OXFORD LECTURE SERIES
IN MATHEMATICS AND ITS APPLICATIONS

1. J. C. Baez (ed.): *Knots and quantum gravity*
2. I. Fonseca and W. Gangbo: *Degree theory in analysis and applications*
3. P. L. Lions: *Mathematical topics in fluid mechanics, Vol. 1: Incompressible models*
4. J. E. Beasley (ed.): *Advances in linear and integer programming*
5. L. W. Beineke and R. J. Wilson (eds): *Graph connections: Relationships between graph theory and other areas of mathematics*
6. I. Anderson: *Combinatorial designs and tournaments*
7. G. David and S. W. Semmes: *Fractured fractals and broken dreams*

Fractured Fractals and Broken Dreams

Self-Similar Geometry through Metric and Measure

Guy David
Université de Paris XI and Institut Universitaire de France

and

Stephen Semmes
Department of Mathematics, Rice University

CLARENDON PRESS • OXFORD
1997

Oxford University Press, Great Clarendon Street, Oxford OX2 6DP
Oxford New York
Athens Auckland Bangkok Bogota Bombay
Buenos Aires Calcutta Cape Town Dar es Salaam
Delhi Florence Hong Kong Istanbul Karachi
Kuala Lumpur Madras Madrid Melbourne
Mexico City Nairobi Paris Singapore
Taipei Tokyo Toronto Warsaw
and associated companies in
Berlin Ibadan

Oxford is a trade mark of Oxford University Press

Published in the United States by
Oxford University Press Inc., New York

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
(Data applied for)

ISBN 0 19 850166 8

Typeset by the authors
Printed in Great Britain by
Bookcraft (Bath) Ltd
Midsomer Norton, Avon

PREFACE

Imagine that a perfectly self-similar fractal set is generated by a simple iteration and then subjected to bending, twisting, breaking, or corrosion. What remains of the initial structure?

Self-similarity is by now a familiar concept. One can repeat combinatorial recipes over and over again to obtain typically "fractal" limiting sets which look exactly the same at all scales and locations. Examples are commonly known and much studied, but often in a way which depends on the particular structure in question. Much less has been attempted by way of general definitions that would incorporate a broad range of phenomena into a single language.

We propose a concept of self-similarity here called *BPI* ("big pieces of itself"). The precise definition is given in the first chapter, and some characterizations and consequences are provided in Chapter 6. Roughly speaking, the definition asks that inside any pair of balls in the space there be pieces of *substantial proportion* which look almost alike in the sense of bilipschitz equivalence (i.e., up to bounded distortion of distances). This provides a framework in which to talk about "all" spaces with some self-similarity, and to make comparisons between them. It also provides a language in which to consider crucial features of a geometry without regard to accidents of a particular realization.

We do not mean to suggest that this framework is definitive – nothing ever will be – but it seems to be rather rich and flexible, and potentially opens up many interesting new opportunities for investigation.

In addition to the usual fractals one should keep more exotic spaces in mind. Examples of BPI spaces come from Heisenberg groups, asymptotic geometry of finitely generated groups, and constructions with doubling measures. (See Chapters 2, 4, and 16.) Doubling measures arise from Riesz products, Riemann surfaces, or more naive probabilistic considerations (as discussed in Chapter 16). A basic problem is to decide whether there is always something like a combinatorial pattern, a group, or a dynamical system behind any given BPI set.

Even in the case of perfectly self-similar sets it can be difficult to "see" precise combinatorial structure through primitive considerations of measure and metric, and this is a basic issue that we seek to address. It is still more problematic in the presence of singularities and distortions.

There are tools available for detecting approximately Euclidean behavior in sets despite singularities and distortions, through the notions of *rectifiability* and *uniform rectifiability*. (See Chapter 3 below.) The BPI condition can be seen as an extension of uniform rectifiability in which a simple model (like Euclidean geometry) is not provided in advance.

One would like to say when two different sets have essentially the same kind of structure. For this purpose there is a notion of *BPI equivalence*, defined in the first chapter and developed in Chapter 7. This condition is weaker than bilipschitz equivalence. (Think of having parts of your self-similar set break off.) BPI equivalence is an equivalence relation, and uniformly rectifiable sets are those BPI sets which are BPI equivalent to a Euclidean space. To ask about the existence of special rules or patterns behind the structure of a BPI space one can look for BPI-equivalent models which are more perfectly self-similar.

There are also natural relations between BPI spaces, of one space being more primitive than another. In Chapter 11 we discuss the concept of one BPI space "looking down" on another. BPI equivalence implies "look-down" equivalence (also discussed in Chapter 11), but the converse remains open. Some geometric consequences of one space looking-down on another are derived in Chapter 12, and a variety of special situations are discussed in Chapters 11 and 13. Related examples are provided in Chapter 14.

Instead of looking at individual BPI sets we can also think about *BPI geometries*, meaning equivalence classes of BPI sets. This leads to many natural questions, e.g., how many BPI geometries are there, how can they be deformed, to what extent are there plenty of continuous deformations or mostly a kind of discreteness, how does the ordering induced by looking down behave, etc. Special cases are treated in Chapters 15, 16, and 17.

We do not manage to go very far in this book towards definitive answers for any of these questions. There are many examples and concepts and basic facts, but no crisp theorems. The subject remains a wilderness, with no central zone, and many paths to try.

The lack of main roadways is also one of the attractions of the subject. It enjoys diverse connections with other aspects of geometry and analysis, including geometric measure theory, real analysis, bilipschitz and quasiconformal mappings, asymptotic behavior of groups and manifolds, and dynamical systems. We hope that the present text will be accessible to researchers in many areas, and also to people interested in entering the field. Many open problems emerge which are largely untouched, even for special cases and examples, and the general language leads to new perspectives and questions about classical constructions.

