha_1(l_1, l_2)$.*

Proof. We modify the analysis of [36],[37]. Rewriting the definition of the Manhattan curve in terms of periodic points for $\sigma : \Sigma^+ \to \Sigma^+$, we have

$$\mathfrak{M}(l_1, l_2) = \partial \left\{ (a, b) \in \mathbb{R}^2 : \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\sigma^n x = x} e^{-ar_1^n(x) - br_2^n(x)} < +\infty \right\}.$$

By (A.2), this may be described as the set $\{(a, b) \in \mathbb{R}^2 : P(-ar_1 - br_2) = 0\}$. Let us define $\mathfrak{q}(t)$ implicitly by

$$P(-tr_1 - \mathfrak{q}(t)r_2) = 0; \tag{5.1}$$

then $\mathfrak{M}(l_1, l_2)$ is the graph of $\mathfrak{q}$. Since

$$\frac{\partial}{\partial s} P(-tr_1 - sr_2) = -\int r_2 \, d\mu_{-tr_1 - sr_2} \neq 0,$$

the Implicit Function Theorem gives that $\mathfrak{q}$ is real analytic, establishing (ii).

If we define $\mathfrak{p}(s) = \mathfrak{p}_{r_1, r_2}(s)$ (i.e. $P(-\mathfrak{p}(s)r_1 + sr_2) = 0$ or, equivalently, $\mathfrak{p}(s) = P_{\sigma^{r_1}}(sR)$) then

$$\mathfrak{q}(t) = -\mathfrak{p}^{-1}(t).$$

Now $P_{\sigma^{r_1}}(sR)$ is strictly convex and $P''_{\sigma^{r_1}}(sR) > 0$ unless R is not cohomologous to a constant, c say. This latter condition gives $\int_\tau R = c\lambda(\tau)$, for all periodic σ^{r_1}-orbits τ, i.e.,

$l_2(w) = cl_1(w)$, for all $w \in \mathcal{C}(F)$. So $\mathfrak{M}(l_1, l_2)$ is strictly convex unless l_2 is a constant multiple of l_1, proving (i).

Let us examine the slope of the normal at a point $(a, b) = (t, \mathfrak{q}(t))$ on $\mathfrak{M}(l_1, l_2)$. At this point, the normal has slope $-1/\mathfrak{q}'(t)$. Now, differentiating (5.1),

$$0 = \frac{d}{dt} P(-tr_1 - \mathfrak{q}(t)r_2) = -\int r_1 d\mu_{-tr_1-\mathfrak{q}(t)r_2} - \left(\int r_2 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2} \right) \mathfrak{q}'(t),$$

i.e.,

$$\mathfrak{q}'(t) = -\frac{\int r_1 d\mu_{-tr_1-\mathfrak{q}(t)r_2}}{\int r_2 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2}}.$$

Thus the normal to $\mathfrak{M}(l_1, l_2)$ at $(t, \mathfrak{q}(t))$ has slope

$$\frac{-1}{\mathfrak{q}'(t)} = \frac{\int r_2 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2}}{\int r_1 d\mu_{-tr_1-\mathfrak{q}(t)r_2}}.$$

Therefore, using Proposition A.3, the set of slopes of normals to $\mathfrak{M}(l_1, l_2)$ is equal to

$$\left\{ \frac{\int r_2 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2}}{\int r_1 d\mu_{-tr_1-\mathfrak{q}(t)r_2}} \; : \; t \in \mathbb{R} \right\} = \mathrm{int}(I_{r_2,r_1}) = \mathrm{int}(D(l_1, l_2)).$$

This proves the first part of (iii) and the statement about the asymptotes follows. It is clear that $\mathfrak{M}(l_1, l_2)$ passes through $(\delta_1, 0)$ at at this point the normal has slope

$$\frac{\int r_2 \; d\mu^1}{\int r_1 \; d\mu^1} = i(l_1, l_2).$$

The statement about $(0, \delta_2)$ follows easily. This proves (iv).

