Introduction

This is a book on geometric measure theory. The main theme is the study of the geometric structure of general Borel sets and Borel measures in the euclidean n-space $\mathbf{R}^n$. There will be emphasis on "small irregular" sets having Lebesgue measure zero but being quite different from smooth curves and surfaces. Examples are Cantor-type sets, non-rectifiable curves having tangent nowhere, etc., in short, sets to which the general descriptive term fractal applies. An abundance of such sets comes from dynamical systems: Julia-sets for rational functions of one complex variable, etc. Very general curve- and surface-like objects are also studied extensively. These are rectifiable sets and measures. They include smooth curves and surfaces and share many of their geometric properties when interpreted in a measure-theoretic sense. They form an optimal class possessing such properties.

Many of the basic ideas developed here originate in the pioneering work done by Besicovitch [1], [4] and [5], by Federer [1], by Marstrand [1] and by Preiss [4]. Besicovitch laid down the foundations of geometric measure theory by describing to an amazing extent the structure of the subsets of the plane having finite one-dimensional Hausdorff measure (i.e. length). Federer extended Besicovitch's work to m-dimensional subsets of $\mathbf{R}^n$, m being an integer, and Marstrand analysed general fractals in the plane whose Hausdorff dimension need not be an integer. Preiss solved one of the most long-standing fundamental open problems, introducing and using effectively tangent measures.

Good introductory texts to the mathematical theory of fractals are the books of Edgar [1] and of Falconer [4], [16]. Closest to this text is Falconer [4]. The relation between this book and those of Falconer is roughly that we shall develop the general theory here beyond Falconer's books but we are not paying much attention to applications, except for the last two chapters, which Falconer does not deal with. Many of the topics discussed here are also treated in the extensive book of Federer [3], often in a more general form. Only Chapters 2 and 3 of Federer [3] are relevant to our present subject. Chapters 4 and 5 there are devoted to currents and their applications to the calculus of variations. This theory is based on rectifiability but we shall not consider it here. More recent texts on this extremely active area of geometric measure theory are L. Simon [1], Hardt and Simon [1] and Giusti [1]. A good survey on geometric measure theory is given in Federer [4]. The book of Morgan [1] serves as an excellent introductory text to many basic concepts and ideas. The books of Evans and Gariepy [1] and Ziemer [1] also deal with some parts of geometric measure theory, for example area and coarea theorems, sets of finite perimeter, which are not considered here. Taylor's obituary on Besicovitch, Taylor [1], is interesting in particular for the historical development of the theory.

Fractals and fractal measures arise in mathematics in many ways; for example in number theory via Diophantine approximation, in probability via Brownian motion and other stochastic processes, in dynamical systems as strange attractors, in complex analysis as limit sets of Kleinian groups, etc. We shall not pay much attention to these relations; discussions on them and further references can be found for example in Barnsley [1], Edgar [1], Falconer [4], [16], Mandelbrot [1] and Peitgen and Richter [1]. Mandelbrot [1] also uses fractals to model many physical phenomena. Computer simulation of fractal images is widely considered in Peitgen and Saupe [1] and Barnsley [1]. Tricot [6] works with many examples and concepts related to curves.

This book splits roughly into three parts. Chapters 1–7 give background in measure theory and develop the required tools and results, mainly in terms of Hausdorff measures and dimension. The second part consists of Chapters 8–14. There sets and measures are considered without dimensional restrictions. Thus this part applies to getting information about sets and measures whose dimension need not be an integer. In the last part, Chapters 15–20, we investigate integral dimensional sets and measures and the unifying concept there is rectifiability.

I shall now briefly describe the topic of each chapter. In Chapter 1 we set up much of the measure-theoretic terminology and notation to be used throughout the rest of the book. We shall mainly prove only the results that cannot be found in standard books of measure theory and real analysis. In Chapter 2 we prove covering theorems of Vitali and Besicovitch and use them to obtain a basic differentiation theorem for measures. In Chapter 3 we introduce and prove some properties of the natural invariant measures on the spaces of orthogonal transformations of $\mathbf{R}^n$ and of linear and affine m-dimensional subspaces of $\mathbf{R}^n$.