The prerequisites for this book are relatively modest, and consist mainly of basic knowledge of metric spaces and measure theory. While additional expertise would sometimes be helpful, basic definitions are recalled when needed, with adequate references to provide the reader with enough information to find out or work out what is needed at the moment. The reader may find the exposition [Se8] a helpful introduction to related aspects of geometry and analysis.

Much of this work was done during visits of the second author to the Institut des Hautes Etudes Scientifiques, to which the authors are very grateful. The second author was partially supported by the U.S. National Science Foundation.

CONTENTS

1 Basic definitions 1

2 Examples 5
2.1 Euclidean spaces 5
2.2 The snowflake functor 5
2.3 Cantor sets 6
2.4 Other fractals 7
2.5 A general procedure 9
2.6 Limit sets of discrete groups 12

3 Comparison with rectifiability 13
3.1 Rectifiable sets in $\mathbf{R}^n$ 13
3.2 Uniform rectifiability 14

4 The Heisenberg group 16

5 Background information 19
5.1 Extending Lipschitz functions 19
5.2 A covering lemma 19
5.3 Ahlfors regular spaces 19
5.4 Assouad's embedding theorem 21
5.5 Dyadic cubes 22
5.6 Semi-regularity 24

6 Stronger self-similarity for BPI spaces 26
6.1 The basic result 26
6.2 A corollary 32
6.3 Regular subsets 32
6.4 Some remarks about EAC 33

7 BPI equivalence 35
7.1 Basic facts 35
7.2 BPI equivalence and uniform rectifiability 36
7.3 A strengthening of BPI equivalence 40
7.4 Mappings with big bilipschitz pieces 49

8 Convergence of metric spaces 52
8.1 Hausdorff convergence 52
8.2 Convergence in Euclidean spaces 52
8.3 Convergence of mappings 53
8.4 Convergence of spaces 54

8.5 Convergence of mappings between spaces 59
8.6 Convergence of measures 61
8.7 Limits of subsets and their measures 63
8.8 Smooth sets 67

9 Weak tangents 71
9.1 The definition 71
9.2 First facts 72
9.3 Limits of BPI spaces 73
9.4 Weak tangents of subsets 75
9.5 Comparisons with rectifiability 77
9.6 BPI spaces which are not BPI equivalent 80
9.7 Weak tangents of mappings 81
9.8 Weak tangents of measures 82

10 Rest stop 84

11 Spaces looking down on other spaces 85
11.1 Definitions and basic facts 85
11.2 Mappings defined everywhere 86
11.3 Cantor sets to Euclidean spaces 87
11.4 Looking down from Euclidean spaces 88
11.5 Euclidean and Heisenberg geometries 89
11.6 A Cantor set with sliding 90
11.7 Iterating patterns with cubes 92
11.8 An observation about snowflakes 93
11.9 Looking down between Cantor sets 94
11.10 Remarks 101

12 Regular mappings 102
12.1 The definition and basic facts 102
12.2 Regular mappings as weak tangents 104
12.3 Looking down from $\mathbf{R}^n$ 109
12.4 Measure-preserving weak tangents 110
12.5 Measure-preserving mappings 118
12.6 Spaces not looking down on each other 119

13 Sets made from nested cubes 122
13.1 Preliminary notions 122
13.2 Convergence of families of cubes 124
13.3 Weak tangents of families of cubes 125
13.4 Mappings 127
13.5 Subsets that block connectedness 128
13.6 Going further 131

14 Big pieces of bilipschitz mappings 136
14.1 Introduction 136

14.2 Some examples 137
14.3 Possibilities for Lipschitz mappings 141
14.4 A stronger example 143

15 Uniformly disconnected spaces 156
15.1 Introduction 156
15.2 Spaces of dimension ≤ 1 157
15.3 Ultrametrics 161
15.4 Uniformization 162
15.5 Making regular mappings 166

16 Doubling measures and geometry 172
16.1 The definition 172
16.2 Deformations of geometry 173
16.3 Uniformization revisited 176
16.4 Doubling measures on Cantor sets 176
16.5 Reasonably self-similar pairs 180
16.6 Riesz products 182
16.7 Riemann surfaces 184
16.8 Remarks 187

17 Deformations of BPI spaces 189

18 Snapshots 194

19 Some sets that are far from BPI 197
19.1 The basic construction 197
19.2 Small variations on the theme 202

20 A few more questions 205

References 207

Index 211

1

BASIC DEFINITIONS

In this chapter we define the notions of *BPI spaces* and *BPI equivalence*, which are central to our study of self-similar geometry.

For the record, a *metric space* is a nonempty set M together with a metric $d(x,y)$, which is to say a symmetric nonnegative function on $M \times M$ that vanishes exactly when $x = y$ and satisfies the triangle inequality

$$d(x,z) \leq d(x,y) + d(y,z) \tag{1.1}$$

for all $x, y, z \in M$. As usual, we write $B(x,r)$ for the open ball in M with center x and radius r, and $\overline{B}(x,r)$ for the closed ball. We shall write $B_M(x,r)$ when we need to make M explicit. We denote by $\operatorname{diam} E$ the diameter of a subset E of M, which is defined by $\operatorname{diam} E = \sup\{d(x,y) : x, y \in E\}$.

Next we define a weak notion of *homogeneity* for a metric space in terms of the distribution of its *mass*.

