Lectures on Analysis on Metric Spaces

Universitext

Juha Heinonen

Springer

Universitext

Springer Science+Business Media, LLC

Universitext

Editors (North America): S. Axler, F.W. Gehring, and K.A. Ribet

Aksoy/Khamsi: Nonstandard Methods in Fixed Point Theory
Andersson: Topics in Complex Analysis
Aupetit: A Primer on Spectral Theory
Bachman/Narici/Beckenstein: Fourier and Wavelet Analysis
Balakrishnan/Ranganathan: A Textbook of Graph Theory
Balser: Formal Power Series and Linear Systems of Meromorphic Ordinary Differential Equations
Bapat: Linear Algebra and Linear Models (2nd ed.)
Berberian: Fundamentals of Real Analysis
Booss/Bleecker: Topology and Analysis
Borkar: Probability Theory: An Advanced Course
Böttcher/Silbermann: Introduction to Large Truncated Toeplitz Matrices
Carleson/Gamelin: Complex Dynamics
Cecil: Lie Sphere Geometry: With Applications to Submanifolds
Chae: Lebesgue Integration (2nd ed.)
Charlap: Bieberbach Groups and Flat Manifolds
Chern: Complex Manifolds Without Potential Theory
Cohn: A Classical Invitation to Algebraic Numbers and Class Fields
Curtis: Abstract Linear Algebra
Curtis: Matrix Groups
DiBenedetto: Degenerate Parabolic Equations
Dimca: Singularities and Topology of Hypersurfaces
Edwards: A Formal Background to Mathematics I a/b
Edwards: A Formal Background to Mathematics II a/b
Foulds: Graph Theory Applications
Friedman: Algebraic Surfaces and Holomorphic Vector Bundles
Fuhrmann: A Polynomial Approach to Linear Algebra
Gardiner: A First Course in Group Theory
Gårding/Tambour: Algebra for Computer Science
Goldblatt: Orthogonality and Spacetime Geometry
Gustafson/Rao: Numerical Range: The Field of Values of Linear Operators and Matrices
Hahn: Quadratic Algebras, Clifford Algebras, and Arithmetic Witt Groups
Heinonen: Lectures on Analysis on Metric Spaces
Holmgren: A First Course in Discrete Dynamical Systems
Howe/Tan: Non-Abelian Harmonic Analysis: Applications of $SL(2, \mathrm{R})$
Howes: Modern Analysis and Topology
Hsieh/Sibuya: Basic Theory of Ordinary Differential Equations
Humi/Miller: Second Course in Ordinary Differential Equations
Hurwitz/Kritikos: Lectures on Number Theory
Jennings: Modern Geometry with Applications
Jones/Morris/Pearson: Abstract Algebra and Famous Impossibilities
Kannan/Krueger: Advanced Analysis
Kelly/Matthews: The Non-Euclidean Hyperbolic Plane
Kostrikin: Introduction to Algebra
Luecking/Rubel: Complex Analysis: A Functional Analysis Approach
MacLane/Moerdijk: Sheaves in Geometry and Logic
Marcus: Number Fields

(continued after index)

Juha Heinonen

Lectures on Analysis on Metric Spaces

Springer

Juha Heinonen
Mathematics Department
East Hall
University of Michigan
Ann Arbor, MI 48109-1109
USA

Mathematics Subject Classification (2000): 28Axx, 43A85, 46Exx

---

Library of Congress Cataloging-in-Publication Data
Heinonen, Juha.
Lectures on analysis on metric spaces / Juha Heinonen.
p. cm — (Universitext)
Includes bibliographical references and index.
ISBN 978-1-4612-6525-2 ISBN 978-1-4613-0131-8 (eBook)
DOI 10.1007/978-1-4613-0131-8
1. Metric spaces. 2. Mathematical analysis. I. Title. II. Series.
QA611.28.H44 2001
514′.32—dc21 00-056266

Printed on acid-free paper.

Production managed by Michael Koy; manufacturing supervised by Jerome Basma.
Typeset by TechBooks, Fairfax, VA.

9 8 7 6 5 4 3 2 1

ISBN 978-1-4612-6525-2

To my teacher Professor Olli Martio

Preface

Analysis in spaces with no a priori smooth structure has progressed to include concepts from first-order calculus. In particular, there have been important advances in understanding the infinitesimal versus global behavior of Lipschitz functions and quasiconformal mappings in rather general settings; abstract Sobolev space theories have been instrumental in this development. The purpose of this book is to communicate some of the recent work in the area while preparing the reader to study more substantial related articles.

The book is a written version of lectures from a graduate course I taught at the University of Michigan in Fall 1996. I have added remarks and references to works that have appeared thereafter. The material can roughly be divided into three different types: classical, standard but sometimes with a new twist, and recent. For instance, the treatment of covering theorems and their applications is classical, with little novelty in the presentation. On the other hand, the ensuing discussion on Sobolev spaces is more modern, emphasizing principles that are valid in larger contexts. Finally, most of the material on quasisymmetric maps is relatively recent and appears for the first time in book format; one can even find a few previously unpublished results. The bibliography is extensive but by no means exhaustive.

The reader is assumed to be familiar with basic analysis such as Lebesgue's theory of integration and the elementary Banach space theory. Some knowledge of complex analysis is helpful but not necessary in the chapters that deal with quasisymmetric maps. Occasionally, some elementary facts from other fields are quoted, mostly in examples and remarks.

The students in the course, many colleagues, and friends made a number of useful suggestions on the earlier versions of the manuscript and found many errors. I thank them all. Special thanks go to four people: Bruce Hanson carefully read the entire manuscript and detected more errors than anyone else; years later, use of Kari Hag's handwritten class notes preserved the spirit of the class; Fred Gehring insisted that I prepare these notes for publication; and Jeremy Tyson provided crucial help in preparing the final version of the text. Finally, I am grateful for having an inspiring audience for my lectures at Michigan.

Juha Heinonen
Ann Arbor, Michigan
July 1999

Contents

1

Covering Theorems

All covering theorems are based on the same idea: from an arbitrary cover of a set in a metric space, one tries to select a subcover that is, in a certain sense, as disjointed as possible. In order to have such a result, one needs to assume that the covering sets are somehow nice, usually balls. In applications, the metric space normally comes with a measure μ, so that if $\mathcal{F} = \{B\}$ is a covering of a set A by balls, then always

$$\mu(A) \leq \sum_{\mathcal{F}} \mu(B)$$

(with proper interpretation of the sum if the collection $\mathcal{F}$ is not countable). What we often would like to have, for instance, is an inequality in the other direction,

$$\mu(A) \geq C \sum_{\mathcal{F}'} \mu(B),$$

for some subcollection $\mathcal{F}' \subset \mathcal{F}$ that still covers A and for some positive constant C that is independent of A and the covering $\mathcal{F}$. There are many versions of this theme.

In the following, we describe three covering theorems, which we call the basic covering theorem, the Vitali covering theorem, and the Besicovitch–Federer covering theorem. (The terms may not be standard.) Each of these three theorems has a slightly different purpose, as well as different domains of validity, and I have chosen them as representatives of the many existing covering theorems. We will not strive for the greatest generality here.

1.1 Convention. Let X be a metric space. When we speak of a ball B in X, it is understood that it comes with a fixed center and radius, although these in general are not uniquely determined by B as a set. Thus, $B = B(x,r)$, and we write $\lambda B = B(x, \lambda r)$ for $\lambda > 0$. In this chapter, unless otherwise stated, a ball B can be either open or closed, with the understanding that λB is of the same type.

We will use the Polish distance notation

$$|x-y|$$

in every metric space, unless there is a specific need to do otherwise.

A family of sets is called *disjointed* if no two sets from the family meet.

Theorem 1.2 (basic covering theorem). *Every family $\mathcal{F}$ of balls of uniformly bounded diameter in a metric space X contains a disjointed subfamily $\mathcal{G}$ such that*

$$\bigcup_{B\in\mathcal{F}} B \subset \bigcup_{B\in\mathcal{G}} 5B.$$

In fact, every ball B from $\mathcal{F}$ meets a ball from $\mathcal{G}$ with radius at least half that of B.

For the proof, recall the following version of Zorn's lemma: If every chain in a nonempty partially ordered set has an upper bound, then the partially ordered set has a maximal element.

Proof. Let Ω denote the partially ordered (by inclusion) set consisting of all disjointed subfamilies ω of $\mathcal{F}$ with the following property: If a ball B from $\mathcal{F}$ meets some ball from ω, then it meets one whose radius is at least half the radius of B. Then, if $C \subset \Omega$ is a chain, it is easy to see that

$$\omega_0 = \bigcup_C \omega$$

belongs to Ω, so there is a maximal element $\mathcal{G}$ in Ω. By construction, $\mathcal{G}$ is disjointed. (Note that Ω is nonempty because the one ball family $\omega = \{B\}$ is in Ω whenever $B \in \mathcal{F}$ has radius close to the supremal one.)

If there is a ball B in $\mathcal{F}$ that does not meet any ball from $\mathcal{G}$, then pick a ball B_0 from $\mathcal{F}$ such that the radius of B_0 is larger than half of the radius of any other ball that does not meet the balls from $\mathcal{G}$. Then, if a ball B from $\mathcal{F}$ meets a ball from the collection $\mathcal{G}' = \mathcal{G} \cup \{B_0\}$, by construction it meets one whose radius is at least half that of B, showing that $\mathcal{G}'$ belongs to Ω. But this contradicts the maximality of $\mathcal{G}$.

Thus, every ball $B = B(x,r)$ from $\mathcal{F}$ meets a ball $B' = B(x', r')$ from $\mathcal{G}$ so that $r \le 2r'$, and the triangle inequality shows that $B \subset 5B'$. The theorem follows. □

Remark 1.3. (a) In Theorem 1.2, $\mathcal{G}$ is not asserted to be countable, although in applications it often will be.

(b) The proof of Theorem 1.2 in [47, 2.8] applies to more general coverings than those by balls.

(c) Zorn's lemma is equivalent to the axiom of choice; see, for example, [91, Chapter 1]. A constructive proof for Theorem 1.2 is possible with mild additional assumptions on X. For instance, in [134, 2.1] one finds a simple geometric argument in the event that the closed balls in X are compact. The idea is that one begins to select a disjointed collection by always picking out nearly the largest ball possible. In this setting, $\mathcal{G}$ will also be countable.

(d) Simple examples show that the assumption "of uniformly bounded diameter" in Theorem 1.2 is necessary.

1.4 Convention. A *measure* in this book means a nonnegative countably subadditive set function defined on all subsets of a metric space. All measures are furthermore assumed to be Borel regular, which means that open sets are measurable and every set is contained in a Borel set with the same measure. If f is an extended real-valued function on a metric measure space (X, μ), we keep using the notation

$$\int_X f\, d\mu$$

even if the measurability of f is not clear; the meticulous reader should interpret this as an upper integral of f [47, 2.4.2].

A measure μ in a metric space is called *doubling* if balls have finite and positive measure and there is a constant $C(\mu) \geq 1$ such that

$$\mu(2B) \leq C(\mu)\mu(B) \tag{1.5}$$

for all balls B. We also call a metric measure space (X, μ) doubling if μ is a doubling measure. Note that condition (1.5) implies

$$\mu(\lambda B) \leq C(\mu, \lambda)\mu(B)$$

for all $\lambda \geq 1$. We adopt the convention that $C(\mu)$ always denotes a constant that depends only on the constant in the doubling condition (1.5) but may not be the same at each occurrence. In general, we denote by C a positive constant whose dependence on some data is clear from the context.

Under rather general circumstances, Borel regular measures have two useful properties, which are often called *inner regularity* and *outer regularity*. More precisely, if μ is a Borel regular measure, then for each Borel set A of finite measure $\mu(A)$ is the supremum of the numbers $\mu(C)$, where C runs through all closed subsets of A; moreover, if the metric balls have finite measure, and A is a Borel set, then $\mu(A)$ is the infimum of the numbers $\mu(U)$, where U runs through all open supersets of A. See [47, 2.2.2].

Theorem 1.6 (Vitali covering theorem). *Let A be a subset in a doubling metric measure space (X, μ), and let $\mathcal{F}$ be a collection of closed balls centered at A such that*

$$\inf\{r > 0 : B(a, r) \in \mathcal{F}\} = 0 \tag{1.7}$$

for each $a \in A$. Then, there is a countable disjointed subfamily $\mathcal{G}$ of $\mathcal{F}$ such that the balls in $\mathcal{G}$ cover μ almost all of A, namely

$$\mu\left(A \Big\backslash \bigcup_{\mathcal{G}} B\right) = 0.$$

Proof. Assume first that A is bounded. Next, we may assume that the balls in $\mathcal{F}$ have uniformly bounded radii; in particular, by the basic covering theorem, we find a disjointed subcollection $\mathcal{G}$ of $\mathcal{F}$ such that A is contained in $\cup_{\mathcal{G}} 5B$. Moreover, the union $\cup_{\mathcal{G}} B$ is contained in some fixed ball, and hence, because μ is doubling, the collection $\mathcal{G} = \{B_1, B_2, \dots\}$ is necessarily countable. We find, by using the doubling property of μ, that

$$\sum_{i\geq 1} \mu(5B_i) \leq C \sum_{i\geq 1} \mu(B_i) < \infty,$$

which implies that

$$\sum_{i>N} \mu(5B_i) \to 0$$

as $N \to \infty$. Therefore, it suffices to show that

$$A \Big\backslash \bigcup_{i=1}^{N} B_i \subset \bigcup_{i>N} 5B_i.$$

To this end, take $a \in A \setminus \cup_{i=1}^{N} B_i$. Because the balls in $\mathcal{F}$ are closed, we can find a small ball $B(a, r) \in \mathcal{F}$ that does not meet any of the balls B_i for $i \leq N$. On the other hand, by the basic covering theorem, the family $\mathcal{G}$ can be chosen so that $B(a, r)$ meets some ball B_j from $\mathcal{G}$ with radius at least $r/2$. Thus, $j > N$ and $B(a, r) \subset 5B_j$, as required.

We leave it to the reader to treat the case where A is unbounded. This completes the proof. □

Classically, the main application of the Vitali covering theorem is the following differentiation theorem of Lebesgue.

Theorem 1.8 (Lebesgue's differentiation theorem). *If f is a nonnegative, locally integrable function on a doubling metric measure space (X, μ), then*

$$\lim_{r\to 0} \fint_{B(x,r)} f \, d\mu = f(x) \tag{1.9}$$

for almost every $x \in X$.

In Eq. (1.9), and throughout the book, we use the notation

$$f_E = \fint_E f\,d\mu = \frac{1}{\mu(E)}\int f\,d\mu \tag{1.10}$$

for the mean value of f in E. A function is said to be *locally integrable* if every point has a neighborhood where the function is integrable. Similarly, we have locally p-integrable functions, and the notation L^p_{loc}, $L^p_{\mathrm{loc}}(\mu)$, $L^p_{\mathrm{loc}}(X), \ldots$ is used.

PROOF. If E denotes the set of points in X where Eq. (1.9) does not hold, cover E by closed balls with centers at E and radii so small that f is integrable in each ball; by the Vitali covering theorem, there is a countable union of balls of this kind containing almost every point of E. Thus, it suffices to show that E has measure zero in a fixed ball B where f is integrable.

To this end, we first claim that if $t > 0$ and if

$$\liminf_{r\to 0} \fint_{B(x,r)} f\,d\mu \le t$$

for each x in a subset A of B, then

$$\int_A f\,d\mu \le t\mu(A). \tag{1.11}$$

To prove this claim, fix $\epsilon > 0$ and choose an open superset U of A such that $\mu(U) \le \mu(A) + \epsilon$ (see 1.4). Then, each point in A has arbitrarily small closed ball neighborhoods contained in U where the mean value of f is less than $t + \epsilon$. The Vitali covering theorem implies that we can pick a countable disjointed collection of such balls covering almost all of A, from which

$$\int_A f\,d\mu \le (t+\epsilon)\mu(U) \le (t+\epsilon)(\mu(A)+\epsilon),$$

and Eq. (1.11) follows upon letting $\epsilon \to 0$. A similar argument shows that if $t > 0$ and if

$$\limsup_{r\to 0} \fint_{B(x,r)} f\,d\mu \ge t$$

for all $x \in A \subset B$, then

$$\int_A f\,d\mu \ge t\mu(A). \tag{1.12}$$

It follows, in particular, that

$$\limsup_{r\to 0} \fint_{B(x,r)} f\,d\mu < \infty$$

for almost every $x \in B$. On the other hand, if $A_{s,t}$ is the set of points x in B for which

$$\liminf_{r\to 0} \fint_{B(x,r)} f\, d\mu \le s < t \le \limsup_{r\to 0} \fint_{B(x,r)} f\, d\mu,$$

then $\mu(A_{s,t}) = 0$, for Eqs. (1.11) and (1.12) together imply

$$t\mu(A_{s,t}) \le \int_{A_{s,t}} f\, d\mu \le s\mu(A_{s,t}).$$

Thus, the limit on the left in Eq. (1.9) exists and is finite almost everywhere in B. Denote this limit by $g(x)$ whenever it exists. It remains to show that $g(x) = f(x)$ almost everywhere in B.

Fix both a set $F \subset B$ and $\epsilon > 0$; for each integer n, denote

$$A_n = \{x \in F : (1+\epsilon)^n \le g(x) < (1+\epsilon)^{n+1}\}.$$

Then, by Eq. (1.12),

$$\begin{aligned} \int_F g\, d\mu &= \sum_n \int_{A_n} g\, d\mu \le \sum_n (1+\epsilon)^{n+1} \mu(A_n) \\ &\le (1+\epsilon) \sum_n \int_{A_n} f\, d\mu = (1+\epsilon) \int_F f\, d\mu, \end{aligned}$$

and similarly, by Eq. (1.11),

$$\begin{aligned} \int_F g\, d\mu &= \sum_n \int_{A_n} g\, d\mu \ge \sum_n (1+\epsilon)^{n} \mu(A_n) \\ &\ge (1+\epsilon)^{-1} \sum_n \int_{A_n} f\, d\mu = (1+\epsilon)^{-1} \int_F f\, d\mu. \end{aligned}$$

By letting $\epsilon \to 0$, we infer that

$$\int_F g\, d\mu = \int_F f\, d\mu$$

and hence that $g = f$ almost everywhere in B. The theorem follows. □

Remark 1.13. The preceding proof of Lebesgue's differentiation theorem does not need the underlying space to be doubling; it is enough that balls in X have finite mass and that the Vitali covering theorem holds. Also note that for the Vitali covering theorem the hypothesis that (X, μ) be doubling is unnecessarily strong.

Thus, let us call a metric measure space (X, μ) a *Vitali space* if balls in X have finite measure and if for every set A in X and for every family $\mathcal{F}$ of closed balls in X satisfying Eq. (1.7) the conclusion of Theorem 1.6 holds. Then, Lebesgue's differentiation theorem (Theorem 1.8) holds in each Vitali space in the sense that Eq. (1.9) holds μ almost everywhere for each nonnegative locally integrable function f on a Vitali space (X, μ).

If μ is any Radon measure[1] in $\mathbb{R}^n$, then $(\mathbb{R}^n, \mu)$ is a Vitali space; in particular, Lebesgue's differentiation theorem holds in $\mathbb{R}^n$ whenever f is a locally μ-integrable function and μ is a Radon measure. (See [134, Theorem 2.8, p. 34].) For a more general statement, see Example 1.15 (f).

In the basic covering theorem, one would often like to have controlled overlap for the balls $5B$ with $B \in \mathcal{G}$. In general, this is impossible. An easy example to this effect is a space consisting of a union of infinitely many line segments of radius 1 emanating from a point in Euclidean space equipped with its internal (path) metric. The problem with this space is that it has too many directions lying too far apart at the common point. Next, we borrow a condition from [47, 2.8.14] on a set $A \subset X$, which guarantees a controlled overlap.

A set A in a metric space X is said to be *directionally* (ϵ, M) *limited* if for each $a \in A$ there are at most M distinct points $b_1, \ldots, b_p$ in $B(a, \epsilon) \cap A$ such that for $i \neq j$ and $x \in X$

$$|a - x| + |x - b_i| = |a - b_i| \quad \text{and} \quad |a - b_j| = |a - x|$$

imply

$$|x - b_j| \geq \frac{1}{4}|a - b_j|.$$

Here, the ball $B(a, \epsilon)$ is assumed to be open.

Theorem 1.14 (Besicovitch–Federer covering theorem). *If A is a directionally (ϵ, M) limited subset of X and if $\mathcal{F}$ is a family of closed balls centered at A with radii bounded by $\epsilon > 0$, then there are $2M + 1$ disjointed subfamilies $\mathcal{G}_1, \ldots, \mathcal{G}_{2M+1}$ of $\mathcal{F}$ such that*

$$A \subset \bigcup_{i=1}^{2M+1} \bigcup_{B \in \mathcal{G}_i} B.$$

Note that $\epsilon > 0$ plays no role in the conclusion; it represents the scale where events are allowed to take place, and it could be large.

For a proof of Theorem 1.14, see [47, 2.8.14]. Because of its rather limited domain of validity, Theorem 1.14 will not be used in this book.

Example 1.15. (a) In a finite-dimensional normed space X, every set is directionally (ϵ, M) limited for all $\epsilon > 0$ and for some M depending only on the dimension of the space. To see this, we may assume that $A = X$ and that $a = 0$. Take two

[1] A *Radon measure* is a Borel regular measure that gives a finite mass to each compact set.

distinct points b_i and b_j with $|b_i| \geq |b_j|$, and write

$$x = \frac{b_i |b_j|}{|b_i|}.$$

Then, suppose

$$\frac{|x - b_j|}{|b_j|} = \left| \frac{b_i |b_j|}{|b_i||b_j|} - \frac{b_j}{|b_j|} \right| = \left| \frac{b_i}{|b_i|} - \frac{b_j}{|b_j|} \right| \geq \frac{1}{4}.$$

But because there are only a finite number (depending on the dimension) of points on the unit sphere in a normed space that lie distance $\frac{1}{4}$ apart,[2] we have a bound for the number of points $b_1, \ldots, b_p$.

(b) An infinite orthonormal set $\{e_1, e_2, \ldots\}$ in a Hilbert space is not directionally (ϵ, M) limited if $\epsilon > 1$.

(c) Any compact subset A of a Riemannian manifold is directionally (ϵ, M) limited for some $\epsilon = \epsilon(A) > 0$ and $M = M(n)$, where n is the dimension of the manifold. This is proved by using the compactness of A and the fact that for each $a \in A$ and for each $\delta > 0$ the exponential map is $(1 + \delta)$-bi-Lipschitz in a sufficiently small neighborhood of the origin in the tangent space at a. See [47, 2.8.9].

(d) In contrast to (c), Chi [33] has proved that if a simply connected nonpositively curved Riemannian manifold covering a compact manifold satisfies a Besicovitch–Federer type covering theorem, then the manifold is isometric to $\mathbb{R}^n$.

(e) The Heisenberg group equipped with its Carnot metric (see Section 9.25) does not satisfy the Besicovitch–Federer covering theorem as it is presented in Theorem 1.16 for $X = \mathbb{R}^n$. See [111, 1.4].

(f) If X is a separable metric space that is a countable union of directionally limited subsets, then (X, μ) is a Vitali space as defined in Remark 1.13 whenever μ is a (Borel) measure such that $\mu(B) < \infty$ for all balls B in X. See [47, 2.8.18].

To finish the chapter, let us record the following version of the classical Besicovitch covering theorem.

Theorem 1.16. *Let A be a bounded set in a metric space X and $\mathcal{F}$ be a collection of balls centered at A. Then, there is a subcollection $\mathcal{G} \subset \mathcal{F}$ such that*

$$A \subset \bigcup_{B \in \mathcal{G}} B$$

and that

$$\frac{1}{5}\mathcal{G} = \left\{ \frac{1}{5}B : B \in \mathcal{G} \right\}$$

[2] Every n-dimensional Banach space is homeomorphic to $\mathbb{R}^n$ by an $\sqrt{n}$-bi-Lipschitz linear map by a theorem of John; see [140, p. 10].

is a disjointed family of balls. Moreover, if X carries a doubling measure, then $\mathcal{G}$ is countable, and if $X = \mathbb{R}^n$, then one can choose $\mathcal{G}$ such that

$$\sum_{B \in \mathcal{G}} \chi_B(x) \leq C(n) < \infty$$

for some dimensional constant $C(n)$, where χ_E denotes the characteristic function of a set E.

The first assertion is merely a special case of the basic covering theorem (Theorem 1.2), and if X carries a doubling measure, it is clear that $\mathcal{G}$ has to be countable. For a simple proof of the last assertion, see, for example, [134, 2.7] or [136, 1.2.1]. Of course, the last assertion also follows from Theorem 1.14.

1.17 Notes to Chapter 1. The material here is classical and can be found in many textbooks. In particular, [47, 2.8] contains a comprehensive study of covering theorems. See also [45] and [134]. Historically, covering theorems were developed in connection with the differentiation theory of real functions, which is one of the triumphs of Lebesgue's theory; see [22] and [91].

2

Maximal Functions

Throughout this chapter, (X, μ) is a doubling metric measure space.

Maximal functions are important everywhere in geometric analysis, and they will play a major role in this book. For a locally integrable real-valued function f in X, define

$$M(f)(x) = \sup_{r>0} \fint_{B(x,r)} |f|\, d\mu, \tag{2.1}$$

where, we recall, the barred integral sign denotes the mean value. A comparable function results if the supremum is taken over all balls containing x rather than those centered at x.

Recall from Chapter 1 that Lebesgue's differentiation theorem holds in the present setting: if f is locally integrable, then

$$\lim_{r\to 0} \fint_{B(x,r)} |f|\, d\mu = |f(x)|$$

for almost every $x \in X$. Thus, the maximal function $M(f)$ is always at least as large as $|f|$ (in the almost everywhere sense), and the important maximal function theorem of Hardy and Littlewood asserts that $M(f)$ is not much larger when measured as an L^p function for $p > 1$; in the case $p = 1$, the situation is different, as explained in Remark 2.5 (a).

Theorem 2.2 (maximal function theorem). *The maximal function maps $L^1(\mu)$ to weak $- L^1(\mu)$ and $L^p(\mu)$ to $L^p(\mu)$ for $p > 1$ in the following precise sense:*

for all $t > 0$, we have

$$\mu(\{M(f) > t\}) \leq \frac{C_1}{t} \int_X |f| \, d\mu \tag{2.3}$$

and

$$\int_X |M(f)|^p d\mu \leq C_p \int_X |f|^p \, d\mu \tag{2.4}$$

whenever f belongs to $L^1(\mu)$ in inequality (2.3) or to $L^p(\mu)$ for $p > 1$ in inequality (2.4). The constants C_1 and C_p depend only on p and on the doubling constant of μ.

Proof. The proof of inequality (2.3) is a simple application of the basic covering theorem. In that theorem, however, one requires that the balls in the covering have uniformly bounded diameter. Therefore, we must consider restricted maximal functions $M_R(f)$, for $0 < R < \infty$, which are defined as in Eq. (2.1) but the supremum is now taken only over the radii $0 < r < R$. The estimates proved in the following will be independent of R, and inequalities (2.3) and (2.4) follow by letting $R \to \infty$. With this understood, we proceed with the notation $M_R(f) = M(f)$.

For each x in the set $\{M(f) > t\}$, we can pick a ball $B(x, r)$ such that

$$\int_{B(x,r)} |f| \, d\mu > t\mu(B(x, r)),$$

and extracting from this collection of balls a countable subcollection $\mathcal{G}$ as in the basic covering theorem (Theorem 1.16), we see that

$$\begin{aligned} \mu(\{M(f) > t\}) &\leq \sum_{\mathcal{G}} \mu(5B) \leq C \sum_{\mathcal{G}} \mu(B) \\ &\leq \frac{C}{t} \sum_{\mathcal{G}} \int_B |f| \, d\mu \leq \frac{C}{t} \int_X |f| \, d\mu, \end{aligned}$$

as desired. Note that both the doubling property of μ and the disjointedness of the family $\mathcal{G}$ were used in the argument.

To prove inequality (2.4), we divide a given $L^p(\mu)$ function f into a good part and a bad part,

$$f = f \cdot \chi_{\{f \leq t/2\}} + f \cdot \chi_{\{f > t/2\}} =: g + b,$$

with $t > 0$ fixed, where χ_* denotes the characteristic function. Then,

$$M(f) \leq M(g) + M(b) \leq t/2 + M(b),$$

which implies that

$$\{M(f) > t\} \subset \{M(b) > t/2\}.$$

Using this observation, the weak estimate (2.3), and writing the formula for the

L^p integral of $M(f)$ in terms of its distribution function, we find that

$$\begin{aligned}\int_X |M(f)|^p \, d\mu &= p \int_0^\infty t^{p-1} \, \mu(\{M(f) > t\})\, dt \\ &\le p \int_0^\infty t^{p-1} \, \mu(\{M(b) > t/2\})\, dt \\ &\le C \int_0^\infty t^{p-1} t^{-1} \int_{\{f > t/2\}} |f| \, d\mu \, dt \\ &= C \int_0^\infty t^{p-2} \left(\frac{t}{2} \mu(\{f > t/2\}) + \int_{t/2}^\infty \mu(\{f > s\})\, ds \right) dt \\ &\le C \int_X |f|^p \, d\mu,\end{aligned}$$

with a simple change of variables in the last two lines. This proves Theorem 2.2. □

It is important to observe that the definition of the maximal function does not enter the proof of inequality (2.4); one only uses the sublinearity of M, the weak estimate (2.3), and the fact that $M(f) \le ||f||_\infty$. The proof of inequality (2.4) is a basic example of ubiquitous *interpolation arguments* in harmonic analysis.

Remark 2.5. (a) It is easy to see that $M(f)$ is never (Lebesgue) integrable in $\mathbb{R}^n$ for a nonzero integrable f in $\mathbb{R}^n$. Thus, inequality (2.4) cannot be extended to the value $p = 1$. On the other hand, one can show that $M(f) \in L^1(\mu)$ if and only if $|f| \log(2 + |f|) \in L^1(\mu)$, provided $\mu(X)$ is finite.

(b) The preceding proof of inequality (2.4) shows that one can take

$$C_p = \frac{C(\mu) 2^p p}{p - 1}, \tag{2.6}$$

where $C(\mu)$ depends on the doubling constant of μ only.

(c) The maximal function theorem is true for (not necessarily doubling) Radon measures in $\mathbb{R}^n$. The proof is similar to the preceding one; the Besicovitch covering theorem is used instead of the basic covering theorem. See, for example, [134, Theorem 2.19, p. 40].

2.7 Lebesgue's theorem revisited. A standard application of the weak-type estimate (2.3) is the following quick proof of Lebesgue's differentiation theorem (Theorem 1.8): if continuous functions are dense in $L^1(\mu)$—which happens for example if X is locally compact [91, p. 197]—and if f is a locally integrable function, then

$$\limsup_{r \to 0} ⨍_{B(x,r)} |f(y) - f(x)| \, d\mu(y) = 0 \tag{2.8}$$

for almost every $x \in X$. Indeed, if we denote by $\Lambda(f)(x)$ the expression on the left-hand side in Eq. (2.8), then Λ is sublinear, vanishes identically on continuous functions, and satisfies

$$\Lambda(f) \leq M(f) + |f|.$$

Thus, for a given locally integrable f, continuous g, and $t > 0$, we have by the weak-type estimate (2.3) that

$$\begin{aligned}\mu(\{\Lambda(f) > t\}) &\leq \mu(\{\Lambda(f-g) > t\}) \\ &\leq \mu(\{M(f-g) > t/2\}) + \mu(\{|f-g| > t/2\}) \\ &\leq \frac{C}{t}\int_X |f-g|\,d\mu.\end{aligned}$$

Because continuous functions are dense in $L^1(\mu)$, the last integral can be made arbitrarily small, from which it follows that $\Lambda(f) = 0$ almost everywhere in X, as desired.

Exercise 2.9. Prove the assertions in Remark 2.5 (a).

Exercise 2.10. Suppose that $\mathcal{B} = \{B_1, B_2, \ldots\}$ is a countable collection of balls in a doubling space (X, μ) and that $a_i \geq 0$ are real numbers. Show that

$$\int_X \left(\sum_{\mathcal{B}} a_i \chi_{\lambda B_i}\right)^p d\mu \leq C(\lambda, p, \mu) \int_X \left(\sum_{\mathcal{B}} a_i \chi_{B_i}\right)^p d\mu$$

for $1 < p < \infty$ and $\lambda > 1$. (Hint: Use the maximal function theorem together with the duality of $L^p(\mu)$ and $L^q(\mu)$ for $p^{-1} + q^{-1} = 1$.)

It follows from Exercise 2.10 that the lack of finite overlap in a family $\mathcal{G}$ in the basic covering theorem is not seen at the L^p level. In fact, the set where many balls overlap is even smaller than Exercise 2.10 indicates, as follows.

Exercise 2.11. In the situation of Exercise 2.10, write

$$T_\lambda(x) = \sum_{\mathcal{B}} \chi_{\lambda B_i}(x)$$

for $\lambda \geq 1$. Show that there is $\epsilon = \epsilon(\mu, \lambda) > 0$ such that

$$\int_{A_\lambda} \exp\{\epsilon T_\lambda\}\,d\mu \leq C(\lambda, \mu)\mu(\cup_{\mathcal{B}} B_i),$$

where A_λ denotes the support of T_λ. (Hint: Use the particular form of the constant C_p in Eq. (2.6) and the power series expansion of e^x.)

2.12 Notes to Chapter 2. The material in this chapter is standard. The best references for maximal functions and their use in analysis are the two books of Stein, [170] and [171]. But see also [15], [45], [58], [59], [134], [164], and [205].

3

Sobolev Spaces

We denote by $\mathbb{R}^n$ Euclidean n-space, $n \geq 1$, and by dx its Lebesgue measure.

A classical way to speak about the "derivative" of a locally integrable function u in $\mathbb{R}^n$ is to view u as a distribution; that is, as a dual element of the space C_0^∞, compactly supported smooth functions in $\mathbb{R}^n$. (By a dual element, we mean a continuous linear functional on the appropriately topologized space C_0^∞; see, for example, [204, Chapter 1].) Thus, the ith partial derivative of u for $i = 1, \ldots, n$ is another distribution $\partial_i u$ defined by

$$\langle \partial_i u, \phi \rangle = -\langle u, \partial_i \phi \rangle = -\int_{\mathbb{R}^n} u \partial_i \phi \, dx$$

for $\phi \in C_0^\infty$. If it so happens that the action of $\partial_i u$ is given by integration against some locally integrable function v_i,

$$\langle \partial_i u, \phi \rangle = \int_{\mathbb{R}^n} v_i \phi \, dx,$$

then v_i is said to be the *weak ith partial derivative* of u and denoted by $\partial_i u$. It is easy to see that the weak derivative v_i in L^1_{loc}, if it exists, is unique, so that the notation $\partial_i u$ is justified.

The vector space of all locally integrable functions u for which locally integrable weak partial derivatives $\partial_i u$ exist for all $i = 1, \ldots, n$ is denoted by

$$W^{1,1}_{\text{loc}} = W^{1,1}_{\text{loc}}(\mathbb{R}^n)$$

and called the *local Sobolev space*. If we consider those u that are globally integrable, with weak derivatives also globally integrable, we have the space

$W^{1,1} = W^{1,1}(\mathbb{R}^n)$; if they all are locally or globally L^p-integrable for $1 \leq p < \infty$, we have the spaces

$$W^{1,p}_{\text{loc}} = W^{1,p}_{\text{loc}}(\mathbb{R}^n), \qquad W^{1,p} = W^{1,p}(\mathbb{R}^n).$$

It is important to note that both u and its weak derivative $\partial_i u$ are elements in the space L^1_{loc}, and thus a priori determined only up to a set of measure zero.

We can similarly define the space $W^{1,\infty}_{\text{loc}}$ as those locally (essentially) bounded functions whose weak derivatives are locally (essentially) bounded. The space $W^{1,\infty}_{\text{loc}}$ and its global version $W^{1,\infty}$ will later be identified as the space of locally or globally Lipschitz functions on $\mathbb{R}^n$.

It is easy to see that the Sobolev spaces $W^{1,p}$ are Banach spaces for all $1 \leq p \leq \infty$, with the norm

$$\|u\|_{1,p} = \|u\|_p + \|\nabla u\|_p.$$

Here, the Lebesgue measure in $\mathbb{R}^n$ is used in the L^p norms, and if u is in $W^{1,p}_{\text{loc}}$, we write

$$\nabla u = (\partial_1 u, \ldots, \partial_n u)$$

for its *weak gradient*. Thus, a locally integrable function u in $\mathbb{R}^n$ is in the Sobolev space $W^{1,1}_{\text{loc}}$ if and only if there is a locally integrable vector field $\mathbf{v} = (v_1, \ldots, v_n)$ such that

$$\langle \mathbf{v}, \phi \rangle = \int_{\mathbb{R}^n} \mathbf{v} \cdot \phi \, dx = -\int_{\mathbb{R}^n} u \nabla \cdot \phi \, dx = -\langle u, \nabla \cdot \phi \rangle \tag{3.1}$$

for each smooth, compactly supported vector field ϕ in $\mathbb{R}^n$; the vector field $\mathbf{v}$ is denoted by ∇u. Note that the operator $\nabla \cdot$ is the *divergence* on vector fields. A similar definition works for globally integrable functions.

Finally, the preceding discussion is valid for functions defined on an open subset Ω of $\mathbb{R}^n$. Thus, we have the Sobolev spaces $W^{1,p}_{\text{loc}}(\Omega)$, $W^{1,p}(\Omega)$.

3.2 Sobolev spaces of differential forms. Definition (3.1) extends to situations where we have operators that are adjoints of each other with respect to some inner product or pairing. For example, consider smooth l-forms on a closed Riemannian n-manifold M^n for $l = 0, \ldots, n$. Then, the exterior d has the adjoint operator δ taking $(l+1)$-forms to l-forms,

$$\delta = (-1)^{nl+1} * d *,$$

where $*$ is the *Hodge star* operator taking l-forms to $(n-l)$-forms isomorphically. The natural inner product in the space of smooth l-forms is

$$\langle \alpha, \beta \rangle = \int_M \alpha \wedge *\beta, \tag{3.3}$$

and to say that δ is the adjoint of d is to say that

$$\langle d\alpha, \phi \rangle = \langle \alpha, \delta\phi \rangle \tag{3.4}$$

for all l-forms α and $(l+1)$-forms ϕ. Formula (3.4) is a simple consequence of the Stokes theorem and the definition for δ. Note that $** = (-1)^{l(n-l)}\mathrm{id}$, and that $*$ is an isometry between the pertinent L^2 spaces of forms determined by the inner product (3.3).

In $\mathbb{R}^n$, the operator δ on 1-forms (naturally identified with vector fields) is the negative of the divergence, that is, $\delta = -\nabla\cdot$. Thus, Definition (3.1) becomes

$$\langle \mathbf{v}, \phi \rangle = \langle u, \delta\phi \rangle.$$

From this and from Eq. (3.4), it is easy to define $d\alpha$ for forms that are not smooth but only integrable: we say that an integrable $(l+1)$-form η is the *weak differential* of an integrable l-form α on M if

$$\langle \eta, \phi \rangle = \langle \alpha, \delta\phi \rangle$$

for all smooth $(l+1)$-forms ϕ. One can again show without difficulty that a weak differential, if it exists, is unique, and we can name it $d\alpha$. This leads to the corresponding Sobolev spaces

$$W_d^{1,p}(M, \Lambda^l)$$

for $1 \le p < \infty$ and $l = 0, 1, \ldots, n$, consisting of all L^p-integrable l-forms α for which $d\alpha$ exists as an L^p-form. With this terminology, $W_d^{1,p}(M, \Lambda^0) = W^{1,p}(M)$ is the Sobolev space of functions on M. The subscript d indicates that we require $d\alpha$ to be in the pertinent Lebesgue class. One could equally well consider spaces

$$W_\delta^{1,p}(M, \Lambda^l)$$

consisting of all L^p-forms α for which the *codifferential* $\delta\alpha$ exists in the weak sense and belongs to L^p. Then, in fact, if the Sobolev spaces $W^{1,p}(M, \Lambda^l)$ of forms on M are defined by using local coordinates and weak derivatives as in Definition (3.1), we have that

$$W^{1,p}(M, \Lambda^l) = W_d^{1,p}(M, \Lambda^l) \cap W_\delta^{1,p}(M, \Lambda^l), \quad 1 < p < \infty. \tag{3.5}$$

For equality (3.5), its history, and related exhaustive discussion, see [159] and [97].

Classically, there are three ways to define the Sobolev spaces $W^{1,p}$ for $1 \le p < \infty$ in $\mathbb{R}^n$. The first is by way of distributional derivatives, as earlier. The second is by way of a completion argument, as follows. For $1 \le p < \infty$, consider the normed space of all smooth functions ϕ in $\mathbb{R}^n$ such that

$$\|\phi\|_{1,p} = \|\phi\|_p + \|\nabla\phi\|_p < \infty, \tag{3.6}$$

and denote its completion in this norm by $H^{1,p}$. Thus, a function u is a member of $H^{1,p}$ if there are smooth functions ϕ_i that converge to u in L^p such that the gradients $\nabla\phi_i$ converge in L^p to some vector-valued function $\mathbf{v}$. Because, in this case,

$$\begin{aligned}\langle \mathbf{v}, \phi\rangle &= \lim_i \langle \nabla\phi_i, \phi\rangle \\ &= -\lim_i \langle \phi_i, \nabla\cdot\phi\rangle \\ &= -\langle u, \nabla\cdot\phi\rangle\end{aligned}$$

for any given compactly supported, smooth vector field ϕ, we see that u is in $W^{1,p}$ and $\mathbf{v}$ is its weak gradient. Thus, the natural inclusion $H^{1,p} \subset W^{1,p}$. The reverse inclusion is also valid. In fact, the convolutions

$$u_\epsilon(x) = \psi_\epsilon * u(x) = \int_{\mathbb{R}^n} \psi_\epsilon(x-y)u(y)\,dy \tag{3.7}$$

for $\epsilon > 0$ are smooth and dense in $W^{1,p}$, as is not hard to see; in Eq. (3.7) the functions ψ_ϵ are smooth, nonnegative, supported in a ball of radius $\epsilon > 0$ about the origin, and have total integral 1.

This approach can be generalized easily: fix any Radon measure μ on $\mathbb{R}^n$ and consider all smooth functions ϕ on $\mathbb{R}^n$ with finite norm,

$$\|\phi\|_{1,p;\mu} = \left(\int_{\mathbb{R}^n} |\phi|\,d\mu\right)^{1/p} + \left(\int_{\mathbb{R}^n} |\nabla\phi|^p\,d\mu\right)^{1/p} < \infty$$

for $1 \le p < \infty$, and denote the completion of this normed space by $H^{1,p}(\mu)$. Thus, a function u belongs to $H^{1,p}(\mu)$ if there is a sequence of smooth functions ϕ_i converging to u in $L^p(\mu)$ such that the functions $\nabla\phi_i$ converge to some vector-valued function $\mathbf{v}$ in $L^p(\mu)$. We could call this the Sobolev space associated with the measure μ (and a number $p \ge 1$), but there is a problem here: in this generality, the limit $\mathbf{v}$ need not be unique, so one cannot legitimately speak about *the* gradient of u. (For an example to this effect, see [46, p. 91].) However, the gradient is known to be unique for many measures μ; for instance, it is unique for the doubling measure

$$d\mu(x) = |x|^a\,dx \tag{3.8}$$

for any $-n < a < \infty$.

