Riemann surfaces and the geometrization of 3-manifolds

Curt McMullen *
Princeton University, Princeton NJ 08544
Current address: University of California, Berkeley CA 94720
16 December, 1991

1 Introduction

About a decade ago Thurston proved that a vast collection of 3-manifolds carry metrics of constant negative curvature. These manifolds are thus elements of hyperbolic geometry, as natural as Euclid’s regular polyhedra. For a closed manifold, Mostow rigidity assures that a hyperbolic structure is unique when it exists [Mos], so topology and geometry mesh harmoniously in dimension 3.

This remarkable theorem applies to all 3-manifolds which can be built up in an inductive way from 3-balls, i.e. Haken manifolds. Thurston’s construction of a hyperbolic structure is also inductive. At the inductive step one must find the right geometry on an open 3-manifold so that its ends may be glued together. Using quasiconformal deformations, the gluing problem can be formulated as a fixed-point problem for a map of Teichmüller space to itself. Thurston proposes to find the fixed point by iterating this map.

Here we outline Thurston’s construction, and sketch a new proof that the iteration converges. Our argument rests on a result entirely in the theory of Riemann surfaces: an extremal quasiconformal mapping can be relaxed (isotoped to a map of lesser dilatation) when lifted to a sufficiently

large covering space (e.g. the universal cover). This contraction gives an immediate estimate for the contraction of Thurston’s iteration.

A detailed account of these results appears in [Mc1] and [Mc2]; therefore we have adopted a more informal tone here. An expository sequel relating hyperbolic manifolds and iterated rational maps appears in [Mc3].

The setup for Thurston’s theorem is discussed at length in [Mor], and Thurston’s proof is contained in the series [Th3], [Th5], [Th6], [Th7], etc.

See also [Th2] for a survey and an introduction to the more general geometrization conjecture, and [Th1] for more on hyperbolic geometry and Kleinian groups, as well as the existence of hyperbolic structures on many manifolds which are not Haken.

Background in Teichmüller theory can be found in e.g. [Bers2], [Bers4], and [Gd].

2 Teichmüller theory and quasiconformal maps.

In this section we state results from [Mc1] concerning contraction of canonical mappings between Teichmüller spaces.

Let \( f : X \to Y \) be an orientation-preserving homeomorphism between two Riemann surfaces \( X \) and \( Y \).

If \( f \) is smooth then we can measure its conformal distortion as follows. First, consider a real-linear isomorphism \( L : \mathbb{C} \to \mathbb{C} \). Think of the complex structure on \( \mathbb{C} \) as being recorded by the family of circles centered at the origin, since these are the orbits of multiplication by the unit complex numbers. The image of these circles under \( L \) is a family of ellipses of constant eccentricity. The ratio of major to minor axis gives a number \( K(L) \geq 1 \) called the dilatation of \( L \).

We define the dilatation \( K(f) \) as the supremum of the dilatation of the derivative

\[
Df : T_x \to T_{f(x)},
\]

over all tangent spaces \( T_x \) to \( X \).

The dilatation of \( f \) is 1 if and only if \( f \) is conformal. If \( K(f) < \infty \), \( f \) is quasiconformal.

Technical remark. The natural degree of smoothness to require is that the distributional derivatives of \( f \) lie in \( L^2 \).

Assuming \( f \) is quasiconformal, we can try to adjust it by isotopy to minimize its dilatation. By compactness of maps with bounded dilatation, at least one extremal (dilatation minimizing) quasiconformal map exists in each isotopy class. For example, if \( X \) and \( Y \) are tori, any extremal map is...
an affine stretch (it takes the form \( z \mapsto Az + Bz + C \) in the universal covers of \( X \) and \( Y \)).

Teichmüller’s theorem describes the extremal maps when \( X \) and \( Y \) are hyperbolic Riemann surfaces of finite area (equivalently, surfaces of negative Euler characteristic obtained from compact surfaces by possibly removing a finite number of points.) In each isotopy class there is a unique extremal Teichmüller map. Away from a finite number of singularities, the map is again an affine stretch in appropriate local coordinates.

Relaxation of quasiconformal maps.

Now consider quasiconformal mappings on the unit disk \( \Delta = \{ z : |z| < 1 \} \).

