LAWRENCE C. EVANS
RONALD F. GARIEPY

Measure Theory and Fine Properties of Functions

CRC PRESS

STUDIES IN ADVANCED MATHEMATICS

Library of Congress Cataloging-in-Publication Data

Evans, Lawrence C., 1949–
Measure theory and fine properties of functions / by Lawrence C. Evans and Ronald F. Gariepy.
p. cm.
Includes bibliographical references and index.
ISBN 0-8493-7157-0
1. Measure Theory. 2. Functional Analysis. I. Gariepy, Ronald F. II. Title.
QA325.E92 1992
515'.42–dc20 91-30277
CIP

This book was formatted with LaTeX by Archetype Publishing Inc., P.O. Box 6567, Champaign, IL 61821.

Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., N.W., Boca Raton, Florida, 33431.

International Standard Book Number 0-8493-7157-0

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

Contents

Preface

These notes gather together what we regard as the essentials of real analysis on $\mathbb{R}^n$.

There are of course many good texts describing, on the one hand, Lebesgue measure for the real line and, on the other, general measures for abstract spaces. But we believe there is still a need for a source book documenting the rich structure of measure theory on $\mathbb{R}^n$, with particular emphasis on integration and differentiation. And so we packed into these notes all sorts of interesting topics that working mathematical analysts need to know, but are mostly not taught. These include Hausdorff measures and capacities (for classifying "negligible" sets for various fine properties of functions), Rademacher's Theorem (asserting the differentiability of Lipschitz functions almost everywhere), Aleksandrov's Theorem (asserting the twice differentiability of convex functions almost everywhere), the Area and Coarea Formulas (yielding change-of-variables rules for Lipschitz maps between $\mathbb{R}^n$ and $\mathbb{R}^m$), and the Lebesgue–Besicovitch Differentiation Theorem (amounting to the Fundamental Theorem of Calculus for real analysis).

This book is definitely not for beginners. We explicitly assume our readers are at least fairly conversant with both Lebesgue measure and abstract measure theory. The expository style reflects this expectation. We do not offer lengthy heuristics or motivation, but as compensation have tried to present *all* the technicalities of the proofs: "God is in the details."

Chapter 1 comprises a quick review of mostly standard real analysis, Chapter 2 introduces Hausdorff measures, and Chapter 3 discusses the Area and Coarea Formulas. In Chapters 4 through 6 we analyze the fine properties of functions possessing weak derivatives of various sorts. Sobolev functions, which is to say functions having weak first partial derivatives in an L^p space, are the subject of Chapter 4; functions of bounded variation, that is, functions having measures as weak first partial derivatives, the subject of Chapter 5. Finally, Chapter 6 discusses the approximation of Lipschitz, Sobolev and BV functions by C^1 functions, and several related subjects.

We have listed in the references the primary sources we have relied upon for these notes. In addition many colleagues, in particular S. Antman, Jo-Ann

Cohen, M. Crandall, A. Damlamian, H. Ishii, N. Owen, P. Souganidis, and S. Spector, have suggested improvements and detected errors. We have also made use of S. Katzenburger's class notes.

Early drafts of the manuscript were typed by E. Hampton, M. Hourihan, B. Kaufman, and J. Slack.

LCE was partially supported by NSF Grants DMS-83-01265, 86-01532, and 89-03328, and by the Institute for Physical Science and Technology at the University of Maryland. RFG was partially supported by NSF Grant DMS-87-04111 and by NSF Grant RII-86-10671 and the Commonwealth of Kentucky through the Kentucky EPSCoR program.

Warnings

Our terminology is occasionally at variance with standard usage. The principal changes are these:

> What we call a *measure* is usually called an *outer measure*.
>
> For us a function is *integrable* if it has an integral (which may equal $\pm\infty$). We call a function f *summable* if $|f|$ has a finite integral.
>
> We do *not* identify two L^p, BV, or Sobolev functions which agree a.e.

The reader should consult as necessary the list of notation, page 261.

1

General Measure Theory

This chapter is primarily a review of standard measure theory, with particular attention paid to Radon measures on $\mathbb{R}^n$.

Sections 1.1 through 1.4 are a rapid recounting of abstract measure theory. In Section 1.5 we establish Vitali's and Besicovitch's Covering Theorems, the latter being the key for the Lebesgue–Besicovitch Differentiation Theorem for Radon measures in Sections 1.6 and 1.7. Section 1.8 provides a vector-valued version of Riesz's Representation Theorem. In Section 1.9 we study weak compactness for sequences of measures and functions.

1.1 Measures and measurable functions

1.1.1 Measures; Approximation by open and compact sets

Although we intend later to work almost exclusively in $\mathbb{R}^n$, it is most convenient to start abstractly.

Let X denote a set, and 2^X the collection of subsets of X.

DEFINITION *A mapping $\mu : 2^X \to [0, \infty]$ is called a measure on X if*

(i) $\mu(\emptyset) = 0$, *and*

(ii) $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$ *whenever* $A \subset \cup_{k=1}^{\infty} A_k$.

Warning: Most texts call such a mapping μ an *outer* measure, reserving the name *measure* for μ restricted to the collection of μ-measurable subsets of X (see below). We will see, however, that there are definite advantages to being able to measure even nonmeasurable sets.

REMARK If μ is a measure on X and $A \subset B \subset X$, then

$$\mu(A) \leq \mu(B). \quad \blacksquare$$

DEFINITION *Let μ be a measure on X and $A \subset X$. Then μ* ***restricted to*** *A, written*

$$\mu \llcorner A,$$

is the measure defined by

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

DEFINITION *A set $A \subset X$ is* ***μ-measurable*** *if for each set $B \subset X$,*

$$\mu(B) = \mu(B \cap A) + \mu(B - A).$$

REMARKS If $\mu(A) = 0$, then A is μ-measurable. Clearly A is μ-measurable if and only if $X - A$ is μ-measurable. Observe also that if A is any subset of X, then any μ-measurable set is also $\mu \llcorner A$-measurable. ∎

THEOREM 1 PROPERTIES OF MEASURABLE SETS
Let $\{A_k\}_{k=1}^\infty$ be a sequence of μ-measurable sets.

(i) *The sets $\cup_{k=1}^\infty A_k$ and $\cap_{k=1}^\infty A_k$ are μ-measurable.*

(ii) *If the sets $\{A_k\}_{k=1}^\infty$ are disjoint, then*

$$\mu\left(\bigcup_{k=1}^\infty A_k\right) = \sum_{k=1}^\infty \mu(A_k).$$

(iii) *If $A_1 \subset \ldots A_k \subset A_{k+1} \ldots$, then*

$$\lim_{k\to\infty} \mu(A_k) = \mu\left(\bigcup_{k=1}^\infty A_k\right).$$

(iv) *If $A_1 \supset \ldots A_k \supset A_{k+1} \ldots$ and $\mu(A_1) < \infty$, then*

$$\lim_{k\to\infty} \mu(A_k) = \mu\left(\bigcap_{k=1}^\infty A_k\right).$$

PROOF

1. Since

$$\mu(B) \le \mu(B \cap A) + \mu(B - A)$$

for all $A, B \subset \mathbb{R}^n$, it suffices to show the opposite inequality in order to prove the set A is μ-measurable.

2. For each set $B \subset \mathbb{R}^n$

$$\begin{aligned}\mu(B) &= \mu(B \cap A_1) + \mu(B - A_1)\\ &= \mu(B \cap A_1) + \mu((B - A_1) \cap A_2) + \mu((B - A_1) - A_2)\\ &\geq \mu(B \cap (A_1 \cup A_2)) + \mu(B - (A_1 \cup A_2)),\end{aligned}$$

and thus $A_1 \cup A_2$ is μ-measurable. By induction the union of finitely many μ-measurable sets is μ-measurable.

3. Since

$$X - (A_1 \cap A_2) = (X - A_1) \cup (X - A_2),$$

the intersection of two, and thus of finitely many, μ-measurable sets is μ-measurable.

4. Assume now the sets $\{A_k\}_{k=1}^\infty$ are disjoint, and write

$$B_j \equiv \bigcup_{k=1}^{j} A_k \qquad (j = 1, 2, \ldots).$$

Then

$$\begin{aligned}\mu(B_{j+1}) &= \mu(B_{j+1} \cap A_{j+1}) + \mu(B_{j+1} - A_{j+1})\\ &= \mu(A_{j+1}) + \mu(B_j) \qquad (j = 1, \ldots);\end{aligned}$$

whence

$$\mu\left(\bigcup_{k=1}^{j} A_k\right) = \sum_{k=1}^{j} \mu(A_k) \qquad (j = 1, \ldots).$$

It follows that

$$\sum_{k=1}^{\infty} \mu(A_k) \leq \mu\left(\bigcup_{k=1}^{\infty} A_k\right),$$

from which inequality assertion (ii) follows.

5. To prove (iii), we note from (ii)

$$\lim_{k\to\infty} \mu(A_k) = \mu(A_1) + \sum_{k=1}^{\infty} \mu(A_{k+1} - A_k) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right).$$

6. Assertion (iv) follows from (iii), since

$$\begin{aligned}\mu(A_1) - \lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \mu(A_1 - A_k) &= \mu\left(\bigcup_{k=1}^{\infty} (A_1 - A_k)\right)\\ &\geq \mu(A_1) - \mu\left(\bigcap_{k=1}^{\infty} A_k\right).\end{aligned}$$

7. Recall that if B is any subset of X, then each μ-measurable set is also $\mu \llcorner B$-measurable. Since $B_j \equiv \cup_{k=1}^{j} A_k$ is μ-measurable by step 2, for each $B \subset X$ with $\mu(B) < \infty$ we have

$$\begin{aligned}
\mu\left(B \bigcap \bigcup_{k=1}^{\infty} A_k\right) + \mu\left(B - \bigcup_{k=1}^{\infty} A_k\right) &= (\mu \llcorner B)\left(\bigcup_{k=1}^{\infty} B_k\right) + (\mu \llcorner B)\left(\bigcap_{k=1}^{\infty}(X - B_k)\right) \\
&= \lim_{k\to\infty} (\mu \llcorner B)(B_k) + \lim_{k\to\infty} (\mu \llcorner B)(X - B_k) \\
&= \mu(B).
\end{aligned}$$

Thus $\cup_{k=1}^{\infty} A_k$ is μ-measurable, as is $\cap_{k=1}^{\infty} A_k$, since

$$X - \bigcap_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty}(X - A_k).$$

This proves (i). ▮

DEFINITION *A collection of subsets $\mathcal{A} \subset 2^X$ is a **σ-algebra** provided*

(i) $\emptyset, X \in \mathcal{A}$;

(ii) $A \in \mathcal{A}$ *implies* $X - A \in \mathcal{A}$;

(iii) $A_k \in \mathcal{A}$ $(k = 1, \dots)$ *implies* $\cup_{k=1}^{\infty} A_k \in \mathcal{A}$.

Thus the collection of all μ-measurable subsets of X forms a σ-algebra.

DEFINITION *A subset $A \subset X$ is **σ-finite with respect to** μ if we can write $A = \cup_{k=1}^{\infty} B_k$, where B_k is μ-measurable and $\mu(B_k) < \infty$ for $k = 1, 2, \dots$.*

DEFINITION *The **Borel σ-algebra** of $\mathbb{R}^n$ is the smallest σ-algebra of $\mathbb{R}^n$ containing the open subsets of $\mathbb{R}^n$.*

Next we introduce certain classes of measures that admit good approximations of various types.

DEFINITIONS

*(i) A measure μ on X is **regular** if for each set $A \subset X$ there exists a μ-measurable set B such that $A \subset B$ and $\mu(A) = \mu(B)$.*

*(ii) A measure μ on $\mathbb{R}^n$ is called **Borel** if every Borel set is μ-measurable.*

(iii) *A measure μ on $\mathbb{R}^n$ is **Borel regular** if μ is Borel and for each $A \subset \mathbb{R}^n$ there exists a Borel set B such that $A \subset B$ and $\mu(A) = \mu(B)$.*

(iv) *A measure μ on $\mathbb{R}^n$ is a **Radon** measure if μ is Borel regular and $\mu(K) < \infty$ for each compact set $K \subset \mathbb{R}^n$.*

THEOREM 2
Let μ be a regular measure on X. If $A_1 \subset \dots A_k \subset A_{k+1} \dots$, then

$$\lim_{k\to\infty} \mu(A_k) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right)$$

REMARK The important point is that the sets $\{A_k\}_{k=1}^{\infty}$ need not be μ-measurable here. ▌

PROOF Since μ is regular, there exist measurable sets $\{C_k\}_{k=1}^{\infty}$, with $A_k \subset C_k$ and $\mu(A_k) = \mu(C_k)$ for each k. Set $B_k \equiv \bigcap_{j\geq k} C_j$. Then $A_k \subset B_k$, each B_k is μ-measurable, and $\mu(A_k) = \mu(B_k)$. Thus

$$\lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \mu(B_k) = \mu\left(\bigcup_{k=1}^{\infty} B_k\right) \geq \mu\left(\bigcup_{k=1}^{\infty} A_k\right).$$

But $A_k \subset \bigcup_{j=1}^{\infty} A_j$, and so also

$$\lim_{k\to\infty} \mu(A_k) \leq \mu\left(\bigcup_{j=1}^{\infty} A_j\right). \quad ▌$$

We demonstrate next that if μ is Borel regular, we can generate a Radon measure by restricting μ to a measurable set of finite measure.

THEOREM 3
Let μ be a Borel regular measure on $\mathbb{R}^n$. Suppose $A \subset \mathbb{R}^n$ is μ-measurable and $\mu(A) < \infty$. Then $\mu \llcorner A$ is a Radon measure.

REMARK If A is a Borel set, then $\mu \llcorner A$ is Borel regular, even if $\mu(A) = \infty$. ▌

$\nu(K) = \mu(A \cap K) \leq \mu(A) < \infty$

PROOF Let $\nu \equiv \mu \llcorner A$. Clearly $\nu(K) < \infty$ for each compact K. Since every μ-measurable set is ν-measurable, ν is a Borel measure.

Claim: ν is Borel regular.

Proof of Claim: Since μ is Borel regular, there exists a Borel set B such that $A \subset B$ and $\mu(A) = \mu(B) < \infty$. Then, since A is μ-measurable,

$$\mu(B - A) = \mu(B) - \mu(A) = 0.$$

Choose $C \subset \mathbb{R}^n$. Then

$$\begin{aligned}(\mu \llcorner B)(C) &= \mu(C \cap B)\\ &= \mu(C \cap B \cap A) + \mu((C \cap B) - A)\\ &\leq \mu(C \cap A) + \mu(B - A)\\ &= (\mu \llcorner A)(C).\end{aligned}$$

Thus $\mu \llcorner B = \mu \llcorner A$, so we may as well assume A is a Borel set.

Now let $C \subset \mathbb{R}^n$. We must show that there exists a Borel set D such that $C \subset D$ and $\nu(C) = \nu(D)$. Since μ is a Borel regular measure, there exists a Borel set E such that $A \cap C \subset E$ and $\mu(E) = \mu(A \cap C)$. Let $D \equiv E \cup (\mathbb{R}^n - A)$. Since A and E are Borel sets, so is D. Moreover, $C \subset (A \cap C) \cup (\mathbb{R}^n - A) \subset D$. Finally, since $D \cap A = E \cap A$,

$$\nu(D) = \mu(D \cap A) = \mu(E \cap A) \leq \mu(E) = \mu(A \cap C) = \nu(C). \quad \blacksquare$$

We consider next the possibility of measure theoretically approximating by open, closed, or compact sets.

LEMMA 1

Let μ be a Borel measure on $\mathbb{R}^n$ and let B be a Borel set.

- *(i) If $\mu(B) < \infty$, there exists for each $\epsilon > 0$ a closed set C such that $C \subset B$ and $\mu(B - C) < \epsilon$.*
- *(ii) If μ is a Radon measure, then there exists for each $\epsilon > 0$ an open set U such that $B \subset U$ and $\mu(U - B) < \epsilon$.*

PROOF

1. Let $\nu \equiv \mu \llcorner B$. Since μ is Borel and $\mu(B) < \infty$, ν is a finite Borel measure. Let

$$\begin{aligned}\mathcal{F} \equiv \{A \subset \mathbb{R}^n \mid &A \text{ is } \mu\text{-measurable and for each } \epsilon > 0\\ &\text{there exists a closed set } C \subset A \text{ such that } \nu(A - C) < \epsilon\}.\end{aligned}$$

Trivially, $\mathcal{F}$ contains all closed sets.

2. Claim #1: If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, then $A \equiv \cap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Proof of Claim #1: Fix $\epsilon > 0$. Since $A_i \in \mathcal{F}$, there exists a closed set $C_i \subset A_i$ with $\nu(A_i - C_i) < \epsilon/2^i$ $(i = 1, 2, \dots)$. Let $C \equiv \cap_{i=1}^{\infty} C_i$. Then C is closed and

$$\begin{aligned}\nu(A - C) &= \nu\left(\bigcap_{i=1}^{\infty} A_i - \bigcap_{i=1}^{\infty} C_i\right)\\ &\leq \nu\left(\bigcup_{i=1}^{\infty}(A_i - C_i)\right)\\ &\leq \sum_{i=1}^{\infty} \nu(A_i - C_i) < \epsilon.\end{aligned}$$

Thus $A \in \mathcal{F}$.

3. Claim #2: If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, then $A \equiv \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Proof of Claim #2: Fix $\epsilon > 0$ and choose C_i as above. Since $\nu(A) < \infty$, we have

$$\begin{aligned}\lim_{m\to\infty} \nu\left(A - \bigcup_{i=1}^{m} C_i\right) &= \nu\left(\bigcup_{i=1}^{\infty} A_i - \bigcup_{i=1}^{\infty} C_i\right)\\ &\leq \nu\left(\bigcup_{i=1}^{\infty}(A_i - C_i)\right)\\ &\leq \sum_{i=1}^{\infty} \nu(A_i - C_i) < \epsilon.\end{aligned}$$

Consequently, there exists an integer m such that

$$\nu\left(A - \bigcup_{i=1}^{m} C_i\right) < \epsilon.$$

But $\cup_{i=1}^{m} C_i$ is closed, and so $A \in \mathcal{F}$.

4. Now, since every open subset of $\mathbb{R}^n$ can be written as a countable union of closed sets, Claim #2 shows that $\mathcal{F}$ contains all open sets. Now consider

$$\mathcal{G} \equiv \{A \in \mathcal{F} \mid \mathbb{R}^n - A \in \mathcal{F}\}.$$

Trivially, if $A \in \mathcal{G}$, then $\mathbb{R}^n - A \in \mathcal{G}$. Note also that $\mathcal{G}$ contains all open sets.

5. Claim #3: If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{G}$, then $A = \cup_{i=1}^{\infty} A_i \in \mathcal{G}$.

Proof of Claim #3: By Claim #2, $A \in \mathcal{F}$. Since also $\{\mathbb{R}^n - A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, Claim #1 implies $\mathbb{R}^n - A = \cap_{i=1}^{\infty}(\mathbb{R}^n - A_i) \in \mathcal{F}$.

6. Thus $\mathcal{G}$ is a σ-algebra containing the open sets and therefore also the Borel sets. In particular, $B \in \mathcal{G}$ and hence given $\epsilon > 0$ there is a closed set $C \subset B$

such that

$$\mu(B - C) = \nu(B - C) < \epsilon.$$

This establishes (i).

7. Write $U_m \equiv U(0, m)$, the open ball with center 0, radius m. Then $U_m - B$ is a Borel set with $\mu(U_m - B) < \infty$, and so we can apply (i) to find a closed set $C_m \subset U_m - B$ such that $\mu((U_m - C_m) - B) = \mu((U_m - B) - C_m) < \epsilon/2^m$. Let $U \equiv \cup_{m=1}^{\infty}(U_m - C_m)$; U is open. Now $B \subset \mathbb{R}^n - C_m$, and thus $U_m \cap B \subset U_m - C_m$. Consequently,

$$B = \bigcup_{m=1}^{\infty} (U_m \cap B) \subset \bigcup_{m=1}^{\infty} (U_m - C_m) = U.$$

Furthermore,

$$\begin{aligned}\mu(U - B) &= \mu\left(\bigcup_{m=1}^{\infty} ((U_m - C_m) - B)\right) \\ &\leq \sum_{m=1}^{\infty} \mu((U_m - C_m) - B) < \epsilon. \quad \blacksquare\end{aligned}$$

THEOREM 4 APPROXIMATION BY OPEN AND COMPACT SETS
Let μ be a Radon measure on $\mathbb{R}^n$. Then

(i) *for each set $A \subset \mathbb{R}^n$,*

$$\mu(A) = \inf\{\mu(U) \mid A \subset U, U \text{ open}\},$$

and

(ii) *for each μ-measurable set $A \subset \mathbb{R}^n$,*

$$\mu(A) = \sup\{\mu(K) \mid K \subset A, K \text{ compact}\}.$$

REMARK Assertion (i) does not require A to be μ-measurable. $\blacksquare$

PROOF

1. If $\mu(A) = \infty$, (i) is obvious, and so let us suppose $\mu(A) < \infty$. Assume first A is a Borel set. Fix $\epsilon > 0$. Then by Lemma 1, there exists an open set $U \supset A$ with $\mu(U - A) < \epsilon$. Since $\mu(U) = \mu(A) + \mu(U - A) < \infty$, (i) holds. Now, let A be an arbitrary set. Since μ is Borel regular, there exists a Borel set $B \supset A$ with $\mu(A) = \mu(B)$. Then

$$\begin{aligned}\mu(A) = \mu(B) &= \inf\{\mu(U) \mid B \subset U, U \text{ open}\} \\ &\geq \inf\{\mu(U) \mid A \subset U, U \text{ open}\}.\end{aligned}$$

The reverse inequality is clear: assertion (i) is proved.

2. Now let A be μ-measurable, with $\mu(A) < \infty$. Set $\nu \equiv \mu \llcorner A$; ν is a Radon measure according to Theorem 3. Fix $\epsilon > 0$. Applying (i) to ν and $\mathbb{R}^n - A$, we obtain an open set U with $\mathbb{R}^n - A \subset U$ and $\nu(U) \leq \epsilon$. Let $C \equiv \mathbb{R}^n - U$. Then C is closed and $C \subset A$. Moreover,

$$\mu(A - C) = \nu(\mathbb{R}^n - C) = \nu(U) \leq \epsilon.$$

Thus

$$0 \leq \mu(A) - \mu(C) \leq \epsilon,$$

and so

$$\mu(A) = \sup\{\mu(C) \mid C \subset A, C \text{ closed}\}. \qquad (\star)$$

Now suppose that $\mu(A) = \infty$. Define $D_k \equiv \{x \mid k - 1 \leq |x| < k\}$. Then $A = \bigcup_{k=1}^{\infty}(D_k \cap A)$; so $\infty = \mu(A) = \sum_{k=1}^{\infty} \mu(A \cap D_k)$. Since μ is a Radon measure, $\mu(D_k \cap A) < \infty$. Then by the above, there exists a closed set $C_k \subset D_k \cap A$ with $\mu(C_k) \geq \mu(D_k \cap A) - 1/2^k$. Now $\bigcup_{k=1}^{\infty} C_k \subset A$ and

$$\begin{aligned}\lim_{n\to\infty} \mu\left(\bigcup_{k=1}^{n} C_k\right) &= \mu\left(\bigcup_{k=1}^{\infty} C_k\right) \\ &= \sum_{k=1}^{\infty} \mu(C_k) \geq \sum_{k=1}^{\infty}\left[\mu(D_k \cap A) - \frac{1}{2^k}\right] = \infty.\end{aligned}$$

But $\bigcup_{k=1}^{n} C_k$ is closed for each n, whence in this case also we have assertion $(\star)$.

Finally, set $B_m = B(0, m)$, the closed ball with center 0, radius m. Let C be closed, $C_m \equiv C \cap B_m$. Each set C_m is compact and $\mu(C) = \lim_{m\to\infty} \mu(C_m)$. Hence for each μ-measurable set A,

$$\sup\{\mu(K) \mid K \subset A, K \text{ compact}\} = \sup\{\mu(C) \mid C \subset A, C \text{ closed}\}. \quad \blacksquare$$

We introduce next a simple and very useful way to verify that a measure is Borel.

THEOREM 5 CARATHEODORY'S CRITERION
Let μ be a measure on $\mathbb{R}^n$. If $\mu(A \cup B) = \mu(A) + \mu(B)$ for all sets $A, B \subset \mathbb{R}^n$ with $\operatorname{dist}(A, B) > 0$, *then μ is a Borel measure.*

PROOF

1. Suppose $C \subset \mathbb{R}^n$ is closed. We must show

$$\mu(A) \geq \mu(A \cap C) + \mu(A - C), \qquad (\star)$$

the opposite inequality following from subadditivity.

If $\mu(A) = \infty$, then $(\star)$ is obvious. Assume instead $\mu(A) < \infty$. Define

$$C_n \equiv \left\{ x \in \mathbb{R}^n \mid \operatorname{dist}(x, C) \leq \frac{1}{n} \right\} \qquad (n = 1, 2, \ldots).$$

Then $\operatorname{dist}(A - C_n, A \cap C) \geq 1/n > 0$. By hypothesis, therefore,

$$\mu(A - C_n) + \mu(A \cap C) = \mu((A - C_n) \cup (A \cap C)) \leq \mu(A). \qquad (\star\star)$$

2. *Claim*: $\lim_{n\to\infty} \mu(A - C_n) = \mu(A - C)$.

Proof of Claim: Set

$$R_k \equiv \left\{ x \in A \mid \frac{1}{k+1} < \operatorname{dist}(x, C) \leq \frac{1}{k} \right\} \qquad (k = 1, \ldots).$$

Then $A - C = (A - C_n) \cup \bigcup_{k=n}^{\infty} R_k$, so that

$$\mu(A - C_n) \leq \mu(A - C) \leq \mu(A - C_n) + \sum_{k=n}^{\infty} \mu(R_k).$$

If we can show $\sum_{k=1}^{\infty} \mu(R_k) < \infty$, we will then have

$$\begin{aligned} \lim_{n\to\infty} \mu(A - C_n) &\leq \mu(A - C) \\ &\leq \lim_{n\to\infty} \mu(A - C_n) + \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(R_k) \\ &= \lim_{n\to\infty} \mu(A - C_n), \end{aligned}$$

thereby establishing the claim.

3. Now $\operatorname{dist}(R_i, R_j) > 0$ if $j \geq i + 2$. Hence by induction we find

$$\sum_{k=1}^{m} \mu(R_{2k}) = \mu\left(\bigcup_{k=1}^{m} R_{2k} \right) \leq \mu(A),$$

and likewise

$$\sum_{k=0}^{m} \mu(R_{2k+1}) = \mu\left(\bigcup_{k=0}^{m} R_{2k+1} \right) \leq \mu(A).$$

Combining these results and letting $m \to \infty$, we discover

$$\sum^{\infty} \mu(R_k) \leq 2\mu(A) < \infty.$$

4. We have

$$\begin{aligned}\mu(A-C)+\mu(A\cap C) &= \lim_{n\to\infty}\mu(A-C_n)+\mu(A\cap C)\\ &\le \mu(A),\end{aligned}$$

according to ($\star\star$), and thus C is μ-measurable. ∎

1.1.2 Measurable functions

We now extend the notion of measurability from sets to functions.

Let X be a set and Y a topological space. Assume μ is a measure on X.

DEFINITION *A function $f : X \to Y$ is called **μ-measurable** if for each open $U \subset Y$, $f^{-1}(U)$ is μ-measurable.*

REMARK If $f : X \to Y$ is μ-measurable, then $f^{-1}(B)$ is μ-measurable for each Borel set $B \subset Y$. Indeed, $\{A \subset Y \mid f^{-1}(A) \text{ is } \mu\text{-measurable}\}$ is a σ-algebra containing the open sets and hence the Borel sets. ∎

DEFINITION *A function $f : X \to [-\infty,\infty]$ is **σ-finite with respect to μ** if f is μ-measurable and $\{x \mid f(x) \neq 0\}$ is σ-finite with respect to μ.*

Measurable functions inherit the good properties of measurable sets.

THEOREM 6 PROPERTIES OF MEASURABLE FUNCTIONS

(i) *If $f, g : X \to \mathbb{R}$ are μ-measurable, then so are $f+g$, fg, $|f|$, $\min(f,g)$, and $\max(f,g)$. The function f/g is also μ-measurable, provided $g \neq 0$ on X.*

(ii) *If the functions $f_k : X \to [-\infty,\infty]$ are μ-measurable $(k = 1,2,\ldots)$, then $\inf_{k\geq 1} f_k$, $\sup_{k\geq 1} f_k$, $\liminf_{k\to\infty} f_k$, and $\limsup_{k\to\infty} f_k$ are also μ-measurable.*

PROOF

1. In view of the remark, we easily check that $f : X \to [-\infty,\infty]$ is μ-measurable if and only if $f^{-1}[-\infty,a)$ is μ-measurable for each $a \in \mathbb{R}$, if and only if $f^{-1}[-\infty,a]$ is μ-measurable for each $a \in \mathbb{R}$.

2. Suppose $f, g : X \to \mathbb{R}$ are μ-measurable. Then

$$(f+g)^{-1}(-\infty,a) = \bigcup_{\substack{r,s \text{ rational}\\ r+s<a}} (f^{-1}(-\infty,r)\cap g^{-1}(-\infty,s)),$$

and so $f+g$ is μ-measurable. Since

$$(f^2)^{-1}(-\infty,a) = f^{-1}(-\infty,a^{\frac{1}{2}}) - f^{-1}(-\infty,-a^{\frac{1}{2}}],$$

for $a \geq 0$, f^2 is μ-measurable. Consequently,

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$$

is μ-measurable as well. Next observe that, if $g(x) \neq 0$ for $x \in X$,

$$\left(\frac{1}{g}\right)^{-1}(-\infty, a) = \begin{cases} g^{-1}(\frac{1}{a}, 0) & \text{if } a < 0 \\ g^{-1}(-\infty, 0) & \text{if } a = 0 \\ g^{-1}(-\infty, 0) \cup g^{-1}(\frac{1}{a}, \infty) & \text{if } a > 0; \end{cases}$$

thus $1/g$ and so also f/g are μ-measurable.

3. Finally,

$$f^+ = f\chi_{\{f \geq 0\}} = \max(f, 0), \qquad f^- = -f\chi_{\{f<0\}} = \max(-f, 0)$$

are μ-measurable, and consequently so are

$$\begin{aligned} |f| &= f^+ + f^-, \\ \max(f, g) &= (f-g)^+ + g, \\ \min(f, g) &= -(f-g)^- + g. \end{aligned}$$

4. Suppose next the functions $f_k : X \to [-\infty, \infty]$ $(k = 1, 2, \ldots)$ are μ-measurable. Then

$$\left(\inf_{k \geq 1} f_k\right)^{-1} [-\infty, a) = \bigcup_{k=1}^{\infty} f_k^{-1}[-\infty, a)$$

and

$$\left(\sup_{k \geq 1} f_k\right)^{-1} [-\infty, a] = \bigcap_{k=1}^{\infty} f_k^{-1}[-\infty, a],$$

so that

$$\inf_{k \geq 1} f_k, \ \sup_{k \geq 1} f_k \text{ are } \mu\text{-measurable.}$$

We complete the proof by noting

$$\liminf_{k \to \infty} f_k = \sup_{m \geq 1} \inf_{k \geq m} f_k,$$

$$\limsup_{k \to \infty} f_k = \inf_{m \geq 1} \sup_{k \geq m} f_k. \quad \blacksquare$$

Next is a simple but useful way to decompose a nonnegative measurable function.

THEOREM 7
Assume $f : X \to [0,\infty]$ is μ-measurable. Then there exist μ-measurable sets $\{A_k\}_{k=1}^{\infty}$ in X such that

$$f = \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}.$$

PROOF Set

$$A_1 \equiv \{x \in X \mid f(x) \geq 1\},$$

and inductively define for $k = 2, 3, \ldots$

$$A_k \equiv \left\{ x \in X \mid f(x) \geq \frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j} \right\}.$$

Clearly,

$$f \geq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}.$$

If $f(x) = \infty$, then $x \in A_k$ for all k. On the other hand, if $0 < f(x) < \infty$, then for infinitely many n, $x \notin A_n$. Hence for infinitely many n

$$0 \leq f(x) - \sum_{k=1}^{n-1} \frac{1}{k}\chi_{A_k} \leq \frac{1}{n}. \quad \blacksquare$$

1.2 Lusin's and Egoroff's Theorems

THEOREM 1
Suppose $K \subset \mathbb{R}^n$ is compact and $f : K \to \mathbb{R}^m$ is continuous. Then there exists a continuous mapping $\bar{f} : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\bar{f} = f \quad \text{on} \quad K.$$

REMARK Extension theorems preserving more of the structure of f will be presented in Sections 3.1.1, 4.4, 5.4, and 6.5. $\blacksquare$

PROOF

1. The assertion for $m > 1$ follows easily from the case $m = 1$, and so we may assume $f : K \to \mathbb{R}$.

2. Let $U \equiv \mathbb{R}^n - K$. For $x \in U$ and $s \in K$, set

$$u_s(x) \equiv \max\left\{2 - \frac{|x-s|}{\operatorname{dist}(x,K)}, 0\right\},$$

so that

$$\begin{cases} x \mapsto u_s(x) \text{ is continuous on } U, \\ 0 \le u_s(x) \le 1, \\ u_s(x) = 0 \text{ if } |x-s| \ge 2\operatorname{dist}(x,K). \end{cases}$$

Now let $\{s_j\}_{j=1}^{\infty}$ be a countable dense subset of K, and define

$$\sigma(x) \equiv \sum_{j=1}^{\infty} 2^{-j} u_{s_j}(x) \text{ for } x \in U.$$

Observe $0 < \sigma(x) \le 1$ for $x \in U$. Now set

$$v_k(x) \equiv \frac{2^{-k} u_{s_k}(x)}{\sigma(x)}$$

for $x \in U$, $k = 1, 2, \ldots$. The functions $\{v_k\}_{k=1}^{\infty}$ form a partition of unity on U. Define

$$\bar{f}(x) \equiv \begin{cases} f(x) & \text{if } x \in K \\ \sum_{k=1}^{\infty} v_k(x) f(s_k) & \text{if } x \in U. \end{cases}$$

By the Weierstrass M-test, $\bar{f}$ is continuous on U.

3. We must show

$$\lim_{\substack{x \to a \\ x \in U}} \bar{f}(x) = f(a)$$

for each $a \in K$. Fix $\epsilon > 0$. There exists $\delta > 0$ such that

$$|f(a) - f(s_k)| < \epsilon$$

for all s_k such that $|a - s_k| < \delta$. Suppose $x \in U$ with $|x - a| < \delta/4$. If $|a - s_k| \ge \delta$, then

$$\delta \le |a - s_k| \le |a - x| + |x - s_k| < \frac{\delta}{4} + |x - s_k|,$$

so that

$$|x - s_k| \ge \frac{3}{4}\delta > 2|x - a| \ge 2 \ \operatorname{dist}(x, K).$$

Thus, $v_k(x) = 0$ whenever $|x - a| < \delta/4$ and $|a - s_k| \ge \delta$. Since

$$\sum_{k=1}^{\infty} v_k(x) = 1$$

if $x \in U$, we calculate for $|x-a| < \delta/4$, $x \in U$,

$$|\bar{f}(x) - f(a)| \leq \sum_{k=1}^{\infty} v_k(x)|f(s_k) - f(x)| < \epsilon. \quad \blacksquare$$

We now show that a measurable function can be measure theoretically approximated by a continuous function.

THEOREM 2 LUSIN'S THEOREM
Let μ be a Borel regular measure on $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ be μ-measurable. Assume $A \subset \mathbb{R}^n$ is μ-measurable and $\mu(A) < \infty$. Fix $\epsilon > 0$. Then there exists a compact set $K \subset A$ such that

(i) $\mu(A - K) < \epsilon$, *and*

(ii) $f\,|_K$ *is continuous.*

PROOF For each positive integer i, let $\{B_{ij}\}_{j=1}^{\infty} \subset \mathbb{R}^m$ be disjoint Borel sets such that $\mathbb{R}^m = \cup_{j=1}^{\infty} B_{ij}$ and diam $B_{ij} < 1/i$. Define $A_{ij} \equiv A \cap f^{-1}(B_{ij})$. Then A_{ij} is μ-measurable and $A = \cup_{j=1}^{\infty} A_{ij}$.

Write $\nu = \mu \llcorner A$; ν is a Radon measure. Theorem 4 in Section 1.1 implies the existence of a compact set $K_{ij} \subset A_{ij}$ with $\nu(A_{ij} - K_{ij}) < \epsilon/2^{i+j}$. Then

$$\begin{aligned}
\mu\left(A - \bigcup_{j=1}^{\infty} K_{ij}\right) &= \nu\left(A - \bigcup_{j=1}^{\infty} K_{ij}\right) \\
&= \nu\left(\bigcup_{j=1}^{\infty} A_{ij} - \bigcup_{j=1}^{\infty} K_{ij}\right) \\
&\leq \nu\left(\bigcup_{j=1}^{\infty} (A_{ij} - K_{ij})\right) < \frac{\epsilon}{2^i}\,.
\end{aligned}$$

As $\lim_{N\to\infty} \mu(A - \cup_{j=1}^{N} K_{ij}) = \mu(A - \cup_{j=1}^{\infty} K_{ij})$, there exists a number $N(i)$ such that

$$\mu\left(A - \bigcup_{j=1}^{N(i)} K_{ij}\right) < \epsilon/2^i.$$

Set $D_i \equiv \cup_{j=1}^{N(i)} K_{ij}$; D_i is compact. For each i and j, we fix $b_{ij} \in B_{ij}$ and then define $g_i : D_i \to \mathbb{R}^m$ by setting $g_i(x) = b_{ij}$ for $x \in K_{ij}$ ($j \leq N(i)$). Since $K_{i1}, \ldots, K_{iN(i)}$ are compact, disjoint sets, and so are a positive distance apart, g_i is continuous. Furthermore, $|f(x) - g_i(x)| < 1/i$ for all $x \in D_i$. Set

$K \equiv \cap_{i=1}^{\infty} D_i$: K is compact and

$$\mu(A - K) \leq \sum_{i=1}^{\infty} \mu(A - D_i) < \epsilon.$$

Since $|f(x) - g_i(x)| < 1/i$ for each $x \in D_i$, we see $g_i \to f$ uniformly on K. Thus $f\,|_K$ is continuous, as required. ▮

COROLLARY 1
Let μ be a Borel regular measure on $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be μ-measurable. Assume $A \subset \mathbb{R}^n$ is μ-measurable and $\mu(A) < \infty$. Fix $\epsilon > 0$. Then there exists a continuous function $\bar{f} : \mathbb{R}^n \to \mathbb{R}^m$ such that $\mu\{x \in A \mid \bar{f}(x) \neq f(x)\} < \epsilon$.

PROOF By Lusin's Theorem there exists a compact set $K \subset A$ such that $\mu(A - K) < \epsilon$ and $f\,|_K$ is continuous. Then by Theorem 1 there exists a continuous function $\bar{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\bar{f}\,|_K = f\,|_K$ and

$$\mu\{x \in A \mid \bar{f}(x) \neq f(x)\} \leq \mu(A - K) < \epsilon. \quad ▮$$

REMARK Compare this with Whitney's Extension Theorem, Theorem 2 in Section 5.6, which identifies conditions ensuring the existence of a C^1 extension $\bar{f}$. ▮

NOTATION The expression "μ a.e." means "almost everywhere with respect the measure μ," that is, except possibly on a set A with $\mu(A) = 0$.

THEOREM 3 EGOROFF'S THEOREM
Let μ be a measure on $\mathbb{R}^n$ and suppose $f_k : \mathbb{R}^n \to \mathbb{R}^m$ $(k = 1, 2, \ldots)$ are μ-measurable. Assume also $A \subset \mathbb{R}^n$ is μ-measurable, with $\mu(A) < \infty$, and $f_k \to g$ μ a.e. on A. Then for each $\epsilon > 0$ there exists a μ-measurable set $B \subset A$ such that

(i) $\mu(A - B) < \epsilon$, *and*

(ii) $f_k \to g$ *uniformly on* B.

PROOF Define $C_{ij} \equiv \cup_{k=j}^{\infty} \{x \mid |f_k(x) - g(x)| > 2^{-i}\}$, $(i, j = 1, 2, \ldots)$. Then $C_{i,j+1} \subset C_{ij}$ for all i, j; and so, since $\mu(A) < \infty$,

$$\lim_{j \to \infty} \mu(A \cap C_{ij}) = \mu\left(A \bigcap \bigcap_{j=1}^{\infty} C_{ij}\right) = 0.$$

Hence there exists an integer $N(i)$ such that $\mu(A \cap C_{i,N(i)}) < \epsilon/2^i$.

Let $B \equiv A - \cup_{i=1}^{\infty} C_{i,N(i)}$. Then

$$\mu(A - B) \leq \sum_{i=1}^{\infty} \mu\left(A \cap C_{i,N(i)}\right) < \epsilon.$$

Then for each i, $x \in B$, and all $n \geq N(i)$, $|f_n(x) - g(x)| \leq 2^{-i}$. Thus $f_n \to g$ uniformly on B. ▮

1.3 Integrals and limit theorems

Now we want to extend calculus to the measure theoretic setting. This section presents integration theory; differentiation theory is harder and will be set forth later in Section 1.6.

NOTATION

$$f^+ = \max(f, 0), \; f^- = \max(-f, 0), \; f = f^+ - f^-.$$

Let μ be a measure on a set X.

DEFINITION *A function $g : X \to [-\infty, \infty]$ is called a* ***simple function*** *if the image of g is countable.*

DEFINITION *If g is a nonnegative, simple, μ-measurable function, we define*

$$\int g \, d\mu \equiv \sum_{0 \leq y \leq \infty} y \mu(g^{-1}\{y\})$$

DEFINITION *If g is a simple μ-measurable function and either $\int g^+ \, d\mu < \infty$ or $\int g^- \, d\mu < \infty$, we call g a* ***μ-integrable simple function*** *and define*

$$\int g \, d\mu \equiv \int g^+ \, d\mu - \int g^- \, d\mu.$$

Thus if g is a μ-integrable simple function,

$$\int g \, d\mu = \sum_{-\infty < y < \infty} y \mu(g^{-1}\{y\}).$$

DEFINITIONS *Let $f : X \to [-\infty, \infty]$. We define the* ***upper integral***

$$\int^{\star} f\, d\mu \equiv \inf\left\{\int g\, d\mu \mid g \text{ a } \mu\text{-integrable simple function with } g \geq f\ \mu \text{ a.e.}\right\}$$

and the ***lower integral***

$$\int_{\star} f\, d\mu \equiv \sup\left\{\int g\, d\mu \mid g \text{ a } \mu\text{-integrable simple function with } g \leq f\ \mu \text{ a.e.}\right\}.$$

DEFINITION *A μ-measurable function $f : X \to [-\infty, \infty]$ is called* ***μ-integrable*** *if $\int^{\star} f\, d\mu = \int_{\star} f\, d\mu$, in which case we write*

$$\int f\, d\mu \equiv \int^{\star} f\, d\mu = \int_{\star} f\, d\mu.$$

Warning: Our use of the term "integrable" differs from most texts. For us, a function is "integrable" provided it has an integral, even if this integral equals $+\infty$ or $-\infty$.

REMARK Note that a nonnegative μ-measurable function is always μ-integrable. ∎

We assume the reader to be familiar with all the usual properties of integrals.

DEFINITIONS

(i) A function $f : X \to [-\infty, \infty]$ is ***μ-summable*** *if f is μ-integrable and*

$$\int |f|\, d\mu < \infty.$$

(ii) We say a function $f : \mathbb{R}^n \to [-\infty, \infty]$ is ***locally μ-summable*** *if $f\mid_K$ is μ-summable for each compact set $K \subset \mathbb{R}^n$.*

DEFINITION *We say ν is a* ***signed measure*** *on $\mathbb{R}^n$ if there exists a Radon measure μ on $\mathbb{R}^n$ and a locally μ-summable function $f : \mathbb{R}^n \to [-\infty, \infty]$ such that*

$$\nu(K) = \int_K f\, d\mu \qquad (\star)$$

for all compact sets $K \subset \mathbb{R}^n$.

NOTATION

(i) We write

$$\nu = \mu \llcorner f$$

provided ($\star$) holds for all compact sets K. Note $\mu \llcorner A = \mu \llcorner \chi_A$.

(ii) We denote by

$$L^1(X, \mu)$$

the set of all μ-summable functions on X, and

$$L^1_{\text{loc}}(\mathbb{R}^n, \mu)$$

the set of all locally μ-summable functions.

The following limit theorems are among the most important assertions in all of analysis.

THEOREM 1 FATOU'S LEMMA
Let $f_k : X \to [0, \infty]$ be μ-measurable $(k = 1, \ldots)$. Then

$$\int \liminf_{k\to\infty} f_k \, d\mu \le \liminf_{k\to\infty} \int f_k \, d\mu.$$

PROOF Take $g \equiv \sum_{j=1}^{\infty} a_j \chi_{A_j}$ to be a nonnegative simple function less than or equal to $\liminf_{k\to\infty} f_k$, and suppose the μ-measurable sets $\{A_j\}_{j=1}^{\infty}$ are disjoint and $a_j > 0$ for $j = 1, \ldots$. Fix $0 < t < 1$. Then

$$A_j = \bigcup_{k=1}^{\infty} B_{j,k},$$

where

$$B_{j,k} \equiv A_j \cap \{x \mid f_l(x) > t a_j \text{ for all } l \ge k\}.$$

Note

$$A_j \supset B_{j,k+1} \supset B_{j,k} \qquad (k = 1, \ldots).$$

Thus

$$\begin{aligned}\int f_k \, d\mu &\ge \sum_{j=1}^{\infty} \int_{A_j} f_k \, d\mu \\ &\ge \sum_{j=1}^{\infty} \int_{B_{j,k}} f_k \, d\mu \ge t \sum_{j=1}^{\infty} a_j \mu(B_{j,k}),\end{aligned}$$

and so

$$\liminf_{k\to\infty}\int f_k\,d\mu \geq t\sum_{j=1}^{\infty}a_j\mu(A_j)$$
$$= t\int g\,d\mu.$$

This estimate holds for each $0 < t < 1$ and each simple function g less than or equal to $\liminf_{k\to\infty} f_k$. Consequently,

$$\liminf_{k\to\infty}\int f_k\,d\mu \geq \int_* \liminf_{k\to\infty} f_k\,d\mu = \int \liminf_{k\to\infty} f_k\,d\mu. \quad \blacksquare$$

THEOREM 2 MONOTONE CONVERGENCE THEOREM
Let $f_k : X \to [0,\infty]$ be μ-measurable $(k = 1,\ldots)$, with $f_1 \leq \ldots \leq f_k \leq f_{k+1} \leq \ldots$. Then

$$\int \lim_{k\to\infty} f_k\,d\mu = \lim_{k\to\infty}\int f_k\,d\mu.$$

PROOF Clearly,

$$\int f_j\,d\mu \leq \int \lim_{k\to\infty} f_k\,d\mu \qquad (j = 1,\ldots),$$

whence

$$\lim_{k\to\infty}\int f_k\,d\mu \leq \int \lim_{k\to\infty} f_k\,d\mu.$$

The opposite inequality follows from Fatou's Lemma. $\blacksquare$

THEOREM 3 DOMINATED CONVERGENCE THEOREM
Let g be μ-summable and f, $\{f_k\}_{k=1}^{\infty}$ be μ-measurable. Suppose $|f_k| \leq g$ and $f_k \to f$ μ a.e. as $k \to \infty$. Then

$$\lim_{k\to\infty}\int |f_k - f|\,d\mu = 0.$$

PROOF By Fatou's Lemma,

$$\int 2g\,d\mu = \int \liminf_{k\to\infty}(2g - |f - f_k|)\,d\mu \leq \liminf_{k\to\infty}\int 2g - |f - f_k|\,d\mu,$$

whence

$$\limsup \int |f - f_k|\,d\mu \leq 0. \quad \blacksquare$$

THEOREM 4 VARIANT OF DOMINATED CONVERGENCE THEOREM
Let g, $\{g_k\}_{k=1}^{\infty}$ be μ-summable and f, $\{f_k\}_{k=1}^{\infty}$ μ-measurable. Suppose $|f_k| \leq g_k$ $(k = 1, \ldots)$, $f_k \to f$ μ a.e., and

$$\lim_{k\to\infty} \int g_k \, d\mu = \int g \, d\mu.$$

Then

$$\lim_{k\to\infty} \int |f_k - f| \, d\mu = 0.$$

PROOF Similar to proof of Theorem 3. ∎

It is easy to verify that $\lim_{k\to\infty} \int |f_k - f| \, d\mu = 0$ does *not* necessarily imply $f_k \to f$ μ a.e. But if we pass to an appropriate subsequence we do obtain a.e. convergence.

THEOREM 5
Assume f, $\{f_k\}_{k=1}^{\infty}$ are μ-summable and

$$\lim_{k\to\infty} \int |f_k - f| \, d\mu = 0.$$

Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ such that

$$f_{k_j} \to f \qquad \mu \text{ a.e.}$$

PROOF We select a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ of the functions $\{f_k\}_{k=1}^{\infty}$ satisfying

$$\sum_{j=1}^{\infty} \int |f_{k_j} - f| \, d\mu < \infty. \tag{$\star$}$$

Fix $\epsilon > 0$. Then

$$\left\{ \limsup_{j\to\infty} |f_{k_j} - f| > \epsilon \right\} \subset \bigcap_{i=1}^{\infty} \bigcup_{j=i}^{\infty} \{|f_{k_j} - f| > \epsilon\}.$$

Hence

$$\begin{aligned} \mu\left(\left\{ \limsup_{j\to\infty} |f_{kj} - f| > \epsilon \right\}\right) &\leq \sum_{j=i}^{\infty} \mu\left(\{|f_{kj} - f| > \epsilon\}\right) \\ &\leq \frac{1}{\epsilon} \sum_{j=i}^{\infty} \int |f_{kj} - f| \, d\mu, \end{aligned}$$

for each $i = 1, \dots$. In view of ($\star$) therefore,

$$\mu\left(\left\{\limsup_{j\to\infty} |f_{kj} - f| > \epsilon\right\}\right) = 0$$

for each $\epsilon > 0$. ∎

1.4 Product measures, Fubini's Theorem, Lebesgue measure

Let X and Y be sets.

DEFINITION *Let μ be a measure on X and ν a measure on Y. We define the measure $\mu \times \nu : 2^{X\times Y} \to [0,\infty]$ by setting for each $S \subset X \times Y$:*

$$(\mu \times \nu)(S) \equiv \inf\left\{\sum_{i=1}^{\infty} \mu(A_i)\nu(B_i)\right\},$$

where the infimum is taken over all collections of μ-measurable sets $A_i \subset X$ and ν-measurable sets $B_i \subset Y$ ($i = 1, \dots$) such that

$$S \subset \bigcup_{i=1}^{\infty}(A_i \times B_i).$$

*The measure $\mu \times \nu$ is called the **product measure** of μ and ν.*

THEOREM 1 FUBINI'S THEOREM
Let μ be a measure on X and ν a measure on Y.

- ***(i)*** *Then $\mu \times \nu$ is a regular measure on $X \times Y$, even if μ and ν are not regular.*
- ***(ii)*** *If $A \subset X$ is μ-measurable and $B \subset Y$ is ν-measurable, then $A \times B$ is $(\mu \times \nu)$-measurable and $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$.*
- ***(iii)*** *If $S \subset X \times Y$ is σ-finite with respect to $\mu \times \nu$, then $S_y \equiv \{x \mid (x,y) \in S\}$ is μ-measurable for ν a.e. y, $S_x \equiv \{y \mid (x,y) \in S\}$ is ν-measurable for μ a.e. x, $\mu(S_y)$ is ν-integrable, and $\nu(S_x)$ is μ-integrable. Moreover,*

$$(\mu \times \nu)(S) = \int_Y \mu(S_y)\, d\nu(y) = \int_X \nu(S_x)\, d\mu(x).$$

- ***(iv)*** *If f is $(\mu \times \nu)$-integrable and f is σ-finite with respect to $\mu \times \nu$ (in particular, if f is $(\mu \times \nu)$-summable), then the mapping*

$$y \mapsto \int_X f(x,y)\, d\mu(x) \text{ is } \nu\text{-integrable},$$

the mapping

$$x \mapsto \int_Y f(x,y)\, d\nu(y) \text{ is } \mu\text{-integrable},$$

and

$$\int_{X\times Y} f\, d(\mu \times \nu) = \int_Y \left[\int_X f(x,y)\, d\mu(x)\right] d\nu(y)$$
$$= \int_X \left[\int_Y f(x,y)\, d\nu(y)\right] d\mu(x).$$

REMARK We will study in Section 3.4 the Coarea Formula, which is a kind of "curvilinear" version of Fubini's Theorem. ▮

PROOF

1. Let $\mathcal{F}$ denote the collection of all sets $S \subset X \times Y$ for which the mapping

$$x \mapsto \chi_S(x,y)$$

is μ-integrable for each $y \in Y$ and the mapping

$$y \mapsto \int_X \chi_S(x,y)\, d\mu(x)$$

is ν-integrable. For $S \in \mathcal{F}$ we write

$$\rho(S) \equiv \int_Y \left[\int_X \chi_S(x,y)\, d\mu(x)\right] d\nu(y).$$

Define

$$\mathcal{P}_0 \equiv \{A \times B \mid A\ \mu\text{-measurable}, B\ \nu\text{-measurable}\},$$
$$\mathcal{P}_1 \equiv \{\cup_{j=1}^\infty S_j \mid S_j \in \mathcal{P}_0\},$$
$$\mathcal{P}_2 \equiv \{\cap_{j=1}^\infty S_j \mid S_j \in \mathcal{P}_1\}.$$

Note $\mathcal{P}_0 \subset \mathcal{F}$ and

$$\rho(A \times B) = \mu(A)\nu(B)\ \ (A \times B \in \mathcal{P}_0).$$

If $A_1 \times B_1, A_2 \times B_2 \in \mathcal{P}_0$, then

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2) \in \mathcal{P}_0,$$

and

$$(A_1 \times B_1) - (A_2 \times B_2) = ((A_1 - A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 - B_2))$$

is a disjoint union of members of $\mathcal{P}_0$. It follows that each member of $\mathcal{P}_1$ is a countable disjoint union of members of $\mathcal{P}_0$ and hence $\mathcal{P}_1 \subset \mathcal{F}$.

2. *Claim #1*: For each $S \subset X \times Y$,

$$(\mu \times \nu)(S) = \inf\{\rho(R) \mid S \subset R \in \mathcal{P}_1\}.$$

Proof of Claim #1: First we note that if $S \subset R = \cup_{i=1}^{\infty}(A_i \times B_i)$, then

$$\rho(R) \leq \sum_{i=1}^{\infty} \rho(A_i \times B_i) = \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i).$$

Thus

$$\inf\{\rho(R) \mid S \subset R \in \mathcal{P}_1\} \leq (\mu \times \nu)(S).$$

Moreover, there exists a disjoint sequence $\{A_j{}' \times B_j{}'\}_{j=1}^{\infty}$ in $\mathcal{P}_0$ such that

$$R = \bigcup_{j=1}^{\infty}(A_j{}' \times B_j{}').$$

Thus

$$\rho(R) = \sum_{j=1}^{\infty} \mu(A_j{}')\nu(B_j{}') \geq (\mu \times \nu)(S).$$

3. Fix $A \times B \in \mathcal{P}_0$. Then

$$(\mu \times \nu)(A \times B) \leq \mu(A)\nu(B) = \rho(A \times B) \leq \rho(R)$$

for all $R \in \mathcal{P}_1$ such that $A \times B \subset R$. Thus Claim #1 implies

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B).$$

Next we must prove $A \times B$ is $(\mu \times \nu)$-measurable. So suppose $T \subset X \times Y$ and $T \subset R \in \mathcal{P}_1$. Then $R - (A \times B)$ and $R \cap (A \times B)$ are disjoint and belong to $\mathcal{P}_1$. Consequently,

$$\begin{aligned}(\mu \times \nu)(T - (A \times B)) &+ (\mu \times \nu)(T \cap (A \times B)) \\ &\leq \rho(R - (A \times B)) + \rho(R \cap (A \times B)) \\ &= \rho(R),\end{aligned}$$

and so, according to Claim #1,

$$(\mu \times \nu)(T - (A \times B)) + (\mu \times \nu)(T \cap (A \times B)) \leq (\mu \times \nu)(T).$$

Thus $(A \times B)$ is $(\mu \times \nu)$-measurable. This proves (ii).

4. *Claim #2*: For each $S \subset X \times Y$ there is a set $R \in \mathcal{P}_2$ such that $S \subset R$ and

$$\rho(R) = (\mu \times \nu)(S).$$

Proof of Claim #2: If $(\mu \times \nu)(S) = \infty$, set $R \equiv X \times Y$. If $(\mu \times \nu)(S) < \infty$, then for each $j = 1, 2, \ldots$ there is according to Claim #1 a set $R_j \in \mathcal{P}_1$ such that $S \subset R_j$ and

$$\rho(R_j) < (\mu \times \nu)(S) + \frac{1}{j}\ .$$

Define

$$R \equiv \bigcap_{j=1}^{\infty} R_j \in \mathcal{P}_2.$$

Then $R \in \mathcal{F}$, and by the Dominated Convergence Theorem,

$$(\mu \times \nu)(S) \leq \rho(R) = \lim_{k\to\infty} \rho\left(\bigcap_{j=1}^{k} R_j\right) \leq (\mu \times \nu)(S).$$

5. From (ii) we see that every member of $\mathcal{P}_2$ is $(\mu \times \nu)$-measurable and thus (i) follows from Claim #2.

6. If $S \subset X \times Y$ and $(\mu \times \nu)(S) = 0$, then there is a set $R \in \mathcal{P}_2$ such that $S \subset R$ and $\rho(R) = 0$; thus $S \in \mathcal{F}$ and $\rho(S) = 0$.

