Riemann surfaces, dynamics and geometry

Course Notes — Harvard University — Math 275

Spring 1998; Spring 2000; Fall 2001; Fall 2009

C. McMullen

August 30, 2010

Contents

1 Introduction

This course will concern the interaction between:

- hyperbolic geometry in dimensions 2 and 3;
- the dynamics of iterated rational maps; and
- the theory of Riemann surfaces and their deformations.

Rigidity of 3-manifolds. A *hyperbolic manifold* M^n is a Riemannian manifold with a metric of constant curvature -1. Almost all our hyperbolic manifolds will be *complete*. There is a unique simply-connected complete hyperbolic manifold $\mathbb{H}^n$ of dimension n, so $M = \mathbb{H}^n/\Gamma$ where $\Gamma \subset \operatorname{Isom}(\mathbb{H}^n)$ is a discrete, torsion-free group isomorphic to $\pi_1(M)$.

In dimensions $n \geq 3$ and higher, closed hyperbolic manifolds are *rigid*. That is, composition of the maps

$$\begin{aligned}\{\text{closed hyperbolic } n\text{-manifolds}\} &\rightarrow \{\text{topological } n\text{-manifolds}\} \\ &\xrightarrow{\pi_1} \{\text{finitely generated groups}\}\end{aligned}$$

is *injective*. This is:

Theorem 1.1 (Mostow rigidity) *Any isomorphism $\iota : \pi_1(M) \to \pi_1(N)$ between the fundamental groups of closed hyperbolic manifolds of dimension 3 or more can be realized as $\iota = \pi_1(f)$ where $f : M \to N$ is an isometry.*

It follows that geometric invariants such as $\operatorname{vol}(M)$, the hyperbolic length $\ell(\gamma)$, $\gamma \in \pi_1(M)$, etc. are (in principle) completely determined by the topological manifold M, or even more combinatorially, by $\pi_1(M)$. Prasad extended this result to finite volume manifolds.

In dimension 3 remarkable results of Thurston suggest that this rigidity coexists with just enough flexibility that *most 3-manifolds are hyperbolic.* Thus the 'forgetful map' above is almost a bijection. For example one has:

Theorem 1.2 (Thurston) *If M is a closed Haken 3-manifold, then M is hyperbolic iff $\pi_1(M)$ is infinite and does not contain $\mathbb{Z} \oplus \mathbb{Z}$.*

One can also compare the situation for manifolds of dimension 2. Closed, orientable surfaces are classified by their genus $g = 0, 1, 2, \ldots$ and they always admit metrics of constant curvature. For the sphere ($g = 0$) this metric is (essentially) unique, but for the torus there is already a moduli space ($\mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$). Any surface of genus $g \geq 2$ admits a complex structure, depending on $3g - 3$ complex parameters, and each complex structure has a unique compatible hyperbolic metric.

Teichmüller space parameterizes these structures and will play a crucial role in the *construction* of rational maps and 3-manifolds with prescribed topology.

For 3-manifolds our goal is to understand some examples of hyperbolic manifolds, prove Mostow rigidity and related results, and give an idea of Thurston's construction of hyperbolic structures on Haken manifolds.

Dynamical systems. Any hyperbolic 3-manifold M gives rise to a *conformal dynamical system* by considering the action of $\pi_1(M)$ on $\widehat{\mathbb{C}}$ thought of as the sphere at infinity for $\mathbb{H}^3$. We have $\mathrm{Isom}^+(\mathbb{H}^3) = \mathrm{Aut}(\widehat{\mathbb{C}}) = \mathrm{PSL}_2(\mathbb{C})$.

Iterated rational maps provide another source of conformal dynamics on the Riemann sphere. These maps exhibit both expanding and contracting features. For example, let

$$f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$$

be a rational map of degree $d > 1$. Then

$$\int_{\widehat{\mathbb{C}}} |(f^n)'(z)|\, |dz|^2 = d^n \,\mathrm{area}(\widehat{\mathbb{C}})$$

tends to infinity exponentially fast as $n \to \infty$. Here $|\cdot|$ denotes the spherical metric. On the other hand, f has $2d-2$ critical points where f is highly contracting.

There are surprisingly many similarities between the theories of rational maps and of Kleinian groups. For example the following rigidity result holds:

Theorem 1.3 (Critically finite rigidity) *Let f and g be rational maps all of whose critical points are preperiodic. Then with rare exceptions, any topological conjugacy between f and g can be deformed to a conformal conjugacy. (In the exceptional cases, f and g are double-covered by an endomorphism of a torus.)*

On the other hand, Thurston has also given a geometrization theorem characterizing rational maps among branched covers of the sphere. The method of proof parallels the more difficult geometrization result for Haken 3-manifolds. Understanding the case of rational maps is good preparation for the 3-dimensional theory, like adaptive excursions from the base camp at 17,000 feet on Mt. Everest.

An exhaustive theory of dynamical systems is probably unachievable. One usually tries to understand the behavior of $f^n(z)$ for most $z \in \widehat{\mathbb{C}}$, and for most $f \in \mathrm{Rat}_d$.

A rational map f is *structurally stable* if all sufficiently nearby maps g are topologically conjugate to f. Despite the many mysteries of general rational maps, one knows:

Theorem 1.4 *The set of structurally stable rational maps is open and dense.*

Many of the components of the structurally stable maps are encoded by critically finite examples, so we are approaching a classification theory in this setting as well.

References. For a brief survey of the iterations on Teichmüller space for rational maps and Kleinian groups, see [Mc3]. For a proof of the density of structural stability in Rat_d, see [MSS] or [Mc4]. Speculations on the role of structural stability in the biology of morphogenesis and other sciences can be found in [Tm].

1.1 Examples of hyperbolic manifolds

A *Kleinian group* Γ is any discrete subgroup of $\mathrm{Isom}(\mathbb{H}^3)$. Its *domain of discontinuity* $\Omega(\Gamma) \subset \widehat{\mathbb{C}}$ is the largest open set on which Γ acts properly discontinuously. The *limit set* $\Lambda(\Gamma) = \widehat{\mathbb{C}} - \Omega(\Gamma)$ can also be defined as $\overline{\Gamma x} \cap \widehat{\mathbb{C}}^{n-1}$ for any $x \in \mathbb{H}^3$.

We say Γ is *elementary* if it is abelian, or more general if it contains an abelian subgroup with finite index. Excluding these cases, Λ is also the closure of the set of repelling fixed-points of $\gamma \in \Gamma$, and the minimal closed Γ-invariant set with $|\Lambda| > 2$.

The quotient $M = \mathbb{H}^3/\Gamma$ is an orbifold, and a manifold if Γ is torsion-free. The *Kleinian manifold*

$$\overline{M} = (\mathbb{H}^3 \cup \Omega(\Gamma))/\Gamma$$

has a complete hyperbolic metric on its interior and a Riemann surface structure (indeed a projective structure) on its boundary.

We now turn to some examples.

1. *Simply-connected surfaces and 3-space.* The unit disk Δ and the upper half-plane $\mathbb{H}$ in $\mathbb{C}$ are models for the hyperbolic plane with metrics $2|dz|/(1-|z|^2)$ and $|dz|/\operatorname{Im}(z)$ respectively. We have $\mathrm{Isom}^+(\mathbb{H}) = \mathrm{PSL}_2(\mathbb{R})$ acting by Möbius transformations.

   The geodesics are circular arcs perpendicular to the boundary, in either model.[1]

   Similarly $\mathbb{H}^{n+1}$ can be modeled on the upper half-space $\{(x_0, \ldots, x_n) : x_0 > 0\}$ in $\mathbb{R}^{n+1}$, with the metric $|dx|/x_0$; or on the unit ball with the metric $2|dx|/(1-|x|^2)$.

   Planes in $\mathbb{H}^3$ are hemispheres perpendicular to the boundary. Reflections through circles on $\widehat{\mathbb{C}}$ prolong to isometric reflections through hyperplane in $\mathbb{H}^3$, and lead to the isomorphism $\mathrm{Isom}^+(\mathbb{H}^3) = \mathrm{Aut}(\widehat{\mathbb{C}}) = \mathrm{PSL}_2(\mathbb{C})$ acting by Möbius transformations.

2. *Domains in* $\widehat{\mathbb{C}}$. By a generalization of the Riemann mapping theorem, the unit disk covers (analytically) any domain $\Omega \subset \widehat{\mathbb{C}}$ with $|\widehat{\mathbb{C}} - \Omega| \geq 3$. Thus any such Ω is a hyperbolic surface.

3. *Examples with* $\pi_1(M) = \mathbb{Z}$.

[1] The unit disk Δ can be taken as a model of either $\mathbb{RH}^2$ or $\mathbb{CH}^1$; for the latter space the natural metric is $|dz|/(1-|z|^2)$ with constant curvature -4. The symmetric space $\mathbb{CH}^n$, $n > 1$ contains copies of both $\mathbb{RH}^2$ and $\mathbb{CH}^1$, with curvatures -1 and -4 respectively. The space $\mathbb{CH}^n$ can be modeled on the unit ball in $\mathbb{C}^n$ with its Hermitian invariant metric, e.g. the Bergman metric.

(a) The punctured disk $\Delta^* = \mathbb{H}/\langle z \mapsto z+1\rangle$; the covering map $\mathbb{H} \to \Delta^*$ is given by $\pi(z) = e^{2\pi i z}$. A neighborhood of the puncture has finite volume. The limit set is $\Lambda = \{\infty\}$.

(b) The annulus $A(r) = \{z : r < |z| < 1\}$. We have $A(r) = \mathbb{H}/\lambda^{\mathbb{Z}}$, $\lambda > 1$, with the covering map $\mathbb{H} \to A(r)$ given by

$$\pi(z) = z^{2\pi i/\log \lambda}.$$

Note that $\pi(\lambda z) = \pi(z)$ and $\pi(\mathbb{R}^+) = S^1$, the unit circle. $\pi(\mathbb{R}^-) = S^1(r)$ where $r = \exp(-2\pi^2/\log\lambda)$, so $\lambda = \exp(2\pi^2/\log(1/r))$.

Thus the length of the core geodesic of $A(r)$ is $2\pi^2/\log(1/r)$. For example this shows $A(r)$ and $A(s)$ cannot be isomorphic if $r \neq s$.

$\Lambda = \{0, \infty\}$.

(c) The torus $X = \mathbb{C}^*/\lambda^{\mathbb{Z}}$ is isomorphic to $\mathbb{C}/(2\pi i\mathbb{Z} \oplus \log\lambda\mathbb{Z})$.

(d) The space $\mathbb{H}^3/\lambda^{\mathbb{Z}}$ is a solid torus with core curve of length $\log\lambda$.

4. *Pairs of pants, genus two and handlebodies.* Now consider 3 disjoint geodesic in Δ symmetric under rotation. The group Γ' generated by reflections through these geodesics gives a quotient orbifold which is a hexagon with alternating edges removed. The limit set is a *Cantor set.*

Let $\Gamma \subset \Gamma'$ be the index two subgroup preserving orientation. This is the familiar subgroup

$$\Gamma \cong \langle AB, AC\rangle \cong \mathbb{Z} * \mathbb{Z} \subset \langle A, B, C \ : \ A^2 = B^2 = C^2 = 1\rangle \cong \Gamma'.$$

Then Δ/Γ' is a pair of pants, i.e. the double of the hexagon.

Next consider the action of Γ on $\widehat{\mathbb{C}} - \Omega$; the quotient X is the double of a pair of pants, namely a surface of genus two. A fundamental domain is the region outside of 4 disjoint disks. Looking at the region above the hemispheres these disks bound, we see $M = \mathbb{H}^3/\Gamma$ is a handlebody of genus two.

Conversely, for any Riemann surface X of genus two, the realization of X as the boundary of a topological handlebody M determines a planar covering space $\Omega \to X$ which can be compactified to the sphere. The action of $\mathbb{Z} * \mathbb{Z} = \pi_1(M)$ on Ω extends to the sphere and gives a Kleinian group.

5. *Surfaces of genus two.* It is familiar from the classification of surfaces that a torus can be obtained from a 4-gon by identifying opposite sides, a surface of genus 2 from an 8-gon, and a surface of genus g from a $4g$-gon. The resulting cell complex has only one vertex, where $4g$ faces come together.

A regular octagon in $\mathbb{H}$ with interior angles of $45°$ serves as a fundamental domain for a Fuchsian group such that $\mathbb{H}/\Gamma = X$ has genus two. Extending the action to $\mathbb{H}^3$ we obtain a compact Kleinian manifold $\overline{M} \cong S \times [0,1]$.

6. *The triply-punctured sphere.* There is only one ideal triangle T in $\mathbb{H}$, and by doubling it we obtain hyperbolic structure on $X = \widehat{\mathbb{C}} - \{0, 1, \infty\}$. The covering map $\pi : \mathbb{H} \to X$ can be constructed by taking the Riemann mapping from T to $\mathbb{H}$, sending the vertices to $\{0, 1, \infty\}$, and then prolonging by Schwarz reflection.

   The corresponding group is arithmetic, indeed $X = \mathbb{H}/\Gamma(2)$.

   We have $\mathrm{Aut}(X) = S_3 = \mathrm{SL}_2(\mathbb{Z}/2)$, which can be seen as the permutation group of the cusps of X.

   The fact that X is covered by $\mathbb{H}$ proves the *Little Picard Theorem*: an entire $f : \mathbb{C} \to \mathbb{C}$ omitting 2 values is constant. Indeed, the omitted values can be taken to be 0 and 1; then f lifts to the universal cover, giving a map $\widetilde{f} : \mathbb{C} \to \mathbb{H}$ which must be constant.

7. *Punctured tori.* To handle a torus we need to delete a point so the Euler characteristic becomes negative. Then we can take as a fundamental domain an ideal quadrilateral, and glue opposite sides. (It is essential to be careful doing the gluing! So the holonomy around a cusp is parabolic.)

   These examples can be perturbed to give Poincaré's examples of quasifuchsian groups, by taking reflections in a necklace of 4 tangent circles.

8. *A compact hyperbolic 3-manifold.* Here is an example to which Mostow rigidity applies: take a regular hyperbolic dodecahedron D with internal dihedral angles $72^\circ = 2\pi/5$. Then identify opposite faces making a twist of 3/10ths of a revolution. The 30 edges of D are identified in six groups of 5 each, so we obtain a manifold structure around each edge. The link of a vertex is orientable and admits a metric of constant curvature 1, so it is a sphere. (In fact there is only one vertex in $D/\sim$, and its link is an S^2 tiled by 20 triangles in the icosahedral pattern.)

9. *The Hopf link.* As a warmup to a hyperbolic link complement, let's look at the Hopf link $L \subset S^3$, which can be thought of as a pair of disjoint geodesics. Then $M = S^3 - L = S^1 \times S^1 \times (0, 1)$; note that we have omitted the boundary, so M is open.

   We claim $S^3 - L$ can be obtained from an octahedron by identifying sides in a suitable pattern, then deleting two vertices (say the north and south poles).

   To see this, first write $L = K_0 \sqcup K_1$. Then $T = S^3 - K_0$ is a solid torus T, and $M = S^3 - L$ is just the complement of the core curve in T, $T - K_1$. To obtain a cell, cut T along a disk to obtain a solid cylinder C; then K_1 becomes an interval I_1 joining the ends of C. Now cut along a rectangle joining the round part of ∂C to I_1, and open it up. The result is like a split log. Shrink the parts of the boundary running along K_0 and K_1 to points, which will become the north and south poles. We obtain a ball with 4 longitudes joining the poles. Adding a square of 4 more edges to form the equator, the result is an octahedron.

Thus the face identifications preserve the northern and southern hemispheres, and match each quadrant to its 'opposite' quadrant in the same hemisphere.

10. *The Whitehead link.* The Whitehead link $W \subset S^3$ is a symmetric link of two unknots, with linking number zero, but one clasp.

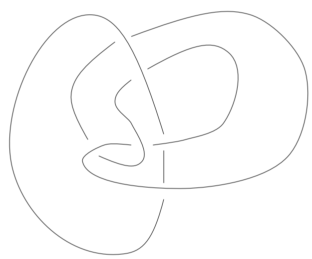

Figure 1. The Whitehead link; $\mathrm{vol}(S^3 - W) = 3.66386\ldots$

Its complement $M = S^3 - W$ has a finite volume hyperbolic metric that can be obtained from a *regular ideal octahedron* in $\mathbb{H}^3$ by a suitable gluing pattern. To see this, first note that for the unknot K_0, $S^3 - K_0$ is a solid torus $D^2 \times S^1$. Thus $M = D^2 \times S^1 - K_1$ for a certain knot K_1 with winding number zero.

Cutting $D^2 \times S^1$ with a disk $D = D^2 \times \{t\}$ meeting K in two points, we obtain a cylinder $D^2 \times [0,1]$ from which a pair of intervals $I_1 \sqcup I_2$ must be removed.

After a little deformation, we can replace $D^2 \times [0,1]$ with a cube $C = [0,1]^3$, and the two intervals with segments joining opposite faces. The original knot K_0 now corresponds to 8 of the 12 edges of C. To obtain a 3-cell, we cut $C - (I_1 \cup I_2)$ along rectangles R_i joining I_i to an adjacent face. Then we obtain a cell structure on $S^2 = \partial D^3$ with 4 pentagons and 4 quadrilaterals. Each pentagon has 3 edges along K_0, and each quadrilateral has 1 edge along K_1. Collapsing these edges, we obtain an octahedron.

The dihedral angles of a regular octahedron are 90°, so it can be reglued to give a hyperbolic structure on $S^3 - W$.

11. *Arithmetic examples.* Let $\mathcal{O}_d$ be the ring of algebraic integers in the quadratic number field $K = \mathbb{Q}(\sqrt{-d})$, $d \geq 1$. Then $\mathcal{O}_d \subset \mathbb{C}$ is discrete, so $\mathrm{SL}_2(\mathcal{O}_d)$ is a Kleinian group. In fact $M_d = \mathbb{H}/\,\mathrm{SL}_2(\mathcal{O}_d)$ is a finite-volume orbifold (Borel), with as many cusps as the class number of K.

The Whitehead link and Borromean rings complements are commensurable $\mathbb{H}^3/\,\mathrm{SL}_2(\mathbb{Z}[i])$, and the figure eight knot to $\mathrm{SL}_2(\mathbb{Z}[\omega])$ (the Gaussian and Eisenstein integers respectively). Both have class number one.

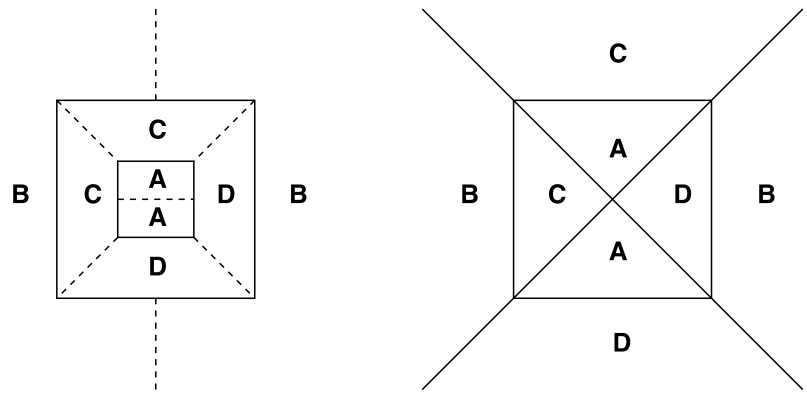

Figure 2. Building the Whitehead link complement from an octahedron.

Theorem 1.5 (Thurston) *A knot complement $M = S^3 - K$ admits a complete hyperbolic structure of finite volume iff M is atoroidal: that is, iff every copy of $\mathbb{Z} \oplus \mathbb{Z}$ in $\pi_1(M)$ comes from a tubular neighborhood of the knot.*

Example. For trefoil knot $T_{2,3}$, and more generally any torus knot, $T_{p,q}$, the complement $M_{p,q} = S^3 - T_{p,q}$ is toroidal. In fact $M_{p,q}$ is a *Seifert fibered manifold*; it admits a nontrivial S^1 action. This action gives rise to many nontrivial copies of $\mathbb{Z} \oplus \mathbb{Z}$ in $\pi_1(M_{p,q})$.

To see the S^1 action, first think of S^3 as the unit sphere in $\mathbb{C}^2$. The circle action on S^3 by

$$(z_1, z_2) \mapsto (e^{pi\theta} z_1, e^{qi\theta} z_2).$$

A generic orbit is a (p, q)-torus knot, so its complement admits an S^1 action.

To visualized an immersed incompressible torus in the complement of the trefoil knot, imagine $T_{2,3}$ as a cable lying on a torus $\Sigma = S^1 \times S^1 \subset S^3$. Try to wrap the torus Σ in paper, avoiding the cable. Passing the paper alternately under and over the cable, it closes up to make an immersed torus.

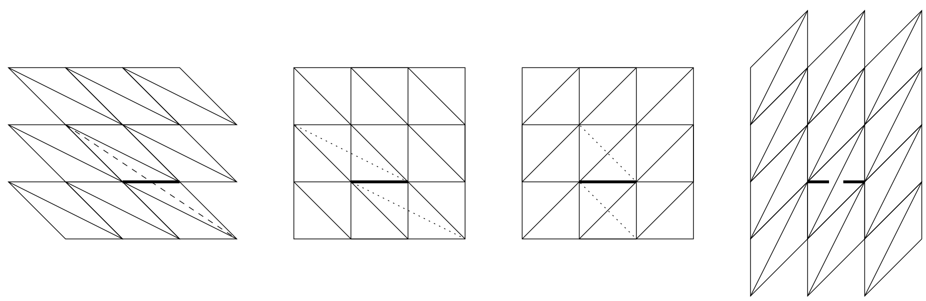

Figure 3. Tiling of the $\mathbb{Z}^2 \ltimes \mathbb{Z}$ cover of the figure eight knot complement.

The figure eight knot. The simplest hyperbolic knot complement is $M = S^3 - K$ where K is the figure eight knot. The hyperbolic manifold M can be built from 2 regular ideal tetrahedra.

One can describe M concretely the complement T^* of the zero section of the torus bundle $T \to S^1$ with monodromy $A = PQ = \left(\begin{smallmatrix} 1 & 1 \\ 2 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. This factorization of A into elementary matrices determines a sequence of triangulations of the plane which are related by cobordisms through tetrahedra (see Figure 3). By inspection, six tetrahedra are adjacent to each edge of the resulting triangulation of the one-point compactification of M. Since the dihedral angles of a regular hyperbolic tetrahedron are all 60°, M can be given a complete hyperbolic structure. There are just 2 edges in the quotient, with six tetrahedra coming together along an edge (like equilateral triangles tiling a hexagon).

We note that the figure eight *knot* (which includes the information of the embedding of M into S^3) can be constructed as the Murasugi sum of two Hopf bands, which makes clear that it is a fibered knot and that its monodromy is the product of two Dehn twists, P and Q. Indeed, the Hopf link complement itself can be regarded as the fibered link obtained by suspending a single Dehn twist on a annulus. Here we have arranged that the monodromy is the identity on the boundary of the fiber, and that the resulting framing of the boundary corresponds to the framing of the Hopf link by meridians and by fiber. If we had, instead, use the identity map on the annulus, we would have obtained the same 3-manifold, but now presented as the complement of a link in $S^2 \times S^1$.

Reflection groups. Quite generally, one can consider any convex polyhedron P in $\mathbb{H}^n$, $\mathbb{R}^n$ or S^n whose dihedral angles are of the form $\theta_i = \pi/n_i$. Then the group Γ generated by reflections in the sides of P is discrete, and P forms a fundamental domain for Γ. The proof is by induction on the dimension of P (see e.g. [Rat, §7]).

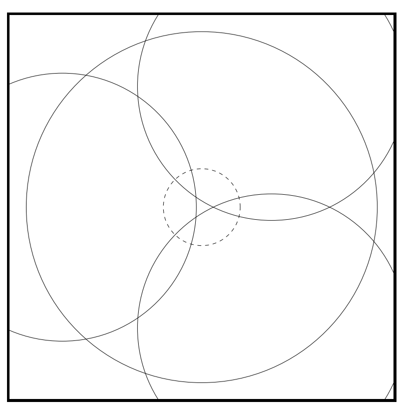

Figure 4. Fundamental polyhedron P_7.

3-dimensional pairs of pants. A pair of pants is associated to each triangle in

$\mathbb{H}^2$ with vertices outside the plane. Similary we can consider regular tetrahedra with vertices outside $\mathbb{H}^3$. For each $n \geq 7$ there is a unique such tetrahedron P_n such that its 'convex core' K_n is bounded by four $2\pi/n$ triangles and by four right hexagons.

The four faces of P_n are simply four circles in the sphere meeting with dihedral angles of π/n. The example P_7 is depicted in Figure 4. Note that 3 of the circles simply determine the $(2, 3, 7)$ triangle group acting on a copy of $\mathbb{H}$, namely the common perpendicular circle (shown as a dotted line). Reflections in the sides of P_n generate a discrete group Γ_n for all n; when n is even, P_n is a fundamental domain for this group.

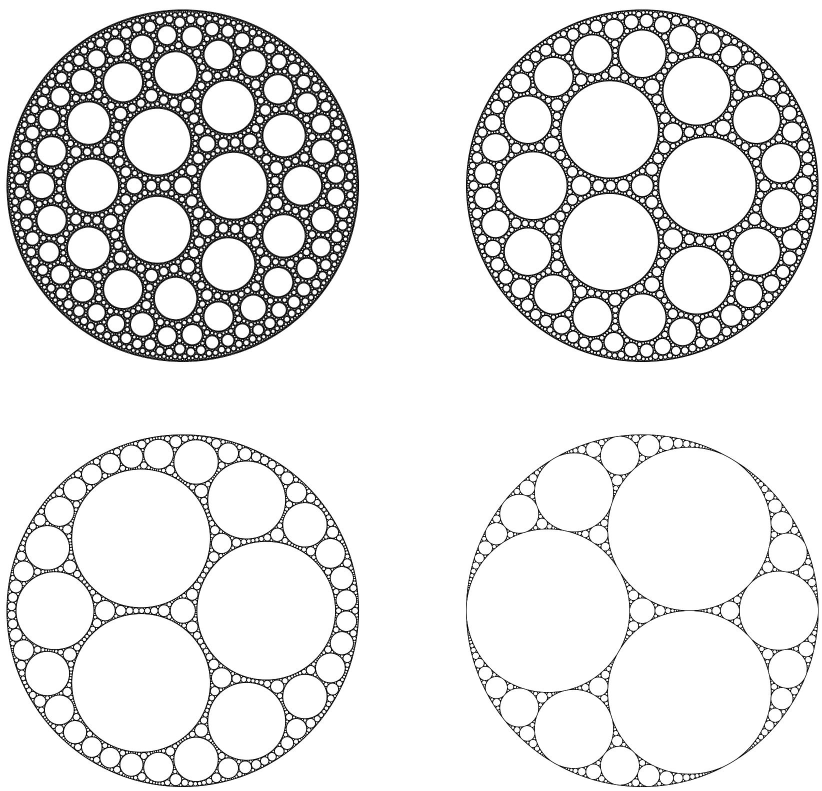

Figure 5. Limit sets from regular tetrahedra with vertices beyond infinity; the cases $n = 7, 8, 12, \infty$.

As $n \to \infty$ the limit sets for these reflection groups tend to the limit set for the Apollonian gasket. See Figure 5 for exaples.

One can glue several copies of K_n, thought of as an orbifold, together along their triangular faces, in a pattern dictated by any 4-valent graph, just as a

3-valent graph gives a gluing pattern for pairs of pants. In this way one obtains many compact, hyperbolic 3-manifolds with $\pi_1(M)$ mapping surjectively to a free group.

1.2 Examples of rational maps

Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a rational map. We will be interested in f as a dynamical system; that is, we will study the iterates

$$f^n(z) = \underbrace{(f \circ f \circ \ldots \circ f)}_{n \text{ times}}(z).$$

Orbits and periodic points. The *forward orbit* of $z \in \widehat{\mathbb{C}}$ is the sequence $\langle z, f(z), f^2(z), \ldots \rangle$. The *backward orbit* is $\bigcup_{n \geq 0} f^{-n}(z)$.

A point $z \in \widehat{\mathbb{C}}$ is *periodic* if $f^n(z) = z$; the least such $n > 0$ is its period.

The *eigenvalue*, *derivative* or *multiplier* of a periodic point is the complex number $\lambda = (f^n)'(z)$. A periodic point is *attracting*, *repelling*, or *indifferent* if its multiplier is < 1, > 1 or $= 1$. A periodic point is *superattracting* if its multiplier is zero. Sometimes attracting is meant to exclude superattracting.

A *cycle* is a finite set cyclically permuted by f, i.e. the forward orbit of a periodic point. A periodic point is superattracting iff its cycle includes a critical point of f.

Conjugacy. Two dynamical systems f_1, f_2 are *conjugate* if $\phi f_1 \phi^{-1} = f_2$. Conjugacies 'preserve the dynamics', e.g. ϕ sends periodic points to periodic points. The quality of ϕ determines the amount of structure of f which is preserved (e.g. topological, measurable, quasiconformal, conformal).

For rational maps, if $\phi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is *conformal*, then f_1 and f_2 are *equivalent* or *isomorphic*. Conformal conjugacy preserves multipliers.

Normal families. A collection of analytic maps $f_\alpha : U \to \widehat{\mathbb{C}}$ is *normal* if every sequence has a convergence subsequence. Here U can be an open subset of $\widehat{\mathbb{C}}$ or more generally a complex manifold.

Theorem 1.6 *Any bounded family of analytic functions is normal. More generally, if $\mathcal{F}$ is not normal then $\bigcup f_\alpha(U)$ is dense in $\widehat{\mathbb{C}}$.*

Proof. First suppose $|f(z)| \leq M$ for all $f \in \mathcal{F}$. Using a chart we reduce to the case $X = \Delta$. By Cauchy's formula, we have

$$|f'(z)| = \left| \frac{1}{2\pi i} \int_{S(z,r)} \frac{f(\zeta)\, d\zeta}{(\zeta - z)^2} \right| \leq \frac{M}{r}$$

for any $r < d(z, \partial\Delta)$. This shows $\mathcal{F}$ is equicontinuous, so by the Arzela-Ascoli theorem $\mathcal{F}$ is normal.

More generally, if $\mathcal{F}$ omits $B(w, r)$, then letting $M(z) = 1/(z - w)$ we see the family

$$M \circ \mathcal{F} = \{M \circ f \ : \ f \in \mathcal{F}\}$$

is bounded by $1/r$, and normality of $\mathcal{F}$ follows from normality of $M \circ \mathcal{F}$. ■

The Julia set. The *Fatou set* $\Omega(f)$ is the largest open set in $\widehat{\mathbb{C}}$ such that the family of iterates $\langle f^n : n \geq 0 \rangle$ forms a normal family when restricted to $\Omega(f)$. Its complement, $J(f) = \widehat{\mathbb{C}} - \Omega(f)$, is the *Julia set.*

Theorem 1.7 *Let $J(f)$ be the Julia set of a rational map f. Then:*

- *The Julia set is closed and totally invariant; that is, $f^{-1}(J(f)) = f(J(f)) = J(f)$.*
- *If an open set U meets $J(f)$ then $\overline{\bigcup f^n(U)} = \widehat{\mathbb{C}}$.*
- *Either $J(f) = \widehat{\mathbb{C}}$ or $J(f)$ is nowhere dense.*

Proof. These assertions follow from the definition of normalization and Theorem 1.6. ■

Order and chaos. Since the iterates $\langle f^n \rangle$ have limit on $\Omega(f)$, orbits of nearby points stay close together. Thus the dynamical behavior of $\langle f^n(z) \rangle$, $z \in \Omega(f)$, is *predictable* — it is not highly sensitive to the exact position of z.

We will show every $z \in \Omega(f)$ is either attracted to a periodic cycle, or lands in a disk or annulus subject to an irrational rotation.

On the other hand, the Julia set is the locus of chaotic behavior, where a small change z can product a vast change in its forward orbit. For example, we will see (Theorem 5.7) the Julia set is the same as the *closure of the repelling periodic points for f.*

Examples of rational maps.

1. *Degree one.* Any Möbius transformation $f(z)$ is hyperbolic, parabolic, elliptic irrational or of finite order.

   In the hyperbolic case up to conjugacy $f(z) = \lambda z$, $|\lambda| > 1$, and all points but $z = 0$ are attracted to infinity. Then $J(f) = \{0, \infty\}$ and
   $$\Omega(f)/f = \mathbb{C}^*/\lambda^{\mathbb{Z}} = \mathbb{C}/(\mathbb{Z}2\pi i \oplus \mathbb{Z}\log\lambda).$$

   In the parabolic case, we can take $f(z) = z + 1$ and then $\Omega(f)/f = \mathbb{C}^*$. This quotient can arise as a limit of degenerating tori, e.g. when $f_n(z) = \lambda_n z + 1$, $|\lambda_n| > 1$ and $\lambda_n \to 1$.

   In the irrational elliptic case, we can put f in the form $f(z) = \exp(2\pi i\theta)z$, $\theta \in \mathbb{R} - \mathbb{Q}$, and the orbits are dense subsets of the circles $|z| = r$.

   In the finite order case, we have $f(z) = \exp(2\pi i p/q)$, and $\widehat{\mathbb{C}}/f$ is the (q, q) orbifold.

2. $f(z) = z^2$. Here $J(f) = S^1$. Clearly $f^n(z) \to 0$ or ∞ when $|z| \neq 1$, and f has a dense set of repelling periodic points on S^1.

3. *Blaschke products.* For $|a| < 1$ let

$$f(z) = z\left(\frac{z+a}{1+\overline{a}z}\right).$$

Theorem 1.8 *The Julia set of $f(z)$ is S^1, the action of f on S^1 is ergodic, and every point outside S^1 is attracted to $z = 0$ or $z = \infty$.*

Proof. Clearly $f(z)$ has an attracting fixed-point of multiplier a at $z = 0$, and since $|f(z)| < |z|$ in the disk we see every $z \in \Delta$ is attracted to the origin under iteration. By symmetry points outside the circle are attracted to infinity. Thus f^n cannot be normal near S^1, so $J(f) = S^1$.

As for ergodicity, let $E \subset S^1$ be a set of positive measure such that $f^{-1}(E) = E$, and let $u : \Delta \to [0,1]$ be the harmonic extension of the indicator function $\chi_E(z)$. Then by invariance, $u(z) = u(f(z))$ and thus

$$u(z) = \lim u(f^n(z)) = u(0)$$

for all $z \in \Delta$. Since u is constant, $E = S^1$. ■

The Blaschke products are strongly reminiscent of Fuchsian groups.

4. *An interval.* Let $f(z) = z^2 - 2$. We claim $J(f) = [-2, 2]$, and every point outside this interval is attracted to infinity.

Indeed, f is a *quotient* of $F(z) = z^2$; setting $p(z) = z + z^{-1}$, we have $f(p(z)) = p(F(z))$. Thus the Julia set is the image of the unit circle, and the rest follows.

5. *A Cantor set.* Let $f(z) = z^2 - 100$. Then the preimages of the critical point $z = 0$ are $z = \pm 10$. The ball $B(0, 20)$ is disjoint from the critical value $z = -100$, so its preimage $f^{-1}(B(0, 20))$ is a pair of disks $D_\pm$ 'centered' at $z = \pm 10$, each of radius about 1, since $|f'(10)| = 20$. The two branches of the inverse of f, mapping $B(0, 20)$ into these disks, are contractions, by a factor of about 20. Thus $J(f)$ is a Cantor set, and every point outside the Cantor set is attracted to infinity.

In this example the fact that $f : (D_- \cup D_+) \to B(0, 20)$ is a covering map is an example of the following basic principle:

> *Let $V \subset \widehat{\mathbb{C}}$ be an open set disjoint from the critical values of f, and let $U = f^{-1}(V)$. Then $f : U \to V$ is a covering map.*

Proof: $f : U \to V$ is a proper local homeomorphism.

6. *Lattès examples.* Here is another example of a quotient dynamical system (like $z^2 - 2$). In this example, $J(f) = \widehat{\mathbb{C}}$.

   Let $X = \mathbb{C}/(\mathbb{Z} \oplus i\mathbb{Z})$ and define $F : X \to X$ by $F(z) = (1+i)z$. Then $|F'| = \sqrt{2}$ everywhere and it is easy to see repelling periodic points of F are dense on X.

   Let $p : X \to \widehat{\mathbb{C}}$ be (essentially) the Weierstrass $\wp$-function, presenting X as a 2-fold cover branched over 0, ± 1 and ∞ (by symmetry). Note that p identifies x and $-x$, so its critical points are the 4 points of order 2 on X (the fixed-points of $x \mapsto -x$). Since $F(-x) = F(x)$, there is a degree 2 rational map $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that

$$\begin{tikzcd} X \arrow[r, "F"] \arrow[d, "p"'] & X \arrow[d, "p"'] \\ \widehat{\mathbb{C}} \arrow[r, "f"] & \widehat{\mathbb{C}} \end{tikzcd}$$

   commutes. From this we find:

$$f(z) = \left(\frac{z-i}{z+i}\right)^2 .$$

   Since repelling points are dense for F, they are also dense for f, and thus $J(f) = \widehat{\mathbb{C}}$.

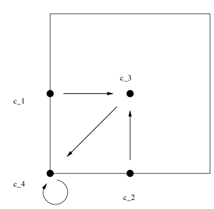

Figure 6. Action of $F(z) = (1+i)z$ on points of order two.

Deriving the formula. Here is how the formula for f was found. The critical points $\{c_1, c_2, c_3, c_4\}$ of p are the points of order 2 on X. Under F, these points map by

$$c_1, c_2 \mapsto c_3 \mapsto c_4 \mapsto c_4.$$

(See Figure 6.)

We can arrange that the 2-fold branched covering $p : X \to \widehat{\mathbb{C}}$ has critical values $e_i = p(c_i)$ with $\{e_1, e_2, e_3, e_4\} = \{0, \infty, 1, -1\}$.

Now $z \in \widehat{\mathbb{C}}$ is a critical value of $f(z) = p \circ F \circ p^{-1}(z)$ exactly when $w = p(z)$ is a critical point of p, but $F^{-1}(w)$ is not. Thus the critical values of f

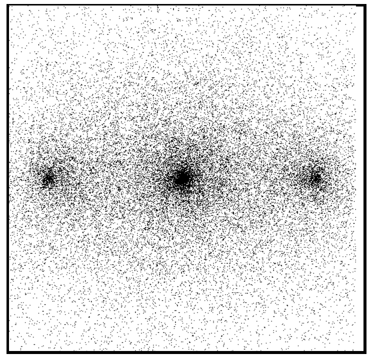

Figure 7. Orbit of Lattés example $f(z) = (z - i)^2/(z + i)^2$.

are $\{p(c_1), p(c_2)\} = \{e_1, e_2\} = \{0, \infty\}$. Therefore $f(z) = M(z)^2$ for some Möbius transformation M.

Since $f(0) = f(\infty) = e_3 = 1$, we have $M(0)^2 = M(\infty)^2 = 1$, so we can write $M(z) = (bz - a)/(bz + a)$. (Of course $M(z)$ and $-M(z)$ give the same map, so M is only determined up to sign.) From $f(1) = -1$ and $f(-1) = -1$ we find $a = i$ and $b = 1$.

Note that F is ergodic on X, preserving the measure $|dz|^2$. Thus f is ergodic on $\widehat{\mathbb{C}}$, with respect to the push-forward of this measure, namely

$$\omega = \frac{|dz|^2}{|z||z-1||z+1|}.$$

Thus the orbits of f concentrate near $\{0, \infty, \pm 1\}$. See Figure 7.

Orbifold picture. The map f has a simple picture in terms of orbifolds. Namely, think of the Riemann sphere $\widehat{\mathbb{C}}$ as the double of a square S. The Euclidean metric on S makes $\widehat{\mathbb{C}}$ into a $(2,2,2,2)$-orbifold X, admitting a symmetry $\iota : X \to X$ obtained by rotating the square 180° around its center. The quotient X/ι also has signature $(2,2,2,2)$, and in the induced Euclidean metric it is similar to X.

Thus we obtain a degree two covering map (of orbifolds) $f : X \to X$, expanding the (singular) Euclidean metric on X by a factor of $\sqrt{2}$. Since f is conformal, it is also a rational map, and f is (conjugate to) the map we have described above.

A mechanical version of this map is well-known in origami: you can fold the corners of a square towards the center to make a smaller square of $1/2$ the original area.

These combinatorial constructions of branched covers are very simple examples of *dessins d'enfants*, cf. [Sn].

7. $f(z) = z^2 - 1$. Here $0 \mapsto -1 \mapsto 0$, so an open set is (super)attracted to a cycle C of order two. What other kinds of behavior can result? For example, can there be another attracting cycle?

   In Figure 8 the points attracted to C are colored gray; the Julia set is black, and $f^n(z) \to \infty$ for z in the white region.

   We will eventually show that whenever $f(z) = z^2 + c$ has an attracting cycle C, all orbits outside $J(f)$ converge to C or to ∞ (Corollary 5.25).

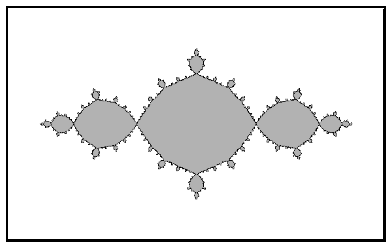

Figure 8. Dynamics of $f(z) = z^2 - 1$.

1.3 Classification of dynamical systems

To put this theory in context, we first mention some general notions in dynamics. Classically dynamics emerges from the theory of differential equations. By taking the flow for a given time, we obtain a diffeomorphism of a manifold, $f : M \to M$. It is also interesting to study groups of maps (e.g. Kleinian groups) and maps that are not invertible (e.g. rational maps with $\deg(f) > 1$.)

Consider the category whose objects are diffeomorphisms of manifolds, $f : M \to M$. These objects can be equipped with various morphisms. For example, we can regard diffeomorphisms $h : M \to N$ such that $h \circ f = g \circ h$ as the morphisms between (M, f) and (N, h). In this category, two objects are isomorphic if they are *smoothly conjugate*.

If we allow h instead to be a homeomorphism, then we obtain the notion of *topological conjugacy*. If we allow h to simply be continuous, then the morphisms are *semiconjugacies*.

Program of dynamics. A central and difficult problem is to classify dynamical systems. Progress can be made in specific families.

Examples.

1. The family of dynamical systems $f_\lambda(x) = \lambda x$, $\lambda \in \mathbb{R}^*$. There are six topological conjugacy classes. A useful observation is that $h(x) = x^\alpha$ conjugates $x \mapsto \lambda x$ to $x \mapsto \lambda^\alpha x$.

2. The family $f_\theta(z) = e^{2\pi i\theta} z$, $|z| = 1$ in $\mathbb{C}$. Here f_θ is topologically conjugate to f_α iff $\alpha = \pm\theta \bmod \mathbb{Z}$. One approach to the proof uses unique ergodicity

of an irrational rotation. Any topological conjugacy h must preserve this measure and so h is an isometry.

In this example there are *uncountably* many topological conjugacy classes.

3. Let $f : S^1 \to S^1$ be a 'generic' C^k diffeomorphism. (Note: C^∞ diffeomorphisms do not form a Baire space.) Then f has rotation number p/q, and the periodic points of f consists of mq pairs of alternating attracting and periodic points (organized into m pairs of cycles). Under iteration, every point converges to one of the m attracting cycles. A model for f is given by taking an mq-fold covering of a hyperbolic Möbius transformation acting on the circle, and composing with a deck transformation. Thus f is determined up to conjugacy its rotation number p/q and the number of attracting cycles m.

Definition. Let M be a compact manifold. A map $f \in \operatorname{Diff}(M)$ is *structurally stable* if there is an open neighborhood U of f such that g is topologically conjugate to f $(f \sim g)$ for all $g \in U$.

By definition, the structurally stable set $\Omega \subset \operatorname{Diff}(M)$ is open. Since $\operatorname{Diff}(M)$ is separable, Ω has at most a countable number of components. Thus there at most a countably number of topological forms for structurally stable dynamical systems.

Program of dynamics. One of the overarching programs in dynamics can be described as follows. Fixing attention on a particular family of dynamical systems, one tries to:

1. Show most maps are structurally stable.
2. Provide models for, and classify topologically, the structurally stable maps.
3. Find continuous moduli to parameterize each component of the structurally stable regime.

Example. This program can be successfully carried out for $\operatorname{Diff}(S^1)$. Namely, (1) the structurally stable maps are indeed dense, (2) the rotation number p/q and number of attracting cycles provide topological invariants, and models are easily constructed; and (3) the eigenvalues at the periodic points provide continuous moduli.

Rational maps. Our main goal is to carry through this program for rational maps $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$. We will find:

1. Structurally stable rational maps are indeed open and dense. Holomorphic motions, the λ-lemma and univalent mappings provide methods of proof.
2. Topological models can be given for rational maps. Indeed, Thurston has characterized rational maps among critically finite branched covers of the sphere. The method here is iteration on Teichmüller space — a toy version of the same discussion that leads to the geometrization of Haken manifolds.

3. Moduli for rational maps can be given in terms of the Teichmüller space of a quotient Riemann surface. Typically these moduli boil down to eigenvalues at attracting cycles, much as in the case of generic diffeomorphisms of S^1.

Expanding maps. We now turn to another example of structural stability. A map $f : S^1 \to S^1$ is *expanding* if there is a $\lambda > 1$ such that $|f'(x)| \geq \lambda > 1$ for all $x \in S^1 \cong \mathbb{R}/\mathbb{Z}$. Then f is a covering map, and $d = \deg(f)$ satisfies $|d| \geq 2$.

Clearly the expanding maps are open in the space of smooth endomorphisms of S^1. We now show that all maps of a given degree are topologically conjugate.

Theorem 1.9 *Any two expanding maps $f : S^1 \to S^1$ of degree d are topologically conjugate.*

Proof. Lifting to the universal cover, we obtain a map $f : \mathbb{R} \to \mathbb{R}$ satisfying $f(x+1) = f(x)+d$. Let $g(x) = dx$. It is enough to construct a homeomorphism h such that $h \circ f = g \circ h$. Since f is expanding, it clearly has a fixed-point, and we can choose coordinates so that $f(0) = 0$. Then $f(k) = dk$ for all $k \in \mathbb{Z}$.

If h exists then it formally satisfies $h = g^{-n} \circ h \circ f^n$. To solve this equation, we replace h on the right with the identity, and then take a limit as $n \to \infty$. More precisely, let

$$h_n(x) = g^{-n} \circ f^n(x) = \frac{f^n(x)}{d^n}.$$

Now if $x \in [k, k+1]$ then $f(x) \in [kd, kd+d]$ and thus $d^{-1}f(x) \in [k, k+1]$ as well. It follows that $|h_{n+1}(x) - h_n(x)| \leq 2d^{-n}$. Thus $h_n(x)$ converges uniformly to a continuous limit $h(x)$. Moreover h is monotone increasing and satisfies $h(x+1) = h(x)+1$, so it descends to a map on the circle.

To complete the proof we must show h is 1-1. Since h is monotone, if it fails to be injective there is a nontrivial interval $I \subset S^1$ such that $h(I)$ is a single point. But since f is expanding, there is an $n > 0$ such that $f^n(I) = S^1$. Then we must also have $g^n(h(I)) = S^1$, which is impossible if I is collapsed by h. Therefore h is a homeomorphism. ■

An analysis of the last step in the proof shows:

1. Even if f is not expanding, it is semiconjugate to $x \mapsto dx$.
2. To get a conjugacy, it suffices that f is LEO (locally eventually onto): for every open interval $I \subset S^1$ there exist an n such that $f^n(I) = S^1$.
3. If f is expanding, then $|h(I)| > C|I|^\alpha$ where $\alpha = \log d / \log \lambda$. Thus the conjugacy and its inverse are Hölder continuous.

Motion of periodic points. An alternative approach to the proof is to consider a family of expanding maps $f_t(z)$ connecting $f(z)$ to $g(z)$. Then one can follow the periodic points along and obtain not just a homeomorphism but an isotopy $h_t(z)$ of conjugacies. This idea leads to the holomorphic motions we use for rational maps.

Failure of structural stability. Although structural stability is dense in $\mathrm{Diff}(S^1)$, it fails to be dense in $\mathrm{Diff}(M)$ for higher-dimensional manifolds. The first result in this direct was [Sm]:

Theorem 1.10 *Structural stability is not dense in* $\mathrm{Diff}(M)$ *for* $M = \mathbb{R}^3/\mathbb{Z}^3$ *the 3-torus.*

Proof. It is known that an Anosov map on the 2-torus $X = \mathbb{R}^2/\mathbb{Z}^2$, such as $f(x) = Ax$ with $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, is structurally stable. Extend f to a map of the 3-torus $M = X \times \mathbb{R}/\mathbb{Z}$ so that f fixes $X_0 = X \times \{0\}$ and is highly contracting normal to X_0. This can be done so that the stable manifolds $W^s(p)$, $p \in X_0$, give a codimension-1 foliation of M.

Next, arrange that f has a hyperbolic fixed-point $q = (0, 0, 1/2)$ that is locally contracting on $X_{1/2}$ and expanding along $(0, 0) \times S^1$. Finally arrange that the unstable manifold $W^u(q)$ has an isolated tangency to $W^s(p)$, some $p \in X_0$.

The entire picture persists for all g in some neighborhood U of f. Let us say g is 'rational' if $W^u(q)$ is tangent to $W^s(p)$ for some periodic point p on X_0. This property is a topological invariant. The set of leaves through periodic points is countable and dense, so the set of rational g is also dense. Similarly, the set of irrational g is dense. Thus structural stability fails throughout U. ■

Shifts. The classical *Smale horseshoe*, with a dense set of saddle points and a basic set isomorphic to the two-side shift, cannot occur in conformal dynamics, even topologically. Indeed the moduli of the two quadrilaterals involved cannot agree. See §2.4.

On the other hand the one-sided shifts (Σ_n, σ) arise frequently, for example inside $z^2 + c$ with $|c|$ large.

Notes. See [AS] and [Wil] for more on the failure of structural stability to be dense. For ideas connecting structural stability and bifurcations to physical sciences and biology, see [Tm].

2 Geometric function theory

This section provides background in geometric methods of complex analysis.

Basic references for Teichmüller theory include [Ah2], [Le].

2.1 The hyperbolic metric

Theorem 2.1 (Schwarz lemma) *Let* $f : (\Delta, 0) \to (\Delta, 0)$ *be holomorphic. Then* $|f'(0)| \le 1$*, and equality holds iff* $f(z) = e^{i\theta} z$.

Proof. Apply the maximum principle to $f(z)/z$. ■

Corollary 2.2 $\mathrm{Aut}(\mathbb{H}) = \mathrm{Isom}^+(\mathbb{H})$, *the isometry group of the hyperbolic metric* $\rho_{\mathbb{H}} = |dz|/y$, $z = x + iy$.

Proof. By the Schwarz lemma, an automorphism fixing the origin in Δ is the restriction of a Möbius transformations. Since the Möbius transformations act transitively on Δ, we see $\mathrm{Aut}(\Delta)$ is the subgroup $PSU(1,1)$ of $\mathrm{Aut}(\widehat{\mathbb{C}})$ preserving the disk.

Similarly, $\mathrm{Aut}(\mathbb{H}) = \mathrm{PSL}_2(\mathbb{R})$. To see that $\mathrm{Aut}(\mathbb{H})$ preserves $\rho_{\mathbb{H}}$, just check invariance under (i) $z \mapsto z+t$, $t \in \mathbb{R}$; (ii) $z \mapsto az$, $a > 0$; and (iii) $z \mapsto 1/z$; and observe these maps generate $\mathrm{PSL}_2(\mathbb{R})$.

Alternatively, note that

$$\rho_\Delta = \frac{2|dz|}{1-|z|^2},$$

so the hyperbolic metric is invariant under the stabilizer S^1 of the origin in $\mathrm{Aut}(\Delta)$; it is also invariant under translation along a geodesic through $z = 0$, since on $\mathbb{H}$ it is invariant $z \mapsto az$; and these two subgroups generate ($\mathrm{PSL}_2(\mathbb{R}) = KAK$). ∎

Corollary 2.3 *Any Riemann surface covered by the disk comes equipped with a canonical metric of constant curvature* -1.

Corollary 2.4 (Schwarz lemma for Riemann surfaces) *Let* $f : X \to Y$ *be a holomorphic map between hyperbolic Riemann surfaces. Then either:*

- f *is a locally isometric covering map, or*
- $\|df\| < 1$ *everywhere, and thus* $d(f(x, fy) < d(x, y)$ *for any pair of distinct points* $x, y \in X$.

Theorem 2.5 *The triply-punctured sphere* $\widehat{\mathbb{C}} - \{0, 1, \infty\}$ *is covered by the disk.*

Proof. Let $T \subset \mathbb{H}$ be the ideal triangle spanning 0, 1 and ∞, and let $f : T \to \mathbb{H}$ be the Riemann mapping normalized to fix 0, 1 and ∞. Using Schwarz reflection through the sides of T in the domain and through the real axis in the range, we can analytically continue f to a covering map $\pi : \mathbb{H} \to \widehat{\mathbb{C}} - \{0, 1, \infty\}$. Thus $\widehat{\mathbb{C}} - \{0, 1, \infty\}$ is isomorphic to $\mathbb{H}/\Gamma(2)$ (the orientation-preserving subgroup of the group of reflections in the sides of T). ∎

Corollary 2.6 (Montel's theorem) *Any family of holomorphic functions omitting 3 fixed values in* $\widehat{\mathbb{C}}$ *is normal.*

Proof. Reduce to the case where $\mathcal{F}$ consists of all maps $f : \Delta \to \mathbb{C} - \{0, 1\}$. Given a sequence $f_n \in \mathcal{F}$, pass to a subsequence such that $f_n(0)$ converges to $z \in \widehat{\mathbb{C}}$. If $z = 0, 1$ or ∞ then f_n converges to a constant map by the Schwarz lemma and completeness of the hyperbolic metric on $\mathbb{C} - \{0, 1\}$. Otherwise we can lift f_n to the universal cover of $\mathbb{C} - \{0, 1\}$, obtaining a sequence $g_n : (\Delta, 0) \to (\Delta, z_n)$, where $z_n \to \widetilde{z}$, a lift of z. Then g_n has a convergent subsequence, so $f_n = \pi \circ g_n$ does too. ∎

Which Riemann surfaces are hyperbolic? It is known that the simply-connected Riemann surfaces are $\widehat{\mathbb{C}}, \mathbb{C}, \mathbb{H}$, and the first two only cover $\mathbb{C}^*$, $\mathbb{C}/\Lambda$ and themselves. All remaining surfaces are hyperbolic.

We will proof a special case:

Theorem 2.7 (Uniformization of planar domains) *Any region $X \subset \mathbb{C}$ whose complement contains 2 or more points is covered by the disk.*

This includes the well-known:

Theorem 2.8 (Riemann mapping theorem) *Any disk $U \subset \mathbb{C}$, $U \neq \mathbb{C}$ is conformally equivalent to the disk.*

Proof of Theorem 2.7. We can arrange by an affine transformation that $X \subset \mathbb{C} - \{0, 1\}$. Pick a basepoint $x \in X$, let $(\widetilde{X}, \widetilde{x})$ be the universal cover of X, and consider the family of maps

$$\mathcal{F} = \{f : (\widetilde{X}, \widetilde{x}) \to (\Delta, 0) \ : \ f \text{ is a covering map to its image}\}.$$

To see $\mathcal{F}$ is nonempty, lift the inclusion $X \subset \mathbb{C} - \{0, 1\}$ to a map between the universal covers, $f : \widetilde{X} \to \Delta$; the image is the covering space of X corresponding to the kernel of $\pi_1(X) \to \pi_1(\mathbb{C} - \{0, 1\})$, and f is a covering map to this image.

Now take a sequence $f_n \in \mathcal{F}$ such that $|f_n'(\widetilde{x})|$ tends to its supremum over $\mathcal{F}$. Since Δ is bounded, f_n is a normal family, and we can take a subsequence $f_n \to f$. It is not hard to check that $f \in \mathcal{F}$.

We claim $f : \widetilde{X} \to \Delta$ is surjective. Indeed, if the image omits a value z, then we can choose a proper degree two map $B : (\Delta, 0) \to (\Delta, 0)$ branched over z. By the Schwarz lemma, $|B'(0)| < 1$; on the other hand, since $z \notin f(\widetilde{X})$, the map $B^{-1} \circ f$ admits a single-valued branched $g : (\widetilde{X}, \widetilde{x}) \to (\Delta, 0)$, we have $g \in \mathcal{F}$, and $|g'(0)| > |f'(0)|$, contrary to the construction of f.

Thus f is surjective; but since it is a covering map, it is an isomorphism. ■

Remark. To prevent cyclic reasoning, one should first prove the Riemann mapping theorem for simply-connected domains, then use it to uniformize $\widehat{\mathbb{C}} - \{0, 1, \infty\}$, and finally extend the uniformization theorem to all planar domains as above.

2.2 Extremal length

Let X be a Riemann surface. A *Borel metric* ρ on X is locally of the form $\rho(z)|dz|$ where $\rho \geq 0$ is a Borel measurable function. If γ is a rectifiable path on X, then its ρ-length is defined by

$$\ell_\rho(\gamma) = \int_\gamma \rho(z)\, |dz|.$$

Similarly the ρ-area of X is given by

$$\mathrm{area}_\rho(X) = \int_X \rho(z)^2\, |dz|^2.$$

Now let Γ be a collection of paths on X. Setting

$$\ell_\rho(\Gamma) = \inf_\Gamma \ell_\rho(\gamma),$$

we define the *extremal length* of Γ by

$$\lambda(\Gamma) = \sup_\rho \frac{\ell_\rho(\Gamma)^2}{\mathrm{area}_\rho(X)}. \tag{2.1}$$

The supremum is take over all Borel metrics of finite area. Clearly $\lambda(\Gamma)$ is a conformal invariant of the pair (X, Γ).

Beurling's criterion for an extremal metric. (See [Ah2, §4.7].) A metric is *extremal* if it realizes the supremum in (2.1).

Theorem 2.9 *Suppose the measure ρ^2 on X lies in the closed convex hull of the measures*

$$\{\rho|\gamma \,:\, \gamma \in \Gamma \text{ and } \ell_\rho(\gamma) = \ell_\rho(\Gamma)\}.$$

Then ρ is extremal for Γ.

Proof. Consider any other Borel metric α. We may assume both α and ρ are normalized to give X area 1.

Now for any $\gamma \in \Gamma$ we have

$$\ell_\alpha(\Gamma) \le \int_\gamma \alpha = \int_\gamma (\alpha/\rho)\rho = \langle \alpha/\rho, \rho|\gamma\rangle$$

where the last expression is the pairing between functions and measures. Since the probability measure ρ^2 is a convex combination of the probability measures $(\rho|\gamma)/\ell_\rho(\Gamma)$, we have

$$\frac{\ell_\alpha(\Gamma)}{\ell_\rho(\Gamma)} \le \langle \alpha/\rho, \rho^2\rangle = \int_X \alpha\rho \le \left(\int \alpha^2 \int \rho^2\right)^{1/2} = 1$$

by Cauchy-Schwarz. Thus $\ell_\alpha(\Gamma) \le \ell_\rho(\Gamma)$, and therefore ρ maximizes the ratio (2.1) of length-squared to area. ■

Examples.

1. A *quadrilateral* $Q \subset \mathbb{C}$ is a Jordan domain with 4 marked points on its boundary, and a distinguished pair of opposite edges. We let Q^* denote the same quadrilateral with the other pair of edges distinguished.

Any quadrilateral is conformally equivalent to a unique Euclidean quadrilateral $Q(a) = [0, a] \times [0, 1]$, with the sides of unit length distinguished. By definition, the *modulus* of Q is a.

Notice that $\mathrm{mod}(Q^*) = 1/\mathrm{mod}(Q)$.

Let $\Gamma(Q)$ be the set of all paths joining the distinguished sides. Since the Euclidean metric on $Q(a)$ satisfies Beurling's criterion (the geodesics are horizontal segments), we find

$$\lambda(\Gamma(Q)) = \mathrm{mod}(Q).$$

By considering any specific metric ρ on Q, we obtain lower bounds on $\mathrm{mod}(Q)$ by considering $\lambda(\Gamma(Q))$ and $\lambda(\Gamma(Q^*))$. Thus one has a powerful method for *estimating* the modulus of a quadrilateral.