3.9 Open problem. Characterize the Radon measures μ in $\mathbb{R}^n$ for which the preceding completion procedure always produces a unique gradient.

It is known that the answer to the uniqueness question is affirmative if μ is doubling and an appropriate Poincaré inequality holds [54].

Incidentally, the Sobolev space $H^{1,p}(\mu)$ associated with the measure (3.8) when $a = p(n+1)$ contains the function $u(x) = |x|^{-n}$, which is not locally integrable against the Lebesgue measure near the origin, so u is not a distribution. The gradient of u in the Sobolev space $H^{1,p}(|x|^{p(n+1)}dx)$ is $-nx|x|^{-n-2}$.

Note that except for the issue of uniqueness of the "gradient," the preceding completion approach to $W^{1,p}$ bypasses the issue of duality, or "integration by parts"; we only need a vector subspace of L^p functions for which a derivative is defined as an L^p function, which then can be closed. See [81] for more on this approach, and the facts just mentioned.

The third standard way to define Sobolev spaces is via Bessel potentials. This approach, while important in defining fractional order derivatives, will not be used in this book. For completeness, it is recalled in the following after some discussion of potentials and embedding theorems.

3.10 Sobolev inequalities. The important inequalities in Sobolev space theory are the following: for a function $u \in W^{1,p}(\mathbb{R}^n)$, we have

$$\|u\|_{\frac{np}{n-p}} \leq C(n,p)\,\|\nabla u\|_p\,, \quad \text{if } 1 \leq p < n; \tag{3.11}$$

if $p > n$, then u has a continuous representative, which satisfies

$$|u(x) - u(y)| \leq C(n,p)|x-y|^{1-n/p}\,\|\nabla u\|_p \tag{3.12}$$

for $x, y \in \mathbb{R}^n$; moreover, there are $\epsilon = \epsilon(n) > 0$ and $C = C(n) \geq 1$ such that

$$\int_\Omega \exp\left\{\left(\frac{|u|}{\epsilon\,\|\nabla u\|_n}\right)^{n/(n-1)}\right\} \leq C|\Omega|, \tag{3.13}$$

if u is compactly supported in an open set Ω, where $|\Omega|$ is the volume of Ω.

The first two inequalities are known as *Sobolev embedding theorems*. For instance, inequality (3.11) expresses the fact that $W^{1,p}$ is continuously embedded into L^{p^*}, where

$$p^* = \frac{np}{n-p}$$

is the *Sobolev conjugate* of p for $1 \leq p < n$. Similarly, inequality (3.12) says that $W^{1,p}$ for $p > n$ is continuously embedded into $C^{0,1-n/p}$, the space of Hölder continuous functions of order $1 - n/p$ in $\mathbb{R}^n$. The third inequality (3.13) is called *Trudinger's inequality*. There are corresponding local versions of the preceding embedding theorems; see Chapter 4.

Notice that Sobolev functions in $W^{1,p}(\mathbb{R}^n)$ are increasingly better integrable when p approaches n and continuous when $p > n$. In the *borderline case* $p = n$, inequality (3.13) is essentially the best one can hope for in general: functions with singularities such as $\log\log|x|$ near the origin are in $W^{1,n}_{\mathrm{loc}}$.

Exercise 3.14. Prove that $H^{1,p} = W^{1,p}$, $1 \leq p < \infty$, by using the convolutions u_ϵ given in Eq. (3.7). Then, prove by a partition of unity argument that

$$W^{1,p}(\mathbb{R}^n) = H_0^{1,p}(\mathbb{R}^n),$$

where $H_0^{1,p}(\mathbb{R}^n)$ denotes the closure of compactly supported smooth functions in the norm (3.6). More generally, prove that smooth functions are dense in $W^{1,p}(\Omega)$ for each open Ω in $\mathbb{R}^n$ and $1 \leq p < \infty$. Show that in general $W^{1,p}(\Omega) \neq H_0^{1,p}(\Omega)$, where $H_0^{1,p}(\Omega)$ denotes the closure in $W^{1,p}(\Omega)$ of smooth functions with compact support in Ω.

There are at least three different ways to prove the Sobolev inequality (3.11). The first one is based on a trick involving integration by parts and is not very apt to generalizations. The second method is based on more general principles of harmonic analysis involving boundedness of certain singular integral operators; we will discuss this proof in detail. The third method is to derive inequality (3.11) from classical isoperimetric inequalities; this method is of great importance in Riemannian geometry and will also be discussed later.

Trudinger's inequality (3.13) will not be used in the book, so we omit the proof. (See [1, Section 3.1] for an exhaustive discussion on inequality (3.13) and related inequalities, plus interesting historical comments.) For a proof of inequality (3.12), see Exercise 4.9.

First proof of inequality (3.11). We only prove inequality (3.11) for $p = 1$. The general case follows from Hölder's inequality by applying the particular case to an appropriate power of $|u|$ (Exercise 3.16). Because smooth functions are dense in $W^{1,1}$, we may assume that u is smooth; in fact, we may assume that u is smooth and compactly supported by Exercise 3.14. Then,

$$|u(x)| \leq \int_{-\infty}^{\infty} |\partial_i u| dx_i$$

for all $i = 1, \dots, n$ and $x \in \mathbb{R}^n$, so that

$$|u(x)|^{n/(n-1)} \leq \left(\int_{-\infty}^{\infty} |\partial_1 u| dx_1 \cdots \int_{-\infty}^{\infty} |\partial_n u| dx_n \right)^{1/(n-1)},$$

from which one easily infers, by using the generalized Hölder inequality

$$\int_{\mathbb{R}^n} f_1 \dots f_k \, dx \leq \|f_1\|_{p_1} \cdots \|f_k\|_{p_k}, \quad p_1^{-1} + \cdots + p_k^{-1} = 1,$$

and Fubini's theorem, the desired inequality

$$\int_{\mathbb{R}^n} |u(x)|^{n/(n-1)} dx \leq \prod_{i=1}^{n} \left(\int_{\mathbb{R}^n} |\partial_i u| \, dx \right)^{1/(n-1)} \leq \left(\int_{\mathbb{R}^n} |\nabla u| \, dx \right)^{n/(n-1)}. \tag{3.15}$$

This was the proof of inequality (3.11) by trick. □

Note that inequality (3.15) gives $C(n, 1) \leq 1$ in inequality (3.11); by using the arithmetic-geometric inequality, one gets the bound $C(n, 1) \leq n^{-1/2}$. The best constant is known to be $C(n, 1) = n^{-1}\Omega_n^{-1/n}$, where Ω_n is the volume of the unit ball in $\mathbb{R}^n$. See, for example, [205, p. 81].

Exercise 3.16. Prove the Sobolev inequality (3.11) for $1 < p < n$.

For the second approach to Sobolev embedding theorems, we first observe that by integrating over the unit sphere the expression

$$u(x) = -\int_0^\infty D_r u(x + r\omega)\, dr,$$

valid for all unit vectors ω in $\mathbb{R}^n$, and by changing variables, we arrive at the pointwise representation

$$u(x) = C(n) \int_{\mathbb{R}^n} \frac{\nabla u(y) \cdot (x - y)}{|x - y|^n}\, dy \tag{3.17}$$

for any smooth, compactly supported function u in $\mathbb{R}^n$. The actual equality in Eq. (3.17) is not as important to us as the following consequential inequality:

$$|u(x)| \le C(n) \int_{\mathbb{R}^n} \frac{|\nabla u(y)|}{|x - y|^{n-1}}\, dy. \tag{3.18}$$

An inequality such as inequality (3.18) is called a *potential estimate.*

Given a nonnegative, locally integrable function f in $\mathbb{R}^n$, its *Riesz potential* (of order 1) is the function

$$I_1(f)(x) = (|y|^{1-n} * f)(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x - y|^{n-1}}\, dy.$$

(The potential of order $\alpha > 0$ would be $I_\alpha(f)(x) = (|y|^{\alpha-n} * f)(x)$.)

The crucial proposition about Riesz potentials is the following.

Proposition 3.19. *The sublinear operator $f \mapsto I_1(|f|)$ maps L^1 to weak $- L^{n/(n-1)}$, and L^p to $L^{np/(n-p)}$ if $1 < p < n$.*

Because inequality (3.18) implies

$$|u(x)| \le C(n) I_1(|\nabla u|)(x), \tag{3.20}$$

the Sobolev inequality (3.11) follows from Proposition 3.19, at least for $1 < p < n$. In fact, Proposition 3.19 can be used, with little extra work, to prove inequality (3.11) also when $p = 1$, as we will see later.

Proof of Proposition 3.19. We may assume that $f \ge 0$. Given $\delta > 0$, we divide the integral defining $I_1(f)$ into a bad part and a good part,

$$\begin{aligned} I_1(f)(x) &= \int_{B(x,\delta)} \frac{f(y)}{|x - y|^{n-1}}\, dy + \int_{\mathbb{R}^n \setminus B(x,\delta)} \frac{f(y)}{|x - y|^{n-1}}\, dy \\ &= b_\delta(x) + g_\delta(x), \end{aligned}$$

where the good part $g_\delta(x)$ can easily be dealt with by Hölder's inequality:

$$g_\delta(x) \le \|f\|_p \left(\int_{\mathbb{R}^n \setminus B(x,\delta)} |x - y|^{q(1-n)} \right)^{1/q},$$

where q is the Hölder conjugate of p. A calculation in polar coordinates shows that the latter integral is a constant times $\delta^{(p-n)/p}$, so that

$$g_\delta(x) \leq C(n, p)\, \|f\|_p\, \delta^{(p-n)/p}.$$

The bad part $b_\delta(x)$ contains a singularity of the convolving kernel and is best dealt with by a maximal function argument: for $j \geq 0$, denote by A_j the annulus

$$B_j \setminus B_{j+1} = B(x, 2^{-j}\delta) \setminus B(x, 2^{-j-1}\delta),$$

and deduce

$$\begin{aligned}
b_\delta(x) &= \sum_{j\geq 0} \int_{A_j} \frac{f(y)}{|x-y|^{n-1}}\, dy \\
&\leq C(n) \sum_{j\geq 0} (2^{-j}\delta)^{1-n} \int_{B_j} f(y)\, dy \\
&\leq C(n)\delta \sum_{j\geq 0} 2^{-j} \fint_{B_j} f(y)\, dy \\
&\leq C(n)\, \delta\, M(f)(x).
\end{aligned}$$

Putting these two estimates together, we find that

$$I_1(f)(x) \leq C(n)(\delta M(f)(x) + \delta^{1-n/p}\, \|f\|_p).$$

Then, observing that the minimum of the right-hand side is attained with

$$\delta = C(n, p) \left(\frac{\|f\|_p}{M(f)(x)} \right)^{p/n},$$

a computation shows that

$$I_1(f)(x) \leq C(n, p)\, \|f\|_p^{p/n}\, M(f)(x)^{1-p/n}. \tag{3.21}$$

Finally, using Eq. (3.21) and the maximal function theorem, we arrive (for $p > 1$) at

$$\begin{aligned}
\int_{\mathbb{R}^n} |I_1(f)(x)|^{np/(n-p)}\, dx &\leq C(n, p)\, \|f\|_p^{p^2/(n-p)} \int_{\mathbb{R}^n} |M(f)(x)|^p\, dx \\
&\leq C(n, p)\, \|f\|_p^{np/(n-p)},
\end{aligned}$$

which is what we wanted to prove. The case $p = 1$ is similar, with the maximal function theorem now giving the weaker conclusion that $I_1(|f|)$ lies in *weak* $- L^{n/(n-1)}$:

$$|\{I_1(|f|) > t\}| \leq C(n)(t^{-1}\, \|f\|_1)^{n/(n-1)},$$

where, again, $|\cdot|$ denotes volume. Proposition 3.19 follows. □

One immediately sees that the argument to prove Proposition 3.19 is very general; in fact, by repeating the preceding steps, we can record the following theorem.

Theorem 3.22. *In a doubling metric measure space (X, μ), define*

$$I_1(f)(x) = \int_X \frac{f(y)|x-y|}{\mu(B(x, |x-y|))}\, d\mu(y) \tag{3.23}$$

for a nonnegative measurable f. If there are constants $s > 1$ and $C \geq 1$ such that

$$\mu(B_r) \geq C^{-1} r^s \tag{3.24}$$

for any ball B_r in X of radius $r <$ diam X, then

$$\|I_1(f)\|_{sp/(s-p),\mu} \leq C(s, p, \mu)\, \|f\|_{p,\mu} \tag{3.25}$$

for $1 < p < s$, and

$$\mu(\{I_1(f) > t\}) \leq C(s, \mu)(t^{-1}\, \|f\|_1)^{s/(s-1)}. \tag{3.26}$$

Condition (3.24) on the mass of balls is not overly restrictive; for example, it applies to many spaces of fractal dimension. However, the condition does exclude many nice Riemannian manifolds such as the cylinder $\mathbb{S}^{n-1} \times \mathbb{R}$.

Finally, let us see how the weak estimate (3.26) can be used to prove the Sobolev inequality in $\mathbb{R}^n$ also for $p = 1$. Pick a smooth, compactly supported function u in $\mathbb{R}^n$; without loss of generality, we assume that u is nonnegative. We can express the support of u as the union of the sets

$$A_j = \{2^j < u \leq 2^{j+1}\}, \quad j \in \mathbb{Z}.$$

By observing that the gradient of the function

$$v_j = \max\{0, \min\{u - 2^j, 2^j\}\}$$

essentially lives on A_j, we use the weak estimate on v_j to conclude that

$$\begin{aligned} |A_{j+1}| &\leq |\{v_j > 2^{j-1}\}| \leq |\{I_1(|\nabla v_j|) > C(n)^{-1} 2^j\}| \\ &\leq C(n) \left(2^{-j} \int_{A_j} |\nabla v_j|\, dx\right)^{n/(n-1)} \\ &= C(n) \left(2^{-j} \int_{A_j} |\nabla u|\, dx\right)^{n/(n-1)}, \end{aligned}$$

where $|E|$ denotes the Lebesgue measure of the set E. This implies that

$$\begin{aligned}\int_{\mathbb{R}^n} |u|^{n/(n-1)}\,dx &= \sum_{j\in\mathbb{Z}} \int_{A_j} |u|^{n/(n-1)}\,dx \\ &\le \sum_{j\in\mathbb{Z}} (2^{j+1})^{n/(n-1)} |A_j| \\ &\le C(n) \sum_{j\in\mathbb{Z}} \left(\int_{A_j} |\nabla u|\,dx\right)^{n/(n-1)} \\ &\le C(n) \left(\sum_{j\in\mathbb{Z}} \int_{A_j} |\nabla u|\,dx\right)^{n/(n-1)} \\ &= C(n) \left(\int_{\mathbb{R}^n} |\nabla u|\,dx\right)^{n/(n-1)},\end{aligned}$$

as desired.

There is one point in the preceding proof that needs a justification, for we applied the pointwise estimate (3.20) to the function v_j, which is not necessarily smooth anymore. However, an approximation argument takes care of this: the functions v_j are Lipschitz and can be approximated uniformly by smooth functions. (We discuss this approximation later in Chapter 6.)

The important point in the preceding argument is that we are dealing with a function/gradient pair rather than an arbitrary L^1 function and its potential. Indeed, it is easy to see that one cannot have Eq. (3.25) at the endpoint $p = 1$. Simply consider functions ϕ_ϵ in $\mathbb{R}^n$ with uniformly bounded L^1 norm such that ϕ_ϵ converges to the delta function at the origin as $\epsilon \to 0$; then, if we had a bound of the type

$$\|I_1(f)\|_{n/(n-1)} \le C(n)\,\|f\|_1, \tag{3.27}$$

by using the approximating sequence ϕ_ϵ we would prove the integrability of the function $x \mapsto |x|^{-n}$ near the origin, which is absurd.

Exercise 3.28. Write details for the preceding discussion showing the falsity of Eq. (3.27).

3.29 Bessel potentials. The connection between Sobolev spaces and convolution with singular kernels is deeper than what the pointwise estimates such as Eq. (3.17) indicate. In fact, the space $W^{1,p}$ coincides with all functions u in $\mathbb{R}^n$ that are convolutions of some L^p function f with a *Bessel potential of order one*,

$$u = g_1 * f,$$

where g_1, or more generally g_α for $\alpha > 0$, is defined via its Fourier transform:

$$\hat{g}_\alpha(x) = (2\pi)^{-n/2}(1 + |x|^2)^{-\alpha/2}.$$

The Sobolev norm of u is comparable to the L^p norm of f. One computes that

$$g_\alpha(x) \sim C(\alpha)|x|^{\alpha-n} + o(|x|^{\alpha-n}), \quad |x| \to 0,$$

and

$$g_\alpha(x) \sim C(\alpha)|x|^{\frac{1}{2}(\alpha-n-1)}e^{-|x|}, \quad |x| \to \infty.$$

After this, it is natural to define Sobolev spaces $W^{\alpha,p}$ for any $\alpha > 0$ as the functions u that are convolutions of L^p functions with the kernel g_α. See [170] or [205] for more details on this approach.

3.30 Sobolev inequalities via isoperimetric inequalities. In Riemannian geometry, Sobolev inequalities play an important role as isoperimetric profile of the manifold. Let us briefly discuss this connection. Let M^n be a complete Riemannian n-manifold, and assume that, for some $\nu > 1$, the following isoperimetric inequality holds: there is a positive constant $I_\nu > 0$ so that

$$|\Omega|^{(\nu-1)/\nu} \le I_\nu |\partial\Omega| \tag{3.31}$$

for each closed, smooth submanifold Ω of M, where $|\cdot|$ denotes both the Riemannian volume and surface area. Then,

$$\|u\|_{\nu/(\nu-1)} \le I_\nu \; \|\nabla u\|_1 \tag{3.32}$$

for all smooth, compactly supported functions u on M.

Conversely, if the Sobolev inequality (3.32) holds on M for some $\nu > 1$ and $I_\nu > 0$, then inequality (3.31) holds with the same constant I_ν.

Note in particular that the dimension n plays no role in this discussion.

To prove the two statements, we need the following *coarea formula*:

$$\int_M |\nabla u|\, dx = \int_0^\infty |\{u = t\}|\, dt \tag{3.33}$$

for any smooth function u on M. (For a proof of Eq. (3.33) in $\mathbb{R}^n$, see, for example, [136] or [205]; for the general Riemannian case, see [57, p. 211].)

Assume now that inequality (3.31) holds on M. Then, by the coarea formula,

$$\begin{aligned}\int_M |\nabla u|\, dx &= \int_0^\infty |\{u = t\}|\, dt \\ &\ge I_\nu^{-1} \int_0^\infty |\{u \ge t\}|^{(\nu-1)/\nu}\, dt,\end{aligned}$$

and upon invoking the following inequality,

$$\int_0^\infty F(t)^\alpha \, dt \geq \left(\frac{1}{\alpha} \int_0^\infty F(t) t^{1/\alpha - 1} \, dt \right)^\alpha , \tag{3.34}$$

whenever $F(t)$ is a decreasing function of t and $0 < \alpha \leq 1$, we conclude that

$$\begin{aligned} \int_M |\nabla u| \, dx &\geq I_\nu^{-1} \left(\frac{\nu}{\nu - 1} \int_0^\infty |\{u \geq t\}| t^{1/(\nu-1)} \, dt \right)^{(\nu-1)/\nu} \\ &= I_\nu^{-1} \left(\int_M |u|^{\nu/(\nu-1)} \, dx \right)^{(\nu-1)/\nu} , \end{aligned}$$

as desired.

Exercise 3.35. Prove inequality (3.34).

That the Sobolev inequality (3.32) implies the isoperimetric inequality (3.31) is best explained in the language of BV functions. A locally integrable function u on a Riemannian manifold is said to be a *function of bounded variation*, or a BV *function*, if its distributional derivatives $\partial_i u$ are measures of finite total mass. This concept generalizes the concept of bounded variation from the real line, and clearly the Sobolev space $W^{1,1}(M)$ belongs to the space of BV functions. One can show, by approximation, that the Sobolev inequality (3.32) holds for BV functions:

$$\|u\|_{\nu/(\nu-1)} \leq I_\nu \, \|\nabla u\|,$$

where now $\|\nabla u\|$ denotes the total mass of the measure, which is the distributional gradient of u. Next, it turns out that the characteristic function of a smooth, bounded domain on M is a BV function and that the corresponding measure is nothing but the surface measure on the boundary of the domain.

This said, it is clear that inequality (3.32) implies the isoperimetric inequality (3.31).

Remark 3.36. If inequality (3.31) holds on a complete Riemannian manifold, then the number $\nu > 1$ is often called the *isoperimetric dimension of* M[1], and the function $s \mapsto s^{(\nu-1)/\nu}$ the *isoperimetric profile of* M. It follows from inequality (3.31) that the *volume growth* of M is at least r^ν; more precisely, if $B(x, r)$ denotes any metric ball on M, then its Riemannian volume $V(x, r)$ satisfies

$$\liminf_{r \to \infty} V(x, r) r^{-\nu} > 0. \tag{3.37}$$

To see this, one notices that $V'(x, r)$ exists for almost every $r > 0$ and

$$V'(x, r) = A(x, r), \tag{3.38}$$

[1] More precisely, ν is assumed to be the largest (alternatively, supremal) number such that (3.31) holds on M.

where $A(x, r)$ denotes the surface area of the boundary of $B(x, r)$; then, we obtain the differential inequality

$$V'(x,r) \geq I_\nu V(x,r)^{(\nu-1)/\nu}$$

valid for almost every $r > 0$, and Eq. (3.37) results by integration. To justify, Eq. (3.38), one can argue as follows: the coarea formula (3.33) remains valid for Lipschitz functions on M, and applying it to the function $u_r(y) = \min\{r, |y - x|\}$ yields

$$\int_{B(x,r)} |\nabla u_r|\, dy = \int_0^r |\{u_r = t\}|\, dt;$$

but $|\nabla u_r| = 1$ almost everywhere on $B(x, r)$, and hence Eq. (3.38) follows.

3.39 Notes to Chapter 3. Most of the material in this chapter is standard and can be found in many books. Some of them are listed in the bibliography. I recommend [45] and [205] as friendly geometric introductions to the subject; these books also contain basics of the theory of BV functions. Extensive treatments of Sobolev spaces are the monographs by Maz'ya [136] and by Adams and Hedberg [1]. Chapter 6 in [198] contains a nice elementary discussion of differential forms and Hodge theory. For the nonsmooth L^p theory, the classical reference is [142]. See [104] and [105] for further properties of weighted Sobolev spaces. The $H = W$ theorem (Exercise 3.14) is often credited to Meyers and Serrin [139], but it was already proved by Deny and Lions in [41]. The elegant proof of Proposition 3.19 is due to Hedberg [78]. The fact that the weak estimate (3.26) can be used to prove the Sobolev embedding theorem for $p = 1$ was apparently first observed by Maz'ya in the early 1960s. A recent abstract application of this idea plus more references can be found in [85]. The connection between Sobolev embedding theorems and isoperimetric inequalities was also emphasized by Fleming and Rishel. See [51] and references in [136]. For the isoperimetric inequalities in Riemannian geometry, see [25] and [30], and for extensions to sub-Riemannian settings, see [29], [36], [60], and [70].

4
Poincaré Inequality

The Sobolev inequality

$$\|u\|_{np/(n-p)} \leq C(n, p)\,\|\nabla u\|_p \tag{4.1}$$

for $1 \leq p < n$ cannot hold for an arbitrary smooth function u that is defined only, say, in a ball B. For instance, if u is a nonzero constant, the right-hand side is zero but the left-hand side is not. However, if we replace the integrand on the left-hand side by $|u - u_B|$, where, we recall, u_B is the mean value of u in the ball B, an appropriate form of inequality (4.1) is salvaged: the inequality

$$\left(\int_B |u - u_B|^{np/(n-p)}\,dx\right)^{(n-p)/np} \leq C(n, p)\left(\int_B |\nabla u|^p\,dx\right)^{1/p} \tag{4.2}$$

holds for all smooth functions u in a ball B in $\mathbb{R}^n$, if $1 \leq p < n$. Consequently, inequality (4.2) holds for all functions u in the Sobolev space $W^{1,p}(B)$. Inequality (4.2) is often called the *Sobolev–Poincaré inequality*, and it will be proved momentarily. Before that, let us derive a weaker inequality (4.4) from inequality (4.2) as follows. By inserting the measure of the ball B into the integrals, we find that

$$\left(\fint_B |u - u_B|^{np/(n-p)}\,dx\right)^{(n-p)/np} \leq C(n, p)(\operatorname{diam} B)\left(\fint_B |\nabla u|^p\,dx\right)^{1/p} \tag{4.3}$$

and hence, by Hölder's inequality, that

$$\fint_B |u - u_B|^p\,dx \leq C(n, p)(\operatorname{diam} B)^p \fint_B |\nabla u|^p\,dx, \tag{4.4}$$

which is customarily known as the *Poincaré inequality*. In fact, inequality (4.4) is valid for all $1 \leq p < \infty$.

To prove inequality (4.2), we use potential estimates as in the previous chapter, except that the pointwise inequality

$$|u(x)| \leq C(n) \int_{\mathbb{R}^n} \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy$$

for compactly supported functions u needs to be replaced with the estimate

$$|u(x) - u_B| \leq C(n) \int_B \frac{|\nabla u(y)|}{|x - y|^{n-1}} \, dy, \tag{4.5}$$

valid for all smooth functions u in a ball B and for all points x in B. To prove estimate (4.5), we use the formula

$$u(x) - u(y) = -\int_0^{|x-y|} D_r u(x + r\omega) \, dr$$

for all x and y in B, where $\omega = \frac{y-x}{|y-x|}$ is a unit vector in $\mathbb{R}^n$; by integrating this expression with respect to y, one arrives at

$$|B|(u(x) - u_B) = -\int_B \int_0^{|x-y|} D_r(x + r\omega) \, dr \, dy,$$

from which estimate (4.5) follows after a change of variables calculation.

Thus, in the language of Riesz potentials, we have that

$$|u(x) - u_B| \leq C(n) I_1(|\nabla u|)(x), \tag{4.6}$$

and thus inequality (4.2) follows from Proposition 3.19. (Note that the endpoint case $p = 1$ must be dealt with by the cutoff argument using a weak-type estimate as in Chapter 3.)

Exercise 4.7. Observe that the preceding argument by change of variables leads to an estimate of the same type as inequality (4.6), and hence to a Poincaré inequality, for smooth functions defined in any *convex* bounded open set Ω in $\mathbb{R}^n$. In this case, the constant in front of inequality (4.6) will depend also on Ω. How? Finally, in what way will the dependence on Ω show in the resulting Poincaré inequality, and is it necessary?

It is often desirable to write the pointwise estimate (4.6) in a symmetric way involving two points:

$$|u(x) - u(y)| \leq C(n)(I_1(|\nabla u|)(x) + I_1(|\nabla u|)(y)), \tag{4.8}$$

valid for smooth functions u defined in a ball B and for all pairs of points x and y in B. Clearly, inequality (4.8) follows from inequality (4.6) and the triangle inequality.

Note also that inequality (4.8) can be used to prove the Sobolev–Poincaré inequality (4.2) by integrating with respect to both x and y.

Exercise 4.9. Prove the Sobolev embedding theorem, inequality (3.12) for $p > n$, by using inequality (4.8). Also observe that inequality (4.4) follows for all $1 \leq p < \infty$.

We next describe a more geometric way (much emphasized by David and Semmes) to obtain a two-point estimate similar to inequality (4.8).

Suppose that u is a smooth function defined, say, in all of $\mathbb{R}^n$. (The ensuing discussion can easily be localized.) Pick two points x and y and observe that if γ is any piecewise smooth curve in $\mathbb{R}^n$ joining x and y, then

$$|u(x) - u(y)| \leq \int_\gamma |\nabla u|\, ds. \tag{4.10}$$

Next, suppose that we can find a family Γ of such curves γ equipped with a probability measure so that the measure

$$A \mapsto \int_\Gamma \int_{A\cap\gamma} ds\, d\gamma \tag{4.11}$$

is majorized by the measure

$$A \mapsto C(n)\left(\int_A \frac{dz}{|x-z|^{n-1}} + \int_A \frac{dz}{|y-z|^{n-1}}\right), \tag{4.12}$$

for each Borel set $A \in \mathbb{R}^n$ and for some constant $C(n) \geq 1$. (We assume that the measure on Γ is nice enough so that the set function (4.11) determines a Borel measure on $\mathbb{R}^n$.) Then, in particular, the function

$$\gamma \mapsto \int_\gamma |\nabla u|\, ds$$

is measurable on Γ, and inequality (4.10) implies

$$\begin{aligned} |u(x) - u(y)| &\leq \textstyle\int_\Gamma \int_\gamma |\nabla u|\, ds\, d\gamma \\ &\leq C(n)\left(\int_{\mathbb{R}^n} \frac{|\nabla u(z)|}{|x-z|^{n-1}}\, dz + \int_{\mathbb{R}^n} \frac{|\nabla u(z)|}{|y-z|^{n-1}}\, dz\right), \end{aligned} \tag{4.13}$$

which is, in essence, estimate (4.8), except that we made the a priori assumption that u is defined everywhere in $\mathbb{R}^n$.

Thus, we have reduced the proof of the Poincaré inequality to a geometric problem about existence of a "thick" family of curves between points, where the thickness requirement is given by the relationship between the two measures (4.11) and (4.12). This is an important theoretical point, because, as we have seen, the step from potential estimates, such as estimate (4.13), to integral inequalities, such as inequality (4.2), is valid quite generally. Moreover, as we will soon see, the compactly supported case of the Sobolev inequality can be deduced from a two-point estimate (4.13).

In $\mathbb{R}^n$, the existence of a required thick family of curves joining any two points x and y is easy to verify. Namely, let Γ be a family of curves that start from x forming a space-angle opening less than or equal to some fixed number and lying symmetrically about the line segment $[x, y]$ from x to y; when these curves reach the hyperplane that is orthogonal to $[x, y]$ and lies half-way between x and y, they are told to go down to y symmetrically with respect to the hyperplane. Thus, Γ is a sort of "pencil" of curves from x to y; the measure we place on Γ is the angular measure properly normalized so that the total mass is 1.

Exercise 4.14. Show that with the preceding choice of Γ, equipped with the normalized angular measure, measure (4.12) majorizes measure (4.11).

The preceding approach to potential estimates works equally well for functions defined on a ball B, except that the pencils must be defined slightly differently for points near the boundary of the ball. Rather than doing this, we observe that for points x and y in the half ball $\frac{1}{2}B$, we can always use pencils of the same kind. Then, the argument leads to an a priori weaker Poincaré inequality, where one has $\frac{1}{2}B$ on the left-hand side. But the estimate is still uniform in that the constant in front is independent of the ball, and this uniformity can be used to iterate to get one back to the strong inequality (4.2), as will be explained next.

There are numerous situations in analysis, where a priori weak inequalities can be shown to self-improve, provided there is a certain uniformity present in these estimates. In connection with Poincaré inequalities, this phenomenon was probably first observed by Jerison [99]. A very simple and at the same time very general formulation of Jerison's result was given by Hajłasz and Koskela [74]. We next describe and prove their result.

Definition 4.15. Given numbers $\lambda \geq 1$, $M \geq 1$, and $a > 1$, a bounded subset A of a metric space X is said to satisfy a *(λ, M, a)-chain condition (with respect to a ball B_0)* if for each point x in A there is a sequence of balls $\{B_i : i = 1, 2, ..\}$ such that

1. $\lambda B_i \subset A$ for all $i \geq 0$;
2. B_i is centered at x for all sufficiently large i;
3. the radius r_i of B_i satisfies

$$M^{-1}a^{-i}\operatorname{diam} A \leq r_i \leq Ma^{-i}\operatorname{diam} A$$

   for all $i \geq 0$; and
4. the intersection $B_i \cap B_{i+1}$ contains a ball B_i' such that $B_i \cup B_{i+1} \subset MB_i'$ for all $i \geq 0$.

For example, a ball in Euclidean space $\mathbb{R}^n$ satisfies the (λ, M, a)-chain condition for any $\lambda \geq 1$ and for some constants $M = M(\lambda)$, $a = a(\lambda)$. More generally, if X

is a *geodesic space*, which means that every pair of points in X can be joined by a curve whose length is equal to the distance between the points, then each ball in X satisfies a $(\lambda, M_\lambda, a_\lambda)$-chain condition for each $\lambda \geq 1$. This is easy to see. (Compare Exercise 9.16.)

We record the following simple consequence of the doubling condition: Let (X, μ) be a doubling measure space and let $C(\mu)$ be the doubling constant of μ. Then, for all balls B_r, centered at a set $A \subset X$ with radius $r < \operatorname{diam} A$, we have that

$$\frac{\mu(B_r)}{\mu(A)} \geq 2^{-s}\left(\frac{r}{\operatorname{diam} A}\right)^s, \tag{4.16}$$

where $s = \log_2 C(\mu) > 0$.

Exercise 4.17. Prove formula (4.16).

Next, we will formulate and prove the Hajłasz–Koskela theorem.

Theorem 4.18. *Let (X, μ) be a doubling space and suppose that A is a subset of X satisfying a (λ, M, a)-chain condition. Suppose further that formula (4.16) holds for some $s > 1$. If u and g are two locally integrable functions on A, with g nonnegative, satisfying*

$$⨍_B |u - u_B|\,d\mu \leq C(\operatorname{diam} B)\left(⨍_{\lambda B} g^p\,d\mu\right)^{1/p} \tag{4.19}$$

for some $1 \leq p < s$, for some $C \geq 1$, and for all balls B in X for which $\lambda B \subset A$, then for each $q < ps/(s-p)$ there is a constant $C' \geq 1$ depending only on q, p, s, λ, M, a, C, and the doubling constant of μ, such that

$$\left(⨍_A |u - u_A|^q\,d\mu\right)^{1/q} \leq C'(\operatorname{diam} A)\left(⨍_A g^p\,d\mu\right)^{1/p}. \tag{4.20}$$

Remark 4.21. Theorem 4.18 allows for a more general formulation, where one assumes that inequality (4.19) holds for a number $p > 0$; then, the conclusion holds for the same p instead of q, and if p is less than one, u_A must be replaced by u_{B_0}. See [74]. Furthermore, if the pair (u, g) satisfies a "truncation property" (satisfied, for example, by $u \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ and $g = |\nabla u|$), then one can choose $q = ps/(p-s)$ in Theorem 4.18. See [75, Theorem 5.1 and Theorem 9.7] for details.

We need the following lemma.

Lemma 4.22. *Let (X, μ) be a measure space and let u be a measurable function on X. If $s > 1$ and if*

$$\mu(\{|u| > t\}) \leq C_0 t^{-s},$$

then for each $q < s$ we have

$$\|u\|_q \le \left(\frac{s}{s-q}\right)^{1/q} C_0^{1/s} \mu(X)^{(s-q)/sq}.$$

Exercise 4.23. Prove Lemma 4.22.

PROOF OF THEOREM 4.18. We may assume that $u_A = 0$. We claim that for any given $0 < \epsilon < 1$, the following weak-type estimate holds:

$$\mu(\{|u| > t\}) \le C(\epsilon)\left(t^{-p} d^p \mu(A)^{-r} \int_A g^p \, d\mu\right)^{1/(1-r)}, \tag{4.24}$$

where $d = \operatorname{diam} A$ and $r = p(1-\epsilon)/s$. Then, the claim follows from Lemma 4.22 upon a computation.

To prove estimate (4.24), pick a point $x \in \{|u| > t\}$ such that x is a Lebesgue point of u; almost every point is such a point. By using the balls $\{B_0, B_1, \ldots\}$ as in the definition for the chain in Definition 4.15, we have

$$\begin{aligned} t < |u(x)| &\le C \sum_{i \ge 0} \fint_{B_i} |u - u_{B_i}| \, d\mu \\ &\le C \sum_{i \ge 0} r_i \Big(\fint_{\lambda B_i} g^p \, d\mu\Big)^{1/p}, \end{aligned} \tag{4.25}$$

because the mean values

$$u_{B_i} = \fint_{B_i} u \, d\mu$$

converge to $u(x)$ as $i \to \infty$. In Eq. (4.25),

$$M^{-1} a^{-i} d \le r_i \le M a^{-i} d,$$

so the sum on the right-hand side is at least

$$t = C(\epsilon) t \sum_{i \ge 0} a^{-i\epsilon} \ge C(\epsilon) t d^{-\epsilon} \sum_{i \ge 0} r_i^{\epsilon}$$

for some positive constant $C(\epsilon)$. In particular, we can find an index i_x such that

$$r_{i_x}^{\epsilon} t d^{-\epsilon} \le C(\epsilon) r_i \left(\fint_{B'_{i_x}} g^p \, d\mu\right)^{1/p}, \tag{4.26}$$

where B'_{i_x} is a ball, concentric with B_{i_x}, but with radius multiplied by a number depending on M such that $x \in B'_{i_x}$; this can be done by the properties of the chain. (Note that B'_{i_x} may intersect the complement of A, but we can think of g being defined everywhere by putting it zero on $X \setminus A$.)

Now, claim (4.24) follows easily by collecting the information from formula (4.16) and inequality (4.26), and by invoking the basic covering theorem.

This completes the proof of Theorem 4.18. □

Exercise 4.27. Suppose that you know the validity of the Poincaré inequality (4.4) in $\mathbb{R}^n$. Explain how Theorem 4.18 and its proof can be used to obtain the Sobolev embedding theorem, inequality (3.11), at least when the true Sobolev conjugate $np/(n-p)$ is replaced by any number $r < np/(n-p)$. Discuss more generally what kind of inequalities one can have for compactly supported functions on a doubling space assuming Theorem 4.18.

4.28 Notes to Chapter 4. The approach to Poincaré inequalities by way of thick curve families has been emphasized by David and Semmes [38], [160]; the method can be used very generally. Jerison's theorem, which was alluded to before Definition 4.15, has been applied widely and has been subject to many extensions. As mentioned in the text, Theorem 4.18 and its proof are due to Hajłasz and Koskela [74]. Chain conditions as in Definition 4.15 are now standard in analysis; they are often known as John-type conditions after F. John. See [14], [24], [36], and [75], for example; [75] in particular contains an extensive bibliography on these matters.

5
Sobolev Spaces on Metric Spaces

We have seen that if u is a smooth function defined on a ball B in $\mathbb{R}^n$ (possibly with infinite radius so that $B = \mathbb{R}^n$), then the inequality

$$|u(x) - u(y)| \le C(n)(I_1|\nabla u|(x) + I_1|\nabla u|(y)) \tag{5.1}$$

holds for each pair of points x, y in B, where I_1 is the Riesz potential. It is easily seen that in the definition of I_1 we can integrate $|\nabla u|$ against the Riesz kernel $|z|^{1-n}$ over a ball whose radius is roughly $|x - y|$ and still retain inequality (5.1); then, $I_1|\nabla u|(x)$ is controlled by a constant $C(n)$ times

$$|x - y|M|\nabla u|(x),$$

as is easily seen by dividing the ball over which the integration occurs into annuli as in the proof of Proposition 3.19. By symmetry, there is a similar bound for $I_1|\nabla u|(y)$, and we therefore conclude that

$$|u(x) - u(y)| \le C(n)|x - y|(M|\nabla u|(x) + M|\nabla u|(y)) \tag{5.2}$$

for each pair of points x, y in B. If u belongs to $W^{1,p}(B)$, so that its gradient is in $L^p(B)$, and if $p > 1$, we conclude from the maximal function theorem that

$$|u(x) - u(y)| \le |x - y|(g(x) + g(y)) \tag{5.3}$$

for each pair of points x, y in B, where $g \in L^p(B)$; in fact, we can choose g in inequality (5.3) to be a constant times the maximal function $M|\nabla u|$. By the density of smooth functions in $W^{1,p}(B)$, we thus obtain that inequality (5.3) continues to hold almost everywhere in B (in the sense that by ruling out a set of measure zero, inequality (5.3) holds for all points x and y outside this set).

This reasoning can be reversed, as we will see momentarily: an L^p function u in a ball B that satisfies inequality (5.3) belongs to $W^{1,p}(B)$, and this assertion holds for all $p \geq 1$.

This observation led Hajłasz to define Sobolev spaces in an arbitrary metric measure space.

5.4 Sobolev spaces of Hajłasz. Let (X, μ) be a metric measure space. For $1 \leq p < \infty$, define

$$M^{1,p}(X) = \{u \in L^p(X) : \text{there exists } g \in L^p(X) \\ \text{so that inequality (5.3) holds a.e.}\}. \tag{5.5}$$

The space $M^{1,p}(X)$ is obviously a linear space[1], and it becomes a normed space with

$$\|u\|_{1,p} = \|u\|_p + \inf \|g\|_p, \tag{5.6}$$

where the infimum is taken over all g satisfying the defining inequality (5.3). Strictly speaking, of course, the members in $M^{1,p}(X)$ are equivalence classes of L^p functions. Also note that upon redefining both u and g in a set of measure zero, we may assume that inequality (5.3) holds everywhere in X.

We will see later that the Hajłasz–Sobolev space $M^{1,p}(X)$ coincides with the standard Sobolev space $W^{1,p}(X)$, if $1 < p < \infty$ and X is a smooth, bounded domain in $\mathbb{R}^n$. However, in general $M^{1,p}(X)$ is a smaller space than $W^{1,p}(X)$ if X is a domain in $\mathbb{R}^n$. Indeed, it follows directly from the definition that, for any metric measure space X, $M^{1,p}(X) = M^{1,p}(X \setminus E)$ whenever E is a measure zero subset of X. No statement like this is true for the ordinary Sobolev space $W^{1,p}$.

Theorem 5.7. *The space $M^{1,p}(X)$ is a Banach space for all $1 \leq p < \infty$.*

PROOF. If (u_n) is a Cauchy sequence in $M^{1,p}(X)$, then (u_n) converges to some function u in L^p, and the burden of proof rests on showing that (u_n) converges to u also in $M^{1,p}(X)$. By passing to a subsequence, we may assume that

$$\|u_{n+1} - u_n\|_{1,p} < 2^{-n}$$

and that $u_n \to u$ almost everywhere. Thus, there exists $g_n \in L^p$ such that

$$|(u_{n+1} - u_n)(x) - (u_{n+1} - u_n)(y)| \leq |x - y|(g_n(x) + g_n(y))$$

almost everywhere, and that

$$\|g_n\|_p < 2^{-n}.$$

---

[1] The letter M is chosen here to designate the Hajłasz–Sobolev space because of its connection with maximal functions.