Now suppose that $S \subset X \times Y$ is $(\mu \times \nu)$-measurable and $(\mu \times \nu)(S) < \infty$. Then there is a $R \in \mathcal{P}_2$ such that $S \subset R$ and

$$(\mu \times \nu)(R - S) = 0;$$

hence

$$\rho(R - S) = 0.$$

Thus

$$\mu\{x \mid (x, y) \in S\} = \mu\{x \mid (x, y) \in R\}$$

for ν a.e. $y \in Y$, and

$$(\mu \times \nu)(S) = \rho(R) = \int \mu\{x \mid (x, y) \in S\}\, d\nu(y).$$

Assertion (iii) follows.

7. Assertion (iv) reduces to (iii) when $f = \chi_S$. If $f \geq 0$, is $(\mu \times \nu)$-integrable and is σ-finite with respect to $\mu \times \nu$, we use Theorem 7, Section 1.1.2, to write

$$f = \sum_{k=1}^{\infty} \frac{1}{k} \chi_{A_k}$$

and note (iv) results for such an f from the Monotone Convergence Theorem. Finally, for general f we write

$$f = f^+ - f^-,$$

and (iv) follows. ∎

DEFINITION ***One-dimensional Lebesgue measure*** $\mathcal{L}^1$ *on* $\mathbb{R}^1$ *is defined by*

$$\overline{\mathcal{L}^1(A)} \equiv \inf\left\{\sum_{i=1}^{\infty} \operatorname{diam} C_i \mid A \subset \bigcup_{i=1}^{\infty} C_i, C_i \subset \mathbb{R}\right\},$$

for all $A \subset \mathbb{R}$.

DEFINITION *We inductively define* n***-dimensional Lebesgue measure*** $\mathcal{L}^n$ *on* $\mathbb{R}^n$ *by*

$$\mathcal{L}^n \equiv \mathcal{L}^{n-1} \times \mathcal{L}^1 = \mathcal{L}^1 \times \cdots \times \mathcal{L}^1 (n \text{ times})$$

Equivalently $\mathcal{L}^n = \mathcal{L}^{n-k} \times \mathcal{L}^k$ *for each* $k \in \{1, \ldots, n-1\}$.

We assume the reader's familiarity with all the usual facts about $\mathcal{L}^n$.

NOTATION We will write "dx," "dy," etc. rather than "$d\mathcal{L}^n$" in integrals taken with respect to $\mathcal{L}^n$. We also write $L^1(\mathbb{R}^n)$ for $L^1(\mathbb{R}^n, \mathcal{L}^n)$, etc.

1.5 Covering theorems

We present in this section the fundamental covering theorems of Vitali and of Besicovitch. Vitali's Covering Theorem is easier and is most useful for investigating $\mathcal{L}^n$ on $\mathbb{R}^n$. Besicovitch's Covering Theorem is much harder to prove, but it is necessary for studying arbitrary Radon measures μ on $\mathbb{R}^n$. The crucial geometric difference is that Vitali's Covering Theorem provides a cover of enlarged balls, whereas Besicovitch's Covering Theorem yields a cover out of the original balls, at the price of a certain (controlled) amount of overlap.

These covering theorems will be employed throughout the rest of these notes, the first and most important application being to the differentiation theorems in Section 1.6.

1.5.1 Vitali's Covering Theorem

NOTATION If B is a closed ball in $\mathbb{R}^n$, we write $\hat{B}$ to denote the concentric closed ball with radius 5 times the radius of B.

DEFINITIONS

(i) *A collection $\mathcal{F}$ of closed balls in $\mathbb{R}^n$ is a* **cover** *of a set $A \subset \mathbb{R}^n$ if*

$$A \subset \bigcup_{B \in \mathcal{F}} B.$$

(ii) *$\mathcal{F}$ is a* **fine cover** *of A if, in addition,*

$$\inf\{\text{diam } B \mid x \in B, B \in \mathcal{F}\} = 0$$

for each $x \in A$.

THEOREM 1 VITALI'S COVERING THEOREM
Let $\mathcal{F}$ be any collection of nondegenerate closed balls in $\mathbb{R}^n$ with

$$\sup\{\text{diam } B \mid B \in \mathcal{F}\} < \infty.$$

Then there exists a countable family $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

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

PROOF
1. Write $D \equiv \sup\{\text{diam } B \mid B \in \mathcal{F}\}$. Set $\mathcal{F}_j \equiv \{B \in \mathcal{F} \mid D/2^j < \text{diam } B \leq D/2^{j-1}\}$, $(j = 1, 2, \ldots)$. We define $\mathcal{G}_j \subset \mathcal{F}_j$ as follows:

(a) Let $\mathcal{G}_1$ be any maximal disjoint collection of balls in $\mathcal{F}_1$.

(b) Assuming $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_{k-1}$ have been selected, we choose $\mathcal{G}_k$ to be any maximal disjoint subcollection of

$$\left\{ B \in \mathcal{F}_k \mid B \cap B' = \emptyset \text{ for all } B' \in \bigcup_{j=1}^{k-1} \mathcal{G}_j \right\}.$$

Finally, define $\mathcal{G} \equiv \cup_{j=1}^{\infty} \mathcal{G}_j$. Clearly $\mathcal{G}$ is a collection of disjoint balls and $\mathcal{G} \subset \mathcal{F}$.

2. Claim: For each $B \in \mathcal{F}$, there exists a ball $B' \in \mathcal{G}$ so that $B \cap B' \neq \emptyset$ and $B \subset \hat{B}'$.

Proof of Claim: Fix $B \in \mathcal{F}$. There then exists an index j such that $B \in \mathcal{F}_j$. By the maximality of $\mathcal{G}_j$, there exists a ball $B' \in \cup_{k=1}^{j} \mathcal{G}_k$ with $B \cap B' \neq \emptyset$. But diam $B' \geq D/2^j$ and diam $B \leq D/2^{j-1}$, so that diam $B \leq 2$ diam B'. Thus $B \subset \hat{B}'$, as claimed. ▮

A technical consequence we will use later is this:

COROLLARY 1
Assume that $\mathcal{F}$ is a fine cover of A by closed balls and

$$\sup\{\operatorname{diam} B \mid B \in \mathcal{F}\} < \infty.$$

Then there exists a countable family $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that for each finite subset $\{B_1, \ldots, B_m\} \subset \mathcal{F}$, we have

$$A - \bigcup_{k=1}^{m} B_k \subset \bigcup_{B \in \mathcal{G} - \{B_1, \ldots, B_m\}} \hat{B}$$

PROOF Choose $\mathcal{G}$ as in the proof of the Vitali Covering Theorem and select $\{B_1, \ldots, B_m\} \subset \mathcal{F}$. If $A \subset \cup_{k=1}^{m} B_k$, we are done. Otherwise, let $x \in A - \cup_{k=1}^{m} B_k$. Since the balls in $\mathcal{F}$ are closed and $\mathcal{F}$ is a fine cover, there exists $B \in \mathcal{F}$ with $x \in B$ and $B \cap B_k = \emptyset$ $(k = 1, \ldots, m)$. But then, from the claim in the proof above, there exists a ball $B' \in \mathcal{G}$ such that $B \cap B' \neq \emptyset$ and $B \subset \hat{B}'$. ∎

Next we show we can measure theoretically "fill up" an arbitrary open set with countably many disjoint closed balls.

COROLLARY 2
Let $U \subset \mathbb{R}^n$ be open, $\delta > 0$. There exists a countable collection $\mathcal{G}$ of disjoint closed balls in U such that diam $B \leq \delta$ *for all $B \in \mathcal{G}$ and*

$$\mathcal{L}^n\left(U - \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

PROOF

1. Fix $1 - 1/5^n < \theta < 1$. Assume first $\mathcal{L}^n(U) < \infty$.

2. Claim: There exists a finite collection $\{B_i\}_{i=1}^{M_1}$ of disjoint closed balls in U such that diam $(B_i) < \delta$ $(i = 1, \ldots, M_1)$, and

$$\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta \mathcal{L}^n(U). \qquad (\star)$$

Proof of Claim: Let $\mathcal{F}_1 \equiv \{B \mid B \subset U, \operatorname{diam} B < \delta\}$. By Theorem 1, there exists a countable disjoint family $\mathcal{G}_1 \subset \mathcal{F}_1$ such that

$$U \subset \bigcup_{B \in \mathcal{G}_1} \hat{B}.$$

Thus

$$\begin{aligned}\mathcal{L}^n(U) &\le \sum_{B\in\mathcal{G}_1} \mathcal{L}^n(\hat{B})\\ &= 5^n \sum_{B\in\mathcal{G}_1} L_n(B)\\ &= 5^n\mathcal{L}^n\left(\bigcup_{B\in\mathcal{G}_1} B\right).\end{aligned}$$

Hence

$$\mathcal{L}^n\left(\bigcup_{B\in\mathcal{G}_1} B\right) \ge \frac{1}{5^n}\mathcal{L}^n(U),$$

so that

$$\mathcal{L}^n\left(U - \bigcup_{B\in\mathcal{G}_1} B\right) \le \left(1-\frac{1}{5^n}\right)\mathcal{L}^n(U).$$

Since $\mathcal{G}_1$ is countable, there exist balls $B_1,\dots,B_{M_1}$ in $\mathcal{G}_1$ satisfying $(\star)$.

3. Now let

$$\begin{aligned}U_2 &\equiv U - \bigcup_{i=1}^{M_1} B_i,\\ \mathcal{F}_2 &\equiv \{B \mid B\subset U_2, \text{diam } B < \delta\},\end{aligned}$$

and, as above, find finitely many disjoint balls $B_{M_1+1},\dots,B_{M_2}$ in $\mathcal{F}_2$ such that

$$\begin{aligned}\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_2} B_i\right) &= \mathcal{L}^n\left(U_2 - \bigcup_{i=M_1+1}^{M_2} B_i\right)\\ &\le \theta\mathcal{L}^n(U_2)\\ &\le \theta^2\mathcal{L}^n(U).\end{aligned}$$

4. Continue this process to obtain a countable collection of disjoint balls such that

$$\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_k} B_i\right) \le \theta^k\mathcal{L}^n(U) \quad (k=1,\dots).$$

Since $\theta^k \to 0$, the corollary is proved if $\mathcal{L}^n(U)<\infty$. Should $\mathcal{L}^n(U)=\infty$, we apply the above reasoning to the sets

$$U_m \equiv \{x\in U \mid m<|x|<m+1\} \qquad (m=0,1,\dots). \quad \blacksquare$$

REMARK See Corollary 1 in the next section, which replaces $\mathcal{L}^n$ in the preceding proof by an arbitrary Radon measure. ∎

1.5.2 Besicovitch's Covering Theorem

If μ is an arbitrary Radon measure on $\mathbb{R}^n$, there is no systematic way to control $\mu(\hat{B})$ in terms of $\mu(B)$. In studying such a measure, Vitali's Covering Theorem is not useful; we need instead a covering theorem that does not require us to enlarge balls.

THEOREM 2 BESICOVITCH'S COVERING THEOREM
There exists a constant N_n, depending only on n, with the following property: If $\mathcal{F}$ is any collection of nondegenerate closed balls in $\mathbb{R}^n$ with

$$\sup\{\text{diam } B \mid B \in \mathcal{F}\} < \infty$$

and if A is the set of centers of balls in $\mathcal{F}$, then there exist $\mathcal{G}_1, \ldots, \mathcal{G}_{N_n} \subset \mathcal{F}$ such that each $\mathcal{G}_i$ $(i = 1, \ldots, N_n)$ is a countable collection of disjoint balls in $\mathcal{F}$ and

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

PROOF
1. First suppose A is bounded. Write $D \equiv \sup\{\text{diam } B \mid B \in \mathcal{F}\}$. Choose any ball $B_1 = B(a_1, r_1) \in \mathcal{F}$ such that $r_1 \geq (3/4)D/2$. Inductively choose B_j, $j \geq 2$, as follows. Let $A_j \equiv A - \cup_{i=1}^{j-1} B_i$. If $A_j = \emptyset$, stop and set $J \equiv j - 1$. If $A_j \neq \emptyset$, choose $B_j = B(a_j, r_j) \in \mathcal{F}$ such that $a_j \in A_j$ and $r_j \geq 3/4 \sup\{r \mid B(a, r) \in \mathcal{F}, a \in A_j\}$. If $A_j \neq \emptyset$ for all j, set $J \equiv \infty$.
2. *Claim #1*: If $j > i$, then $r_j \leq (4/3)r_i$.

Proof of Claim #1: Suppose $j > i$. Then $a_j \in A_i$ and so

$$r_i \geq \frac{3}{4} \sup\{r \mid B(A, r) \in \mathcal{F}, a \in A_i\} \geq \frac{3}{4} r_j.$$

3. Claim #2: The balls $\{B(a_j, r_j/3)\}_{j=1}^{J}$ are disjoint.

Proof of Claim #2: Let $j > i$. Then $a_j \notin B_i$; hence

$$|a_i - a_j| > r_i = \frac{r_i}{3} + \frac{2r_i}{3} \geq \frac{r_i}{3} + \left(\frac{2}{3}\right)\left(\frac{3}{4}\right) r_j > \frac{r_i}{3} + \frac{r_j}{3}.$$

4. *Claim #3*: If $J = \infty$, then $\lim_{j\to\infty} r_j = 0$.

Proof of Claim #3: By Claim #2 the balls $\{B(a_j, r_j/3)\}_{j=1}^{J}$ are disjoint. Since $a_j \in A$ and A is bounded, $r_j \to 0$.

5. *Claim #4*: $A \subset \cup_{j=1}^{J} B_j$.

Proof of Claim #4: If $J < \infty$, this is trivial. Suppose $J = \infty$. If $a \in A$, there exists an $r > 0$ such that $B(a,r) \in \mathcal{F}$. Then by Claim #3, there exists an r_j with $r_j < (3/4)r$, a contradiction to the choice of r_j if $a \notin \cup_{i=1}^{j-1} B_i$.

6. Fix $k > 1$ and let $I \equiv \{j \mid 1 \le j < k, B_j \cap B_k \ne \emptyset\}$. We need to estimate the cardinality of I. Set $K \equiv I \cap \{j \mid r_j \le 3r_k\}$.

7. *Claim #5*: Card $(K) \le 20^n$.

Proof of Claim #5: Let $j \in K$. Then $B_j \cap B_k \ne \emptyset$ and $r_j \le 3r_k$. Choose any $x \in B(a_j, r_j/3)$. Then

$$\begin{aligned} |x - a_k| &\le |x - a_j| + |a_j - a_k| \le \frac{r_j}{3} + r_j + r_k \\ &= \frac{4}{3} r_j + r_k \le 4r_k + r_k = 5r_k, \end{aligned}$$

so that $B(a_j, r_j/3) \subset B(a_k, 5r_k)$. Recall from Claim #2 that the balls $B(a_i, r_i/3)$ are disjoint. Thus

$$\begin{aligned} \alpha(n) 5^n {r_k}^n &= \mathcal{L}^n(B(a_k, 5r_k)) \\ &\ge \sum_{j\in K} \mathcal{L}^n(B(a_j, \frac{r_j}{3})) \\ &= \sum_{j\in K} \alpha(n) \left(\frac{r_j}{3}\right)^n \\ &\ge \sum_{j\in K} \alpha(n) \left(\frac{r_k}{4}\right)^n \qquad \text{by Claim \#1} \\ &= \text{Card } (K) \alpha(n) \frac{{r_k}^n}{4^n} \ . \end{aligned}$$

Consequently,

$$5^n \ge \text{Card } (K) \frac{1}{4^n} \ .$$

8. We must now estimate Card $(I - K)$.

Let $i, j \in I - K$, with $i \ne j$. Then $1 \le i, j < k$, $B_i \cap B_k \ne \emptyset$, $B_j \cap B_k \ne \emptyset$, $r_i > 3r_k$, $r_j > 3r_k$. For simplicity of notation, we take (without loss of generality) $a_k = 0$. Let $0 \le \theta \le \pi$ be the angle between the vectors a_i and a_j.

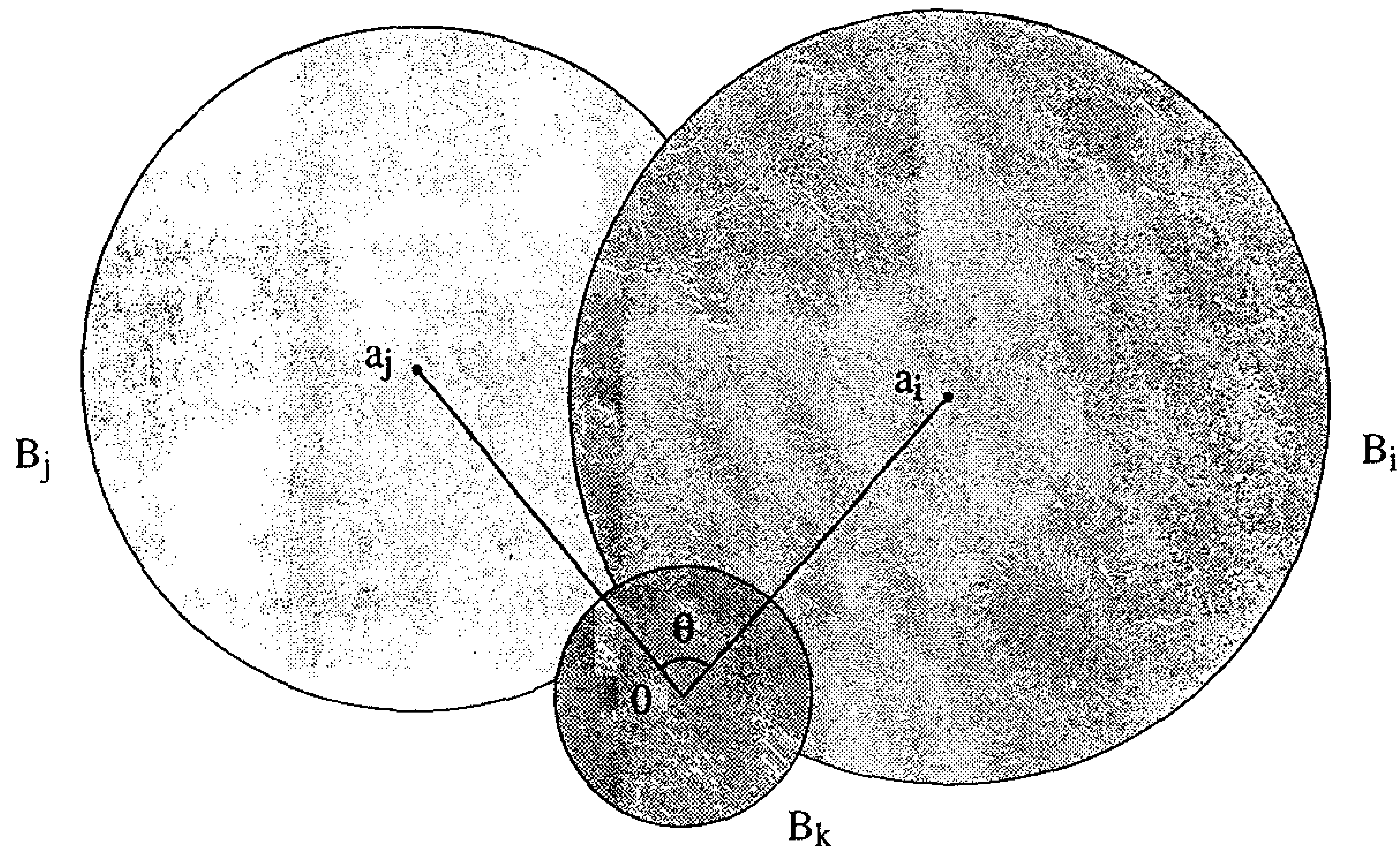

FIGURE 1.1
Illustration of Claim #6.

We intend to find a lower bound on θ, and to this end we first assemble some facts:

Since $i, j < k$, $0 = a_k \notin B_i \cup B_j$. Thus $r_i < |a_i|$ and $r_j < |a_j|$. Since $B_i \cap B_k \neq \emptyset$ and $B_j \cap B_k \neq \emptyset$, $|a_i| \leq r_i + r_k$ and $|a_j| \leq r_j + r_k$. Finally, without loss of generality we can take $|a_i| \leq |a_j|$. In summary,

$$\begin{cases} 3r_k < r_i < |a_i| \leq r_i + r_k \\ 3r_k < r_j < |a_j| \leq r_j + r_k \\ |a_i| \leq |a_j|. \end{cases}$$

9. *Claim #6a*: If $\cos\theta > 5/6$, then $a_i \in B_j$.

Proof of Claim #6a: Suppose $|a_i - a_j| \geq |a_j|$; then the Law of Cosines gives

$$\begin{aligned} \cos\theta &= \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \\ &\leq \frac{|a_i|^2}{2|a_i||a_j|} = \frac{|a_i|}{2|a_j|} \leq \frac{1}{2} < \frac{5}{6} . \end{aligned}$$

Suppose now instead $|a_i - a_j| \le |a_j|$ and $a_i \notin B_j$. Then $r_j < |a_i - a_j|$ and

$$\begin{aligned}
\cos\theta &= \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \\
&= \frac{|a_i|}{2|a_j|} + \frac{(|a_j| - |a_i - a_j|)(|a_j| + |a_i - a_j|)}{2|a_i||a_j|} \\
&\le \frac{1}{2} + \frac{(|a_j| - |a_i - a_j|)(2|a_j|)}{2|a_i||a_j|} \\
&\le \frac{1}{2} + \frac{r_j + r_k - r_j}{r_i} = \frac{1}{2} + \frac{r_k}{r_i} \le \frac{5}{6}\,.
\end{aligned}$$

10. Claim #6b: If $a_i \in B_j$, then

$$0 \le |a_i - a_j| + |a_i| - |a_j| \le |a_j|\epsilon(\theta),$$

for

$$\epsilon(\theta) \equiv \frac{8}{3}(1 - \cos\theta).$$

Proof of Claim #6b: Since $a_i \in B_j$, we must have $i < j$; hence $a_j \notin B_i$ and so $|a_i - a_j| > r_i$. Thus

$$\begin{aligned}
0 &\le \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \\
&\le \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \cdot \frac{|a_i - a_j| - |a_i| + |a_j|}{|a_i - a_j|} \\
&= \frac{|a_i - a_j|^2 - (|a_j| - |a_i|)^2}{|a_j||a_i - a_j|} \\
&= \frac{|a_i|^2 + |a_j|^2 - 2|a_i||a_j|\cos\theta - |a_i|^2 - |a_j|^2 + 2|a_i||a_j|}{|a_j||a_i - a_j|} \\
&= \frac{2|a_i|(1 - \cos\theta)}{|a_i - a_j|} \\
&\le \frac{2(r_i + r_k)(1 - \cos\theta)}{r_i} \\
&\le \frac{2(1 + \frac{1}{3})r_i(1 - \cos\theta)}{r_i} = \epsilon(\theta).
\end{aligned}$$

11. Claim #6c: If $a_i \in B_j$, then $\cos\theta \le 61/64$.

Proof of Claim #6c: Since $a_i \in B_j$ and $a_j \notin B_i$, we have $r_i < |a_i - a_j| \le r_j$. Since $i < j$, $r_j \le (4/3)r_i$. Therefore,

$$\begin{aligned}
|a_i - a_j| + |a_i| - |a_j| &\ge r_i + r_i - r_j - r_k \\
&\ge \frac{3}{2} r_j - r_j - r_k \\
&= \frac{1}{2} r_j - r_k \ge \frac{1}{6} r_j \\
&= \frac{1}{6}\left(\frac{3}{4}\right)\left(r_j + \frac{1}{3} r_j\right) \ge \frac{1}{8}(r_j + r_k) \\
&\ge \frac{1}{8}|a_j|.
\end{aligned}$$

Then, by Claim #6b,

$$\frac{1}{8}|a_j| \le |a_i - a_j| + |a_i| - |a_j| \le |a_j|\epsilon(\theta).$$

Hence $\cos\theta \le 61/64$.

12. We combine Claims #6a–c to obtain

Claim #6: For all $i, j \in I - K$ with $i \neq j$, let θ denote the angle between $a_i - a_k$ and $a_j - a_k$. Then $\theta \ge \arccos 61/64 \equiv \theta_0 > 0$.

13. Claim #7: There exists a constant L_n depending only on n such that Card $(I - K) \le L_n$.

Proof of Claim #7: First, fix $r_0 > 0$ such that if $x \in \partial B(0,1)$ and $y, z \in B(x, r_0)$, then the angle between y and z is less than the constant θ_0 from Claim #6. Choose L_n so that $\partial B(0,1)$ can be covered by L_n balls with radius r_0 and centers on $\partial B(0,1)$, but cannot be covered by $L_n - 1$ such balls.

Then ∂B_k can be covered by L_n balls of radius $r_0 r_k$, with centers on ∂B_k. By Claim #6, if $i, j \in I - K$ with $i \neq j$, then the angle between $a_i - a_k$ and $a_j - a_k$ exceeds θ_0. Thus by the construction of r_0, the rays $a_j - a_k$ and $a_i - a_k$ cannot both go through the same ball on ∂B_k. Consequently, Card $(I - K) \le L_n$.

14. Finally, set $M_n \equiv 20^n + L_n + 1$. Then by Claims #5 and #7,

$$\begin{aligned}
\text{Card }(I) &= \text{Card }(K) + \text{Card }(I - K) \\
&\le 20^n + L_n < M_n.
\end{aligned}$$

15. We next define $\mathcal{G}_1, \ldots, \mathcal{G}_{M_n}$.

First define $\sigma : \{1, 2, \ldots\} \to \{1, \ldots, M_n\}$ as follows:

(a) Let $\sigma(i) = i$ for $1 \le i \le M_n$.

(b) For $k \ge M_n$ inductively define $\sigma(k+1)$ as follows. According to the

calculations above,

$$\text{Card}\,\{j \mid 1 \leq j \leq k, B_j \cap B_{k+1} \neq \emptyset\} < M_n,$$

so there exists $l \in \{1, \ldots, M_n\}$ such that $B_{k+1} \cap B_j = \emptyset$ for all j such that $\sigma(j) = l$ $(1 \leq j \leq k)$. Set $\sigma(k+1) = l$.

Now, let $\mathcal{G}_j \equiv \{B_i \mid \sigma(i) = j\}$, $1 \leq j \leq M_n$. By the construction of $\sigma(i)$, each $\mathcal{G}_j$ consists of disjoint balls from $\mathcal{F}$. Moreover, each B_i is in some $\mathcal{G}_j$, so that

$$A \subset \bigcup_{i=1}^{J} B_i = \bigcup_{i=1}^{M_n} \bigcup_{B \in \mathcal{G}_i} B.$$

16. Next, we extend the result to general (unbounded) A.

For $l \geq 1$, let $A_l \equiv A \cap \{x \mid 3D(l-1) \leq |x| < 3Dl\}$ and set $\mathcal{F}^l \equiv \{B(a,r) \in \mathcal{F} \mid a \in A_l\}$. Then by step 15, there exist countable collections $\mathcal{G}_1^l, \ldots, \mathcal{G}_{M_n}^l$ of disjoint closed balls in $\mathcal{F}^l$ such that

$$A_l \subset \bigcup_{i=1}^{M_n} \bigcup_{B \in \mathcal{G}_i^l} B.$$

Let

$$\mathcal{G}_j \equiv \bigcup_{l=1}^{\infty} \mathcal{G}_j^{2l-1} \text{ for } 1 \leq j \leq M_n,$$

$$\mathcal{G}_{j+M_n} = \bigcup_{l=1}^{\infty} \mathcal{G}_j^{2l} \text{ for } 1 \leq j \leq M_n.$$

Set $N_n \equiv 2M_n$. ∎

We now see as a consequence of Besicovitch's Theorem that we can "fill up" an arbitrary open set with a countable collection of disjoint balls in such a way that the remainder has μ-measure zero.

COROLLARY 1

Let μ be a Borel measure on $\mathbb{R}^n$, and $\mathcal{F}$ any collection of nondegenerate closed balls. Let A denote the set of centers of the balls in $\mathcal{F}$. Assume $\mu(A) < \infty$ and $\inf\{r \mid B(a,r) \in \mathcal{F}\} = 0$ for each $a \in A$. Then for each open set $U \subset \mathbb{R}^n$, there exists a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

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

and

$$\mu\left((A \cap U) - \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

REMARK The set A need not be μ-measurable here. Compare this assertion with Corollary 2 of Vitali's Covering Theorem, above. ▮

PROOF Fix $1 - 1/N_n < \theta < 1$.

1. Claim: There exists a finite collection $\{B_1, \dots, B_{M_1}\}$ of disjoint closed balls in U such that

$$\mu\left((A \cap U) - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta\mu(A \cap U). \tag{$\star$}$$

Proof of Claim: Let $\mathcal{F}_1 = \{B \mid B \in \mathcal{F}, \text{diam } B \leq 1, B \subset U\}$. By Theorem 2, there exist families $\mathcal{G}_1, \dots, \mathcal{G}_{N_n}$ of disjoint balls in $\mathcal{F}_1$ such that

$$A \cap U \subset \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{G}_i} B.$$

Thus

$$\mu(A \cap U) \leq \sum_{i=1}^{N_n} \mu\left(A \cap U \bigcap \bigcup_{B \in \mathcal{G}_i} B\right).$$

Consequently, there exists an integer j between 1 and N_n for which

$$\mu\left(A \cap U \bigcap \bigcup_{B \in \mathcal{G}_j} B\right) \geq \frac{1}{N_n}\mu(A \cap U).$$

By Theorem 2 in Section 1.1, there exist balls $B_1, \dots, B_{M_1} \in \mathcal{G}_j$ such that

$$\mu\left(A \cap U \bigcap \bigcup_{i=1}^{M_1} B_i\right) \geq (1 - \theta)\mu(A \cap U).$$

But

$$\mu(A \cap U) = \mu\left(A \cap U \bigcap \bigcup_{i=1}^{M_1} B_i\right) + \mu\left(A \cap U - \bigcup_{i=1}^{M_1} B_i\right),$$

since $\cup_{i=1}^{M_1} B_i$ is μ-measurable, and hence ($\star$) holds.

2. Now let $U_2 \equiv U - \cup_{i=1}^{M_1} B_i$, $\mathcal{F}_2 \equiv \{B \mid B \in \mathcal{F}, \text{diam } B \leq 1, B \subset U_2\}$, and as above, find finitely many disjoint balls $B_{M_1+1}, \dots, B_{M_2}$ in $\mathcal{F}_2$ such that

$$\begin{aligned}\mu\left((A \cap U) - \bigcup_{i=1}^{M_2} B_i\right) &= \mu\left((A \cap U_2) - \bigcup_{i=M_1+1}^{M_2} B_i\right) \\ &\leq \theta\mu(A \cap U_2) \\ &\leq \theta^2\mu(A \cap U).\end{aligned}$$

3. Continue this process to obtain a countable collection of disjoint balls from $\mathcal{F}$ and within U such that

$$\mu\left((A \cap U) - \bigcup_{i=1}^{M_k} B_i\right) \leq \theta^k \mu(A \cap U).$$

Since $\theta^k \to 0$ and $\mu(A) < \infty$, the corollary is proved. ∎

1.6 Differentiation of Radon measures

We now utilize the covering theorems of the previous section to study the differentiation of Radon measures on $\mathbb{R}^n$.

1.6.1 Derivatives

Let μ and ν be Radon measures on $\mathbb{R}^n$.

DEFINITION *For each point $x \in \mathbb{R}^n$, define*

$$\overline{D}_\mu \nu(x) \equiv \begin{cases} \limsup_{r\to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0, \end{cases}$$

$$\underline{D}_\mu \nu(x) \equiv \begin{cases} \liminf_{r\to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0. \end{cases}$$

DEFINITION *If $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) < +\infty$, we say ν is* ***differentiable*** *with respect to μ at x and write*

$$D_\mu \nu(x) \equiv \overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x).$$

$D_\mu \nu$ is the ***derivative*** *of ν with respect to μ. We also call $D_\mu \nu$ the* ***density*** *of ν with respect to μ.*

Our goals are to study (a) when $D_\mu \nu$ exists and (b) when ν can be recovered by integrating $D_\mu \nu$.

LEMMA 1
Fix $0 < \alpha < \infty$. Then

(i) $A \subset \{x \in \mathbb{R}^n \mid \underline{D}_\mu \nu(x) \leq \alpha\}$ *implies* $\nu(A) \leq \alpha\mu(A)$,

(ii) $A \subset \{x \in \mathbb{R}^n \mid \overline{D}_\mu \nu(x) \geq \alpha\}$ *implies* $\nu(A) \geq \alpha\mu(A)$.

REMARK The set A need not be μ- or ν-measurable here. ∎

PROOF We may assume $\mu(\mathbb{R}^n)$, $\nu(\mathbb{R}^n) < \infty$, since we could otherwise consider μ and ν restricted to compact subsets of $\mathbb{R}^n$.

Fix $\epsilon > 0$. Let U be open, $A \subset U$, where A satisfies the hypothesis of (i). Set

$$\mathcal{F} \equiv \{B \mid B = B(a,r), a \in A, B \subset U, \nu(B) \leq (\alpha + \epsilon)\mu(B)\}.$$

Then $\inf\{r \mid B(a,r) \in \mathcal{F}\} = 0$ for each $a \in A$, and so Corollary 1 in Section 1.5.2 provides us with a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\nu\left(A - \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

Then

$$\nu(A) \leq \sum_{B \in \mathcal{G}} \nu(B) \leq (\alpha + \epsilon) \sum_{B \in \mathcal{G}} \mu(B) \leq (\alpha + \epsilon)\mu(U).$$

This estimate is valid for each open set $U \supset A$, so that Theorem 4 in Section 1.1 implies $\nu(A) \leq (\alpha + \epsilon)\mu(A)$. This proves (i). The proof of (ii) is similar. ∎

THEOREM 1

Let μ and ν be Radon measures on $\mathbb{R}^n$. Then $D_\mu \nu$ exists and is finite μ a.e. Furthermore, $D_\mu \nu$ is μ-measurable.

PROOF We may assume $\nu(\mathbb{R}^n), \mu(\mathbb{R}^n) < \infty$, as we could otherwise consider μ and ν restricted to compact subsets of $\mathbb{R}^n$.

1. Claim #1: $D_\mu \nu$ exists and is finite μ a.e.

Proof of Claim #1: Let $I \equiv \{x \mid \overline{D}_\mu \nu(x) = +\infty\}$, and for all $0 < a < b$, let $R(a,b) \equiv \{x \mid \underline{D}_\mu \nu(x) < a < b < \overline{D}_\mu \nu(x) < \infty\}$. Observe that for each $\alpha > 0$, $I \subset \{x \mid \overline{D}_\mu \nu(x) \geq \alpha\}$. Thus by Lemma 1,

$$\mu(I) \leq \frac{1}{\alpha}\nu(I).$$

Send $\alpha \to \infty$ to conclude $\mu(I) = 0$, and so $\overline{D}_\mu \nu$ is finite μ a.e.

Again using Lemma 1, we see

$$b\mu(R(a,b)) \leq \nu(R(a,b)) \leq a\mu(R(a,b)),$$

whence $\mu(R(a,b)) = 0$, since $b > a$. Furthermore,

$$\{x \mid \underline{D}_{\mu}\nu(x) < \overline{D}_{\mu}\nu(x) < \infty\} = \bigcup_{\substack{0<a<b \\ a,b \text{ rational}}} R(a,b),$$

and consequently $D_\mu \nu$ exists and is finite μ a.e.

2. *Claim #2*: For each $x \in \mathbb{R}^n$ and $r > 0$,

$$\limsup_{y \to x} \mu(B(y,r)) \leq \mu(B(x,r)).$$

A similar assertion holds for ν.

Proof of Claim #2: Choose $y_k \in \mathbb{R}^n$ with $y_k \to x$. Set $f_k \equiv \chi_{B(y_k,r)}$, $f = \chi_{B(x,r)}$. Then

$$\limsup_{k \to \infty} f_k \leq f$$

and so

$$\liminf_{k \to \infty} (1 - f_k) \geq (1 - f).$$

Thus by Fatou's Lemma,

$$\begin{aligned} \int_{B(x,2r)} (1-f)\, d\mu &\leq \int_{B(x,2r)} \liminf_{k\to\infty} (1-f_k)\, d\mu \\ &\leq \liminf_{k\to\infty} \int_{B(x,2r)} (1-f_k)\, d\mu, \end{aligned}$$

that is,

$$\mu(B(x,2r)) - \mu(B(x,r)) \leq \liminf_{k\to\infty} (\mu(B(x,2r)) - \mu(B(y_k,r))).$$

Now since μ is a Radon measure, $\mu(B(x,2r)) < \infty$; the assertion follows.

3. *Claim #3*: $D_\mu \nu$ is μ-measurable.

Proof of Claim #3: According to Claim #2, for all $r > 0$, the functions $x \mapsto \mu(B(x,r))$ and $x \mapsto \nu(B(x,r))$ are upper semicontinuous and thus Borel measurable. Consequently, for every $r > 0$,

$$f_r(x) \equiv \begin{cases} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \end{cases}$$

is μ-measurable. But

$$D_\mu \nu = \lim_{r\to 0} f_r = \lim_{k\to\infty} f_{\frac{1}{k}} \qquad \mu\text{- a.e.}$$

and so $D_\mu \nu$ is μ-measurable. ∎

1.6.2 Integration of derivatives; Lebesgue decomposition

DEFINITION *The measure ν is* ***absolutely continuous*** *with respect to μ, written*

$$\nu \ll \mu,$$

provided $\mu(A) = 0$ implies $\nu(A) = 0$ for all $A \subset \mathbb{R}^n$.

DEFINITION *The measures ν and μ are* ***mutually singular****, written*

$$\nu \perp \mu,$$

if there exists a Borel subset $B \subset \mathbb{R}^n$ such that

$$\mu(\mathbb{R}^n - B) = \nu(B) = 0.$$

THEOREM 2 DIFFERENTIATION THEOREM FOR RADON MEASURES
Let ν, μ be Radon measures on $\mathbb{R}^n$, with $\nu \ll \mu$. Then

$$\nu(A) = \int_A D_\mu \nu \, d\mu$$

for all μ-measurable sets $A \subset \mathbb{R}^n$.

REMARK This is a version of the ***Radon–Nikodym Theorem***. Observe we prove not only that ν has a density with respect to μ, but also that this density $D_\mu \nu$ can be computed by "differentiating" ν with respect to μ. These assertions comprise in effect the Fundamental Theorem of Calculus for Radon measures on $\mathbb{R}^n$. ▮

PROOF

1. Let A be μ-measurable. Then there exists a Borel set B with $A \subset B$, $\mu(B - A) = 0$. Thus $\nu(B - A) = 0$ and so A is ν-measurable. Hence each μ-measurable set is also ν-measurable.

2. Set

$$Z \equiv \{x \in \mathbb{R}^n \mid D_\mu \nu(x) = 0\},$$
$$I \equiv \{x \in \mathbb{R}^n \mid D_\mu \nu(x) = +\infty\};$$

Z and I are μ-measurable. By Theorem 1, $\mu(I) = 0$ and so $\nu(I) = 0$. Also, Lemma 1 implies $\nu(Z) \leq \alpha\mu(Z)$ for all $\alpha > 0$; thus $\nu(Z) = 0$. Hence

$$\nu(Z) = 0 = \int_Z D_\mu \nu \, d\mu$$

and

$$\nu(I) = 0 = \int_I D_\mu \nu \, d\mu.$$

3. Now let A be μ-measurable and fix $1 < t < \infty$. Define for each integer m

$$A_m \equiv A \cap \{x \in \mathbb{R}^n \mid t^m \le D_\mu \nu(x) < t^{m+1}\}.$$

Then A_m is μ-, and so also ν-, measurable. Moreover,

$$A - \bigcup_{m=-\infty}^{\infty} A_m \subset Z \cup I \cup \{x \mid \overline{D}_\mu \nu(x) \ne \underline{D}_\mu \nu(x)\},$$

and hence

$$\mu\left(A - \bigcup_{m=-\infty}^{\infty} A_m\right) = \nu\left(A - \bigcup_{m=-\infty}^{\infty} A_m\right) = 0.$$

Consequently,

$$\begin{aligned}
\nu(A) &= \sum_{m=-\infty}^{\infty} \nu(A_m) \\
&\le \sum_m t^{m+1} \mu(A_m) \qquad \text{(by Lemma 1)} \\
&= t \sum_m t^m \mu(A_m) \\
&\le t \sum_m \int_{A_m} D_\mu \nu \, d\mu \\
&= t \int_A D_\mu \nu \, d\mu.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\nu(A) &= \sum_m \nu(A_m) \\
&\ge \sum_m t^m \mu(A_m) \qquad \text{(by Lemma 1)} \\
&= \frac{1}{t} \sum_m t^{m+1} \mu(A_m) \\
&\ge \frac{1}{t} \sum_m \int_{A_m} D_\mu \nu \, d\mu \\
&= \frac{1}{t} \int_A D_\mu \nu \, d\mu.
\end{aligned}$$

Thus $1/t \int_A D_\mu\nu \, d\mu \le \nu(A) \le t\int_A D_\mu\nu \, d\mu$ for all $1 < t < \infty$. Send $t \to 1^+$. ∎

THEOREM 3 LEBESGUE DECOMPOSITION THEOREM
Let ν, μ be Radon measures on $\mathbb{R}^n$.

(i) Then $\nu = \nu_{\mathrm{ac}} + \nu_{\mathrm{s}}$, where $\nu_{\mathrm{ac}}, \nu_{\mathrm{s}}$ are Radon measures on $\mathbb{R}^n$ with

$$\nu_{\mathrm{ac}} \ll \mu \qquad \textit{and} \qquad \nu_{\mathrm{s}} \perp \mu.$$

(ii) Furthermore,

$$D_\mu\nu = D_\mu\nu_{\mathrm{ac}} \qquad \textit{and} \qquad D_\mu\nu_{\mathrm{s}} = 0 \qquad \mu \text{ a.e.},$$

and consequently

$$\nu(A) = \int_A D_\mu\nu \, d\mu + \nu_{\mathrm{s}}(A)$$

for each Borel set $A \subset \mathbb{R}^n$.

DEFINITION *We call ν_{ac} the* ***absolutely continuous*** *part, and ν_{s} the* ***singular*** *part, of ν with respect to μ.*

PROOF

1. As before, we may as well assume $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < \infty$.

2. Define

$$\mathcal{E} \equiv \{A \subset \mathbb{R}^n \mid A \text{ Borel}, \mu(\mathbb{R}^n - A) = 0\},$$

and choose $B_k \in \mathcal{E}$ such that, for $k = 1 \ldots,$

$$\nu(B_k) \le \inf_{A\in\mathcal{E}} \nu(A) + \frac{1}{k}.$$

Write $B \equiv \cap_{k=1}^\infty B_k$. Since

$$\mu(\mathbb{R}^n - B) \le \sum_{k=1}^\infty \mu(\mathbb{R}^n - B_k) = 0,$$

we have $B \in \mathcal{E}$, and so

$$\nu(B) = \inf_{A\in\mathcal{E}} \nu(A). \tag{$\star$}$$

Define

$$\nu_{\mathrm{ac}} \equiv \nu \llcorner B,$$
$$\nu_{\mathrm{s}} \equiv \nu \llcorner (\mathbb{R}^n - B);$$

these are Radon measures according to Theorem 3 in Section 1.1.

3. Now suppose $A \subset B$, A is a Borel set, $\mu(A) = 0$, but $\nu(A) > 0$. Then $B - A \in \mathcal{E}$ and $\nu(B - A) < \nu(B)$, a contradiction to $(\star)$. Consequently, $\nu_{\text{ac}} \ll \mu$. On the other hand, $\mu(\mathbb{R}^n - B) = 0$, and thus $\nu_{\text{s}} \perp \mu$.

4. Finally, fix $\alpha > 0$ and set

$$C \equiv \{x \in B \mid D_\mu \nu_{\text{s}}(x) \geq \alpha\}.$$

According to Lemma 1,

$$\alpha \mu(C) \leq \nu_{\text{s}}(C) = 0,$$

and therefore $D_\mu \nu_{\text{s}} = 0$ μ a.e. This implies

$$D_\mu \nu_{\text{ac}} = D_\mu \nu \ \ \mu \text{ a.e.} \quad \blacksquare$$

1.7 Lebesgue points; Approximate continuity

1.7.1 Lebesgue–Besicovitch Differentiation Theorem

NOTATION We denote the *average* of f over the set E with respect to μ by

$$\fint_E f \, d\mu \equiv \frac{1}{\mu(E)} \int_E f \, d\mu,$$

provided $0 < \mu(E) < \infty$ and the integral on the right is defined.

THEOREM 1 LEBESGUE–BESICOVITCH DIFFERENTIATION THEOREM
Let μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mu)$. Then

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

for μ a.e. $x \in \mathbb{R}^n$.

PROOF For Borel $B \subset \mathbb{R}^n$, define $\nu^{\pm}(B) \equiv \int_B f^{\pm} \, d\mu$, and for arbitrary $A \subset \mathbb{R}^n$, $\nu^{\pm}(A) \equiv \inf\{\nu^{\pm}(B) \mid A \subset B,\ B \text{ Borel}\}$. Then ν^+ and ν^- are Radon measures, and so by Theorem 2 in Section 1.6, $\nu^{\pm} \ll \mu$

$$\nu^+(A) = \int_A D_\mu \nu^+ \, d\mu = \int_A f^+ \, d\mu$$

and

$$\nu^-(A) = \int_A D_\mu \nu^- \, d\mu = \int_A f^- \, d\mu$$

for all μ-measurable A. Thus $D_\mu \nu^\pm = f^\pm$ μ a.e. Consequently,

$$\begin{aligned}
\lim_{r\to 0} \fint_{B(x,r)} f\, d\mu &= \lim_{r\to 0} \frac{1}{\mu(B(x,r))}[\nu^+(B(x,r)) - \nu^-(B(x,r))] \\
&= D_\mu \nu^+(x) - D_\mu \nu^-(x) \\
&= f^+(x) - f^-(x) = f(x) \text{ for } \mu \text{ a.e. } x. \quad \blacksquare
\end{aligned}$$

COROLLARY 1
Let μ be a Radon measure on $\mathbb{R}^n$, $1 \le p < \infty$, and $f \in L^p_{\text{loc}}(\mathbb{R}^n, \mu)$. Then

$$\lim_{r\to 0} \fint_{B(x,r)} |f - f(x)|^p\, d\mu = 0 \tag{$\star$}$$

for μ a.e. x.

DEFINITION *A point x for which ($\star$) holds is called a **Lebesgue point** of f with respect to μ.*

PROOF Let $\{r_i\}_{i=1}^\infty$ be a countable dense subset of $\mathbb{R}$. By Theorem 1,

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

for μ a.e. x and $i = 1, 2, \ldots$. Thus there exists a set $A \subset \mathbb{R}^n$ such that $\mu(A) = 0$, and $x \in \mathbb{R}^n - A$ implies

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

for all i. Fix $x \in \mathbb{R}^n - A$ and $\epsilon > 0$. Choose r_i such that $|f(x) - r_i|^p < \epsilon/2^p$. Then

$$\begin{aligned}
&\limsup_{r\to 0} \fint_{B(x,r)} |f - f(x)|^p\, d\mu \\
&\qquad \le 2^{p-1}\left[\limsup_{r\to 0} \fint_{B(x,r)} |f - r_i|^p\, d\mu + \fint_{B(x,r)} |f(x) - r_i|^p\, d\mu\right] \\
&\qquad = 2^{p-1}[|f(x) - r_i|^p + |f(x) - r_i|^p] < \epsilon. \quad \blacksquare
\end{aligned}$$

For the case $\mu = \mathcal{L}^n$, this stronger assertion holds:

COROLLARY 2
If $f \in L^p_{\text{loc}}$ for some $1 \le p < \infty$, then

$$\lim_{B \downarrow \{x\}} \fint_B |f - f(x)|^p\, dy = 0 \textit{ for } \mathcal{L}^n \textit{ a.e. } x.,$$

where the limit is taken over all closed balls B ***containing*** *x as* $\operatorname{diam} B \to 0$.

The point is that the balls need not be centered at x.

PROOF We show that for each sequence of closed balls $\{B_k\}_{k=1}^{\infty}$ with $x \in B_k$ and $d_k \equiv \operatorname{diam} B_k \to 0$,

$$\fint_{B_k} |f - f(x)|^p \, dy \to 0$$

as $k \to \infty$, at each Lebesque point of f. Choose balls $\{B_k\}_{k=1}^{\infty}$ as above. Then $B_k \subset B(x, d_k)$, and consequently,

$$\fint_{B_k} |f - f(x)|^p \, dy \leq 2^n \fint_{B(x,d_k)} |f - f(x)|^p \, dy.$$

The right-hand side goes to zero if x is a Lebesque point. ∎

COROLLARY 3
Let $E \subset \mathbb{R}^n$ be $\mathcal{L}^n$-measurable. Then

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 1 \quad \textit{for } \mathcal{L}^n \textit{ a.e. } x \in E$$

and

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 0 \quad \textit{for } \mathcal{L}^n \textit{ a.e. } x \in \mathbb{R}^n - E.$$

PROOF Set $f = \chi_E$, $\mu = \mathcal{L}^n$ in Theorem 1. ∎

DEFINITION *Let $E \subset \mathbb{R}^n$. A point $x \in \mathbb{R}^n$ is a* ***point of density 1*** *for E if*

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 1$$

and a ***point of density 0*** *for E if*

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 0.$$

REMARK We regard the set of points of density 1 of E as comprising the *measure theoretic interior* of E; according to Corollary 3, $\mathcal{L}^n$ a.e. point in an $\mathcal{L}^n$-measurable set E belongs to its measure theoretic interior. Similarly, the points of density 0 for E make up the *measure theoretic exterior* of E. In Section 5.8 we will define and investigate the *measure theoretic boundary* of certain sets E. See also Section 5.11. ∎

DEFINITION *Assume $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then*

$$f^*(x) \equiv \begin{cases} \lim_{r\to 0} ⨍_{B(x,r)} f\,dy & \textit{if this limit exists} \\ 0 & \textit{otherwise} \end{cases}$$

*is the **precise representative** of f.*

REMARK Note that if $f,\ g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, with $f = g$ $\mathcal{L}^n$ a.e., then $f^* = g^*$ for *all* points $x \in \mathbb{R}^n$. In view of Theorem 1 with $\mu = \mathcal{L}^n$, $\lim_{r\to 0} ⨍_{B(x,r)} f\ dy$ exists $\mathcal{L}^n$ a.e. In Chapters 4 and 5, we will prove that if f is a Sobolev or BV function, then $f^* = f$, except possibly on a "very small" set of appropriate capacity or Hausdorff measure zero.

Observe also that it is possible for the above limit to exist even if x is not a Lebesgue point of f; cf. Theorem 3 and Corollary 1 in Section 5.9. ▮

1.7.2 Approximate limits, approximate continuity

DEFINITION *Let $f : \mathbb{R}^n \to \mathbb{R}^m$. We say $l \in \mathbb{R}^m$ is the **approximate limit** of f as $y \to x$, written*

$$\operatorname{ap}\lim_{y\to x} f(y) = l,$$

if for each $\epsilon > 0$,

$$\lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r) \cap \{|f - l| \geq \epsilon\})}{\mathcal{L}^n(B(x,r))} = 0.$$

So if l is the approximate limit of f at x, for each $\epsilon > 0$ the set $\{|f - l| \geq \epsilon\}$ has density zero at x.

THEOREM 2
An approximate limit is unique.

PROOF Assume for each $\epsilon > 0$ that both

$$\frac{\mathcal{L}^n(B(x,r) \cap \{|f - l| \geq \epsilon\})}{\mathcal{L}^n(B(x,r))} \to 0 \qquad (\star)$$

and

$$\frac{\mathcal{L}^n(B(x,r) \cap \{|f - l'| \geq \epsilon\})}{\mathcal{L}^n(B(x,r))} \to 0 \qquad (\star\star)$$

as $r \to 0$. Then if $l \neq l'$, we set $\epsilon \equiv |l - l'|/3$ and observe for each $y \in B(x,r)$

$$3\epsilon = |l - l'| \leq |f(y) - l| + |f(y) - l'|.$$

Thus

$$B(x,r) \subset \{|f - l| \geq \epsilon\} \cup \{|f - l'| \geq \epsilon\}.$$

Therefore

$$\begin{aligned}\mathcal{L}^n(B(x,r)) \leq \mathcal{L}^n(&B(x,r) \cap \{|f - l| \geq \epsilon\}) \\ &+\mathcal{L}^n(B(x,r) \cap \{|f(y) - l'| \geq \epsilon\}),\end{aligned}$$

a contradiction to $(\star)$, $(\star\star)$. ▮

DEFINITION *Let $f : \mathbb{R}^n \to \mathbb{R}$. We say l is the* ***approximate lim sup*** *of f as $y \to x$, written*

$$\text{ap} \limsup_{y \to x} f(y) = l,$$

if l is the infimum of the real numbers t such that

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap \{f > t\})}{\mathcal{L}^n(B(x,r))} = 0.$$

Similarly, l is the ***approximate lim inf*** *of f as $y \to x$, written*

$$\text{ap} \liminf_{y \to x} f(y) = l,$$

if l is the supremum of the real numbers t such that

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap \{f < t\})}{\mathcal{L}^n(B(x,r))} = 0.$$

DEFINITION *$f : \mathbb{R}^n \to \mathbb{R}^m$ is* ***approximately continuous*** *at $x \in \mathbb{R}^n$ if*

$$\text{ap} \lim_{y \to x} f(y) = f(x).$$

THEOREM 3
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be $\mathcal{L}^n$-measurable. Then f is approximately continuous $\mathcal{L}^n$ a.e.

REMARK Thus a measurable function is "practically continuous at practically every point." The converse is also true; see Federer [F, Section 2.9.13]. ▮

PROOF

1. Claim: There exist disjoint, compact sets $\{K_i\}_{i=1}^{\infty} \subset \mathbb{R}^n$ such that

$$\mathcal{L}^n\left(\mathbb{R}^n - \left(\bigcup_{i=1}^{\infty} K_i\right)\right) = 0$$

and for each $i = 1, 2, \ldots$

$$f|_{K_i} \text{ is continuous.}$$

Proof of Claim: For each positive integer m, set $B_m \equiv B(0, m)$. By Lusin's Theorem, there exists a compact set $K_1 \subset B_1$ such that $\mathcal{L}^n(B_1 - K_1) \leq 1$ and $f|_{K_1}$ is continuous. Assuming now $K_1, \ldots, K_m$ have been constructed, there exists a compact set

$$K_{m+1} \subset B_{m+1} - \bigcup_{i=1}^{m} K_i$$

such that

$$\mathcal{L}^n\left(B_{m+1} - \bigcup_{i=1}^{m+1} K_i\right) \leq \frac{1}{m+1}$$

and $f|_{K_{m+1}}$ is continuous.

2. For $\mathcal{L}^n$ a.e. $x \in K_i$,

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) - K_i)}{\mathcal{L}^n(B(x,r))} = 0. \tag{$\star$}$$

Define $A \equiv \{x \mid \text{for some } i,\ x \in K_i \text{ and } (\star) \text{ holds}\}$; then $\mathcal{L}^n(\mathbb{R}^n - A) = 0$. Let $x \in A$, so that $x \in K_i$ and $(\star)$ holds for some fixed i. Fix $\epsilon > 0$. There exists $s > 0$ such that $y \in K_i$ and $|x - y| < s$ imply $|f(x) - f(y)| < \epsilon$. Then if $0 < r < s$, $B(x,r) \cap \{y \mid |f(y) - f(x)| \geq \epsilon\} \subset B(x,r) - K_i$. In view of $(\star)$, we see

$$\operatorname{ap}\lim_{y \to x} f(y) = f(x). \quad \blacksquare$$

REMARK If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, the proof is much easier. Indeed, for each $\epsilon > 0$

$$\frac{\mathcal{L}^n(B(x,r) \cap \{|f - f(x)| > \epsilon\})}{\mathcal{L}^n(B(x,r))} \leq \frac{1}{\epsilon} \fint_{B(x,r)} |f - f(x)|\, dy,$$

and the right-hand side goes to zero for $\mathcal{L}^n$ a.e. x. In particular *a Lebesgue point is a point of approximate continuity.* $\blacksquare$

REMARK In Section 6.1.3 we will define and discuss the related notion of *approximate differentiability.* $\blacksquare$

1.8 Riesz Representation Theorem

In these notes there will be two primary sources of measures to which we will apply the foregoing abstract theory: these are (a) Hausdorff measures, constructed in Chapter 2, and (b) Radon measures characterizing certain linear functionals, generated as follows.

THEOREM 1 RIESZ REPRESENTATION THEOREM
Let $L : C_c(\mathbb{R}^n; \mathbb{R}^m) \to \mathbb{R}$ be a linear functional satisfying

$$\sup\{L(f) \mid f \in C_c(\mathbb{R}^n; \mathbb{R}^m), |f| \leq 1, \ spt(f) \subset K\} < \infty \qquad (\star)$$

for each compact set $K \subset \mathbb{R}^n$. Then there exists a Radon measure μ on $\mathbb{R}^n$ and a μ-measurable function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ such that

(i) $|\sigma(x)| = 1$ *for μ-a.e. x, and*
(ii) $L(f) = \int_{\mathbb{R}^n} f \cdot \sigma \, d\mu$

for all $f \in C_c(\mathbb{R}^n; \mathbb{R}^m)$.

DEFINITION *We call μ the* **variation measure**, *defined for each open set $V \subset \mathbb{R}^n$ by*

$$\mu(V) \equiv \sup\{L(f) \mid f \in C_c(\mathbb{R}^n; \mathbb{R}^m), |f| \leq 1, \ spt(f) \subset V\}.$$

PROOF

1. Define μ on open sets V as above and then set

$$\mu(A) \equiv \inf\{\mu(V) \mid A \subset V \text{ open}\}$$

for arbitrary $A \subset \mathbb{R}^n$.