Any quasiconformal map \( f : \Delta \to \Delta \) extends to a homeomorphism of \( S^1 \), the boundary of the disk. Of course \( f \) is always isotopic to the identity; to obtain an interesting extremal problem, we require that the isotopy fix the values of \( f \) on \( S^1 \).

Our first result states that a Teichmüller mapping is no longer extremal when lifted to the universal cover.

**Theorem 2.1** Let \( f : X \to Y \) be a Teichmüller mapping between hyperbolic Riemann surfaces of finite type. Then the map \( \tilde{f} : \Delta \to \Delta \) obtained by lifting \( f \) to the universal covers of \( X \) and \( Y \) is not extremal among quasiconformal maps with the same boundary values (unless \( f \) is conformal).

This result can be reformulated in terms of natural maps between Teichmüller spaces. Given a Riemann surface \( X \), the Teichmüller space \( \text{Teich}(X) \) consists of equivalence classes of data \((f : X \to X_1)\) where \( X_1 \) is another Riemann surface and \( f \) is a quasiconformal map. Given two points \((f : X \to X_1)\) and \((g : X \to X_2)\), the Teichmüller distance \( d(X_1, X_2) \) is defined to be \( \inf \log(K(h)) \) where \( h \) ranges over all quasiconformal maps \( h : X_1 \to X_2 \) which are isotopic to \( g \circ f^{-1} \) (rel ideal boundary). Thus \( d(X_1, X_2) = 0 \) exactly when there is a conformal isomorphism in the correct isotopy class, and in this case we consider \((f, X_1)\) and \((g, X_2)\) to represent the same point in Teichmüller space.

Now given a covering \( Y \to X \), there is a natural inclusion \( \text{Teich}(X) \hookrightarrow \text{Teich}(Y) \); one simply forms the corresponding covering space of each Riemann surface quasiconformally equivalent to \( X \). For the special case of the universal covering of a Riemann surface of finite type, Theorem 2.1 says the map

\[ \text{Teich}(X) \to \text{Teich}(\Delta) \]

is a contraction for the Teichmüller metric.
Poincaré series.

To check that a map between Teichmüller spaces is a contraction, it suffices to show the derivative of the map is a contracting operator. It is actually more convenient to work with the coderivative, dual to the derivative operator.

Let $Q(X)$ denote the space of holomorphic quadratic differentials $\phi(z)dz^2$ on $X$, such that

$$||\phi|| = \int_X |\phi(z)| |dz|^2 < \infty.$$  

With the above norm, $Q(X)$ is a Banach space.

$Q(X)$ may be naturally identified with the cotangent space to Teichmüller space at $X$; its norm is the infinitesimal form of the Teichmüller cometric. (As $Q(X)$ generally fails to be a Hilbert space, this is a Finsler metric rather than a Riemannian metric. Technically, $Q(X)$ is the predual to the tangent space.)

If $Y \to X$ is a covering space, then there is a natural push-forward operator

$$\Theta : Q(Y) \to Q(X).$$

This operator is the coderivative at $X$ of the inclusion $\text{Teich}(X) \hookrightarrow \text{Teich}(Y)$.

In the case of the universal covering, the operator $\Theta$ is identical with the classical Poincaré series operator

$$\Theta(\phi) = \sum_{\gamma \in \Gamma} \gamma^* \phi$$

which takes an integrable quadratic differential $\phi$ on the disk and converts it into an automorphic form for the Fuchsian group $\Gamma = \pi_1(X)$ (and therefore an element of $Q(X)$) [Poin].

A more precise formulation of Theorem 2.1 is:

**Theorem 2.2 (Kra’s Theta conjecture) $||\Theta|| < 1$ for classical Poincaré series.**

This means the inclusion of Teichmüller spaces is a contraction even at the infinitesimal level.

Amenability.

One can characterize those coverings for which contraction is obtained in terms of the purely combinatorial notion of amenability.

To introduce this notion, first consider the case of a graph (1-complex) $G$. For any set $V$ of vertices of $G$, let $\partial V$ denote the set of vertices at
distance 1 from \( V \). (A vertex at distance 1 is connected to \( V \) by an edge but does not itself lie in \( V \)). Then the \textit{expansion} \( \gamma \) of \( G \) is given by

\[
\gamma = \inf \frac{|\partial V|}{|V|}
\]

over all finite sets of vertices \( V \).