Definition 1.1 (Ahlfors regularity) *A metric space $(M, d(x,y))$ is said to be* (Ahlfors) regular of dimension d *(or simply* regular*) if it is complete, has positive diameter, and if there is a constant $C > 0$ so that*

$$C^{-1}\, r^d \leq H^d(B(x,r)) \leq C\, r^d \tag{1.2}$$

for all $x \in M$ and $0 < r \leq \operatorname{diam} M$, where H^d denotes d-dimensional Hausdorff measure on M.

Recall that d-dimensional Hausdorff measure H^d is defined as follows. Given $A \subseteq M$ and $\delta > 0$ we set

$$H^d_\delta(A) = \inf\{\textstyle\sum_j (\operatorname{diam} E_j)^d : \{E_j\} \text{ is a sequence of sets in } M \text{ which covers } A \text{ and satisfies } \operatorname{diam} E_j < \delta \text{ for all } j\} \tag{1.3}$$

and $H^d(A) = \lim_{\delta \to 0} H^d_\delta(A)$. This defines an outer measure on M which is additive when restricted to Borel sets. A good general reference for Hausdorff measures is provided by [Ma].

A set A is said to have Hausdorff dimension d if $H^s(A) = 0$ when $s > d$ and $H^s(A) = \infty$ when $s < d$. Ahlfors regularity determines the Hausdorff dimension d, and it does so in a uniform and scale-invariant manner.

In practice the following observation makes it easier to check and use Ahlfors regularity.

Lemma 1.2 *Let $(M, d(x,y))$ be a complete metric space, and suppose that μ is a Borel measure on M with the property that there are positive constants K and d such that*

$$K^{-1} r^d \leq \mu(B(x,r)) \leq K r^d \tag{1.4}$$

for all $x \in M$ and $0 < r \leq \operatorname{diam} M$. Then $(M, d(x,y))$ is regular with dimension d, and there is a constant C depending only on K and d so that $C^{-1} \mu(E) \leq H^d(E) \leq C \mu(E)$ for all Borel sets $E \subseteq M$.

We shall explain this in Section 5.3.

The next definition will be used to make comparisons between metric spaces.

Definition 1.3 (Lipschitz and bilipschitz mappings) *Let $(M, d(x,y))$ and $(N, \rho(u,v))$ be metric spaces. A mapping $f : M \to N$ is said to be* Lipschitz *if there is a constant C such that*

$$\rho(f(x), f(y)) \leq C\, d(x,y) \tag{1.5}$$

*for all $x, y \in M$. We might say that f is C-*Lipschitz *to make the constant explicit. We say that f is* bilipschitz *if there is a C so that*

$$C^{-1} d(x,y) \leq \rho(f(x), f(y)) \leq C\, d(x,y) \tag{1.6}$$

*for all $x, y \in M$, and again we might say that f is C-*bilipschitz *to be more explicit. Two spaces are said to be* bilipschitz equivalent *if there is a bilipschitz mapping of one onto the other.*

Thus Lipschitz mappings do not expand distances very much, while bilipschitz mappings do not expand or contract distances by more than a bounded factor. If $f : M \to N$ is K-Lipschitz, then

$$H^d(f(A)) \leq K^d H^d(A) \tag{1.7}$$

for all $A \subseteq M$. This follows directly from the definitions. Note that the two uses of H^d here are not quite the same, since one refers to M, the other to N. Similarly, if f is K-bilipschitz, then

$$K^{-d} H^d(A) \leq H^d(f(A)) \leq K^d H^d(A). \tag{1.8}$$

The next definition introduces terminology which is convenient but not standard.

Definition 1.4 (Conformally bilipschitz mappings) *Given a pair of metric spaces $(M, d(x,y))$ and $(N, \rho(u,v))$ and a mapping $f : M \to N$ between them, we say that f is C-*conformally bilipschitz *if there is a $\lambda > 0$ such that*

$$C^{-1} \lambda\, d(x,y) \leq \rho(f(x), f(y)) \leq C \lambda\, d(x,y) \tag{1.9}$$

for all $x, y \in M$. We call λ the scale factor. *(In practice we may say that f is C-*conformally bilipschitz with scale factor *λ.)*

Of course a conformally bilipschitz mapping is bilipschitz, but with a constant that depends on the scale factor λ. The reason for introducing this terminology is that we shall often try to compare pieces of metric spaces which lie within balls of very different radii, and we shall want uniform estimates which make sense independently of the choice of radii.

Note that the composition of two conformally bilipschitz mappings is conformally bilipschitz, with suitable behavior the constants.

The following is our basic notion of "self-similarity". It asks that for any pair of balls in the given space there be substantial subsets inside them which look approximately the same in terms of conformal bilipschitz equivalence.

Definition 1.5 (BPI spaces) *A metric space $(M, d(x,y))$ is a* BPI space *("big pieces of itself") if it is regular of some dimension d and if there exist constants $\theta, C > 0$ so that for each pair of balls $B(x_1, r_1)$ and $B(x_2, r_2)$ in M with $0 < r_1, r_2 \leq \operatorname{diam} M$ there is a closed set $A \subseteq B(x_1, r_1)$ with $H^d(A) \geq \theta r_1^d$ and a mapping $h : A \to B(x_2, r_2)$ which is C-conformally bilipschitz with scale factor r_2/r_1.*

Note that there is a uniform lower bound for $r_2^{-d} H^d(h(A))$ under the conditions above.

For us a ball $B(x,r)$ has always finite radius. It is sometimes convenient to write ranges of radii that theoretically permit $r = \infty$, as in the definition above, but whenever we write $B(x,r)$ we implicitly restrict ourselves to finite radii.