We now find that

$$|(u_{n+k} - u_n)(x) - (u_{n+k} - u_n)(y)| \leq |x - y| \left(\sum_{i=n}^{\infty} g_i(x) + \sum_{i=n}^{\infty} g_i(y) \right)$$

for all $k \geq 1$, so that by letting $k \to \infty$, we arrive at the inequality

$$|(u - u_n)(x) - (u - u_n)(y)| \leq |x - y| \left(\sum_{i=n}^{\infty} g_i(x) + \sum_{i=n}^{\infty} g_i(y) \right),$$

valid almost everywhere. Because the L^p norm of the sum

$$\sum_{i=n}^{\infty} g_n$$

is no more than 2^{-n+1}, we obtain that $u \in M^{1,p}(X)$ and that $u_n \to u$ in $M^{1,p}(X)$. This completes the proof. □

Remark 5.8. Standard reasoning of functional analysis shows that if $p > 1$, then the infimum in Eq. (5.6) is attained for a unique function g_u in $L^p(X)$, which still satisfies inequality (5.3) almost everywhere. In fact, if (g_i) is a minimizing sequence, then it is a bounded sequence in a reflexive Banach space and hence contains a weakly convergent subsequence [204, Theorem 1, p. 126]. A theorem of Mazur [204, Theorem 2, p. 120] then asserts that a sequence of convex combinations of the functions g_i converges in L^p to a function g_0, and because inequality (5.3) continues to hold for each convex combination, we conclude that it holds for g_0 as well. Moreover, this limit g_0 is unique as L^p spaces are uniformly convex for $p > 1$, and we rename it g_u.

The aforementioned fails if $p = 1$. For example, in a two-point space, with unit distance and counting measure, each function g from the family $\{(t, 1 - t) : 0 \leq t \leq 1\} \subset \mathbb{R}^2$ minimizes Eq. (5.6) when u is the characteristic function of either point.

Note that if $X = B$, a ball in $\mathbb{R}^n$ of perhaps infinite radius, and if $u \in W^{1,p}(B)$ with $p > 1$, then

$$\|g_u\|_p \leq C(n, p) \|\nabla u\|_p$$

by inequality (5.2) so that the imbedding $W^{1,p}(B) \hookrightarrow M^{1,p}(B)$ is continuous. We show next that actually, in this case, $W^{1,p}(B)$ and $M^{1,p}(B)$ are isomorphic.

Proposition 5.9. *For any open set Ω in $\mathbb{R}^n$, equipped with the Euclidean metric, the Sobolev space $M^{1,p}(\Omega)$ embeds continuously to $W^{1,p}(\Omega)$ for all $1 < p < \infty$.*

Corollary 5.10. *If B is a ball in $\mathbb{R}^n$, with possibly infinite radius, and if $1 < p < \infty$, then*

$$M^{1,p}(B) = W^{1,p}(B)$$

in the sense that an L^p function u belongs to $M^{1,p}(B)$ if and only if it belongs to $W^{1,p}(B)$, and the corresponding Banach norms are comparable.

Proof of Proposition 5.9. Take a function $u \in M^{1,p}(\Omega)$ with $p > 1$. We must show that, as a distribution, each of the partial derivatives $\partial_i u$ corresponds to an L^p function v_i with an appropriate norm bound; that is, for each $i = 1, \dots, n$ there should exist $v_i \in L^p(\Omega)$ such that

$$\|v_i\|_p \leq C \, \|g_u\|_p \,,$$

where $C \geq 1$ is independent of u, and that

$$\langle \partial_i u, \phi \rangle = -\int_\Omega u \partial_i \phi \, dx = \int_\Omega v_i \phi \, dx \tag{5.11}$$

for all compactly supported smooth functions ϕ in Ω. For the latter, it suffices to show that $\partial_i u$ determines a continuous functional on the dense subspace $C_0^\infty(\Omega)$ of the space $L^{p/(p-1)}(\Omega)$ by the formula in the middle in Eq. (5.11). To see this, consider the convolution approximations $u_\epsilon = \psi_\epsilon * u$ in Ω; the functions u_ϵ are smooth and satisfy

$$\begin{aligned} -\int_\Omega u \partial_i \phi \, dx &= -\lim_{\epsilon \to 0} \int_\Omega u_\epsilon \partial_i \phi \, dx \\ &= \lim_{\epsilon \to 0} \int_\Omega \partial_i u_\epsilon \phi \, dx \\ &= \lim_{\epsilon \to 0} \int_\Omega (\partial_i \psi_\epsilon) * u \phi \, dx. \end{aligned} \tag{5.12}$$

On the other hand,

$$((\partial_i \psi_\epsilon) * u)(x) = (\partial_i \psi_\epsilon) * (u - u_{B(x,\epsilon)})(x),$$

so that Eq. (5.12) and the definition of $M^{1,p}(\Omega)$ imply

$$\begin{aligned} \left| \int_\Omega u \partial_i \phi \, dx \right| &\leq \max |\partial_i \psi_\epsilon| \int_\Omega \left| \int_{B(x,\epsilon)} \right| u - u_{B(x,\epsilon)} \, |dy \phi(x)| \, dx + o(\epsilon) \\ &\leq C \epsilon^{-n-1} \int_\Omega \int_{B(x,\epsilon)} |u - u_{B(x,\epsilon)}| \, dy \, |\phi(x)| \, dx + o(\epsilon) \\ &\leq C \int_\Omega \fint_{B(x,\epsilon)} g_u \, dy \, |\phi(x)| \, dx + o(\epsilon) \\ &\leq C \int_\Omega M g_u(x) |\phi(x)| \, dx + o(\epsilon) \end{aligned}$$

by the properties of the convolution kernels ψ_ϵ. Because Mg belongs to $L^p(\Omega)$ by the maximal function theorem, the last integral is less than a constant times

$$\|g_u\|_p \, \|\phi\|_{p/(p-1)} \,.$$

This shows both that $u \in W^{1,p}(\Omega)$ and that $\|\nabla u\|_p \leq C\,\|g_u\|_p$, as required. The proof is thereby complete. □

5.13 Remark on the case $p = 1$. The preceding proof for Proposition 5.9 does not work for $p = 1$, but the assertion still remains true. Namely, an alternative way to prove Proposition 5.9 is to invoke a characterization of $W^{1,p}(\Omega)$ as those L^p functions u in Ω that have a representative that is absolutely continuous on almost every line parallel to coordinate axes; then, the partial derivatives $\partial_i u$ exist almost everywhere, and part of the requirement here is that these partials be L^p-integrable in Ω. For this characterization of $W^{1,p}(\Omega)$, which is valid for *any* open set Ω and $1 \leq p < \infty$, see, for example, [205, p. 44].

Now, the proof of Proposition 5.9 for all $p \geq 1$ can be reduced to the one-dimensional case, as is easily seen. Thus, pick a finite line segment $I \subset \mathbb{R}$ and assume that u is an integrable function on I such that

$$|u(x) - u(y)| \leq |x - y|(g(x) + g(y))$$

everywhere on I for some integrable function g on I. (Recall that we may assume without loss of generality that inequality (5.3) holds for each pair of points x and y.) We want to show that u is absolutely continuous on I, and for this we show that

$$|u(a) - u(b)| \leq 4 \int_a^b g(x)\,dx \tag{5.14}$$

for almost every a and b on I, with $a < b$.

To this end, fix a and b such that $g(a)$ and $g(b)$ are finite numbers, and divide $[a, b]$ into n subintervals $I_1, \dots, I_n$ of equal length $(b-a)/n$; now each I_i contains a point x_i such that

$$g(x_i) \leq \fint_{I_i} g\,dx,$$

so that (with the convention $a = x_0$ and $b = x_{n+1}$)

$$\begin{aligned}
|u(a) - u(b)| &\leq \sum_{i=0}^{n} |u(x_i) - u(x_{i+1})| \\
&\leq \frac{2(b-a)}{n} \sum_{i=0}^{n} (g(x_i) + g(x_{i+1})) \\
&\leq \frac{4(b-a)}{n} \sum_{i=0}^{n} \fint_{I_i} g\,dx + \frac{2(b-a)}{n}(g(a) + g(b)),
\end{aligned}$$

which implies inequality (5.14) upon letting $n \to \infty$.

Note that the preceding proof gives the estimate

$$\|\nabla u\|_p \leq 4\sqrt{n}\,\|g_u\|_p$$

for all $1 \leq p < \infty$ and $u \in M^{1,p}(\Omega)$.

It is not true, however, that $M^{1,1}(B) = W^{1,1}(B)$, even for balls. To see this, consider the interval $I = (-1/4, 1/4)$ in $\mathbb{R}$. It is easy to see that the function

$$u(x) = \frac{-x}{|x| \log |x|}$$

belongs to $W^{1,1}(I)$ but not to $M^{1,1}(I)$.

Next, we show that a Poincaré inequality is valid for functions in the Sobolev space $W^{1,p}(X)$; this turns out to be a direct consequence of the definition.

Theorem 5.15. *Let (X, μ) be a metric measure space with $\mu(X) < \infty$. Then, for all $p \geq 1$ and for all functions $u \in M^{1,p}(X)$, we have that*

$$\int_X |u - u_X|^p \, d\mu \leq 2^p (\operatorname{diam} X)^p \int_X g^p \, d\mu \tag{5.16}$$

whenever $g \geq 0$ is a function such that inequality (5.3) holds.

Proof. We simply integrate inequality (5.3) first with respect to y,

$$\begin{aligned} |u(x) - u_X| &= \left| ⨍_X (u(x) - u(y)) \, d\mu(y) \right| \\ &\leq \operatorname{diam} X ⨍_X (g(x) + g(y)) \, dy \\ &= \operatorname{diam} X \, (g(x) + g_X), \end{aligned}$$

and then with respect to x,

$$\begin{aligned} \int_X |u(x) - u_X|^p \, d\mu(x) &\leq 2^{p-1} (\operatorname{diam} X)^p \left(\int_X g^p \, d\mu + \int_X g^p \, d\mu \right) \\ &\leq 2^p (\operatorname{diam} X)^p \int_X g^p \, d\mu. \end{aligned}$$

This proves the theorem. □

Note that the Poincaré inequality (5.16) is interesting only if X has finite diameter.

Remark 5.17. (a) Earlier in this chapter, we mentioned that in general the spaces $M^{1,p}(\Omega)$ and $W^{1,p}(\Omega)$ for a domain Ω in $\mathbb{R}^n$ need not be the same. This can also be seen by the aid of Theorem 5.15, because there are many domains Ω where the Poincaré inequality does not hold. For example, if Ω is a domain consisting of two squares of fixed size in the plane, connected by a very narrow passage of width ϵ, then the Poincaré inequality holds but with a constant that blows up as $\epsilon \to 0$. By combining infinitely many such square pairs with appropriately narrowing passages, one constructs a domain Ω where the Poincaré inequality fails.

The (non)validity of a Poincaré inequality in a domain does not necessarily tell the whole truth about the (non)equality of the spaces $M^{1,p}(\Omega)$ and $W^{1,p}(\Omega)$. Namely, the Poincaré inequality (4.2) holds for $W^{1,p}$ functions in a disk with a radius removed; that is, in a "slit disk" B'. This can be seen as in Chapter 4 by using integration by parts. In fact, a global Poincaré inequality holds in any bounded *starlike domain* in $\mathbb{R}^n$. However, by the reasoning in Remark 5.13, every function in the space $M^{1,p}(B')$ must be absolutely continuous along almost every vertical line in the disk B, no matter if these lines cross the slit or not. This shows, in fact, that $M^{1,p}(B') = M^{1,p}(B)$. On the other hand, the Sobolev space $W^{1,p}(B')$ contains many more functions than $W^{1,p}(B)$, as is easily seen.

More to the point, if Ω is of the form $\Omega = B \setminus E$, where B is a ball and E is a closed set of measure zero in B such that E is not a removable set for Sobolev functions, then $M^{1,p}(\Omega)$ is strictly contained in $W^{1,p}(\Omega)$. Although not apparent, there is a close connection between Poincaré inequalities and removable sets for Sobolev functions; see [112].

(b) One possible approach to Sobolev spaces in an arbitrary metric measure space is to consider functions u for which an L^p function g can be found so that a Poincaré inequality such as inequality (5.16) holds (not just globally but uniformly on all balls in the space). This approach has been pursued in [75].

A domain $\Omega \subset \mathbb{R}^n$ is called a *$W^{1,p}$-extension domain* if for every Sobolev function $u \in W^{1,p}(\Omega)$ there is a Sobolev function $\tilde{u} \in W^{1,p}(\mathbb{R}^n)$ such that $u = \tilde{u}$ almost everywhere in Ω and that

$$\|\tilde{u}\|_{1,p} \le C(\Omega)\, \|u\|_{1,p}\,,$$

where the norm on the left-hand side is the Sobolev norm in $W^{1,p}(\mathbb{R}^n)$ and that on the right-hand side is the Sobolev norm in $W^{1,p}(\Omega)$. Every bounded domain with smooth boundary is an extension domain, but an extension domain can be quite complicated. See Example 8.24 (f).

Exercise 5.18. Verify that $M^{1,p}(\Omega)$ and $W^{1,p}(\Omega)$ are isomorphic for any $W^{1,p}$-extension domain Ω if $p > 1$ in the sense of Corollary 5.10.

We close this chapter by proving an "$H = W$ theorem" for spaces $M^{1,p}(X)$. The ensuing theorem also states that functions in $M^{1,p}(X)$ agree with Lipschitz functions outside a set of arbitrarily small measure; this is a Lusin type theorem for Sobolev functions. The discussion assumes an extension theorem for Lipschitz functions, proved in the next chapter. (For the definition of Lipschitz functions, also see Chapter 6.)

Theorem 5.19. *Given a function $u \in M^{1,p}(X)$, $1 \le p < \infty$, and $\epsilon > 0$, there is a Lipschitz function ϕ on X such that*

$$\mu(\{u \neq \phi\}) < \epsilon$$

and that

$$\|u - \phi\|_{1,p} < \epsilon.$$

PROOF. Let $g \geq 0$ be a function satisfying inequality (5.3) with L^p norm essentially the infimum, as in the definition (5.6). Then,

$$\lambda^p \mu(X \setminus E_\lambda) \to 0$$

as $\lambda \to \infty$, where

$$E_\lambda = \{x \in X : |u(x)| \leq \lambda,\ g(x) \leq \lambda\}.$$

Moreover, on E_λ the function u is 2λ-Lipschitz, and by Theorem 6.2 it can be extended to a 2λ-Lipschitz function in all of X. Denote this extension by u_λ, and put

$$v_\lambda = \operatorname{sgn} u_\lambda \min\{|u_\lambda|, \lambda\}.$$

Then, v_λ is still a Lipschitz function in all of X with the same Lipschitz constant 2λ as u_λ, and v_λ continuous to agree with u on E_λ. It is easy to see that

$$v_\lambda \to u$$

in $L^p(X)$ as $\lambda \to \infty$. It remains to observe that the function

$$g_\lambda = g\chi_{E_\lambda} + 2\lambda\chi_{X\setminus E_\lambda}$$

satisfies

$$|v_\lambda(x) - v_\lambda(y)| \leq |x - y|(g_\lambda(x) + g_\lambda(y))$$

almost everywhere and that g_λ belongs to $L^p(X)$ and converges to g in $L^p(X)$. The theorem follows. □

Remark 5.20. Although, for a given $u \in M^{1,p}(X)$, the function g_u (assume here $p > 1$ for simplicity so that the minimizer g_u for Eq. (5.6) exists) behaves like a derivative of u in many ways, the similarities are always in the global behavior. For instance, it is easy to see that g_u hardly ever vanishes on places where u is constant, which is an important feature of an honest derivative. To alleviate this defect, one can try to modify the definition for $M^{1,p}(X)$ by making requirement (5.3) local.

For a metric measure space (X, μ), and for $1 \leq p < \infty$, let $M^{1,p}_{\mathrm{loc}}(X)$ denote the collection of functions $u \in L^p_{\mathrm{loc}}$ for which there exists $g \in L^p_{\mathrm{loc}}$ so that inequality (5.3) holds locally almost everywhere. Here, "locally almost everywhere" means that every point in X has a neighborhood where inequality (5.3) holds almost everywhere. Recall that $L^p_{\mathrm{loc}} = L^p_{\mathrm{loc}}(X)$ consists of those functions u such that every point in X has a neighborhood where u is p-integrable.

Another possibility is an even weaker requirement: define a "weak local $M^{1,p}$ space" by letting $M^{1,p}_{w,\mathrm{loc}}(X)$ consist of those functions $u \in L^p_{\mathrm{loc}}$ for which every point in X has a neighborhood U such that $u \in M^{1,p}(U)$. It is obvious that

$$M^{1,p}_{\mathrm{loc}}(X) \subset M^{1,p}_{w,\mathrm{loc}}(X),$$

and an exercise shows that these two spaces are in fact equal; see Exercise 5.22.

Exercise 5.21. Prove that for an open set $\Omega \subset \mathbb{R}^n$, and for $1 < p < \infty$,

$$M^{1,p}_{w,\mathrm{loc}}(\Omega) = M^{1,p}_{\mathrm{loc}}(\Omega) = W^{1,p}_{\mathrm{loc}}(\Omega),$$

in the sense that if a function $u \in L^p_{\mathrm{loc}}(\Omega)$ belongs to one of the spaces it belongs to the two others. (Recall that $W^{1,p}_{\mathrm{loc}}(\Omega)$ is the set of functions in $L^p_{\mathrm{loc}}(\Omega)$ that belong to $W^{1,p}(U)$ for every open set U whose closure lies in Ω.)

Exercise 5.22. Show that $M^{1,p}_{w,\mathrm{loc}}(X) = M^{1,p}_{\mathrm{loc}}(X)$ always. (Hint: Every metric space is paracompact, which means that every open covering of the space has a locally finite refinement.)

Exercises 5.21 and 5.22 were suggested by Marius Dabija and Nageswari Shanmugalingam.

Even with the preceding localized definitions for the space $M^{1,p}$, one cannot eliminate the problem that inequality (5.3) inevitably has a global character. This is demonstrated by the following example.

Example 5.23. One can choose two sequences of pairwise disjoint closed balls (B_i) and (B_i') in $\mathbb{R}^n$ converging to the origin, and a function u in the Sobolev space $W^{1,p}(\mathbb{R}^n)$ for each $1 \leq p \leq n$, such that u is continuous outside the origin and takes the value 1 on each B_i and value -1 on each B_i'. Thus, in no neighborhood of the origin can a function g as in inequality (5.3) vanish on the balls B_i and B_i'.

Exercise 5.24. Show that a function u as in Example 5.23 exists.

5.25 Notes to Chapter 5. This chapter is taken for the most part from the two papers [71], [72] of Hajłasz. The theory of Hajłasz–Sobolev spaces is currently taking shape, and it would be unreasonable to go deeper into this study here. Papers on the topic include [73], [75], [76], [106], [108], [113], and [167].

6

Lipschitz Functions

Lipschitz functions are the smooth functions of metric spaces. A real-valued function f on a metric space X is said to be *L-Lipschitz* if there is a constant $L \geq 1$ such that

$$|f(x) - f(y)| \leq L|x - y| \tag{6.1}$$

for all x and y in X. Of course, there is nothing special about having the real line as a target, and in general we call a map $f : X \to Y$ between metric spaces *Lipschitz*, or *L-Lipschitz* if the constant $L \geq 1$ deserves to be mentioned, if condition (6.1) holds.

It is important to observe that in every metric space there are plenty of nontrivial real-valued Lipschitz functions (unless the space itself is somehow trivial such as a point). Namely, for each point x_0 in a metric space X, the function $x \mapsto |x_0 - x|$ is 1-Lipschitz, as readily follows from the triangle inequality. More examples are obtained by postcomposing the distance function by a Lipschitz function $\mathbb{R} \to \mathbb{R}$.

We begin with the following classical, extremely useful theorem of McShane.

Theorem 6.2. *If $f : A \to \mathbb{R}$ is an L-Lipschitz function on a subset A of a metric space X, then f extends to an L-Lipschitz function $F : X \to \mathbb{R}$.*

Theorem 6.2 requires a lemma whose simple proof we leave to the reader as an exercise.

Lemma 6.3. *If $\mathcal{F}$ is a family of L-Lipschitz functions on a metric space X, then the function*

$$F(x) = \inf\{f(x) : f \in \mathcal{F}\}$$

is L-Lipschitz on X if it is finite at one point in X.

Assuming Lemma 6.3, Theorem 6.2 follows upon defining an extension of f by the formula

$$F(x) = \inf\{f(a) + L|x - a| : a \in A\}.$$

Indeed, each function of the form $x \mapsto f(a) + L|x - a|$ is L-Lipschitz on X, as is easily seen by the triangle inequality, and the rest follows from Lemma 6.3 because clearly $F(a) = f(a)$ for $a \in A$.

If we replace the target real line by an arbitrary metric space, then Theorem 6.2 need not be true. A trivial example to this effect is a nonconstant map from the endpoints of an interval to a two point space; such a map cannot be extended even continuously to the interval. But even if there are no topological obstructions, Lipschitz extensions need not exist.

Exercise 6.4. Observe that every continuous extension of the identity map $\mathbb{S}^1 \to \mathbb{S}^1$ to the annulus $A_\epsilon = \{\epsilon \leq |z| \leq 1\}$ is at best ϵ^{-1} Lipschitz, and then create a compact, connected subset X of the plane with a compact subset $A \subset X$ such that the identity map $A \to A$ does not extend to a Lipschitz map $X \to A$ although it extends to a continuous map $X \to A$.

Note, however, that if $f : A \to \mathbb{R}^n$ is L-Lipschitz, where $A \subset X$, then f extends to an $L\sqrt{n}$-Lipschitz function $X \to \mathbb{R}^n$, as can be seen by applying Theorem 6.2 to the components of f. (For subsets A of $\mathbb{R}^m$, it is true that an L-Lipschitz extension can be found. This theorem of Kirzsbraun's is much more difficult to prove [47, 2.10.43].)

Another easy but useful observation is that every *Hölder continuous* real- or $\mathbb{R}^n$-valued function f defined on a subset A of a metric space X extends to a Hölder continuous function to all of X. Recall that a map $f : X \to Y$ between two metric spaces is *Hölder continuous of order* $0 < \alpha \leq 1$ if there is a constant $L \geq 1$ so that

$$|f(x) - f(y)| \leq L|x - y|^\alpha \tag{6.5}$$

for all x and y in X. The observation follows from Theorem 6.2 by applying it to the metric space $(X, |x - y|^\alpha)$. Thus, Lipschitz maps are just Hölder maps of order 1, and Hölder maps of order α are Lipschitz maps on the space $(X, |x - y|^\alpha)$.

Remark 6.6. A topic of current research interest is to determine for what pairs of metric spaces (X, Y) McShane's result remains valid in the sense that each Lipschitz map $f : A \to Y$ extends to a Lipschitz map $F : X \to Y$ whenever $A \subset X$. (See the notes to this chapter for references.) The answer is negative already for $Y = \ell^2$, the Hilbert space of square summable sequences; see [9, Chapter 2]. On the other hand, McShane's result implies that the answer is affirmative when Y is the Banach space $\ell^\infty(Z)$ of bounded functions on a set Z.

Exercise 6.7. Prove that if $f : A \to \ell^\infty$ is an L-Lipschitz map of a subset A of a metric space X, then f extends to an L-Lipschitz map $F : X \to \ell^\infty$.

In $\mathbb{R}^n$, the convolution approximations of a continuous function converge locally uniformly to the function. Similarly, a (uniformly) continuous function on an arbitrary metric space is a uniform limit of Lipschitz functions, as follows.

Theorem 6.8. *Every uniformly continuous bounded function in a metric space is a uniform limit of Lipschitz functions.*

PROOF. Let $f : X \to \mathbb{R}$ be a bounded uniformly continuous function; then,

$$|f(x) - f(y)| \leq \omega(|x - y|) \tag{6.9}$$

for some modulus of continuity ω. Define, for each $j = 1, 2, \ldots$, a function f_j on X by

$$f_j(x) = \inf\{f(y) + j|x - y| : y \in X\};$$

we observe that

$$-\infty < \inf f \leq f_j(x),$$

so that f_j is j-Lipschitz by Lemma 6.3. Also,

$$f_j(x) \leq f(x),$$

for all $x \in X$, so

$$j|x - y| > 2\sup|f|$$

implies

$$f_j(x) \leq f(x) - f(y) + f(y) \leq 2\sup|f| + f(y) < j|x - y| + f(y),$$

which means that in the definition for $f_j(x)$ the infimum can be taken over those y that lie within distance $j^{-1}2\sup|f|$ from x. In conclusion, for all $x \in X$ it holds that

$$0 \leq f(x) - f_j(x) \leq f(x) - f(y) \leq \omega(j^{-1}2\sup|f|).$$

This shows that $f_j \to f$ uniformly, and the theorem follows. □

With a little extra work, the following (suggested by Mattias Jonsson) can be proved.

Exercise 6.10. Show that if $f : X \to (c, \infty]$, $-\infty < c < \infty$, is a lower semicontinuous function on a metric space X, then there is an increasing sequence (f_j) of Lipschitz functions on X such that $\lim f_j(x) = f(x)$ for all $x \in X$.

Recall that a function $f : X \to (\infty, \infty]$ is *lower semicontinuous* if

$$\liminf_{x \to y} f(x) \geq f(y)$$

for each $y \in X$.

6.11 Lipschitz functions and Sobolev spaces. In our discussion of Sobolev spaces, whether the classical ones or those of Hajłasz, we have so far deliberately avoided the case $p = \infty$, although formally every definition makes sense in that case, too. For instance, the space $W^{1,\infty}(\Omega)$, for $\Omega \subset \mathbb{R}^n$ open, consists of all (essentially) bounded functions u in Ω whose distributional derivatives are (essentially) bounded functions as well. The following theorem identifies the space $W^{1,\infty}(\Omega)$ as a space of Lipschitz functions.

Theorem 6.12. *The space $W^{1,\infty}(\Omega)$ consists precisely of bounded functions that are locally uniformly Lipschitz on Ω.*

Proof. The formulation of the theorem requires an explanation. Theorem 6.12 states that every function $u \in W^{1,\infty}(\Omega)$ has a continuous representative that is L-Lipschitz continuous locally in Ω for some L. Note that u need not satisfy a global Lipschitz condition if, for instance, Ω is the slit disk B' as in Remark 5.17 (a). Conversely, a bounded function u in Ω that is L-Lipschitz continuous near every point in Ω for some fixed $L \geq 1$ is in $W^{1,\infty}(\Omega)$.

With this understood, it suffices to prove the theorem when $\Omega = B$, a ball. First, if u is in $W^{1,\infty}(B)$, then it is in every Sobolev space $W^{1,p}(B)$ for $p < \infty$, and in particular has a continuous representative by the Sobolev embedding theorem (Eq. (3.12)). The fact that this representative is Lipschitz continuous can be proved similarly; see Exercise 4.9.

Conversely, let u be a Lipschitz function on a ball B. Then, the convolution approximations u_ϵ converge to u locally uniformly in B; in fact, just the continuity of u is enough for this statement. Moreover, locally all functions u_ϵ, for small enough ϵ, satisfy a Lipschitz condition with a uniform Lipschitz constant, say L. This means, in particular, that $|\nabla u_\epsilon|$ is uniformly bounded by L (locally and for small enough ϵ), and it follows from weak* compactness of the closed balls in $L^\infty(B)$ (Alaoglu's theorem, see [154, Theorem 3.15]) that, given a smooth ϕ with compact support in B,

$$\begin{aligned} \int_B v_i \phi \, dx &= \lim_{\epsilon \to 0} \int_B \partial_i u_\epsilon \phi \, dx \\ &= -\lim_{\epsilon \to 0} \int_B u_\epsilon \partial_i \phi \, dx \\ &= -\int_B u \partial_i \phi \, dx \end{aligned} \tag{6.13}$$

for some bounded measurable function v_i, for each $i = 1, \ldots, n$. (Strictly speaking, a priori we only know that a subsequence of $(\partial_i u_\epsilon)$ converges in the weak* topology to some element in the dual space of $L^1(B)$, which is $L^\infty(B)$, but it follows from Eq. (6.13) that any such limit has to be the weak derivative $\partial_i u$, which is unique.)

Theorem 6.12 is thereby proved. □

Exercise 6.14. Prove that in the setting of Theorem 6.12 the best (local) Lipschitz constant for a function $u \in W^{1,\infty}_{loc}(\Omega)$ is $||\nabla u||_\infty$.

By the aid of Theorem 6.12, we prove the following famous theorem of Rademacher.

Theorem 6.15. *Every Lipschitz function on an open set in $\mathbb{R}^n$ is differentiable almost everywhere.*

A strengthening of Rademacher's theorem, due to Stepanoff, arrives at the same conclusion with a weaker hypothesis on the function: it suffices to assume that f satisfies

$$\limsup_{y \to x} \frac{|f(x) - f(y)|}{|x - y|} < \infty \tag{6.16}$$

for almost every x in the domain of definition, and when properly interpreted, the conclusion still remains valid if f is defined on an arbitrary set in $\mathbb{R}^n$. See [47, 3.1.8 and 3.1.9].

We derive Theorem 6.15 from the following result of Calderón, which is at the same time both more general (it requires less than Lipschitz) and more restrictive (it does not deal with the weaker assumption (6.16)) than the original Rademacher–Stepanoff theorem. In any case, the proof of Theorem 6.17 is short and elegant.

Theorem 6.17. *Every function in the Sobolev space $W^{1,p}(\Omega)$ for $p > n$ is almost everywhere differentiable.*

Proof. Because the question is purely local, we may assume that u belongs to $W^{1,p}(B)$ for some ball B. The obvious candidate for the derivative of u is the weak gradient ∇u, which is defined almost everywhere in B. We claim that ∇u is indeed the derivative of u at each Lebesgue point of ∇u; this suffices, as almost every point is a Lebesgue point. Thus, pick a point $x_0 \in B$ such that

$$\lim_{r \to 0} \fint_{B(x_0,r)} |\nabla u(x_0) - \nabla u(x)|^p \, dx = 0. \tag{6.18}$$

Then, consider the function

$$v(x) = u(x) - u(x_0) - \nabla u(x_0) \cdot (x - x_0).$$

Clearly, $v \in W^{1,p}(B)$ as well with

$$\nabla v(x) = \nabla u(x) - \nabla u(x_0)$$

for almost every $x \in B$. If we now apply the Sobolev embedding theorem, inequality (3.12) for the function v, we obtain, with $r = |y - x_0|$, that

$$
\begin{aligned}
|u(y) - u(x_0) - \nabla u(x_0)(y - x_0)| &= |v(y)| = |v(y) - v(x_0)| \\
&\leq C(n, p) r \left(\fint_{B(x,r)} |\nabla v(x)|^p \, dx \right)^{1/p} \\
&\leq C(n, p) r \left(\fint_{B(x,r)} |\nabla u(x) - \nabla u(x_0)|^p \, dx \right)^{1/p}.
\end{aligned}
$$

But this shows that

$$
\frac{|u(y) - u(x_0) - \nabla u(x_0)(y - x_0)|}{|y - x_0|} \to 0
$$

as $y \to x_0$, so that u is differentiable at x_0 with derivative equal to $\nabla u(x_0)$. The theorem follows. □

6.19 Notes to Chapter 6. The material in this chapter is classical. McShane's paper [138] on Lipschitz extensions is from the 1930s. A careful discussion of extensions can be found in [160]. Lipschitz extensions for maps with nonlinear or infinite-dimensional targets have been studied in [118] and [9], for example. For an interesting account on the Hölder condition in metric spaces, see [164]. Calderón's theorem (Theorem 6.17) perhaps is not as well known as it should be; I learned it from [15] (the original paper [27] appeared in 1951). Recently, a far-reaching extension of Rademacher's theorem to metric measure spaces has been obtained by Cheeger [31].

7

Modulus of a Curve Family, Capacity, and Upper Gradients

In Chapter 5, we discussed a possible definition for a Sobolev space in a metric measure space; this was the space $M^{1,p}$ of Hajłasz. While often convenient, it does not capture the geometry of the underlying space. This is seen, for example, in the fact that a Poincaré inequality always holds for functions in $M^{1,p}$ (Theorem 5.15). On the other hand, the validity of a Sobolev–Poincaré-type inequality in a space should tell us something about the geometry of the space, as in the discussion of the isoperimetric inequalities in Section 3.30. Now $M^{1,p}$ fails in this task because the definition of the "derivative" g of a function u in $M^{1,p}$ already is global. There is necessarily a loss of information. (In some sense, $M^{1,p}$ precisely consists of those functions for which a Poincaré inequality is satisfied [75].)

In this chapter, we discuss the notion of an "upper gradient" of a function in a metric space. This definition captures the infinitesimal behavior of the given function much as $|\nabla u|$ does in $\mathbb{R}^n$. Its main defect is that the definition is useful only in a space with plenty of rectifiable curves in it. This restriction rules out many spaces where the Hajłasz space is interesting; for example, it rules out many fractal-like spaces.

Before that, we discuss the notion of modulus of a curve family, which can be used to turn the phrase "plenty of curves" into a mathematical statement. The modulus can also be used to define important conformal invariants of a space.

Let us start with a brief discussion of line integration in a metric space.

7.1 Line integrals. A *curve* in a metric space X is a continuous map γ of an interval $I \subset \mathbb{R}$ into X. We usually abuse terminology and call γ both the map and the image $\gamma(I)$. If $I = [a, b]$ is a closed interval, then the *length* of a curve

$\gamma : I \to X$ is

$$\ell(\gamma) = \text{length}(\gamma) = \sup \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i+1})|,$$

where the supremum is over all sequences $a = t_1 \le t_2 \le \cdots \le t_n \le t_{n+1} = b$. If I is not closed, then we define the length of γ to be the supremum of the lengths of all closed subcurves of γ. A curve γ is *rectifiable* if its length is finite, and it is *locally rectifiable* if all its closed subcurves are rectifiable.

We will not distinguish among open, closed, or half-open intervals when they are domains of a curve; such distinction makes no difference in the ensuing discussion. Notice, however, that if $\gamma : I \to X$ is rectifiable and I is not closed, then γ has a unique extension to a rectifiable curve defined at the endpoints of I, where "endpoints" should be understood in the generalized sense if I is unbounded. (Strictly speaking, this extension takes values in the completion of X, but we ignore such issues here.)

Any rectifiable curve γ factors

$$\gamma = \gamma_s \circ s_\gamma, \tag{7.2}$$

where $s_\gamma : I \to [0, \ell(\gamma)]$ is the associated *length function* and $\gamma_s : [0, \ell(\gamma)] \to X$ is the unique 1-Lipschitz continuous map such that the factorization in Eq. (7.2) holds. The curve γ_s is the *arc length parametrization* of γ.

If γ is a rectifiable curve in X, the line integral over γ of a Borel function $\rho : X \to [0, \infty]$ is

$$\int_\gamma \rho\, ds = \int_0^{\ell(\gamma)} \rho \circ \gamma_s(t)\, dt.$$

If γ is only locally rectifiable, we set

$$\int_\gamma \rho\, ds = \sup \int_{\gamma'} \rho\, ds,$$

where the supremum is taken over all rectifiable subcurves γ' of γ. If γ is not locally rectifiable, no line integrals are defined.

A detailed treatment of line integrals in the case $X = \mathbb{R}^n$ can be found in [182, Chapter 1], and the general case follows with obvious modifications.

7.3 Modulus of a curve family. Let (X, μ) be a metric measure space. (Recall that all measures in this book are assumed to be Borel regular.) For a given curve family Γ in X and a real number $p \ge 1$, we define the *p-modulus* of Γ by

$$\text{mod}_p \Gamma = \inf \int_X \rho^p\, d\mu,$$

where the infimum is taken over all nonnegative Borel functions $\rho : X \to [0, \infty]$ satisfying

$$\int_\gamma \rho\, ds \ge 1 \tag{7.4}$$

for all locally rectifiable curves $\gamma \in \Gamma$. Functions ρ that satisfy inequality (7.4) are called *admissible functions*, or *metrics*, for the family Γ.

By definition, the modulus of all curves in X that are not locally rectifiable is zero. If Γ contains a constant curve, then there are no admissible functions and the modulus is infinite. Furthermore, the following properties are easily verified:

$$\operatorname{mod}_p \emptyset = 0; \tag{7.5}$$

$$\operatorname{mod}_p \Gamma_1 \leq \operatorname{mod}_p \Gamma_2 \tag{7.6}$$

if $\Gamma_1 \subset \Gamma_2$;

$$\operatorname{mod}_p \left(\bigcup_{i=1}^{\infty} \Gamma_i \right) \leq \sum_{i=1}^{\infty} \operatorname{mod}_p \Gamma_i; \tag{7.7}$$

and

$$\operatorname{mod}_p \Gamma \leq \operatorname{mod}_p \Gamma_0 \tag{7.8}$$

if Γ_0 and Γ are two curve families such that each curve $\gamma \in \Gamma$ has a subcurve $\gamma_0 \in \Gamma_0$. Thus, mod_p is an outer measure on the set of all curves in X. (In general, there are no nontrivial measurable curve families [133].)

The requirement that ρ be an arbitrary Borel function, and not for instance continuous, is essential. In fact, as we shall soon see, the n-modulus of the family of all (nonconstant) curves in $\mathbb{R}^n$ that pass through a given point is zero, but clearly there are no admissible locally bounded metrics for this family.

If X is locally compact, it follows from the *Vitali–Carathéodory theorem* [153, p. 57] that every function f in $L^p(X)$ can be approximated in $L^p(X)$ by a lower semicontinuous function g with $g \geq f$. Thus, in locally compact spaces one could restrict to lower semicontinuous test functions ρ. This observation comes in handy sometimes.

7.9 Conformal modulus. In $\mathbb{R}^n$, the most important modulus from the point of view of (quasi)conformal geometry is the n-modulus $\operatorname{mod}_n \Gamma$, which is conformally invariant. This is true more generally on Riemannian n-manifolds.

A diffeomorphism $f : M^n \to N^n$ between two Riemannian manifolds is *conformal* if at each point its tangent map is a homothety; that is,

$$\langle Df(x)X, Df(x)Y \rangle_{f(x)} = \lambda(x) \langle X, Y \rangle_x$$

at each point $x \in M$ for all tangent vectors X and Y in $T_x M$, where λ is a continuous function on M. In $\mathbb{R}^n$, this translates to the property that the derivative matrix of f is a multiple of an orthogonal transformation.

Theorem 7.10. *If $f : M^n \to N^n$ is conformal, then*

$$\operatorname{mod}_n \Gamma = \operatorname{mod}_n f\Gamma$$

for all curve families Γ in M.

PROOF. If ρ is an admissible function for $f\Gamma$, then it is easily seen that

$$\int_\gamma \rho \circ f(x)|Df(x)|\,|dx| \geq \int_{f\circ\gamma} \rho\, ds \geq 1$$

for all $\gamma \in \Gamma$, so that $\rho \circ f(x)|Df(x)|$ is admissible for Γ. Thus,

$$\operatorname{mod}_n \Gamma \leq \int_M \rho \circ f(x)^n |Df(x)|^n\, dx. \tag{7.11}$$

But the conformality of f precisely means that $|Df(x)|^n = |\det Df(x)|$, so that the integral on the right-hand side in inequality (7.11) is, by change of variables,

$$\int_N \rho(y)^n\, dy.$$

This shows that $\operatorname{mod}_n \Gamma \leq \operatorname{mod}_n f\Gamma$, and the claim follows by symmetry. □

The n-modulus on a Riemannian n-manifold is often called the *conformal modulus.*

By relaxing the invariance in Theorem 7.10 to a "quasi-invariance," we arrive at the concept of quasiconformality, as follows.

Definition 7.12. A homeomorphism $f : M^n \to N^n$ between two Riemannian n-manifolds, $n \geq 2$, is called *K-quasiconformal* if there is $K \geq 1$ so that

$$K^{-1}\operatorname{mod}_n f\Gamma \leq \operatorname{mod}_n \Gamma \leq K \operatorname{mod}_n f\Gamma \tag{7.13}$$

for all curve families Γ in M.

There are many equivalent ways to define quasiconformal maps, and the one given in Definition 7.12 is in some sense the strongest, as well as the least satisfactory. This means that many strong properties of quasiconformal maps can be derived rather directly from the definition, and that it is impractical to check quasiconformality of a given map by using the formula (7.13). We discuss other definitions later, including those that appropriately generalize the notion of quasiconformality to spaces where modulus is an empty tool.

7.14 Basic example. In $\mathbb{R}^n, n \geq 2$, the n-modulus of the curve family Γ consisting of all curves joining the two boundary components of an annular region

$$B(x_0, R) \setminus \overline{B}(x_0, r) = \{x : r < |x_0 - x| < R\}, \quad 0 < r < R,$$

is

$$\operatorname{mod}_n \Gamma = \omega_{n-1} \left(\log \frac{R}{r}\right)^{1-n}, \tag{7.15}$$

where ω_{n-1} is the area of the unit sphere in $\mathbb{R}^n$. This can be proved by first observing that the function

$$\rho(x) = \left(\log \frac{R}{r}\right)^{-1} |x_0 - x|^{-1}$$

restricted to $B(x_0, R) \setminus B(x_0, r)$ is admissible, so that the modulus is at most

$$\begin{aligned}\int_{B(x_0,R)\setminus B(x_0,r)} \rho^n(x)\,dx &= \left(\log \frac{R}{r}\right)^{-n} \int_{\mathbb{S}^{n-1}} \int_r^R t^{-1}\,dt\,d\omega \\ &= \omega_{n-1} \left(\log \frac{R}{r}\right)^{1-n}.\end{aligned}$$

On the other hand, if ρ is any admissible function, supported in $B(x_0, R)\setminus B(x_0, r)$, then for each point ω on the unit sphere $\mathbb{S}^{n-1}$ we have that

$$\begin{aligned}1 \le \int_r^R \rho(t\omega)\,dt &= \int_r^R \rho(t\omega) t^{(n-1)/n} t^{(1-n)/n}\,dt \\ &\le \left(\int_r^R \rho(t\omega)^n t^{n-1}\,dt\right)^{1/n} \left(\int_r^R t^{-1}\,dt\right)^{(n-1)/n},\end{aligned}$$

and hence

$$\int_{B(x_0,R)\setminus B(x_0,r)} \rho^n(x)\,dx \ge \omega_{n-1} \left(\log \frac{R}{r}\right)^{1-n}.$$

This proves the equality in Eq. (7.15).

Exercise 7.16. Let Γ be the curve family in Example 7.14. Show that

$$\operatorname{mod}_p \Gamma = \omega_{n-1} \left(\frac{|n-p|}{p-1}\right)^{p-1} \left| R^{\frac{p-n}{p-1}} - r^{\frac{p-n}{p-1}} \right|^{1-p} \tag{7.17}$$

if $p \neq n$, $p > 1$.

The logarithmic behavior of Eq. (7.15) persists in more general situations.

Lemma 7.18. *Suppose that (X, μ) is a metric measure space such that*

$$\mu(B_R) \le C R^n \tag{7.19}$$

for some constant $C \ge 1$, exponent $n > 1$, and for all balls B_R of radius $R > 0$. Then, for all $x_0 \in X$ and $2r < R$, we have that

$$\operatorname{mod}_n \Gamma \le C' \left(\log \frac{R}{r}\right)^{1-n},$$

where Γ denotes the family of all curves joining $\overline{B}(x_0, r)$ to $X \setminus B(x_0, R)$ and C' is a constant depending only on the value C in inequality (7.19) and on the dimension n.