Example: let $Q_0 = [0, 1] \times [0, 1]$ with the vertical sides distinguished; we have $\mathrm{mod}(Q_0) = 1$. Now construct Q by adding a 'roof' to the house, i.e. adding a right isosceles triangle of hypotenuse 1 to the top of the square. Then $\rho = |dz|$ on Q gives $\mathrm{mod}(Q) \geq 2/3$ (the area has increased to $3/2$), while $\rho = |dz|$ restricted to Q_0 gives $\mathrm{mod}(Q^*) \geq 1$. Thus

$$2/3 < \mathrm{mod}(Q) < 1;$$

in particular, adding the roof brings the walls closer together.

2. An annulus. Any Riemann surface with $\pi_1(X) = \mathbb{Z}$ is isomorphic to $\mathbb{C}^*$, Δ^* or
$$A(R) = \{z \,:\, 1 < |z| < R\}$$
for a unique R. In the last case we define $\mathrm{mod}(X) = \log(R)/(2\pi)$; and by convention, $\mathrm{mod}(X) = \infty$ in the first two cases.

   Fixing R, let Γ be the family of 'topological radii', that is curves joining the two boundary components of $A(R)$. These curves are geodesics for the cylindrical metric $\rho = |dz|/|z|$, which is extremal by Beurling's criterion. Thus
$$\lambda(\Gamma) = \frac{\log(R)}{2\pi} = \mathrm{mod}(A(R)).$$

   Letting Γ^* denote the family of simple essential loops in $A(R)$, we find $\lambda(\Gamma^*) = 1/\lambda(\Gamma) = 1/\mathrm{mod}(A(R))$.

   Thus any metric ρ on an annular Riemann surface X yields upper and lower bounds for $\mathrm{mod}(X)$, in terms of the ρ-area of X and the shortest curves in Γ and Γ^*.

3. *The real projective plane.* Let $X = \mathbb{RP}^2$. Then X is canonically a Riemann surface, in the sense that its universal cover S^2 has a unique conformal structure, and this structure is preserved (up to orientation) by the deck transformations of S^2/X.

Let Γ be the family of all loops generating $\pi_1(X) \cong \mathbb{Z}/2$. We claim $\lambda(\Gamma) = \pi/2$.

To see this, let ρ be the round metric on X, making its universal cover S^2 into the sphere of radius 1. If we average linear measure on a great circle over the rotation group of S^2, we obtain an invariant measure which must be the usual area form. Thus ρ satisfies Beurling's criterion. The minimal length of a curve joining antipodal points on S^2 is π, and the area of $\mathbb{RP}^2$ is 2π, so $\lambda(\Gamma) = \pi^2/(2\pi) = \pi/2$.

4. Simple curves on a torus. Consider the torus $X = \mathbb{C}/\mathbb{Z} \oplus \mathbb{Z}\tau$, $\operatorname{Im}\tau > 0$. Let Γ be the family of loops on X in the homotopy class of $[0, 1]$. By Beurling's criterion, the flat metric $|dz|$ on X is extremal. The area of X is $\operatorname{Im}\tau$ and the length of the geodesic $[0, 1]$ is 1, so we find

$$\lambda(\Gamma) = \frac{1}{\operatorname{Im}\tau}.$$

Remark. For completeness we recall a direct proof that the quadrilaterals $Q(a)$ and $Q(b)$ are conformally isomorphic iff $a = b$. Namely, given a conformal map $f : Q(a) \to Q(b)$, one can develop f by Schwarz reflection through the sides of $Q(a)$ and $Q(b)$ to obtain an automorphism $F : \mathbb{C} \to \mathbb{C}$, which must be of the form $F(z) = \alpha z + \beta$. Since f fixes $[0, 1]i$, F is the identity.

A similar argument shows the annuli $A(R)$ and $A(S)$ are conformally isomorphic iff $R = S$.

2.3 Extremal length and quasiconformal mappings

Theorem 2.10 *If Γ and Γ' are related by a K-quasiconformal mapping, then*

$$\frac{1}{K}\lambda(\Gamma') \le \lambda(\Gamma) \le K\lambda(\Gamma').$$

Proof. Let $f : X \to X'$ be a quasiconformal mapping sending Γ to Γ'. For each conformal metric ρ on the domain, we get a metric $\rho' = f_*(\rho)$ on X' with the same lengths and areas as ρ; in particular, $\operatorname{area}_{\rho'}(X') = \operatorname{area}_\rho(X)$ and $\ell_{\rho'}(\Gamma') = \ell_\rho(\Gamma)$.

However ρ' is generally not conformal. To make it conformal while we expand its infinitesimal unit balls (which are ellipses of oblateness at most K) to round balls. This does not decrease the ρ'-length of Γ', and it increases the ρ'-area of X by at most K. Thus $\lambda(\Gamma') \ge \lambda(\Gamma)/K$, and the Theorem follows by symmetry. ∎

Cf. [LV, §IV.3.3].

Corollary 2.11 *If $f : X \to Y$ is a K-quasiconformal map, then* $\operatorname{mod} f(Q) \le K \operatorname{mod} Q$ *for every quadrilateral Q on X.*

In fact the converse is true: a map which distorts the modulus of every quadrilateral by at most K is K-quasiconformal. The converse is clear for linear maps, so it follows easily for smooth quasiconformal maps. The general case depends on the a.e. differentiability of quasiconformal maps. Using this fact, we have:

Theorem 2.12 *Any homeomorphism f which is a uniform limit of K-quasiconformal mappings is itself K-quasiconformal.*

2.4 Aside: the Smale horseshoe

To study celestial mechanics, Poincaré investigated the Hamiltonian flow on phase space corresponding to time evolution of the planets. Since the orbit of a particular planet is approximately periodic, it is natural to take a transversal M to the flow (e.g. the configurations of the sun-earth-moon system as the earth passes through a window transverse to its orbit), and study the *first return map* $f : M \to M$.

The map f is volume-preserving. Indeed, the symplectic volume form $\Omega = \omega^n$ on phase space is preserved by the Hamiltonian vector field v, as is v, and thus the $2n - 1$ form

$$\alpha = i_v(\Omega)$$

is also preserved by the flow. Since α vanishes any hypersurface tangent to the flow, the volume form $\alpha|M$ is f-invariant.

The existence of very complicated behavior in these dynamical systems was known to Poincaré and Birkhoff.

Smale provided a simple picture, the horseshoe, that sums up in an immediate and geometric form a mechanism leading to infinitely many periodic cycles. Namely a square S is stretched to a long, thin rectangle, then laced through itself as in Figure 9. The thick edges of S map to the thick edges of $f(S)$.

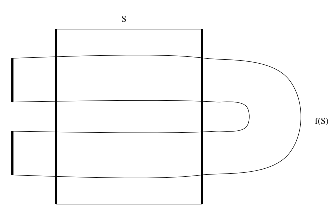

Figure 9. The Smale horseshoe.

Within the two rectangles forming $f(S) \cap S$, one finds a totally invariant Cantor set E. The dynamics of $f|E$ is conjugate to the action of $\mathbb{Z}$ on the shift space $(\mathbb{Z}/2)^{\mathbb{Z}}$ of all functions $\phi : \mathbb{Z} \to \mathbb{Z}/2$.

Note that f can be realized as an area-preserving map. But by basic properties of extremal length, the modulus of $f(S)$ (the extremal length of the paths joining the thick sides) is greater than the modulus of S. Thus f cannot be made conformal, and the horseshoe does not occur in conformal dynamics!

On the other hand, by flattening the horseshoe we obtain a 2-to-1 map of an interval over itself, the critical point being the 'bend' in the horseshoe. As the horseshoe is created (in a family of diffeomorphisms), the bend pushes through the square, much as the critical point pushes through the interval. Thus understanding the dynamics of $f_c(x) = x^2 + c$ is a natural prerequisite to understanding the bifurcations leading to a horseshoe, and one is back to complex dynamics again.

2.5 The Ahlfors-Weill extension

In this section we discuss the extension of conformal maps $f : \mathbb{H} \to \widehat{\mathbb{C}}$ to quasiconformal maps on the whole Riemann sphere. This extension is useful for Teichmüller theory and holomorphic motions, as well as the theory of structural stability for rational maps and Kleinian groups.

Norms. The natural L^p-norm on holomorphic quadratic differentials on a hyperbolic Riemann surface X is given by

$$\|\phi\|_p = \left(\int_X \rho^{2-2p}|\phi|^p\right)^{1/p}, \qquad \|\phi\|_\infty = \sup_X \frac{|\phi|}{\rho^2},$$

where $\rho = \rho(z)\,|dz|$ denotes the hyperbolic metric. Note that the L^1 norm does not involve ρ at all. Similarly for Beltrami differentials we have

$$\|\mu\|_p \;=\; \left(\int_X \rho^2|\mu|^p\right)^{1/p}, \qquad \|\mu\|_\infty \;=\; \sup_X |\mu|.$$

A Beltrami differential μ on X is *harmonic* if $\mu = \overline{\phi}/\rho^2$ for some holomorphic quadratic differential ϕ. This means that μ formally minimizes the L^2 norm among equivalent differentials $\mu + \overline{\partial} v$. Note that $\|\mu\|_\infty = \|\phi\|_\infty$.

Theorem 2.13 *Let $f : \mathbb{H} \to \widehat{\mathbb{C}}$ be a holomorphic map, and suppose $\|Sf\|_\infty < 1/2$. Then there exists a unique extension of f to a quasiconformal map on the whole Riemann sphere such that the dilatation μ of f on the lower halfplane is a harmonic Beltrami differential. In fact we have $\mu(z) = -2y^2\phi(\overline{z})$.*

Proof. Here is a quick construction of the map g on the lower halfplane extending f. Given $w \in \mathbb{H}$, let $M_w(z)$ be the unique Möbius transformation whose 2-jet at $z = w$ matches the 2-jet of $f(z)$ at $z = w$. Then set $g(z) = M_{\overline{z}}(z)$.

To verify that g has the required properties, it is useful to know how to reconstruct f from its Schwarzian derivative $\phi = (f''/f')' - (1/2)(f''/f')^2$. To do this, one can begin with two linearly independent solutions f_1, f_2 to the differential equation

$$y'' + (1/2)\phi y = 0.$$

After scaling one can assume the Wronskian $f_1 f_2' - f_2 f_1' = 1$. From these solutions we obtain a map $z \mapsto (f_1(z), f_2(z))$ from $\mathbb{H}$ into $\mathbb{C}^2$. Projectivizing, we obtain a map $f : \mathbb{H} \to \widehat{\mathbb{C}}$ given by $f = f_2/f_1$.

Because of our normalization of the Wronskian, this map satisfies

$$\begin{aligned} f' &= (f_1 f_2' - f_2 f_1')/f_1^2 = 1/f_1^2, \\ f'' &= -2f_1'/f_1^3, \quad \text{and} \\ (f''/f') = (\log f')' &= -2f_1'/f_1. \end{aligned}$$

From this we find

$$Sf = -2(f_1 f_1'' - (f_1')^2)/f_1^2 - 2(f_1')^2/f_1^2 = -2f_1''/f_1^2 = \phi.$$

The choice of basis for the space of solutions of the different equation accounts for the fact that Sf only determines f up to a fractional linear transformation.

Once we have $f = f_1/f_2$ it is easy to see that

$$M_z(z+\epsilon) = \frac{f_2(z) + \epsilon f_2'(z)}{f_1(z) + \epsilon f_1'(z)}. \tag{2.2}$$

To check this one simply examines the power series in ϵ for the expression above, up to the term ϵ^2, and compares the terms to (f, f', f'') computed above.

Thus the extension of f is given by

$$g(z) = \frac{f_2(\overline{z}) + \epsilon f_2'(\overline{z})}{f_1(\overline{z}) + \epsilon f_1'(\overline{z})},$$

with $\epsilon = z - \overline{z} = 2iy$, if $z = x + iy$. To compute the Beltrami differential $\overline{\partial} g/\partial g$, we first observe that both terms with involve the square of the denominator of the expression above. As for the numerator, for ∂g it is simply

$$(f_1 + \epsilon f_1')f_2' - (f_2 + \epsilon f_2')f_1' = f_1 f_2' - f_2 f_1' = 1.$$

For $\overline{\partial} g$ we first note that:

$$\overline{\partial}(f_i + \epsilon f_i') = f_i' + \epsilon f_i'' - f_i' = -(1/2)\epsilon\phi f_i.$$

Thus its numerator is given by

$$(-1/2)\epsilon\phi((f_1 + \epsilon f_1')f_2 - (f_2 + \epsilon f_2')f_1) = (1/2)\epsilon^2\phi.$$

So altogether the Beltrami differential of g is given by

$$\mu(z) = -2y^2\phi(\overline{z}).$$

Since $\overline{\phi}(\overline{z})$ is a holomorphic quadratic differential, we have shown that μ is a harmonic Beltrami differential.

Finally we verify uniqueness. Suppose we have another extension G to the lower halfplane with $\mu(G) = -2\rho^{-2}\overline{\psi}$ harmonic. Then G arises as the Ahlfors-Weill extension of a univalent map $F : \mathbb{H} \to \widehat{\mathbb{C}}$ with $SF = \psi$. By uniqueness of the solution to the Beltrami equation (up to a Möbius transformation), we have $F = f$ and thus $\psi = \phi$, which implies $g = G$. ■

Corollary 2.14 *A conformal immersion $f : \mathbb{H} \to \widehat{\mathbb{C}}$ is univalent if $\|Sf\| \le 1/2$.*

Area theorem. Let $f(z) = z + \sum_1^\infty b_n z^{-n}$ be a univalent map on $\widehat{\mathbb{C}} - \overline{\Delta} = \{z : |z| > 1\}$. Then f sends the outside of the unit disk to the outside of a full, compact set $K_f \subset \mathbb{C}$ of capacity one. One can compute the area of K_f directly from the coefficients of f: namely we have

$$\text{area}(K_f) = \pi(1 - \sum n|b_n|^2).$$

In particular we have:

Proposition 2.15 (Area theorem) *If $f(z) = z + \sum b_n z^{-n}$ is univalent, then $\sum n|b_n|^2 \le 1$.*

Using the area theorem one can prove:

Corollary 2.16 (Nehari) *If $f : \mathbb{H} \to \widehat{\mathbb{C}}$ is univalent, then $\|Sf\| \le 3/2$.*

Geometry of the Schwarzian derivative. The Schwarzian derivative can also be interpreted as the rate of change of the osculating Möbius transformation M_z, and as the curvature of a surface in $\mathbb{H}^3$ naturally associated to $f : \mathbb{H} \to \widehat{\mathbb{C}}$. See [Th], [Ep1], [Ep2].

Let $\pi : \mathbb{H}^3 \to \mathbb{H}$ be the 'nearest point' projection, obtained by following the normals to the hyperplane spanned by $\widehat{\mathbb{R}}$. Then the Ahlfors-Weill map can be further prolonged to a map

$$F : \overline{\mathbb{H}}^3 \to \overline{\mathbb{H}}^3$$

by $F(p) = M_{\pi(p)}(p)$. This map is a diffeomorphism and a quasi-isometry when $\|Sf\| < 1/2$.

On the other hand, it is easy to see that if $Q \subset \widehat{\mathbb{C}}$ is a K-quasicircle with K near 1, then $\|Sf\| < 1/2$ for the Riemann map to one side of Q. Thus any 'mild' quasicircle can be realized as $Q = F(\widehat{\mathbb{R}})$ for a *canonical* map of the closed hyperbolic ball to itself (unique up to pre-composition with an isometry stabilizing $\widehat{\mathbb{R}}$ and preserving its orientation).

3 Teichmüller theory via geometery

In this section we discuss Teichmüller space from the perspective of hyperbolic geometry.

3.1 Teichmüller space

Let S be a closed, oriented surface of genus $g \ge 2$. A *marked hyperbolic surface* is a pair (ϕ, X) consisting of an oriented compact hyperbolic surface $X \cong \mathbb{H}/\Gamma$ and an orientation-preserving homeomorphism $\phi : S \to X$.

Two marked surfaces (ϕ_i, X_i), $i = 1, 2$ are *equivalent* if there exists an isometry $\alpha : X_1 \to X_2$ such that $\phi_2^{-1} \circ \alpha \circ \phi_1 = \psi$ is isotopic to the identity.

The space of such equivalence classes is the *Teichmüller space* $\mathcal{T}_g = \mathrm{Teich}(S)$.

For any essential closed loop $\alpha \subset S$, there is a unique closed geodesic $\overline{\alpha} \subset X$ freely isotopic to $\phi(\alpha)$. We denote its hyperbolic length by $L_\alpha(X)$. We give Teichmüller space the weakest topology which makes all such length functions continuous.

Representations. Since each surface in Teichmüller space is equipped with a complete hyperbolic metric, from $\phi : S \to X = \mathbb{H}/\Gamma$ we obtain a homomorphism

$$\rho : \pi_1(S) \to \Gamma \subset \mathrm{Isom}^+(\mathbb{H}) = \mathrm{PSL}_2(\mathbb{R}),$$

well-defined up to conjugacy. Thus Teichmüller space admits an embedding

$$\mathrm{Teich}(S) \hookrightarrow \mathrm{Hom}(\pi_1(S), \mathrm{PSL}_2(\mathbb{R}))/(\text{conjugation}).$$

Since traces recover lengths, the topology induced by this embedding is the same as that defined above.

Mapping-class group. We let $\mathrm{Mod}(S)$ denote the group of orientation-preserving homeomorphisms $\psi : S \to S$, modulo those isotopic (equivalently, homotopic) to the identity. It acts on $\mathrm{Teich}(S)$ by $\psi \cdot (\phi, X) = (\phi \circ \psi^{-1}, X)$. The quotient space is the *moduli space*

$$\mathcal{M}_g = \mathcal{M}(S) = \mathrm{Teich}(S)/\mathrm{Mod}(S).$$

3.2 Fenchel-Nielsen coordinates

Theorem 3.1 *The Teichmüller space* $\mathrm{Teich}(S)$ *is homeomorphic to a ball* B^n*, and* $n = \dim_{\mathbb{R}} Q(X)$ *for any* $X \in \mathrm{Teich}(S)$.

Proof. (Fenchel-Nielsen) Suppose S is a closed surface of genus $g \geq 2$. Then S can be decomposed along $3g-3$ simple closed curves into pairs of pants. For any $X \in \mathrm{Teich}(S)$, these curves are canonically represented by geodesics, whose lengths determine each pair of pants up to isometry. To recover X in addition we need a twist parameter when gluing pants together. Thus altogether $\mathrm{Teich}(S)$ is parameterized by $\mathbb{R}_+^{3g-3} \times \mathbb{R}^{3g-3}$, and $6g - 6 = \dim Q(X)$.

The case of surfaces with boundary or of smaller genus is similar. ■

Construction of pants and triangles. Note that the construction of a pair of pants with cuffs of length $2L_i$, $i = 1, 2, 3$ is tantamount to the construction of three disjoint geodesics in $\mathbb{H}^2$ with $d(\gamma_i, \gamma_{i+1}) = L_{i+2}$. Now an (oriented) geodesic, in the Minkowski model $\mathbb{R}^{2,1}$, corresponds to a vector with $v^2 = 1$, and the oriented distance between two them satisfies $-\cosh d(\gamma_1, \gamma_2) = \langle v_1, v_2 \rangle$. Thus our pair of pants corresponds to a basis for $\mathbb{R}^{2,1}$ with the quadratic form

$$B = \begin{pmatrix} 1 & -\cosh L_3 & -\cosh L_2 \\ -\cosh L_3 & 1 & -\cosh L_1 \\ -\cosh L_2 & -\cosh L_1 & 1 \end{pmatrix}.$$

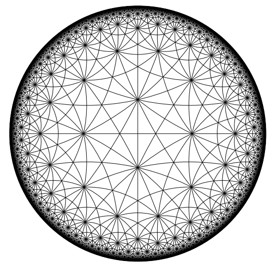

Figure 10. Tiling by $(2,3,7)$ triangles

It is readily verified that B has signature $(2,1)$, for all choices of (L_i); this shows such a configuration exists and is unique up to isometry.

Similarly, to construct an (α,β,γ) triangle, one can use the form

$$B = \begin{pmatrix} 1 & -\cos\gamma & -\cos\beta \\ -\cos\gamma & 1 & -\cos\alpha \\ -\cos\beta & -\cos\alpha & 1 \end{pmatrix}.$$

Note that

$$\det B = 1 - 2\cos\alpha\cos\beta\cos\gamma + \cos^2\alpha + \cos^2\beta + \cos^2\gamma = 0$$

iff $S = \alpha+\beta+\gamma = \pi$. If $S < \pi$ then B has signature $(2,1)$, while if $S > \pi$ then it has signature $(3,0)$. (Consider the extreme cases where all angles are 0 or all are $\pi/2$. Note that when all angles go to π the determinant goes to 0, since the vertices become colinear on the sphere.)

As another example, suppose we wish to construct $(2,p,q)$ triangle in $\mathbb{H}^2 \cong \Delta$. We can arrange that the x and y axes form two of the sides, i.e $v_1 = e_1$ and $v_2 = e_2$ in $\mathbb{R}^{2,1}$. Then the third side is a circle centered at $(x/z, y/z)$ in $\mathbb{C}$, where $v_3 = (x,y,z)$ satisfies $v_3^2 = x^2 + y^2 - z^2 = 1$ and $\langle v_3, e_1\rangle = x = \cos\pi/p$, $\langle v_3, e_2\rangle = y = \cos\pi/q$. This allows one to easily locate the center; moreover, the radius satisfies $1/r^2 = z^2 = x^2 + y^2 - 1$.

The case of a $(2,3,7)$ triangle is shown in Figure 10.

Trivalent graphs. We remark that up to homeomorphism, there are only finitely many decomposition of a surface of genus g into pairs of pants, and these decompositions correspond to trivalent graphs with g loops (first Betti number g).

Limits of pants. Interesting geometric limits can arise as the lengths (L_1, L_2, L_3) of the cuffs of a pair of pants P tend to infinity. There are 3 basic cases: $(0,0,\infty)$: P splits into a pair of punctured monogons. (∞,∞,∞): P collapses (after rescaling) to a trivalent graph, or (before rescaling) to a pair of ideal

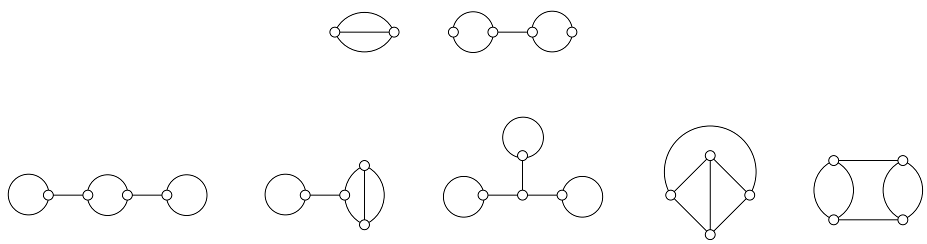

Figure 11. The trivalent graphs of genus 2 and 3.

triangles. $(0, \infty, \infty)$: P collapses to a punctured bigon, which has a nontrivial modulus (it is bounded by a pair of geodesics from 1 to $\tau \in S^1$ in Δ^*).

Symplectic structure. Fixing a pair of pants decomposition P of Σ_g, we obtain twist and length coordinates (ℓ_i, τ_i) for $\mathcal{T}_g$. (Note that the twist parameter is measured in units of length, not angle.) In these coordinates, a Dehn twist around the ith element of P acts by $\tau_i \mapsto \tau_i + \ell_i$. An important result is that the symplectic form

$$\omega = \sum_i d\ell_i \wedge d\tau_i$$

is invariant under the full mapping-class group. It coincides with the symplectic form on twisted cohomology coming from the intersection pairing on $H^1(\Sigma, \mathrm{sl}_2(\mathbb{R})_\rho)$.

Case of genus one. In the case of a torus, normalize to have total area 1, we have just 2 coordinates ℓ and τ, giving the lattice $\mathbb{Z}(\ell, 0) \oplus \mathbb{Z}(\tau, 1/\ell)$. In terms of $\tau = x + iy \in \mathbb{H}$, this lattice has coordinates $(x, y) = (\tau/\ell, 1/\ell^2)$. We then find that $d\ell d\tau$ is a constant multiple of the hyperbolic area form $dxdy/y^2$.

Characteristic classes. It is known that the Weil-Petersson symplectic form has the property that $[\omega/\pi^2]$ generates $H^2(\mathcal{M}_g, \mathbb{Q})$. Consequently the Weil-Petersson volume of moduli space is a rational multiple of π^{6g-6} [Wol1]. It can also be shown that $[\omega/2\pi^2] = \kappa_1$ as a class in $H^2(\overline{\mathcal{M}}_g, \mathbb{Q})$, where $\kappa_1 = \pi_*(c_1(V)^2)$ is the pushfoward of the square of first Chern class of the relative tangent bundle to the universal curve $\pi : \overline{\mathcal{C}}_g \to \overline{\mathcal{M}}_g$; see [Wol2].

3.3 Geodesic currents

The circle at infinity. Fix $g \geq 2$ and let $X, Y \in \mathcal{T}_g$ be a pair of marked hyperbolic surfaces. Then there is a unique homotopy class of homeomorphism $f : X \to Y$ compatible with markings.

Theorem 3.2 *The lift $\widetilde{f} : \mathbb{H} \to \mathbb{H}$ of f to the universal covers of X and Y extends to a homeomorphism of S^1_∞. This extension depends only on the homotopy class of f.*

Thus the circles at infinity for all points in $\mathcal{T}_g$ are canonically identified. This circle, together with an action of $\pi_1(Z_g)$ on it, can be constructed topologically by taking the metric completion of a conformal rescaling of the word metric on $\pi_1(Z_g)$.

For concreteness, we fix a particular $X = \mathbb{H}/\Gamma$ in $\mathcal{T}_g$, but observe that the topological dynamics of (Γ, S^1) is independent of X.

Theorem 3.3 *Γ has dense orbits on $S^1 \times S^1$.*

Proof. By ergodicity, there exists a dense geodesic γ in T_1X. Thus there exists a geodesic $\widetilde{\gamma}$ in $\mathbb{H}$, with endpoints (a, b), whose Γ-orbit is dense in $T_1(\mathbb{H})$. Consequently $\Gamma \cdot (a, b)$ is dense in $S^1 \times S^1$. ∎

On the other hand, 'periodic orbits' are also dense:

Theorem 3.4 *The closed geodesics on X correspond to a dense subset of $S^1 \times S^1$.*

The space of currents. Let $\mathcal{G} = (S^1 \times S^1 - \Delta)/(\mathbb{Z}/2)$ denote the Möbius band forming the space of unoriented geodesics in $\mathbb{H}$. Then $\mathcal{G}/\Gamma$ is the space of geodesics on X. (Since Γ has a dense orbit in $\mathcal{G}$, this is a 'quantum space' in the sense of Connes; nevertheless it carries many interesting closed sets and measures.)

A *geodesic current* is a locally finite, Γ-invariant measure μ on $\mathcal{G}$. *We require that* $\mu \neq 0$. The space of all geodesic currents on X will be denoted $\mathcal{C}(X)$.

Examples of currents. (i) An unoriented closed geodesic $\gamma \subset X$ gives a point $p_\gamma \in \mathcal{G}$ such that $\Gamma \cdot p$ is discrete. Thus we can form an invariant measure by putting a δ mass on each point in its orbit. This is the current $[\gamma]$ associated to a closed loop.

(ii) A positive weighted sum of closed curves $\sum C_i \gamma_i$ determines a current $\sum C_i[\gamma_i]$.

For simplicity we suppressed the brackets in the future.

Integral geometry. The *integral geometric* measure on $\mathcal{G}$ is defined as follows. Fix an *oriented* geodesic $\gamma \subset \mathbb{H}$. Then there is an injective map $\gamma \times S^1 \to \mathcal{G}$ sending (x, θ) to the unique unoriented geodesic $\delta(x, \theta)$ through x making angle θ with γ. (Since the geodesics are unoriented, θ ranges in $[0, \pi]$.) The measure

$$\mu = (1/2) \sin \theta d\theta \, dx$$

can be shown to be independent of the choice of chart. It is characterized by the following property.

Theorem 3.5 *For any geodesic segment $S \subset \mathbb{H}$, the measure of the set of geodesics $\gamma \in \mathcal{G}$ meeting S is equal to $\ell(S)$.*

Proof. We have $\int_0^\pi (1/2) \sin \theta \, d\theta = 1$. ∎

Clearly μ is invariant under all isometries of $\mathbb{H}$, so it defines an element $L_X \in \mathcal{C}(X)$. This *Liouville current* depends very much on X; in fact, as we will see below, the natural identification $\mathcal{C}(X) = \mathcal{C}(Y)$ sends L_X to L_Y iff $X = Y$ in $\mathcal{T}_g$.

Constructing $T_1(X)$. Now let $\Delta \subset \mathcal{G} \times S^1$ be the set of pairs (γ, p) where p is an endpoint of γ. Then we have a natural isomorphism

$$(\mathcal{G} \times S^1 - \Delta) \cong T_1(\mathbb{H}).$$

Indeed, the orthogonal projection of p to γ determines a point $x \in \mathbb{H}$, and there is a unique unit vector $v \in T_x(\mathbb{H})$ tangent to γ, oriented so that p lies to the right. The foliation of $T_1(\mathbb{H})$ by geodesics is obtained simply by varying the S^1 factor p in this product. Taking the quotient by Γ, we find:

Theorem 3.6 *The topological space $T_1(X)$ and its foliation by geodesics can be reconstructed from the action of Γ on S^1.*

In particular, this foliated 3-manifold does not depend on the choice of $X \in \mathcal{T}_g$.

Theorem 3.7 *The space of currents $\mathcal{C}(X)$ is isomorphic to the space of invariant measures for the geodesic flow on $T_1(X)$ that are also invariant under time reversal.*

Proof. The lift of an invariant measure to $T_1(\mathbb{H})$ gives a measure on $\mathcal{G} \times S^1$ that is locally of the form $\mu \times ds$ for length measure along geodesics. By invariance under Γ, μ is a geodesic current. Conversely, the product of a geodesic current with length measure gives an invariant measure for the geodesic flow. ■

Remark. The geodesic *flow* cannot be reconstructed from the topological action of Γ on S^1, since its time parameterization determines the lengths of closed geodesics.

Intersection number. Let $\mathcal{I} \subset \mathcal{G} \times \mathcal{G}$ be the set of pairs of geodesics (α, β) that cross one another. Note that we can identify $\mathcal{I}$ with the bundle $I_1(\mathbb{H})$ whose fiber over $x \in \mathbb{H}$ consists of ordered pairs $(\pm v, \pm w)$ of unoriented, linearly independent unit vectors.

Similarly, we have $\mathcal{I}/\Gamma = I_1(X)$. We define the *intersection number* $i(\alpha, \beta)$ of a pair of geodesic currents to be the total measure of $I_1(X)$ with respect to the measure $\alpha \times \beta$.

Theorem 3.8 *If α, β are closed geodesics on X, then $i(\alpha, \beta)$ is the number of transverse intersections of α and β.*

Theorem 3.9 *We have $i(\alpha, L_X) = \ell_\alpha(X)$.*

Proof. We have a natural projection $I_1(X) \rightarrow P_1(X) = T_1(X)/(\pm 1)$ by $(\pm v, \pm w) \mapsto (\pm v)$. By the characteristic property of the integral geometry measure, the pushforward of $\alpha \times L_X$ to $P_1(X)$ gives the product of α with length measure along geodesics; i.e. it gives the measure invariant under the geodesic flow corresponding to α. The total mass of this measure is then just the length of α. ■

Remark. The proof shows that $i(\alpha, L_X)$ is simply the total mass of the corresponding measure on $P_1(X)$ invariant under the geodesic flow.

Theorem 3.10 *We have* $i(L_X, L_X) = \pi^2|\chi(X)|$.

Proof. As above, the pushforward of $L_X \times L_X$ to $P_1(X)$ is the invariant measure μ for the geodesic flow attached to L_X. We wish to relate this measure to the usual Liouville measure $dx\,dy\,d\theta$. Since the computation is at a small scale, it suffices to work in Euclidean space.

Thus we let L_X be the measure on the space of unoriented lines in $\mathbb{R}^2$ given by the 2-form

$$\omega = (1/2)\sin\theta\, d\theta\, du.$$

Here the line $L(\theta, u)$ passes through $(u, 0)$ and has slope $\theta \in [0, \pi]$. Letting (θ, x, y) denote coordinates on $P_1(\mathbb{R}^2)$, we have

$$\tan\theta = y/(x-u),$$

and thus $u = x - y/\tan\theta$. Using the fact that $d\theta$ already appears in ω, we then have

$$\omega = (1/2)\sin\theta\, d\theta(dx - dy/\tan\theta) = (1/2)d\,\theta(\sin\theta\, dx - \cos\theta\, dy)$$

as a form pulled back to $P_1(\mathbb{R}^2)$. Since the form $\alpha = \cos\theta\, dx + \sin\theta\, dy$ restricts to arclength along any geodesic in $P_1(\mathbb{R})$, we find:

$$\mu = \omega \wedge \alpha = (1/2)d\,\theta(\sin^2\theta + \cos^2\theta)\, dx\, dy = (1/2)dx\, dy\, d\theta.$$

Consequently $i(L_X, L_X)$ is one-half the volume of $P_1(X)$ with respect to the standard measure. The standard measure has total volume $= \pi\,\mathrm{area}(X) = 2\pi^2|\chi(X)|$, completing the proof. ■

Topology of $\mathcal{C}(X)$. The space $\mathcal{C}(X)$ is equipped with the *weak topology.* That is, we have $\mu_n \rightarrow \mu$ as measures on $\mathcal{G} \times \mathcal{G}$ if and only if

$$\int \phi\mu_n \rightarrow \int \phi\mu$$

for every $\phi \in C_0(\mathcal{G} \times \mathcal{G})$.

Using the fact that closed curves are dense in $\mathcal{G}$, it is straightforward to show:

Theorem 3.11 *The set of weighted closed curves is dense in* $\mathcal{C}(X)$*.*

The next result is a little more subtle, since $\mathcal{I}$ is an open set. It depends on the fact that pairs of geodesics near ∂I are nearly parallel.

Theorem 3.12 *The intersection number* $i : \mathcal{C}(X) \times \mathcal{C}(X) \to \mathbb{R}$ *is continuous.*

Corollary 3.13 *A map* $f : S \to \mathcal{C}(X)$ *is continuous iff* $i(f(s), \alpha)$ *is continuous for every closed curve* α*.*

Corollary 3.14 *The map* $\mathcal{T}_g \to \mathcal{C}(X)$ *given by* $Y \mapsto L_Y$ *is a proper homeomorphism to its image.*

Proof. We can find a finite set of simple closed curves whose lengths determine Fenchel-Nielsen coordinates for Y. ■

Compactness. We say a current α *binds* X if every geodesic on X crosses a geodesic in the support of α. For example, if $\alpha = \sum C_i \alpha_i$ is a weighted sum of closed curves with $C_i > 0$, and if every component of $X - \bigcup \alpha_i$ is a topological disk, then α binds X. Similarly, L_X binds X.

Theorem 3.15 *If* α *binds* X*, then for any* $M > 0$ *the set*

$$K = \{\beta \,:\, i(\alpha, \beta) \leq M\} \subset \mathcal{C}(X)$$

is compact.

Proof. It suffices to show that every $p \in \mathcal{G}$ has a neighborhood U such that $\beta(U) \leq C_U$ for all $\beta \in K$. Given p, pick $q \in \operatorname{supp} \alpha$ such that $(p, q) \in \mathcal{I}$. Since $\mathcal{I}$ is open, we can find a neighborhood $U \times V$ of (p, q) within it. Then for all $\beta \in K$, we have

$$\beta(U)\alpha(V) \leq i(\alpha, \beta) \leq M,$$

and thus $\beta(U) \leq M/\alpha(V) = c_U$. ■

Corollary 3.16 *Let* $(\alpha_i)_1^N$ *be a set of closed curves that bind* X*. Then the set of points* $Y \in \mathcal{T}_g$ *such that* $\ell_Y(\alpha_i) \leq M$ *is compact.*

Random geodesics. Putting these results together, we see that any point in Teichmüller space can be specified by a sequence of weighted simple closed curves $C_n \gamma_n \to L_X$. These curves can be constructed by choosing a random vector in $T_1(X)$ and closing longer and longer segments of the resulting geodesic. Thus one sometimes refers to L_X as a 'random geodesic' on X. This is made more precise by the following result.

Theorem 3.17 *For suitable closed geodesics* γ_n*, we have*

$$L_X(\delta) = \frac{\pi}{2} \operatorname{area}(X) \lim \frac{i(\gamma_n, \delta)}{L_X(\gamma_n)}. \tag{3.1}$$

Proof. Choose γ_n such that $C_n\gamma_n \to L_X$; then we have $L_X(\delta) = \lim C_n i(\gamma_n, \delta)$ and the equation

$$i(C_n\gamma_n, L_X) = C_n L_X(\gamma_n) \to i(L_X, L_X) = \pi^2|\chi(X)| = (\pi/2)\,\mathrm{area}(X)$$

gives the behavior of the constants C_n. ∎

Alternatively, for small r one can consider the locus $U_r \subset \mathrm{T}_1(X)$ of tangents to segments of length $2r$ centered on points of δ. Then the angular measure of U_r through a point p at distance t from δ is given by $4\cos^{-1}(t/r)$ (which tends to 2π as $t \to 0$). Thus

$$m(U_r) \sim L_X(\delta) \int_{-r}^{r} 4\cos^{-1}(|t|/r)\,dt = 8rL_X(\delta).$$

Now a random geodesic γ_n will meet U_r in $i(\gamma_n, \delta)$ segments, each of length $2r$; hence we have:

$$\frac{\ell(\gamma_n \cap U_r)}{\ell(\gamma_n)} = \frac{2ri(\gamma_n, \delta)}{L_X(\gamma_n)} \to \frac{m(U_r)}{m(T_1(X))} = \frac{8rL_X(\delta)}{2\pi\,\mathrm{area}(X)}.$$

This shows:

$$i(\gamma_n, \delta) \sim \frac{2L_X(\gamma_n)L_X(\delta)}{\pi\,\mathrm{area}(X)},$$

which also gives (3.1).

Laminations. A *geodesic lamination* $\lambda \subset X$ is a closed set that is a union of simple geodesics. Through each point of λ there passes a unique complete simple geodesic contained in λ. (One might regard the foliation of λ by simple geodesics as part of the structure of the lamination; for surfaces of finite volume this structure is redundant, but for surfaces of infinite volume (such as $\mathbb{H}$ itself) it is not.)

A *transverse measure* μ for λ is an assignment of a positive measure to each transversal τ to λ, supported on $\tau \cap \lambda$ and invariant under homotopy. A *measured lamination* is a geodesic lamination equipped with a transverse measure of full support.

We say $\alpha \in \mathcal{C}(X)$ is a *measured lamination* if $i(\alpha, \alpha) = 0$. This condition implies that the set of geodesics in the support of α form a geodesic lamination λ, and that α yields a transverse invariant measure of full support. The converse also holds, and thus:

Theorem 3.18 *The set $\mathcal{ML}(X)$ of measured geodesic laminations can be identified with the space of currents such that $i(\alpha, \alpha) = 0$.*

Examples of laminations. The simplest example of a measured lamination is a simple closed curve α.

For a more interesting example, let $\mathcal{F}$ be an irrational measured foliation of a torus T. Cut T open along a single leaf, and insert a bigon. The result is a punctured torus X with a measure lamination λ whose transversals are Cantor sets.

This example can be compared to the construction of the Cantor set by cutting the interval $[0, 1]$ open at every dyadic rational $p/2^n$, and inserting an interval of length 3^{-n}.

Theorem 3.19 (Nielsen) *Any geodesic lamination of a surface of finite volume has measure zero.*

Proof. Apply Gauss-Bonnet to the surface obtained by doubling $X - \lambda$ across its boundary. ■

Compactification. Let $\mathbb{P}\mathcal{C}(X) = \mathcal{C}(X)/\mathbb{R}_+$ with the quotient topology.

Theorem 3.20 *The projective space $\mathbb{P}\mathcal{C}(X)$ is compact.*

Proof. Any sequence of projective currents $[\alpha_n]$ lifts to a sequence normalized to that $\ell_X(\alpha_n) = 1$, and these currents form a compact set by Theorem 3.15. ■

Theorem 3.21 *The Teichmüller space $\mathcal{T}_g$ embeds into $\mathbb{P}\mathcal{C}(X)$, and the union*

$$\mathcal{T}_g \cup \mathbb{P}\mathcal{ML}(X)$$

is compact.

Proof. Embedding is immediate because $i(L_Y, L_Y)$ is constant for all $Y \in \mathcal{T}_g$. Now suppose $Y_n \to \infty$ in $\mathcal{T}_g$. Then there exists a simple closed curve α such that $i(\alpha, L_{Y_n}) \to \infty$. Consequently L_{Y_n} goes to infinity in the space of currents. Thus if we rescale L_{Y_n} so it remains in a compact set, its self-intersection number tends to zero and any limit is a measured lamination. ■

Pinching. As an example, let $(\ell_i, \tau_i)_1^{3g-3}$ be Fenchel-Nielsen coordinates associated to a pair-of-pants decomposition (α_i) of Z_g. Set $\tau_i = 0$ and consider the surfaces $Y_n = Y(\ell_i^n)$ where $\ell_i^n \le 1$ and $\ell_i^n \to 0$ for some i. Each corresponding geodesic γ_i^n on Y_n is enclosed in a standard collar neighborhood of width approximately $\log(1/\ell_i^n)$. Passing to a subsequence, we can assume that $[\log(1/\ell_i^n)] \to [w_i]$ in the projective space $\mathbb{P}(\mathbb{R}^n)$. Then $Y_n \to \sum w_i\alpha_i \in \mathbb{P}\mathcal{ML}_g$, because

$$\ell_\beta(Y_n) = \sum \log(1/\ell_i^n) i(\beta, \alpha_i) + O_\beta(1)$$

for every closed geodesic β on Y_n.

This shows every weight sum of disjoint simple closed curves actually occurs in $\partial\mathcal{T}_g$. We will soon see these sums are dense in $\mathbb{P}\mathcal{ML}_g$, and thus $\mathbb{P}\mathcal{ML}_g = \partial\mathcal{T}_g$.

Twisting. Similarly, let $\psi \in \mathrm{Mod}_g$ be a product $\psi = T_{C_i}^{n_i}$ of Dehn twists about disjoint simple closed curves, and let $\lambda = \sum |n_i| C_i$. Then $\psi^n(X) \to [\lambda]$ for all $X \in \mathcal{T}_g$.

As Bers once put it: there are two ways to send a Riemann surface to infinity in Teichmüller space: by pinching it, and by wringing its neck.

3.4 Laminations

The space $\mathcal{ML}_g$ turns out to be a rather concrete space: it is a finite-dimensional manifold, homeomorphic to a $\mathbb{R}^{6g-6}$, with combinatorial charts that depend only on the topology of Z_g.

Example: braids and simple closed curves. To motivate our approach to $\mathcal{ML}_g$, and more generally to $\mathcal{ML}_{g,n}$, we consider a particular sequence of simple closed curves on the 4-times punctured sphere.

Let $X = \mathbb{C} - \{-1, 0, 1\}$. Choosing a marking, we can consider X as an element of the Teichmüller space $\mathcal{M}_{0,4}$.

Let A, B be simple closed curves on X, represented by round circles, such that A encloses $\{-1, 0\}$ and B encloses $\{0, 1\}$. Let $\alpha, \beta \in \mathrm{Mod}(X)$ be given by the left Dehn twist on A and a right Dehn twist on B. Note that α and β are conjugate under the reflection $\rho(x + iy) = -x + iy$.

We now wish to examine the simple closed curves given by

$$(C_0, C_1, C_2, C_3, \ldots) = (A, \beta(A), \alpha\beta(A), \beta\alpha\beta(A), \ldots).$$

These curves can be visualized by enclosing the top two strands of an alternating, 3-stranded braid in a rubber band, and then pushing the band downwards, one crossing at a time. See Figure 12.

By the time one has drawn C_4, it has become evident that (although they are very long) these curvesi consist mainly of many parallel strands, joined together in a simple branching pattern.

The branching pattern for the curves with odd indices is the *train track* τ shown in Figure 13. Any collection of non-negative integral *weights* (a, x, b, y, c) on the edges of τ, satisfying the *switching conditions*

$$a = b + x, \ a + x = c + y, \ b + c = y$$

at its vertices, determines a simple closed curve $C(a, x, b, y, c)$ *carried* by τ. To construct $C(a, x, b, y, c)$, replace each edge of τ by a number of parallel strands determined by its weight, and join them without crossing at the vertices.

For simplicity, we eliminate the weights x and y (since they are determined by the switching conditions); we then have a curve $C(a, b, c)$ defined by integers satisfying $a = b + c$. In this notation, $C_0 = C(1, 1, 0)$, $C_2 = (2, 1, 1)$ and $C_4 = (5, 2, 3)$, and in fact all the curves C_{2i} are carried by τ. The weights (a, b, c) are simply the number of strands crossing $(-\infty, -1)$, $(0, -i\infty)$ and $(1, \infty)$ respectively.

The curves C_{2i+1} are similarly carried by the train track $\tau' = \rho(\tau)$. Using the same crossing numbers, we then have simple closed curves $C'(a, b, c)$ carried by τ'. On τ' the weight relation becomes $c = a + b$.

It is now easy to check:

Theorem 3.22 *We have $\beta(C(a, b, c)) = C'(a, c, b + 2c)$ and $\alpha(C'(a, b, c)) = C(b + 2a, a, c)$.*

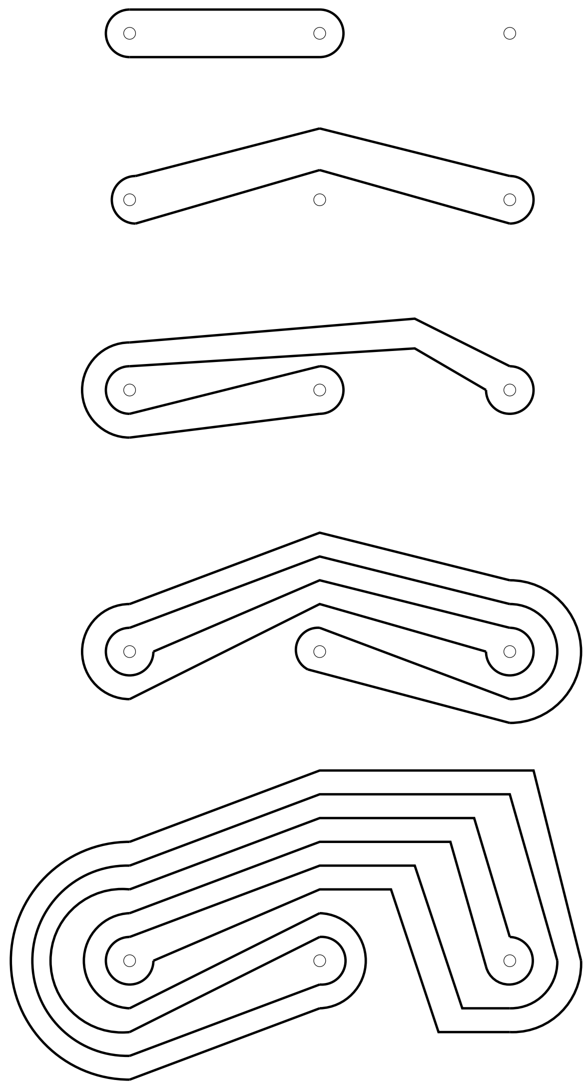

Figure 12. Simple closed curves C_0, C_1, C_2, C_3, C_4.

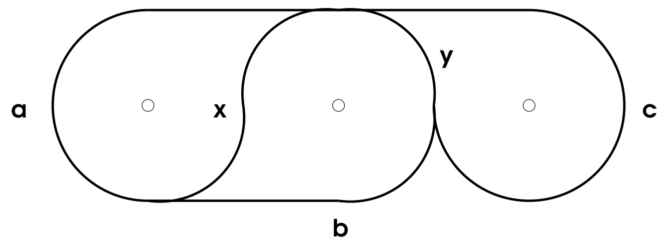

Figure 13. A train track for C_{2i}.

Since the weight transformations are linear, it is easy to compute that the action of $\psi = \alpha\beta$ is given by $\psi(C(a,b,c)) = C(2a+c, a, b+2c)$. We can eliminate $b = a - c$ and obtain

$$\psi(C(a,c)) = C(2a+c, a+c) = C\left(\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ c \end{pmatrix}\right).$$

Starting with $C_0 = C(1,0) = C(f_1, f_0)$, we then find $C_{2i} = (f_{2i+1}, f_{2i})$, where $(f_0, f_1, f_2, f_3, \ldots) = (0, 1, 1, 2, 3, 5, 8, \ldots)$ are the Fibonacci numbers.

It is well-known that $f_{n+1}/f_n \to \gamma = (1+\sqrt{5})/2$ (the golden ratio). Thus $[C_{2i}]$ converges, in the space $\mathbb{P}\mathcal{ML}$, to the projective measured lamination $[C(\gamma, 1)] = [\lambda]$. Moreover, this measured lamination satisfies $\psi(\lambda) = \gamma \cdot \lambda$.

Note that $[\lambda]$ is *not* transversally orientable! Thus it does not represent a cohomology class on X.

We now proceed to a more formal discussion.

Train tracks. A *train track* $\tau \subset X$ is a finite 1-complex such that

(i) every $x \in \tau$ lies in the interior of a smooth arc embedded in τ,

(ii) any two such arcs are tangent at x, and

(iii) for each component U of $X - \tau$, the double of U along the smooth part of ∂U has negative Euler characteristic.

Example. The complementary regions of our train track τ are all punctured monogons. The double of such a surface is a triply-punctured sphere.

The *vertices* (or switches) of a train track, $V \subset \tau$, are the points where 3 or more smooth arcs come together. The *edges* E of τ are the components of $\tau - V$; some 'edges' may be closed loops.

The module of a train track. Let $T(\tau)$ denote the $\mathbb{Z}$-module generated by the edges E of τ, modulo the relations

$$[e_1] + \cdots + [e_r] \quad = \quad [e'_1] + \cdots + [e'_s]$$

for each vertex $v \in V$ with incoming edges (e_i) and outgoing edges (e'_j). (The distinction between incoming and outgoing edges depends on the choice a direction along τ at v.) Since there is one relation for each vertex, we obtain a presentation for $T(\tau)$ of the form:

$$\mathbb{Z}^V \xrightarrow{D} \mathbb{Z}^E \rightarrow T(\tau) \rightarrow 0. \tag{3.2}$$

The space of 1-cycles on τ with values in B is defined by

$$Z_1(\tau, B) = \mathrm{Hom}(T(\tau), B).$$

Laminations A geodesic lamination λ is *carried* by a train track τ if there is a continuous *collapsing map* $f : \lambda \to \tau$ such that for each leaf $\lambda_0 \subset \lambda$,

(i) $f|\lambda_0$ is an immersion, and

(ii) λ_0 is the geodesic representative of the path or loop $f : \lambda_0 \to X$.

Collapsing maps between train tracks are defined similarly.

Theorem 3.23 *Every geodesic lamination is carried by some train track.*

See [HP, 1.6.5].

A collapsing map $\lambda \to \tau$ sends transverse measures on λ to elements of $Z_1(\tau, \mathbb{R}_+)$.

Theorem 3.24 *The set of measured laminations carried by τ corresponds to the set of positive 1-cycles $Z_1(\tau, \mathbb{R}_+)$.*

The set of systems of simple closed curves with rational weights carried by τ corresponds to $Z^1(\tau, \mathbb{Q}_+)$.

Corollary 3.25 *Weighted systems of simple closed curves are dense in $\mathcal{ML}(X)$.*

Corollary 3.26 *The closure of $\mathcal{T}_g$ is all of $\mathbb{P}\mathcal{ML}(X)$.*

Using train tracks as charts, one finds:

Theorem 3.27 *The space $\mathcal{ML}(X)$ is a PL-manifold of dimension $6g - 6$.*

For example, if we take $3g-3$ curves forming a pair of pants decomposition of Σ_g (with $2g-2$ pants), then add 3 more arcs to each pair of pants to obtain a train track, then the original curves give $4(3g-3) = 12g-12$ edges, the new arcs give $3(2g-2) = 6g-6$ more edges; and each of the original curves now carries 4 vertices, so we get $|E| = 18-18$, $|V| = 4(3g-3) = 12g-12$, and the difference is $6g-6$.

With more work (using e.g. the classification of surface diffeomorphisms), one can show more precisely:

Theorem 3.28 *The space $\mathbb{P}\mathcal{ML}(X)$ is homeomorphic to a sphere of dimension $6g-7$.*

Example. Using our train track τ, one sees that $\mathbb{P}\mathcal{ML}_{0,4}$ is a compact 1-manifold, hence a circle. It is instructive to find a finite collection of train tracks whose charts cover $\mathbb{P}\mathcal{ML}_{0,4}$.

3.5 Symplectic geometry of Teichmüller space

Hamiltonian formalism. Recall that a *symplectic manifold* (M^{2n},ω) is a smooth manifold equipped with a closed 2-form such that $\omega^n \neq 0$. A typical example is $\mathbb{C}^n \cong \mathbb{R}^{2n}$ with $\omega = \sum dx_i \wedge dy_i$.

Any smooth function $H : M \to \mathbb{R}$ determines a *Hamiltonian vector field* X_H, characterized by $\omega(X_H, Y) = dH(Y)$. The flow generated by X_H preserves both ω and H; for example,

$$\mathcal{L}_X(\omega) = di_X(\omega) + i_X d(\omega) = d(dH) + i_X(0) = 0.$$

Fenchel-Nielsen coordinates. Recall that any pair of pants decomposition of Z_g, determined by simple closed curves $\alpha_1, \ldots, \alpha_{3g-3}$, gives a coordinate system (τ_i, ℓ_i) for $\mathcal{T}_g$.

We emphasize that τ_i, like ℓ_i, has units of length, and that the vector field $\partial/\partial\tau_i$ represents a *right* twist along the corresponding simple closed curve.

Theorem 3.29 *The symplectic form $\omega = \sum d\ell_i \wedge d\tau_i$ is independent of the choice of coordinate system.*

Kähler form. This symplectic form has many other manifestations; for example, when coupled to the complex structure it yields the *Weil-Petersson metric*, given on the cotangent space by $\|\phi\|_2^2 = \int_X |\phi|^2/\rho^2$.

In the case of $\mathbb{H} = \mathcal{T}_{1,1}$, we have $\tau \sim x/y$ and $\ell \sim 1/y$, and thus $\omega = d\ell \wedge d\tau \sim dxdy/y^3$ as $y \to \infty$. Thus the Weil-Petersson metric is approximately $|dz|/y^{3/2}$, the distance to the cusp is finite and hence $\mathcal{M}_1$ has finite Weil-Petersson diameter.

In particular, the Weil-Petersson metric is not equivalent to the Bergman metric, since the latter is complete and in fact comparable to the Teichmüller metric [Ha].

Earthquakes. For simple closed curves we have:

Corollary 3.30 *The Fenchel-Nielsen twist $-\partial/\partial\tau_i$ is the Hamiltonian vector field generated by the length function $\ell_i(X)$.*

Based on this observation, for any $\lambda \in \mathcal{ML}_g$ we define the *twist deformation* $\mathrm{tw}_\lambda : \mathcal{T}_g \to \mathcal{T}_g$ as the unit-time diffeomorphism generated by the Hamiltonian vector field $-X_H$, where $H(X) = \ell_\lambda(X)$. The result is a *right earthquake* along the lamination λ.

Theorem 3.31 *Any two points in Teichmüller space are connected by a unique right earthquake.*

Convexity of lengths. Given simple closed curves α and β, and $p \in \alpha \cap \beta$, let $\theta_p \in [0,\pi]$ denote the angle through which β must be rotated to line up with α. This angle changes to $\pi - \theta_p$ if the roles of α and β are reversed.

Theorem 3.32 *For any pair of laminations α, β the right twist along β satisfies*

$$\frac{d}{dt}\ell_\alpha(\mathrm{tw}_{t\beta})(X) = \sum_{p \in \alpha \cap \beta} \cos(\theta_p).$$

Corollary 3.33 *Geodesic lengths are convex along earthquake paths, and strictly so if α meets β.*

Proof. It is also easy to see that θ_p is decreasing under the earthquake flow along β, and thus

$$\frac{d^2}{dt^2}\ell_\alpha(\mathrm{tw}_{t\beta})(X) = \sum_{p \in \alpha \cap \beta} -\sin(\theta_p)\,\frac{d\theta_p}{dt} > 0.$$

■

Following [Ker], we then have:

Theorem 3.34 (Nielsen realization conjecture) *Any finite subgroup $G \subset \mathrm{Mod}_g$ can be lifted to a finite subgroup of $\mathrm{Diff}(Z_g)$.*

Equivalently, any finite subgroup G of Mod_g has a fixed-point in $\mathcal{T}_g$.

Proof. Pick a closed curves α that binds Z_g, and let $\beta = \sum_G g \cdot \alpha \in \mathcal{C}(Z_g)$. Then $\ell_\beta(X)$ is a proper function on Teichmüller space, so it achieves its minimum at some point X_0. If the minimum is also achieved at X_1, then $X_0 = X_1$ because $\ell_\beta(X_t)$ is strictly convex along the earthquake path from X_0 to X_1. Since $\ell_\beta(g \cdot X) = \ell_\beta(X)$, we have $g \cdot X_0 = X_0$ for all $g \in G$. ■

4 Teichmüller theory via complex analysis

This section provides an introduction to the *complex analytic theory* of Teichmüller space. It centers around the geometric meaning of the Teichmüller metric.

Basic references for Teichmüller theory include [Gd], [IT] and [Nag].

4.1 Teichmüller space

Definitions. A Riemann surface X is of *finite type* (g, n) if $X = \overline{X} - E$ for some compact Riemann surface $\overline{X}$ of genus g and finite set E with $|E| = n$. Each end of X is isomorphic to a punctured disk.

Let S be a compact connected oriented surface. and let X be a Riemann surface of finite type homeomorphic to $\mathrm{int}(S)$. A *marking* of X by S is an orientation-preserving homeomorphism

$$f : \mathrm{int}(S) \to X.$$

The *Teichmüller metric* on Riemann surfaces marked by S is defined by

$$d((f,X),(g,Y)) \;=\; \frac{1}{2}\inf\{\log K(h) \,:\, h : X \to Y \text{ is homotopic to } g\circ f^{-1}\}.$$

Here h ranges over all quasiconformal maps respecting markings, and

$$K(h) = \sup_X \frac{|h_z| + |h_{\overline{z}}|}{|h_z| - |h_{\overline{z}}|} \geq 1$$

is its dilatation.

For any two Riemann surfaces X, Y marked by S, the compactifications $\overline{X}$ and $\overline{Y}$ are diffeomorphic. Thus their exists a quasiconformal map $f : X \to Y$ respecting markings, and therefore the Teichmüller distance between X and Y is finite.

The *Teichmüller space* $\mathrm{Teich}(S)$ is obtained from the space of all marked Riemann surfaces (f, X) by identifying those at distance zero. Equivalently, two marked surfaces (f_1, X_1) and (f_2, X_2) define the same point in $\mathrm{Teich}(S)$ if there is a *conformal map* $h : X_1 \to X_2$ homotopic to $f_2 \circ f_1^{-1}$.

The *mapping class group* $\mathrm{Mod}(S)$ consists of all homotopy classes of orientation-preserving homeomorphisms $h : S \to S$. (In fact homotopy and isotopy define the same equivalence relation here.) There is a natural action of $\mathrm{Mod}(S)$ on $\mathrm{Teich}(S)$ by

$$h \cdot (f, X) = (f \circ h^{-1}, X).$$

This action is an isometry for the Teichmüller metric. (In fancier language, the map $S \mapsto \mathrm{Teich}(S)$ is a functor from the category of surfaces and isotopy classes of homeomorphisms, to the category of metric spaces and isometries.)

The notation $T_{g,n}$ is often used for the Teichmüller space of Riemann surfaces of genus g with n punctures.

4.2 The Teichmüller space of a torus

Theorem 4.1 *The Teichmüller space of a torus S is naturally isometric to $\mathbb{H}$, with the conformal metric $\frac{1}{2}|dz|/\,\mathrm{Im}(z)$ of constant curvature -4, and with $\mathrm{Mod}(S) \cong \mathrm{SL}_2(\mathbb{Z})$ acting by Möbius transformations.*

Proof. A marked torus $f : S \to X = \mathbb{C}/\Lambda$ determines a homomorphism $f_* : \pi_1(S) \to \Lambda$. Choosing oriented generators $\langle a, b\rangle$ for $\pi_1(S) \cong \mathbb{Z}\oplus\mathbb{Z}$, we can normalize by scaling in $\mathbb{C}$ so that $f_*(a, b) = (1, \tau)$ with $\tau \in \mathbb{H}$. This gives the desired map $\mathrm{Teich}(S) \cong \mathbb{H}$.