By asking only that pairs of balls contain substantial subsets which look alike, rather than whole replicas of each other, we have allowed a greater role for measure theory in this notion of self-similarity than is customary. We have increased the possibility for repetition, distortion, rupture, and clutter. The following notion of *BPI equivalence* takes these possibilities into account. It says that two BPI spaces are considered to be equivalent if for any ball in the first space and any ball in the second we can find substantial subsets which are conformally bilipschitz equivalent to each other.

Definition 1.6 (BPI equivalence) *Two BPI metric spaces $(M, d(x,y))$ and $(N, \rho(u,v))$ with the same dimension d are said to be* BPI equivalent *if there exist constants $K, \alpha > 0$ so that for any $x \in M$, $0 < r \leq \operatorname{diam} M$, $u \in N$, and $0 < t \leq \operatorname{diam} N$ there exists a closed subset A of $B_M(x,r)$ with $H^d(A) \geq \alpha r^d$ and a K-conformally bilipschitz mapping $\phi : A \to B_N(u,t)$ with scale factor t/r.*

We shall see that BPI equivalence defines an equivalence relation on BPI spaces, and that two BPI spaces are BPI equivalent as soon as they contain a pair of subsets of positive measure which are bilipschitz equivalent. We can think of equivalence classes of BPI spaces as representing the same basic "geometry".

In Chapter 11 we shall discuss other relations between BPI spaces connected to the idea that one space is more "primitive" than another.

Our first task will be to give some examples BPI geometries that arise naturally in mathematics. Afterwards we establish basic facts about BPI spaces and

BPI equivalence. In Chapters 8 and 9 we develop the themes of compactness and convergence of metric spaces in connection with the BPI property and BPI equivalence.

2

EXAMPLES

2.1 Euclidean spaces

The first example is Euclidean space $\mathbf{R}^n$ equipped with the Euclidean metric. This is certainly Ahlfors regular of dimension n (Hausdorff measure coincides with Lebesgue measure in this case), and it is also BPI, because all balls are conformally isometric to each other.

A union of two n-planes in some $\mathbf{R}^m$ is also regular and BPI. In $\mathbf{R}^n$ let $\{x_i\}_i$ be a sequence of points such that $|x_i| = 2^{-i}$ for each i. Then

$$E = \{0\} \cup \bigcup_i \overline{B}(x_i, 2^{-i}) \tag{2.1}$$

is a regular set of dimension n which is BPI.

These two examples illustrate the kind of singularities, or failures of true self-similarity, that the BPI condition allows. On the other hand all of these sets are BPI equivalent to each other. This also illustrates the idea that one might have a BPI space that is not completely beautiful but for which there is a much nicer space that represents the same BPI equivalence class.

2.2 The snowflake functor

Let $(M, d(x,y))$ be a metric space. Given a real number $0 < s < 1$ consider the space $(M, d(x,y)^s)$. This is still a metric space, as is well known. We call this transformation of $(M, d(x,y))$ into $(M, d(x,y)^s)$ the *snowflake functor*.

It is important here that $s \leq 1$; if $s > 1$, then in general $D(x,y) = d(x,y)^s$ is only a *quasimetric*, which means that the triangle inequality should be weakened to

$$D(x,z) \leq K\,\{D(x,y) + D(y,z)\} \tag{2.2}$$

for some $K > 0$ and all $x, y, z \in M$. Conversely it turns out that if $D(x,y)$ is a quasimetric, then there exists a metric $d(x,y)$ and constants $C, s \geq 1$ such that

$$C^{-1}\, d(x,y)^s \leq D(x,y) \leq C\, d(x,y)^s \tag{2.3}$$

for all x and y. This is essentially the content of the proof of Theorem 2 in [MS].

The term *snowflake* here stems from the classical examples of snowflake curves in the plane, e.g., the Von Koch snowflake. When these constructions are made in a sufficiently regular manner they turn out to be bilipschitz equivalent to the standard line, interval, or circle (as appropriate), but with the metric deformed as in the snowflake functor for a suitable power s. The snowflake functor simply makes this idea abstract and applicable to any metric space. One should think of the classical snowflake pictures though, and their very crinkled shapes.

If $(M, d(x,y))$ is regular of dimension d, then $(M, d(x,y)^s)$ is regular of dimension d/s. If $(M, d(x,y))$ is BPI, then so is $(M, d(x,y)^s)$. This is an easy consequence of the definitions.

2.3 Cantor sets

Rather than discussing Cantor sets as explicitly constructed subsets of Euclidean spaces, we shall prefer to work with them in a more symbolic and abstract manner. Let F be a finite set with $k \geq 2$ elements. Let F^∞ denote the set of sequences $\{x_i\}_{i=1}^\infty$ with $x_i \in F$ for each i. F^∞ is our Cantor set, which we might call the *k-Cantor* set to be precise. It is a compact Hausdorff space when one gives it the usual product topology coming from the discrete topology on each factor of F.

We want our Cantor sets to be metric spaces, though. Given $x = \{x_i\}$ and $y = \{y_i\}$ in F^∞, let $L(x,y) = l$ if l is the largest integer such that $x_i = y_i$ when $1 \leq i \leq l$, and set $L(x,y) = \infty$ when $x = y$. ($L(x,y) = 0$ is allowed.) Given $0 < a < 1$ set

$$d_a(x,y) = a^{L(x,y)}, \tag{2.4}$$

where the right side is interpreted to be 0 when $L(x,y) = \infty$. This defines a metric on F^∞, and in fact it is an *ultrametric*, which means that

$$d_a(x,z) \leq \max\{d_a(x,y), d_a(y,z)\} \tag{2.5}$$

for all x, y, z. These metrics are all related by the snowflake functor, i.e., $d_a(x,y) = d_b(x,y)^s$ when $a = b^s$.