PROOF. As in the proof of Eq. (7.15), define

$$\rho(x) = \left(\log \frac{R}{r}\right)^{-1} |x_0 - x|^{-1}$$

for $r < |x_0 - x| < R$ and $\rho(x) = 0$ elsewhere. Then, ρ is readily seen to be admissible for Γ. Now, if $k \geq 1$ is the least integer so that $2^{k+1} r \geq R$, we find by using inequality (7.19) that

$$\begin{aligned}\int_X \rho^n \, d\mu &\leq \left(\log \frac{R}{r}\right)^{-n} \sum_{j=0}^{k} \int_{\{2^j r \leq |x_0 - x| < 2^{j+1} r\}} |x_0 - x|^{-n} \, d\mu(x) \\ &\leq C' \left(\log \frac{R}{r}\right)^{1-n}.\end{aligned}$$

This proves the lemma. □

By invoking properties (7.7) and (7.8), we easily obtain the following corollary.

Corollary 7.20. *In a metric measure space as in Lemma 7.18, in particular in $\mathbb{R}^n$, the n-modulus of the family of all nonconstant curves passing through a fixed point is zero.*

The conclusion of Corollary 7.20 need not hold in general. For instance, the p-modulus of the family of all curves in $\mathbb{R}^n$ passing through a point is not zero if $p > n$ by Eq. (7.17).

Lemma 7.18 gives an upper bound for modulus. Such bounds are easier to obtain than lower bounds, for the obvious reason that every nontrivial choice of an admissible function ρ gives rise to an upper bound. Another obvious reason is that it may, in general, be the case that no (nonconstant) curve family in X has positive modulus; this happens, for instance, if X admits no rectifiable curves. On the other hand, quantitative lower bounds for modulus, if they exist, lead to important applications. We discuss this more fully in the next chapter.

7.21 Historical remark. The conformal modulus has been an important tool in function theory since the 1920s, first used by Grötzsch and Teichmüller, and then, extensively, by Beurling and Ahlfors. The concept itself is much older, having its roots in electromagnetism. For example, the discussion by Maxwell in his "Treatise on Electricity and Magnetism"[135, Art. 306, 307] over a hundred years ago comes very close to that in Section 7.3[1].

[1] I am thankful for Alex Eremenko for this reference.

7.22 Upper gradients. If u is a smooth function on $\mathbb{R}^n$, or on any Riemannian manifold, then

$$|u(x) - u(y)| \leq \int_\gamma |\nabla u|\, ds \tag{7.23}$$

for each rectifiable curve γ joining x and y. This property of gradients can be isolated: given a real-valued function u in a metric space X, a Borel function $\rho : X \to [0, \infty]$ is said to be an *upper gradient* of u if

$$|u(x) - u(y)| \leq \int_\gamma \rho\, ds \tag{7.24}$$

for each rectifiable curve γ joining x and y in X.

Every function has an upper gradient[2], namely $\rho \equiv \infty$, and upper gradients are never unique. The constant function $\rho \equiv L$ is an upper gradient of every L-Lipschitz function, but this is rarely the best choice.

Exercise 7.25. Suppose that u is a Lipschitz function in a metric space X. Show that

$$\rho(x) = \liminf_{r \to 0} r^{-1} \sup_{|x-y|<r} |u(x) - u(y)|$$

is an upper gradient of u.

On the other hand, if X contains no nontrivial rectifiable curves, then $\rho \equiv 0$ is an upper gradient of any function. It follows that upper gradients are potentially useful objects only if the underlying space has plenty of rectifiable curves; in particular, we are ruling out many fractal or disconnected spaces. This is but one difference from Hajłasz's approach, which is nonvacuous in (almost) every metric measure space. On the positive side, no measure is needed to define the upper gradient; it is a purely metric concept. Moreover, it is not difficult to see that ρ can be chosen to be zero in every closed set where u is constant.

Exercise 7.26. Suppose that ρ is an upper gradient of a function u in a metric space X. Show that the function

$$\rho'(x) = \rho(x), \quad x \in X \setminus E, \quad \rho'(x) = 0, \quad x \in E,$$

is also an upper gradient of u if u is constant on a closed set E.

7.27 Sobolev spaces based on upper gradients. It now seems natural to define a "Sobolev space" on a metric measure space as those L^p functions that possess an L^p upper gradient. This space can be equipped with a (semi)norm

$$\|u\|_{1,p} = \|u\|_p + \inf \|\rho\|_p\,, \tag{7.28}$$

[2]The term "upper gradient" was suggested by John Garnett; originally, the name "very weak gradient" was used in [83] and [84].

where the infimum is taken over all upper gradients ρ of u. Equivalently, Shanmugalingam [166], [167] defines the *Newtonian space* $N^{1,p}(X)$, $1 \leq p < \infty$, in a metric measure space X to be the space of those L^p functions u in X for which there exists an L^p function ρ such that inequality (7.24) holds except for a family of curves $\Gamma = \{\gamma\}$ of p-modulus zero in X.[3] Expression (7.28) only defines a seminorm, but Shanmugalingam shows that the associated normed space is always Banach. It is also true that $N^{1,p}(\Omega) = W^{1,p}(\Omega)$ for each open set Ω in $\mathbb{R}^n$, and for all $1 \leq p < \infty$.

Cheeger [31] took a somewhat different approach and defined a Sobolev space $H^{1,p}(X)$ in an arbitrary metric measure space X as follows. The space $H^{1,p}(X)$ consists of those L^p functions u for which there exists a sequence (u_i) of L^p functions converging to u in L^p and a corresponding sequence of upper gradients ρ_i of u_i having uniformly bounded L^p norm. The norm in $H^{1,p}(X)$ is defined as in Eq. (7.28), but the infimum is now taken over all possible upper gradients ρ_i associated with all possible sequences (u_i) that converge to u in L^p. Shanmugalingam [166] has shown that $N^{1,p}(X) = H^{1,p}(X)$ isometrically always when $p > 1$.

It is not true in general that the Hajłasz–Sobolev space $M^{1,p}(X)$ is the same space as $N^{1,p}(X)$. For instance, if X has no (nonconstant) rectifiable curves, $N^{1,p}(X) = L^p(X)$. Shanmugalingam [167] has shown that if the metric measure space is doubling and admits a Poincaré inequality (see Chapter 9), then the two spaces are isomorphic as Banach spaces. Under similar hypotheses, Cheeger [31] proved the important theorem that the spaces $H^{1,p}(X)$ are reflexive provided $p > 1$.

It is an open problem whether the spaces $M^{1,p}(X)$ or $N^{1,p}(X)$ are always reflexive when $p > 1$.

We will not go into the study of the spaces $N^{1,p}$ here but refer the reader to [31], [166], and [167]. (See also the notes to this chapter.)

7.29 Capacity. In certain situations, the modulus of a curve family can be defined somewhat differently. The concept is known as the capacity of a condenser and is defined in a general space as follows. Let E and F be subsets of a metric measure space X, and let $1 \leq p < \infty$; the *p-capacity* of the pair (E, F) is the number

$$\operatorname{cap}_p(E, F) = \inf \int_X \rho^p \, d\mu, \tag{7.30}$$

where the infimum is taken over all upper gradients of all real-valued functions u on X such that $u \mid E \leq 0$ and $u \mid F \geq 1$. Notice that no regularity assumption is made on u. In this generality, the capacity appears rather useless, but on the other hand, the following fundamental equality between modulus and capacity becomes almost a tautology.

[3] The letter N in $N^{1,p}$ stands for Newton because the definition of the space is essentially based on the fundamental theorem of calculus. A more continental notation would involve L for Leibniz, but the spaces $L^{1,p}$ are already known in the literature as "Dirichlet spaces" (figure this out!) of locally integrable functions with weak first derivatives in L^p.

Theorem 7.31. *We have that*

$$\operatorname{cap}_p(E, F) = \operatorname{mod}_p(E, F),$$

where the modulus on the right-hand side is the modulus of all curves joining the sets E and F in X.

PROOF. If u is a function on X with $u \mid E \leq 0$ and $u \mid F \geq 1$, and if ρ is any upper gradient of u, then

$$1 \leq |u(x) - u(y)| \leq \int_\gamma \rho\, ds$$

for any rectifiable curve γ joining a point $x \in E$ and a point $y \in F$. Therefore,

$$\operatorname{mod}_p(E, F) \leq \operatorname{cap}_p(E, F).$$

On the other hand, if ρ is an admissible function for the family (E, F), then define

$$u(x) = \inf \int_{\gamma_x} \rho\, ds, \tag{7.32}$$

where the infimum is taken over all curves γ_x joining E to the point x in X. Then, $u \mid E = 0$, $u \mid F \geq 1$, and ρ is easily seen to be an upper gradient of u. This implies that $\operatorname{cap}_p(E, F) \leq \operatorname{mod}_p(E, F)$, and the theorem follows. □

Notice that we do not assume in the preceding argument that there are rectifiable curves in our space; if rectifiable curves do not exist, the discussion simply becomes vacuous and one must observe the usual convention that the infimum over the empty set is ∞.

It is through capacity that most estimates for modulus can be made. Because we want lower estimates, the following result appears useful.

Theorem 7.33. *Suppose that E and F are two compact disjoint subsets of a compact and locally quasiconvex metric measure space (X, μ). Then, in the definition of* $\operatorname{cap}_p(E, F)$, *one can restrict to locally Lipschitz functions u.*

A metric space is called *quasiconvex* if there is a constant $C \geq 1$ so that every pair of points x and y in the space can be joined by a curve whose length does not exceed $C|x - y|$. A space is *locally quasiconvex* if every point in it has a quasiconvex neighborhood.

The assumptions that X be compact and locally quasiconvex in Theorem 7.33 are qualitative assumptions, and it is not clear to what extent they really are necessary. (See Remark 7.34.) Traditionally, one defines capacity by using smooth or Lipschitz functions at the start, and then the issue becomes how to prove Theorem 7.31. This result in increasing generality is due to Gehring, Ziemer, Hesse, and others. Theorem 7.33 follows from [84, Proposition 2.17] and will not be proved here.

Remark 7.34. It is an open problem whether Theorem 7.33 is true without the assumption that X is compact. Hesse [90] showed that it is true if X is an open connected set in $\mathbb{R}^n$ and μ is Lebesgue measure. Recently, Kallunki and Shanmugalingam [101] have generalized Hesse's theorem substantially by showing that Theorem 7.33 is true when $p > 1$ and X is an open connected set in a proper and locally quasiconvex doubling metric measure space that in addition admits a $(1, p)$-Poincaré inequality (as defined in Chapter 9). A metric space is said to be *proper* if its closed balls are compact.

It would be interesting to know the optimal conditions for Theorem 7.33. One can also ask when the capacity can be taken over continuous test functions. (Compare [85] and [101].)

7.35 Notes to Chapter 7. Väisälä's lecture notes [182] have stood the test of time remarkably and continue to be the best source for the general theory of quasiconformal maps in Euclidean n-space. In particular, several equivalent definitions for quasiconformal maps are given there; for a more recent addition, see [82]. An abstract and elegant treatment of modulus was given by Fuglede in [56]. A good source for quantitative estimates for modulus is Vuorinen's monograph [197]. For an analytic approach to quasiconformality, see [15]. More recent advances in the area can be read about in [44] and [95]. More general *quasiregular maps* are studied in the monographs [148] and [149]. A comprehensive treatment of quasiconformal and related mappings and their role in nonlinear analysis is the recent monograph [96].

Upper gradients were introduced in [83] and [84] as a tool to study quasiconformal maps. After these lectures were given, Cheeger [31] and Shanmugalingam [167] independently developed a theory of Sobolev spaces based on the idea of upper gradient. Earlier work on this Sobolev space concept can be found in [113]. Cheeger furthermore develops an interesting theory of minimal upper gradients and shows that under quite general circumstances one can define an exterior differential for functions in the Sobolev space $H^{1,p}(X)$. In this connection, see also [199]. Removable sets for the Sobolev spaces in metric measure spaces have been studied in [114]. The idea of an abstract Poincaré inequality defined in terms of upper gradients, introduced in [83] and [84], has an important role in these studies (see Chapter 9). The fact that $N^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$ is implicit in the works of Fuglede and Ohtsuka, starting from the 1950s, but the idea is old and can be traced back to the Italian school of analysis in the early part of the 1900s.

8

Loewner Spaces

Let (X, μ) be a metric measure space. For each real number $n > 1$, we define the *Loewner function* $\phi_n : (0, \infty) \to [0, \infty)$ of X by

$$\phi_{X,n}(t) = \phi_n(t) = \inf\{\operatorname{mod}_n(E, F; X) : \Delta(E, F) \le t\},$$

where E and F are disjoint nondegenerate continua in X with

$$\Delta(E, F) = \frac{\operatorname{dist}(E, F)}{\min\{\operatorname{diam} E, \operatorname{diam} F\}}$$

designating their relative position in X. Here $(E, F; X)$ denotes the family of all curves joining E and F in X. If X is understood from the context, we usually write $(E, F; X) = (E, F)$. If one cannot find two disjoint continua in X, it is understood that $\phi_{X,n}(t) \equiv 0$. Recall that a *continuum* is a compact connected set, and a continuum is *nondegenerate* if it is not a point; henceforth, we tacitly assume that all continua are nondegenerate.

We preferred to assume $n > 1$ above to avoid some uninteresting special cases. Also, one should remember that we do not require n to be an integer.

By definition, the function ϕ_n is decreasing.

Definition 8.1. A pathwise connected metric measure space (X, μ) is said to be a *Loewner space of exponent* n, or an n*-Loewner space*, if the Loewner function $\phi_{X,n}(t)$ is positive for all $t > 0$.

Note that the positivity of the Loewner function alone does not imply that the space X in question is pathwise connected; for instance, X can be the disjoint union of a Loewner space and a point. It is natural to stipulate that a Loewner space be pathwise connected.

In a Loewner space, one finds lots of rectifiable curves joining two disjoint continua, and the plentitude of curves is being quantified by the function ϕ_n. In particular, a space without rectifiable curves, such as $(\mathbb{R}^n, |x-y|^{1/2})$, cannot be a Loewner space. Also, notice the scale invariance that is built into the condition.

The following theorem was proved by Loewner [122] in 1959, whence the terminology.

Theorem 8.2. *Euclidean space $\mathbb{R}^n$ is an n-Loewner space.*

We will prove Theorem 8.2 in Chapter 9. In this chapter, we give a general discussion of Loewner spaces.

To digress a bit, we recall some basic facts about Hausdorff measures[1].

8.3 Hausdorff measures. Given a metric space X and a real number $\alpha > 0$, the *α-Hausdorff measure* of a subset A of X is defined as follows. First, for each $\delta > 0$, we define the premeasure

$$\mathcal{H}_\alpha^\delta(A) = \inf \sum_{B \in \mathcal{B}} (\operatorname{diam} B)^\alpha, \tag{8.4}$$

where the infimum is taken over all covers $\mathcal{B}$ of A by (closed) balls of diameter at most δ. When δ gets smaller, the number of possible covers decreases, so the limit

$$\lim_{\delta \to 0} \mathcal{H}_\alpha^\delta(A) = \mathcal{H}_\alpha(A) \in [0, \infty]$$

exists and is called the α-Hausdorff measure of A. The resulting set function $A \mapsto \mathcal{H}_\alpha(A)$ is a Borel regular measure on X.

Usually, the infimum in Eq. (8.4) is taken over all possible coverings of A by sets of diameter at most δ; however, that procedure leads to a measure that is comparable to the one defined through closed balls.

As defined here, $\mathcal{H}_n$ in $\mathbb{R}^n$ is a constant multiple of Lebesgue measure.

It is not difficult to see that if $\mathcal{H}_\alpha(X)$ is finite for some $\alpha > 0$, then $\mathcal{H}_{\alpha'}(X) = 0$ for each $\alpha' > \alpha$. Therefore, the number

$$\dim_H(X) = \inf\{\alpha > 0 : \mathcal{H}_\alpha(X) = 0\}$$

exists and is called the *Hausdorff dimension* of X. Note that

$$0 \leq \dim_H(X) \leq \infty.$$

For example, $\mathbb{R}^n$ has Hausdorff dimension n, the Cantor ternary set has Hausdorff dimension $\log 2/\log 3$, and the sequence $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\} \subset \mathbb{R}$ has Hausdorff dimension

[1] See [47, 2.10] or [134, Chapter 4] for a careful discussion of Hausdorff measures, as well as some of the unproven facts to follow.

0. The actual α-Hausdorff measure of a space at the critical exponent $\alpha = \dim_H(X)$ may be zero, finite but nonzero, or infinite. For instance, $\mathbb{R}^n$ has infinite n-measure, and the Cantor ternary set has finite and positive $\log 2/\log 3$-measure.

A somewhat different concept results when no restriction is put on the covers. The α-dimensional *Hausdorff content* of a set A in a metric space X is

$$\mathcal{H}_\alpha^\infty(A) = \inf \sum_{B \in \mathcal{B}} (\operatorname{diam} B)^\alpha, \tag{8.5}$$

where the infimum is taken over all covers $\mathcal{B}$ of A by closed balls. Then, clearly

$$\mathcal{H}_\alpha^\infty(A) \leq \mathcal{H}_\alpha(A),$$

and the α content is finite for any bounded set. However, these two set functions vanish simultaneously, a fact that often makes estimating the Hausdorff dimension easier.

Exercise 8.6. Show that $\mathcal{H}_\alpha(X) = 0$ if and only if $\mathcal{H}_\alpha^\infty(X) = 0$ for any space X.

8.7 Ways to estimate Hausdorff dimension. Suppose that μ is a Borel measure on a metric space X that is finite on bounded sets and satisfies

$$\mu(B_R) \geq C^{-1} R^n \tag{8.8}$$

for some constant $C \geq 1$, for some exponent $n > 0$, and for all closed balls B_R of radius $R < \operatorname{diam} X$. Then, the Hausdorff dimension of X is at most n.

To see this, assume first that X is bounded. Then, cover X by closed balls $\{B_i\}$ of radius R_i at most δ such that

$$\frac{1}{5}B_i \cap \frac{1}{5}B_j = \emptyset$$

whenever $i \neq j$ (see Chapter 1) to conclude that

$$\begin{aligned}
\mathcal{H}_n^\delta(X) &\leq \sum (\operatorname{diam} B_i)^n \leq \sum (2R_i)^n \\
&= 10^n \sum \left(\frac{1}{5}R_i\right)^n \\
&\leq C \sum \mu\left(\frac{1}{5}B_i\right) \leq C\mu(X) < \infty.
\end{aligned}$$

This means that $\dim_H X$ is at most n. If X is not bounded, one does the same calculation for each ball of X and arrives at the same conclusion.

Similarly, if μ is a (nontrivial) measure on X such that

$$\mu(B_R) \leq C R^n \tag{8.9}$$

for some constant $C \geq 1$, for some exponent $n > 0$, and for all closed balls B_R of radius $R > 0$, then the Hausdorff dimension of X is at least n. In fact, $\mathcal{H}_n(X) > 0$.

In particular, if X is a metric space admitting a Borel regular measure μ such that

$$C^{-1}R^n \leq \mu(B_R) \leq CR^n \tag{8.10}$$

for some constant $C \geq 1$, for some exponent $n > 0$, and for all closed balls B_R of radius $0 < R < \operatorname{diam} X$, then X has Hausdorff dimension precisely n.

Exercise 8.11. Prove that if μ is a Borel regular measure on a metric space X satisfying formula (8.10), then there is a constant $C' \geq 1$ such that

$$C'^{-1}\mathcal{H}_n(E) \leq \mu(E) \leq C'\mathcal{H}_n(E)$$

for all Borel sets E in X. (Hint: Use [47, 2.2.2].)

Metric measure spaces where formula (8.10) holds are called *(Ahlfors) n-regular*. By Exercise 8.11, we can replace μ by the Hausdorff n-measure in an n-regular space without essential loss of information.

8.12 Frostman's lemma. There is an interesting converse to one of the preceding statements. Namely, if X is a (Borel measurable) subset of Euclidean space and if $\mathcal{H}_n(X) > 0$ for some positive (not necessarily integral) number n, then there is a nontrivial Borel measure μ with support in X such that inequality (8.9) holds. See [134, Chapter 8] for a proof as well as for more general statements.

8.13 Topological dimension. Hausdorff dimension of a space is a metric concept; the underlying topological space can have metrics with different dimension. For instance, $\mathbb{R}$ with the standard metric has Hausdorff dimension one, but $(\mathbb{R}, |x - y|^{1/2})$ has Hausdorff dimension two.

On the other hand, every separable metric space has its associated *topological dimension*, which is an integer and a topological invariant[2]. The definition is inductive: the empty set has dimension -1, and the dimension of an arbitrary space is the least integer n with the property that every point in the space has arbitrarily small neighborhoods whose boundaries have dimension $n - 1$ or less.

For example, $\mathbb{R}^n$ has dimension n, and the Cantor ternary set (more generally, every totally disconnected space) has dimension zero.

It turns out that the topological dimension of a metric space bounds its Hausdorff dimension from below.

Theorem 8.14. *The Hausdorff dimension of a metric space is at least its topological dimension.*

Theorem 8.14 clearly is implied by the following.

---

[2]For topological spaces that are not separable metric, the story of dimension gets more complicated and will not be touched upon here.

Theorem 8.15. *If $\mathcal{H}_{n+1}(X) = 0$ for a metric space X and for some integer n, then the topological dimension of X is at most n.*

Proof. We claim that the assumption $\mathcal{H}_{n+1}(X) = 0$ implies that the Hausdorff n-measure of the "sphere" $\partial B(x_0, r)$ is zero for each $x_0 \in X$ and for almost every $r > 0$. In fact, the theorem easily follows from this claim by induction on the dimension.

To prove the claim, fix $x_0 \in X$. Then, observe that if B is a closed ball in X and if

$$r_2 = \sup_{x \in B} \operatorname{dist}(x_0, x), \qquad r_1 = \inf_{x \in B} \operatorname{dist}(x_0, x),$$

then

$$\begin{aligned} \int_0^\infty (\operatorname{diam}(\partial B(x_0, r) \cap B))^n \, dr &= \int_{r_1}^{r_2} (\operatorname{diam}(\partial B(x_0, r) \cap B))^n \, dr \\ &\leq (\operatorname{diam} B)^n (r_2 - r_1) \\ &\leq (\operatorname{diam} B)^{n+1}. \end{aligned} \tag{8.16}$$

Because, for each $\epsilon > 0$, we can cover X by closed balls such that

$$\sum (\operatorname{diam} B)^{n+1} < \epsilon,$$

we easily infer from Eq. (8.16) that

$$\int_0^\infty \mathcal{H}_n(\partial B(x_0, r)) \, dr = 0.$$

The proof of Theorem 8.14 is thereby complete. □

There are many equivalent ways to define the topological dimension of a separable metric space. The beautiful classic treatise on dimension is [94].

Exercise 8.17. Prove that if X is a connected metric space, then $\operatorname{diam} X \leq \mathcal{H}_1^\infty(X)$.

8.18 Properties of Loewner spaces. A given metric measure space can be a Loewner space for more than one exponent. For example, if X is an n-Loewner space of finite total measure, $\mu(X) < \infty$, then X is a p-Loewner space for each $p \geq n$ simply by Hölder's inequality. It turns out that the Hausdorff dimension bounds these exponents from below, as follows.

Proposition 8.19. *If (X, μ) is an n-Loewner space, then there is a constant $C = C(\phi, n) \geq 1$ so that*

$$\mu(B_R) \geq C^{-1} R^n \tag{8.20}$$

for all balls B_R of radius $R < \operatorname{diam} X$. In particular, the Hausdorff dimension of X is at most n.

Corollary 8.21. *No space can be a Loewner space for an exponent less than its topological dimension.*

Corollary 8.22. *An n-Loewner space of topological dimension n has Hausdorff dimension n.*

PROOF OF PROPOSITION 8.19. Let $B_R = B(x, R)$ be a ball in X of radius $R < \operatorname{diam} X/2$. Then there is a point y in $X \setminus B_R$. Let σ be a path joining x and y; then σ has subcurves σ_1 and σ_2 joining x to $\partial B(x, R/4)$ inside $B(x, R/4)$ and $\partial B(x, R/2)$ to $\partial B(x, R)$ inside $B(x, R) \setminus B(x, R/2)$, respectively. We have that

$$\Delta(\sigma_1, \sigma_2) \le \frac{R/2}{R/4} = 2,$$

so

$$\operatorname{mod}_n(\sigma_1, \sigma_2) \ge \phi_n(2).$$

On the other hand, the function $\rho(z) = 4/R$, if $z \in B(x, R)$, and $\rho(z) = 0$ elsewhere, is admissible for the curve family (σ_1, σ_2), which implies that

$$\phi_n(2) \le \int_{B(x,R)} 4^n R^{-n}\, d\mu \le 4^n R^{-n} \mu(B_R),$$

and inequality (8.8) follows with $C = \max\{1, 4^n/\phi_n(2)\}$. This proves the proposition. □

I do not know what more can be said about the geometry of Loewner spaces without further assumptions. If the volume-growth condition (8.9) is assumed, Loewner spaces can be shown to enjoy some strong geometric properties.

A metric space is said to be *linearly locally connected* if there is a constant $C \ge 1$ so that every pair of points in $B(x, R)$ can be joined in $B(x, CR)$ and every pair of points in $X \setminus B(x, R)$ can be joined in $X \setminus B(x, R/C)$, whenever $B(x, R)$ is a ball in X. By joining, here we mean joining by a continuum.

Theorem 8.23. *Suppose that (X, μ) is an n-Loewner space where the volume growth (8.9) holds. Then*

1. *X is linearly locally connected;*
2. *X is quasiconvex;*
3. *the Loewner function ϕ_n of X behaves asymptotically as follows:*

$$\phi_n(t) \approx (\log t)^{1-n}, \quad t \to \infty,$$

and

$$\phi_n(t) \approx \log(1/t), \quad t \to 0.$$

The statement is quantitative in that all the relevant constants depend only on the Loewner function, on n, and on the constant in inequality (8.9).

We will not prove Theorem 8.23 here, but refer to [84, Section 3]. Properties (1) and (2) are relatively easy to prove in view of Lemma 7.18, but the proof for property (3) is more involved. I urge the reader to prove properties (1) and (2) as an exercise before consulting [84].

Example 8.24. (a) The open unit ball $\mathbb{B}^n$ in $\mathbb{R}^n$ is an n-Loewner space. This follows from the fact that $\mathbb{R}^n$ is an n-Loewner space (Theorem 8.1) together with the fact that the n-modulus of any curve family in $\mathbb{B}^n$ is at least half of the modulus of the same family when measured in all of $\mathbb{R}^n$. (The second fact can be proved easily by using a conformal reflection (inversion) across $\partial\mathbb{B}^n$, except for one technical point. Compare [197, p. 55].)

(b) The open unit ball $\mathbb{B}^n$ in $\mathbb{R}^n$ is in fact a p-Loewner space for all $p \geq n$. This follows from the remark made before Proposition 8.19 because $\mathbb{B}^n$ has finite Lebesgue measure. Perhaps more to the point is a proof that uses the Sobolev embedding (3.12), together with the fact that one can restrict to Sobolev functions in defining capacity (see Remark 7.34).

On the other hand, $\mathbb{R}^n$ is a Loewner space only with exponent n. In fact, Corollary 8.21 gives only large p's a chance, while Eq. (7.17) implies that the p-modulus of the curves joining the sets $\partial B(x_0, 2r)$ and $\partial B(x_0, r)$ is $c(n, p)r^{n-p}$, which tends to 0 as $r \to \infty$ if $p > n$.

(c) The real hyperbolic space $\mathbb{H}^n$ is an n-Loewner space. This follows fairly easily from Theorem 7.10. In fact, by using the ball model for $\mathbb{H}^n$, the identity map $i : \mathbb{B}^n \to \mathbb{H}^n$ is conformal. (That is, we identify $\mathbb{H}^n$ with the open unit ball in $\mathbb{R}^n$ equipped with the metric $(1 - |x|^2)^{-2}ds^2$.) Because the n-modulus remains invariant, one only needs to check how the relative distances $\Delta(E, F)$ are affected under this conformal change of metric. Notice that the Loewner function of the hyperbolic space satisfies $\lim_{t\to 0+} \phi_n(t) < \infty$, so that something like assumption (8.9) is necessary for Theorem 8.23 (3). The volume growth of the hyperbolic space is of course exponential.

(d) The volume-growth restriction (8.9) is necessary also for (1) and (2) in Theorem 8.23. To see this, consider the region $X = \{z = (x, y) \in \mathbb{R}^2 : x^2 \leq y^2 + 1\}$. This domain X with Euclidean metric and Lebesgue measure cannot be 2-Loewner because it is not linearly locally connected. However, if we take as our measure $d\mu(z) = P(|z|)\,dz$, where P is a sufficiently fast-growing polynomial, we easily verify that the space (X, μ) is 2-Loewner. What happens here is that the modulus between two continua that are far away from the x axis, but at different sides of it can be increased appropriately by increasing the volume growth.

It is left to the reader to construct similar examples where (2) in Theorem 8.23 fails.

(e) Generalizing (a) above, every Euclidean domain Ω that has the $L^{1,n}$-extension property is an n-Loewner space (with Euclidean distance and Lebesgue measure). A domain Ω is said to have the *$L^{1,n}$-extension property* if every continuous function u in Ω, with distributional gradient $|\nabla u|$ in $L^n(\Omega)$, has an extension $\tilde{u}$ to a locally integrable function in $\mathbb{R}^n$ whose weak gradient is in $L^n(\mathbb{R}^n)$ such that

$$\|\nabla\tilde{u}\|_{n,\mathbb{R}^n} \leq C\,\|\nabla u\|_{n,\Omega}$$

for some constant $C \geq 1$ depending only on Ω. Indeed, it follows then that

$$\operatorname{mod}_n(E, F; \mathbb{R}^n) = \operatorname{cap}_n(E, F; \mathbb{R}^n) \leq C^n \operatorname{cap}_n(E, F; \Omega) = C^n \operatorname{mod}_n(E, F; \Omega),$$

which gives the Loewner property of Ω. (Compare Remark 7.34.)

(f) A domain in $\mathbb{R}^n$ that is quasiconformally homeomorphic to a round ball is an n-Loewner space if and only if it is a uniform domain. This follows by combining Theorem 8.23 (1) with works of Gehring, Martio, Jones, and Väisälä, as explained later.

A domain Ω in $\mathbb{R}^n$ is *uniform* if there is a constant $C \geq 1$ such that every pair of points in Ω can be joined by a curve γ in Ω whose length does not exceed $C|x-y|$ (that is, Ω is quasiconvex), and moreover, for each point $z \in \gamma$ the distance $\operatorname{dist}(z, \partial\Omega)$ is at least $C^{-1}\min\{|x-z|, |y-z|\}$.

A uniform domain has the $L^{1,n}$-extension property by a theorem of Jones [100] and hence is a Loewner space by the discussion in (e). On the other hand, if a domain is a Loewner space, then it is linearly locally connected by Theorem 8.23, and by the works [64] and [184], every quasiconformal image of a uniform domain is itself uniform if it is linearly locally connected. (Compare Theorem 11.19.)

More generally, by the same argument, a domain in $\mathbb{R}^n$ that is quasiconformally equivalent to a uniform domain is a Loewner space if and only if it is itself a uniform domain.

(g) It is often natural to equip a Euclidean domain Ω with the *internal distance*, in which $|x-y|$ is defined to be the infimum of the diameters of paths joining x and y in Ω. Then, a domain in $\mathbb{R}^n$ that is quasiconformally equivalent to a ball is a Loewner space with the internal metric (and Lebesgue measure) if and only if it is a John domain. This follows by combining Theorem 8.23 (1) with works of Näkki, Väisälä, and this author [145], [186], [79]. The reasoning is similar to that in (f).

Recall that a domain is a *John domain* if it satisfies the second requirement of uniformity; that is, we give away the quasiconvexity.

The general statement in the end of (f) has no direct counterpart here, for it is not true that all John domains are Loewner (in the internal metric). The subclass of John domains that are Loewner has been identified in [16]. See also [192].

(h) Väisälä in [186] introduced the concept of a *broad domain* in $\mathbb{R}^n$ with a definition similar to Definition 8.1; he used the internal metric of the domain and required that the function $\phi_n(t)$ tend to ∞ as $t \to 0$. By Theorem 8.23, a domain is broad if and only if it is Loewner in the internal metric quantitatively.

On the other hand, it is not clear whether the *quasiextremal distance*, or *QED, domains* of Gehring and Martio [64] coincide with domains that are Loewner in the Euclidean metric. The difference here is that QED domains are defined by a linear relation: a domain Ω is a QED domain if there is a constant $K \geq 1$ so that $\operatorname{mod}_n(E, F; \mathbb{R}^n) \leq K \operatorname{mod}_n(E, F; \Omega)$ for all pairs of disjoint continua E and F in Ω.

Exercise 8.25. Prove all the assertions made in Example 8.24.

A characterization of Loewner spaces in terms of Poincaré type inequalities will be given in Chapter 9. We close this chapter by mentioning that there are

n-regular n-Loewner spaces for each real number $n > 1$. Bourdon and Pajot in [21] showed that the classical Menger curve admits n-regular metrics for infinitely many nonintegral n such that the resulting metric space is geodesic and n-Loewner. Soon afterwards, Laakso [116] exhibited geodesic and n-regular n-Loewner spaces for each $n > 1$. Laakso's spaces are topologically quotients of spaces of the form $K \times [0, 1]$, where K is an appropriate self-similar Cantor set in Euclidean space.

Other curious examples of n-Loewner spaces satisfying expression (8.10) are certain nilpotent Lie groups equipped with their Carnot metrics. For these groups, the number n, although always an integer, is larger than the topological dimension of the underlying space. (Compare the example in Section 9.25.) Using Carnot groups, one can construct n-regular Loewner spaces with different topological dimensions at different parts of the space. See [84, Section 6] for more information.

8.26 Notes to Chapter 8. Abstract Loewner spaces were introduced in [83] and [84] in connection with quasiconformal mappings, and the observations about Loewner spaces given in this chapter are at least implicitly in those papers. The idea of axiomatizing a Loewner-type condition had appeared earlier in the quasiconformal literature, but always in special situations. See, for example, [144], [64], and [186]. Subsequent works on Loewner spaces and quasiconformal mappings include [7] and [177].

The material on Hausdorff measures and dimension is classical. The book by Hurewicz and Wallman [94] reads beautifully after 50 years.

The concept of linear local connectivity was introduced by Gehring in the 1960s, also in connection with quasiconformal mapping problems. Far-reaching extensions of the idea can be found in the recent work of Väisälä [190].

An elementary introduction to the Carnot geometry is the lecture notes [80] by the author; for more references, see the notes to Chapter 9. An interesting account on spaces with "good calculus" is [165].

9

Loewner Spaces and Poincaré Inequalities

In Chapter 4, we proved the Poincaré inequality in $\mathbb{R}^n$ by using the fact that points in $\mathbb{R}^n$ can be joined by a thick "pencil" of curves. In the previous chapter, we defined the Loewner function of a metric space that detects quantitatively whether or not the space contains rectifiable curves. In this chapter, we will see that these two concepts, the Poincaré inequality and the Loewner condition, are related.

We begin by defining what is meant by a Poincaré inequality in a metric measure space, and then show that the validity of such an inequality can be characterized in many different ways. The concept of an upper gradient as defined in Chapter 7 is crucial here.

9.1 Poincaré inequalities. Let (X, μ) be a metric measure space and let $1 \le p < \infty$. We say that X admits a *weak* $(1, p)$*-Poincaré inequality* if there are constants $0 < \lambda \le 1$ and $C \ge 1$ so that

$$\fint_{\lambda B} |u - u_{\lambda B}|\, d\mu \le C(\operatorname{diam} B)\left(\fint_B \rho^p \, d\mu\right)^{1/p} \tag{9.2}$$

for all balls B in X, for all bounded continuous functions u on B, and for all upper gradients ρ of u. To avoid some unpleasant pathologies, we make the further assumption that each ball in X has finite and positive measure.

The terminology is explained by the fact that the ball on the left-hand side in inequality (9.2) is allowed to be slightly smaller than the ball on the right-hand side, and by the fact that the L^1 norm is used on the left-hand side in inequality (9.2); we could speak about a (q, p)-Poincaré inequality with a self-explanatory meaning.[1] One should remember the Hajłasz–Koskela theorem (Theorem 4.18), which asserts that under certain conditions the weak inequality (9.2) implies an inequality that is a priori stronger. We say that X admits a $(1, p)$-*Poincaré inequality* if inequality (9.2) holds with $\lambda = 1$.

It follows from Hölder's inequality that condition (9.2) becomes weaker when p becomes larger, and one can show in fact that the family of spaces admitting a $(1, p)$-Poincaré inequality grows strictly with p. See [84, Section 7].

Next we will describe how the validity of a Poincaré inequality relates to the validity of various pointwise estimates.

Recall that the *Riesz potential* of a nonnegative function f on a metric measure space (X, μ) was defined in Eq. (3.23) as follows:

$$I_1(f)(x) = \int_X \frac{f(y)|x-y|}{\mu(B(x, |x-y|))}\, d\mu(y). \tag{9.3}$$

We will also use the notation

$$I_{1,A}(f)(x) = \int_A \frac{f(y)|x-y|}{\mu(B(x, |x-y|))}\, d\mu(y) \tag{9.4}$$

for $A \subset X$. The *restricted maximal function* (of a nonnegative f) is defined by

$$M_R f(x) = \sup_{r<R} \fint_{B(x,r)} f(y)\, d\mu.$$

Now, consider the following five conditions that may or may not be satisfied by a metric measure space (X, μ) with a given $p \geq 1$; the conditions are assumed to hold for all balls B in X, for all bounded continuous functions u on B, and for all upper gradients ρ of u. The constants $C_1, \ldots, C_6$ are assumed to be independent of B, u, ρ, and the points x and y.

(1) $$|u(x) - u_B|^p \leq C_1 (\operatorname{diam} B)^{p-1} I_{1,B}\rho^p(x)$$
for all $x \in \frac{1}{2}B$;

(2) $$|u(x) - u(y)|^p \leq C_2 |x-y|^{p-1} (I_{1,B}\rho^p(x) + I_{1,B}\rho^p(y))$$
for all $x, y \in (4C_3)^{-1}B$;

(3) $$|u(x) - u(y)|^p \leq C_4 |x-y|^p (M_{C_5|x-y|}\rho^p(x) + M_{C_5|x-y|}\rho^p(y))$$
for all $x, y \in (4C_6)^{-1}B$;

(4) X admits a weak $(1, p)$-Poincaré inequality;

(5) X admits a $(1, p)$-Poincaré inequality.

Theorem 9.5. *If (X, μ) is doubling, then we have*

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4).$$

[1] A shorter term for inequality (9.2), the *p-Poincaré inequality*, is commonly used in the recent literature.

If, in addition, X is geodesic, then also

$$(4) \Rightarrow (5) \Rightarrow (1).$$

All implications are quantitative in the usual sense.

Recall that a metric space is called geodesic if every pair of points in it can be joined by a curve whose length is the distance between the points; also recall that if the length of the curve can only be chosen to be at most a fixed constant multiple of the distance, then the space is called quasiconvex.

To tie this all up with the Loewner condition, we have the following two theorems.

Theorem 9.6. *Let (X, μ) be a quasiconvex, proper, and doubling metric measure space where the volume growth*

$$\mu(B_R) \geq C^{-1} R^n \tag{9.7}$$

is satisfied for some $n > 1$ and for all $R < \operatorname{diam} X$. If X admits a weak $(1, n)$-Poincaré inequality, then X is an n-Loewner space.

Recall that a metric space is proper if its closed balls are compact.

Theorem 9.8. *If (X, μ) is an n-Loewner space with the volume-growth condition*

$$C^{-1} R^n \leq \mu(B_R) \leq C R^n \tag{9.9}$$

satisfied for some $n > 1$ and for all $R < \operatorname{diam} X$, then (X, μ) admits a weak $(1, n)$-Poincaré inequality.

In conclusion, we have the following characterization of the Loewner condition in terms of Poincaré inequalities.

Theorem 9.10. *An n-regular, geodesic, and proper metric measure space (X, μ) is an n-Loewner space if and only if one (hence all) of the conditions (1) – (5) holds with $p = n$.*

The statements in Theorems 9.6, 9.8, and 9.10 are quantitative in that all the relevant constants and functions depend only on each other and on n.

We will not prove Theorem 9.8 here, but refer to [84, Lemma 5.16], where it is shown that property (3) holds in an n-regular Loewner space. Theorem 9.6 will be proved later under the stronger assumption that X is geodesic and admits a $(1, 1)$-Poincaré inequality. For the more general statement of Theorem 9.6, see [84, Theorem 5.7].

Remark 9.11. Many spaces of interest automatically satisfy the assumptions of Theorem 9.10, with the possible exception of geodesicity. On the other hand, we have seen in Theorem 8.23 that every Loewner space is quasiconvex if it is

n-regular; that is, if condition (9.9) holds. Then, if X is in addition proper, we can turn it into a geodesic space by defining a new metric,

$$|x-y|_{\text{new}} = \inf \operatorname{length} \gamma, \tag{9.12}$$

where the infimum is taken over all rectifiable curves γ joining x and y. Because X is proper, it is easily seen that there is a minimizing curve, which is necessarily a geodesic. On the other hand, by quasiconvexity, we have that

$$|x-y|_{\text{old}} \leq |x-y|_{\text{new}} \leq C|x-y|_{\text{old}}$$

so that the identity map

$$(X, |x-y|_{\text{old}}) \to (X, |x-y|_{\text{new}})$$

is bi-Lipschitz. Most of the properties of metric spaces we care about in this book are bi-Lipschitz invariant, so no essential harm is done by considering on X the geodesic metric defined in Eq. (9.12).

Exercise 9.13. Prove that geodesics always exist in the metric $|x-y|_{\text{new}}$ as defined in Eq. (9.12) if $(X, |x-y|_{\text{old}})$ is a proper metric space where every pair of points can be joined by at least one rectifiable curve.

A metric space is called a *length space* if the distance between any two points is the infimum of the lengths of all curves joining the two points. Clearly, a geodesic space is a length space, but a length space need not be geodesic.

Exercise 9.14. Show that a locally compact and complete length space is proper, and hence geodesic. Observe that a locally compact length space need not be geodesic.

Proof of Theorem 9.5. (1) $\Rightarrow$ (2): This follows from the triangle inequality with $C_2 = 8^{p-1}C_1$ and $C_3 = 3$. In fact, if $x, y \in \frac{1}{12}B$, then the ball $B_x = B(x, 2|x-y|)$ is contained in $B_y = B(y, 3|x-y|)$, and both of these balls in turn are contained in B, where u is defined. Because $x, y \in \frac{1}{2}B_x$, we thus have

$$\begin{aligned}|u(x)-u(y)|^p &\leq 2^{p-1}(|u(x)-u_{B_x}|^p + |u(y)-u_{B_x}|^p) \\ &\leq 2^{p-1}C_1(4|x-y|)^{p-1}\left(\int_B \frac{\rho^p(z)|x-z|}{\mu(B(x,|x-z|))}\,d\mu(z)\right. \\ &\quad \left.+ \int_B \frac{\rho^p(z)|y-z|}{\mu(B(y,|y-z|))}\,d\mu(z)\right),\end{aligned}$$

as desired.