The main theme of Chapter 4 is the introduction of one of our basic tools, s-dimensional Hausdorff measures $\mathcal{H}^s$ and Hausdorff dimension, dim, although we also give a general construction leading to many other measures as well. We study several examples and briefly consider self-similar and related sets. In Chapter 5 we discuss other concepts of dimension and related measures, in particular Minkowski dimension, packing dimension and packing measures. In Chapter 6 we prove the basic density estimates for Hausdorff and packing measures. For instance, they say that if s is a positive number and A an $\mathcal{H}^s$ measurable subset of $\mathbf{R}^n$ with $\mathcal{H}^s(A) < \infty$, then at $\mathcal{H}^s$ almost all points $x \in A$,

$$(1) \qquad 2^{-s} \leq \limsup_{r \downarrow 0} (2r)^{-s} \mathcal{H}^s\big(A \cap B(x,r)\big) \leq 1,$$

where $B(x,r)$ is the closed ball with centre x and radius r.

Chapter 7 gives a brief treatment of Lipschitz maps. For example we prove Rademacher's theorem on their differentiability almost everywhere and a simple Sard-type theorem.

In Chapter 8 we introduce some potential-theoretic methods and concepts to study Hausdorff dimension, that is, we use the s-energies

$$I_s(\mu) = \iint |x-y|^{-s} \, d\mu x \, d\mu y$$

for Radon measures μ on $\mathbf{R}^n$ and the capacities related to them. We prove Frostman's lemma stating that a Borel set has positive s-dimensional Hausdorff measure if and only if it supports a non-zero Radon measure μ such that

$$(2) \qquad \mu(B(x,r)) \leq r^s \quad \text{for all } x \in \mathbf{R}^n \text{ and } r > 0.$$

Since (2) is closely related to the condition $I_s(\mu) < \infty$ this leads to a definition of the Hausdorff dimension in terms of capacities. In fact, these relations mean that a large part of this book could be interpreted as a study of geometric properties of Radon measures μ on $\mathbf{R}^n$ satisfying either (2) or the inequality $I_s(\mu) < \infty$. We shall also use Howroyd's new technique in general compact metric spaces to prove Frostman's lemma and the theorem on the existence of subsets with positive and finite Hausdorff measure inside a given set with infinite measure.

Chapter 9 studies how Hausdorff dimension transforms under orthogonal projections. The main results, essentially due to Marstrand, say that a given Borel subset of $\mathbf{R}^n$ with Hausdorff dimension s projects into a set of Hausdorff dimension s on almost all linear m-dimensional subspaces of $\mathbf{R}^n$ provided $s \leq m$. In the case $s > m$, the projections have generically positive m-dimensional measure. In Chapter 10 we show that such an s-dimensional set intersects "usually" $(n-m)$-dimensional affine subspaces of $\mathbf{R}^n$ in a set of Hausdorff dimension $\max\{0, s-m\}$. In both of these chapters we use a potential-theoretic approach. Thus we prove similar and sharper results for capacities and measures with finite s-energy.

The density theorems for Hausdorff measures of Chapter 6, such as the inequalities (1), give the first information as to how much measure we can expect to find in small balls. In Chapter 11 we find out more about how this measure is distributed in narrow cones. For example, if $n-1 < s \leq n$, $\alpha > 0$ and A is an $\mathcal{H}^s$ measurable subset of $\mathbf{R}^n$ with $\mathcal{H}^s(A) < \infty$, then at $\mathcal{H}^s$ almost all points $x \in A$

$$\limsup_{r \downarrow 0} r^{-s}\mathcal{H}^s\big(A \cap B(x,r) \cap C(\alpha,x)\big) \geq c(\alpha) > 0,$$

where $C(\alpha, x)$ is a cone with vertex x and opening angle α. Again we work with general Radon measures and their s-energies.

In Chapter 12 we bring in another effective tool to study Hausdorff dimension, capacities and energy-integrals; this is the Fourier transform. We develop some preliminary results and as an application give a simple proof of Falconer for estimating the Hausdorff dimension of distance sets. Other applications will be presented in Chapter 13 where we study the generic Hausdorff dimension of the intersection of two Borel sets A and B moving in $\mathbf{R}^n$. It turns out that

$$\dim(A \cap fB) \geq \dim A + \dim B - n \tag{3}$$

for many euclidean motions f, provided $\dim B > (n+1)/2$; this assumption may be superfluous. We also give conditions which guarantee that equality holds in (3).

