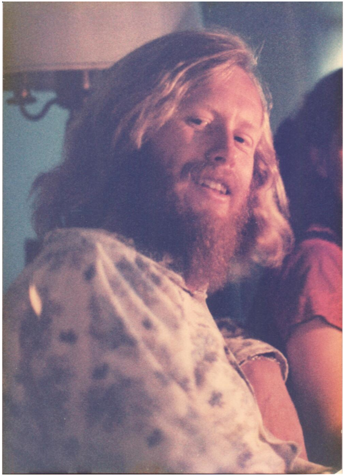
The Bowen-Series coding

Caroline Series

THE UNIVERSITY OF WARWICK



Rufus Bowen: 1947 - 1978

The set-up

• A Fuchsian group G ($=$ discrete group of hyperbolic isometries) acting in the hyperbolic disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

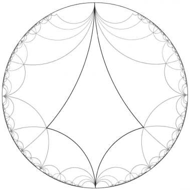
• A fundamental domain $\mathcal{R}$ and tessellation $\mathcal{T}$ of $\mathbb{D}$.

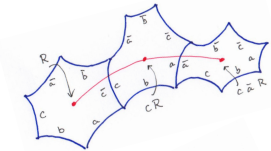
• A set of generators G_0 of G given by the elements of G which pair the sides of $\mathcal{R}$. This labels the sides of $\mathcal{R}$.

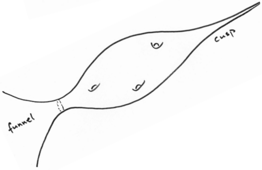
• A path from $\mathcal{R}$ to $g\mathcal{R}$ across $\mathcal{T}$ can be recorded by the sequence of generators labelling the sides which it cuts. This gives an expression for g as a product of generators. In particular, a circuit round a vertex gives a relation in G.

• The quotient $S = \mathbb{D}/G$ is a surface of constant negative curvature. It is compact; finite area; infinite area as $\overline{\mathcal{R}}$ is: (a) inside $\mathbb{D}$; (b) has vertices on $\mathbb{D}$; (c) has intervals on $\mathbb{D}$.

• The limit set $\Lambda \subset \partial\mathbb{D} = S^1$ is the set of accumulation points of G orbits, aka copies of $\mathcal{R}$.







The Bowen-Series map

For simplicity we will state the results for the case when $S = \mathbb{D}/G$ is compact; they extend with minor tweaking to all cases.

A map $f : S^1 \to S^1$ is called *Markov* if S^1 is partitioned into a finite union of intervals I_j, disjoint except at their endpoints, such that $f(I_j) = \cup_{i=1}^{k} I_{j_i}$.

THEOREM [BOWEN-S. 1979] There exists a Markov map $f : S^1 \to S^1$ such that the restriction of f to I_j equals some element $e \in G_0$. Moreover f is eventually expanding and closely related to the G action on S^1.

If $\xi \in S^1$ we define its f-expansion as follows: For $e \in G_0$ let $J_e = \cup\{I_{i_j} : f_{|I_{i_j}} = e^{-1}\}$. Then $\xi \sim (e_{i_1} e_{i_2} \dots e_{i_k} \dots)$ if $\xi \in \cap_{k=0}^{\infty} f^{-k}(J_{i_k})$. Let Σ_G^+ denote the set of f-expansions which occur. Let $\mathcal{I}$ be the set of intervals and let $\mathbb{A}$ be the $\mathcal{I} \times \mathcal{I}$ transition matrix (a_{ij}) with $a_{ij} = 1$ if $f(I_i) \supset I_j$ and 0 otherwise. Let $\Sigma_{\mathbb{A}}^+$ be the associated finite type shift space.

COROLLARY The maps $\Sigma_{\mathbb{A}}^+ \to \Sigma_G^+$ and $\Sigma_G^+ \to S^1$ are both well defined and $1-1$ except on a countable set of points and conjugate the shifts on $\Sigma_{\mathbb{A}}^+, \Sigma_G^+$ with the map f on S^1.