(2) $\Rightarrow$ (3): We perform the standard trick:

$$\begin{aligned}I_{1,B}\rho^p(x) &= \int_B \frac{\rho^p(z)|x-z|}{\mu(B(x,|x-z|))}\,d\mu(z) \\ &= \sum_{j\geq -1}\int_{A_j} \frac{\rho^p(z)|x-z|}{\mu(B(x,|x-z|))}\,d\mu(z),\end{aligned} \tag{9.15}$$

where

$$A_j = \{z : 2^{-j-1}R \leq |z - x| < 2^{-j}R\}.$$

(The reason why we start the summing from -1 is that $B(x, R)$ may not cover the whole ball $B = B(x_0, R)$, but $B(x, 2R)$ does.) We may clearly assume that $|x - y| \geq (100C_3)^{-1}R$. Thus, we obtain from Eq. (9.15), by using the doubling property of μ, that

$$\begin{aligned} I_{1,B}\rho^p(x) &\leq C \sum_{j\geq -1} 2^{-j}R \fint_{\{z:|z-x|<2^{-j}R\}} \rho^p(z)\,d\mu(z) \\ &\leq CRM_{2R}\rho^p(x), \end{aligned}$$

from which (3) follows with appropriate choices of C_5 and C_6.

(3) $\Rightarrow$ (4): The proof of this is very similar to the argument after Theorem 3.22, where a weak-type estimate is shown to imply a strong estimate, as soon as the weak-type estimate concerns a function/gradient pair; the fact that we have an upper gradient here serves equally well. We omit the details here; they can be found in [84, Section 5] or in [160, Appendix C]. (The assumption that X is locally compact is made both in [84] and in [160]; this is, however, redundant in light of Exercise 7.26.) I urge the reader to try to produce the proof of this implication, given the argument in Chapter 3, before consulting the literature.

(4) $\Rightarrow$ (5): Now, we assume X is geodesic. Under this assumption, balls in X can be shown to satisfy a (λ, M)-chain condition as in Definition 4.15. In particular, it follows from Theorem 4.18 that the validity of a weak Poincaré inequality at all scales implies that the strong inequality, where $\lambda = 1$ in inequality (9.2), holds in X. □

Exercise 9.16. Show that in a geodesic metric space balls satisfy a (λ, M, a)-chain condition of Definition 4.15 for some universal λ, M, and a.

(5) $\Rightarrow$ (1): Pick $x \in \frac{1}{2}B$. We may clearly assume that $X \setminus B$ is nonempty. Then, choose $y \in B$ such that

$$r_1 := \frac{|x - y|}{2} = \frac{1}{16}\operatorname{diam} B.$$

Let γ be a geodesic from y to x, and let $\{B_1, B_2, \ldots\}$ be a sequence of balls centered on γ as follows: x_1 is the midpoint of γ,

$$B_1 = B\left(x_1, \frac{r_1}{2}\right),$$

and, in general, if x_i, for $i \geq 2$, is the point on γ with $|x_i - x| = 2^{-i+1}r_1 = r_i$, then

$$B_i = B\left(x_i, \frac{r_i}{2}\right).$$

Define, moreover, $B_0 = B$. We find

$$\begin{aligned}
|u(x) - u_B| &\leq \sum_{i \geq 0} \left| u_{B_i} - u_{B_{i+1}} \right| \\
&\leq \sum_{i \geq 0} \left| u_{B_i} - u_{\frac{3}{2}B_i} \right| + \left| u_{\frac{3}{2}B_i} - u_{B_{i+1}} \right| \\
&\leq C \sum_{i \geq 0} r_i \left(\fint_{\frac{3}{2}B_i} \rho^p \, d\mu \right)^{1/p} \\
&\leq C \left(\sum_{i \geq 0} r_i \right)^{(p-1)/p} \left(\sum_{i \geq 0} r_i \fint_{\frac{3}{2}B_i} \rho^p \, d\mu \right)^{1/p} \\
&\leq C (\operatorname{diam} B)^{(p-1)/p} \left(\sum_{i \geq 0} r_i \fint_{\frac{3}{2}B_i} \rho^p \, d\mu \right)^{1/p}.
\end{aligned} \tag{9.17}$$

Notice here the careful selection of the balls B_i and their enlargements $\frac{3}{2}B_i$; there is only a fixed amount of overlap among the balls $\frac{3}{2}B_i$, and

$$\operatorname{dist}\left(x, \frac{3}{2}B_i\right) = \frac{1}{4} r_i = \frac{1}{6} \operatorname{diam} \frac{3}{2} B_i. \tag{9.18}$$

Therefore, Eqs. (9.17) and (9.18) imply

$$\begin{aligned}
|u(x) - u_B|^p &\leq C (\operatorname{diam} B)^{p-1} \sum_{i \geq 0} \int_{\frac{3}{2}B_i} \frac{\rho(z)^p |x - z|}{\mu(B(x, |x - z|))} \, d\mu(z) \\
&\leq C (\operatorname{diam} B)^{p-1} \int_B \frac{\rho(z)^p |x - z|}{\mu(B(x, |x - z|))} \, d\mu(z),
\end{aligned}$$

where in the last inequality the finite overlap property was used.

This completes the proof of Theorem 9.5.

As mentioned earlier, for the full proof of Theorem 9.6 we refer to [84]. But in order to illustrate the principle of how various Poincaré inequalities and the Loewner condition are related, we give a simple proof of the following weaker result (which nevertheless contains Loewner's Theorem 8.2).

Theorem 9.19. *Let (X, μ) be a geodesic, proper, and doubling metric measure space that satisfies the lower mass bound (9.7) for some $n > 1$, and where the conclusion of Theorem 7.33 holds. If X furthermore admits a weak $(1, 1)$-Poincaré inequality, then X is an n-Loewner space, quantitatively.*

The proof shows that it is in fact enough to assume that the conclusion of Theorem 7.33 holds, with the word "Lipschitz" replaced by "continuous." See Remark 7.34 for examples of spaces where Theorem 9.19 is applicable.

PROOF. We begin by observing that property (2) holds in X, for $p = 1$, by Theorem 9.5. Consider a maximal function

$$M_{1,n;R}f(x) = \sup_{0<r<R}\left(r^{-1}\int_{B(x,r)} f(z)^n \, d\mu(z)\right)^{1/n}$$

of a locally integrable nonnegative function f in X. The standard argument by splitting a ball $B(x,r)$ into annuli (as in the proof of Proposition 3.19, for example) gives that

$$I_1 f(x) \leq C r^{1/n} M_{1,n;R} f(x) \tag{9.20}$$

whenever $0 < r < R$ and f is supported in the ball $B(x,r)$. The mass bound (9.7) was used here, too.

Now, fix $t > 0$ and fix two continua E and F in X so that $\Delta(E, F) \leq t$; we may assume both that t is large and that

$$e = \operatorname{diam} E \leq \operatorname{diam} F \leq 2e < \frac{1}{10}d,$$

where $d = \operatorname{dist}(E, F)$. (That we may so assume involves cutting the continua into smaller connected pieces, and a theorem [93, 2-16] from point-set topology is required.) Fix a point $x_0 \in E$ and a point $y_0 \in F$ such that $d = |x_0 - y_0|$, and note that E and F are both contained in $B = B(x_0, 2d)$. Let u be a continuous function in X that satisfies $u \mid E \leq 0$ and $u \mid F \geq 1$, and let ρ be an upper gradient of u in X. By property (2) and by inequality (9.20), either E or F has to belong to the set

$$\{z \in B : M_{1,n;2d}\rho(z) > \epsilon_0 d^{-1/n}\},$$

where $\epsilon_0 > 0$ depends only on the data. Say E does. By the standard covering argument (see Chapter 1), we can cover E by balls B_i of radius r_i such that the family $(\frac{1}{5}B_i)$ is disjointed and such that

$$r_i < Cd\int_{\frac{1}{5}B_i} \rho^n \, d\mu.$$

By summing over all such radii, we obtain the following bound for the Hausdorff 1-content of E:

$$\mathcal{H}_1^\infty(E) \leq Cd\int_B \rho^n \, d\mu. \tag{9.21}$$

But the Hausdorff 1-content of every connected set in a metric space is at least its diameter [47, 2.10.25] (Exercise 8.17), so that inequality (9.21) implies, upon taking the infimum over all u and ρ, that

$$C^{-1}\frac{1}{t} \leq C^{-1}\frac{e}{d} \leq \operatorname{cap}_n(E, F) = \operatorname{mod}_n(E, F),$$

as desired. The proof of Theorem 9.19 is complete. □

In the beginning of this chapter, we postulated that a space admits a (weak) $(1, p)$-Poincaré inequality if inequality (9.2) holds for all continuous functions u and for all their upper gradients ρ. One could equally well demand that inequality (9.2) holds for all (measurable) functions u or that it holds for Lipschitz functions only. Remember that every function has an upper gradient. Thus, we have three a priori different classes of metric spaces, depending on which type of Poincaré equality is required to hold. To make the distinction clear, if needed, let us say that X *admits a Poincaré inequality for Lipschitz functions*, or *for measurable functions*, the meaning being self-explanatory. An example in [112] shows that a space can admit a Poincaré inequality for Lipschitz functions without admitting one for continuous functions. On the other hand, the following theorem was proved in [85].

Theorem 9.22. *Let (X, μ) be a quasiconvex, proper and doubling metric measure space, and let $1 \leq p < \infty$. Then, X admits a (weak) $(1, p)$-Poincaré inequality for measurable functions if and only if it admits a (weak) $(1, p)$-Poincaré inequality for Lipschitz functions, quantitatively.*

I do not know of an example of a space that admits a Poincaré inequality for continuous functions but does not admit one for measurable functions.

Although Theorem 9.6 can be used to give many examples of Loewner spaces, much work still remains, for to establish the validity of a Poincaré inequality in a space is hardly an easy task. Fortunately, many interesting examples can be found in the recent literature. To finish this chapter, let us briefly discuss two prominent examples. (Recall also the examples of Bourdon-Pajot and Laakso mentioned in Chapter 8.)

9.23 Spaces with nonnegative Ricci curvature. Buser [26] proved that a $(1, 2)$-Poincaré inequality is valid on each complete Riemannian manifold whose Ricci curvature is everywhere nonnegative. On each such manifold, the Riemannian measure μ is doubling and satisfies the upper mass bound (8.9) on metric balls, where $n \geq 2$ is the dimension of the manifold. We therefore obtain from Proposition 8.19 and Theorem 9.10 the following result:

Theorem 9.24. *A complete Riemannian n-manifold X^n, $n \geq 2$, with everywhere nonnegative Ricci curvature, is a Loewner space if and only if it has maximal volume growth; that is, the Riemannian measure μ on X satisfies*

$$\mu(B(x_0, R)) \geq C^{-1} R^n$$

for some (hence all) $x_0 \in X$ and for all $0 < R < \operatorname{diam} X$, and for some constant $C \geq 1$ independent of the point x_0. The statement is quantitative in the usual sense.

Poincaré inequalities on spaces that are limits of Riemannian manifolds with (Ricci) curvature bounds have recently been studied by Cheeger and Colding [32].

9.25 The Heisenberg group. Euclidean 3-space $\mathbb{R}^3$ can be equipped with the noncommutative group law

$$(x, y, z) \cdot (x', y', z') = \left(x + x', y + y', z + z' + \frac{1}{2}(xy' - x'y)\right),$$

and with the *homogeneous norm*

$$\|(x, y, z)\| = ((x^2 + y^2)^2 + z^2)^{1/4},$$

where by *norm* it is meant simply that $\|\cdot\|$ takes on nonnegative values on $\mathbb{R}^3$, vanishing only at 0, and the following relations are satisfied:

$$\|a\| = \|a^{-1}\|, \quad \|\delta_t(a)\| = t\,\|a\|, \quad a \in \mathbb{R}^3,$$

where a^{-1} denotes the inverse element of $a \in \mathbb{R}^3$, and where $\delta_t : \mathbb{R}^3 \to \mathbb{R}^3$ for $t > 0$ is a *dilation* defined by $\delta_t((x, y, z)) = (tz, ty, t^2 z)$.

One can show that the symmetric distance function

$$d(a, b) = \|a^{-1} \cdot b\| \tag{9.26}$$

defines a metric in $\mathbb{R}^3$, and the space $(\mathbb{R}^3, d)$ is called the *first Heisenberg group* equipped with a *Carnot metric*. We denote this space by $\mathbf{H}_1$.

Left multiplication, $a \mapsto b \cdot a$, for $b \in \mathbf{H}_1$, is a distance-preserving map, and it preserves the Lebesgue 3-measure of each (measurable) set in $\mathbb{R}^3$.

The important facts about $\mathbf{H}_1$ are listed in the following theorem.

Theorem 9.27. *The metric measure space* $(\mathbf{H}_1, \mathcal{L}_3)$ *is a proper, geodesic, and* 4*-regular space admitting a* $(1, 1)$*-Poincaré inequality.*

In particular, we see that $\mathbf{H}_1$ is a Loewner space of exponent 4, which is strictly larger than the topological dimension of $\mathbf{H}_1$. In the preceding case, $\mathcal{L}_3$ denotes Lebesgue 3-measure in $\mathbb{R}^3$.

The first Heisenberg group $\mathbf{H}_1$ is an example from the large family of spaces commonly known as *Carnot groups*. Even a superficial digression to the theory of Carnot groups would take us too far a field, but I could not resist giving a quick definition of the first Heisenberg group just to show how easily one can expose an extremely intriguing geometric structure.

The first three claims in Theorem 9.27 are easy to establish starting from the definitions. The Poincaré inequality was proved by Jerison [99]. (In fact, Jerison proves more than is needed for Theorem 9.27; a simple proof in this case has been given by Varopoulos [194].)

9.28 Notes to Chapter 9. A thorough study of the Loewner condition and related Poincaré inequalities appears in [84]. There is a large body of literature on implications as in Theorem 9.5, in varying generality. The reader can start with [29], [55], [60], [75], and the references in these papers.

There is also extensive literature on Poincaré-type inequalities on Riemannian manifolds. See [30], [32], [34], [35], [158], and references therein.

Carnot groups are also known as *homogeneous groups* or *stratified groups*. The literature here is plentiful, and I list [52], [70], [109], [110], [111], [147], [172], and [195] among the more authoritative references. An elementary introduction can be found in [80]. Poincaré inequalities have been studied under yet more general sub-Riemannian conditions, where no group structure is present, and interesting metric space structures emerge therein. See [143] and [141] for the geometry of these spaces, and [53], [99], [29], and [123] for Poincaré-type inequalities.

Exotic examples of spaces admitting a Poincaré inequality can be found in [160], [21], [116], and [77].

10

Quasisymmetric Maps: Basic Theory I

In this chapter, we develop a basic theory of quasisymmetric embeddings in metric spaces, following for the most part the paper by Tukia and Väisälä [176]. We use the notation $f : X \to Y$ for an embedding f of a metric space X in a metric space Y. Thus, note in particular that in this notation f is not supposed to be onto. Recall that an *embedding* is a map that is a homeomorphism onto its image.

An embedding $f : X \to Y$ is called *quasisymmetric* if there is a homeomorphism $\eta : [0, \infty) \to [0, \infty)$ so that

$$|x - a| \leq t|x - b| \qquad \text{implies } |f(x) - f(a)| \leq \eta(t)|f(x) - f(b)| \tag{10.1}$$

for all triples a, b, x of points in X, and for all $t > 0$. Thus, f is quasisymmetric if it distorts relative distances by a bounded amount. We also say that f is η-*quasisymmetric* if the function η needs to be mentioned.

An embedding $f : X \to Y$ is said to be *bi-Lipschitz* if both f and f^{-1} are Lipschitz. The term L-*bi-Lipschitz* is also used, if necessary (see Chapter 6). Notice the difference between quasisymmetric maps and bi-Lipschitz maps: the latter distort absolute distances by a bounded amount, which is a much stronger condition. It is easy to see that an L-bi-Lipschitz embedding is L^2t-quasisymmetric. On the other hand, for any metric space $(X, |x - y|)$ and for any $\epsilon \in (0, 1)$, the identity map

$$(X, |x - y|) \to (X, |x - y|^{\epsilon}) \tag{10.2}$$

is always quasisymmetric but rarely Lipschitz. Also, the map $x \mapsto \lambda x$ in $\mathbb{R}^n$ is quasisymmetric with $\eta(t) = t$ for each $\lambda > 0$, but its bi-Lipschitz constant is $\max\{\lambda, 1/\lambda\}$.

Informally, one can define an embedding to be bi-Lipschitz if it distorts both the shape and size of an object by a bounded amount, while quasisymmetry only preserves the approximate shape. The quasisymmetry condition (10.1) is very subtle; it is a flexible condition and also guarantees strong properties of embeddings. Many basic questions about quasisymmetric maps remain unanswered.

Quasisymmetry, or its local versions, is the right generalization of quasiconformality to a general metric space setting. Although this is not obvious from Definition 7.12, given for quasiconformal mappings in Chapter 7, it is true that a self-homeomorphism of $\mathbb{R}^n$, $n \geq 2$, is quasiconformal if and only if it is quasisymmetric quantitatively. The necessity part of this claim is not too difficult to show in view of the fact that $\mathbb{R}^n$ is a Loewner space. The sufficiency is more involved, see [182, Section 34] or [177].

Exercise 10.3. Show that the map $x \mapsto x^p$, $[0, \infty) \to [0, \infty)$, is quasisymmetric for each $p > 0$.

If you did Exercise 10.3, you noticed that in order to verify the quasisymmetry of a given embedding, one is often led to a tedious case study that analyzes the various possible mutual locations of the points a, b, x. Therefore, it would be helpful to have a result to the effect that only certain values of t need to be checked.

We call an embedding $f : X \to Y$ *weakly (H-)quasisymmetric* if there is a constant $H \geq 1$ so that

$$|x - a| \leq |x - b| \quad \text{implies } |f(x) - f(a)| \leq H|f(x) - f(b)| \tag{10.4}$$

for all triples a, b, x of points in X.

Weakly quasisymmetric maps need not be quasisymmetric in general. Let $X = \mathbb{N} \times \{0, -\frac{1}{4}\} \subset \mathbb{R}^2$ and let $f : X \to \mathbb{R}^2$ be the embedding defined by $f(n, 0) = (n, 0)$, and $f(n, -\frac{1}{4}) = (n, -\frac{1}{4n})$. Then, f is weakly quasisymmetric but not quasisymmetric.

Exercise 10.5. Verify the preceding example.

It will be shown in Theorem 10.19 that in many important instances weak quasisymmetry implies quasisymmetry. Before this, let us further develop the basic theory.

The next two propositions list some elementary but important properties of quasisymmetric maps; they will be used without further mention later.

Proposition 10.6. *If $f : X \to Y$ is η-quasisymmetric, then $f^{-1} : f(X) \to X$ is η'-quasisymmetric, where $\eta'(t) = 1/\eta^{-1}(t^{-1})$ for $t > 0$. Moreover, if $f : X \to Y$ and $g : Y \to Z$ are η_f- and η_g-quasisymmetric, respectively, then $g \circ f : X \to Z$ is $(\eta_g \circ \eta_f)$-quasisymmetric.*

Proposition 10.7. *The restriction to a subset of an η-quasisymmetric map is η-quasisymmetric.*

Väisälä in [188] has studied the question of to what extent the converse to Proposition 10.7 is true; that is, under what assumptions a "piecewise quasisymmetric map" is quasisymmetric.

The next proposition describes how the relative distortion of points gets transferred to subsets.

Proposition 10.8. *Quasisymmetric embeddings map bounded spaces to bounded spaces. More quantitatively, if $f : X \to Y$ is η-quasisymmetric and if $A \subset B \subset X$ are such that* $0 < \operatorname{diam} A \le \operatorname{diam} B < \infty$, *then* $\operatorname{diam} f(B)$ *is finite and*

$$\frac{1}{2\eta\left(\frac{\operatorname{diam} B}{\operatorname{diam} A}\right)} \le \frac{\operatorname{diam} f(A)}{\operatorname{diam} f(B)} \le \eta\left(\frac{2\operatorname{diam} A}{\operatorname{diam} B}\right). \tag{10.9}$$

Proof. To see that $\operatorname{diam} f(B)$ is finite, choose two sequences (b_n) and (b'_n) of points in B such that

$$\frac{1}{2}\operatorname{diam} B \le |b_n - b'_n| \to \operatorname{diam} B, \quad \text{as } n \to \infty.$$

Then, if $b \in B$, we have

$$|b - b_1| \le \operatorname{diam} B \le 2|b_1 - b'_1|,$$

implying

$$|f(b) - f(b_1)| \le \eta(2)|f(b_1) - f(b'_1)|,$$

from which it follows that $\operatorname{diam} f(B) < \infty$.

To prove inequality (10.9), pick x, a from A. Then,

$$|b_n - b'_n| \le |b_n - a| + |b'_n - a|,$$

from which, by symmetry, we may assume that

$$|b_n - a| \ge \frac{1}{2}|b_n - b'_n|,$$

which implies

$$\begin{aligned} |f(x) - f(a)| &\le \eta\left(\frac{|x - a|}{|b_n - a|}\right)|f(b_n) - f(a)| \\ &\le \eta\left(\frac{2\operatorname{diam} A}{|b_n - b'_n|}\right)\operatorname{diam} f(B). \end{aligned}$$

Because the right-hand side of this inequality converges to

$$\eta\left(\frac{2\operatorname{diam} A}{\operatorname{diam} B}\right)\operatorname{diam} f(B)$$

as $n \to \infty$, we conclude that the second inequality in (10.9) is valid.

Finally, the first inequality of (10.9) follows from the second by applying it to f^{-1} (see Proposition 10.6), and the proof of Proposition 10.8 is complete. □

The first of the next two propositions follows readily from Proposition 10.8. The second requires some straightforward arguments using only the definitions. We leave the proofs to the reader as an exercise.

Proposition 10.10. *Quasisymmetric maps take Cauchy sequences to Cauchy sequences. In particular, every quasisymmetric image of a complete space is complete.*

Proposition 10.11. *An η-quasisymmetric embedding of one space in another always extends to an η-quasisymmetric embedding to the completions.*

Exercise 10.12. Prove Propositions 10.10 and 10.11.

The Hausdorff dimension of a metric space is a bi-Lipschitz invariant, but not a quasisymmetric invariant. For instance, the quasisymmetric map in (10.2) changes the Hausdorff dimension α of X to $\epsilon^{-1}\alpha$. However, for metric concepts of dimension, it is often true that their positivity or finiteness is a quasisymmetric invariant.

10.13 Doubling spaces. In Chapter 1, we termed a metric measure space (X, μ) doubling if the measure μ satisfies the doubling condition (1.5) on balls. Next, we consider a similar definition without the presence of measure.

A metric space is called *doubling* if there is a constant $C_1 \geq 1$ so that every set of diameter d in the space can be covered by at most C_1 sets of diameter at most $d/2$. It is clear that the definition is hereditary: subsets of doubling spaces are doubling (with the same constant).

Alternate but equivalent definitions for doubling spaces are often used. For instance, in the definition one may replace arbitrary sets by balls. Moreover, doubling spaces have the following seemingly stronger covering property: there is a function $C_1 : (0, \frac{1}{2}] \to (0, \infty)$ such that every set of diameter d can be covered by at most $C_1(\epsilon)$ sets of diameter at most ϵd. The function C_1, called a *covering function of* X, can be chosen to be of the form

$$C_1(\epsilon) = C\epsilon^{-\beta} \tag{10.14}$$

for some $C \geq 1$ and $\beta > 0$.

Definition 10.15. Given a doubling metric space X, the infimum of all numbers $\beta > 0$ such that a covering function of the form (10.14) can be found is called the *Assouad dimension* of X.

Thus, doubling spaces are precisely the spaces of *finite Assouad dimension.*[1]

---

[1]The terms *(metric) covering dimension* and *uniform metric dimension* appear in the literature for the same concept.

Exercise 10.16. Show that the Hausdorff dimension of a metric space does not exceed its Assouad dimension. Then, show that the Assouad dimension of the compact set $\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ in $\mathbb{R}$ is one and conclude that the two dimensions are not equal in general.

Exercise 10.17. Show that the Assouad dimension of a space can be defined equivalently as the infimum of all numbers $\beta > 0$ with the property that every ball of radius $r > 0$ has at most $C\epsilon^{-\beta}$ disjoint points of mutual distance at least ϵr, for some $C \geq 1$ independent of the ball.

It is easy to see that $\mathbb{R}^n$ is doubling with a constant depending only on n, and in fact the Assouad dimension of $\mathbb{R}^n$ is n. Thus, every subset of Euclidean space is doubling.

If X carries a doubling measure as defined in Chapter 1, then it is easy to see that X is doubling according to the definition in Section 10.13. In the other direction, every complete doubling space carries a doubling measure. We will prove this for compact spaces in Chapter 13. The concept of a doubling space is, however, a genuine extension of the case where measures are present; there are doubling spaces, even open subsets of the real line, that do not carry doubling measures. (See Remark 13.20 (d).)

Quasisymmetric maps and doubling spaces and measures are closely related mathematical objects. This will become apparent in the remainder of the book.

The first basic observation is that the doubling condition (equivalently, finiteness of the Assouad dimension) is a quasisymmetric invariant. See also Exercise 13.22.

Theorem 10.18. *A quasisymmetric image of a doubling space is doubling, quantitatively.*

Proof. Let $f : X \to Y$ be an η-quasisymmetric homeomorphism. It suffices to show that every ball B of diameter d in Y can be covered by at most some fixed number C_2 of sets of diameter at most $d/4$. Let $B = B(y, R)$ and let

$$L = \sup_{z \in B} |f^{-1}(y) - f^{-1}(z)|.$$

Then, we can cover $f^{-1}(B)$ by at most $C_1(\epsilon)$ sets of diameter at most $\epsilon 2L$ for any $\epsilon \leq 1/2$, where C_1 is a covering function of X. Let $A_1, \ldots, A_p$ be such sets, so that $p = p(\epsilon) \leq C_1(\epsilon)$. We may clearly assume that $A_i \subset f^{-1}(B)$ for all $i = 1, \ldots, p$. Thus, $f(A_1), \ldots, f(A_p)$ cover B, and are contained in B, so that by Proposition 10.8 we compute that their diameters satisfy

$$\begin{aligned} \operatorname{diam} f(A_i) &\leq \operatorname{diam} B\, \eta\left(\frac{2\operatorname{diam} A_i}{\operatorname{diam} f^{-1}(B)}\right) \\ &\leq d\eta\left(\frac{4\epsilon L}{L}\right) \leq d\eta(4\epsilon). \end{aligned}$$

The theorem now follows upon choosing $\epsilon = \epsilon(\eta) > 0$ so small that $\eta(4\epsilon) \leq 1/4$. $\square$

We now return to the question of when weak quasisymmetry implies quasisymmetry.

Theorem 10.19. *A weakly quasisymmetric embedding of a connected doubling space in a doubling space is quasisymmetric, quantitatively.*

Proof. Let $f : X \to Y$ be weakly H-quasisymmetric, where both X and Y are doubling, and X in addition is connected. We may clearly assume that X is not a singleton. Pick three distinct points a, b, x from X, and write

$$t = \frac{|a-x|}{|b-x|}, \qquad t' = \frac{|f(a)-f(x)|}{|f(b)-f(x)|}.$$

We need to show that $t' \leq \eta(t)$, where $\eta(t) \to 0$ as $t \to 0$.

First, assume that $t > 1$. Denote by C_1 the covering function of X as in Eq. (10.14). Set $r = |b-x|$ and $\epsilon = \min\{r/2C_1(1/6t), r/2\}$. Join x to a by an ϵ-chain $x = x_0, \dots, x_N = a$ (that is, $|x_{i-1} - x_i| \leq \epsilon$ for all $i = 1, \dots, N$) and choose a subsequence $a_0, a_1, \dots$ of (x_i) so that $a_0 = x_0 = x$ and that a_{i+1} is the last point x_j with $|x_j - a_i| \leq r - i\epsilon$. Because $\epsilon \leq r/2$, the point a_{i+1} is well defined with $a_{i+1} \neq a_i$ whenever $\{a_0, \dots, a_i\} \subset B(x, tr)$ and $i\epsilon \leq r/2$. We will show that under these conditions, we have that $(i+1)\epsilon \leq r/2$.

To this end, we use the doubling property of X and cover $B(x, tr)$ by sets $A_1, \dots, A_m$, all of diameter not exceeding $r/3$, such that $m \leq C_1(1/6t)$. Because the mutual distance between each pair of points from $\{a_0, \dots, a_i\}$ is at least $r - i\epsilon \geq r/2 > r/3$, we have that $i + 1 \leq m$. In particular, $(i+1)\epsilon \leq r/2$, as desired.

It follows that there is an integer s such that the points $a_0, \dots, a_{s+1}$ are well defined, $|a_j - x| < tr$ for $0 \leq j \leq s$, and $|a_{s+1} - x| \geq tr = |a - x|$. Moreover, $s \leq C_1(1/6t)$ and $|a_{i+1} - a_i| \leq r - i\epsilon \leq |a_i - a_{i-1}|$ for $1 \leq i \leq s$. Therefore, we have that

$$|f(a) - f(x)| \leq H|f(a_{s+1}) - f(x)|,$$

and that

$$|f(a_{i+1}) - f(a_i)| \leq H^{i+1}|f(b) - f(x)|$$

for $0 \leq i \leq s$. Hence,

$$\begin{aligned} |f(a) - f(x)| &\leq H(H + H^2 + \cdots + H^{s+1})|f(b) - f(x)| \\ &\leq (s+1)H^{s+2}|f(b) - f(x)|. \end{aligned}$$

This proves the assertion in the case $t > 1$.

Next, assume that $t \leq 1$. This case is slightly more tedious because we must take care of the correct asymptotic behavior of $\eta(t)$ when $t \to 0$. In other words, using the same notation as previously, we must show that $t' \leq \eta(t)$, where $\eta(t) \to 0$ as

$t \to 0$. For this, we can assume that t is small to start with, say $t < 1/3$. (There is also the issue of making η homeomorphic, but this is easily achieved.)

By using the connectedness of X, one finds points $b = b_0, b_1, \ldots, b_{s-1}$ such that

$$|x - b_i| = 3^{-i}|b - x|$$

for $i = 0, \ldots, s-1$, where $s \geq 2$ is the least integer such that

$$3^{-s}|b - x| \leq |a - x|.$$

Thus,

$$\frac{\log 1/t}{\log 3} \leq s. \tag{10.20}$$

The inequalities

$$|f(a) - f(b_j)| \leq H|f(b_i) - f(b_j)|$$

and

$$|f(x) - f(b_j)| \leq H|f(b_i) - f(b_j)|$$

for $0 \leq i < j \leq s-1$ follow from the choices of the points and from the weak quasisymmetry of f; in particular,

$$|f(a) - f(x)| \leq 2H|f(b_i) - f(b_j)| \tag{10.21}$$

for all such i, j. Because also $|b_i - x| \leq |b - x|$, all the points $f(b_0), \ldots, f(b_{s-1})$ lie in the closed ball $B(f(x), H|f(b) - f(x)|)$, so that by the doubling property of Y and by inequality (10.21),

$$s \leq C'\left(\frac{1}{t'}\right)^{\beta'},$$

where $C' \geq 1$ and $\beta' > 0$ depend only on H and the doubling data of Y. By combining this with inequality (10.20), we see that $t' \to 0$ as $t \to 0$.

This completes the proof of Theorem 10.19. $\square$

Corollary 10.22. *A weakly quasisymmetric embedding of a connected subset of Euclidean space in another Euclidean space is quasisymmetric. In particular, a weakly quasisymmetric embedding $\mathbb{R}^p \to \mathbb{R}^n$, $1 \leq p \leq n$, is quasisymmetric.*

The statements in Corollary 10.22 are quantitative in the usual sense.

Remark 10.23. Although sufficient in many applications, in this book in particular, Theorem 10.19 does not cover infinite-dimensional situations. Väisälä has proved in [193, Theorem 6.6] that every weakly quasisymmetric embedding $f : X \to Y$ is quasisymmetric provided both X and $f(X)$ are quasiconvex. The proof is similar to the proof of Theorem 10.19, and the result is quantitative in the usual sense.

Exercise 10.24. Prove the result described in Remark 10.23.

10.25 Compactness properties of quasisymmetric embeddings. A family of η-quasisymmetric maps need not be compact. For instance, the family $\mathcal{F} = \{f_\lambda(x) = \lambda x : \lambda > 0\}$ consists of t-quasisymmetric maps, but it is not compact; infinite sequences of maps from the family need not subconverge to any embedding. Next, we describe what mild normalizing conditions need to be added in order to secure compactness.

Proposition 10.26. *Given metric spaces X and Y, the family*

$$\{f : X \to Y \text{ is } \eta\text{-quasisymmetric and } |f(a) - f(b)| \leq M\},$$

where a, b are some fixed distinct points in X, is equicontinuous.

Proof. Fix $x_0 \in X$. We can assume $a \neq x_0$ by interchanging the roles of a and b if necessary. Thus, for $x \in X$,

$$\begin{aligned} |f(x) - f(x_0)| &\leq \eta\left(\frac{|x - x_0|}{|a - x_0|}\right)|f(a) - f(x_0)| \\ &\leq \eta\left(\frac{|x - x_0|}{|a - x_0|}\right)\eta\left(\frac{|a - x_0|}{|a - b|}\right)|f(a) - f(b)| \\ &\leq \eta\left(\frac{|x - x_0|}{|a - x_0|}\right)\eta\left(\frac{|a - x_0|}{|a - b|}\right)M, \end{aligned}$$

which shows the equicontinuity at x_0, and the theorem follows. □

Corollary 10.27. *The family of all η-quasisymmetric embeddings of X in Y is equicontinuous if Y is bounded.*

Next, recall the following form of Ascoli's theorem [152, p. 179].

Theorem 10.28 (Ascoli's theorem). *An equicontinuous family $\mathcal{F}$ of maps from a separable topological space X to a metric space Y is normal provided the closure of the set*

$$\{f(x) : f \in \mathcal{F}\}$$

is compact in Y for each $x \in X$.

Recall that a family of maps from a topological space to a metric space is normal if every sequence from the family subconverges uniformly on compacta.

Exercise 10.29. Suppose that a sequence (f_n) of η-quasisymmetric maps from X to Y converges pointwise to a (not necessarily continuous) map $f : X \to Y$. Show that f is either a constant or an η-quasisymmetric embedding, and that the convergence is uniform on each compact set.

Ascoli's theorem together with the preceding equicontinuity results can be used in many situations to show that a given family of quasisymmetric maps is compact. The fact that quasisymmetric maps so easily form compact families is one of their most important properties. Many strong results about quasisymmetric maps can be proved by employing this property only.

Corollary 10.30. *Let $\mathcal{F}$ be a family of η-quasisymmetric embeddings of X in Y. Assume either that Y is compact or that Y is proper and there is $x_0 \in X$ and $y_0 \in Y$ such that $f(x_0) = y_0$ for each $f \in \mathcal{F}$. If*

$$C^{-1} \leq |f(a) - f(b)| \leq C$$

for some $C \geq 1$, for some pair of points $a, b \in X$, and for all $f \in \mathcal{F}$, then $\mathcal{F}$ is a sequentially compact family of embeddings; that is, every sequence (f_n) in $\mathcal{F}$ subconverges (uniformly on compacta) to an element in $\mathcal{F}$.

In other words, under the hypotheses of Corollary 10.30, the family $\mathcal{F}$ is a sequentially compact set in the topological space $\mathcal{C}(X, Y)$ consisting of all continuous maps from X to Y equipped with the *compact-open topology*. Recall that a proper space is one whose closed balls are compact.

Exercise 10.31. Verify Corollary 10.30.

There is one particularly important situation, where one in fact obtains a characterization of quasisymmetric embeddings based on compactness. Fix a real number $H \geq 1$, integers $1 \leq p \leq n$, and a unit vector $e \in \mathbb{R}^p \subset \mathbb{R}^n$, and denote by $\mathcal{Q}_H$ the collection of weakly H-quasisymmetric embeddings $f : \mathbb{R}^p \to \mathbb{R}^n$ that satisfy $f(0) = 0$ and $f(e) = e$. Note that in this case weak quasisymmetry quantitatively implies quasisymmetry (Corollary 10.22); the weak notation is used just for convenience. For an arbitrary embedding $f : \mathbb{R}^p \to \mathbb{R}^n$, denote by $\mathcal{W}_f$ the family of all embeddings of the form $g = \alpha \circ f \circ \beta$, where α and β are similarities of the pertinent Euclidean spaces so that $g(0) = 0$ and $g(e) = e$. Recall that a similarity in Euclidean space is a map that is a composition of translations, dilations, and rotations.

Theorem 10.32. *The family $\mathcal{Q}_H$ is compact. An embedding $f : \mathbb{R}^p \to \mathbb{R}^n$ is quasisymmetric if and only if $\mathcal{W}_f$ belongs to some compact family of embeddings.*

Proof. That $\mathcal{Q}_H$ is compact follows from the preceding discussion (Corollary 10.30). It thus remains to show that f is (weakly) quasisymmetric if $\mathcal{W}_f$ belongs to some compact family of embeddings. To this end, pick three points $a, b, x \in \mathbb{R}^p$ so that

$$|a - x| \leq |b - x|.$$

By pre- and postcomposing with appropriate similarities, we may assume that $x = 0 = f(x)$ and $b = e = f(b)$. It suffices to show that $f(a)$ lies in some fixed ball whose radius is independent of f. But this is clear in view of the fact that a lies in the closed unit ball of $\mathbb{R}^p$ and the fact that $\mathcal{W}_f$ lies in a compact family of embeddings. The theorem follows. $\square$

10.33 Notes to Chapter 10. Quasisymmetric maps first appeared in the celebrated paper [10] by Beurling and Ahlfors. The authors identified quasisymmetric maps of the real line as the boundary values of quasiconformal self-maps of the upper half-plane. This fact makes one-dimensional quasisymmetric maps important in the theory of Teichmüller spaces [120].

The concept of quasisymmetry was implicitly used in the early papers on quasiconformal mappings in Euclidean space. The term "quasisymmetric" was apparently first used by Kelingos in [103]. Tukia and Väisälä in [176] took up a systematic study of quasisymmetric embeddings in general spaces, and the material in this chapter is taken for the most part directly from that paper. Theorem 10.19, courtesy of Jussi Väisälä, is previously unpublished. The current literature on quasisymmetric maps is large. For starters, see [183], [184], [185], [193], [161], [84], [88], and the references therein. More remarks and references follow in the ensuing chapters.

Assouad dimension (but not the name) was introduced by Assouad [4]. Although related concepts had been studied earlier by many authors, it seems that Assouad was the first to employ explicitly the definition as given in Definition 10.15. Larman [119] studied metric spaces of finite Assouad dimension (under a different name), and sometimes such spaces are called *finite dimensional in the sense of Larman*. For an excellent survey of the concept, plus a historical account, see the paper by Luukkainen [125].

11

Quasisymmetric Maps: Basic Theory II

In this chapter, we first introduce an important quasisymmetrically invariant class of metric spaces and then prove that quasisymmetric maps between spaces from this class are Hölder continuous. Quantitative bounds for the change in Hausdorff dimension then follow. As a second topic, we discuss the relationship between quasisymmetry and quasiconformality for maps between Euclidean domains. Finally, as an example, we show how quasisymmetric maps naturally arise in one-variable complex dynamics.

11.1 Uniformly perfect spaces. A metric space X is called *uniformly perfect* if there is a constant $C \geq 1$ so that for each $x \in X$ and for each $r > 0$ the set $B(x, r) \setminus B(x, r/C)$ is nonempty whenever the set $X \setminus B(x, r)$ is nonempty. (For the sake of definiteness, assume here that the balls are open.)

The condition of uniform perfectness forbids isolated islands in a uniform manner. We call a set of the form $B(x, R) \setminus B(x, r)$ an *annulus* if $0 < r < R$ and if $X \setminus B(x, R)$ is nonempty. Thus, a space is uniformly perfect if for every empty annulus in the space the ratio R/r between the outer and inner radii is bounded from above.

Connected spaces are uniformly perfect, and so are many classical totally disconnected fractals; for instance, the Cantor ternary set is uniformly perfect. It is also easy to see that every Ahlfors regular space (see expression (8.10)) is uniformly perfect. For many problems in analysis, the condition of being uniformly perfect is just as good as being connected.

There is a rich theory of uniformly perfect sets in $\mathbb{R}^2$. Compact planar uniformly perfect sets can be characterized in terms of harmonic measure, Brownian motion, hyperbolic geometry, subharmonic functions, and modulus of curve families. The Julia sets of rational functions are uniformly perfect, and so are the limit sets of nonelementary, finitely generated Kleinian groups. (See the notes to this chapter for references.)

It is not difficult to see that the property of uniform perfectness is a quasisymmetric invariant.

Exercise 11.2. Show that the uniform perfectness is a quasisymmetrically invariant property of a metric space quantitatively.

We will prove the following result.

Theorem 11.3. *A quasisymmetric embedding f of a uniformly perfect space X is η-quasisymmetric with η of the form*

$$\eta(t) = C \max\{t^{\alpha}, t^{1/\alpha}\}, \tag{11.4}$$

where $C \geq 1$ and $\alpha \in (0, 1]$ depend only on the data associated with f and X.

Corollary 11.5. *Quasisymmetric embeddings of uniformly perfect spaces are Hölder continuous on bounded sets.*

Quasisymmetric embeddings whose control function η can be chosen to be of the form (11.4) were termed *power quasisymmetric* by Trotsenko and Väisälä in [174]. Such a choice is not always possible: the map $f(x) = -(\log x)^{-1}$ is quasisymmetric from the set $\{e^{-n!} : n = 2, 3, \dots\}$ into the reals but is not Hölder continuous. The authors of [174] characterize metric spaces whose quasisymmetric maps are power quasisymmetric.

In view of Corollary 11.5, the preceding quasisymmetric homeomorphism between compact subsets of $\mathbb{R}$,

$$X = \{0, e^{-n!} : n = 2, 3, \dots\} \to Y = \{0, (n!)^{-1} : n = 2, 3, \dots\}, \ x \mapsto -(\log x)^{-1},$$

extended to 0 by continuity, cannot be extended to a quasisymmetric self-homeomorphism of any Euclidean or Hilbert space. Every topological embedding $f : A \to \mathbb{R}^n$, $A \subset \mathbb{R}^k$, can be extended to a homeomorphism $F : \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$, by the so-called *Klee trick* (see [155, pp. 74–75]). If f is L-bi-Lipschitz, then F can be chosen to be $\sqrt{7}L^2$-bi-Lipschitz by [2, Theorem 5.5]. The failure of a similar extension result in the quasisymmetric category can be attributed to the fact that quasisymmetric maps do not in general behave well in products.

Väisälä in [191] posed the following question.

11.6 Open problem. Can every quasisymmetric homeomorphism between connected sets in Euclidean space be extended to a quasisymmetric self-homeomorphism of some larger-dimensional Euclidean space?

For the proof of Theorem 11.3, it is convenient to use the following equivalent description of uniformly perfect spaces.