2. *Claim #1*: μ is a measure.

Proof of Claim #1: Let $V, \{V_i\}_{i=1}^{\infty}$ be open subsets of $\mathbb{R}^n$, with $V \subset \cup_{i=1}^{\infty} V_i$. Choose $g \in C_c(\mathbb{R}^n; \mathbb{R}^m)$ such that $|g| \leq 1$ and $\text{spt}(g) \subset V$. Since $\text{spt}(g)$ is compact, there exists an index k such that $\text{spt}(g) \subset \cup_{j=1}^{k} V_j$. Let $\{\zeta_j\}_{j=1}^{k}$ be a finite sequence of smooth functions such that $\text{spt}(\zeta_j) \subset V_j$ for $1 \leq j \leq k$ and $\sum_{j=1}^{k} \zeta_j = 1$ on $\text{spt}(g)$. Then $g = \sum_{j=1}^{k} g\zeta_j$, and so

$$|L(g)| = \left|\sum_{j=1}^{k} L(g\zeta_j)\right| \leq \sum_{j=1}^{k} |L(g\zeta_j)| \leq \sum_{j=1}^{\infty} \mu(V_j).$$

Then, taking the supremum over g, we find $\mu(V) \leq \sum_{j=1}^{\infty} \mu(V_j)$. Now let $\{A_j\}_{j=1}^{\infty}$ be arbitrary sets with $A \subset \cup_{j=1}^{\infty} A_j$. Fix $\epsilon > 0$. Choose open sets V_j

such that $A_j \subset V_j$ and $\mu(A_j) + \epsilon/2^j \geq \mu(V_j)$. Then

$$\mu(A) \leq \mu\left(\bigcup_{j=1}^{\infty} V_j\right) \leq \sum_{j=1}^{\infty} \mu(V_j)$$
$$\leq \sum_{j=1}^{\infty} \mu(A_j) + \epsilon.$$

3. Claim #2: μ is a Radon measure.

Proof of Claim #2: Let U_1 and U_2 be open sets with $\operatorname{dist}(U_1, U_2) > 0$. Then $\mu(U_1 \cup U_2) = \mu(U_1) + \mu(U_2)$ by definition of μ. Hence if $A_1, A_2 \subset \mathbb{R}^n$ and $\operatorname{dist}(A_1, A_2) > 0$, then $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$. According to Caratheordory's Criterion (Section 1.1.1), μ is a Borel measure. Furthermore, by its definition, μ is Borel regular; indeed, given $A \subset \mathbb{R}^n$, there exist open sets V_k such that $A \subset V_k$ and $\mu(V_k) \leq \mu(A) + 1/k$ for all k. Thus $\mu(A) = \mu\left(\cap_{k=1}^{\infty} V_k\right)$. Finally, condition $(\star)$ implies $\mu(K) < \infty$ for all compact K.

4. Now, let $C_c^+(\mathbb{R}^n) \equiv \{f \in C_c(\mathbb{R}^n) \mid f \geq 0\}$, and for $f \in C_c^+(\mathbb{R}^n)$, set

$$\lambda(f) \equiv \sup\{|L(g)| \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \leq f\}.$$

Observe that for all $f_1, f_2 \in C_c^+(\mathbb{R}^n)$, $f_1 \leq f_2$ implies $\lambda(f_1) \leq \lambda(f_2)$. Also $\lambda(cf) = c\lambda(f)$ for all $c \geq 0$, $f \in C_c^+(\mathbb{R}^n)$.

5. Claim #3: For all $f_1, f_2 \in C_c^+(\mathbb{R}^n)$, $\lambda(f_1 + f_2) = \lambda(f_1) + \lambda(f_2)$.

Proof of Claim #3: If $g_1, g_2 \in C_c(\mathbb{R}^n; \mathbb{R}^m)$ with $|g_1| \leq f_1$ and $|g_2| \leq f_2$, then $|g_1 + g_2| \leq f_1 + f_2$. We can furthermore assume $L(g_1), L(g_2) \geq 0$. Therefore,

$$|L(g_1)| + |L(g_2)| = L(g_1 + g_2) = |L(g_1 + g_2)| \leq \lambda(f_1 + f_2).$$

Taking suprema over g_1 and g_2 with $g_1, g_2 \in C_c(\mathbb{R}^n; \mathbb{R}^m)$ gives

$$\lambda(f_1) + \lambda(f_2) \leq \lambda(f_1 + f_2).$$

Now fix $g \in C_c(\mathbb{R}^n; \mathbb{R}^m)$, with $|g| \leq f_1 + f_2$. Set

$$g_i \equiv \begin{cases} \frac{f_i g}{f_1 + f_2} & \text{if } f_1 + f_2 > 0 \\ 0 & \text{if } f_1 + f_2 = 0 \end{cases}$$

for $i = 1, 2$. Then $g_1, g_2 \in C_c(\mathbb{R}^n; \mathbb{R}^m)$ and $g = g_1 + g_2$. Moreover, $|g_i| \leq f_i$, $(i = 1, 2)$, so that

$$|L(g)| \leq |L(g_1)| + |L(g_2)| \leq \lambda(f_1) + \lambda(f_2).$$

Consequently,

$$\lambda(f_1 + f_2) \leq \lambda(f_1) + \lambda(f_2).$$

6. *Claim #4:* $\lambda(f) = \int_{\mathbb{R}^n} f\, d\mu$ for all $f \in C_c^+(\mathbb{R}^n)$.

Proof of Claim #4: Let $\epsilon > 0$. Choose $0 = t_0 < t_1 < \cdots < t_N$ such that $t_N \equiv 2\|f\|_{L^\infty}$, $0 < t_i - t_{i-1} < \epsilon$, and $\mu(f^{-1}\{t_i\}) = 0$ for $i = 1, \ldots, N$. Set $U_j = f^{-1}((t_{j-1}, t_j))$; U_j is open and $\mu(U_j) < \infty$.

By Theorem 4 in Section 1.1, there exist compact sets K_j such that $K_j \subset U_j$ and $\mu(U_j - K_j) < \epsilon/N$, $j = 1, 2, \ldots, N$. Furthermore there exist functions $g_j \in C_c(\mathbb{R}^n; \mathbb{R}^m)$ with $|g_j| \le 1$, spt $(g_j) \subset U_j$, and $|L(g_j)| \ge \mu(U_j) - \epsilon/N$. Note also that there exist functions $h_j \in C_c^+(\mathbb{R}^n)$ such that spt $(h_j) \subset U_j$, $0 \le h_j \le 1$, and $h_j = 1$ on the compact set $K_j \cup \text{spt}\,(g_j)$. Then

$$\lambda(h_j) \ge |L(g_j)| \ge \mu(U_j) - \epsilon/N$$

and

$$\begin{aligned}
\lambda(h_j) &= \sup\{|L(g)| \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \le h_j\} \\
&\le \sup\{|L(g)| \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \le 1,\ \text{spt}\,(g) \subset U_j\} \\
&= \mu(U_j),
\end{aligned}$$

whence $\mu(U_j) - \epsilon/N \le \lambda(h_j) \le \mu(U_j)$.

Define

$$A \equiv \left\{ x \mid f(x) \left(1 - \sum_{j=1}^N h_j(x) \right) > 0 \right\};$$

A is open. Next, compute

$$\begin{aligned}
\lambda\left(f - f \sum_{j=1}^N h_j \right) &= \sup \left\{ |L(g)| \ \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \le f - f \sum_{j=1}^N h_j \right\} \\
&\le \sup\{|L(g)| \ \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \le \|f\|_{L^\infty} \chi_A\} \\
&= \|f\|_{L^\infty} \sup\{L(g) \mid g \in C_c(\mathbb{R}^n; \mathbb{R}^m), |g| \le \chi_A\} \\
&= \|f\|_{L^\infty} \mu(A) \\
&= \|f\|_{L^\infty} \mu\left(\bigcup_{j=1}^N (U_j - \{h_j = 1\}) \right) \\
&\le \|f\|_{L^\infty} \sum_{j=1}^N \mu(U_j - K_j) \le \epsilon \|f\|_{L^\infty}.
\end{aligned}$$

Hence

$$\begin{aligned}\lambda(f) &= \lambda\left(f - f\sum_{j=1}^{N} h_j\right) + \lambda\left(f\sum_{j=1}^{N} h_j\right)\\ &\leq \epsilon\|f\|_{L^\infty} + \sum_{j=1}^{N}\lambda(fh_j)\\ &\leq \epsilon\|f\|_{L^\infty} + \sum_{j=1}^{N} t_j\mu(U_j)\end{aligned}$$

and

$$\begin{aligned}\lambda(f) &\geq \sum_{j=1}^{N}\lambda(fh_j)\\ &\geq \sum_{j=1}^{N} t_{j-1}(\mu(U_j) - \epsilon/N)\\ &\geq \sum_{j=1}^{N} t_{j-1}\mu(U_j) - t_N\epsilon.\end{aligned}$$

Finally, since

$$\sum_{j=1}^{N} t_{j-1}\mu(U_j) \leq \int_{\mathbb{R}^n} f\,d\mu \leq \sum_{j=1}^{N} t_j\mu(U_j),$$

we have

$$\begin{aligned}\left|\,\lambda(f) - \int f\,d\mu\,\right| &\leq \sum_{j=1}^{N}(t_j - t_{j-1})\mu(U_j) + \epsilon\|f\|_{L^\infty} + \epsilon t_N\\ &\leq \epsilon\mu(\text{ spt }(f)) + 3\epsilon\|f\|_{L^\infty}.\end{aligned}$$

7. *Claim #5*: There exists a μ-measurable function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ satisfying (ii).

Proof of Claim #5: Fix $e \in \mathbb{R}^m$, $|e| = 1$. Define $\lambda_e(f) \equiv L(fe)$ for $f \in C_c(\mathbb{R}^n)$. Then λ_e is linear and

$$\begin{aligned}|\lambda_e(f)| &= |L(fe)|\\ &\leq \sup\{|L(g)| \mid g \in C_c(\mathbb{R}^n;\mathbb{R}^m), |g| \leq |f|\}\\ &= \lambda(|f|) = \int_{\mathbb{R}^n} |f|\,d\mu;\end{aligned}$$

thus we can extend λ_e to a bounded linear functional on $L^1(\mathbb{R}^n;\mu)$. Hence there exists $\sigma_e \in L^\infty(\mu)$ such that

$$\lambda_e(f) = \int_{\mathbb{R}^n} f\sigma_e \, d\mu \qquad (f \in C_c(\mathbb{R}^n)).$$

Let $e_1, \ldots, e_m$ be the standard basis for $\mathbb{R}^m$ and define $\sigma \equiv \sum_{j=1}^m \sigma_{e_j} e_j$. Then if $f \in C_c(\mathbb{R}^n;\mathbb{R}^m)$, we have

$$\begin{aligned} L(f) &= \sum_{j=1}^m L((f\cdot e_j)e_j) \\ &= \sum_{j=1}^m \int (f \cdot e_j)\sigma_{e_j} \, d\mu \\ &= \int f \cdot \sigma \, d\mu. \end{aligned}$$

8. *Claim #6*: $|\sigma| = 1$ μ a.e.

Proof of Claim #6: Let $U \subset \mathbb{R}^n$ be open, $\mu(U) < \infty$. By definition,

$$\mu(U) = \sup\left\{ \int f \cdot \sigma \, d\mu \,\middle|\, f \in C_c(\mathbb{R}^n;\mathbb{R}^m)\,,\ |f| \le 1,\ \operatorname{spt}(f) \subset U \right\}. \qquad (\star\star)$$

Now take $f_k \in C_c(\mathbb{R}^n;\mathbb{R}^m)$ such that $|f_k| \le 1$, spt $(f_k) \subset U$, and $f_k \cdot \sigma \to |\sigma|$ μ a.e.; such functions exist by Corollary 1 in Section 1.2. Then

$$\int_U |\sigma| \, d\mu = \lim_{k\to\infty} \int f_k \cdot \sigma \, d\mu \le \mu(U)$$

by $(\star\star)$.

On the other hand, if $f \in C_c(\mathbb{R}^n;\mathbb{R}^m)$ with $|f| \le 1$ and spt $(f) \subset U$, then

$$\int f \cdot \sigma \, d\mu \le \int_U |\sigma| \, d\mu.$$

Consequently $(\star\star)$ implies

$$\mu(U) \le \int_U |\sigma| \, d\mu.$$

Thus $\mu(U) = \int_U |\sigma| \, d\mu$ for all open $U \subset \mathbb{R}^n$; hence $|\sigma| = 1$ μ a.e. ∎

An immediate and very useful application is the following characterization of nonnegative linear functionals.

COROLLARY 1

Assume $L : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ is linear and nonnegative, so that

$$L(f) \ge 0 \text{ for all } f \in C_c^\infty(\mathbb{R}^n), f \ge 0. \qquad (\star)$$

Then there exists a Radon measure μ on $\mathbb{R}^n$ such that

$$L(f) = \int_{\mathbb{R}^n} f \, d\mu \qquad \textit{for all } f \in C_c^\infty(\mathbb{R}^n).$$

PROOF Choose any compact set $K \subset \mathbb{R}^n$, and select a smooth function ζ such that ζ has compact support, $\zeta \equiv 1$ on K, $0 \leq \zeta \leq 1$. Then for any $f \in C_c^\infty(\mathbb{R}^n)$ with spt $(f) \subset K$, set $g \equiv ||f||_{L^\infty}\zeta - f \geq 0$. Therefore $(\star)$ implies

$$0 \leq L(g) = ||f||_{L^\infty} L(\zeta) - L(f),$$

and so

$$L(f) \leq C||f||_{L^\infty}$$

for $C \equiv L(\zeta)$. L thus extends to a linear mapping from $C_c(\mathbb{R}^n)$ into $\mathbb{R}$, satisfying the hypothesis of the Riesz Representation Theorem. Hence there exist μ, σ as above so that

$$L(f) = \int_{\mathbb{R}^n} f\sigma \, d\mu \qquad (f \in C_c^\infty(\mathbb{R}^n))$$

with $\sigma = \pm 1$ μ a.e. But then $(\star)$ implies $\sigma = 1$ μ a.e. ∎

1.9 Weak convergence and compactness for Radon measures

We introduce next a notion of weak convergence for measures.

THEOREM 1
Let μ, μ_k $(k = 1, 2, \ldots)$ be Radon measures on $\mathbb{R}^n$. The following three statements are equivalent:

- *(i)* $\lim_{k\to\infty} \int_{\mathbb{R}^n} f \, d\mu_k = \int_{\mathbb{R}^n} f \, d\mu$ *for all* $f \in C_c(\mathbb{R}^n)$.
- *(ii)* $\limsup_{k\to\infty} \mu_k(K) \leq \mu(K)$ *for each compact set* $K \subset \mathbb{R}^n$ *and* $\mu(U) \leq \liminf_{k\to\infty} \mu_k(U)$ *for each open set* $U \subset \mathbb{R}^n$.
- *(iii)* $\lim_{k\to\infty} \mu_k(B) = \mu(B)$ *for each bounded Borel set* $B \subset \mathbb{R}^n$ *with* $\mu(\partial B) = 0$.

DEFINITION *If (i) through (iii) hold, we say the measures μ_k **converge weakly** to the measure μ, written*

$$\mu_k \rightharpoonup \mu.$$

PROOF

1. Assume (i) holds and fix $\epsilon > 0$. Let $U \subset \mathbb{R}^n$ be open and choose a compact set $K \subset U$. Next, choose $f \in C_c(\mathbb{R}^n)$ such that $0 \leq f \leq 1$, spt $(f) \subset U$, $f \equiv 1$ on K. Then

$$\mu(K) \leq \int_{\mathbb{R}^n} f\, d\mu = \lim_{k\to\infty} \int_{\mathbb{R}^n} f\, d\mu_k \leq \liminf_{k\to\infty} \mu_k(U).$$

Thus

$$\mu(U) = \sup\{\mu(K) \mid K \text{ compact }, K \subset U\} \leq \liminf_{k\to\infty} \mu_k(U).$$

This proves the second part of (ii); the proof of the other part is similar.

2. Suppose now (ii) holds, $B \subset \mathbb{R}^n$ is a bounded Borel set, $\mu(\partial B) = 0$. Then

$$\begin{aligned}\mu(B) = \mu(B^o) &\leq \liminf_{k\to\infty} \mu_k(B^o)\\ &\leq \limsup_{k\to\infty} \mu_k(\overline{B})\\ &\leq \mu(\overline{B}) = \mu(B).\end{aligned}$$

3. Finally, assume (iii) holds. Fix $\epsilon > 0$, $f \in C_c^+(\mathbb{R}^n)$. Let $R > 0$ be such that $\operatorname{spt}(f) \subset B(0,R)$ and $\mu(\partial B(0,R)) = 0$. Choose $0 = t_0 < t_1 < \cdots < t_N$ such that $t_N \equiv 2\|f\|_{L^\infty}$, $0 < t_i - t_{i-1} < \epsilon$, and $\mu(f^{-1}\{t_i\}) = 0$ for $i = 1, \ldots, N$. Set $B_i = f^{-1}(t_{i-1}, t_i]$; then $\mu(\partial B_i) = 0$ for $i \geq 2$. Now

$$\sum_{i=2}^N t_{i-1}\mu_k(B_i) \leq \int_{\mathbb{R}^n} f\, d\mu_k \leq \sum_{i=2}^N t_i\mu_k(B_i) + t_1\mu_k(B(0,R))$$

and

$$\sum_{i=2}^N t_{i-1}\mu(B_i) \leq \int_{\mathbb{R}^n} f\, d\mu \leq \sum_{i=2}^N t_i\mu(B_i) + t_1\mu(B(0,R));$$

so (iii) implies

$$\limsup_{k\to\infty} \left| \int_{\mathbb{R}^n} f d\mu_k - \int_{\mathbb{R}^n} f\, d\mu \right| \leq 2\epsilon\mu(B(0,R)). \quad ∎$$

The great advantage in studying the weak convergence of measures is that compactness is had relatively easily.

THEOREM 2 WEAK COMPACTNESS FOR MEASURES

Let $\{\mu_k\}_{k=1}^\infty$ be a sequence of Radon measures on $\mathbb{R}^n$ satisfying

$$\sup_k \mu_k(K) < \infty \textit{ for each compact set } K \subset \mathbb{R}^n.$$

Then there exists a subsequence $\{\mu_{k_j}\}_{j=1}^\infty$ and a Radon measure μ such that

$$\mu_{k_j} \rightharpoonup \mu.$$

PROOF

1. Assume first

$$\sup_k \mu_k(\mathbb{R}^n) < \infty. \tag{$\star$}$$

2. Let $\{f_k\}_{k=1}^\infty$ be a countable dense subset of $C_c(\mathbb{R}^n)$. As $(\star)$ implies $\int f_1 \, d\mu_j$ is bounded, we can find a subsequence $\{\mu_j^1\}_{j=1}^\infty$ and $a_1 \in \mathbb{R}$ such that

$$\int f_1 \, d\mu_j^1 \to a_1.$$

Continuing, we choose a subsequence $\{\mu_j^k\}_{j=1}^\infty$ of $\{\mu_j^{k-1}\}_{j=1}^\infty$ and $a_k \in \mathbb{R}$ such that

$$\int f_k \, d\mu_j^k \to a_k.$$

Set $\nu_j \equiv \mu_j^j$; then

$$\int f_k \, d\nu_j \to a_k$$

for all $k \geq 1$. Define $L(f_k) \equiv a_k$, and note that L is linear and $|L(f_k)| \leq ||f_k||_{L^\infty} M$ by $(\star)$, for $M \equiv \sup_k \mu_k(\mathbb{R}^n)$. Thus L can be uniquely extended to a bounded linear functional $\overline{L}$ on $C_c(\mathbb{R}^n)$. Then according to the Riesz Representation Theorem (Section 1.8) there exists a Radon measure μ on $\mathbb{R}^n$ such that

$$\overline{L}(f) = \int f \, d\mu$$

for all $f \in C_c(\mathbb{R}^n)$.

3. Choose any $f \in C_c(\mathbb{R}^n)$. The denseness of $\{f_k\}_{k=1}^\infty$ implies the existence of a subsequence $\{f_i\}_{i=1}^\infty$ such that $f_i \to f$ uniformly. Fix $\epsilon > 0$ and then choose i so large that

$$||f - f_i||_{L^\infty} < \frac{\epsilon}{4M}.$$

Next choose J so that for all $j > J$

$$\left| \int f_i \, d\nu_j - \int f_i \, d\mu \right| < \epsilon/2.$$

Then for $j > J$

$$\begin{aligned} \left| \int f \, d\nu_j - \int f \, d\mu \right| &\leq \left| \int (f - f_i) \, d\nu_j \right| + \left| \int (f - f_i) \, d\mu \right| \\ &\quad + \left| \int f_i \, d\nu_j - \int f_i \, d\mu \right| \\ &\leq 2M ||f - f_i||_{L^\infty} + \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

4. In the general case that $(\star)$ fails to hold, but

$$\sup_k \mu_k(K) < \infty$$

for each compact $K \subset \mathbb{R}^n$, we apply the reasoning above to the measures

$$\mu_k^l \equiv \mu_k \llcorner B(0,l) \qquad (k,l = 1,2,\ldots)$$

and use a diagonal argument. ∎

Assume now that $U \subset \mathbb{R}^n$ is open, $1 \leq p < \infty$.

DEFINITION *A sequence $\{f_k\}_{k=1}^\infty$ in $L^p(U)$ converges **weakly** to $f \in L^p(U)$, written*

$$f_k \rightharpoonup f \qquad \text{in } L^p(U),$$

provided

$$\lim_{k\to\infty} \int_U f_k g \, dx = \int_U f g \, dx$$

for each $g \in L^q(U)$, where $1/p + 1/q = 1$, $1 < q \leq \infty$.

THEOREM 3 WEAK COMPACTNESS IN L^p
Suppose $1 < p < \infty$. Let $\{f_k\}_{k=1}^\infty$ be a sequence of functions in $L^p(U)$ satisfying

$$\sup_k ||f_k||_{L^p(U)} < \infty. \tag{$\star$}$$

Then there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ and a function $f \in L^p(U)$ such that

$$f_{k_j} \rightharpoonup f \qquad \text{in } L^p(U).$$

REMARK This assertion is in general false for $p = 1$. ∎

PROOF
1. If $U \neq \mathbb{R}^n$ we extend each function f_k to all of $\mathbb{R}^n$ by setting it equal to zero on $\mathbb{R}^n - U$. This done, we may with no loss of generality assume $U = \mathbb{R}^n$. Furthermore, we may as well suppose

$$f_k \geq 0 \qquad \mathcal{L}^n \text{ a.e.};$$

for we could otherwise apply the following analysis to f_k^+ and f_k^-.
2. Define the Radon measures

$$\mu_k \equiv \mathcal{L}^n \llcorner f_k \qquad (k = 1,2,\ldots).$$

Then for each compact set $K \subset \mathbb{R}^n$

$$\mu_k(K) = \int_K f_k \, dx \leq \left(\int_K f_k^p \, dx \right)^{\frac{1}{p}} \mathcal{L}^n(K)^{1-\frac{1}{p}},$$

and so

$$\sup_k \mu_k(K) < \infty.$$

Accordingly, we may apply Theorem 2 to find a Radon measure μ on $\mathbb{R}^n$ and a subsequence $\mu_{k_j} \rightharpoonup \mu$.

3. *Claim #1*: $\mu \ll \mathcal{L}^n$.

Proof of Claim #1: Let $A \subset \mathbb{R}^n$ be bounded, $\mathcal{L}^n(A) = 0$. Fix $\epsilon > 0$ and choose an open, bounded set $V \supset A$ such that $\mathcal{L}^n(V) < \epsilon$. Then

$$\begin{aligned}
\mu(V) &\leq \liminf_{j\to\infty} \mu_{k_j}(V) \\
&= \liminf_{j\to\infty} \int_V f_{k_j} \, dx \\
&\leq \liminf_{j\to\infty} \left(\int_V f_{k_j}^p \, dx \right)^{\frac{1}{p}} \mathcal{L}^n(V)^{1-\frac{1}{p}} \\
&\leq C\epsilon^{1-\frac{1}{p}}.
\end{aligned}$$

Thus $\mu(A) = 0$.

4. In view of Theorem 2 in Section 1.6.2, there exists an L^1_{loc} function f satisfying

$$\mu(A) = \int_A f \, dx$$

for all Borel sets $A \subset \mathbb{R}^n$.

5. *Claim #2*: $f \in L^p(\mathbb{R}^n)$.

Proof of Claim #2: Let $\varphi \in C_c(\mathbb{R}^n)$. Then

$$\begin{aligned}
\int_{\mathbb{R}^n} \varphi f \, dx &= \int_{\mathbb{R}^n} \varphi \, d\mu \\
&= \lim_{j\to\infty} \int_{\mathbb{R}^n} \varphi \, d\mu_{k_j} \\
&= \lim_{j\to\infty} \int_{\mathbb{R}^n} \varphi f_{k_j} \, dx \\
&\leq \sup_k ||f_k||_{L^p} ||\varphi||_{L^q} \\
&\leq C||\varphi||_{L^q}
\end{aligned}$$

Thus

$$||f||_{L^p} = \sup_{\substack{\varphi \in C_c(\mathbb{R}^n) \\ ||\varphi||_{L^q} \leq 1}} \int_{\mathbb{R}^n} \varphi f \, dx < \infty.$$

6. *Claim #3*: $f_{k_j} \rightharpoonup f$ in $L^p(\mathbb{R}^n)$.

Proof of Claim #3: As noted above,

$$\int_{\mathbb{R}^n} f_{k_j} \varphi \, dx \to \int_{\mathbb{R}^n} f \varphi \, dx$$

for all $\varphi \in C_c(\mathbb{R}^n)$. Given $g \in L^q(\mathbb{R}^n)$, we fix $\epsilon > 0$ and then choose $\varphi \in C_c(\mathbb{R}^n)$ with

$$||g - \varphi||_{L^q(\mathbb{R}^n)} < \epsilon.$$

Then

$$\int_{\mathbb{R}^n} f_{k_j} g \, dx = \int_{\mathbb{R}^n} f_{k_j} \varphi \, dx + \int_{\mathbb{R}^n} f_{k_j} (g - \varphi) \, dx,$$

and the last term is estimated by

$$||f_{k_j}||_{L^p} ||g - \varphi||_{L^q} \leq C\epsilon. \quad \blacksquare$$

2

Hausdorff Measure

We introduce next certain "lower dimensional" measures on $\mathbb{R}^n$, which allow us to measure certain "very small" subsets of $\mathbb{R}^n$. These are the Hausdorff measures $\mathcal{H}^s$, defined in terms of the diameters of various efficient coverings. The idea is that A is an "s-dimensional subset" of $\mathbb{R}^n$ if $0 < \mathcal{H}^s(A) < \infty$, even if A is very complicated geometrically.

Section 2.1 provides the definitions and basic properties of Hausdorff measures. In Section 2.2 we prove n-dimensional Lebesgue and n-dimensional Hausdorff measure agree on $\mathbb{R}^n$. Density theorems for lower dimensional Hausdorff measures are established in Section 2.3. Section 2.4 records for later use some easy facts concerning the Hausdorff dimension of graphs and the sets where a summable function is large.

2.1 Definitions and elementary properties; Hausdorff dimension

DEFINITIONS

(i) Let $A \subset \mathbb{R}^n$, $0 \leq s < \infty$, $0 < \delta \leq \infty$. *Define*

$$\mathcal{H}^s_\delta(A) \equiv \inf\left\{\sum_{j=1}^{\infty} \alpha(s)\left(\frac{\operatorname{diam} C_j}{2}\right)^s \mid A \subset \bigcup_{j=1}^{\infty} C_j, \operatorname{diam} C_j \leq \delta\right\},$$

where

$$\alpha(s) \equiv \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)}.$$

Here $\Gamma(s) \equiv \int_0^\infty e^{-x} x^{s-1}\, dx$, $(0 < s < \infty)$, *is the usual gamma function.*

(ii) *For A and s as above, define*

$$\mathcal{H}^s(A) \equiv \lim_{\delta \to 0} \mathcal{H}^s_\delta(A) = \sup_{\delta > 0} \mathcal{H}^s_\delta(A).$$

We call $\mathcal{H}^s$ ***s-dimensional Hausdorff measure*** *on $\mathbb{R}^n$.*

REMARKS

(i) Our requiring $\delta \to 0$ forces the coverings to "follow the local geometry" of the set A.

(ii) Observe

$$\mathcal{L}^n(B(x,r)) = \alpha(n)r^n$$

for all balls $B(x,r) \subset \mathbb{R}^n$. We will see later in Chapter 3 that if $s = k$ is an integer, $\mathcal{H}^k$ agrees with ordinary "k-dimensional surface area" on nice sets; this is the reason we include the normalizing constant $\alpha(s)$ in the definition. ∎

THEOREM 1
$\mathcal{H}^s$ is a Borel regular measure $(0 \leq s < \infty)$.

Warning: $\mathcal{H}^s$ is *not* a Radon measure if $0 \leq s < n$, since $\mathbb{R}^n$ is not σ-finite with respect to $\mathcal{H}^s$.

PROOF
1. Claim #1: $\mathcal{H}^s_\delta$ is a measure.

Proof of Claim #1: Choose $\{A_k\}_{k=1}^\infty \subset \mathbb{R}^n$ and suppose $A_k \subset \cup_{j=1}^\infty C_j^k$, diam $C_j^k \leq \delta$; then $\{C_j^k\}_{j,k=1}^\infty$ covers $\cup_{k=1}^\infty A_k$. Thus

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^\infty A_k\right) \leq \sum_{k=1}^\infty \sum_{j=1}^\infty \alpha(s)\left(\frac{\text{diam } C_j^k}{2}\right)^s.$$

Taking infima we find

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^\infty A_k\right) \leq \sum_{k=1}^\infty \mathcal{H}^s_\delta(A_k).$$

2. Claim #2: $\mathcal{H}^s$ is a measure.

Proof of Claim #2: Select $\{A_k\}_{k=1}^{\infty} \subset \mathbb{R}^n$. Then

$$\mathcal{H}_\delta^s\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mathcal{H}_\delta^s(A_k) \leq \sum_{k=1}^{\infty} \mathcal{H}^s(A_k).$$

Let $\delta \to 0$.

3. *Claim #3*: $\mathcal{H}^s$ is a Borel measure.

Proof of Claim #3: Choose $A, B \subset \mathbb{R}^n$ with $\operatorname{dist}(A, B) > 0$. Select $0 < \delta < 1/4 \operatorname{dist}(A, B)$. Suppose $A \cup B \subset \cup_{k=1}^{\infty} C_k$ and diam $C_k \leq \delta$.

Write $\mathcal{A} \equiv \{C_j \mid C_j \cap A \neq \emptyset\}$, and let $\mathcal{B} \equiv \{C_j \mid C_j \cap B \neq \emptyset\}$. Then $A \subset \cup_{C_j \in \mathcal{A}} C_j$ and $B \subset \cup_{C_j \in \mathcal{B}} C_j$, $C_i \cap C_j = \emptyset$ if $C_i \in \mathcal{A}$, $C_j \in \mathcal{B}$: Hence

$$\begin{aligned}\sum_{j=1}^{\infty} \alpha(s)\left(\frac{\operatorname{diam} C_j}{2}\right)^s &\geq \sum_{C_j \in \mathcal{A}} \alpha(s)\left(\frac{\operatorname{diam} C_j}{2}\right)^s + \sum_{C_j \in \mathcal{B}} \alpha(s)\left(\frac{\operatorname{diam} C_j}{2}\right)^s \\ &\geq \mathcal{H}_\delta^s(A) + \mathcal{H}_\delta^s(B).\end{aligned}$$

Taking the infimum over all such sets $\{C_j\}_{j=1}^{\infty}$, we find $\mathcal{H}_\delta^s(A \cup B) \geq \mathcal{H}_\delta^s(A) + \mathcal{H}_\delta^s(B)$, provided $0 < 4\delta < \operatorname{dist}(A, B)$. Letting $\delta \to 0$, we obtain $\mathcal{H}^s(A \cup B) \geq \mathcal{H}^s(A) + \mathcal{H}^s(B)$. Consequently,

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B)$$

for all $A, B \subset \mathbb{R}^n$ with $\operatorname{dist}(A, B) > 0$. Hence Caratheodory's Criterion, Section 1.1.1., implies $\mathcal{H}^s$ is a Borel measure.

4. *Claim #4*: $\mathcal{H}^s$ is a Borel regular measure.

Proof of Claim#4: Note that diam $\overline{C} =$ diam C for all C; hence

$$\mathcal{H}_\delta^s(A) = \inf\left\{\sum_{j=1}^{\infty} \alpha(s)\left(\frac{\operatorname{diam} C_j}{2}\right)^s \mid A \subset \bigcup_{j=1}^{\infty} C_j, \operatorname{diam} C_j \leq \delta, C_j \text{ closed}\right\}$$

Choose $A \subset \mathbb{R}^n$ such that $\mathcal{H}^s(A) < \infty$; then $\mathcal{H}_\delta^s(A) < \infty$ for all $\delta > 0$. For each $k \geq 1$, choose closed sets $\{C_j^k\}_{j=1}^{\infty}$ so that diam $C_j^k \leq 1/k$, $A \subset \cup_{j=1}^{\infty} C_j^k$, and

$$\sum_{j=1}^{\infty} \alpha(s)\left(\frac{\operatorname{diam} C_j^k}{2}\right)^s \leq \mathcal{H}_{1/k}^s(A) + \frac{1}{k}.$$

Let $A_k \equiv \cup_{j=1}^{\infty} C_j^k$, $B \equiv \cap_{k=1}^{\infty} A_k$; B is Borel. Also $A \subset A_k$ for each k, and so $A \subset B$. Furthermore,

$$\mathcal{H}^s_{1/k}(B) \le \sum_{j=1}^{\infty} \alpha(s) \left(\frac{\text{diam } C_j^k}{2} \right)^s \le \mathcal{H}^s_{1/k}(A) + \frac{1}{k}.$$

Letting $k \to \infty$, we discover $\mathcal{H}^s(B) \le \mathcal{H}^s(A)$. But $A \subset B$, and thus $\mathcal{H}^s(A) = \mathcal{H}^s(B)$. ▌

THEOREM 2 ELEMENTARY PROPERTIES OF HAUSDORFF MEASURE

(i) *$\mathcal{H}^0$ is counting measure.*

(ii) *$\mathcal{H}^1 = \mathcal{L}^1$ on $\mathbb{R}^1$.*

(iii) *$\mathcal{H}^s \equiv 0$ on $\mathbb{R}^n$ for all $s > n$.*

(iv) *$\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ for all $\lambda > 0$, $A \subset \mathbb{R}^n$.*

(v) *$\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$ for each affine isometry $L : \mathbb{R}^n \to \mathbb{R}^n$, $A \subset \mathbb{R}^n$.*

PROOF

1. Statements (iv) and (v) are easy.

2. First observe $\alpha(0) = 1$. Thus obviously $\mathcal{H}^0(\{a\}) = 1$ for all $a \in \mathbb{R}^n$, and (i) follows.

3. Choose $A \subset \mathbb{R}^1$ and $\delta > 0$. Then

$$\begin{aligned} \mathcal{L}^1(A) &= \inf \left\{ \sum_{j=1}^{\infty} \text{diam } C_j \mid A \subset \bigcup_{j=1}^{\infty} C_j \right\} \\ &\le \inf \left\{ \sum_{j=1}^{\infty} \text{diam } C_j \mid A \subset \bigcup_{j=1}^{\infty} C_j, \text{diam } C_j \le \delta \right\} \\ &= \mathcal{H}^1_\delta(A). \end{aligned}$$

On the other hand, set $I_k \equiv [k\delta, (k+1)\delta]$ $(k = \ldots -1, 0, 1, \ldots)$. Then diam $(C_j \cap I_k) \le \delta$ and

$$\sum_{k=-\infty}^{\infty} \text{diam } (C_j \cap I_k) \le \text{diam } C_j.$$

Hence

$$\begin{aligned} \mathcal{L}^1(A) &= \inf \left\{ \sum_{j=1}^{\infty} \text{diam } C_j \mid A \subset \bigcup_{j=1}^{\infty} C_j \right\} \\ &\ge \inf \left\{ \sum_{j=1}^{\infty} \sum_{k=-\infty}^{\infty} \text{diam } (C_j \cap I_k) \mid A \subset \bigcup_{j=1}^{\infty} C_j \right\} \\ &\ge \mathcal{H}^1_\delta(A). \end{aligned}$$

Thus $\mathcal{L}^1 = \mathcal{H}^1_\delta$ for all $\delta > 0$, and so $\mathcal{L}^1 = \mathcal{H}^1$ on $\mathbb{R}^1$.

4. Fix an integer $m \geq 1$. The unit cube Q in $\mathbb{R}^n$ can be decomposed into m^n cubes with side $1/m$ and diameter $n^{1/2}/m$. Therefore

$$\mathcal{H}^s_{\sqrt{n}/m}(Q) \leq \sum_{i=1}^{m^n} \alpha(s)(n^{1/2}/m)^s = \alpha(s) n^{s/2} m^{n-s},$$

and the last term goes to zero as $m \to \infty$, if $s > n$. Hence $\mathcal{H}^s(Q) = 0$, and so $\mathcal{H}^s(\mathbb{R}^n) = 0$. ∎

A convenient way to verify that $\mathcal{H}^s$ vanishes on a set is this.

LEMMA 1
Suppose $A \subset \mathbb{R}^n$ and $\mathcal{H}^s_\delta(A) = 0$ for some $0 < \delta \leq \infty$. Then $\mathcal{H}^s(A) = 0$.

PROOF The conclusion is obvious for $s = 0$, and so we may assume $s > 0$.

Fix $\epsilon > 0$. There then exist sets $\{C_j\}_{j=1}^\infty$ such that $A \subset \cup_{j=1}^\infty C_j$, and

$$\sum_{j=1}^\infty \alpha(s)\left(\frac{\text{diam } C_j}{2}\right)^s \leq \epsilon.$$

In particular for each i,

$$\text{diam } C_i \leq 2\left(\frac{\epsilon}{\alpha(s)}\right)^{1/s} \equiv \delta(\epsilon).$$

Hence

$$\mathcal{H}^s_{\delta(\epsilon)}(A) \leq \epsilon.$$

Since $\delta(\epsilon) \to 0$ as $\epsilon \to 0$, we find

$$\mathcal{H}^s(A) = 0. \quad ∎$$

We want next to define the Hausdorff dimension of a subset of $\mathbb{R}^n$.

LEMMA 2
Let $A \subset \mathbb{R}^n$ *and* $0 \le s < t < \infty$.

(i) *If* $\mathcal{H}^s(A) < \infty$, *then* $\mathcal{H}^t(A) = 0$.
(ii) *If* $\mathcal{H}^t(A) > 0$, *then* $\mathcal{H}^s(A) = +\infty$.

PROOF Let $\mathcal{H}^s(A) < \infty$ and $\delta > 0$. Then there exist sets $\{C_j\}_{j=1}^{\infty}$ such that diam $C_j \le \delta$, $A \subset \cup_{i=1}^{\infty} C_j$ and

$$\sum_{j=1}^{\infty} \alpha(s)\left(\frac{\text{diam } C_j}{2}\right)^s \le \mathcal{H}^s_\delta(A) + 1 \le \mathcal{H}^s(A) + 1.$$

Then

$$\begin{aligned}\mathcal{H}^t_\delta(A) &\le \sum_{j=1}^{\infty} \alpha(t)\left(\frac{\text{diam } C_j}{2}\right)^t \\ &= \frac{\alpha(t)}{\alpha(s)} 2^{s-t} \sum_{j=1}^{\infty} \alpha(s)\left(\frac{\text{diam } C_j}{2}\right)^s (\text{diam } C_j)^{t-s} \\ &\le \frac{\alpha(t)}{\alpha(s)} 2^{s-t} \delta^{t-s} (\mathcal{H}^s(A) + 1).\end{aligned}$$

We send $\delta \to 0$ to conclude $\mathcal{H}^t(A) = 0$. This proves assertion (i). Assertion (ii) follows at once from (i). ∎

*DEFINITION The **Hausdorff dimension** of a set* $A \subset \mathbb{R}^n$ *is defined to be*

$$\mathcal{H}_{\text{dim}}(A) \equiv \inf\{0 \le s < \infty \mid \mathcal{H}^s(A) = 0\}.$$

REMARK Observe $\mathcal{H}_{\text{dim}}(A) \le n$. Let $s = \mathcal{H}_{\text{dim}}(A)$. Then $\mathcal{H}^t(A) = 0$ for all $t > s$ and $\mathcal{H}^t(A) = +\infty$ for all $t < s$; $\mathcal{H}^s(A)$ may be any number between 0 and ∞, inclusive. Furthermore, $\mathcal{H}_{\text{dim}}(A)$ need not be an integer. Even if $\mathcal{H}_{\text{dim}}(A) = k$ is an integer and $0 < \mathcal{H}^k(A) < \infty$, A need not be a "k-dimensional surface" in any sense; see Falconer [FA] or Federer [F] for examples of extremely complicated Cantor-like subsets A of $\mathbb{R}^n$, with $0 < \mathcal{H}^k(A) < \infty$. ∎

2.2 Isodiametric Inequality; $\mathcal{H}^n = \mathcal{L}^n$

Our goal in this section is to prove $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$. This is nontrivial: $\mathcal{L}^n$ is defined as the n-fold product of one-dimensional Lebesgue measure $\mathcal{L}^1$, whence

$\mathcal{L}^n = \inf\{\sum_{i=1}^{\infty} \mathcal{L}^n(Q_i) \mid Q_i \text{ cubes}, A \subset \cup_{i=1}^{\infty} Q_i\}$. On the other hand $\mathcal{H}^n(A)$ is computed in terms of *arbitrary* coverings of small diameter.

LEMMA 1

Let $f : \mathbb{R}^n \to [0, \infty]$ be $\mathcal{L}^n$-measurable. Then the region "under the graph of f,"

$$A \equiv \{(x, y) \mid x \in \mathbb{R}^n, y \in \mathbb{R}, 0 \leq y \leq f(x)\},$$

is $\mathcal{L}^{n+1}$-measurable.

PROOF Let $B \equiv \{x \in \mathbb{R}^n \mid f(x) = \infty\}$ and $C \equiv \{x \in \mathbb{R}^n \mid 0 \leq f(x) < \infty\}$. In addition, define

$$C_{jk} \equiv \left\{x \in C \mid \frac{j}{k} \leq f(x) < \frac{j+1}{k}\right\} \qquad (j = 0, \ldots; k = 1, \ldots),$$

so that $C = \cup_{j=0}^{\infty} C_{jk}$. Finally, set

$$D_k \equiv \bigcup_{j=0}^{\infty} \left(C_{jk} \times \left[0, \frac{j}{k}\right]\right) \cup (B \times [0, \infty]),$$

$$E_k \equiv \bigcup_{j=0}^{\infty} \left(C_{jk} \times \left[0, \frac{j+1}{k}\right]\right) \cup (B \times [0, \infty]).$$

Then D_k and E_k are $\mathcal{L}^{n+1}$-measurable and $D_k \subset A \subset E_k$. Write $D \equiv \cup_{k=1}^{\infty} D_k$ and $E \equiv \cap_{k=1}^{\infty} E_k$. Then also $D \subset A \subset E$, with D and E both $\mathcal{L}^{n+1}$-measurable. Now

$$\mathcal{L}^{n-1}((E - D) \cap B(0, R)) \leq \mathcal{L}^{n+1}((E_k - D_k) \cap B(0, R)) \leq \frac{1}{k}\mathcal{L}^n(B(0, R)),$$

and the last term goes to zero as $k \to \infty$. Thus, $\mathcal{L}^{n+1}((E - D) \cap B(0, R)) = 0$, and so $\mathcal{L}^{n+1}(E - D) = 0$. Hence $\mathcal{L}^{n+1}(A - D) = 0$, and consequently A is $\mathcal{L}^{n-1}$-measurable. ▮

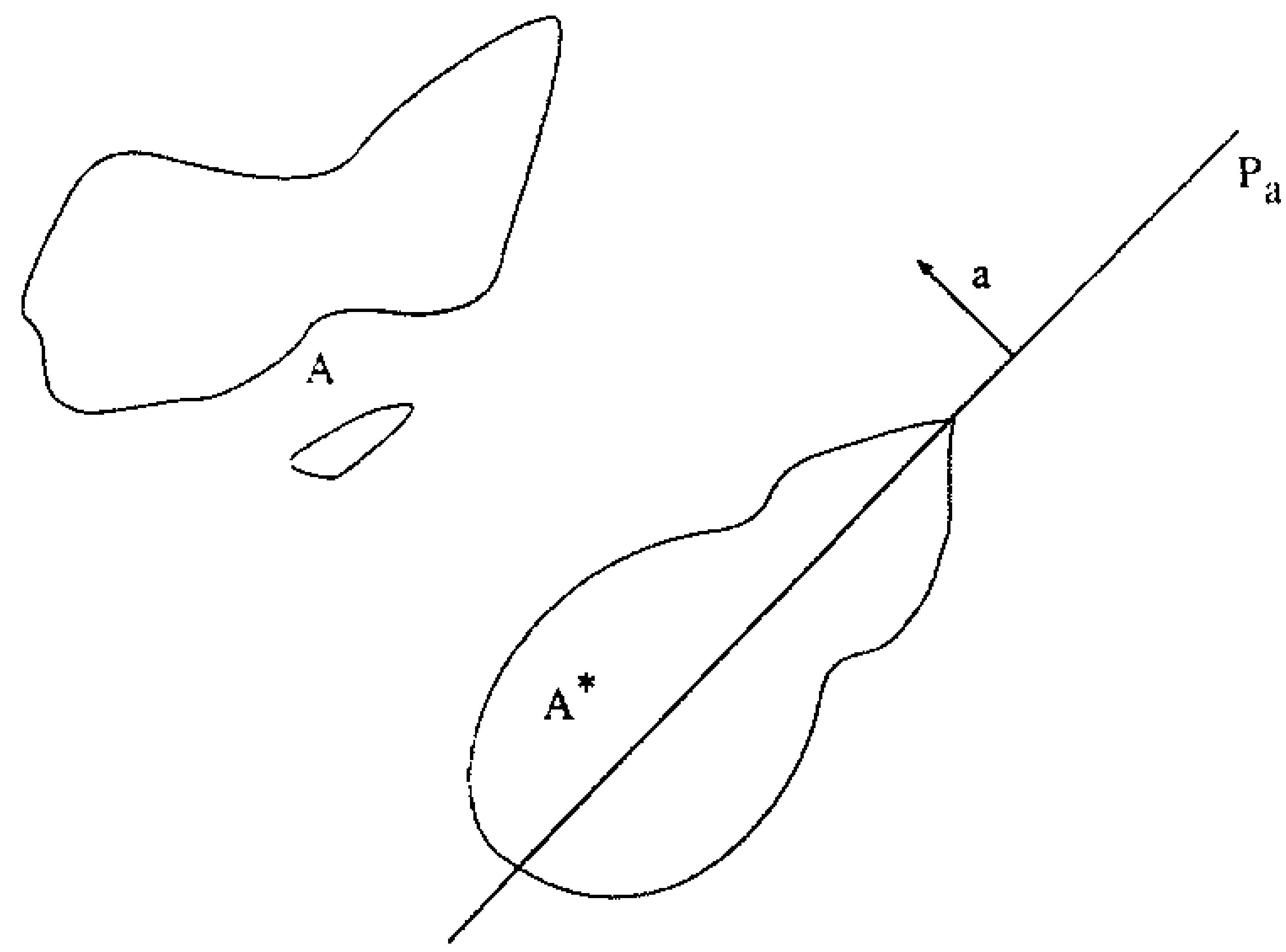

FIGURE 2.1
Steiner symmetrization.

NOTATION Fix $a, b \in \mathbb{R}^n$, $|a| = 1$. We define

$L_b^a \equiv \{b + ta \mid t \in \mathbb{R}\}$, the line through b in the direction a,

$P_a \equiv \{x \in \mathbb{R}^n \mid x \cdot a = 0\}$, the plane through the origin perpendicular to a.

DEFINITION Choose $a \in \mathbb{R}^n$ with $|a| = 1$, and let $A \subset \mathbb{R}^n$. We define the ***Steiner symmetrization*** *of A with respect to the plane P_a to be the set*

$$S_a(A) \equiv \bigcup_{\substack{b\in P_a \\ A\cap L_b^a \neq \emptyset}} \left\{ b + ta \,\middle|\, |t| \leq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a) \right\}.$$

LEMMA 2 PROPERTIES OF STEINER SYMMETRIZATION

(i) diam $S_a(A) \leq$ diam A.

(ii) If A is $\mathcal{L}^n$-measurable, then so is $S_a(A)$; and $\mathcal{L}^n(S_a(A)) = \mathcal{L}^n(A)$.

PROOF

1. Statement (i) is trivial if diam $A = \infty$; assume therefore diam $A < \infty$.

We may also suppose A closed. Fix $\epsilon > 0$ and select $x, y \in S_a(A)$ such that

$$\text{diam } S_a(A) \leq |x - y| + \epsilon.$$

Write $b \equiv x - (x \cdot a)a$ and $c \equiv y - (y \cdot a)a$; then $b, c \in P_a$. Set

$$r \equiv \inf\{t \mid b + ta \in A\},$$

$$s \equiv \sup\{t \mid b + ta \in A\},$$

$$u \equiv \inf\{t \mid c + ta \in A\},$$

$$v \equiv \sup\{t \mid c + ta \in A\}.$$

Without loss of generality, we may assume $v - r \geq s - u$. Then

$$\begin{aligned} v - r &\geq \frac{1}{2}(v - r) + \frac{1}{2}(s - u) \\ &= \frac{1}{2}(s - r) + \frac{1}{2}(v - u) \\ &\geq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a) + \frac{1}{2}\mathcal{H}^1(A \cap L_c^a). \end{aligned}$$

Now, $|x \cdot a| \leq 1/2\ \mathcal{H}^1(A \cap L_b^a)$, $|y \cdot a| \leq 1/2\ \mathcal{H}^1(A \cap L_c^a)$, and consequently,

$$v - r \ \geq\ |x \cdot a| + |y \cdot a| \ \geq\ |x \cdot a - y \cdot a|.$$

Therefore,

$$\begin{aligned} (\text{diam } S_a(A) - \epsilon)^2 &\leq |x - y|^2 \\ &= |b - c|^2 + |x \cdot a - y \cdot a|^2 \\ &\leq |b - c|^2 + (v - r)^2 \\ &= |(b + ra) - (c + va)|^2 \\ &\leq (\text{diam } A)^2, \end{aligned}$$

since A is closed and so $b + ra, c + va \in A$. Thus diam $S_a(A) - \epsilon \leq$ diam A. This establishes (i).

2. As $\mathcal{L}^n$ is rotation invariant, we may assume $a = e_n = (0, \ldots, 0, 1)$. Then $P_a = P_{e_n} = \mathbb{R}^{n-1}$. Since $\mathcal{L}^1 = \mathcal{H}^1$ on $\mathbb{R}^1$, Fubini's Theorem implies the map $f : \mathbb{R}^{n-1} \to \mathbb{R}$ defined by $f(b) = \mathcal{H}^1(A \cap L_b^a)$ is $\mathcal{L}^{n-1}$-measurable and $\mathcal{L}^n(A) = \int_{\mathbb{R}^{n-1}} f(b)\, db$. Hence

$$S_a(A) \equiv \left\{(b, y) \mid \frac{-f(b)}{2} \leq y \leq \frac{f(b)}{2}\right\} - \{(b, 0) \mid L_b^a \cap A = \emptyset\}$$

is $\mathcal{L}^n$-measurable by Lemma 1, and

$$\mathcal{L}^n(S_a(A)) = \int_{\mathbb{R}^{n-1}} f(b)\, db = \mathcal{L}^n(A). \quad \blacksquare$$

REMARK In proving $\mathcal{H}^n = \mathcal{L}^n$ below, observe we only use statement (ii) above in the special case that a is a standard coordinate vector. Since $\mathcal{H}^n$ is obviously rotation invariant, we therefore in fact prove $\mathcal{L}^n$ is rotation invariant. $\blacksquare$

THEOREM 1 ISODIAMETRIC INEQUALITY
For all sets $A \subset \mathbb{R}^n$,

$$\mathcal{L}^n(A) \le \alpha(n)\left(\frac{\operatorname{diam} A}{2}\right)^n.$$

REMARK This is interesting since it is not necessarily the case that A is contained in a ball of diameter diam A. $\blacksquare$

PROOF If diam $A = \infty$, this is trivial; let us therefore suppose diam $A < \infty$. Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$. Define $A_1 \equiv S_{e_1}(A)$, $A_2 \equiv S_{e_2}(A_1), \ldots, A_n \equiv S_{e_n}(A_{n-1})$. Write $A^* = A_n$.

1. Claim #1: A^* is symmetric with respect to the origin.

Proof of Claim #1: Clearly A_1 is symmetric with respect to P_{e_1}. Let $1 \le k < n$ and suppose A_k is symmetric with respect to $P_{e_1}, \ldots, P_{e_k}$. Clearly $A_{k+1} = S_{e_{k+1}}(A_k)$ is symmetric with respect to $P_{e_{k+1}}$. Fix $1 \le j \le k$ and let $S_j : \mathbb{R}^n \to \mathbb{R}^n$ be reflection through P_{e_j}. Let $b \in P_{e_{k+1}}$. Since $S_j(A_k) = A_k$,

$$\mathcal{H}^1(A_k \cap L_b^{e_{k+1}}) = \mathcal{H}^1(A_k \cap L_{S_j b}^{e_{k+1}});$$

consequently

$$\{t \mid b + te_{k+1} \in A_{k+1}\} = \{t \mid S_j b + te_{k+1} \in A_{k+1}\}.$$

Thus $S_j(A_{k+1}) = A_{k+1}$; that is, A_{k+1} is symmetric with respect to P_{e_j}. Thus $A^* = A_n$ is symmetric with respect to $P_{e_1}, \ldots, P_{e_n}$ and so with respect to the origin.

2. Claim #2: $\mathcal{L}^n(A^*) \le \alpha(n)\left(\dfrac{\operatorname{diam} A^*}{2}\right)^n$.

Proof of Claim #2: Choose $x \in A^*$. Then $-x \in A^*$ by Claim #1, and so diam $A^* \geq 2|x|$. Thus $A^* \subset B(0, \text{diam } A^*/2)$ and consequently

$$\mathcal{L}^n(A^*) \leq \mathcal{L}^n\left(B\left(0, \frac{\text{diam } A^*}{2}\right)\right) = \alpha(n)\left(\frac{\text{diam } A^*}{2}\right)^n.$$

3. *Claim #3:* $\mathcal{L}^n(A) \leq \alpha(n)(\text{diam } A/2)^n$.

Proof of Claim #3: $\overline{A}$ is $\mathcal{L}^n$-measurable, and thus Lemma 2 implies

$$\mathcal{L}^n((\overline{A})^*) = \mathcal{L}^n(\overline{A}) \quad , \quad \text{diam } (\overline{A})^* \leq \text{diam } \overline{A}.$$

Hence

$$\begin{aligned}
\mathcal{L}^n(A) \leq \mathcal{L}^n(\overline{A}) &= \mathcal{L}^n((\overline{A})^*) \\
&\leq \alpha(n)\left(\frac{\text{diam } (\overline{A})^*}{2}\right)^n \qquad \text{by Claim \#2} \\
&\leq \alpha(n)\left(\frac{\text{diam } \overline{A}}{2}\right)^n \\
&= \alpha(n)\left(\frac{\text{diam } A}{2}\right)^n. \quad \blacksquare
\end{aligned}$$

THEOREM 2
$\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$.

1. *Claim #1:* $\mathcal{L}^n(A) \leq \mathcal{H}^n(A)$ for all $A \subset \mathbb{R}^n$.

Proof of Claim #1: Fix $\delta > 0$. Choose sets $\{C_j\}_{j=1}^\infty$ such that $A \subset \cup_{j=1}^\infty C_j$ and diam $C_j \leq \delta$. Then by the Isodiametric Inequality,

$$\mathcal{L}^n(A) \leq \sum_{j=1}^\infty \mathcal{L}^n(C_j) \leq \sum_{j=1}^\infty \alpha(n)\left(\frac{\text{diam } C_j}{2}\right)^n.$$

Taking infima, we find $\mathcal{L}^n(A) \leq \mathcal{H}^n_\delta(A)$, and thus $\mathcal{L}^n(A) \leq \mathcal{H}^n(A)$.

2. Now, from the definition of $\mathcal{L}^n$ as $\mathcal{L}^1 \times \cdots \times \mathcal{L}^1$, we see that for all $A \subset \mathbb{R}^n$ and $\delta > 0$,

$$\mathcal{L}^n(A) = \inf\left\{\sum_{i=1}^\infty \mathcal{L}^n(Q_i) \mid Q_i \text{ cubes}, A \subset \bigcup_{i=1}^\infty Q_i, \text{diam } Q_i \leq \delta\right\}.$$

Here and afterwards we consider only cubes parallel to the coordinate axes in $\mathbb{R}^n$.

3. *Claim #2:* $\mathcal{H}^n$ is absolutely continuous with respect to $\mathcal{L}^n$.

Proof of Claim #2: Set $C_n \equiv \alpha(n)(\sqrt{n}/2)^n$. Then for each cube $Q \subset \mathbb{R}^n$,

$$\alpha(n)\left(\frac{\text{diam } Q}{2}\right)^n = C_n \mathcal{L}^n(Q).$$

Thus

$$\mathcal{H}^n_\delta(A) \leq \inf\left\{\sum_{i=1}^\infty \alpha(n)\left(\frac{\text{diam } Q_i}{2}\right)^n \mid Q_i \text{ cubes, } A \subset \bigcup_{i=1}^\infty Q_i, \text{diam } Q_i \leq \delta\right\}$$
$$= C_n \mathcal{L}^n(A).$$

Let $\delta \to 0$.