First consequences

The maps $\Sigma_{\mathbb{A}}^+ \to \Sigma_G^+ \to S^1$ allow us to more or less identify the f-action on S^1 as a subshift of finite type to which we can apply the Bowen theory of Gibbs states and thermodynamic formalism.

Remark: If $\mathbb{D}/G$ has infinite volume then the image of the infinite admissible words is the limit set Λ of G.

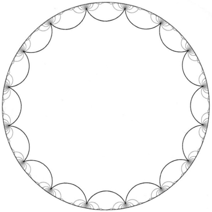
The connection to the G action is very important:
PROPOSITION The actions of G and f on S^1 are *orbit equivalent* :
$\xi = g\eta, g \in G$ iff their f-expansions have the same tails. Moreover each element of G has a unique shortest representation as a word w which occurs as an admissible sequence in Σ_G^+.

Thus for example, there is a bijection between purely periodic sequences in $\Sigma_{\mathbb{A}}^+$ and cyclically shortest 'cyclic words', aka conjugacy classes, in G.

Origins

Any surface of constant negative curvature can be represented as $\mathbb{D}/G$ for some Fuchsian group G. The geodesic flow on such surfaces was proved to be ergodic in a famous theorem of E.Hopf (1936). However in a very special symmetrical case, G. Hedlund (1934) had already given a completely different proof based on a *boundary coding* invented by J. Nielsen (1927).

Nielsen's boundary coding is defined relative to a group G whose fundamental domain is a regular $4g$-gon. Any genus g surface carries a hyperbolic metric obtained in this way. The coding expands points on S^1 as infinite sequences of generators of G. Rufus and I observed that the estimates used by Hedlund look very like those of Gibbs state theory in Rufus' Springer volume *Equilibrium States*. Another example is the classical result that $\xi, \eta \in \mathbb{R}$ are equivalent under the action of $GL(2, \mathbb{Z})$ if and only if their continued fraction expansions have the same tails.

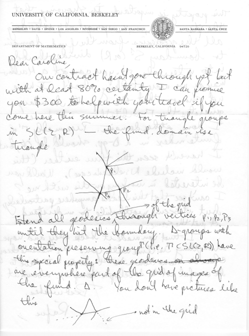


Gluing opposite sides of a 16-gon gives a genus 4 surface.

Thus we conceived the idea of trying to generalise the coding to arbitrary Fuchsian groups.

Communication

I started discussing this question with Rufus sometime in 1977. I was in Cambridge, and in those days people communicated by letter.



UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO   SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF MATHEMATICS   BERKELEY, CALIFORNIA 94720

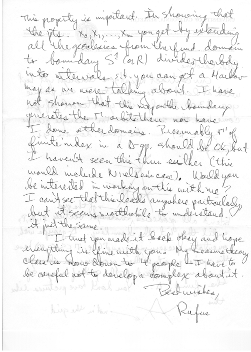
Dear Caroline,

Our contract hasn't gone through yet but with at least 80% certainty I can promise you $300 to help with your travel if you come here this summer. For triangle groups in $SL(2,R)$ — the fund. domain is a triangle

p_1 p_2 p_3

Extend all geodesics of the grid through vertices p_1, p_2, p_3 until they hit the boundary. Δ-groups with orientation preserving group Γ (i.e. $\Gamma \subset SL(2,R)$) have this special property: these geodesics ~~or always~~ are everywhere part of the grid of images of the fund. Δ. You don't have pictures like this

not in the grid



This property is important. In showing that the pts. $x_0, x_1, \ldots, x_m$ you get by extending all the geodesics from the fund. domain to boundary S^1 (or R) divides the bdy. into intervals s.t. you can get a Markov map as we were talking about. I have not shown that this map on the boundary generates the Γ-orbits there nor have I done other domains. Presumably Γ' of finite index in a Δ-gp. should be OK, but I haven't seen this thru either (this would include Nielsen's case). Would you be interested in working on this with me? I can't see that this leads anywhere particularly, but it seems worthwhile to understand it just the same.