Lemma 11.7. *A metric space X is uniformly perfect if and only if there are numbers $0 < \lambda_1 \le \lambda_2 < 1$ so that for each pair of points $a, b \in X$ there is a point $x \in X$ with*

$$\lambda_1|a-b| \le |x-a| \le \lambda_2|a-b|. \tag{11.8}$$

The assertion is quantitative in the usual sense.

Proof. If X is C-uniformly perfect, then obviously a point $x \in X$ can be found that satisfies

$$\frac{1}{2C}|a-b| \le |a-x| \le \frac{1}{2}|a-b|$$

for any given pair of points a and b in X, so that X satisfies condition (11.8) with $\lambda_1 = (2C)^{-1}$ and $\lambda_2 = \frac{1}{2}$.

To prove the converse, let $B(a,r)$ be a ball such that $X \setminus B(a,r)$ contains a point b. We will show that the annulus $B(a,r) \setminus B(a,\lambda_1 r)$ is not empty, so that X is uniformly perfect with constant $C = \lambda_1^{-1}$. This is achieved by iterating condition (11.8): pick a point $x_0 \in X$ such that

$$\lambda_1|a-b| \le |a-x_0| \le \lambda_2|a-b|,$$

and observe that if $|a-x_0| < r$, then x_0 lies in the annulus $B(a,r) \setminus B(a,\lambda_1 r)$ as desired; on the other hand, if $|a-x_0| \ge r$, we repeat the preceding reasoning with b replaced with x_0 to find a point x_1 with

$$\lambda_1|a-x_0| \le |a-x_1| \le \lambda_2|a-x_0| \le \lambda_2^2|a-b|,$$

and so on, eventually reaching a point x_k for which

$$\lambda_1 r \le \lambda_1|a-x_{k-1}| \le |a-x_k| \le \lambda_2^{k+1}|a-b| < r.$$

This proves Lemma 11.7. $\square$

Proof of Theorem 11.3. Fix three distinct points $a, b, x \in X$ and write

$$t = \frac{|a-x|}{|b-x|}.$$

We separate two cases depending on whether t is less than 1 or not.

Let us first assume $t \ge 1$. By condition (11.8), we find points $x_0 = a, x_1, \dots, x_s$, $s \ge 0$, such that

$$\lambda_1|a-x| \le |x_1-x| \le \lambda_2|a-x|,$$

that

$$\lambda_1|x_i-x| \le |x_{i+1}-x| \le \lambda_2|x_i-x| \le \lambda_2^{i+1}|a-x|$$

for $i = 0, \dots, s-1$, and that

$$\lambda_1|x_s-x| \le |b-x| \le |x_s-x| \le \lambda_2^s|a-x|.$$

This implies that

$$|f(a) - f(x)| \leq H^{s+1}|f(b) - f(x)|, \tag{11.9}$$

where $H = \eta(\lambda_1^{-1})$. On the other hand,

$$|b - x| \leq \lambda_2^s |a - x| = \lambda_2^s t |b - x|,$$

so that

$$t \geq \lambda_2^{-s}.$$

By combining this with Eq. (11.9), we obtain

$$|f(a) - f(x)| \leq Ht^{\alpha}|f(b) - f(x)|,$$

where $\alpha = (\log H)(\log \lambda_2^{-1}) \geq 1$. This proves the claim in the case $t \geq 1$.

Assume now that $t < 1$. We use Exercise 11.2 and observe that $f(X)$ is uniformly perfect; in particular, by Lemma 11.7, $f(X)$ satisfies condition (11.8) for some appropriate positive constants $0 < \lambda_1 \leq \lambda_2 < 1$. It is no loss of generality to assume that $\eta(\lambda_2) \leq \frac{1}{2}$, for we can always replace the pair (λ_1, λ_2) by $(\lambda_1^n, \lambda_2^n)$ for any positive integer n. In particular, as earlier, we can find points $x_0 = b, x_1, \dots, x_s$, $s \geq 0$, such that

$$\lambda_1 |x_i - x| \leq |x_{i+1} - x| \leq \lambda_2 |x_i - x|$$

for $i = 0, \dots, s - 1$, and

$$\lambda_1 |x - x_s| \leq |a - x| \leq |x - x_s|.$$

Thus,

$$|f(x_{i+1}) - f(x)| \leq \eta(\lambda_2)|f(x_i) - f(x)| \leq \frac{1}{2}|f(x_i) - f(x)|,$$

which implies

$$|f(x_s) - f(x)| \leq 2^{-s}|f(b) - f(x)|.$$

Furthermore, we have

$$|f(a) - f(x)| \leq H|f(x_s) - f(x)|,$$

where $H = \eta(1)$, and hence

$$\frac{|f(a) - f(x)|}{|f(b) - f(x)|} = t' \leq H2^{-s}.$$

On the other hand,

$$|a - x| \geq \lambda_1 |x - x_s| \geq \lambda_1^2 |x - x_{s-1}| \geq \cdots \geq \lambda_1^{s+1}|x - b|,$$

so that

$$t' \leq 2H t^{\alpha},$$

where $\alpha = \log 2 / \log \lambda_1^{-1} > 0$. This completes the proof of Theorem 11.3. $\square$

Corollary 11.10. *If $f : X \to Y$ is an η-quasisymmetric embedding of a C-uniformly perfect metric space in a metric space Y, then for each subset A of X we have that*

$$\alpha \dim_H A \leq \dim_H f(A) \leq \frac{1}{\alpha} \dim_H A$$

for some $0 < \alpha \leq 1$ depending only on η and C, where $\dim_H$ *denotes the Hausdorff dimension.*

There is a corresponding result for the Assouad dimension; see Exercise 13.22.

If X is a locally compact, uniformly perfect metric space, then it is not difficult to construct a quasisymmetric embedding of a compact subset A of X onto the standard Cantor ternary set. Thus, we can conclude the following result.

Corollary 11.11. *The Hausdorff dimension of a locally compact, uniformly perfect metric space is bounded from below by a positive constant that depends only on the constant of uniform perfectness.*

Exercise 11.12. Prove Corollary 11.11 by following the instructions given.

One can also deduce Corollary 11.11 from the results in Chapter 13. (See Exercise 13.1 and Theorem 13.3.)

11.13 Quasisymmetry in Euclidean domains. According to Definition 7.12, a homeomorphism $f : D \to D'$ between domains in $\mathbb{R}^n$, $n \geq 2$, is quasiconformal if it distorts the moduli of curve families according to expression (7.13). Quasiconformal homeomorphisms can also be characterized as locally quasisymmetric maps in the following precise sense.

Theorem 11.14. *A homeomorphism $f : D \to D'$ between domains in $\mathbb{R}^n$, $n \geq 2$, is K-quasiconformal if and only if there is η such that f is η-quasisymmetric in each ball $B(x, \frac{1}{2} dist(x, \partial D))$ for $x \in D$. The statement is quantitative involving K, η, and the dimension n.*

It is a somewhat long story to show that (locally) quasisymmetric maps are quasiconformal according to the modulus definition 7.12; we omit the proof here. See [182] for a complete account. (See also [178] for a recent new approach.)

The necessity part of Theorem 11.14 can be shown rather painlessly, since we have the Loewner studies of Chapter 9 behind us. Observe that if $D = D' = \mathbb{R}^n$ in Theorem 11.14, then f is quasiconformal if and only if f is quasisymmetric quantitatively. It is also true that a homeomorphism between compact Riemannian manifolds is quasiconformal if and only if it is quasisymmetric, but this statement is not quantitative without some additional assumptions. For example, conformal transformations on a sphere are not uniformly quasisymmetric.

The necessity part of Theorem 11.14 is sometimes called the *egg yolk principle*[1]. Namely, if f is a K-quasiconformal map of a ball $B(x, 2r)$ (the *egg*) in $\mathbb{R}^n$, then f is quasisymmetric in the smaller ball $B(x, r)$ (the *yolk*) with quasisymmetry function η that depends only on K and n. In particular, the image of the yolk $B(x, r)$ remains somewhat round and (see Exercise 11.18) has diameter roughly the distance to the boundary, no matter how badly and widely the image of the egg $B(x, 2r)$ spreads and spills near the boundary. In the theory of univalent functions, the egg yolk principle is better known as the *Koebe distortion theorem* [43, p. 32].

PROOF OF NECESSITY IN THEOREM 11.14. Let $B = B(x, r)$ for some $x \in D$, where $r = \frac{1}{2}\operatorname{dist}(x, \partial D)$. By Corollary 10.22, it suffices to show that f is weakly quasisymmetric in B.

To this end, pick three distinct points a, z, w from B such that $|a - z| \leq |a - w|$. We must show that there is $H = H(n, K) \geq 1$ such that

$$|f(a) - f(z)| \leq H|f(a) - f(w)|. \tag{11.15}$$

In what follows, H will denote any positive constant that depends on n and K only. First, we will show that there is a path γ' joining $f(a)$ and $f(w)$ in D' such that

$$\operatorname{diam}\gamma' \leq H\,|f(a) - f(w)|. \tag{11.16}$$

Indeed, if the line segment from $f(a)$ to $f(w)$ does not meet the boundary of D', there is nothing to prove. Thus, let σ' be the closed subsegment of $[f(a), f(w)]$ that connects $f(a)$ to $\partial D'$ inside D'. (To simplify the discussion, we ignore the fact that one of the endpoints of σ' lies on the boundary of D', where f^{-1} may not be defined. A simple limiting argument can be used to make the proof completely rigorous.) Now, the preimage σ of σ' connects a to the boundary of D. Assume next that y' is a point on $\gamma' = f([a, w])$ such that $|y' - f(a)| > 2|f(a) - f(w)|$; if no such point exists, then γ' satisfies expression (11.16) with $H = 4$. Consider the half line from $f(a)$ that passes through y', and let L' be the part of that line that lies in D' but outside the ball $B(f(a), |f(a) - y'|)$ and connects y' to the boundary of D'. Note that L' may not meet any finite part of $\partial D'$. By the standard modulus estimate in Example 7.14, we find that the curve family Γ' that joins σ' to L' inside D' has modulus at most

$$\omega_{n-1}\left(\log\frac{|y' - f(a)|}{|f(a) - f(w)|}\right)^{1-n}. \tag{11.17}$$

On the other hand, in the domain D, both preimages $\sigma = f^{-1}(\sigma')$ and $L = f^{-1}(L')$ connect the ball $B = B(x, r)$ to the complement of $B(x, 2r)$. Because $B(x, 2r)$ lies in D, the modulus of the preimage family $\Gamma = f^{-1}\Gamma'$ has a lower bound that depends only on n, by Loewner-type estimates as explained in Chapter 8.

[1]The term has been attributed to Curt McMullen.

The quasi-invariance of modulus thus implies, in view of expression (11.17), that $|y' - f(a)| \leq H|f(a) - f(w)|$, and expression (11.16) follows.

Once we have the curve γ' satisfying expression (11.16), then expression (11.15) can be established by an argument that is very similar to the preceding one. Namely, let γ' take the role of σ', so if $|f(a) - f(z)|$ is much larger than $|f(a) - f(w)|$, it is also much larger than $\operatorname{diam} \gamma'$. Then, a part L' of the half line from $f(a)$ through $f(z)$ can be used to obtain appropriate modulus estimates in D'. On the other hand, the assumption $|a - z| \leq |a - w|$ about the mutual location of the points allows for Loewner-type estimates in $B(x, 2r) \subset D$. We leave the details to the reader. The theorem follows. □

In the preceding proof, we only used an estimate that bounds modulus from below. The full version of the egg yolk principle can be established by invoking the asymptotic behavior of the Loewner function $\phi_n(t)$ in $\mathbb{R}^n$ for small values of t. See Theorem 8.23 (3). The following is a good exercise in the art of modulus method.

Exercise 11.18. Let $f : D \to D'$ be a K-quasiconformal homeomorphism between domains D and D' in $\mathbb{R}^n$, $n \geq 2$, and let $B = B(x, r)$ be a ball in D, where $r = \frac{1}{2}\operatorname{dist}(x, \partial D)$. Show that

$$H^{-1}\operatorname{dist}(f(B), \partial D') \leq \operatorname{diam} f(B) \leq H \operatorname{dist}(f(B), \partial D')$$

for some constant $H \geq 1$ depending only on K and n.

A modest elaboration of the preceding modulus argument yields some interesting global conclusions about quasiconformal maps of uniform domains. In Example 8.24 (f), it was mentioned that uniform domains are Loewner spaces, by Jones's Sobolev extension theorem. Thus, if we could join *each* pair of points in the image domain by a continuum that is not too big relative to the distance between the points, and by another continuum that avoids a given small (relative to the distance between the points) ball, then the argument for Theorem 11.14 would work globally. The condition of linear local connectivity, defined just before Theorem 8.23, is precisely what one needs here. In the end, the following theorem is not difficult to establish.

Theorem 11.19. *A quasiconformal map from a bounded uniform domain onto a bounded, linearly locally connected domain in $\mathbb{R}^n$ is quasisymmetric.*

There is a subtlety in Theorem 11.19, which is why we made the safe assumption that the domains be bounded. Namely, the assertion is not quantitative. One only has to think of the conformal self-automorphism group of the unit disk; although each individual map in the group is quasisymmetric, there is no η that works for all of them. One could remedy this defect by considering *quasimöbius maps* instead. These are maps that are defined similarly to quasisymmetric maps by using a four-point condition rather than a three-point condition, and in some sense they form a more natural class of embeddings to consider. To keep matters simple, in this book we restrict ourselves to quasisymmetry and triples instead of quasimöbius maps and quadruples. (For a basic theory of quasimöbius maps, including a proof of Theorem 11.19, see [184].)

Some important corollaries can be drawn from Theorem 11.19. For example, it follows from it and from Theorem 11.3 that a quasiconformal map from the unit disk in the plane onto a bounded, linearly locally connected Jordan domain is Hölder continuous. This fact is nontrivial already for conformal maps. Simply connected, linearly locally connected domains in the plane can be characterized as *quasidisks* [62], [63].

One can also prove results that are analogous to Theorem 11.19 by using the internal metric of the domain. This leads to various John-type conditions; see Example 8.24 (g) and the references there.

To mention one last application, it follows from Theorem 11.19 that a quasiconformal mapping cannot map a uniform domain onto a linearly locally connected domain unless the target itself is uniform. This is because uniform domains are preserved under quasisymmetric (or quasimöbius) maps, which is not difficult to see. In particular, we obtain the well-known example of Gehring and Väisälä [65]: a ball in $\mathbb{R}^n$, $n \geq 3$, cannot be mapped quasiconformally onto a domain that contains an inward-directed spike or an outward-directed wedge. In contrast, maps to domains with outward-directed spikes and inward-directed wedges are constructed in [65]. Thus, there are topological surfaces in $\mathbb{R}^n$, $n \geq 3$, that are smooth except at one point dividing $\mathbb{R}^n$ into two components, only one of which is quasiconformally equivalent to a ball. Another corollary is that a domain of the form $\mathbb{B}^k \times \mathbb{R}^{n-k}$ is quasiconformally equivalent to the open n-ball $\mathbb{B}^n$ if and only if $k = n - 1$. For the construction of a map from $\mathbb{B}^n$ to the cylinder $\mathbb{B}^{n-1} \times \mathbb{R}$, see [182, pp. 50–51].

11.20 Holomorphic motions and quasisymmetric maps. We close this chapter with an example from complex dynamics. It shows how naturally the concept of quasisymmetry arises in mathematics.

A *holomorphic motion* of a set A in the complex plane $\mathbb{C}$ is a map $f : A \times \mathbb{B}^2 \to \mathbb{C}$ such that

(1) $\lambda \mapsto f(\lambda, a)$ is holomorphic for each $a \in A$;
(2) $a \mapsto f(\lambda, a)$ is injective for each $\lambda \in \mathbb{B}^2$;
(3) $f(0, a) = a$ for each $a \in A$.

Here, $\mathbb{B}^2$ denotes the open unit disk in $\mathbb{C}$. Note in particular that no continuity assumption is made in (2), nor is it assumed that the set A has any particular structure.

Both in the theory of iteration of polynomials and in the theory of Kleinian groups, natural situations arise where sets depend holomorphically on complex parameters. If $\{\Gamma_\lambda : \lambda \in \mathbb{B}^2\}$ is an analytic family of Kleinian groups, then the loxodromic fixed points move holomorphically. If c moves in a hyperbolic component of the Mandelbrot set, then the repelling periodic points of the polynomials $z \mapsto z^2 + c$ move holomorphically. In both cases, the holomorphically moving set is dense in a compact set of interest, either in the limit set of a Kleinian group or in the Julia set of a polynomial.

Each holomorphic motion is in fact a quasisymmetric map of A and moreover can be extended to a holomorphic motion of the entire plane. (The proper language here would involve quasimöbius maps and the Riemann sphere rather than quasisymmetric maps and the finite plane.) The first assertion is the following celebrated λ-lemma of Mañé, Sad, and Sullivan.

Theorem 11.21. *If f is a holomorphic motion of a set A, where $\operatorname{card} A \geq 3$, then the map $a \mapsto f(\lambda, a)$ is an η-quasisymmetric embedding of A into $\mathbb{C} = \mathbb{R}^2$, with η depending only on $|\lambda| < 1$.*

Proof. Fix three distinct points a, z, w from A. Then, consider the holomorphic function

$$F(\lambda) = \frac{f(\lambda, a) - f(\lambda, z)}{f(\lambda, a) - f(\lambda, w)}.$$

Because $f(\lambda, \cdot)$ is injective, F is indeed holomorphic in $\mathbb{B}^2$, and it omits the points 0 and 1. The twice-punctured finite plane admits a hyperbolic metric $\rho_{0,1}$; that is, $\rho_{0,1}$ is a complete metric of constant negative curvature in $\mathbb{C} \setminus \{0, 1\}$. By the generalized Schwartz lemma [115, Chapter 2, p. 71], the map F is distance nonincreasing in the hyperbolic metrics $\rho_{\mathbb{B}^2}$ and $\rho_{0,1}$; that is,

$$\rho_{0,1}(F(\lambda), F(0)) \leq \rho_{\mathbb{B}^2}(\lambda, 0) = \log \frac{1 + |\lambda|}{1 - |\lambda|}. \tag{11.22}$$

For the required quasisymmetry, we need to show that

$$|F(\lambda)| \leq \eta(|\lambda|, t),$$

where

$$t = \frac{|a - z|}{|a - w|} = F(0)$$

and $\eta(|\lambda|, t) \to 0$ as $t \to 0$. This is achieved easily by using Eq. (11.22) and some elementary properties of the hyperbolic metric $\rho_{0,1}$. The λ-lemma follows. □

Slodkowski [168] proved that each holomorphic motion can in fact be extended to the entire plane as a quasisymmetric map. A proof of this result in this book would take us too far.

11.23 Notes to Chapter 11. Papers on uniformly perfect sets and their role in analysis include [3], [8], [49], [68], [92], [98], [107], [121], [129], and [200]. Theorem 11.3 and its corollaries are due to Tukia and Väisälä [176]. Corollary 11.11

is part of folklore. Condition (11.8) was introduced in [176], where spaces satisfying it were called *homogeneously dense*.

The "only if" part of Theorem 11.14 was first proved by Väisälä in [183]. Distortion results such as the one in Exercise 11.18 were known to the early quasiconformal mappers almost forty years ago. Theorem 11.19 was essentially proved by Gehring and Martio in [64], but in its present formulation the result appeared in [184]. Elaborations of the result led to the so-called *subinvariance principle*, first published in [50], which has turned out to be an extremely useful tool in mapping problems. For applications and generalizations of the subinvariance principle, see [186], [187], [189], [192], [79], and [87].

Very little is known about the general mapping problem, which asks what domains in Euclidean n-space, $n \geq 3$, are quasiconformally equivalent. For interesting characterizations in special cases, see [61] and [186].

The original paper on holomorphic motions and quasisymmetry is [130]. For an account of Slodkowski's theorem, as well as more history, background, and applications, see the beautiful paper [6]. For further discussion, see [5], [137], and the references therein.

12

Quasisymmetric Embeddings of Metric Spaces in Euclidean Space

As in topology, where one wants to understand the homeomorphism type of a given space, a basic question in the theory of quasisymmetric maps asks which metric spaces are quasisymmetrically homeomorphic. This question is extremely difficult in general. There are closed (compact without boundary) four-dimensional Riemannian manifolds that are homeomorphic but not quasisymmetrically homeomorphic [42]. An easier question asks which spaces can be embedded quasisymmetrically in a given space or in a space from a given collection. A beautiful and complete answer to this problem can be given in the case where the receiving space is Euclidean.

Theorem 12.1. *A metric space is quasisymmetrically embeddable in some Euclidean space if and only if it is doubling, quantitatively.*

The necessity in Theorem 12.1 follows from Theorem 10.18. As for the sufficiency, Assouad [4] proved a stronger statement, which we describe next.

If $(X, |x-y|)$ is a metric space, then its *snowflaking* is a metric space $(X, |x-y|^{\epsilon})$, where $0 < \epsilon < 1$. We also say that $(X, |x-y|^{\epsilon})$ is a *snowflaked version* of $(X, |x-y|)$. The snowflaking identity map $(X, |x-y|) \to (X, |x-y|^{\epsilon})$ is t^{ϵ}-quasisymmetric.

Theorem 12.2. *Each snowflaked version of a doubling metric space admits a bi-Lipschitz embedding in some Euclidean space.*

Theorem 12.2 is quantitative in the sense that the dimension of the receiving space as well as the quasisymmetry control function η depend both on the doubling constant of the source space and on the amount ϵ of snowflaking.

It is not true that one can take $\epsilon = 1$ in Theorem 12.2. The Heisenberg group with its Carnot metric as described in Section 9.25 cannot be bi-Lipschitzly embedded in any Euclidean space. This follows from a work of Pansu [147], which implies that every Lipschitz map from the Heisenberg group to $\mathbb{R}^n$ is almost everywhere differentiable, with differential that is an algebra homomorphism from the Heisenberg Lie algebra to $\mathbb{R}^n$; because the latter algebra is Abelian and the former is not, such a differential must have a nontrivial kernel, which is an impossible conclusion in the case of a bi-Lipschitz map. Neither the Bourdon–Pajot spaces nor the Laakso spaces (see Chapter 8) can be bi-Lipschitzly embedded in a Euclidean space [116], [31].

In contrast to Theorem 12.1, only a few nontrivial criteria that would guarantee bi-Lipschitz embeddability of a metric space in some $\mathbb{R}^n$ are known. (See the notes to this Chapter.)

12.3 Open problem. Characterize the metric spaces that can be embedded bi-Lipschitzly in some Euclidean space.

Recall that every metric space embeds isometrically in some Banach space. This so-called *Kuratowski embedding* embeds a given space X in the Banach space $\ell^\infty(X)$, consisting of all bounded functions on X with the (uniform) sup-norm, via

$$x \mapsto \{y \mapsto |x_0 - y| - |x - y|\}, \tag{12.4}$$

where x_0 is a fixed point in X.

Exercise 12.5. Check that the embedding $X \to \ell^\infty(X)$ given in expression (12.4) is an isometry to its image.

For separable spaces, a better embedding is possible, as follows.

Exercise 12.6. Prove that every separable metric space can be embedded isometrically in the Banach space $\ell^\infty = \ell^\infty(\mathbb{N})$, where $\mathbb{N}$ is the set of natural numbers.

The embedding results in Exercises 12.5 and 12.6 are not needed in this book. We begin the proof of Theorem 12.2 with the following lemma.

Lemma 12.7. *Let X be a metric space. Suppose that E is a finite-dimensional Hilbert space and that there exist positive numbers A, B, and $0 < \tau < 1$, and a sequence of maps*

$$\varphi_j : X \to E, \qquad j \in \mathbb{Z},$$

such that the following two conditions hold:

(1) $|\varphi_j(s) - \varphi_j(t)| \geq A$ *if* $s, t \in X$ *satisfy* $\tau^{j+1}C < |s - t| \leq \tau^j C$;
(2) $|\varphi_j(s) - \varphi_j(t)| \leq B \min\{\tau^{-j}|s - t|, 1\}$ *for all* $s, t \in X$.

Then, for each $0 < \epsilon < 1$, there is a bi-Lipschitz embedding f of the metric space $(X, |x - y|^\epsilon)$ in some $\mathbb{R}^N$; that is, $f : X \to \mathbb{R}^N$ is an embedding that *satisfies*

$$L^{-1}|s-t|^{\epsilon} \leq |f(s)-f(t)| \leq L\,|s-t|^{\epsilon} \tag{12.8}$$

for some $L \geq 1$ and all $s, t \in X$. The assertion is quantitative in that the bi-Lipschitz constant L and the dimension N depend only on the data: $\dim E$*, A, B, τ, and ϵ.*

PROOF. Fix a Hilbert space F of dimension $2d$, where $d \geq 1$ will be chosen later, and let $e_1, \ldots, e_{2d}$ be an orthonormal basis of F. We extend (e_j) periodically to all integers j, $e_{2d+j} = e_j$, and define

$$f(s) = \sum_{j\in\mathbb{Z}} \tau^{j\epsilon}\varphi_j(s) \otimes e_j,$$

where $\otimes$ denotes the tensor product. (To guarantee that the preceding sum converges, we can replace φ_j by $\varphi_j - \varphi_j(x_0)$ for some fixed $x_0 \in X$, if necessary.) We claim that $f: X \to E \otimes F$ satisfies expression (12.8) for d large enough, depending on the data. Because the tensor product $E \otimes F$ is isomorphic to $\mathbb{R}^N$ for $N = \dim E \cdot \dim F$, this suffices.

To prove the claim, we first estimate $|f(s)-f(t)|$ from above for given $s, t \in X$. Let l be the unique integer such that $\tau^{l+1} < |s-t| \leq \tau^l$. Then,

$$\begin{aligned}
|f(s)-f(t)| &\leq \sum_{j\geq l} \tau^{j\epsilon}|\varphi_j(s)-\varphi_j(t)| + \sum_{j<l} \tau^{j\epsilon}|\varphi_j(s)-\varphi_j(t)| \\
&\leq B\sum_{j\geq l}\tau^{j\epsilon} + B\sum_{j<l}\tau^{j\epsilon}|s-t|\tau^{-j} \\
&\leq C\left(\tau^{l\epsilon} + |s-t|\tau^{l(\epsilon-1)}\right) \leq C\,|s-t|^{\epsilon},
\end{aligned}$$

where $C \geq 1$ is a constant depending only on the data. Note that assumption (2) was used in the second inequality, and that no restriction on d was needed here.

To obtain a lower bound, we estimate, with a choice of integer l as earlier,

$$\begin{aligned}
|f(s)-f(t)| \geq{} & \Big|\sum_{-d\leq j-l<d} \tau^{j\epsilon}(\varphi_j(s)-\varphi_j(t))\otimes e_j\Big| \\
& - \sum_{j\geq d+l}\tau^{j\epsilon}|\varphi_j(s)-\varphi_j(t)| - \sum_{j<-d+l}\tau^{j\epsilon}|\varphi_j(s)-\varphi_j(t)|.
\end{aligned} \tag{12.9}$$

The sum in the middle is at least

$$-B\sum_{j\geq d+l}\tau^{j\epsilon} \geq -C\,\tau^{(d+l)\epsilon},$$

while the last sum is at least

$$-B\sum_{j<-d+l}\tau^{j\epsilon}\tau^{-j}|s-t| \geq -C\tau^{(-d+l)(\epsilon-1)}|s-t|,$$

by assumption (2), for some $C > 0$ depending only on the data. Next, the terms in the first sum in expression (12.9) are mutually orthogonal, so its absolute value is at least

$$\tau^{l\epsilon}|\varphi_l(s) - \varphi_l(t)| \geq \tau^{l\epsilon} A.$$

By combining the last three estimates with expression (12.9), we have

$$\begin{aligned} |f(s) - f(t)| &\geq \tau^{l\epsilon} A - C\,\tau^{(d+l)\epsilon} - C\tau^{(-d+l)(\epsilon-1)}|s-t|, \\ &\geq \tau^{l\epsilon} A - C\,\tau^{d\epsilon}|s-t|^{\epsilon} - C\tau^{-d(\epsilon-1)}|s-t|^{\epsilon} \end{aligned}$$

for some $C > 0$ depending only on the data. We can choose $d \geq 1$ large enough, depending only on the data, so that the desired lower bound in expression (12.8) follows. This proves the lemma. □

PROOF OF THEOREM 12.2. Fix $\tau = \frac{1}{2}$, and an integer $j \in \mathbb{Z}$. We construct a map $\varphi = \varphi_j : X \to \mathbb{R}^M$ such that conditions (1) and (2) of Lemma 12.7 will be satisfied with $A = \frac{1}{2}$ and $B = 8C_0$, where M and C_0 depend only on the doubling constant of X and will be determined in the course of the proof. This suffices, by Lemma 12.7.

First, set $c = \frac{1}{4}\tau^{j+1}$ and choose a c-net Y in X. By a *c-net in X* we mean a subset of X whose c-neighborhood is all of X and with the property that any two distinct points from the subset lie at least distance c apart. Thus, we require that

$$\bigcup_{y\in Y} B(y, c) = X,$$

where $B(y, c)$ is an open ball of radius c about y, and that $|y - y'| \geq c$ whenever $y, y' \in Y$ are distinct. By Zorn's lemma, one can easily show that such a net Y exists.

Exercise 12.10. Show that c-nets exist in any metric space for each $c > 0$.

Now, for each $y \in Y$, the set $Y \cap \{x \in X : |x - y| \leq 12c\}$ has at most M elements in it, where M depends on the doubling constant of X only. Let $\{e_1, \dots, e_M\}$ be an orthonormal basis in $\mathbb{R}^M$. We think of the vectors $e_1, \dots, e_M$ as *colors*. A *coloring* of Y is a map

$$k : Y \to \{1, \dots, M\}$$

that satisfies $k(y) \neq k(y')$ if $|y - y'| \leq 12c$.

Exercise 12.11. Show that a coloring k as in Exercise 12.10 exists. More generally, show that if Z is a metric space with the property that the cardinality of each closed ball of radius $b > 0$ in it is at most M, then there exists a map $k : Z \to \{1, \dots, M\}$ such that $k(z) \neq k(z')$ if $|z - z'| \leq b$. (Hint: Use the fact that every set can be well ordered and transfinite induction.)

Fix a coloring k as previously, and define, for $s \in X$,

$$\varphi(s) = \sum_{y \in Y} g_y(s) e_{k(y)}, \tag{12.12}$$

where

$$g_y(s) = \max\{2c - |s - y|, 0\}(2c)^{-1}.$$

We claim that the map $\varphi(s)$ is the desired map.

To prove the claim, observe first that $g_y(s) \neq 0$ if and only if $|y - s| < 2c$. In particular, there is a constant C_0 such that in the sum in the definition of the map φ in Eq. (12.12) at most C_0 terms are nonzero. One also easily checks that

$$|g_y(s) - g_y(t)| \leq (2c)^{-1}|s - t| = 4\tau^{-j}|s - t| \tag{12.13}$$

for all $s, t \in X$. It follows, therefore, that

$$|\varphi(s) - \varphi(t)| \leq |\varphi(s)| + |\varphi(t)| \leq 2\sqrt{C_0}$$

and that

$$|\varphi(s) - \varphi(t)| \leq 2C_0 4\tau^{-j}|s - t|$$

for all $s, t \in X$. In particular, condition (2) in Lemma 12.7 is satisfied with $B = 8C_0$.

Finally, let $s, t \in X$ be such that $4c = \tau^{j+1} < |s - t| \leq \tau^j = 8c$. In this case, the vectors $\varphi(s)$ and $\varphi(t)$ are orthogonal, so that

$$|\varphi(s) - \varphi(t)|^2 = \sum_{y \in Y} |g_y(s)|^2 + \sum_{y \in Y} |g_y(t)|^2.$$

Because we must have $|s - y| < c$ for some $y \in Y$, the corresponding term satisfies

$$|g_y(s)| = (2c - |s - y|)(2c)^{-1} \geq \frac{1}{2}.$$

Thus, condition (1) in Lemma 12.7 holds as well.

The proof of Theorem 12.2 is thereby complete. □

12.14 Notes to Chapter 12. Theorem 12.2 and its proof are due to Assouad [4]. The problem of bi-Lipschitz embeddability of metric spaces in Euclidean space has been studied by Luukkainen and Movahedi-Lankarani [126], Luosto [124], Semmes [163], and Lang, Pavlović, and Schroeder [118]. See [125] for a survey of these and other related issues. Semmes has raised the question whether it is easier to find *regular maps* from a metric space to some $\mathbb{R}^n$ than bi-Lipschitz maps. Regular maps were introduced by David in [37], in connection with his studies of singular integral operators on surfaces, and they can be thought of as multivalued analogs of bi-Lipschitz maps. For the definitions and further studies, see [39] and [86].

13

Existence of Doubling Measures

If μ is a doubling measure on a metric space X, then it is easy to see that X is doubling as defined in Section 10.13. On the other hand, not every doubling space carries a doubling measure.

Exercise 13.1. Let μ be a doubling measure on a uniformly perfect space X. Prove that there are constants $C \geq 1$ and $\alpha > 0$, depending only on the doubling constant and on the uniform perfectness constant, such that

$$\frac{\mu(B(x,r))}{\mu(B(x,R))} \leq C\left(\frac{r}{R}\right)^{\alpha} \tag{13.2}$$

for all $x \in X$ and $0 < r \leq R < \operatorname{diam} X$. Conclude that the Hausdorff dimension of X is bounded from below by a positive constant that depends only on the constants associated with μ and the uniform perfectness.

It follows from the exercise that no uniformly perfect doubling space of Hausdorff dimension zero can support a doubling measure. For example, the set of all rational points on $\mathbb{R}$ with the induced metric is a doubling space but does not carry a doubling measure.

However, we have the following theorem.

Theorem 13.3. *A complete doubling space carries a doubling measure.*

There is a stronger statement, which we describe next. A measure μ on a metric space X is said to be *α-homogeneous* if there are constants $\alpha > 0$ and $C \geq 1$ such that

$$\frac{\mu(B(x,r))}{\mu(B(x,R))} \geq C^{-1}\left(\frac{r}{R}\right)^{\alpha} \tag{13.4}$$

for all $x \in X$ and $0 < r \leq R$. (Assume here that balls are open.) Clearly, α-homogeneous measures are doubling, and it follows from formula (4.16) that every doubling measure is α-homogeneous for some $\alpha > 0$.

Theorem 13.5. *A complete doubling space X carries an α-homogeneous measure for each α larger than the Assouad dimension of X.*

Recall the definition of the Assouad dimension from Definition 10.15.

Theorem 13.5 for compact spaces is due to Vol'berg and Konyagin [196]; the extension to the general complete case was provided by Luukkainen and Saksman [127]. The statement is quantitative in that the constants in condition (13.4) depend only on the constant in the doubling condition and on α.

It is not true that every locally compact doubling space admits a doubling measure, nor is it true that one can choose α to be exactly the Assouad dimension of X in general; see Remark 13.20 (d) and (e) to follow. Note that if X carries an α-homogeneous measure, then the Assouad dimension of X is at most α.

Proof of Theorem 13.5. For simplicity, we prove only the case when X is compact, and refer to [127] for the general case. The idea of the proof is familiar: one makes a sequence of discrete approximations of the space, distributes mass appropriately between the points in each approximation, and passes to a limit measure. The difficulty lies in distributing the mass appropriately at each stage so that the limiting measure has the desired properties; the choices must be made with care.

We assume without loss of generality that $\operatorname{diam} X = 1$. Let α_0 be the Assouad dimension of X, and fix $\alpha > \beta > \alpha_0$. Then the cardinality of each ϵr-net in a ball of radius r in X is at most $C_0\epsilon^{-\beta}$ for all $0 < \epsilon \leq 1/2$, for some $C_0 \geq 1$. Henceforth, *data* will refer to the triple (C_0, β, α). We fix $A \geq 10$ large enough, but depending only on the data, so that

$$C_0 A^{\beta} \leq A^{\alpha}. \tag{13.6}$$

Then, fix for each $k = 0, 1, \ldots$ a maximal A^{-k}-net S_k in X such that $S_0 \subset S_1 \subset S_2 \subset \ldots$. It follows from the choice of A in formula (13.6) that

$$\# S_{k+1} \cap B(x, A^{-k}) \leq A^{\alpha} \tag{13.7}$$

for each ball $B(x, A^{-k})$ in X of radius A^{-k}.

Next, for each k, fix partitions $T_{k,j}$, $j = 1, \ldots, n_k$, of the net S_{k+1} such that

$$S_{k+1} \cap B(x_{k,j}, A^{-k}/2) \subset T_{k,j} \subset S_{k+1} \cap B(x_{k,j}, A^{-k}) \tag{13.8}$$

for each point $x_{k,j}$ in the net S_k. (It is easy to see that such partitions exist.) Thus, the set $T_{k,j}$ is associated with the net point $x_{k,j}$. The points in $T_{k,j}$ can be thought of as the *children* of $x_{k,j}$. It is also clear what is meant by an *ancestor* and a *descendant* of a net point, with the understanding that each net point is both an ancestor and a descendant of itself. It follows from the second inclusion in formula (13.8) that no descendant $x_{m,*}$ lies further than distance $A^{-k+1}(A-1)^{-1}$ away from its ancestor $x_{k,*}$; that is,

$$|x_{m,*} - x_{k,*}| < \frac{A^{-k+1}}{A-1} \leq \frac{1}{9} A^{-k+1} \tag{13.9}$$

if $x_{m,*}$ is a descendant of $x_{k,*}$. Similarly, it follows from the first inclusion in formula (13.8), and from formula (13.9), that every net point in S_m for $m > k$ that lies within distance $A^{-k}/3$ from a net point $x_{k,*}$ is a descendant of $x_{k,*}$. Indeed, if $|x_{m,*} - x_{k,*}| < \frac{1}{3}A^{-k}$ and $x_{m,*} \neq x_{k,*}$, then $|x_{m,*} - x_{k+1,*}| < \frac{1}{9}A^{-k}$ for an ancestor $x_{k+1,*}$, which implies by way of the triangle inequality that $|x_{k,*} - x_{k+1,*}| < \frac{1}{2}A^{-k}$.

We make one more set of choices before describing the atomic measures on the nets S_k. Let $w_{k,j}$ be positive numbers, or *weights*, associated with points $x_{k,j}$ such that the following three conditions are satisfied:

$$A^{-\alpha} \leq w_{k,j} \leq 1, \tag{13.10}$$

$$w_{k,j} \equiv w_k \qquad \text{on } T_{k-1,j} \setminus \{x_{k-1,j}\}, \tag{13.11}$$

$$\sum_{T_{k,j}} w_{k+1,i} = 1, \tag{13.12}$$

where the sum is taken over all weights associated with the points $x_{k+1,i} \in T_{k,j}$. (By formulas (13.7) and (13.8) such weights can be found.) One can think of the weight $w_{k+1,i}$, associated with a point $x_{k+1,i} \in T_{k,j}$, as the amount that the corresponding child $x_{k+1,i}$ inherits from its parent $x_{k,j}$. The children of a fixed parent $x_{k,j}$ (excluding $x_{k,j}$ itself) all get the same amount, namely w_k. The parents keep something for themselves, but the total sum that remains in the family is always one.

Now, we define measures μ_k on the nets S_k as follows. There is only one point $x_{0,1}$ in S_0, and we let μ_0 be the unit point mass at that point. The measure μ_1 is defined by distributing the mass according to the weights $w_{1,j}$:

$$\mu_1(\{x_{1,j}\}) = w_{1,j}, \quad j = 1, \dots, n_1.$$

In general, if μ_k has been defined, μ_{k+1} is determined by the weights $w_{k+1,i}$ on each point $x_{k+1,i}$:

$$\mu_{k+1}(\{x_{k+1,i}\}) = w_{k+1,i}\,\mu_k(\{x_{k,j}\}) \geq A^{-\alpha}\mu_k(\{x_{k,j}\}) \tag{13.13}$$

for each $x_{k+1,i} \in T_{k,j}$. Obviously, $\mu_k(X) = 1$ for all $k \geq 0$.

We claim that each weak* limit μ of the sequence (μ_k) satisfies condition (13.4). To this end, we need the following lemma.

Lemma 13.14. *If* $|x_{k,j} - x_{k,i}| < \frac{2}{9}A^{-k+3}$, *then*

$$\mu_k(\{x_{k,i}\}) \leq A^{3\alpha}\mu_k(\{x_{k,j}\}) \tag{13.15}$$

for all $k \geq 1$.

We shall prove Lemma 13.14 later. At this juncture, we assume that inequality (13.15) holds and proceed with the proof of Theorem 13.5.

Fix a point $x \in X$, and fix a radius $r > 0$ and a real number $\lambda > 1$. We want to show that

$$\frac{\mu(B(x,r))}{\mu(B(x,\lambda r))} \geq C^{-1}\lambda^{-\alpha}, \tag{13.16}$$

where $C \geq 1$ depends only on the data, and where μ is a fixed weak* limit of the sequence (μ_k). To accomplish this, let $n < N$ be integers such that

$$\lambda r \leq A^{-n} < A\lambda r, \quad r/A \leq A^{-N} < r. \tag{13.17}$$

(Without loss of generality, we may assume that $1 \leq n$.) Then,

$$\mu(B(x,\lambda r)) \leq \mu_n(B(x, \lambda r + C_1 A^{-n})) \tag{13.18}$$

and

$$\mu_{N+1}(\{x_{N+1,p}\}) \leq \mu(B(x,r)), \tag{13.19}$$

where $x_{N+1,p}$ is such that $|x - x_{N+1,p}| < A^{-N-1}$. Indeed, inequality (13.18) follows from the choices and from formula (13.9); no mass distribution after μ_n from points outside $B(x, \lambda r + C_1 A^{-n})$ will trickle down to $B(x, \lambda r)$, for an appropriate constant $C_1 > 0$ that depends only on the data. Similarly, inequality (13.19) follows because all the descendants of $x_{N+1,p}$ remain in $B(x,r)$.

Because X is a doubling space, there are at most C_2 elements in $S_n \cap B(x, \lambda r + C_1 A^{-n})$ by the choice of n, where C_2 only depends on the data. Thus, by the mass-transfer inequality (13.13), and by inequality (13.15), we have that

$$\begin{aligned}\mu_{N+1}(\{x_{N+1,p}\}) &\geq A^{\alpha(n-N-1)}\mu_n(\{x_{n,i}\}) \\ &\geq A^{\alpha(n-N-1)}A^{-3\alpha}C_2^{-1}\mu_n(B(x,\lambda r + C_1 A^{-n})),\end{aligned}$$

where $x_{n,i}$ is the ancestor of $x_{N+1,p}$ in S_n. (Note that $x_{n,i}$ belongs to $B(x, \lambda r + C_1 A^{-n})$ for an appropriate C_1.) By combining this with inequalities (13.18) and (13.19), we arrive at

$$\mu(B(x,\lambda r)) \leq C_3 A^{\alpha(N-n)}\mu(B(x,r)).$$

Because $A^{\alpha(N-n)} < A^{2\alpha}\lambda^{\alpha}$, we obtain inequality (13.16).