4. *Claim #3*: $\mathcal{H}^n(A) \leq \mathcal{L}^n(A)$ for all $A \subset \mathbb{R}^n$.

Proof of Claim #3: Fix $\delta > 0$, $\epsilon > 0$. We can select cubes $\{Q_i\}_{i=1}^\infty$ such that $A \subset \cup_{i=1}^\infty Q_i$, diam $Q_i < \delta$, and

$$\sum_{i=1}^\infty \mathcal{L}^n(Q_i) \leq \mathcal{L}^n(A) + \epsilon.$$

By Corollary 1 in Section 1.5, for each i there exist disjoint closed balls $\{B^i_k\}_{k=1}^\infty$ contained in Q_i^o such that

$$\text{diam } B^i_k \leq \delta, \quad \mathcal{L}^n\left(Q_i - \bigcup_{k=1}^\infty B^i_k\right) = \mathcal{L}^n\left(Q_i^o - \bigcup_{k=1}^\infty B^i_k\right) = 0.$$

By Claim #2, $\mathcal{H}^n\left(Q_i - \cup_{k=1}^\infty B^i_k\right) = 0$. Thus

$$\mathcal{H}^n_\delta(A) \leq \sum_{i=1}^\infty \mathcal{H}^n_\delta(Q_i) = \sum_{i=1}^\infty \mathcal{H}^n_\delta\left(\bigcup_{k=1}^\infty B^i_k\right) \leq \sum_{i=1}^\infty\sum_{k=1}^\infty \mathcal{H}^n_\delta(B^i_k)$$
$$\leq \sum_{i=1}^\infty\sum_{k=1}^\infty \alpha(n)\left(\frac{\text{diam } B^i_k}{2}\right)^n = \sum_{i=1}^\infty\sum_{k=1}^\infty \mathcal{L}^n(B^i_k) = \sum_{i=1}^\infty \mathcal{L}^n\left(\bigcup_{k=1}^\infty B^i_k\right)$$
$$= \sum_{i=1}^\infty \mathcal{L}^n(Q_i) \leq \mathcal{L}^n(A) + \epsilon. \quad \blacksquare$$

2.3 Densities

We proved in Section 1.7

$$\lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap E)}{\alpha(n)r^n} = \begin{cases} 1 & \text{for } \mathcal{L}^n \text{ a.e. } x \in E \\ 0 & \text{for } \mathcal{L}^n \text{ a.e. } x \in \mathbb{R}^n - E, \end{cases}$$

provided $E \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable. This section develops some analogous statements for lower dimensional Hausdorff measures. We assume throughout $0 < s < n$.

THEOREM 1
Assume $E \subset \mathbb{R}^n$, E is $\mathcal{H}^s$-measurable, and $\mathcal{H}^s(E) < \infty$. Then

$$\lim_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} = 0$$

for $\mathcal{H}^s$ a.e. $x \in \mathbb{R}^n - E$.

PROOF Fix $t > 0$ and define

$$A_t \equiv \left\{ x \in \mathbb{R}^n - E \mid \limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

Now $\mathcal{H}^s \llcorner E$ is a Radon measure, and so given $\epsilon > 0$, there exists a compact set $K \subset E$ such that

$$\mathcal{H}^s(E - K) \leq \epsilon. \tag{$\star$}$$

Set $U \equiv \mathbb{R}^n - K$; U is open and $A_t \subset U$. Fix $\delta > 0$ and consider

$$\mathcal{F} \equiv \left\{ B(x,r) \mid B(x,r) \subset U, 0 < r < \delta, \ \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

By the Vitali Covering Theorem, there exists a countable disjoint family of balls $\{B_i\}_{i=1}^{\infty}$ in $\mathcal{F}$ such that

$$A_t \subset \bigcup_{i=1}^{\infty} \hat{B}_i.$$

Write $B_i = B(x_i, r_i)$. Then

$$\begin{aligned} \mathcal{H}^s_{10\delta}(A_t) &\leq \sum_{i=1}^{\infty} \alpha(s)(5r_i)^s \leq \frac{5^s}{t} \sum_{i=1}^{\infty} \mathcal{H}^s(B_i \cap E) \\ &\leq \frac{5^s}{t} \mathcal{H}^s(U \cap E) = \frac{5^s}{t} \mathcal{H}^s(E - K) \leq \frac{5^s}{t} \epsilon, \end{aligned}$$

by ($\star$). Let $\delta \to 0$ to find $\mathcal{H}^s(A_t) \leq 5^s t^{-1} \epsilon$. Thus $\mathcal{H}^s(A_t) = 0$ for each $t > 0$, and the theorem follows. ∎

THEOREM 2
Assume $E \subset \mathbb{R}^n$, E is $\mathcal{H}^s$-measurable, and $\mathcal{H}^s(E) < \infty$. Then

$$\frac{1}{2^s} \leq \limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} \leq 1$$

for $\mathcal{H}^s$ a.e. $x \in E$.

REMARK It is possible to have

$$\limsup_{r\to 0} \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s)r^s} < 1$$

and

$$\liminf_{r\to 0} \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s)r^s} = 0$$

for $\mathcal{H}^s$ a.e. $x \in E$, even if $0 < \mathcal{H}^s(E) < \infty$. ∎

REMARK Pf.

1. Claim #1: $\limsup_{r\to 0} \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s)r^s} \le 1$ for $\mathcal{H}^s$ a.e. $x \in E$.

Proof of Claim #1: Fix $\epsilon > 0$, $t > 1$ and define

$$B_t \equiv \left\{ x \in E \mid \limsup_{r\to 0} \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s)r^s} > t \right\}.$$

Since $\mathcal{H}^s \llcorner E$ is Radon, there exists an open set U containing B_t, with

$$\mathcal{H}^s(U \cap E) \le \mathcal{H}^s(B_t) + \epsilon. \tag{$\star$}$$

Define

$$\mathcal{F} \equiv \left\{ B(x,r) \mid B(x,r) \subset U,\ 0 < r < \delta,\ \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s)r^s} > t \right\}.$$

By Corollary 1 of the Vitali Covering Theorem in Section 1.5.1, there exists a countable disjoint family of balls $\{B_i\}_{i=1}^\infty$ in $\mathcal{F}$ such that

$$B_t \subset \bigcup_{i=1}^m B_i \cup \bigcup_{i=m+1}^\infty \hat{B}_i$$

for each $m = 1, 2, \ldots$. Write $B_i = B(x_i, r_i)$. Then

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(B_t) &\le \sum_{i=1}^m \alpha(s)r_i^s + \sum_{i=m+1}^\infty \alpha(s)(5r_i)^s \\
&\le \frac{1}{t}\sum_{i=1}^m \mathcal{H}^s(B_i\cap E) + \frac{5^s}{t}\sum_{i=m+1}^\infty \mathcal{H}^s(B_i \cap E) \\
&\le \frac{1}{t}\mathcal{H}^s(U\cap E) + \frac{5^s}{t}\mathcal{H}^s\left(\bigcup_{i=m+1}^\infty B_i \cap E\right).
\end{aligned}$$

This estimate is valid for $m = 1, \ldots$; thus our sending m to infinity yields the estimate

$$\mathcal{H}^s_{10\delta}(B_t) \le \frac{1}{t}\mathcal{H}^s(U\cap E) \le \frac{1}{t}(\mathcal{H}^s(B_t) + \epsilon),$$

by ($\star$). Let $\delta \to 0$ and then $\epsilon \to 0$:

$$\mathcal{H}^s(B_t) \leq \frac{1}{t}\mathcal{H}^s(B_t).$$

Since $\mathcal{H}^s(B_t) \leq \mathcal{H}^s(E) < \infty$, this implies $\mathcal{H}^s(B_t) = 0$ for each $t > 1$.

2. *Claim #2*: $\limsup_{r\to 0} \dfrac{\mathcal{H}^s_\infty(B(x,r)\cap E)}{\alpha(s)r^s} \geq \dfrac{1}{2^s}$ for $\mathcal{H}^s$ a.e. $x \in E$.

Proof of Claim #2: For $\delta > 0$, $1 > \tau > 0$, denote by $E(\delta,\tau)$ the set of points $x \in E$ such that

$$\mathcal{H}^s_\delta(C\cap E) \leq \tau\alpha(s)\left(\frac{\text{diam } C}{2}\right)^s$$

whenever $C \subset \mathbb{R}^n$, $x \in C$, diam $C \leq \delta$. Then if $\{C_i\}_{i=1}^\infty$ are subsets of $\mathbb{R}^n$ with diam $C_i \leq \delta$, $E(\delta,\tau) \subset \cup_{i=1}^\infty C_i$, $C_i \cap E(\delta,\tau) \neq \emptyset$, we have

$$\begin{aligned}\mathcal{H}^s_\delta(E(\delta,\tau)) &\leq \sum_{i=1}^\infty \mathcal{H}^s_\delta(C_i\cap E(\delta,\tau))\\ &\leq \sum_{i=1}^\infty \mathcal{H}^s_\delta(C_i\cap E)\\ &\leq \tau\sum_{i=1}^\infty \alpha(s)\left(\frac{\text{diam } C_i}{2}\right)^s.\end{aligned}$$

Hence $\mathcal{H}^s_\delta(E(\delta,\tau)) \leq \tau\mathcal{H}^s_\delta(E(\delta,\tau))$, and so $\mathcal{H}^s_\delta(E(\delta,\tau)) = 0$, since $0 < \tau < 1$ and $\mathcal{H}^s_\delta(E(\delta,\tau)) \leq \mathcal{H}^s_\delta(E) \leq \mathcal{H}^s(E) < \infty$. In particular,

$$\mathcal{H}^s(E(\delta, 1-\delta)) = 0. \tag{$\star$}$$

Now if $x \in E$ and

$$\limsup_{r\to 0}\frac{\mathcal{H}^s_\infty(B(x,r)\cap E)}{\alpha(s)r^s} < \frac{1}{2^s},$$

there exists $\delta > 0$ such that

$$\frac{\mathcal{H}^s_\infty(B(x,r)\cap E)}{\alpha(s)r^s} \leq \frac{1-\delta}{2^s} \tag{$\star\star$}$$

for all $0 < r \leq \delta$. Thus if $x \in C$ and diam $C \leq \delta$,

$$\begin{aligned}\mathcal{H}^s_\delta(C\cap E) &= \mathcal{H}^s_\infty(C\cap E)\\ &\leq \mathcal{H}^s_\infty(B(x,\text{diam } C)\cap E)\\ &\leq (1-\delta)\alpha(s)\left(\frac{\text{diam } C}{2}\right)^s\end{aligned}$$

by $(\star\star)$; consequently $x \in E(\delta, 1-\delta)$. But then

$$\left\{x \in E \mid \limsup_{r \to 0} \frac{\mathcal{H}^s_\infty(B(x,r) \cap E)}{\alpha(s) r^s} < \frac{1}{2^s}\right\} \subset \bigcup_{k=1}^\infty E(1/k, 1-1/k),$$

and so $(\star)$ finishes the proof of Claim #2.

3. Since $\mathcal{H}^s(B(x,r) \cap E) \geq \mathcal{H}^s_\infty(B(x,r) \cap E)$, Claim #2 at once implies the lower estimate in the statement of the theorem. ∎

2.4 Hausdorff measure and elementary properties of functions

In this section we record for later use some simple properties relating the behavior of functions and Hausdorff measure.

2.4.1 Hausdorff measure and Lipschitz mappings

DEFINITIONS

(i) *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called* ***Lipschitz*** *if there exists a constant C such that*

$$|f(x) - f(y)| \leq C|x-y| \text{ for all } x, y \in \mathbb{R}^n.$$

(ii) $\operatorname{Lip}(f) \equiv \sup\left\{\frac{|f(x)-f(y)|}{|x-y|} \mid x, y \in \mathbb{R}^n, x \neq y\right\}.$

THEOREM 1

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $A \subset \mathbb{R}^n$, $0 \leq s < \infty$. Then

$$\mathcal{H}^s(f(A)) \leq (\operatorname{Lip}(f))^s \mathcal{H}^s(A).$$

PROOF Fix $\delta > 0$ and choose sets $\{C_i\}_{i=1}^\infty \subset \mathbb{R}^n$ such that $\operatorname{diam} C_i \leq \delta$, $A \subset \cup_{i=1}^\infty C_i$. Then $\operatorname{diam} f(C_i) \leq \operatorname{Lip}(f) \operatorname{diam} C_i \leq \operatorname{Lip}(f)\delta$ and $f(A) \subset \cup_{i=1}^\infty f(C_i)$. Thus

$$\begin{aligned}\mathcal{H}^s_{\operatorname{Lip}(f)\delta}(f(A)) &\leq \sum_{i=1}^\infty \alpha(s)\left(\frac{\operatorname{diam} f(C_i)}{2}\right)^s \\ &\leq (\operatorname{Lip}(f))^s \sum_{i=1}^\infty \alpha(s)\left(\frac{\operatorname{diam} C_i}{2}\right)^s.\end{aligned}$$

Taking infima over all such sets $\{C_i\}_{i=1}^{\infty}$, we find

$$\mathcal{H}^s_{\mathrm{Lip}\,(f)\delta}(f(A)) \leq (\mathrm{Lip}\,(f))^s \mathcal{H}^s_\delta(A).$$

Send $\delta \to 0$ to finish the proof. ∎

COROLLARY 1
Suppose $n > k$. Let $P : \mathbb{R}^n \to \mathbb{R}^k$ be the usual projection, $A \subset \mathbb{R}^n$, $0 \leq s < \infty$. Then

$$\mathcal{H}^s(P(A)) \leq \mathcal{H}^s(A).$$

PROOF $\mathrm{Lip}\,(P) = 1$. ∎

2.4.2 Graphs of Lipschitz functions

DEFINITION *For $f : \mathbb{R}^n \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, write*

$$G(f; A) \equiv \{(x, f(x)) \mid x \in A\} \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m};$$

*$G(f; A)$ is the **graph** of f over A.*

THEOREM 2
Assume $f : \mathbb{R}^n \to \mathbb{R}^m$, $\mathcal{L}^n(A) > 0$.

(i) *Then $\mathcal{H}_{\mathrm{dim}}(G(f; A)) \geq n$.*
(ii) *If f is Lipschitz, $\mathcal{H}_{\mathrm{dim}}(G(f; A)) = n$.*

REMARK We thus see the graph of a Lipschitz function f has the expected Hausdorff dimension. We will later discover from the Area Formula in Section 3.3 that $\mathcal{H}^n(G(f; A))$ can be computed according to the usual rules of calculus. ∎

PROOF
1. Let $P : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be the projection. Then $\mathcal{H}^n(G(f; A)) \geq \mathcal{H}^n(A) > 0$ and thus $\mathcal{H}_{\mathrm{dim}}(G(f; A)) \geq n$.
2. Let Q denote any cube in $\mathbb{R}^n$ of side length 1. Subdivide Q into k^n subcubes of side length $1/k$. Call these subcubes $Q_1, \ldots, Q_{k^n}$. Note diam $Q_i = \sqrt{n}/k$. Define

$$a^i_j \equiv \min_{x \in Q_j} f^i(x) \quad \text{and} \quad b^i_j \equiv \max_{x \in Q_j} f^i(x) \qquad (i = 1, \ldots, m; j = 1, \ldots, k^n).$$

Since f is Lipschitz,

$$|b_j^i - a_j^i| \leq \mathrm{Lip}\,(f)\mathrm{diam}\; Q_j = \mathrm{Lip}\,(f)\frac{\sqrt{n}}{k}\,.$$

Next, let $C_j \equiv Q_j \times \prod_{i=1}^m (a_j^i, b_j^i)$. Then

$$\{(x, f(x)) \mid x \in Q_j \cap A\} \subset C_j$$

and diam $C_j \leq C/k$. Since $G(f; A \cap Q) \subset \cup_{j=1}^{k^n} C_j$, we have

$$\begin{aligned}
\mathcal{H}^n_{C/k}(G(f; A \cap Q)) &\leq \sum_{j=1}^{k^n} \alpha(n) \left(\frac{\mathrm{diam}\; C_j}{2}\right)^n \\
&\leq k^n \alpha(n) \left(\frac{C}{2k}\right)^n = \alpha(n) \left(\frac{C}{2}\right)^n .
\end{aligned}$$

Then, letting $k \to \infty$, we find $\mathcal{H}^n(G(f; A \cap Q)) < \infty$, and so $\mathcal{H}_{\mathrm{dim}}(G(f; A \cap Q)) \leq n$. This estimate is valid for each cube Q in $\mathbb{R}^n$ of side length 1, and consequently $\mathcal{H}_{\mathrm{dim}}(G(f; A)) \leq n$. ∎

2.4.3 The set where a summable function is large

If a function is locally summable, we can estimate the Hausdorff measure of the set where it is locally large.

THEOREM 3
Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, suppose $0 \leq s < n$, and define

$$\Lambda_s \equiv \left\{ x \in \mathbb{R}^n \mid \limsup_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f|\, dy > 0 \right\}.$$

Then

$$\mathcal{H}^s(\Lambda_s) = 0.$$

PROOF We may as well assume $f \in L^1(\mathbb{R}^n)$. By the Lebesgue–Besicovitch Differentiation Theorem (Section 1.7.1)

$$\lim_{r \to 0} ⨍_{B(x,r)} |f|\, dy = |f(x)|,$$

and thus

$$\lim_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f|\, dy = 0$$

for $\mathcal{L}^n$ a.e. x, since $0 \le s < n$. Hence

$$\mathcal{L}^n(\Lambda_s) = 0.$$

Now fix $\epsilon > 0$, $\delta > 0$, $\sigma > 0$. As f is $\mathcal{L}^n$-summable, there exists $\eta > 0$ such that

$$\mathcal{L}^n(U) \le \eta \ \text{ implies } \ \int_U |f| \, dx < \sigma.$$

Define

$$\Lambda_s^\epsilon \equiv \left\{ x \in \mathbb{R}^n \mid \limsup_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f| \, dy > \epsilon \right\};$$

by the preceding

$$\mathcal{L}^n(\Lambda_s^\epsilon) = 0.$$

There thus exists an open subset U with $U \supset \Lambda_s^\epsilon$, $\mathcal{L}^n(U) < \eta$. Set

$$\mathcal{F} \equiv \left\{ B(x,r) \mid x \in \Lambda_s^\epsilon, 0 < r < \delta, B(x,r) \subset U, \int_{B(x,r)} |f| \, dy > \epsilon r^s \right\}.$$

By the Vitali Covering Theorem, there exist disjoint balls $\{B_i\}_{i=1}^\infty$ in $\mathcal{F}$ such that

$$\Lambda_s^\epsilon \subset \bigcup_{i=1}^\infty \hat{B}_i.$$

Hence, writing r_i for the radius of B_i, we compute

$$\begin{aligned} \mathcal{H}_{10\delta}^s(\Lambda_s^\epsilon) &\le \sum_{i=1}^\infty \alpha(s)(5r_i)^s \\ &\le \frac{\alpha(s)5^s}{\epsilon} \sum_{i=1}^\infty \int_{B_i} |f| \, dy \\ &\le \frac{\alpha(s)5^s}{\epsilon} \int_U |f| \, dy \\ &\le \frac{\alpha(s)5^s}{\epsilon} \sigma. \end{aligned}$$

Send $\delta \to 0$, and then $\sigma \to 0$, to discover

$$\mathcal{H}^s(\Lambda_s^\epsilon) = 0. \quad \blacksquare$$

3

Area and Coarea Formulas

In this chapter we study Lipschitz mappings

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

and derive corresponding "change of variables" formulas. There are two essentially different cases depending on the relative size of n, m.

If $m \geq n$, the *Area Formula* asserts that the n-dimensional measure of $f(A)$, counting multiplicity, can be calculated by integrating the appropriate Jacobian of f over A.

If $m \leq n$, the *Coarea Formula* states that the integral of the $n - m$ dimensional measure of the level sets of f is computed by integrating the Jacobian. This assertion is a far-reaching generalization of Fubini's Theorem. (The word "coarea" is pronounced, and sometimes spelled, "co-area.")

We begin in Section 3.1 with a detailed study of the differentiability properties of Lipschitz functions and prove Rademacher's Theorem. In Section 3.2 we discuss linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ and introduce Jacobians. The Area Formula is proved in Section 3.3, the Coarea Formula in Section 3.4.

3.1 Lipschitz functions, Rademacher's Theorem

3.1.1 Lipschitz functions

We recall and extend slightly some terminology from Section 2.4.1.

DEFINITION

(i) *Let $A \subset \mathbb{R}^n$. A function $f : A \to \mathbb{R}^m$ is called* ***Lipschitz*** *provided*

$$|f(x) - f(y)| \leq C|x - y| \qquad (\star)$$

for some constant C and all $x, y \in A$. The smallest constant C such that $(\star)$ holds for all x, y is denoted

$$\operatorname{Lip}(f) \equiv \sup\left\{\frac{|f(x)-f(y)|}{|x-y|} \mid x, y \in A, x \neq y\right\}.$$

(ii) *A function $f : A \to \mathbb{R}^m$ is called* ***locally Lipschitz*** *if for each compact $K \subset A$, there exists a constant C_K such that*

$$|f(x)-f(y)| \leq C_K|x-y|$$

for all $x, y \in K$.

THEOREM 1 EXTENSION OF LIPSCHITZ FUNCTIONS
Assume $A \subset \mathbb{R}^n$, and let $f : A \to \mathbb{R}^m$ be Lipschitz. There exists a Lipschitz function $\bar{f} : \mathbb{R}^n \to \mathbb{R}^m$ such that

(i) $\bar{f} = f$ *on* A.

(ii) $\operatorname{Lip}(\bar{f}) \leq \sqrt{m}\operatorname{Lip}(f)$.

PROOF

1. First assume $f : A \to \mathbb{R}$. Define

$$\bar{f}(x) \equiv \inf_{a \in A}\{f(a) + \operatorname{Lip}(f)|x-a|\}.$$

If $b \in A$, then we have $\bar{f}(b) = f(b)$. This follows since for all $a \in A$,

$$f(a) + \operatorname{Lip}(f)|b-a| \geq f(b),$$

whereas obviously $\bar{f}(b) \leq f(b)$. If $x, y \in \mathbb{R}^n$, then

$$\begin{aligned}\bar{f}(x) &\leq \inf_{a \in A}\{f(a) + \operatorname{Lip}(f)(|y-a| + |x-y|)\}\\ &= \bar{f}(y) + \operatorname{Lip}(f)|x-y|,\end{aligned}$$

and similarly

$$\bar{f}(y) \leq \bar{f}(x) + \operatorname{Lip}(f)|x-y|.$$

2. In the general case $f : A \to \mathbb{R}^m$, $f = (f^1, \ldots, f^m)$, define $\bar{f} \equiv (\bar{f}^1, \ldots, \bar{f}^m)$ Then

$$|\bar{f}(x) - \bar{f}(y)|^2 = \sum_{i=1}^m |\bar{f}^i(x) - \bar{f}^i(y)|^2 \leq m(\operatorname{Lip}(f))^2|x-y|^2. \quad \blacksquare$$

REMARK *Kirszbraun's Theorem* (Federer [F, Section 2.10.43]) asserts that there in fact exists an extension $\bar{f}$ with $\operatorname{Lip}(\bar{f}) = \operatorname{Lip}(f)$. ∎

3.1.2 Rademacher's Theorem

We next prove Rademacher's remarkable theorem that a Lipschitz function is differentiable $\mathcal{L}^n$ a.e. This is surprising since the inequality

$$|f(x) - f(y)| \leq \mathrm{Lip}(f)|x - y|$$

apparently says nothing about the possibility of locally approximating f by a linear map.

*DEFINITION The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at $x \in \mathbb{R}^n$ if there exists a linear mapping*

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$\lim_{y \to x} \frac{|f(y) - f(x) - L(x - y)|}{|x - y|} = 0,$$

or, equivalently,

$$f(y) = f(x) + L(y - x) + o(|y - x|) \quad \textit{as} \quad y \to x.$$

NOTATION If such a linear mapping L exists, it is clearly unique, and we write

$$Df(x)$$

for L. We call $Df(x)$ the *derivative* of f at x.

THEOREM 2 RADEMACHER'S THEOREM
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. Then f is differentiable $\mathcal{L}^n$ a.e.

PROOF

1. We may assume $m = 1$. Since differentiability is a local property, we may as well also suppose f is Lipschitz.

2. Fix any $v \in \mathbb{R}^n$ with $|v| = 1$, and define

$$D_v f(x) \equiv \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \qquad (x \in \mathbb{R}^n),$$

provided this limit exists.

3. *Claim #1*: $D_v f(x)$ exists for $\mathcal{L}^n$ a.e. x.

Proof of Claim #1: Since f is continuous,

$$\overline{D}_v f(x) \equiv \limsup_{t\to 0} \frac{f(x+tv)-f(x)}{t}$$
$$= \lim_{k\to\infty} \sup_{\substack{0<|t|<1/k \\ t \text{ rational}}} \frac{f(x+tv)-f(x)}{t}$$

is Borel measurable, as is

$$\underline{D}_v f(x) \equiv \liminf_{t\to 0} \frac{f(x+tv)-f(x)}{t}.$$

Thus

$$A_v \equiv \{x \in \mathbb{R}^n \mid D_v f(x) \text{ does not exist}\}$$
$$= \{x \in \mathbb{R}^n \mid \underline{D}_v f(x) < \overline{D}_v f(x)\}$$

is Borel measurable.

Now, for each $x, v \in \mathbb{R}^n$, with $|v| = 1$, define $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t) \equiv f(x+tv) \qquad (t \in \mathbb{R}).$$

Then φ is Lipschitz, thus absolutely continuous, and thus differentiable $\mathcal{L}^1$ a.e. Hence

$$\mathcal{H}^1(A_v \cap L) = 0$$

for each line L parallel to v. Fubini's Theorem then implies

$$\mathcal{L}^n(A_v) = 0.$$

4. As a consequence of Claim #1, we see

$$\text{grad } f(x) \equiv \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)$$

exists for $\mathcal{L}^n$ a.e. x.

5. *Claim #2*: $D_v f(x) = v \cdot \text{grad } f(x)$ for $\mathcal{L}^n$ a.e. x.

Proof of Claim #2: Let $\zeta \in C_c^\infty(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} \left[\frac{f(x+tv)-f(x)}{t}\right] \zeta(x)\, dx = -\int_{\mathbb{R}^n} f(x) \left[\frac{\zeta(x)-\zeta(x-tv)}{t}\right] dx.$$

Let $t = 1/k$ for $k = 1, \ldots$ in the above equality and note

$$\left|\frac{f(x+\frac{1}{k}v)-f(x)}{\frac{1}{k}}\right| \le \text{Lip}\,(f)|v| = \text{Lip}\,(f).$$

Thus the Dominated Convergence Theorem implies

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\zeta(x)\,dx &= -\int_{\mathbb{R}^n} f(x) D_v\zeta(x)\,dx \\
&= -\sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} f(x)\frac{\partial \zeta}{\partial x_i}(x)\,dx \\
&= \sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x)\zeta(x)\,dx \\
&= \int_{\mathbb{R}^n} (v \cdot \text{grad } f(x))\zeta(x)\,dx,
\end{aligned}$$

where we used Fubini's Theorem and the absolute continuity of f on lines. The above equality holding for each $\zeta \in C_c(\mathbb{R}^n)$ implies $D_v f = v \cdot \text{grad } f$ $\mathcal{L}^n$ a.e.

6. Now choose $\{v_k\}_{k=1}^{\infty}$ to be a countable, dense subset of $\partial B(0,1)$. Set

$$A_k \equiv \{x \in \mathbb{R}^n \mid D_{v_k} f(x),\ \text{grad } f(x) \text{ exist and } D_{v_k} f(x) = v_k \cdot \text{grad } f(x)\}$$

for $k = 1, 2, \ldots$, and define

$$A \equiv \bigcap_{k=1}^{\infty} A_k.$$

Observe

$$\mathcal{L}^n(\mathbb{R}^n - A) = 0.$$

7. Claim #3: f is differentiable at each point $x \in A$.

Proof of Claim #3: Fix any $x \in A$. Choose $v \in \partial B(0,1)$, $t \in \mathbb{R}$, $t \neq 0$, and write

$$Q(x, v, t) \equiv \frac{f(x+tv) - f(x)}{t} - v \cdot \text{grad } f(x).$$

Then if $v' \in \partial B(0,1)$, we have

$$\begin{aligned}
|Q(x,v,t) - Q(x,v',t)| &\le \left|\frac{f(x+tv) - f(x+tv')}{t}\right| + |(v - v') \cdot \text{grad } f(x)| \\
&\le \text{Lip}(f)|v - v'| + |\text{grad } f(x)|\,|v - v'| \\
&\le (\sqrt{n} + 1)\text{Lip}(f)|v - v'|. \qquad (\star)
\end{aligned}$$

Now fix $\epsilon > 0$, and choose N so large that if $v \in \partial B(0,1)$, then

$$|v - v_k| \le \frac{\epsilon}{2(\sqrt{n}+1)\text{Lip}(f)} \quad \text{for some} \quad k \in \{1, \ldots, N\}. \qquad (\star\star)$$

Now

$$\lim_{t\to 0} Q(x, v_k, t) = 0 \qquad (k = 1, \dots, N),$$

and thus there exists $\delta > 0$ so that

$$|Q(x, v_k, t)| < \frac{\epsilon}{2} \text{ for all } 0 < |t| < \delta,\ k = 1, \dots, N. \qquad (\star\star\star)$$

Consequently, for each $v \in \partial B(0, 1)$, there exists $k \in \{1, \dots, N\}$ such that

$$|Q(x, v, t)| \leq |Q(x, v_k, t)| + |Q(x, v, t) - Q(x, v_k, t)| < \epsilon$$

if $0 < |t| < \delta$, according to $(\star)$ through $(\star\star\star)$. Note the same $\delta > 0$ works for all $v \in \partial B(0, 1)$.

Now choose any $y \in \mathbb{R}^n$, $y \neq x$. Write $v \equiv (y-x)/|y-x|$, so that $y = x+tv$, $t \equiv |x - y|$. Then

$$\begin{aligned} f(y) - f(x) - \operatorname{grad} f(x) \cdot (y - x) &= f(x + tv) - f(x) - tv \cdot \operatorname{grad} f(x) \\ &= o(t) \\ &= o(|x - y|), \text{ as } y \to x. \end{aligned}$$

Hence f is differentiable at x, with

$$Df(x) = \operatorname{grad} f(x). \quad \blacksquare$$

REMARK See Theorem 2 in Section 6.2 for another proof of Rademacher's Theorem and Theorem 1 in Section 6.2 for a generalization. In Section 6.4 we prove Aleksandrov's Theorem, stating that a convex function is twice differentiable a.e. ▮

We next record a technical lemma for use later.

COROLLARY 1

(i) *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz, and*

$$Z \equiv \{x \in \mathbb{R}^n \mid f(x) = 0\}.$$

Then $Df(x) = 0$ for $\mathcal{L}^n$ a.e. $x \in Z$.

(ii) *Let $f, g : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz, and*

$$Y \equiv \{x \in \mathbb{R}^n \mid g(f(x)) = x\}.$$

Then

$$Dg(f(x))Df(x) = I \quad \textit{for} \quad \mathcal{L}^n \text{ a.e. } x \in Y.$$

PROOF

1. We may assume $m = 1$ in assertion (i).

2. Choose $x \in Z$ so that $Df(x)$ exists, and

$$\lim_{r \to 0} \frac{\mathcal{L}^n(Z \cap B(x,r))}{\mathcal{L}^n(B(x,r))} = 1; \qquad (\star)$$

$\mathcal{L}^n$ a.e. $x \in Z$ will do. Then

$$f(y) = Df(x) \cdot (y - x) + o(|y - x|) \text{ as } y \to x \qquad (\star\star)$$

Assume $Df(x) \equiv a \neq 0$, and set

$$S \equiv \left\{ v \in \partial B(0,1) \mid a \cdot v \geq \frac{1}{2}|a| \right\}.$$

For each $v \in S$ and $t > 0$, set $y = x + tv$ in $(\star\star)$:

$$\begin{aligned} f(x + tv) &= a \cdot tv + o(|tv|) \\ &\geq \frac{t|a|}{2} + o(t) \qquad \text{as } t \to 0. \end{aligned}$$

Hence there exists $t_0 > 0$ such that

$$f(x + tv) > 0 \text{ for } 0 < t < t_0,\ v \in S,$$

a contradiction to $(\star)$. This proves assertion (i).

3. To prove assertion (ii), first define

$$\text{dmn } Df \equiv \{x \mid Df(x) \text{ exists }\},$$

$$\text{dmn } Dg \equiv \{x \mid Dg(x) \text{ exists }\}.$$

Let

$$X \equiv Y \cap \text{dmn } Df \cap f^{-1}(\text{dmn } Dg).$$

Then

$$Y - X \subset (\mathbb{R}^n - \text{dmn } Df) \cup g(\mathbb{R}^n - \text{dmn } Dg). \qquad (\star\star\star)$$

This follows since

$$x \in Y - f^{-1}(\text{dmn } Dg)$$

implies

$$f(x) \in \mathbb{R}^n - \text{dmn } Dg,$$

and so

$$x = g(f(x)) \in g(\mathbb{R}^n - \text{dmn } Dg).$$

According to $(\star\star\star)$ and Rademacher's Theorem,

$$\mathcal{L}^n(Y - X) = 0.$$

Now if $x \in X$, $Dg(f(x))$ and $Df(x)$ exist, and so

$$Dg(f(x))Df(x) = D(g \circ f)(x)$$

exists. Since $(g \circ f)(x) - x = 0$ on Y, assertion (i) implies

$$D(g \circ f) = I \qquad \mathcal{L}^n \text{ a.e. on } Y. \quad \blacksquare$$

3.2 Linear maps and Jacobians

We next review some basic linear algebra. Our goal thereafter will be to define the Jacobian of a map $f : \mathbb{R}^n \to \mathbb{R}^m$.

3.2.1 Linear maps

DEFINITIONS

(i) *A linear map $O : \mathbb{R}^n \to \mathbb{R}^m$ is **orthogonal** if $(Ox) \cdot (Oy) = x \cdot y$ for all $x, y \in \mathbb{R}^n$.*

(ii) *A linear map $S : \mathbb{R}^n \to \mathbb{R}^n$ is **symmetric** if $x \cdot (Sy) = (Sx) \cdot y$ for all $x, y \in \mathbb{R}^n$.*

(iii) *A linear map $D : \mathbb{R}^n \to \mathbb{R}^n$ is **diagonal** if there exist $d_1, \ldots, d_n \in \mathbb{R}$ such that $Dx = (d_1x_1, \ldots, d_nx_n)$ for all $x \in \mathbb{R}^n$.*

(iv) *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be linear. The **adjoint** of A is the linear map $A^\star : \mathbb{R}^m \to \mathbb{R}^n$ defined by $x \cdot (A^\star y) = (Ax) \cdot y$ for all $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.*

First we recall some routine facts from linear algebra.

THEOREM 1

(i) $A^{\star\star} = A$.

(ii) $(A \circ B)^\star = B^\star \circ A^\star$.

(iii) $O^\star = O^{-1}$ *if $O : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal.*

(iv) $S^\star = S$ *if $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric.*

(v) *If $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric, there exists an orthogonal map $O : \mathbb{R}^n \to \mathbb{R}^n$ and a diagonal map $D : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$S = O \circ D \circ O^{-1}.$$

(vi) *If* $O : \mathbb{R}^n \to \mathbb{R}^m$ *is orthogonal, then* $n \le m$ *and*

$$O^\star \circ O = I \qquad \text{on } \mathbb{R}^n,$$
$$O \circ O^\star = I \qquad \text{on } O(\mathbb{R}^n).$$

THEOREM 2 POLAR DECOMPOSITION
Let $L : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear mapping.*

(i) *If* $n \le m$, *there exists a symmetric map* $S : \mathbb{R}^n \to \mathbb{R}^n$ *and an orthogonal map* $O : \mathbb{R}^n \to \mathbb{R}^m$ *such that*

$$L = O \circ S.$$

(ii) *If* $n \ge m$, *there exists a symmetric map* $S : \mathbb{R}^m \to \mathbb{R}^m$ *and an orthogonal map* $O : \mathbb{R}^m \to \mathbb{R}^n$ *such that*

$$L = S \circ O^\star.$$

PROOF

1. First suppose $n \le m$. Consider $C \equiv L^\star \circ L : \mathbb{R}^n \to \mathbb{R}^n$. Now

$$\begin{aligned}(Cx) \cdot y = (L^\star \circ Lx) \cdot y &= Lx \cdot Ly \\ &= x \cdot L^\star \circ Ly \\ &= x \cdot Cy\end{aligned}$$

and also

$$(Cx) \cdot x = Lx \cdot Lx \ \ge \ 0.$$

Thus C is symmetric, nonnegative definite. Hence there exist $\mu_1, \dots, \mu_n \ge 0$ and an orthogonal basis $\{x_k\}_{k=1}^n$ of $\mathbb{R}^n$ such that

$$Cx_k = \mu_k x_k \qquad (k = 1, \dots, n).$$

Write $\mu_k \equiv \lambda_k^2$, $\lambda_k \ge 0$ $(k = 1, \dots, n)$.

2. Claim: There exists an orthonormal set $\{z_k\}_{k=1}^n$ in $\mathbb{R}^m$ such that

$$Lx_k = \lambda_k z_k \qquad (k = 1, \dots, n).$$

Proof of Claim: If $\lambda_k \ne 0$, define

$$z_k \equiv \frac{1}{\lambda_k} Lx_k.$$

Then if $\lambda_k, \lambda_l \ne 0$,

$$\begin{aligned}z_k \cdot z_l &= \frac{1}{\lambda_k \lambda_l} Lx_k \cdot Lx_l = \frac{1}{\lambda_k \lambda_l}(Cx_k) \cdot x_l \\ &= \frac{\lambda_k^2}{\lambda_k \lambda_l} x_k \cdot x_l = \frac{\lambda_k}{\lambda_l}\delta_{kl}.\end{aligned}$$

Thus the set $\{z_k \mid \lambda_k \neq 0\}$ is orthonormal. If $\lambda_k = 0$, define z_k to be any unit vector such that $\{z_k\}_{k=1}^n$ is orthonormal.

3. Now define

$$S : \mathbb{R}^n \to \mathbb{R}^n \text{ by } Sx_k = \lambda_k x_k \qquad (k = 1, \ldots, n)$$

and

$$O : \mathbb{R}^n \to \mathbb{R}^m \text{ by } Ox_k = z_k \qquad (k = 1, \ldots, n).$$

Then $O \circ Sx_k = \lambda_k Ox_k = \lambda_k z_k = Lx_k$, and so

$$L = O \circ S.$$

The mapping S is clearly symmetric, and O is orthogonal since

$$Ox_k \cdot Ox_l = z_k \cdot z_l = \delta_{kl}.$$

4. Assertion (ii) follows from our applying (i) to $L^* : \mathbb{R}^m \to \mathbb{R}^n$. ▮

DEFINITION *Assume $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear.*

(i) *If $n \leq m$, we write $L = O \circ S$ as above, and we define the* ***Jacobian*** *of L to be*

$$[\![L]\!] = |\det S|.$$

(ii) *If $n \geq m$, we write $L = S \circ O^*$ as above, and we define the* ***Jacobian*** *of L to be*

$$[\![L]\!] = |\det S|.$$

REMARKS

(i) It follows from Theorem 3 below that the definition of $[\![L]\!]$ is independent of the particular choices of O and S.

(ii) Clearly,

$$[\![L]\!] = [\![L^*]\!]. \quad ▮$$

THEOREM 3

(i) *If $n \leq m$,*

$$[\![L]\!]^2 = \det(L^* \circ L).$$

(ii) *If $n \geq m$,*

$$[\![L]\!]^2 = \det(L \circ L^*).$$

PROOF

1. Assume $n \leq m$ and write

$$L = O \circ S, L^\star = S^\star \circ O^\star = S \circ O^\star;$$

then

$$L^\star \circ L = S \circ O^\star \circ O \circ S = S^2,$$

since O is orthogonal, and thus $O^\star \circ O = I$. Hence

$$\det(L^\star \circ L) = (\det S)^2 = [\![L]\!]^2.$$

2. The proof of (ii) is similar. ▮

Theorem 3 provides us with a useful method for computing $[\![L]\!]$, which we augment with the Binet–Cauchy formula below.

DEFINITIONS

(i) *If $n \leq m$, we define*

$$\Lambda(m,n) = \{\lambda : \{1,\dots,n\} \to \{1,\dots,m\} \mid \lambda \textit{ is increasing}\}.$$

(ii) *For each $\lambda \in \Lambda(m,n)$, we define $P_\lambda : \mathbb{R}^m \to \mathbb{R}^n$ by*

$$P_\lambda(x_1,\dots,x_m) \equiv (x_{\lambda(1)},\dots,x_{\lambda(n)}).$$

REMARK For each $\lambda \in \Lambda(m,n)$, there exists an n-dimensional subspace

$$S_\lambda \equiv \operatorname{span}\{e_{\lambda(1)},\dots,e_{\lambda(n)}\} \subset \mathbb{R}^m$$

such that P_λ is the projection of $\mathbb{R}^m$ onto S_λ. ▮

THEOREM 4 BINET–CAUCHY FORMULA
Assume $n \leq m$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear. Then

$$[\![L]\!]^2 = \sum_{\lambda \in \Lambda(m,n)} (\det(P_\lambda \circ L))^2.$$

REMARK

(i) *Thus to calculate $[\![L]\!]^2$, we compute the sums of the squares of the determinants of each $(n \times n)$-submatrix of the $(m \times n)$-matrix representing L (with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$).*

(ii) In view of Lemma 1 in Section 3.3.1, this is a kind of higher dimensional version of the Pythagorean Theorem. ▮

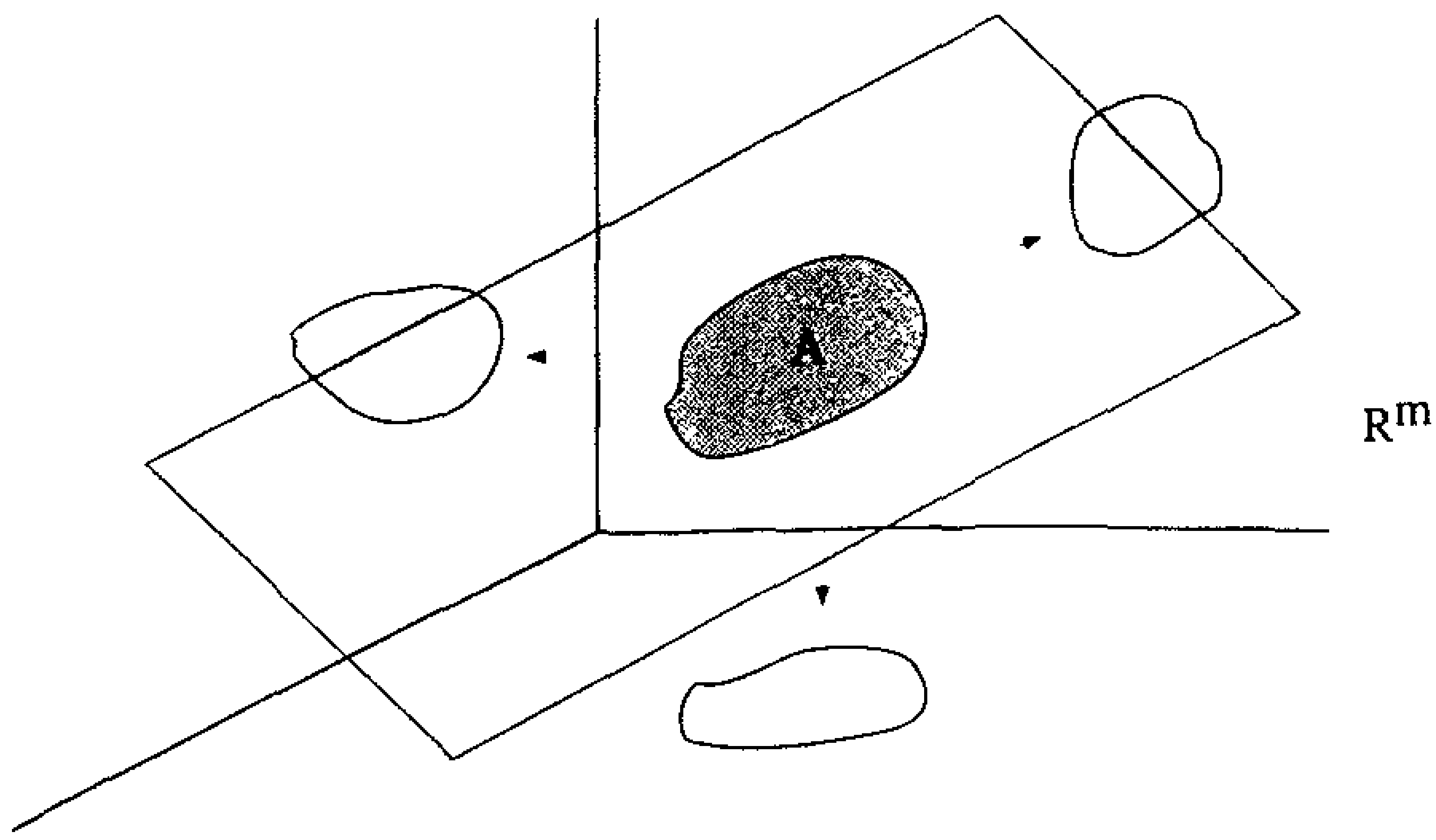

FIGURE 3.1
The square of the $\mathcal{H}^n$-measure of A equals the sum of the squares of the $\mathcal{H}^n$-measure of the projections of A onto the coordinate planes.

PROOF

1. Identifying linear maps with their matrices with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, we write

$$L = ((l_{ij}))_{m\times n},\ A = L^\star \circ L = ((a_{ij}))_{n\times n};$$

so that

$$a_{ij} = \sum_{k=1}^{m} l_{ki} l_{kj} \qquad (i,j = 1,\dots,n).$$

2. Then

$$[\![L]\!]^2 = \det A = \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)},$$

Σ denoting the set of all permutations of $\{1,\dots,n\}$. Thus

$$\begin{aligned}[\![L]\!]^2 &= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \sum_{k=1}^{m} l_{ki} l_{k\sigma(i)} \\ &= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\varphi\in\Phi} \prod_{i=1}^{n} l_{\varphi(i)i} l_{\varphi(i)\sigma(i)},\end{aligned}$$

Φ denoting the set of all one-to-one mappings of $\{1,\dots,n\}$ into $\{1,\dots,m\}$.

3. For each $\varphi \in \Phi$, we can uniquely write $\varphi = \lambda \circ \theta$, where $\theta \in \Sigma$ and $\lambda \in \Lambda(m,n)$. Consequently,

$$
\begin{aligned}
[\![L]\!]^2 &= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \prod_{i=1}^{n} l_{\lambda\circ\theta(i),i}\, l_{\lambda\circ\theta(i),\sigma(i)} \\
&= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \prod_{i=1}^{n} l_{\lambda(i),\theta^{-1}(i)}\, l_{\lambda(i),\sigma\circ\theta^{-1}(i)} \\
&= \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)}\, l_{\lambda(i),\sigma\circ\theta(i)} \\
&= \sum_{\lambda\in\Lambda(m,n)} \sum_{\rho\in\Sigma} \sum_{\theta\in\Sigma} \operatorname{sgn}(\theta)\operatorname{sgn}(\rho) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)}\, l_{\lambda(i),\rho(i)}
\end{aligned}
$$

(where we set $\rho = \sigma \circ \theta$)

$$
\begin{aligned}
&= \sum_{\lambda\in\Lambda(m,n)} \left(\sum_{\theta\in\Sigma} \operatorname{sgn}(\theta) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)} \right)^2 \\
&= \sum_{\lambda\in\Lambda(m,n)} (\det(P_\lambda \circ L))^2. \quad \blacksquare
\end{aligned}
$$

3.2.2 Jacobians

Now let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz. By Rademacher's Theorem, f is differentiable $\mathcal{L}^n$ a.e., and therefore $Df(x)$ exists and can be regarded as a linear mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ for $\mathcal{L}^n$ a.e. $x \in \mathbb{R}^n$.

NOTATION If $f : \mathbb{R}^n \to \mathbb{R}^m$, $f = (f^1, \ldots, f^m)$, we write the gradient matrix

$$
Df = \begin{bmatrix} \dfrac{\partial f^1}{\partial x_1} & \cdots & \dfrac{\partial f^1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f^m}{\partial x_1} & \cdots & \dfrac{\partial f^m}{\partial x_n} \end{bmatrix}_{m\times n}
$$

DEFINITION *The* ***Jacobian*** *of f is*

$$
Jf(x) \equiv [\![Df(x)]\!] \qquad (\mathcal{L}^n \text{ a.e. } x).
$$

3.3 The Area Formula

Throughout this section, we assume

$$n \le m.$$

3.3.1 Preliminaries

LEMMA 1
Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, $n \le m$. Then

$$\mathcal{H}^n(L(A)) = [\![L]\!]\mathcal{L}^n(A)$$

for all $A \subset \mathbb{R}^n$.

PROOF

1. Write $L = O \circ S$ as in Section 3.1; $[\![L]\!] = |\det S|$.

2. If $[\![L]\!] = 0$, then $\dim S(\mathbb{R}^n) \le n-1$ and so $\dim L(\mathbb{R}^n) \le n-1$. Consequently, $\mathcal{H}^n(L(\mathbb{R}^n)) = 0$.

3. If $[\![L]\!] > 0$, then

$$\begin{aligned}\frac{\mathcal{H}^n(L(B(x,r))}{\mathcal{L}^n(B(x,r))} &= \frac{\mathcal{L}^n(O^\star \circ L(B(x,r))}{\mathcal{L}^n(B(x,r))} = \frac{\mathcal{L}^n(O^\star \circ O \circ S(B(x,r))}{\mathcal{L}^n(B(x,r))}\\ &= \frac{\mathcal{L}^n(S(B(x,r))}{\mathcal{L}^n(B(x,r))} = \frac{\mathcal{L}^n(S(B(0,1))}{\alpha(n)}\\ &= |\det S| = [\![L]\!].\end{aligned}$$

4. Define $\nu(A) \equiv \mathcal{H}^n(L(A))$ for all $A \subset \mathbb{R}^n$. Then ν is a Radon measure, $\nu \ll \mathcal{L}^n$, and

$$D_{\mathcal{L}^n}\nu(x) = \lim_{r\to 0} \frac{\nu(B(x,r))}{\mathcal{L}^n(B(x,r))} = [\![L]\!].$$

Thus for all Borel sets $B \subset \mathbb{R}^n$, Theorem 2 in Section 1.6.2 implies

$$\mathcal{H}^n(L(B)) = [\![L]\!]\mathcal{L}^n(B).$$

Since ν and $\mathcal{L}^n$ are Radon measures, the same formula holds for all sets $A \subset \mathbb{R}^n$. ∎

Henceforth we assume $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz.

LEMMA 2
Let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$-measurable. Then

(i) *$f(A)$ is $\mathcal{H}^n$-measurable,*

(ii) *the mapping $y \mapsto \mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$-measurable on $\mathbb{R}^m$, and*

(iii) $\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n \leq (\text{Lip}\,(f))^n \mathcal{L}^n(A)$.

REMARK The mapping $y \mapsto \mathcal{H}^0(A \cap f^{-1}\{y\})$ is called the *multiplicity function*. ∎

PROOF

1. We may assume with no loss of generality that A is bounded.

2. By Theorem 5 in Section 1.1.1, there exist compact sets $K_i \subset A$ such that

$$\mathcal{L}^n(K_i) \geq \mathcal{L}^n(A) - \frac{1}{i} \qquad (i = 1, 2, \ldots).$$

As $\mathcal{L}^n(A) < \infty$ and A is $\mathcal{L}^n$-measurable, $\mathcal{L}^n(A - K_i) < 1/i$. Since f is continuous, $f(K_i)$ is compact and thus $\mathcal{H}^n$-measurable. Hence $f\left(\cup_{i=1}^\infty K_i\right) = \cup_{i=1}^\infty f(K_i)$ is $\mathcal{H}^n$-measurable. Furthermore,

$$\begin{aligned}\mathcal{H}^n\left(f(A) - f\left(\bigcup_{i=1}^\infty K_i\right)\right) &\leq \mathcal{H}^n\left(f\left(A - \bigcup_{i=1}^\infty K_i\right)\right)\\ &\leq (\text{Lip}\,(f))^n \mathcal{L}^n\left(A - \bigcup_{i=1}^\infty K_i\right) = 0.\end{aligned}$$

Thus $f(A)$ is $\mathcal{H}^n$-measurable: this proves (i).

3. Let

$$\begin{aligned}\mathcal{B}_k \equiv \Big\{Q \mid Q &= (a_1, b_1] \times \cdots \times (a_n, b_n],\\ a_i &= \frac{c_i}{k}\,,\ b_i = \frac{c_i + 1}{k}\,,\ c_i \text{ integers, } i = 1, 2, \ldots, n\Big\},\end{aligned}$$

and note

$$\mathbb{R}^n = \bigcup_{Q \in \mathcal{B}_k} Q.$$

Now

$$g_k \equiv \sum_{Q \in \mathcal{B}_k} \chi_{f(A \cap Q)} \text{ is } \mathcal{H}^n\text{-measurable by (i),}$$

and

$$g_k(y) = \text{number of cubes } Q \in \mathcal{B}_k \text{ such that } f^{-1}\{y\} \cap (A \cap Q) \neq \emptyset.$$

Thus

$$g_k(y) \uparrow \mathcal{H}^0(A \cap f^{-1}\{y\}) \qquad \text{as } k \to \infty$$

for each $y \in \mathbb{R}^m$, and so $y \mapsto \mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$-measurable.

4. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n &= \lim_{k\to\infty} \int_{\mathbb{R}^m} g_k \, d\mathcal{H}^n \\ &= \lim_{k\to\infty} \sum_{Q\in\mathcal{B}_k} \mathcal{H}^n(f(A\cap Q)) \\ &\le \limsup_{k\to\infty} \sum_{Q\in\mathcal{B}_k} (\mathrm{Lip}\,(f))^n \mathcal{L}^n(A\cap Q) \\ &= (\mathrm{Lip}\,(f))^n \mathcal{L}^n(A). \quad \blacksquare\end{aligned}$$

LEMMA 3
Let $t > 1$ and $B \equiv \{x \mid Df(x) \text{ exists}, Jf(x) > 0\}$. Then there is a countable collection $\{E_k\}_{k=1}^\infty$ of Borel subsets of $\mathbb{R}^n$ such that

(i) $B = \cup_{k=1}^\infty E_k$;

(ii) $f\,|_{E_k}$ *is one-to-one* $(k = 1, 2, \ldots)$; *and*

(iii) *for each $k = 1, 2, \ldots$, there exists a symmetric automorphism $T_k : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$\mathrm{Lip}\,((f\,|_{E_k}) \circ T_k^{-1}) \le t, \qquad \mathrm{Lip}\,(T_k \circ (f\,|_{E_k})^{-1}) \le t,$$

$$t^{-n}|\det T_k| \le Jf\,|_{E_k} \le t^n |\det T_k|.$$

PROOF
1. Fix $\epsilon > 0$ so that

$$\frac{1}{t} + \epsilon < 1 < t - \epsilon.$$

Let C be a countable dense subset of B and let S be a countable dense subset of symmetric automorphisms of $\mathbb{R}^n$.

2. Then, for each $c \in C$, $T \in$ S, and $i = 1, 2, \ldots$, define $E(c, T, i)$ to be the set of all $b \in B \cap B(c, 1/i)$ satisfying

$$\left(\frac{1}{t} + \epsilon\right)|Tv| \le |Df(b)v| \le (t - \epsilon)|Tv| \qquad (\star)$$

for all $v \in \mathbb{R}^n$ and

$$|f(a) - f(b) - Df(b)\cdot(a - b)| \le \epsilon|T(a - b)| \qquad (\star\star)$$

for all $a \in B(b, 2/i)$. Note that $E(c, T, i)$ is a Borel set since Df is Borel measurable. From ($\star$) and ($\star\star$) follows the estimate

$$\frac{1}{t}|T(a-b)| \leq |f(a)-f(b)| \leq t|T(a-b)| \qquad (\star\star\star)$$

for $b \in E(c,T,i)$, $a \in B(b,2/i)$.

3. Claim: If $b \in E(c,T,i)$, then

$$\left(\frac{1}{t}+\epsilon\right)^n |\det T| \leq Jf(b) \leq (t-\epsilon)^n |\det T|.$$

Proof of Claim: Write $Df(b) = L = O \circ S$, as above;

$$Jf(b) = [\![Df(b)]\!] = |\det S|.$$

By ($\star$),

$$\left(\frac{1}{t}+\epsilon\right)|Tv| \leq |(O \circ S)v| = |Sv| \leq (t-\epsilon)|Tv|$$

for $v \in \mathbb{R}^n$, and so

$$\left(\frac{1}{t}+\epsilon\right)|v| \leq |(S \circ T^{-1})v| \leq (t-\epsilon)|v| \qquad (v \in \mathbb{R}^n).$$

Thus

$$(S \circ T^{-1})(B(0,1)) \subset B(0,t-\epsilon);$$

whence

$$|\det(S \circ T^{-1})|\alpha(n) \leq \mathcal{L}^n(B(0,t-\epsilon)) = \alpha(n)(t-\epsilon)^n,$$

and hence

$$|\det S| \leq (t-\epsilon)^n |\det T|.$$

The proof of the other inequality is similar.