To see the Teichmüller metric d_T is half the hyperbolic metric $d_{\mathbb{H}}$, first note that the natural real affine map $f : X_1 \to X_2$, between a pair of tori $X_i = \mathbb{C}/(\mathbb{Z}\oplus\mathbb{Z}\tau_i)$, has dilatation $\log K(f) = d_{\mathbb{H}}(\tau_1, \tau_2)$. This is particularly clear when $\tau_i = iy_i$, since then $f(x, y) = (x, y_2 y/y_1)$ has $\log K(f) = |\log(y_2) - \log(y_1)|$.

Therefore

$$d_T(\tau_1, \tau_2) \leq d_{\mathbb{H}}(\tau_1, \tau_2)/2.$$

To see f is extremal, we look at the distortion of extremal length. Letting $\Gamma(1,0)$ denote the family of loops on S in the isotopy class $(1,0) \in \pi_1(S)$, we have seen that

$$\lambda(\Gamma(1,0), \mathbb{C}/(\mathbb{Z} \oplus \tau\mathbb{Z})) = \frac{1}{\operatorname{Im}\tau}.$$

Since extremal length is distorted by at most the factor $K(f)$, we have

$$d_T(\tau_1, \tau_2) \geq \frac{1}{2}|\log \operatorname{Im}(\tau_1) - \log \operatorname{Im}(\tau_2)|.$$

Now the right-hand side is just the hyperbolic distance between the two horocycles resting on $\infty \in \widehat{\mathbb{R}} = \partial\mathbb{H}$ that pass through τ_1 and τ_2. By replacing $\Gamma(1,0)$ with $\Gamma(p,q)$, we see in the same way that

$$d_T(\tau_1, \tau_2) \geq \frac{1}{2} d_{\mathbb{H}}(H_1(p/q), H_2(p/q)),$$

where $H_i(p/q)$ is the horocycle resting on p/q and passing through τ_i. Letting p/q tend to an endpoint of the geodesic between τ_1 and τ_2, the right-hand side tends to $d_{\mathbb{H}}(\tau_1, \tau_2)/2$, and thus $d_T = d_{\mathbb{H}}/2$. ■

Cross-ratios. Any 4-tuple $E \subset \widehat{\mathbb{C}}$ naturally determines a 2-fold covering space $\pi : X(E) \to \widehat{\mathbb{C}}$ branched over E, where $X = \mathbb{C}/\Lambda$ is a torus and the critical points of π are the points of order two on $X(E)$. Conversely, the quotient of a torus X by the involution $x \mapsto -x$ gives the sphere with 4 branch points E.

The map $E \mapsto X(E)$ gives a natural bijection between $T_{0,4}$, T_1. Since the extremal quasiconformal map $F : X(E) \to X(E')$ can be chosen to be a group homomorphism, it commutes with $x \mapsto -x$ and so it descends to a map $f : (\widehat{\mathbb{C}} - E) \to (\widehat{\mathbb{C}} - E')$. Thus $T_{0,4}$ and T_1 are isometric.

Let $\widetilde{\mathcal{M}}_{0,4}$ denote the moduli space of ordered 4-tuples of points on the Riemann sphere. Then $\widetilde{\mathcal{M}}_{0,4} \cong \widehat{\mathbb{C}} - \{0, 1, \infty\}$; the isomorphism is given by taking the cross-ratio. There is a regular covering map $\widetilde{\mathcal{M}_{0,4}} \to \mathcal{M}_{0,4}$ with Galois group S_3. From the discussion above, we have:

Corollary 4.2 *The Teichmüller metric on the space of cross-ratios $\widetilde{\mathcal{M}}_{0,4} \cong \widehat{\mathbb{C}} - \{0, 1, \infty\}$ coincides with the hyperbolic metric of constant curvature -4.*

Corollary 4.3 *A K-quasiconformal mapping distorts the cross-ratio of any 4 points by a bounded amount, as measured in the hyperbolic metric on $\widehat{\mathbb{C}} - \{0, 1, \infty\}$.*

Corollary 4.4 *A K-quasiconformal map f is $(1/K)$-Hölder continuous.*

Proof. Assume f fixes $\{0, 1, \infty\}$. The hyperbolic metric on $X = \widehat{\mathbb{C}} - \{0, 1, \infty\}$ near $z = 0$ is bounded below by $|dz|/(|z| \log |Az|)$ for some $A > 0$, and

$$\int_a^b \frac{dx}{x \log(Ax)} = \log \log Ax,$$

so for z near zero we have

$$\log K(f) \geq d_X(z, f(z)) \geq d_X(|z|, |f(z)|) \geq |\log\log|Az| - \log\log|Af(z)||,$$

which gives $|f(z)| \leq O(|z|^{1/K})$.

This proves Hölder continuity of f at $z = 0$. By replacing f with AfB for Möbius transformations A and B, the same proof applies near any point. ■

Corollary 4.5 *The space of K-quasiconformal homeomorphisms of the sphere is compact, up to the action of* $\operatorname{Aut}(\widehat{\mathbb{C}})$.

Proof. Once normalized to fix 3 points, the space of K-quasiconformal maps is uniformly Hölder continuous by the preceding corollary, and it is closed because the property $\operatorname{mod}(fQ) \leq K \operatorname{mod}(Q)$ persists under uniform limits. By Arzela-Ascoli, we have compactness. ■

4.3 Quadratic differentials

Let X be a Riemann surface. A *quadratic differential* ϕ on X is a tensor locally given by $\phi = \phi(z)dz^2$, where $\phi(z)$ is holomorphic. In other words, ϕ is a section of the square of the canonical bundle on X.

If $\phi(p) \neq 0$, then we can find a local chart near p in which $\phi = dz^2$. This chart is unique up to

$$z \mapsto \pm z + a; \tag{4.1}$$

it is given locally by

$$z(q) = \int_p^q \sqrt{\phi}.$$

(Note that $\sqrt{\phi}$ is a holomorphic 1-form.)

A quadratic differential determines a *flat metric* $|\phi|$ on X and a *foliation* $\mathcal{F}$ tangent to the vectors with $\phi(v) > 0$. For the differential $\phi = dz^2$ on $\mathbb{C}$, the metric is just the Euclidean metric on the foliation is by horizontal lines. Note that these structures are preserved by the transformations (4.1).

If $\phi(p) = 0$, then there is a local chart with $\phi(z) = z^d dz^2$. The metric $|\phi|$ has a cone-like singularity with $(2+d)\pi$ degrees at p (and thus negative curvature); the foliation has $d + 1$ leaves coming together at p.

A quadratic differential is *integrable* if

$$\|\phi\| = \int_X |\phi| < \infty.$$

Let $Q(X)$ denote the Banach space of all integrable quadratic differentials with the L^1-norm above. The norm of ϕ is simply the total area of X in the $|\phi|$-metric.

Proposition 4.6 *Let X be a Riemann surface of finite type. Then $Q(X)$ consists of the holomorphic quadratic differentials on X with at worst simple poles at the punctures of X.*

Proof. Near $z = 0$, $\int |dz|^2/|z|^n = \int r^{-n} r\, dr\, d\theta$ converges iff $n < 2$. ∎

Example. The surface $X = \mathbb{C}^*$ becomes an infinite cylinder in the metric determined by $\phi = dz^2/z^2$, so $\phi \notin Q(X)$. On the other hand, for $X = (\mathbb{C} - \{0, 1, a\})$, the differential

$$\phi = \frac{dz^2}{z(z-1)(z-a)}$$

lies in $Q(X)$. When $a \in \mathbb{R}$, X is the double of a rectangle (a pillowcase) in the $|\phi|$-metric.

Corollary 4.7 *If X is a Riemann surface of finite type (g, n), then*

$$\dim Q(X) = \begin{cases} n-3 & g = 0, \\ 1 & g = 1, n = 0, \\ n-1 & g = 1, n > 0, \\ 3g-3+n & g \geq 2. \end{cases}$$

4.4 Measured foliations

A *measured foliation* $\mathcal{F}$ on a manifold M is a foliation equipped with a measure on the space of leaves. That is, $\mathcal{F}$ comes equipped with a measure α on all transversals T to $\mathcal{F}$, such that the natural maps between transversals (following the leaves of the foliation) are measure-preserving. Morally, α is a measure on the *leaf space* of $\mathcal{F}$.

Basic examples.

1. The foliation by horizontal lines in $\mathbb{C}$ with the transverse measure $|dy|$ determines a measured foliation $\mathcal{F}$, such that the mass of a transversal T is its total variation in the vertical direction.

2. Since $\mathcal{F}$ is translation invariant, it descends to a measured foliation on any complex torus $X = \mathbb{C}/\Lambda$.

3. On a surface, a measured foliation is determined by a degenerate metric g of rank one on X. The null-geodesics of g are the leaves of the foliation, and the measure of a transversal is its g-length. Thus $g(x, y) = y^2$ describes the standard foliation in the plane.

4. A nowhere vanishing holomorphic quadratic differential ϕ on X similarly determines a measured foliation $\mathcal{F}(\phi)$. In a chart where $\phi = dz^2$, the

measured foliation becomes the usual one whose leaves are horizontal lines. The degenerate metric g defining $\mathcal{F}(\phi)$ is given by

$$g = |\operatorname{Im}\sqrt{\phi}|^2.$$

The measured foliation $\mathcal{F}(-\phi)$ has leaves orthogonal to those of $\mathcal{F}(\phi)$, and the product of the transverse measures gives the area form determined by the quadratic differential:

$$|\phi| = \alpha_{\mathcal{F}(\phi)} \times \alpha_{\mathcal{F}(-\phi)}.$$

5. By convention, one broadens the definition of a measured foliation on a surface to allow singularities like those of quadratic differentials at their zeros. Then *any* holomorphic quadratic differential determines a measured foliation.

6. A *closed 1-form* ω with generic zeros determines a measured foliation with singularities smoothly equivalent to the standard model $(z\,dz)^2$ at $z = 0$. Namely $\omega = df$ locally, the foliation $\mathcal{F}$ is by the level sets of f (or the integral curves of the kernel of ω), and the f is the integral of the transverse measure. Conversely, any *transversely oriented* measured foliation corresponds to a closed 1-form.

   The special feature of holomorphic quadratic differentials is that they locally correspond to the square of a *harmonic* 1-form. Geometrically this means that the foliation obtained by rotating the leaves of $\mathcal{F}(\phi)$ by $90°$ (or any other angle) also comes equipped with a transverse measure; or that $\mathcal{F}(\phi)$ is locally defined by a 1-form ω such that $d * \omega = 0$.

Now given a measured foliation $\mathcal{F}$ and a conformal metric ρ, we can form a measure $\rho \times \alpha_{\mathcal{F}}$ on X that is locally the product of the transverse measure of $\mathcal{F}$ with ρ-length along leaves. We define the *length* of $\mathcal{F}$ by

$$\ell_\rho(\mathcal{F}) = \int_{\mathcal{F}} \rho(z)|dz| = \int_X \rho \times \alpha_{\mathcal{F}}.$$

Let us say $\mathcal{F} \sim \mathcal{F}'$ if there is a homeomorphism of X, isotopic to the identity, sending $\mathcal{F}$ to $\mathcal{F}'$. This relation is meant to generalize the idea of moving a simple curve by isotopy. Then the *extremal length* of $\mathcal{F}$ on X is defined by

$$\lambda(\mathcal{F}, X) = \sup_\rho \frac{\inf_{\mathcal{F}\sim\mathcal{F}'} \ell_\rho(\mathcal{F}')^2}{\operatorname{area}_\rho(X)}.$$

A measured foliation is *geodesic* for ρ if it minimizes $\ell_\rho(\mathcal{F})$ in its isotopy class, and ρ is *extremal* if it maximizes the extremal length quotient above.

Theorem 4.8 *Let $\phi \in Q(X)$ be a nonzero holomorphic quadratic differential on a Riemann surface X of finite type, and let $\rho = |\phi|^{1/2}$. Then*

$$\ell_\tau(\mathcal{F}(\phi)) \le \ell_\rho(\mathcal{F}(\phi))$$

for any metric τ with $\operatorname{area}_\tau(X) = \operatorname{area}_\rho(X)$. *Equality holds iff $\tau = \rho$.*

Proof. By definition, the restriction of $\rho = |\phi|^{1/2}$ to the orthogonals of $\mathcal{F} = \mathcal{F}(\phi)$ gives its transverse measure. Thus

$$\ell_\rho(\mathcal{F})^2 = \left(\int_X \rho^2\right)^2 = \mathrm{area}_\rho(X)^2.$$

On the other hand, by Cauchy-Schwarz,

$$\ell_\tau(\mathcal{F})^2 = \left(\int_{\mathcal{F}} \tau\right)^2 = \left(\int_X \tau\rho\right)^2 \le \int \tau^2 \int \rho^2 = (\mathrm{area}_\rho(X))^2\,,$$

by our assumption that τ and ρ give X the same area. If equality holds in Cauchy-Schwarz, then τ is a constant multiple of ρ, so $\tau = \rho$ by our normalization of the area. ■

Corollary 4.9 *We have $\lambda(\mathcal{F}(\phi)) = \|\phi\|$, and $\rho = |\phi|^{1/2}$ is the unique extremal metric for $\mathcal{F}(\phi)$, up to scale.*

Proof. The foliation $\mathcal{F} = \mathcal{F}(\phi)$ is the unique geodesic in its equivalence class because ρ is a metric of non-positive curvature. More precisely, suppose $\mathcal{F}' = f(\mathcal{F})$, where f is isotopic to the identity. Define a map $g : X \to X$ by sending each leaf L' of $\mathcal{F}'$ to the corresponding leaf $L = f^{-1}(L)$ of F by the nearest-point projection, in the ρ-metric. Since L is a ρ-geodesic and ρ is non-positively curved, this projection does not increase distances. Thus $\int_{\mathcal{F}'} \rho \ge \int_{\mathcal{F}} \rho$ and therefore $\mathcal{F}$ is geodesic. It follows that

$$\lambda(\mathcal{F}) \ge \frac{\inf_{\mathcal{F}'} \ell_\rho(\mathcal{F}')^2}{\mathrm{area}_\rho(X)} \ge \frac{\ell_\rho(\mathcal{F})^2}{\mathrm{area}_\rho(X)} = \mathrm{area}_\rho(X) = \|\phi\|.$$

On the other hand,

$$\lambda(\mathcal{F}) = \sup_\tau \frac{\inf_{\mathcal{F}'} \ell_\tau(\mathcal{F}')^2}{\mathrm{area}_\tau(X)} \le \sup_\tau \frac{\ell_\tau(\mathcal{F})^2}{\mathrm{area}_\tau(X)} \le \frac{\ell_\rho(\mathcal{F})^2}{\mathrm{area}_\rho(X)} = \|\phi\|$$

by the preceding theorem, and for equality to hold we must have τ a constant multiple of ρ. ■

Unique ergodicity. The ρ-geodesic representative of $\mathcal{F}(\phi)$ may not be unique as a *measured foliation.* For example, let $X = \mathbb{C}/\mathbb{Z} \oplus \tau\mathbb{Z}$ be a torus, let $\phi = [dz^2] \in Q(X)$, and let $\mathcal{F} = \mathcal{F}(\phi)$ and $\rho = |\phi|^{1/2}$. Then projection to the y-axis gives a fibration

$$\pi : X \to S^1 = \mathbb{R}/\mathbb{Z}\,\mathrm{Im}\,\tau,$$

whose fibers are the leaves of $\mathcal{F}$. The transverse measure to $\mathcal{F}$ pushes forward to give a measure α on S^1, and conversely any measure α' on S^1 determines a measured foliation $\mathcal{F}'$ with the same leaves as $\mathcal{F}$. If the total mass of α and α'

is the same, then $\mathcal{F}' \sim \mathcal{F}$, and both are geodesic for the ρ-metric, even though they are distinct as measured foliations.

On the other hand, we say a measured foliation is *uniquely ergodic* if it admits a *unique* transverse invariant measure up to scale. For example, an *irrational* foliation of X is uniquely ergodic. If $\mathcal{F}(\phi)$ is uniquely ergodic, then it is the *unique* geodesic measured foliation in the $|\phi|^{1/2}$-metric.

4.5 Teichmüller's theorem

In this section we continue to assume $\chi(S) < 0$, so the Riemann surfaces in $\mathrm{Teich}(S)$ are hyperbolic.

Definition. Let $X, Y \in \mathrm{Teich}(S)$ be marked Riemann surfaces. A *Teichmüller mapping*

$$f : X \to Y$$

is a quasiconformal map, respecting markings, such that

$$\mu(f) = \frac{\overline{\partial} f}{\partial f} = k\frac{\overline{\phi}}{|\phi|}$$

for some $\phi \in Q(X)$ and $0 \le k < 1$. The map f has dilatation $K = (1+k)/(1-k)$.

Equivalently: there is a quadratic differentials $\phi \in Q(X)$ such that, away form its zeros, there are charts on X and Y in which $\phi = dz^2$ and $f(x + iy) = Kx + iy$.

This means f maps $\mathcal{F}(\phi)$ to $\mathcal{F}(\psi)$ for some $\psi \in Q(Y)$, stretching the leaves of $\mathcal{F}(\phi)$ by a constant factor K in the flat metric, and sending the orthogonal leaves to the orthogonal leaves by an isometry.

Theorem 4.10 *Let $f : X \to Y$ be a Teichmüller mapping between hyperbolic Riemann surfaces. Then f is the unique extremal quasiconformal map in its homotopy class. That is, $K(g) \ge K(f)$ for any g homotopic to f, and equality holds iff $f = g$. In particular,*

$$d_T(X, Y) = \frac{1}{2}\log K(f).$$

Proof. First note that the identity map is the only conformal map homotopic to the identity. This can be seen by lifting f to the universal cover so it induces the identity on the circle at infinity. So we may assume $K = K(f) > 1$.

Let $\phi \in Q(X)$ and $\psi \in Q(Y)$ be the quadratic differentials (unique up to a positive multiple) such that $\|\psi\| = K\|\phi\|$ and $f(\mathcal{F}(\phi)) = \mathcal{F}(\psi)$. That is, f stretches the leaves of $\mathcal{F}(\phi)$ by the factor K.

Consider any quasiconformal mapping $F : X \to Y$ homotopic to f. Let

$$\rho = |\psi|^{1/2}$$

be the extremal metric for $\mathcal{F}(\psi)$ on Y, let

$$g = F^*(\rho)$$

be the Riemannian metric obtained by pulling it back, and let

$$\gamma \le g \le K(F)\gamma$$

be the 'rounding down' of g, i.e. the largest conformal metric on X lying below g. Then:

$$\begin{aligned} K(f)\|\phi\| &= \|\psi\| = \lambda(\mathcal{F}(\psi)) = \frac{\ell_\rho(\mathcal{F}(\psi))^2}{\operatorname{area}_\rho(Y)} \le \frac{\ell_\rho(F(\mathcal{F}(\phi)))^2}{\operatorname{area}_\rho(Y)} \\ &= \frac{\ell_g(\mathcal{F}(\phi))^2}{\operatorname{area}_g(X)} \le K(F)\frac{\ell_\gamma(\mathcal{F}(\phi))^2}{\operatorname{area}_\gamma(X)} \le K(F)\lambda(\mathcal{F}(\phi)) = K(F)\|\phi\|. \end{aligned}$$

Therefore $K(F) \ge K(f)$ and thus f is extremal.

To check uniqueness, suppose equality holds above. Then the process of rounding down, that is replacing g with γ, must decrease the area of X by exactly $K(F)$ while holding the length of $\mathcal{F}(\phi)$ constant. It follows that the major axis of g must be tangent to $\mathcal{F}(\phi)$, and the eccentricity of g must be $K(F) = K(f)$ a.e. But then $\mu(f) = \mu(g)$, so $g^{-1} \circ f$ is a conformal map isotopic to the identity, and thus $f = g$. ∎

Theorem 4.11 (Teichmüller's theorem) *There is a unique Teichmüller mapping $f : X \to Y$ between any pair of Riemann surfaces $X, Y \in \operatorname{Teich}(S)$.*

Proof. The proof is by the 'method of continuity'.

Let $Q(X)_1$ denote the open unit ball in $Q(X)$. We will define a map

$$\pi : Q(X)_1 \to \operatorname{Teich}(S)$$

such that the theorem holds for all Y in the image; then we will show π is a bijection.

To define π, let ϕ belong to $Q(X)_1$ and let $k = \|\phi\| < 1$. For $\phi = 0$ we set $\pi(0) = X$. Otherwise, we construct a Teichmüller mapping $f : X \to Y$ with $\mu(f) = k\overline{\phi}/|\phi|$, and set $\pi(\phi) = (f, Y)$.

To construct Y, we just apply a stretch by the factor $K = (1+k)/(1-k)$ along the foliation $\mathcal{F}(\phi)$, to obtain charts for a new Riemann surface. The complex structure fills in over the isolated zeros of ϕ by the removable singularities theorem. Then by construction we have a Teichmüller mapping $f : X \to Y$ stretching along $\mathcal{F}(\phi)$.

Since f is extremal, we have $d(X, \pi(X)) = \frac{1}{2}\log K$. Since the extremal is unique, and f determines ϕ, we can recover ϕ from $\pi(X)$; and thus π is injective. But π is also continuous, and the dimensions of domain and range are equal by Theorem 3.1. Therefore π is an proper local homeomorphism, and thus a covering map. Since π is injective, it is a global homeomorphism.

In particular, π is surjective, so any $Y \in \operatorname{Teich}(S)$ is the target of a Teichmüller mapping with domain X. ∎

The case of a torus. When S is a torus, Teichmüller's theorem still holds up to the action of translations on X. That is, the Teichmüller mapping is only unique up to composition with a conformal automorphism of X isotopic to the identity.

Real and complex Teichmüller geodesics. Associated to a nonzero quadratic differential $\phi \in Q(X)$ is an *isometric, holomorphic* map

$$\Delta \to \mathrm{Teich}(S)$$

of the unit disk (with curvature -4) into Teichmüller space. This map $t \mapsto X_t$ is defined by letting X_t be X endowed with the complex structure coming from the Beltrami differential

$$\mu_t = t\phi/|\phi|,$$

i.e. for $t = ke^{i\theta}$, the surface X_t is the terminus of the Teichmüller mapping

$$f_t : X \to X_t$$

with quadratic differential $e^{-i\theta}\phi$ and with dilatation $K(f_t) = (1+k)/(1-k)$.

The image of the geodesic $(-1, 1) \subset \Delta$ is the *real Teichmüller geodesic* through X determined by ϕ; the image of Δ itself is the corresponding *complex* geodesic.

4.6 The tangent and cotangent spaces to Teichmüller space

On any complex manifold M, a deformation of complex structure can be defined by moving charts relative to one another. Thus the deformations are naturally isomorphic to $H^1(M, \Theta)$, where Θ is the sheaf of holomorphic vector fields on M. This point of view is the basis of Kodaira-Spencer deformation theory.

On a Riemann surface X, Θ is the sheaf of sections of K^*, the dual of the canonical bundle. By Serre duality,

$$H^1(X, \Theta)^* \cong H^0(X, K - \Theta) = H^0(X, 2K) = Q(X).$$

In other words, $Q(X)$ is naturally the *cotangent space* to the Teichmüller space $\mathrm{Teich}(S)$ at X.

From the quasiconformal point of view, any complex structure on X can be specified by a Beltrami differential of norm less than one, and thus we have a surjective holomorphic map

$$\pi : M(X)_1 \to \mathrm{Teich}(S)$$

sending 0 to X.

Here is a concrete description of this map π. First present X as the quotient $\mathbb{H}/\Gamma$ of the upper half-plane by a Fuchsian group Γ. Given $\mu \in M(X)_1$, lift it to $\mathbb{H}$ and extend it by reflection to the lower half-plane, obtaining a differential $\widetilde{\mu} \in M(\widehat{\mathbb{C}})_1$. Now let F denote the solution to the Beltrami equation

$$\frac{F_{\overline{z}}}{F_z} = \widetilde{\mu}$$

that fixes $\{0,1,\infty\}$. Then by symmetry, F sends $\mathbb{H}$ to itself. Since $\widetilde{\mu}$ is Γ-invariant, F conjugates Γ to a new Fuchsian group Γ'. Let $X' = \mathbb{H}/\Gamma'$. Then F descends to a quasiconformal map $f : X \to X'$, and $\pi(\mu) = (f, X')$.

The tangent space $T_X \operatorname{Teich}(S)$ is the quotient of $M(X)$ (the tangent space to $M(X)_1$ at the origin) by the space of trivial deformations of X. Here μ is *trivial* if the complex structure it specifies is the same as the original structure on X up to isotopy. On an infinitesimal level, this means $\mu = \overline{\partial} v$ for some quasiconformal vector field on X. (The space of such v can be thought of as the tangent space at the identity to the group $QC(X)$ of quasiconformal homeomorphisms of X to itself.)

We claim $\mu = \overline{\partial} v$ iff $\int \mu\phi = 0$ for every $\phi \in Q(X)$. To see this, we note that the space $M_0(X)$ of trivial μ is a closed subspace of $M(X)$, by compactness of quasiconformal vector fields. As Banach spaces we have

$$L^1(X, dz^2)^* = L^\infty(X, d\overline{z}/dz) = M(X),$$

so $M_0(X) = P^\perp$ for some subspace P of the space of L^1 measurable quadratic differentials. But if $\phi \in P$, then

$$\int_X \phi \overline{\partial} v = -\int_X (\overline{\partial}\phi) v = 0$$

for every smooth vector field v, and thus $\overline{\partial}\phi = 0$ as a distribution. Thus $P^\perp \subset Q(X)$, and the reverse inclusion is obvious.

Now think of $\mu \in M(X)$ as representing an infinitesimal quasiconformal mapping $f : X \to X'$ with dilatation $\epsilon\mu$. Then

$$d(X, X') = \frac{1}{2}\log K(f) = \frac{1}{2}\log\frac{1+\epsilon\|\mu\|}{1-\epsilon\|\mu\|} = \epsilon\|\mu\| + O(\epsilon^2).$$

Thus the L^∞ norm on $M(X)$ gives the dilatation of f. If we vary μ by elements of $M_0(X)$, we obtain all maps that are infinitesimally isotopy to f. Thus the infinitesimal form of the Teichmüller metric is just the quotient norm on $M(X)/M_0(X)$. Similarly, the cometric on $Q(X)$ is just the induced norm from $L^1(X, dz^2)$.

Summarizing, we have:

Theorem 4.12 *For any $X \in \operatorname{Teich}(S)$, we have*

$$\begin{aligned} T_X \operatorname{Teich}(S) &\cong M(X)/Q(X)^\perp \quad \textit{and} \\ T_X^* \operatorname{Teich}(S) &\cong Q(X). \end{aligned}$$

The quotient L^∞-norm on $M(X)/Q(X)^\perp$ is the Teichmüller metric, and the L^1-norm metric on $Q(X)$ is the Teichmüller cometric.

Note that these Banach spaces are generally *not* Hilbert spaces, and thus the Teichmüller metric is generally not a Riemannian metric.

The factors 1/2 and 4. The definition of the Teichmüller metric as $d_T(X,Y) = (1/2)\inf \log K(f)$ is compatible with the *natural* L^1 and L^∞ norms on the tangent and cotangent space used above. The same factor arises if we use the Euclidean metric on $\mathbb{C}$ to give a metric ρ on $T_0\Delta$; then ρ agrees at $z=0$ with the metric of constant curvature -4, not -1.

Geodesics and cotangents. At first sight it seems strange that a Teichmüller geodesic is associated to a quadratic differential $\phi \in Q(X)$, because $Q(X)$ is the *co*tangent space to Teichmüller space (rather than the tangent space). The explanation is that the Teichmüller metric, while not Riemannian, gives a nonlinear duality between $\mathbb{P}Q(X)$ and $\mathbb{P}M(X)/Q(X)^\perp$. Namely to each line $\mathbb{R}\phi \subset Q(X)$ there is associated the supporting hyperplane H_ϕ for the unit ball $B_X = \{\phi : \|\phi\| \le 1\}$. This hyperplane is the kernel of the Beltrami differential $\mu = \overline{\phi}/|\phi|$, and μ is tangent to the Teichmüller geodesic that stretches along ϕ.

4.7 A novel formula for the Poincaré metric

Theorem 4.13 *The hyperbolic metric $\rho(t)|dt|$ on $\widehat{\mathbb{C}} - \{0,1,\infty\}$ is given by*

$$\rho^{-1}(t) = 2|t(t-1)| \int_{\widehat{\mathbb{C}}} \frac{|dz|^2}{|z||z-1||z-t|}.$$

Proof. Think of $t \in \widehat{\mathbb{C}} - \{0,1,\infty\}$ as determining a Riemann surface $X_t = \widehat{\mathbb{C}} - \{0,1,\infty,t\}$ in the moduli space of a 4-times punctured sphere. The Teichmüller metric has constant curvature -4 so up to a factor of 2 it agrees with the Poincaré metric $\rho(t)$. On the other hand, a vector $v \in T_t\widehat{\mathbb{C}}$ determines a deformation of X_t whose Teichmüller length is given by $|\operatorname{Res}_t(v\phi)|/\|\phi\|$, where $\phi \in Q(X_t)$ is any nonzero holomorphic quadratic differential. Taking $\phi(z) = dz^2/(z(z-1)(z-t))$, we obtain the formula above. ∎

4.8 The Kobayashi metric

Why is the Teichmüller metric important? Given any complex manifold M, the *Kobayashi metric* on M is defined as the largest metric such that every holomorphic map

$$f : \Delta \to M$$

satisfies $\|f'(0)\| \le 1$. For example, on the disk itself, the Kobayashi metric is $|dz|^2/(1-|z|^2)$, the multiple of the hyperbolic metric that gives constant curvature -4.

The Kobayashi metric is generally not a Riemannian metric; rather, it is a *Finsler metric* defined by a norm on each tangent space to M. The manifold M is *Kobayashi hyperbolic* if this metric is nowhere degenerate. A nice criterion for hyperbolicity is:

Theorem 4.14 *A compact complex manifold M is Kobayashi hyperbolic iff every holomorphic map $f : \mathbb{C} \to M$ is constant.*

Cf. [La, III, §2]. For example, if M is covered by a bounded domain in $\mathbb{C}^n$, then it is Kobayashi hyperbolic.

From the definition it is easy to see:

Theorem 4.15 *Any holomorphic map between complex manifolds is distance non-increasing for the Kobayashi metric.*

Now for Teichmüller space we have:

Theorem 4.16 (Royden) *The Teichmüller metric on* $\mathrm{Teich}(S)$ *coincides with the Kobayashi metric.*

Corollary 4.17 *Any holomorphic map* $f : \mathrm{Teich}(S) \to \mathrm{Teich}(S)$ *satisfies*

$$d(fX, fY) \le d(X, Y)$$

in the Teichmüller metric.

This Corollary is a Schwarz lemma for Teichmüller space. It indicates that the Teichmüller metric is well-adapted to the problem of finding the *fixed-point* of an iteration on Teichmüller space.

See [La] for more details on Kobayashi hyperbolicity.

4.9 Moduli space

The *moduli space* $\mathcal{M}(S) = \mathrm{Teich}(S)/\mathrm{Mod}(S)$ is obtained from Teichmüller space by forgetting the marking of S. Moduli space is an orbifold with exactly one point for each isomorphism class of Riemann surface of the type specified by S. The orbifold structure arises because $\mathrm{Mod}(S)$ does not act freely; for any $X \in \mathrm{Teich}(S)$, the stabilizer of X in $\mathrm{Mod}(S)$ is isomorphic to the conformal automorphism group $\mathrm{Aut}(X)$.

Example. For a torus, $\mathrm{Mod}(S)$ is the $(2, 3, \infty)$ orbifold $\mathbb{H}/\mathrm{SL}_2(\mathbb{Z})$; it is isomorphic to the quotient of $\widehat{\mathbb{C}} - \{0, 1, \infty\}$ by its S_3 automorphism group.

Theorem 4.18 (Mumford) *Let* $M_L \subset \mathcal{M}(S)$ *be the set of Riemann surfaces whose shortest closed hyperbolic geodesic is of length* $\ge L > 0$. *Then* M_L *is compact.*

Proof. For any $X \in \mathcal{M}(S)$, we have

$$\mathrm{area}(X) = 2\pi|\chi(S)|$$

by Gauss-Bonnet. If X belongs to M_L, then balls of radius $L/10$ on X are embedded, and by the area bound there is an $N = N(L, S)$ such that X is covered by N closed balls of radius $L/10$.

Given a sequence $X_n \in M_L$, choose such a covering $\mathcal{B}_n$ for each n. Passing to a subsequence, we can assume the nerve of $\mathcal{B}_n$ is constant. (The *nerve* is the finite complex with a vertex for each ball and a k-simplex for each k-tuple of balls with nonempty intersection).

Now choose a baseframe at the center of each ball. Whenever two balls meet, we can lift them uniquely to $\mathbb{H}^2$ so the first frame goes to a fixed standard frame (e.g. at the origin in the ball model for $\mathbb{H}^2$); then there is a unique isometry of $\mathbb{H}^2$ moving the first frame to the second. In this way the oriented edges of the nerve of $\mathcal{B}_n$ are naturally labeled by hyperbolic isometries.

These isometries range in a compact subset of $\mathrm{Isom}(\mathbb{H}^2)$, so after passing to a subsequence we can assume they converge. Along the way a cocycle condition is satisfied along the boundary of any 2 simplex of the nerve, so the cocycle condition holds in the limit by continuity. Thus the limiting data determines charts for a hyperbolic manifold X_∞, and X_∞ is closed since it is covered by a finite number of balls $\mathcal{B}_\infty$.

Now each ball in $\mathcal{B}_\infty$ is naturally associated to a ball in $\mathcal{B}_n$ for $n \gg 0$. Using baseframes, we obtain isometries between associated balls. Using a partition of unity on X_∞, we can piece these isometries together to obtain a continuous map $f_n : X_\infty \to X_n$ for all $n \gg 0$. Since the gluing data for X_n converges to that of X_∞, the isometries relating adjacent balls are nearly the same, so f_n is a nearly isometric diffeomorphism for all $n \gg 0$. In particular, $K(f_n) \to 1$, and thus the Teichmüller distance from X_n to X_∞ tends to zero.

Thus $X_\infty \in M_L$ and we have shown every sequence in M_L has a convergent subsequence. ∎

4.10 The mapping-class group

In this section, following [Bers2], we give the proof of:

Theorem 4.19 (Thurston) *Any mapping-class $[f] \in \mathrm{Mod}(S)$ has a reprentative which is: finite order, reducible, or pseudo-Anosov.*

Pseudo-Anosov mappings. Here f is *reducible* if there is a multicurve $C \subset S$ whose components are permuted by f.

A mapping $f : X \to X$ on a Riemann surface is *pseudo-Anosov* if f is a Teichmüller mapping of dilatation $K^2 > 1$ and its initial and terminal quadratic differentials (normalized so $\int |q| = 1$) coincide. Note that the Riemann surface X is not unique — it can be chosen as any point on the loop in $\mathcal{M}_g$ determined by f.

With this in mind, we say $f : S \to S$ is pseudo-Anosov it is preserves a pair of transverse, measured foliations, multiplying the transverse measure on one by K and on the other by $1/K$.

This structure permits a detailed analysis of the topological and measure-theoretic properties of f. For example:

Proposition 4.20 *A pseudo-Anosov mapping is ergodic.*

Proof. (Hopf.) The ergodic sums of T^n for $n > 0$ are clearly constant along the leaves of one foliation, and the sums for $n < 0$ are constant along the leaves of the other. Thus their limit is constant. ■

Proposition 4.21 *The period points for f are dense.*

Proof. Take a small square $S \subset X$. By Poincaré recurrence, we can find a large n such that $f^n(S)$ meets S. Thus $f^n(3S)$ passes near the center of $3S$, giving a long thin rectangle cutting through two of its sides. It follows easily that $f^n|3S$ has an invariant leaf and there is a fixed point on this leaf. ■

A more detailed analysis shows:

Proposition 4.22 *The map f is mixing, and each of its invariant foliations is uniquely ergodic.*

Proof of Theorem 4.19. Connect X to fX by a path in $\mathcal{T}_g$ and project to obtain a loop $\gamma \subset \mathcal{M}_g$ of finite length. Now try to shrink this loop to make it as short as possible. If it shrinks to a point, f is of finite order; if it goes out the end of moduli space, f is reducible; and if it shrinks to a geodesic, then f is pseudo-Anosov.

For a more precise analysis, observe that in the last case we can find an $X \in \mathcal{T}_g$ minimizing $d(X, fX)$. Let Y be the midpoint of the geodesic from X to fX; then we must have

$$d(Y, fY) = d(Y, fX) + d(fX, fY).$$

But for a pair of linear maps, we have $K(AB) < K(A)K(B)$ unless the expanding and contracting directions coincide. This shows the initial and terminal foliations of f coincide. ■

4.11 Counterexamples

1. *There exist normalized K-quasiconformal maps $f_n : \mathbb{C} \to \mathbb{C}$ such that $\mu(f_n) \to 0$ weakly but f_n does not converge uniformly to the identity.*

   Fix $K > 1$ and let $f(x + iy) = x + s(y)$, where s is a piecewise-linear function with $s(0) = 0$ and

   $$s'(y) = \begin{cases} K & \text{on } [2n, 2n+1) \text{ and} \\ 1/K & \text{on } [2n+1, 2n+2). \end{cases}$$

   Then the dilatation $\mu(f)$ is constant on horizontal strips of unit width, but alternating in sign. Thus if we set $f_n(z) = nf(z/n)$, then $\mu(f_n)$ alternates sign on strips of width $1/n$. Moreover, f_n fixes $\{0, 1, \infty\}$ and $\mu(f_n) \to 0$ in the weak topology on $L^\infty(\mathbb{C})$ (that is, $\int \mu(f_n)g \to 0$ for any $f \in C_0^\infty(\mathbb{C})$). But f_n converges uniformly to the real-linear map $F(x + iy) = x + iLy$, where $L = (K + 1/K)/2$.

2. *There exist locally affine maps $f : X \to Y$ between Riemann surfaces of finite type which are not Teichmüller mappings.*

   Fix $\lambda > 1$ and let $X = \mathbb{C}^*/\lambda^{\mathbb{Z}}$ be a complex torus. Define $f_t : \mathbb{C}^* \to \mathbb{C}^*$ by
   $$f_t(x + iy) = x + e^t y.$$
   Then $f_t(\lambda z) = \lambda f_t(z)$, and $F = f_1 : X \to X$ is a locally affine map. But $f_0(z) = z$, so F is isotopic to the identity, and thus it is not the Teichmüller mapping in its homotopy class.

   The vertical lines of maximal stretch of F $(t > 0)$ descend to give a Reeb foliation $\mathcal{F}$ of X; there are two closed leaves, coming from the imaginary axis, and all other leaves of $\mathcal{F}$ spiral from one closed leaf to the other. Because of the spiraling, this foliation admits no transverse invariant measures other than δ-masses on the closed leaves.

   Note that $F : X \to X$, while locally affine, is *not* Anosov — it preserves a splitting of the tangent space, but the splitting is not along the stable and unstable manifolds of F.

4.12 Bers embedding

In this section we discuss the Teichmüller space of a quite general Riemann surface X, and its embedding as a domain in a complex Banach space.

Ideal boundary. Let $X = \mathbb{H}/\Gamma_X$ be a hyperbolic Riemann surface, presented as the quotient of the upper halfplane by a Fuchsian group. We do not assume that X has finite volume.

The group Γ_X acts on the whole sphere, with limit set $\Lambda \subset \widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. In particular the quotient of the lower halfplane,
$$X^* = (-\mathbb{H})/\Gamma_X,$$
is the *complex conjugate* of X. The map $z \mapsto \overline{z}1$ descends to a natural anticonformal isomorphism $X \to \overline{X}$, identifying their fundamental groups.

Let $\Omega = \widehat{\mathbb{R}} - \Lambda$ be the domain of discontinuity of Γ_X acting on $\widehat{\mathbb{R}}$. Then we can form a partial compactification of X by taking the quotient
$$\overline{X} = (\mathbb{H} \cup \Omega)/\Gamma_X.$$
We refer to ideal-$\partial X = \overline{X} - X$ as the *ideal boundary* of X.

As a special case, ideal-$\partial\mathbb{H} = \widehat{\mathbb{R}}$. Any quasiconformal $f : \mathbb{H} \to \mathbb{H}$ extends to a homeomorphism of $\overline{\mathbb{H}}$, and thus any quasiconformal map $f : X \to Y$ extends to a homeomorphism $f : \overline{X} \to \overline{Y}$.

Homotopy and isotopy. The key to defining the Teichmüller space is to know which quasiconformal maps $f : X \to X$ are to be considered trivial. Luckily several natural notions coincide.

Theorem 4.23 *Let $X = \mathbb{H}/\Gamma_X$ and let $f : X \to X$ be a quasiconformal map. The following are equivalent.*

1. *There is a lift of f to a map $\mathbb{H} \to \mathbb{H}$ that extends to the identity on $\widehat{\mathbb{R}}$.*

2. *The map f is homotopic to the identity* rel *ideal boundary.*

3. *The map f is isotopic to the identity* rel *ideal boundary, through uniformly quasiconformal maps.*

See [EaM].

A *Riemann surface marked by X* is a pair (Y, g) where $g : X \to Y$ is a quasiconformal map. Two marked Riemann surfaces (Y, g) and (Z, h) are *equivalent* if there is a conformal isomorphism $\alpha : Y \to Z$ such that

$$f = h^{-1} \circ \alpha \circ g : X \to X$$

is isotopic to the identity *rel* ideal boundary. (Of course any of the other conditions above give the same definition).

The *Teichmüller space* of X is the space of equivalence classes of Riemann surfaces (Y, g) marked by X.

Example: universal Teichmüller space. The case where Γ_X is the trivial group, and $X = \mathbb{H}$¡ is called *universal Teichmüller space*. In this case we have

$$\operatorname{Teich}(\mathbb{H}) \cong \operatorname{QS}(\widehat{\mathbb{R}}) / \operatorname{PSL}_2(\mathbb{R}),$$

where $\operatorname{QS}(\widehat{\mathbb{R}})$ is the group of quasi-symmetric maps $g : \widehat{\mathbb{R}} \to \widehat{\mathbb{R}}$. These are exactly the homeomorphisms of $\widehat{\mathbb{R}}$ that arise as boundary values of quasiconformal maps.

Quasifuchsian groups and univalent maps. Fix a model $X = \mathbb{H}/\Gamma_X$ for the universal cover of X. The Bers embedding of $\operatorname{Teich}(X)$ is based on a construction that canonically associates to any $(Y, g) \in \operatorname{Teich}(X)$:

- A Kleinian group Γ_Y acting on the Riemann sphere; and
- A quasiconformal map $F : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, fixing $\{0, 1, \infty\}$ and conjugating Γ to Γ_Y; such that
- F is conformal in $(-\mathbb{H})$, $Y = F(\mathbb{H})/\Gamma_Y$ and $X^* = F(-\mathbb{H})/\Gamma_Y$.

Because of the last property, the group Γ_Y is said to *simultaneously uniformize* X^* and Y.

The construction is the following. Let $g : X \to Y$ be a Riemann surface marked by X. Let $\mu = \mu(g)$, lifted to $\mathbb{H}$ and extended to $(-\mathbb{H})$ by 0. Let $F : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be the normalized solution to the Beltrami equation $\mu(F) = \mu$. Since μ is Γ_X-invariant, we have $\mu(F \circ \gamma) = \mu(F)$ for all $\gamma \in \Gamma_X$. But the solution to the Beltrami equation is unique up to post-composition with a Möbius transformation, so there is an isomorphism

$$\rho : \Gamma_X \to \Gamma_Y \subset \operatorname{Aut} \widehat{\mathbb{C}}$$

such that

$$F \circ \gamma = \rho(\gamma) \circ F$$

for all $\gamma \in \Gamma_X$.

Thus F conjugates Γ_X to Γ_Y, and by construction F is conformal in $(-\mathbb{H})$.

Since F intertwines Γ_X and Γ_Y, the Schwarzian derivative $\phi_Y = SF|(-\mathbb{H})$ is Γ_X-invariant. Since F is univalent, we have $\phi_Y \in P(X^*)$; in fact $\|\phi_Y\| < 3/2$.

The map $\beta : (Y, g) \mapsto \phi_Y$ gives the *Bers embedding*

$$\beta : \mathrm{Teich}(X) \to P(X^*).$$

Image of Bers embedding. Let us write $S(X^*) \subset P(X^*)$ for the set of Γ_X-invariant Schwarzians of univalent (*schlicht*) maps, and $T(X^*)$ for the image of Bers embedding. Then we have

$$T(X^*) \subset S(X^*) \subset P(X^*).$$

In fact, $T(X^*)$ is exactly the set of Schwarzians of univalent maps to quasidisks. By the Ahlfors-Weill extension, the ball of radius $1/2$ is contained in $T(X^*)$. With similar reasoning, one can also show that $T(X^*)$ is open. Thus the Bers embedding provides a model for $T(X^*)$ as a *complex manifold*, in facts as a bounded domain in a complex Banach space with

$$B(1/2) \subset T(X^*) \subset B(3/2).$$

Role of the Ahlfors-Weill extension. If Y is close enough to X that $\|\phi_Y\| < 1/2$, then we can extend $F|(-\mathbb{H})$ to $\mathbb{H}$ using the Ahlfors-Weill construction. The result is a *canonical* quasiconformal map $g_Y : X \to Y$, characterized by having a harmonic Beltrami coefficient

$$\mu(g_Y) = \rho^{-2}\phi_Y(\overline{z}).$$

Notice that $\mu(g_Y) \in M(X)$ is a *holomorphic* function of Y, since it is a holomorphic function of ϕ_Y (for each fixed z). In other words, the Ahlfors-Weill extension provides a holomorphic $\mathrm{Teich}(X) \dashrightarrow M(X)$ defined near X and lifting the natural holomorphic projection $M(X) \to \mathrm{Teich}(X)$.

Finite sets on the sphere. The Teichmüller space $T_{0,n}$ of a sphere with $n \geq 3$ punctures has an alternative and very simple complex model: we have

$$T_{0,n}/G_n = \{z \in \mathbb{C}^{n-3} \;:\; z_i \neq z_j \neq 0, 1, \infty \;\forall i, j\},$$

where G_n is the pure mapping-class group of n points on the sphere (the pure braid group modulo its center). This group acts without fixed-points by automorphism of $T_{0,n}$, so Teichmüller space arises as the universal cover of the complement of suitable hyperplanes in $\mathbb{C}^{n-3}$.

The complex structure on $T_{0,n}$ derived from $\mathbb{C}^{n-3}$ agrees with that coming from the Bers embedding. To see this consider a Beltrami differential μ on $X = \widehat{\mathbb{C}} - E$. Then the solution to the Beltrami equation at any point, $f_\mu(z)$, is holomorphic in μ; thus $Y = \widehat{\mathbb{C}} - f_\mu(E)$ is the complement of a holomorphically moving finite set. At the same time, pulling μ back to $-\mathbb{H}$, the quadratic differential $\phi_Y = SF_\mu$ is a holomorphic function of μ as well. In other words, the pair of maps $M(X)_1 \to \mathbb{C}^{n-3}$ and $M(X)_1 \to T_{0,n}$ are both holomorphic submersions, so the transition between them is also holomorphic.

4.13 Conjectures on the Bers embedding

In general one hopes to show that all Kleinian groups are limits of geometrically finite ones. A precise conjecture in this direction for quasifuchsian groups was made by Bers.

Conjecture 4.24 *For all finite-volume hyperbolic surfaces X, we have $\overline{T(X)} = S(X)$.*

K. Bromberg has recently announced a proof of this central conjecture.

This conjecture can also be formulated for other Riemann surfaces, including the unit disk.

It is false for universal Teichmüller space. Let us say a compact quasiarc $Q \subset \mathbb{C}$ is *incorrigible* if the Hausdorff limits of $\mathrm{Aut}(\widehat{\mathbb{C}}) \cdot Q$ in the space of compact subsets of $\widehat{\mathbb{C}}$ include no circle. Then we have [Th]:

Theorem 4.25 (Thurston) *There exist isolated points in $S(\Delta)$. In fact, the Riemann mapping $f : \Delta \to \widehat{\mathbb{C}} - Q$ is isolated in $S(\Delta)$ for any incorrigible quasiarc Q.*

Such quasiarcs Q are also called zippers, because they lock two sides of the plane together. The 'Bers density conjecture' can be formulated as saying there are no group-invariant zippers, at least for finitely-generated nonelementary Kleinian groups.

4.14 Quadratic differentials and interval exchanges

Let $q \in Q(X)$ be a holomorphic quadratic differential on a compact Riemann surface X. In local coordinates where $q = dz^2$, we have a foliation by vertical lines together with a transverse invariant measure $|dx|$. These objects patch together to yield a measured *vertical foliation* $\mathcal{F}(q)$ on X. The corresponding *horizontal foliation* is $\mathcal{F}(-q)$.

Intrinsically, the tangent space to the vertical $\mathcal{F}(q)$ is given by the vectors $v \in TX$ such that $q(v, v) < 0$. If $q = \omega^2$ happens to be the square of a holomorphic 1-form, then $\mathcal{F}(q)$ is defined by the closed harmonic form $\rho = \mathrm{Re}(\omega)$. That is, $T\mathcal{F} = \mathrm{Ker}\,\rho$ and $|\int_I \rho| = \mu(I)$ gives the transverse measure.

In this section we discuss the dynamics of the foliation $\mathcal{F}$. In other words, we discuss the asymptotic distribution of very long leaves of $\mathcal{F}$. The usual ergodic theorems apply to the undirected 'flow' along the leaves of $\mathcal{F}$. Indeed, after passing to a branched double cover of (X, q) we can assume $q = \omega^2$, in which case the leaves are coherently oriented by $\mathrm{Im}\,\omega$ and we obtain a flow in the usual sense.

See [MT] for more details.

Cylinders and saddle connections. A *saddle connection* is a leaf of $\mathcal{F}$ joining a pair of its zeros. A *cylinder* is an open annulus $C \subset X$ foliated by closed leaves of $\mathcal{F}$. So long as X is not a torus, it carries a canonical collection of disjoint maximal cylinders C_i whose boundaries are unions of saddle connections. The locus $\bigcup C_i$ coincides with the union of the (smooth) closed leaves of $\mathcal{F}$.

The foliation $\mathcal{F}$ is *aperiodic* if it has no cylinders. It is *minimal* if every leaf disjoint from the zeros of q is dense.

Interval exchange. The first return map $f : I \to I$ to any transversal arc to the vertical foliation $\mathcal{F}$ (often taken to be an interval along the horizontal foliation) is an *interval exchange transformation*: there is a partition $I = \bigcup I_i$ into disjoint subintervals such that $f(x) = x + t_i$ on I_i.

Here the metric $|q|$ is used to identify I with an interval in $\mathbb{R}$. The discontinuities of f come from the zeros of q and the endpoints of I. Note that there are only finitely many vertical arcs the start on I and terminate on a zero of q without meeting any other zeros en route. It is useful to choose a convention for continuing the flow past such points, e.g. by always turning to the right, to make f well-defined even at the endpoints of the intervals I_i.

The map f is *aperiodic* if it has no periodic points, and *minimal* if every orbit is dense. A transversal is *full* if it meets every leaf of $\mathcal{F}$; in this case it inherits aperiodicity and minimality from $\mathcal{F}$.

Invariant measures. By definition, f preserves Lebesgue measure $dx|I$. It may however have other invariant measures ν. These correspond to other transverse invariant measures for $\mathcal{F}$. If ν has full support and no atoms, we can integrate it to obtain a topological conjugacy from f to another interval exchange g, such that nu becomes dx.

Lack of mixing.

Theorem 4.26 *An interval exchange is never mixing.*

Proof. Let $f : I \to I$ be an interval exchange with n subintervals. We may assume f is aperiodic, otherwise it is clearly not mixing.

Let $J \subset I$ be a subinterval; then the first return map gives an induced interval exchange $f^{n_i} : J_i \to J$, $\bigcup J_i = J$. Since the discontinuities of the first return map come from the endpoints of I and the discontinuities of f, the induced map has at most $n+2$ subintervals.

Note that I is partitioned into intervals of the form $J_i^m = f^m(J_i)$.

Repeat the process by taking the first return map to each J_i. The result is an interval exchange $f^{n_{ij}} : J_{ij} \to J_i$, where again j takes on at most $(n+2)$ values.

Now the key point is that we have

$$T^{n_{ij}}(J_{ij}^m) \subset J_i^m,$$

which implies

$$J_i^m \subset \bigcup T^{-n_{ij}}(J_i^m).$$

Note that there are at most $(n+2)$ terms on the right. Since preimages respect set operations, if A is any set that is a union of interval of the form J_i^m, we have

$$A \subset \bigcup T^{-n_{ij}}(A).$$

But this implies

$$\mu(A \cap T^{-n_{ij}}(A)) > \mu(A)/(n+2)^2 \tag{4.2}$$

for some ij.

Now let $B \subset I$ be an interval of length ϵ. Then if we have mixing, there is an N such that $\mu(B \cap T^{-n}(B)) \approx \epsilon^2 \ll \mu(B)/(n+2)^2$ for all $n > N$. Choose J such that all $n_{ij} > N$, and so that the union A of the intervals $J_i^m \subset B$ has measure at least $\epsilon/2$. Then we obtain a contradiction to the frequent return guaranteed by (4.2). ∎

Corollary 4.27 *The geodesic flow for (X, q), even restricted to geodesics of a fixed slope, is never mixing.*

Theorem 4.28 *An aperiodic interval exchange $f : I \to I$ with n subintervals has at most n ergodic invariant probability measures.*

Proof. An invariant measure μ is determined by $\mu(I_i)$, and the ergodic invariant measures map to linearly independent vectors $\nu_j(I_i)$ in $\mathbb{R}^n$. ∎

4.15 Unique ergodocitiy for quadratic differentials

In this section we will establish:

Theorem 4.29 (Masur) *Given $q \in Q(X)$, suppose the Teichmüller ray from X to $[\mathcal{F}(q)] \in \mathbb{P}\mathcal{ML}_g$ is recurrent when projected to $\mathcal{M}_g$. Then $\mathcal{F}(q)$ is uniquely ergodic.*

Pseudo-Anosov case. Let $\psi : X \to X$ be a pseudo-Anosov diffeomorphism, adapted to the quadratic differential q. Let $\lambda > 1$ be its expansion factor, so the the leaves of $\mathcal{F} = \mathcal{F}(q)$ are *contracted* by λ under ψ. We begin by showing:

Theorem 4.30 *The vertical invariant foliation $\mathcal{F}$ for a pseudo-Anosov transformation is uniquely ergodic.*

Invariant measures. Let $\nu_1, \ldots, \nu_n$ be the ergodic invariant probability measures for $\mathcal{F}$. These are measures on X, invariant under parallel transport along leaves, such that for ν_i-almost every $x \in X$ and every $\phi \in C(X)$ we have

$$\frac{1}{L} \int \phi(x)|dq| \to \int_X \phi\, \nu_i \tag{4.3}$$

as $|L| \to \infty$. Here L is any segment of a leaf of $\mathcal{F}$ with $x \in L$.

Note that we can write the Lebesgue measure $|q| = \sum a_i \nu_i$ for some $a_i \geq 0$. We will exploit the property that $|q|$ is invariant under ψ; note that we do not know, *a priori*, if this is true for any individual ν_i.

Let $E_i \subset X$ be the set of points x where (4.3) holds. We say such points are 'generic' for ν_i. Note that $E = \bigcup E_i$ has full Lebesgue measure, and the $E_i \cap E_j = \emptyset$ if $i \neq j$.

Renormalization. The next idea is to use ψ to *renormalize* the geodesic flow along the leaves of $\mathcal{F}$. The point is that the behavior of a long segment L can be related to the behavior of the shorter segment $\psi^n(L)$, which satisfies $|\psi^n(L)| = \lambda^{-n} L$.

A *rectangle* $R \subset (X, |q|)$ is a closed region isometric to a Euclidean rectangle, bounded by segments of the vertical and horizontal foliations of q. Note that every point of X (including the zeros of q) lies in at least one rectangle, and that any two points can be joined by a chain of rectangles.

Lemma 4.31 *Given $x_1, x_2 \in E$, suppose there is a sequence $N \subset \mathbb{N}$ such that $\psi^n(x_i) \to y_i$ as $n \to \infty$ in N, y_1, y_2 lie in a rectangle R, and $\psi^n(x_i) \in R$ for all $n \gg 0$ in N. Then x_1 and x_2 are generic for the same measure ν_j.*

Proof. Let S_i be the leaf of $\mathcal{F}$ running the length of R and passing through y_i. Then for $n \in N$, the parallel leaves $L_n(i) = \psi^{-n}(S_i)$ passing through x_i have length $\asymp \lambda^n$ and are separated by $O(\lambda^{-n})$ in the $|q|$-metric. Thus the averges in (4.3) for $L = L_n(i)$ converge to the same value as $n \to \infty$, and therefore the ergodic measures for x_1 and x_2 are the same. ∎

Lemma 4.32 *Given a sequence $N \subset \mathbb{N}$ and a nonempty open set $U \subset X$, there is an $x \in \bigcup E_i$ and a further subsequence $N' \subset N$ along which $\psi^n(x) \to y \in U$.*

Proof. Let $K \subset U$ be a compact set of positive measure, and let $K_n = \psi^{-n}(K)$. Then since ψ is measure preserving, the sum over $n \in N$ of the $|q|$-measure of K_n diverges, and thus (by Borel-Cantelli) there is a set of $A \subset E$ of positive Lebesgue measure consisting of points with $\psi^n(x) \in K$ for infinitely many $n \in A$. Passing to a convergent subsequence yields the lemma. ∎

Proof of Theorem 4.29. Let u_1 and u_2 be generic for a pair of ergodic invariant measures ν_1 and ν_2. We will show $\nu_1 = \nu_2$.

First, pass to a subsequence such that $\psi^n(u_i) \to v_i \in X$. Next, choose a chain of overlapping rectangles connecting v_1 to v_2. In case v_1 or v_2 is a zero of q, choose the initial or terminal rectangle so it also contains $\psi^n(u_i)$ for infinitely many n.

Let $U_1, \ldots, U_m$ be open sets in the overlap of adjacent rectangles. Then by the lemma above, we can find $y_j \in U_j$ and $x_j \in E$ such that $\psi^n(x_j) \to y_j$ along a further subsequence. By Lemma 4.31, v_1 is generic for the same measure as y_1. Similarly y_i and y_{i+1} are generic for the same measure, as are y_m and v_2. Thus $\nu_1 = \nu_2$, and therefore $\mathcal{F}$ is uniquely ergodic. ∎

4.16 Hodge theory

Real multiplication in higher genus. [Mo], [Sch].

Theorem 4.33 *Let $V \subset \mathcal{M}_g$ be a Teichmüller curve generated by (X,ω), where the trace field K of $\mathrm{SL}(X,\omega)$ has degree g over $\mathbb{Q}$. Then $\mathrm{Jac}(X)$ admits real multiplication by K with ω as an eigenform.*

Proof. We will use a basic fact about variations of Hodge structures: if $\sigma = \sum \sigma^{p,q}$ is a flat section of a Hodge bundle (over a quasiprojective base such as V), then its components $\sigma^{p,q}$ are also flat. (See e.g. Schmid, Invent. math., 1973).

Let $H \to V$ be the Hodge bundle whose fiber over t is $H^1(X_t,\mathbb{Z}) \otimes \mathbb{C}$. Let H_0 be the fiber over a basepoint $t = 0$. Then under the action of $\pi_1(V)$, the space H_0 decomposes into 2-dimensional irreducible subspaces

$$H_0 = S_1 \oplus \cdots \oplus S_g,$$

where S_1 is spanned by ω and $\overline{\omega}$. The full commutant of $\pi_1(V)$ is the abelian algebra $\mathbb{C}^g$ acting diagonally and preserving each S_i.