Suppose that $l_1, l_2 \in CV(F)$, $l_1 \neq l_2$. Then, by Lemma 4.2, $1 \in \mathrm{int}(D(l_1, l_2))$. If the normal has slope 1 at $(a, b) = (t, \mathfrak{q}(t))$ then

$$\frac{\int r_2 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2}}{\int r_1 \; d\mu_{-tr_1-\mathfrak{q}(t)r_2}} = 1$$

and

$$a + b = t + \mathfrak{q}(t) = \mathfrak{p}(-\mathfrak{q}(t)) + \mathfrak{q}(t) = \mathfrak{p}(\xi) - \xi,$$

where ξ is chosen so that $\mathfrak{p}'(\xi) = 1$. Then

$$\mathfrak{p}(\xi) - \xi = \frac{h(\mu_{-\mathfrak{p}(\xi)-\xi r_2})}{\int r_1 \; d\mu_{-\mathfrak{p}(\xi)r_1-\xi r_2}} = \mathfrak{h}_{r_1,r_2}(1)$$

and so

$$a + b = \mathfrak{h}_{r_1,r_2}(1) = \alpha_1(l_1, l_2),$$

as required. This proves (v). $\square$

Appendix: Shifts of Finite Type and Thermodynamic Formalism

We shall study length functions on free groups via a dynamical system called a subshift of finite type. In this appendix we review some background material about the ergodic theory of these systems. For basic facts about ergodic theory, see [39].

We begin with the definition of a subshift of finite type. Let A be finite matrix, indexed by a set $\mathcal{I}$, with entries zero and one. We define the shift space

$$\Sigma_A^+ = \{(x_n)_{n=0}^\infty \in \mathcal{I}^{\mathbb{Z}^+} \,:\, A(x_n, x_{n+1}) = 1 \ \forall n \in \mathbb{Z}^+\}$$

and the (one-sided) subshift of finite type $\sigma : \Sigma_A^+ \to \Sigma_A^+$ by $(\sigma x)_n = x_{n+1}$. We give $\mathcal{I}$ the discrete topology, $\mathcal{I}^{\mathbb{Z}^+}$ the product topology and Σ_A^+ the subspace topology. A compatible metric is given by

$$d((x_n)_{n=0}^\infty, (y_n)_{n=0}^\infty) = \sum_{n=0}^{\infty} \frac{1 - \delta_{x_n y_n}}{2^n},$$

where δ_{ij} is the Kronecker symbol.

We say that A is irreducible if, for each $(i,j) \in \mathcal{I}^2$, there exists $n(i,j) \geq 1$ such that $A^{n(i,j)}(i,j) > 0$ and aperiodic if there exists $n \geq 1$ such that, for each $(i,j) \in \mathcal{I}^2$, $A^n(i,j) > 0$. The latter statement is equivalent to $\sigma : \Sigma_A^+ \to \Sigma_A^+$ being topologically mixing (i.e. that there exists $n \geq 1$ such that for any two non-empty open sets $U, V \subset \Sigma_A^+$, $\sigma^{-m}(U) \cap V \neq \varnothing$, for all $m \geq n$).

If A is aperiodic then it has a positive simple eigenvalue β which is strictly maximal in modulus (i.e. every other eigenvalue has modulus strictly less that β) and the topological entropy $h(\sigma)$ of σ is equal to $\log \beta$.

If an ordered n-tuple $(x_0, x_1, \ldots, x_{n-1}) \in \mathcal{I}^n$ is such that $A(x_m, x_{m+1}) = 1$, $m = 0, 1, \ldots, n-2$ then we say that $(x_0, x_1, \ldots, x_{n-1})$ is an allowed word of length n in Σ_A^+; the set of these is denoted $W_A^{(n)}$. If $\sigma^n x = x$ then we say that $\{x, \sigma x, \ldots, \sigma^{n-1} x\}$ is a periodic orbit for σ. Clearly any such an x is obtained by repeating a word $(x_0, x_1, \ldots, x_{n-1}) \in W_A^{(n)}$ with the additional property that $A(x_{n-1}, x_0) = 1$. Note that we regard the periodic orbits $\{x, \sigma x, \ldots, \sigma^{n-1} x\}$, $\{x, \sigma x, \ldots, \sigma^{n-1} x\}, x, \ldots, \sigma^{n-1} x$, etc., as distinct objects (even though they are identical as point sets). If $\sigma^n x = x$ but $\sigma^m x \neq x$ for $0 < m < n$ then we say that $\{x, \sigma x, \ldots, \sigma^{n-1} x\}$ is a prime periodic orbit.