In Chapter 14 we introduce the tangent measures in the sense of Preiss. They contain information about the local structure of a given Radon measure μ in a similar but often more complicated way as the derivative of a function tells us about the local behaviour. The tangent measures of μ at a point a consist of all non-zero locally finite weak limits of the sequences of measures

$$A \mapsto c_i\mu(r_iA + a) \quad \text{where } r_i \downarrow 0 \text{ and } 0 < c_i < \infty.$$

As the first application of tangent measures we prove Marstrand's theorem according to which for any non-integral number s there exists no non-zero Borel measure μ in $\mathbf{R}^n$ such that the positive and finite limit $\lim_{r\downarrow 0} r^{-s}\mu(B(x,r))$ would exist for μ almost all $x \in \mathbf{R}^n$.

Then we start the last, integral dimensional, part of the book. First, in Chapter 15 we define m-rectifiable sets as a natural and convenient generalization of nice m-dimensional surfaces, such as C^1 submanifolds, Lipschitz graphs, etc. They are sets which except possibly for a set of $\mathcal{H}^m$ measure zero lie on countably many C^1 submanifolds. We give a characterization of rectifiability in terms of the almost everywhere existence of approximate tangent planes. In Chapter 16 we continue the study of the tangential properties in connection with rectifiability in a more technical manner. This leads to a characterization of rectifiability using only "weak, rotating" tangent planes; the approximating plane is allowed to depend on the scale. We also formulate such results in terms of tangent measures. As side-products we derive information about the density and projection properties of rectifiable sets.

Chapter 17 discusses the theorem of Preiss characterizing rectifiability in terms of the existence of densities. The main part of this is the following statement: if m is a positive integer and μ is a Borel measure on $\mathbf{R}^n$ such that the positive and finite limit $\lim_{r\downarrow 0} r^{-m}\mu(B(x,r))$ exists for μ almost all $x \in \mathbf{R}^n$, then μ is m-rectifiable in the sense that there exist m-dimensional C^1 submanifolds $M_1, M_2, \ldots$ such that $\mu(\mathbf{R}^n \setminus \bigcup_{i=1}^{\infty} M_i) = 0$. The tangent measures play a fundamental role in the proof. The complete proof is very complicated and we shall give only parts of it and derive a weaker result.

Chapter 18 is mainly devoted to the proof of the fundamental theorem of Besicovitch and Federer characterizing rectifiability with projection properties. More precisely, let A be an $\mathcal{H}^m$ measurable subset of $\mathbf{R}^n$ with $\mathcal{H}^m(A) < \infty$. Then A meets every m-dimensional C^1 submanifold of $\mathbf{R}^n$ in a zero $\mathcal{H}^m$ measure if and only if $\mathcal{H}^m(P_V A) = 0$ for almost all orthogonal projections $P_V : \mathbf{R}^n \to V$ onto m-dimensional linear subspaces V of $\mathbf{R}^n$.

The last two chapters involve relations of rectifiability to complex and harmonic analysis. In Chapter 19 we discuss a classical problem of complex analysis: what are the null-sets for analytic capacity, or, in other words, which compact subsets of the complex plane are removable for bounded analytic functions? We try to explain how this open problem is related to rectifiability and we prove some partial results. In Chapter 20 we study the behaviour of certain natural singular integrals with respect to measures. It has turned out that here too there are many connections to rectifiability. We prove some results concerning the almost everywhere existence of principal values and discuss briefly some others, like the boundedness on L^2.

A sufficient prerequisite for reading this book is the knowledge of basic theory of measure and integration. On some occasions we shall use the Hahn–Banach theorem and once the Krein–Milman theorem, but they are only needed for certain specific results. In Chapter 12 we shall take for granted some properties of Fourier transforms and distributions.

After that they will only be needed in Chapter 13. Chapter 19 assumes some very basic facts from complex analysis. The complete understanding of the last part of Chapter 20 would require a great deal from the theory of singular integrals. We shall not treat the theory of Suslin (i.e. analytic) sets as there are many good sources for that, see e.g. Carleson [1], Dellacherie [1], Federer [3], Hayman and Kennedy [1] or Rogers [1]. This would not be needed if we were to restrict the formulation of the results to closed sets, but often generalization to Borel sets seems to require the theory of Suslin sets. In particular, we shall prove Frostman's lemma, Theorem 8.8, only for closed sets and many of the results of Chapters 9–13 depend on it. Most of the results which we state for the more familiar Borel sets actually hold for Suslin sets.