Let $P_0 : H_0 \to H_0$ be projection onto S_i. Since P_0 commutes with the action of $\pi_1(V)$, it extends to a flat section P of the (weight zero) Hodge bundle $\mathrm{Hom}(H,H)$. By Schmid's theorem, the Hodge components P^{ij} of P are also flat. Thus $P_0^{ij} = P|H_0$ commutes with $\pi_1(V)$, and hence it is diagonal. But $P^{-1,1}$ maps $H^{1,0}$ into $H^{0,1}$, so it is "off-diagonal"; for example, its square is zero. Consequently $P^{-1,1} = P^{1,-1} = 0$, and thus $P = P^{0,0}$. This implies that P preserves $H^{1,0}$, and therefore each subspace S_i decomposes as a direct sum $S_i^{1,0} \oplus S_i^{0,1}$. Thus $\mathrm{Jac}(X_0)$ admits real multiplication by the trace field of $\mathrm{SL}(X,\omega)$. ∎

5 Dynamics of rational maps

This chapter develops the basic picture of the dynamics of a single rational map, and the Teichmüller theory of deformations of its deformations.

5.1 Dynamical applications of the hyperbolic metric

All rational maps will be of degree $d > 1$.

Exceptional points. A set $E \subset \widehat{\mathbb{C}}$ is *exceptional* if E is finite and $f^{-1}(E) \subset E$. A single point z is exceptional if it belongs to an exceptional set; equivalently, if its inverse orbit $\bigcup f^{-n}(z)$ is finite.

Theorem 5.1 *Any rational map f has a maximal exceptional E, $|E| \leq 2$ and $E \cap J(f) = \emptyset$.*

Proof. Consider any exceptional set E. Then f maps $X = \widehat{\mathbb{C}} - E$ into itself, so $\chi(X) \leq \deg(f)\chi(X)$ by Riemann-Hurwitz. Therefore $\chi(X) = 2 - |E| \geq 0$, and $|E| \leq 2$. Thus the number of exceptional points is at most two, and the union of all exceptional points gives the maximal exceptional set.

Since every $z \in \widehat{\mathbb{C}}$ has d preimages counted with multiplicity, any $z \in E$ is a critical value, and therefore E consists of superattracting cycles. In particular E is disjoint from the Julia set. ∎

Since $J(f)$ is totally invariant we have:

Corollary 5.2 *The Julia set is infinite.*

Examples. The maximal exceptional set is $E = \{0, \infty\}$ for $f(z) = z^d$, and $E = \{\infty\}$ if $f(z)$ is a polynomial not conjugate to z^d. These rational maps represent all the cases where $E \neq \emptyset$ (up to conjugacy).

Attractors and critical points. We next use the Schwarz lemma to demonstrate a close connection between critical points and attractors of f. In particular we shall see all but finitely many cycles of a rational map are repelling. (On the other hand, it is known that smooth dynamical systems — even polynomial maps in dimension two — can have infinitely many attracting cycles.)

Theorem 5.3 *A rational map of degree d has $2d - 2$ critical points, counted with multiplicity.*

Proof. (1) (Riemann-Hurwitz): Letting N be the number of branch points counted with multiplicity. We have $\chi(\widehat{\mathbb{C}}) = \deg(f) \cdot \chi(\widehat{\mathbb{C}}) - N$, so $N = 2d - 2$.

(2) Write $f(z) = P(z)/Q(z)$ with $\deg P = \deg Q = d$ (after a generic conjugation). Then the critical points are zeros of $PQ' - P'Q$. The leading terms of this polynomial cancel, resulting in a polynomial of degree $2d - 2$.

(3) Let V be a holomorphic vector field on $\widehat{\mathbb{C}}$; then V has 2 zeros and no poles. Now consider $f^*(V)$: it has zeros at the $2d$ preimages of these zeros, and poles at the critical points of f. Since the number of zeros minus the number of poles of $f^*(V)$ is still two, we have $2 = 2d - |C(f)|$. ∎

Theorem 5.4 *Every attracting periodic cycle of f attracts a critical point.*

Proof 1. Let $U \subset \widehat{\mathbb{C}}$ be the open set of points attracted to a given cycle. Since $J(f)$ is infinite and disjoint from U, the set U is hyperbolic. If U contains no critical point, then $f : U \to U$ is a covering map, hence an isometry for the hyperbolic metric. But then $|f'|_U = 1$, contrary to the existence of an attracting cycle.

Proof 2. For a superattracting cycle the result is immediate. Otherwise f acts on the immediate basin U of the cycle by a proper local homeomorphism, hence a covering map. Clearly U/f is a torus, so U is a covering space of a torus. But then U is isomorphic to $\mathbb{C}$ or $\mathbb{C}^*$, and these possibilities are easily ruled out. ∎

Corollary 5.5 *The number of attracting cycles is at most* $2\deg(f) - 2$.

Example. A quadratic polynomial $f_c(z) = z^2 + c$ has at most one attracting cycle in the complex plane, and this cycle can be located by iterating f_c on $z = 0$.

Corollary 5.6 *A rational map f of degree d has at most $4d - 4$ non-repelling cycles.*

Proof. Consider the holomorphic family of maps $f_t(z) = (1-t)f(z) + tz^d$, $t \in \mathbb{C}$. By transversality, any periodic point of multiplier $\lambda \neq 1$ can be locally labeled (near $t = 0$) by a holomorphic function p_t, with $f_t^n(p_t) = p_t$. In the case of a multiple periodic point (multiplier 1), the same is true after making a base change by $t \mapsto t^n$.

Now consider the multiplier $(f^n)'(p_t) = \lambda_t$ as t varies, and assume p_0 is indifferent. Since $z \mapsto z^d$ has no indifferent cycle, λ_t is nonconstant. The set of directions $t \in S^1$ such that $|\lambda_{\epsilon t}| < 1$ (and thus p_t becomes an attractor for $f_{\epsilon t}$) has measure $1/2$.

Applying the same reasoning the any collection of N indifferent cycles, we conclude there is some direction t such that $N/2$ of them become attracting for $f_{\epsilon t}$. Under this perturbation, the attracting cycles for f_0 remain attracting. Since $f_{\epsilon t}$ has at most $2d - 2$ attracting cycles, we find:

$$(\text{attracting}) + (\text{indifferent})/2 \leq 2d - 2,$$

so at most $4d - 4$ cycles of f_0 are not repelling. ∎

Remarks. It is easy to construct maps of a fixed degree where the *period* of one or more attracting cycles is very long. For example, there are $c_n \to -2$ such that the critical point of $f_n(z) = z^2 + c_n$ has period n.

5.2 Basic properties of the Julia set.

Theorem 5.7 *Let $J(f)$ be the Julia set of a rational map. Then:*

1. *Repelling periodic points are dense in $J(f)$.*

2. *The inverse orbit of every point in $J(f)$ is dense in $J(f)$.*

3. *More generally, the inverse orbit of $z \in \widehat{\mathbb{C}}$ accumulates on $J(f)$ unless z is exceptional.*

4. *If an open set U meets $J(f)$, then $J(f) \subset f^n(U)$ for some n.*

5. *The Julia set is perfect: it has no isolated points.*

Proof. For (1), consider $p \in J(f)$. By replacing p by a point in its inverse orbit, we can assume p is neither a fixed-point nor a critical value of f. On any small neighborhood U of p we can find maps $g_i : U \to \widehat{\mathbb{C}}$ with disjoint graphs, $i = 1, 2, 3$, such that $g_1(z) = z$, and $f(g_i(z)) = z$ for $i = 2, 3$.

If $f^n(z) = g_i(z)$ for some $n > 0$ and $z \in U$, then either $f^n(z) = z$ or $f^{n+1}(z) = z$, and thus p is approximated by a periodic point. Otherwise the sequence of functions

$$h_n(z) = \frac{f^n(z) - g_0(z)}{f^n(z) - g_1(z)} \left(\frac{g_3(z) - g_0(z)}{g_3(z) - g_2(z)} \right)^{-1} = M_z \circ f^n$$

omits $\{0, 1, \infty\}$, and is hence normal on U by Montel's theorem. But then $\langle f^n | U \rangle$ is normal, contradicting the definition of $J(f)$.

This shows p is a limit of periodic points of f; and all but $4d - 4$ of these are repelling. Once p is so approximated, so is any point in the forward orbit of p, and thus $J(f)$ is the closure of the repelling periodic points of f.

(2): For $p \in J(f)$ consider any 3 points z_1, z_2, z_3 in the inverse orbit of p. Let U be an open set meeting $J(f)$. Since $\langle f^n | U \rangle$ is not a normal family, it cannot omit all 3 values $\{z_1, z_2, z_3\}$, and thus $f^n(z) = z_i$ for some $z \in U$. Thus the inverse orbit of p enters U.

The same argument applies to any point p with an infinite inverse orbit, yielding (3).

(4): By (1), U contains a repelling periodic point, so after shrinking U and replacing f with f^n we can assume $U \subset f(U)$. Then $f^n(U)$ is an increasing sequence of open sets, and by (3) any non-exceptional point is eventually covered by $f^n(U)$. Since the exceptional points do not belong to $J(f)$, by compactness we have $J(f) \subset f^n(U)$ for some finite n.

(5): If $J(f)$ has an isolated point, then by (4) $J(f)$ is a finite set. Since $f^{-1}(J(f)) = J(f)$, the Julia set must consist of exceptional points, which implies $J(f) = \emptyset$. But $\deg(f^n) \to \infty$, so the iterates of f cannot form a normal family on the whole sphere. ∎

Corollary 5.8 *The Julia set is the smallest totally invariant closed set on the sphere such that* $|J(f)| \geq 3$.

Remarks.

1. To apply the Schwarz lemma to a rational map, we want to find a hyperbolic open set $\Omega \subset \widehat{\mathbb{C}}$ such that $f(\Omega) \subset \Omega$. Then $F = \widehat{\mathbb{C}} - \Omega$ is a closed, backward invariant set with $|F| \geq 3$. By Corollary 5.8, the Fatou set $\Omega(f)$ is the largest hyperbolic open set which is mapped into itself.

2. The argument to prove (1) in Theorem 5.7 illustrates a useful generalization of Montel's theorem: for any bundle of hyperbolic Riemann surfaces $E \to U$, the space of all holomorphic sections $s : U \to E$ is normal. In the case at hand, the bundle has fibers $E_z = \widehat{\mathbb{C}} - \{g_1(z), g_2(z), g_3(z)\}$.

3. The density of periodic cycles in the 'chaotic locus' is not known for a generic C^2 diffeomorphism of a manifold. Thus the density of periodic cycles in $J(f)$ is one of many indications that one-dimensional complex dynamical systems are better behaved than higher-dimensional smooth dynamical systems. (For C^1 diffeomorphisms, the generic density of periodic cycles in the non-wandering set is known by Pugh's 'closing lemma' [Pu]).

4. Property (4) is sometimes called LEO (locally eventually onto).

5.3 Univalent maps

We now introduce another very useful tool. Let S be the set of all univalent analytic maps

$$f : \Delta \to \mathbb{C},$$

normalized so that $f(0) = 0$ and $f'(0) = 1$. We give S the topology of uniform convergence on compact sets.

Theorem 5.9 (Koebe distortion theorem) *The space of normalized univalent maps S is compact.*

Proof. By the Schwarz Lemma, $f(\Delta)$ cannot contain $B(0, r)$ for any $r > 1$. So given a sequence $f_n \in S$, we can find a sequence of points $a_n, b_n \notin f_n(\Delta)$ with $|a_n| = 1$ and $|b_n| = 2$. Letting $A_n(z) = (z - a_n)/(b_n - a_n)$, we see

$$A_n \circ f_n : \Delta \to \mathbb{C} - \{0, 1\}.$$

By Montel's theorem, we can pass to a subsequence converging on the disk. Along a further subsequence we have $a_n \to a$ and $b_n \to b$ with $a \neq b$, so $A_n \to A$. Then $A_n \circ f_n \to A \circ f$ for some $f \in S$, and thus $f_n \to f$. ■

Corollary 5.10 *Writing $f(z) = \sum_1^\infty a_n z^n$, we have $|a_n| \leq C_n$ for universal constants C_n.*

In fact we can take $C_n = n$; this is the *Bieberbach conjecture*, proved by deBranges.

Corollary 5.11 *There is an $r > 0$ such that $f(\Delta) \supset B(0, r)$ for any $f \in S$.*

In fact we can take $r = 1/4$; this is the *Koebe 1/4 theorem.*

Corollary 5.12 *If $f : B(z, r) \to \mathbb{C}$ is univalent on a ball, then f has bounded distortion on the smaller ball $B(z, r/2)$. That is, all ratios of distance are distorted by a bounded factor; f is a quasi-similarity.*

Keep in mind that the Riemann mapping theorem allows a conformal map to send Δ to an arbitrarily wild region. The point of the Corollary is that a small baller has a controlled image. *The yolk of the egg stays good.*

5.4 Periodic points

Before analyzing the global dynamics of a rational map, we discuss the local dynamics of a holomorphic map near one of its fixed-points. Let f be an analytic mapping fixing $z = 0$, with multiplier $\lambda = f'(0)$. We distinguish several cases, depending on the multiplier.

Attracting case: $\mathbf{0 < |\boldsymbol{\lambda}| < 1}$.

Theorem 5.13 (Linearization) *There is a holomorphic chart $\phi : (U, 0) \to (\mathbb{C}, 0)$ such $f(w) = \lambda w$ where $w = \phi(z)$.*

Proof 1. Let $B(0, s)$ be a small enough ball that $|f(z)| \le r|z| < |z|$ for all $z \in B(0, s) - \{0\}$. Then $|z_n| = |f^n(z)| = O(r^n)$. Defining $\phi_n(z) = \lambda^{-n} f^n(z)$, we have

$$\frac{\phi_{n+1}(z)}{\phi_n(z)} = \frac{\lambda^{-1} f(z_n)}{z_n} = \frac{z_n + O(z_n^2)}{z_n} = 1 + O(r^n).$$

Since $\sum r^n < \infty$, $\phi_n(z)$ converges uniformly to a nonzero holomorphic map $\phi : (U, 0) \to (\mathbb{C}, 0)$. Since $\phi'(0) = \lim \phi_n'(0) = 1$, the map ϕ is a chart on a small enough neighborhood of $z = 0$.

Proof 2. Consider the annulus $A = B(0, s) - f(B(0, s))$ for s small. Then $X = A/f$ is a Riemann surface, indeed a complex torus. The subgroup of $\pi_1(X)$ corresponding to f determines a cyclic covering space of X isomorphic to $\mathbb{C}^*$. Uniformization of this covering space gives the linearizing map. ∎

Superattracting case: $\boldsymbol{\lambda} = \mathbf{0}$.

Theorem 5.14 *Suppose $f(z) = z^d + O(z^{d+1})$, $d > 1$. Then there is a holomorphic chart $\phi : (U, 0) \to (\mathbb{C}, 0)$ such $f(w) = w^d$, where $w = \phi(z)$.*

Proof. For z small enough we have

$$|z_n| = |f^n(z)| \le \frac{1}{2^{d^n}}.$$

Let $\phi_n(z) = (f^n(z))^{1/d^n}$, where the d^n-th root is chosen so that $\phi_n'(0) = 1$. Then

$$\frac{\phi_{n+1}(z)}{\phi_n(z)} = \left(\frac{f(z_n)^{1/d}}{z_n}\right)^{1/d^n} = (1 + O(z_n))^{1/d^n} = 1 + O(z^n) = 1 + O(2^{-d^n}).$$

Since $\sum 2^{-d^n}$ converges (very rapidly!), ϕ_n converges uniformly to the desired chart ϕ. ∎

Parabolic case: $\boldsymbol{\lambda = 1}$. Assuming f is not the identity, we can make a linear change of coordinates so that

$$f(z) = z + z^{p+1} + O(z^{p+2}),$$

$p \geq 1$. The behavior of f is controlled by the leading term in $f(z) - z$.

Indeed, the dynamics of f is nicely modeled by the holomorphic vector field

$$V = z^{p+1}\frac{\partial}{\partial z}.$$

For $p = 1$, V is an infinitesimal parabolic Möbius transformation, and its orbits are circles tangent to the real axis at $z = 0$. Every orbit converges to $z = 0$, although those near the real axis make a long excursion away from $z = 0$ before returning to converge.

For $p > 1$, the picture is just a p-fold cover of the parabolic picture.

Theorem 5.15 (Leau-Fatou flower theorem) *For $p = 1$, there is a petal U tangent to the positive real axis such that $f(U) \subset U$ and $f^n(z) \to 0$ uniformly on compact subsets of U. The quotient Riemann surface U/f is isomorphic to $\mathbb{C}^*$.*

For $p > 1$ there are p such petals, $U_1, \ldots, U_p$, each invariant under f and tangent to the lines $\arg(z) = 2\pi k/p$, $k = 1, 2, \ldots, p$. We have $U_i/f \cong \mathbb{C}^*$.

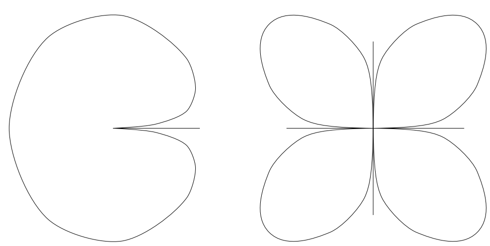

Figure 14. One and several petals

Proof. First suppose $p = 1$. The discussion is mostly easily carried out by moving the fixed-point to infinity with the coordinate change $w = -1/z$. Then

$$f(w) = w + 1 + O(1/w),$$

from which we see $f^n(w) = w + n + O(\log n)$ if $\operatorname{Re} w \gg 0$. It follows that the basin of attraction of $w = \infty$ contains the half-plane $H = \{w : \operatorname{Re} w > R\}$ for some large R. By considering the action of f on the edges of the strip $H - f(H)$, we see that $H/f \cong \mathbb{C}^*$.

Now if $\operatorname{Re} w$ is less than R, we can still insure that the orbit of w enters H so long as $|\operatorname{Im} w|$ is large enough and a gradual drift towards $w = 0$ coming from the $O(1/w)$ terms is controlled. If $w = x + iy$, it will take approximately $n = R - x$

iterations to reach H, and the potential vertical drift is $O(\log n) = O(\log(2+|x|)$, so we set

$$U = \{w = x + iy \,:\, x > R \text{ or } |y| > C \log(2 + |x|)\}.$$

Then $f(U) \subset U$ and $f^n(w) \to \infty$ in U as desired. Since the slope of the boundary of U tends to zero as $x \to -\infty$, in the z-coordinate U is tangent to the positive real axis.

The case of $p > 1$ is similar. It is useful to make the coordinate change $w = -1/z^p$; then f is a multivalued map, spread out over the w-plane, with the form

$$f(w) = w + 1 + O(w^{-1/p}).$$

■

Parabolic case: $\boldsymbol{\lambda^q = 1}$. If the multiplier is a primitive qth root of unity, then one can conjugate f to the form $f(z) = \lambda z + O(z^{q+1})$ by a holomorphic change of coordinates. Thus $f^q(z) = z + O(z^{q+1})$ has *at least* q petals. Generically there are exactly q, but in general there may be $p = kq$ altogether. These petals fall into k orbits under the action of f.

Corollary 5.16 *Every cycle of parabolic petals contains a critical point of f.*

Proof. Otherwise $f : U \to U$ is a covering map. But from the local picture we see $U/f \cong \mathbb{C}^*$. Thus $U \cong \mathbb{C}^*$ or $\mathbb{C}$, and neither of these spaces can embed in the complement of the Julia set. ■

The snail lemma. For later use we record the following criterion for a fixed-point to be parabolic.

Theorem 5.17 *Let f have an indifferent fixed-point at $z = 0$, and suppose there is a domain $U \subset \widehat{\mathbb{C}}$ with $f(U) \subset U$ and $f^n(z) \to 0$ uniformly on compact subsets of U. Then $f'(0) = 1$.*

Intuitively, if $f'(0) \neq 1$, then U must wind around 0, like a snail shell; then a short cross-cut for U forms a disk mapping properly into itself, showing $|f'(0)| < 1$.

Proof. We can assume f is univalent on Δ. Let $B \subset U$ be a ball and let $K \subset U$ be a compact *connected* set containing B and $f(B)$. Since $f^n \to 0$ uniformly on K, we can assume $f^n(K) \subset \Delta$ for all n, and thus $f^n|K$ is univalent for all n.

Suppose the multiplier $\lambda = f'(0) \neq 1$. Then $d(f^n(K), 0) = O(\operatorname{diam} f^n(K))$, for otherwise $f^n(K)$ would not be big enough to join $f^n(B)$ to $f^{n+1}(B) \approx \lambda f^n(B)$. By the Koebe distortion theorem, we also have $d(f^n(B), 0) = O(\operatorname{diam} f^n(B))$.

Thus the visual size of $f^n(B)$ as seen from $z = 0$ is bounded below. But then there is a universal N such that

$$\langle \lambda^i(f^n(B)) \,:\, i = 1, \ldots, N\rangle$$

forms a necklace encircling $z = 0$. Once $f^n(B)$ is close enough to zero, $f(z) \approx \lambda z$, and therefore

$$\langle f^{n+i}(B) \, : \, i = 1, \dots N \rangle$$

also encircles $z = 0$. The union of these balls is in U, so there is a loop $L \subset U$ encircling zero. Since $f^n|L \to 0$ uniformly, the maximum principle implies that $f^n \to 0$ uniformly on the disk enclosed by L. Therefore $|f'(0)| < 1$, contrary to our assumption of indifference.

Therefore $f'(0) = 1$. ∎

Irrationally indifferent case: $\boldsymbol{\lambda = \exp(2\pi i \theta)}$, $\boldsymbol{\theta \in \mathbb{R} - \mathbb{Q}}$.

First we remark that f may not be linearizable. This is not surprising since a parabolic point is not linearizable (unless f has finite order), so a nearly parabolic point should not be either. More concretely we have:

Theorem 5.18 *There is a dense G_δ of multipliers $\lambda \in S^1$ such that $f(z) = \lambda z + z^2$ cannot be linearized near $z = 0$.*

Proof. Note that

$$f^n(z) - z = (\lambda^n - 1)z + a_2 z^2 + \cdots + z^{2^n}.$$

Thus the product of the nonzero roots of $f^n(z) = z$ is $\lambda^n - 1$, and therefore f has a periodic point p_n with

$$|p_n| \le |\lambda^n - 1|^{1/2^n} = r_n.$$

The set of λ such that $\liminf r_n = 0$ is a dense G_δ on S^1 (containing the roots of unity), and for any such λ there are periodic points $p_{n_k} \to 0$, so f is not linearizable. ∎

Theorem 5.19 *There is a full-measure set of multipliers $\lambda \in S^1$ such that $f(z) = \lambda z + z^2$* can *be linearized near $z = 0$.*

Proof. Let $f_\lambda = \lambda z + z^2$, let $U_\lambda = \{z \, : \, f_\lambda^n(z) \to 0\}$. For $\lambda \in \Delta^*$, let $\phi_\lambda : U_\lambda \to \mathbb{C}$ be the linearizing map, normalized so that $\phi_\lambda'(0) = 1$. Finally, noting that the critical value of f_λ is at $-\lambda^2/2$, set

$$R_\lambda = \phi_\lambda(-\lambda^2/2).$$

Next note that ϕ_λ^{-1} admits a univalent branch defined on $B(0, |R_\lambda|)$. By the Koebe 1/4 theorem, this means $U_\lambda \supset B(0, |R_\lambda|/4)$. But U_λ is bounded, so R_λ is a bounded analytic function on Δ^* (and indeed on Δ, since the singularity at $z = 0$ is removable).

By a general result, from complex analysis, R_λ has nonzero radial limits at almost every $\lambda \in S^1$. Thus for almost every $\lambda \in S^1$, there is a ball of positive radius $B(0, |R_\lambda|/4)$ contained in the Fatou set of f_λ.

Thus the component U of $\Omega(f_\lambda)$ containing $z = 0$ is a nonempty disk, and $f : U \to U$ satisfies $|f'(0)| = 1$. By the Schwarz lemma, f is an isometry in the hyperbolic metric, the Riemann mapping $(U, 0) \to (\Delta, 0)$ conjugates f to a rotation. ∎

The above argument is from [Y, §II.2].

Conditions of Diophantus and Brjuno. Here is a more precise statement about linearizing irrational multipliers. An irrational number θ is *Diophantine* if there exist $C, d > 0$ such that

$$\left|\theta - \frac{p}{q}\right| > \frac{C}{q^d}$$

for all rationals p/q. Almost every number is Diophantine, and Siegel showed any holomorphic germ of the form $f(z) = e^{2\pi i\theta}z + O(z^2)$ with θ Diophantine is linearizable.

It is known that the set of θ for which $f(z) = e^{2\pi i\theta}z + z^2$ is linearizable is *exactly* those satisfying the Brjuno condition:

$$\sum \frac{\log q_{n+1}}{q_n} < \infty,$$

where p_n/q_n are the continued fraction approximants of θ. See [Y] for details.

5.5 Classification of periodic regions

Theorem 5.20 (Classification of stable regions) *A component Ω_0 of period p in the Fatou set of a rational map f is of exactly one of the following five types:*

1. *An* attractive basin*: there is a point x_0 in Ω_0, fixed by f^p, with $0 < |(f^p)'(x_0)| < 1$, attracting all points of Ω_0 under iteration of f^p.*
2. *A* superattractive basin*: as above, but x_0 is a critical point of f^p, so $(f^p)'(x_0) = 0$.*
3. *A* parabolic basin*: there is a point x_0 in $\partial\Omega_0$ with $(f^p)'(x_0) = 1$, attracting all points of Ω_0.*
4. *A* Siegel disk*: Ω_0 is conformally isomorphic to the unit disk, and f^p acts by an irrational rotation.*
5. *A* Herman ring*: Ω_0 is isomorphic to an annulus, and f^p acts again by an irrational rotation.*

Proof. Replacing f with f^p, we can assume f maps Ω_0 to itself. Then f is non-expanding for the hyperbolic metric on Ω_0 (which we denote by $d(\cdot)$); that is, $d(fx, fy) \le d(x, y)$ for all $x, y \in \Omega_0$.

Consider an arbitrary $z \in \Omega_0$ and its forward orbit $z_n = f^n(z)$. Then $d(z_{n+1}, z_n) \le d(z_1, z_0)$.

If $z_n \to \infty$ (meaning the orbit leaves every compact set of Ω_0), then $d_{\widehat{\mathbb{C}}}(z_n, z_{n+1}) \to 0$, since the ratio of the spherical to hyperbolic metrics on Ω_0 tends to zero at the boundary. Thus the orbit accumulates on a connected set $E \subset \partial\Omega_0$ on which

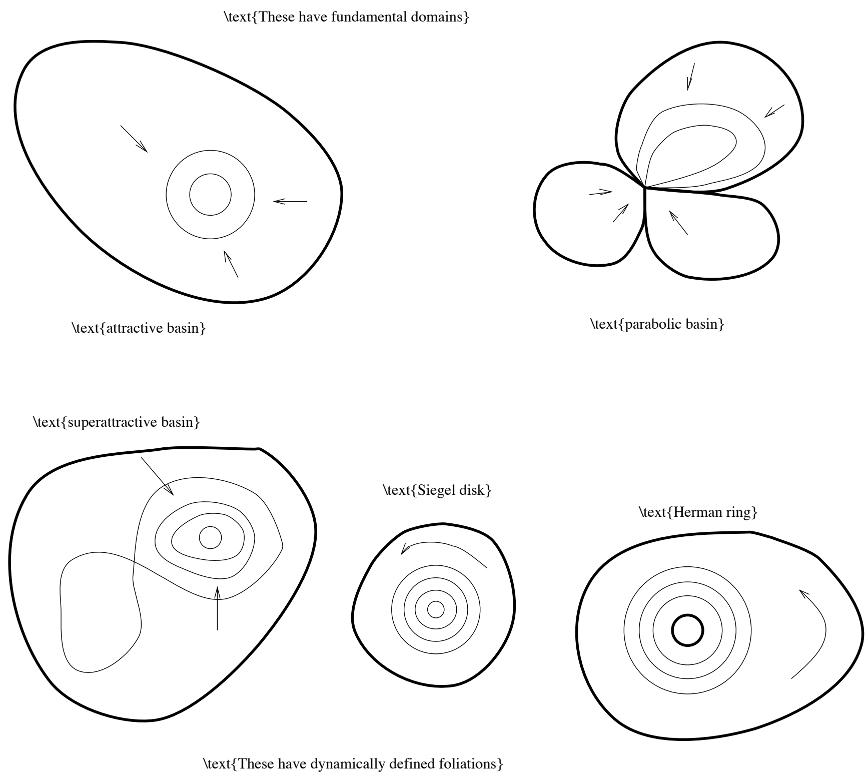

Figure 15. The five types of stable regions.

$f(z) = z$. Since f is not the identity map, E reduces to a single fixed-point $E = \{x_0\}$. Since $x_0 \in J(f)$, it is an indifferent fixed-point, and $f'(x_0) = 1$ by the Snail Lemma (Theorem 5.17).

Now suppose z_n is recurrent: that is, the orbit returns infinitely often to a fixed compact set $K \subset \Omega_0$. We distinguish two cases, depending on whether or not $f : \Omega_0 \to \Omega_0$ is a covering map.

If f is *not* a covering map, then $|f'| < c < 1$ on K (in the hyperbolic metric); therefore $d(z_n, z_{n+1}) \to 0$, and any accumulation point x_0 of z_n in K is a fixed-point of f. By contraction, this fixed-point is unique, $|f'(x_0)| < 1$, and we have the attracting or superattracting case.

If f *is* a covering map, then by recurrence there are self-coverings (of the form f^n) arbitrarily close to the identity (on compact sets). More precisely, if v_0 is a unit tangent vector at z_0, $v_n = Df^n(v_0)$, and $v_k \approx v_{k+n}$, then $f^n \approx \mathrm{id}$ on a large ball about z_k.

If Ω_0 carries a nontrivial closed hyperbolic geodesic γ, then $f^n(\gamma) = \gamma$ for arbitrarily large n. Since $f^n \neq \mathrm{id}$, we conclude that Ω_0 is an annulus around γ and f is an irrational rotation. This is the case of a Herman ring.

Finally if Ω_0 carries no closed geodesics, then it must be a disk (it cannot be a punctured disk since the Julia set is perfect). Then $f|\Omega_0$ is an irrational rotation; any other isometry of infinite order would fail to be recurrent. This is the case of a Siegel disk. ■

Cf. [McS].

5.6 The postcritical set

Although the Schwarz lemma gives contraction in the hyperbolic metric, it can also be used to reveal expanding properties of rational maps.

Definition. The *post-critical* set is given by

$$P(f) = \overline{\bigcup_{n=1}^{\infty} \bigcup_{f'(c)=0} f^n(c)}.$$

Thus $P(f)$ is the smallest forward-invariant closed set containing the critical values of f.

Once easily checks $P(f) = P(f^n)$ for any $n \geq 1$, and $|P(f)| \leq 2$ iff $f(z)$ is conjugate to $z \mapsto z^d$.

Setting the (easily analyzed) cases where $f(z) = z^d$, we can assume $|P(f)| \geq 3$ and thus its complement $\widehat{\mathbb{C}} - P(f)$ is hyperbolic (in the sense that each component is hyperbolic).

Since $f(P(f)) \subset P(f)$, the pre-image of the post-critical set contains itself. Thus we obtain the diagram:

$$\begin{array}{ccc}
\widehat{\mathbb{C}} - f^{-1}P(f) & \xrightarrow{\iota} & \widehat{\mathbb{C}} - P(f) \\
{\scriptstyle f}\big\downarrow & & \\
\widehat{\mathbb{C}} - P(f) & &
\end{array}$$

where along the vertical arrow f is a covering map, and along the horizontal arrow ι is an inclusion. Imposing the hyperbolic metric on $\widehat{\mathbb{C}}-P(f)$, we conclude that the vertical arrow is an *isometry*, while the inclusion ι is *non-expanding*. Thus the composition

$$f \circ \iota^{-1} : \widehat{\mathbb{C}} - P(f) \to \widehat{\mathbb{C}} - P(f)$$

is a *non-expanding* where it is defined.

Theorem 5.21 *For any $z \in J(f)$, we have $\|(f^n)'(z)\| \to \infty$ in the hyperbolic metric on $\widehat{\mathbb{C}} - P(f)$.*

(If $f^n(z)$ happens to land in $P(f)$, we set $\|(f^n)'(z)\| = \infty$.)

Proof. Let $X_n = \widehat{\mathbb{C}} - f^{-n}(P(f))$ and

$$\iota_n : X_n \hookrightarrow X_0$$

the inclusion. Since $|P(f)| \geq 3$, $f^{-n}(P(f))$ accumulates on the Julia set, and thus the hyperbolic metric on X_n at $z \in J(f)$ tends to infinity. In other words, $\|\iota_n'(z)\| \to 0$ with respect to the hyperbolic metrics on domain and range. Since $f^n : X_n \to X_0$ is a covering map, we have $\|(f^n)'(z)\| \to \infty$ with respect to the hyperbolic metric on X_0. ∎

Every periodic component of the Fatou set is related to the post-critical set. More precisely:

Theorem 5.22 *The post-critical set $P(f)$ contains every attracting, superattracting and parabolic cycle, every indifferent point in the Julia set, and the full boundary of every Herman ring and Siegel disk.*

Proof. Assume $|P(f)| > 2$, since the theorem clearly holds for $f(z) = z^d$.

Consider any ball $B = B(w, r)$ disjoint from $P(f)$. Then all branches of $f^{-n}|B$ are well-defined and univalent on B. Moreover, these inverse branches form a normal family, since they omit $P(f)$ from their range.

If w is a periodic point, of period p, then by normality its multiplier satisfies $|\lambda| \geq 1$. If $|\lambda| = 1$, then by univalence the forward iterates of f are also normal near w, so w is the center of a Siegel disk. Summing up, any periodic point outside $P(f)$ must be either repelling or the center of a Siegel disk.

Similarly, if w is in the boundary of a Siegel disk or Herman ring, we can choose a subsequence of $f^{-n}|B$ that converges to the identity on an open subset of B (namely its intersection with the rotation domain). Thus along a subsequence, f^{-n} converges to the identity on all of B. Thus B contains no repelling cycles, so B is disjoint from the Julia set, a contradiction. ∎

5.7 Expanding rational maps

Definition. A rational map $f(z)$ is *expanding* if for some n, the derivative of f^n in the spherical metric satisfies $|(f^n)'(z)|_\sigma > 1$ for all $z \in J(f)$.

Theorem 5.23 *The following are equivalent:*

1. *The rational map f is expanding.*
2. $P(f) \cap J(f) = \emptyset$.
3. *The forward orbit of every critical point tends to an attracting cycle.*
4. *There are no critical points in the Julia set, and every point in the Fatou set converges to an attracting cycle.*

Here 'attracting' includes 'superattracting'.

Proof. The maps $f(z) = z^d$ clearly satisfy all the above conditions, so we may assume $|P(f)| \geq 3$.

(1) $\implies$ (4). By expansion, there is an $r > 0$ such that for any $z \in \Omega(f)$, $\sup_n d(f^n(z), J(f)) > r$. Since only finitely many components of $\Omega(f)$ can contain a ball of radius r, we see the orbit of z eventually lands in a periodic component. Expansion rules out Siegel disks, Herman rings and parabolic basins, so $f^n(z)$ must tend to an attracting cycle.

(4) $\implies$ (3) $\implies$ (2). These implications are immediate.

(2) $\implies$ (1). By Theorem 5.21, under iteration f expands the hyperbolic metric on $\widehat{\mathbb{C}} - P(f)$ at any point of $J(f)$. If $P(f) \cap J(f)$ is empty, then by compactness of $J(f)$, some fixed iterate f^n strictly expands the hyperbolic metric everywhere on $J(f)$. A further iterate f^{nk} then expands the spherical metric, since the spherical and hyperbolic metrics have a bounded ratio of $J(f)$. ■

Theorem 5.24 *The Julia set of an expanding map is quasi-self-similar, and* $\mathrm{H.dim}\, J(f) < 2$.

Proof. By expansion and the distortion lemma for univalent maps, any small piece of $J(f)$ can be blown up to definite size with bounded distortion. Thus $J(f)$ is quasi-self-similar, and there is no ball in which $J(f)$ is very dense. It follows that $\mathrm{H.dim}\, J(f) < 2$. ■

The preceding two results, taken together, completely determine the behavior of a typical point $z \in \widehat{\mathbb{C}}$ under the iteration of an expanding map f. For example we can now see:

Corollary 5.25 *Suppose $f(z) = z^d + c$ has an attracting cycle $C \subset \mathbb{C}$. Then for any $z \in \widehat{\mathbb{C}}$, either:*

1. *$f^n(z) \to \infty$; or*

2. $d(f^n(z), C) \to 0$; *or*

3. $z \in J(f)$, *a compact set of Hausdorff dimension* < 2.

Proof. Since f has only two critical points, it has exactly two attracting cycles, C and ∞. These cycles must each attract a critical point, so f is expanding. Therefore every point in $\Omega(f)$ tends to C or ∞, and the remaining points $J(f)$ have dimension less than two. ∎

5.8 Density of expanding dynamics

We can now state a central conjecture:

Conjecture 5.26 *For each $d \geq 2$, the expanding maps are open and dense in* Rat_d *and in* Poly_d, *the spaces of rational maps and polynomials of degree d.*

Clearly the expanding maps form an open set, but their density is still unknown, even in the case of Poly_2 (quadratic polynomials).

Recall from §1.3 that in the early history of smooth dynamical systems, it was suspected that Axiom A systems (an analogue of expanding) should be open and dense in $\mathrm{Diff}(M^n)$. This was soon proved to be false for all manifolds of dimension $n \geq 3$, and finally even for all surfaces.

For $n = 1$ Axiom A dynamics is trivially seen to be dense. There are good reasons (both theoretical and experimental) to think that expanding maps may be dense for conformal dynamical systems. The most naive reason is that polynomials and rational maps are 1-dimensional dynamical systems over $\mathbb{C}$.

5.9 Quasiconformal maps and vector fields

We now introduce a new technique.

Let $f : U \to V$ be a diffeomorphism between regions in $\mathbb{C}$. Recall the differential operators

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right), \qquad \frac{\partial f}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right).$$

These operators behave like differentiation with respect to the independent variables $(z, \overline{z})$. For example,

$$f(z+t) = f(z) + t\frac{\partial f}{\partial z} + \overline{t}\frac{\partial f}{\partial \overline{z}} + O(t^2).$$

The map f is *conformal* iff $\partial f / \partial \overline{z} = 0$.

We say f is *quasiconformal* if for some $0 < k < 1$, we have

$$\left|\frac{\partial f}{\partial \overline{z}}\right| \leq k \left|\frac{\partial f}{\partial z}\right|.$$

Noting that for $|t| = 1$,

$$|f_z| - |f_{\overline{z}}| \le |t f_z + \overline{t} f_{\overline{z}}| \le |f_z| + |f_{\overline{z}}|,$$

we see Df maps circles to ellipses of oblateness bounded by

$$K = \frac{1+k}{1-k} \ge 1;$$

we say f is *K-quasiconformal* and note that 1-quasiconformal mappings are conformal.

The Jacobian determinant is given by the product of the maximum and minimum stretchings:

$$\det Df = (|f_z| + |f_{\overline{z}}|)(|f_z| - |f_{\overline{z}}|) = |f_z|^2 - |f_{\overline{z}}|^2.$$

We also have $|f_x|^2$, $|f_y|^2 \le K \det Df$. Thus for any region U,

$$\operatorname{area}(f(U)) = \int_U \det Df \, |dz|^2 \ge \frac{1}{K} \int_U \left| \frac{\partial f}{\partial x} \right|^2 |dz|^2,$$

and similarly for $\partial f / \partial y$.

This suggest the more general definition: a homeomorphism $f : U \to V$ is *K-quasiconformal* iff f has distributional first derivatives f_x and f_y (or f_z and $f_{\overline{z}}$) in L^2, and $Df(z)$ is K-quasiconformal for almost every z.

The distortion of a general quasiconformal map is measured by the *complex dilatation*

$$\mu = \mu(z) \frac{d\overline{z}}{dz} = \frac{\overline{\partial} f}{\partial f}.$$

This *Beltrami differential* μ is a measurable $(-1, 1)$-form, so its absolute value is natural, and we have

$$\|\mu\|_\infty \le k = \frac{K-1}{K+1} < 1.$$

The space of complex structures. Here is a more intrinsic discussion of the Beltrami differential. For any Riemann surface X, we let $M(X)$ be the space of L^∞ Beltrami differentials, equipped with the sup-norm. We claim the unit ball in $M(X)$ is naturally identified with the space of complex structures on X at a bounded distance form the given structure.

To see this, recall that the space of conformal structures on a 2-dimensional real vector space V is $\mathrm{SL}(V)/\mathrm{SO}(V) \cong \mathbb{H}$. Thus the space of complex structures on a surface S is naturally a bundle of hyperbolic disks $CS = \mathrm{SL}(TS)/\mathrm{SO}(TS)$.

On a Riemann surface X, this bundle comes equipped with a section s, so we have a bundle of pointed hyperbolic disks. The associated bundle of tangent spaces, $T_s CX$, is naturally isomorphic to the Beltrami line bundle $T^{-1,1}X$. Then CX is naturally realized as the space of unit disks *in* the Beltrami line bundle, and a Beltrami differential with $|\mu| < 1$ is simply a measurable section of this bundle of unit disks, at a bounded hyperbolic distance from the zero section.

The Beltrami differential as a connection. A complex structure is the same as a $\overline{\partial}$-operator on functions, and we can also associate to μ the operator

$$\overline{\partial}_\mu = \overline{\partial} - \mu\partial$$

to specify the Riemann surface (X, μ).

Complex structures and self-adjoint maps. Alternatively, given a *pair* of a complex structures J_1, J_2 on V, there is a unique map $A : V \to V$ such that

(i) A is positive with respect to J_1 (i.e. A has a set of positive eigenvectors, orthogonal with respect to J_1);

(ii) $A^*(J_2) = J_1$ (i.e. $J_1 = A^{-1}J_2A$); and

(iii) $\det(A) = 1$.

Thus we can identify $C(V)$ with the space $P(V, J_1) \subset \mathrm{SL}(V)$ of positive endomorphisms of determinant one.

Now consider the special case where $V = \mathbb{C}$ with its standard conformal structure. Then a $\mathbb{R}$-linear map $A : \mathbb{C} \to \mathbb{C}$ given by

$$A(z) = az + b\overline{z}$$

is self-adjoint iff $a \in \mathbb{R}$. When self-adjoint, its eigenvalues are $a \pm |b|$. Thus A is positive iff we have $a \in \mathbb{R}$ and $a > |b|$.

Now return to the setting of a Riemann surface X. A Beltrami differential $\mu \in T_p^{-1,1}X$ with $|\mu| < 1$ determines a positive map $A : T_pX \to T_pX$ by

$$A(v) = \frac{1 + \mu(v)}{1 - |\mu|^2} \cdot v.$$

This map $A \in P(T_pX, J_p)$ corresponds canonically to the complex structure determined by μ. For example, if $\mu = a\, d\overline{z}/dz$ on $\mathbb{C}$, then under the identification $T_p\mathbb{C} = \mathbb{C}$ we have $\mu(\xi) = a\overline{\xi}/\xi$ and

$$A(\xi) = \frac{\xi + a\overline{\xi}}{1 - |a|^2}.$$

When a is real, the eigendirection of A are along the real and imaginary axes. In general, they correspond to the lines along which $\mu(v) > 0$ and $\mu(v) < 0$.

Note: if $|\mu(p)| = 1$, the unnormalized map $A(v) = (1 + \mu(v)) \cdot v$ has rank one, so its kernel determines a real line $L_p \subset T_pX$.

Theorem 5.27 (Measurable Riemann mapping theorem) *For any $\mu \in M(\widehat{\mathbb{C}})$ with $\|\mu\|_\infty < 1$, there is a unique quasiconformal homeomorphism $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $f_{\overline{z}} = \mu f_z$ and f fixes $\{0, 1, \infty\}$.*

Moreover, the mapping $f_t(z)$ with dilatation $t\mu$, $t \in \Delta$, varies holomorphically with respect to t.

The first statement is due to Morrey; the second is due to Ahlfors and Bers. Bers also introduced the use of the Beltrami equation in the theory of deformations of Kleinian groups.

Examples.

1. For $|\alpha| < 1$ let
$$f(z) \;=\; \begin{cases} z + \alpha\overline{z}, & |z| \le 1, \\ z + \alpha/z & \text{otherwise.} \end{cases}$$
Then $f_{\overline{z}}/f_z = \alpha$ on the unit disk and zero elsewhere. The mapping f is K-quasiconformal with $K = (1 + |\alpha|)/(1 - |\alpha|)$.

2. For $1 > \alpha > 0$ let $f(z) = z|z|^{\alpha-1}$. In polar coordinates we have $f(re^{i\theta}) = r^\alpha e^{i\theta}$.
Writing $f(z) = z^{(\alpha+1)/2}\overline{z}^{(\alpha-1)/2}$, we can easily compute
$$\mu = \frac{f_{\overline{z}}}{f_z} = \frac{(\alpha-1)}{(\alpha+1)}\frac{z}{\overline{z}}.$$
Thus f is K-quasiconformal with $K = 1/\alpha$.
Note that f is only α-Hölder continuous at the origin. It is a general fact that a K-quasiconformal map is $1/K$-Hölder continuous.

Distortion of balls. An alternative definition of K-quasiconformal mappings f is that
$$H(z) = \limsup_{r\to 0} \frac{\max_{S^1(z,r)} |f(z) - f(w)|}{\min_{S^1(z,r)} |f(z) - f(w)|} \le K$$
for almost every z. Thus roughly speaking, the ratio of inradius to outradius of the image of a ball is bounded by K.

This sharp bound on distortion of balls holds *only* a.e., and only at a microscopic level. For example, the K-quasiconformal map $f(re^{i\theta}) = r^K e^{i\theta}$ sends the ball $B(1,1)$ to a region U containing $[0, 2^K]$. The ratio of inradius to outradius of U at $z = 1$ is about $2^K \gg K$ when K is large.

The distortion of macroscopic balls under f *is* bounded in terms of $K(f)$; it's just that the best bound is not K.

Vector fields. A vector field $v(z)\partial/\partial z$ is *quasiconformal* if its Beltrami differential $\mu = \overline{\partial} v$ is in L^∞ as a distribution. A quasiconformal vector field is continuous; in fact $v \in C^{1-\epsilon}$ for any $\epsilon > 0$.

For many purposes we can get by with the infinitesimal form of the Measurable Riemann Mapping Theorem.

Theorem 5.28 (Infinitesimal version) *For any $\mu \in M(\widehat{\mathbb{C}})$, there is a continuous vector field v on the sphere such that $\overline{\partial} v = \mu$.*

For the proof we recall that $1/(\pi z)$ is a fundamental solution to the $\overline{\partial}$-equation; that is, $(1/z)_{\overline{z}} = \pi\delta$, a multiple of the δ-function at $z = 0$.

To see this, consider any $f \in C_0^\infty(\mathbb{C})$. In terms of the pairing between smooth functions and distributions, we have

$$\begin{aligned}
\langle (1/z)_{\overline{z}}, f\rangle &= \int_{\mathbb{C}} f(z) \frac{\partial}{\partial \overline{z}} \frac{1}{z} \, |dz|^2 = -\int \frac{1}{z} \frac{\partial f}{\partial \overline{z}} \, |dz|^2 \\
&= \frac{1}{2i} \int \frac{1}{z} \frac{\partial f}{\partial \overline{z}} \, dz \wedge d\overline{z} = \frac{1}{2i} \int \frac{dz}{z} \wedge \overline{\partial} f = \frac{-1}{2i} \int d\left(\frac{f(z)\, dz}{z}\right) \\
&= \lim_{r \to 0} \frac{1}{2i} \int_{S^1(r)} f(z) \, \frac{dz}{z} = \frac{1}{2i} 2\pi i f(0) = \pi f(0) = \langle \pi\delta, f\rangle,
\end{aligned}$$

and thus $(1/z)_{\overline{z}} = -\pi\delta$.

Proof of Theorem 5.28. It suffices to handle the case where μ is compactly supported in $\mathbb{C}$, since up to the action of $\mathrm{Aut}(\mathbb{C})$ any Beltrami differential is a sum of two of this form.

To prove the theorem for compactly supported μ, just let $v(z)$ be the convolution of μ with the fundamental solution $-1/(\pi z)$ the $\overline{\partial}$ equation.

Since $1/z$ is in L^1 on compact sets, it is easy to see directly that v is continuous. In fact v is $C^{1-\epsilon}$ for any $\epsilon > 0$. ■

Examples. Given complex numbers $(a_1, \ldots, a_n)$, let

$$\mu(z) = \begin{cases} \displaystyle\sum_1^n k a_k \overline{z}^{k-1}, & |z| \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\overline{\partial} v = \mu$ where

$$v = \begin{cases} \displaystyle\sum_1^n a_k \overline{z}^k, & |z| \le 1, \\ \displaystyle\sum_1^n a_k \frac{1}{z^k} & |z| > 1. \end{cases}$$

Trivial deformations. A Beltrami differential $\mu \in M(\Delta)$ is *trivial* if there is a solution to $\overline{\partial} v = \mu$ such that $v = 0$ on S^1.

Using the preceding examples we can see:

Proposition 5.29 *There exists an infinite-dimensional space of compactly supported Beltrami differentials $V \subset M(\Delta)$ such that $\mu \in V$ is trivial iff $\mu = 0$.*

Proof. Let V be the span of $\langle \overline{z}^k d\overline{z}/dz \,:\, k \ge 0\rangle$. Suppose $\mu \in V$ is trivial; that is, there is a solution to $\overline{\partial} w = \mu$ with $w = 0$ on S^1. Let v be the solution to $\overline{\partial} v = \mu$ given in the example above. Then $\overline{\partial}(v - w) = 0$ and thus $v - w$ is holomorphic on the disk. Since $w = 0$ on S^1, we see $v|S^1$ admits a holomorphic extension to the unit disk. But $v|S^1$ is a polynomial in negative powers of z, so we conclude $v = 0$. Therefore $\mu = 0$.

To obtain compact support, restrict the elements of V to the ball $B(0, 1/2)$. If $\overline{\partial} v = \mu \in V$, then $v = 0$ on $|z| = 1$ and so $v = 0$ on $|z| = 1/2$, since v is holomorphic on $1/2 < |z| < 1$. Then $\mu = 0$ by the argument above. ∎

For applications we need to recognize triviality from the behavior of a solution to $\overline{\partial} v = \mu$ on the sphere. This recognition is provided by:

Theorem 5.30 *Let v be a quasiconformal vector field on $\widehat{\mathbb{C}}$, and let*

$$\pi : \Delta \to \Omega \subset \widehat{\mathbb{C}}$$

be a Riemann mapping to a disk. Suppose $\mu = \overline{\partial} v|\Omega$ is compactly supported, and $v = 0$ on $\partial\Omega$. Then $\pi^*(v) = 0$ on $\partial\Delta$, and $\pi^*(\mu)$ is trivial in $M(\Delta)$.

Proof. Write $v = v(z)(\partial/\partial z)$. Then $v(z)$ is holomorphic outside a compact subset of Ω, and $v(z) = 0$ on $\partial\Omega$. Therefore

$$\pi^*(v) = \frac{v(\pi(z))}{\pi'(z)} \frac{\partial}{\partial z} = \frac{w(z)}{\pi'(z)} \frac{\partial}{\partial z}$$

with $w(z)$ holomorphic outside a compact subset of Δ, and $w(z) \to 0$ on $\partial\Delta$. By Schwarz reflection, $w(z)$ is identically zero near the boundary of Δ, and thus $\pi^*(v)$ is a compactly supported vector field in Δ. Since $\overline{\partial}\pi^*(v) = \pi^*(\mu)$, we find μ is trivial in $M(\Delta)$. ∎

5.10 Deformations of rational maps

We now explain how f-invariant Beltrami differentials determine deformations of f.

Conformal deformations. The space Rat_d of all rational maps of degree d is a complex manifold, isomorphic to an open subset of $\mathbb{P}^{2d+1}$. (If $f(z) = P(z)/Q(z)$, then $\deg(P), \deg(Q) \leq d$ so we have $2d + 2$ coefficients, giving homogeneous coordinates on projective space. The condition that P and Q are relatively prime, and at least one is of degree d, determines a Zariski open subset.)

Suppose a rational map $f_0(z)$ varies smoothly in a family $f_t(z)$. Then the derivative $w = df_t(z)/dt$ is a vector field whose value $w(z)$ lies in the tangent space $T_{f(z)}\widehat{\mathbb{C}}$. In other words, w is a holomorphic section of the bundle $f^*(T\widehat{\mathbb{C}})$. Thus

$$T_f \,\mathrm{Rat}_d \;\; = \;\; \Gamma(\widehat{\mathbb{C}}, f^*(T\widehat{\mathbb{C}})).$$

The pullback of the tangent bundle has degree $2d$, so it is isomorphic to $\mathcal{O}(2d)$, and thus

$$\dim T_f \,\mathrm{Rat}_d = 2d + 1,$$

consistent with our description of Rat_d as an open subset of $\mathbb{P}^{2d+1}$.

A continuous vector field v on $\widehat{\mathbb{C}}$ gives a *deformation* of f iff

$$\delta v = Df(v) - v \circ f$$

is a *holomorphic* section of $f^*(T\widehat{\mathbb{C}})$. Since

$$\overline{\partial}\delta v = f_*\mu - \mu \circ f,$$

we see v gives a deformation iff $\mu = \overline{\partial} v$ is f-invariant, i.e. $f^*\mu = \mu$.

The deformation produced by v can be interpreted as df_t/dt, where

$$f_t = \phi_t^{-1} \circ f \phi_t$$

for a isotopy ϕ_t of $\widehat{\mathbb{C}}$ with $d\phi_t/dt = v$. Thus deformations give directions in which f is (formally) varying by conjugacy.

In particular, we should only expect the full tangent $T_f \mathrm{Rat}_d$ to be spanned by deformations when f is structurally stable.

A 3-dimensional subspace of deformations arises from the action of $\mathrm{PSL}_2(\mathbb{C})$ on Rat_d by conjugation. Namely, if $v \in sl_2(\mathbb{C})$ is a holomorphic vector field on the sphere, then δv is clearly holomorphic.

A *trivial deformation* is a vector field with $\delta v = 0$, which is equivalent to the condition $f^*(v) = v$. If $f(z) = z$ and $f'(z) \neq 1$, we can conclude that $f(v) = 0$. Applying the same reasoning to the repelling periodic points, by continuity of v we have:

Proposition 5.31 *If a continuous vector field v gives a trivial deformation of f, then v vanishes identically on the Julia set of f.*

Corollary 5.32 *If v is holomorphic and $\delta v = 0$ then $v = 0$.*

Intuitively, the Proposition says that a small deformation commuting with f fixes the repelling cycles, and so it is the identity on the Julia set. The Corollary says the group $\mathrm{Aut}(f)$ of Möbius transformations commuting with f is discrete. (In fact $\mathrm{Aut}(f)$ is finite.)

Because of this Corollary, $sl_2(\mathbb{C})$ maps injectively into the space of deformation of f. Taking the quotient, we obtain the 'cohomology group'

$$H^1(f, \mathrm{T}\widehat{\mathbb{C}}) = H^0(\widehat{\mathbb{C}}, f^*(\mathrm{T}\widehat{\mathbb{C}}))/sl_2(\mathbb{C}).$$

This group measures deformations modulo those that come from conformal conjugacy. It is naturally the tangent space at f to the variety $V_d = \mathrm{Rat}_d / \mathrm{PSL}_2(\mathbb{C})$. We have

$$\dim H^1(f, \mathrm{T}\widehat{\mathbb{C}}) = 2d - 2.$$

Quasiconformal deformations. Let $M(\widehat{\mathbb{C}})^f \subset M(\widehat{\mathbb{C}})$ be the space of f-invariant Beltrami differentials, i.e. those satisfying $f^*(\mu) = \mu$. Each invariant μ determines an almost-complex structure on $\widehat{\mathbb{C}}$ with respect to which f is holomorphic. Usually $\dim M(\widehat{\mathbb{C}})^f = \infty$. For example, if an open set U is disjoint from its forward images, and $f^n|U$ is injective, then we can propagate any $\mu \in M(U)$ to an f-invariant differential on the sphere, and thus $M(U) \hookrightarrow M(\widehat{\mathbb{C}})^f$.

We have a natural map

$$\delta : M(\widehat{\mathbb{C}})^f \to H^1(f, \mathrm{T}\widehat{\mathbb{C}})$$

defined as follows: given an f-invariant μ, solve the equation $\overline{\partial} v = \mu$ and then map μ to the deformation δv. The solution to the $\overline{\partial}$-equation is only well-defined up to elements of $sl_2(\mathbb{C})$, which is why the image lies in the cohomology group.

To under δ, note that $\mu \in M(\widehat{\mathbb{C}})^f$ determines a 1-parameter family of quasiconformal maps ϕ_t with complex dilatation $t\mu$. The rational maps $f_t = \phi_t \circ f \circ \phi_t^{-1}$ vary holomorphically with respect to t, and by differentiating at $t = 0$ we obtain the deformation $[df_t/dt] = \delta\mu$.

5.11 No wandering domains

Using the fact that a rational map admits only a finite-dimensional space of deformations, we can finally prove:

Theorem 5.33 (No wandering domains) *Every component of the Fatou set eventually cycles.*

We begin with an observation due to N. Baker.

Lemma 5.34 *If $\Omega(f)$ has a wandering domain, then it has a wandering disk.*

Proof. Let U be a wandering component of $\Omega(f)$, and let $U_n = f^n(U)$ be its distinct forward images (each a component of $\Omega(f)$). Since f has just finitely many critical values, $f : U_n \to U_{n+1}$ is a covering map, hence a hyperbolic isometry, for all $n \gg 0$.

We claim U_n is a disk for all $n \gg 0$. Otherwise, for all $n \gg 0$, U_n carries a hyperbolic geodesic γ_n such that $f(\gamma_n) = \gamma_{n+1}$. (Note that U_n cannot be a punctured disk because $J(f)$ is perfect.) Since f is a covering, the length of γ_n in the hyperbolic metric on U_n is the same for all $n \gg 0$. Now the diameter of the largest spherical ball contained in U_n must tend to zero, since these components are distinct. By comparing the hyperbolic and spherical metrics, we see $\mathrm{diam}_\sigma(\gamma_n) \to 0$. Since f is Lipschitz in the spherical metric, once the spherical size of γ_n is small, any small disk bounded by γ_n must map to a small disk bounded by γ_{n+1}. It follows that f^{n+i} is normal on this small disk, which is impossible since each component of $\widehat{\mathbb{C}} - \gamma_n$ must meet $J(f)$. Thus U_n is eventually a disk. ∎

Proof of Theorem 5.33. Let U be a wandering disk, or more precisely a simply-connected component of $\Omega(f)$ such that $f^n|U$ is injective for all $n > 0$. Let $\pi : \Delta \to U$ be a Riemann mapping. Then π allows us to transfer Beltrami differentials from Δ to U. Let $V \subset M(\Delta)$ be the infinite-dimensional space of compactly supported Beltrami differentials guaranteed by Proposition 5.29, such that $\mu \in V$ is trivial iff $\mu = 0$. Then we have

$$V \subset M(\Delta) \overset{\pi_*}{\cong} M(U) \hookrightarrow M(\widehat{\mathbb{C}})^f \xrightarrow{\delta} H^1(f, \mathrm{T}\widehat{\mathbb{C}}).$$

Now suppose $\mu \in V$ maps to $0 \in H^1(f, \mathrm{T}\widehat{\mathbb{C}})$. Then the associated vector field v vanishes on $J(f)$, and in particular $v = 0$ on ∂U. By Theorem 5.30, this implies μ is a trivial deformation of the disk, and thus $\mu = 0$.

Thus V maps injectively into $T_f(\mathrm{Rat}_d)$. But $\dim(V) = \infty > \dim H_1(f, \mathrm{T}\widehat{\mathbb{C}}) = 2d - 2$, so this is impossible. Therefore f has no wandering domain. ■

5.12 Finiteness of periodic regions

Theorem 5.35 *Every component of the Fatou set $\Omega(f)$ eventually cycles, and there are only finitely many periodic components.*

Proof. We have seen there are no wandering domains. Apart from Herman rings, the periodic components of $\Omega(f)$ are associated to superattracting, attracting, parabolic or indifferent periodic points, so they are finite in number.