A function $f : \Sigma_A^+ \to \mathbb{R}$ is Hölder continuous if there exists $\alpha > 0$ and $C(f, \alpha) \geq 0$ such that $|f(x) - f(y)| \leq C(f, \alpha) d(x,y)^\alpha$, for all $x, y \in \Sigma_A^+$. We say that a function $f : \Sigma_A^+ \to \mathbb{R}$ is locally constant if there exists $N \geq 0$ such that if $x = (x_n)_{n=0}^\infty, y = (x_n)_{n=0}^\infty$ have $x_n = y_n$ for all $n \geq N$ then $f(x) = f(y)$. Clearly, if f is locally constant then f is Hölder continuous (for any choice of exponent $\alpha > 0$).

Let $\mathcal{M}$ denote the space of all Borel probability measures on Σ_A^+, equipped with the weak* topology, and let $\mathcal{M}_\sigma$ denote the subspace consisting of σ-invariant probability measures. For $\mu \in \mathcal{M}_\sigma^1$, write $h(\mu)$ for the measure theoretic entropy of μ. There is a unique measure $\mu_0 \in \mathcal{M}_\sigma^1$, called the measure of maximal entropy, for which

$$h(\mu_0) = \sup_{\mu \in \mathcal{M}_\sigma^1} h(\mu)$$

and this value coincides with the topological entropy $h(\sigma)$.

For a continuous function $f : \Sigma_A^+ \to \mathbb{R}$, we define the pressure $P(f)$ of f by the formula

$$P(f) = \sup_{\mu \in \mathcal{M}_\sigma} \left(h(\mu) + \int f \, d\mu \right) \tag{A.1}$$

and call any measure for which the supremum is attained an equilibrium state for f. If f is Hölder continuous then f has a unique equilibrium state which we denote by μ_f. The latter is fully supported and $h(\mu_f) > 0$. The equilibrium state of the zero function is the measure of maximal entropy, so this is consistent with our earlier notation. The pressure of f also has the following characterization in terms of periodic points:

$$P(f) = \lim_{n \to +\infty} \frac{1}{n} \log \sum_{\sigma^n x = x} e^{f^n(x)}, \tag{A.2}$$

where $f^n = f + f \circ \sigma + \cdots + f \circ \sigma^{n-1}$.

We say that two functions $f, g : \Sigma_A^+ \to \mathbb{R}$ are (continuously) cohomologous if there is a continuous function $u : \Sigma_A^+ \to \mathbb{R}$ such that $f = g + u \circ \sigma - u$. The cohomology class of a Hölder continuous function is determined by its values around periodic orbits. More precisely, we have the following.

Proposition A.1 (Livsic Theorem). *Two Hölder continuous functions $f, g : \Sigma_A^+ \to \mathbb{R}$ are cohomologous if and only if $f^n(x) = g^n(x)$ whenever $\sigma^n x = x$.*

If f and g are cohomologous then $P(f) = P(g)$ and if c is a real number then $P(f+c) = P(f) + c$. Now suppose that f and g are Hölder continuous and, for $t \in \mathbb{R}$, consider the function $t \mapsto P(tf)$ This function is convex and real analytic and

$$P'(tf + g) = \int f \, d\mu_{tf+g}. \tag{A.3}$$

Furthermore, if f is not cohomologous to a constant then $P(tf + g)$ is strictly convex and $P''(tf + g) > 0$ everywhere. (If f is cohomologous to a constant c then clearly $P(tf + g) = P(g) + tc$.)

Suppose that a Hölder continuous function $r : \Sigma_A^+ \to \mathbb{R}$ is cohomologous to a strictly positive function. Since we are only interested in sums over periodic orbits and integrals with respect to invariant measures, we shall, in fact, assume that r itself is strictly positive. If $\delta > 0$ satisfies $P(-\delta r) = 0$ then, since $\mu_{-\delta r}$ attains the supremum in (A.1), we have the relation

$$\delta = \frac{h(\mu_{-\delta r})}{\int r \, d\mu_{-\delta r}}. \tag{A.4}$$

More generally, given another Hölder continuous function $r : \Sigma_A^+ \to \mathbb{R}$, we define $\mathfrak{p}_{r,f}(s)$ by $P(-\mathfrak{p}_{r,f}(s) r + sf) = 0$. Note that, by (A.3),

$$P'(-tr + f) = \int r \, d\mu_{-tr+f} > 0.$$

Therefore, by the Implicit Function Theorem, $\mathfrak{p}_{r,f}(s)$ is real analytic and

$$\mathfrak{p}'_{r,f}(s) = \frac{\int f \, d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}}{\int r \, d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}}. \tag{A.5}$$