If the expansion is 0, \( G \) is \textit{amenable}; otherwise it is nonamenable, and the boundary of any vertex set is comparable in size to the set itself. For example, there is a unique infinite tree with degree \( d \) at each vertex; it is amenable when \( d = 2 \) (the tree is an infinite line) and nonamenable for \( d > 2 \) (in fact the expansion constant is \( d - 2 \)).

Let \( X \) be a hyperbolic Riemann surface of finite type; then we can choose a finite graph \( G \subset X \) such that \( \pi_1(G) \) surjects onto \( \pi_1(X) \). An \textit{amenable covering} \( p : Y \to X \) is one for which \( p^{-1}(G) \) is an amenable graph. It is easy to check that the definition is independent of the choice of \( G \).

\textbf{Remark.} Alternatively, a covering is amenable if there exists a linear functional (a \textit{mean})

\[
m : L^\infty(\pi_1(X)/\pi_1(Y)) \to \mathbb{R},
\]

invariant under the left-action of \( \pi_1(X) \), such that \( \inf(f) \leq m(f) \leq \sup(f) \) for all \( f \). A normal covering is amenable if the deck transformations form an amenable group; cf. [Gre], [Pier].

\textbf{Theorem 2.3} Let \( Y \to X \) be a covering of a hyperbolic Riemann surface of finite type. Then either

\begin{enumerate}
\item The covering is amenable, \( ||\Theta|| = 1 \) and the induced map \( \text{Teich}(X) \to \text{Teich}(Y) \) is a global isometry for the Teichmüller metric, or
\item The covering is nonamenable, \( ||\Theta|| < 1 \) and \( \text{Teich}(X) \to \text{Teich}(Y) \) is contracting.
\end{enumerate}

The universal covering is easily shown to be nonamenable, so this theorem contains the previous two.

The proof of this result is technical, but the relation between relaxation and nonamenability is easy to describe.

It is related to the idea of a chain letter, or pyramid game. To join the game on a given round you must (a) pay $1 and (b) get two other people to join on the next round. Ten rounds later you leave the game, collecting $1,024 from the “descendents” of your two new members. In
real life someone eventually loses, but on a nonamenable graph (which is
necessarily infinite), it is possible to enrich every vertex by drawing capital
at a steady rate from infinity.

Similarly, when a quasiconformal map is lifted to a cover, we can begin
relaxing it on some compact set while creating a certain amount of additional
stress near the boundary. Using nonamenability, the stress can be entirely
dissipated to infinity.

**Bibliographical remarks.** Theorem 2.1 was checked in many examples
by Strebel [Str]; Ohtake proved $||\Theta|| = 1$ for abelian coverings [Oh]. Other
relations between amenability and function theory appear in [Gr], [Br], [LS].

**Dependence on moduli.**

Let us focus again on the case of the universal covering $\Delta \to X$. For
application to iteration, it is useful to know how much the map $\text{Teich}(X) \to
\text{Teich}(\Delta)$ contracts the Teichmüller metric, because a uniformly contracting
iteration has a fixed point.

**Theorem 2.4** For the universal covering of a Riemann surface $X$ of genus
g with $n$ punctures,

$$||\Theta|| \leq C(L, g, n) < 1,$$

where $L$ is the length of the shortest geodesic on $X$.

This theorem is immediate once $||\Theta||$ is shown to be a continuous function
on the moduli space $M_{g,n}$, by compactness of the set of Riemann surfaces
without short geodesics [Mum].

Moreover, $||\Theta|| \to 1$ as $L \to 0$; there is no uniform bound on all of
moduli space. Intuitively, a Riemann surface with a short geodesic is de-
genrating towards an infinite cylinder, whose fundamental group is $\mathbb{Z}$ and
whose universal cover is amenable.

## 3 Hyperbolic 3-manifolds

In this section we describe Thurston’s theorem on hyperbolic 3-manifolds,
and how the preceding results give a new approach to the proof.

A **hyperbolic 3-manifold** $N$ is a complete Riemannian 3-manifold with
a metric of constant curvature $-1$. We will only consider $N$ with finitely
generated fundamental groups.