For Herman rings, one can use the quasiconformal deformation theory to show each cycle of rings contributes one complex parameter to the moduli space of f. Since $\dim \mathcal{M}(f) \leq 2d - 2$, there are at most $2d - 2$ Herman rings. ■

Remarks. More sophisticated arguments show the number of cycles of components of the Fatou set is at most $2d - 2$ [Shi].

5.13 The Teichmüller space of a dynamical system

In preparation to discuss the Teichmüller space of a rational map, we outline the more general theory of the Teichmüller space of a dynamical system.

Definitions. A *holomorphic dynamical system* (X, f) consists of a 1-dimensional complex manifold X, possibly disconnected, and a holomorphic map $f : X \to X$. (The definition can be extended in a natural way to replace f by a group action or a semigroup or a collection of holomorphic correspondences.)

An *isomorphism* $\alpha : (X_1, g_1) \to (X_2, g_2)$ is given by a conformal map $\alpha : X_1 \to X_2$ such that $\alpha \circ g_1 = g_2 \circ \alpha$.

A quasiconformal conjugacy $\phi : (X, f) \to (Y, g)$ gives a *marking* of (Y, g) by (X, f). Two such marked dynamical systems are isomorphic if there is an isomorphism $\alpha : (Y_1, g_1) \to (Y_2, g_2)$ respecting markings (i.e. such that $\phi_2 = \alpha \circ \phi_1$).

The *deformation space* $\mathrm{Def}(X, f)$ is the space of isomorphism classes of dynamical systems marked by (X, f). It is easy to see the deformation space may be identified with the unit ball in the space of f-invariant Beltrami differentials; that is, we have:

$$\mathrm{Def}(X, f) = M(X)_1^f.$$

The quasiconformal conjugacies $\phi : (X, f) \to (X, f)$ form a group we denote $\mathrm{QC}(X, f)$. It acts on $\mathrm{Def}(X, f)$ by changing the marking. The normal subgroup $\mathrm{QC}_0(X, f)$ consists of quasiconformal conjugacies isotopic to the identity, *rel* the ideal boundary of X and through uniformly quasiconformal conjugacies.

The *Teichmüller space* $\mathrm{Teich}(X, f)$ is the quotient space of dynamical systems marked up to isotopy:

$$\mathrm{Teich}(X, f) = \mathrm{Def}(X, f)/\,\mathrm{QC}_0(X, f).$$

The *mapping-class group*

$$\mathrm{Mod}(X, f) = \mathrm{QC}(X, f)/\,\mathrm{QC}_0(X, f)$$

acts on Teichmüller space, yielding as quotient the *moduli space*

$$\mathcal{M}(X, f) = \mathrm{Teich}(X, f)/\,\mathrm{Mod}(X, f)$$

of isomorphism classes of dynamical systems quasiconformally conjugate to f.

The subgroup of conformal conjugacies will be denoted $\mathrm{Aut}(X, f) \subset \mathrm{QC}(X, f)$. Its image in $\mathrm{Mod}(X, f)$ coincides with the stabilizer of $[(X, f)] \in \mathrm{Teich}(X, f)$.

Rotation of an annulus. To make these ideas concrete, let us consider the Teichmüller space of $(X, f) = (A(R), f)$, where $A(R) = \{1 < |z| < R\}$ is the standard annulus of modulus $\log(R)/2\pi$, and

$$f(z) = e^{2\pi i\alpha} z$$

is an irrational rotation.

1. We have $\mathrm{Aut}(X, f) = S^1$ acting by rotations. Moreover, the dynamical system generated by f is dense S^1.

2. The space $\mathrm{Def}(X, f) = M(A(R))_1^f$ can be identified with $L^\infty([1, R])_1$. Indeed, any f-invariant Beltrami differential μ on $A(R)$ is also S^1-invariant, by ergodicity of the irrational rotation of a circle. Thus μ is determined by its values on the radius $[1, R]$.

3. In terms of marked dynamical systems, $\mathrm{Def}(X, f)$ can be identified with the set of triples $(Y, g, \phi) = (A(S), f, \phi)$ such that $\phi : A(R) \to A(S)$ is an S^1-equivariant quasiconformal map and

   $$\phi | S^1(1) = \mathrm{id}\,.$$

   To see this, just note that any Riemann surface Y quasiconformally equivalent to X is of the form $A(S)$, that the rotation number of f is a topological invariant, and that by composing with an automorphism of $A(S)$ we can normalize the conjugacy so that $\phi(1) = 1$.

4. The group $\mathrm{QC}(X, f)$ consists of all quasiconformal maps $\phi : A(R) \to A(R)$ that commute with rotations. It is conveniently to visualize the orbits of S^1 on $A(R)$ as circles of constant radius in the cylinderical metric $|dz|/|z|$; then ϕ preserves this foliation and maps leaves to leaves by isometries.

   Any $\phi \in \mathrm{QC}(X, f)$ is actually Lipschitz and is determined by its values on $[1, R]$. Thus ϕ has the form
$$\phi(z) = e^{2\pi i\theta(|z|)} \cdot z,$$
   where $\theta : [1, R] \to \mathbb{C}$ is a Lipschitz function such that $\theta(1)$ and $\theta(R)$ are real. Note that θ and $\theta + n$, $n \in \mathbb{Z}$, determine the same map ϕ.

   Moreover, the condition that ϕ as above is quasiconformal is equivalent to the condition that θ is Lipschitz and $r \mapsto |\phi(r)| = r \exp(2\pi \operatorname{Im} \theta(r))$ is a bilipschitz map. In other words, if we write $\theta(r) = \alpha(r) + i\beta(r)$ then we need α and β to be Lipschitz and we need a definite gap betwen $1/r$ and $2\pi\beta'(r)$; that is, we need
$$\inf_r((1/r) - 2\pi\beta'(r)) > 0$$
   to insure $|\phi(r)|$ increases at a definite rate.

5. The group $\mathrm{QC}_0(X, f)$ consists of those ϕ that can be deformed to the identity while keeping the values of ϕ on $S^1(1)$ and $S^1(R)$ fixed. This means that ϕ is the identity on $\partial A(R)$ and moreover that there is no relative twisting. In terms of $\theta(r)$, the condition $\phi \in \mathrm{QC}_0(R)$ is equivalent to the condition
$$\theta(0) = \theta(R) \in \mathbb{Z}.$$

6. A linear combination of Lipschitz functions is again Lipschitz. Using this fact, it is easy to see that the map $\phi \mapsto (\theta(1), \theta(R))$ gives an isomorphism
$$\mathrm{Mod}(X, f) = \mathrm{QC}(X, f)/\,\mathrm{QC}_0(X, f) \cong \mathbb{R}^2/\mathbb{Z},$$
   where $\mathbb{Z} = \{(n, n) \in \mathbb{R}^2\}$.

7. Now we have seen that points $(A(S), f, \phi) \in \mathrm{Def}(X, f)$ can be normalized so that $\phi(1) = 1$, i.e. such that $\theta(1) = 0$. Therefore ϕ is determined up to isotopy (rel ideal boundary) by the value of $\theta(R) \in \mathbb{R}$. Thus the map
$$\begin{array}{rcl} (A(S), f, \phi) & \mapsto & -(\log \phi(R))/2\pi i \\ & = & \theta(R) + i \operatorname{mod}(A(S)) \end{array}$$
   establishes an isomorphism
$$\mathrm{Teich}(X, f) \cong \mathbb{H}.$$
   It can be verified that this isomorphism is complex-analytic.

8. Finally $\mathrm{Mod}(X, f) = (\mathbb{R}^2/\mathbb{Z})$ acts on $\mathrm{Teich}(X, f)$ by changing the boundary values of ϕ. This actions is on $z \in \mathbb{H}$ given by:

$$(s, t)(z) = z + (t - s).$$

   Note that any point $z \in \mathbb{H}$ is stabilized by the diagonal subgroup $\mathbb{R}/\mathbb{Z} \cong S^1 \subset \mathbb{R}^2/\mathbb{Z}$, consistent with the fact that $\mathrm{Aut}(Y, f) = S^1$ for every $(Y, f) \in \mathrm{Teich}(X, f)$.

9. Taking the quotient by the group of all translations of $\mathbb{H}$, we find the map $(Y, g) \mapsto \mathrm{mod}(Y)$ (or equivalently $z \mapsto \mathrm{Im}\, z$) gives an isomorphism

$$\mathcal{M}(X, f) = \mathrm{Teich}(X, f)/\, \mathrm{Mod}(X, f) \cong \mathbb{R}_+ = (0, \infty).$$

   In other words, the rotations of an annulus Y through irrational angle θ are classified up to conjugacy by $\mathrm{mod}(Y)$.

Attracting fixed-points. As a second example that is important for the theory of attracting fixed-points, let $(X, f) = (\mathbb{C}, f)$ where

$$f(z) = \lambda z$$

with $0 < |\lambda| < 1$. For this map, a central role in the Teichmüller theory is placed by the quotient torus $T = \mathbb{C}^*/\langle f \rangle$.

1. The automorphism group of $(\mathbb{C}, f)$ is the multiplicative group $\mathbb{C}^*$.

2. The space $\mathrm{Def}(X, f)$ is naturally identified with $M(T)^1$. (Invariant Beltrami differential come from the quotient T.) An invariant Beltrami differential is determined by its values in the fundamental annulus $|\lambda| < |z| < 1$.

3. In terms of dynamical systems, any $(Y, g, \phi) \in \mathrm{Def}(X, f)$ has the form $(Y, g) = (\mathbb{C}, \kappa z)$ with $0 < |\kappa| < 1$. The map ϕ can be normalized so that $\phi(1) = 1$.

4. The dynamical system (Y, g) has its own quotient torus $T' = \mathbb{C}^*/\langle \kappa \rangle$, naturally marked by ϕ, giving an isomorphism

$$\mathrm{Teich}(X, f) \cong \mathrm{Teich}(T) \cong \mathbb{H}.$$

   In terms of $(Y, g) = (\mathbb{C}, \kappa z)$, this isomorphism is given by

$$(\mathbb{C}, \kappa z) \mapsto \frac{\log \kappa}{2\pi i} \in \mathbb{H}.$$

   The value of $\log \kappa$ can be made well-defined using the marking of (Y, g) by (X, f). To use the marking, first choose an arc $\gamma \subset X = \mathbb{C}^*$ connecting 1 to λ. Then $\phi(\gamma) \subset Y$ connects 1 to κ. There is a unique continous branch of the logarithm along $\phi(\gamma)$ with $\log(1) = 0$, and this branch defines $\log \kappa$.

5. The torus T has a distinguished map $\pi_1(T) \to \mathbb{Z}$ defining the covering space $X = \mathbb{C}^* \to T$. In other words, T has a distinguished cohomology class $C \in H^1(T, \mathbb{Z})$. It is not hard to see that $\mathrm{Mod}(X, f) \cong \mathbb{Z}$ coincides with the subgroup of $\mathrm{Mod}(T) \cong \mathrm{SL}_2(\mathbb{Z})$ stabilizing C. Thus we have the isomorphism

$$\mathrm{Mod}(X, f) \cong \mathbb{Z} \cong \{\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})\}.$$

6. The modular group acts on $\mathbb{H}$ in the natural way — by sending z to $z+n$. Thus we have

$$\mathcal{M}(X, f) \cong \mathbb{H}/\mathbb{Z} \cong \Delta^*.$$

   This isomorphism simply sends $(\mathbb{C}^*, \kappa z)$ to κ.

7. In summary, the Teichmüller space of an attracting fixed-point coincides with the Teichmüller space of its quotient torus; the torus carries a distinguished cohomology class; a marking determines a logarithmic lift of the multiplier; and the moduli space is the punctured disk.

5.14 The Teichmüller space of a rational map

We are now in a position to describe the space of all rational maps quasiconformally conjugate to a given one. See [McS] for details.

The quotient Riemann surface. To relate $\mathrm{Teich}(f)$ to more traditional Teichmüller spaces, we introduce a quotient Riemann surface for f.

The *grand orbits* of f are the equivalence classes of the relation $x \sim y$ if $f^n(x) = f^m(y)$ for some $n, m \geq 0$. Let $\widehat{J}$ denote the closure of the grand orbits of all periodic points and all critical points of f. Let $\widehat{F} = \widehat{\mathbb{C}} - \widehat{J} \subset \Omega$. Then $f : \widehat{F} \to \widehat{F}$ is a covering map without periodic points. Let $\widehat{F} = \Omega^{\mathrm{dis}} \sqcup \Omega^{\mathrm{fol}}$ denote the partition into open sets where the grand orbit equivalence relation is discrete and where it is indiscrete.

Theorem 5.36 *The Teichmüller space of a rational map f of degree d is naturally isomorphic to*

$$M_1(J, f) \times \mathrm{Teich}(\Omega^{fol}, f) \times \mathrm{Teich}(\Omega^{dis}/f),$$

where Ω^{dis}/f is a complex manifold.

Here $M_1(J, f)$ denotes the unit ball in the space of f-invariant Beltrami differentials on J. In particular, $M_1(J, f) = 0$ if the Julia set has measure zero.

Proof. Since $\widehat{J}$ contains a dense countable dynamically distinguished subset, $\omega|\widehat{J} = \mathrm{id}$ for all $\omega \in \mathrm{QC}_0(\widehat{\mathbb{C}}, f)$, and so

$$\mathrm{QC}_0(\widehat{\mathbb{C}}, f) = \mathrm{QC}_0(\Omega^{\mathrm{fol}}, f) \times \mathrm{QC}_0(\Omega^{\mathrm{dis}}, f).$$

This implies the theorem with the last factor replaced by $\mathrm{Teich}(\Omega^{\mathrm{dis}}, f)$, using the fact that J and $\widehat{J}$ differ by a set of measure zero. To complete the proof, write Ω^{dis} as $\bigcup \Omega_i^{\mathrm{dis}}$, a disjoint union of totally invariant open sets such that each quotient $\Omega_i^{\mathrm{dis}}/f$ is connected. Then we have

$$\mathrm{Teich}(\Omega^{\mathrm{dis}}, f) = \prod{}' \mathrm{Teich}(\Omega_i^{\mathrm{dis}}, f) = \prod{}' \mathrm{Teich}(\Omega_i^{\mathrm{dis}}/f) = \mathrm{Teich}(\Omega^{\mathrm{dis}}/f).$$

Note that Ω^{dis}/f is a complex manifold (rather than an orbifold), because $f|\Omega^{\mathrm{dis}}$ has no periodic points. ■

The dimension of the Teichmüller space of a rational map of degree d is at most $2d-2$, i.e. at most the dimension of $\mathrm{Rat}_d \,/\, \mathrm{Aut}\,\widehat{\mathbb{C}}$.

Next we give a more concrete description of the factors appearing in Theorem 5.36.

By the classification of stable regions, it is easy to see that $\widehat{J}$ is the union of:

1. The Julia set of f;
2. The grand orbits of the attracting and superattracting cycles and the centers of Siegel disks (a countable set);
3. The grand orbits of the critical points that land in attracting and parabolic basins (a countable set); and
4. The leaves of the canonical foliations which meet the grand orbit of the critical points (a countable union of one-dimensional sets).

The superattracting basins, Siegel disks and Herman rings of a rational map are canonical foliated by the components of the *closures* of the grand orbits. In the Siegel disks, Herman rings, and near the superattracting cycles, the leaves of this foliation are real-analytic circles. In general countably many leaves may be singular. Thus Ω^{dis} contains the points which eventually land in attracting or parabolic basins, while Ω^{fol} contains those which land in Siegel disks, Herman rings and superattracting basins.

Theorem 5.37 *The quotient space Ω^{dis}/f is a finite union of Riemann surfaces, one for each cycle of attractive or parabolic components of the Fatou set of f.*

An attractive basin contributes an n-times punctured torus to Ω^{dis}/f, while a parabolic basin contributes an $(n+2)$-times punctured sphere, where $n \geq 1$ is the number of grand orbits of critical points landing in the corresponding basin.

Proof. Every component of Ω^{dis} is preperiodic, so every component X of Ω^{dis}/f can be represented as the quotient Y/f^p, where Y is obtained from a parabolic or attractive basin U of period p be removing the grand orbits of critical points and periodic points.

First suppose U is attractive. Let x be the attracting fixed point of f^p in U, and let $\lambda = (f^p)'(x)$. After a conformal conjugacy if necessary, we can assume $x \in \mathbb{C}$. Then by classical results, there is a holomorphic linearizing map

$$\psi(z) = \lim \lambda^{-n}(f^{pn}(z) - x)$$

mapping U onto $\mathbb{C}$, injective near x and satisfying $\psi(f^p(z)) = \lambda(\psi(z))$.

Let U' be the complement in U of the grand orbit of x. Then the space of grand orbits in U' is isomorphic to $\mathbb{C}^*/ < z \mapsto \lambda z >$, a complex torus. Deleting the points corresponding to critical orbits in U', we obtain Y/f^p.

Now suppose U is parabolic. Then there is a similar map $\psi : U \to \mathbb{C}$ such that $\psi(f^p(z)) = z + 1$, exhibiting $\mathbb{C}$ as the small orbit quotient of U. Thus the space of grand orbits in U is the infinite cylinder $\mathbb{C}/ < z \mapsto z+1 > \cong \mathbb{C}^*$. Again deleting the points corresponding to critical orbits in U, we obtain Y/f^p.

In both cases, the number n of critical orbits to be deleted is at least one, since the immediate basin of an attracting or parabolic cycle always contains a critical point. Thus the number of components of $\Omega^{\rm dis}/f$ is bounded by the number of critical points, namely $2d - 2$. ■

For a detailed development of attracting and parabolic fixed points, see e.g. [CG, Chapter II].

Definitions. A critical point is *acyclic* if its forward orbit is infinite. Two points x and y in the Fatou set are in the same *foliated equivalence class* if the closures of their grand orbits agree. For example, if x and y are on the same leaf of the canonical foliation of a Siegel disk, then they lie in a single foliated equivalence class. On the other hand, if x and y belong to an attracting or parabolic basin, then to lie in the same foliated equivalence class they must have the same grand orbit.

Theorem 5.38 *The space* $\operatorname{Teich}(\Omega^{fol}, f)$ *is a finite-dimensional polydisk, whose dimension is given by the number of cycles of Herman rings plus the number of foliated equivalence classes of acyclic critical points landing in Siegel disks, Herman rings or superattracting basins.*

Proof. As for $\Omega^{\rm dis}$, we can write $\Omega^{\rm fol} = \bigcup \Omega_i^{\rm fol}$, a disjoint union of totally invariant open sets such that $\Omega_i^{\rm fol}/f$ is connected for each i. Then

$$\operatorname{Teich}(\Omega^{\rm fol}, f) = \prod{}' \operatorname{Teich}(\Omega_i^{\rm fol}, f).$$

Each factor on the right is either a complex disk or trivial. Each disk factor can be lifted to $\operatorname{Def}(\widehat{\mathbb{C}}, f)$, so by finiteness of the space of rational maps the number of disk factors is finite. A disk factor arises whenever $\Omega_i^{\rm fol}$ has an annular component. A cycle of foliated regions with n critical leaves gives n periodic annuli in the Siegel disk case, $n + 1$ in the case of a Herman ring, and n wandering annuli in the superattracting case. If two critical points account for the same leaf, then they lie in the same foliated equivalence class. ■

Definition. An *invariant line field* on a positive-measure totally invariant subset E of the Julia set is the choice of a real 1-dimensional subspace $L_e \subset T_e\widehat{\mathbb{C}}$, varying measurably with respect to $e \in E$, such that f' transforms L_e to $L_{f(e)}$ for almost every $e \in E$.

Equivalently, an invariant line field is given by a measurable Beltrami differential μ supported on E with $|\mu| = 1$, such that $f^*\mu = \mu$. The correspondence is given by $L_e = \{v \in T_e\widehat{\mathbb{C}} \,:\, \mu(v) = 1 \text{ or } v = 0\}$.

Theorem 5.39 *The space $M_1(J, f)$ is a finite-dimensional polydisk, whose dimension is equal to the number of ergodic components of the maximal measurable subset of J carrying an invariant line field.*

Corollary 5.40 *On the Julia set there are finitely many positive measure ergodic components outside of which the action of the tangent map of f is irreducible.*

Theorem 5.41 (Number of moduli) *The dimension of the Teichmüller space of a rational map is given by $n = n_{AC} + n_{HR} + n_{LF} - n_P$, where*

- *n_{AC} is the number of foliated equivalence classes of acyclic critical points in the Fatou set,*
- *n_{HR} is the number of cycles of Herman rings,*
- *n_{LF} is the number of ergodic line fields on the Julia set, and*
- *n_P is the number of cycles of parabolic basins.*

Proof. The Teichmüller space of an n-times punctured torus has dimension n, while that of an $n+2$-times punctured sphere has dimension $n-1$. Thus the dimension of $\mathrm{Teich}(\Omega^{\mathrm{dis}}/f)$ is equal to the number of grand orbits of acyclic critical points in Ω^{dis}, minus n_P. We have just seen the number of remaining acyclic critical orbits (up to foliated equivalence), plus n_{HR}, gives the dimension of $\mathrm{Teich}(\Omega^{\mathrm{fol}}, f)$. Finally n_{LF} is the dimension of $M_1(J, f)$. ∎

Remark. The number n_P can exceed the number of parabolic cycles. For example, a parabolic fixed point can have many petals attached, and these petals may fall into several distinct cycles under the dynamics.

5.15 The modular group of a rational map

Recall that $\mathrm{Mod}(\widehat{\mathbb{C}}, f) = \mathrm{QC}(\widehat{\mathbb{C}}, f)/\mathrm{QC}_0(\widehat{\mathbb{C}}, f)$ is the group of quasiconformal automorphisms of f, modulo those isotopic (through conjugacies) to the identity. In this section we will prove:

Theorem 5.42 (Discreteness of the modular group) *The group* $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ *acts properly discontinuously by holomorphic automorphisms of* $\mathrm{Teich}(\widehat{\mathbb{C}}, f)$.

Proof of Theorem 5.42 (Discreteness of the modular group). The group $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ acts isometrically on the finite-dimensional complex manifold $\mathrm{Teich}(\widehat{\mathbb{C}}, f)$ with respect to the Teichmüller metric. The stabilizer of a point $([\phi], \widehat{\mathbb{C}}, g)$ is isomorphic to $\mathrm{Aut}(g)$ and hence finite; thus the quotient of $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ by a finite group acts faithfully. By compactness of quasiconformal maps with bounded dilatation, $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ maps to a closed subgroup of the isometry group; thus $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ is a Lie group. If $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ has positive dimension, then there is an arc $(\phi_t, \widehat{\mathbb{C}}, f)$ of inequivalent markings of f in Teichmüller space; but such an arc can be lifted to the deformation space, which implies each ϕ_t is in $\mathrm{QC}_0(f)$, a contradiction.

Therefore $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ is discrete. ∎

Remark. Equivalently, we have shown that for any $K > 1$, there are only a finite number of non-isotopic quasiconformal automorphisms of f with dilatation less than K.

Let Rat_d denote the space of all rational maps $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ of degree d. This space can be realized as the complement of a hypersurface in projective space $\mathbb{P}^{2d+1}$ by considering $f(z) = p(z)/q(z)$ where p and q are relatively prime polynomials of degree d in z. The group of Möbius transformations $\mathrm{Aut}(\widehat{\mathbb{C}})$ acts on Rat_d by sending f to its conformal conjugates.

A *complex orbifold* is a space which is locally a complex manifold divided by a finite group of complex automorphisms.

Corollary 5.43 (Uniformization of conjugacy classes) *There is a natural holomorphic injection of complex orbifolds*

$$\mathrm{Teich}(\widehat{\mathbb{C}}, f)/\,\mathrm{Mod}(\widehat{\mathbb{C}}, f) \to \mathrm{Rat}_d\,/\,\mathrm{Aut}(\widehat{\mathbb{C}})$$

parameterizing the rational maps g quasiconformal conjugate to f.

Corollary 5.44 *If the Julia set of a rational map is the full sphere, then the group $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ maps with finite kernel into a discrete subgroup of $PSL_2(\mathbb{R})^n \ltimes S_n$ (the automorphism group of the polydisk).*

Proof. The Teichmüller space of f is isomorphic to $\mathbb{H}^n$. ∎

Corollary 5.45 (Finiteness theorem) *The number of cycles of stable regions of f is finite.*

Proof. Let d be the degree of f. By Corollary 5.43, the complex dimension of $\mathrm{Teich}(\widehat{\mathbb{C}}, f)$ is at most $2d - 2$. This is also the number of critical points of f, counted with multiplicity.

By Theorem 5.41 f has at most $2d-2$ Herman rings, since each contributes at least a one-dimensional factor to $\mathrm{Teich}(\widehat{\mathbb{C}}, f)$ (namely the Teichmüller space of a foliated annulus. By a classical argument, every attracting, superattracting or parabolic cycle attracts a critical point, so there are at most $2d-2$ cycles of stable regions of these types. Finally the number of Siegel disks is bounded by $4d-4$. (The proof, which goes back to Fatou, is that a suitable perturbation of f renders at least half of the indifferent cycles attracting; cf. [Mon, §106].)

Consequently the total number of cycles of stable regions is at most $8d-8$. ∎

Remark. The sharp bound of $2d-2$ (conjectured in [Sul4]) has been achieved by Shishikura and forms an analogue of Bers' area theorem for Kleinian groups [Shi], [Bers1].

6 Hyperbolic 3-manifolds

6.1 Kleinian groups and hyperbolic manifolds

A *hyperbolic manifold* M^n is a connected, complete Riemannian manifold of constant sectional curvature -1.

There is a unique simply-connected hyperbolic manifold $\mathbb{H}^n$ of dimension n, up to isometry. Thus any hyperbolic manifold can be regarded as a quotient $M^n = \mathbb{H}^n/\Gamma$ where $\Gamma \subset \mathrm{Isom}(\mathbb{H}^n)$ is a discrete group.

Two explicit models for hyperbolic space are the *upper half-space model*,

$$\mathbb{H}^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \, : \, x_n > 0\}$$

with the metric $\rho = |dx|/x_n$; and the *Poincaré ball model*

$$\mathbb{H}^n = \mathbb{B}^n = \{x \in \mathbb{R}^n \, : \, |x| < 1\}$$

with $\rho = 2|dx|/(1-|x|^2)$. Hyperbolic space has a natural *sphere at infinity* S^{n-1}_∞, corresponding to $\mathbb{R}^{n-1} \cup \{\infty\}$ in the upper half-space model and to $\partial\mathbb{B}^n$ in the Poincaré ball model.

The points on S^{n-1}_∞ can be naturally interpreted as endpoints of geodesics.

Theorem 6.1 *For $n > 1$, every hyperbolic isometry extends continuously to a conformal automorphism of the sphere at infinity, establishing an isomorphism*

$$\mathrm{Isom}(\mathbb{H}^n) \cong \mathrm{Aut}(S^{n-1}_\infty).$$

Proof. First note that reflection through a hyperplane $P^{n-1} \subset \mathbb{H}^n$ extends continuously to (conformal) reflection through the sphere $S^{n-2} = \partial P^{n-1} \subset S^{n-1}_\infty$. Conversely, reflection through a sphere extends to reflection through a hyperplane in hyperbolic space. Since reflections generate both groups, we see the boundary values of any hyperbolic isometry are conformal, and any conformal map extends to an isometry.

Finally any isometry inducing the identity on S^{n-1}_∞ must be the identity, since it stabilizes every geodesic. ∎

A *Kleinian group* is a discrete subgroup $\Gamma \subset \mathrm{Isom}(\mathbb{H}^n)$. The *limit set* $\Lambda \subset S^{n-1}_\infty$ of Γ is defined by $\Lambda = \overline{\Gamma x} \cap S^{n-1}_\infty$ for any $x \in \mathbb{H}^n$. The complement $\Omega = S^{n-1}_\infty - \Lambda$ is the *domain of discontinuity* for Γ.

Theorem 6.2 *The action of a Kleinian group on its domain of discontinuity is properly discontinuous. That is, for any compact set $K \subset \Omega$, the set of $\gamma \in \Gamma$ such that $\gamma(K) \cap K \neq \emptyset$ is finite.*

A Kleinian group is *elementary* if it contains an abelian subgroup of finite index; equivalently, if $|\Lambda| \leq 2$.

Theorem 6.3 *If Γ is nonelementary, then Λ is the smallest nonempty closed Γ-invariant subset of S^{n-1}_∞.*

A *baseframe* ω for a hyperbolic manifold M is simply a point in the frame bundle of M. There is a natural bijection:

$$\{\text{Baseframed hyperbolic manifolds } (M^n, \omega)\}$$
$$\leftrightarrow$$
$$\{\text{Torsion-free Kleinian groups } \Gamma \subset \mathrm{Isom}(\mathbb{H}^n)\}.$$

General Kleinian groups correspond to hyperbolic orbifolds. Forgetting the baseframe amounts to only knowing Γ up to conjugacy in $\mathrm{Isom}(\mathbb{H}^n)$.

6.2 Ergodicity of the geodesic flow

Theorem 6.4 *The geodesic flow on $M = \mathbb{H}^n/\Gamma$ is ergodic if and only if Γ acts ergodically on $S^{n-1}_\infty \times S^{n-1}_\infty$.*

Theorem 6.5 *The geodesic flow is ergodic on any finite-volume hyperbolic manifold M.*

Proof. (Hopf) Let $f \in C_0(T_1 M)$ be a compactly supported continuous function on the unit tangent bundle. Let g_t denote the geodesic flow, and $I \subset L^2(T_1 M)$ the subspace of functions invariant under g_t. To prove ergodicity we need to show I consists of the constant functions.

By the ergodic theorem,

$$f_+(v) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(g_t(v))\, dt$$

exists for almost every $v \in T_1 M$, and converges to the L^2-projection F of f to I. On the other hand, the negative time average $f_-(v)$ converges to the same thing, so $F(v) = f_+(v) = f_-(v)$ for almost every v.

Now if v and w are vectors converging to the same point on S^{n-1}_∞ in positive time, then the geodesic rays through v and w are asymptotic, so $f_+(v) = f_+(w)$ by uniform continuity of f. In other words, $F(v)$ is constant along the $(n-$

1)-spheres of the positive horocycle foliation of $T_1(M)$. Applying the same argument to $f_-(v)$, we see F is also constant along the negative horocycle foliation. Finally $F(v)$ is invariant under the geodesic flow. By Fubini's theorem, we conclude that $F(v)$ is constant.

Since $C_0(T_1 M)$ is dense in $L^2(T_1 M)$, we have shown I consists only of the constant functions, and thus the geodesic flow is ergodic. ■

6.3 Quasi-isometry

Let X and Y be complete metric spaces. A map $f : X \to Y$ is a K-*quasi-isometry* if for some $R > 0$ we have

$$R + Kd(x, x') \geq d(f(x), f(x')) \geq \frac{d(x, x')}{K} - R$$

for all $x, x' \in X$. In other words, on a large scale, f gives a bi-Lipschitz map to its image.

We say $f : X \to X$ is *close to the identity* if $\sup_X d(x, f(x)) < \infty$.

We say $f : X \to Y$ is a quasi-isometric *isomorphism* if there is a quasi-isometry $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are close to the identity. If $f : X \to Y$ is a quasi-isometry and $B(f(X), R) = Y$ for some R, then in fact f is an isomorphism.

Example. The inclusion $f : \mathbb{Z}^n \to \mathbb{R}^n$ is a quasi-isometric isomorphism. An inverse is the map $g : \mathbb{R} \to \mathbb{Z}$ defined by taking the integer part, $g(x) = [x]$.

Groups. Let G be a finitely-generated group. Choosing a finite set of generators $\langle g_i \rangle$, we can construct the *Cayley graph* $C(G)$ by taking G as the vertices and connecting g and h by an edge if $g = g_i h$ for some g_i. Taking the edges to be of unit length, we obtain a metric d on G.

Alternatively, one can define $d(\mathrm{id}, g) = n$ where n is the length of a minimal word expressing g in terms of $\langle g_i^{\pm 1} \rangle$, and then extend d to $G \times G$ so it is right-invariant.

Another choice of generators $\langle g_i' \rangle$ determines another metric d' on G. By expressing each g_i as a word in $\langle g_i' \rangle$ and vice-versa, one easily sees that (G, d) and (G, d') are quasi-isometric.

Theorem 6.6 *For any compact Riemannian manifold M, the universal cover $\widetilde{M}$ and the group $\pi_1(M)$ are quasi-isometric.*

Proof. Realize $\pi_1(M, *) = G$ as a group of deck transformations acting on $(\widetilde{M}, *)$, and define $f : G \to \widetilde{M}$ by $f(g) = g*$. The map f can be extended to the Cayley graph by mapping each edge to a geodesic segment. Edges corresponding to the same generator of G map to segments of the same length, so f is Lipschitz.

To see f is a quasi-isometry, note that by compactness of M there is an $R > 0$ such that every point of the universal cover is within distance R of the orbit $G*$. Consider a geodesic segment $\gamma \subset \widetilde{M}$ of length $L = d(*, g*)$ joining $*$ to $g*$. Cut

γ into L segments of about unit length, and assign to each a point h_i* within distance R. Then $d(h_i*, h_{i+1}*) \leq 2R+1$). Now by discreteness of $G*$, there are only finitely many h such that $d(*, h*) \leq 2R+1$; letting $C = \max d(\mathrm{id}, g)$ over such g, we have $d(h_i, h_{i+1}) \leq C$, and thus $d(\mathrm{id}, g) \leq CL = Cd(*, g*)$. Thus f is a quasi-isometry. ■

Corollary 6.7 (Milnor, Švarc) *If M is a closed manifold with a metric of negative curvature, then $\pi_1(M)$ has exponential growth.*

Quasi-geodesics. A *quasi-geodesic* in a metric space X is a quasi-isometric map

$$\gamma : [a, b] \to X.$$

Often we will take $[a, b] = \mathbb{R}$.

Examples.

1. If $\gamma : [a, b] \to X$ is a geodesic and $f : X \to Y$ is a quasi-isometry, then $f \circ \gamma$ is a quasi-geodesic.

2. Let $\gamma : \mathbb{R} \to C(G)$ be a geodesic in the Cayley graph of a group $G = \pi_1(M)$, and let
$$f : (C(G), \mathrm{id}) \to (\widetilde{M}, *)$$
be defined by $f(g) = g*$ and by sending edges to geodesic segments. Then $f \circ \gamma$ is a quasi-geodesic.

3. Efficient taxis take quasi-geodesics along the grid of 'streets' connecting $\mathbb{Z}^2 \subset \mathbb{R}^2$. These *Manhattan geodesics* are far from unique; e.g. there are $\binom{2n}{n}$ geodesics from $(0, 0)$ to (n, n). This example comes from the universal cover of a 2-torus.

4. If $\gamma : \mathbb{R} \to \mathbb{H}^n$ is a C^2 curve parameterized by arclength, with geodesic curvature $k(s) < k_0 < 1$ at every point, then $\gamma(s)$ is a quasigeodesic. Indeed, the normal planes $P(s)$ through $\gamma(s)$ advance at a uniform pace C, so
$$d(\gamma(s), \gamma(t)) \geq d(P(s), P(t)) \geq C(k_0)|s - t|.$$
The borderline case is a horocycle, which is *not* a quasi-geodesic (in fact $d(\gamma(s), \gamma(0))$ grows like $\log s$). On the other hand, a curve at constant distance D from a geodesic in $\mathbb{H}^2$ has curvature $k_0(D) < 1$ (and is obviously a quasi-geodesic).

5. The loxodromic spiral $\gamma : [0, \infty) \to \mathbb{C}$ given by $\gamma(s) = s^{1+i}$ is a quasi-geodesic ray with no definite direction; that is, $\arg \gamma(s) = \log(s)$ moves around the circle an infinite number of times as $s \to \infty$.

   This γ is not within a bounded distance of any Euclidean geodesic.

Lemma 6.8 *For any closed convex set $K \subset \mathbb{H}^n$, nearest-point projection $\pi : \mathbb{H}^n \to K$ contracts by a factor of at least $\cosh(r)$ at distance r from K.*

Proof. First calculate $\|d\pi(z)\|$ in the case where K is a geodesic in $\mathbb{H}^2$. We can normalize coordinates so K is the imaginary axis and z lies on the unit circle; then $\pi(i) = 1$. As a hyperbolic geodesic, the arclength parameterization of the unit circle is given by $(\tanh s, \operatorname{sech} s)$, so $\operatorname{Im} z = \operatorname{sech} r$ where $r = d(z, K)$. Now the geodesics normal to K are Euclidean circles, so $\|d\pi\| = 1$ in the Euclidean metric. Thus the hyperbolic contraction is by a factor of $\operatorname{Im}(\pi(z))/\operatorname{Im}(z) = \cosh r$.

The general case reduces to this one by considering a supporting hyperplane to K. ∎

Theorem 6.9 *Let $\gamma : \mathbb{R} \to \mathbb{H}^n$ be a quasi-geodesic. Then γ is within a bounded distance of a unique hyperbolic geodesic.*

Proof. For convenience, assume γ is continuous. For $T \gg 0$, let δ_T be the complete geodesic passing through $\gamma(-T)$ and $\gamma(T)$. Consider the cylinder $B(\delta_T, r)$ of radius r about δ_T, and suppose (a, b) is a maximal interval for which $\gamma(a, b)$ is outside the cylinder. Then $\gamma(a), \gamma(b)$ lie on the boundary of the cylinder, and the nearest point projection of $\gamma(a, b)$ to δ_T is contracting by a factor of $\cosh(r)$. Thus, if $|a - b| > R$, we have

$$\frac{|a-b|}{K} \leq d(\gamma(a), \gamma(b)) \leq 2r + \frac{d(\gamma(a), \gamma(b))}{\cosh(r)} \leq 2r + \frac{K|a-b|}{\cosh(r)}.$$

If we take r large enough that $\cosh(r) \gg K^2$, then the inequality above implies $|a - b|$ is not too large, and hence $\gamma[a, b]$ stays close to δ_T.

In other words, there is an absolute constant D such that $\gamma(-T, T) \subset B(\delta_T, D)$. Since the space of geodesics within distance D of $\gamma(0)$ is compact, we can pass to a convergent subsequence and obtain a geodesic δ with $\gamma \subset B(\delta, D)$.

In hyperbolic space, distinct geodesics diverge, so δ is unique. ∎

6.4 Quasiconformal maps

We now revisit the notion of a quasiconformal map from a geometric point of view.

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a homeomorphism (often required to preserve orientation). For each sphere $S = S(x, r)$, let

$$K(S) = \frac{\max\{d(f(y), f(x)) \ : \ y \in S\}}{\min\{d(f(y), f(x)) \ : \ y \in S\}}.$$

We say f is *geometrically quasiconformal* if $\sup K(S)$ is finite. A homeomorphism $f : S^n \to S^n$ is geometrically quasiconformal if $\sup K(S)$ is bounded in the spherical metric.

Theorem 6.10 *Let f be a quasiconformal homeomorphism of $\mathbb{R}^n$ or S^n with $n \geq 2$. Then:*

- *f is analytically quasiconformal (f has derivatives in L^2 satisfying $\|Df\|^n \leq K|\det Df|$ almost everywhere);*
- *f is absolutely continuous ($f(E)$ has measure zero iff E has measure zero);*
- *f is differentiable almost everywhere; and*
- *if Df is conformal a.e., then f is a Möbius transformation.*

For proofs, see [LV]. To convey the spirit of the passage between the geometric and analytic definitions, we will prove:

Theorem 6.11 *A quasiconformal map $f : \mathbb{C} \to \mathbb{C}$ is absolutely continuous on lines (ACL).*

This means that for any line $L \subset \mathbb{C}$, the real and imaginary parts of f are absolutely continuous functions on $L + t$ for almost every $t \in \mathbb{C}$.

Proof. By making a linear change of coordinates in the domain of f, it suffices to show that $\operatorname{Re} f(x + iy)$ is an absolutely continuous function of $x \in [0, 1]$ for almost every y.

Consider the function $A(y) = \operatorname{area} f([0, 1] \times [0, y])$. Since $A(y)$ is monotone increasing, it has a finite derivative a.e. Choose y such that $A'(y)$ exists; we will show $F(x) = \operatorname{Re} f(x + iy)$ is absolutely continuous for $x \in [0, 1]$.

Consider a collection of disjoint intervals $I_1, \ldots, I_n$ in $[0, 1]$, with $\sum |I_i| < \epsilon$. We must show that $\sum |F(I_i)| < \delta(\epsilon) \to 0$ as $\epsilon \to 0$.

By subdividing the intervals, we can assume $|I_i| = h \ll \epsilon$ for all i, so $nh = \epsilon$. Let S_i be the $h \times h$ square resting on I_i. Since f is quasiconformal, we have $\operatorname{area}(f(S_i)) \asymp |J_i|^2$. Thus we have:

$$\begin{aligned} \left(\sum |J_i|\right)^2 &\leq n \sum |J_i|^2 \asymp n \sum \operatorname{area}(f(S_i)) \\ &\leq n \operatorname{area}(f([0, 1] \times [y, y + h])) \approx nhA'(y). \end{aligned}$$

Therefore $\sum |J_i| = O(\sqrt{nh}) = O(\epsilon^{1/2})$. ■

Measuring quasiconformality. The natural measure of distortion for a homeomorphism $f : \mathbb{C} \to \mathbb{C}$ is the *dilatation* $K(f)$.

If f is $\mathbb{R}$-linear, it maps circles to ellipses with major and minor axes M and m, and we define $K(f) = M/m$. For a general quasiconformal map, we define $K(f)$ to be the least constant such that $K(Df) \leq K(f)$ almost everywhere.

Perhaps surprisingly, $K(f)$ is not the same as $\sup K(S)$ over all spheres. For example, if f is given in polar coordinates by $f(r, \theta) = (r^\alpha, \theta)$, with $\alpha > 1$, then $K(f) = \alpha$ even though $K(S) = 2^\alpha$ for the sphere $S = \{z \,:\, |z - 1| = 1\}$.

The proper geometric definition of the dilatation is that $K(f)$ is the least constant such that $\limsup_{r \to 0} K(S(x, r)) \leq K(f)$ almost everywhere.

The 1-dimensional case. The geometric definition of quasiconformality makes sense even in dimension one, and yields the useful family of *k-quasisymmetric* homeomorphisms $f : \mathbb{R} \to \mathbb{R}$, satisfying $\sup K(S) \leq k$ for *every* sphere S. Of course a quasisymmetric map f is differentiable a.e. (since it is a monotone function) , but f need *not* be absolutely continuous.

6.5 Quasi-isometries become quasiconformal at infinity

Let $f : \mathbb{H}^{n+1} \to \mathbb{H}^{n+1}$ be a quasi-isometry.

The *extension* $F : S^n_\infty \to S^n_\infty$ of f is defined as follows. Given $x \in S^n_\infty$, take a geodesic ray γ landing at x, and let δ be a geodesic ray that shadows the quasi-geodesic $f \circ \gamma$; then $F(x)$ is the endpoint of δ.

It is easy to see:

If f is close to the identity, then F is the identity.

Thus if f is an isomorphism with quasi-inverse g, the extensions of f and g satisfy $F \circ G = G \circ F = \mathrm{id}$, so F is bijective. In fact we have:

Theorem 6.12 *The extension $F : S^n_\infty \to S^n_\infty$ of a quasi-isometry $f : \mathbb{H}^{n+1} \to \mathbb{H}^{n+1}$ is a homeomorphism.*

Proof. It suffices to show F is injective and continuous. Let us work in the Poincaré unit ball model, with $\mathbb{H}^{n+1} = \mathbb{B}^{n+1} \subset \mathbb{R}^{n+1}$. Composing with an isometry, we can assume $f(0) = 0$.

Consider $x, y \in S^n_\infty$ with $|x - y| = \epsilon > 0$. Let γ be the geodesic joining x and y, and let δ be the geodesic shadowing $f(\gamma)$. Then the endpoints are δ are $F(x)$ and $F(y)$. Since the endpoints are distinct, F is injective.

Moreover, $d(0, \gamma) = |\log \epsilon| + O(1)$ in the hyperbolic metric, so we have $d(0, \delta) \geq |\log \epsilon|/K + O(1)$. Therefore $|F(x) - F(y)| = O(\epsilon^{1/K})$, showing F is even Hölder continuous. ■

Corollary 6.13 *Any quasi-isometry $f : \mathbb{H}^{n+1} \to \mathbb{H}^{n+1}$ is an isomorphism.*

Proof. Let F be the extension of f. Given $z \in \mathbb{H}^{n+1}$, choose a geodesic δ through z with endpoints $F(x), F(y)$, and let γ be the geodesic from x to y. Then $f(\gamma)$ comes within a bounded distance of z, so f is essentially surjective and therefore it admits a quasi-inverse. ■

Theorem 6.14 *The extension of a quasi-isometry f on $\mathbb{H}^{n+1}$ is a quasiconformal homeomorphism F on S^n_∞.*

Proof. To see F is quasi-conformal, we will show $F(S(a, r))$ has a bounded ratio of inradius to outradius for any small sphere $S(a, r) \subset S^n_\infty$. For convenience we normalize so $a = 0$ and F fixes 0 and ∞. Let $|b| = |c| = r$ maximize the ratio

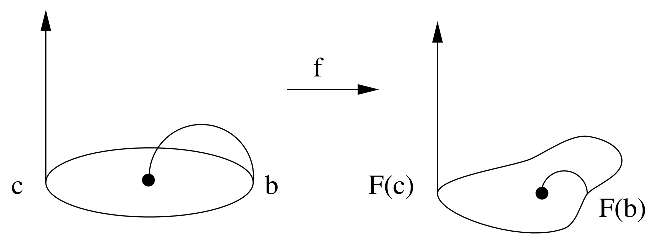

Figure 16. Geodesics, inradius and outradius.

$K = |F(c)/F(b)|$. Then the geodesics $[0, b]$ and $[c, \infty]$ are at distance $O(1)$, while the geodesics $[0, F(b)]$ and $[F(c), \infty]$ are at distance about $\log K$ (see Figure 16). Since f is a quasi-isometry, $f([0, b])$ and $f([\infty, c])$ lie at a bounded distance from $[0, F(b)]$ and $[\infty, F(c)]$, and from each other. Thus $\log K = O(1)$ and F is quasiconformal. ∎

6.6 Mostow rigidity

In this section we prove that a compact hyperbolic manifold of dimension 3 or more can be reconstructed from its fundamental group.

Lemma 6.15 *Let $f : M \to N$ be a homotopy equivalence between compact Riemannian manifolds. Then the lift*

$$\widetilde{f} : \widetilde{M} \to \widetilde{N}$$

of f gives a quasi-isometric isomorphism between the universal covers of M and N.

Proof. Let $g : N \to M$ be a homotopy inverse to f, and let $\widetilde{g} : \widetilde{N} \to \widetilde{M}$ be a lift of g compatible with $\widetilde{f}$. Then the homotopy h_t of $g \circ f$ to the identity lifts to a homotopy $\widetilde{h}_t$ of $\widetilde{g} \circ \widetilde{f}$ to the identity.

We can assume that f and g are smooth, so by compactness of M and N their lifts are K-Lipschitz for some K. Similarly, since $\widetilde{h}_t$ is a lift of a homotopy on M, there is a $D > 0$ such that

$$d(x, \widetilde{g} \circ \widetilde{f}(x)) \quad \leq \quad \operatorname{diam} \widetilde{h}_{0,1}(x) \quad \leq \quad D \tag{6.1}$$

for all $x \in \widetilde{M}$. Therefore $\widetilde{g} \circ \widetilde{f}$ is close to the identity. Similarly, $\widetilde{f} \circ \widetilde{g}$ is close to the identity. It follows that $\widetilde{f}$ is a quasi-isometry. Indeed, we have the upper bound

$$d(\widetilde{f}(x), \widetilde{f}(y)) \leq K d(x, y)$$

since $\widetilde{f}$ is Lipschitz, and the lower bound

$$d(x, y) \leq d(\widetilde{g} \circ \widetilde{f}(x), \widetilde{g} \circ \widetilde{f}(y)) - D \leq K d(\widetilde{f}(x), \widetilde{f}(y)) - D$$

by (6.1). Similarly, $\widetilde{g}$ is a quasi-inverse for $\widetilde{f}$. ∎

Remark. One can also argue that $\widetilde{M}$ and $\widetilde{N}$ are quasi-isometry to $\pi_1(M)$ and $\pi_1(N)$ in such a way that $\widetilde{f}$ is close to the quasi-isometric isomorphism $f_* : \pi_1(M) \to \pi_1(N)$.

Theorem 6.16 (Mostow) *Let M^n and N^n be compact hyperbolic n-manifolds, $n \geq 3$, and let*

$$\iota : \pi_1(M^n) \to \pi_1(N^n)$$

be an isomorphism. Then there is an isometry $I : M^n \to N^n$ such that $\iota = I_*$.

Proof. The manifolds M and N are $K(\pi, 1)$'s, so the isomorphism ι between their fundamental groups can be realized by a homotopy equivalence $f : M \to N$. By the preceding lemma, the lift $\widetilde{f} : \mathbb{H}^n \to \mathbb{H}^n$ is a quasi-isometry, so its extension is a quasiconformal map $F : S^{n-1}_\infty \to S^{n-1}_\infty$ conjugating the action of $\pi_1(M)$ to that of $\pi_1(N)$. The map F is differentiable almost everywhere, by fundamental results on quasiconformal mappings.

If DF is conformal almost everywhere then, since $n > 2$, F is a Möbius transformation. (This step fails when $n = 2$). Then F extends to an isometry $\widetilde{I} : \mathbb{H}^n \to \mathbb{H}^n$ which descends to the desired isometry $I : M \to N$.

Otherwise, DF fails to be conformal on a set of positive measure in S^{n-1}_∞. By ergodicity, DF is nonconformal almost everywhere.

Now for concreteness suppose $n = 3$. Then the conformal distortion of $DF(x)$ defines an ellipse in the tangent space $\mathrm{T}_x S^2_\infty$ for almost every x. Let $L_x \subset T_x S^2_\infty$ be the line through the major axis of this ellipse.

Define $\theta : S^2_\infty \times S^2_\infty \to S^1$ as follows: given x, y on the sphere, use parallel transport along the geodesic joining x to y to identify $\mathrm{T}_x S^2_\infty$ with $\mathrm{T}_y S^2_\infty$, and let θ be the angle between the lines L_x and L_y.

Then θ is invariant under the action of $\pi_1(M)$, so by ergodicity of the geodesic flow it is a constant a.e. This means that if we choose coordinates on S^2_∞ so $x = \infty$, then the lines L_y have constant slope for $y \in \mathbb{R}^2_\infty$. But almost any point can play the role of x, while it is clearly impossible to arrange the linefield L_y to have constant slope in more than one affine chart on the sphere.

The proof for $n > 2$ is similar. ∎

Note. Mostow rigidity also holds for finite volume manifolds.

6.7 Rigidity in dimension two

Here is a version of Mostow rigidity that works for hyperbolic surfaces.

Theorem 6.17 *Let $f : S^1_\infty \to S^1_\infty$ be an orientation-preserving homeomorphism conjugating Γ to Γ', where $X = \mathbb{H}/\Gamma$ and $X' = \mathbb{H}/\Gamma'$ are finite-volume hyperbolic surfaces. Then f is either singular or absolutely continuous. In the absolutely continuous case, f must be a Möbius transformation.*

Proof. For convenience we treat the case where X and X' are compact. By hypothesis, there is an isomorphism $\iota : \Gamma \to \Gamma'$ induced by f. We first observe that if $\Gamma = \Gamma'$ and ι is the identity, then f is the identity. This is because f must fix the attracting fixed-point of every $g \in \Gamma$, and such fixed points are dense on S^1_∞.

By ergodicity of the action of Γ on the circle, f is either absolutely continuous or singular. We will show that in the former case, Γ is conjugate to Γ' inside $G = \operatorname{Isom} \mathbb{H}$; in other words, $X \cong X'$. To this end, we identify S^1_∞ with $\widehat{Q}$ and choose coordinates so that $f(\infty) = \infty$. Then $f : \mathbb{R} \to \mathbb{R}$ is a homeomorhism; in particular, f' exists a.e.

Our hypothesis of absolutely continuity implies that $f = \int f'$; in particular, $f'(x) \neq 0$ for some x. We can assume $x = 0$ and $f'(0) = 1$. Let $A_n(x) = nx$. Then $f_n = A_n f A_n^{-1}$ converges to the identity map $f_\infty(x) = x$ uniformly on compact sets. Moreover f_n intertwines the actions of certain conjugates Γ_n and Γ'_n of our original Fuchsian groups. Since G/Γ and G/Γ' are compact, we can pass to a subsequence so these conjugates converge, say to Γ_∞ and Γ'_∞. Since $f_\infty(x) = x$, we have $\Gamma_\infty = \Gamma'_\infty$ and hence $X \cong X'$.

The same proof shows the isomorphism $\iota : \Gamma \to \Gamma'$ induced by the original map f is given by conjugation by an element $g \in G$. Thus our initial remarks show $f = g$. ∎

6.8 Geometric limits

For another viewpoint on Mostow rigidity, we introduce in this section the *geometric topology* on baseframed hyperbolic manifolds.

For any topological space X, let $\mathrm{Cl}(X)$ be the set of all closed subsets of X. When X is a compact Hausdorff space, we introduce a topology on $\mathrm{Cl}(X)$ by defining $F_\alpha \to F$ iff

- for any open set $U \supset F$, we have $F_\alpha \subset U$ for all $\alpha \gg 0$; and
- for any open set with $U \cap F \neq \emptyset$, we have $F_\alpha \cap U \neq \emptyset$ for all $\alpha \gg 0$.

Theorem 6.18 *If X is a compact Hausdorff space, then so is* $\mathrm{Cl}(X)$.

Next suppose X is only locally compact and Hausdorff. Then the one-point compactification $X' = X \cup \{\infty\}$ is compact, and there is a natural map $\mathrm{Cl}(X) \to \mathrm{Cl}(X')$ by sending F to $F \cup \{\infty\}$. Under this inclusion, $\mathrm{Cl}(X)$ is closed, so it becomes a compact Hausdorff space with the induced topology.

Example. In $\mathrm{Cl}(\mathbb{R})$, the intervals $I_n = [n, \infty)$ converge to the empty set as $n \to \infty$.

Now suppose G is a Lie group. The set of all closed subgroups of G forms a compact subset of $\mathrm{Cl}(G)$. If H, H_n are subgroups, we say $H_n \to H$ *geometrically* if we have convergence in the Hausdorff topology.

Finally let $G = \mathrm{Isom}(\mathbb{H}^n)$. Recall that every baseframed hyperbolic manifold (M, ω) determines a torsion-free discrete group $\Gamma \subset G$, and vice-versa. We say

(M_n, ω_n) converges *geometrically* to (M, ω) if the corresponding Kleinian groups satisfy $\Gamma_n \to \Gamma$ in $\mathrm{Cl}(G)$.

Theorem 6.19 *The set $\mathcal{H}^n_r$ of baseframed hyperbolic manifold (M, ω) with injectivity radius $\geq r$ at the basepoint is compact in the geometric topology.*

If we fix a hyperbolic manifold M, then we have a natural map $FM \to \mathrm{Cl}(G)$ sending each ω in the frame bundle FM to the subgroup $\Gamma(M, \omega)$. This map is continuous. In particular, if M is compact, then the set of baseframed manifolds (M, ω) that can be formed from M is compact in the geometric topology.

We can now offer a proof of the last step of Mostow rigidity that does not make reference to ergodic theory of the geodesic flow.

Theorem 6.20 *Let $M = \mathbb{H}^3/\Gamma$ be a compact hyperbolic 3-manifold, and let $\mu \in M(\widehat{\mathbb{C}})$ be an L^∞ Γ-invariant Beltrami differential. Then $\mu = 0$.*

Proof. Suppose $\mu \neq 0$. Then, by a variant of the Lebesgue density theorem, there exists a $p \in \mathbb{C}$ such that $\mu(p) \neq 0$ and μ is almost continuous at p. That is, if we write $\mu = \mu(z)\, d\overline{z}/dz$, then for each $\epsilon > 0$ we have

$$\lim_{r \to 0} \frac{m\{x \in B(p,r) \ : \ |\mu(x) - \mu(p)| > \epsilon\}}{m(B(p,r))} = 0.$$

By a change of coordinates, we can assume that $p = 0$. Let $\mu(0) = a$. Then for $g_t(z) = tz$, we have the weak* limit

$$g_t^* \mu = \mu(tz) \frac{d\overline{z}}{dz} \to \nu = a \frac{d\overline{z}}{dz}$$

as $t \to 0$. Concretely, this means

$$\int_{\mathbb{C}} (g_t^* \mu) \phi = \int \mu(tz) \phi(z)\, |dz|^2 \to \int a \cdot \phi(z)\, |dz|^2$$

for every L^1 measurable quadratic differential $\phi = \phi(z)\, dz^2$.

Since μ is Γ-invariant, $g_t^*(\mu)$ is invariant under

$$\Gamma_t = g_t^*(\Gamma) = g_t^{-1} \Gamma g_t.$$

These conjugates Γ_t correspond to the baseframed manifolds (M, ω_t) as ω_t moves along a geodesic. Since M is compact, we can pass to a subsequence such that $\Gamma_t \to \Gamma'$, where Γ' is a conjugate of Γ'. Indeed, Γ_t only depends on the value of $[g_t]$ in the compact space $\Gamma \backslash G$.

Then ν is invariant under Γ'. This implies Γ' fixes the point $z = \infty$, since ∞ is the only point at which ν is discontinuous. Therefore Γ' is an elementary group — which is impossible, since in fact every orbit of Γ' on $\widehat{\mathbb{C}}$ is dense. ∎

Pushing this argument further, one can show:

Theorem 6.21 *Let $M^3 = \mathbb{H}^3/\Gamma$ be a hyperbolic manifold whose injectivity radius is bounded above. Then M^3 is quasiconformally rigid: the only measurable Γ-invariant Beltrami differential $\mu \in M(\widehat{\mathbb{C}})$ is $\mu = 0$.*

See [Mc5, Thm. 2.9].

6.9 Promotion

The basic mechanism of Mostow rigidity is *promotion*: we can use the expanding dynamics of a cocompact group to promote a point of measurable continuity to a point of topological continuity.

Here are two simpler results with the same promotion principle at work.

Theorem 6.22 *Let $M = \mathbb{H}^n/\Gamma$ be a compact hyperbolic manifold. Then the action of Γ on S^{n-1}_∞ is ergodic.*

Proof 1. Let $A \subset S^{n-1}_\infty$ be a Γ-invariant set of positive measure. Then A has a point of Lebesgue density $p \in \mathbb{R}^{n-1}_\infty$; that is,

$$\lim_{r\to 0} \frac{m(B(p,r)\cap A)}{m(B(p,r))} = 1.$$

We may assume $p = 0$. Let $g_n(x) = x/n$; then $g_n \in G = \mathrm{Isom}(\mathbb{H}^n)$. Since 0 is a point of density, we have $g_n^*(\chi_A) \to 1$ in the weak* topology on $L^\infty(S^{n-1}_\infty)$. Since G/Γ is compact, we can write $g_n = \gamma_n h_n$ with $\gamma_n \in \Gamma$ and h_n in a compact subset of G. Then passing to a subsequence, we have $h_n \to h$, and therefore

$$g_n^*(\chi_A) = h_n^*(\gamma_n^* \chi_A) = h_n^*(\chi_A) \to h^*(\chi_A).$$

Therefore $h^*(\chi_A) = 1$, which shows $\chi_A = 1$ and A has full measure. Thus Γ acts ergodically. ∎

Proof 2 (Ahlfors). Let $A \subset S^{n-1}_\infty$ be a Γ-invariant set of positive measure. Then the harmonic extension of χ_A to $\mathbb{H}^n$ descends to a harmonic function $u : M \to \mathbb{R}$. By the maximum principle, u is constant, and thus $A = S^{n-1}_\infty$. ∎

Theorem 6.23 *Let $f : X \to Y$ be a homotopy equivalence between a pair of compact hyperbolic surfaces, and let*

$$F : S^1_\infty \to S^1_\infty$$

be the boundary values of $\widetilde{f}$. Then either $F' = 0$ almost everywhere, or f is homotopic to an isometry.

Proof. Write $X = \mathbb{H}/\Gamma$ and $Y = \mathbb{H}/\Gamma'$. Suppose $F'(p) \neq 0$. By a change of coordinates, we can assume $p = F(p) = 0 \in \mathbb{R}^1_\infty$. Let $a = F'(0)$ and $g_n(x) = x/n$. Then we have

$$F_n(x) = g_n^{-1} \circ F \circ g_n(x) = nF(x/n) \to F_\infty(x) = ax$$

uniformly on compact sets. That is, the blowups of F yield in the limit the boundary values F_∞ of an isometry of $\mathbb{H}^2$.