We write

$$I_{f,r} := \left\{ \frac{\int f d\mu}{\int r d\mu} \, : \, \mu \in \mathcal{M}_\sigma \right\}.$$

Since $\int r d\mu > 0$, for any $\mu \in \mathcal{M}_\sigma$, $\mu \mapsto \int f d\mu / \int r d\mu$ is continuous with respect to the weak* topology on $\mathcal{M}_\sigma$. So, as $\mathcal{M}_\sigma$ is compact and convex, $I_{f,r}$ is a closed interval.

Proposition A.2.

(i) *We have $I_{r,f} = \{c\}$ if and only if f is cohomologous to cr.*

(ii) *If $\mathrm{int}(I_{r,f}) \neq \varnothing$ and $\int f d\mu / \int r d\mu$ is an endpoint of $I_{r,f}$ then μ does not have full support.*

(iii) *If $\mathrm{int}(I_{r,f}) \neq \varnothing$ then*

$$\mathrm{int}(I_{r,f}) = \{\mathfrak{p}'_{r,f}(s) \, : \, s \in \mathbb{R}\} = \left\{ \frac{\int f \, d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}}{\int r \, d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}} \, : \, s \in \mathbb{R} \right\}.$$

(iv) *We have $\mathrm{int}(I_{r,f}) \neq \varnothing$ if and only if $\mathfrak{p}''_{r,f}(s) > 0$ for some (or, equivalently, all) $s \in \mathbb{R}$.*

Proof. (i) The statement $I_{r,f} = \{c\}$ is equivalent to

$$\int f \, d\mu - \int cr \, d\mu = 0,$$

for all $\mu \in \mathcal{M}_\sigma$. In particular, this is equivalent to $f^n(x) = cr^n(x)$, whenever $\sigma^n x = x$. By Proposition A.1, this is equivalent to f and cr being cohomologous.

(ii) Suppose that $I_{r,f} = [a, b]$ and that $\int f d\mu / \int r d\mu = b$. Then μ is a *maximizing measure* for $f - br$. Then $\mathrm{supp}\,\mu$ is contained in a closed subset of Σ_A^+ on which $f - br$ is cohomologous to 0 [15, Lemma 2]. By (i), this closed set cannot be Σ_A^+ itself.

(iii) Since $\mathfrak{p}_{r,f}$ is strictly convex, $\{\mathfrak{p}'_{r,f}(s) \, : \, s \in \mathbb{R}\}$ is an open interval. The following argument is taken from [15]. From the definition of $\mathfrak{p}_{r,f}(s)$, for any $\mu \in \mathcal{M}_\sigma$,

$$\mathfrak{p}_{r,f}(s) \geq \frac{h(\mu)}{\int r \, d\mu} + s \frac{\int f \, d\mu}{\int r \, d\mu},$$

for all $s \in \mathbb{R}$. In particular, the graph of the convex function $\mathfrak{p}_{r,f}$ lies above a line with slope $\int f d\mu / \int r d\mu$ in $\mathbb{R}^2$ (possibly touching it tangentially) and so $\int f d\mu / \int r d\mu \in \overline{\{\mathfrak{p}'_{r,f}(s) \, : \, s \in \mathbb{R}\}}$. Thus, since μ is arbitrary,

$$I_{r,f} \subset \overline{\{\mathfrak{p}'_{r,f}(s) \, : \, s \in \mathbb{R}\}}.$$

By (A.5), $\{\mathfrak{p}'_{r,f}(s) : s \in \mathbb{R}\} \subset I_{r,f}$. Therefore, $\mathrm{int}(I_{r,f}) = \{\mathfrak{p}'_{r,f}(s) : s \in \mathbb{R}\}$, as required.