I trust you made it back okay and hope everything is fine with you. My measure theory class is now down to 4 people — I have to be careful not to develop a complex about it.

Best wishes,

Rufus

In summer 1978, I revisited Berkeley. We started working seriously together, and then suddenly one weekend Rufus was dead. I would like to quote from the introduction to our paper:

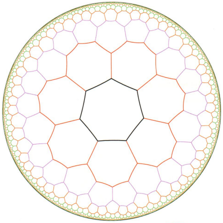
The writing of this paper has been overshadowed by Rufus' untimely death in July. We had intended to write jointly: most of the main ideas were worked out together and I have done my best to complete them. In sorrow, I dedicate this work to his memory. Berkeley, September 1978.

Key to the construction

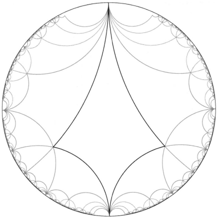
I want to explain the construction in a slightly more general way than done in our orginal paper. The starting point is a condition on $\mathcal{R}$ called *even corners*: the extended sides of $\mathcal{R}$ form part of the tessellation $\mathcal{T}$.

Each side s of $\mathcal{R}$ is labelled on the outside by some $e \in G_0$. Let $H(e)$ be the closure in $\bar{\mathbb{D}} = \mathbb{D} \cup \mathcal{S}^1$ of the half space cut off by the extension of s not containing $\mathcal{R}$. Even corners implies $H(e) \cap H(e') \neq \emptyset$ for at most one other $e' \in G_0$. Define $f : \overline{\mathbb{D}} \setminus \mathcal{R} \to \overline{\mathbb{D}}$ as follows. On $H(e) \setminus \cup_{e' \neq e} H(e')$ define $f(x) = e^{-1}(x)$; on $H(e) \cap H(e')$ make one or other choice.

Why is $f|_{S^1}$ Markov? Consider the union $\mathcal{L}$ of *all* the sides of $\mathcal{T}$ which go through any vertex of $\mathcal{R}$. Their endpoints partition S^1 into finitely many intervals. Moreover the image of line in $\mathcal{L}$ under either of the possible definitions of f is again in $\mathcal{L}$. Thus the set of endpoints maps into itself. From this we see that f is Markov.



Tessellation without even corners.



Tessellation with even corners.

Properties of f

PROPOSITION [S. 1991*] • The restriction of f to S^1 has all the properties of a Bowen-Series coding.

• The restriction of f to $\mathbb{D}$ reduces length in the sense that if $f(g\mathcal{R}) = h\mathcal{R}$ then $|h| = |g| - 1$.

• Let $B(e) = \{x \in H(e) : f(x) = e^{-1}\}$. Then $\cap_{r=0}^{k} f^{-r} B(e_{i_r}) \neq \emptyset$ iff $e_{i_1} e_{i_2} \dots e_{i_k}$ is a shortest word in G_0; moreover each $g \in G$ has a unique shortest representation of this form.

Thus we have simultaneously constructed a B-S map and solved the word problem in G *in a finite type way*.

GENERALIZATIONS

PROPOSITION [S. 1981] The B-S construction can be extended to general Fuchsian groups by using f-expansions to define paths in the Cayley graph which end in suitable endpoints for the intervals $\mathcal{I}$.

PROPOSITION [ROCHA 1996] Let G be a Kleinian group whose fundamental domain in $\mathbb{H}^3$ is such that every extended face lies in the tessellation $\mathcal{T}$. Then one can construct a Markov map f with similar properties to the above.

**Ergodic theory, symbolic dynamics and hyperbolic spaces*, Eds. Bedford, Keane, S. OUP 1991.

Uses of the coding

1. Hausdorff dimension of the limit set.
2. Relationship to the geodesic flow; zeta functions and counting problems.
3. Simple geodesics and counting intersection numbers.
4. Generalizations of continued fractions and Diophantine approximation.
5. Random walks on Cayley graphs.
6. Bounded cohomology.
7. Growth functions of groups.