Now F_n conjugates $\Gamma_n = g_n^{-1}\Gamma g_n$ to $\Gamma'_n = g_n^{-1}\Gamma' g_n$. By compactness of X and Y, we can pass to a subsequence such that Γ_n and Γ'_n converge geometrically to groups Γ_∞ and Γ'_∞ that are conjugates of Γ and Γ'. Then F_∞ conjugates Γ_∞ to Γ'_∞, so X and Y are isometric. With more care one can check that the isometry is in the homotopy class of f. ∎

Notes. Bowen used this fact to prove that the limit set of a quasifuchsian group is either a round circle or a Jordan curve of Hausdorff dimension $d > 1$ [Bo].

6.10 Ahlfors' finiteness theorem

We now turn to results which control much more general hyperbolic 3-manifolds, with the constraint that $M^3 = \mathbb{H}^3/\Gamma$ is compact replaced by the assumption that $\pi_1(M^3)$ is finitely-generated.

Recall that the action of a Kleinian group on its domain of discontinuity Ω is properly discontinuous.

Theorem 6.24 (Ahlfors' finiteness theorem) *Let $\Gamma \subset \mathrm{Isom}^+(\mathbb{H}^3)$ be a finitely-generated torsion-free Kleinian group. Then the quotient complex 1-manifold $X = \Omega/\Gamma$ is isomorphic to the complement of a finite set in a finite union of compact Riemann surfaces.*

Corollary 6.25 *The components of Ω fall into finitely many orbits under the action of Γ.*

Corollary 6.26 *The domain of discontinuity has no wandering domain; in fact, the stabilizer Γ^U of any component U is infinite.*

If Γ is elementary, then Ω/Γ is isomorphic to $\widehat{\mathbb{C}}$, $\mathbb{C}$, $\mathbb{C}^*$ or a torus $\mathbb{C}/L$, so the conclusion of the finiteness theorem holds.

Let us assume from now on that $\Gamma \subset \mathrm{Aut}(\widehat{\mathbb{C}})$ is an N-generator nonelementary Kleinian group. This condition implies that the centralizer of Γ in $\mathrm{Aut}(\widehat{\mathbb{C}})$ is finite, and some of the results below will hold under that slightly weaker hypothesis.

We will start by proving the slightly weaker statement established in Ahlfors' original paper [Ah1]:

Theorem 6.27 *If Γ has N generators, then* $\dim \mathrm{Teich}(\Omega/\Gamma) \le 3N - 3$.

Remarks. The bound is sharp: an N-generator Schottky group has quotient surface X of genus $g = N$, satisfying $\dim \mathrm{Teich}(X) = 3g - 3$.

While it is true that any *connected* Riemann surface of infinite hyperbolic area has an infinite-dimensional Teichmüller space, the result above does not exclude the possibility that X contains infinitely many components isomorphic to the triply-punctured sphere $Y = \widehat{\mathbb{C}} - \{0, 1, \infty\}$ (since $\dim \mathrm{Teich}(Y) = 0$).

The proof has two main ingredients — the cohomology of deformations, and an estimate for quasiconformal vector fields.

Cocycles and group cohomology. Let G be a group, and A a G-module. A *1-cocycle* is a map $\xi : G \to A$ such that

$$\xi(gh) = \xi(g) + g \cdot \xi(h).$$

Such a cocycle is also called a *crossed homomorphism*, since it coincides with a homomorphism when G acts trivially on A. A *1-coboundary* is a cocycle of the form

$$\xi(g) = g \cdot a - a$$

for some $a \in A$. The quotient space, (cocycles)/(coboundaries), gives the group cohomology $H^1(G, A)$.

One can view the group $H^1(G, A)$ as classifying affine actions of G on A of the form $g(a) = g \cdot a + \xi(a)$, up to conjugacy by translation. This explains the cocycle rule: we need to have

$$(gh)(a) = gh \cdot a + \xi(gh) = g \cdot (h \cdot a + \xi(h)) + \xi(g) = gh \cdot a + g \cdot \xi(h) + \xi(g).$$

The action $g(a) = g \cdot a$ is considered trivial. It is conjugate to $g(a+b) - b = g \cdot a + (g \cdot b - b)$, so the coboundary $\xi(g) = g \cdot b - b$ is also considered trivial.

Holomorphic vector fields and deformations. Let $G = \mathrm{PSL}_2(\mathbb{C})$, and let $A = sl_2(\mathbb{C})$ be the Lie algebra of holomorphic vector fields on the sphere. We have $\dim sl_2(\mathbb{C}) = \dim G = 3$. The adjoint action makes A into a G-module. We can regard an element $X \in sl_2(\mathbb{C})$ as a matrix or as a vector field, in terms of which the action can be written

$$g \cdot X = gXg^{-1} = g_*(X).$$

Lemma 6.28 *If Γ is a nonelementary N-generator Kleinian group, then*

$$\dim H^1(\Gamma, sl_2(\mathbb{C})) \leq 3N - 3.$$

Proof. The space of cocycles is at most $3N$ dimensional since a cocycle is determined by its values on generators. Since the centralizer of Γ is trivial, the space of coboundaries is isomorphic to $sl_2(\mathbb{C})$, hence 3-dimensional. The difference gives the bound $3N - 3$. ■

The group $H^1(\Gamma, sl_2(\mathbb{C}))$ can be interpreted as the tangent space to the variety of homomorphisms from Γ into G, modulo conjugacy, *at* the inclusion. That is, if $\rho_t : \Gamma \to G$ is a 1-parameter family of representations, with $\rho_0 = \mathrm{id}$, and we set

$$\xi(g) = \frac{d}{dt}\rho_t(g)g^{-1},$$

then $\xi(g)$ gives a cocycle with values in $sl_2(\mathbb{C})$. This cocycle is a coboundary iff to first order we have $\rho_t(g) = \gamma_t g \gamma_t^{-1}$, i.e. if the deformation is by conjugacy.

From Beltrami differentials to cocycles. Now let $M(X)$ be the space of L^∞ Beltrami differentials on X, or equivalently the space of Γ-invariant μ on $\widehat{\mathbb{C}}$ supported on Ω. We now define a natural map

$$\delta : M(X) \to H^1(\Gamma, sl_2(\mathbb{C})).$$

Namely, we solve the equation $\overline{\partial} v = \mu$ and set $\delta\mu = \xi$ where

$$\xi(g) = g_*(v) - v.$$

Since μ is Γ-invariant, we have $\overline{\partial}\xi(g) = 0$, so $\xi(g)$ is indeed a holomorphic vector field and therefore an element of $sl_2(\mathbb{C})$. Note that the solution to $\overline{\partial}v = \mu$ is only well-defined modulo $sl_2(\mathbb{C})$, but changes v by an element of $sl_2(\mathbb{C})$ only changes ξ by a coboundary.

The idea of the construction is that $\mu \in M(X)$ should correspond to Riemann surface X_μ complex structure deformed infinitesimally in the direction μ, plus a quasiconformal map $f : X \to X_\mu$ with complex dilatation an infinitesimal multiple of μ. Since f is infinitely close to the identity, it should be represented by a vector field v. But the vector field does not quite live on X, because the target of f is X_μ rather than X. On the universal cover, v is really well-defined, and its failure to live on X is measured by the cocycle $\xi(g)$.

Quadratic differentials. Let $Q(X)$ be the Banach space of holomorphic quadratic differentials on X with finite L^1-norm: $\|\phi\| = \int_X |\phi|$. There is a natural pairing $M(X) \times Q(X) \to \mathbb{C}$ given by

$$\langle \phi, \mu \rangle = \int \phi\mu.$$

Since $\langle \phi, \overline{\phi}/|\phi| \rangle = \int |\phi|$, the pairing descends to a perfect pairing on $M(X)/Q(X)^\perp \times Q(X)$. The quotient pairing is exactly that between the tangent space $\mathrm{T}_X \operatorname{Teich}(X)$ and the cotangent space $Q(X) = \mathrm{T}^*_X \operatorname{Teich}(X)$.

Let us say $\mu \in M(X)$ is *trivial* if $\mu \in Q(X)^\perp$; that is, if $\int \mu\phi = 0$ for all $\phi \in Q(X)$, or equivalently if $[\mu]$ represents the zero tangent vector to Teichmüller space.

Let $V(X)$ be the space of quasiconformal vector fields on X, and let $\|v\|_X = \sup \rho_X(z)|v(z)|$ denote the supremum of the hyperbolic length of v.

Lemma 6.29 *If $\delta\mu = 0$, then $\overline{\partial}v = \mu$ has a Γ-invariant solution vanishing on the limit set.*

Proof. If $\delta\mu = 0$, then we can modify any solution v by a holomorphic vector field to obtain $\xi(g) = 0$ for all g; then v is Γ-invariant. Now if $z \in \Lambda$ is a hyperbolic fixed-point of an element $g \in \Gamma$, then $g'(z) \neq 1$, so the condition $g_*(v) = v$ implies $v(z) = 0$. Such points are dense in the limit set, so $v|\Lambda = 0$. ∎

Lemma 6.30 *Let v be a quasiconformal vector field on $Z = \widehat{\mathbb{C}} - \{0, 1, \infty\}$, vanishing at 0, 1 and ∞. Then $\|v\|_Z$, the maximum speed of v in the hyperbolic metric on Z, is finite.*

Proof 1. The vector field v has an $|x \log x|$ modulus of continuity, which exactly balances the $1/|x \log x|$ singularity of the hyperbolic metric at the punctures $0, 1, \infty$.

Proof 2. The Teichmüller space of the 4-times punctured sphere is isometric to Z by the cross-ratio map. Thus the hyperbolic length of $v(z)$ is the same as the Teichmüller length of the deformation of $\widehat{\mathbb{C}} - \{0, 1, \infty, z\}$ defined by $\overline{\partial} v$, which is controlled by $\|\overline{\partial} v\|_\infty$.

Proof 3. Let $\mu = \overline{\partial} v$. By linearity we can assume $\|\mu\|_\infty = 1$. Let $\phi_t : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be the unique holomorphic motion fixing 0, 1 and ∞ and with $\mu(\phi_t) = t\mu$. Then $f(t) = \phi_t(z)$ gives a holomorphic map $f : \Delta \to Z$. By the Schwarz lemma, with $w = d/dz$, we have

$$\|v(z)\|_X = \|Df(w)\|_X \leq \|w\|_\Delta = 2.$$

■

Corollary 6.31 *If v is a quasiconformal vector field on a domain $\Omega \subset \widehat{\mathbb{C}}$, and $v|\partial\Omega = 0$, then $\|v\|_\Omega$ is finite.*

Proof. At any point $z \in \Omega$ we can find three points $\{a, b, c\}$ in $\partial\Omega$ such that the hyperbolic metric on Ω at z is comparable to the hyperbolic metric on $\widehat{\mathbb{C}} - \{a, b, c\}$. But the length of $v(z)$ in the hyperbolic metric on $\widehat{\mathbb{C}} - \{a, b, c\}$ is bounded. ■

Lemma 6.32 *If $v \in V(X)$ has $\|v\|_X < \infty$ then $\mu = \overline{\partial} v$ is trivial. In other words, if v has bounded hyperbolic speed, then $\int \mu\phi = 0$ for all $\phi \in Q(X)$.*

Proof. Let ρ be the hyperbolic metric on X, and fix $\phi \in Q(X)$. Let X_0 be a component of X, and consider a large ball $B(p, r) \subset X_0$. Then we have

$$\int_{X_0} |\phi| \quad = \quad \int_0^\infty dr \int_{\partial B(p,r)} |\phi|/\rho < \infty.$$

Integrating by parts and using the fact that $\overline{\partial}\phi = 0$, we find:

$$\int_{B(p,r)} \phi(\overline{\partial} v) = \int_{\partial B(p,r)} \phi v.$$

Now the sup-norm of ρv is bounded, while the L^1-norm of $|\phi|/\rho$ tends to zero as $r \to \infty$, so we can conclude that $\int_{X_0} \phi\overline{\partial} v = 0$. Since X_0 and ϕ were arbitrary, we find $\mu = \overline{\partial} v \in M(X)$ is trivial. ■

Proof of Theorem 6.27. By the preceding lemmas, we have

$$\operatorname{Ker} \delta \subset Q(X)^\perp.$$

Indeed, if $\delta\mu = 0$, then $\mu = \overline{\partial} v$ for a Γ-invariant v vanishing on $\partial\Omega = \Lambda$. Then v descends to a quasiconformal vector field on X with boundary hyperbolic speed, and hence $\mu = \overline{\partial} v$ is trivial, i.e. μ belongs to $Q(X)^\perp$.

Since the pairing between $Q(X)$ and $M(X)/Q(X)^{\perp}$ is perfect, we have:

$$\begin{aligned}\dim Q(X) &= \dim M(X)/Q(X)^{\perp} \\ &\leq \dim M(X)/\operatorname{Ker}\delta \\ &\leq \dim H^1(\Gamma, sl_2(\mathbb{C})) \leq 3N-3.\end{aligned}$$

■

Notes. Ahlfors' finiteness theorem *fails* for hyperbolic 4-manifolds; see [KP].

For another viewpoint on boundedness of $\|v\|_{\Omega}$ when $v|\partial\Omega = 0$, one can use the fact that a quasiconformal vector field has modulus of continuity $x|\log x|$, while the hyperbolic metric near a puncture is at worst like $|dz|/|z||\log z|$.

6.11 Bers' area theorem

Theorem 6.33 (Bers) *Let Γ be a nonelementary N-generator Kleinian group. Then the hyperbolic area of $X = \Omega(\Gamma)/\Gamma$ is finite; in fact we have*

$$\operatorname{area}(X) \leq 4\pi(N-1).$$

Remark. By Gauss-Bonnet, once we know X has finite hyperbolic area, we have $\operatorname{area}(X) = 2\pi|\chi(X)|$, and thus $|\chi(X)| \leq 2N-2$. Again this inequality is sharp for a handlebody of genus g: we have $N = g$ and $\chi(X) = 2-2g$.

Proof. We follow the same lines as Ahlfors' finiteness theorem, but replace the space $sl_2(\mathbb{C}) = H^0(\widehat{\mathbb{C}}, \mathcal{O}(2))$ of holomorphic vector fields with the space $V_d = H^0(\widehat{\mathbb{C}}, \mathcal{O}(2d))$ of holomorphic sections of the dth power of the tangent bundle to $\widehat{\mathbb{C}}$.

Then a typical element of V_d has the form $v = v(z)(\partial/\partial z)^d$, and we have $\dim V_d = 2d+1$. As before, the space of 1-coboundaries is isomorphic to V_d, so we have

$$\dim H^1(\Gamma, \mathcal{O}(2d)) \leq (2d+1)(N-1).$$

Then $\overline{\partial} v = \mu$ is a $(-d,1)$-form.

For the analytical part, let $M_d(X)$ denote the measurable $(-d,1)$-forms on X which are in L^{∞} with respect to the hyperbolic metric. Similarly let $Q_d(X)$ denote the L^1 holomorphic sections $\phi(z)\,dz^{d+1}$ of the $(d+1)$st power of the canonical bundle on X. Then there is a natural pairing between $M_d(X)$ and $Q_d(X)$ as before.

We define

$$\delta_d : M_d(X) \to H^1(\Gamma, V_d)$$

as before: by solving $\delta v = \mu$ and taking the resulting cocycle. Note that the lift of μ to Ω satisfies

$$\sup \rho^{d-1}(z)|\mu(z)| < \infty$$

where the hyperbolic metric $\rho(z)|dz|$ satisfies $\rho(z) \to \infty$ near $\partial\Omega$. Thus μ is locally in L^{∞} and the $\overline{\partial}$-equation is solvable. (For example, any $\mu \in M_d(\mathbb{H})$ satisfies $\mu(z) = O(y^{d-1})$.)

We have $\delta\mu = 0$ iff there is a Γ-invariant solution to $\overline{\partial} v = \mu$. In this case, $v = 0$ on Λ and one can again show v is bounded in the hyperbolic metric. Then integration by parts can be justified, showing $\mu \in Q_d(X)^\perp$. In conclusion, we find $\dim Q_d(X) \le (2d+1)(N-1)$.

Now let us examine the integrability condition on $\phi = \phi(z)\, dz^{d+1}$. On the punctured disk, the hyperbolic metric is given by $\rho = |dz|/|z \log z|$. So if we have

$$\int_U \rho^{1-d} |\phi| < \infty$$

for some neighborhood U of $z = 0$, then ϕ has at worst a pole of order d at $z = 0$.

Thus if $X = \overline{X} - P$, where $\overline{X}$ is a compact surface of genus g and $|P| = n$, then $Q_d(X) = H^0(\overline{X}, \mathcal{O}((d+1)K + dP))$, where K is a canonical divisor. We have $\deg(K + P) = 2g - 2 + n = |\chi(X)|$. By Riemann-Roch, the dimension of $Q_d(X)$ agrees with the degree of the divisor $d(K+P)$, up to an additive constant, so we have

$$\dim Q_d(X) = d|\chi(X)| + O(1).$$

Comparing with the dimension of the group cohomology, we find

$$d|\chi(X)| \le (2d+1)(N-1) + O(1).$$

Dividing by d and letting $d \to \infty$ we get the area theorem. ∎

Notes. Bers' theorem is proved in [Bers1]. The case of a Schottky group again shows the bound is sharp.

It is desirable to have a geometric interpretation of the cohomology $H^1(\Gamma, \mathcal{O}(d))$, generalizing the case $H^1(\Gamma, \mathcal{O}(2))$ which measures deformations of Γ inside $\mathrm{PSL}_2(\mathbb{C})$ and which appears in Ahlfors' finiteness theorem. Such an interpretation has been developed by Anderson. The idea is to use the embedding $\mathbb{P}^1 \to \mathbb{P}^n$ as a rational normal curve to extend the action of Γ to $\mathbb{P}^n$, and then investigate its deformations inside $PGL_{n+1}(\mathbb{C})$. See [And].

For example, $\mathrm{SL}_2(\mathbb{Z})$ acts on both $\mathbb{RP}^1$ and $\mathbb{RP}^2 = \mathbb{PR}^{2+1}$. The latter action comes from the Klein and Minkowski models for hyperbolic space. The action of $\mathbb{P}^1$ is rigid but the action on $\mathbb{P}^2$ admits deformations; see [Sc].

6.12 No invariant linefields

Let $\Gamma \subset \mathrm{Isom}(\mathbb{H}^3)$ be a finitely-generated Kleinian group. In this section we will show:

Theorem 6.34 (Sullivan) *The limit set $\Lambda(\Gamma)$ supports no measurable, Γ-invariant field of tangent lines.*

Equivalently, if $\mu \in M(\widehat{\mathbb{C}})$ is a Γ-invariant Beltrami differential, and if $\mu = 0$ outside Λ, then $\mu = 0$ a.e.

Let $Jg(x) = |\gamma'(x)|^2_\sigma$ denote the Jacobian determinant of g at x for the spherical metric. If $Jg(x) = 1$ on a set of positive measure, then g is an isometry in the spherical metric and g has a fixed-point in $\mathbb{H}^3$. Since Γ is torsion-free, we conclude that for almost every x, the map

$$g \mapsto Jg(x)$$

gives an injection of Γ into $(0, \infty)$.

We begin with a general analysis of the action of Γ on an invariant set $E \subset \widehat{\mathbb{C}}$ of positive measure. The measurable dynamical system (Γ, E) is *conservative* if for any $A \subset E$ of positive measure, we have $m(A \cap gA) > 0$ for infinitely many $g \in \Gamma$. At the other extreme, (Γ, E) is *dissipative* if it has a 'fundamental domain' $F \subset E$, meaning $E = \bigcup_\Gamma gF$ and $m(F \cap gF) = 0$ for all $g \neq e$.

Example: $f(z,t) = (e^{2\pi i\theta}x, t)$ gives a conservative action of $\mathbb{Z}$ on $E = S^1 \times [0,1]$. Note that f is far from ergodic.

Lemma 6.35 *Let $C = \{x \in E \,:\, \sum_\Gamma Jg(x) = \infty\}$ and let $D = E - C$. Then (Γ, C) is conservative and (Γ, D) is dissipative.*

Proof. Suppose $A \subset C$ has positive measure but $m(A \cap gA) = 0$ for all g outside a finite set $G_0 \subset G$. Then any point of C belongs to at most $n = |G_0|$ translates of A. It follows that

$$\int_A \sum_\Gamma Jg(x)\, dm \quad \leq \quad n \cdot m(C) < \infty,$$

contrary to the definition of C. Thus (Γ, C) is conservative.

To show (Γ, D) is dissipative, let $F \subset D$ be the set of y such that $Jg(y) < 1$ for all $g \in \Gamma$. Clearly gF is disjoint from F for $g \neq e$, because $Jg^{-1}(g(y)) = 1/Jg(y) > 1$ for all $g(y) \in gF$.

Now for almost any $x \in D$, the values $\{Jg(x) \,:\, g \in \Gamma\} \subset \mathbb{R}$ are discrete (by summability) and correspond bijectively to the elements of Γ. Thus there is a unique g maximizing $Jg(x)$, and therefore $y = g(x)$ belongs to F. Thus $\bigcup_\Gamma gF = D$ and (Γ, D) is dissipative. ■

Lemma 6.36 *Let $\Gamma \subset \mathrm{Isom}(\mathbb{H}^3)$ be a finitely-generated Kleinian group with limit set Λ. Then (Γ, Λ) is conservative.*

Proof. Suppose not. Then there is a set $D \subset \Lambda$ of positive measure such that (Γ, D) is dissipative. Let $F \subset D$ be a measurable fundamental domain for Γ, and let $M(F) \subset M(\widehat{\mathbb{C}})$ denote the space of L^∞ Beltrami differentials supported on F. Each $\mu \in F$ can be freely translated by Γ to give a Γ-invariant Beltrami differential μ' supported on $D \subset \Lambda$. Thus the space $M(\Lambda)^\Gamma$ of Γ-invariant differentials supported on the limit set is infinite-dimensional.

On the other hand, the natural map

$$\delta : M(\Lambda)^\Gamma \to H^1(\Gamma, sl_2(\mathbb{C}))$$

is injective. Indeed, if $\delta\mu = 0$, then the equation $\overline{\partial} v = \mu$ has a solution vanishing on Λ, which implies $\mu = 0$ on Λ. Since Γ is finitely-generated, the cohomology group on the right is finite-dimensional, so we obtain a contradiction. ∎

Lemma 6.37 *Let (Γ, E) be conservative and fix $\epsilon > 0$. Then there is a $g \in \Gamma - \{e\}$ such that $|Jg(x) - 1| < \epsilon$ on a set of positive measure.*

Proof. Suppose not. Then we can find $\epsilon > 0$ such that $|Jg(x) - 1| < \epsilon \implies g = e$ a.e. Therefore the image of $g \mapsto Jg(x)$ is discrete a.e.: since if $Jg(x)$ and $Jh(x)$ are close, then $J(hg^{-1})(y)$ is close to one at $y = g(x)$.

Let $E_+ = \{x : Jg(x) \geq 1 \ \forall g \in \Gamma\}$. Then $g(E_+) \cap E_+ = \emptyset$ for any $g \neq e$, since $J(g^{-1})(g(x)) = 1/Jg(x) < 1$ for $x \in E_+$. By conservativity, we have $m(E_+) = 0$. By similar reasoning, we conclude that for a.e. x there are unique elements $g_-, g_+ \in \Gamma$ (depending on x) such that

$$Jg_-(x) < 1 < Jg_+(x)$$

and no elements have Jacobians closer to 1 at x. Then the maps $F_\pm : E \to E$ defined by $F_\pm(x) = g_\pm(x)$ are inverses of one another, so both are injective. But F_+ is uniformly expanding, which is impossible since the total measure of E is finite. ∎

Inducing. The preceding result is obvious if $E = S^2_\infty$. Indeed, for any $g \in \operatorname{Isom} \mathbb{H}^3$, the Jacobian $Jg(x)$ is continuous and $\int_{S^2_\infty} Jg = 1$, so the Jacobian is close to one on a set of positive measure.

For applications it is useful to have the following stronger statement.

Lemma 6.38 *Let (Γ, E) be conservative and fix $\epsilon > 0$ and a set $F \subset E$ of positive measure. Then there is a $g \in \Gamma - \{e\}$ and a set of positive measure $A \subset F$ such that $g(A) \subset F$ and $|Jg(x) - 1| < \epsilon$ for all $x \in A$.*

The most conceptual formulation of this result is via *inducing*. First we must generalize our setting. A *partial automorphism* $g : E \dashrightarrow E$ is an invertible measurable map whose domain and range are in E. A collection of partial automorphisms G forms a *pseudo-group* if whenever the composition $g \circ h$ of $g, h \in G$ is defined on a set of positive measure A, we have $g \circ h = k|A$ for some $k \in G$. We say (G, E) is *conservative* if whenever $m(A) > 0$, we have $m(A \cap g(A)) > 0$ for infinitely many $g \in G$.

Now suppose Γ is a group acting on an invariant set E, and suppose $F \subset E$ has positive measure. Let

$$G = \Gamma|F = \{g|F \cap g^{-1}(F) \ : \ g \in \Gamma\}.$$

Then G is a pseudo-group acting on F. If (Γ, E) is conservative, so is (G, F).

It is then easy to see that Lemma 6.37 applies just as well to pseudo-group actions. Applying the Lemma to $\Gamma|F$ yields Lemma 6.38.

Theorem 6.39 *Let Γ be any Kleinian group, and suppose (Γ, E) is conservative, $E \subset \widehat{\mathbb{C}}$. Then E supports no Γ-invariant linefield.*

Remark. For this Theorem we do not assume Γ is finitely generated. This very general statement shows that the Γ-invariant complex structure is unique on the conservative part of the dynamics; any variation must be supported on the dissipative part, where it can be freely specified in a measurable fundamental domain.

Proof. Suppose to the contrary that E supports a Γ-invariant line field, specified by a Beltrami differential with $|\mu| = 1$. Fix a small $\epsilon > 0$. Choose coordinates so that $z = \infty$ is not fixed by any $g \neq e$ in Γ. Writing $\mu = \mu(z)d\overline{z}/dz$, we can find a compact set of positive measure $K \subset \mathbb{C}$ on which the linefield has nearly constant slope. Rotating the linefield, we can assume the slope is nearly one, i.e. $|\mu(z) - 1| < \epsilon$ for all $z \in K$.

Almost every point of K is a point of Lebesgue density. So shrinking K, we can assume there exists an $r_0 > 0$ such that for any $z \in K$ and $r < r_0$, we have $|\mu(w) - 1| < \epsilon$ for 99% of the points $w \in B(z, r)$

By the measure-theory Lemma above, for any $\delta > 0$ there is a $z \in K$ and $g \neq e$ in Γ such that $g(z) \in K$ and $|Jg(z) - 1| < \delta$. (In fact there is a positive measure set of such z, but only one is necessary to get a contradiction.) Since $g(\infty) \neq \infty$, we can write $g = I \circ R$ where $I : \mathbb{C} \to \mathbb{C}$ is a Euclidean isometry and $R : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is a reflection through a circle $S = S(p, s)$.

Now the Möbius transformation g is determined by its 2-jet, $g(z)$, $g'(z)$ and $g''(z)$. The first two terms range in a compact set. By taking K small, we can exclude g from any finite subset of Γ, so $g''(z)$ must be large. But this means the radius s of S is small. So with a suitable choice of δ, we can assume $s \ll r$. Since $Jg(z) = JR(z) = s^2|z-c|^{-2}$ is close to one, the point z lies in the annulus $A = \{w : s/2 < |w - c| < 2s\}$ about S, which satisfies $R(A) = A$.

But it is clear that R sends the parallel linefield on A to a linefield that is very far from parallel. On the other hand, $g(A) = I(R(A)) = I(A)$ is isometric to A, so it still has diameter $2s \ll r$ and it contains $g(z) \in K$. Thus $\mu|A$ and $\mu|g(A)$ are both nearly 1, which is impossible. ∎

Notes.

1. Sullivan's proof appears in [Sul2]; see also [Ot, Ch. 7]. One of its first applications, as explained in [Ot], was in the endgame of Thurston's construction of hyperbolic structures on 3-manifolds that fiber over the circle. For a survey of the ergodic theory of Kleinian groups, see [Sul3].

2. Attached to any measurable dynamical system (Γ, E) there is a von Neumann algebra A. To construct A, one first forms a bundle of Hilbert spaces $H \to E/\Gamma$ whose fiber over Γx is $\ell^2(\Gamma x)$. Then A is the space of L^∞ sections of the bundle of operator algebras $B(H)$.

   One can also describe A as a space of matrices of the form $T = (T_{xy})$ with $x, y \in E$, such that $T_{xy} = 0$ unless x and y lie in the same orbit. Composition is defined by $(ST)_{xz} = \sum_y S_{xy}T_{yz}$.

The space $L^\infty(E)$ forms the commutative subalgebra of diagonal operators in A. That is, for each $x \in E$ and $f \in L^\infty(E)$, we have a bounded operator

$$F_x : \ell^2(\Gamma x) \to \ell^2(\Gamma x)$$

defined by $F_x(a_{\gamma x}) = f(\gamma x) a_{\gamma x}$, and these fit together to give a section of $B(H)$. The group Γ also embeds in A, acting by a unitary shift of each Hilbert space $\ell^2(\Gamma x)$.

The center of A corresponds to $L^\infty(E/\Gamma)$, so A is a *factor* (the center is trivial) iff (Γ, E) is ergodic. It can also be shown that A admits a trace (a linear functional satisfying $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ and $\mathrm{tr}(1) = 1$) iff (Γ, E) admits an invariant measure. In this case, (Γ, E) is said to be a factor of type II_1.

The proof that Λ supports no Γ-invariant line field (when Γ is finitely generated) also shows that (Γ, Λ) is an algebra of type III_1 — it admits no invariant measure, and the 'ratio set' of Radon-Nikodym derivatives is dense in $\mathbb{R}^*$ for measure equivalent to Lebesgue.

3. A nice example of an infinitely-generated group Γ which is quasiconformally rigid is the group generated by reflections in a hexagonal packing of circles in $\mathbb{C}$. Here the quotient Riemann surface Ω/Γ is a countable union of triply-punctured spheres (corresponding to the interstices in the packing), so its Teichmüller space is trivial.

   A fundamental domain F for Γ is the region above the hemispheres resting on the circles in the hexagonal packing. The set $E = \overline{F} \cap S^2_\infty$ is just the closure of the union of the interstices, and it serves as a fundamental domain for the dissipative part of the action on S^2_∞. Since $m(E \cap \Lambda) = 0$, the action on the limit set is conservative, so $M(\Lambda)^\Gamma = 0$. Therefore Γ is rigid.

   Using the fact, one can show that circle packings furnish an algorithm for the construction of Riemann mappings [RS]. This algorithm has been used to apply conformal mappings to the human brain.

6.13 Sullivan's bound on cusps

Theorem 6.40 (Sullivan) *Let Γ be a nonelementary N generator Kleinian group. Then the number of cusps of $M = \mathbb{H}^3/\Gamma$ is at most $5N - 5$.*

The idea of the proof is similar to the proof of Bers' area theorem, but with an analytical twist: we allow *distributional* Beltrami coefficients.

More precisely, let $C \subset \Lambda$ be the countable set of *cusps* of Γ, i.e. fixed-points of parabolic elements. Let $M_d(C)$ be the vector space of Γ-invariant $(-d, 1)$-forms on C with the quality of *measures*. That is, $\mu \in M_d(C)$ is locally of the form $\mu(z) d\overline{z}/(dz)^d$, where $\mu(z)$ is a measure supported on C.

Theorem 6.41 *For $d \geq 2$, we have* $\dim M_d(C) = |C/\Gamma|$.

Proof. We claim that for any cusp $c \in \widehat{\mathbb{C}}$,

$$\sum_{\Gamma/\Gamma_c} \|\gamma'(c)\|^3 < \infty.$$

Here Γ_c is the stabilizer of c. Note that $\gamma'(c)$ is well-defined since $\gamma'(c) = 1$ for $\gamma \in \Gamma_c$. Then $\|\gamma'(c)\|$ is bounded above by the Euclidean volume of the horoball (in the ball model for $\mathbb{H}^3$) resting on $\gamma(c)$. Since these are disjoint we get convergence.

Now consider, for $c \in C$, the measure-differential $\mu_c = \delta(z-c)d\overline{z}/(dz)^d$. Then $|\mu_c|$ transforms by $|\gamma'(c)|^{2+d}$ (the 2 comes from the action on area densities the sphere), so

$$\mu = \sum \sum_{\Gamma_c} \gamma^* \mu_c$$

converges to an element of $M_d(C)$ supported on Γc. Since μ's with different supports are linearly independent, we obtain the bound on $\dim M_d(C)$. ∎

Proof of Theorem 6.40. Solving the $\overline{\partial}$-equation, we have as usual a cocycle map $\delta : M_d(C) \to H^1(\Gamma, V_d)$. We claim δ is injective. Indeed, if $\delta\mu = 0$, then $\overline{\partial}v = \mu$ for some Γ-invariant d-field $v = v(z)(d/dz)^d$. Since the fundamental solution $1/z$ to the $\overline{\partial}$-equation μ is in L^1_{loc}, so is v. Now v is holomorphic on Ω, so it vanishes there — the Riemann surface Ω/Γ admits no holomorphic vector fields or d-fields. On the other hand, v also vanishes on the limit set by a variant of the no-invariant linefields theorem. Thus $v = 0$.

Taking $d = 2$, we find

$$|C/\Gamma| = \dim M_2(C) \leq \dim H^1(\Gamma, V_2) \leq 5N - 5.$$

∎

Note. Sullivan's bound on the number of cusps appears in [Sul1]. (In this reference, the bound of $5N-5$ is misstated as $5N-4$.)

6.14 The Teichmüller space of a 3-manifold

Let M be a compact 3-manifold. A *convex* hyperbolic structure on M is a Riemann metric g of constant curvature -1 such that ∂M is locally convex. Equivalently, for any homotopy class of path γ between two points in $\mathrm{int}(M)$, the shortest representative of γ also lies in $\mathrm{int}(M)$.

Given such a metric, the developing map $\widetilde{M} \to \mathbb{H}^n$ is injective and its image is convex. Thus (M, g) can be extended to a complete hyperbolic manifold N in a unique way. The manifold $N = \mathbb{H}^n/\Gamma$ is *geometrically finite* and indeed *convex cocompact*: that is, the convex core of N, given by

$$\mathrm{core}(N) = \mathrm{hull}(\Lambda)/\Gamma \subset N,$$

is compact.

Conversely, for any convex cocompact hyperbolic manifold N, a unit neighborhood M of its convex core carries a (strictly) convex hyperbolic structure.

We can complete N to a *Kleinian manifold*

$$\overline{N} = (\mathbb{H}^n \cup \Omega)/\Gamma,$$

with a complete hyperbolic metric on its interior and a conformal structure on its boundary. The manifolds M and $\overline{N}$ are homeomorphic.

In analogy with the Teichmüller space of a Riemann surface, given a compact oriented 3-manifold M, we can consider the space $GF(M)$ of all geometrically finite hyperbolic manifolds *marked* by M.

A point $(f, \overline{N}) \in GF(M)$ is represented by a homeomorphism

$$f : M \to \overline{N},$$

where $\overline{N}$ is an oriented Kleinian manifold, and f respects orientations. As usual, two pairs $(f_1, \overline{N}_1)$, $(f_2, \overline{N}_2)$ are *equivalent* if there is an orientation-preserving isometry $\iota : \overline{N}_1 \to \overline{N}_2$ and an $h : M \to M$ homotopic to the identity such that

$$\begin{array}{ccc} M & \xrightarrow{f_1} & \overline{N}_1 \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle \iota} \\ M & \xrightarrow{f_2} & \overline{N}_2 \end{array}$$

commutes.

Theorem 6.42 *Let M be a compact 3-manifold admitting at least one convex hyperbolic structure. Then there is a naturally isomorphism*

$$GF(M) \cong \mathrm{Teich}(\partial M)$$

given by

$$(N, f) \mapsto (\partial N, f|\partial M).$$

6.15 Hyperbolic volume

Let $л(\theta) = -\int_0^\theta \log|2\sin t|\,dt$ be the Lobachevsky function. An ideal tetrahedron $T(\alpha, \beta, \gamma)$ is determined by its dihedral angles, which sum to π.

Theorem 6.43 *The hyperbolic volume of $T(\alpha, \beta, \gamma)$ is given by $л(\alpha) + л(\beta) + л(\gamma)$.*

Idea of the proof.

Corollary 6.44 *The regular tetrahedron has maximum volume.*

Proof. By Lagrange multipliers, at the maximum of $л(\alpha)+л(\beta)+л(\gamma)$ subject to $\alpha + \beta + \gamma = \pi$ we have $л'(\alpha) = л'(\beta) = л'(\gamma)$. Since $л'(\theta)$ determines $|\sin\theta|$, this means $\sin(\alpha) = \sin(\beta) = \sin(\gamma)$. By the law of sines, this means the associated triangle is equilateral. ∎

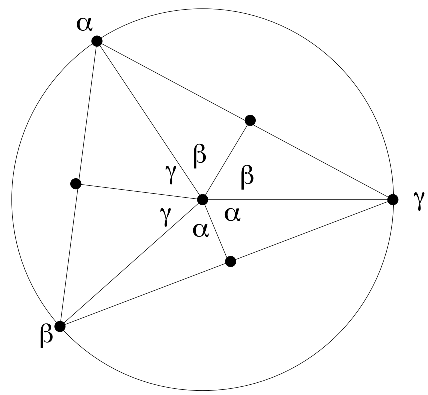

Figure 17. Dissection of an ideal tetrahedron into 6 pieces.

Ideal octahedron, Whitehead link, Apollonian gasket, Scott's theorem.

Theorem 6.45 *Surface groups are LERF.*

This means every finitely generated subgroup of $\pi_1(\Sigma_g)$ comes from a subsurface of a finite-sheeted covering space.

Corollary 6.46 *Every closed geodesic on a closed surface is covered by a simple geodesic on a finite cover.*

The proof is based in part on:

Theorem 6.47 $\mathbb{H}$ *is exhausted by convex regions tiled by right pentagons.*

We note that in fact the layers of pentagons surrounding a given one always form convex regions. Indeed, the total number of edges e_n and right angles a_n around the nth layer satisfies

$$\begin{pmatrix} a_{n+1} \\ e_{n+1} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} a_n \\ e_n \end{pmatrix}.$$

By Gauss-Bonnet the total number of pentagons p_n out to generation n is $a_n - 4$. Thus one can easily compute $p_1, p_2, \ldots = 1, 11, 51, 201, 761, 2851, 10651, \ldots$ with $p_i \sim C\lambda^i$, $\lambda = 2 + \sqrt{3}$.

Three dimensions. It is also known that the Bianchi groups $\mathrm{SL}_2(\mathcal{O}_D)$ are LERF, and in particular that the figure eight and Whitehead link complements are LERF. However the proof is much more difficult in 3-dimensions, because a finitely-generated subgroup of a geometrically finite group need not be geometrically finite. In particular, the proof for the Whitehead link complement does

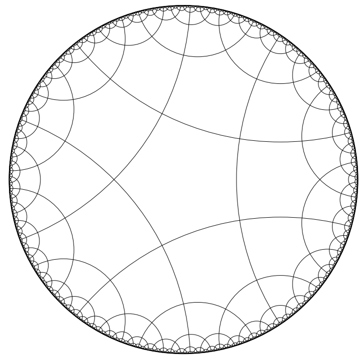

Figure 18. Tiling by right pentagons.

not just follow from commensurability with a reflection group, as indicated in [Scott].

There are also now examples of 3-manifold groups which are *not* LERF; however these examples are torus-reducible.

Question. Is the fundamental group of every compact hyperbolic 3-manifold LERF?

Question. Is every surface subgroup of a compact hyperbolic 3-manifold represented by a virtually embedded surface?

Remark: Apollonian gasket. The submanifold of M obtained by cutting along a totally geodesic triply-punctured sphere (the disk spanning one component of W) has, as its limit sets, the Apollonian gasket.

Remark: Vol3 and volume coincidences. There is a *closed* hyperbolic 3-manifold whose volume is a rational multiple of the volume of the figure eight knot complement [JR]. In general, for an arithmetic hyperbolic manifold, $\mathrm{vol}(M)$ is a rational multiple of $\zeta_k(2)$, where k is the invariant trace field of M; but the commensurability class of M depends on the associated quaternion algebra over k.

In general a quaternion algebra over a quadratic imaginary field is uniquely determined by specifying the even number of places where it ramifies. It gives a cocompact lattice *unless* this set of places is empty, in which case $Q \cong \mathrm{M}_2(k)$.

7 Holomorphic motions and structural stability

In this section we study holomorphic motions, and their use in constructing conjugacy for holomorphic dynamical systems.

7.1 The notion of motion

A *holomorphic motion* of a set $A \subset \widehat{\mathbb{C}}$ over a pointed complex manifold (Λ, λ_0) is a map $\phi : \Lambda \times A \to \widehat{\mathbb{C}}$, written $(\lambda, a) \mapsto \phi_\lambda(a)$, such that:

1. $\phi_\lambda(a)$ is injective for each fixed λ;
2. $\phi_\lambda(a)$ is holomorphic for each fixed a; and
3. $\phi_{\lambda_0}(a) = a$.

Usually we take $(\Lambda, \lambda_0) = (\Delta, 0)$.

Examples.

1. Given a torsion-free Fuchsian group $\Gamma \subset \mathrm{Aut}(\Delta)$, the map $\phi_\lambda(\gamma(0)) = \gamma(\lambda)$ gives a holomorphic motion of $A = \Gamma \cdot 0$ over $(\Delta, 0)$.
2. The natural map sending $A = \mathbb{Z} \oplus \mathbb{Z}i$ to $\mathbb{Z} \oplus \lambda\mathbb{Z}$ gives a holomorphic motion over $(\mathbb{H}, i)$.
3. Let $E \subset \widehat{\mathbb{C}}$ be a finite set, $a \notin E$, and $A = E \cup \{a\}$. Let $\phi_\lambda | E = \mathrm{id}$ and $\phi_\lambda(a) = \pi(\lambda)$, where $\pi : (\Delta, 0) \to (\widehat{\mathbb{C}} - E, a)$ is the universal covering map. Then ϕ is a holomorphic motion of A over $(\Delta, 0)$.
4. Let $f : \mathbb{H} \to \mathbb{H}$ be an earthquake, and let $\Gamma \subset \mathrm{Aut}(\mathbb{H})$ be the set of Möbius transformations that match f on at least one geodesic. Then Γ is not in general a group, but it does have the property that $\gamma_1 \circ \gamma_2^{-1}$ is never elliptic if $\gamma_1 \neq \gamma_2$ in Γ. Therefore $\lambda \mapsto \gamma(\lambda)$ is a holomorphic motion of $\Gamma \cdot 0$ over $(\Delta, 0)$.
5. Let μ be a Beltrami differential on $\widehat{\mathbb{C}}$ with $\|\mu\|_\infty < 1$, and let $\phi_\lambda : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be the unique quasiconformal map with dilatation $\lambda\mu$ fixing $(0, 1, \infty)$. Then ϕ gives a holomorphic motion of the whole sphere.

Theorem 7.1 *A holomorphic motion of A over $(\Delta, 0)$ extends uniquely to a holomorphic and continuous motion of $\overline{A}$. For each fixed λ, the map $\phi_\lambda : A \to \widehat{\mathbb{C}}$ extends to a $(1+k)/(1-k)$-quasiconformal homeomorphism of the sphere, $k = |\lambda|$.*

Topological remarks. Given $A \subset Y$, a *continuous motion* of A over (T, t_0) is a continuous map $\phi : T \times A \to Y$ such that ϕ_t is injective for each t, and $\phi_{t_0} = \mathrm{id}$.

A continuous motion of A need *not* extend to a continuous motion of $\overline{A}$; for example, the original might not be uniformly continuous, even though $T \times \overline{A}$ is compact. A simple counter-example is providing by sliding an infinite set of counters from one end of an abacus to another, moving the later counters faster and faster.

Also, a continuous motion of a compact set $A \subset S^2$ need not extend to a motion of S^2. An counter-example is provided by infinitely braiding $A =$

$\{0, 1, 2^{-1}, 3^{-1}, 4^{-1}, \ldots\}$. If we twist the strands through $(1, 1/2)$ once around during time $[0, 1/2]$, then those around $(1/3, /1/4)$ once around during time $[1/3, 1/4]$, etc. we obtain a 'wild braid' in $S^2 \times [0,1]$ and hence a motion that cannot be extended to the sphere. (For a tameness criterion for infinite braids, see [CJ].)

On the other hand, any continuous motion of a finite set $A \subset S^2$ does extend — indeed a 'radial extension' can be constructed for small time near each point. Also it is worth noting that there are no "knotted" Cantor sets in S^2; if $A \subset S^2$ is a Cantor set, then any homeomorphism $f : A \to S^2$ extends to a homeomorphism of S^2.

7.2 Stability of the Julia set

Let $f : X \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a holomorphic family over rational maps over a complex manifold X.

Theorem 7.2 *The following conditions are equivalent.*

1. *The number of attracting cycles $N(f_t)$ is constant.*
2. *The periods of the attracting cycles of f_t are locally bounded.*
3. *Every periodic cycles of f_t is attracting, repelling or persistently indifferent.*
4. *The number of critical points in the Julia set is locally constant.*
5. *The Julia set $J(f_t)$ moves continuously in the Hausdorff topology.*
6. *The Julia set $J(f_t)$ moves locally by a conjugating holomorphic motion.*

If the critically points are locally given by holomorphic functions $c_i(t)$, then these conditions are equivalent to the functions

$$\mathcal{F} = \{t \mapsto f_t^n(c_i(t))\}$$

forming a normal family.

The proof is based on extending the holomorphic motion of the repelling periodic points to a holomorphic motion of the whole Julia set [Mc4, §4]; see also [MSS] [Ly].

When any of these conditions hold, we say $f_t(z)$ is a stable family of rational maps. The largest open set $X^{\text{stable}} \subset X$ on which these conditions hold is the *stable regime*. Since the set where $N(f_t)$ achieves its maximum is open and dense, we have:

Theorem 7.3 *The stable regime is open and dense for any holomorphic family of rational maps.*

Remark. The unstable set $X - U$ can have positive measure by a result of Rees; see [Rs].

Evidentally the number of attracting cycles is locally constant at f_0 if f_0 is expanding, since all critical points are already accounted for. This shows:

Theorem 7.4 *If f is expanding then f is stable in* Rat_d*, and hence in any holomorphic family.*

Examples.

1. $f(z) = z^2$ is expanding; thus $J(z^2 + \epsilon)$ is a quasicircle. Note that z^2 and $z^2 + \epsilon$ are not topologically conjugate on the whole sphere, only on their Julia sets.

2. $f(z) = z^2 - 2$ is unstable. For example the critical point $z = 0$ lies in $J(f)$ but escapes for infinity for $z^2 - 2 - \epsilon$. One can also perturb f slightly to create attracting cycles of arbitrarily high order.

3. $f(z) = z^2 + 1/4$ is unstable. In fact under a slight perturbation any point in $K(f)$ can be made to lie in the Julia set.

4. $f(z) = z^2 - 1$ is stable and hyperbolic.

Problem. In the expanding case, reconstruct the topological dynamical system $f : J(f) \to J(f)$ from a finite amount of combinatorial data.

Alternative proof that expanding $\implies$ stable. Let f_0 be an expanding map, $f_0 : U_0 \to V$ a covering map as before with $J(f_0) \subset U_0$. We can arrange that $\overline{V} - U_0$ is a smoothly bounded domain. We wish to construct a topological conjugacy between f_0 and a small perturbation f_t.

To do this, set $U_t = f_t^{-1}(V)$; then $f_t : U_t \to V_t$ is also a covering map. Choose a homeomorphism $\phi_0 : V \to V$ such that $\phi_0 \circ f_0(z) = f_t \circ \phi(z)$ when $f_0(z) \in V - U$. Using the lifting property of covering maps, define $\phi_{n+1} = f_t^{-1} \circ \phi_n \circ f_0$ on U, and keep its old values on $V - U$. Then using expansion it is easy to see that ϕ_n is uniformly Hölder continuous. Taking a limit and using the fact that $\bigcap f_0^{-n}(V) = J(f)$ is nowhere dense, we obtain a conjugacy ϕ. ∎

Conjecture 7.5 *A rational map $f \in$* Rat_d *or* Poly_d *is stable iff f is expanding.*

In the case of degree 2 polynomial, this conjecture implies $f_c^n(0)$ converges to an attracting cycle whenever c lies in the interior of the Mandelbrot set and $f_c(z) = z^2 + c$.

Theorem 7.6 *Expanding dynamics is dense in* Poly_2 *iff there is no c such that $J(f_c)$ has positive measure and carries an f_c-invariant linefield.*

Note: a Beltrami differential $\mu(z)d\overline{z}/dz$ determines a function on the tangent space at each point, homogeneous of degree zero, by

$$\mu(v) = \mu(p)\frac{d\overline{z}}{dz}\left(v(p)\frac{\partial}{\partial z}\right) = \mu(p)\frac{\overline{v(p)}}{v(p)}.$$

Thus $\mu(v) > 0$ iff $v(p) \in \mathbb{R} \cdot \sqrt{\overline{\mu}(p)}$. The linefield associated to a Beltrami differential μ is defined by

$$L_p = \{v \in T_p\widehat{\mathbb{C}} \,:\, \mu(v) > 0\}$$

on the set where $\mu(p) \neq 0$.

Conjecture 7.7 *Any rational map that carries an invariant line field on its Julia set is double-covered by an integral endomorphism of a complex torus.*

7.3 Extending holomorphic motions

We now discuss the problem of extending a holomorphic motion to the whole sphere. We will prove a theorem from [BR].

Theorem 7.8 (Bers-Royden) *Let $\phi : \Delta \times A \to \widehat{\mathbb{C}}$ be a holomorphic motion. Then after restricting to the disk of radius $1/3$, there is an extension of ϕ to a holomorphic motion*

$$\Phi : \Delta(1/3) \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}.$$

The extended motion can be chosen such that $\mu(\Phi_t)$ is a harmonic Beltrami differential on $\widehat{\mathbb{C}} - \overline{A}$, in which case the extension is unique.

Recognizing holomorphic motions.

Theorem 7.9 *Let $\phi : \Delta \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a continuous motion of the sphere such that ϕ_λ is quasiconformal for each λ. Then ϕ_λ is a holomorphic motion iff $\mu_\lambda = \mu(\phi_\lambda)$ varies holomorphically as a function of λ.*

Moreover, any holomorphically varying $\|\mu_\lambda\| < 1$ arises from a holomorphic motion, unique up to composition with a holomorphically varying Möbius transformation.

Example. Define

$$\phi_\lambda(z) = \begin{cases} z + \lambda/z & |z| \geq 1, \\ z + \lambda\overline{z} & |z| < 1. \end{cases}$$

Note that ϕ_λ maps the unit circle to an ellipse. We have $\mu_\lambda = \mu(\phi_\lambda) = \lambda d\overline{z}/dz$ on Δ on $\mu = 0$ elsewhere, so μ_λ varies holomorphically.

The Ahlfors-Weill extension as a holomorphic motion. Let X be the ball of radius $1/2$ in the space of L^∞ holomorphic quadratic differential ϕ on $\mathbb{H}$. Then X has the structure of an infinite-dimensional complex manifold. To each $\phi \in X$ we can associate the unique univalent map $f_\phi : \mathbb{H} \to \widehat{\mathbb{C}}$ which fixes $(0, 1, \infty)$ and satisfies $Sf = \phi$. Thus we obtain a holomorphic motion

$$F : X \times \mathbb{H} \to \widehat{\mathbb{C}}$$

over $(X, 0)$ defined by $F(\phi, z) = f_\phi(z)$.

Now the Ahlfors-Weill extension prolongs F to a motion of the whole sphere, which we continue to denote by F. Examining equation 2.2 we see that for fixed $z \in -\mathbb{H}$, the extension $g(z) = M_{\overline{z}}(z)$ is a holomorphic function of f, and hence of ϕ. Thus the Ahlfors-Weill extension not only prolongs f, it prolongs the 'universal' holomorphic motion of $\mathbb{H}$, by (suitably bounded) univalent functions, to a motion of the whole sphere.

Proof of the Bers-Royden Theorem 7.8. First consider the case where the moving set A is a finite. Then $Y_\lambda = \widehat{\mathbb{C}} - \phi_\lambda(A)$ defines a holomorphic path in $\operatorname{Teich}(X)$, $X = Y_0$. By the Schwarz lemma, $\Delta(1/3)$ maps under the Bers embedding to a holomorphically varying family of quadratic differentials ϕ_λ in the ball of radius $1/2$ in $P(\overline{X})$. Thus there is a canonical quasiconformal map $f_\lambda : X \to Y_\lambda$. Its dilatation is given by

$$\mu(f_\lambda) = \rho_X^{-2} \phi_\lambda(\overline{z}),$$

so it varies holomorphically and thus f_λ defines a holomorphic motion of $\widehat{\mathbb{C}}$.

To handle the general case, exhaust A by a sequence of dense and denser finite sets. ∎

Sections of the universal curve. The *universal curve* over Teichmüller space is the complex fiber bundle $\mathcal{C}_g \to \mathcal{T}_g$ such that the fiber over $[Y]$ is Y itself. More precisely, the universal curve is the quotient of $\mathcal{T}_{g,1}$ by the subgroup $\pi_1(\Sigma_g) \subset \operatorname{Mod}_{g,1}$ coming from the kernel of the map $\operatorname{Mod}_{g,1} \to \operatorname{Mod}_g$. Alternatively, $\mathcal{C}_g$ is the pullback of the natural fibration $\mathcal{M}_{g,1} \to \mathcal{M}_g$ under the covering map (of orbifolds) $\mathcal{T}_g \to \mathcal{M}_g$.

From [Hub] we have following result.

Theorem 7.10 *There is no holomorphic section of the universal curve over $\mathcal{C}_g \to \mathcal{T}_g$ over Teichmüller space when $g \geq 3$.*

Sketch of the proof. A section of the universal curve lifts to a section $s : \mathcal{T}_g \to \mathcal{T}_{g,1}$ inverting the natural projection. Since the Teichmüller metric and the Kobayashi metric coincide (Royden), s is an isometry. It follows that there exists, for any $X \in \mathcal{T}_g$ and $p = s(X) \in X$, a unit-norm projection

$$P : Q(X - \{p\}) \to Q(X)$$

(dual to the derivative of s at X). (Note that $Q(X)$ naturally forms a subspace of $Q(X - \{p\})$, namely the subspace of differentials holomorphic at p.) Using Riemann-Roch and a differentiability argument, one shows such a projection cannot exist. ∎

Remark. For genus 2, the 6 Weierstrass points provide sections.

Corollary 7.11 *There exists a holomorphic motion*

$$\phi : U \times (-\mathbb{H}) \to \widehat{\mathbb{C}}$$

where $U \subset \mathbb{C}^n$ is a bounded domain of holomorphy, that cannot be extended to a holomorphic motion of the sphere.

Proof. Let $U = \mathcal{T}_g \cong \operatorname{Teich}(X)$. Using the construction of the Bers embedding, for each $Y \in \mathcal{T}_g$ we obtain a univalent map $f_Y : (-\mathbb{H}) \to \widehat{\mathbb{C}}$ with continuous boundary values. This gives the desired holomorphic motion, by setting $\phi(Y, z) = f_Y(z)$. If it could be extended to the whole sphere, then for $z \in \mathbb{H}$ we would have $s(Y) = \phi(Y, z) \in \Omega_Y$. Passing to the quotient $Y = \Omega_Y/\Gamma_Y$, we would then obtain a section of the universal curve $\mathcal{C}_g \to \mathcal{T}_g$, a contradiction. ■

Słodkowski's theorem.

Theorem 7.12 *A holomorphic motion $\phi : \Delta \times A \to \widehat{\mathbb{C}}$ extends to a holomorphic motion of the whole sphere. The extension can be made equivariant in a suitable sense.*

See [Sl], [Dou]

7.4 Stability of Kleinian groups

We now turn to the notion of structural stability in families of Kleinian groups. See [Bers3] and [Sul5] for more details.

Families of groups. A *holomorphic family of representations* (into $\operatorname{Aut}\widehat{\mathbb{C}}$) over a complex manifold X is a map

$$\rho : X \times G \to \operatorname{Aut}\widehat{\mathbb{C}},$$

where G is an abstract group, $\rho_t(g)$ is a holomorphic function of t for each $g \in G$, and $\rho_t : G \to \operatorname{Aut}\widehat{\mathbb{C}}$ is a homomorphism of groups for each $t \in X$.

Two representations ρ_s and ρ_t are *quasiconformally conjugate* if there is a quasiconformal map $\phi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $\phi \circ \rho_s(g) = \rho_t(g) \circ \phi$ for every $g \in G$.

The next result shows faithfulness is enough to prove a strong form of structural stability.

Theorem 7.13 *Let $\rho : X \times G \to \operatorname{Aut}\widehat{\mathbb{C}}$ be a holomorphic family of faithful representations over a connected complex manifold X. Then ρ_s and ρ_t are quasiconformally conjugate for all $s, t \in X$.*

Lemma 7.14 *Let $g, h \in \operatorname{Aut}\widehat{\mathbb{C}}$ be elements other than the identity, and let $G = \langle g, h\rangle$ be the subgroup they generate. Then exactly one of the following possibilities holds.*

1. *$G \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. Then the fixed-points of g and h are disjoint.*

2. *G is an abelian group other than $\mathbb{Z}/2 \times \mathbb{Z}/2$. Then the fixed points sets of g and h coincide.*

3. *$[G, G]$ is an infinite, torsion-free group of parabolics. Then g and h share exactly one fixed-point.*

4. *G is none of the above. Then the fixed-points of g and h are disjoint.*

Proof. First suppose the fixed-points of g and h coincide. Then it is easy to see G is abelian and not isomorphic to $\mathbb{Z}/2 \times \mathbb{Z}/2$. So the Lemma holds in this case.

Now suppose g and h share one but not all fixed-points. (In particular they are not both parabolic). Normalizing so the common fixed-point is at $z = \infty$, we can assume G lies in the solvable subgroup of affine linear maps $\operatorname{Aut} \mathbb{C}$. In particular, G is solvable, and $[G, G]$ is an infinite torsion-free subgroup of parabolic transformations. So the Lemma holds in this case as well.

Finally suppose g and h share no fixed-points. If G is abelian then g must interchange the fixed-points of h, and similarly h must interchange the fixed-points of g. Thus $g^2 = h^2 = 1$ and $G \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. Next suppose $[G, G]$ is an infinite, torsion-free abelian group. Then $[G, G]$ cannot consist entirely of parabolics; otherwise they would all share the same fixed-point p, which would also be fixed by g and h since $[G, G]$ is normal. Thus G falls into the case 'none of the above', and the Lemma is true in the case as well. ■

Isomorphic groups with different actions. The groups

$$\begin{aligned} G_1 &= \langle z \mapsto z + 1, z \mapsto -z \rangle, \\ G_2 &= \langle z \mapsto 2z, z \mapsto 1/z \rangle \end{aligned}$$

are both isomorphic to $\mathbb{Z}/2 \ltimes \mathbb{Z}$, but they have different fixed-point structures. The generators of G_1 have a common fixed-point and those of G_2 do not. This is the one instance where the isomorphism type of G does not determine the fixed-point structure.

Proof of Theorem 7.13. Since X is connected, it suffices to prove the theorem locally. Thus we can restrict to the case where $X = \Delta$.

Notice $\rho_t(g)$ is parabolic for one value of t iff it is parabolic for all values. Otherwise $\operatorname{tr} \rho_t(g)$ would be a nonconstant holomorphic function passing through the value 2; but then for some s and s we would have $\operatorname{tr} \rho_s(g) = 2\cos(\pi/n)$, which implies $\rho_s(g^n) = 1$, contradicting faithfulness.

Let Λ_t be the closure of the fixed points of the elements of $\Gamma_t = \rho_t(G)$ other than the identity. Then Λ_t is a closed, Γ_t-invariant set. Since $\rho_t(g)$ is parabolic iff $\rho_0(g)$ is parabolic, the fixed-points of individual elements move holomorphically. By the preceding Lemma, fixed-points of g and h coincide for ρ_0 iff they coincide for all ρ_t. Thus Λ_t moves by a holomorphic motion, which is also a conjugacy for the action of Γ_t.