(iv) If $\mathfrak{p}''_{r,f}(s) > 0$ for some $s \in \mathbb{R}$ then the graph of $\mathfrak{p}_{r,f}$ cannot be a straight line, so $\mathrm{int}(I_{r,f}) \neq \varnothing$. On the other hand, suppose that $\mathrm{int}(I_{r,f}) \neq \varnothing$ and choose $\xi \in \mathbb{R}$. Write $\rho = \mathfrak{p}'_{r,f}(\xi)$. Note that $\mathfrak{p}_{r,f-\rho r}(s) = \mathfrak{p}_{r,f}(s) - s\rho$, so $\mathfrak{p}'_{r,f-\rho r}(\xi) = 0$. Differentiating the identity $P(-\mathfrak{p}_{r,f-\rho r}(s)r + s(f-\rho r)) = 0$ twice at $s = \xi$ gives

$$\mathfrak{p}''_{r,f}(\xi) = \mathfrak{p}''_{r,f-\rho r}(\xi) = \left(\int r \, d\mu_{\mathfrak{p}_{r,f-\rho r}(\xi)r+\xi(f-\rho r)}\right)^{-1} \frac{\partial^2 P(-\mathfrak{p}_{r,f-\rho r}(\xi)r + s(f-\rho r))}{\partial s^2},$$

which is positive unless $f - \rho r$ is cohomologous to a constant. By the choice of ρ, this constant must necessarily be zero, so $\mathfrak{p}''_{r,f}(\xi) > 0$ unless f is cohomologous to ρr, which contradicts the assumption that $\mathrm{int}(I_{r,f}) \neq \varnothing$. $\square$

We also define a weighted entropy function $\mathfrak{h}_{r,f} : I_{r,f} \to \mathbb{R}^+$ by

$$\mathfrak{h}_{r,f}(\rho) = \sup\left\{\frac{h(\mu)}{\int r \, d\mu} : \mu \in \mathcal{M}_\sigma \text{ and } \int f \, d\mu = \rho\right\}. \tag{A.6}$$

From its definition,

$$\begin{aligned}
\mathfrak{p}_{r,f}(s) &= \sup\left\{\frac{h(\mu)}{\int r \, d\mu} + s\frac{\int f \, d\mu}{\int r \, d\mu} : \mu \in \mathcal{M}_\sigma\right\} \\
&= \sup_{\rho \in I_{r,f}} \sup\left\{\frac{h(\mu)}{\int r \, d\mu} + s\frac{\int f \, d\mu}{\int r \, d\mu} : \mu \in \mathcal{M}_\sigma, \ \frac{\int f \, d\mu}{\int r \, d\mu} = \rho\right\} \\
&= \sup_{\rho \in I_{r,f}} \left(\sup\left\{\frac{h(\mu)}{\int r \, d\mu} : \mu \in \mathcal{M}_\sigma, \ \frac{\int f \, d\mu}{\int r \, d\mu} = \rho\right\} + s\rho\right) \\
&= \sup_{\rho \in I_{r,f}} (\mathfrak{h}_{r,f}(\rho) + s\rho).
\end{aligned}$$

In the language of convex analysis, $-\mathfrak{h}_{r,f}$ is the convex conjugate of $\mathfrak{p}_{r,f}$ [33,p.104]. Assuming that f is not cohomologous to a constant multiple of r, $\mathfrak{p}_{r,f}$ is strictly convex, so $\mathfrak{h}_{r,f}$ is strictly concave. Since the entropy map $\mu \mapsto h(\mu)$ is upper semi-continuous, this implies that $\mathfrak{h}_{r,f} : I_{r,f} \to \mathbb{R}$ is continuous. The next result follows from the theory of convex functions [33].

Proposition A.3. *Suppose that f is not cohomologous to a constant multiple of r. If, for $\rho \in \mathrm{int}(I_{r,f})$, $\xi \in \mathbb{R}$ is the unique solution to $\mathfrak{p}'(\xi) = \rho$ then $\mathfrak{p}_{r,f}(\xi) = \mathfrak{h}_{r,f}(\rho) + \xi\rho$, so that*

$$\mathfrak{h}_{r,f}(\rho) = \frac{h(\mu_{-\mathfrak{p}_{r,f}(\xi)r+\xi f})}{\int r \, d\mu_{-\mathfrak{p}_{r,f}(\xi)r+\xi f}}$$

and $\mu_{-\mathfrak{p}_{r,f}(\xi)r+\xi f}$ is the only element of $\mathcal{M}_\sigma$ realizing the supremum in (A.6). Furthermore, $\mathfrak{h}_{r,f} : \mathrm{int}(I_{r,f}) \to \mathbb{R}$ is real analytic.