Up to isometry, there is a unique simply-connected hyperbolic 3-manifold,
**hyperbolic space** $\mathbb{H}^3$, which is topologically a 3-ball. Hyperbolic geometry
tends to conformal geometry at infinity; for example, $\mathbb{H}^3$ may be compactified by the Riemann sphere $\mathbb{C}\hat{}$ in such a way that hyperbolic isometries extend to conformal maps.

Since the universal cover of $N$ is isometric to $\mathbb{H}^3$, we can think of $N$ as $\mathbb{H}^3/\Gamma$ where $\Gamma$ is a Kleinian group, i.e. a discrete group of hyperbolic isometries. There is a maximal open set $\Omega \subset \mathbb{C}\hat{}$ on which $\Gamma$ acts properly discontinuously, and from this we form the Kleinian manifold $(\mathbb{H}^3 \cup \Omega)/\Gamma$.

Thus any hyperbolic manifold $N$ is provided with a natural Riemann surface boundary $\partial N = \Omega/\Gamma$.

There are two important topological properties of any hyperbolic 3-manifold. First, $N$ is irreducible: any 2-sphere in $N$ bounds a ball, or equivalently $\pi_2(N) = 0$. This is immediate from contractibility of $\mathbb{H}^3$, the universal cover of $N$.

The second property is that $N$ is atoroidal; this means any incompressible torus $T^2 \subset N$ is peripheral (homotopic to an end of $N$). Equivalently, a discrete group of hyperbolic isometries isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$ is parabolic, i.e. it is conjugate to a lattice of translations acting on $\mathbb{C}\hat{}$ by

$$< z \mapsto z + 1, \quad z \mapsto z + \tau >$$

for some choice of $\tau$. Geometrically, such a subgroup determines a finite volume end of $N$ (a cusp), diffeomorphic to $T^2 \times [0, \infty)$ but rapidly narrowing.

Thurston’s theorem states that in a large category of 3-manifolds (including, for example, all 3-manifolds with non-empty boundary), these are the only obstructions to existence of a hyperbolic structure.

**Theorem 3.1 (Thurston)** An atoroidal Haken 3-manifold is hyperbolic.

A Haken manifold is one built up inductively from 3-balls by gluing along incompressible submanifolds of the boundary.

A piece of the boundary is compressible if there is a simple curve which is essential in the boundary but bounds a disk in the manifold. When gluing together 3-manifolds, there is danger of creating a reducible manifold by gluing two compressible curves together; the two disks the curve bounds join to form an essential 2-sphere. A Haken manifold is always irreducible: by assumption, the gluing locus is incompressible so this danger never arises.

**Remark.** For simplicity we will suppress consideration of the parabolic locus $P \subset \partial M$; in general one designates a portion of the boundary which is to be realized as cusps. For example, a knot complement frequently carries a hyperbolic structure; the boundary of a tubular neighborhood of the knot is a torus corresponding to a rank 2 cusp.
Sketch of Thurston’s proof of Theorem 3.1.
A finite collection of disjoint balls obviously carries a hyperbolic structure. Start gluing them together along incompressible submanifolds of their boundary. By an orbifold trick (using Andreev’s theorem, see [Mor]), one need only deal with the case of gluing along entire boundary components. At the inductive step, one has a hyperbolic realization $N$ of a 3-manifold $M$ with incompressible boundary, and gluing instructions encoded by an orientation-reversing involution

$$\tau : \partial M \to \partial M.$$ 

The construction is completed by the following result:

**Theorem 3.2** $M/\tau$ has a hyperbolic structure if and only if the quotient is atoroidal.

This key result is intimately related to iteration on Teichmüller space, and it this relation which allows the results of §2 to be brought to bear.

Let $M$ denote a topological (or equivalently smooth) 3-manifold. Let $GF(M)$ denote the space of hyperbolic 3-manifolds $N$ which are homeomorphic to $M$.

**Technical remark.** $N$ is required to be geometrically finite, and each $N$ is equipped with a choice of homeomorphism to $M$ up to homotopy equivalence rel boundary. Then, as in Teichmüller theory, $N_1$ and $N_2$ are regarded as the same point in $GF(M)$ if there is an isometry $N_1 \to N_2$ in the appropriate homotopy class.