Modulo Lemma 13.14, this completes the proof of Theorem 13.5. □

Proof of Lemma 13.14. If $k = 1$, the claim follows from the mass-transfer inequality (13.13), so assume that $k \geq 2$. Fix $x_{k,j}$ and $x_{k,i}$, and let $x_{k_0,*}$ be their first common ancestor; this means that if $x_{s,*}$ is a different common ancestor, then $s < k_0$. If $k_0 \geq k-3$, then inequality (13.15) again follows from the mass-transfer inequality. Thus, assume that $k_0 \leq k-4$.

Let $x_{k_0,*} = a_{k_0}, a_{k_0+1}, \ldots, a_{k-1}, a_k = x_{k,j}$ be the direct ancestral line from $x_{k_0,*}$ to $x_{k,j}$, and similarly for $x_{k_0,*} = b_{k_0}, b_{k_0+1}, \ldots, b_{k-1}, b_k = x_{k,i}$. There cannot be repetition among the points $a_{k_0+1}, \ldots, a_{k-2}$, for otherwise, if $a_m = a_{m-1}$, the point $x_{k,j}$ is stuck in the ball $B(a_m, \frac{1}{9}A^{-m+1})$ by estimate (13.9), while $x_{k,i}$ must lie outside the ball $B(a_{m-1}, \frac{1}{3}A^{-m+1})$ (see the discussion after estimate (13.9)); this will put the mutual distance between $x_{k,j}$ and $x_{k,i}$ larger than $\frac{2}{9}A^{-m+1} \geq \frac{2}{9}A^{-k+3}$, which contradicts the assumption. The assertion (13.15) now easily follows from the mass transfer rules (13.10), (13.11), and (13.13) (all new net points in S_k inherit the same amount w_k). We thereby conclude the proof of the lemma. □

Remark 13.20. (a) The proof of Theorem 13.5 as presented in this book is essentially due to Wu [203], but the idea of an α-homogeneous measure is taken from [196]. Wu further showed in [203], by elaborating the preceding argument, that given a compact doubling metric space X and a number $\alpha > 0$, there exists a doubling measure on X that gives full measure to a set of Hausdorff dimension less than α. This result is nontrivial already when $X = [0, 1]$, in which case it follows from an earlier work of Tukia [175].

(b) Quasisymmetric maps and doubling measures on the real line are in one-to-one correspondence in the following sense: if $f : \mathbb{R} \to \mathbb{R}$ is quasisymmetric, then the pull back measure $\mu_f(E) = |f(E)|$ is doubling, and conversely every doubling measure μ on $\mathbb{R}$ defines a quasisymmetric map by the formula

$$f(x) = \int_0^x d\mu. \tag{13.21}$$

(Here, $|\cdot|$ denotes Lebesgue measure.) Thus, the result of Tukia mentioned in (a) implies not only that quasisymmetric maps need not be absolutely continuous but that they can map an arbitrarily small (in the sense of Hausdorff dimension) set to a set of full measure; compare Corollary 11.10.

(c) The fact that quasisymmetric homeomorphisms $\mathbb{R} \to \mathbb{R}$ need not be absolutely continuous was first demonstrated by Beurling and Ahlfors [10]. They wrote an explicit expression for a singular quasisymmetric map. It was later realized that singular quasisymmetric maps are commonplace in complex analysis. For example, if X and Y are two conformally inequivalent compact Riemann surfaces of genus at least two, then the lift to the upper half-plane (the universal cover) of each diffeomorphism between the two surfaces has singular quasisymmetric boundary values. If Γ is a fractal Jordan curve on the Riemann sphere, as often arises in complex dynamics, or in the theory of Kleinian groups, then the map $f_+^{-1} \circ f_-$ of the upper half-plane $\mathbb{H}$ onto itself has singular quasisymmetric boundary values, where f_+ and f_- are two conformal maps from $\mathbb{H}$ onto the two complementary components of Γ. Moreover, in this case the maps are strongly singular in the sense that they take a set of small Hausdorff dimension on $\mathbb{R}$ to a set of full measure, akin to Tukia's theorem. See [120] and [150].

(d) In light of the discussion in (b), it is not difficult to show that there are open subsets of $\mathbb{R}$ that do not carry doubling measures. Indeed, it is not difficult to show that there exist compact, totally disconnected sets K such that $|f(K)| > 0$ for all quasisymmetric maps $f : \mathbb{R} \to \mathbb{R}$. Then, $U = \mathbb{R} \setminus K$ cannot carry a doubling measure, for if it did, the formula (13.21) would provide a quasisymmetric map f

with $|f(K)| = 0$. Saksman [157] has exhibited connected open sets in $\mathbb{R}^n$, $n \geq 2$, that do not carry doubling measures.

Compact sets $K \subset \mathbb{R}$ with the property that $|f(K)| > 0$ for all quasisymmetric maps $f : \mathbb{R} \to \mathbb{R}$ were termed *quasisymmetrically thick* by Staples and Ward [169]; they also gave nontrivial examples of quasisymmetrically thick sets. Buckley, Hanson, and MacManus [23] later characterized the quasisymmetrically thick regular Cantor sets. (See also [201].) The complete picture is still missing.

In Section 14.37, we extend the notion of a quasisymmetrically thick set to general (uniformly perfect) metric spaces.

(e) In [196, Theorem 4], Vol'berg and Konyagin give an example of a compact set K in $\mathbb{R}^n$, $n \geq 1$, that has Assouad dimension α_0, but no doubling measure on K satisfies condition (13.4) with $\alpha = \alpha_0$. Here, α_0 can be chosen to be any positive number less than n.

If X is an (Ahlfors) n-regular metric space, then it is easy to see that the Assouad dimension of X is n and that the Hausdorff measure in X is n-homogeneous. It seems that not much is known about the question of which spaces carry homogeneous measures of dimension equal to the Assouad dimension.

(f) Doubling measures also arise in the theory of elliptic partial differential equations, where the Dirichlet problem (in the upper half space in $\mathbb{R}^n$, say) can be solved by integrating the boundary data against a "harmonic measure" associated with the equation at hand. Quasisymmetric maps were first used to show that such measures can be singular with respect to Lebesgue measure. See [48], [81], [201], [202], and the references therein.

The following exercise was suggested by Jeremy Tyson.

Exercise 13.22. Use Proposition 10.11 and Theorem 13.5 to prove the following analog of Corollary 11.10 for the Assouad dimension: if $f : X \to Y$ is a power quasisymmetric embedding (see the discussion following Corollary 11.5) with control function $\eta(t) = C \max\{t^\alpha, t^{1/\alpha}\}$ for some $C > 0$ and $0 < \alpha \leq 1$, then

$$\alpha \dim_A E \leq \dim_A f(E) \leq \frac{1}{\alpha} \dim_A E$$

for all $E \subset X$.

13.23 Notes to Chapter 13. This chapter is mostly based on the papers [196] and [203] by Vol'berg and Konyagin and Wu. For further interesting constructions of doubling measures, see [102] and [151]. Remark 13.20 contains other relevant references.

14

Doubling Measures and Quasisymmetric Maps

In this chapter, we further explore the close connection between doubling measures and quasisymmetric maps. The main theme is that doubling measures can be used to deform the underlying metric space and that such deformations are quasisymmetric. This point of view has been advocated by David and Semmes; see [39], [164], and [165].

14.1 Quasimetric spaces. For the ensuing discussion, it is convenient to introduce the concept of a quasimetric space. A *quasimetric* on a set X is a function $q : X \times X \to [0, \infty)$ that is symmetric[1], vanishes if and only if $x = y$, and satisfies, for some $K \geq 1$,

$$q(x, y) \leq K\,(q(x, z) + q(z, y)) \tag{14.2}$$

for all $x, y, z \in X$. Then, a *quasimetric space* is a set X together with a quasimetric q. Quasimetric spaces are not quite as well behaved as metric spaces, but many concepts make sense in this context, too. For example, balls can be defined as usual: if q is a quasimetric on a set X, then for $x \in X$ and $r > 0$, a *quasimetric ball* centered at x with radius r is the set

$$B(x, r) = \{y \in X : q(x, y) < r\}. \tag{14.3}$$

Note that X has no fixed topology, and even if it does, we do not require balls as defined in Eq. (14.3) to be open. (However, see Proposition 14.5 to follow and the remark after it.)

---

[1]Some authors do not require that a quasimetric be symmetric but rather that $q(x, y) \leq Cq(y, x)$ for some constant $C \geq 1$ independent of x, y.

If μ is a measure on a quasimetric space X, we call μ *doubling* if

$$\mu(B(x, 2r)) \leq C\mu(B(x, r))$$

for all $x \in X$ and $r > 0$, and for some constant $C \geq 1$ independent of x and r. As in Convention 1.4, we understand that μ is an outer measure defined on all subsets of X. We also assume that μ is Borel regular in the metric topology given by Proposition 14.5.

One can speak about quasisymmetric maps between quasimetric spaces. We call an injection $f : (X, q) \to (Y, p)$ from one quasimetric space into another η-*quasisymmetric*, where $\eta : [0, \infty) \to [0, \infty)$ is a homeomorphism, if

$$\frac{p(f(x), f(a))}{p(f(x), f(b))} \leq \eta\left(\frac{q(x, a)}{q(x, b)}\right) \tag{14.4}$$

for all triples x, a, b of distinct points in X. Because there is no preassigned topology given on X or Y, we cannot stipulate any topological conditions on f.

To give an example of a quasimetric that is not a metric, consider any metric space (X, d), and define $d_\epsilon(x, y) = d(x, y)^\epsilon$. Then, d_ϵ is a quasimetric for all $\epsilon > 0$, but rarely a metric if $\epsilon > 1$. Note that the identity map from (X, d) to (X, d_ϵ) is quasisymmetric.

A useful fact about quasimetric spaces is given in the following proposition.

Proposition 14.5. *Let q be a quasimetric on a set X. Then, there is $\epsilon_0 > 0$, depending only on the constant K in formula (14.2), such that $q_\epsilon(x, y) = q(x, y)^\epsilon$ is bi-Lipschitz equivalent to a metric for each $0 < \epsilon \leq \epsilon_0$. That is, for each $0 < \epsilon \leq \epsilon_0$ there is a metric d_ϵ on X and a constant $C = C(\epsilon, K) \geq 1$ such that*

$$C^{-1} q_\epsilon(x, y) \leq d_\epsilon(x, y) \leq C q_\epsilon(x, y) \tag{14.6}$$

for all $x, y \in X$.

Note that the metrics d_ϵ in Proposition 14.5 determine a topology on X that is independent of ϵ. If necessary, we can always assume that each quasimetric space is equipped with this (canonically defined) topology.

Proof of Proposition 14.5. If q is a quasimetric on X, then

$$q(x, y) \leq K \max\{q(x, z), q(z, y)\} \tag{14.7}$$

for all $x, y, z \in X$, and for some $K \geq 1$ independent of the points. Thus, q is a "quasiultrametric"[2]. What we will show in fact is that each small enough power of a quasiultrametric is bi-Lipschitz equivalent to a metric.

[2] An *ultrametric* is a metric d that satisfies the following strong form of the triangle inequality: $d(x, y) \leq \max\{d(x, z), d(z, y)\}$.

To that end, fix $\epsilon > 0$ and define, for $x, y \in X$,

$$d_\epsilon(x, y) = \inf \sum_{i=0}^{k} q_\epsilon(x_i, x_{i+1}), \tag{14.8}$$

where the infimum is taken over all finite sequences $x = x_0, x_1, \ldots, x_{k+1} = y$ of points in X. Clearly, d_ϵ is symmetric, satisfies the triangle inequality, and does not exceed q_ϵ. We claim that

$$q_\epsilon(x, y) \leq K^{2\epsilon} d_\epsilon(x, y), \tag{14.9}$$

provided $K^{2\epsilon} \leq 2$. It then follows that d_ϵ is a metric, bi-Lipschitz equivalent to q_ϵ. To prove inequality (14.9), it suffices to show that

$$q_\epsilon(x, y) \leq K^{2\epsilon} \sum_{i=0}^{k} q_\epsilon(x_i, x_{i+1}) \tag{14.10}$$

for all sequences as earlier. We prove this by induction on k. Because

$$q_\epsilon(x, y) \leq K^\epsilon \max\{q_\epsilon(x, z), q_\epsilon(z, y)\} \leq K^{2\epsilon} \max\{q_\epsilon(x, z), q_\epsilon(z, y)\},$$

formula (14.10) holds for $k = 1$. Next, assume that formula (14.10) holds for some $k \geq 1$, and let

$$x = x_0, x_1, \ldots, x_{k+1}, x_{k+2} = y$$

be a sequence of points in X. Let

$$R = \sum_{i=0}^{k+1} q_\epsilon(x_i, x_{i+1}),$$

and let x_p, $0 \leq p \leq k + 1$, be the last point in the (ordered) sequence such that

$$\sum_{i=0}^{p-1} q_\epsilon(x_i, x_{i+1}) \leq R/2.$$

(To be rigorous here and in the following, we define $x_{-1} = x_0 = x$ and $x_{k+3} = x_{k+2} = y$.) Then,

$$\sum_{i=p+1}^{k+2} q_\epsilon(x_i, x_{i+1}) \leq R/2,$$

and hence

$$\begin{aligned} q_\epsilon(x, y) &\leq K^\epsilon \max\{q_\epsilon(x, x_p), q_\epsilon(x_p, y)\} \\ &\leq K^\epsilon \max\{q_\epsilon(x, x_p), K^\epsilon q_\epsilon(x_p, x_{p+1}), K^\epsilon q_\epsilon(x_{p+1}, y)\} \\ &\leq K^\epsilon \max\{K^{2\epsilon} R/2, K^\epsilon R, K^{2\epsilon} R/2\} \leq K^{2\epsilon} R, \end{aligned}$$

provided $K^{2\epsilon} \leq 2$. The proof of the proposition is complete. □

It follows from the preceding proof that if q is a quasimetric on X with constant $K \geq 1$ in formula (14.2), then $q_\epsilon(x, y) = q(x, y)^\epsilon$ is bi-Lipschitz equivalent to a metric for all $\epsilon > 0$ such that $(2K)^{2\epsilon} \leq 2$. Moreover, one can choose $C = (2K)^{2\epsilon}$ for the bi-Lipschitz constant in formula (14.6).

Note that the identity map from the quasimetric space (X, q) onto the metric space (X, d_ϵ) is quasisymmetric as defined in formula (14.4).

14.11 Quasimetrics determined by doubling measures. Each doubling measure μ on a metric space (X, d)[3] determines a family of quasimetrics on X by the formula

$$q_{\mu,\epsilon}(x, y) = \mu(B_{x,y})^\epsilon, \quad \epsilon > 0, \tag{14.12}$$

where

$$B_{x,y} = B(x, d(x, y)) \cup B(y, d(x, y)). \tag{14.13}$$

For definiteness, we assume that the balls in Eq. (14.13) are closed. The quasimetric constant for $q_{\mu,\epsilon}$ depends only on the doubling constant of μ and on ϵ. It follows from Proposition 14.5 that there exists a constant $\epsilon_0 = \epsilon_0(\mu) > 0$ such that $q_{\mu,\epsilon}$ is bi-Lipschitz equivalent to a metric $d_{\mu,\epsilon}$ for each $\epsilon \leq \epsilon_0$.

The next proposition easily follows from the definitions and from Exercise 13.1. Recall the definition for Ahlfors regular spaces from formula (8.10).

Proposition 14.14. *Let (X, d) be a uniformly perfect metric space and let μ be a doubling measure on X. Then, the metric measure spaces $(X, d_{\mu,\epsilon}, \mu)$ for $0 < \epsilon \leq \epsilon_0(\mu)$ are Ahlfors regular of dimension $1/\epsilon$ and the identity map $(X, d) \to (X, d_{\mu,\epsilon})$ is quasisymmetric.*

Because Ahlfors regular spaces are uniformly perfect, because uniform perfectness is a quasisymmetric invariant (Exercise 11.2), and because doubling measures can be pushed forward to doubling measures under quasisymmetric maps (this is easy to see), we have the following characterization of metric spaces that are quasisymmetrically equivalent to Ahlfors regular spaces.

Corollary 14.15. *A metric space is quasisymmetrically equivalent to an Ahlfors regular space if and only if it is uniformly perfect and carries a doubling measure.*

In particular, because complete doubling spaces carry doubling measures (Theorem 13.3), we find that every uniformly perfect complete doubling space is quasisymmetrically equivalent to an Ahlfors regular space. Assouad's theorem (Theorem 12.2) further implies that one can realize the target as a subspace of some Euclidean space. It is of interest to estimate the smallest, or infimal, regularity dimension of the image. This amounts to estimating the number $\epsilon_0(\mu)$. Although

[3] In this chapter, we often depart from the convention that metrics are generically denoted by $|x - y|$.

an explicit bound can be derived from the proof of Proposition 14.6 (see the remark after the proof), it is hardly optimal. On the other hand, if μ is α-homogeneous as defined in condition (13.4), then one easily checks, by using definition (14.8), that

$$q_{\mu,\alpha}(x, y) = \mu(B_{x,y})^{1/\alpha}$$

is bi-Lipschitz equivalent to a metric. Because every complete doubling space admits doubling measures that are homogeneous of each (preassigned) exponent larger than the Assouad dimension of the space (Theorem 13.5), we can record the following theorem.

Theorem 14.16. *Let X be a complete, uniformly perfect metric space of finite Assouad dimension α_0. Then, for each $\alpha > \alpha_0$ there exists a quasisymmetric homeomorphism of X onto a closed Ahlfors α-regular subset of some $\mathbb{R}^N$.*

Corollary 14.17. *Let X be a closed, uniformly perfect subset of $\mathbb{R}^n$, $n \geq 1$, and let $\alpha > n$. Then, there exists $N \geq 1$ and a quasisymmetric embedding $f : X \to \mathbb{R}^N$ such that $f(X)$ is Ahlfors α-regular.*

The statement in Theorem 14.16 is quantitative in the sense that both the quasisymmetry function η and the dimension N depend only on α_0 and α. Corollary 14.17 exhibits a tremendous flexibility in quasisymmetric maps: sets with very mild metric and topological properties in Euclidean space (all nondegenerate continua, for example) can be mapped quasisymmetrically onto sets with strong measure-theoretic symmetry.

14.18 Metric doubling measures. The deformation of a metric space by a doubling measure, as described earlier, is especially interesting if the underlying metric space (X, d) itself is Ahlfors n-regular for some real number $n > 0$. In this case, a doubling measure μ is said to be a *metric doubling measure* if $q_{\mu,1/n}$ is bi-Lipschitz equivalent to a metric; that is, μ is a metric doubling measure if there is a constant $C \geq 1$ and a metric $d_{\mu,1/n} = d_\mu$ on X such that

$$C^{-1} d_\mu(x, y) \leq \mu(B_{x,y})^{1/n} \leq C d_\mu(x, y) \tag{14.19}$$

for all $x, y \in X$. Note that X is necessarily uniformly perfect if it is n-regular for some $n > 0$. It is easy to see that a doubling measure μ is a metric doubling measure if and only if

$$\inf \sum_{i=0}^{k} \mu\big(B_{x_i,x_{i+1}}\big)^{1/n} > 0, \tag{14.20}$$

whenever $x \neq y$, and the infimum is taken over all sequences $x = x_0, x_1, \ldots, x_k$, $x_{k+1} = y$ in X. (Compare the proof of Proposition 14.5.)

If $f : X \to Y$ is a quasisymmetric homeomorphism from a metric space X onto an n-regular metric space Y, then the pullback measure

$$\mu_f(E) = \mathcal{H}_n(f(E)) \tag{14.21}$$

for $E \subset X$ is a metric doubling measure on X. Indeed, it is easy to see that μ_f is doubling (note that Y, and hence X, are uniformly perfect) and that the metric

$$d_f(a,b) = d_Y(f(a), f(b)) \tag{14.22}$$

satisfies

$$C^{-1} d_f(a,b) \le \mu_f(B_{a,b})^{1/n} \le C d_f(a,f), \tag{14.23}$$

where $C \ge 1$ depends only on the data associated with f and Y. It is an interesting problem to determine which metric doubling measures μ on an n-regular space X arise in this manner from self-maps of X. This problem is open even when $X = \mathbb{R}^n$, $n \ge 2$. In $\mathbb{R}$, all doubling measures are metric doubling measures, and as explained in Remark 13.20 (b), we have a one-to-one correspondence between doubling measures and (normalized) quasisymmetric maps. In higher dimensions, $n \ge 2$, there are doubling measures in $\mathbb{R}^n$ that cannot be metric doubling measures; for instance, the doubling measure $d\mu(x) = |x_1|dx$ in $\mathbb{R}^2 = \{x = (x_1, x_2)\}$ cannot be a metric doubling measure because of formula (14.20). However, one could have hoped that in $\mathbb{R}^n$, $n \ge 2$, metric doubling measures would correspond to quasisymmetric maps via the pullback formula (14.21). This hope, however, turned out to be false at least in dimensions $n \ge 3$. We describe a counterexample, due to Semmes, in Example 14.32.

14.24 Uniformly disconnected spaces. A metric space is said to be *uniformly disconnected* if there is $\epsilon_0 > 0$ so that no pair of distinct points $x, y \in X$ can be connected by an ϵ_0-chain, where an *ϵ_0-chain connecting x and y* is a sequence of points $x = x_0, x_1, \ldots, x_m, x_{m+1} = y$ in the space satisfying $|x_i - x_{i+1}| \le \epsilon_0 |x - y|$ for all $i = 0, \ldots, m$.

For example, the Cantor ternary set is uniformly disconnected, whereas the sequence $\{\frac{1}{n} : n = 1, 2, \ldots\} \subset \mathbb{R}$ is not.

Exercise 14.25. Show that uniform disconnectedness is a quasisymmetrically invariant property of metric spaces quantitatively.

Uniformly disconnected spaces were defined in [39] in a different but equivalent way, as follows.

Exercise 14.26. Prove that a metric space X is uniformly disconnected if and only if for each $x \in X$ and $r > 0$ there is a subset A of X, containing the ball $B(x, r/C)$ and contained in the ball $B(x, r)$, such that $\operatorname{dist}(A, X \setminus A) \ge r/C$, where $C \ge 1$ is independent of x and r. The statement is quantitative in the usual sense.

Exercise 14.27. (See Proposition 15.7 of [39].) A metric space (X, d) is uniformly disconnected if and only if there is an ultrametric d' on X that is bi-Lipschitz equivalent to d. (Hint: For distinct points $x, y \in X$, let $d'(x, y)$ be the infimum of those values γ for which there exists a $d(x, y)^{-1}\gamma$-chain connecting x and y.)

We can use uniformly disconnected sets to construct interesting metric doubling measures.

Lemma 14.28. *If A is a closed, uniformly disconnected set in a connected Ahlfors n-regular metric space X, then the measure $d\mu(x) = \min\{dist(x, A)^a, 1\}$ $d\mathcal{H}_n(x)$ is a metric doubling measure in X for each $a > 0$.*

Proof. We employ characterization (14.20) for metric doubling measures. Fix two distinct points x and y in X, and let $x_0 = x, x_1, \ldots, x_k, x_{k+1} = y$ be points in X. First, consider the case where

$$|x_i - x_{i+1}| < \frac{\epsilon_0}{1000}|x - y| \tag{14.29}$$

for all $i = 0, \ldots, k$, where $\epsilon_0 > 0$ is a constant that works for the uniform disconnectedness condition; we may assume that $\epsilon_0 \leq \frac{1}{4}$. In the ordered sequence $x_0, \ldots, x_{k+1}$, let x_{i_0} be the last point inside the ball $B(x, \frac{1}{4}|x - y|)$, and let $x_{i_{p+1}}$ be the first point outside the ball $B(x, \frac{1}{2}|x - y|)$. Thus, the points in between, $x_{i_1}, \ldots, x_{i_p}$, all lie at least distance $\frac{1}{4}|x - y|$ from both points x and y. Because of formula (14.29), and because $\epsilon_0 \leq \frac{1}{4}$, we must have that $p \geq 3$.

Next, consider balls $B_j = B(x_{i_j}, \epsilon_0|x-y|/100)$, for $j = 1, \ldots, p$. First, assume that for each $j = 1, \ldots, p$ there is a point $a_j \in B_j \cap A$. Then, we have by the triangle inequality that

$$|a_1 - a_p| \geq \left|x_{i_1} - x_{i_p}\right| - \left|a_1 - x_{i_1}\right| - \left|a_p - x_{i_p}\right| > \frac{1}{10}|x - y|,$$

so

$$|a_j - a_{j+1}| \leq \frac{\epsilon_0}{50}|x - y| + \frac{\epsilon_0}{1000}|x - y| < \frac{\epsilon_0}{10}|x - y| < \epsilon_0|a_1 - a_p|,$$

which contradicts the uniform disconnectedness of A.

It follows that there is a ball $B_s = B(x_s, \epsilon_0|x - y|/100)$ that does not meet A for some x_s that lies in $B(x, |x - y|/2) \setminus B(x, |x - y|/4)$ and is part of the chain from x to y as fixed earlier. Let $x_s = x_{s_1}, x_{s_2}, \ldots, x_{s_t}$ be the points from our (ordered) sequence such that each x_{s_i} lies in $\frac{1}{2}B_s$ but the successor $x_{s_{t+1}}$ of x_{s_t} lies outside $\frac{1}{2}B_s$. Then, $t \geq 2$ and

$$B_{x_{s_i}, x_{s_{i+1}}} \subset \frac{3}{4}B_s$$

for each $i = 1, \ldots, t$ by formula (14.29). Thus,

$$\begin{aligned} \sum_{i=1}^{t} \mu\left(B_{x_{s_i}, x_{s_{i+1}}}\right)^{1/n} &= \sum_{i=1}^{t} \left(\int_{B_{x_{s_i}, x_{s_{i+1}}}} \min\{\text{dist}(x, A)^a, 1\}\, d\mathcal{H}_n(x)\right)^{1/n} \\ &\geq C^{-1} \sum_{i=1}^{t} \min\{(\epsilon_0|x - y|/400)^a, 1\}^{1/n} \left|x_{s_i} - x_{s_{i+1}}\right| \\ &\geq C^{-1} \min\{(\epsilon_0|x - y|/400)^a, 1\}^{1/n} |x - y|, \end{aligned} \tag{14.30}$$

where the constant $C \geq 1$ depends only on the Ahlfors regularity constant, on ϵ_0, and on n. This shows that formula (14.20) holds, and the proof is complete under the assumption (14.29).

On the other hand, if the assumption (14.29) does not hold for some pair x_i, x_{i+1} from the chain, we can find a ball B' inside the set $B_{x_i,x_{i+1}}$ that does not meet A and has radius comparable to $|x-y|$ (see Exercise 14.31). The proof then can be completed as in Eq. (14.30) by using the Ahlfors regularity. □

Exercise 14.31. Prove that a uniformly disconnected set A in a connected Ahlfors regular space X is *porous*, quantitatively; that is, prove that there is a constant $c>0$ such that for each $x \in X$, and $0<r<\operatorname{diam} X$, the ball $B(x,r)$ contains a ball $B(y,cr)$ that does not meet A.

Example 14.32. Let us accept the following two facts: (1) there exists, in each $\mathbb{R}^n$, $n \geq 3$, a compact, uniformly disconnected set A whose complement $\mathbb{R}^n \setminus A$ is not simply connected; and (2) any compact set A in $\mathbb{R}^n$, $n \geq 3$, with $\mathbb{R}^n \setminus A$ not simply connected must have Hausdorff dimension at least $n-2$.

One can construct a set A as in (1) by appropriately modifying the classical constructions of Antoine and Blankenship (see [93], [156], and [162]). For the second fact, see [132] and Exercise 14.36.

Next, fix a set A in $\mathbb{R}^n$ as in (1), and consider the measure

$$d\mu(x) = w(x)dx, \quad w(x) = \min\{\operatorname{dist}(x,A)^a, 1\}, \; a > n(n-1). \tag{14.33}$$

Then, μ is a metric doubling measure by Lemma 14.28. Let $d_\mu(x,y)$ be a metric on $\mathbb{R}^n$ comparable to $\mu(B_{x,y})^{1/n}$ as in formula (14.19). One easily computes that the Hausdorff dimension of A in the metric space $(\mathbb{R}^n, d_\mu)$ is less than one. On the other hand, if, for a quasisymmetric map $f:\mathbb{R}^n \to \mathbb{R}^n$ (in the standard metrics), the measure μ were comparable to the pullback of Lebesgue n-measure,

$$C^{-1}\mu(E) \leq |f(E)| \leq C\mu(E), \tag{14.34}$$

then $f:(\mathbb{R}^n, d_\mu) \to \mathbb{R}^n$ would be a bi-Lipschitz map. In particular, the Hausdorff dimension of $f(A)$ would be less than $n-2$, contradicting fact (2) (see Exercise 14.36).

We conclude that there exist, for each $n \geq 3$, metric doubling measures in $\mathbb{R}^n$ that are not comparable to the pullback measure $E \mapsto |f(E)|$ under any quasisymmetric map $f:\mathbb{R}^n \to \mathbb{R}^n$. (As always, $|\cdot|$ denotes Lebesgue measure of appropriate dimension.)

In dimension $n=2$, the situation is unknown.

14.35 Open problem. Is every metric doubling measure in $\mathbb{R}^2$ comparable to the pull-back measure $E \mapsto |f(E)|$ under a quasisymmetric map $f:\mathbb{R}^2 \to \mathbb{R}^2$?[4]

Exercise 14.36. Prove that if A is a closed set in $\mathbb{R}^n$, $n \geq 3$, such that $\mathbb{R}^n \setminus A$ is not simply connected, then the Hausdorff dimension of A is at least $n-2$.

[4] Added in Proof: Tomi Laakso has recently answered this question in the negative [117].

In fact, show that $\mathcal{H}_{n-2}(A) > 0$. (Hint: Given a loop $\gamma : \mathbb{S}^1 \to \mathbb{R}^n \setminus A$, take its polygonal approximation, extend piecewise linearly to the disk $\hat{\gamma} : \mathbb{B}^2 \to \mathbb{R}^n$, and show that the image $\hat{\gamma}(\mathbb{B}^2)$ can be moved off of A by Fubini's theorem, if $\mathcal{H}_{n-2}(A) = 0$.)

14.37 Doubling measures and quasisymmetrically thick sets. Next, we generalize the concept of a quasisymmetrically thick set from the real line to more general metric spaces. (Compare Remark 13.20 (d).)

Let X be a uniformly perfect, complete and doubling metric space. By Theorem 13.3, X carries a doubling measure. We are interested in characterizing the dense subsets of X with the same property. Recall that a characterization in the case $X = \mathbb{R}$ was given in Remark 13.20 (d) in terms of quasisymmetric maps: $A \subset \mathbb{R}$ carries a doubling measure if and only if there exists a quasisymmetric map $f : \mathbb{R} \to \mathbb{R}$ such that $|f(\mathbb{R} \setminus A)| = 0$. In other words, a subset of the real line carries a doubling measure if and only if its complement is not quasisymmetrically thick.

Let us call a subset E of X *quasisymmetrically thick* if $\mathcal{H}_n(f(E)) > 0$ whenever $f : X \to Y$ is a quasisymmetric homeomorphism from X onto an Ahlfors n-regular space Y. Note that n is allowed to depend on Y, and it can be any positive real number.

Proposition 14.38. *The two notions of quasisymmetric thickness agree on* $\mathbb{R}$.

Proof. Let $E \subset \mathbb{R}$ be a set such that $|g(E)| > 0$ for all quasisymmetric maps $g : \mathbb{R} \to \mathbb{R}$, and let $f : \mathbb{R} \to Y$ be a quasisymmetric map onto some n-regular space Y. Thus, Y is a topological line with a metric such that the Hausdorff n-measure of each arc in Y of diameter d is comparable to d^n. (Note that this description of Ahlfors regularity via arcs is not precisely the one given in formula (8.10) but easily seen to be equivalent by the quasisymmetry of f.) Now, such a line Y can be parametrized by a homeomorphism $h : \mathbb{R} \to Y$ that satisfies

$$C^{-1}|x - y|^{1/n} \leq |h(x) - h(y)| \leq C|x - y|^{1/n} \qquad (14.39)$$

for each pair of points $x, y \in \mathbb{R}$, and for some $C \geq 1$ independent of the points. (See [66].) Then, h is quasisymmetric, so that $|h^{-1} \circ f(E)| > 0$ by assumption. On the other hand, it is easy to see from formula (14.39) that $\mathcal{H}_n(h(F)) > 0$ if and only if $|F| > 0$ for $F \subset \mathbb{R}$, from which $\mathcal{H}_n(f(E)) > 0$, as required. The proposition follows. □

Remark 14.40. Quasisymmetric maps $\mathbb{R}^n \to \mathbb{R}^n$, $n \geq 2$, are absolutely continuous in that they carry sets of n-measure zero to sets of n-measure zero. (See [182, Section 33].) Thus, a straightforward generalization of the definition for quasisymmetric thickness, using maps from $\mathbb{R}^n$ to itself, would be uninteresting.

The next proposition ties up quasisymmetric thickness with the existence of doubling measures.

Proposition 14.41. *Let X be a uniformly perfect, complete, and doubling metric space, and let A be a dense subset of X. Then, A carries a doubling measure if and only if $X \setminus A$ is not quasisymmetrically thick.*

Proof. Let A be a dense subset of X, and assume that A carries a doubling measure μ. Then, we can extend μ to be a doubling measure on X by setting $\mu(X \setminus A) = 0$. By Proposition 14.14, there is $\epsilon_0 > 0$ such that the metric space $(X, d_{\mu,\epsilon})$ is Ahlfors regular of dimension $1/\epsilon$ for all $0 < \epsilon \leq \epsilon_0$, and that the identity map $X \to (X, d_{\mu,\epsilon})$ is quasisymmetric. Moreover, μ is comparable to the $1/\epsilon$-dimensional Hausdorff measure in $(X, d_{\mu,\epsilon})$ (Exercise 8.11). Therefore, we have $\mathcal{H}_{1/\epsilon}(X \setminus A) = 0$ and hence $X \setminus A$ is not quasisymmetrically thick.

Conversely, assume that there exists a quasisymmetric map $f : X \to Y$ onto some n-regular space Y such that $\mathcal{H}_n(f(X \setminus A)) = 0$ for a dense subset A of X. Then, the pullback measure $\mu(E) = \mathcal{H}_n(f(E))$ is easily seen to be doubling (by the quasisymmetry of f), and clearly $\mu(X \setminus A) = 0$, so that $\mu|A$ is a doubling measure on A. The proposition follows. □

In light of the preceding discussion, the procedure in Section 14.11, where doubling measures on a space X generate quasisymmetric maps, can be thought of as "integrating" the measure μ; see Eq. (13.21). If X is Ahlfors regular, for instance $X = \mathbb{R}^n$, metric doubling measures are precisely those doubling measures for which this procedure produces a space with the same Ahlfors regularity dimension.

Saksman [157] has exhibited examples of open dense subsets of $\mathbb{R}^n$, $n \geq 2$, that do not carry doubling measures. These are complements of certain thick Cantor sets. These Cantor sets, in turn, are examples of quasisymmetrically thick sets by way of Proposition 14.41.

14.42 Notes to Section 14. The material on quasimetric spaces is standard; see [18, Chapters 5–10]. More properties of quasimetric spaces along these lines can be found in [131]. The idea of deforming metric spaces by doubling measures is due to David and Semmes; see [38] and [39]. Theorem 14.16 and its corollary seem to be new observations, but as the proof shows, they follow straightforwardly from a number of known theorems. David and Toro [40] have shown that one can snowflake hyperplanes in $\mathbb{R}^n$, inside $\mathbb{R}^n$, for each $n \geq 2$. See also [11]. See [88] for further open problems.

The material on uniformly disconnected spaces and metric doubling measures originates from the paper [162] by Semmes, as does Example 14.32. Exercise 14.36 is from [132, Lemma 3.3]; see also [128, Section 6]. The general definition for quasisymmetric thickness seems to be new. Recent papers on quasisymmetric thickness on the real line include [169] and [23]. It would be interesting to know to what extent the one-dimensional results have analogs in more general settings. Even in $\mathbb{R}^n$, not much is known about quasisymmetrically thick sets. (See, however, the examples of Saksman in [157].)

15

Conformal Gauges

In this final chapter, we introduce a convenient language to discuss quasisymmetrically invariant properties of metric spaces.

15.1 Conformal gauges. A *conformal gauge* on a set X is a maximal collection of metrics on X such that each pair of metrics from the collection is quasisymmetrically related by the identity map.[1] Note that each given conformal gauge fixes a topology on X. If X is a topological space to begin with, it is natural to consider only those gauges that yield the same topology.

Two conformal gauges are said to be *equivalent* if there is a quasisymmetric map between one (hence each) pair of metrics from the two collections. The terms *locally equivalent*, *embeddable*, and *locally embeddable* make sense in the context of gauges. If (X, d) is a fixed metric space, it determines a unique conformal gauge on X, and the problem of characterizing this gauge is the same as the problem of characterizing metric spaces quasisymmetrically equivalent to (X, d).

Every topologically invariant property is obviously a gauge property, so we can speak about connected gauges, compact gauges, and finite-dimensional gauges, for example. On the other hand, while a property such as completeness is not a topological invariant, it is a quasisymmetric invariant, so it is fair to ask, for example, whether a conformal gauge is complete. (If we adopt the self-explanatory term *topological gauge*, then a classical theorem of Urysohn can be rephrased by saying that a topological gauge is complete if and only if it is compact.) In addition, a conformal gauge can be doubling, uniformly perfect, uniformly disconnected,

[1] The terminology here was suggested by Dennis Sullivan.

or linearly locally connected. If a measure μ on X is doubling with respect to one metric from a conformal gauge, then it is doubling with respect to all the others; thus, we may say that μ is a *doubling measure on a conformal gauge*. Theorem 13.3 on the existence of doubling measures can equivalently be stated as follows.

Theorem 15.2. *A complete and doubling conformal gauge admits a doubling measure.*

In the context of Gromov hyperbolic spaces, conformal gauges arise in a natural manner. A *Gromov hyperbolic space* is a geodesic[2] metric space where geodesic triangles are *thin*. By definition, this means that there exists $\delta \geq 0$ so that for every collection of three geodesics $[x, y]$, $[y, z]$, and $[x, z]$ in the space, each one of the three is contained in the δ-neighborhood of the union of the two others. Without going into the details, there is canonically associated with every proper Gromov hyperbolic space X a compact conformal gauge $\mathcal{G}_{\partial X}$ on the "Gromov boundary" ∂X of X. In general, there is no preferred metric in this gauge, and even if there is, it may not be easy to locate. A major conjecture in three-dimensional topology states that if $X = \Gamma$ is a Gromov hyperbolic group that is also a fundamental group of a closed 3-manifold, and if the boundary gauge $\mathcal{G}_{\partial\Gamma}$ is a topological 2-sphere, then the gauge $\mathcal{G}_{\partial\Gamma}$ is equivalent to the standard conformal gauge determined by $\mathbb{S}^2$. (See [28].) For the construction of the boundary gauge $\mathcal{G}_{\partial X}$ and for more discussion on Gromov hyperbolic spaces, see [67]. The boundary gauges of Gromov hyperbolic spaces have been studied in [16] and [17].

Very few conformal gauges have been characterized. The author is familiar with the following two characterizations.

Theorem 15.3. *A conformal gauge on a topological circle is equivalent to that determined by the standard circle $\mathbb{S}^1$ if and only if the gauge is both doubling and of bounded turning. The same two conditions characterize the conformal gauge of the real line $\mathbb{R}$, up to equivalence.*

A metric space is said to be of *C-bounded turning*, $C \geq 1$, if every pair of points x, y in the space can be joined by a curve whose diameter does not exceed $C|x - y|$. This is obviously a quasisymmetrically invariant property. There are bounded turning circles, and lines, in infinite-dimensional Banach spaces that are not doubling; thus both conditions in Theorem 15.3 are needed. (See Exercise 15.4.) Theorem 15.3 is due to Tukia and Väisälä, and we refer to [176] for a proof.

Exercise 15.4. Equip $\mathbb{R}^\infty = \{x = (x_1, x_2, \ldots) : x_i \in \mathbb{R}\}$ with the norm $\|x\| = \max_i |x_i|$, and let $e_1, e_2, \ldots$ be the standard basis. Let $v_i = e_i/i$ and let γ be the topological circle in $\mathbb{R}^\infty$ that is the union of the lines $[v_i, v_{i+1}]$ for $i \geq 0$, where

---

[2] There is a more general definition for Gromov hyperbolicity that does not require this assumption [67].

$v_0 = 0$. Show that γ is of bounded turning but not doubling. In particular, γ cannot be quasisymmetrically equivalent to the standard circle $\mathbb{S}^1$.

Theorem 15.5. *A conformal gauge is equivalent to that determined by the standard Cantor ternary set if and only if it is compact, doubling, uniformly perfect, and uniformly disconnected.*

Theorem 15.5 is due to David and Semmes, and we refer to [39] for a proof.

To have some idea how difficult the characterization problem can be, we recall that there are closed (compact, no boundary) smooth 4-manifolds that carry infinitely many inequivalent, locally Euclidean conformal gauges. (See [42].) In contrast, there is a unique locally Euclidean conformal gauge on any given closed topological n-manifold for $n \neq 4$ by [173].

Another difficult open problem is the following.

15.6 Open problem. Characterize conformal gauges that are locally Euclidean.

In effect, Problem 15.6 asks for an intrinsic characterization of quasiconformal manifolds. Problem 15.6 in the context of metric gauges has been studied in [89]. A *metric gauge* on a set is defined similarly to the conformal gauge but by using bi-Lipschitz maps instead of quasisymmetric maps.

Interestingly, the Euclidean embedding problem for conformal gauges is solved completely by way of Assouad's theorem (Theorem 12.2): a conformal gauge is embeddable in some Euclidean space if and only if the gauge is doubling. On the other hand, little is known about the same problem for metric gauges. Let us rephrase Problem 12.3 as follows.

15.7 Open problem. Characterize metric gauges that are embeddable in some Euclidean space.

15.8 Conformal dimension. One cannot unambiguously define the Hausdorff dimension of a conformal gauge, except in some trivial cases. For example, the snowflake transformation $d \mapsto d^\epsilon$, $0 < \epsilon < 1$, raises the Hausdorff dimension by the (multiplicative) factor $1/\epsilon$. On the other hand, if a conformal gauge is uniformly perfect, then it has either positive and finite, infinite, or zero Hausdorff dimension by Corollary 11.11; that is, if one of the three alternatives holds for one metric in the gauge, then it holds for all the other metrics as well.

The snowflake transformation also shows that there is, in general, no finite supremal Hausdorff dimension for the metrics in a conformal gauge. However, it turns out to be interesting to study how *small* the Hausdorff dimension can be. If $\mathcal{G}$ is a given conformal gauge on a set X, we define its *conformal (Hausdorff) dimension*, $\operatorname{confdim}\mathcal{G}$, to be the infimum of the Hausdorff dimensions of the metric spaces (X, d), $d \in \mathcal{G}$. Given a metric space X, we define its conformal dimension to be the conformal dimension of the conformal gauge determined by X.