In the remainder of this appendix, we shall reinterpret $I_{r,f}$, $\mathfrak{p}_{r,f}$ and $\mathfrak{h}_{r,f}$ in terms of a suspended semi-flow over Σ_A^+. The reason for this is that it will allow us to quote a number of results from the theory of suspended and hyperbolic flows (which also hold automatically for semi-flows.)

We define the r-suspension space

$$\Sigma_A^r = \{(x,s) : x \in \Sigma_A^+,\ 0 \le s \le r(x)\}/\sim,$$

where $(x, r(x)) \sim (\sigma x, 0)$. (It was for this construction to make sense that we took r to be strictly positive, rather than merely cohomologous to a strictly positive function.) This supports the suspended semi-flow $\sigma_t^r : \Sigma_A^r \to \Sigma_A^r$ defined (for small $t > 0$) by $\sigma_t^r(x,s) = (x, s+t)$ and respecting the identifications. A more familiar object of study is the suspended flow over the corresponding two-sided shift space $\Sigma_A = \{(x_n)_{n=-\infty}^{\infty} \in \mathcal{I}^{\mathbb{Z}} : A(x_n, x_{n+1}) = 1\ \forall n \in \mathbb{Z}\}$. Such flows and their periodic orbits are well understood [22][29],[35] and it is easy to see that the same results hold for semi-flows.

We let $\mathcal{M}_{\sigma^r}$ denote the space of σ_t^r-invariant probability measures on Σ_A^r. Every $m \in \mathcal{M}_{\sigma^r}$ may be written locally as

$$m = \frac{\mu \times \mathrm{Leb}}{\int r\ d\mu}, \tag{A.7}$$

where $\mu \in \mathcal{M}_\sigma$ and Leb is one dimensional Lebesgue measure, and has entropy

$$h_{\sigma^r}(m) = \frac{h(\mu)}{\int r\ d\mu}. \tag{A.8}$$

Given a Hölder continuous function $F : \Sigma_A^r \to \mathbb{R}$, we may associate a Hölder continuous function $f = \mathfrak{I}(F) : \Sigma_A^+ \to \mathbb{R}$ by

$$\mathfrak{I}(F) := \int_0^{r(x)} F(x,t)\ dt.$$

One easily sees that if m and μ are related by (A.7) then

$$\int F dm = \frac{\int f\ d\mu}{\int r\ d\mu}.$$

Given a Hölder continuous function $f : \Sigma_A^+ \to \mathbb{R}$, it is easy to construct a Hölder continuous function $F : \Sigma_A^r \to \mathbb{R}$ such that $\mathfrak{I}(F) = f$. For example, choose a smooth function $\Delta : [0,1] \to \mathbb{R}^+$ with $\Delta(0) = \Delta(1) = 0$ and $\int \Delta(y) dy = 1$. Then set

$$F(x,s) = \Delta\left(\frac{s}{r(x)}\right) \frac{f(x)}{r(x)}, \quad 0 \le s \le r(x).$$

(The function Δ is introduced so that F respects the equivalence relation $(x, r(x)) \sim (\sigma x, 0)$.) In particular,

$$I_{r,f} = \left\{ \int F dm : m \in \mathcal{M}_{\sigma^r} \right\}.$$

We define the pressure $P_{\sigma^r}(F)$ of a function $F : \Sigma_A^r \to \mathbb{R}$ by the formula

$$P_{\sigma^r}(F) = \sup_{m \in \mathcal{M}_{\sigma^r}} \left(h(m) + \int F\ dm \right)$$

and call any measure for which the supremum is attained an equilibrium state for F. If F is Hölder continuous then F has a unique equilibrium state which we denote by m_F.

Proposition A.4. *We have $P(-P_{\sigma^r}(F)r+f)=0$, where $f=\mathfrak{I}(F)$, and, locally,*

$$m_F = \frac{\mu_{-P(F)r+f}\times \mathrm{Leb}}{\int r\ d\mu_{-P(F)r+f}}.$$

Proof. The identity $P(-cr+f)=0$ is equivalent to

$$\sup_{\mu\in\mathcal{M}_\sigma}\left(h(\mu)+\int(-cr+f)\ d\mu\right)=0,$$

which may be rewritten as

$$\sup_{\mu\in\mathcal{M}_\sigma}\left(\frac{h(\mu)}{\int r\ d\mu}+\frac{\int f\ d\mu}{\int r\ d\mu}\right)=c.$$