4. Relabel the countable collection $\{E(c,T,i) \mid c \in C, T \in \mathbf{S}, i = 1,2,\dots\}$ as $\{E_k\}_{k=1}^{\infty}$. Select any $b \in B$, write $Df(b) = O \circ S$ as above, and choose $T \in \mathbf{S}$ such that

$$\operatorname{Lip}(T \circ S^{-1}) \leq \left(\frac{1}{t}+\epsilon\right)^{-1}, \qquad \operatorname{Lip}(S \circ T^{-1}) \leq t-\epsilon.$$

Now select $i \in \{1,2,\dots\}$ and $c \in C$ so that $|b-c| < 1/i$,

$$|f(a)-f(b)-Df(b)\cdot(a-b)| \leq \frac{\epsilon}{\operatorname{Lip}(T^{-1})}|a-b| \leq \epsilon|T(a-b)|$$

for all $a \in B(b,2/i)$. Then $b \in E(c,T,i)$. As this conclusion holds for all $b \in B$, statement (i) is proved.

5. Next choose any set E_k, which is of the form $E(c, T, i)$ for some $c \in C$, $T \in \mathrm{S}$, $i = 1, 2, \ldots$. Let $T_k = T$. According to $(\star\star\star)$,

$$\frac{1}{t}|T_k(a-b)| \leq |f(a) - f(b)| \leq t|T_k(a-b)|$$

for all $b \in E_k$, $a \in B(b, 2/i)$. As $E_k \subset B(c, 1/i) \subset B(b, 2/i)$, we thus have

$$\frac{1}{t}|T_k(a-b)| \leq |f(a) - f(b)| \leq t|T_k(a-b)| \qquad (\star\star\star\star)$$

for all a, $b \in E_k$; hence $f\,|_{E_k}$ is one-to-one.

6. Finally, notice $(\star\star\star\star)$ implies

$$\mathrm{Lip}\,((f\,|_{E_k}) \circ T_k^{-1}) \leq t, \qquad \mathrm{Lip}\,(T_k \circ (f\,|_{E_k})^{-1}) \leq t,$$

whereas the claim provides the estimate

$$t^{-n}|\det T_k| \leq Jf\,|_{E_k} \leq t^n|\det T_k|.$$

Assertion (iii) is proved. ∎

3.3.2 Proof of the Area Formula

THEOREM 1 AREA FORMULA
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \leq m$. Then for each $\mathcal{L}^n$-measurable subset $A \subset \mathbb{R}^n$,

$$\int_A Jf\,dx = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\,d\mathcal{H}^n(y).$$

PROOF

1. In view of Rademacher's Theorem, we may as well assume $Df(x)$ and $Jf(x)$ exist for all $x \in A$. We may also suppose $\mathcal{L}^n(A) < \infty$.

2. Case 1. $A \subset \{Jf > 0\}$. Fix $t > 1$ and choose Borel sets $\{E_k\}_{k=1}^\infty$ as in Lemma 3. We may assume the sets $\{E_k\}_{k=1}^\infty$ are disjoint. Define $\mathcal{B}_k$ as in the proof of Lemma 2. Set

$$F_j^i = E_j \cap Q_i \cap A \qquad (Q_i \in \mathcal{B}_k, j = 1, 2, \ldots).$$

Then the sets F_j^i are disjoint and $A = \cup_{i,j=1}^\infty F_j^i$.

3. Claim #1:

$$\lim_{k\to\infty} \sum_{i,j=1}^\infty \mathcal{H}^n(f(F_j^i)) = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\,d\mathcal{H}^n.$$

Proof of Claim #1: Let

$$g_k \equiv \sum_{i,j=1}^\infty \chi_{f(F_j^i)}$$

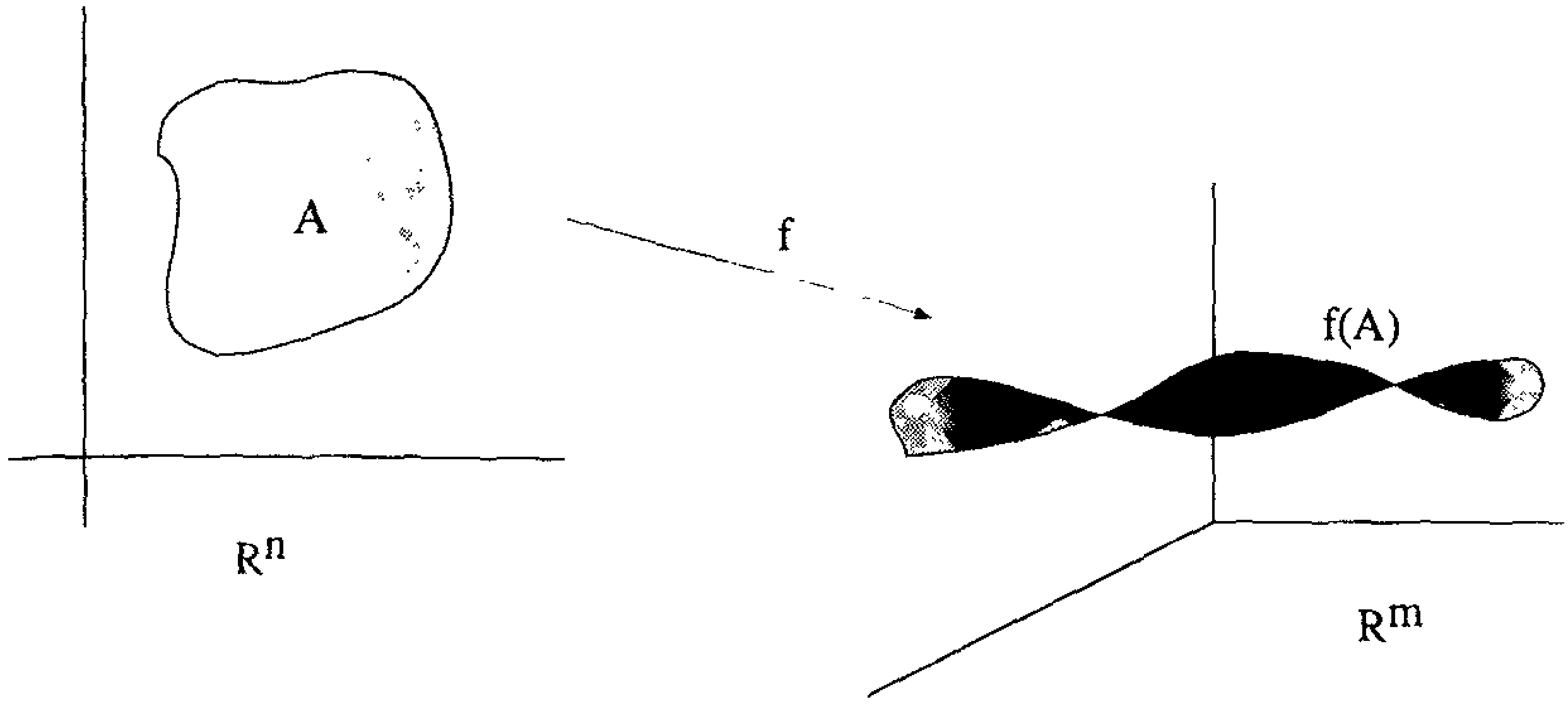

FIGURE 3.2
The Area Formula.

so that $g_k(y)$ is the number of the sets $\{F_j^i\}$ such that $F_j^i \cap f^{-1}\{y\} \neq \emptyset$. Then $g_k(y) \uparrow \mathcal{H}^0(A \cap f^{-1}\{y\})$ as $k \to \infty$. Apply the Monotone Convergence Theorem.

4. Note

$$\mathcal{H}^n(f(F_j^i)) = \mathcal{H}^n(f\,|_{E_j} \circ T_j^{-1} \circ T_j(F_j^i)) \leq t^n \mathcal{L}^n(T_j(F_j^i))$$

and

$$\mathcal{L}^n(T_j(F_j^i)) = \mathcal{H}^n(T_j \circ (f\,|_{E_j})^{-1} \circ f(F_j^i)) \leq t^n \mathcal{H}^n(f(F_j^i))$$

by Lemma 3. Thus

$$\begin{aligned} t^{-2n}\mathcal{H}^n(f(F_j^i)) &\leq t^{-n}\mathcal{L}^n(T_j(F_j^i)) \\ &= t^{-n}|\det T_j|\mathcal{L}^n(F_j^i) \\ &\leq \int_{F_i^j} Jf\,dx \\ &\leq t^n|\det T_j|\mathcal{L}^n(F_j^i) \\ &= t^n\mathcal{L}^n(T_j(F_j^i)) \\ &\leq t^{2n}\mathcal{H}^n(f(F_j^i)), \end{aligned}$$

where we repeatedly used Lemmas 1 and 3. Now sum on i and j:

$$t^{-2n}\sum_{i,j=1}^{\infty}\mathcal{H}^n(f(F_j^i)) \leq \int_A Jf\,dx \leq t^{2n}\sum_{i,j=1}^{\infty}\mathcal{H}^n(f(F_j^i)).$$

Now let $k \to \infty$ and recall Claim #1:

$$t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n \leq \int_A Jf\, dx$$
$$\leq t^{2n} \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n.$$

Finally, send $t \to 1^+$.

5. *Case 2.* $A \subset \{Jf = 0\}$. Fix $\epsilon > 0$. We factor $f = p \circ g$, where

$$g : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n, \qquad g(x) \equiv (f(x), \epsilon x) \text{ for } x \in \mathbb{R}^n,$$

and

$$p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m, \qquad p(y,z) = y \text{ for } y \in \mathbb{R}^m,\ z \in \mathbb{R}^n.$$

6. *Claim #2*: There exists a constant C such that

$$0 < Jg(x) \leq C\epsilon$$

for $x \in A$.

Proof of Claim #2: Write $g = (f^1, \dots, f^m, \epsilon x_1, \dots, \epsilon x_n)$; then

$$Dg(x) = \begin{pmatrix} Df(x) \\ \epsilon I \end{pmatrix}_{(n+m)\times n}.$$

Since $Jf(x)^2$ equals the sum of the squares of the $(n \times n)$-subdeterminants of $Df(x)$ according to the Binet–Cauchy formula, we see

$$Jg(x)^2 = \text{ sum of squares of } (n \times n)\text{-subdeterminants of } Dg(x) \ \geq\ \epsilon^{2n} > 0.$$

Furthermore, since $|Df| \leq \text{Lip}\ (f) < \infty$, we may employ the Binet–Cauchy Formula to compute

$$Jg(x)^2 = Jf(x)^2 + \left\{ \begin{matrix} \text{sum of squares of terms each} \\ \text{involving at least one } \epsilon \end{matrix} \right\} \leq C\epsilon^2$$

for each $x \in A$.

7. Since $p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ is a projection, we can compute, using Case 1 above,

$$\begin{aligned} \mathcal{H}^n(f(A)) &\leq \mathcal{H}^n(g(A)) \\ &\leq \int_{\mathbb{R}^{n+m}} \mathcal{H}^0(A \cap g^{-1}\{y,z\})\, d\mathcal{H}^n(y,z) \\ &= \int_A Jg(x)\, dx \\ &\leq \epsilon C \mathcal{L}^n(A). \end{aligned}$$

Let $\epsilon \to 0$ to conclude $\mathcal{H}^n(f(A)) = 0$, and thus

$$\int_{\mathbb{R}^n} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n = 0,$$

since spt $\mathcal{H}^0(A \cap f^{-1}\{y\}) \subset f(A)$. But then

$$\int_{\mathbb{R}^n} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n = 0 = \int_A Jf\, dx.$$

8. In the general case, write $A = A_1 \cup A_2$ with $A_1 \subset \{Jf > 0\}$, $A_2 \subset \{Jf = 0\}$, and apply Cases 1 and 2 above. ∎

3.3.3 Change of variables formula

THEOREM 2
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \leq m$. Then for each $\mathcal{L}^n$-summable function $g : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} g(x) Jf(x)\, dx = \int_{\mathbb{R}^m} \left[\sum_{x \in f^{-1}\{y\}} g(x) \right] d\mathcal{H}^n(y).$$

REMARK Using the Area Formula, we see $f^{-1}\{y\}$ is at most countable for $\mathcal{H}^n$ a.e. $y \in \mathbb{R}^m$. ∎

PROOF
1. Case 1. $g \geq 0$. According to Theorem 7 in Section 1.1.2 we can write

$$g = \sum_{i=1}^{\infty} \frac{1}{i} \chi_{A_i}$$

for appropriate $\mathcal{L}^n$-measurable sets $\{A_i\}_{i=1}^{\infty}$. Then the Monotone Convergence Theorem implies

$$\begin{aligned}
\int_{\mathbb{R}^n} gJf\, dx &= \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^n} \chi_{A_i} Jf\, dx \\
&= \sum_{i=1}^{\infty} \frac{1}{i} \int_{A_i} Jf\, dx \\
&= \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^m} \mathcal{H}^0(A_i \cap f^{-1}\{y\})\, d\mathcal{H}^n(y)
\end{aligned}$$

$$= \int_{\mathbb{R}^m} \sum_{i=1}^{\infty} \frac{1}{i} \sum_{x \in f^{-1}\{y\}} \chi_{A_i}(x)\, d\mathcal{H}^n(y)$$

$$= \int_{\mathbb{R}^m} \sum_{x \in f^{-1}\{y\}} \sum_{i=1}^{\infty} \frac{1}{i} \chi_{A_i}(x)\, d\mathcal{H}^n(y)$$

$$= \int_{\mathbb{R}^m} \left[\sum_{x \in f^{-1}\{y\}} g(x) \right] d\mathcal{H}^n(y).$$

2. *Case 2.* g is any $\mathcal{L}^n$-summable function. Write $g = g^+ - g^-$ and apply Case 1. ∎

3.3.4 Applications

A. Length of a curve ($n = 1, m \geq 1$). Assume $f : \mathbb{R} \to \mathbb{R}^m$ is Lipschitz and one-to-one. Write

$$f = (f^1, \ldots, f^m), \qquad Df = (\dot{f}^1, \ldots, \dot{f}^m),$$

so that

$$Jf = |Df| = |\dot{f}| \qquad \left(\cdot = \frac{d}{dt} \right).$$

For $-\infty < a < b < \infty$, define the curve

$$C \equiv f([a,b]) \subset \mathbb{R}^m.$$

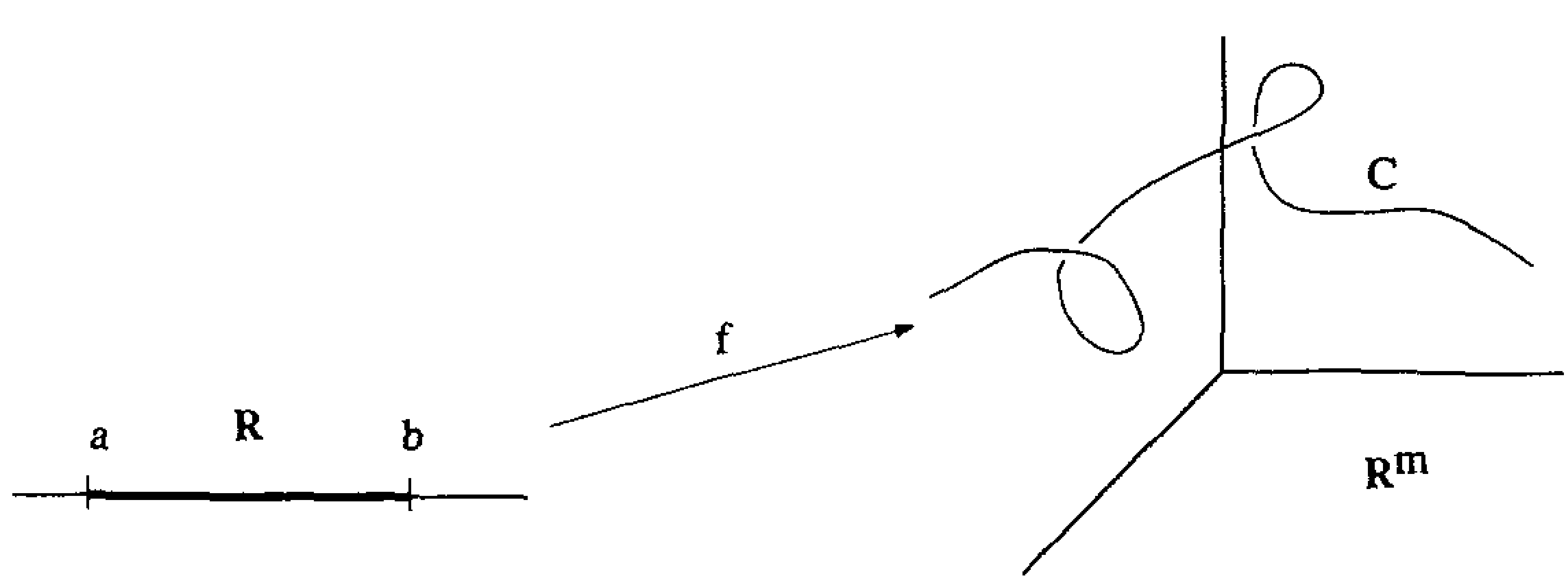

FIGURE 3.3
Length of a curve.

Then

$$\mathcal{H}^1(C) = \text{“length” of } C = \int_a^b |\dot{f}|\, dt.$$

B. Surface area of a graph ($n \geq 1, m = n + 1$). Assume $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz and define $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$f(x) \equiv (x, g(x)).$$

Then

$$Df = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \\ \dfrac{\partial g}{\partial x_1} & \cdots & \dfrac{\partial g}{\partial x_n} \end{bmatrix}_{(n+1)\times n};$$

so that

$$\begin{aligned} (Jf)^2 &= \text{ sum of squares of } (n \times n)\text{-subdeterminants} \\ &= 1 + |Dg|^2. \end{aligned}$$

For each open set $U \subset \mathbb{R}^n$, define the graph of g over U,

$$G = G(g; U) \equiv \{(x, g(x)) \mid x \in U\} \subset \mathbb{R}^{n+1}.$$

Then

$$\mathcal{H}^n(G) = \text{“surface area” of } G = \int_U (1 + |Dg|^2)^{\frac{1}{2}}\, dx.$$

C. Surface area of a parametric hypersurface ($n \geq 1, m = n + 1$). Suppose $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ is Lipschitz and one-to-one. Write

$$f = (f^1, \ldots, f^{n+1}),$$

$$Df = \begin{bmatrix} \dfrac{\partial f^1}{\partial x_1} & \cdots & \dfrac{\partial f^1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f^{n+1}}{\partial x_1} & \cdots & \dfrac{\partial f^{n+1}}{\partial x_n} \end{bmatrix}_{(n+1)\times n}$$

so that

$$\begin{aligned} (Jf)^2 &= \text{ sum of squares of } (n \times n)\text{-subdeterminants} \\ &= \sum_{k=1}^{n+1} \left[\frac{\partial(f^1, \ldots, f^{k-1}, f^{k+1}, \ldots, f^{n+1})}{\partial(x_1, \ldots, x_n)} \right]^2. \end{aligned}$$

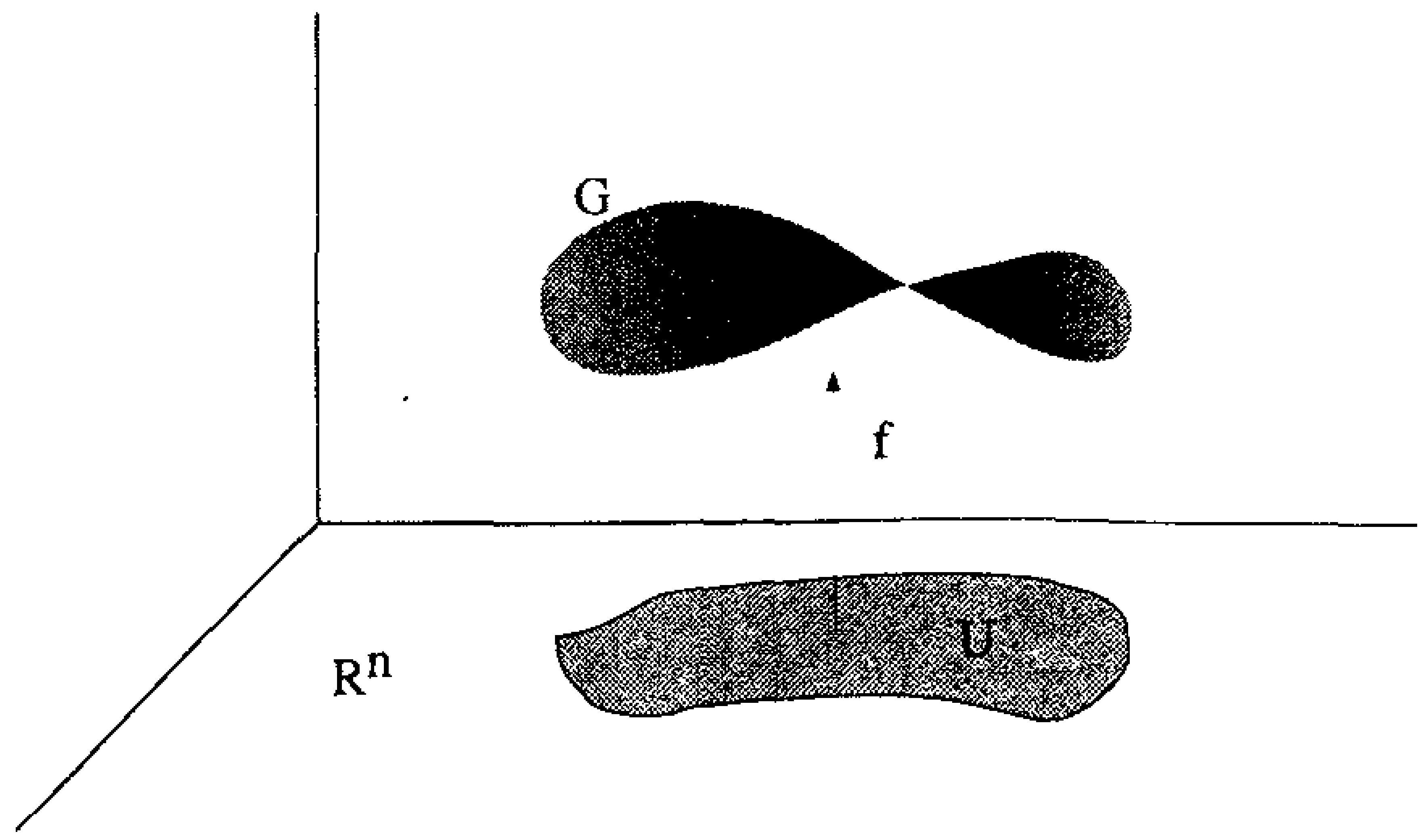

FIGURE 3.4
Surface area of a graph.

For each open set $U \subset \mathbb{R}^n$, write

$$S \equiv f(U) \subset \mathbb{R}^{n+1}.$$

Then

$$\begin{aligned}\mathcal{H}^n(S) &= \text{“surface area” of } S \\ &= \int_U \left(\sum_{k=1}^{n+1} \left[\frac{\partial(f^1, \ldots, f^{k-1}, f^{k+1}, \ldots, f^{n+1})}{\partial(x_1, \ldots, x_n)} \right]^2 \right)^{\frac{1}{2}} dx.\end{aligned}$$

D. Submanifolds. Let $M \subset \mathbb{R}^m$ be a Lipschitz, n-dimensional embedded submanifold. Suppose that $U \subset \mathbb{R}^n$ and $f : U \to M$ is a chart for M. Let $A \subset f(U)$, A Borel, $B \equiv f^{-1}(A)$. Define

$$g_{ij} \equiv \frac{\partial f}{\partial x_i} \cdot \frac{\partial f}{\partial x_j} \qquad (1 \le i, j \le n),$$
$$g \equiv \det((g_{ij})).$$

Then

$$(Df)^* \circ Df = ((g_{ij}))_{n \times n},$$

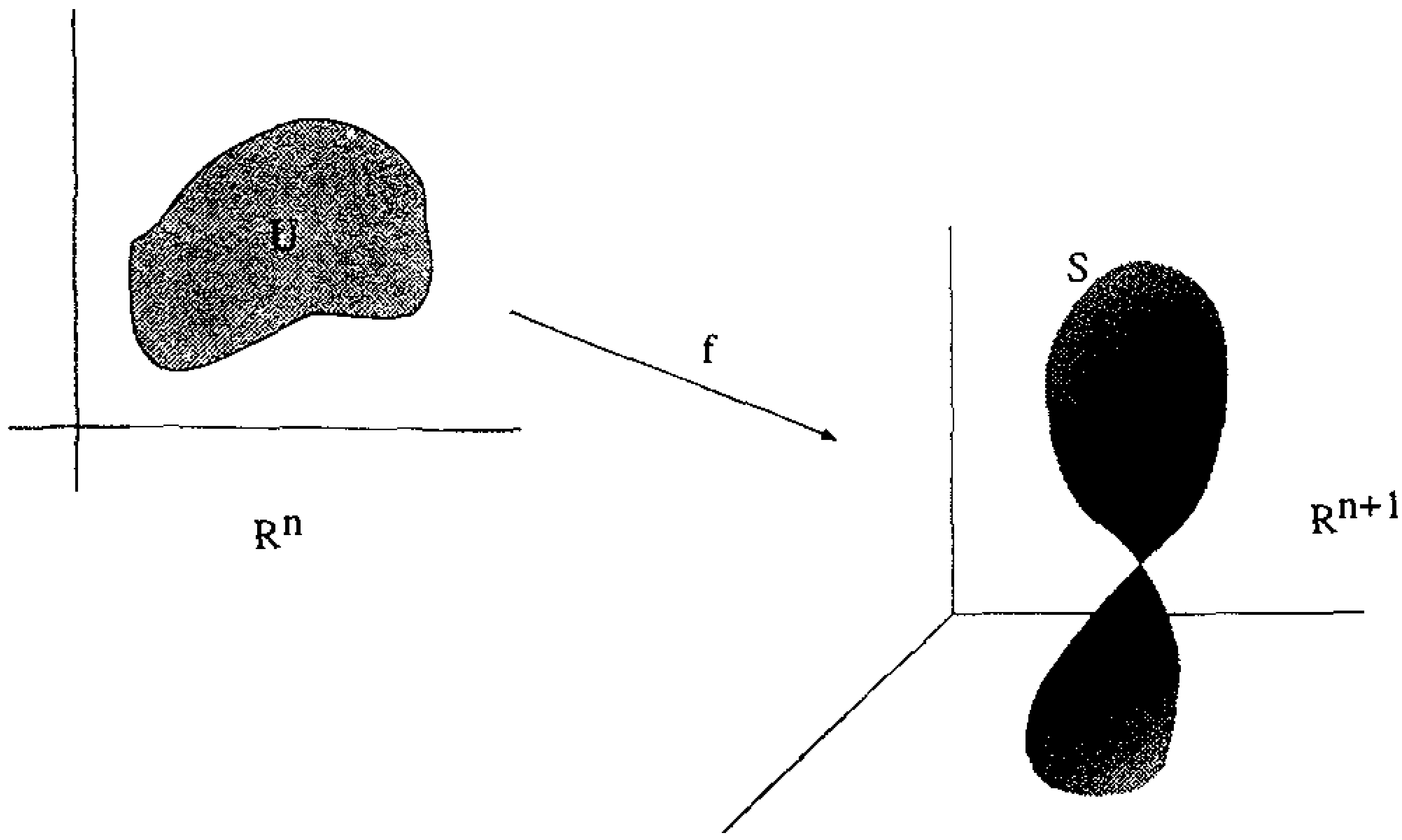

FIGURE 3.5
Surface area of a parametric hypersurface.

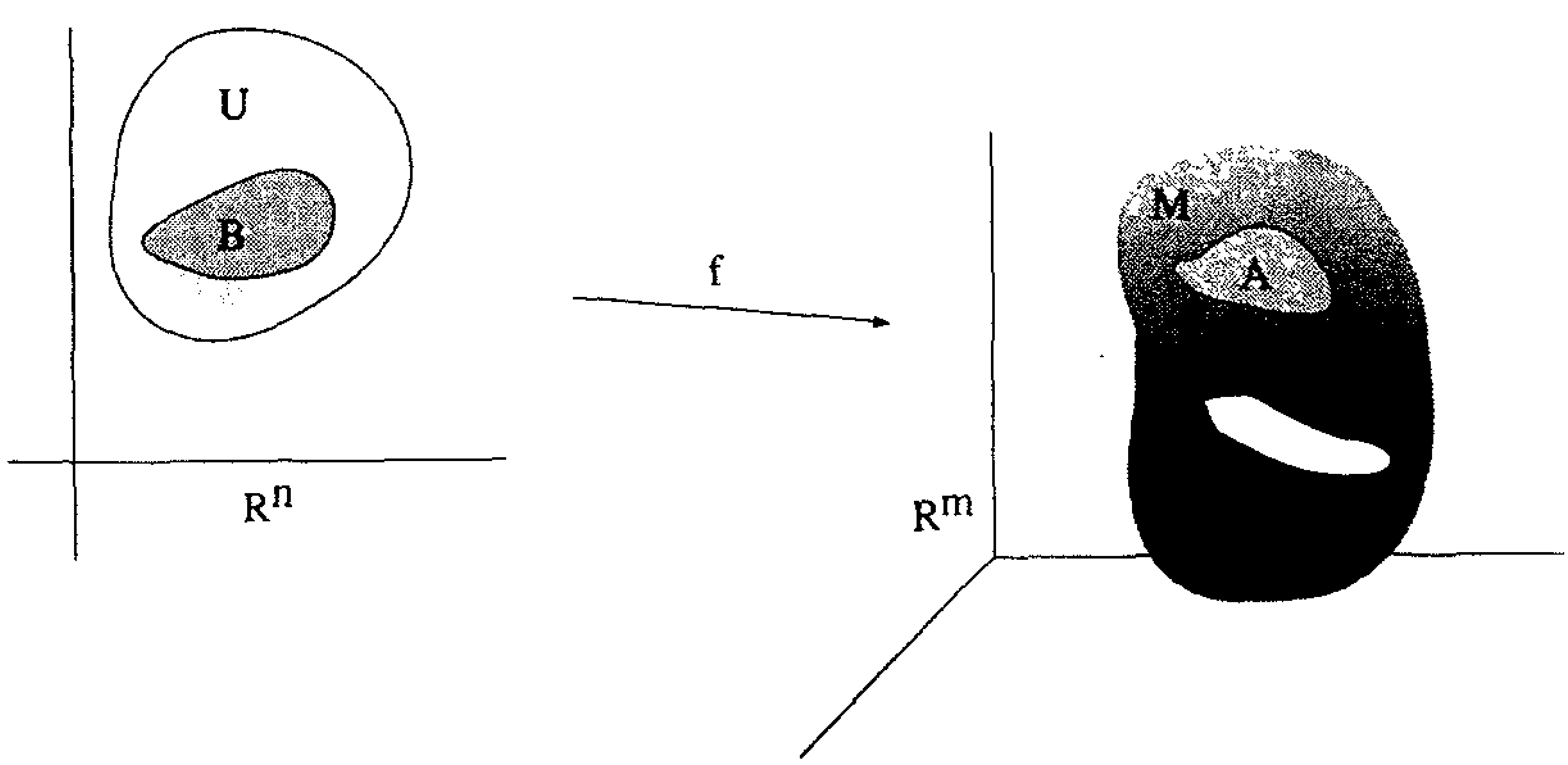

FIGURE 3.6
Volume of a submanifold.

and so

$$Jf = g^{\frac{1}{2}}.$$

Thus

$$\mathcal{H}^n(A) = \text{"volume" of } A \text{ in } M = \int_B g^{\frac{1}{2}}\, dx$$

3.4 The Coarea Formula

Throughout this section we assume

$$n \geq m.$$

3.4.1 Preliminaries

LEMMA 1

Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, $n \geq m$, and $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable. Then

(i) *the mapping $y \mapsto \mathcal{H}^{n-m}(A \cap L^{-1}\{y\})$ is $\mathcal{L}^m$-measurable and*

(ii) $\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}\{y\})\, dy = [\![L]\!]\mathcal{L}^n(A)$.

PROOF

1. *Case 1.* $\dim L(\mathbb{R}^n) < m$.

Then $A \cap L^{-1}\{y\} = \emptyset$ and consequently $\mathcal{H}^{n-m}(A \cap L^{-1}\{y\}) = 0$ for $\mathcal{L}^m$ a.e. $y \in \mathbb{R}^n$. Also, if we write $L = S \circ O^\star$ as in the Polar Decomposition Theorem (Section 3.2.1), we have $L(\mathbb{R}^n) = S(\mathbb{R}^m)$. Thus $\dim S(\mathbb{R}^m) < m$ and hence $[\![L]\!] = |\det S| = 0$.

2. *Case 2.* $L = P =$ orthogonal projection of $\mathbb{R}^n$ onto $\mathbb{R}^m$.

Then for each $y \in \mathbb{R}^m$, $P^{-1}\{y\}$ is an $(n-m)$-dimensional affine subspace of $\mathbb{R}^n$, a translate of $P^{-1}\{0\}$. By Fubini's Theorem,

$$y \mapsto \mathcal{H}^{n-m}(A \cap P^{-1}\{y\}) \quad \text{is } \mathcal{L}^m\text{-measurable}$$

and

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap P^{-1}\{y\})\, dy = \mathcal{L}^n(A). \qquad (\star)$$

3. *Case 3.* $L : \mathbb{R}^n \to \mathbb{R}^m$, $\dim L(\mathbb{R}^n) = m$.

Using the Polar Decomposition Theorem, we can write

$$L = S \circ O^\star$$

where

$$S : \mathbb{R}^m \to \mathbb{R}^m \text{ is symmetric,}$$

$$O : \mathbb{R}^m \to \mathbb{R}^n \text{ is orthogonal,}$$

$$[\![L]\!] = |\det S| > 0.$$

4. *Claim*: $O^\star = P \circ Q$, where P is the orthogonal projection of $\mathbb{R}^n$ onto $\mathbb{R}^m$ and $Q : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal.

Proof of Claim: Let Q be any orthogonal map of $\mathbb{R}^n$ onto $\mathbb{R}^n$ such that

$$Q^*(x_1, \ldots, x_m, 0, \ldots, 0) = O(x_1, \ldots, x_m)$$

for all $x \in \mathbb{R}^m$. Note

$$P^*(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0) \in \mathbb{R}^n$$

for all $x \in \mathbb{R}^m$. Thus $O = Q^* \circ P^*$ and hence $O^* = P \circ Q$.

5. $L^{-1}\{0\}$ is an $(n-m)$-dimensional subspace of $\mathbb{R}^n$ and $L^{-1}\{y\}$ is a translate of $L^{-1}\{0\}$ for each $y \in \mathbb{R}^m$. Thus by Fubini's Theorem, $y \mapsto \mathcal{H}^{n-m}(A \cap L^{-1}\{y\})$ is $\mathcal{L}^m$-measurable, and we may calculate

$$\begin{aligned}
\mathcal{L}^n(A) &= \mathcal{L}^n(Q(A)) \\
&= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(Q(A) \cap P^{-1}\{y\})\, dy \qquad \text{by } (\star) \\
&= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1}\{y\})\, dy.
\end{aligned}$$

Now set $z = Sy$ to compute using Theorem 2 in Section 3.3.3

$$|\det S|\, \mathcal{L}^n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap Q^{-1} \circ P^{-1} \circ S^{-1}\{z\})\, dz.$$

But $L = S \circ O^* = S \circ P \circ Q$, and so

$$[\![L]\!]\mathcal{L}^n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap L^{-1}\{z\})\, dz. \quad \blacksquare$$

Henceforth we assume $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz.

LEMMA 2
Let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$-measurable, $n \geq m$. Then

(i) *$f(A)$ is $\mathcal{L}^m$-measurable,*

(ii) *$A \cap f^{-1}\{y\}$ is $\mathcal{H}^{n-m}$ measurable for $\mathcal{L}^m$ a.e. y,*

(iii) *the mapping $y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$ is $\mathcal{L}^m$-measurable, and*

(iv) *$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy \leq (\alpha(n-m)\alpha(m))/\alpha(n)(\mathrm{Lip}\, f)^m \mathcal{L}^n(A)$.*

PROOF

1. Statement (i) is proved exactly like the corresponding statement of Lemma 2 in Section 3.3.1.

2. For each $j = 1, 2, \ldots$, there exist closed balls $\{B_i^j\}_{i=1}^\infty$ such that

$$A \subset \bigcup_{i=1}^\infty B_i^j, \qquad \mathrm{diam}\, B_i^j \leq \frac{1}{j},$$

and

$$\sum_{i=1}^{\infty} \mathcal{L}^n(B_i^j) \leq \mathcal{L}^n(A) + \frac{1}{j}.$$

Define

$$g_i^j \equiv \alpha(n-m)\left(\frac{\operatorname{diam} B_i^j}{2}\right)^{n-m} \chi_{f(B_i^j)}.$$

By (i), g_i^j is $\mathcal{L}^m$-measurable. Note also for all $y \in \mathbb{R}^m$,

$$\mathcal{H}_{1/j}^{n-m}(A \cap f^{-1}\{y\}) \leq \sum_{i=1}^{\infty} g_i^j(y).$$

Thus, using Fatou's Lemma and the Isodiametric Inequality (Section 2.2), we compute

$$\begin{aligned}
&\int_{\mathbb{R}^m}^{*} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy \\
&\quad = \int_{\mathbb{R}^m}^{*} \lim_{j\to\infty} \mathcal{H}_{1/j}^{n-m}(A \cap f^{-1}\{y\})\, dy \\
&\quad \leq \int_{\mathbb{R}^m} \liminf_{j\to\infty} \sum_{i=1}^{\infty} g_i^j\, dy \\
&\quad \leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \int_{\mathbb{R}^m} g_i^j\, dy \\
&\quad = \liminf_{j\to\infty} \sum_{i=1}^{\infty} \alpha(n-m)\left(\frac{\operatorname{diam} B_i^j}{2}\right)^{n-m} \mathcal{L}^m(f(B_i^j)) \\
&\quad \leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \alpha(n-m)\left(\frac{\operatorname{diam} B_i^j}{2}\right)^{n-m} \alpha(m)\left(\frac{\operatorname{diam} f(B_i^j)}{2}\right)^{m} \\
&\quad \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}(\operatorname{Lip} f)^m \liminf_{j\to\infty} \sum_{i=1}^{\infty} \mathcal{L}^n(B_i^j) \\
&\quad \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}(\operatorname{Lip} f)^m \mathcal{L}^n(A).
\end{aligned}$$

Thus

$$\int_{\mathbb{R}^m}^{*} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}(\operatorname{Lip} f)^m \mathcal{L}^n(A). \qquad (\star)$$

This will prove (iv) once we establish (iii).

3. Case 1: A compact.

Fix $t \geq 0$, and for each positive integer i, let U_i consist of all points $y \in \mathbb{R}^m$ for which there exist finitely many open sets $S_1, \dots, S_l$ such that

$$\begin{cases} A \cap f^{-1}\{y\} \subset \bigcup_{j=1}^{l} S_j, \\ \operatorname{diam} S_j \leq \dfrac{1}{i} \qquad (j = 1, 2, \dots, l), \\ \displaystyle\sum_{j=1}^{l} \alpha(n-m) \left(\frac{\operatorname{diam} S_j}{2}\right)^{n-m} \leq t + \frac{1}{i}\,. \end{cases}$$

4. Claim #1: U_i is open.

Proof of Claim #1: Assume $y \in U_i$, $A \cap f^{-1}\{y\} \subset \bigcup_{j=1}^{l} S_j$, as above. Then, since f is continuous and A is compact,

$$A \cap f^{-1}\{z\} \subset \bigcup_{j=1}^{l} S_j$$

for all z sufficiently close to y.

5. Claim #2.

$$\{y \mid \mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) \leq t\} = \bigcap_{i=1}^{\infty} U_i$$

and hence is a Borel set.

Proof of Claim #2: If $\mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) \leq t$, then for each $\delta > 0$,

$$\mathcal{H}_\delta^{n-m}(A \cap f^{-1}\{y\}) \leq t.$$

Given i, choose $\delta \in (0, 1/i)$. Then there exist sets $\{S_j\}_{j=1}^{\infty}$ such that

$$\begin{cases} A \cap f^{-1}\{y\} \subset \displaystyle\bigcup_{j=1}^{\infty} S_j, \\ \operatorname{diam} S_j \leq \delta < \dfrac{1}{i}, \\ \displaystyle\sum_{j=1}^{\infty} \alpha(n-m) \left(\frac{\operatorname{diam} S_j}{2}\right)^{n-m} < t + \frac{1}{i}\,. \end{cases}$$

We may assume the S_j are open. Since $A \cap f^{-1}\{y\}$ is compact, a finite subcollection $\{S_1, \dots, S_l\}$ covers $A \cap f^{-1}\{y\}$, and hence $y \in U_i$. Thus

$$\{y \mid \mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) \leq t\} \subset \bigcap_{i=1}^{\infty} U_i.$$

On the other hand, if $y \in \cap_{i=1}^{\infty} U_i$, then for each i,

$$\mathcal{H}^{n-m}_{1/i}(A \cap f^{-1}\{y\}) \leq t + \frac{1}{i} ,$$

and so

$$\mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) \leq t.$$

Thus

$$\bigcap_{i=1}^{\infty} U_i \subset \{y \mid \mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) \leq t\}.$$

6. According to Claim #2, for compact A the mapping

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$$

is a Borel function.

7. Case 2: A open.

There exist compact sets $K_1 \subset K_2 \subset \cdots \subset A$ such that

$$A = \bigcup_{i=1}^{\infty} K_i.$$

Thus, for each $y \in \mathbb{R}^m$,

$$\mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) = \lim_{i \to \infty} \mathcal{H}^{n-m}(K_i \cap f^{-1}\{y\}),$$

and hence the mapping

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$$

is Borel measurable.

8. Case 3. $\mathcal{L}^n(A) < \infty$.

There exist open sets $V_1 \supset V_2 \supset \cdots \supset A$ such that

$$\lim_{i \to \infty} \mathcal{L}^n(V_i - A) = 0, \qquad \mathcal{L}^n(V_1) < \infty.$$

Now

$$\mathcal{H}^{n-m}(V_i \cap f^{-1}\{y\}) \leq \mathcal{H}^{n-m}(A \cap f^{-1}\{y\}) + \mathcal{H}^{n-m}((V_i - A) \cap f^{-1}\{y\}),$$

and thus by $(\star)$,

$$\begin{aligned} &\limsup_{i \to \infty} \int^{\star}_{\mathbb{R}^m} \left|\mathcal{H}^{n-m}(V_i \cap f^{-1}\{y\}) - \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\right| \, dy \\ &\qquad \leq \limsup_{i \to \infty} \int^{\star}_{\mathbb{R}^n} \mathcal{H}^{n-m}((V_i - A) \cap f^{-1}\{y\}) \, dy \\ &\qquad \leq \limsup_{i \to \infty} \frac{\alpha(n-m)\alpha(m)}{\alpha(n)} (\mathrm{Lip}\, f)^m \mathcal{L}^n(V_i - A) = 0. \end{aligned}$$

Consequently,

$$\mathcal{H}^{n-m}(V_i \cap f^{-1}\{y\}) \to \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$$

$\mathcal{L}^m$ a.e., and so according to Case 2,

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$$

is $\mathcal{L}^m$-measurable. In addition, we see $\mathcal{H}^{n-m}((V_i - A) \cap f^{-1}\{y\}) \to 0$ $\mathcal{L}^m$ a.e. and so $A \cap f^{-1}\{y\}$ is $\mathcal{H}^{n-m}$ measurable for $\mathcal{L}^m$ a.e. y.

9. *Case 4.* $\mathcal{L}^n(A) = \infty$.

Write A as a union of an increasing sequence of bounded $\mathcal{L}^n$-measurable sets and apply Case 3 to prove

$$A \cap f^{-1}\{y\} \text{ is } \mathcal{H}^{n-m} \text{ measurable for } \mathcal{L}^m \text{ a.e. } y,$$

and

$$y \mapsto \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$$

is $\mathcal{L}^m$-measurable.

This proves (ii) and (iii), and (iv) follows from ($\star$). ∎

REMARK A proof similar to that of (iv) shows

$$\int_{\mathbb{R}^m}^{*} \mathcal{H}^k(A \cap f^{-1}\{y\})\, d\mathcal{H}^l \le \frac{\alpha(k)\alpha(l)}{\alpha(k+l)}(\text{Lip } f)^l \mathcal{H}^{k+l}(A)$$

for each $A \subset \mathbb{R}^m$; see Federer [F, Sections 2.10.25 and 2.10.26]. ∎

LEMMA 3

Let $t > 1$, assume $h : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz, and set

$$B = \{x \mid Dh(x) \text{ exists, } Jh(x) > 0\}.$$

Then there exists a countable collection $\{D_k\}_{k=1}^{\infty}$ of Borel subsets of $\mathbb{R}^n$ such that

(i) $\mathcal{L}^n\left(B - \cup_{k=1}^{\infty} D_k\right) = 0$;

(ii) $h\,|_{D_k}$ *is one-to-one for* $k = 1, 2, \ldots$; *and*

(iii) *for each* $k = 1, 2, \ldots$, *there exists a symmetric automorphism* $S_k : \mathbb{R}^n \to \mathbb{R}^n$ *such that*

$$\text{Lip }(S_k^{-1} \circ (h\,|_{D_k})) \le t, \qquad \text{Lip }((h\,|_{D_k})^{-1} \circ S_k) \le t,$$

$$t^{-n}\,|\det S_k| \le Jh|_{D_k} \le t^n |\det S_k|.$$

PROOF

1. Apply Lemma 3 of Section 3.3.1 with h in place of f to find Borel sets $\{E_k\}_{k=1}^{\infty}$ and symmetric automorphisms $T_k : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\text{(a)} \qquad B = \bigcup_{k=1}^{\infty} E_k,$$

$$\text{(b)} \qquad h\,|_{E_k} \text{ is one–to–one,}$$

$$\text{(c)} \qquad \begin{cases} \operatorname{Lip}\left((h\,|_{E_k}) \circ T_k^{-1}\right) \le t, \ \operatorname{Lip}\left(T_k \circ (h\,|_{E_k})^{-1}\right) \le t \\ t^{-n}|\det T_k| \le Jh\,|_{E_k} \le t^n|\det T_k| \qquad (k = 1, 2, \ldots). \end{cases}$$

According to (c), $(h\,|_{E_k})^{-1}$ is Lipschitz and thus by Theorem 1 in Section 3.1.1, there exists a Lipschitz mapping $h_k : \mathbb{R}^n \to \mathbb{R}^n$ such that $h_k = (h\,|_{E_k})^{-1}$ on $h(E_k)$.

2. Claim #1: $Jh_k > 0$ $\mathcal{L}^n$ a.e. on $h(E_k)$.

Proof of Claim #1: Since $h_k \circ h(x) = x$ for $x \in E_k$, Corollary 1 in Section 3.1.2 implies

$$Dh_k(h(x)) \circ Dh(x) = I \qquad \mathcal{L}^n \text{ a.e. on } E_k,$$

and so

$$Jh_k(h(x))Jh(x) = 1 \qquad \mathcal{L}^n \text{ a.e. on } E_k.$$

In view of (c), this implies $Jh_k(h(x)) > 0$ for $\mathcal{L}^n$ a.e. $x \in E_k$, and the claim follows since h is Lipschitz.

3. Now apply Lemma 3 of Section 3.3.1 to h_k: there exist Borel sets $\{F_j^k\}_{j=1}^{\infty}$ and symmetric automorphisms $\{R_j^k\}_{j=1}^{\infty}$ such that

$$\text{(d)} \qquad \mathcal{L}^n\left(h(E_k) - \bigcup_{j=1}^{\infty} F_j^k\right) = 0;$$

$$\text{(e)} \qquad h_k\,|_{F_j^k} \text{ is one-to-one;}$$

$$\text{(f)} \qquad \begin{cases} \operatorname{Lip}\left((h_k\,|_{F_j^k}) \circ (R_j^k)^{-1}\right) \le t, \ \operatorname{Lip}\left(R_j^k \circ (h_k\,|_{F_j^k})^{-1}\right) \le t \\ t^{-n}|\det R_j^k| \le Jh_k\,|_{F_j^k} \le t^n|\det R_j^k| \ (k = 1, 2, \ldots). \end{cases}$$

Set

$$D_j^k \equiv E_k \cap h^{-1}(F_j^k), \ S_j^k \equiv (R_j^k)^{-1} \qquad (k = 1, 2, \ldots).$$

4. Claim #2: $\mathcal{L}^n\left(B - \cup_{k,j=1}^{\infty} D_j^k\right) = 0.$

Proof of Claim #2: Note

$$h_k\left(h(E_k) - \bigcup_{j=1}^{\infty} F_j^k\right) = h^{-1}\left(h(E_k) - \bigcup_{j=1}^{\infty} F_j^k\right)$$
$$= E_k - \bigcup_{j=1}^{\infty} D_j^k.$$

Thus, by (d),

$$\mathcal{L}^n\left(E_k - \bigcup_{j=1}^{\infty} D_j^k\right) = 0 \qquad (k = 1, \ldots).$$

Now recall (a).

5. Clearly (b) implies $h\,|_{D_j^k}$ is one-to-one.

6. Claim #3: For $k, j = 1, 2, \ldots$, we have

$$\text{Lip}\,((S_j^k)^{-1} \circ (h\,|_{D_j^k})) \le t, \qquad \text{Lip}\,((h\,|_{D_j^k})^{-1} \circ S_j^k) \le t$$

$$t^{-n}|\det S_j^k| \le Jh\,|_{D_j^k} \le t^n|\det S_j^k|.$$

Proof of Claim #3:

$$\text{Lip}\,((S_j^k)^{-1} \circ (h\,|_{D_j^k})) = \text{Lip}\,(R_j^k \circ (h\,|_{D_j^k}))$$
$$\le \text{Lip}\,(R_j^k \circ (h_k\,|_{F_j^k})^{-1}) \le t$$

by (f); similarly,

$$\text{Lip}\,((h\,|_{D_j^k})^{-1} \circ S_j^k) = \text{Lip}\,((h\,|_{D_j^k})^{-1} \circ (R_j^k)^{-1})$$
$$\le \text{Lip}\,((h_k\,|_{F_j^k}) \circ (R_j^k)^{-1}) \le t.$$

Furthermore, as noted above,

$$Jh_k(h(x))Jh(x) = 1 \qquad \mathcal{L}^n \text{ a.e. on } D_j^k.$$

Thus (f) implies

$$t^{-n}|\det S_j^k| = t^{-n}|\det R_j^k|^{-1} \le Jh\,|_{D_j^k} \le t^n|\det R_j^k|^{-1} = t^n|\det S_j^k|. \quad \blacksquare$$

3.4.2 Proof of the Coarea Formula

/

THEOREM 1 COAREA FORMULA
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \geq m$. Then for each $\mathcal{L}^n$-measurable set $A \subset \mathbb{R}^n$,

$$\int_A Jf\,dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\,dy.$$

REMARKS

(i) Observe that the Coarea Formula is a kind of "curvilinear" generalization of Fubini's Theorem.

(ii) Applying the Coarea Formula to $A = \{Jf = 0\}$, we discover

$$\mathcal{H}^{n-m}(\{Jf = 0\} \cap f^{-1}\{y\}) = 0 \qquad (\star)$$

for $\mathcal{L}^m$ a.e. $y \in \mathbb{R}^m$. This is a weak variant of the *Morse–Sard Theorem*, which asserts

$$\{Jf = 0\} \cap f^{-1}\{y\} = \emptyset$$

for $\mathcal{L}^m$ a.e. y, provided $f \in C^k(\mathbb{R}^n; \mathbb{R}^m)$ for

$$k = 1 + n - m.$$

Observe, however, $(\star)$ only requires that f be Lipschitz. ∎

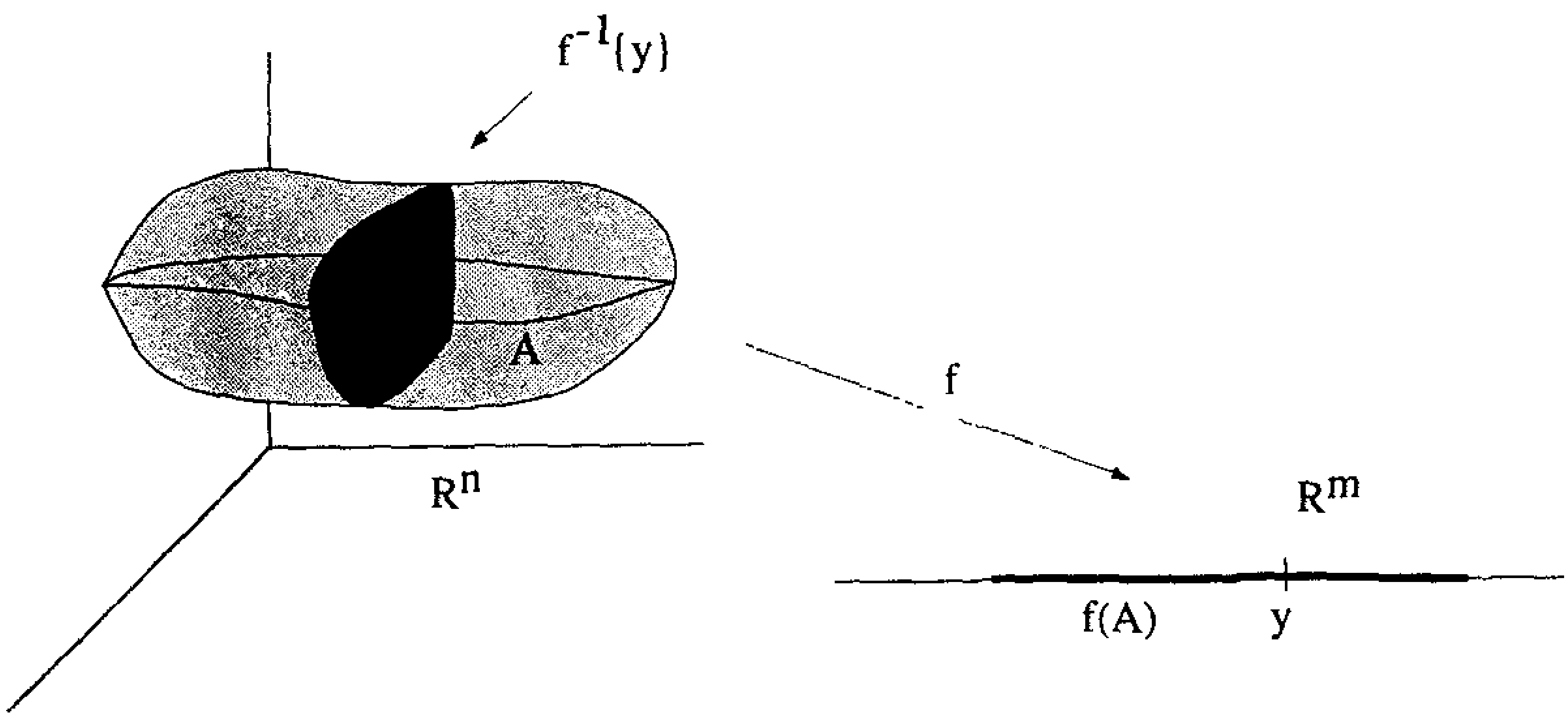

FIGURE 3.7
The Coarea Formula.

PROOF

1. In view of Lemma 2, we may assume that $Df(x)$, and thus $Jf(x)$, exist for all $x \in A$ and that $\mathcal{L}^n(A) < \infty$.

2. Case 1. $A \subset \{Jf > 0\}$.

For each $\lambda \in \Lambda(n, n-m)$, write

$$f = q \circ h_\lambda,$$

where

$$h_\lambda : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^{n-m}, \quad h_\lambda(x) \equiv (f(x), P_\lambda(x)) \quad (x \in \mathbb{R}^n)$$

$$q : \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m, \quad q(y,z) \equiv y \qquad (y \in \mathbb{R}^m, z \in \mathbb{R}^{n-m}),$$

and P_λ is the projection defined in Section 3.2.1. Set

$$\begin{aligned} A_\lambda &\equiv \{x \in A \mid \det Dh_\lambda \neq 0\} \\ &= \{x \in A \mid P_\lambda \mid_{[Df(x)]^{-1}(0)} \text{ is injective}\}. \end{aligned}$$

Now $A = \cup_{\lambda \in \Lambda(n,n-m)} A_\lambda$; therefore we may as well for simplicity assume $A = A_\lambda$ for some $\lambda \in \Lambda(n, n-m)$

3. Fix $t > 1$ and apply Lemma 3 to $h = h_\lambda$ to obtain disjoint Borel sets $\{D_k\}_{k=1}^\infty$ and symmetric automorphisms $\{S_k\}_{k=1}^\infty$ satisfying assertions (i)–(iii) in Lemma 3. Set $G_k \equiv A \cap D_k$.

4. Claim #1: $t^{-n}[\![q \circ S_k]\!] \le Jf \mid_{G_k} \le t^n [\![q \circ S_k]\!]$.

Proof of Claim #1: Since $f = q \circ h$, we have $\mathcal{L}^n$ a.e.

$$\begin{aligned} Df &= q \circ Dh \\ &= q \circ S_k \circ S_k^{-1} \circ Dh \\ &= q \circ S_k \circ D(S_k^{-1} \circ h) \\ &= q \circ S_k \circ C, \end{aligned}$$

where $C \equiv D(S_k^{-1} \circ h)$.

By Lemma 3,

$$t^{-1} \le \text{Lip}\,(S_k^{-1} \circ h) = \text{Lip}\,(C) \le t \text{ on } G_k. \qquad (\star)$$

Now write

$$Df = S \circ O^\star$$
$$q \circ S_k = T \circ P^\star$$

for symmetric $S, T : \mathbb{R}^m \to \mathbb{R}^m$ and orthogonal $O, P : \mathbb{R}^m \to \mathbb{R}^n$.

We have then

$$S \circ O^\star = T \circ P^\star \circ C. \qquad (\star\star)$$

Consequently,

$$S = T \circ P^\star \circ C \circ O.$$

As $G_k \subset A \subset \{Jf > 0\}$, $\det S \neq 0$ and so $\det T \neq 0$.