There is a natural probability measure μ on F^∞, defined as follows. Let ν denote the probability measure on F that assigns mass $1/k$ to each element of F. We take μ to be the infinite product of copies of ν.

Given $x = \{x_i\} \in F^\infty$ and a nonnegative integer j, the ball $B_a(x, a^j)$ with respect to the metric $d_a(x,y)$ consists of the points $y = \{y_i\} \in F^\infty$ such that $y_i = x_i$ when $i \leq j$. Thus

$$\mu(B_a(x, a^j)) = k^{-j}. \tag{2.6}$$

Using this it is easy to check that $(F^\infty, d_a(x,y))$ is regular of dimension d, where d is determined by the equation

$$a^d = k^{-1}. \tag{2.7}$$

Notice that we get all possible $0 < d < \infty$ by varying a in $(0,1)$.

It is easy to see that each space $(F^\infty, d_a(x,y))$ is BPI. Again, like Euclidean spaces, these spaces are much better behaved than that, they have much more symmetry, but one can modify them and still get BPI spaces.

These symbolic Cantor sets are bilipschitz equivalent to the standard self-similar Cantor sets constructed in Euclidean spaces. The original middle-thirds

set, for instance, is bilipschitz equivalent to $(F^\infty, d_a(x,y))$ with $k = 2$ and $a = 1/3$. This is easy to check.

We have described here the simplest and most classical geometries to put on Cantor sets, but they are not the only possibilities. Instead of defining a metric on F^∞ as above, using a single number a, we can use a function $a : F \to (0,1)$, as follows. Let $x = \{x_i\}$ and $y = \{y_i\}$ in F^∞ be given, and let $L(x,y)$ be as above. Define the distance between x and y to be

$$\prod_{i=1}^{L(x,y)} a(x_i). \tag{2.8}$$

This agrees with the previous definition when a is constant. One can check that this defines a metric on F^∞, in fact an ultrametric, in such a way that F^∞ is again Ahlfors regular and BPI.

There is a general recipe for "twisting" the geometry of Cantor sets that we shall discuss in detail later, and which will be shown to be exhaustive in a certain sense (see Chapter 16). It is not clear in general which of these geometries is BPI. Bilipschitz and BPI equivalence turn out to be tricky even in very self-similar situations.

2.4 Other fractals

This time we do construct our sets inside of the plane $\mathbf{R}^2$ for simplicity. Let Q denote the unit square $[0,1] \times [0,1]$ in $\mathbf{R}^2$. We can decompose Q into 9 (closed) subsquares Q_i in the obvious manner, so that the Q_i's have disjoint interiors and sidelength 1/3.

Now suppose that we take Q and replace it with the union of all the Q_i's except the one in the center of Q. This produces a set K_1 which is the union of 8 squares. We can now apply this process to each of these squares, and then repeat the procedure indefinitely, to get a sequence of compact sets $\{K_j\}$ in the plane, where each K_j consists of 8^j squares of sidelength 3^{-j}, and where these squares have disjoint interiors. Set $K = \bigcap_j K_j$. This defines a compact subset of $\mathbf{R}^2$ which is called *the Sierpinski carpet* (see Figure 2.1).

There is a natural probability measure μ on K, which can be defined as follows. For each j let μ_j denote the probability measure on K_j which is just $(9/8)^j$ times Lebesgue measure. Then $\{\mu_j\}$ converges in the usual weak topology to a probability measure μ on K. One can also realize K as the continuous image of an 8-Cantor set F^∞ in the obvious manner, and obtain the measure from the one on the Cantor set. One of the main properties of our measure μ on K is that if S is one of the 8^j squares of which K_j is composed, then the piece of K inside S has mass 8^{-j}. It is not hard to check that K is regular of dimension d, where $3^d = 8$, and that K is BPI. (Each $S \cap K$ is an exact replica of the original.)

We can modify this construction to get a fractal tree instead. For this we replace the initial square Q with the union of the five squares among the Q_i's which consist of the four in the corners and the one in the middle. This also gives

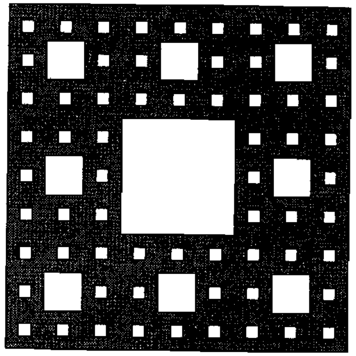

FIG. 2.1. The Sierpinski carpet

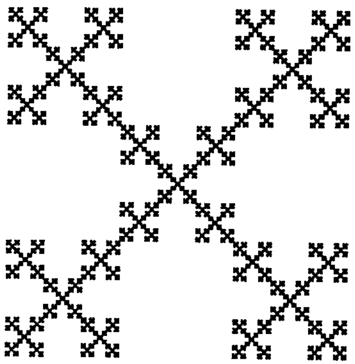

FIG. 2.2. A fractal tree

an Ahlfors regular BPI set of dimension d determined by $3^d = 5$. In this case the set is a fractal tree (see Figure 2.2).