Philosophical remarks Some applications are very specific and depend on the detail of the coding. Some are not so specific but do depend on properties of the coding, for example that f is piecewise complex analytic and can be extended to a neighborhood of each interval of definition.
Some results were orginally proved as extensions of results for free groups using the specifics of the coding, but have since been extended to much wider classes of groups.
The coding reveals the *finite type* or recursive nature of the tessellation and associated Cayley graph.

Hausdorff dimension

By applying the theory of Gibbs states as in his Springer book *Equilibrium states ...*, Bowen proved:

Theorem[Bowen 1979] The Hausdorff dimension of the limit set of a quasifuchsian group is ≥ 1 with equality if and only if G is Fuchsian, that is, $\Lambda(G)$ is a round circle.

Remark *This paper was written by Rufus before our work on coding was complete. According to Dennis Sullivan, it was through studying this paper and seeing that similar methods could be applied to Julia sets which inspired him to envisage and develop his famous 'dictionary' between Kleinian groups and holomorphic dynamics.*

Theorem[Anderson-Rocha 1997] For a large class of Kleinian groups, the Hausdorff dimension of the limit sets varies real analytically as the groups vary holomophically.

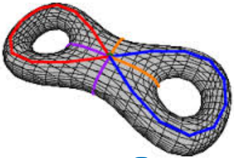
Geodesic flow

There is an intimate relationship between the B-S coding and geodesic flow on the unit tangent bundle of a surface $S = \mathbb{D}/G$ of constant negative curvature.

Here are two ways of encoding a geodesic γ on S by a bi-infinite string of generators of G:

Method 1 The labelled sides of $\mathcal{R}$ project to lines on S which can be used as a Poincaré section for the flow. Recording the labels of the sides cut in order along γ leads to a natural representation of the flow as a suspension over a shift on a certain subset of infinite strings of generators.

Method 2 Lift γ to $\mathbb{D}$ and record the f-expansions of its two end-points on S^1. This method is less obviously related to the flow but has the advantage of being finite type.



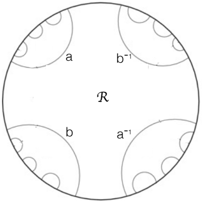
Sides of $\mathcal{R}$ projected to S.

THEOREM [S. 1981, 1986] These two coding methods are sufficiently close, in a precise sense, that the geodesic flow can be represented by a suspension flow over $\Sigma_{\mathbb{A}}$, so that smooth closed geodesics correspond to periodic sequences.

This effectively gives a concrete geometrical Markov partition for the flow. It has been exploited by Mark Pollicott, Richard Sharp and others in problems about counting closed geodesics, the Ruelle zeta function etc.

Extending known results for free groups

The free group F_k on k generators can be viewed as a Fuchsian group G, giving a special case of the coding. To see this, take a fundamental domain with $2k$ sides and no interior vertices which meets S^1 in intervals. The side pairings generate F_k and Λ is identified with the space of infinite reduced sequences in the generators.



EXAMPLE 1: BOUNDED COHOMOLOGY This is the cohomology $H_b^r(G, \mathbb{R})$ formed from the complex of continuous functions $C_b^r = \{\phi : G^r \to \mathbb{R}\}$ with bounded image and boundary map $\delta : C^r \to C^{r+1}$: $\delta(\phi)(g_0, \ldots, g_r)$
$= f(g_1 \ldots g_r) + \sum_{i=1}^r (-1)^r \phi(g_0, \ldots, g_{i-1}g_i, \ldots g_r) + (-1)^{r+1}\phi(g_0 \ldots g_{r-1})$.

THEOREM [BROOKS 1980] $H_b^2(F_k, \mathbb{R})$ is infinitely generated.