Thus, if $v \in \mathbb{R}^m$,

$$\begin{aligned} |T^{-1} \circ Sv| &= |P^\star \circ C \circ Ov| \\ &\leq |C \circ Ov| \\ &\leq t|Ov| \qquad \text{by } (\star) \\ &= t|v|. \end{aligned}$$

Therefore

$$(T^{-1} \circ S)(B(0,1)) \subset B(0,t),$$

and so

$$Jf = |\det S| \leq t^n |\det T| = t^n [\![q \circ S_k]\!].$$

Similarly, if $v \in \mathbb{R}^m$, we have from $(\star\star)$

$$\begin{aligned} |S^{-1} \circ Tv| &= |O^\star \circ C^{-1} \circ Pv| \\ &\leq |C^{-1} \circ Pv| \\ &\leq t|Pv| \text{ by } (\star) \\ &= t|v|. \end{aligned}$$

Thus

$$[\![q \circ S_k]\!] = |\det T| \leq t^n |\det S| = t^n Jf.$$

5. Now calculate:

$$\begin{aligned} & t^{-3n+m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(G_k \cap f^{-1}\{y\})\, dy \\ &= t^{-3n+m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}(h(G_k) \cap q^{-1}\{y\}))\, dy \\ &\leq t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1}(h(G_k) \cap q^{-1}\{y\}))\, dy \\ &= t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1} \circ h(G_k) \cap (q \circ S_k)^{-1}\{y\})\, dy \\ &= t^{-2n} [\![q \circ S_k]\!] \mathcal{L}^n(S_k^{-1} \circ h(G_k)) \qquad \text{(by Lemma 1)} \\ &\leq t^{-n} [\![q \circ S_k]\!] \mathcal{L}^n(G_k) \\ &\leq \int_{G_k} Jf\, dx \end{aligned}$$

$$
\begin{aligned}
&\leq t^n [\![q \circ S_k]\!] \mathcal{L}^n(G_k) \\
&\leq t^{2n} [\![q \circ S_k]\!] \mathcal{L}^n(S_k^{-1} \circ h(G_k)) \\
&= t^{2n} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(S_k^{-1} \circ h(G_k) \cap (q \circ S_k)^{-1}\{y\})\, dy \\
&\leq t^{3n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(h^{-1}(h(G_k) \cap q^{-1}\{y\}))\, dy \\
&= t^{3n-m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(G_k \cap f^{-1}\{y\})\, dy.
\end{aligned}
$$

Since

$$\mathcal{L}^n\left(A - \bigcup_{k=1}^{\infty} G_k\right) = 0,$$

we can sum on k, use Lemma 2, and let $t \to 1^+$ to conclude

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy = \int_A Jf\, dx.$$

6. *Case 2.* $A \subset \{Jf = 0\}$.
Fix $\epsilon > 0$ and define

$$
\begin{aligned}
&g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m, \quad g(x,y) \equiv f(x) + \epsilon y \\
&p : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m, \quad p(x,y) \equiv y \qquad\qquad (x \in \mathbb{R}^n,\ y \in \mathbb{R}^m).
\end{aligned}
$$

Then

$$Dg = (Df,\ \epsilon I)_{m \times (n+m)},$$

and

$$\epsilon^m \leq Jg = [\![Dg]\!] = [\![Dg^*]\!] \leq C\epsilon.$$

7. Observe

$$
\begin{aligned}
&\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy \\
&\qquad = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y - \epsilon w\})\, dy \quad \text{for all } w \in \mathbb{R}^m \\
&\qquad = \frac{1}{\alpha(m)} \int_{B(0,1)} \int_{\mathbb{R}^m} \mathcal{H}^{n-1}(A \cap f^{-1}\{y - \epsilon w\})\, dy\, dw.
\end{aligned}
$$

8. Claim #2: Fix $y \in \mathbb{R}^m$, $w \in \mathbb{R}^m$, and set $B \equiv A \times B(0,1) \subset \mathbb{R}^{n+m}$. Then

$$B \cap g^{-1}\{y\} \cap p^{-1}\{w\} = \begin{cases} \emptyset & \text{if } w \notin B(0,1) \\ (A \cap f^{-1}\{y - \epsilon w\}) \times \{w\} & \text{if } w \in B(0,1). \end{cases}$$

Proof of Claim #2: We have $(x,z) \in B \cap g^{-1}\{y\} \cap p^{-1}\{w\}$ if and only if

$$x \in A, \quad z \in B(0,1), \quad f(x) + \epsilon z = y, \quad z = w;$$

if and only if

$$x \in A, \quad z = w \in B(0,1), \quad f(x) = y - \epsilon w;$$

if and only if

$$w \in B(0,1), \quad (x,z) \in (A \cap f^{-1}\{y - \epsilon w\}) \times \{w\}.$$

9. Now use Claim #2 to continue the calculation from step 7:

$$\begin{aligned} &\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy \\ &\quad = \frac{1}{\alpha(m)} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(B \cap g^{-1}\{y\} \cap p^{-1}\{w\})\, dw\, dy \\ &\quad \leq \frac{\alpha(n-m)}{\alpha(n)} \int_{\mathbb{R}^m} \mathcal{H}^n(B \cap g^{-1}\{y\})\, dy \qquad \text{(by the Remark after Lemm} \\ &\quad = \frac{\alpha(n-m)}{\alpha(n)} \int_B Jg\, dx\, dz \\ &\quad \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)} \mathcal{L}^n(A) \sup_B Jg \\ &\quad \leq C\mathcal{L}^n(A)\epsilon. \end{aligned}$$

Let $\epsilon \to 0$ to obtain

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, dy = 0 = \int_A Jf\, dx.$$

10. In the general case we write $A = A_1 \cup A_2$ where $A_1 \subset \{Jf > 0\}$, $A_2 \subset \{Jf = 0\}$, and apply Cases 1 and 2 above. ▮

3.4.3 Change of variables formula

THEOREM 2
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \geq m$. Then for each $\mathcal{L}^n$-summable function $g : \mathbb{R}^n \to \mathbb{R}$,

$$g\,|_{f^{-1}\{y\}} \quad \text{is } \mathcal{H}^{n-m} \text{ summable for } \mathcal{L}^m \text{ a.e. } y$$

and

$$\int_{\mathbb{R}^n} g(x) Jf(x)\, dx = \int_{\mathbb{R}^m} \left[\int_{f^{-1}\{y\}} g \; d\mathcal{H}^{n-m} \right] dy.$$

REMARK For each $y \in \mathbb{R}^m$, $f^{-1}\{y\}$ is closed and thus $\mathcal{H}^{n-m}$-measurable. ▮

PROOF

1. Case 1. $g \geq 0$.

Write $g = \sum_{i=1}^{\infty} (1/i)\chi_{A_i}$ for appropriate $\mathcal{L}^n$-measurable sets $\{A_i\}_{i=1}^{\infty}$; this is possible owing to Theorem 7 in Section 1.1.2. Then the Monotone Convergence Theorem implies

$$\begin{aligned}
\int_{\mathbb{R}^n} g\, Jf\, dx &= \sum_{i=1}^{\infty} \frac{1}{i} \int_{A_i} Jf\, dx \\
&= \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A_i \cap f^{-1}\{y\})\, dy \\
&= \int_{\mathbb{R}^n} \sum_{i=1}^{\infty} \frac{1}{i} \mathcal{H}^{n-m}(A_i \cap f^{-1}\{y\})\, dy \\
&= \int_{\mathbb{R}^n} \left[\int_{f^{-1}\{y\}} g\, d\mathcal{H}^{n-m} \right] dy.
\end{aligned}$$

2. Case 2. g is any $\mathcal{L}^n$-summable function. Write $g = g^+ - g^-$ and use Case 1. ▮

3.4.4 Applications

A. Polar coordinates.

PROPOSITION 1
Let $g : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{L}^n$-summable. Then

$$\int_{\mathbb{R}^n} g\, dx = \int_0^\infty \left(\int_{\partial B(0,r)} g\, d\mathcal{H}^{n-1} \right) dr.$$

In particular, we see

$$\frac{d}{dr}\left(\int_{B(0,r)} g\, dx \right) = \int_{\partial B(0,r)} g\, d\mathcal{H}^{n-1}$$

for $\mathcal{L}^1$ a.e. $r > 0$.

PROOF Set $f(x) = |x|$; then

$$Df(x) = \frac{x}{|x|}, \quad Jf(x) = 1 \quad (x \neq 0). \quad \blacksquare$$

B. Level sets.

PROPOSITION 2
Assume $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz. Then

$$\int_{\mathbb{R}^n} |Df|\, dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\{f = t\})\, dt.$$

PROOF $Jf = |Df|$. $\blacksquare$

REMARK Compare this with the Coarea Formula for BV functions proved in Section 5.5. $\blacksquare$

PROPOSITION 3
Let $f : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz, with

$$\operatorname{ess\,inf} |Df| > 0.$$

Suppose also $g : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{L}^n$-summable. Then

$$\int_{\{f>t\}} g\, dx = \int_t^\infty \left(\int_{\{f=s\}} \frac{g}{|Df|}\, d\mathcal{H}^{n-1} \right) ds.$$

In particular, we see

$$\frac{d}{dt}\left(\int_{\{f>t\}} g\,dx\right) = -\int_{\{f=t\}} \frac{g}{|Df|}\,d\mathcal{H}^{n-1}$$

for $\mathcal{L}^1$ a.e. t.

PROOF As above, $Jf = |Df|$. Write $E_t \equiv \{f > t\}$ and use Theorem 2 to calculate

$$\begin{aligned}\int_{\{f>t\}} g\,dx &= \int_{\mathbb{R}^n} \chi_{E_t} \frac{g}{|Df|} Jf\,dx \\ &= \int_{-\infty}^{\infty}\left(\int_{\partial E_s} \frac{g}{|Df|}\chi_{E_t}\,d\mathcal{H}^{n-1}\right) ds \\ &= \int_t^{\infty}\left(\int_{\partial E_s} \frac{g}{|Df|}\,d\mathcal{H}^{n-1}\right) ds. \quad \blacksquare\end{aligned}$$

4

Sobolev Functions

In this chapter we study *Sobolev functions* on $\mathbb{R}^n$, functions whose weak first partial derivatives belong to some L^p space. The various Sobolev spaces have good completeness and compactness properties and consequently are often proper settings for the applications of functional analysis to, for instance, linear and nonlinear PDE theory.

Now, as we will see, by definition, integration-by-parts is valid for Sobolev functions. It is, however, far less obvious to what extent the other rules of calculus are valid. We intend to investigate this general question, with particular emphasis on pointwise properties of Sobolev functions.

Section 4.1 provides basic definitions. In Section 4.2 we derive various ways of approximating Sobolev functions by smooth functions. Section 4.3 interprets boundary values of Sobolev functions using traces, and Section 4.4 discusses extending such functions off Lipschitz domains. We prove the fundamental Sobolev-type inequalities in Section 4.5, an immediate application of which is the compactness theorem in Section 4.6. The key to understanding fine properties of Sobolev functions is capacity, introduced in Section 4.7 and utilized in Sections 4.8 and 4.9.

4.1 Definitions and elementary properties

Throughout this chapter, let U denote an open subset of $\mathbb{R}^n$.

DEFINITION *Assume $f \in L^1_{\mathrm{loc}}(U)$, $1 \leq i \leq n$. We say $g_i \in L^1_{\mathrm{loc}}(U)$ is the **weak partial derivative** of f with respect to x_i in U if*

$$\int_U f \frac{\partial \varphi}{\partial x_i}\, dx = -\int_U g_i \varphi \, dx \qquad (\star)$$

for all $\varphi \in C^1_c(U)$.

NOTATION It is easy to check that the weak partial derivative with respect to x_i, if it exists, is uniquely defined $\mathcal{L}^n$ a.e. We write

$$\frac{\partial f}{\partial x_i} \equiv g_i \qquad (i = 1, \dots, n)$$

and

$$Df \equiv \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right),$$

provided the weak derivatives $\partial f / \partial x_1, \dots, \partial f / \partial x_n$ exist.

DEFINITIONS *Let $1 \leq p \leq \infty$.*

(i) *The function f belongs to the* ***Sobolev space***

$$W^{1,p}(U)$$

if $f \in L^p(U)$ and the weak partial derivatives $\partial f / \partial x_i$ exist and belong to $L^p(U)$, $i = 1, \dots, n$.

(ii) *The function f belongs to $W^{1,p}_{\text{loc}}(U)$ if $f \in W^{1,p}(V)$ for each open set $V \subset\subset U$.*

(iii) *We say f is a* ***Sobolev function*** *if $f \in W^{1,p}_{\text{loc}}(U)$ for some $1 \leq p \leq \infty$.*

REMARK Note carefully: if f is a Sobolev function, then *by definition* the integration-by-parts formula

$$\int_U f \frac{\partial \varphi}{\partial x_i} \, dx = - \int_U \frac{\partial f}{\partial x_i} \varphi \, dx$$

is valid for all $\varphi \in C^1_c(U)$, $i = 1, \dots n$. ▮

NOTATION If $f \in W^{1,p}(U)$, define

$$||f||_{W^{1,p}(U)} \equiv \left(\int_U |f|^p + |Df|^p \, dx \right)^{1/p}$$

for $1 \leq p < \infty$, and

$$||f||_{W^{1,\infty}(U)} \equiv \text{ess} \sup_U (|f| + |Df|).$$

DEFINITION *We say*

$$f_k \to f \ \textit{ in } \ W^{1,p}(U)$$

provided

$$\|f_k - f\|_{W^{1,p}(U)} \to 0,$$

and

$$f_k \to f \ \ in \ \ W^{1,p}_{\mathrm{loc}}(U)$$

provided

$$\|f_k - f\|_{W^{1,p}(V)} \to 0 \textit{ for each } V \subset\subset U.$$

4.2 Approximation

4.2.1 Approximation by smooth functions

NOTATION *(i)* If $\epsilon > 0$, we write $U_\epsilon \equiv \{x \in U \mid \operatorname{dist}(x, \partial U) > \epsilon\}$.

(ii) Define the C^∞-function $\eta : \mathbb{R}^n \to \mathbb{R}$ as follows:

$$\eta(x) \equiv \begin{cases} c \exp\left(\dfrac{1}{|x|^2 - 1}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

the constant c adjusted so

$$\int_{\mathbb{R}^n} \eta(x)\, dx = 1.$$

Next define

$$\eta_\epsilon(x) \equiv \frac{1}{\epsilon^n} \eta\left(\frac{x}{\epsilon}\right) \qquad (\epsilon > 0, x \in \mathbb{R}^n);$$

η_ϵ is the *standard mollifier.*

(iii) If $f \in L^1_{\mathrm{loc}}(U)$, define

$$f^\epsilon \equiv \eta_\epsilon * f;$$

that is,

$$f^\epsilon(x) \equiv \int_U \eta_\epsilon(x - y) f(y)\, dy \qquad (x \in U_\epsilon).$$

Mollification provides us with a systematic technique for approximating Sobo lev functions by C^∞ functions.

THEOREM 1

(i) *For each* $\epsilon > 0$, $f^\epsilon \in C^\infty(U_\epsilon)$.

(ii) *If* $f \in C(U)$, *then*

$$f^\epsilon \to f$$

uniformly on compact subsets of U.

(iii) *If* $f \in L^p_{\mathrm{loc}}(U)$ *for some* $1 \le p < \infty$, *then*

$$f^\epsilon \to f \text{ in } L^p_{\mathrm{loc}}(U).$$

(iv) *Furthermore,* $f^\epsilon(x) \to f(x)$ *if* x *is a Lebesgue point of* f*; in particular,*

$$f^\epsilon \to f \qquad \mathcal{L}^n \text{ a.e.}$$

(v) *If* $f \in W^{1,p}_{\mathrm{loc}}(U)$ *for some* $1 \le p \le \infty$, *then*

$$\frac{\partial f^\epsilon}{\partial x_i} = \eta_\epsilon * \frac{\partial f}{\partial x_i} \qquad (i = 1, \dots, n)$$

on U_ϵ.

(vi) *In particular, if* $f \in W^{1,p}_{\mathrm{loc}}(U)$ *for some* $1 \le p < \infty$, *then*

$$f^\epsilon \to f \text{ in } W^{1,p}_{\mathrm{loc}}(U).$$

PROOF

1. Fix any point $x \in U_\epsilon$, choose $1 \le i \le n$, and write e_i to denote the ith coordinate vector $(0, \dots, 1, \dots, 0)$. Then for $|h|$ small enough, $x + he_i \in U_\epsilon$, and we may compute

$$\begin{aligned}\frac{f^\epsilon(x+he_i) - f^\epsilon(x)}{h} &= \frac{1}{\epsilon^n}\int_U \frac{1}{h}\left[\eta\left(\frac{x+he_i-y}{\epsilon}\right) - \eta\left(\frac{x-y}{\epsilon}\right)\right] f(y)\,dy \\ &= \frac{1}{\epsilon^n}\int_V \frac{1}{h}\left[\eta\left(\frac{x+he_i-y}{\epsilon}\right) - \eta\left(\frac{x-y}{\epsilon}\right)\right] f(y)\,dy\end{aligned}$$

for some $V \subset\subset U$. The difference quotient converges as $h \to 0$ to

$$\frac{1}{\epsilon}\frac{\partial\eta}{\partial x_i}\left(\frac{x-y}{\epsilon}\right) = \epsilon^n \frac{\partial \eta_\epsilon}{\partial x_i}(x-y)$$

for each $y \in V$. Furthermore, the absolute value of the integrand is bounded by

$$\frac{1}{\epsilon}||D\eta||_{L^\infty}|f| \in L^1(V).$$

Hence the Dominated Convergence Theorem implies

$$\frac{\partial f^\epsilon}{\partial x_i}(x) = \lim_{h\to 0}\frac{f^\epsilon(x+he_i) - f^\epsilon(x)}{h}$$

exists and equals

$$\int_U \frac{\partial \eta_\epsilon}{\partial x_i}(x-y) f(y)\, dy.$$

A similar argument demonstrates that the partial derivatives of f^ϵ of all orders exist and are continuous at each point of U_ϵ; this proves (i).

2. Given $V \subset\subset U$, we choose $V \subset W \subset U$. Then for $x \in V$,

$$f^\epsilon(x) = \frac{1}{\epsilon^n} \int_{B(x,\epsilon)} \eta\left(\frac{x-y}{\epsilon}\right) f(y)\, dy = \int_{B(0,1)} \eta(z) f(x - \epsilon z)\, dz.$$

Thus, since $\int_{B(0,1)} \eta(z)\, dz = 1$,

$$|f^\epsilon(x) - f(x)| \leq \int_{B(0,1)} \eta(z) |f(x-\epsilon z) - f(x)|\, dz.$$

If f is uniformly continuous on W, we conclude from this estimate that $f^\epsilon \to f$ uniformly on V. Assertion (ii) follows.

3. Assume $1 \leq p < \infty$ and $f \in L^p_{\mathrm{loc}}(U)$. Then for $V \subset\subset W \subset\subset U$, $x \in V$, and $\epsilon > 0$ small enough, we calculate in case $1 < p < \infty$

$$\begin{aligned}
|f^\epsilon(x)| &\leq \int_{B(0,1)} \eta(z)^{1-\frac{1}{p}} \eta(z)^{\frac{1}{p}} |f(x-\epsilon z)|\, dz \\
&\leq \left(\int_{B(0,1)} \eta(z)\, dz\right)^{1-\frac{1}{p}} \left(\int_{B(0,1)} \eta(z) |f(x-\epsilon z)|^p\, dz\right)^{\frac{1}{p}} \\
&= \left(\int_{B(0,1)} \eta(z) |f(x-\epsilon z)|^p\, dz\right)^{\frac{1}{p}}.
\end{aligned}$$

Hence for $1 \leq p < \infty$ we find

$$\begin{aligned}
\int_V |f^\epsilon(x)|^p\, dx &\leq \int_{B(0,1)} \eta(z) \left(\int_V |f(x-\epsilon z)|^p\, dx\right) dz \\
&\leq \int_W |f(y)|^p\, dy \qquad (\star)
\end{aligned}$$

for $\epsilon > 0$ small enough.

Now fix $\delta > 0$. Since $f \in L^p(W)$, there exists $g \in C(\overline{W})$ such that

$$\|f - g\|_{L^p(W)} \leq \delta.$$

This implies, according to estimate $(\star)$,

$$\|f^\epsilon - g^\epsilon\|_{L^p(V)} \leq \delta.$$

Consequently,

$$\|f^\epsilon - f\|_{L^p(V)} \le 2\delta + \|g^\epsilon - g\|_{L^p(V)} \le 3\delta$$

provided $\epsilon > 0$ is small enough, owing to assertion (ii). Assertion (iii) is proved.

4. To prove (iv), let us suppose $f \in L^1_{\mathrm{loc}}(U)$ and $x \in U$ is a Lebesgue point of f. Then, by the calculation above, we see

$$\begin{aligned}|f^\epsilon(x) - f(x)| &\le \frac{1}{\epsilon^n}\int_{B(x,\epsilon)} \eta\left(\frac{x-y}{\epsilon}\right)|f(y) - f(x)|\,dy\\ &\le \alpha(n)\|\eta\|_{L^\infty} ⨍_{B(x,\epsilon)} |f - f(x)|\,dy\\ &= o(1) \qquad \text{as } \epsilon \to 0.\end{aligned}$$

5. Now assume $f \in W^{1,p}_{\mathrm{loc}}(U)$ for some $1 \le p \le \infty$. Consequently, as computed above,

$$\begin{aligned}\frac{\partial f^\epsilon}{\partial x_i}(x) &= \int_U \frac{\partial \eta_\epsilon}{\partial x_i}(x-y)f(y)\,dy\\ &= -\int_U \frac{\partial \eta_\epsilon}{\partial y_i}(x-y)f(y)\,dy\\ &= \int_U \eta_\epsilon(x-y)\frac{\partial f}{\partial y_i}(y)\,dy\\ &= \eta_\epsilon * \frac{\partial f}{\partial x_i}(x)\end{aligned}$$

for $x \in U_\epsilon$. This establishes assertion (v), and (vi) follows at once from (iii). ∎

THEOREM 2 LOCAL APPROXIMATION BY SMOOTH FUNCTIONS

Assume $f \in W^{1,p}(U)$ for some $1 \le p < \infty$. Then there exists a sequence $\{f_k\}_{k=1}^\infty \subset W^{1,p}(U) \cap C^\infty(U)$ such that

$$f_k \to f \quad in \quad W^{1,p}(U).$$

Note that we do not assert $f_k \in C^\infty(\overline{U})$: see Theorem 3 below.

PROOF

1. Fix $\epsilon > 0$ and define

$$\begin{cases} U_k \equiv \left\{x \in U \mid \mathrm{dist}(x, \partial U) > \dfrac{1}{k}\right\} \cap U(0,k) & (k = 1, 2, \ldots),\\ U_0 \equiv \emptyset. \end{cases}$$

Set

$$V_k \equiv U_{k+1} - \overline{U}_{k-1} \qquad (k = 1, 2, \ldots),$$

and let $\{\zeta_k\}_{k=1}^{\infty}$ be a sequence of smooth functions such that

$$\begin{cases} \zeta_k \in C_c^{\infty}(V_k), & 0 \le \zeta_k \le 1, \qquad (k = 1, 2, \ldots), \\ \sum_{k=1}^{\infty} \zeta_k \equiv 1 & \text{on } U. \end{cases}$$

For each $k = 1, 2, \ldots$, $f\zeta_k \in W^{1,p}(U)$, with spt $(f\zeta_k) \subset V_k$; hence there exists $\epsilon_k > 0$ such that

$$\begin{cases} \text{spt } (\eta_{\epsilon_k} * (f\zeta_k)) \subset V_k \\ \left(\int_U | \eta_{\epsilon_k} * (f\zeta_k) - f\zeta_k |^p \, dx \right)^{\frac{1}{p}} < \frac{\epsilon}{2^k} \\ \left(\int_U | \eta_{\epsilon_k} * (D(f\zeta_k)) - D(f\zeta_k) |^p \, dx \right)^{\frac{1}{p}} < \frac{\epsilon}{2^k} . \end{cases} \tag{$\star$}$$

Define

$$f_\epsilon \equiv \sum_{k=1}^{\infty} \eta_{\epsilon_k} * (f\zeta_k).$$

In some neighborhood of each point $x \in U$, there are only finitely many nonzero terms in this sum; hence

$$f_\epsilon \in C^{\infty}(U).$$

2. Since

$$f = \sum_{k=1}^{\infty} f\zeta_k,$$

($\star$) implies

$$||f_\epsilon - f||_{L^p(U)} \le \sum_{k=1}^{\infty} \left(\int_U |\eta_{\epsilon_k} * (f\zeta_k) - f\zeta_k|^p \, dx \right)^{\frac{1}{p}} < \epsilon$$

and

$$||Df_\epsilon - Df||_{L^p(U)} \le \sum_{k=1}^{\infty} \left(\int_U |\eta_{\epsilon_k} * (D(f\zeta_k)) - D(f\zeta_k)|^p \, dx \right)^{\frac{1}{p}} < \epsilon.$$

Consequently $f_\epsilon \in W^{1,p}(U)$ and

$$f_\epsilon \to f \text{ in } W^{1,p}(U) \text{ as } \epsilon \to 0. \quad \blacksquare$$

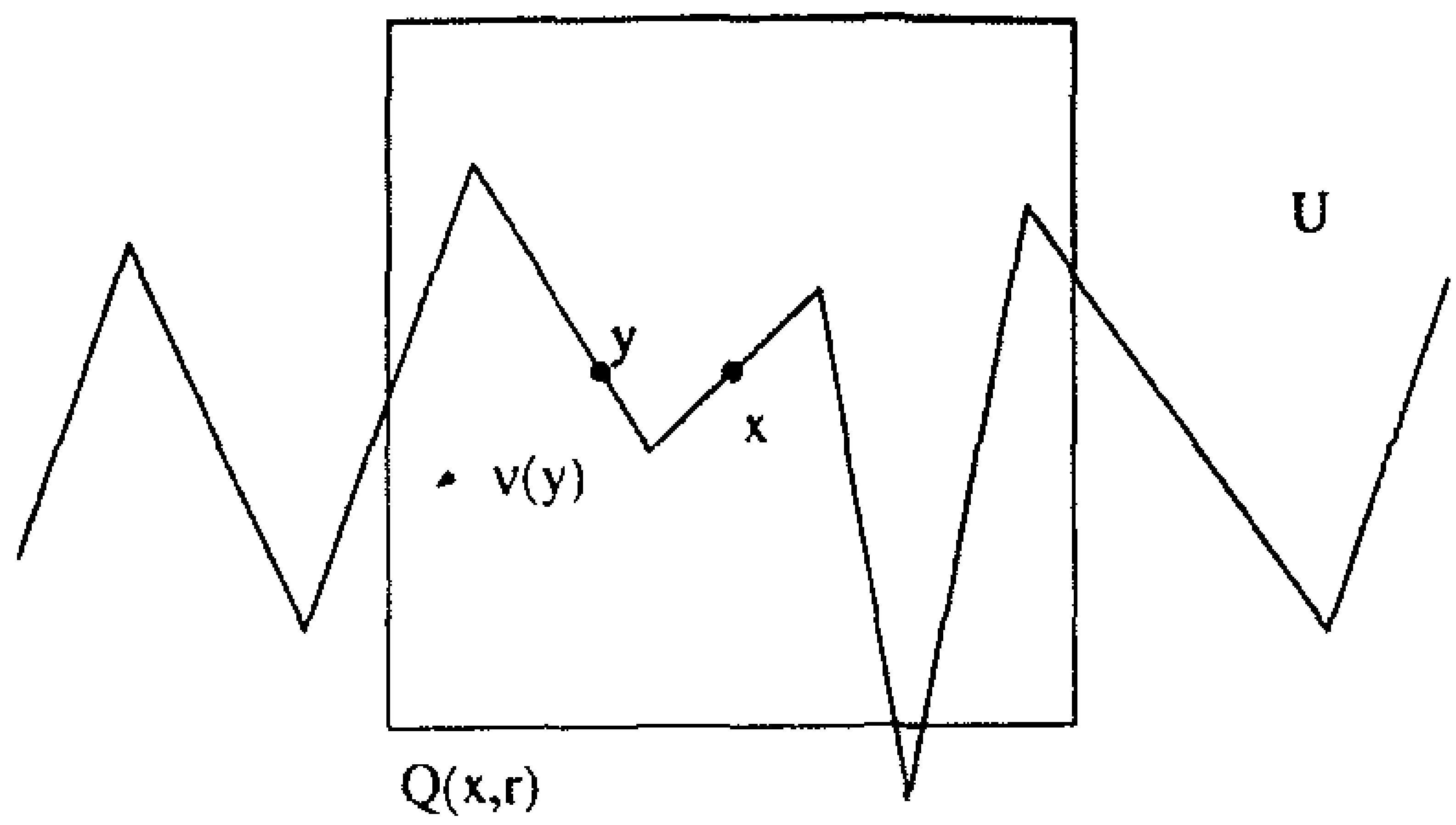

FIGURE 4.1
A Lipschitz boundary.

Our intention next is to approximate a Sobolev function by functions smooth all the way up to the boundary. This necessitates some hypothesis on the geometric behavior of ∂U.

DEFINITION *We say ∂U is* ***Lipschitz*** *if for each point $x \in \partial U$, there exist $r > 0$ and a Lipschitz mapping $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that — upon rotating and relabeling the coordinate axes if necessary — we have*

$$U \cap Q(x,r) = \{y \mid \gamma(y_1, \dots, y_{n-1}) < y_n\} \cap Q(x,r),$$

where $Q(x,r) \equiv \{y \mid |y_i - x_i| < r, i = 1, \dots, n\}$.

In other words, near x, ∂U is the graph of a Lipschitz function.

REMARK By Rademacher's Theorem, Section 3.1.2, the outer unit normal $\nu(x)$ to U exists for $\mathcal{H}^{n-1}$ a.e. $x \in \partial U$. ▮

THEOREM 3 GLOBAL APPROXIMATION BY SMOOTH FUNCTIONS
Assume U is bounded, ∂U Lipschitz. Then if $f \in W^{1,p}(U)$ for some $1 \leq p < \infty$, there exists a sequence $\{f_k\}_{k=1}^{\infty} \subset W^{1,p}(U) \cap C^{\infty}(\bar{U})$ such that $f_k \to f$ in $W^{1,p}(U)$.

PROOF

1. For $x \in \partial U$, take $r > 0$ and $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ as in the definition above. Also write $Q \equiv Q(x,r)$, $Q' = Q(x, r/2)$.

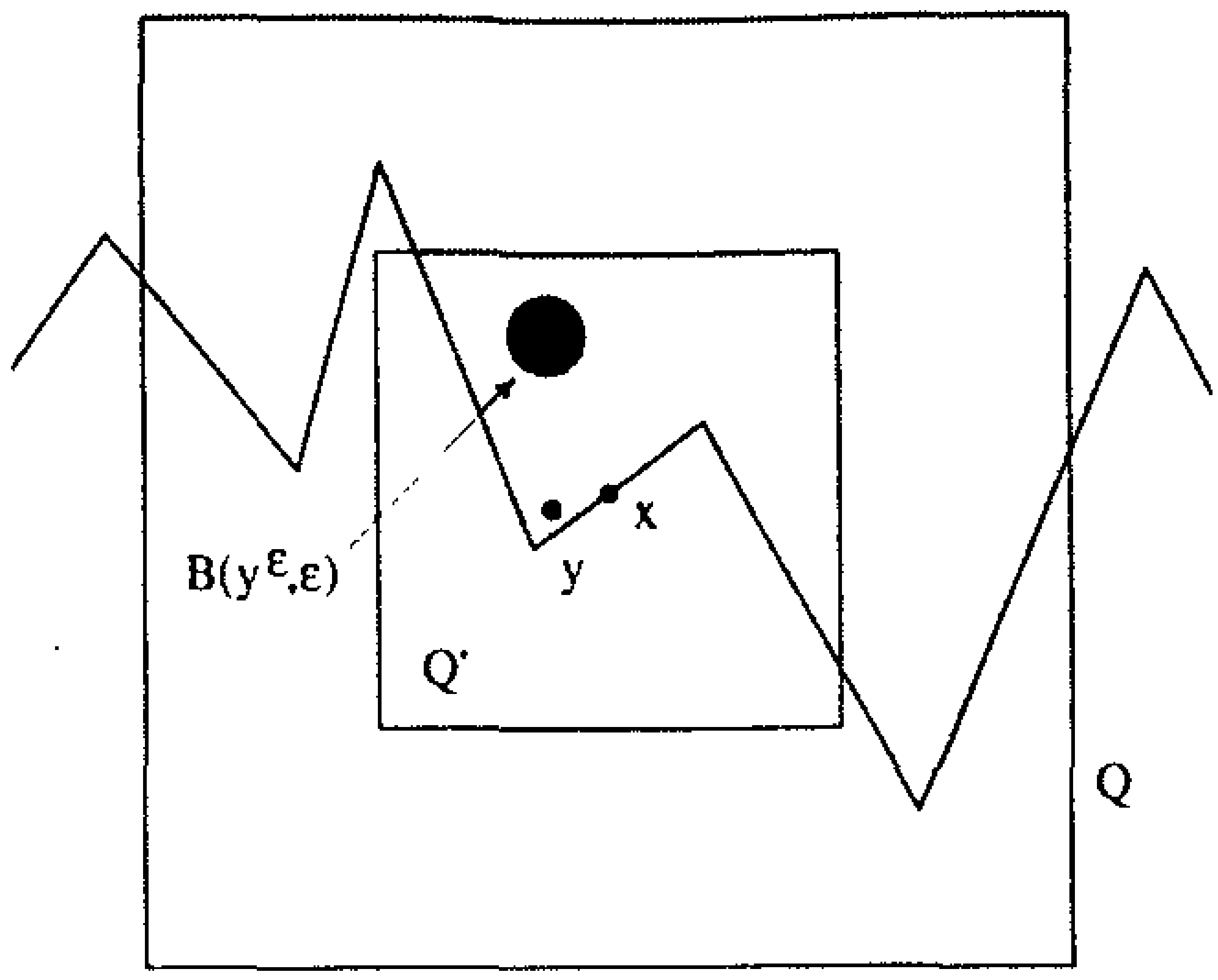

FIGURE 4.2
The small ball $B(y^\varepsilon, \varepsilon)$ lies in $U \cap Q$.

2. Suppose first f vanishes near $\partial Q' \cap U$. For $y \in U \cap Q'$, $\epsilon > 0$ and $\alpha > 0$, we define

$$y^\epsilon \equiv y + \epsilon\alpha e_n.$$

Observe $B(y^\epsilon, \epsilon) \subset U \cap Q$ for all ϵ sufficiently small, provided α is large enough, say $\alpha \equiv \text{Lip}\,(\gamma) + 2$. See Figure 4.2.

We define

$$\begin{aligned} f_\epsilon(y) &\equiv \frac{1}{\epsilon^n} \int_U \eta\left(\frac{z}{\epsilon}\right) f(y^\epsilon - z)\, dz \\ &= \frac{1}{\epsilon^n} \int_{B(y^\epsilon,\epsilon)} \eta\left(\frac{y-w}{\epsilon} + \alpha e_n\right) f(w)\, dw \end{aligned}$$

for $y \in U \cap Q'$.

3. As in the proof of Theorem 1, we check

$$f_\epsilon \in C^\infty(\overline{U \cap Q'})$$

and

$$f_\epsilon \to f \text{ in } W^{1,p}(U \cap Q').$$

Furthermore, since $f = 0$ near $\partial Q' \cap U$, we have $f_\epsilon = 0$ near $\partial Q' \cap U$ for sufficiently small $\epsilon > 0$; we can thus extend f_ϵ to be 0 on $U - Q'$.

4. Since ∂U is compact, we can cover ∂U with finitely many cubes $Q_i' = Q(x_i, r_i/2)$ $(i = 1, 2, \ldots, N)$, as above. Let $\{\zeta_i\}_{i=0}^N$ be a sequence of smooth functions such that

$$\begin{cases} 0 \le \zeta_i \le 1, & \operatorname{spt} \zeta_i \subset Q_i' \qquad (i = 1, \dots, N) \\ 0 \le \zeta_0 \le 1, & \operatorname{spt} \zeta_0 \subset U \\ \sum_{i=0}^{N} \zeta_i \equiv 1 & \text{on } U \end{cases}$$

and set

$$f^i \equiv f\zeta_i \qquad (i = 0, 1, 2, \dots, N).$$

Fix $\delta > 0$. Construct as in step 3 functions $g^i \equiv (f^i)_{\epsilon_i} \in C^\infty(\overline{U})$ satisfying

$$\begin{cases} \operatorname{spt}(g^i) \subset \overline{U} \cap Q_i \\ \|g^i - f^i\|_{W^{1,p}(U \cap Q_i)} < \dfrac{\delta}{2N} \end{cases}$$

for $i = 1, \dots, N$. Mollify f^0 as in the proof of Theorem 2 to produce $g^0 \in C_c^\infty(U)$ such that

$$\|g^0 - f^0\|_{W^{1,p}(U)} < \frac{\delta}{2}\,.$$

Finally, set

$$g \equiv \sum_{i=0}^{N} g^i \in C^\infty(\overline{U})$$

and compute

$$\|g - f\|_{W^{1,p}(U)} \le \|g^0 - f^0\|_{W^{1,p}(U)} + \sum_{i=1}^{N} \|g^i - f^i\|_{W^{1,p}(U \cap Q_i)} < \delta. \qquad \blacksquare$$

4.2.2 Product and chain rules

In view of Section 4.2.1 we can approximate Sobolev functions by smooth functions, and consequently we can now verify that many of the usual calculus rules hold for weak derivatives.

Assume $1 \le p < \infty$.

THEOREM 4

(i) ***(Product rule)*** *If $f, g \in W^{1,p}(U) \cap L^\infty(U)$, then $fg \in W^{1,p}(U) \cap L^\infty(U)$ and*

$$\frac{\partial (fg)}{\partial x_i} = \frac{\partial f}{\partial x_i} g + f \frac{\partial g}{\partial x_i} \qquad \mathcal{L}^n \text{ a.e. } (i = 1, 2, \dots, n).$$

(ii) **(Chain rule)** *If $f \in W^{1,p}(U)$ and $F \in C^1(\mathbb{R})$, $F' \in L^\infty(\mathbb{R})$, $F(0) = 0$, then $F(f) \in W^{1,p}(U)$ and*

$$\frac{\partial F(f)}{\partial x_i} = F'(f)\frac{\partial f}{\partial x_i} \qquad \mathcal{L}^n \text{ a.e. } (i = 1, \ldots, n).$$

(If $\mathcal{L}^n(U) < \infty$, the condition $F(0) = 0$ is unnecessary.)

(iii) *If $f \in W^{1,p}(U)$, then f^+, f^-, $|f| \in W^{1,p}(U)$ and*

$$Df^+ = \begin{cases} Df & \mathcal{L}^n \text{ a.e. on } \{f > 0\} \\ 0 & \mathcal{L}^n \text{ a.e. on } \{f \leq 0\} \end{cases}$$

$$Df^- = \begin{cases} 0 & \mathcal{L}^n \text{ a.e. on } \{f \geq 0\} \\ -Df & \mathcal{L}^n \text{ a.e. on } \{f < 0\} \end{cases}$$

$$D|f| = \begin{cases} Df & \mathcal{L}^n \text{ a.e. on } \{f > 0\} \\ 0 & \mathcal{L}^n \text{ a.e. on } \{f = 0\} \\ -Df & \mathcal{L}^n \text{ a.e. on } \{f < 0\}. \end{cases}$$

(iv) $Df = 0$ *$\mathcal{L}^n$ a.e. on $\{f = 0\}$.*

REMARK Assertion (iv) generalizes Corollary 1(i) in Section 3.1.2. If F is only Lipschitz, the chain rule is valid but more subtle. ∎

PROOF

1. To establish (i), choose $\varphi \in C_c^1(U)$ with spt $(\varphi) \subset V \subset\subset U$. Let $f^\epsilon \equiv \eta_\epsilon * f$, $g^\epsilon \equiv \eta_\epsilon * g$ as in Section 4.2.1. Then

$$\begin{aligned}
\int_U fg\frac{\partial \varphi}{\partial x_i} &= \int_V fg\frac{\partial \varphi}{\partial x_i}\, dx \\
&= \lim_{\epsilon\to 0}\int_V f^\epsilon g^\epsilon \frac{\partial \varphi}{\partial x_i}\, dx \\
&= -\lim_{\epsilon\to 0}\int_V \left(\frac{\partial f^\epsilon}{\partial x_i}g^\epsilon + f^\epsilon\frac{\partial g^\epsilon}{\partial x_i}\right)\varphi\, dx \\
&= -\int_V \left(\frac{\partial f}{\partial x_i}g + f\frac{\partial g}{\partial x_i}\right)\varphi\, dx \\
&= -\int_U \left(\frac{\partial f}{\partial x_i}g + f\frac{\partial g}{\partial x_i}\right)\varphi\, dx,
\end{aligned}$$

according to Theorem 1.

2. To prove (ii), choose φ, V, and f' as above. Then

$$\begin{aligned}
\int_U F(f)\frac{\partial\varphi}{\partial x_i}\,dx &= \int_V F(f)\frac{\partial\varphi}{\partial x_i}\,dx \\
&= \lim_{\epsilon\to 0}\int_V F(f^\epsilon)\frac{\partial\varphi}{\partial x_i}\,dx \\
&= -\lim_{\epsilon\to 0}\int_V F'(f^\epsilon)\frac{\partial f^\epsilon}{\partial x_i}\varphi\,dx \\
&= -\int_V F'(f)\frac{\partial f}{\partial x_i}\varphi\,dx \\
&= -\int_U F'(f)\frac{\partial f}{\partial x_i}\varphi\,dx,
\end{aligned}$$

where again we have repeatedly used Theorem 1.

3. Fix $\epsilon > 0$ and define

$$F_\epsilon(r) \equiv \begin{cases} (r^2+\epsilon^2)^{\frac{1}{2}} - \epsilon & \text{if } r \geq 0 \\ 0 & \text{if } r < 0. \end{cases}$$

Then $F_\epsilon \in C^1(\mathbb{R})$, $F'_\epsilon \in L^\infty(\mathbb{R})$, and so assertion (ii) implies for $\varphi \in C^1_c(U)$

$$\int_U F_\epsilon(f)\frac{\partial\varphi}{\partial x_i}\,dx = -\int_U F'_\epsilon(f)\frac{\partial f}{\partial x_i}\varphi\,dx.$$

Now let $\epsilon \to 0$ to find

$$\int_U f^+\frac{\partial\varphi}{\partial x_i}\,dx = -\int_{U\cap\{f>0\}}\frac{\partial f}{\partial x_i}\varphi\,dx.$$

This proves the first part of (iii); the other assertions follow from the formulas

$$f^- = (-f)^+, \qquad |f| = f^+ + f^-.$$

4. Assertion (iv) follows at once from (iii), since

$$Df = Df^+ - Df^-. \quad \blacksquare$$

4.2.3 $W^{1,\infty}$ and Lipschitz functions

THEOREM 5
Let $f : U \to \mathbb{R}$. Then f is locally Lipschitz in U if and only if $f \in W^{1,\infty}_{\text{loc}}(U)$.

PROOF

1. First suppose f is locally Lipschitz. Fix $1 \leq i \leq n$. Then for each $V \subset\subset W \subset\subset U$, pick $0 < h < \text{dist}\,(V, \partial W)$, and define

$$g_i^h(x) \equiv \frac{f(x + he_i) - f(x)}{h} \qquad (x \in V).$$

Now

$$\sup_{h>0} |g_i^h| \leq \text{Lip}\,(f\,|_W) < \infty,$$

so that according to Theorem 3 in Section 1.9 there is a sequence $h_j \to 0$ and a function $g_i \in L^\infty_{\text{loc}}(U)$ such that

$$g_i^{h_j} \rightharpoonup g_i \quad \text{weakly in } L^p_{\text{loc}}(U)$$

for all $1 < p < \infty$. But if $\varphi \in C_c^1(V)$, we have

$$\int_U f(x) \frac{\varphi(x + he_i) - \varphi(x)}{h}\, dx = -\int_U g_i^h(x) \varphi(x + he_i)\, dx.$$

We set $h = h_j$ and let $j \to \infty$:

$$\int_U f \frac{\partial \varphi}{\partial x_i}\, dx = -\int_U g_i \varphi\, dx.$$

Hence g_i is the weak partial derivative of f with respect to x_i $(i = 1, \ldots n)$, and thus $f \in W^{1,\infty}_{\text{loc}}(U)$.

2. Conversely, suppose $f \in W^{1,\infty}_{\text{loc}}(U)$. Let $B \subset\subset U$ be any closed ball contained in U. Then by Theorem 1 we know

$$\sup_{0<\epsilon<\epsilon_0} ||Df^\epsilon||_{L^\infty(B)} < \infty$$

for ϵ_0 sufficiently small, where $f^\epsilon \equiv \eta_\epsilon * f$ is the usual mollification. Since f^ϵ is C^∞, we have

$$f^\epsilon(x) - f^\epsilon(y) = \int_0^1 Df^\epsilon(y + t(x - y))\, dt \cdot (x - y)$$

for $x, y \in B$, whence

$$|f^\epsilon(x) - f^\epsilon(y)| \leq C|x - y|,$$

the constant C independent of ϵ. Thus

$$|f(x) - f(y)| \leq C|x - y| \qquad (x, y \in B).$$

Hence $f\,|_B$ is Lipschitz for each ball $B \subset\subset U$, and so f is locally Lipschitz in U. ∎

4.3 Traces

THEOREM 1
Assume U is bounded, ∂U is Lipschitz, $1 \le p < \infty$.

(i) *There exists a bounded linear operator*

$$T : W^{1,p}(U) \to L^p(\partial U; \mathcal{H}^{n-1})$$

such that

$$Tf = f \quad \text{on} \quad \partial U$$

for all $f \in W^{1,p}(U) \cap C(\overline{U})$.

(ii) *Furthermore, for all $\varphi \in C^1(\mathbb{R}^n; \mathbb{R}^n)$ and $f \in W^{1,p}(U)$,*

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U Df \cdot \varphi \, dx + \int_{\partial U} (\varphi \cdot \nu) Tf \, d\mathcal{H}^{n-1},$$

ν denoting the unit outer normal to ∂U.

DEFINITION *The function Tf, which is uniquely defined up to sets of $\mathcal{H}^{n-1} \llcorner \partial U$ measure zero, is called the* **trace** *of f on ∂U. We interpret Tf as the "boundary values" of f on ∂U.*

REMARK We will see in Section 5.3 that for $\mathcal{H}^{n-1}$ a.e. point $x \in \partial U$,

$$\lim_{r\to 0} \fint_{B(x,r)\cap U} |f - Tf(x)| \, dy = 0,$$

and so

$$Tf(x) = \lim_{r\to 0} \fint_{B(x,r)\cap U} f \, dy. \quad \blacksquare$$

PROOF

1. Assume first $f \in C^1(\overline{U})$. Since ∂U is Lipschitz, we can for any point $x \in \partial U$ find $r > 0$ and a Lipschitz function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that — upon rotating and relabeling the coordinate axes if necessary —

$$U \cap Q(x,r) = \{y \mid \gamma(y_1, \dots, y_{n-1}) < y_n\} \cap Q(x,r).$$

Write $Q \equiv Q(x,r)$ and suppose temporarily $f \equiv 0$ on $U - Q$. Observe

$$-e_n \cdot \nu \ \ge\ (1 + (\operatorname{Lip}(\gamma))^2)^{-\frac{1}{2}} > 0 \qquad \mathcal{H}^{n-1} \text{ a.e. on } Q \cap \partial U. \qquad (\star)$$

2. Fix $\epsilon > 0$, set

$$\beta_\epsilon(t) \equiv (t^2 + \epsilon^2)^{\frac{1}{2}} - \epsilon \qquad (t \in \mathbb{R}),$$

and compute

$$\begin{aligned}
\int_{\partial U} \beta_\epsilon(f)\, d\mathcal{H}^{n-1} &= \int_{Q\cap\partial U} \beta_\epsilon(f)\, d\mathcal{H}^{n-1} \\
&\leq C \int_{Q\cap\partial U} \beta_\epsilon(f)(-e_n\cdot\nu)\, d\mathcal{H}^{n-1} && \text{by } (\star) \\
&= -C \int_{Q\cap U} \frac{\partial}{\partial y_n}(\beta_\epsilon(f))\, dy
\end{aligned}$$

(by the Gauss–Green Theorem; cf. Section 5.8)

$$\begin{aligned}
&\leq C \int_{Q\cap U} |\beta_\epsilon{}'(f)||Df|\, dy \\
&\leq C \int_U |Df|\, dy,
\end{aligned}$$

since $|\beta_\epsilon{}'| \leq 1$. Now let $\epsilon \to 0$ to discover

$$\int_{\partial U} |f|\, d\mathcal{H}^{n-1} \leq C \int_U |Df|\, dy. \qquad (\star\star)$$

3. We have established $(\star\star)$ under the assumption that $f \equiv 0$ on $U - Q$ for some cube $Q = Q(x,r)$, $x \in \partial U$. In the general case, we can cover ∂U by a finite number of such cubes and use a partition of unity as in the proof of Theorem 3 in Section 4.2.1 to obtain

$$\int_{\partial U} |f|\, d\mathcal{H}^{n-1} \leq C \int_U |Df| + |f|\, dy \qquad (\star\star\star)$$

for all $f \in C^1(\overline{U})$. For $1 < p < \infty$, we apply estimate $(\star\star\star)$ with $|f|^p$ replacing $|f|$ to obtain

$$\begin{aligned}
\int_{\partial U} |f|^p\, d\mathcal{H}^{n-1} &\leq C \int_U |Df|\,|f|^{p-1} + |f|^p\, dy \\
&\leq C \int_U |Df|^p + |f|^p\, dy \qquad (\star\star\star\star)
\end{aligned}$$

for all $f \in C^1(\overline{U})$.

4. Thus if we define

$$Tf \equiv f\,|_{\partial U}$$

for $f \in C^1(\overline{U})$, we see from $(\star\star\star\star)$ and Theorem 3 in Section 4.2.1 that T uniquely extends to a bounded linear operator from $W^{1,p}(U)$ to $L^p(\partial U; \mathcal{H}^{n-1})$.

Clearly,

$$Tf = f\,|_{\partial U}$$

for all $f \in W^{1,p}(U) \cap C(\bar{U})$. This proves assertion (i); assertion (ii) follows by a routine approximation argument from the Gauss–Green Theorem. ∎

4.4 Extensions

THEOREM 1
Assume U is bounded, ∂U Lipschitz, $1 \le p < \infty$. Let $U \subset\subset V$. There exists a bounded linear operator

$$E : W^{1,p}(U) \to W^{1,p}(\mathbb{R}^n)$$

such that

$$Ef = f \quad on \quad U$$

and

$$spt\,(Ef) \subset V$$

for all $f \in W^{1,p}(U)$.

DEFINITION *Ef is called an* ***extension*** *of f to $\mathbb{R}^n$.*

PROOF
1. First we introduce some notation:

(a) Given $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let us write $x = (x', x_n)$ for $x' = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$, $x_n \in \mathbb{R}$. Similarly, we write $y = (y', y_n)$.

(b) Given $x \in \mathbb{R}^n$, and $r, h > 0$, define the open cylinder

$$C(x, r, h) \equiv \{y \in \mathbb{R}^n \mid |y' - x'| < r, |y_n - x_n| < h\}.$$

Since ∂U is Lipschitz, for each $x \in \partial U$ there exist — upon rotating and relabeling the coordinate axes if necessary — $r, h > 0$ and a Lipschitz function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$\begin{cases} \max\limits_{|x'-y'|<r} |\gamma(y') - x_n| < \dfrac{h}{4}, \\ U \cap C(x, r, h) = \{y \mid |x' - y'| < r, \gamma(y') < y_n < x_n + h\}, \\ C(x, r, h) \subset V. \end{cases}$$

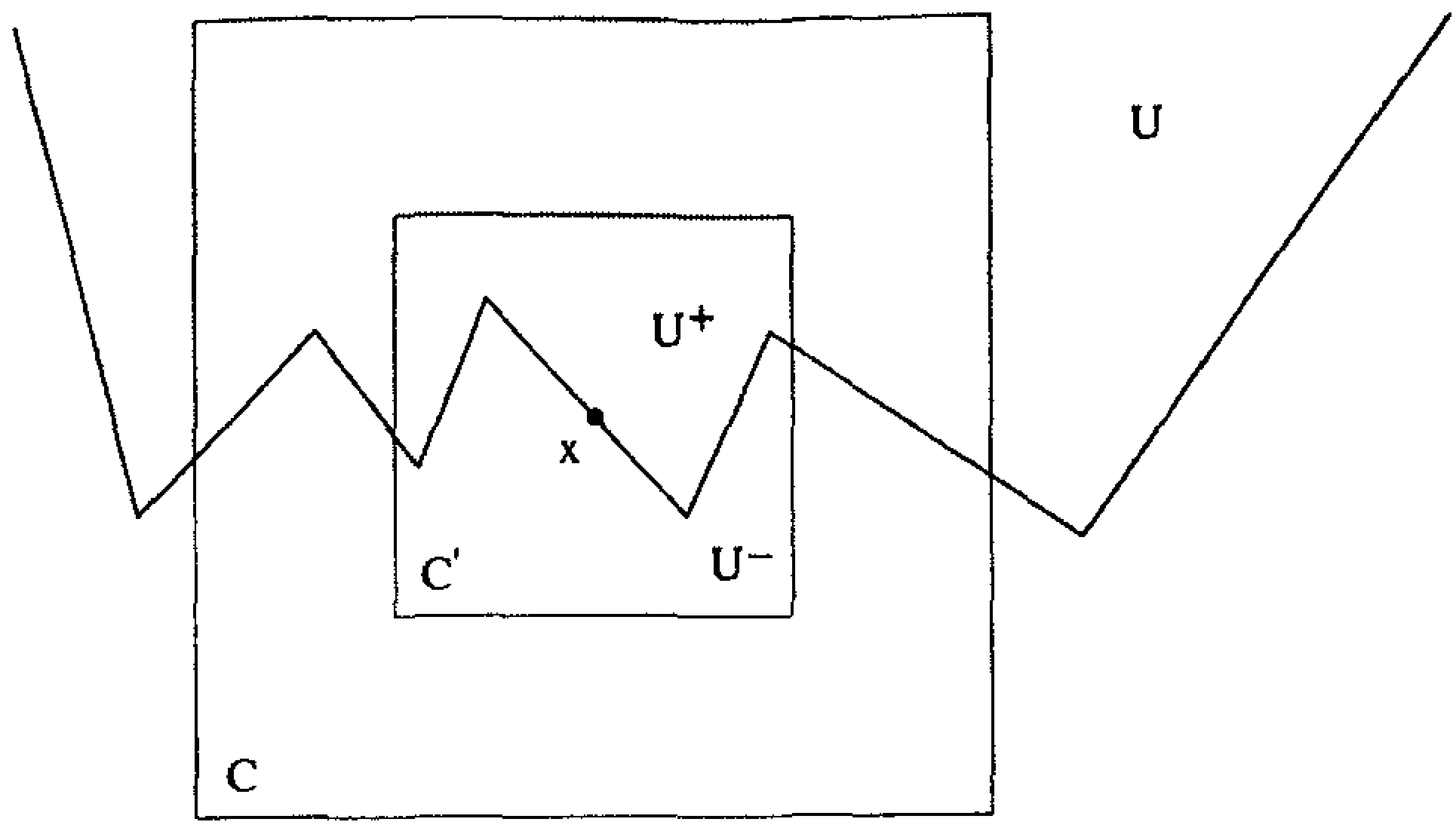

FIGURE 4.3
A region U^+ above, and a region U^- below, a Lipschitz boundary.

2. Fix $x \in \partial U$ and with r, h, γ as above, write

$$\begin{aligned} C &\equiv C(x,r,h), \quad & C' &\equiv C(x,r/2,h/2) \\ U^+ &\equiv C' \cap U, \quad & U^- &\equiv C' - \overline{U}. \end{aligned}$$

3. Let $f \in C^1(\overline{U})$ and suppose for the moment spt $(f) \subset C' \cap \overline{U}$. Set

$$\begin{aligned} f^+(y) &= f(y) \text{ if } y \in \overline{U^+}, \\ f^-(y) &= f(y', 2\gamma(y') - y_n) \text{ if } y \in \overline{U^-}. \end{aligned}$$

Note $f^- = f^+ = f$ on $\partial U \cap C'$.

4. *Claim #1*: $||f^-||_{W^{1,p}(U^-)} \leq C||f||_{W^{1,p}(U)}$.

Proof of Claim #1: Let $\varphi \in C_c^1(U^-)$ and let $\{\gamma_k\}_{k=1}^\infty$ be a sequence of C^∞ functions such that

$$\begin{cases} \gamma_k \geq \gamma & \\ \gamma_k \to \gamma & \text{uniformly} \\ D\gamma_k \to D\gamma & \mathcal{L}^{n-1} \text{ a.e.,} \\ \sup_k ||D\gamma_k||_{L^\infty} < \infty. & \end{cases}$$

Then, for $1 \leq i \leq n-1$,

$$\begin{aligned}
&\int_{U^-} f^- \frac{\partial \varphi}{\partial y_i}\, dy \\
&= \int_{U^-} f(y', 2\gamma(y') - y_n) \frac{\partial \varphi}{\partial y_i}\, dy \\
&= \lim_{k\to\infty} \int_{U^-} f(y', 2\gamma_k(y') - y_n) \frac{\partial \varphi}{\partial y_i}\, dy \\
&= -\lim_{k\to\infty} \int_{U^-} \left(\frac{\partial f}{\partial y_i}(y', 2\gamma_k(y') - y_n) + 2\frac{\partial f}{\partial y_n}(y', 2\gamma_k(y') - y_n) \frac{\partial \gamma_k}{\partial y_i}(y') \right) \varphi\, dy \\
&= -\int_{U^-} \left(\frac{\partial f}{\partial y_i}(y', 2\gamma(y') - y_n) + 2\frac{\partial f}{\partial y_n}(y', 2\gamma(y') - y_n) \frac{\partial \gamma}{\partial y_i}(y') \right) \varphi\, dy.
\end{aligned}$$

Similarly,

$$\int_{U^-} f^- \frac{\partial \varphi}{\partial y_n}\, dy = \int_{U^-} \frac{\partial f}{\partial y_n}(y', 2\gamma(y') - y_n) \varphi\, dy.$$

Now recall

$$\|D\gamma\|_{L^\infty} < \infty,$$

and thus

$$\int_{U^-} |Df(y', 2\gamma(y') - y_n)|^p\, dy \leq C \int_U |Df|^p\, dy < \infty$$

by the change of variables formula (Theorem 2 in Section 3.3.3).