A remarkable feature of dimension 3 is the following:

**Theorem 3.3** As long as $M$ admits at least one hyperbolic realization, there is a 1–1 correspondence between hyperbolic structures on $M$ and conformal structures on $\partial M$, i.e.

$$GF(M) \cong \text{Teich}(\partial M).$$

(The isomorphism is by $N \mapsto \partial N$).

**Remarks.**

1. If $\partial M$ is empty then the hyperbolic structure on $M$ is unique; this is *Mostow rigidity.*
2. Let $M = S \times [0,1]$ where $S$ is a surface of genus 2; then $M$ has at least one hyperbolic realization (consider a Fuchsian group). The theorem states that given any two Riemann surfaces $X, Y \in \text{Teich}(S)$ there is a unique 3-manifold interpolating between them; this is Bers’ simultaneous uniformization theorem [Bers1].

3. The general theorem stated above was developed by Bers, Maskit and others, and put into a final form by Sullivan [Sul1].

To prove Theorem 3.2, we must find in $GF(M)$ the correct geometry so that the ends to be glued together have compatible shape. This can be formulated as a fixed-point problem in Teichmüller space, as follows.

First, the skinning map

$$\sigma : \text{Teich}(\partial M) \to \text{Teich}(\overline{\partial M})$$

is defined by forming, for each $N \in GF(M)$, the quasifuchsian covering spaces for each component of $\partial N$, and recording the conformal structures on the new ends which appear. Then, the gluing instructions determine an isometry

$$\tau : \text{Teich}(\overline{\partial M}) \to \text{Teich}(\partial M),$$

and a fixed point for $\tau \circ \sigma$ corresponds to a hyperbolic manifold $N \in GF(M)$ whose ends fit together isometrically under the gluing instructions.

**Properties of the skinning map.**

A detailed treatment of the skinning map would take us deeper into the realm of Kleinian groups than we wish to venture at the moment; here we simply report what we would have learned upon our return.

1. The skinning map typically contracts the Teichmüller metric ($||d\sigma_X|| < 1$ at each point $X$.) Thus if $\tau \circ \sigma$ has a fixed point $X$, it can be found by iteration:

$$(\tau \circ \sigma)^n(Y) \to X$$

for every $Y \in \text{Teich}(\partial M)$. (An important exception is the case of a 3-manifold which fibers over the circle; this can be described as $(M = F \times [0,1])/\tau$, where the surface $F$ is a fiber, and $\tau$ is the monodromy map. In this case $\sigma$ is an isometry and a different approach is required — cf. [Th5].)

2. In general, the skinning map does not contract *uniformly*. This is sensible, because $M/\tau$ might not admit a hyperbolic structure. A
potential obstruction is a cylinder $S^1 \times [0, 1]$ in $M$ whose ends are glued together by $\tau$ to form a nonperipheral torus. Then (as noted above) $M/\tau$ cannot be hyperbolic.

3. A 3-manifold is \textit{acylindrical} if every cylinder

$$(S^1 \times [0, 1], S^1 \times \{0, 1\}) \subset (M, \partial M)$$

with its ends resting on essential curves in $\partial M$ can be deformed into the boundary. Such a manifold will remain atoroidal for any gluing instructions, so there is no evident obstruction to the hyperbolicity of $M/\tau$. One expects the iteration $(\tau \circ \sigma)$ to be robust in this case, and our first result (see [Mc2]) is:

\textbf{Theorem 3.4} \textit{If $M$ is acylindrical, the skinning map is uniformly contracting:}

$$||d\sigma_X|| < C(M) < 1.$$ 

\textit{Thus $M/\tau$ has a hyperbolic structure for any choice of gluing map $\tau$.}

\textbf{Remark.} Thurston has proved the stronger assertion that the image of the skinning map is \textit{bounded} in the acylindrical case.

4. We will briefly sketch how the results of the preceding section bear on the study of $||d\sigma||$. Suppose $M$ is acylindrical. Here is how the map $\sigma$ appears from the point of view of Riemann surfaces. Starting with $X$ in $\text{Teich}(\partial M)$, one forms countably many copies of its universal cover $\tilde{X}$. These are then patched together in a complicated but canonical pattern (depending on the topology of $M$) to form $\sigma(X)$, the image of $X$ under the skinning map.