By (A.7) and (A.8), this becomes

$$c=\sup_{m\in\mathcal{M}_{\sigma^r}}\left(h_{\sigma^r}(m)+\int F\ dm\right)=P_{\sigma^r}(F).$$

This calculations also give the formula for m_F. □

Remark. In particular, the equilibrium state of the zero function m_0 is the measure of maximal entropy for σ_t^r and, locally,

$$m_0=\frac{\mu_{-h(\sigma)r}\times \mathrm{Leb}}{\int r\ d\mu_{-h(\sigma)r}}.$$

Applying the lemma to sF, gives $P_{\sigma^r}(sF)=\mathfrak{p}_{r,f}(s)$. Hence, by (A.5),

$$P'_{\sigma^r}(sF)=\mathfrak{p}'_{r,f}(s)=\frac{\int f\ d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}}{\int r\ d\mu_{-\mathfrak{p}_{r,f}(s)r+sf}}=\int F\ dm_{sF}. \tag{A.9}$$

There is clearly an exact correspondence between periodic orbits for σ_t^r and σ. If τ is the prime periodic σ^r-orbit passing through the prime periodic σ-orbit $x,\sigma x,\dots,\sigma^{n-1}x$ ($\sigma^n x = x$) then the period of τ is

$$\lambda(\tau):=\inf\{t>0\,:\,\sigma_t^r(x,0)=(x,0)\}=r^n(x).$$

Furthermore, if we let m_τ denote the σ^r-invariant probability measure supported on τ then, for $F:\Sigma_A^r\to\mathbb{R}$,

$$\int F\ dm_\tau:=\frac{1}{\lambda(\tau)}\int_0^{\lambda(\tau)}F(\sigma_t^r(x,0))\ dt=\frac{f^n(x)}{r^n(x)}, \tag{A.10}$$

where $f=\mathfrak{I}(F)$. Since the periodic orbit measures m_τ are weak* dense in $\mathcal{M}_{\sigma^r}$, we have

$$I_{f,r}=\overline{\left\{\int F dm_\tau\,:\,\tau \text{ periodic } \sigma^r\text{-orbit}\right\}}=\overline{\left\{\frac{f^n(x)}{r^n(x)}\,:\,\sigma^n x=x\right\}}. \tag{A.11}$$