The action of Γ_t on $\widehat{\mathbb{C}} - \Lambda_t = \Omega_t$ is free. Thus if $|\Lambda_t| < 3$, we can extend it by an orbit of Γ_t until $|\Lambda_t| \geq 3$ and it is still moving by a holomorphic conjugacy.

Now let $\phi_t : \Delta(1/3) \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be the Bers-Royden extension of the motion of Λ_t. For any $g \in G$, the holomorphic motion

$$\psi_t = \rho_t(g)^{-1} \circ \phi_t \circ \rho_0(g)$$

is also an extension of the motion of Λ_t, with $\mu(\psi_t)$ harmonic. Since the Bers-Royden extension is unique subject to this condition, we have $\psi_t = \phi_t$.

In other words, by naturality the Bers-Royden extension commutes with the action of G, and therefore it provides a quasiconformal conjugacy between ρ_0 and ρ_t whenever $|t| < 1/3$. This proves the Theorem locally, and the global version follows from connectedness of X. ∎

Here is a variant:

Theorem 7.15 *Let $\rho_t : G \to \operatorname{Aut} \widehat{\mathbb{C}}$ be a holomorphic family of representations over a connected base X. If $\rho_t(G)$ is discrete and nonelementary for all t, then all the representations are quasiconformally conjugate.*

Representation varieties. Fix a finitely-generate group G. Let us suppose G is non-elementary; that is, any abelian subgroup has infinite index.

Let $\mathcal{V}(G)$ denote the algebraic variety of irreducible representations $\rho : G \to \operatorname{Aut} \widehat{\mathbb{C}}$ modulo conjugacy. The functions $g \mapsto \operatorname{tr} \rho(g)$ provide a smooth embedding of $\mathcal{V}(G)$ into affine space. Let $AH(G) \subset \mathcal{V}(G)$ denote the subset of discrete, faithful representations. Let $CC(G) \subset AH(G)$ denote the set of faithful representations whose image is a convex cocompact Kleinian group. It is easy to see that $CC(G)$ is open, and in fact each component of $CC(G)$ is parameterized by an appropriate Teichmüller space.

We can now formulate a central conjecture, analogous to the density of expanding rational maps in Rat_d.

Conjecture 7.16 *Suppose G does not contain $\mathbb{Z} \oplus \mathbb{Z}$. Then $CC(G)$ is dense in $AH(G)$.*

It is certainly not true in general that $AH(G)$ is dense in $\mathcal{V}(G)$. Indeed, if any component of $\mathcal{V}(G)$ has positive dimension, then a trace is non-constant there, which leads to an open set of indiscrete groups. In fact we have the following basic result:

Theorem 7.17 *The set of discrete faithful representations $AH(G) \subset \mathcal{V}(G)$ is closed.*

Structural stability implies hyperbolicity. Because of the quasiconformal conjugacy result, it is natural to say that a Kleinian group $\Gamma \subset \operatorname{Aut} \widehat{\mathbb{C}}$ is *structurally stable* if all representations $\rho : \Gamma \to \operatorname{Aut} \widehat{\mathbb{C}}$ close to the identity are injective — there are no new relations.

Equivalently, the structurally stable groups correspond to the interior of $AH(G)$. We then have:

Theorem 7.18 (Sullivan) *A structurally stable finitely-generated Kleinian group Γ is convex cocompact.*

Corollary 7.19 *The set $CC(G)$ coincides with the interior of $AH(G)$.*

In other words, we do not yet know that structural stability is dense (in $AH(G)$), but we do know that structural implies hyperbolicity (i.e. the expanding property of G on its limit set).

Cusps. We remark that many important groups — since as the fundamental groups of knot complements — do contain $\mathbb{Z} \oplus \mathbb{Z}$. For these groups $CC(G)$ is empty, even though $AH(G)$ may not be. To formulate an appropriate result in this setting, one must restrict to the representations where the $\mathbb{Z} \oplus \mathbb{Z}$ subgroups are parabolic.

Quasifuchsian groups. To explain Sullivan's result, we will treat the case of a surface group, $G = \pi_1(S)$ where S is a closed surface of genus $g \geq 2$.

Theorem 7.20 *A structurally stable Kleinian group isomorphic to $\pi_1(S)$ is quasifuchsian.*

First we need some topological arguments to recognize a quasifuchsian group.

Theorem 7.21 *The limit set of any Kleinian group $\Gamma \cong \pi_1(S)$ is connected.*

Proof. Write $S = \mathbb{H}^2/\Gamma_S$, take a smooth homotopy equivalent $S \to M = \mathbb{H}^3/\Gamma$ and lift it to an equivariant map $f : \mathbb{H}^2 \to \mathbb{H}^3$. Pick any basepoint $p \in \mathbb{H}^2$. Since f has a compact fundamental domain, the limit points in S^2_∞ of $f(\mathbb{H})$ coincide exactly with the limit set $\Lambda(\Gamma)$. Since f is proper, we have

$$\Lambda(\Gamma) = \bigcap_{R>0} \overline{f(\mathbb{H} - B(p,R))} \subset \mathbb{H}^3 \cup S^2_\infty.$$

This expression presents $\Lambda(\Gamma)$ as a nested intersection of connected sets, so it is connected. ■

Theorem 7.22 *Suppose $\Gamma \cong \pi_1(S)$ has two invariant components in its domain of discontinuity. Then Γ is quasifuchsian.*

Proof. Let $\overline{M} = (\mathbb{H}^3 \cup \Omega)/\Gamma$. Since the limit set is connected, $\partial\overline{M}$ is incompressible. By assumption, there are two compact components X_0 and X_1 in $\partial\overline{M}$ such that the inclusion of each is a homotopy equivalence. By taking a homotopy from X_0 to X_1, we obtain a degree-one homotopy equivalence $f : S \times [0,1] \to \overline{M}$ sending $S \times \{i\}$ to X_i. Thus $\overline{M}$ is compact and $\partial\overline{M} = X_0 \cup X_1$.

At this point we know Γ is convex cocompact. To prove Γ is quasifuchsian, one can appeal to theorems in 3-dimensional topology that show $\overline{M}$ is homeomorphic to $S \times I$ (cf. [Hem, Ch. 10]). ■

Compare [Msk].

Proof of Theorem 7.20. Let $\rho_0 : G = \pi_1(S) \to \Gamma$ be a point in $\mathcal{V}(G)$ representing Γ. By assumption, all ρ_t close enough to ρ are quasiconformally conjugate to ρ_0.

By counting dimensions ($2g$ generators, one relation, and one more conjugating element), we have

$$\dim_{\mathbb{C}} \mathcal{V}(S) = (2g-2) \dim \operatorname{Aut} \widehat{\mathbb{C}} = 6g - 6.$$

Since ρ_0 is structurally stable, $6g-6$ agrees with the dimension of the Teichmüller space of Γ. Thus, letting $X = \Omega/\Gamma$, we have

$$6g - 6 = \dim \operatorname{Teich}(\widehat{\mathbb{C}}, \Gamma) = \dim \operatorname{Teich}(X).$$

By structurally stability, Γ has no parabolics. By the Ahlfors finiteness theorem, X is a closed surface. Therefore $\dim \operatorname{Teich}(X) = 3|\chi(X)|$, which implies $\chi(X) = 2\chi(S)$.

Each component $X_i \subset X$ can be expressed as Ω_i/Γ_i, where Γ_i is the stabilizer of a component of Ω_i of Ω. Since $\Gamma_i \cong \pi_1(X_i)$ is not a free group, it must have finite index d_i in Γ. Then $\sum d_i = 2$ since $\chi(X) = 2\chi(S)$.

In the case $1+1=2$ we have a quasifuchsian group by the preceding Theorem. To rule out the case $2=2$, note that in this case Γ_1 is a quasifuchsian subgroup of index two in Γ. Then $\overline{M}$ is homeomorphic to a properly twisted I-bundle over S. But $\overline{M}$ is orientable, as is S, so the I-bundle must be trivial. ■

Components of the $CC(G)$. In general the compact hyperbolic manifold $\overline{M}$ has a constant *homeomorphism* type on each component U of $CC(G) = \operatorname{int} AH(G)$, while manifolds in different components of $CC(G)$ are *homotopy equivalent* but not homeomorphic. Thus the classification of components of $CC(G)$ is bound up in the classification of homeomorphism types of 3-manifolds with a fixed homotopy type.

For $G = \pi_1(S)$ only one (orientable) homeomorphism type arises ($S \times I$), but in general there are many possibilities, coming from different ways in which $\pi_1(\partial M)$ may be deployed in $\pi_1(M)$.

A similar situation arises already for surfaces; a pair of pants and a torus with one boundary component are also homotopy equivalent but not homeomorphic.

7.5 Cusped tori

In this section we give an example of the shape of the domain of locally faithful representations in a holomorphic family.

The representation variety and simple closed curves. Let S be a compact surface of genus $g=1$ with $n=1$ boundary component. We will study representations of the free-group

$$G = \langle a, b \rangle \cong \pi_1(S),$$

subject to the relation that $\rho([a,b])$ is parabolic. Geometrically, this condition forces ∂S to be realized as a cusp.

The traces (α, β, γ) of

$$(A, B, AB) = (\rho(a), \rho(b), \rho(ab))$$

determine ρ up conjugacy (when it is irreducible). In particular, the trace of the commutator can be calculated from the relation

$$\operatorname{tr}(A)^2 + \operatorname{tr}(B)^2 + \operatorname{tr}(AB)^2 = \operatorname{tr}(A)\operatorname{tr}(B)\operatorname{tr}(AB) + \operatorname{tr}[A,B] + 2.$$

Imposing the condition $\operatorname{tr}[A,B] = -2$ reduces one to a hypersurface

$$\mathcal{V}_0(G) \subset \mathbb{C}^3$$

defined by the equation

$$\alpha^2 + \beta^2 + \gamma^2 = \alpha\beta\gamma. \tag{7.1}$$

This slice of the representation variety of the group G contains a copy of the quasifuchsian space $\mathrm{QF}(S)$, where S is a torus with one boundary component.

Note that once α and β are specified, there are two choices for γ. But these two choices are related by the change of basis $A \mapsto A^{-1}$. This change of basis preserves $\operatorname{tr} A$ and $\operatorname{tr} B$ but changes $\operatorname{tr} AB$ to $\operatorname{tr} A^{-1}B$. These two traces are related by the important equation

$$\operatorname{tr} AB + \operatorname{tr} A^{-1}B = \operatorname{tr} A \operatorname{tr} B, \tag{7.2}$$

coming from the relation $A + A^{-1} = (\operatorname{tr} A)I$.

It is a useful coincidence that $\mathrm{SL}_2(\mathbb{Z})$ is both a Fuchsian group and the mapping-class group of a punctured torus. Thus $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on pairs of generators of $\pi_1(S)$. The generators themselves correspond bijectively to simple closed curves on S, which in turn correspond to lines of rational slope in $H_1(S,\mathbb{R})$. Passing to projective space, we can identify the space of slopes with the rational points (union infinity) on the boundary of the upper halfplane. Then the action of $\mathrm{Mod}(S)$ on $\mathbb{P}H_1(S,\mathbb{R})$ goes over to the usual action of $\mathrm{SL}_2(\mathbb{Z})$ on $\widehat{\mathbb{R}}$.

The standard generators $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$ for $\mathrm{SL}_2(\mathbb{Z})$ act by the automorphism of G given by:

$$\langle a,b\rangle \to \langle a, ab\rangle \quad \text{and} \langle a,b\rangle \to \langle ab, a\rangle.$$

Using equation (7.2), we obtain simple expressions for the action of $\mathrm{SL}_2(\mathbb{Z}) = \mathrm{Mod}(S)$ on $\mathcal{V}(S) \subset \mathbb{C}^3$, namely:

$$(\alpha,\beta,\gamma) \mapsto (\alpha,\gamma,\alpha\gamma-\beta) \quad \text{and} (\alpha,\beta,\gamma) \mapsto (\gamma,\beta,\beta\gamma-\alpha).$$

Now for any slope $s \in \mathbb{Q} \cup \{\infty\}$, there is a unique pair of conjugacy classes $w \in G \cong \pi_1(S)$ represented by simple closed curves of slope s. (There is a pair because w and w^{-1} have the same slope.) It turns out that we have

$$\operatorname{Tr}\rho(w) = P_s(\alpha,\beta,\gamma),$$

where P_s is a polynomial. This polynomial, and the words w representing a given slope, are easily computed recursively using the action of $\mathrm{SL}_2(\mathbb{Z})$. Namely, to compute P_s we find an element $T \in \mathrm{SL}_2(\mathbb{Z})$ sending s to 0; then we compute the polynomial action of T on $\mathcal{V}(S)$; and finally we observe that $P_s(\alpha, \beta, \gamma)$ is nothing more than the first coordinate of T.

The standard tiling of the unit disk by ideal triangles is a convenient way to organize this computation. Starting with a pair of triangles forming a quadrilateral, we recursively fill out the hyperbolic plane by inserting reflection through the edges. Each time a new triangle is added along an edge E, the element w of the free group for the new vertex v is the product of the elements at the vertices of E. The trace of the new element is determined by the (previously calculated) traces of other three elements on the vertices of the quadrilateral with diagonal E. Finally the slope p/q of w is determined from the slopes a/b and c/d of the words at the vertices of E by Farey addition:

$$\frac{p}{q} = \frac{a+c}{b+d}\cdot$$

In is thus straightforward to calculate $P_s(\alpha, \beta, \gamma)$ for any rational slope $s = p/q$.

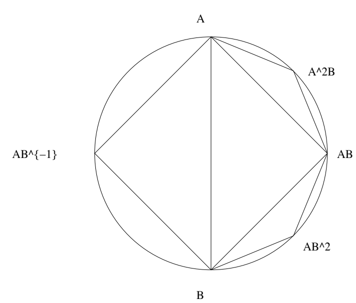

Figure 19. Farey triangles. We have $\operatorname{tr} AB^{-1} + \operatorname{tr} AB = \operatorname{tr} A \operatorname{tr} B$.

The Maskit slice. To reduce to a 1-dimensional space, let us impose the condition $\beta = \operatorname{tr}(B) = 2$. Then the only essential remaining free parameter is $\alpha = \operatorname{tr}(A) \in \mathbb{C}$. More precisely, for every value of α there are two possible values of γ, but for $\beta = 2$ equation (7.1) factors as

$$(\gamma - \alpha)^2 = -4,$$

so we can uniformly choose $\gamma = \alpha + 2i$.

Indeed, once we have chosen

$$\beta = 2 \quad \text{and} \quad \gamma = \alpha + 2i,$$

we can go further and obtain an explicit family of representations

$$\rho_\alpha : G \to \operatorname{Aut} \mathbb{C},$$

$\alpha \in X = \mathbb{C}$, given by the matrices:

$$A = \begin{pmatrix} 0 & i \\ i & \alpha \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$
$$AB = \begin{pmatrix} 0 & i \\ i & \alpha + 2i \end{pmatrix}, \qquad [A,B] = \begin{pmatrix} 1 & -2 \\ 2 & -3 \end{pmatrix}.$$

Note that B and $[A,B]$ together generate a Fuchsian subgroup of $\Gamma_\alpha = \rho_\alpha(G)$. In fact they generate the congruence subgroup $\Gamma(2) \subset \mathrm{SL}_2(\mathbb{Z})$, which uniformizes $X = \widehat{\mathbb{C}} - \{0, 1, \infty\}$. This comes from the fact that the condition $\mathrm{tr}(B) = 2$ pinches a loop on the torus to a cusp, resulting in one boundary component becoming a triply-punctured sphere. Thus the limit set $\Lambda(\Gamma_\alpha)$ contains a tree of circles, one for each conjugate of $\Gamma(2)$ inside Γ_α. Its domain of discontinuity Ω consists of a single invariant component together with a countable collection of round disks (see Figure 20).

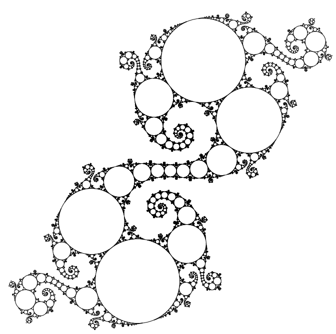

Figure 20. Limit set for Γ_α in Maskits' embedding of $\mathcal{T}_{1,1}$. The unbounded domain uniformizes a punctured torus; the round disks, triply-punctured spheres.

The regime $T \subset \mathbb{C}$ where ρ_α is locally faithful is parameterized by the Teichmüller space of $X = \Omega/\Gamma$. It turns out that X consists of a triply-punctured sphere (with no moduli) and a once-punctured torus. Then T is a natural model for $T_{1,1}$, namely *Maskit's embedding* of Teichmüller space.

What does the domain T look like? To find T, we can look for solutions to the polynomial equation $Q_s(\alpha) = P_s(\alpha, \beta, \gamma)$ as s ranges over $\mathbb{Q} \cup \{\infty\}$. The roots of $Q_s(\alpha)$ are definitely outside T, since near these points $\rho_\alpha(w)$ becomes parabolic for a word $w \in G$ of slope s.

A plot of those points is shown in Figure 21. The zeros of $P_s(\alpha)$ as s varies over many rational slopes, give rise to the points below the cusped curve in Figure 21. The cusped curve itself, which forms the boundary of the stable regime T, is obtained by connecting together the 'highest' zeros. (One should take this terminology with a grain of salt, since the curve is not a graph — in

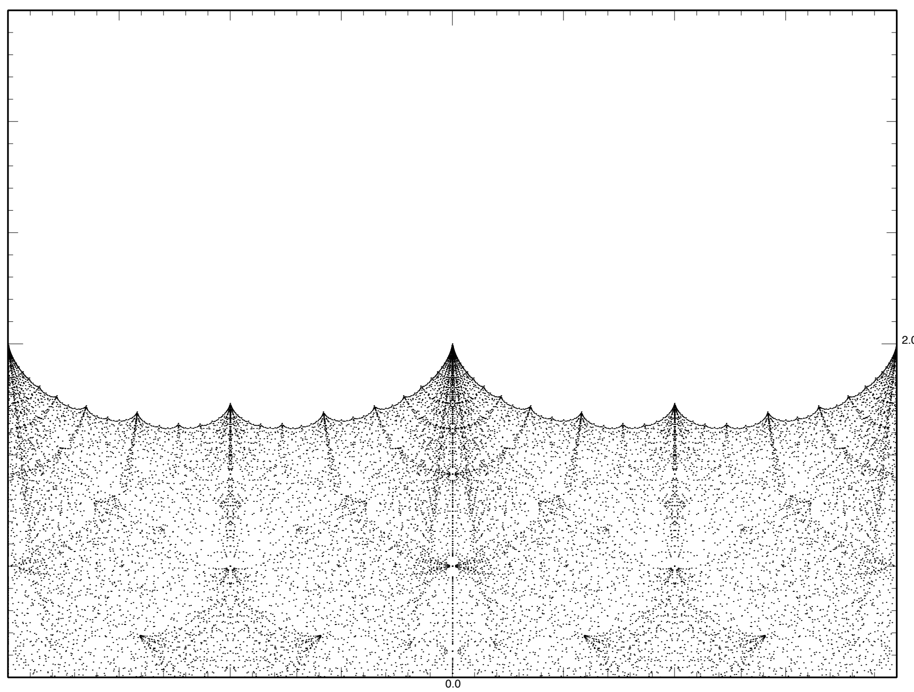

Figure 21. Faithful representations. Dots represent relations.

fact it winds an infinite amount around most boundary points that correspond to quadratic irrationals.)

Notes. The case of punctured-torus groups we examine above has been much-studied. See, for example, [Wr], [MMW], [Mc5, §3.7], [KS1], [KS2].

7.6 Structural stability of rational maps

We now turn to the study of stability for families of rational maps on the Riemann sphere.

Definitions. Let X be a complex manifold. A *holomorphic family* of rational maps $f_\lambda(z)$ over X is a holomorphic map $X \times \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, given by $(\lambda, z) \mapsto f_\lambda(z)$.

Let $X^{\mathrm{top}} \subset X$ be the set of *topologically stable* parameters. That is, $\alpha \in X^{\mathrm{top}}$ if and only if there is a neighborhood U of α such that f_α and f_β are topologically conjugate for all $\beta \in U$.[2]

The space $X^{\mathrm{qc}} \subset X^{\mathrm{top}}$ of *quasiconformally stable* parameters is defined similarly, except the conjugacy is required to be quasiconformal.

Let $X_0 \subset X$ be the set of parameters such that the number of critical points of f_λ (counted *without* multiplicity) is locally constant. For $\lambda \in X_0$ the critical points can be locally labeled by holomorphic functions $c_1(\lambda), \ldots, c_n(\lambda)$. Indeed, X_0 is the maximal open set over which the projection $C \to X$ is a covering space, where C is variety of critical points

$$\{(\lambda, c) \in X \times \widehat{\mathbb{C}} \,:\, f'_\lambda(c) = 0\}.$$

A *critical orbit relation* of f_λ is a set of integers (i, j, a, b) such that $f^a(c_i(\lambda)) = f^b(c_j(\lambda))$; here $a, b \geq 0$.

The set $X^{\mathrm{post}} \subset X_0$ of *postcritically stable* parameters consists of those λ such that the set of critical orbit relations is locally constant. That is, $\lambda \in X^{\mathrm{post}}$ if any coincidence between the forward orbits of two critical points persists under a small change in λ. This is clearly necessary for topologically stability, so $X^{\mathrm{top}} \subset X^{\mathrm{post}}$.

The main result of this section, whose proof is completed in the next, is:

Theorem 7.23 *In any holomorphic family of rational maps, the topologically stable parameters are open and dense.*

Moreover the structurally stable, quasiconformally stable and postcritically stable parameters coincide ($X^{top} = X^{qc} = X^{post}$).

This result was anticipated and nearly established in [MSS, Theorem D]. It is completed in [McS].

Definition. Let $f_\lambda(z)$ be a holomorphic family of rational maps over (X, x). A holomorphic motion *respects the dynamics* if it is a conjugacy: that is, if

$$\phi_\lambda(f_x(a)) = f_\lambda(\phi_\lambda(a))$$

whenever a and $f_x(a)$ both belong to A.

[2] In the terminology of smooth dynamics, these are the *structurally stable* parameters in the family X.

Theorem 7.24 *For any $x \in X^{post}$, there is a neighborhood U of x and a holomorphic motion of the sphere over (U, x) respecting the dynamics.*

Proof. Choose a polydisk neighborhood V of x in X^{post}. Then over V the critical points of f_λ can be labeled by distinct holomorphic functions $c_i(\lambda)$. In other words,

$$\{c_1(\lambda), \ldots, c_n(\lambda)\}$$

defines a holomorphic motion of the critical points of f_x over V.

The condition of constant critical orbit relations is exactly what we need to extend this motion to the forward orbits of the critical points. That is, if we specify a correspondence between the forward orbits of the critical points of f_x and f_λ by

$$f_x^a(c_i(x)) \mapsto f_\lambda^a(c_i(\lambda)),$$

the mapping we obtain is well-defined, injective and depends holomorphically on λ.

Next we extend this motion to the *grand* orbits of the critical points. There is a unique extension compatible with the dynamics. Indeed, consider a typical point where the motion has already been defined, say by $q(\lambda)$. Let

$$Z = \{(\lambda, p) \,:\, f_\lambda(p) = q(\lambda)\} \subset V \times \widehat{\mathbb{C}} \xrightarrow{\ \pi\ } V$$

be the graph of the multivalued function $f_\lambda^{-1}(q(\lambda))$. If $q(\lambda)$ has a preimage under f_λ which is a critical point $c_i(\lambda)$, then this critical point has constant multiplicity and the graph of c_i forms one component of Z. The remaining preimages of $q(\lambda)$ have multiplicity one, and therefore $\pi^{-1}(\lambda)$ has constant cardinality as λ varies in V. Consequently Z is a union of graphs of single-valued functions giving a holomorphic motion of $f_x^{-1}(q(x))$.

The preimages of $q(\lambda)$ under f_x^n are treated similarly, by induction on n, giving a holomorphic motion of the grand orbits compatible with the dynamics.

By the λ-lemma, this motion extends to one sending $\widehat{P}(x)$ to $\widehat{P}(\lambda)$, where $\widehat{P}(\lambda)$ denotes the closure of the grand orbits of the critical points of f_λ.

If $|\widehat{P}(x)| \leq 2$, then f_λ is conjugate to $z \mapsto z^n$ for all $\lambda \in V$; the theorem is easy in this special case. Otherwise $\widehat{P}(x)$ contains the Julia set of f_x, and its complement is a union of hyperbolic Riemann surfaces.

To conclude the proof, we apply the Bers-Royden Harmonic λ-lemma to extend the motion of $\widehat{P}(x)$ to a *unique* motion $\phi_\lambda(z)$ of the whole sphere, such that the Beltrami coefficient $\mu_\lambda(z)$ of $\phi_\lambda(z)$ is harmonic on $\widehat{\mathbb{C}} - \widehat{P}(x)$. This motion is defined on a polydisk neighborhood U of x of one-third the size of V.

For each λ in U the map

$$f_\lambda \,:\, (\widehat{\mathbb{C}} - \widehat{P}(\lambda)) \to (\widehat{\mathbb{C}} - \widehat{P}(\lambda))$$

is a covering map. Define another extension of the motion to the whole sphere by

$$\psi_\lambda(z) \;=\; f_\lambda^{-1} \circ \phi_\lambda \circ f_x(z),$$

where $z \in \widehat{\mathbb{C}} - \widehat{P}(x)$ and the branch of the inverse is chosen continuously so that $\psi_x(z) = z$. (On $\widehat{P}(x)$ we set leave the motion the same, since it already respects the dynamics).

The rational maps f_λ and f_x are conformal, so the Beltrami coefficient of ψ_λ is simply $f_x^*(\mu_\lambda)$. But f_x is a holomorphic local isometry for the hyperbolic metric on $\widehat{\mathbb{C}} - \widehat{P}(x)$, so it pulls back harmonic Beltrami differentials to harmonic Beltrami differentials. By uniqueness of the Bers-Royden extension, we have $\psi_\lambda = \phi_\lambda$, and consequently the motion ϕ respects the dynamics. ■

Corollary 7.25 *The postcritically stable, quasiconformally stable and topologically stable parameters coincide.*

Proof. It is clear that $X^{\mathrm{qc}} \subset X^{\mathrm{top}} \subset X^{\mathrm{post}}$. By the preceding theorem, if $x \in X^{\mathrm{post}}$ then for all λ in a neighborhood of x we have $\phi_\lambda \circ f_\lambda = f_x \circ \phi_\lambda$, where $\phi_\lambda(z)$ is a holomorphic motion of the sphere. By the λ-lemma, $\phi_\lambda(z)$ is quasiconformal, so $X^{\mathrm{post}} \subset X^{\mathrm{qc}}$. ■

Remark. The monodromy of holomorphic motions in non-simply connected families of rational maps can be quite interesting; see [Mc1], [Mc2], [GK], [BDK] and [Br].

7.7 Postcritical stability

To establish the density of structural stability, we must bridge the gap between J-stability (discussed in §7.2) and postcritical stability. We will show that X^{post} is only slightly smaller than X^{stable}, so it too is dense. Since $X^{\mathrm{post}} = X^{\mathrm{top}} = X^{\mathrm{qc}}$, the proof of Theorem 7.23 is completed by:

Theorem 7.26 *The postcritically stable parameters are open and dense in the set of J-stable parameters.*

Proof. Using Corollary 7.25, we have $X^{\mathrm{post}} \subset X^{\mathrm{stable}}$ because $X^{\mathrm{post}} = X^{\mathrm{top}}$ and topological conjugacy preserves the number of attracting cycles. By definition X^{post} is open, so it only remains to prove it is dense in X^{stable}.

Let $V \cong \Delta^m \subset X^{\mathrm{stable}}$ be any polydisk on which number of critical points of f_λ is constant. The advantage of working on V is that we can label the critical points of $f_\lambda(z)$ by holomorphic functions $c_i : V \to \widehat{\mathbb{C}}$, $i = 1, \ldots, n$.

Fixing i and j, we will show there is an open dense subset of V on which the critical orbit relations between i and j are constant. This will suffice to complete the proof, since the intersection of a finite number of open dense sets is again open and dense.

Since V is a simply-connected subset of the J-stable regime, over V the dynamics on the Julia set is canonically trivialized. In particular, the critical points in the Julia set remain there and their critical orbit relations are constant. So we may assume c_i and c_j lie outside the Julia set.

For any $\kappa, \lambda \in V$, there is a canonical correspondence between the *components* of the Fatou sets $\Omega(f_\kappa)$ and $\Omega(f_\lambda)$. This correspondence commutes with the induced dynamics on the set of components. It also preserves the *types* of periodic components, except that an attracting component may become superattracting or vice-versa.

We begin by considering the case $i = j$. Suppose the relation $f^a(c_i) = f^b(c_i)$ for some $a < b$ holds throughout V. Then there are only a finite number of possibilities for the set of all critical self-relations of c_i. Each self-relation either holds throughout V or on a proper complex-analytic subset of V. Thus there is an open dense set — the complement of a complex subvariety — on which the self-relations of c_i are constant.

Next suppose the relation $f^a(c_i) = f^b(c_i)$ holds at $\lambda_0 \in V$ but not throughout V. Then there exists a periodic point $p(\lambda)$ such that $f^b(c_i) = p$ for $\lambda = \lambda_0$. The point p is the center of an attracting or superattracting basin, or a Siegel disk. By assumption, the relation $f^b(c_i) = p$ only holds on a proper analytic subvariety $W \subset V$. For all $\lambda \in V - W$ near λ_0, the forward orbit of the critical point lands near but not on the center of the basin. Thus c_i has an infinite forward orbit for such λ, and in particular its self-relations are constant on an open dense set.

Finally it may be that no relation of the form $f^a(c_i) = f^b(c_i)$ ever hold in V. Then the forward orbit of c_i is always infinite and the set of self-relations is constant in this case as well.

Since we have shown the regime where the self-relations among critical points are constant is open and dense, we may now replace V with a polydisk contained in this regime.

Now consider two critical points c_i, c_j, $i \neq j$, lying outside of the Julia set. Suppose there is a critical point relation $f^a(c_i) = f^b(c_j)$ for $\lambda = \lambda_0$. Then c_i either both have infinite orbits or both have finite orbits. In the latter case, only a finite number of patterns of orbit relations are possible. In the former case, if the given relation holds throughout V then again only a finite number of patterns are possible. In either case on such pattern holds outside a complex subvariety of V, and hence on an open dense set.

So we are finally reduced to the case where c_i and c_j both have infinite forward orbits, and the relation $f^a(c_i) = f^b(c_j)$ holds only on a proper subvariety $W \subset V$ containing λ_0. To complete the proof, we will show there is an open set U with $\lambda_0 \in \overline{U}$ such that c_i and c_j have no orbit relations for $\lambda \in U$.

Increasing a and b if necessary, we can assume the point

$$p = f^a_{\lambda_0}(c_i(\lambda_0)) = f^b_{\lambda_0}(c_j(\lambda_0))$$

lies in a periodic component of the Fatou set. If this periodic component is an attracting, superattracting or parabolic basin, we may assume that p lies close to the corresponding periodic cycle.

In the attracting and parabolic cases, we claim there is a ball B containing p such that for all λ sufficiently close to λ_0, the sets

$$\langle f^n_\lambda(B) \,:\, n = 0, 1, 2, \ldots \rangle$$

are disjoint and $f_\lambda^n|B$ is injective for all $n > 0$. This can be verified using the local models of attracting and parabolic cycles. Now for λ near λ_0 but outside the subvariety where $(*)$ holds, $f_\lambda^a(c_i(\lambda))$ and $f_\lambda^b(c_j(\lambda))$ are distinct points in B. Thus $c_i(\lambda)$ and $c_j(\lambda)$ have disjoint forward orbits and we have shown λ_0 is in the closure of the interior of V_{ij}.

In the superattracting case, this argument breaks down because we cannot obtain injectivity of all iterates of f_λ on a neighborhood of p. Instead, we will show that near λ_0, c_i and c_j lie on different leaves of the canonical foliation of the superattracting basin, and therefore have distinct grand orbits.

To make this precise, choose a local coordinate with respect to which the dynamics takes the form $Z \mapsto Z^d$. More precisely, if the period of the superattracting cycle is k, let $Z_\lambda(z)$ be a holomorphic function of (λ, z) in a neighborhood of (λ_0, p), which is a homeomorphism for each fixed λ and which satisfies

$$Z_\lambda(f_\lambda^k(z)) = Z_\lambda^d(z).$$

(The existence of Z follows from classical results on superattracting cycles; cf. [CG, §II.4].) Since $(*)$ does not hold throughout V, there is a neighborhood V of λ_0 on which

$$Z_\lambda(f_\lambda^a(c_i(\lambda))) \neq Z_\lambda(f_\lambda^b(c_j(\lambda)))$$

unless $\lambda = \lambda_0$. Note too that neither quantity vanishes since each critical point has an infinite forward orbit. Shrinking V if necessary we can assume

$$|Z_\lambda(f_\lambda^a(c_i(\lambda)))|^d < |Z_\lambda(f_\lambda^b(c_j(\lambda)))| < |Z_\lambda(f_\lambda^a(c_i(\lambda)))|^{1/d}$$

for $\lambda \in V$. If we remove from V the proper real-analytic subset where

$$|Z_\lambda(f_\lambda^a(c_i(\lambda)))| = |Z_\lambda(f_\lambda^b(c_j(\lambda)))|,$$

we obtain an open subset of V_{ij} with λ_0 in its closure. (Here we use the fact that two points where $\log\log(1/|Z|)$ differs by more than zero and less than $\log d$ cannot be in the same grand orbit.)

Finally we consider the case of a Siegel disk or Herman ring of period k. In this case, for all $\lambda \in V$, the forward orbit of $c_i(\lambda)$ determines a dense subset of $C_i(\lambda)$, a union of k invariant real-analytic circles. This dynamically labeled subset moves injectively as λ varies, so the λ-lemma gives a holomorphic motion

$$\phi : V \times C_i(\lambda_0) \to \widehat{\mathbb{C}},$$

which respects the dynamics. For each fixed λ, $\phi_\lambda(z)$ is a holomorphic function of z as well, since f_λ is holomorphically conjugate to a linear rotation in domain and range.

By assumption, $f_\lambda^b(c_j(\lambda)) \in C_i(\lambda)$ when $\lambda = \lambda_0$. If this relation holds on an open subset V of V, then

$$g(\lambda) = \phi_\lambda^{-1}(f_\lambda^b(c_j(\lambda)))$$

is a holomorphic function on V with values in $C_i(\lambda_0)$, hence constant. It follows that $(*)$ holds on V, and hence on all of V and we are done. Otherwise, $c_i(\lambda)$

and $c_j(\lambda)$ have disjoint forward orbits for all λ outside the proper real-analytic subset of V where $f_\lambda^b(c_j(\lambda)) \in C_i(\lambda)$. ∎

From the proof we have a good qualitative description of the set of points that must be removed to obtain X^{top} from X^{stable}.

Corollary 7.27 *The set X^{stable} is the union of the open dense subset X^{top} and a countable collection of proper complex and real-analytic subvarieties. Thus X^{top} has full measure in X^{stable}.*

The real-analytic part occurs only when f_λ has a foliated region (a Siegel disk, Herman ring or a persistent superattracting basin) for some $\lambda \in X^{stable}$.

Corollary 7.28 *If X is a connected J-stable family of rational maps with no foliated regions, then X^{post} is connected and $\pi_1(X^{post})$ maps surjectively to $\pi_1(X)$.*

Proof. The complement $X - X^{\text{post}}$ is a countable union of proper complex subvarieties, which have real codimension two. ∎

7.8 No invariant line fields

In this section we formulate a conjecture about the ergodic theory of a *single* rational map which implies the density of hyperbolicity.

Let $X = \mathbb{C}/\Lambda$ be a complex torus; then X also has a group structure coming from addition on $\mathbb{C}$. Let $\wp : X \to \widehat{\mathbb{C}}$ be a degree two holomorphic map to the Riemann sphere such that $\wp(-z) = \wp(z)$; such a map is unique up to automorphisms of $\widehat{\mathbb{C}}$ and can be given by the Weierstrass $\wp$-function.

Let $F : X \to X$ be the endomorphism $F(z) = nz$ for some integer $n > 1$. Since $n(-z) = -(nz)$, there is a unique rational map f on the sphere such that the diagram

$$\begin{array}{ccc} \mathbb{C}/\Lambda & \xrightarrow{F} & \mathbb{C}/\Lambda \\ {\scriptstyle\wp}\downarrow & & {\scriptstyle\wp}\downarrow \\ \widehat{\mathbb{C}} & \xrightarrow{f} & \widehat{\mathbb{C}} \end{array}$$

commutes. (Compare [Lat].)

Definition. A rational map f is *double covered by an integral torus endomorphism* if it arises by the above construction.

It is easy to see that repelling periodic points of F are dense on the torus, and therefore the Julia set of f is equal to the whole sphere. Moreover, F and $z \mapsto -z$ preserve the line field tangent to geodesics of a constant slope on X (or more formally, the Beltrami differential $\mu = d\overline{z}/dz$), and therefore f has an invariant line field on $\widehat{\mathbb{C}}$.

Conjecture 7.29 (No invariant line fields) *A rational map f carries no invariant line field on its Julia set, except when f is double covered by an integral torus endomorphism.*

Theorem 7.30 *The no invariant line fields conjecture implies the density of hyperbolic dynamics in the space of all rational maps.*

Proof. Let $X = \mathrm{Rat}_d$ be the space of all rational maps of a fixed degree $d > 1$, and let X^{qc} be the open dense set of quasiconformally stable maps in this universal family. Let $f \in X^{\mathrm{qc}}$. Then $\mathrm{Teich}(\widehat{\mathbb{C}}, f)/\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ parameterizes the component of $X^{\mathrm{qc}}/\mathrm{Aut}(\widehat{\mathbb{C}})$ containing f. Since the modular group is discrete and $\mathrm{Aut}(\widehat{\mathbb{C}})$ acts with finite stabilizers, we have

$$\dim \mathrm{Teich}(\widehat{\mathbb{C}}, f) = \dim \mathrm{Rat}_d - \dim \mathrm{Aut}(\widehat{\mathbb{C}}) = 2d - 2.$$

Clearly f has no indifferent cycles, since by J-stability these would have to persist on an open neighborhood of f in Rat_d and then on all of Rat_d. Therefore f has no Siegel disks or parabolic basins. Similarly f has no periodic critical points, and therefore no superattracting basins. By a Theorem of Mañé, f has no Herman rings [Me]. Thus all stable regions are attracting basins. Finally f is not covered by an integral torus endomorphism because such rational maps form a proper subvariety of Rat_d.

By Theorem 5.41, the dimension of $\mathrm{Teich}(\widehat{\mathbb{C}}, f)$ is given by $n_{AC} + n_{LF}$, the number of grand orbits of acyclic critical points in the Fatou set plus the number of independent line fields on the Julia set. The no invariant line fields conjecture then implies $n_{LF} = 0$, so $n_{AC} = 2d - 2$. Thus all critical points of f lie in the Fatou set and converge to attracting periodic cycles, and therefore f is hyperbolic. ∎

By a similar argument one may establish:

Theorem 7.31 *The no invariant line fields conjecture implies the density of hyperbolic maps in the space of polynomials of any degree.*

Remarks. If f is covered by an integral torus endomorphism, then $\mathrm{Mod}(\widehat{\mathbb{C}}, f)$ contains $\mathrm{PSL}_2(\mathbb{Z})$ with finite index (compare [Her]). It seems likely that the modular group is finite for any other rational map whose Julia set is the sphere. This finiteness would follow from the no invariant line fields conjecture as well, since then $\mathrm{Mod}(\widehat{\mathbb{C}}, f) = \mathrm{Aut}(f)$.

7.9 Centers of hyperbolic components

A rational map f is *critically finite* if every critical point of f has a finite forward orbit.

Theorem 7.32 *Let $U \subset \mathrm{Rat}_d / \mathrm{Aut}(\widehat{\mathbb{C}})$ be a component of the space of expanding rational maps of degree d. Suppose $J(f)$ is connected for $f \in U$. Then there exists a unique critically finite map $f_0 \in U$.*

Compare [Mc1, §3].

8 Iteration on Teichmüller space

In this section we will develop Thurston's topological characterization of critically finite rational maps. Of central interest is the *construction* of a rational map with given combinatorics, via iteration on Teichmüller space. This iteration is similar to, but simpler than, the iteration leading to a hyperbolic structure on an atoroidal Haken 3-manifold.

8.1 Critically finite rational maps

A rational map f is *critically finite* if $|P(f)| < \infty$; that is, if every critical point of f has a finite forward orbit.

Theorem 8.1 *Let f be a critically finite rational map, and A its set of periodic critical points. Then either*

- *$A \neq \emptyset$, the Julia set $J(f)$ has measure zero, and every $z \in \widehat{\mathbb{C}} - J(f)$ is attracted to A; or*
- *$A = \emptyset$, $J(f) = \widehat{\mathbb{C}}$ and the action of f on the sphere is ergodic.*

Proof. Assume $A = \emptyset$. Then $J(f) = \widehat{\mathbb{C}}$ by the classification of stable regions; for example, there is no Siegel disks or Herman rings U since we would have $\partial U \subset P(f)$, and this is impossible because $P(f)$ is finite.

Similarly, all periodic cycles of f are repelling (since an indifferent cycle in $J(f)$ must be a limit point of $P(f)$). Thus under iteration, every critical point of f lands on a repelling cycle. It follows that $\limsup d(f^n z, P(f)) > 0$ for all z outside the grand orbit of the critical points, a countable set. In particular, the forward orbit of almost every $z \in J(f)$ is a definite distance from $P(f)$ infinitely often; that is, $\limsup d(f^n(z), P(f)) > 0$.

Now let $E \subset \widehat{\mathbb{C}}$ be an f-invariant set of positive measure. Let $z \in E$ be a point of Lebesgue density, outside the grand orbit of $P(f)$. Then $\|(f^n)'(z)\| \to \infty$ in the hyperbolic metric on $\widehat{\mathbb{C}} - P(f)$, so small balls about z can be blown up to definite size in the hyperbolic metric. But $f^n(z)$ lands a definite distance from $P(f)$ infinitely often, so these balls can also be blown up to definite spherical size. Taking a limit, we conclude there exists a spherical ball B such that $m(B \cap E) = m(B)$. But then $E = \widehat{\mathbb{C}}$ a.e., since $f^k(B) = \widehat{\mathbb{C}}$ for some k.

The analysis of the case where $A \neq \emptyset$ is similar. ■

8.2 Rigidity of critically finite rational maps

The orbifold of a critically finite map. Let $f : S^2 \to S^2$ be a critically finite rational map. Let $\deg(f, p)$ denote the local degree of f at $p \in S^2$.

For each $p \in S^2$, let

$$N(p) = \operatorname{lcm}_{f^n(q)=p} \deg(f^n, q). \tag{8.1}$$

Note that $N(p) > 1$ iff $p \in P(f)$. The *orbifold of* f, denote $\mathcal{O}_f$, has underlying space S^2 and a singular point of order $N(p)$ at each point in $P(f)$. If $N(p) = \infty$ then we introduce a puncture at p. The list of values of $N(p)$ along $P(f)$ is the *signature* of $\mathcal{O}_f$.

The *Euler characteristic* of the orbifold is given by

$$\chi(\mathcal{O}_f) = \chi(S^2) - \sum_{P(f)} \left(1 - \frac{1}{N(p)}\right).$$

It is easy to prove that $\chi(\mathcal{O}_f) \leq 0$. We say $\mathcal{O}_f$ is *Euclidean* if $\chi(\mathcal{O}_f) = 0$; else $\mathcal{O}_f$ is *hyperbolic*. For example, if $|P(f)| > 4$ then clearly f has a hyperbolic orbifold.

The Euclidean orbifolds are easily classified; the possible signatures are:

$$(\infty, \infty),\ (\infty, 2, 2),\ (2, 2, 2, 2),\ (2, 4, 4),\ (3, 3, 3) \text{ and } (2, 3, 6).$$

The complex structure on the sphere $S^2 \cong \widehat{\mathbb{C}}$ gives $\mathcal{O}_f$ the structure of a one-dimensional *complex* orbifold. Just as for Riemann surfaces, these orbifolds can be uniformized; see [Mc4, Appendix A].

Theorem 8.2 (Uniformization) *A complex orbifold* $\mathcal{O}$ *is covered by* $\mathbb{C}$ *if* $\chi(\mathcal{O}) = 0$*, and by* $\mathbb{H}$ *if* $\chi(\mathcal{O}) < 0$*.*

Corollary 8.3 *Let* f *be a critically finite rational map; then* $\|f'z\| > 1$ *for all* z *with* $f(z) \in \mathcal{O}_f$*.*

Proof. The main case occurs when $\mathcal{O}_f$ is hyperbolic, so we treat this case first. Let $\pi : \mathbb{H} \to \mathcal{O}_f$ be the universal covering map. Condition (8.1) guarantees that for all $q = f(p)$, we have

$$\deg(f, p) N(p) | N(q).$$

Thus f^{-1} can be lifted to the local manifold coverings of $\mathcal{O}_f$ at p and q: that is, for uniformizing parameters z and w near p and q, we have

$$f(z^{N(p)}) = z^{N(p) \deg(f,p)} = w^{N(q)}$$

and therefore

$$z = \widetilde{f^{-1}}(w) = w^{N(q)/(N(p) \deg(f,p))}.$$

Because of this local lifting, on the universal cover we obtain a map $\widetilde{f^{-1}}$ such that the diagram

$$\begin{tikzcd} \mathbb{H} \arrow[d, "\pi"'] & \mathbb{H} \arrow[l, "\widetilde{f^{-1}}"'] \arrow[d, "\pi"'] \\ \widehat{\mathbb{C}} \arrow[r, "f"] & \widehat{\mathbb{C}} \end{tikzcd}$$

commutes. If $\widetilde{f^{-1}}$ were an isometry, then f would be too, contradicting the abundance of repelling cycles. Thus by the Schwarz lemma, $\widetilde{f^{-1}}$ contracts the hyperbolic metric, and consequently $\|f'\| > 1$.

In the Euclidean case one finds $f : \mathcal{O}_f \to \mathcal{O}_f$ is a covering map, and $\|f'\| = \deg(f)$ or $\deg(f)^{1/2}$ depending on whether or not $\mathcal{O}_f$ is compact. ∎

Examples.

1. The orbifold for $f(z) = z^n$ is Euclidean, with signature (∞, ∞).
2. The orbifold for $f(z) = z^2 - 1$ in hyperbolic with signature (∞, ∞, ∞).
3. The map $f(z) = z^2 + i$ has a hyperbolic orbifold with signature $(\infty, 2, 2, 2)$ (here $P(f) = \{\infty, i, -1+i, -i\}$).
4. The Lattès example
$$f(z) = \left(\frac{z-i}{z+i}\right)^2$$
has $P(f) = \{0, \infty, -1, 1\}$ with signature $(2, 2, 2, 2)$ and is thus Euclidean.

Affine maps. The orbifold provides a clean way to isolate exceptional branched covers related to toral endomorphisms. Suppose $\mathcal{O}_f$ has signature $(2, 2, 2, 2)$. Then $\mathcal{O}_f$ is canonically double-covered by a torus T, and f lifts to an endomorphism $F : T \to T$. The action of F on $H_1(T, \mathbb{Z})$ determines an element $A(f) \in GL_2(\mathbb{Z})/\{\pm 1\}$ with $\det(A(f)) = \deg(f)^2$, and F is isotopic to the affine linear map
$$A(f) : \mathbb{R}^2/\mathbb{Z}^2 \to \mathbb{R}^2/\mathbb{Z}^2.$$

If $A(f) = nI \in \mathbb{Z}$, we say f is an *affine* branched cover. Thus:

Theorem 8.4 *A rational map is affine iff it is double-covered by an integral torus endomorphism.*

Here is another nice characterization.

Theorem 8.5 *A rational map f is affine iff $f^*\phi = \deg(f)\phi$ for some nonzero quadratic differential $\phi \in Q(\widehat{\mathbb{C}} - P(f))$.*

Proof. If f is affine, say double-covered by $F(z) = nz$ on $\mathbb{C}/\Lambda$, then $F^*(dz^2) = n^2\, dz^2$ and so the pushforward of dz^2 to $\widehat{\mathbb{C}}$ defines the require quadratic differential ϕ.

For the converse, suppose $f^*\phi = \deg(f)\phi$. For $z \in \widehat{\mathbb{C}}$ let $N(z)$ be the number of leaves of $\mathcal{F}(\phi)$ through z. Then $N(z) = 2$ except at the finitely many zeros and poles of ϕ. Since the foliation of ϕ is f-invariant we have
$$N(f(z)) \deg(f, z) \;\;=\;\; N(z).$$
In particular, $N(z)$ is does not decrease under backwards iteration.

Now $\phi(z) = 0$ iff $N(z) \geq 3$, and thus a zero at z implies a zero along the full inverse orbit of z, which is impossible since ϕ has only finitely many zeros. (If z is exceptional, then ϕ must have a zero of infinite order at z, which is also impossible.) Therefore ϕ has no zeros, and thus $\phi(z)$ has exactly 4 poles. These poles are located in $P(f)$.

Moreover every point in $P(f)$ must be a pole of ϕ. For if $N(z) \geq 2$ and $z \in P(f)$, then $f^n(w) = z$ for some critical point w, and thus $N(w) = \deg(f^n, w)N(z) > 2$, contrary to the fact that ϕ has no zeros.

It now straightforward to see that $f : \mathcal{O}_f \to \mathcal{O}_f$ is a covering map of orbifolds, and thus f lifts to an endomorphism F of the characteristic torus covering $T \to \mathcal{O}_f$. Then F preserves the pullback of ϕ, so F is an integral endomorphism.

■

Theorem 8.6 (Rigidity of rational maps) *Let f and g be topologically conjugate critically finite rational maps. Then either*

- *f and g are conformally conjugate; or*
- *f and g are double-covered by integral torus endomorphisms.*

Proof. Let $Q(f) = f^{-1}(P(f))$ and similarly for $Q(g)$. Let $\phi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a topological conjugacy from f to g; then ϕ sends $P(f)$ and $Q(f)$ to $P(g)$ and $Q(g)$, since these sets are topologically defined. Thus we have a commutative diagram

$$\begin{array}{ccc} \widehat{\mathbb{C}} - Q(f) & \xrightarrow{\phi} & \widehat{\mathbb{C}} - Q(g) \\ {\scriptstyle f}\Big\downarrow & & {\scriptstyle g}\Big\downarrow \\ \widehat{\mathbb{C}} - P(f) & \xrightarrow{\phi} & \widehat{\mathbb{C}} - P(g) \end{array}$$

where the vertical maps f and g are covering maps between multiply-punctured spheres.

Now tighten the lower arrow ϕ as much as possible relative to the post-critical set. That is, deform ϕ to the Teichmüller mapping

$$\psi_0 : \widehat{\mathbb{C}} - P(f) \to \widehat{\mathbb{C}} - P(g)$$

in the homotopy class of ϕ.

Lifting ψ_0 to ψ_1, using the theory of covering spaces, we obtain the diagram

$$\begin{array}{ccc} \widehat{\mathbb{C}} - Q(f) & \xrightarrow{\psi_1} & \widehat{\mathbb{C}} - Q(g) \\ {\scriptstyle f}\Big\downarrow & & {\scriptstyle g}\Big\downarrow \\ \widehat{\mathbb{C}} - P(f) & \xrightarrow{\psi_0} & \widehat{\mathbb{C}} - P(g). \end{array}$$

Since f and g are conformal maps, the dilatation satisfies $K(\psi_1) = K(\psi_0)$. But ψ_0 and ψ_1 are homotopic rel $P(f)$, since they are both homotopic to ϕ (and $Q(f) \supset P(f)$).

By uniqueness of the Teichmüller mapping, we have $\psi_0 = \psi_1$. So we will remove the subscript and simply denote the Teichmüller map by ψ.

If ψ is conformal, then it provides a conformal conjugacy between g and f, so we have proved rigidity.

Now suppose ψ is strictly quasiconformal, and let α be its associated quadratic differential. Then $f^*(\alpha) = \deg(f)\alpha$ because f and ψ commute. More geometrically, the foliations of α are invariant under f, so $f^*\alpha = R\alpha$ for some $R > 0$, and

$$\|f^*\alpha\| = \int_{\widehat{\mathbb{C}}} |f^*\alpha| = \deg(f) \int_{\widehat{\mathbb{C}}} |\alpha| = \deg(f)\|\alpha\|$$

determines $R = \deg(f)$.

By Theorem 8.5, f is affine, so we are done. ■

Corollary 8.7 (The Monotonicity Conjecture) *The topological entropy $h(t)$ of the real quadratic map $f_t(x) = tx(1-x)$ is a monotone function, increasing from 0 to $\log 2$ as t increases from 0 to 4.*

Here the entropy can be defined as

$$h(f_t) = \lim_{n\to\infty} \frac{\log(\text{number of fixed-points of } f_t^n)}{\log n}.$$

The idea of the proof is that if monotonicity fails, then the same finite kneading sequence must occur twice in the quadratic family. But this would contradict the uniqueness of a rational map with a given combinatorial type [MeSt, II.10].

8.3 Branched coverings

Let $F : S^2 \to S^2$ be a smooth map of positive degree. We say F is a *branched cover* if near any point p, we can find smooth charts ϕ, ψ sending p and $F(p)$ to $0 \in \mathbb{C}$, preserving orientation, such that

$$\phi \circ F \circ \psi^{-1}(z) = z^d$$

for some $d \geq 1$.

We wish to recognize rational maps among branched coverings of the sphere. In the absence of dynamics, all branched coverings are representec by rational maps.

Theorem 8.8 (Thom) *Any branched covering F between a pair of spheres is equivalent to a rational map f. That is, there is a rational map $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ and homeomorphisms h_0, h_1 such that the diagram*

$$\begin{array}{ccc} S^2 & \xrightarrow{F} & S^2 \\ {\scriptstyle h_0}\downarrow & & {\scriptstyle h_1}\downarrow \\ \widehat{\mathbb{C}} & \xrightarrow{f} & \widehat{\mathbb{C}} \end{array}$$

commutes.

Proof. Take the standard complex structure on S^2, pull it back by f and apply the uniformization theorem.

In more detail: identify S^2 with $\widehat{\mathbb{C}}$ in the usual way, let $h_1 = \mathrm{id}$ and let h_0 solve the Beltrami equation

$$\frac{(h_0)_{\overline{z}}}{(h_0)_z} = \frac{F_{\overline{z}}}{F_z} = \mu.$$

Note that $\|\mu\|_\infty < 1$ because F is everywhere locally the composition of a holomorphic map and a diffeomorphism. Since h_0 and F have the same complex dilatation, the mapping f making the diagram above commute is holomorphic. ■

For dynamical applications, we want to identify the spheres in the domain and range of F so iteration makes sense. That is, we would want to have $h_0 = h_1$ so that F is *conjugate* to a rational map. Equivalently, we would want to find a complex structure on S^2 that is *preserved* by F.

As one result in this direction we note:

Theorem 8.9 (Sullivan) *A branched cover F is quasiconformally conjugate to a rational map iff the iterates of F are uniformly quasiregular; that is, $K(F^n) \leq K_0 < \infty$.*

Proof. If $F = \phi \circ f \circ \phi^{-1}$ with f rational and ϕ quasiconformal, then $F^n = \phi \circ f^n \circ \phi^{-1}$ and thus $K(F^n) \leq K(\phi)^2 < \infty$ for all n.

For the converse, let for each $z \in \widehat{\mathbb{C}}$ let $\mathbb{H}_z = \mathrm{SL}_2(T_z)/\,\mathrm{SO}_2(T_z)$ denote the hyperbolic plane of conformal structures on the tangent space $T_z\widehat{\mathbb{C}}$. For almost every z, all forward and backward iterates of F are smooth at z, and thus all tangent spaces along the grand orbit of z can be identified with T_z use DF^n. Use this identification to transport the standard conformal structure along the grand orbit to a set of conformal structures $E_z \subset \mathbb{H}_z$. (The set E_z can also be though of as the set of all Beltrami coefficients at z that arise from forward and backwards iterates of f.)

Since F^n is uniformly quasiconformal, the set E_z is bounded. Let $\mu_z \in \mathbb{H}_z$ be the center of the smallest hyperbolic ball containing E_z. Then μ_z is preserved by the dynamics, so it gives an F-invariant measurable complex structure at bounded distance from the standard structure. Solving the Beltrami equation $\phi_{\overline{z}}/\phi_z = \mu$, we obtain a rational map f by setting $f = \phi \circ F \circ \phi^{-1}$. ■

Remark. By the same method one can show a uniformly quasiconformal group Γ acting on $\widehat{\mathbb{C}}$ is quasiconformally conjugate to a subgroup of $\mathrm{PSL}_2(\mathbb{C})$.

On the other hand, Freedman and Skora have constructed a uniformly quasiconformal group $\Gamma \subset \mathrm{Diff}(S^3)$ that is not even topologically conjugate to a Möbius group [FS]. Their proof uses the fact that a finite link of circles $\bigcup C_i \subset S^3$ is unlinked iff every pair of circles is unlinked. For example, the Borromean rings cannot be made out of round rings.

8.4 Combinatorial equivalence and Teichmüller space

Let $f : S^2 \to S^2$ be a branched covering. Just as for a rational map, we say f is *critically finite* if $|P(f)| < \infty$, in which case the orbifold $\mathcal{O}_f$ with singularities or punctures along $P(f)$ is defined. For a branched covering, $\mathcal{O}_f$ is simply a smooth orbifold — it has no complex structure.

Let f and g be critically finite branched covers of the sphere. We say f and g are *combinatorially equivalent* if there are homeomorphisms ϕ_0, ϕ_1 such that

$$\begin{array}{ccc} (S^2, P(f)) & \xrightarrow{\phi_1} & (S^2, P(g)) \\ {\scriptstyle f}\big\downarrow & & {\scriptstyle g}\big\downarrow \\ (S^2, P(f)) & \xrightarrow{\phi_0} & (S^2, P(g)) \end{array}$$

commutes, and ϕ_1 is isotopic to ϕ_0 *rel* $P(f)$. This means there is a continuous family of homeomorphisms

$$\phi_t : (S^2, P(f)) \to (S^2, P(g))$$

connecting ϕ_0 and ϕ_1.

Roughly speaking, two branched coverings are combinatorial equivalent if they are isotopic *rel* their postcritical sets.

If f and g are rational, we can pull ϕ_0 taut (take its Teichmüller representative); then its lift is the taut representative of ϕ_1 *rel* $Q(f)$, but generally it can be *relaxed* by only pinning down its values on $P(f) \subset Q(f)$. Thus the argument proving the Rigidity Theorem 8.6 also shows:

Theorem 8.10 *A critically finite branched covering of the sphere with hyperbolic orbifold is combinatorially equivalent to at most one rational map (up to conformal conjugacy).*

The goal of the remainder of this section is to answer the question:

> *Which branched coverings of the sphere are combinatorial rational maps?*

8.5 Iteration on Teichmüller space

We now show the classification of critically finite rational maps reduces to a fixed-point problem on Teichmüller space.

Convention. For any finite set $A \subset S^2$, we denote by $\mathrm{Teich}(S^2, A)$ the Teichmüller space of the sphere with the points in A marked. This space is the same as $\mathrm{Teich}(S^2 - \mathcal{N}(A))$, where $\mathcal{N}(A)$ is a regular neighborhood of A.

Since there is only one complex structure on S^2, any point in $\mathrm{Teich}(S^2, A)$ is represented by a finite set $B \subset \widehat{\mathbb{C}}$ together with a marking homeomorphism

$$f : (S^2, A) \to (\widehat{\mathbb{C}}, B).$$

The cotangent space is $Q(\widehat{\mathbb{C}}, B) = Q(\widehat{\mathbb{C}} - B)$, the space of meromorphic quadratic differentials on $\widehat{\mathbb{C}}$ with at worst simple poles on B and holomorphic elsewhere.

Iteration. Now let $F : S^2 \to S^2$ be a critically finite branched cover. Starting with a complex structure $(\widehat{\mathbb{C}}, P_0) \in \mathrm{Teich}(S^2, P(F))$, and use the covering

$$F : (S^2, Q(F)) \to (S^2, P(F)),$$

we can form a new Riemann surface

$$F^*(\widehat{\mathbb{C}}, P_0) = (\widehat{\mathbb{C}}, Q_0) \in \mathrm{Teich}(S^2, Q(F))$$

by pulling back the complex structure on $(S^2, P(F))$. The Riemann surface $\widehat{\mathbb{C}} - Q_0$ is just the covering space of $\widehat{\mathbb{C}} - P_0$ dictated by F.