The list of references is long but there is no attempt at completeness. In particular concerning the topics which are related to the material of this book, but are not developed here, the choice of the references has been to some extent arbitrary. There are many more works on self-similarity, dynamical systems, etc. which could, and perhaps should, have been mentioned. However, I hope that the remarks in the text and the references given open the way to the interested reader to discover more about the literature.

1. General measure theory

In this chapter we shall introduce some general measure-theoretic concepts, terminology, and results which will be needed later on. But we shall also assume that the reader is familiar with basic measure theory. Most of the material needed can be found in several standard books such as Halmos [1], Hewitt and Stromberg [1], Munroe [1], Royden [1], Rudin [1], and many others including Federer [3] and Rogers [1]. Many of the proofs will be omitted. In this chapter we shall also introduce a great deal of notation and terminology to be used throughout the book. We shall generally follow the most standardized terminology of measure theory with one notable exception. Following Federer and Rogers we shall call measure what is usually called outer measure.

Some basic notation

We shall work in a metric space X with a metric d, although most of the measure theory presented here goes through in more general settings. Later on we shall however mainly stay in the euclidean n-space $\mathbf{R}^n$. Here are the basic notations used in metric spaces throughout this book.

The closed and open balls with centre $x \in X$ and radius r, $0 < r < \infty$, are denoted by

$$B(x,r) = \{y \in X : d(x,y) \leq r\},$$
$$U(x,r) = \{y \in X : d(x,y) < r\}.$$

In $\mathbf{R}^n$ we also set

$$B(r) = B(0,r),\ U(r) = U(0,r),\ S(x,r) = \partial B(x,r) \text{ and } S(r) = S(0,r).$$

The diameter of a non-empty subset A of X is

$$d(A) = \sup\{d(x,y) : x,y \in A\}.$$

We agree $d(\emptyset) = 0$. If $x \in X$ and A and B are non-empty subsets of X, the distance from x to A and the distance between A and B are, respectively,

$$d(x,A) = \inf\{d(x,y) : y \in A\},$$
$$d(A,B) = \inf\{d(x,y) : x \in A, y \in B\}.$$

For $\varepsilon > 0$ the closed ε-neighbourhood of A is

$$A(\varepsilon) = \{x \in X : d(x,A) \leq \varepsilon\}.$$

Measures

A measure for us will be a non-negative, monotonic, subadditive set function vanishing for the empty set.

1.1. Definition. *A set function* $\mu\colon \{A : A \subset X\} \to [0,\infty] = \{t : 0 \leq t \leq \infty\}$ *is called a measure if*

(1) $\mu(\emptyset) = 0$,
(2) $\mu(A) \leq \mu(B)$ *whenever* $A \subset B \subset X$,
(3) $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$ *whenever* $A_1, A_2, \dots \subset X$.

Usually in measure theory a measure means a non-negative countably additive set function defined on some σ-algebra of subsets of X, which need not be the whole power set $\{A : A \subset X\}$. However, considering measures in the sense of Definition 1.1 is a convenience rather than a restriction. That is, if ν is a countably additive non-negative set function on a σ-algebra $\mathcal{A}$ of subsets of X, it can be extended to a measure ν^* on X (in the sense of Definition 1.1) by

$$\nu^*(A) = \inf\{\nu(B) : A \subset B \in \mathcal{A}\}, \tag{1.2}$$

see Exercise 1. On the other hand, a measure μ gives a countably additive set function when restricted to the σ-algebra of μ measurable sets.

1.3. Definition. *A set* $A \subset X$ *is* μ *measurable if*

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A) \quad \text{for all } E \subset X.$$

We collect the well-known basic properties of measurable sets in the following theorem.

1.4. Theorem. *Let* μ *be a measure on* X *and let* $\mathcal{M}$ *be the family of all* μ *measurable subsets of* X.

(1) $\mathcal{M}$ *is a* σ*-algebra, that is,*

(i) $\emptyset \in \mathcal{M}$ *and* $X \in \mathcal{M}$,
(ii) *if* $A \in \mathcal{M}$, *then* $X \setminus A \in \mathcal{M}$,
(iii) *if* $A_1, A_2, \dots \in \mathcal{M}$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$.