By using Exercise 14.27, it is easy to see that the conformal dimension of every uniformly disconnected gauge is zero. For example (by Theorem 15.5), the conformal dimension of the gauge determined by the Cantor ternary set is zero. Jeremy Tyson has proposed the interesting conjecture that the conformal dimension of each compact conformal gauge is either zero or at least one. (Compare Exercise 15.21.)

The conformal dimension of $\mathbb{R}^n$ is n for purely topological reasons. (See Theorem 8.14.) It is of interest to find gauges whose conformal dimension is somehow "large." Let us call a conformal gauge $\mathcal{G}$ *fractal* if its conformal dimension is larger than its topological dimension. (The term "fractal" here should not be taken as an indication that fractal gauges have nonintegral conformal dimension. In fact, this need not be the case, as follows from Corollary 15.16.)

The following theorem identifies the conformal dimension for many gauges.

Theorem 15.9. *If a conformal gauge $\mathcal{G}$ contains an n-regular n-Loewner metric, then* $\operatorname{confdim} \mathcal{G} = n$.

In fact, the following stronger result is true.

Theorem 15.10. *Let X be a compact Ahlfors n-regular metric space, $n > 1$, with nontrivial n-modulus. Then, the conformal dimension of X is n.*

We say that a metric measure space X has *nontrivial p-modulus*, $p \geq 1$, if there exists a family of curves in X with positive p-modulus. (Recall the definition for the p-modulus of a curve family from Chapter 7.) Thus, it is clear that Theorem 15.9 follows from Theorem 15.10.

Proof of Theorem 15.10. We may assume that there is a curve family Γ in X with positive n-modulus such that the diameter of each curve γ in Γ is at least a fixed positive constant d, independent of γ. (See formula (7.7).) Suppose now that $f : X \to Y$ is a quasisymmetric homeomorphism onto a metric space Y with Hausdorff dimension less than n. Because X is compact, f is uniformly continuous, from which we have a uniform positive lower bound, say $d' > 0$, for the diameters of the curves $f(\gamma)$, $\gamma \in \Gamma$.

Next, fix $\epsilon > 0$ and fix a disjointed collection of balls $B'_1, B'_2, \dots$ with radii $r'_1, r'_2, \dots$ in Y such that $Y = \cup_i 5B'_i$ and that

$$\sum_i r_i'^n < \epsilon. \tag{15.11}$$

This is possible by the basic covering theorem (Theorem 1.2) and by the contrapositive assumption that the Hausdorff dimension of Y is less than n. Then, choose balls $B_1, B_2, \dots$ with radii $r_1, r_2, \dots$ in X so that

$$B_i \subset f^{-1}(B'_i) \subset f^{-1}(5B'_i) \subset HB_i$$

for some fixed $H \geq 1$ depending on the quasisymmetry of f only; we may assume, by choosing ϵ sufficiently small, that $8Hr_i < d$. Define a Borel function

$\rho : X \to [0, \infty]$ by

$$\rho(x) = \sum_i \frac{r_i'}{r_i} \chi_{2HB_i}(x). \tag{15.12}$$

If $\gamma \in \Gamma$, then

$$\int_\gamma \rho \, ds = \sum_i \frac{r_i'}{r_i} \mathrm{length}(\gamma \cap 2HB_i) \geq \frac{1}{10} \sum_{i, f(\gamma) \cap 5B_i' \neq \emptyset} 10 r_i' \frac{\mathrm{length}(\gamma \cap 2HB_i)}{r_i}. \tag{15.13}$$

Since $\mathrm{diam} 2HB_i \leq 4Hr_i < d/2 < \mathrm{diam}\gamma$, we know that γ cannot lie in $2HB_i$. If i is an index such that $f(\gamma) \cap 5B_i' \neq \emptyset$, then also $\gamma \cap HB_i \neq \emptyset$, from which $\mathrm{length}(\gamma \cap 2HB_i) \geq Hr_i$. It therefore follows from Eq. (15.13) that

$$\int_\gamma \rho \, ds \geq \frac{H}{10} \sum_{i, f(\gamma) \cap 5B_i' \neq \emptyset} 10 r_i' \geq \frac{H}{10} \mathcal{H}_1^\infty(f(\gamma)) \geq \frac{H}{10} \mathrm{diam} f(\gamma) \geq \frac{H}{10} d'. \tag{15.14}$$

Here, $\mathcal{H}_1^\infty$ denotes the Hausdorff 1-content as defined in Eq. (8.5), and we also used the fact that the diameter of every connected space bounds its 1-content from below (Exercise 8.17).

In particular, we have from formula (15.14) that the function $10\rho/Hd'$ is admissible for the curve family Γ. Using this, Exercise 2.10, Ahlfors regularity, and inequality (15.11), we obtain

$$\begin{aligned} \mathrm{mod}_n \Gamma &\leq C \int_X \rho^n d\mu = C \int_X \left(\sum_i \frac{r_i'}{r_i} \chi_{2HB_i} \right)^n d\mu \\ &\leq C \int_X \left(\sum_i \frac{r_i'}{r_i} \chi_{B_i} \right)^n d\mu \leq C \sum_i \frac{r_i'^n}{r_i{}^n} \mu(B_i) \\ &\leq C \sum_i r_i'^n < C\epsilon, \end{aligned}$$

where $C \geq 1$ is independent of ϵ. This is a contradiction, and the theorem follows. □

Remark 15.15. Upon inspecting the proof, one sees that the assumption of Ahlfors regularity in Theorem 15.10 can be replaced by the assumption that X is a doubling metric measure space where the upper mass bound $\mu(B_R) \leq CR^n$ holds uniformly for small balls. The conclusion then is that the conformal dimension of X is at least n. (Compare formula (8.9).)

Recall that the Heisenberg group, or more generally all Carnot groups that are different from $\mathbb{R}^n$, are Ahlfors regular Loewner spaces with regularity dimension larger than the topological dimension. (See Section 9.25.) Moreover, Bourdon and Pajot [21] and Laakso [116] have exhibited one-dimensional spaces that are Ahlfors regular for noninteger exponents larger than 1, and Loewner. Laakso in particular exhibits such spaces for each real number $n > 1$. The presentation of these examples here would take us too far, but nevertheless we can record the following corollary.

Corollary 15.16. *For each real number $n > 1$, there exists a fractal conformal gauge of conformal dimension n.*

Another observation is the following.

Corollary 15.17. *Two conformal gauges are inequivalent if they contain Ahlfors regular Loewner spaces of different Hausdorff dimension.*

It follows from the aforementioned examples of Bourdon, Pajot, and Laakso that a set can have infinitely many inequivalent conformal gauges, each of which contains an Ahlfors regular Loewner metric.

Remark 15.18. In the context of Theorem 15.9, one can prove a stronger quantitative statement: if X is a metric space that is quasisymmetrically equivalent to an Ahlfors n-regular Loewner space, then

$$\mathcal{H}_n(B_R) \geq C^{-1} R^n \tag{15.19}$$

for each ball B_R in X of radius $0 < R < \operatorname{diam} X$. See [179].

The following conjecture was suggested by Jeremy Tyson (compare the discussion after Theorem 12.2).

Conjecture 15.20. *If a conformal gauge is fractal and contains an Ahlfors regular Loewner metric, then it cannot be embedded bi-Lipschitzly in Euclidean space.*

Tyson [181] has exhibited more examples of fractal gauges. In fact, he shows that if K is a compact Ahlfors n-regular metric space, $n > 0$, then the conformal dimension of the Ahlfors $(n+1)$-regular metric space $X \times [0, 1]$ (equipped with the usual product metric) is $n + 1$. Tyson's result can be used to prove Corollary 15.16 as well.

Another interesting question is whether the conformal dimension of a given gauge is realized by some metric in the gauge. If the answer is affirmative, let us call the gauge in question *flat*. Note that the conformal dimension of the gauge determined by the Cantor ternary set is not attained by a metric in the gauge; thus, the gauge is nonflat. I do not know of examples of nonflat conformal gauges of conformal dimension at least one.[3]

[3]Bishop and Tyson [13], [12] have recently exhibited nonflat conformal gauges with conformal dimension any prescribed real number at least one.

In a similar manner, we may define the *conformal Assouad dimension*, $\mathrm{confdim}_A \mathcal{G}$, of a conformal gauge $\mathcal{G}$ (or a metric space) as the infimal Assouad dimension of the metrics in $\mathcal{G}$. It follows from Theorem 14.16 that if $\mathcal{G}$ is complete and doubling, then $\mathrm{confdim}_A \mathcal{G}$ is attained through Ahlfors regular metrics.

Naturally, one can use any dimension concept (Minkowski, packing, and others – see [134]) and define the corresponding conformal dimension. Each choice of dimension further gives rise to questions about fractal and flat gauges.

Exercise 15.21. (See [180].) Show that the conformal Assouad dimension of every metric space is either zero or at least one. (Hint: First, show that every metric space with Assouad dimension strictly less than one is uniformly disconnected; see [39, Lemma 15.2]. Then, use Exercise 14.27.)

There are many concrete metric spaces whose conformal dimension has not been identified. As an example, let us record the following.

15.22 Open problem. What is the conformal dimension of the Sierpinski carpet?

By a *Sierpinski carpet* we mean the classical fractal obtained by subdividing the square $[0, 1] \times [0, 1]$ in $\mathbb{R}^2$ into nine congruent squares, removing the middle one, and continuing in a similar manner. The resulting compact set Σ is a subset of the plane that is Ahlfors regular of dimension $\log 8/\log 3 \approx 1.893$. Because Σ contains the product of the Cantor ternary set with an interval, the conformal dimension of Σ is at least $1 + \frac{\log 2}{\log 3} = 1.6309\ldots$ by the results in [181]. I do not know of a better lower bound.

15.23 Notes to Chapter 15. The concept of conformal dimension is due to Pansu [146], who studied the conformal dimensions of the boundary gauges of certain negatively curved (Gromov hyperbolic) spaces. See also the work of Bourdon [19], [20] and Tyson [178], [181]. For the basics of Gromov hyperbolic spaces, see [67] and [69]. In [16], Gromov hyperbolic spaces are related to uniform domains and spaces.

In contrast to Corollary 15.9, Luukkainen [125] has shown that every finite-dimensional metric space X admits a topologically equivalent metric whose Hausdorff dimension equals the topological dimension of X.

Theorem 15.10 is due to Tyson and is a consequence of the theory developed in [178]. The short proof given here is previously unpublished and due to Mario Bonk and Jeremy Tyson.

References

[1] ADAMS, D.R. AND HEDBERG, L.I. *Function Spaces and Potential Theory*. Springer-Verlag, Berlin, 1996.

[2] ALESTALO, P. AND VÄISÄLÄ, J. Uniform domains of higher order. III. *Ann. Acad. Sci. Fenn. Ser. A I Math. 22* (1997), 445–464.

[3] ANCONA, A. On strong barrier and inequality of Hardy for domains in $\boldsymbol{R}^n$. *J. London Math. Soc. 34* (1986), 274–290.

[4] ASSOUAD, P. Plongements lipschitziens dans $\boldsymbol{R}^n$. *Bull. Soc. Math. France 111* (1983), 429–448.

[5] ASTALA, K. Planar quasiconformal mappings; deformations and interactions. In *Quasiconformal Mappings and Analysis (Ann Arbor, MI, 1995)*. Springer, New York, 1998, pp. 33–54.

[6] ASTALA, K. AND MARTIN, G. Holomorphic motions. Preprint, University of Helsinki, 1992.

[7] BALOGH, Z. AND KOSKELA, P. Quasiconformality, quasisymmetry, and removability in Loewner spaces. With an appendix by Jussi Väisälä. *Duke Math. J. 101* (2000), 554–577.

[8] BEARDON, A.F. AND POMMERENKE, C. The Poincaré metric of plane domains. *J. London Math. Soc. 18* (1978), 475–483.

[9] BENYAMINI, Y. AND LINDENSTRAUSS, J. *Geometric Nonlinear Functional Analysis, Volume I*. Volume 48 of Colloquium Publications. American Mathematical Society, Providence, R. I., 2000.

[10] BEURLING, A. AND AHLFORS, L. The boundary correspondence under quasiconformal mappings. *Acta Math. 96* (1956), 125–142.

[11] BISHOP, C.J. A quasisymmetric surface with no rectifiable curves. *Proc. Amer. Math. Soc. 127* (1999), 2035–2040.

[12] BISHOP, C.J. AND TYSON, J.T. The conformal dimension of the antenna set. *Proc. Amer. Math. Soc.* (to appear).

[13] Bishop, C.J. and Tyson, J.T. Locally minimal sets for conformal dimension. *Ann. Acad. Sci. Fenn. Ser. A I Math.* (to appear).

[14] Bojarski, B. Remarks on Sobolev imbedding inequalities. In *Complex Analysis, Joensuu 1987.* Volume 1351 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1988, pp. 52–68.

[15] Bojarski, B. and Iwaniec, I. Analytical foundations of the theory of quasiconformal mappings in $\boldsymbol{R}^n$. *Ann. Acad. Sci. Fenn. Ser. AI Math. 8* (1983), 257–324.

[16] Bonk, M., Heinonen, J. and Koskela, P. Uniformizing Gromov hyperbolic spaces. Astérisque (to appear).

[17] Bonk, M. and Schramm, O. Embeddings of Gromov hyperbolic spaces. *Geom. Funct. Anal. 10* (2000), 266–306.

[18] Bourbaki, N. *Topologie générale.* Hermann, Paris, 1974.

[19] Bourdon, M. Au bord de certains polyèdres hyperboliques. *Ann. Inst. Fourier (Grenoble) 45* (1995), 119–141.

[20] Bourdon, M. Immeubles hyperboliques, dimension conforme et rigidité de Mostow. *Geom. Funct. Anal. 7* (1997), 245–268.

[21] Bourdon, M. and Pajot, H. Poincaré inequalities and quasiconformal structure on the boundaries of some hyperbolic buildings. *Proc. Amer. Math. Soc. 127* (1999), 2315–2324.

[22] Bruckner, A., Bruckner, J. and Thomson, B. *Real Analysis.* Prentice-Hall, Englewood Cliffs, N.J., 1997.

[23] Buckley, S., Hanson, B. and MacManus, P. Doubling for general sets. *Math. Scand.* (to appear).

[24] Buckley, S., Koskela, P. and Lu, G. Boman equals John. In *XVIth Rolf Nevanlinna Colloquium (Joensuu, 1995).* de Gruyter, Berlin, 1996, pp. 91–99.

[25] Burago, Y.D. and Zalgaller, V.A. *Geometric inequalities.* Springer-Verlag, New York, 1988.

[26] Buser, P. A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup. 15 (4)* (1982), 213–230.

[27] Calderón, A.P. On differentiability of absolutely continuous functions. *Riv. Mat. Univ. Parma 2* (1951), 203–213.

[28] Cannon, J.W., Floyd, W.J., Kenyon, R. and Parry, W.R. Hyperbolic geometry. In *Flavors of Geometry.* Cambridge University Press, Cambridge, 1997, pp. 59–115.

[29] Capogna, L., Danielli, D. and Garofalo, N. An isoperimetric inequality and the Sobolev embedding theorem for vector fields. *Math. Res. Lett. 1* (1994), 263–268.

[30] Chavel, I. *Riemannian Geometry: A Modern Introduction.* Cambridge University Press, Cambridge, 1993.

[31] Cheeger, J. Differentiability of Lipschitz functions on metric spaces. *Geom. Funct. Anal. 9* (1999), 428–517.

[32] Cheeger, J. and Colding, T.H. On the structure of spaces with Ricci curvature bounded below. III. Preprint, 1999.

[33] Chi, Q.-S. Besicovitch covering lemma, Hadamard manifolds, and zero entropy. *J. Geom. Anal. 1* (1991), 373–382.

[34] Colding, T.H. and Minicozzi, W.P., II. Harmonic functions on manifolds. *Ann. Math. 146 (2)* (1997), 725–747.

[35] Coulhon, T. and Saloff-Coste, L. Variétés riemanniennes isométriques à l'infini. *Rev. Mat. Iberoamer. 11* (1995), 687–726.

[36] Danielli, D., Garofalo, N. and Nhieu, D.-M. Trace inequalities for Carnot–Carathéodory spaces and applications. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. 27 (4)* (1998), 195–252.

[37] David, G. Morceaux de graphes lipschitziennes et intégrales singulières sur un surface. *Rev. Mat. Iberoamer. 4* (1988), 73–114.

[38] David, G. and Semmes, S. Strong A_∞ weights, Sobolev inequalities and quasiconformal mappings. In *Analysis and Partial Differential Equations.* Volume 122 of Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, New York, 1990, pp. 101–111.

[39] David, G. and Semmes, S. *Fractured Fractals and Broken Dreams: Self-similar Geometry through Metric and Measure.* Volume 7 of Oxford Lecture Series in Mathematics and its Applications. Clarendon Press, Oxford, 1997.

[40] David, G. and Toro, T. Reifenberg flat metric spaces, snowballs, and embeddings. *Math. Ann. 315* (1999), 641–710.

[41] Deny, J. and Lions, J.L. Les espaces du type de Beppo Levi. *Ann. Inst. Fourier 5* (1953/1954), 305–370.

[42] Donaldson, S.K. and Sullivan, D.P. Quasiconformal 4-manifolds. *Acta Math. 163* (1989), 181–252.

[43] Duren, P.L. *Univalent Functions.* Springer-Verlag, New York, 1983.

[44] Duren, P.L., et al., Eds. *Quasiconformal Mappings and Analysis.* A collection of papers honoring F. W. Gehring; Papers from the International Symposium held in Ann Arbor, MI, August 1995. Springer-Verlag, New York, 1998.

[45] Evans, L.C. and Gariepy, R.F. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, Fl., 1992.

[46] Fabes, E.B., Kenig, C.E. and Serapioni, R. The local regularity of solutions to degenerate elliptic equations. *Comm. PDE 7* (1982), 77–116.

[47] Federer, H. *Geometric Measure Theory.* Volume 153 of Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, New York, 1969.

[48] Fefferman, R.A., Kenig, C.E. and Pipher, J. The theory of weights and the Dirichlet problem for elliptic equations. *Ann. Math. 134* (1991), 65–124.

[49] Fernández, J.L. Domains with strong barrier. *Rev. Mat. Iberoamer. 5* (1989), 47–65.

[50] Fernández, J.L., Heinonen, J. and Martio, O. Quasilines and conformal mappings. *J. Anal. Math. 52* (1989), 117–132.

[51] Fleming, W.H. and Rishel, R. An integral formula for total gradient variation. *Arch. Math. 11* (1960), 218–222.

[52] Folland, G.B. and Stein, E.M. *Hardy Spaces on Homogeneous Groups.* Princeton University Press, Princeton, N.J., 1982.

[53] Franchi, B., Gallot, S. and Wheeden, R.L. Sobolev and isoperimetric inequalities for degenerate metrics. *Math. Ann. 300* (1994), 557–571.

[54] Franchi, B., Hajłasz, P. and Koskela, P. Definitions of Sobolev classes on metric spaces. *Ann. Inst. Fourier 49* (1999), 1903–1924.

[55] Franchi, B., Lu, G. and Wheeden, R.L. A relationship between Poincaré-type inequalities and representation formulas in spaces of homogeneous type. *Int. Math. Res. Not.*, 1 (1996), 1–14.

[56] Fuglede, B. Extremal length and functional completion. *Acta Math. 98* (1957), 171–219.

[57] Gallot, S., Hulin, D. and Lafontaine, J. *Riemannian Geometry.* Springer-Verlag, New York, 1990.

[58] García-Cuerva, J. and Rubio De Francia, J.L. *Weighted Norm Inequalities and Related Topics.* North-Holland, Amsterdam, 1985.

[59] Garnett, J.B. *Bounded Analytic Functions.* Academic Press, New York, 1981.

[60] Garofalo, N. and Nhieu, D.-M. Isoperimetric and Sobolev inequalities for Carnot–Carathéodory spaces and the existence of minimal surfaces. *Comm. Pure Appl. Math. 49* (1996), 1081–1144.

[61] Gehring, F.W. Quasiconformal mappings of slit domains in three space. *J. Math. Mech. 18* (1969), 689–703.

[62] Gehring, F.W. Characteristic properties of quasidisks. Les Presses de l'Université de Montréal, Montréal, 1982.

[63] Gehring, F.W. Uniform domains and the ubiquitous quasidisk. *Jahresber. Deutsch. Math.-Verein. 89* (1987), 88–103.

[64] Gehring, F.W. and Martio, O. Quasiextremal distance domains and extension of quasiconformal mappings. *J. Anal. Math. 45* (1985), 181–206.

[65] Gehring, F.W. and Väisälä, J. The coefficients of quasiconformality of domains in space. *Acta Math. 114* (1965), 1–70.

[66] Ghamsari, M. and Herron, D.A. Higher dimensional Ahlfors regular sets and chordarc curves in $\boldsymbol{R}^n$. *Rocky Mount. J. Math. 28* (1998), 191–222.

[67] Ghys, E. and de la Harpe, P. *Sur les Groupes Hyperboliques d'après Mikhael Gromov.* Progress in Mathematics. Birkhäuser, Boston, 1990.

[68] González, M.J. Uniformly perfect sets, Green's function, and fundamental domains. *Rev. Mat. Iberoamer. 8* (1992), 239–269.

[69] Gromov, M. Hyperbolic groups. In *Essays in Group Theory.* S. Gersten, Ed. Volume 8 of MSRI Publications. Springer-Verlag, New York, 1987, pp. 75–265.

[70] Gromov, M. Carnot–Carathéodory spaces seen from within. In *Sub-Riemannian Geometry.* Volume 144 of Progress in Mathematics. Birkhäuser, Basel, 1996, pp. 79–323.

[71] Hajłasz, P. Geometric approach to Sobolev spaces and badly degenerated elliptic equations. In *Nonlinear Analysis and Applications (Warsaw, 1994).* Volume 7 of GAKUTO International Series Mathematical Sciences and Applications. Gakkōtosho, Tokyo, 1996, pp. 141–168.

[72] Hajłasz, P. Sobolev spaces on an arbitrary metric space. *Potential Anal. 5* (1996), 403–415.

[73] Hajłasz, P. and Kinnunen, J. Hölder quasicontinuity of Sobolev functions on metric spaces. *Rev. Mat. Iberoamer. 14* (1998), 601–622.

[74] Hajłasz, P. and Koskela, P. Sobolev meets Poincaré. *C. R. Acad. Sci. Paris Sér. A Math. 320* (1995), 1211–1215.

[75] Hajłasz, P. and Koskela, P. Sobolev met Poincaré. *Mem. Amer. Math. Soc. 145* (2000).

[76] Hajłasz, P. and Martio, O. Traces of Sobolev functions on fractal type sets and characterization of extension domains. *J. Funct. Anal. 143* (1997), 221–246.

[77] Hanson, B. and Heinonen, J. An n-dimensional space that admits a Poincaré inequality but has no manifold points. *Proc. Amer. Math. Soc.* (to appear).

[78] Hedberg, L.I. On certain convolution inequalities. *Proc. Amer. Math. Soc. 36* (1972), 505–510.

[79] Heinonen, J. Quasiconformal mappings onto John domains. *Rev. Mat. Iberoamer. 5* (1989), 97–123.

[80] Heinonen, J. Calculus on Carnot groups. In *Fall School in Analysis (Jyväskylä, 1994).* Volume 68 of Report, Univ. Jyväskylä 1995, pp. 1–31.

[81] Heinonen, J., Kilpeläinen, T. and Martio, O. *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Oxford Science Publications. Clarendon Press/ Oxford University Press, New York, 1993.

[82] Heinonen, J. and Koskela, P. Definitions of quasiconformality. *Invent. Math. 120* (1995), 61–79.

[83] Heinonen, J. and Koskela, P. From local to global in quasiconformal structures. *Proc. Natl. Acad. Sci. USA 93* (1996), 554–556.

[84] Heinonen, J. and Koskela, P. Quasiconformal maps in metric spaces with controlled geometry. *Acta Math. 181* (1998), 1–61.

[85] Heinonen, J. and Koskela, P. A note on Lipschitz functions, upper gradients, and the Poincaré inequality. *N.Z. Math. J. 28* (1999), 37–42.

[86] Heinonen, J. and Rickman, S. Geometric branched covers between generalized manifolds. Preprint 240, University of Helsinki, 1999.

[87] Heinonen, J. and Rohde, S. Koenigs functions, quasicircles and BMO. *Duke Math. J. 78* (1995), 301–313.

[88] Heinonen, J. and Semmes, S. Thirty-three *yes* or *no* questions about mappings, measures, and metrics. *Conform. Geom. Dyn. 1* (1997), 1–12.

[89] Heinonen, J. and Sullivan, D. On the locally Euclidean metric gauge. in preparation.

[90] Hesse, J. A p-extremal length and p-capacity equality. *Ark. Mat. 13* (1975), 131–144.

[91] Hewitt, E. and Stromberg, K. *Real and Abstract Analysis*. Springer-Verlag, New York, 1965.

[92] Hinkkanen, A. Julia sets of rational functions are uniformly perfect. *Math. Proc. Cambridge Philos. Soc. 113* (1993), 543–559.

[93] Hocking, J.G. and Young, G.S. *Topology*. Academic Press, New York, 1959.

[94] Hurewicz, W. and Wallman, H. *Dimension Theory*. Volume 4 of Princeton Mathematical Series. Princeton University Press, Princeton, N.J., 1941.

[95] Iwaniec, T. Current advances in quasiconformal geometry and nonlinear analysis. In *XVIth Rolf Nevanlinna Colloquium (Joensuu, 1995)*. de Gruyter, Berlin, 1996, pp. 59–80.

[96] Iwaniec, T. and Martin, G. Geometric function theory and nonlinear analysis. Oxford Univ. Press, to appear.

[97] Iwaniec, T., Scott, C. and Stroffolini, B. Nonlinear Hodge theory on manifolds with boundary. *Ann. Mat. Pura Appl. 177* (1999), 37–115.

[98] Järvi, P. and Vuorinen, M. Uniformly perfect sets and quasiregular mappings. *J. London Math. Soc. 54* (1996), 515–529.

[99] Jerison, D. The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J. 53* (1986), 503–523.

[100] Jones, P.W. Quasiconformal mappings and extendability of functions in Sobolev spaces. *Acta Math. 147* (1981), 71–88.

[101] Kallunki, S. and Shanmugalingam, N. Modulus and continuous capacity. Ann. Acad. Sci. Fenn. Ser. A I Math., to appear.

[102] Kaufman, R. and Wu, J.-M. Two problems on doubling measures. *Rev. Mat. Iberoamer. 11* (1995), 527–545.

[103] Kelingos, J.A. Boundary correspondence under quasiconformal mappings. *Mich. Math. J. 13* (1966), 235–249.

[104] Kilpeläinen, T. Weighted Sobolev spaces and capacity. *Ann. Acad. Sci. Fenn. Ser. AI Math. 19* (1994), 95–113.

[105] Kilpeläinen, T. Smooth approximation in weighted Sobolev spaces. *Comment. Math. Univ. Carolin. 38* (1997), 29–35.

[106] Kilpeläinen, T., Kinnunen, J. and Martio, O. Sobolev spaces with zero boundary values on metric spaces. *Potential Anal. 12* (2000), 233–247.

[107] Kilpeläinen, T. and Koskela, P. Global integrability of the gradients of solutions to partial differential equations. *Nonlinear Anal. 23* (1994), 899–909.

[108] Kinnunen, J. and Martio, O. The Sobolev capacity on metric spaces. *Ann. Acad. Sci. Fenn. Ser. AI Math. 21* (1996), 367–382.

[109] Korányi, A. Geometric properties of Heisenberg-type groups. *Adv. Math. 56* (1985), 28–38.

[110] Korányi, A. and Reimann, H.M. Quasiconformal mappings on the Heisenberg group. *Invent. Math. 80* (1985), 309–338.

[111] Korányi, A. and Reimann, H.M. Foundations for the theory of quasiconformal mappings on the Heisenberg group. *Adv. Math. 111* (1995), 1–87.

[112] Koskela, P. Removable sets for Sobolev spaces. *Ark. Mat. 37* (1999), 291–304.

[113] Koskela, P. and MacManus, P. Quasiconformal mappings and Sobolev spaces. *Studia Math. 131* (1998), 1–17.

[114] Koskela, P., Shanmugalingam, N. and Tuominen, H. Removable sets for the Poincaré inequality on metric spaces. *Indiana Univ. Math. J.* to appear.

[115] Krantz, S.G. *Complex analysis: The Geometric Viewpoint*. Mathematical Association of America, Washington, D.C., 1990.

[116] Laakso, T. Ahlfors Q-regular spaces with arbitrary Q admitting weak Poincaré inequalities. *Geom. Funct. Anal. 10* (2000), 111–123.

[117] Laakso, T. Plane with A_∞-weighted metric not bilipschitz embeddable to $\boldsymbol{R}^n$. Preprint 258, University of Helsinki, 2000.

[118] Lang, U., Pavlović, B. and Schroeder, V. Extensions of lipschitz maps into hadamard spaces. Preprint (1999).

[119] Larman, D.G. A new theory of dimension. *Proc. London Math. Soc. 17* (1967), 178–192.

[120] Lehto, O. *Univalent Functions and Teichmüller Spaces*. Springer-Verlag, New York.

[121] Lewis, J.L. Uniformly fat sets. *Trans. Amer. Math. Soc. 830* (1988), 177–196.

[122] Loewner, C. On the conformal capacity in space. *J. Math. Mech. 8* (1959), 411–414.

[123] Lu, G. Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications. *Rev. Mat. Iberoamer. 8* (1992), 367–439.

[124] Luosto, K. Ultrametric spaces bi-Lipschitz embeddable in $\boldsymbol{R}^n$. *Fundam. Math. 150* (1996), 25–42.

[125] Luukkainen, J. Assouad dimension: antifractal metrization, porous sets, and homogeneous measures. *J. Korean Math. Soc. 35* (1998), 23–76.

[126] Luukkainen, J. and Movahedi-Lankarani, H. Minimal bi-Lipschitz embedding dimension of ultrametric spaces. *Fundam. Math. 144* (1994), 181–193.

[127] Luukkainen, J. and Saksman, E. Every complete doubling metric space carries a doubling measure. *Proc. Amer. Math. Soc. 126* (1998), 531–534.

[128] Luukkainen, J. and Väisälä, J. Elements of Lipschitz topology. *Ann. Acad. Sci. Fenn. Ser. AI Math. 3* (1977), 85–122.

[129] Mañé, R. and da Rocha, L.F. Julia sets are uniformly perfect. *Proc. Amer. Math. Soc. 116* (1992), 251–257.

[130] Mañé, R., Sad, P. and Sullivan, D. On the dynamics of rational maps. *Ann. Sci. École Norm. Sup. 16 (4)* (1983), 193–217.

[131] Maćias, R.A. and Segovia, C. Lipschitz functions on spaces of homogeneous type. *Adv. Math. 33* (1979), 257–270.

[132] Martio, O., Rickman, S. and Väisälä, J. Topological and metric properties of quasiregular mappings. *Ann. Acad. Sci. Fenn. Ser. AI Math.*, 488 (1971), 1–31.

[133] Mattila, P. Integration in a space of measures. *Ann. Acad. Sci. Fenn. Ser. AI*, 555 (1973), 1–37.

[134] Mattila, P. *Geometry of Sets and Measures in Euclidean Spaces*. Volume 44 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1995.

[135] Maxwell, J.C. *A Treatise on Electricity and Magnetism*, third ed. Clarendon Press, Oxford, 1892.

[136] Maz'ja, V.G. *Sobolev Spaces*. Springer-Verlag, Berling, 1985.

[137] McMullen, C.T. and Sullivan, D.P. Quasiconformal homeomorphisms and dynamics III: the Teichmüller space of a holomorphic dynamical system. *Adv. Math. 135* (1998), 351–395.

[138] McShane, E.J. Extension of range of functions. *Bull. Amer. Math. Soc. 40* (1934), 837–842.

[139] Meyers, N. and Serrin, J. $H = W$. *Proc. Natl. Acad. Sci. USA 51* (1964), 1055–1056.

[140] Milman, V.D. and Schechtman, G. *Asymptotic Theory of Finite Dimensional Normed Spaces*. Volume 1200 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1986.

[141] Mitchell, J. On Carnot–Carathéodory metrics. *J. Differential Geom. 21* (1985), 35–45.

[142] Morrey, C.B. *Multiple Integrals in the Calculus of Variations*. Springer-Verlag, Berlin, 1966.

[143] Nagel, A., Stein, E.M. and Wainger, S. Balls and metrics defined by vector fields I: Basic properties. *Acta Math. 155* (1985), 103–147.

[144] Näkki, R. Extension of Loewner's capacity theorem. *Trans. Amer. Math. Soc. 180* (1973), 229–236.

[145] Näkki, R. and Väisälä, J. John disks. *Exposition. Math. 9* (1991), 3–43.

[146] Pansu, P. Dimension conforme et sphère à l'infini des variétés à courbure négative. *Ann. Acad. Sci. Fenn. Ser. AI Math. 14* (1989), 177–212.

[147] Pansu, P. Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. Math. 129 (2)* (1989), 1–60.

[148] Reshetnyak, Y.G. *Space Mappings with Bounded Distortion*. Volume 73 of Translations of Mathematical Monographs. American Mathematical Society, Providence, R.I., 1989. Translated from the Russian by H.H. McFaden.

[149] Rickman, S. *Quasiregular Mappings*. Springer-Verlag, Berlin, 1993.

[150] Rohde, S. On conformal welding and quasicircles. *Mich. Math. J. 38* (1991), 111–116.

[151] Rohde, S. Quasicircles modulo bi-Lipschitz maps. Preprint, 1999.

[152] Royden, H.L. *Real Analysis*, third ed. Macmillan, New York, 1988.

[153] Rudin, W. *Real and Complex Analysis*, third ed. McGraw-Hill, New York, 1987.

[154] Rudin, W. *Functional Analysis*, second ed. McGraw-Hill, New York, 1991.

[155] Rushing, T.B. *Topological Embeddings*. Volume 52 of Pure and Applied Mathematics. Academic Press, New York, 1973.

[156] Rushing, T.B. Hausdorff dimension of wild fractals. *Trans. Amer. Math. Soc. 334* (1992), 597–613.

[157] Saksman, E. Remarks on the nonexistence of doubling measures. *Ann. Acad. Sci. Fenn. Ser. AI Math.* (1999), 155–163.

[158] Saloff-Coste, L. Uniformly elliptic operators on Riemannian manifolds. *J. Differential Geom. 36* (1992), 417–450.

[159] SCOTT, C. L^p-theory of differential forms on manifolds. *Trans. Amer. Math. Soc. 347* (1995), 2075–2096.

[160] SEMMES, S. Finding curves on general spaces through quantitative topology, with applications to Sobolev and Poincaré inequalities. *Selecta Math. 2* (1996), 155–295.

[161] SEMMES, S. Good metric spaces without good parameterizations. *Rev. Mat. Iberoamer. 12* (1996), 187–275.

[162] SEMMES, S. On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights. *Rev. Mat. Iberoamer. 12* (1996), 337–410.

[163] SEMMES, S. Bi-Lipschitz embeddings of metric spaces into Euclidean spaces. *Publ. Mat. 43* (1999), 571–653.

[164] SEMMES, S. Metric spaces and mappings seen at many scales (appendix). In *Metric Structures in Riemannian Spaces. M. Gromov*. Ed. Progress in Mathematics. Birkhäuser, Boston, 1999.

[165] SEMMES, S. Some novel types of fractal geometry. Oxford University Press, to appear.

[166] SHANMUGALINGAM, N. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. Ph.D. thesis, University of Michigan, 1999.

[167] SHANMUGALINGAM, N. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamer.* (to appear).

[168] SLODKOWSKI, Z. Holomorphic motions and polynomial hulls. *Proc. Amer. Math. Soc. 111* (1991), 347–355.

[169] STAPLES, S.G. AND WARD, L. Quasisymmetrically thick sets. *Ann. Acad. Sci. Fenn. Ser. AI Math. 23* (1998), 151–168.

[170] STEIN, E.M. *Singular Integrals and Differentiability Properties of Functions*. Volume 30 of Princeton Mathematical Series. Princeton University Press, Princeton, N.J., 1970.

[171] STEIN, E.M. (With the assistance of Timothy S. Murphy). *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Volume 3 of Monographs in Harmonic Analysis. Princeton University Press, Princeton, NJ, 1993.

[172] STRICHARTZ, R.S. Sub-Riemannian geometry. *J. Differential Geom. 24* (1986), 221–263.

[173] SULLIVAN, D. Hyperbolic geometry and homeomorphisms. In *Geometric Topology*. Proceedings of the Georgia Topology Conference, Athens, Ga., 1977. Academic Press, New York, 1979, pp. 543–555.

[174] TROTSENKO, D.A. AND VÄISÄLÄ, J. Upper sets and quasisymmetric maps. *Ann. Acad. Sci. Fenn. Ser. AI Math.* (1999), 465–488.

[175] TUKIA, P. Hausdorff dimension and quasisymmetric mappings. *Math. Scand. 65* (1989), 152–160.

[176] TUKIA, P. AND VÄISÄLÄ, J. Quasisymmetric embeddings of metric spaces. *Ann. Acad. Sci. Fenn. Ser. AI Math. 5* (1980), 97–114.

[177] TYSON, J. Quasiconformality and quasisymmetry in metric measure spaces. *Ann. Acad. Sci. Fenn. Ser. AI Math. 23* (1998), 525–548.

[178] TYSON, J. Geometric and analytic applications of a generalized definition of the conformal modulus. Ph.D. thesis, University of Michigan, 1999.

[179] TYSON, J. Analytic properties of locally quasisymmetric mappings from Euclidean domains. Indiana Univ. Math. J.

[180] TYSON, J. Lowering the Assouad dimension by quasisymmetric mappings. *Ill. J. Math.* (to appear).

[181] TYSON, J. Sets of minimal Hausdorff dimension for quasiconformal maps. *Proc. Amer. Math. Soc.* (to appear).

[182] VÄISÄLÄ, J. *Lectures on n-dimensional Quasiconformal Mappings.* Volume 229 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1971.

[183] VÄISÄLÄ, J. Quasi-symmetric embeddings in Euclidean spaces. *Trans. Amer. Math. Soc. 264* (1981), 191–204.

[184] VÄISÄLÄ, J. Quasimöbius maps. *J. Anal. Math. 44* (1984/1985), 218–234.

[185] VÄISÄLÄ, J. Invariants for quasisymmetric, quasi-Möbius and bi-Lipschitz maps. *J. Anal. Math. 50* (1988), 201–223.

[186] VÄISÄLÄ, J. Quasiconformal maps of cylindrical domains. *Acta Math. 162* (1989), 201–225.

[187] VÄISÄLÄ, J. Free quasiconformality in Banach spaces. I. *Ann. Acad. Sci. Fenn. Ser. AI Math. 15* (1990), 355–379.

[188] VÄISÄLÄ, J. Quasisymmetry and unions. *Manuscripta Math. 68* (1990), 101–111.

[189] VÄISÄLÄ, J. Domains and maps. In *Quasiconformal Space Mappings*, A Collection of Surveys 1960–1990. M. Vuorinen, Ed. Springer-Verlag, Berlin, 1992.

[190] VÄISÄLÄ, J. Metric duality in Euclidean spaces. *Math. Scand. 80* (1997), 249–288.

[191] VÄISÄLÄ, J. Questions on quasiconformal maps in space. In *Quasiconformal Mappings and Analysis (Ann Arbor, MI, 1995)*. Springer, New York, 1998, pp. 369–374.

[192] VÄISÄLÄ, J. Relatively and inner uniform domains. *Conform. Geom. Dyn. 2* (1998), 56–88.

[193] VÄISÄLÄ, J. The free quasiworld. Freely quasiconformal and related maps in Banach spaces. In Quasiconformal Geometry and Dynamics, Banach Center Publications, Volume 48, 55–118, Polish Academy of Sciences, Warszawa 1999.

[194] VAROPOULOS, N.T. Fonctions harmoniques sur les groupes de Lie. *C. R. Acad. Sci. Paris Sér. A Math. 304* (1987), 519–521.

[195] VAROPOULOS, N.T., SALOFF-COSTE, L. AND COULHON, T. *Analysis and Geometry on Groups.* Volume 100 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1992.

[196] VOL´BERG, A.L. AND KONYAGIN, S.V. On measures with the doubling condition. *Izv. Akad. Nauk SSSR Ser. Mat. 51* (1987), 666–675. English translation: *Math. USSR-Izv., 30*(1988), 629–638.

[197] VUORINEN, M. *Conformal Geometry and Quasiregular Mappings.* Springer-Verlag, Berlin, 1988.

[198] WARNER, F.W. *Foundations of Differentiable Manifolds and Lie Groups.* Springer-Verlag, New York, 1983.

[199] WEAVER, N. Lipschitz algebras and derivations II: exterior differentiation. Preprint, 1998.

[200] WU, J.-M. Boundary density and the Green function. *Mich. Math. J. 38* (1991), 175–189.

[201] WU, J.-M. Null sets for doubling and dyadic doubling measures. *Ann. Acad. Sci. Fenn. Ser. A1 Math. 18* (1993), 77–91.

[202] WU, J.-M. Porous sets and null sets for elliptic harmonic measures. *Trans. Amer. Math. Soc. 346* (1994), 455–473.

[203] WU, J.-M. Hausdorff dimension and doubling measures on metric spaces. *Proc. Amer. Math. Soc. 126* (1998), 1453–1459.

[204] YOSIDA, K. *Functional Analysis.* Springer-Verlag, New York, 1980.

[205] ZIEMER, W.P. *Weakly Differentiable Functions.* Volume 120 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1989.

Index

Universitext (continued)

McCarthy: Introduction to Arithmetical Functions
Meyer: Essential Mathematics for Applied Fields
Mines/Richman/Ruitenburg: A Course in Constructive Algebra
Moise: Introductory Problems Course in Analysis and Topology
Morris: Introduction to Game Theory
Poizat: A Course In Model Theory: An Introduction to Contemporary Mathematical Logic
Polster: A Geometrical Picture Book
Porter/Woods: Extensions and Absolutes of Hausdorff Spaces
Radjavi/Rosenthal: Simultaneous Triangularization
Ramsay/Richtmyer: Introduction to Hyperbolic Geometry
Reisel: Elementary Theory of Metric Spaces
Ribenboim: Classical Theory of Algebraic Numbers
Rickart: Natural Function Algebras
Rotman: Galois Theory
Rubel/Colliander: Entire and Meromorphic Functions
Sagan: Space-Filling Curves
Samelson: Notes on Lie Algebras
Schiff: Normal Families
Shapiro: Composition Operators and Classical Function Theory
Simonnet: Measures and Probability
Smith: Power Series From a Computational Point of View
Smith, Kahanpää, Kekäläinen, Traves: An Invitation to Algebraic Geometry
Smoryński: Self-Reference and Modal Logic
Stillwell: Geometry of Surfaces
Stroock: An Introduction to the Theory of Large Deviations
Sunder: An Invitation to von Neumann Algebras
Tondeur: Foliations on Riemannian Manifolds
Wong: Weyl Transforms
Zhang: Matrix Theory: Basic Results and Techniques
Zong: Sphere Packings
Zong: Strange Phenomena in Convex and Discrete Geometry