THEOREM [BROOKS-S. 1982] Same result with F_k replaced by an arbitrary Fuchsian group.

THEOREM [EPSTEIN AND FUJIWARA 1997] Same result with F_k replaced by an arbitrary Gromov hyperbolic group.

Example 2: Martin boundaries of random walks

The *Cayley graph* $\mathcal{G}$ of (G, G_0) is the graph whose vertices are elements of G with an edge from g to h whenever $g^{-1}h \in G_0$. Given a (finite support) probability measure p on G, the random walk on $\mathcal{G}$ is the walk which jumps from a vertex v with probability $p(g)$ to the vertex gv. A function on $\mathcal{G}$ is harmonic if $h(x) = \sum_{g\in G} p(g)h(gx)$.

The *Martin boundary* is a compactification $\bar{\mathcal{G}}$ of $\mathcal{G}$ such that all positive harmonic functions can be represented as an integral over the boundary $\partial\mathcal{G} = \bar{\mathcal{G}} \setminus \mathcal{G}$ with respect to the *Martin kernel* $K(x, \xi), x \in \mathcal{G}, \xi \in \partial\mathcal{G}$.

Theorem [Dynkin and Malyutov 1969] Identify the free group F_k on $k > 1$ generators as a Fuchsian group. Then the Martin boundary of a random walk on F_k is its limit set Λ.

Theorem [S. 1983] Same result with F_k replaced by an arbitrary Fuchsian group without parabolics (and a minor modification if it has).

Theorem [Kaimanovich 1994] Same result with F_k replaced by an arbitrary Gromov hyperbolic group G and Λ replaced by the Gromov boundary of G.

Some recent work on ergodic averaging

Joint with Sasha Bufetov and Alexey Klimenko

Suppose given a group G, a symmetric set of generators G_0, a measure space (X, ν) and a measure preserving action of G on X.

Let $\phi : X \to \mathbb{C}$. For $g \in G$, define $T_g(\phi)(x) = \phi(g^{-1}x)$. Let s_n be the number of words of length n in the generators G_0 and define:

$$\sigma_n(\phi)(x) = 1/s_n \sum_{g \in G, |g|=n} \phi(T_g(x)) \quad \text{and} \quad c_n(\phi)(x) = 1/n \sum_{m \leq n} \sigma_n(\phi)(x)$$

THEOREM [BUFETOV 2002] Let G be the free group, $F_k, k > 1$. Then if $\int |\phi| \log^+ |\phi| d\nu < \infty$, the sequence $\sigma_{2n}(\phi)$ converges ν a.e. and in L^1 to a $G^{(2)}$-invariant function.

Remark Guivarc'h (1969) proved L^2-convergence. For $p > 1$, Nevo and Stein (1994) proved convergence in L^p and a.e.

THEOREM [BUFETOV-S. 2011] Let G be a Fuchsian group. For $\phi \in L^1$, the Cesàro averages $c_n(\phi)$ converge a.e. and in L^1 to a $G^{(2)}$-invariant function.

The proof uses Bufetov's results on Markov operators. Note $T_g = T_{e_{i_1}} \dots T_{e_{i_k}}$ where we express $g = e_{i_1} \dots e_{i_k}$ using the B-S coding. The proof required checking some additional properties of the transition matrix $\mathbb{A}$. Recent results by Bufetov et al; and Pollicott & Sharp extend to word hyperbolic groups.

Convergence of spherical averages

The problem with extending the result on spherical averages $\sigma_n(\phi)(x) = \sum_{g \in G, |g|=n} \phi(T_g(x))$ is that the proof depends heavily on symmetry of inversion in F_k: namely, if $e_{i_1} \dots e_{i_k}$ is a reduced word in the generators, then so is $e_{i_k}^{-1} \dots e_{i_1}^{-1}$. This is used to define a self-adjoint operator integral to the proof.

At first sight this difficulty seems insurmountable, since defining a canonical word paths in the Cayley graph round a vertex requires a choice between right and left, while the inverse path goes round the opposite direction.