(2) *If* $\mu(A) = 0$, *then* $A \in \mathcal{M}$.
(3) *If* $A_1, A_2, \dots \in \mathcal{M}$ *are pairwise disjoint, then*

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i).$$

(4) *If* $A_1, A_2, \dots \in \mathcal{M}$, *then*

(i) $\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \lim_{i\to\infty} \mu(A_i)$ *provided* $A_1 \subset A_2 \subset \dots$,

(ii) $\mu\Big(\bigcap_{i=1}^{\infty} A_i\Big) = \lim_{i\to\infty} \mu(A_i)$ *provided* $A_1 \supset A_2 \supset \dots$ *and* $\mu(A_1) < \infty$.

It is also good to remember that the first statement of (4) holds without the measurability assumption if μ is *regular*, that is, for every $A \subset X$ there is a μ measurable set $B \subset X$ such that $A \subset B$ and $\mu(A) = \mu(B)$.

Recall that the family of *Borel sets* in X is the smallest σ-algebra containing the open (or equivalently closed) subsets of X. We shall often consider measures with some of the following properties.

1.5. Definition. *Let* μ *be a measure on* X.

(1) μ *is locally finite if for every* $x \in X$ *there is* $r > 0$ *such that* $\mu(B(x,r)) < \infty$.
(2) μ *is a Borel measure if all Borel sets are* μ *measurable.*
(3) μ *is Borel regular if it is a Borel measure and if for every* $A \subset X$ *there is a Borel set* $B \subset X$ *such that* $A \subset B$ *and* $\mu(A) = \mu(B)$.
(4) μ *is a Radon measure if it is a Borel measure and*

(i) $\mu(K) < \infty$ *for compact sets* $K \subset X$,
(ii) $\mu(V) = \sup\{\mu(K) : K \subset V \text{ is compact}\}$ *for open sets* $V \subset X$,
(iii) $\mu(A) = \inf\{\mu(V) : A \subset V,\ V \text{ is open}\}$ *for* $A \subset X$.

We shall give a few simple examples. Many others will be encountered later on.

1.6. Examples.

(1) The Lebesgue measure $\mathcal{L}^n$ on $\mathbf{R}^n$ is a Radon measure.

(2) The Dirac measure δ_a at a point $a \in X$ is defined by $\delta_a(A) = 1$, if $a \in A$, $\delta_a(A) = 0$, if $a \notin A$ (that is, $\delta_a(A) = \chi_A(a)$). It is a Radon measure on any metric space X.
(3) The counting measure n on X is defined by letting $n(A)$ be the number of elements in A, possibly ∞. It is Borel regular on any metric space X, but it is a Radon measure only if every compact subset of X is finite, that is, X is discrete.

In general, Radon measures are always Borel regular as a rather immediate consequence of the definition. The converse is not true as the above example (3) shows. However, locally finite Borel regular measures in complete separable metric spaces are Radon measures, see e.g. Jacobs [1, Theorem V.5.3] or Schwartz [1, Part I, § II.3]. In $\mathbf{R}^n$ this will be stated in Corollary 1.11. Clearly in $\mathbf{R}^n$ the local finiteness means that compact sets have finite measure.

Borel measures in metric spaces are often called metric (outer) measures, because the following, Carathéodory's, criterion gives a very convenient necessary and sufficient metric condition for the measurability of Borel sets.

1.7. Theorem. *Let μ be a measure on X. Then μ is a Borel measure if and only if*

$$\mu(A \cup B) = \mu(A) + \mu(B) \quad \textit{whenever } A, B \subset X \textit{ with } d(A, B) > 0.$$

The proof of the more essential "if" part is given in many text-books, e.g. Munroe [1], Falconer [4], Federer [3], L. Simon [1]. The easier "only if" part is left as an exercise.

Given a measure μ and a subset A of X we can form a new measure by restricting μ to A.

1.8. Definition. *The restriction of a measure μ to a set $A \subset X$, $\mu \llcorner A$, is defined by*

$$(\mu \llcorner A)(B) = \mu(A \cap B) \quad \textit{for } B \subset X.$$

It is clear that $\mu \llcorner A$ is a measure. Many of the relations between μ and $\mu \llcorner A$ are easy to derive. For example,

1.9. Theorem.

(1) *Every μ measurable set is also $\mu \llcorner A$ measurable.*
(2) *If μ is Borel regular and A is μ measurable with $\mu(A) < \infty$, then $\mu \llcorner A$ is Borel regular.*