Now suppose $X_1$ and $X_2$ are in $\text{Teich}(\partial M)$, and let $\phi : X_1 \rightarrow X_2$ be an extremal quasiconformal map. We can form a new quasiconformal map

$$\tilde{\phi} : \sigma(X_1) \rightarrow \sigma(X_2)$$

with the same dilatation as $\phi$, as follows: lift $\phi$ to a map between corresponding copies of the universal covers of $X_1$ and $X_2$, and complete by continuity.

By Theorem 2.1, $\tilde{\phi}$ can be relaxed to a map of lesser dilatation on each copy of the universal cover without changing its boundary values. Thus the relaxed maps still fit together, and we have shown that $\sigma$ contracts the Teichmüller metric. Moreover, the refined version of this relaxation result — Theorem 2.4 — yields:
Theorem 3.5 \[ |d\sigma_X| < c(L) < 1 \] where \( L \) is the length of the shortest geodesic on \( X \).

This result even holds in the cylindrical case, so long as \( M \) is not an interval bundle over a surface. The cylindrical case requires a discussion of the nonamenability of covers of \( X \) other than the universal cover.

5. The preceding does not yet yield uniform contraction; rather, it reduces the problem to a study of short geodesics.

This is a significant simplification, however, because a short geodesic controls the geometry of a hyperbolic manifold over a large distance, by the Margulis lemma. When combined with the theory of geometric limits of quadratic differentials (cf. Appendix to [Mc1]), one finds a qualitative picture which again forces contraction unless the short geodesic lies on one end of a compressing cylinder in the 3-manifold.

This cannot occur in the acylindrical case, so then the contraction is uniform.

In general we find:

\textbf{Theorem 3.6} For any initial guess \( Y \in \text{Teich}(\partial M) \), either

\[ Y_n = (\tau \circ \sigma)^n(Y) \to \text{a (unique) fixed point} \ X, \]

or \( Y_n \) develops short geodesics, bounding cylinders in \( M \) linked by \( \tau \) to form a nonperipheral torus in \( M/\tau \).

Thus the gluing problem is solvable if and only if \( M/\tau \) is atoroidal.

\section{Epilogue}

Two other iterations on Teichmüller space have been analyzed with a similar paradigm, and indeed one of our motivations was to give a parallel treatment of the skinning map. One iteration is the action on \( \text{Teich}(S) \) of an element of the mapping class group of the surface \( S \) [Th4], [FLP]; the other arises in the construction of critically finite rational maps [Th8], [DH].

The paradigm is to play (a) geometric control in the absence of short geodesics, against (b) topological information in their presence.
For example, a mapping class whose minimum translation distance is achieved is **geometric**; it either fixes a point in Teichmüller space (and can be represented by an **isometry**), or it fixes a geodesic (this is the **pseudo-Anosov** case.) Otherwise the minimum translation distance is not achieved, so there are Riemann surfaces tending to infinity in moduli space and translated a bounded distance. In the end of moduli space the surfaces have short geodesics, which are necessarily permuted; thus we have the topological conclusion that the mapping class is **reducible**. This approach appears in [Bers3].

Similarly, given a smooth branched cover $f: S^2 \rightarrow S^2$ whose critical points eventually cycle, one seeks a conformal structure preserved up to isotopy by $f$ (rel the post-critical set); then $f$ can be geometrized as a rational map. This desired conformal structure can again be described as a fixed point for an iteration on Teichmüller space. The iteration contracts uniformly in the absence of short geodesics; in their presence, one locates a topological obstruction to geometrization.

Finally, in retrospect it is promising to think of the skinning map as an instance of **renormalization**; a similar approach to the Feigenbaum phenomenon (using infinite dimensional Teichmüller spaces) has been proposed by Douady and Hubbard, and pursued by Sullivan [Sul2].

**Addendum October 1991.** The connection with renormalization is now understood more precisely: construction of the Feigenbaum fixed-point closely resembles the construction hyperbolic structures on 3-manifolds which fiber over the circle, the one case omitted from the discussion above. This analogy and others (which relate the construction of rational maps to that of hyperbolic 3-manifolds) are discussed in more detail in [Mc3]. Abundant progress towards understanding renormalization of quadratic polynomials appears in [Sul3].

A new approach to the Theta conjecture and its variants is discussed in [BD] and [Mc4].

**References**