5. Define

$$Ef \equiv \bar{f} \equiv \begin{cases} f^+ & \text{on } \overline{U}^+ \\ f^- & \text{on } \overline{U}^- \\ 0 & \text{on } \mathbb{R}^n - (\overline{U}^+ \cup \overline{U}^-), \end{cases}$$

and note $\bar{f}$ is continuous on $\mathbb{R}^n$.

6. *Claim #2:* $E(f) \in W^{1,p}(\mathbb{R}^n)$, spt $(E(f)) \subset C' \subset V$, and

$$\|E(f)\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|f\|_{W^{1,p}(U)}.$$

Proof of Claim #2: Let $\varphi \in C^1_c(C')$. For $1 \leq i \leq n$

$$\begin{aligned}
\int_{C'} \bar{f} \frac{\partial \varphi}{\partial y_i}\, dy &= \int_{U^+} f^+ \frac{\partial \varphi}{\partial y_i}\, dy + \int_{U^-} f^- \frac{\partial \varphi}{\partial y_i}\, dy \\
&= -\int_{U^+} \frac{\partial f^+}{\partial y_i} \varphi\, dy - \int_{U^-} \frac{\partial f^-}{\partial y_i} \varphi\, dy \\
&\quad + \int_{\partial U} (T(f^+) - T(f^-)) \varphi \nu_i\, d\mathcal{H}^{n-1}
\end{aligned}$$

by Theorem 1 in Section 4.3. But $T(f^+) = T(f^-) = f\,|_{\partial U}$, and so the last term vanishes.

This calculation and Claim #1 complete the proof in case f is C^1, with support in $C' \cap \overline{U}$.

7. Now assume $f \in C^1(\overline{U})$, but drop the restriction on its support. Since ∂U is compact, we can cover ∂U with finitely many cylinders $C_k = C(x_k, r_k, h_k)$ $(k = 1, \dots, N)$ for which assertions analogous to the foregoing hold. Let $\{\zeta_k\}_{k=0}^N$ be a partition of unity as in the proof of Theorem 3 in Section 4.2.1, define $E(\zeta_k f)$ $(k = 1, 2, \dots, N)$ as above and set

$$Ef \equiv \sum_{k=1}^{N} E(\zeta_k f) + \zeta_0 f.$$

8. Finally, if $f \in W^{1,p}(U)$, we approximate f by functions $f_k \in W^{1,p}(U) \cap C^1(\overline{U})$ and set

$$Ef \equiv \lim_{k \to \infty} Ef_k. \quad \blacksquare$$

4.5 Sobolev inequalities

4.5.1 Gagliardo–Nirenberg–Sobolev inequality

We prove next that if $f \in W^{1,p}(\mathbb{R}^n)$ for some $1 \le p < n$, then in fact f lies in $L^{p^*}(\mathbb{R}^n)$.

DEFINITION *For $1 \le p < n$, define*

$$p^* \equiv \frac{np}{n-p};$$

p^* is called the ***Sobolev conjugate*** *of p. Note $1/p^* = 1/p - 1/n$.*

THEOREM 1 GAGLIARDO–NIRENBERG–SOBOLEV INEQUALITY
Assume $1 \le p < n$. There exists a constant C_1, depending only on p and n, such that

$$\left(\int_{\mathbb{R}^n} |f|^{p^*} dx \right)^{1/p^*} \le C_1 \left(\int_{\mathbb{R}^n} |Df|^p \, dx \right)^{1/p}$$

for all $f \in W^{1,p}(\mathbb{R}^n)$.

PROOF

1. According to Theorem 2 in Section 4.2.1, we may as well assume $f \in C^1_c(\mathbb{R}^n)$. Then for $1 \leq i \leq n$,

$$f(x_1, \dots, x_i, \dots, x_n) = \int_{-\infty}^{x_i} \frac{\partial f}{\partial x_i}(x_1, \dots, t_i, \dots, x_n)\, dt_i$$

and so

$$|f(x)| \leq \int_{-\infty}^{\infty} |Df(x_1, \dots, t_i, \dots, x_n)|\, dt_i \qquad (1 \leq i \leq n).$$

Thus

$$|f(x)|^{\frac{n}{n-1}} \leq \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |Df(x_1, \dots, t_i, \dots, x_n)|\, dt_i \right)^{\frac{1}{n-1}}.$$

Integrate with respect to x_1:

$$\begin{aligned}
\int_{-\infty}^{\infty} |f|^{1^*}\, dx_1 &\leq \left(\int_{-\infty}^{\infty} |Df|\, dt_1 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |Df|\, dt_i \right)^{\frac{1}{n-1}} dx_1 \\
&\leq \left(\int_{-\infty}^{\infty} |Df|\, dt_1 \right)^{\frac{1}{n-1}} \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df|\, dx_1\, dt_i \right)^{\frac{1}{n-1}}.
\end{aligned}$$

Next integrate with respect to x_2 to find

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f|^{1^*}\, dx_1\, dx_2 \\
&\qquad \leq \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df|\, dx_1\, dt_2 \right)^{\frac{1}{n-1}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df|\, dt_1\, dx_2 \right)^{\frac{1}{n-1}} \\
&\qquad \times \prod_{i=3}^{n} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Df|\, dx_1\, dx_2\, dt_i \right)^{\frac{1}{n-1}}.
\end{aligned}$$

We continue and eventually discover

$$\begin{aligned}
\int_{\mathbb{R}^n} |f|^{1^*}\, dx &\leq \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |Df|\, dx_1 \dots dt_i \dots dx_n \right)^{\frac{1}{n-1}} \\
&= \left(\int_{\mathbb{R}^n} |Df|\, dx \right)^{\frac{n}{n-1}}.
\end{aligned}$$

This immediately gives

$$\left(\int_{\mathbb{R}^n} |f|^{1^*}\, dx\right)^{\frac{1}{1^*}} \le \int_{\mathbb{R}^n} |Df|\, dx, \tag{$\star$}$$

and so proves the theorem for $p = 1$.

2. If $1 < p < n$, set $g = |f|^\gamma$ with $\gamma > 0$ as selected below. Applying ($\star$) to g we find

$$\begin{aligned}\left(\int_{\mathbb{R}^n} |f|^{\frac{\gamma n}{n-1}}\, dx\right)^{\frac{n-1}{n}} &\le \gamma \int_{\mathbb{R}^n} |f|^{\gamma-1}|Df|\, dx \\ &\le \gamma \left(\int_{\mathbb{R}^n} |f|^{\frac{(\gamma-1)p}{p-1}}\, dx\right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^n} |Df|^p\, dx\right)^{\frac{1}{p}}.\end{aligned}$$

Choose γ so that

$$\frac{\gamma n}{n-1} = (\gamma - 1)\frac{p}{p-1}.$$

Then

$$\frac{\gamma n}{n-1} = (\gamma - 1)\frac{p}{p-1} = \frac{np}{n-p} = p^\star.$$

Thus

$$\left(\int_{\mathbb{R}^n} |f|^{p^*}\, dx\right)^{\frac{n-1}{n}} \le C\left(\int_{\mathbb{R}^n} |f|^{p^*}\, dx\right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^n} |Df|^p\, dx\right)^{\frac{1}{p}}$$

and so

$$\left(\int_{\mathbb{R}^n} |f|^{p^*}\, dx\right)^{\frac{1}{p^*}} \le C\left(\int_{\mathbb{R}^n} |Df|^p\, dx\right)^{\frac{1}{p}}$$

where C depends only on n and p. ∎

4.5.2 Poincaré's inequality on balls

We next derive a local version of the preceding inequality.

LEMMA 1
For each $1 \le p < \infty$ there exists a constant C, depending only on n and p, such that

$$\int_{B(x,r)} |f(y) - f(z)|^p\, dy \le Cr^{n+p-1} \int_{B(x,r)} |Df(y)|^p |y - z|^{1-n}\, dy$$

for all $B(x,r) \subset \mathbb{R}^n$, $f \in C^1(B(x,r))$, and $z \in B(x.r)$.

PROOF

If $y, z \in B(x,r)$, then

$$f(y) - f(z) = \int_0^1 \frac{d}{dt} f(z + t(y-z))\, dt = \int_0^1 Df(z + t(y-z))\, dt \cdot (y - z),$$

and so

$$|f(y) - f(z)|^p \le |y - z|^p \int_0^1 |Df(z + t(y-z))|^p\, dt.$$

Thus, for $s > 0$,

$$\begin{aligned}
&\int_{B(x,r)\cap\partial B(z,s)} |f(y) - f(z)|^p \, d\mathcal{H}^{n-1}(y) \\
&\quad \le s^p \int_0^1 \int_{B(x,r)\cap\partial B(z,s)} |Df(z + t(y-z))|^p \, d\mathcal{H}^{n-1}(y)\, dt \\
&\quad \le s^p \int_0^1 \frac{1}{t^{n-1}} \int_{B(x,r)\cap\partial B(z,ts)} |Df(w)|^p \, d\mathcal{H}^{n-1}(w)\, dt \\
&\quad = s^{n+p-1} \int_0^1 \int_{B(x,r)\cap\partial B(z,ts)} |Df(w)|^p |w - z|^{1-n} \, d\mathcal{H}^{n-1}(w)\, dt \\
&\quad = s^{n+p-2} \int_{B(x,r)\cap B(z,s)} |Df(w)|^p |w - z|^{1-n} \, dw.
\end{aligned}$$

Hence Proposition 1 in Section 3.4.4 implies

$$\int_{B(x,r)} |f(y) - f(z)|^p \, dy \le C r^{n+p-1} \int_{B(x,r)} |Df(w)|^p |w - z|^{1-n} \, dw. \quad \blacksquare$$

THEOREM 2 POINCARÉ'S INEQUALITY

For each $1 \le p < n$ there exists a constant C_2, depending only on p and n, such that

$$\left(⨍_{B(x,r)} |f - (f)_{x,r}|^{p^*} \, dy \right)^{1/p^*} \le C_2 r \left(⨍_{B(x,r)} |Df|^p \, dy \right)^{1/p}$$

for all $B(x,r) \subset \mathbb{R}^n$, $f \in W^{1,p}(U(x,r))$.

Recall $(f)_{x,r} = ⨍_{B(x,r)} f\, dy$.

PROOF

1. In view of Theorem 2 in Section 4.2.1 we may assume $f \in C^1(B(x,r))$. We recall Lemma 1 to compute

$$\begin{aligned}
\fint_{B(x,r)} |f - (f)_{x,r}|^p \, dy &= \fint_{B(x,r)} \left| \fint_{B(x,r)} f(y) - f(z) \, dz \right|^p dy \\
&\le \fint_{B(x,r)} \fint_{B(x,r)} |f(y) - f(z)|^p \, dz \, dy \\
&\le C \fint_{B(x,r)} r^{p-1} \int_{B(x,r)} |Df(z)|^p |y - z|^{1-n} \, dz \, dy \\
&\le C r^p \fint_{B(x,r)} |Df|^p \, dz. \qquad (\star)
\end{aligned}$$

2. Claim: There exists a constant $C = C(n,p)$ such that

$$\left(\fint_{B(x,r)} |g|^{p^*} dy \right)^{\frac{1}{p^*}} \le C \left(r^p \fint_{B(x,r)} |Dg|^p \, dy + \fint_{B(x,r)} |g|^p \, dy \right)^{\frac{1}{p}}$$

for all $g \in W^{1,p}(U(x,r))$.

Proof of Claim: First observe that, upon replacing $g(y)$ by $(1/r)g(ry)$ if necessary, we may assume $r = 1$. Similarly we may suppose $x = 0$. We next employ Theorem 1 in Section 4.4 to extend g to $\bar{g} \in W^{1,p}(\mathbb{R}^n)$ satisfying

$$||\bar{g}||_{W^{1,p}(\mathbb{R}^n)} \le C||g||_{W^{1,p}(U(0,1))}. \qquad (\star\star)$$

Then Theorem 1 implies

$$\begin{aligned}
\left(\int_{B(0,1)} |g|^{p^*} dy \right)^{\frac{1}{p^*}} &\le \left(\int_{\mathbb{R}^n} |\bar{g}|^{p^*} dy \right)^{\frac{1}{p^*}} \\
&\le C_1 \left(\int_{\mathbb{R}^n} |D\bar{g}|^p \, dy \right)^{\frac{1}{p}} \\
&\le C \left(\int_{B(0,1)} |Dg|^p + |g|^p \, dy \right)^{\frac{1}{p}},
\end{aligned}$$

according to $(\star\star)$.

3. We use $(\star)$ and the claim with $g \equiv f - (f)_{x,r}$ to complete the proof of the theorem. ▌

4.5.3 Morrey's inequality

DEFINITION *Let $0 < \alpha < 1$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is* ***Hölder continuous with exponent*** α *provided*

$$\sup_{\substack{x,y\in\mathbb{R}^n \\ x\neq y}} \frac{|f(x) - f(y)|}{|x - y|^\alpha} < \infty.$$

THEOREM 3 MORREY'S INEQUALITY

(i) For each $n < p < \infty$ there exists a constant C_3, depending only on p and n, such that

$$|f(y) - f(z)| \leq C_3 r \left(⨍_{B(x,r)} |Df|^p \, dw \right)^{1/p}$$

for all $B(x,r) \subset \mathbb{R}^n$, $f \in W^{1,p}(U(x,r))$, and $\mathcal{L}^n$ a.e. $y, z \in U(x,r)$.

(ii) In particular, if $f \in W^{1,p}(\mathbb{R}^n)$, then the limit

$$\lim_{r\to 0} (f)_{x,r} \equiv f^\star(x)$$

exists for all $x \in \mathbb{R}^n$, and $f^\star$ is Hölder continuous with exponent $1 - n/p$.

REMARK See Section 4.2.3 for the case $p = \infty$. ∎

PROOF

1. First assume f is C^1 and recall Lemma 1 with $p = 1$ to calculate

$$\begin{aligned}
&|f(y) - f(z)| \\
&\quad \leq ⨍_{B(x,r)} |f(y) - f(w)| + |f(w) - f(z)| \, dw \\
&\quad \leq C \int_{B(x,r)} |Df(w)|(|y-w|^{1-n} + |z-w|^{1-n}) \, dw \\
&\quad \leq C \left(\int_{B(x,r)} (|y-w|^{1-n} + |z-w|^{1-n})^{\frac{p}{p-1}} \, dw \right)^{\frac{p-1}{p}} \left(\int_{B(x,r)} |Df|^p \, dw \right)^{\frac{1}{p}} \\
&\quad \leq C r^{\left(n-(n-1)\frac{p}{p-1}\right)\frac{p-1}{p}} \left(\int_{B(x,r)} |Df|^p \, dw \right)^{\frac{1}{p}} \\
&\quad = C r^{1-\frac{n}{p}} \left(\int_{B(x,r)} |Df|^p \, dw \right)^{\frac{1}{p}}.
\end{aligned}$$

2. By approximation, we see that if $f \in W^{1,p}(U(x,r))$, the same estimate holds for $\mathcal{L}^n$ a.e. $y, z \in U(x,r)$. This proves (i).

3. Now suppose $f \in W^{1,p}(\mathbb{R}^n)$. Then for $\mathcal{L}^n$ a.e. x, y we can apply the estimate of (i) with $r = |x-y|$ to obtain

$$\begin{aligned}|f(y) - f(x)| &\leq C|x-y|^{1-\frac{n}{p}} \left(\int_{B(x,r)} |Df|^p \, dw \right)^{\frac{1}{p}} \\ &\leq C||Df||_{L^p(\mathbb{R}^n)} |x-y|^{1-\frac{n}{p}}.\end{aligned}$$

Thus f is equal $\mathcal{L}^n$ a.e. to a Hölder-continuous function $\bar{f}$. Clearly $f^* = \bar{f}$ everywhere in $\mathbb{R}^n$. ∎

4.6 Compactness

THEOREM 1

Assume U is bounded, ∂U is Lipschitz, $1 < p < n$. Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence in $W^{1,p}(U)$ satisfying

$$\sup_k ||f_k||_{W^{1,p}(U)} < \infty.$$

Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ and a function $f \in W^{1,p}(U)$ such that

$$f_{k_j} \to f \ \ in \ \ L^q(U).$$

for each $1 \leq q < p^*$.

PROOF

1. Fix a bounded open set V such that $U \subset\subset V$ and extend each f_k to $\bar{f}_k \in W^{1,p}(\mathbb{R}^n)$, spt $(\bar{f}_k) \subset V$, with

$$\sup_k ||\bar{f}_k||_{W^{1,p}(\mathbb{R}^n)} \leq C \sup_k ||f_k||_{W^{1,p}(U)} < \infty. \tag{$\star$}$$

2. Let $\bar{f}_k^\epsilon \equiv \eta_\epsilon * \bar{f}_k$ be the usual mollification, as described in Section 4.2.1.

3. *Claim #1*: $||\bar{f}_k^\epsilon - \bar{f}_k||_{L^p(\mathbb{R}^n)} \leq C\epsilon$, uniformly in k.

Proof of Claim #1: First suppose the functions $\bar{f}_k$ are smooth, and calculate

$$\begin{aligned}|\bar{f}_k^\epsilon(x) - \bar{f}_k(x)| &\leq \int_{B(0,1)} \eta(z)|\bar{f}_k(x - \epsilon z) - \bar{f}_k(x)|\, dz \\ &= \int_{B(0,1)} \eta(z)\left|\int_0^1 \frac{d}{dt}\bar{f}_k(x - t\epsilon z)\, dt\right| dz \\ &\leq \epsilon \int_{B(0,1)} \eta(z) \int_0^1 |D\bar{f}_k(x - \epsilon t z)|\, dt\, dz.\end{aligned}$$

Thus

$$\begin{aligned}\|\bar{f}_k^\epsilon - \bar{f}_k\|^p_{L^p(\mathbb{R}^n)} &\leq C\epsilon^p \int_{B(0,1)} \eta(z) \int_0^1 \left(\int_{\mathbb{R}^n} |D\bar{f}_k(x - \epsilon t z)|^p\, dx\right) dt\, dz \\ &\leq C\epsilon^p \|\bar{f}_k\|^p_{W^{1,p}(\mathbb{R}^n)} \\ &\leq C\epsilon^p \qquad \text{by } (\star).\end{aligned}$$

The general case follows by approximation.

4. *Claim #2*: For each $\epsilon > 0$, the sequence $\{\bar{f}_k^\epsilon\}_{k=1}^\infty$ is bounded and equicontinuous on $\mathbb{R}^n$.

Proof of Claim #2: We calculate

$$\begin{aligned}|\bar{f}_k^\epsilon(x)| &\leq \int_{B(x,\epsilon)} \eta_\epsilon(x - y)|\bar{f}_k(y)|\, dy \\ &\leq C\epsilon^{-n}\|\bar{f}_k\|_{L^p(\mathbb{R}^n)} \\ &\leq C\epsilon^{-n}\end{aligned}$$

and

$$\begin{aligned}|D\bar{f}_k^\epsilon(x)| &\leq \int_{B(x,\epsilon)} |D\eta_\epsilon(x - y)||\bar{f}_k(y)|\, dy \\ &\leq C\epsilon^{-n-1}.\end{aligned}$$

5. *Claim #3*: For each $\delta > 0$ there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty \subset \{f_k\}_{k=1}^\infty$ such that

$$\limsup_{i,j\to\infty} \|f_{k_i} - f_{k_j}\|_{L^p(U)} \leq \delta.$$

Proof of Claim #3: Recalling Claim #1, we choose $\epsilon > 0$ so small that

$$\sup_k \|\bar{f}_k^\epsilon - \bar{f}_k\|_{L^p(\mathbb{R}^n)} \leq \frac{\delta}{3}\,.$$

Next we use Claim #2 and the Arzela–Ascoli Theorem to find a subsequence $\{\bar{f}^{\epsilon}_{k_j}\}_{j=1}^{\infty}$ which converges uniformly on $\mathbb{R}^n$. Then

$$\begin{aligned}
&||f_{k_j} - f_{k_i}||_{L^p(U)} \\
&\quad \leq ||\bar{f}_{k_j} - \bar{f}_{k_i}||_{L^p(\mathbb{R}^n)} \\
&\quad \leq ||\bar{f}_{k_j} - \bar{f}^{\epsilon}_{k_j}||_{L^p(\mathbb{R}^n)} + ||\bar{f}^{\epsilon}_{k_j} - \bar{f}^{\epsilon}_{k_i}||_{L^p(\mathbb{R}^n)} + ||\bar{f}^{\epsilon}_{k_i} - \bar{f}_{k_i}||_{L^p(\mathbb{R}^n)} \\
&\quad \leq \frac{2\delta}{3} + ||\bar{f}^{\epsilon}_{k_j} - \bar{f}^{\epsilon}_{k_i}||_{L^p(\mathbb{R}^n)} \\
&\quad \leq \delta
\end{aligned}$$

for i, j large enough.

6. We use a diagonal argument and Claim #3 with $\delta = 1, 1/2, 1/4$, etc. to obtain a subsequence, also denoted $\{f_{k_j}\}_{j=1}^{\infty}$, converging to f in $L^p(U)$. We observe also for $1 \leq q < p^\star$,

$$||f_{k_j} - f||_{L^q(U)} \leq ||f_{k_j} - f||^{\theta}_{L^p(U)} ||f_{k_j} - f||^{1-\theta}_{L^{p^\star}(U)},$$

where $1/q = \theta/p + (1-\theta)/p^\star$ and hence $\theta > 0$. Since $\{f_k\}_{k=1}^{\infty}$ is bounded in $L^{p^\star}(U)$, we see

$$\lim_{j \to \infty} ||f_{k_j} - f||_{L^q(U)} = 0$$

for each $1 \leq q < p^\star$. Since $p > 1$, it follows from Theorem 3 in Section 1.9 that $f \in W^{1,p}(U)$. ∎

REMARK The compactness assertion is false for $q = p^\star$. In case $p = 1$, the above argument shows that there is a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ and $f \in L^{1^\star}(U)$ such that

$$\lim_{j \to \infty} ||f_{k_j} - f||_{L^q(U)} = 0$$

for each $1 \leq q < 1^\star$. It follows from Theorem 1 in Section 5.2 that $f \in BV(U)$. ∎

4.7 Capacity

We next introduce *capacity* as a way to study certain "small" subsets of $\mathbb{R}^n$. We will later see that in fact capacity is precisely suited for characterizing the fine properties of Sobolev functions. For this section, fix $1 \leq p < n$.

4.7.1 Definitions and elementary properties

DEFINITION $K^p \equiv \{f : \mathbb{R}^n \to \mathbb{R} \mid f \geq 0, f \in L^{p^*}(\mathbb{R}^n), Df \in L^p(\mathbb{R}^n; \mathbb{R}^n)\}$.

DEFINITION *If* $A \subset \mathbb{R}^n$, *set*

$$\mathrm{Cap_p}(A) \equiv \inf\left\{\int_{\mathbb{R}^n} |Df|^p\, dx \mid f \in K^p, A \subset \{f \geq 1\}^o\right\}.$$

We call $\mathrm{Cap_p}(A)$ *the* ***p-capacity*** *of* A.

REMARK

(i) Note carefully the requirement that A lie in the *interior* of the set $\{f \geq 1\}$.

(ii) Using regularization, we see

$$\mathrm{Cap_p}(K) = \inf\left\{\int_{\mathbb{R}^n} |Df|^p\, dx \mid f \in C_c^\infty(\mathbb{R}^n), f \geq \chi_K\right\}$$

for each compact set $K \subset \mathbb{R}^n$.

(iii) Clearly, $A \subset B$ implies

$$\mathrm{Cap_p}(A) \leq \mathrm{Cap_p}(B). \quad \blacksquare$$

LEMMA 1

(i) *If* $f \in K^p$ *for some* $1 \leq p < n$, *there exists a sequence* $\{f_k\}_{k=1}^\infty \subset W^{1,p}(\mathbb{R}^n)$ *such that*

$$||f - f_k||_{L^{p^*}(\mathbb{R}^n)} \to 0$$

and

$$||Df - Df_k||_{L^p(\mathbb{R}^n)} \to 0$$

as $k \to \infty$.

(ii) *If* $f \in K^p$, *then*

$$||f||_{L^{p^*}(\mathbb{R}^n)} \leq C_1 ||Df||_{L^p(\mathbb{R}^n)},$$

where C_1 *is the constant from Theorem 1 in Section 4.5.1.*

PROOF Select $\zeta \in C_c^1(\mathbb{R}^n)$ so that

$$\begin{cases} 0 \le \zeta \le 1, & \zeta \equiv 1 \text{ on } B(0,1) \\ \operatorname{spt}(\zeta) \subset B(0,2), & |D\zeta| \le 2. \end{cases}$$

For each $k = 1, 2, \ldots$, set $\zeta_k(x) \equiv \zeta(x/k)$.

Given $f \in K^p$, write $f_k \equiv f\zeta_k$. Then $f_k \in W^{1,p}(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} |f - f_k|^{p^*} \, dy \le \int_{\mathbb{R}^n - B(0,k)} |f|^{p^*} \, dy,$$

and

$$\begin{aligned}
&\int_{\mathbb{R}^n} |Df - Df_k|^p \, dy \\
&\quad \le 2^{p-1} \left\{ \int_{\mathbb{R}^n} |(1 - \zeta_k) Df|^p + |f D\zeta_k|^p \, dy \right\} \\
&\quad \le 2^{p-1} \left\{ \int_{\mathbb{R}^n - B(0,k)} |Df|^p \, dy + \frac{2^p}{k^p} \int_{B(0,2k) - B(0,k)} |f|^p \, dy \right\} \\
&\quad \le C \left\{ \int_{\mathbb{R}^n - B(0,k)} |Df|^p \, dy + 4^p \left(\int_{\mathbb{R}^n - B(0,k)} |f|^{p^*} \, dy \right)^{1 - \frac{p}{n}} \right\}.
\end{aligned}$$

This proves assertion (i). Assertion (ii) follows from (i) and Theorem 1 in Section 4.5.1. ∎

LEMMA 2

(i) *Assume $f, g \in K^p$. Then*

$$h \equiv \max\{f, g\} \in K^p$$

and

$$Dh = \begin{cases} Df & \mathcal{L}^n \text{ a.e. on } \{f \ge g\} \\ Dg & \mathcal{L}^n \text{ a.e. on } \{f \le g\}. \end{cases}$$

An analogous assertion holds for $\min\{f, g\}$.

(ii) *If $f \in K^p$ and $t \ge 0$,*

$$h \equiv \min\{f, t\} \in K^p.$$

(iii) *Given a sequence $\{f_k\}_{k=1}^\infty \subset K^p$, define*

$$g \equiv \sup_{1 \le k < \infty} f_k$$

and

$$h \equiv \sup_{1 \le k < \infty} |Df_k|.$$

If $h \in L^p(\mathbb{R}^n)$, then $g \in K^p$ and $|Dg| \le h$ $\mathcal{L}^n$ a.e.

PROOF

1. To prove (i) we note

$$h = \max\{f, g\} = f + (g - f)^+.$$

Hence Theorem 4 in Section 4.2.2 implies

$$Dh = \begin{cases} Df & \text{a.e. on } \{f \ge g\} \\ Dg & \text{a.e. on } \{f \le g\} \end{cases}$$

Thus $Dh \in L^p(\mathbb{R}^n)$. Since $0 \le h \le f + g$, we have $h \in L^{p^*}(\mathbb{R}^n)$ as well.

2. The proof of (ii) is similar; we need only observe

$$0 \le h = \min\{f, t\} \le f,$$

and so $h \in L^{p^*}(\mathbb{R}^n)$.

3. To prove (iii) let us set

$$g_l \equiv \sup_{1 \le k \le l} f_k.$$

Using assertion (i) we see $g_l \in K^p$ and

$$|Dg_l| \le \sup_{1 \le k \le l} |Df_k| \le h.$$

Since $g_l \to g$ monotonically, we have

$$\begin{aligned} ||g||_{L^{p^*}(\mathbb{R}^n)} &= \lim_{l \to \infty} ||g_l||_{L^{p^*}(\mathbb{R}^n)} \\ &\le C_1 \liminf_{l \to \infty} ||Dg_l||_{L^p(\mathbb{R}^n)} \qquad \text{by Lemma 1} \\ &\le C_1 ||h||_{L^p(\mathbb{R}^n)}. \end{aligned}$$

Thus $g \in L^{p^*}(\mathbb{R}^n)$. Now, for each $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$,

$$\begin{aligned} \int_{\mathbb{R}^n} g \operatorname{div} \varphi \, dy &= \lim_{l \to \infty} \int_{\mathbb{R}^n} g_l \operatorname{div} \varphi \, dy \\ &= - \lim_{l \to \infty} \int_{\mathbb{R}^n} \varphi \cdot Dg_l \, dy \\ &\le \int_{\mathbb{R}^n} |\varphi| h \, dy. \end{aligned}$$

It follows that the linear functional L defined by

$$L(\varphi) \equiv \int_{\mathbb{R}^n} g \operatorname{div} \varphi \, dy \qquad (\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n))$$

has a unique extension $\bar{L}$ to $C_c(\mathbb{R}^n; \mathbb{R}^n)$ such that

$$\bar{L}(\varphi) \leq \int_{\mathbb{R}^n} |\varphi| h \, dy,$$

for $\varphi \in C_c(\mathbb{R}^n; \mathbb{R}^n)$. We apply Theorem 1 in Section 1.8 and note the measure μ constructed there satisfies

$$\mu(A) \leq \int_A h \, dy$$

for any Lebesgue measurable set $A \subset \mathbb{R}^n$. It follows that

$$\bar{L}(\varphi) = \int_{\mathbb{R}^n} \varphi \cdot k \, dy$$

where $k \in L^p(\mathbb{R}^n; \mathbb{R}^n)$ and $|k| \leq h$ $\mathcal{L}^n$ a.e. Thus $g \in K^p$ and $|Dg| = |k| \leq h$ $\mathcal{L}^n$ a.e. ∎

THEOREM 1
Cap_p *is a measure on* $\mathbb{R}^n$.

Warning: Cap_p is *not* a Borel measure. In fact, if $A \subset \mathbb{R}^n$ and $0 < \mathrm{Cap}_p(A) < \infty$, then A is *not* Cap_p-measurable. Remember also that what we call a measure in these notes is usually called an "outer measure" in other texts.

PROOF Assume $A \subset \bigcup_{k=1}^{\infty} A_k$, $\sum_{k=1}^{\infty} \mathrm{Cap}_p(A_k) < \infty$. Fix $\epsilon > 0$. For each $k = 1, \ldots,$ choose $f_k \in K^p$ so that

$$A_k \subset \{f_k \geq 1\}^o$$

and

$$\int_{\mathbb{R}^n} |Df_k|^p \, dx \leq \mathrm{Cap}_p(A_k) + \frac{\epsilon}{2^k} \, .$$

Define $g \equiv \sup_{1 \leq k < \infty} f_k$. Then $A \subset \{g \geq 1\}^o$, $g \in K^p$ by Lemma 2, and

$$\begin{aligned}
\int_{\mathbb{R}^n} |Dg|^p \, dx &\leq \int_{\mathbb{R}^n} \sup_{1 \leq k < \infty} |Df_k|^p \, dx \\
&\leq \sum_{k=1}^{\infty} \int_{\mathbb{R}^n} |Df_k|^p \, dx \\
&\leq \sum_{k=1}^{\infty} \mathrm{Cap}_p(A_k) + \epsilon.
\end{aligned}$$

Thus

$$\mathrm{Cap}_{\mathrm{p}}(A) \leq \sum_{k=1}^{\infty} \mathrm{Cap}_{\mathrm{p}}(A_k) + \epsilon. \quad \blacksquare$$

THEOREM 2 PROPERTIES OF CAPACITY
Assume $A, B \subset \mathbb{R}^n$.

(i) $\mathrm{Cap}_{\mathrm{p}}(A) = \inf\{\mathrm{Cap}_{\mathrm{p}}(U) \mid U$ *open*, $A \subset U\}$.

(ii) $\mathrm{Cap}_{\mathrm{p}}(\lambda A) = \lambda^{n-p}\mathrm{Cap}_{\mathrm{p}}(A) \qquad (\lambda > 0)$.

(iii) $\mathrm{Cap}_{\mathrm{p}}(L(A)) = \mathrm{Cap}_{\mathrm{p}}(A)$ *for each affine isometry* $L : \mathbb{R}^n \to \mathbb{R}^n$.

(iv) $\mathrm{Cap}_{\mathrm{p}}(B(x,r)) = r^{n-p}\mathrm{Cap}_{\mathrm{p}}(B(0,1))$.

(v) $\mathrm{Cap}_{\mathrm{p}}(A) \leq C\mathcal{H}^{n-p}(A)$ *for some constant C depending only on p and n.*

(vi) $\mathcal{L}^n(A) \leq C\mathrm{Cap}_{\mathrm{p}}(A)^{n/n-p}$ *for some constant C depending only on p and n.*

(vii) $\mathrm{Cap}_{\mathrm{p}}(A \cup B) + \mathrm{Cap}_{\mathrm{p}}(A \cap B) \leq \mathrm{Cap}_{\mathrm{p}}(A) + \mathrm{Cap}_{\mathrm{p}}(B)$.

(viii) *If $A_1 \subset \dots A_k \subset A_{k+1} \dots$, then*

$$\lim_{k\to\infty} \mathrm{Cap}_{\mathrm{p}}(A_k) = \mathrm{Cap}_{\mathrm{p}}\left(\bigcup_{k=1}^{\infty} A_k\right).$$

(ix) *If $A_1 \supset \dots A_k \supset A_{k+1} \dots$ are compact, then*

$$\lim_{k\to\infty} \mathrm{Cap}_{\mathrm{p}}(A_k) = \mathrm{Cap}_{\mathrm{p}}\left(\bigcap_{k=1}^{\infty} A_k\right).$$

REMARK Assertion (ix) may be false if the sets $\{A_k\}_{k=1}^{\infty}$ are not compact. See Theorem 3 in Section 4.7.2 for an improvement of (v). $\blacksquare$

PROOF

1. Clearly $\mathrm{Cap}_{\mathrm{p}}(A) \leq \inf\{\mathrm{Cap}_{\mathrm{p}}(U) \mid U \text{ open}, U \supset A\}$. On the other hand, for each $\epsilon > 0$, there exists $f \in K^p$ such that $A \subset \{f \geq 1\}^o \equiv U$ and

$$\int_{\mathbb{R}^n} |Df|^p \, dx \leq \mathrm{Cap}_{\mathrm{p}}(A) + \epsilon.$$

But then

$$\mathrm{Cap}_{\mathrm{p}}(U) \leq \int_{\mathbb{R}^n} |Df|^p \, dx,$$

and so statement (i) holds.

2. Fix $\epsilon > 0$ and choose $f \in K^p$ as above. Let $g(x) \equiv f(x/\lambda)$. Then $g \in K^p$, $\lambda A \subset \{g \geq 1\}^o$ and

$$\int_{\mathbb{R}^n} |Dg|^p \, dx = \lambda^{n-p} \int_{\mathbb{R}^n} |Df|^p \, dx.$$

Thus $\mathrm{Cap_p}(\lambda A) \leq \lambda^{n-p}(\mathrm{Cap_p}(A) + \epsilon)$. The other inequality is similar, and so (ii) is verified.

3. Assertion (iii) is clear.

4. Statement (iv) is a consequence of (ii), (iii).

5. To prove (v), fix $\delta > 0$ and suppose

$$A \subset \bigcup_{k=1}^{\infty} B(x_k, r_k)$$

where $2r_k < \delta$, $(k = 1, \ldots)$. Then

$$\mathrm{Cap_p}(A) \leq \sum_{k=1}^{\infty} \mathrm{Cap_p}(B(x_k, r_k)) = \mathrm{Cap_p}(B(0,1)) \sum_{k=1}^{\infty} r_k^{n-p}.$$

Hence

$$\mathrm{Cap_p}(A) \leq C\mathcal{H}^{n-p}(A).$$

6. Choose $\epsilon > 0$, $f \in K^p$ as in part 1 of the proof. Then by Lemma 1

$$\begin{aligned} \mathcal{L}^n(A)^{1/p^*} &\leq \left(\int_{\mathbb{R}^n} f^{p^*} \, dx \right)^{1/p^*} \\ &\leq C_1 \left(\int_{\mathbb{R}^n} |Df|^p \, dx \right)^{1/p} \\ &\leq C_1 (\mathrm{Cap_p}(A) + \epsilon)^{1/p}. \end{aligned}$$

Consequently,

$$\mathcal{L}^n(A) \leq C \mathrm{Cap_p}(A)^{p^*/p};$$

this is (vi).

7. Fix $\epsilon > 0$, select $f \in K^p$ as above, and choose also $g \in K^p$ so that

$$B \subset \{g \geq 1\}^o, \qquad \int_{\mathbb{R}^n} |Dg|^p \, dx \leq \mathrm{Cap_p}(B) + \epsilon.$$

Then

$$\max\{f, g\}, \min\{f, g\} \in K^p$$

and

$$|D(\max\{f,g\})|^p + |D(\min\{f,g\})|^p = |Df|^p + |Dg|^p \qquad \mathcal{L}^n \text{ a.e.},$$

according to Lemma 2. Furthermore,

$$A \cup B \subset \{\max\{f,g\} \geq 1\}^o,$$

$$A \cap B \subset \{\min\{f,g\} \geq 1\}^o.$$

Thus

$$\begin{aligned}\mathrm{Cap_p}(A \cup B) + \mathrm{Cap_p}(A \cap B) &\leq \int_{\mathbb{R}^n} |D(\max\{f,g\})|^p + |D(\min\{f,g\})|^p \, dx \\ &= \int_{\mathbb{R}^n} |Df|^p + |Dg|^p \, dx \\ &\leq \mathrm{Cap_p}(A) + \mathrm{Cap_p}(B) + 2\epsilon\end{aligned}$$

and assertion (vii) is proved.

8. We will prove statement (viii) for the case $1 < p < n$ only; see Federer and Ziemer [FZ] for $p = 1$. Assume $\lim_{k\to\infty} \mathrm{Cap_p}(A_k) < \infty$ and $\epsilon > 0$. Then for each $k = 1, 2, \ldots$, choose $f_k \in K^p$ such that

$$A_k \subset \{x \mid f_k(x) \geq 1\}^o$$

and

$$\int_{\mathbb{R}^n} |Df_k|^p \, dx < \mathrm{Cap_p}(A_k) + \frac{\epsilon}{2^k}.$$

Define

$$h_m \equiv \max\{f_k \mid 1 \leq k \leq m\}, \qquad h_0 \equiv 0$$

and notice from Lemma 2 that $h_m = \max(h_{m-1}, f_m) \in K^p$,

$$A_{m-1} \subset \{x \mid \min(h_{m-1}, f_m) \geq 1\}^o.$$

We compute

$$\begin{aligned}\int_{\mathbb{R}^n} |Dh_m|^p \, dx + \mathrm{Cap_p}(A_{m-1}) &\leq \int_{\mathbb{R}^n} |D(\max(h_{m-1}, f_m))|^p \, dx \\ &\quad + \int_{\mathbb{R}^n} |D(\min(h_{m-1}, f_m))|^p \, dx \\ &= \int_{\mathbb{R}^n} |Dh_{m-1}|^p + |Df_m|^p \, dx \\ &\leq \int_{\mathbb{R}^n} |Dh_{m-1}|^p \, dx + \mathrm{Cap_p}(A_m) + \frac{\epsilon}{2^m}.\end{aligned}$$

Consequently,

$$\int_{\mathbb{R}^n} |Dh_m|^p dx - \int_{\mathbb{R}^n} |Dh_{m-1}|^p \, dx \leq \mathrm{Cap_p}(A_m) - \mathrm{Cap_p}(A_{m-1}) + \frac{\epsilon}{2^m},$$

from which it follows by adding that

$$\int_{\mathbb{R}^n} |Dh_m|^p \, dx \leq \mathrm{Cap}_p(A_m) + \epsilon \qquad (m = 1, 2, \ldots).$$

Set $f \equiv \lim_{m\to\infty} h_m$. Then $\bigcup_{k=1}^{\infty} A_k \subset \{x \mid f(x) \geq 1\}^0$. Furthermore,

$$\begin{aligned} ||f||_{L^{p^*}(\mathbb{R}^n)} &= \lim_{m\to\infty} ||h_m||_{L^{p^*}(\mathbb{R}^n)} \\ &\leq C_1 \liminf_{m\to\infty} ||Dh_m||_{L^p(\mathbb{R}^n)} \\ &\leq C \left[\lim_{m\to\infty} \mathrm{Cap}_p(A_m) + \epsilon \right]^{1/p}. \end{aligned}$$

Since $p > 1$, a subsequence of $\{Dh_m\}_{m=1}^{\infty}$ converges weakly to Df in $L^p(\mathbb{R}^n)$ (cf. Theorem 3 in Section 1.9); thus $f \in K^p$. Consequently,

$$\mathrm{Cap}_p\left(\bigcup_{k=1}^{\infty} A_k\right) \leq ||Df||^p_{L^p(\mathbb{R}^n)} \leq \lim_{m\to\infty} \mathrm{Cap}_p(A_m) + \epsilon.$$

9. We prove (ix) by first noting

$$\mathrm{Cap}_p\left(\bigcap_{k=1}^{\infty} A_k\right) \leq \lim_{k\to\infty} \mathrm{Cap}_p(A_k).$$

On the other hand, choose any open set U with $\bigcap_{k=1}^{\infty} A_k \subset U$. As $\bigcap_{k=1}^{\infty} A_k$ is compact, there exists a positive integer m such that $A_k \subset U$ for $k \geq m$. Thus

$$\lim_{k\to\infty} \mathrm{Cap}_p(A_k) \leq \mathrm{Cap}_p(U).$$

Recall (i) to complete the proof of (ix). ∎

4.7.2 Capacity and Hausdorff dimension

As noted earlier, we are interested in capacity as a way of characterizing certain "very small" subsets of $\mathbb{R}^n$. Obviously Hausdorff measures provide another approach, and so it is important to understand the relationships between capacity and Hausdorff measure.

We begin with a refinement of assertion (v) from Theorem 2:

THEOREM 3
If $\mathcal{H}^{n-p}(A) < \infty$, *then* $\mathrm{Cap}_p(A) = 0$ $(1 < p < n)$.

PROOF

1. We may assume A is compact.

2. *Claim*: There exists a constant C, depending only on n and A, such that if V is any open set containing A, there exists an open set W and $f \in K^p$ such that

$$\begin{cases} A \subset W \subset \{f = 1\}, \\ \text{spt}\ (f) \subset V, \\ \int_{\mathbb{R}^n} |Df|^p \, dx \leq C. \end{cases}$$

Proof of Claim: Let V be an open set containing A and let $\delta \equiv 1/2 \operatorname{dist}(A, \mathbb{R}^n - V)$. Since $\mathcal{H}^{n-p}(A) < \infty$ and A is compact, there exists a finite collection $\{U(x_i, r_i)\}_{i=1}^m$ of open balls such that $2r_i < \delta$, $U(x_i, r_i) \cap A \neq \emptyset$, $A \subset \bigcup_{i=1}^m U(x_i, r_i)$, and

$$\sum_{i=1}^{m} \alpha(n-p) r_i^{n-p} \leq C\mathcal{H}^{n-p}(A) + 1.$$

for some constant C.

Now set $W \equiv \bigcup_{i=1}^m U(x_i, r_i)$ and define $f_i \in K^p$ by

$$f_i(x) = \begin{cases} 1 & \text{if } |x - x_i| \leq r_i \\ 2 - \dfrac{|x - x_i|}{r_i} & \text{if } r_i \leq |x - x_i| \leq 2r_i \\ 0 & \text{if } 2r_i \leq |x - x_i|. \end{cases}$$

Then

$$\int_{\mathbb{R}^n} |Df_i|^p \, dx \leq C r_i^{n-p}.$$

Let $f \equiv \max_{1 \leq i \leq m} f_i$. Then $f \in K^p$, $W \subset \{f = 1\}$, spt $(f) \subset V$, and

$$\int_{\mathbb{R}^n} |Df|^p \, dx \leq \sum_{i=1}^{m} \int_{\mathbb{R}^n} |Df_i|^p \, dx \leq C \sum_{i=1}^{m} r_i^{n-p} \leq C(\mathcal{H}^{n-p}(A) + 1).$$

3. Using the claim inductively, we can find open sets $\{V_k\}_{k=1}^\infty$ and functions $f_k \in K^p$ such that

$$\begin{cases} A \subset V_{k+1} \subset V_k, \\ \overline{V}_{k+1} \subset \{f_k = 1\}^o, \\ \text{spt}\ (f_k) \subset V_k, \\ \int_{\mathbb{R}^n} |Df_k|^p \, dx \leq C. \end{cases}$$

Set

$$S_j \equiv \sum_{k=1}^{j} \frac{1}{k}$$

and

$$g_j \equiv \frac{1}{S_j} \sum_{k=1}^{j} \frac{f_k}{k} .$$

Then $g_j \in K^p$, $g_j \geq 1$ on V_{j+1}. Since $\operatorname{spt}(|Df_k|) \subset V_k - \overline{V}_{k+1}$, we see

$$\begin{aligned}
\mathrm{Cap}_p(A) \leq \int_{\mathbb{R}^n} |Dg_j|^p \, dx &= \frac{1}{S_j^p} \sum_{k=1}^{j} \frac{1}{k^p} \int_{\mathbb{R}^n} |Df_k|^p \, dx \\
&\leq \frac{C}{S_j^p} \sum_{k=1}^{j} \frac{1}{k^p} \to 0 \text{ as } j \to \infty,
\end{aligned}$$

since $p > 1$. ▮

THEOREM 4
Assume $A \subset \mathbb{R}^n$ and $1 \leq p < \infty$. If $\mathrm{Cap}_p(A) = 0$, *then $\mathcal{H}^s(A) = 0$ for all $s > n - p$.*

REMARK We will prove later in Section 5.6.3 that $\mathrm{Cap}_1(A) = 0$ if and only if $\mathcal{H}^{n-1}(A) = 0$. ▮

PROOF
1. Let $\mathrm{Cap}_p(A) = 0$ and $n - p < s < \infty$. Then for all $i \geq 1$, there exists $f_i \in K^p$ such that $A \subset \{f_i \geq 1\}^o$ and

$$\int_{\mathbb{R}^n} |Df_i|^p \, dx \leq \frac{1}{2^i} .$$

Let $g \equiv \sum_{i=1}^{\infty} f_i$. Then

$$\left(\int_{\mathbb{R}^n} |Dg|^p \, dx \right)^{1/p} \leq \sum_{i=1}^{\infty} \left(\int_{\mathbb{R}^n} |Df_i|^p \, dx \right)^{1/p} < \infty$$

and by the Gagliardo–Nirenberg–Sobolev inequality,

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} |g|^{p^*} \, dx \right)^{1/p^*} &\leq \sum_{i=1}^{\infty} \left(\int_{\mathbb{R}^n} |f_i|^{p^*} \, dx \right)^{1/p^*} \\
&\leq \sum_{i=1}^{\infty} C_1 \left(\int_{\mathbb{R}^n} |Df_i|^p \, dx \right)^{1/p} < \infty.
\end{aligned}$$

Thus $g \in K^p$.

2. Note $A \subset \{g \geq m\}^o$ for all $m \geq 1$. Fix any $a \in A$. Then for r small enough that $B(a,r) \subset \{g \geq m\}^o$, $(g)_{a,r} \geq m$; therefore $(g)_{a,r} \to \infty$ as $r \to 0$.

3. *Claim*: For each $a \in A$,

$$\limsup_{r \to 0} \frac{1}{r^s} \int_{B(a,r)} |Dg|^p \, dx = +\infty.$$

Proof of Claim: Let $a \in A$ and suppose

$$\limsup_{r \to 0} \frac{1}{r^s} \int_{B(a,r)} |Dg|^p \, dx < \infty.$$

Then there exists a constant $M < \infty$ such that

$$\frac{1}{r^s} \int_{B(a,r)} |Dg|^p \, dx \leq M$$

for all $0 < r \leq 1$. Then for $0 < r \leq 1$,

$$⨍_{B(a,r)} |g - (g)_{a,r}|^p \, dx \leq C_2 r^p ⨍_{B(a,r)} |Dg|^p \, dx \leq C r^\theta,$$

where $\theta = s - (n - p) > 0$. Thus

$$\begin{aligned}
|(g)_{a,r/2} - (g)_{a,r}| &= \frac{1}{\mathcal{L}^n(B(a,r/2))} \left| \int_{B(a,r/2)} g - (g)_{a,r} \, dx \right| \\
&\leq 2^n ⨍_{B(a,r)} |g - (g)_{a,r}| \, dx \\
&\leq 2^n \left(⨍_{B(a,r)} |g - (g)_{a,r}|^p \, dx \right)^{1/p} \\
&= C r^{\frac{\theta}{p}}.
\end{aligned}$$

Hence if $k > j$,

$$\begin{aligned}
|(g)_{a,1/2^k} - (g)_{a,1/2^j}| &\leq \sum_{l=j+1}^{k} |(g)_{a,1/2^l} - (g)_{a,1/2^{l-1}}| \\
&\leq C \sum_{l=j+1}^{k} \left(\frac{1}{2^{l-1}} \right)^{\frac{\theta}{p}}.
\end{aligned}$$

This last sum is the tail of a geometric series, and so $\{(g)_{a,1/2^k}\}_{k=1}^\infty$ is a Cauchy sequence. Thus $(g)_{a,1/2^k} \not\to \infty$, a contradiction.

4. Consequently,

$$A \subset \left\{ a \in \mathbb{R}^n \mid \limsup_{r \to 0} \frac{1}{r^s} \int_{B(a,r)} |Dg|^p \, dx = +\infty \right\}$$
$$\subset \left\{ a \in \mathbb{R}^n \mid \limsup_{r \to 0} \frac{1}{r^s} \int_{B(a,r)} |Dg|^p \, dx > 0 \right\} \equiv \Lambda_s.$$

But since $|Dg|^p$ is $\mathcal{L}^n$-summable, $\mathcal{H}^s(\Lambda_s) = 0$, according to Theorem 3 in Section 2.4.3. ∎

4.8 Quasicontinuity; Precise representatives of Sobolev functions

This section studies the fine properties of Sobolev functions.

LEMMA 1
Assume $f \in K^p$ and $\epsilon > 0$. Let

$$A \equiv \{x \in \mathbb{R}^n \mid (f)_{x,r} > \epsilon \text{ for some } r > 0\}.$$

Then

$$\operatorname{Cap}_p(A) \leq \frac{C}{\epsilon^p} \int_{\mathbb{R}^n} |Df|^p \, dx \tag{$\star$}$$

where C depends only on n and p.

REMARK This is a kind of capacity variant of the usual estimate

$$\mathcal{L}^n\{x \in \mathbb{R}^n \mid f(x) > \epsilon\} \leq \frac{1}{\epsilon^p} \int_{\mathbb{R}^n} |f|^p \, dx. \quad ∎$$

PROOF For the moment we set $\epsilon = 1$ and observe that if $x \in A$ and $(f)_{x,r} > 1$, then

$$\alpha(n) r^n \leq \int_{B(x,r)} f \, dy \leq (\alpha(n) r^n)^{1 - \frac{1}{p^*}} \left(\int_{B(x,r)} f^{p^*} \, dy \right)^{\frac{1}{p^*}}$$

so that

$$r \leq C$$

for some constant C.

According to the Besicovitch Covering Theorem (Section 1.5.2), there exist an integer N_n and countable collections $\mathcal{F}_1, \ldots, \mathcal{F}_{N_n}$ of disjoint closed balls such that

$$A \subset \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{F}_i} B$$

and

$$(f)_B > 1 \quad \text{for each} \quad B \in \bigcup_{i=1}^{N_n} \mathcal{F}_i.$$

Denote by B_i^j the elements of $\mathcal{F}_i$ $(i = 1, \ldots, N_n;\ j = 1, \ldots)$. Choose $h_{ij} \in K^p$ such that

$$h_{ij} = ((f)_{B_i^j} - f)^+ \quad \text{on} \quad B_i^j$$

and

$$\int_{\mathbb{R}^n} |Dh_{ij}|^p \, dx \le C \int_{B_i^j} |Df|^p \, dx \qquad (i = 1, \ldots N_n;\ j = 1, 2, \ldots)$$

where C depends only on n and p. This is possible according to Theorem 1 in Section 4.4 and Poincaré's inequality in Section 4.5.2. Note that

$$f + h_{ij} \ \ge \ (f)_{B_i^j} \ \ge \ 1 \ \text{ in } \ B_i^j$$

and hence, setting

$$h \equiv \sup\{h_{ij} \mid \ i = 1, \ldots, N_n,\ j = 1, \ldots\} \in K^p,$$

that

$$f + h \ \ge \ 1 \ \text{ on } \ A. \tag{$\star\star$}$$

Now

$$\begin{aligned}\int_{\mathbb{R}^n} |D(f+h)|^p \, dx &\le C \left\{ \int_{\mathbb{R}^n} |Df|^p \, dx + \sum_{i=1}^{N_n} \sum_{j=1}^{\infty} \int_{\mathbb{R}^n} |Dh_{ij}|^p \, dx \right\} \\ &\le C \int_{\mathbb{R}^n} |Df|^p \, dx.\end{aligned}$$

Consequently, since A is open and so $(\star\star)$ implies

$$A \subset \{f + h \ \ge \ 1\}^o,$$

we have

$$\operatorname{Cap}_p(A) \le \int_{\mathbb{R}^n} |D(f+h)|^p \, dx \le C \int_{\mathbb{R}^n} |Df|^p \, dx.$$

In case $0 < \epsilon \neq 1$, we set $g = \epsilon^{-1} f \in K^p$, so that

$$
\begin{aligned}
A &\equiv \{x \mid (f)_{x,r} > \epsilon \text{ for some } r > 0\} \\
&= \{x \mid (g)_{x,r} > 1 \text{ for some } r > 0\}.
\end{aligned}
$$

Thus

$$\mathrm{Cap}_p(A) \leq C \int_{\mathbb{R}^n} |Dg|^p \, dx = \frac{C}{\epsilon^p} \int_{\mathbb{R}^n} |Df|^p \, dx. \quad ∎$$

We now study the fine structure properties of Sobolev functions, using capacity to measure the size of the "bad" sets.

DEFINITION *A function f is* ***p-quasicontinuous*** *if for each $\epsilon > 0$, there exists an open set V such that*

$$\mathrm{Cap}_p(V) \leq \epsilon$$

and

$$f\,|_{\mathbb{R}^n - V} \quad \text{is continuous.}$$

THEOREM 1 FINE PROPERTIES OF SOBOLEV FUNCTIONS
Suppose $f \in W^{1,p}(\mathbb{R}^n)$, $1 \leq p < n$.

(i) There is a Borel set $E \subset \mathbb{R}^n$ such that

$$\mathrm{Cap}_p(E) = 0$$

and

$$\lim_{r \to 0} (f)_{x,r} \equiv f^\star(x)$$

exists for each $x \in \mathbb{R}^n - E$.

(ii) In addition,

$$\lim_{r \to 0} ⨍_{B(x,r)} |f - f^\star(x)|^{p^\star} \, dy = 0$$

for each $x \in \mathbb{R}^n - E$.

(iii) The precise representative $f^\star$ is p-quasicontinuous.