Another example is given by the *Sierpinski gasket*. In this case we start with an equilateral triangle, we subdivide it into 4 subtriangles of equal size, and we keep the three in the corners and remove the one in the middle. This again leads to an Ahlfors regular BPI set (see Figure 2.3)

These examples are different in their geometry in terms of the number of ways that there are to connect points within the sets by curves of finite length. There are a lot more curves of finite length in the Sierpinski carpet than in the other two examples. There are many more rectifiable curves in Euclidean spaces of dimension > 1 than in the Sierpinski carpet, even if one accounts for the difference in the dimension. In Cantor sets there are no nontrivial curves whatsoever. In snowflakes there are plenty of nontrivial curves but no nontrivial rectifiable curves.

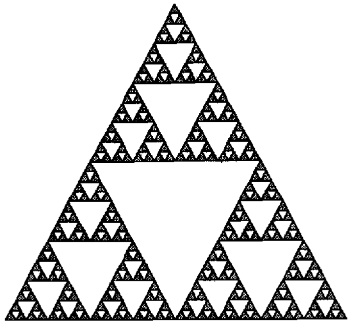

FIG. 2.3. The Sierpinski gasket

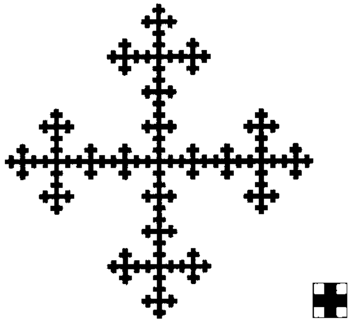

FIG. 2.4. Another fractal tree

2.5 A general procedure

Let Q denote the unit cube in $\mathbf{R}^n$. ("Cubes" should be interpreted as closed cubes here.) Fix an integer M, and subdivide Q into M^n subcubes of size M^{-1} in the obvious manner. Let $\mathcal{S}$ denote the collection of these subcubes.

Let $\mathcal{R}$ be a proper subset of $\mathcal{S}$. We think of $\mathcal{R}$ as being a "rule" for defining a (probably fractal) subset of Q. That is, we start with Q, we replace it with the cubes in $\mathcal{R}$, we identify each with Q in order to replace it with a collection of subcubes like $\mathcal{R}$, etc. If $\mathcal{R}$ has k elements, then this procedure produces a decreasing sequence of compact sets $\{A_j\}$, where A_j consists of k^j cubes of size M^{-j}. We take A to be the intersection of all the A_j's.

This procedure includes traditional Cantor sets and the examples of the previous section. Additional examples and their rules are given in Figures 2.4-2.8.

The limiting set A obtained in this way is Ahlfors regular of dimension d, where $M^d = k$. This is not hard to check. One can do the following, for instance. Let μ_j denote the probability measure on $\mathbf{R}^n$ supported on A_j which is uniformly distributed on A_j (with respect to Lebesgue measure). Thus μ_j assigns measure k^{-j} to each of the constituent cubes in A_j. It is not very difficult to see that $\{\mu_j\}$ converges weakly on $\mathbf{R}^n$ to a probability measure μ. If Q is one of the constituent cubes of A_j, then one can easily check that

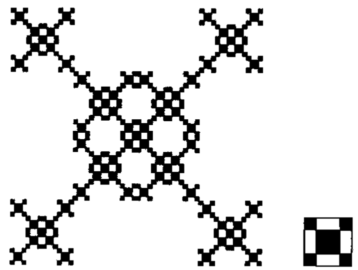

FIG. 2.5. More cubes in the middle

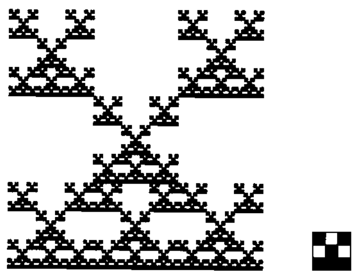

FIG. 2.6. More cubes near the bottom

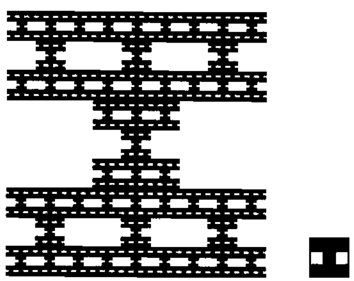

FIG. 2.7. Fractal train tracks

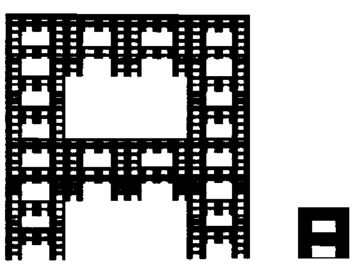

FIG. 2.8. The gates to knowledge

$$k^{-j} \leq \mu(Q) \leq C(n)\, k^{-j}, \tag{2.9}$$

where $C(n)$ depends only on n. In fact one can be careful and get $C(n) = 1$, but in related situations one has to think more about the possibility of mass accumulating on common faces of cubes.

This set is clearly BPI. Indeed, the set look the same modulo translations and dilations in each of the constituent cubes in A_j.

Fix M and k, and consider the collection of rules that we have described as defining a small universe of their own. This collection is typically not so small after all, since we are talking about choosing k elements from a set of cardinality M^n, and the number of ways of doing this can be quite large. Even with modest choices of the parameters there can be substantial variety in the phenomena that can occur, as indicated by the examples shown in this section and in the preceding one.

Let us think of this as a small universe of BPI sets and ask about the way that they compare with each other. Many rules will give practically the same result. In some cases this will be true in a very strong way, for instance if two rules are just translations of each other. In other cases it will be true in a more flexible manner. If a cube in a rule is isolated from the others, and it is moved to another position that is isolated from the others, then the change is not very important. There are many ways to get a set which amounts to a standard Cantor set. Even if cubes are not isolated from each other, it will frequently make no difference whether a chain of them points to the right or to the left.