However a recent result of Matthew Wroten gets round a similar problem in a new way. Recall that smooth closed geodesics on $\mathbb{D}/G$ correspond bijectively to conjugacy classes in G.

Theorem [Wroten 2014] Let G be a Fuchsian group. Then with respect to the uniform distribution on conjugacy classes of length n, the number of self-intersections approaches a Gaussian distribution as $n \to \infty$.

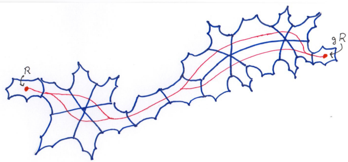
This generalises results of Lalley and Chas (2012) for F_k. A main additional ingredient is to use an extension of the B-S coding devised by Mark Lustig (1987) to count intersection numbers.

Wroten's idea

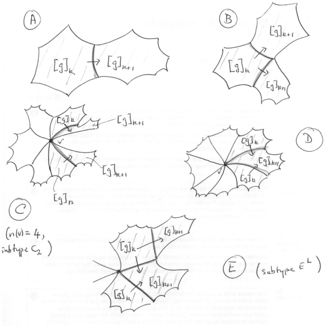
To prove his result, Wroten devised a coding which has strong properties under inversion, as follows.

Let $g \in G$ and consider the 'snake' consisting of all fundamental domains which occur in *any* shortest path from $\mathcal{R}$ to $g\mathcal{R}$. Denote the union of all regions in subpaths of length k by $[g]_k$. The additional regions forming $[g]_{k+1}$ are attached to those in $[g]_k$ across certain 'leading edges' of $[g]_k$. The key point is that there are only finitely many ways in which such attachments can occur; AND that the process is self-inverse.

It turns out that the new regions are attached across one, two or three edges. The states will consist of the possible configurations, together with the labels of the attaching edges.



A typical snake



Possible configurations

The self inverse transition matrix

A state is a possible configuration of 'end' regions in a snake of length k, to which can be attached regions forming a snake of length $k+1$, together with the ordered labels of the sides along which the attaching is made.

PROPOSITION [S. 2017] The possible transitions are of finite type with associated $0-1$ transition matrix $\mathbb{A}$. Let $\mathcal{S}$ denote the set of states, among which are initial states $I \subset \mathcal{S}$ and final states $F \subset \mathcal{S}$. Let $\mathcal{B}_n$ denote the set of finite admissible sequences of length n whose first and last elements belong to I, F respectively. Then there is a projection $\Pi : \mathcal{B}_n \to G$ such that:

(1) Π is a bijection between elements of G of length n and $\mathcal{B}_n$.
(2) there is an involution $\tau : \mathcal{S} \to \mathcal{S}$ such that $\tau(I) = F$ and such that for $i, j \in \mathcal{S}$ we have $a_{i,j} = 1$ iff $a_{\tau(j),\tau(i)} = 1$.
(3) There are two special projections $\pi_R, \pi_L : \mathcal{S} \to G_0$ which when extended to $\Sigma_{\mathbb{A}}^{+}$ give the two usual B-S codings. Moreover $\pi_L(i)^{-1} = \pi_R(\tau(i))$.

WORK IN PROGRESS Bufetov's Markov operator P has suitable self-adjointness and transitivity properties to enable one to apply the method of [Bufetov 2002]. Hence spherical averages for a Fuchsian group converge as claimed.

Remark on Wroten's idea:

Floyd and Plotnick (1987) computed the growth functions $f(t) = \sum_n t^n s_n$ of many Fuchsian groups by showing that the process of passing from $\cup_{g \in G:|g|=n} g\mathcal{R}$ to $\cup_{g \in G:|g|=n+1} g\mathcal{R}$ is finite type. (Their constructions can be simplified by use of the B-S coding.)They observed that their process could be reversed led to the additional result $f(t) = f(1/t)$. This was the origin of the idea that the coding might be reversible; it took Wroten's construction to implement.