References

1. R. Alperin and H. Bass, *Length functions of group actions on Λ-trees*, Math. Ann. (1987), Princeton Univ. Press, Princeton, NJ,, 265–378.
2. M. Babillot and F. Ledrappier, *Lalley's theorem on periodic orbits of hyperbolic flows*, Ergodic Theory Dyn. Syst. **18** (1998), 17-39.
3. F. Bonahon, *Bouts de variétés hyperboliques de dimension 3*, Ann. of Math. **124** (1986), 71-158.
4. F. Bonahon, *The geometry of Teichmüller space via geodesic currents*, Invent. math. **92** (1988), 139-162.
5. M. Bourdon, *Actions quasi-convexes d'un groupe hyperbolique, flot géodésique*, PhD thesis (Orsay) 1993.
6. M. Burger, *Intersection, the Manhattan curve, and the Patterson-Sullivan theory in rank 2*, Int. Math. Res. Notices **7** (1993), 217-225.
7. I. Chiswell, *Abstract length functions in groups*, Math. Proc. Cambridge Philos. Soc. **80** (1976), 451-463.
8. C. Croke and A. Fathi, *An inequality between energy and intersection*, Bull London Math. Soc. **22** (1990), 489-494.
9. M. Culler and J. Morgan, *Group actions on $\mathbb{R}$-trees*, Proc. London Math. Soc. **55** (1987), 571-604.
10. M. Culler and K. Vogtmann, *Moduli of graphs and automorphisms of free groups*, Invent. Math. **84** (1986), 91-119.
11. A. Fathi and L. Flaminio, *Infinitesimal conjugacies and Weil-Petersson metric*, Ann. Inst. Fourier (Grenoble) **43** (1993), 279-299.
12. E. Ghys and P. de la Harpe, *Sur les groupes hyperboliques d'après Mikhael Gromov*, Birkhauser., Boston, 1990.
13. L. Guillopé, *Entropies et spectres*, Osaka J. Math. **31** (1994), 247-289.
14. M. Horsham and R. Sharp, *Lengths, quasi-morphisms and statistics for free groups*, to appear, "Spectral Analysis in Geometry and Number Theory", Contemporary Mathematics.
15. O. Jenkinson, *Rotation, entropy, and equilibrium states*, Trans. Amer. Math. Soc. **353** (2001), 3713-3739.
16. V. Kaimanovich, I. Kapovich and P. Schupp, *The subadditive ergodic theorem and generic stretching factors for free group automorphisms*, Israel J. Math. **157** (2007), 1-46.
17. I. Kapovich, *The frequency space of a free group*, Internat. J. Algebra Comput. **15** (2005), 939-969.
18. I. Kapovich, *Currents on free groups*, Topological and asymptotic aspects of group theory, Contemp. Math., 394, Amer. Math. Soc., Providence, RI, 2006, pp. 149-176.
19. I. Kapovich and T. Nagnibeda, *The Patterson-Sullivan embedding and minimal volume entropy for outer space*, Geom. Funct. Anal. **17** (2007), 1201-1236.
20. Y. Kifer, *Large deviations, averaging and periodic orbits of dynamical systems*, Comm. Math. Phys. **162** (1994), 33-46.
21. G. Knieper, *Volume growth, entropy and the geodesic stretch*, Math. Res. Lett. **2** (1995), 39-58.
22. S. Lalley, *Distribution of periodic orbits of symbolic and Axiom A flows*, Adv. Appl. Math. **8** (1987), 154-193.
23. S. Lalley, *Renewal theorems in symbolic dynamics, with applications to geodesic flows, noneuclidean tessellations and their fractal limits*, Acta Math. **163** (1989), 1-55.
24. R. Lyndon, *Length functions in groups*, Math. Scand. **12** (1963), 209-234.
25. I. Morris, *Maximizing measures of generic Hölder functions have zero entropy*, Nonlinearity **21** (2008), 993-1000.
26. A. G. Myasnikov and V. Shpilrain, *Some metric properties of automorphisms of groups*, J. Algebra **304** (2006), 782-792.
27. Walter Parry, *Axioms for translation length functions*, Arboreal group theory (Berkeley, CA, 1988), Math. Sci. Res. Inst. Publ., 19, Springer, New York, 1991, pp. 295-330.
28. William Parry, *An analogue of the prime number theorem for closed orbits of shifts of finite type and their suspensions*, Israel J. Math. **45** (1983), 41-52.
29. William Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque **187-188** (1990), 1-268.

30. M. Pollicott and R. Sharp, *Poincaré series and zeta functions for surface group actions on $\mathbb{R}$-trees*, Math. Zeit. **226** (1997), 335-347.
31. M. Pollicott and R. Sharp, *Comparison theorems and orbit counting in hyperbolic geometry*, Trans. Amer. Math. Soc. **350** (1998), 473-499.
32. M. Pollicott and R. Sharp, *Poincaré series and comparison theorems for variable negative curvature*, Topology, ergodic theory, real algebraic geometry, Amer. Math. Soc. Transl. Ser. 2, 202, Amer. Math. Soc., Providence, RI, 2001, pp. 229-240.
33. R. Rockafeller, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
34. R. Schwartz and R. Sharp, *The correlation of length spectra of two hyperbolic surfaces*, Comm. Math. Phys. **153** (1993), 423-430.
35. R. Sharp, *Prime orbit theorems with multi-dimensional constraints for Axiom A flows*, Monat. Math. **114** (1992), 261-304.
36. R. Sharp, *The Manhattan curve and the correlation of length spectra on hyperbolic surfaces*, Math. Zeit. **228** (1998), 745-750.
37. R. Sharp, *Distortion and entropy for automorphisms of free groups*, to appear, Discrete Contin. Dyn. Syst..
38. K. Vogtmann, *Automorphisms of free groups and outer space. Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part I (Haifa, 2000)*, Geom. Dedicata **94** (2002), 1-31.
39. P. Walters, *An introduction to ergodic theory*, Springer-Verlag, Berlin, 1982.
40. T. White, *The Geometry of the Outer Space*, PhD Thesis, UCLA (1991).
41. S. Wolpert, *Thurston's Riemannian metric for Teichmller space*, J. Differential Geom. **23** (1986), 143-174.

School of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, U.K.