One can formulate natural notions of combinatorial equivalence. A basic question then is to what extent the combinatorics can be recaptured from the metric geometry. One can make combinatorial criteria for two sets generated by these rules to be bilipschitz equivalent, but what about the converse? How much of the combinatorial structure of the rules has to be the same if the resulting sets are bilipschitz equivalent?

This is not so clear.

Here is a simple example. Take $n = 1$, $M = 5$, and $k = 3$. A rule consists of choosing three of the five equal pieces of the unit interval. Define one rule by taking the first, third, and fifth subintervals, and define a second rule by taking the first, fourth, and fifth. Consider the two sets that result from these two rules. Are they bilipschitz equivalent? We do not know. There is an obvious mapping from the first set onto the second which is Lipschitz but far from bilipschitz, but the absence of *any* bilipschitz mapping is not clear. We shall discuss this example more later, beginning in Section 11.6.

This example illustrates another point that we shall discuss later, the idea that one set can be more scattered than another, in the sense that there is a Lipschitz mapping from one onto the other. We shall discuss this further in Chapter 11. In our small universe of rules we see again that there can be simple combinatorial reasons for one of our sets to map onto another in a Lipschitz manner. For instance, if all the cubes in the rule are isolated from each other, then the resulting set can be mapped onto any other (in this universe) in a simple way. Again it is not clear what combinatorial conditions are necessary for the existence of such a mapping.

Thus in this small universe we can ask how the various rules are related combinatorially and in terms of the geometry of the sets that they generate. The collections of rules are rich enough to allow for some nontrivial relationships.

In this class of examples we have a nice illustration of the ideas of comparing BPI geometries, and having families of BPI geometries, rather than looking at BPI spaces as isolated objects.

One can also think about what happens when one builds sets by applying one rule after another, but maybe not the same rule each time. With a fixed choice of M and k, though, to have Ahlfors regularity anyway. This provides a way to build a large collection of sets in a way that one can describe – "parameterize" – and perhaps even analyze. It is natural to ask whether a set which is constructed in this manner and which has BPI should have some combinatorial structure behind it. We do not know the answer to this question. One can consider this question more generally for BPI spaces, but it is more crisp here.

We shall discuss these constructions further in Chapter 13.

2.6 Limit sets of discrete groups

For simplicity of exposition we shall permit ourselves to use more specialized terminology in this section without much explanation. We hope that the reader not acquainted with these topics will still get a flavor of the examples, for which we shall provide references.

Let $\mathbf{H}^{d+1}$ denote d-dimensional hyperbolic space, viewed as the unit ball in $\mathbf{R}^{d+1}$ equipped with the familiar metric. Let Γ be a discrete group of isometries on $\mathbf{H}^{d+1}$, which are Möbius transformations that can be extended also to the unit sphere $\mathbf{S}^d$. Under the "convex cocompact" assumption, Sullivan [Su1] has shown that the limit set of Γ is an Ahlfors regular set. (See also [Su2], especially Theorem 10.) One can use the convex cocompactness to show that the group visits all scales and locations on the limit set (this is better imagined in terms of the orbit of γ in hyperbolic space), in such a way that the geometry of Möbius transformations implies that the limit set is BPI. (In this case there is not a question of finding clever subsets of the space, one can just work with balls themselves.)

Sullivan's work was extended by Coornaert [Co] to groups acting on spaces of negative curvature in the sense of Gromov. In this case one has to be careful about what plays the role of the sphere at infinity, but there is a version of this. Under a convex-cocompactness assumption, Coornaert also establishes Ahlfors regularity of the limit set in the space at infinity. If we have understood this correctly, this limit set is BPI, and this method applies for instance to the case case of any compact negatively curved Riemannian manifold, acted on by the group of deck transformations. In this case the group is cocompact, and the limit set is the whole space at infinity.

3

COMPARISON WITH RECTIFIABILITY

3.1 Rectifiable sets in $\mathbf{R}^n$

Let d and n be integers, $0 < d < n$, fixed.

Definition 3.1 (Rectifiable sets) *A subset E of $\mathbf{R}^n$ is said to be rectifiable if there exists a sequence of subsets $\{E_j\}$ of E, each bilipschitz equivalent to a subset of $\mathbf{R}^d$, such that $H^d(E \setminus \bigcup_j E_j) = 0$.*

Strictly speaking we should say something like "rectifiable of dimension d", but it is simpler to let the dimension be understood implicitly.

Rectifiable sets are approximately Euclidean in their geometry. They need not be BPI, because no quantitative estimates are required, but the principle is similar. If E is measurable and rectifiable, then almost all points of E lie in an E_j, and therefore almost all points are points of density for some E_j. This means that there is a thick copy of a piece of a Euclidean space near the given point.

One of the nice features of rectifiability is that it is a very stable condition. Instead of asking that the E_j's be bilipschitz equivalent to subsets of $\mathbf{R}^d$ we could ask that they be Lipschitz images of subsets of $\mathbf{R}^d$, and we would get an equivalent condition. If instead we asked that they lie on C^1 manifolds of dimension d we would also get an equivalent condition. At bottom these equivalences come from the fact that Lipschitz functions on Euclidean spaces are differentiable almost everywhere. A convenient consequence of this fact is that a Lipschitz function f on a Euclidean space can be modified on a set of arbitrarily small measure to get a C^1 function.