The inclusion

$$I : (S^2, P(F)) \to (S^2, Q(F))$$

permits us to mark a subset of Q_0 by $P(F)$ and obtain a point

$$I^*(\widehat{\mathbb{C}}, Q_0) = (\widehat{\mathbb{C}}, P_1) \in \mathrm{Teich}(S^2, P(F)).$$

The covering $\widehat{\mathbb{C}} - Q_0 \to \widehat{\mathbb{C}} - P_0$ then prolongs to a rational map

$$f_0 : (\widehat{\mathbb{C}}, P_1) \to (\widehat{\mathbb{C}}, P_0).$$

On the level of marked surfaces the composition $I^* \circ F^*$ gives a map

$$T_F : \mathrm{Teich}(S^2, P(F)) \to \mathrm{Teich}(S^2, P(F)).$$

Now suppose $(\widehat{\mathbb{C}}, P_0) = (\widehat{\mathbb{C}}, P_1)$ in $\mathrm{Teich}(S^2, P(F))$. Then after adjusting by a Möbius transformation we can assume $P_0 = P_1$, and thus f_0 is a rational map with $P(f_0) = P_0$. Moreover the marking of domain and range gives a combinatorial equivalence of f_0 to F. The converse is also easy to check, so we have:

Theorem 8.11 *F is combinatorially equivalent to a rational map iff T_F has a fixed-point on Teichmüller space.*

In general, even if we do not start with a fixed-point, we obtain a sequence of marked spheres $(\widehat{\mathbb{C}}, P_i) = T_F^i(\widehat{\mathbb{C}}, P_0)$ and of rational maps f_i such that $f(P_{i+1}) \subset P_i$; that is, we have the tower of maps:

$$\ldots (\widehat{\mathbb{C}}, P_{i+1}) \xrightarrow{f_i} (\widehat{\mathbb{C}}, P_i) \xrightarrow{f_{i-1}} \ldots \xrightarrow{f_1} (\widehat{\mathbb{C}}, P_1) \xrightarrow{f_0} (\widehat{\mathbb{C}}, P_0).$$

If $(\widehat{\mathbb{C}}, P_i)$ converges in Teichmüller space, we will still obtain a fixed-point, and indeed the mappings f_i will converge to the desired rational map f.

Theorem 8.12 (Contraction) *Let $F : S^2 \to S^2$ be a critically finite branched covering. Then either*

- *Some iterate T_F^k contracts the Teichmüller metric; that is,*

$$\|DT_F^k\| < 1$$

at each point of $\mathrm{Teich}(S^2, P(F))$*; or*

- *$\mathcal{O}_F$ has signature $(2,2,2,2)$, and T_F is an isometry.*

Proof. Suppose $T_F(\widehat{\mathbb{C}}, P_0) = (\widehat{\mathbb{C}}, P_1)$. Then we have a rational map

$$f : (\widehat{\mathbb{C}}, P_1) \to (\widehat{\mathbb{C}}, P_0)$$

in the combinatorial class of F. A tangent vector to $\mathrm{Teich}(S^2, P(F))$ at $(\widehat{\mathbb{C}}, P_0)$ is specified by a Beltrami differential μ_0, and

$$DT_F(\mu_0) = \mu_1 = f^*(\mu_0).$$

Of course $\|\mu_0\|_\infty = \|\mu_1\|_\infty$; to compute $\|DT_F\|$, we must use the fact that the tangent space is a *quotient* of the space of Beltrami differentials. To do this, we consider the *coderivative* on quadratic differentials,

$$(DT_F)^* : Q(\widehat{\mathbb{C}} - P_1) \to Q(\widehat{\mathbb{C}} - P_0)$$

which is given by

$$(DT_F)^*(\phi_1) = f_*(\phi_1) = \phi_0.$$

(Note that $\langle \phi_1, f^*\mu_0 \rangle = \langle f_*\phi_1, \mu_0 \rangle$.) Under pushforward the total mass of the area form $|\phi_1|$ cannot increase, but there may be a decrease due to cancellation between corresponding sheets of f. Thus

$$\|(DT_F)^*\| \leq 1.$$

Now suppose $\|DT_F\| = 1$. Then $\|(DT_F)^*\| = 1$. Since $Q(\widehat{\mathbb{C}} - P_1)$ is finite dimensional, its unit ball is compact, and thus there exist $\phi_1 \in Q(\widehat{\mathbb{C}} - P_1)$ and $\phi_0 \in Q(\widehat{\mathbb{C}} - P_0)$ such that $\phi_0 = f_*(\phi_1)$ and $\|\phi_0\| = \|\phi_1\| = 1$. Because there is no cancellation in the pushforward, $\phi_1(z)$ must be a positive real multiple of $\phi_0(w)$ whenever $f(z) = w$. In other words, their foliations agree under f.

Thus $\phi_1/f^*\phi_0$ is a real-valued holomorphic function, hence a constant, and therefore $f^*(\phi_0) = \deg(f)\phi_1$. There are only finitely many possibilities for the poles of ϕ_0 and ϕ_1 as a subset of $P(f)$. Thus upon replacing T_f with a finite iterate, we can assume the singularity structure (zeros and poles) of ϕ_0 is the same as the of ϕ_1. By the same method as that used in the proof of Theorem 8.5, we conclude that $\mathcal{O}_F$ is the $(2,2,2,2)$-orbifold.

Since F preserves a foliation, the matrix $A(F) \in \mathrm{End}(\mathbb{Z}^2)$ giving the action on the homology of the torus is hyperbolic; that is, it has real eigenvalues. It is easy to see that T_F acts on $\mathrm{Teich}(S^2, P(F)) \cong \mathbb{H}$ by the Möbius transformation with matrix $A(F)$; in particular, T_F is an isometry. ∎

Remark. There are examples where T_F is not contracting but an iterate is. For example, suppose F is a mating of two strictly pre-periodic quadratic polynomials. Then there is a set $E \subset P(f)|$ with $|E| = 4$ such that $F^{-1}(E) \subset P(f) \cup C(f)$. Namely we can take E to consist of the critical values together with the first periodic point in each of their forward orbits. A differential with poles only at E has the property that its pullback has poles only on $P(f)$, and then its pushforward is not contracted.

The map T_F is *never* uniformly contracting.

8.6 Thurston's algorithm for real quadratics

Before discussing the general case of T_F, we mention its relation to a practical algorithm for constructing real quadratic polynomials with given kneading sequences.

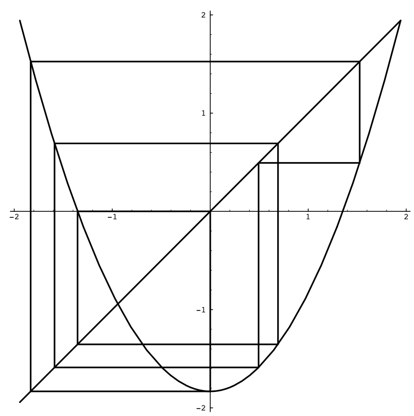

Figure 22. A critically finite quadratic polynomial of period 7.

Let $f : \mathbb{R} \to \mathbb{R}$ be a critically finite polynomial $f(x) = x^2 + c$, $c \in \mathbb{R}$. Then $P(f) \subset \mathbb{R}$ is a finite subset of the line. Furthermore we can write $P(f) = \{p_1, \ldots, p_n\}$ such that $p_i = f^i(0)$. In particular $p_1 = c$. (For ease of notation, we do not include ∞ in $\mathbb{R}$, nor do we include ∞ in $P(f)$.)

Now

$$Q(f) = f^{-1}(P(f)) = \{\pm\sqrt{p_i - p_1}, i = 1, \ldots, n\} \subset \mathbb{R}. \tag{8.2}$$

(Since every point in $P(f)$ has one real preimage, the other preimage is also real.)

The combinatorial type or *kneading sequence of f* determines which points in $Q(f)$ correspond to which points in $P(f)$. That is, if we write

$$Q(f) = \{q_{-n} < q_{-n+1} < \ldots < q_{-1} = 0 = q_1 < q_2 < \ldots < q_n\},$$

then there is a unique index $k(i)$ such that

$$p_i = q_{k(i)}, \ i = 1, \ldots, n. \tag{8.3}$$

Figure 22 shows the example $f(x) = x^2 - 1.83232\ldots$ with $|P(f)| = 7$. In this example k assumes the values $\langle -7, 5, 2, -6, 3, -4, 1\rangle$ on $1, 2, \ldots, 7$.

We also know $P(f) \subset J(f) \subset [-2, 2]$. Using just the kneading data $k(i)$, we can define a map

$$T_k : [-2, 2]^n \to [-2, 2]^n$$

by sending a candidate $(p_1, \ldots, p_n)$ for $P(f)$ to a candidate $Q(f)$ by (8.2), then re-indexing a subset by (8.3) to obtain a new candidate $P(f) \subset [-2, 2]$. (The new candidate lies in $[-2, 2]$ since $|\sqrt{p_i - p_1}| \le \sqrt{4} = 2$.)

If the kneading data k conforms to an actual quadratic polynomial f, then $T_k(P(f)) = P(f)$ so T_k has a fixed-point. Conversely, if T_k fixes $(p_1, \ldots, p_n)$, then $f(x) = x^2 + p_1$ has the kneading sequence k. By the rigidity theorem proved before, we can conclude:

Theorem 8.13 *The map T_k has at most one fixed-point $(p_1, \ldots, p_n)$ consisting of n distinct points.*

On the other hand, by Brouwer's fixed-point theorem, T_k has *at least* one fixed-point in $[-2, 2]^n$. The only problem is that this fixed-point might involve certain p_i coalescing. Indeed, the point $(0, \ldots, 0)$ is *always* fixed by T_k.

Scaling. The iteration on $[-2, 2]^n$ is not quite the same as iteration on Teichmüller space, since the rescaling $(p_i) \mapsto (\alpha p_i)$, $\alpha > 0$, is trivial on $\mathrm{Teich}(S^2, P(f))$. However if $T_k(p_i) = p_i$, then $T_k(\alpha p_i) = (\sqrt{\alpha} p_i)$; since $\alpha^{1/2^n} \to 1$, T_k contracts the rescaling direction.

8.7 Annuli in Euclidean and hyperbolic geometry

In this section we summarize some results on annuli and geodesics to be used below.

Lemma 8.14 *If $f : A \to B$ is a degree d covering map between annuli, then* $\mathrm{mod}(A) = \mathrm{mod}(B)/d$.

Proof. Reduce to the case $A = \{1 < |z| < r\}$, $B = \{1 < |z| < s\}$ and $f(z) = z^d$; then $r = s^{1/d}$. ∎

Lemma 8.15 *Let $\sqcup A_i \subset B$ be a collection of disjoint annuli A_i nested inside in an annulus B (so $\pi_1(A_i) \cong \pi_1(B)\ \forall i$). Then*

$$\mathrm{mod}(B) \ge \sum \mathrm{mod}(A_i).$$

Equality holds iff, in the extremal metric on B, each A_i is a right subcylinder and $\bigcup \overline{A_i} = B$.

Proof. Let Γ be the set of arcs joining the ends of B. Let ρ_i be the extremal metric on A_i, making A_i into a right unit cylinder of height $\operatorname{mod}(A_i)$; and let $\rho = \sum \rho_i$. Then $\operatorname{area}_\rho(B) = 2\pi \sum \operatorname{mod}(A_i)$, and $\ell_\rho(\gamma) \geq \sum \operatorname{mod}(A_i)$ for any arc $\gamma \in \Gamma$, so

$$\operatorname{mod}(B) = 2\pi\lambda(\Gamma) \geq \sum \operatorname{mod}(A).$$

For equality to hold, ρ must be the extremal metric on B, and thus the A_i must form a partition of B into right subcylinders. ■

Lemma 8.16 *Let $\gamma \subset X$ be a closed geodesic on a hyperbolic Riemann surface, and let $X_\gamma \to X$ be the cyclic covering space corresponding to $\langle\gamma\rangle \subset \pi_1(X)$. Then X_γ is an annulus with*

$$\operatorname{mod}(X_\gamma) = \frac{2\pi^2}{\ell_X(\gamma)}.$$

Proof. We have $X_\gamma \cong \mathbb{H}/\langle z \mapsto e^L z\rangle$ where $L = \ell_X(\gamma)$. The invariant metric $|dz|/|z|$ makes X_γ into a right cylinder of height π and circumference L. Rescaling to obtain circumference 2π, the height becomes $2\pi^2/L$. ■

Corollary 8.17 *For any closed curve γ on S, $\ell_\gamma(X)$ is a continuous function on* Teich(S). *In fact*

$$\frac{1}{K}\ell_X(\gamma) \leq \ell_Y(\gamma) \leq K\ell_X(\gamma)$$

if X and Y are related by a K-quasiconformal map.

Proof. A K-quasiconformal map between X and Y lifts to a map between their covering annuli, so $\operatorname{mod}(Y_\gamma)/\operatorname{mod}(X_\gamma) \in [1/K, K]$. ■

Corollary 8.18 *If $\gamma \subset X$ is a simple geodesic, then*

$$\operatorname{mod}(A) \leq \frac{2\pi^2}{\ell_X(\gamma)}$$

for any annulus $A \subset X$ embedded in the homotopy class of γ.

Proof. We can lift A to an annulus nested inside X_γ, so $\operatorname{mod}(A) \leq \operatorname{mod}(X_\gamma)$. ■

Lemma 8.19 *Let $A \subset \mathbb{C}$ be an annulus. Then if* $\mathrm{mod}(A)$ *is large enough, there is a round annulus $B = \{z : a < |z - c| < c\}$ nested in A such that* $\mathrm{mod}(A) = \mathrm{mod}(B) + O(1)$.

Proof. Consider any univalent map

$$f : \{z : R^{-1} < |z| < R\} \to \mathbb{C}.$$

If $R = \infty$ then $f(z) = az + b$. By a normal families argument, once R is large enough, $f(S^1)$ is convex and nearly round. Thus for $A \cong \{z : S^{-1} < |z| < S\}$ with $S \gg R$, the part of A corresponding to $|z| \in [(R^2/S), (S/R^2)]$ is bounded by nearly round curves, and hence A contains a round annulus B with $\mathrm{mod}(B) \geq \mathrm{mod}(A) - 4 \log R$. ■

Corollary 8.20 *Let $\iota : X \hookrightarrow Y$ be an inclusion between hyperbolic Riemann surfaces of genus zero, where X is n-times punctured sphere. Then any short geodesic loop γ on Y satisfies*

$$\frac{1}{\ell_Y(\gamma)} \leq \sum_{\iota(\alpha) \sim \gamma} \frac{1}{\ell_X(\alpha)} + O(1),$$

where the sum is over all short geodesics α on X with $\iota(\alpha)$ homotopic to γ. Here short *means of length less than $\epsilon_n > 0$, and the constant in $O(1)$ also depends on n.*

Proof. Up to isomorphism we can assume X and Y are complements of finite sets $E \supset F$ in $\mathbb{C}$, and ι is the identity map. Then when $\ell_\gamma(Y)$ is sufficiently short, the loop γ is in the same homotopy class as a maximal round annulus

$$B \subset Y = \mathbb{C} - F$$

with

$$\mathrm{mod}(B) = \ell_Y(\gamma)^{-1} + O(1) \gg 1.$$

Now the circles through $E \cap B$ cut B into at most n concentric annuli $A_1, \ldots, A_m$, each in the homotopy class of a geodesic α_i on $X = \mathbb{C} - E$. Then

$$\mathrm{mod}(B) = \sum \mathrm{mod}(A_i) = \sum \frac{1}{\ell_X(\alpha_i)} + O(1).$$

The geodesics which are not short contribute $O(1)$ to the sum, so we have the Corollary. ■

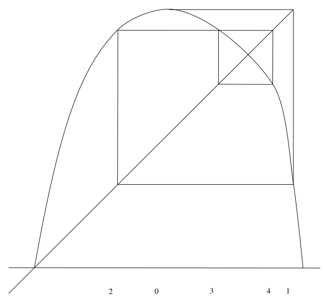

Figure 23. An impossible kneading sequence.

8.8 Invariant curve systems

We begin by detailing a potential *obstruction* to realizing a branched covering F as a rational map.

Consider the combinatorial quadratic map f with $|P(f)| = 4$ depicted in Figure 23. This kneading sequence cannot be realized by a quadratic polynomial.

For a proof, consider the annulus

$$A = (\mathbb{C} - \mathbb{R}) \cup (p_2, p_4) \cup (p_3, p_1) \subset \widehat{\mathbb{C}} - P(f).$$

It is not hard to see that A is the unique annulus of maximum modulus in its homotopy class on $\widehat{\mathbb{C}} - P(f)$. (For a proof note that A lifts to an annulus bounded by geodesics on the flat torus branched over (p_2, p_4, p_3, p_1).) On the other hand, if f is a real quadratic then $f^{-1}(A) = B \cup B'$ will be a pair of disjoint annuli, with $\mathrm{mod}(B) = \mathrm{mod}(B') = \mathrm{mod}(A)$ and with B and A in the same homotopy class on $\widehat{\mathbb{C}} - P(f)$. Clearly $B \neq A$ and this contradicts uniqueness of the annulus of maximum modulus.

The eigenvalue of a curve system. We now formulate a general obstruction.

A simple closed curve $\gamma \subset S^2 - P(F)$ is *essential* if it does not bound a disk, and *nonperipheral* if it does not bound a punctured disk. Two simple curves are *parallel* if they bound an annulus. A *curve system* Γ on $S^2 - P(F)$ is a nonempty collection of *disjoint* simple closed curves γ on $S^2 - P(F)$, each essential and nonperipheral and with no pair of curves parallel. By Euler characteristic considerations we have $|\Gamma| \leq |P(F)| - 3$.

A curve system determines a *transition matrix* $M(\Gamma) : \mathbb{R}^\Gamma \to \mathbb{R}^\Gamma$ by

$$M_{\gamma\delta} = \sum_\alpha \frac{1}{\deg(f : \alpha \to \delta)}$$

where the sum is taken over components α of $f^{-1}(\delta)$ which are isotopic to γ.

Let $\lambda(\Gamma) \geq 0$ denote the spectral radius of $M(\Gamma)$. By the Perron-Frobenius theorem, $\lambda(\Gamma)$ is an eigenvalue for $M(\Gamma)$ with non-negative eigenvector.

Theorem 8.21 *Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a critically finite rational map. Then either:*

- *f is affine, and $\lambda(\Gamma) = 1$ for any curve system; or*
- *f is not affine, and $\lambda(\Gamma) < 1$ for any curve system on $\widehat{\mathbb{C}} - P(f)$.*

Proof. Let $\langle A_\gamma \rangle$ be any collection disjoint annuli in the homotopy classes represented by Γ on $\widehat{\mathbb{C}} - P(f)$. Fixing $\gamma \in \Gamma$, consider the curves $\{\alpha\} \subset f^{-1}(\bigcup \Gamma)$ with α isotopic to γ. For each such α, with $f(\alpha) = \delta \in \Gamma$, there is a corresponding annulus $B_\alpha \subset f^{-1}(A_\delta)$ with

$$\operatorname{mod}(B_\alpha) = \frac{\operatorname{mod}(A_\delta)}{\deg(f : \alpha \to \delta)}.$$

These B_α are disjoint and nested, so we can form a single annulus $A'_\gamma \supset \bigcup B_\alpha$ in the same homotopy class. Then

$$\operatorname{mod}(A'_\gamma) \geq \sum_\alpha \operatorname{mod}(B_\alpha) = \sum_\delta M_{\gamma\delta} \operatorname{mod}(A_\delta).$$

Now suppose $\lambda(\Gamma) > 1$. Starting with any system of annuli $\mathcal{A}^0 = \langle A^0_\gamma \rangle$, by pulling back and regrouping as above we obtain systems $\mathcal{A}^i$ with the vector of moduli satisfying

$$\operatorname{mod}(\mathcal{A}^i) \geq M(\Gamma)^i \operatorname{mod}(\mathcal{A}^0).$$

If $\lambda(\Gamma) > 1$ then this vector grows without bound, contrary to the fact that $\operatorname{mod}(\mathcal{A}^i_\gamma)$ is bounded above in terms of the hyperbolic length of γ.

Thus $\lambda(\Gamma) \leq 1$ for any rational map f.

In case $\lambda(\Gamma) = 1$, there is a vector of moduli (m_γ) fixed by $M(\Gamma)$. Using a result of Strebel, one finds the extremal metrics for annuli of maximum moduli proportional to (m_γ) piece together to give a metric $|\phi|$ from a holomorphic quadratic differential [Str2]. Because of equality the pullback annuli must dissect the extremal ones into right cylinders, we find $f^*\phi = \deg(f)\phi$, and thus f is affine.

So in fact $\lambda(\Gamma) < 1$ unless f is affine. For an affine map, $|P(f)| = 4$ and thus Γ consists of a single curve γ. This curve has n pre-images, each mapping by degree n, as can be seen by lifting γ to a geodesic on the torus double cover of $\widehat{\mathbb{C}}$. Thus $M(\Gamma) = 1$. ∎

8.9 Characterization of rational maps

Lemma 8.22 *The moduli space $\mathcal{M}_{0,n}$ is an algebraic variety.*

Proof. Let $E \subset \mathbb{C}$ be a set of $n-1$ distinct points. Normalizing by a translation we can assume $\sum E = 0$. Then there is a unique monic polynomial

$$p(z) = z^{n-1} + a_1 z^{n-3} + \ldots + a_{n-2}$$

whose roots coincide with E. Conversely, if the discriminant $D(p) \neq 0$ then p has $n-1$ distinct roots. If we send E to λE, $\lambda \in \mathbb{C}^*$, then $a_i \mapsto \lambda^{i+1} a_i$. Modulo this action, the coefficients of p determine a point $[p]$ in a weighted projective space $P = \mathbb{P}^{n-3}_w$. The discriminant vanishes on a subvariety $D \subset P$, and we see:

$$\begin{aligned} P - D &\cong (n-1\text{-tuples } E \subset \mathbb{C}) / \operatorname{Aut}(\mathbb{C}) \\ &\cong (n\text{-tuples } F \subset \widehat{\mathbb{C}} \text{ with one distinguished point}) / \operatorname{Aut}(\widehat{\mathbb{C}}). \end{aligned}$$

By forgetting which point is distinguished, we obtain an n-to-1 covering map $P - D \to \mathcal{M}_{0,n}$, and thus the moduli space of n-tuples of distinct points on the sphere is an algebraic variety (indeed an affine variety). ■

Remark. In fact $\mathcal{M}_{g,n}$ is a quasiprojective variety for every (g, n). It is generally not affine.

Lemma 8.23 *The graph of T_F covers an algebraic subvariety*

$$R \subset \mathcal{M}_{0,n} \times \mathcal{M}_{0,n},$$

$n = |P(F)|$. *The variety R is the graph of a correspondence, and $|R(X)|$ is finite for every $X \in \mathcal{M}_{0,n}$. In fact*

$$|R(X)| \leq \binom{nd}{n} (d!)^{n-1},$$

where $d = \deg(F)$.

Proof. We regard points $X \in \mathcal{M}_{0,n}$ as Riemann surfaces of genus 0 with n points removed. Let $V_d \subset \mathcal{M}_{0,n}$ consists of those pairs (X_0, X_1) such that X_1 contains an embedded degree d (connected) covering of X_0. By general principles, V_d is an algebraic subvariety, and clearly the graph of T_F covers a component R of V_d. Thus R itself is a subvariety, in fact an irreducible component of V_d.

To bound $R(X_0)$, note first that a covering space $Y \to X_0$ of degree d is specified by a map $\pi_1(X_0) \to S_d$. Since $\pi_1(X_0)$ is a free group on $n-1$ generators, there are $(d!)^{n-1}$ such maps. Each Y has at most nd punctures, so Y is contained in at most $\binom{nd}{n}$ surfaces X_0 with n punctures. ■

Lemma 8.24 *We have $\|DT_F(X)\| < C([X]) < 1$ where $C([X])$ is a continuous function depending only on the location $[X]$ of X in the moduli space $\mathcal{M}_{0,n}$, $n = |P(F)|$.*

Proof. As in the preceding lemma, given $(\widehat{\mathbb{C}}, P_0)$ representing a point in $\mathcal{M}_{0,n}$, there are only finitely maps

$$f_0 : (\widehat{\mathbb{C}}, P_1) \to (\widehat{\mathbb{C}}, P_0)$$

that arise from T_F. Since $\|DT_F\| = \|f_*\| < 1$ for each of these maps, we obtain a bound depending only on the location of $(\widehat{\mathbb{C}}, P_0)$ in $\mathcal{M}_{0,n}$. ■

Remark. Alternatively, let $R_0 \subset R$ denote the component of R covered by the graph of T_F. Then at any point $([X],[Y]) \in R_0$, the derivative DR_0 of the local correspondence in the Teichmüller metric on $\mathcal{M}_{0,n}$ is well-defined, and agrees with the derivative of T_F over this pair. Thus $\|DT_F(X_i)\| \to 1$ implies $([X_i],[T_F(X_i)]) \to \infty$ in R_0, which implies $X_i \to \infty$ in $\mathcal{M}_{0,n}$.

Theorem 8.25 (Thurston) *Let $F : S^2 \to S^2$ be a critically finite branched covering. Then F is combinatorially equivalent to a rational map f if and only if:*

(Torus case) $\mathcal{O}_F$ has signature $(2,2,2,2)$ and $T_F : \mathbb{H} \to \mathbb{H}$ is an elliptic Möbius transformation; or

(General case) $\mathcal{O}_F$ does not have signature $(2,2,2,2)$ and

$$\lambda(\Gamma) < 1$$

for every F-invariant curve system Γ on $S^2 - P(F)$.

In the second case f is unique up to conformal conjugacy.

Proof. If $\mathcal{O}_F$ has signature $(2,2,2,2)$, then T_F is an isometry of $\mathrm{Teich}(S^2, P(F)) \cong \mathbb{H}$, and F is rational $\iff$ T_F has a fixed-point $\iff$ T_F is elliptic (or the identity). So we may suppose $\mathcal{O}_F$ does not have signature $(2,2,2,2)$.

If F is equivalent to a rational map f, then $\lambda(\Gamma) < 1$ for every invariant curve system, by Theorem 8.21. Uniqueness of f follows from contraction of T_F.

For the converse, assume F is *not* equivalent to any rational map. Then F has no fixed point on Teichmüller space. To complete the proof, we will show

$$\lambda(\Gamma) \geq 1$$

for some curve system Γ on $S^2 - P(F)$

To construct Γ, pick any $X_0 \in \mathrm{Teich}(S^2, P(F))$, and set

$$X_i = T_F^i(X_0).$$

Then $\|DT_F(X_i)| < C(X_i) < 1$ where $C(X_i)$ depends only on the location of $[X_i]$ in moduli space $\mathcal{M}_{0,n}$. Let

$$p_0 \subset \mathrm{Teich}(S^2, P(F))$$

be the geodesic segment joining X_0 to X_1, and let $p_i = T_F^i(p_0)$. If $[X_i]$ is contained in a compact subset of $\mathcal{M}_{0,n}$, then so is $[\bigcup p_i]$; then T_F is uniformly contracting on $\bigcup p_i$, so X_i converges to a fixed-point, contrary to assumption.

Therefore $[X_i]$ eventually leaves any compact subset of $\mathcal{M}_{0,n}$ (although it may return). By Mumford's Theorem 4.18, this means for any $\epsilon > 0$ there is an i such that

$$L(X_i) = (\text{length of the shortest geodesic on } X_i) < \epsilon.$$

We will now show that for suitable $\epsilon > 0$, the short curves Γ on X_i satisfy

$$[\ell_\gamma(X_{i+1})^{-1}] \approx M(\Gamma)[\ell_\gamma(X_i)^{-1}],$$

and thus $M(\Gamma)$ gives a linear approximation to the action of T_F. If $\lambda(\Gamma) < 1$, then $M(\Gamma)$ is contracting and the short geodesics are forced to become long again. Since $\liminf L(X_i) = 0$, we will eventually find a Γ with $\lambda(\Gamma) \geq 1$ and this will complete the proof.

To choose ϵ, first pick an $\epsilon_0 > 0$ such that ϵ_0-short curves on any $X \in \mathrm{Teich}(S^2, P(F))$ are disjoint simple geodesics, and such that Corollary 8.20 holds for ϵ_0-short curves on an nd-times punctured sphere. Next let K be the dilatation of the extremal quasiconformal map from X_0 to X_1. Finally pick $M \gg \max(K^2, n)$ whose exact size will be fixed later, and set

$$\epsilon = \epsilon_0 / M^n.$$

Now consider any X_i with $L(X_i) < \epsilon$. Since X_i has at most $(n-3)$ geodesics of length less than ϵ_0, there must be a $\delta \in [\epsilon, \epsilon_0]$ so the length spectrum of X_i avoids the interval $[\delta, M\delta]$. Let Γ denote the nonempty set of geodesic loops on X_i of length less than δ. Then we have

$$\ell_\gamma(X_i) < \frac{\ell_\delta(X_i)}{M} \ll \ell_\delta(X_i)$$

for any geodesic $\delta \notin \Gamma$.

We will now refer to geodesics of length less than $\sqrt{M}\delta$ as short. Since $d(X_i, X_{i+1}) \leq d(X_1, X_0)$, there is a K-quasiconformal map from X_i to X_{i+1}, and hence the lengths of geodesics on X_i and X_{i+1} agree to within a factor of K. Since $\sqrt{M} \geq K$, the length spectrum of X_i avoids $[\delta, K^2\delta]$, so Γ also represents the set of all short geodesics on X_{i+1}.

On the other hand, we have a diagram

$$\begin{array}{ccc} Y_i & \xrightarrow{\iota} & X_{i+1} \\ {\scriptstyle f}\big\downarrow & & \\ X_i & & \end{array}$$

where f is a degree $d = \deg(F)$ rational map and ι is an inclusion of genus zero Riemann surfaces. By Corollary 8.20, for each $\gamma \in \Gamma$ we have

$$\frac{1}{\ell_\gamma(X_{i+1})} \leq \sum_{\iota(\alpha) \sim \gamma} \frac{1}{\ell_\alpha(Y_i)} + O(1),$$

where the sum includes only curves α with $\ell_{Y_i}(\alpha) \leq \epsilon_{nd}$ (a constant depending only on the number of punctures of Y_i). The loops α which cover loops Γ on X_i contribute most to the sum, since any other loop is at least M/d times longer, and there are less than n other loops. Thus only a small error is committed if we leave out these other loops, and retain α only if it covers some $\delta \in \Gamma$. More precisely we have:

$$\frac{1}{\ell_\gamma(X_{i+1})} \leq \left(1 + \frac{nd^2}{M}\right) \sum_{\iota(\alpha)\sim\gamma, f(\alpha)=\delta\in\Gamma} \frac{1}{\deg(f : \alpha \to \delta)} \cdot \frac{1}{\ell_\alpha(X_i)} + O(1).$$

The vector of inverse lengths

$$v_i = [\ell_\gamma(X_i)^{-1}]$$

therefore satisfies

$$v_{i+1} \leq (1+\eta)M(\Gamma)v_i + O(1) \tag{8.4}$$

where $\eta \to 0$ as $M \to \infty$.

Although Γ ranges among the countably many different collections of curve systems on $S^2 - P(F)$, since $n = |P(F)|$ and $d = \deg(F)$ are fixed, there are only a finite number of possibilities for the matrix $M(\Gamma)$. Thus there is a λ_0 such that $\lambda(\Gamma) \geq 1$ or $\lambda(\Gamma) < \lambda_0 < 1$. Since

$$\lambda(\Gamma) = \lim \|M(\Gamma)^p\|^{1/p},$$

we can choose p such that $\|M(\Gamma)^p\| < 1/2$ whenever $\lambda(\Gamma) < 1$.

Now if $\|v_i\|$ is large enough, we have $L(X_i) \leq \epsilon/K^p$ and thus $L(X_j) \leq \epsilon$ for $j = i, i+1, \ldots, i+p$. For $M \geq K^{2p}$, the short geodesics Γ on X_i agree with the short geodesics on X_j for $i \leq j \leq i+p$. Finally if $\lambda(\Gamma) < 1$ then by (8.4) we find

$$\|v_{i+p}\| \leq \frac{(1+\eta)^p}{2}\|v_i\| + O(1) \leq \frac{2}{3}\|v_i\|$$

once M is large enough that η is negligible.

In other words, if $\|v_i\|$ gets large then it is forced to get small again within p iterates. Noting that $\|v_{i+1}\| \leq K\|v_i\|$, we conclude that $\sup_i \|v_i\| < \infty$. But $\|v_i\| \geq 1/L(X_i)$, so we have contradicted the fact that $\liminf L(X_i) = 0$.

Thus to have $\|v_i\| \to \infty$, we must at some point obtain a curve system Γ with

$$\lambda(\Gamma) \geq 1.$$

This Γ is the desired topological obstruction to realizing F as a rational map. ■

8.10 Notes

Thom's result is in [Thom]; Douady and Hubbard present Thurston's characterization of combinatorially rational branched covers in [DH]. For more on the orbifold approach to critically finite rational maps, see [Mc4, Appendix A and B].

{ 2,0,4,3,1}	3
{ 2,0,3,4,1}	3
{ 2,3,0,5,4,1}	4
{ 2,3,0,4,5,1}	4
{ 2,3,4,0,6,5,1}	5
{ 2,3,4,0,5,6,1}	5
{ 2,5,3,0,6,4,1}	3
{ 2,3,5,0,4,6,1}	3
{ 2,3,4,5,0,7,6,1}	6
{ 2,3,4,5,0,6,7,1}	6
{ 2,3,6,4,0,7,5,1}	4
{ 2,3,4,6,0,5,7,1}	4
{ 2,5,3,0,7,6,4,1}	6
{ 2,5,3,0,6,7,4,1}	6
{ 2,5,7,0,3,6,4,1}	4
{ 2,5,0,3,7,6,4,1}	6
{ 2,5,0,3,6,7,4,1}	6
{ 2,7,5,0,3,4,6,1}	4

Table 24. Impossible kneading sequences.

8.11 Appendix: Kneading sequences for real quadratics

Notation. The post-critical set is labeled with the integers $0 =$ critical point, $n = f^n$(critical point). To describe a real unimodal critically finite map, we list the postcritical set in the order it appears on the real line, and give a single digit indicating the image of the point with the highest label. For example, $2, 0, 1\ 2$ indicates that 2, the image of the critical value, is a fixed point.

The following kneading sequences cannot be realized by quadratic polynomials. In each case there is a finite collection of disjoint intervals I_k, with endpoints in the post-critical set, permuted homeomorphically by the mapping.

The remaining kneading sequences can be realized as quadratic polynomials. This table gives the kneading sequence and the value of t such that $f(x) = tx(1-x)$ has the same kneading behavior.

{ 0}	0	2.
{ 0,1}	0	3.23607
{ 2,0,1}	2	4.
{ 2,0,1}	0	3.83187
{ 2,3,0,1}	0	3.96027
{ 2,0,3,1}	0	3.49856
{ 2,0,3,1}	3	3.67857
{ 2,3,4,0,1}	0	3.99027
{ 2,3,0,4,1}	3	3.89087
{ 2,3,0,4,1}	0	3.90571
{ 2,3,0,4,1}	4	3.92774
{ 2,0,3,4,1}	0	3.73891
{ 2,3,4,5,0,1}	0	3.99758
{ 2,3,4,0,5,1}	3	3.96792
{ 2,3,4,0,5,1}	4	3.97459
{ 2,3,4,0,5,1}	0	3.97777
{ 2,3,4,0,5,1}	5	3.98257
{ 2,5,3,0,4,1}	0	3.84457
{ 2,3,0,4,5,1}	3	3.94935
{ 2,3,0,4,5,1}	0	3.93754
{ 2,0,4,3,5,1}	0	3.62756
{ 2,5,0,3,4,1}	4	3.76495
{ 2,3,4,5,6,0,1}	0	3.9994
{ 2,3,4,5,0,6,1}	3	3.99125
{ 2,3,4,5,0,6,1}	4	3.99228
{ 2,3,4,5,0,6,1}	5	3.99378
{ 2,3,4,5,0,6,1}	0	3.99454
{ 2,3,4,5,0,6,1}	6	3.99569
{ 2,3,6,4,0,5,1}	0	3.96898
{ 2,3,4,0,5,6,1}	3	3.9891
{ 2,3,4,0,5,6,1}	4	3.98734
{ 2,3,4,0,5,6,1}	0	3.98475
{ 2,5,3,0,6,4,1}	0	3.88605
{ 2,5,3,0,6,4,1}	6	3.87654
{ 2,5,3,0,6,4,1}	4	3.8568
{ 2,3,0,5,4,6,1}	3	3.91694
{ 2,3,0,5,4,6,1}	0	3.92219
{ 2,6,3,0,4,5,1}	0	3.95103
{ 2,3,6,0,4,5,1}	5	3.94282
{ 2,6,0,4,3,5,1}	5	3.59257
{ 2,5,0,3,6,4,1}	0	3.77421
{ 2,5,0,3,6,4,1}	3	3.78088
{ 2,5,0,3,6,4,1}	6	3.7911
{ 2,0,5,3,4,6,1}	0	3.70177
{ 2,3,4,5,6,7,0,1}	0	3.99985
{ 2,3,4,5,6,0,7,1}	3	3.99771
{ 2,3,4,5,6,0,7,1}	4	3.99785
{ 2,3,4,5,6,0,7,1}	5	3.99809
{ 2,3,4,5,6,0,7,1}	6	3.99845
{ 2,3,4,5,6,0,7,1}	0	3.99864
{ 2,3,4,5,6,0,7,1}	7	3.99893
{ 2,3,4,7,5,0,6,1}	0	3.99252
{ 2,3,4,5,0,6,7,1}	3	3.99745
{ 2,3,4,5,0,6,7,1}	4	3.99726
{ 2,3,4,5,0,6,7,1}	5	3.99685
{ 2,3,4,5,0,6,7,1}	0	3.99622
{ 2,6,3,7,4,0,5,1}	0	3.96094
{ 2,3,6,4,0,7,5,1}	0	3.97372
{ 2,3,6,4,0,7,5,1}	7	3.97221
{ 2,3,6,4,0,7,5,1}	5	3.97032
{ 2,3,4,6,0,5,7,1}	3	3.97643
{ 2,3,4,0,6,5,7,1}	3	3.97898
{ 2,3,4,0,6,5,7,1}	4	3.98034
{ 2,3,4,0,6,5,7,1}	0	3.98141
{ 2,3,7,4,0,5,6,1}	0	3.98775
{ 2,3,4,7,0,5,6,1}	6	3.98597
{ 2,5,3,0,7,6,4,1}	4	3.88184
{ 2,5,3,0,6,7,4,1}	3	3.86828
{ 2,5,3,0,6,7,4,1}	0	3.87054
{ 2,7,3,5,0,4,6,1}	5	3.9006
{ 2,7,3,5,0,4,6,1}	0	3.89947
{ 2,7,3,0,5,4,6,1}	0	3.91205
{ 2,7,3,0,5,4,6,1}	5	3.91346
{ 2,3,7,0,5,4,6,1}	6	3.91879
{ 2,6,3,7,0,4,5,1}	5	3.9523
{ 2,3,6,0,4,7,5,1}	0	3.94421
{ 2,3,6,0,4,7,5,1}	4	3.94539
{ 2,3,6,0,4,7,5,1}	7	3.94655
{ 2,3,6,0,4,5,7,1}	3	3.94126
{ 2,3,0,6,4,5,7,1}	3	3.93342
{ 2,3,0,6,4,5,7,1}	0	3.93047
{ 2,6,0,4,3,7,5,1}	0	3.55465
{ 2,0,6,4,3,5,7,1}	0	3.66219
{ 2,5,0,3,6,7,4,1}	0	3.80077
{ 2,5,0,3,6,7,4,1}	3	3.79812
{ 2,7,0,5,3,4,6,1}	4	3.72029
{ 2,7,0,5,3,4,6,1}	6	3.71255

Table 25. Quadratic kneading sequences.

9 Geometrization of 3-manifolds

9.1 Topology of hyperbolic manifolds

Let M be a closed orientable smooth 3-manifold. When can we expect M to carry a hyperbolic structure? That is, when can we hope to write $M = \mathbb{H}^3/\Gamma$? Here are some obvious constraints:

- First, M must be irreducible — that is, every embedded $S^2 \subset M$ must bound a ball $B^3 \subset M$. To see this, just lift to the universal cover: we find a collection $\Gamma \cdot S^2$ of disjoint 2-spheres in $\mathbb{H}^3$, and the balls $\Gamma \cdot B^3$ they bound are disjoint, so they determine an embedded ball in $\mathbb{H}^3$.

  Irreducibility means if $M = A \# B$ then A or B is S^3.

- Second, $\pi_1(M)$ must be infinite — since the universal cover of M is not compact.

- Finally, M must be atoroidal: that is, $\pi_1(M)$ cannot contain a copy of $\mathbb{Z} \oplus \mathbb{Z}$.

  For a proof, note that the centralizer of a hyperbolic element in $\mathrm{Isom}^+(\mathbb{H}^3)$ is isomorphic to $\mathbb{C}^* = S^1 \times \mathbb{R}$. If $G \subset \mathbb{C}^*$ is a discrete, torsion-free group, then G maps injectively by absolute value to a discrete subgroup of $\mathbb{R}^*$, and hence G is trivial or $\mathbb{Z}$.

Conjecture 9.1 *A closed, irreducible orientable 3-manifold is hyperbolic iff $\pi_1(M)$ is infinite and does not contain $\mathbb{Z} \oplus \mathbb{Z}$.*

Now suppose M is an oriented manifold with boundary, admitting a convex hyperbolic structure. Let $\tau : \partial M \to \partial M$ be a fixed-point free, orientation-reversing involution. Then M/τ is also an orientable manifold. We will allow M to be disconnected but require M/τ to be connected.

When can we expect the closed manifold M/τ to be hyperbolic? Let us apply the criteria above.

1. First, M/τ must be irreducible. There is danger that a pair of properly embedded 2-disks $D_1 \sqcup D^2 \subset M$ might get glued together by τ to form a 2-sphere. If ∂D_i is nontrivial in $\pi_1(\partial M)$, then $D_1 \cup D_2$ cannot bound a ball and so M/τ is reducible.

   To guard against this possibility we will assume ∂M is incompressible. This means
   $$\pi_1(\partial M, *) \hookrightarrow \pi_1(M, *)$$
   is injective for every choice of basepoint. Then M/τ is guaranteed to be irreducible.

2. Since M is hyperbolic, each of its boundary components has genus $g \geq 2$, and by van Kampen's theorem $\pi_1(\partial M/\tau)$ injects into $\pi_1(M)$. So $\pi_1(M)$ is always infinite.

3. Finally M/τ must be atoroidal. Here the danger is that M might contain two or more cylinders C_i — properly embedded copies of $S^1 \times [0,1]$ — which are glued together by τ to form an incompressible torus in M.

   To guard against this possibility, one can require that M is *acylindrical* — i.e., that every cylinder $C : S^1 \times [0,1] \to M$ is homotopic, rel boundary, into ∂M.

   More liberally, one can allow cylinders but require that they not be glued together. That is, we can simply require that M/τ is atoroidal.

Theorem 9.2 (Thurston) *Let M admit a convex hyperbolic structure. Then M/τ is hyperbolic iff it is atoroidal.*

9.2 The skinning map

Let M be an oriented compact 3-manifold with incompressible boundary, admitting a convex hyperbolic structure. Let $\tau : M \to M$ be an orientation-reversing fixed-point free involution.

The *skinning map*

$$\sigma : \mathrm{Teich}(\partial M) \cong GF(M) \to \mathrm{Teich}(\overline{\partial M})$$

is defined as follows. Given a Riemann surface $X \in \mathrm{Teich}(\partial M)$, construct the unique complete hyperbolic 3-manifold $N \in GF(M)$ with $\partial N \cong X$. Next, form the covering space $Q \to N$ corresponding to ∂N. Topologically we have:

$$Q \cong (\partial N) \times \mathbb{R}.$$

The manifold Q is convex cocompact, and naturally marked by ∂M, so we have

$$Q \in GF(\partial M \times [0,1]) \cong \mathrm{Teich}(\partial M) \times \mathrm{Teich}(\overline{\partial M}).$$

The projection to the second factor above is the skinning map. That is,

$$\partial Q = \partial N \sqcup \sigma(\partial N).$$

Theorem 9.3 *The manifold M/τ is hyperbolic if*

$$\tau \circ \sigma : \mathrm{Teich}(\partial M) \to \mathrm{Teich}(\partial M)$$

has a fixed-point.

Remark. This theorem is not quite iff. If $M = S \times [0,1]$, then σ is the reflection map, and M/τ is hyperbolic iff τ comes from a pseudo-Anosov automorphism of S. But in the pseudo-Anosov case, $\tau \circ \sigma$ has no fixed-point in Teichmüller space. Indeed, $\tau \circ \sigma$ has a fixed-point iff M/τ is finitely covered by $S \times S^1$ — in which M admits an $\mathbb{H}^2 \times \mathbb{R}$ structure but no $\mathbb{H}^3$ structure.

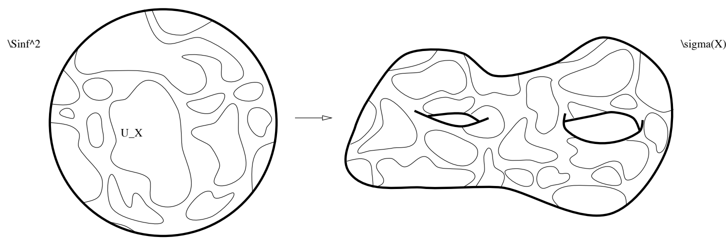

Figure 26. The skinning map.

9.3 The Theta conjecture

Theorem 9.4 *Let $X = \Delta/\Gamma$ be a finite-area hyperbolic surface presented as the quotient of the unit disk by a Fuchsian group. Then the natural map*

$$\theta : \mathrm{Teich}(X) \to \mathrm{Teich}(\Delta)$$

is a contraction for the Teichmüller metric. In fact

$$\|d\theta^*\| = \|\Theta_{\Delta/X}\| < C(X) < 1$$

where $C(X)$ depends only on the location of X in moduli space, and

$$\Theta_{\Delta/X} : Q(\Delta) \to Q(X)$$

is the Poincaré series operator given by

$$\Theta_{\Delta/X}(\phi) = \sum_{\Gamma} \gamma^* \phi.$$

Corollary 9.5 *Let $f : X \to Y$ be a Teichmüller mapping between two points in* $\mathrm{Teich}(S)$. *Then the lift $\widetilde{f} : \Delta \to \Delta$ to the universal covers is not extremal among all quasiconformal maps with the same boundary values, unless f is conformal.*

Example [Str1]. Let $f : X \to Y$ be an affine stretch from a square torus to a rectangular torus, respecting the origin. The punctured surfaces X^* and Y^* obtained by removing the origin are hyperbolic, and f restricts to a Teichmüller mapping between them.

The universal cover of X^* (Y^*) can be thought of as a countable collection of squares (rectangles) with their vertices removed, and glued together along their edges in the pattern of a free group $Z * Z$ to form a thickened tree. The lifted map $\widetilde{f}$ sends each square to the corresponding rectangle.

To make $\widetilde{f}$ more nearly conformal, we want to make the target more nearly square. A single rectangle R_0 can be made more square (its modulus can be moved towards 1) by bending its short edges in and its long edges out. Now one edge E_i on each adjacent rectangle R_i has been bent the wrong way. The remaining 3 edges, however, can be bent the right way, resulting in an overall improvement in shape (Figure 27). The next layer of 12 rectangles each have only one edge committed, and similarly for each following layer of $4 \cdot 3^n$ rectangles... so we can continue modifying $\widetilde{f}$ to reduce its dilatation on each square, in the end obtaining a map g with the same boundary values but $K(g) < K(f)$.

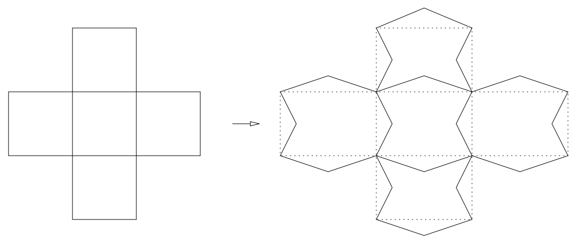

Figure 27. Strebel's idea of relaxation.

Corollary 9.6 *If M is an acylindrical manifold, then the skinning map*

$$\sigma_M : \mathrm{Teich}(\partial M) \to \mathrm{Teich}(\overline{\partial M})$$

satisfies

$$\|d\sigma_M(X)\| \leq C(X) < 1$$

with $C(\cdot)$ as above.

References

[AS] R. Abrahams and S. Smale. Nongenericity of Ω-stability. In *Global Analysis*, pages 5–8. Am. Math. Soc., 1970.

[Ah1] L. Ahlfors. Finitely generated Kleinian groups. *Amer. J. of Math.* **86**(1964), 413–429.

[Ah2] L. Ahlfors. *Conformal Invariants: Topics in Geometric Function Theory.* McGraw-Hill Book Co., 1973.

[And] G. Anderson. *Projective Structures on Riemann Surfaces and Developing Maps to $\mathbb{H}^3$ and $\mathbb{CP}^n$.* Mem. Amer. Math. Soc., To appear.

[Bers1] L. Bers. Inequalities for finitely generated Kleinian groups. *J. d'Analyse Math.* **18**(1967), 23–41.

[Bers2] L. Bers. An extremal problem for quasiconformal maps and a theorem by Thurston. *Acta Math.* **141**(1978), 73–98.

[Bers3] L. Bers. Holomorphic families of isomorphisms of Möbius groups. *J. of Math. of Kyoto University* **26**(1986), 73–76.

[BR] L. Bers and H. L. Royden. Holomorphic families of injections. *Acta Math.* **157**(1986), 259–286.

[BDK] P. Blanchard, R. Devaney, and L. Keen. The dynamics of complex polynomials and automorphisms of the shift. *Inv. Math.* **104**(1991), 545–580.

[Bo] R. Bowen. Hausdorff dimension of quasi-circles. *Publ. Math. IHES* **50**(1978), 11–25.

[Br] B. Branner. Cubic polynomials: turning around the connectedness locus. In L. R. Goldberg and A. V. Phillips, editors, *Topological Methods in Modern Mathematics*, pages 391–427. Publish or Perish, Inc., 1993.

[CG] L. Carleson and T. Gamelin. *Complex Dynamics*. Springer-Verlag, 1993.

[CJ] A. Casson and D. Jungreis. Convergence groups and Seifert fibered 3-manifolds. *Invent. math.* **118**(1994), 441–456.

[Dou] A. Douady. Prolongement de mouvements holomorphes (d'après Słodkowski et autres). In *Séminaire Bourbaki, 1993/94*, pages 7–20. Astérisque, vol. 227, 1995.

[DH] A. Douady and J. Hubbard. A proof of Thurston's topological characterization of rational maps. *Acta Math.* **171**(1993), 263–297.

[EaM] C. Earle and C. McMullen. Quasiconformal isotopies. In *Holomorphic Functions and Moduli I*, pages 143–154. Springer-Verlag: MSRI publications volume 10, 1988.

[Ep1] C. L. Epstein. The hyperbolic Gauss map and quasiconformal reflections. *J. reine angew. Math.* **372**(1986), 96–135.

[Ep2] C. L. Epstein. Univalence criteria and surfaces in hyperbolic space. *J. reine angew. Math.* **380**(1987), 196–214.

[FS] M. H. Freedman and R. Skora. Strange actions of groups on spheres. *J. Differential Geom.* **25**(1987), 75–98.

[Gd] F. Gardiner. *Teichmüller Theory and Quadratic Differentials*. Wiley Interscience, 1987.

[GK] L. R. Goldberg and L. Keen. The mapping class group of a generic quadratic rational map and automorphisms of the 2-shift. *Invent. math.* **101**(1990), 335–372.

[Ha] K. T. Hahn. On completeness of the Bergman metric and its subordinate metric. *Proc. Natl. Acad. Sci. USA* **73**(1976).

[HP] J. L. Harer and R. C. Penner. *Combinatorics of Train Tracks*, volume 125 of *Annals of Math. Studies*. Princeton University Press, 1992.

[Hem] J. Hempel. *3-Manifolds*, volume 86 of *Annals of Math. Studies*. Princeton University Press, 1976.

[Her] M. Herman. Exemples de fractions rationelles ayant une orbite dense sur la sphere de Riemann. *Bull. Soc. Math. de France* **112**(1984), 93–142.

[Hub] J. H. Hubbard. *Sur les sections analytiques de la courbe universelle de Teichmüller*. Mem. Amer. Math. Soc., No. 166, 1976.

[IT] Y. Imayoshi and M. Taniguchi. *An Introduction to Teichmüller Spaces*. Springer-Verlag, 1992.

[JR] K. N. Jones and A. W. Reid. Vol3 and other exceptional hyperbolic 3-manifolds. *Proc. Amer. Math. Soc.* **129**(2001), 2175–2185.

[KP] M. E. Kapovich and L. D. Potyagaĭlo. On the absence of finiteness theorems of Ahlfors and Sullivan for Kleinian groups in higher dimensions. *Sibirsk. Mat. Zh.* **32**(1991), 61–73.

[KS1] L. Keen and C. Series. Pleating coordinates for the Maskit embedding of the Teichmüller space of a punctured torus. *Topology* **32**(1993), 719–749.

[KS2] L. Keen and C. Series. How to bend pairs of punctured tori. In *Lipa's Legacy (New York, 1995)*, pages 359–387. Amer. Math. Soc., 1997.

[Ker] S. Kerckhoff. The Nielsen realization problem. *Ann. of Math.* **177**(1983), 235–265.

[La] S. Lang. *Introduction to Complex Hyperbolic Spaces*. Springer-Verlag, 1987.

[Lat] S. Lattès. Sur l'iteration des substitutions rationelles et les fonctions de Poincaré. *C. R. Acad. Sci. Paris* **166**(1918), 26–28.

[Le] O. Lehto. *Univalent functions and Teichmüller spaces*. Springer-Verlag, 1987.

[LV] O. Lehto and K. J. Virtanen. *Quasiconformal Mappings in the Plane*. Springer-Verlag, 1973.

[Ly] M. Lyubich. An analysis of the stability of the dynamics of rational functions. *Selecta Math. Sov.* **9**(1990), 69–90.

[Me] R. Mañé. On the instability of Herman rings. *Invent. math.* **81**(1985), 459–471.

[MSS] R. Mañé, P. Sad, and D. Sullivan. On the dynamics of rational maps. *Ann. Sci. Éc. Norm. Sup.* **16**(1983), 193–217.

[Msk] B. Maskit. On a class of Kleinian groups. *Ann. Acad. Sci. Fenn. Ser. A I* **442**(1969), 8 pp.

[MT] H. Masur and S. Tabachnikov. Rational billiards and flat structures. In *Handbook of Dynamical Systems, Vol. 1A*, pages 1015–1089. North–Holland, 2002.

[Mc1] C. McMullen. Automorphisms of rational maps. In *Holomorphic Functions and Moduli I*, pages 31–60. Springer-Verlag, 1988.

[Mc2] C. McMullen. Braiding of the attractor and the failure of iterative algorithms. *Invent. math.* **91**(1988), 259–272.

[Mc3] C. McMullen. Rational maps and Kleinian groups. In *Proceedings of the International Congress of Mathematicians (Kyoto, 1990)*, pages 889–900. Springer-Verlag, 1991.

[Mc4] C. McMullen. *Complex Dynamics and Renormalization*, volume 135 of *Annals of Math. Studies*. Princeton University Press, 1994.

[Mc5] C. McMullen. *Renormalization and 3-Manifolds which Fiber over the Circle*, volume 142 of *Annals of Math. Studies*. Princeton University Press, 1996.

[McS] C. McMullen and D. Sullivan. Quasiconformal homeomorphisms and dynamics III: The Teichmüller space of a holomorphic dynamical system. *Adv. Math.* **135**(1998), 351–395.

[MeSt] W. de Melo and S. van Strien. *One-Dimensional Dynamics*. Springer-Verlag, 1993.

[Mo] M. Möller. Variations of Hodge structures of a Teichmüller curve. *J. Amer. Math. Soc.* **19**(2006), 327–344.

[Mon] P. Montel. *Familles Normales*. Gauthiers-Villars, 1927.

[MMW] D. Mumford, C. McMullen, and D. Wright. Limit sets of free two-generator kleinian groups. *Preprint, 1990*.

[Nag] S. Nag. *The Complex Analytic Theory of Teichmüller Space*. Wiley, 1988.

[Ot] J.-P. Otal. *Le théorème d'hyperbolisation pour les variétés fibrées de dimension trois*. Astérisque, vol. 235, 1996.

[Pu] C. C. Pugh. The closing lemma. *Amer. J. Math.* **89**(1967), 956–1009.

[Rat] J. G. Ratcliffe. *Foundations of Hyperbolic Manifolds*. Springer-Verlag, 1994.

[Rs] M. Rees. Positive measure sets of ergodic rational maps. *Ann. scient. Éc. Norm. Sup.* **19**(1986), 383–407.

[RS] B. Rodin and D. Sullivan. The convergence of circle packings to the Riemann mapping. *J. Differential Geom.* **26**(1987), 349–360.

[Sch] W. Schmid. Variation of Hodge structure: the singularities of the period mapping. *Invent. math.* **22**(1973), 211–319.

[Sn] L. Schneps, editor. *The Grothendieck theory of dessins d'enfants (Luminy, 1993)*, volume 200 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, 1994.

[Sc] R. Schwartz. Pappus' theorem and the modular group. *Publ. Math. IHES* **78**(1993), 187–206.

[Scott] P. Scott. Subgroups of surface groups are almost geometric. *J. London Math. Soc.* **17**(1978), 555–565.

[Shi] M. Shishikura. On the quasiconformal surgery of rational functions. *Ann. Sci. Éc. Norm. Sup.* **20**(1987), 1–30.

[Sl] Z. Słodkowski. Holomorphic motions and polynomial hulls. *Proc. Amer. Math. Soc.* **111**(1991), 347–355.

[Sm] S. Smale. Structurally stable systems are not dense. *Am. J. of Math.* **88**(1966), 491–496.

[Str1] K. Strebel. On lifts of extremal quasiconformal mappings. *J. d'Analyse Math.* **31**(1977), 191–203.

[Str2] K. Strebel. *Quadratic Differentials.* Springer-Verlag, 1984.

[Sul1] D. Sullivan. A finiteness theorem for cusps. *Acta Math.* **147**(1981), 289–299.

[Sul2] D. Sullivan. On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions. In I. Kra and B. Maskit, editors, *Riemann Surfaces and Related Topics: Proceedings of the 1978 Stony Brook Conference*, volume 97 of *Annals of Math. Studies.* Princeton University Press, 1981.

[Sul3] D. Sullivan. Discrete conformal groups and measurable dynamics. *Bull. Amer. Math. Soc.* **6**(1982), 57–74.

[Sul4] D. Sullivan. Quasiconformal homeomorphisms and dynamics I: Solution of the Fatou-Julia problem on wandering domains. *Annals of Math.* **122**(1985), 401–418.

[Sul5] D. Sullivan. Quasiconformal homeomorphisms and dynamics II: Structural stability implies hyperbolicity for Kleinian groups. *Acta Math.* **155**(1985), 243–260.

[Thom] R. Thom. L'equivalence d'une fonction différentiable et d'un polynôme. *Topology* **3**(1965), 297–307.

[Tm] R. Thom. *Structural stability and morphogenesis.* Addison-Wesley, 1989.

[Th] W. P. Thurston. Zippers and univalent functions. In A. Baernstein et al, editor, *The Bieberbach conjecture*, pages 185–197. Amer. Math. Soc., 1986.

[Wil] R. F. Williams. The "DA" maps of Smale and structural stability. In *Global Analysis*, pages 329–334. Am. Math. Soc., 1970.

[Wol1] S. Wolpert. On the homology of the moduli space of stable curves. *Annals of Math.* **118**(1983), 491–523.

[Wol2] S. Wolpert. Geodesic length functions and the Nielsen problem. *J. Diff. Geom.* **25**(1987), 275–296.

[Wr] D. Wright. The shape of the boundary of Maskit's embedding of the Teichmüller space of once-punctured tori. *Preprint, 1990.*

[Y] J.-C. Yoccoz. *Petits diviseurs en dimension* 1. Astérisque, vol. 231, 1995.