These rigidity properties of Lipschitz mappings lead to stability properties of geometry, i.e., approximately Euclidean geometry in a weak sense implies approximately Euclidean behavior that appears to be much stronger. Covering almost everywhere by Lipschitz images of $\mathbf{R}^d$ implies covering almost everywhere by C^1 manifolds. Similarly, there are theorems which characterize rectifiability in terms of the requirement that for almost all points p in E the part of E near p lies mostly in a cone about a d-plane, and then this condition implies the apparently stronger property that the part of E near almost any point lies asymptotically close to a d-plane.

These stability properties help to make for nontrivial theorems about rectifiable sets. One can then ask whether similar phenomena hold for other kinds of geometries, such as the snowflakes and Cantor sets and fractals described in Chapter 2. The basic answer seems to be no, because of the absence of any rigidity theorems akin to differentiability almost everywhere. We say "seems" because it is not always clear how to formulate such rigidity theorems. Although the most obvious formulations fail, there does appear to be some room for structure. For instance, one might expect that there is more structure to a Sierpinski carpet because it has lots of rectifiable curves, even if there is not so much structure as for Euclidean spaces.

At any rate, in general one cannot expect results of the type "bounded equivalence implies asymptotic equivalence", as one has in the context of rectifiability and the differentiability of Lipschitz functions almost everywhere.

One reason for defining the concept of BPI sets is to provide a setting where we can ask ourselves such questions. Even if there are not so many rigidity results to be found, the lack of rigidity ought to have some interesting consequences for geometry. If we have a geometry which is not very rigid, then it should be easier to deform it, for example. One of the points of BPI is to try to have a language for formulating this idea in a precise way. More generally, one would like to be able to talk about different geometries at once, and moving between them.

Although this kind of incredibly strong rigidity does not work for most other spaces, there are other types of rigidity to look for, as we shall see. One can say that Euclidean geometry is incredibly un-rigid in terms of being able to map spaces into $\mathbf{R}^n$ easily, while it is very rigid in the opposite way, in the difficulty with mapping $\mathbf{R}^n$ into other spaces without having some very special structure. One can hope that there is some nontrivial "total" rigidity for other spaces, that the extent to which mappings on a given space do not have to have special structure is balanced by the special structure that is required for mappings into the space.

See [F1, Fe, Ma] for more information about rectifiability.

3.2 Uniform rectifiability

Uniform rectifiability is a variant of the notion of rectifiability which comes with uniform and scale-invariant estimates. Again let d and n be fixed integers with $0 < d < n$.

Definition 3.2 (Uniform rectifiability) *A subset E of $\mathbf{R}^n$ is said to be* uniformly rectifiable *if it is Ahlfors regular (of dimension d) and if there exist constants $M, \theta > 0$ so that for each $x \in E$ and $0 < r \leq \operatorname{diam} E$ there is a subset A of $E \cap B(x,r)$ such that*

$$H^d(A) \geq \theta\, r^d \tag{3.1}$$

and

$$A \text{ is } M\text{-bilipschitz equivalent to a subset of } \mathbf{R}^d. \tag{3.2}$$

This condition implies ordinary rectifiability (this is not to hard to show), but it is much stronger, because of the uniform bounds. This particular definition of uniform rectifiability should not be taken too seriously, in the sense that there are many equivalent conditions. See [D4, DS2, DS4, Se3] for more information. We could obviously extend the definition of uniform rectifiability to metric spaces

instead of just subsets of Euclidean spaces, but some results about uniform rectifiability do not make sense in the abstract setting, and some are just not known. (Similar considerations apply to ordinary rectifiability.)

Proposition 3.3 *Uniformly rectifiable sets are BPI.*

To prove this we use the following.

Lemma 3.4 *Suppose that A_1, A_2 are two measurable subsets of the unit ball $B(0,1)$ in $\mathbf{R}^d$, with $|A_i| \geq \delta$ for some $\delta > 0$. (Here $|A|$ denotes the Lebesgue measure of A.) Then there is a point $z \in B(0,2)$ such that*

$$|\tau_z(A_1) \cap A_2| \geq C^{-1}\delta^2, \tag{3.3}$$

where $\tau_z : \mathbf{R}^d \to \mathbf{R}^d$ is the translation $\tau_z(x) = x - z$, and where C can be taken to be the volume of $B(0,2)$.

Indeed, Fubini's theorem implies that

$$\int_{B(0,2)} \left(\int_{B(0,1)} \mathbf{1}_{\tau_z(A_1)}(y)\, \mathbf{1}_{A_2}(y)\, dy \right) dz \geq \delta \int_{B(0,1)} \mathbf{1}_{A_2}(y)\, dy \geq \delta^2, \tag{3.4}$$

where $\mathbf{1}_A$ denotes the characteristic function of A. Hence

$$\int_{B(0,1)} \mathbf{1}_{\tau_z(A_1)}(y)\, \mathbf{1}_{A_2}(y)\, dy \geq \frac{\delta^2}{|B(0,2)|} \tag{3.5}$$

for some $z \in B(0,2)$, as desired.

The proposition is an easy consequence of the lemma. Indeed the lemma says that two bounded subsets of a Euclidean space of definite size always have subsets of definite size which are isometric (one is the image of the other by a translation). There is an obvious extension of this to subsets of balls of different size (with dilations used too), and the proposition follows from this extended version.

Of course we (the authors) began this story with uniform rectifiability and the question of whether there was something like it for other geometries. As in the comments at the end of the preceding section, we cannot expect anything like uniform rectifiability in its full glory for general BPI geometries, too much structure is missing. But again BPI provides a setting in which to ask certain questions, and to look at how geometries behave differently when one does not have so much structure as in uniform rectifiability.