REMARK Notice that if f is a Sobolev function and $f = g$ $\mathcal{L}^n$ a.e., then g is also a Sobolev function. Consequently if we wish to study the fine properties of f, we must turn our attention to the precise representative $f^\star$, defined in Section 1.7.1. ∎

PROOF

1. Set

$$A \equiv \left\{ x \in \mathbb{R}^n \mid \limsup_{r\to 0} \frac{1}{r^{n-p}} \int_{B(x,r)} |Df|^p \, dy > 0 \right\}.$$

By Theorem 3 in Section 2.4.3 and Theorem 3 in Section 4.7.2,

$$\mathcal{H}^{n-p}(A) = 0, \qquad \mathrm{Cap}_p(A) = 0.$$

Now, according to Poincaré's inequality,

$$\lim_{r\to 0} \fint_{B(x,r)} |f - (f)_{x,r}|^{p^*} \, dy = 0 \tag{$\star$}$$

for each $x \notin A$. Choose functions $f_i \in W^{1,p}(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} |Df - Df_i|^p \, dy \le \frac{1}{2^{(p+1)i}} \qquad (i = 1, 2, \ldots),$$

and set

$$B_i \equiv \left\{ x \in \mathbb{R}^n \mid \fint_{B(x,r)} |f - f_i| \, dy > \frac{1}{2^i} \text{ for some } r > 0 \right\}.$$

According to Lemma 1,

$$\frac{\mathrm{Cap}_p(B_i)}{2^{pi}} \le C \int_{\mathbb{R}^n} |Df - Df_i|^p \, dy \le \frac{C}{2^{(p+1)i}}\ .$$

Consequently, $\mathrm{Cap}_p(B_i) \le C/2^i$. Furthermore,

$$\begin{aligned} |(f)_{x,r} - f_i(x)| \le &\fint_{B(x,r)} |f - (f)_{x,r}| \, dy + \fint_{B(x,r)} |f - f_i| \, dy \\ &+ \fint_{B(x,r)} |f_i - f_i(x)| \, dy. \end{aligned}$$

Thus ($\star$) and the definition of B_i imply

$$\limsup_{r\to 0} |(f)_{x,r} - f_i(x)| \le \frac{1}{2^i} \qquad (x \notin A \cup B_i). \tag{$\star\star$}$$

Set $E_k \equiv A \cup (\bigcup_{j=k}^\infty B_j)$. Then

$$\mathrm{Cap}_p(E_k) \le \mathrm{Cap}_p(A) + \sum_{j=k}^\infty \mathrm{Cap}_p(B_j) \le C \sum_{j=k}^\infty \frac{1}{2^j}\ .$$

Furthermore, if $x \in \mathbb{R}^n - E_k$ and $i, j \geq k$, then

$$\begin{aligned}|f_i(x) - f_j(x)| &\leq \limsup_{r\to 0} |(f)_{x,r} - f_i(x)| \\ &\quad + \limsup_{r\to 0} |(f)_{x,r} - f_j(x)| \\ &\leq \frac{1}{2^i} + \frac{1}{2^j} \qquad \text{by } (\star\star).\end{aligned}$$

Hence $\{f_j\}_{j=1}^{\infty}$ converges uniformly on $\mathbb{R}^n - E_k$ to some continuous function g. Furthermore,

$$\limsup_{r\to 0} |g(x) - (f)_{x,r}| \leq |g(x) - f_i(x)| + \limsup_{r\to 0} |f_i(x) - (f)_{x,r}|,$$

so that $(\star\star)$ implies

$$g(x) = \lim_{r\to 0} (f)_{x,r} = f^\star(x) \qquad (x \in \mathbb{R}^n - E_k).$$

Now set $E \equiv \bigcap_{k=1}^{\infty} E_k$. Then $\mathrm{Cap}_p(E) \leq \lim_{k\to\infty} \mathrm{Cap}_p(E_k) = 0$ and

$$f^\star(x) = \lim_{r\to 0} (f)_{x,r} \quad \text{exists for each } x \in \mathbb{R}^n - E.$$

This proves (i).

2. To prove (ii), note $A \subset E$ and so $(\star)$ implies for $x \in \mathbb{R}^n - E$ that

$$\begin{aligned}&\lim_{r\to 0} \left(\fint_{B(x,r)} |f - f^\star(x)|^{p^\star} dy \right)^{\frac{1}{p^\star}} \\ &\quad \leq \lim_{r\to 0} |(f)_{x,r} - f^\star(x)| + \lim_{r\to 0} \left(\fint_{B(x,r)} |f - (f)_{x,r}|^{p^\star} dy \right)^{\frac{1}{p^\star}} \\ &\quad = 0.\end{aligned}$$

3. Finally, we prove (iii) by fixing $\epsilon > 0$ and then choosing k such that $\mathrm{Cap}_p(E_k) < \epsilon/2$. According to Theorem 2 in Section 4.7, there exists an open set $U \supset E_k$ with $\mathrm{Cap}_p(U) < \epsilon$. Since the $\{f_i\}_{i=1}^{\infty}$ converge uniformly to $f^\star$ on $\mathbb{R}^n - U$, $f^\star|_{\mathbb{R}^n - U}$ is continuous. ▮

4.9 Differentiability on lines

We will study in this section the properties of a Sobolev function f, or more exactly its precise representative $f^\star$, restricted to lines.

4.9.1 Sobolev functions of one variable

NOTATION If $h : \mathbb{R} \to \mathbb{R}$ is absolutely continuous on each compact subinterval, we write h' to denote its derivative (which exists $\mathcal{L}^1$ a.e.)

THEOREM 1
Assume $1 \leq p < \infty$.

(i) *If* $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R})$, *then its precise representative* f^* *is absolutely continuous on each compact subinterval of* $\mathbb{R}$ *and* $(f^*)' \in L^p_{\mathrm{loc}}(\mathbb{R})$.

(ii) *Conversely, suppose* $f \in L^p_{\mathrm{loc}}(\mathbb{R})$ *and* $f = g$ $\mathcal{L}^1$ *a.e., where* g *is absolutely continuous on each compact subinterval of* $\mathbb{R}$ *and* $g' \in L^p_{\mathrm{loc}}(\mathbb{R})$. *Then* $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R})$.

PROOF

1. First assume $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R})$ and let $(d/dt)f$ denote its weak derivative. For $0 < \epsilon \leq 1$ define $f^\epsilon \equiv \eta_\epsilon * f$, as before. Then

$$f^\epsilon(y) = f^\epsilon(x) + \int_x^y (f^\epsilon)'(t)\, dt. \qquad (\star)$$

Let x_0 be a Lebesgue point of f and ϵ, $\delta \in (0, 1)$. Since

$$|f^\epsilon(x) - f^\delta(x)| \leq \int_{x_0}^x |(f^\epsilon)'(t) - (f^\delta)'(t)|\, dt + |f^\epsilon(x_0) - f^\delta(x_0)|$$

for $x \in \mathbb{R}$, it follows from Theorem 1 in Section 4.2.1 that $\{f^\epsilon\}_{\epsilon>0}$ converges uniformly on compact subsets of $\mathbb{R}$ to a continuous function g with $g = f$ $\mathcal{L}^1$ a.e. From $(\star)$ we see

$$g(x) = g(x_0) + \int_{x_0}^x \frac{d}{dt} f(t)\, dt$$

and hence g is locally absolutely continuous with $g' = (d/dt)f$ $\mathcal{L}^1$ a.e.

Finally, since $(f)_{x,r} = (g)_{x,r} \to g(x)$ for each $x \in \mathbb{R}$, we see $g = f^*$. This proves (i).

2. On the other hand, assume $f = g$ $\mathcal{L}^1$ a.e., g is absolutely continuous and $g' \in L^p_{\mathrm{loc}}(\mathbb{R})$. Then for each $\varphi \in C^1_c(\mathbb{R})$,

$$\int_{-\infty}^\infty f\varphi'\, dx = \int_{-\infty}^\infty g\varphi'\, dx = -\int_{-\infty}^\infty g'\varphi\, dx,$$

and thus g' is the weak derivative of f. Since $g' \in L^p_{\mathrm{loc}}(\mathbb{R})$, we conclude $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R})$. ∎

4.9.2 Differentiability on a.e. line

THEOREM 2

(i) *If $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, then for each $k = 1, \dots, n$ the functions*

$$f^*_k(x', t) \equiv f^*(\dots, x_{k-1}, t, x_{k+1}, \dots)$$

are absolutely continuous in t on compact subsets of $\mathbb{R}$, for $\mathcal{L}^{n-1}$ a.e. point $x' = (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in \mathbb{R}^{n-1}$. In addition, $(f^*_k)' \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$.

(ii) *Conversely, suppose $f \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ and $f = g$ $\mathcal{L}^n$ a.e., where for each $k = 1, \dots, n$, the functions*

$$g_k(x', t) \equiv g(\dots, x_{k-1}, t, x_{k+1}, \dots)$$

are absolutely continuous in t on compact subsets of $\mathbb{R}$ for $\mathcal{L}^{n-1}$ a.e. point $x' = (x_1, \dots, x_{k-1}, x_{k+1}, \dots x_n) \in \mathbb{R}^{n-1}$, and $g'_k \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$. Then $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$.

PROOF

1. It suffices to prove assertion (i) for the case $k = n$. Define $f^\epsilon \equiv \eta_\epsilon * f$ as before, and recall

$$f^\epsilon \to f \text{ in } W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n).$$

By Fubini's Theorem, for each $L > 0$ and $\mathcal{L}^{n-1}$ a.e. $x' = (x_1, \dots, x_{n-1})$, the expression

$$\int_{-L}^{L} |f^\epsilon(x', t) - f(x', t)|^p + \left| \frac{\partial f^\epsilon}{\partial x_n}(x', t) - \frac{\partial f}{\partial x_n}(x', t) \right|^p dt$$

goes to zero as $\epsilon \to 0$. Thus the functions

$$f^\epsilon_n(t) \equiv f^\epsilon(x', t) \qquad (t \in \mathbb{R})$$

converge in $W^{1,p}_{\mathrm{loc}}(\mathbb{R})$, and so locally uniformly, to a locally absolutely continuous function f_n, with $f'_n(t) = (\partial f / \partial x_n)(x', t)$ for $\mathcal{L}^1$ a.e. $t \in \mathbb{R}$. On the other hand, Theorem 1 in Section 4.8, Theorem 2 in Section 5.6.3, and Theorem 4 in Section 4.7.2 imply

$$f^\epsilon \to f^* \qquad \mathcal{H}^{n-1} \text{ a.e.}$$

Therefore, in view of Corollary 1 in Section 2.4.1, for $\mathcal{L}^{n-1}$ a.e. x'

$$f^\epsilon_n(t) \to f^*(x', t)$$

for *all* $t \in \mathbb{R}$. Hence for $\mathcal{L}^{n-1}$ a.e. x' and all $t \in \mathbb{R}$,

$$f_n(t) = f^*(x', t).$$

This proves statement (i).

2. Assume now the hypothesis of assertion (ii). Then for each $\varphi \in C_c^1(\mathbb{R}^n)$,

$$\begin{aligned}
\int_{\mathbb{R}^n} f \frac{\partial \varphi}{\partial x_k}\, dx &= \int_{\mathbb{R}^n} g \frac{\partial \varphi}{\partial x_k}\, dx \\
&= \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} g_k(x', t) \varphi'(x', t)\, dt \right) dx' \\
&= - \int_{\mathbb{R}^{n-1}} \left(\int_{-\infty}^{\infty} g_k'(x', t) \varphi(x', t)\, dt \right) dx' \\
&= - \int_{\mathbb{R}^n} g_k' \varphi\, dx.
\end{aligned}$$

Thus $\partial f / \partial x_k = g_k'$ $\mathcal{L}^n$ a.e., $k = 1, \ldots, n$, and hence $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$. ∎

5

BV Functions and Sets of Finite Perimeter

In this chapter we introduce and study functions on $\mathbb{R}^n$ of *bounded variation*, which is to say functions whose weak first partial derivatives are Radon measures. This is essentially the weakest measure theoretic sense in which a function can be differentiable. We also investigate sets E having *finite perimeter*, which means the indicator function χ_E is BV.

It is not so obvious that any of the usual rules of calculus apply to functions whose first derivatives are merely measures. The principal goal of this chapter is therefore to study this problem, investigating in particular the extent to which a BV function is "measure theoretically C^1" and a set of finite perimeter has "a C^1 boundary measure theoretically."

Our study initially, in Sections 5.1 through 5.4, parallels the corresponding investigation of Sobolev functions in Chapter 4. Section 5.5 extends the Coarea Formula to the BV setting and Section 5.6 generalizes the Gagliardo–Nirenberg–Sobolev Inequality. Sections 5.7, 5.8, and 5.11 analyze the measure theoretic boundary of a set of finite perimeter, and most importantly establish a version of the Gauss–Green Theorem. This study is carried over in Sections 5.9 and 5.10 to study the fine, pointwise properties of BV functions.

5.1 Definitions; Structure Theorem

Throughout this chapter, U denotes an open subset of $\mathbb{R}^n$.

DEFINITION *A function $f \in L^1(U)$ has **bounded variation** in U if*

$$\sup\left\{\int_U f \operatorname{div}\varphi \, dx \mid \varphi \in C_c^1(U;\mathbb{R}^n), |\varphi| \leq 1\right\} < \infty.$$

We write

$$BV(U)$$

to denote the space of functions of bounded variation.

DEFINITION *An $\mathcal{L}^n$-measurable subset $E \subset \mathbb{R}^n$ has* ***finite perimeter*** *in U if*

$$\chi_E \in BV(U).$$

It is convenient to introduce also local versions of the above concepts:

DEFINITION *A function $f \in L^1_{\text{loc}}(U)$ has* ***locally bounded variation*** *in U if for each open set $V \subset\subset U$,*

$$\sup\left\{\int_V f \operatorname{div}\varphi \, dx \mid \varphi \in C^1_c(V;\mathbb{R}^n),\ |\varphi| \le 1\right\} < \infty.$$

We write

$$BV_{\text{loc}}(U)$$

to denote the space of such functions.

DEFINITION *An $\mathcal{L}^n$-measurable subset $E \subset \mathbb{R}^n$ has* ***locally finite perimeter*** *in U if*

$$\chi_E \in BV_{\text{loc}}(U).$$

Some examples will be presented later, after we establish this general structure assertion.

THEOREM 1 STRUCTURE THEOREM FOR BV_{loc} FUNCTIONS
Let $f \in BV_{\text{loc}}(U)$. Then there exists a Radon measure μ on U and a μ-measurable function $\sigma : U \to \mathbb{R}^n$ such that

(i) $|\sigma(x)| = 1$ *μ a.e., and*

(ii) $\int_U f \operatorname{div}\varphi \, dx = -\int_U \varphi \cdot \sigma \, d\mu$

for all $\varphi \in C^1_c(U;\mathbb{R}^n)$.

As we will discuss in detail later, the Structure Theorem asserts that the weak first partial derivatives of a BV function are Radon measures.

PROOF Define the linear functional

$$L : C_c^1(U;\mathbb{R}^n) \to \mathbb{R}$$

by

$$L(\varphi) \equiv -\int_U f \operatorname{div}\varphi\, dx$$

for $\varphi \in C_c^1(U;\mathbb{R}^n)$. Since $f \in BV_{\mathrm{loc}}(U)$, we have

$$\sup\left\{L(\varphi) \mid \varphi \in C_c^1(V;\mathbb{R}^n), |\varphi| \le 1\right\} \equiv C(V) < \infty$$

for each open set $V \subset\subset U$, and thus

$$|L(\varphi)| \le C(V)\|\varphi\|_{L^\infty} \tag{$\star$}$$

for $\varphi \in C_c^1(V;\mathbb{R}^n)$.

Fix any compact set $K \subset U$, and then choose an open set V such that $K \subset V \subset\subset U$. For each $\varphi \in C_c(U;\mathbb{R}^n)$ with $\operatorname{spt}\varphi \subset K$, choose $\varphi_k \in C_c^1(V;\mathbb{R}^n)$ $(k = 1,\dots)$ so that $\varphi_k \to \varphi$ uniformly on V. Define

$$\bar{L}(\varphi) \equiv \lim_{k\to\infty} L(\varphi_k);$$

according to $(\star)$ this limit exists and is independent of the choice of the sequence $\{\varphi_k\}_{k=1}^\infty$ converging to φ. Thus L uniquely extends to a linear functional

$$\bar{L} : C_c(U;\mathbb{R}^n) \to \mathbb{R}$$

and

$$\sup\{\bar{L}(\varphi) \mid \varphi \in C_c(U;\mathbb{R}^n), |\varphi| \le 1, \operatorname{spt}\varphi \subset K\} < \infty$$

for each compact set $K \subset U$. The Riesz Representation Theorem, Section 1.8, now completes the proof. ∎

NOTATION

(i) If $f \in BV_{\mathrm{loc}}(U)$, we will henceforth write

$$\|Df\|$$

for the measure μ, and

$$[Df] \equiv \|Df\| \llcorner \sigma.$$

Hence assertion (ii) in Theorem 1 reads

$$\int_U f \operatorname{div}\varphi\, dx = -\int_U \varphi\cdot\sigma\, d\|Df\| = -\int_U \varphi\cdot\, d[Df]$$

for all $\varphi \in C_c^1(U;\mathbb{R}^n)$.

(ii) Similarly, if $f = \chi_E$, and E is a set of locally finite perimeter in U, we will hereafter write

$$\|\partial E\|$$

for the measure μ, and

$$\nu_E \equiv -\sigma.$$

Consequently,

$$\int_E \operatorname{div}\varphi\, dx = \int_U \varphi \cdot \nu_E \, d\|\partial E\|$$

for all $\varphi \in C^1_c(U;\mathbb{R}^n)$.

MORE NOTATION If $f \in BV_{\mathrm{loc}}(U)$, we write

$$\mu^i = \|Df\| \llcorner \sigma^i \qquad (i = 1, \dots, n)$$

for $\sigma = (\sigma^1, \dots, \sigma^n)$. By Lebesgue's Decomposition Theorem (Theorem 3 in Section 1.6.2), we may further set

$$\mu^i = \mu^i_{\mathrm{ac}} + \mu^i_{\mathrm{s}},$$

where

$$\mu^i_{\mathrm{ac}} \ll \mathcal{L}^n, \qquad \mu^i_{\mathrm{s}} \perp \mathcal{L}^n.$$

Then

$$\mu^i_{\mathrm{ac}} = \mathcal{L}^n \llcorner f_i$$

for some function $f_i \in L^1_{\mathrm{loc}}(U)$ $(i = 1, \dots, n)$. Write

$$\begin{cases} \dfrac{\partial f}{\partial x_i} & \equiv \quad f_i \qquad (i = 1, \dots, n) \\ Df & \equiv \quad \left(\dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_n}\right), \\ [Df]_{\mathrm{ac}} & \equiv \quad (\mu^1_{\mathrm{ac}}, \dots, \mu^n_{\mathrm{ac}}) = \mathcal{L}^n \llcorner Df, \\ [Df]_{\mathrm{s}} & \equiv \quad (\mu^1_{\mathrm{s}}, \dots, \mu^n_{\mathrm{s}}). \end{cases}$$

Thus

$$[Df] = [Df]_{\mathrm{ac}} + [Df]_{\mathrm{s}} = \mathcal{L}^n \llcorner Df + [Df]_{\mathrm{s}},$$

so that $Df \in L^1_{\mathrm{loc}}(U;\mathbb{R}^n)$ is the density of the absolutely continuous part of $[Df]$.

REMARK Compare this with the notation for convex functions set forth in Section 6.3. ∎

REMARK

(i) $\|Df\|$ is the *variation measure* of f; $\|\partial E\|$ is the *perimeter measure* of E; $\|\partial E\|(U)$ is the *perimeter of E in U*.

(ii) If $f \in BV_{\mathrm{loc}}(U) \cap L^1(U)$, then $f \in BV(U)$ if and only if $\|Df\|(U) < \infty$, in which case we define

$$\|f\|_{BV(U)} \equiv \|f\|_{L^1(U)} + \|Df\|(U).$$

(iii) From the proof of the Riesz Representation Theorem, we see

$$\|Df\|(V) = \sup\left\{\int_V f \operatorname{div}\varphi\, dx \mid \varphi \in C_c^1(V;\mathbb{R}^n), |\varphi| \leq 1\right\},$$

$$\|\partial E\|(V) = \sup\left\{\int_E \operatorname{div}\varphi\, dx \mid \varphi \in C_c^1(V;\mathbb{R}^n), |\varphi| \leq 1\right\}$$

for each $V \subset\subset U$. ∎

Example 1

Assume $f \in W^{1,1}_{\mathrm{loc}}(U)$. Then, for each $V \subset\subset U$ and $\varphi \in C_c^1(V;\mathbb{R}^n)$, with $|\varphi| \leq 1$, we have

$$\int_U f \operatorname{div}\varphi\, dx = -\int_U Df \cdot \varphi\, dx \leq \int_V |Df|\, dx < \infty.$$

Thus $f \in BV_{\mathrm{loc}}(U)$. Furthermore,

$$\|Df\| = \mathcal{L}^n \llcorner |Df|,$$

and

$$\sigma = \begin{cases} \dfrac{Df}{|Df|} & \text{if } Df \neq 0 \\ 0 & \text{if } Df = 0. \end{cases} \qquad \mathcal{L}^n \text{ a.e.}$$

Hence

$$W^{1,1}_{\mathrm{loc}}(U) \subset BV_{\mathrm{loc}}(U),$$

and similarly

$$W^{1,1}(U) \subset BV(U).$$

In particular,

$$W^{1,p}_{\mathrm{loc}}(U) \subset BV_{\mathrm{loc}}(U) \text{ for } 1 \leq p \leq \infty.$$

Hence, *each Sobolev function has locally bounded variation.* □

Example 2

Assume E is a smooth, open subset of $\mathbb{R}^n$ and $\mathcal{H}^{n-1}(\partial E \cap K) < \infty$ for each compact set $K \subset U$. Then for V and φ as above,

$$\int_E \operatorname{div} \varphi \, dx = \int_{\partial E} \varphi \cdot \nu \, d\mathcal{H}^{n-1},$$

ν denoting the outward unit normal along ∂E.

Hence

$$\int_E \operatorname{div} \varphi \, dx = \int_{\partial E \cap V} \varphi \cdot \nu \, d\mathcal{H}^{n-1} \leq \mathcal{H}^{n-1}(\partial E \cap V) < \infty.$$

Thus E has locally finite perimeter in U. Furthermore,

$$\|\partial E\|(U) = \mathcal{H}^{n-1}(\partial E \cap U)$$

and

$$\nu_E = \nu \qquad \mathcal{H}^{n-1} \text{ a.e. on } \partial E \cap U.$$

Thus $\|\partial E\|(U)$ measures the "size" of ∂E in U. Since $\chi_E \notin W^{1,1}_{\mathrm{loc}}(U)$ (according, for instance, to Theorem 2 in Section 4.9.2), we see

$$W^{1,1}_{\mathrm{loc}}(U) \subsetneq BV_{\mathrm{loc}}(U),$$
$$W^{1,1}(U) \subsetneq BV(U).$$

That is, *not every function of locally bounded variation is a Sobolev function.* □

REMARK Indeed, if $f \in BV_{\mathrm{loc}}(U)$, we can write as above

$$[Df] = [Df]_{\mathrm{ac}} + [Df]_{\mathrm{s}} = \mathcal{L}^n \llcorner Df + [Df]_{\mathrm{s}}.$$

Consequently, $f \in BV_{\mathrm{loc}}(U)$ *belongs to* $W^{1,p}_{\mathrm{loc}}(U)$ *if and only if*

$$f \in L^p_{\mathrm{loc}}(U), \qquad [Df]_{\mathrm{s}} = 0, \qquad Df \in L^p_{\mathrm{loc}}(U). \quad \blacksquare$$

The study of BV functions is for the most part more subtle than the study of Sobolev functions since we must always keep track of the singular part $[Df]_{\mathrm{s}}$ of the vector measure Df.

5.2 Approximation and compactness

5.2.1 Lower semicontinuity

THEOREM 1 LOWER SEMICONTINUITY OF VARIATION MEASURE
Suppose $f_k \in BV(U)$ $(k = 1, \ldots)$ and $f_k \to f$ in $L^1_{\mathrm{loc}}(U)$. Then

$$||Df||(U) \le \liminf_{k\to\infty} ||Df_k||(U).$$

PROOF Let $\varphi \in C^1_c(U;\mathbb{R}^n)$, $|\varphi| \le 1$. Then

$$\begin{aligned}\int_U f \operatorname{div}\varphi\,dx &= \lim_{k\to\infty} \int_U f_k \operatorname{div}\varphi\,dx \\ &= -\lim_{k\to\infty} \int_U \varphi\cdot\sigma_k\,d||Df_k|| \\ &\le \liminf_{k\to\infty} ||Df_k||(U).\end{aligned}$$

Thus

$$\begin{aligned}||Df||(U) &= \sup\left\{\int_U f\operatorname{div}\varphi\,dx \mid \varphi \in C^1_c(U;\mathbb{R}^n), |\varphi| \le 1\right\} \\ &\le \liminf_{k\to\infty} ||Df_k||(U). \quad \blacksquare\end{aligned}$$

5.2.2 Approximation by smooth functions

THEOREM 2 LOCAL APPROXIMATION BY SMOOTH FUNCTIONS
Assume $f \in BV(U)$. There exist functions $\{f_k\}_{k=1}^\infty \subset BV(U) \cap C^\infty(U)$ such that

(i) $f_k \to f$ in $L^1(U)$ and
(ii) $||Df_k||(U) \to ||Df||(U)$ as $k \to \infty$.

REMARK Compare with Theorem 2 in Section 4.2.1. Note we do not assert $||D(f_k - f)||(U) \to 0$. $\blacksquare$

PROOF
1. Fix $\epsilon > 0$. Given a positive integer m, define the open sets

$$U_k \equiv \left\{x \in U \mid \operatorname{dist}(x, \partial U) > \frac{1}{m+k}\right\} \cap U(0, k+m) \qquad (k = 1, \ldots)$$

and then choose m so large

$$||Df||(U - U_1) < \epsilon. \tag{$\star$}$$

Set $U_0 \equiv \emptyset$ and define

$$V_k \equiv U_{k+1} - \overline{U}_{k-1} \qquad (k = 1, \dots).$$

Let $\{\zeta_k\}_{k=1}^{\infty}$ be a sequence of smooth functions such that

$$\begin{cases} \zeta_k \in C_c^\infty(V_k) \quad 0 \le \zeta_k \le 1 \quad (k = 1, \dots) \\ \sum_{k=1}^{\infty} \zeta_k \equiv 1 \qquad \text{on } U. \end{cases}$$

Fix the mollifier η, as described in Section 4.2.1. Then for each k, select $\epsilon_k > 0$ so small that

$$\begin{cases} \operatorname{spt}(\eta_{\epsilon_k} * (f\zeta_k)) \subset V_k \\ \int_U |\eta_{\epsilon_k} * (f\zeta_k) - f\zeta_k| \, dx < \frac{\epsilon}{2^k}, \\ \int_U |\eta_{\epsilon_k} * (f\, D\zeta_k) - f\, D\zeta_k| \, dx < \frac{\epsilon}{2^k}. \end{cases} \tag{$\star\star$}$$

Define

$$f_\epsilon \equiv \sum_{k=1}^{\infty} \eta_{\epsilon_k} * (f\zeta_k).$$

In some neighborhood of each point $x \in U$ there are only finitely many nonzero terms in this sum; hence

$$f_\epsilon \in C^\infty(U).$$

2. Since also

$$f = \sum_{k=1}^{\infty} f\zeta_k,$$

($\star\star$) implies

$$||f_\epsilon - f||_{L^1(U)} \le \sum_{k=1}^{\infty} \int_U |\eta_{\epsilon_k} * (f\zeta_k) - f\zeta_k| \, dx < \epsilon.$$

Consequently,

$$f_\epsilon \to f \text{ in } L^1(U), \text{ as } \epsilon \to 0.$$

3. According to Theorem 1,

$$||Df||(U) \le \liminf_{\epsilon \to 0} ||Df_\epsilon||(U). \tag{$\star\star\star$}$$

4. Now let $\varphi \in C_c^1(U;\mathbb{R}^n)$, $|\varphi| \leq 1$. Then

$$\begin{aligned}
\int_U f_\epsilon \operatorname{div}\varphi\, dx &= \sum_{k=1}^\infty \int_U \eta_{\epsilon_k} * (f\,\zeta_k) \operatorname{div}\varphi\, dx \\
&= \sum_{k=1}^\infty \int_U f\,\zeta_k \operatorname{div}(\eta_{\epsilon_k} * \varphi)\, dx \\
&= \sum_{k=1}^\infty \int_U f \operatorname{div}(\zeta_k(\eta_{\epsilon_k} * \varphi))\, dx \\
&\quad - \sum_{k=1}^\infty \int_U f\, D\zeta_k \cdot (\eta_{\epsilon_k} * \varphi)\, dx \\
&= \sum_{k=1}^\infty \int_U f \operatorname{div}(\zeta_k(\eta_{\epsilon_k} * \varphi))\, dx \\
&\quad - \sum_{k=1}^\infty \int_U \varphi \cdot (\eta_{\epsilon_k} * (f\, D\zeta_k) - f\, D\zeta_k)\, dx \\
&\equiv I_1^\epsilon + I_2^\epsilon.
\end{aligned}$$

Here we used the fact $\sum_{k=1}^\infty D\zeta_k = 0$ in U. Now $|\zeta_k(\eta_{\epsilon_k} * \varphi)| \leq 1$ $(k = 1, \ldots)$, and each point in U belongs to at most three of the sets $\{V_k\}_{k=1}^\infty$. Thus

$$\begin{aligned}
|I_1^\epsilon| &= \left| \int_U f \operatorname{div}(\zeta_1(\eta_{\epsilon_1} * \varphi))\, dx + \sum_{k=2}^\infty \int_U f \operatorname{div}(\zeta_k \eta_{\epsilon_k} * \varphi)\, dx \right| \\
&\leq \|Df\|(U) + \sum_{k=2}^\infty \|Df\|(V_k) \\
&\leq \|Df\|(U) + 3\|Df\|(U - U_1) \\
&\leq \|Df\|(U) + 3\epsilon, \qquad \text{by } (\star).
\end{aligned}$$

On the other hand, $(\star\star)$ implies

$$|I_2^\epsilon| < \epsilon.$$

Therefore

$$\int_U f_\epsilon \operatorname{div}\varphi\, dx \leq \|Df\|(U) + 4\epsilon,$$

and so

$$\|Df_\epsilon\|(U) \leq \|Df\|(U) + 4\epsilon.$$

This estimate and $(\star\star\star)$ complete the proof. ∎

THEOREM 3 WEAK APPROXIMATION OF DERIVATIVES
For each function f_k as in the statement of Theorem 2, define the (vector-valued) Radon measure

$$\mu_k(B) \equiv \int_{B\cap U} Df_k \, dx$$

for each Borel set $B \subset \mathbb{R}^n$. Set also

$$\mu(B) \equiv \int_{B\cap U} d[Df].$$

Then

$$\mu_k \rightharpoonup \mu$$

weakly in the sense of (vector-valued) Radon measures on $\mathbb{R}^n$.

PROOF Fix $\varphi \in C_c^1(\mathbb{R}^n;\mathbb{R}^n)$ and $\epsilon > 0$. Define $U_1 \subset\subset U$ as in the previous proof and choose a smooth cutoff function ζ satisfying

$$\begin{cases} \zeta \equiv 1 \text{ on } U_1, \text{ spt } (\zeta) \subset U, \\ 0 \le \zeta \le 1. \end{cases}$$

Then

$$\int_{\mathbb{R}^n} \varphi \, d\mu_k = \int_U \varphi \cdot Df_k \, dx = \int_U \zeta\varphi \cdot Df_k \, dx + \int_U (1-\zeta)\varphi \cdot Df_k \, dx$$

$$= -\int_U \operatorname{div}(\zeta\varphi) f_k \, dx + \int_U (1-\zeta)\varphi \cdot Df_k \, dx. \qquad (\star)$$

Since $f_k \to f$ in $L^1(U)$, the first term in $(\star)$ converges to

$$-\int_U \operatorname{div}(\zeta\varphi) f \, dx = \int_U \zeta\varphi \cdot \, d[Df]$$
$$= \int_U \varphi \cdot \, d[Df] + \int_U (\zeta - 1)\varphi \cdot \, d[Df]. \qquad (\star\star)$$

The last term in $(\star\star)$ is estimated by

$$||\varphi||_{L^\infty} ||Df||(U - U_1) \le C\epsilon.$$

Using Theorem 1 in Section 5.2.1, we see that for k large enough, the last term in $(\star)$ is estimated by

$$||\varphi||_{L^\infty} ||Df_k||(U - U_1) \le C\epsilon.$$

Hence

$$\left| \int_{\mathbb{R}^n} \varphi \, d\mu_k - \int_{\mathbb{R}^n} \varphi \, d\mu \right| \le C\epsilon$$

for all sufficiently large k. ∎

5.2.3 Compactness

THEOREM 4

Let $U \subset \mathbb{R}^n$ be open and bounded, with ∂U Lipschitz. Assume $\{f_k\}_{k=1}^{\infty}$ is a sequence in $BV(U)$ satisfying

$$\sup_k ||f_k||_{BV(U)} < \infty.$$

Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ and a function $f \in BV(U)$ such that

$$f_{k_j} \to f \text{ in } L^1(U)$$

as $j \to \infty$.

PROOF For $k = 1, 2, \dots$, choose $g_k \in C^\infty(U)$ so that

$$\begin{cases} \int_U |f_k - g_k| \, dx < \frac{1}{k}, \\ \sup_k \int_U |Dg_k| \, dx < \infty; \end{cases} \qquad (\star)$$

such functions exist according to Theorem 2. By the remark following Theorem 1 in Section 4.6 there exist $f \in L^1(U)$ and a subsequence $\{g_{k_j}\}_{j=1}^{\infty}$ such that $g_{k_j} \to f$ in $L^1(U)$. But then $(\star)$ implies also $f_{k_j} \to f$ in $L^1(U)$. According to Theorem 1, $f \in BV(U)$. ∎

5.3 Traces

Assume for this section that U is open and bounded, with ∂U Lipschitz. Observe that since ∂U is Lipschitz, the outer unit normal ν exists $\mathcal{H}^{n-1}$ a.e. on ∂U, according to Rademacher's Theorem.

We now extend to BV functions the notion of trace, defined in Section 4.3 for Sobolev functions.

THEOREM 1
Assume U is open and bounded, with ∂U Lipschitz. There exists a bounded linear mapping

$$T : BV(U) \to L^1(\partial U; \mathcal{H}^{n-1})$$

such that

$$\int_U f \operatorname{div} \varphi \, dx = -\int_U \varphi \cdot d[Df] + \int_{\partial U} (\varphi \cdot \nu) Tf \, d\mathcal{H}^{n-1} \qquad (\star)$$

for all $f \in BV(U)$ and $\varphi \in C^1(\mathbb{R}^n; \mathbb{R}^n)$.

The point is that we do not now require φ to vanish near ∂U.

DEFINITION *The function Tf, which is uniquely defined up to sets of $\mathcal{H}^{n-1} \llcorner \partial U$ measure zero, is called the **trace** of f on ∂U.*

We interpret Tf as the "boundary values" of f on ∂U.

REMARK If $f \in W^{1,1}(U) \subset BV(U)$, the definition of trace above and that from Section 4.3 agree. ▮

PROOF
1. First we introduce some notation:

(a) Given $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, let us write $x = (x', x_n)$ for $x' \equiv (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$, $x_n \in \mathbb{R}$. Similarly we write $y = (y', y_n)$.

(b) Given $x \in \mathbb{R}^n$ and $r, h > 0$, define the open cylinder

$$C(x, r, h) \equiv \{y \in \mathbb{R}^n \mid |y' - x'| < r, |y_n - x_n| < h\}.$$

Now since ∂U is Lipschitz, for each point $x \in \partial U$ there exist $r, h > 0$ and a Lipschitz function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$\max_{|x'-y'| \le r} |\gamma(y') - x_n| \le \frac{h}{4}$$

and — upon rotating and relabeling the coordinate axes if necessary —

$$U \cap C(x, r, h) = \{y \mid |x' - y'| < r, \gamma(y') < y_n < x_n + h\}.$$

2. Assume for the time being $f \in BV(U) \cap C^\infty(U)$. Pick $x \in \partial U$ and choose r, h, γ, etc., as above. Write

$$C \equiv C(x, r, h).$$

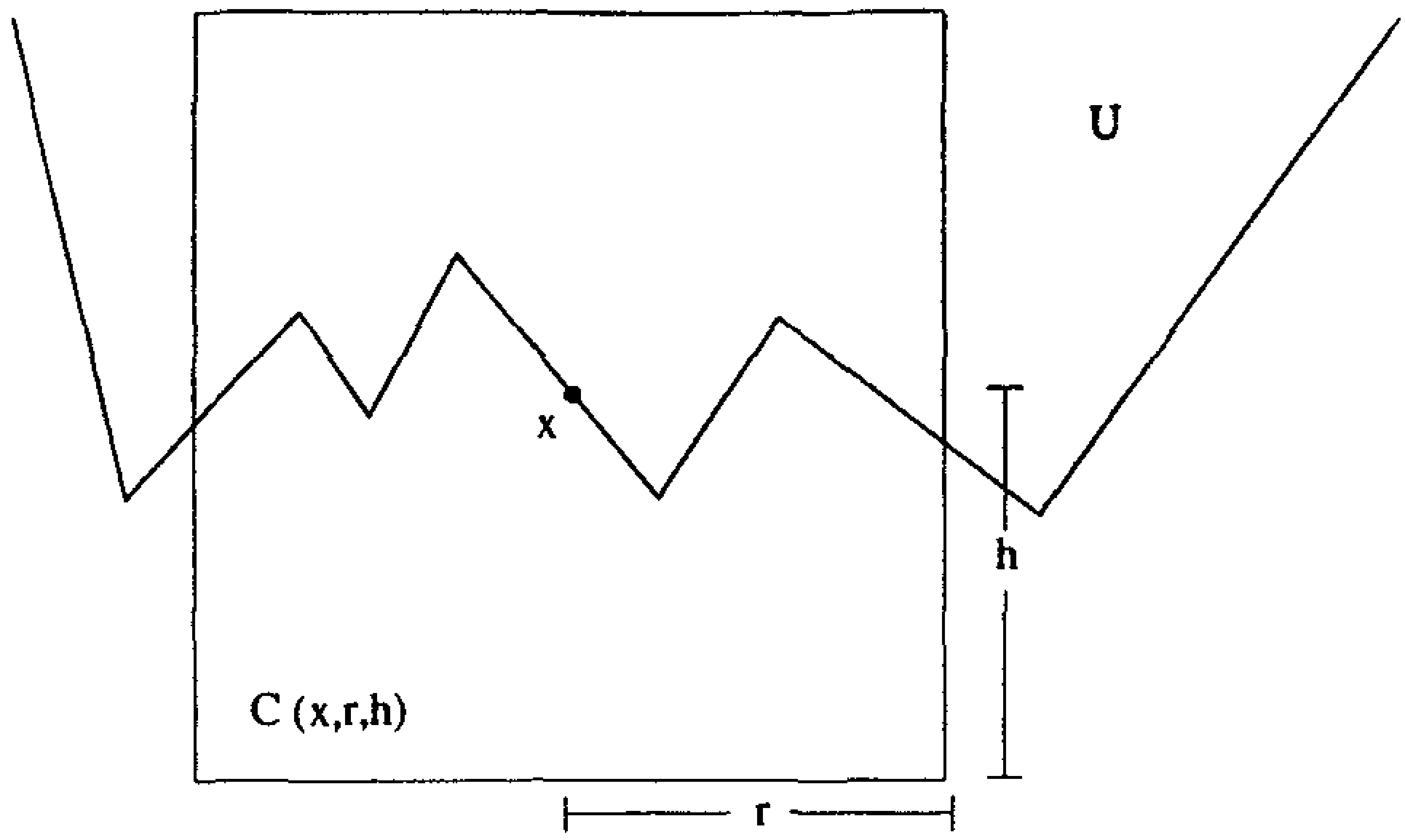

FIGURE 5.1
A Lipschitz boundary within a cylinder.

If $0 < \epsilon < h/2$ and $y \in \partial U \cap C$, we define

$$f_\epsilon(y) \equiv f(y', \gamma(y') + \epsilon).$$

Let us also set

$$C_{\delta,\epsilon} \equiv \{y \in C \mid \gamma(y') + \delta < y_n < \gamma(y') + \epsilon\}$$

for $0 \le \delta < \epsilon < h/2$, and define $C_\epsilon \equiv C_{0,\epsilon}$. Write $C^\epsilon \equiv (C \cap U) - C_\epsilon$.

Then

$$\begin{aligned} |f_\delta(y) - f_\epsilon(y)| &\le \int_\delta^\epsilon \left| \frac{\partial f}{\partial x_n}(y', \gamma(y') + t) \right| dt \\ &\le \int_\delta^\epsilon |Df(y', \gamma(y') + t)|\, dt, \end{aligned}$$

and consequently, since γ is Lipschitz, the Area Formula, Section 3.3, implies

$$\int_{\partial U \cap C} |f_\delta - f_\epsilon|\, d\mathcal{H}^{n-1} \le C \int_{C_{\delta,\epsilon}} |Df|\, dy = C\|Df\|(C_{\delta,\epsilon}).$$

Therefore $\{f_\epsilon\}_{\epsilon>0}$ is Cauchy in $L^1(\partial U \cap C; \mathcal{H}^{n-1})$, and thus the limit

$$Tf \equiv \lim_{\epsilon \to 0} f_\epsilon$$

exists in this space. Furthermore, our passing to limits as $\delta \to 0$ in the foregoing

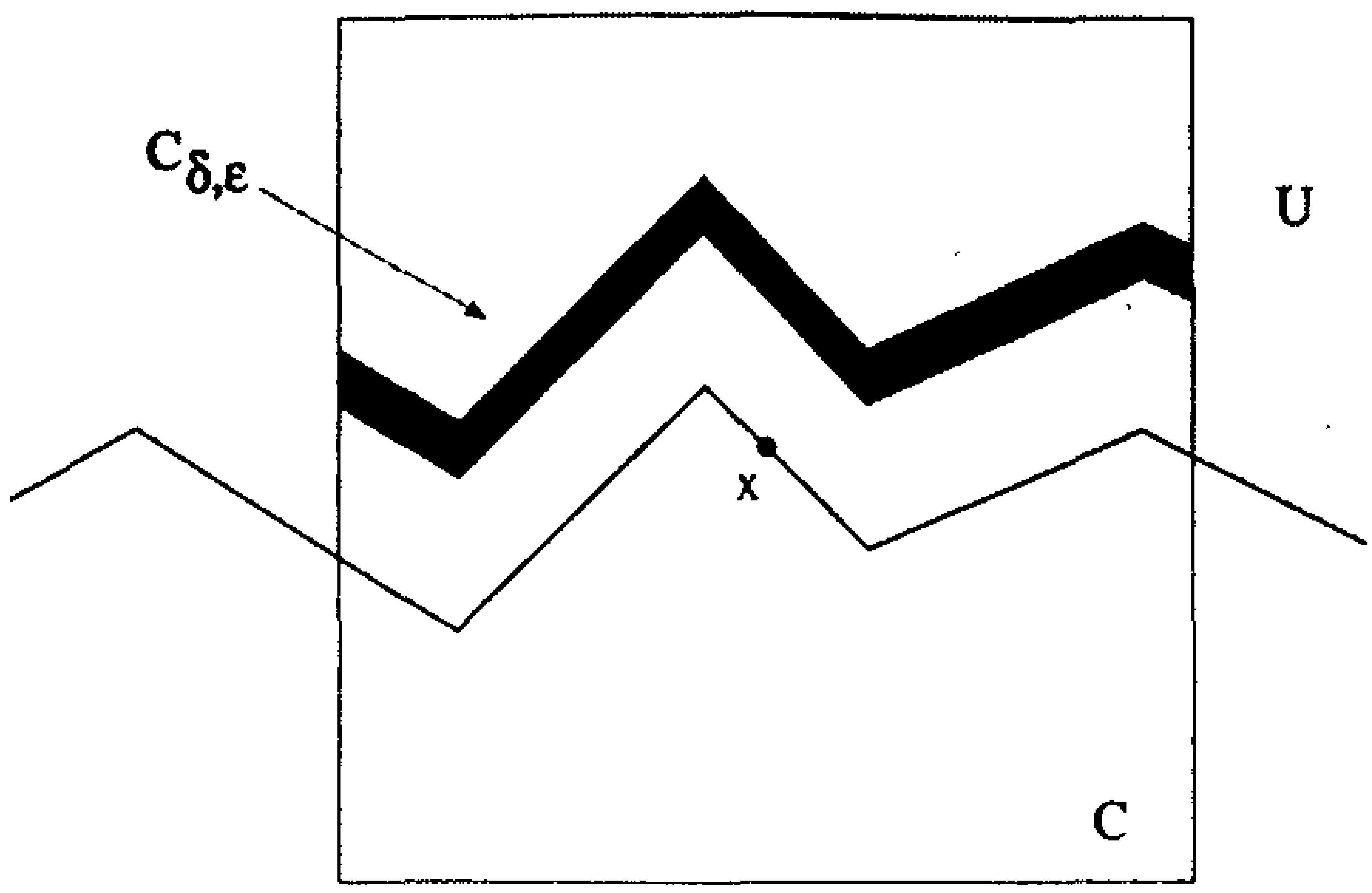

FIGURE 5.2
The $||Df||$ measure of the shaded region $C_{\delta,\epsilon}$ goes to zero as $\epsilon, \delta \to 0$.

inequality yields also

$$\int_{\partial U \cap C} |Tf - f_\epsilon| \, d\mathcal{H}^{n-1} \leq C||Df||(C_\epsilon). \tag{$\star\star$}$$

Next fix $\varphi \in C_c^1(C; \mathbb{R}^n)$. Then

$$\int_{C^\epsilon} f \operatorname{div} \varphi \, dy = -\int_{C^\epsilon} \varphi \cdot Df \, dy + \int_{\partial U \cap C} f_\epsilon \varphi_\epsilon \cdot \nu \, d\mathcal{H}^{n-1}.$$

Let $\epsilon \to 0$ to find

$$\int_{U \cap C} f \operatorname{div} \varphi \, dy = -\int_{U \cap C} \varphi \cdot \sigma \, d||Df|| + \int_{\partial U \cap C} Tf \varphi \cdot \nu \, d\mathcal{H}^{n-1}. \tag{$\star\star\star$}$$

3. Since ∂U is compact, we can cover ∂U with finitely many cylinders $C_i = C(x_i, r_i, h_i)$ $(i = 1, \ldots, N)$ for which assertions analogous to ($\star\star$) and ($\star\star\star$) hold. A straightforward argument using a partition of unity subordinate to the $\{C_i\}_{i=1}^{\infty}$ then establishes formula ($\star$). Observe also that ($\star\star\star$) shows the definition of "Tf" to be the same (up to sets of $\mathcal{H}^{n-1} \llcorner \partial U$ measure zero) on any part of ∂U that happens to lie in two or more of the cylinders C_i.

4. Now assume only $f \in BV(U)$. In this general case, choose $f_k \in BV(U) \cap C^\infty(U)$ $(k = 1, 2, \ldots)$ such that

$$f_k \to f \text{ in } L^1(U), \qquad ||Df_k||(U) \to ||Df||(U)$$

and

$$\mu_k \rightharpoonup \mu \text{ weakly,}$$

where the measures $\{\mu_k\}_{k=1}^{\infty}$, μ are defined as in Theorem 3 of Section 5.2.2.

5. *Claim*: $\{Tf_k\}_{k=1}^{\infty}$ is a Cauchy sequence in $L^1(\partial U; \mathcal{H}^{n-1})$.

Proof of Claim: Choose a cylinder C as in the previous part of the proof. Fix $\epsilon > 0$, $y \in \partial U \cap C$, and then define

$$\begin{aligned} f_k^\epsilon(y) &\equiv \frac{1}{\epsilon}\int_0^\epsilon f_k(y', \gamma(y') + t)\, dt \\ &= \frac{1}{\epsilon}\int_0^\epsilon (f_k)_t(y)\, dt. \end{aligned}$$

Then $(\star\star)$ implies

$$\begin{aligned} \int_{\partial U \cap C} |Tf_k - f_k^\epsilon|\, d\mathcal{H}^{n-1} &\leq \frac{1}{\epsilon}\int_0^\epsilon \int_{\partial U \cap C} |Tf_k - (f_k)_t|\, d\mathcal{H}^{n-1}\, dt \\ &\leq C\|Df_k\|(C_\epsilon). \end{aligned}$$

Thus

$$\begin{aligned} \int_{\partial U \cap C} |Tf_k - Tf_l|\, d\mathcal{H}^{n-1} &\leq \int_{\partial U \cap C} |Tf_k - f_k^\epsilon|\, d\mathcal{H}^{n-1} \\ &\quad + \int_{\partial U \cap C} |Tf_l - f_l^\epsilon|\, d\mathcal{H}^{n-1} \\ &\quad + \int_{\partial U \cap C} |f_k^\epsilon - f_l^\epsilon|\, d\mathcal{H}^{n-1} \\ &\leq C(\|Df_k\| + \|Df_l\|)(C_\epsilon) \\ &\quad + \frac{C}{\epsilon}\int_{C_\epsilon} |f_k - f_l|\, dy. \end{aligned}$$

and so

$$\limsup_{k,l\to\infty} \int_{\partial U \cap C} |Tf_k - Tf_l|\, d\mathcal{H}^{n-1} \leq C\|Df\|(\overline{C}_\epsilon \cap U).$$

Since the quantity on the right-hand side goes to zero as $\epsilon \to 0$, the claim is proved.

6. In view of the claim, we may define

$$Tf \equiv \lim_{k\to\infty} Tf_k;$$

this definition does not depend on the particular choice of approximating sequence.

Finally, formula (⋆) holds for each f_k and thus also holds in the limit for f. ∎

THEOREM 2
Assume U is open, bounded, with ∂U Lipschitz. Suppose also $f \in BV(U)$. Then for $\mathcal{H}^{n-1}$ a.e. $x \in \partial U$,

$$\lim_{r\to 0} \fint_{B(x,r)\cap U} |f - Tf(x)|\, dy = 0,$$

and so

$$Tf(x) = \lim_{r\to 0} \fint_{B(x,r)\cap U} f\, dy.$$

REMARK Thus in particular if $f \in BV(U)\cap C(\overline{U})$, then

$$Tf = f|_{\partial U} \qquad \mathcal{H}^{n-1} \text{ a.e.} \quad ∎$$

PROOF
1. Claim: For $\mathcal{H}^{n-1}$ a.e. $x \in \partial U$,

$$\lim_{r\to 0} \frac{\|Df\|(B(x,r)\cap U)}{r^{n-1}} = 0.$$

Proof of Claim: Fix $\gamma > 0$, $\delta > \epsilon > 0$, and let

$$A_\gamma \equiv \left\{ x \in \partial U \mid \limsup_{r\to 0} \frac{\|Df\|(B(x,r)\cap U)}{r^{n-1}} > \gamma \right\}.$$

Then for each $x \in A_\gamma$, there exists $0 < r < \epsilon$ such that

$$\frac{\|Df\|(B(x,r)\cap U)}{r^{n-1}} \geq \gamma. \tag{$\star$}$$

Using Vitali's Covering Theorem, we obtain a countable collection of disjoint balls $\{B(x_i,r_i)\}_{i=1}^{\infty}$ satisfying (⋆), such that

$$A_\gamma \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_i).$$

Then

$$\begin{aligned}\mathcal{H}^{n-1}_{10\delta}(A_\gamma) &\leq \sum_{i=1}^{\infty} \alpha(n-1)(5r_i)^{n-1}\\ &\leq \frac{C}{\gamma}\sum_{i=1}^{\infty} ||Df||(B(x_i,r_i)\cap U)\\ &\leq C||Df||(U^\epsilon),\end{aligned}$$

where

$$U^\epsilon \equiv \{x \in U \mid \text{dist}(x, \partial U) < \epsilon\}.$$

Send $\epsilon \to 0$ to find $\mathcal{H}^{n-1}_{10\delta}(A_\gamma) = 0$ for all $\delta > 0$.

2. Now fix a point $x \in \partial U$ such that

$$\lim_{r\to 0} \frac{||Df||(B(x,r)\cap U)}{r^{n-1}} = 0,$$

$$\lim_{r\to 0} ⨍_{B(x,r)\cap\partial U} |Tf - Tf(x)|\, d\mathcal{H}^{n-1} = 0.$$

According to the claim and the Lebesgue–Besicovitch Differentiation Theorem, $\mathcal{H}^{n-1}$ a.e. $x \in \partial U$ will do. Let $h = h(r) \equiv 2\max(1, 4\text{Lip }(\gamma))r$, and consider the cylinders

$$C(r) = C(x, r, h).$$

Observe that for sufficiently small r, the cylinders $C(r)$ work in place of the cylinder C in the previous proof. Thus estimates similar to those developed in that proof show

$$\int_{\partial U\cap C(r)} |Tf - f_\epsilon|\, d\mathcal{H}^{n-1} \leq C||Df||(C(r)\cap U),$$

where

$$f_\epsilon(y) \equiv f(y', \gamma(y') + \epsilon) \qquad \left(y \in C(r)\cap\partial U,\ 0 < \epsilon < \frac{h(r)}{2}\right).$$

Consequently, we may employ the Coarea Formula to estimate

$$\int_{B(x,r)\cap U} |Tf(y', \gamma(y')) - f(y)|\, dy \leq Cr||Df||(C(r)\cap U).$$

Hence we compute

$$\begin{aligned}\fint_{B(x,r)\cap U} |f(y) - Tf(x)|\, dy &\le \frac{C}{r^{n-1}} \int_{C(r)\cap\partial U} |Tf - Tf(x)|\, d\mathcal{H}^{n-1} \\ &\quad + \frac{C}{r^n} \int_{B(x,r)\cap U} |Tf(y', \gamma(y')) - f(y)|\, dy \\ &\le o(1) + \frac{C}{r^{n-1}} \|Df\|(C(r)\cap U) \\ &= o(1) \text{ as } r \to 0, \qquad \text{by } (\star\star). \quad \blacksquare\end{aligned}$$

5.4 Extensions

THEOREM 1

Assume $U \subset \mathbb{R}^n$ is open and bounded, with ∂U Lipschitz. Let $f_1 \in BV(U)$, $f_2 \in BV(\mathbb{R}^n - \bar{U})$.

Define

$$\bar{f}(x) \equiv \begin{cases} f_1(x) & x \in U \\ f_2(x) & x \in \mathbb{R}^n - \bar{U}. \end{cases}$$

Then

$$\bar{f} \in BV(\mathbb{R}^n)$$

and

$$\|D\bar{f}\|(\mathbb{R}^n) = \|Df_1\|(U) + \|Df_2\|(\mathbb{R}^n - \bar{U}) + \int_{\partial U} |Tf_1 - Tf_2|\, d\mathcal{H}^{n-1}.$$

REMARK In particular, under the stated assumptions on U,

(i) the extension

$$Ef \equiv \begin{cases} f & \text{on } U \\ 0 & \text{on } \mathbb{R}^n - U \end{cases}$$

belongs to $BV(\mathbb{R}^n)$ provided $f \in BV(U)$, and

(ii) the set U has finite perimeter and $\|\partial U\|(\mathbb{R}^n) = \mathcal{H}^{n-1}(\partial U)$. $\blacksquare$

PROOF

1. Let $\varphi \in C_c^1(\mathbb{R}^n, \mathbb{R}^n)$, $|\varphi| \le 1$. Then

$$\begin{aligned}
\int_{\mathbb{R}^n} \bar{f} \operatorname{div} \varphi \, dx &= \int_U f_1 \operatorname{div} \varphi \, dx + \int_{\mathbb{R}^n - \overline{U}} f_2 \operatorname{div} \varphi \, dx \\
&= -\int_U \varphi \cdot d[Df_1] - \int_{\mathbb{R}^n - \overline{U}} \varphi \cdot d[Df_2] \\
&\quad + \int_{\partial U} (Tf_1 - Tf_2) \varphi \cdot \nu \, d\mathcal{H}^{n-1} \\
&\le \|Df_1\|(U) + \|Df_2\|(\mathbb{R}^n - \overline{U}) + \int_{\partial U} |Tf_1 - Tf_2| \, d\mathcal{H}^{n-1}.
\end{aligned}$$

Thus $\bar{f} \in BV(\mathbb{R}^n)$ and

$$\|D\bar{f}\|(\mathbb{R}^n) \le \|Df_1\|(U) + \|Df_2\|(\mathbb{R}^n - \overline{U}) + \int_{\partial U} |Tf_1 - Tf_2| \, d\mathcal{H}^{n-1}.$$

2. To show equality, observe

$$\begin{aligned}
-\int_{\mathbb{R}^n} \varphi \cdot d[D\bar{f}] &= -\int_U \varphi \cdot d[Df_1] - \int_{\mathbb{R}^n - \overline{U}} \varphi \cdot d[Df_2] \\
&\quad + \int_{\partial U} (Tf_1 - Tf_2) \varphi \cdot \nu \, d\mathcal{H}^{n-1} \qquad (\star)
\end{aligned}$$

for all $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$. Thus

$$[D\bar{f}] = \begin{cases} [Df_1] & \text{on } U \\ [Df_2] & \text{on } \mathbb{R}^n - \overline{U}. \end{cases}$$

Consequently, $(\star)$ implies

$$-\int_{\partial U} \varphi \cdot d[D\bar{f}] = \int_{\partial U} (Tf_1 - Tf_2) \varphi \cdot \nu \, d\mathcal{H}^{n-1},$$

and so

$$\|D\bar{f}\|(\partial U) = \int_{\partial U} |Tf_1 - Tf_2| \, d\mathcal{H}^{n-1}. \quad \blacksquare$$

5.5 Coarea Formula for BV functions

Next we relate the variation measure of f and the perimeters of its level sets.

NOTATION For $f : U \to \mathbb{R}$ and $t \in \mathbb{R}$, define

$$E_t \equiv \{x \in U \mid f(x) > t\}.$$

LEMMA 1
If $f \in BV(U)$, the mapping

$$t \mapsto \|\partial E_t\|(U) \qquad (t \in \mathbb{R})$$

is $\mathcal{L}^1$-measurable.

PROOF The mapping

$$(x,t) \mapsto \chi_{E_t}(x)$$

is $(\mathcal{L}^n \times \mathcal{L}^1)$-measurable, and thus, for each $\varphi \in C_c^1(U;\mathbb{R}^n)$, the function

$$t \mapsto \int_{E_t} \operatorname{div} \varphi \, dx = \int_U \chi_{E_t} \operatorname{div} \varphi \, dx$$

is $\mathcal{L}^1$-measurable. Let D denote any countable dense subset of $C_c^1(U;\mathbb{R}^n)$
Then

$$t \mapsto \|\partial E_t\|(U) = \sup_{\substack{\varphi \in D \\ |\varphi| \le 1}} \int_{E_t} \operatorname{div} \varphi \, dx$$

is $\mathcal{L}^1$-measurable. ∎

THEOREM 1 COAREA FORMULA FOR BV FUNCTIONS
Let $f \in BV(U)$. Then

(i) *E_t has finite perimeter for $\mathcal{L}^1$ a.e. $t \in \mathbb{R}$ and*
(ii) $\|Df\|(U) = \int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt$.
(iii) *Conversely, if $f \in L^1(U)$ and*

$$\int_{-\infty}^{\infty} \|\partial E_t\|(U)\, dt < \infty,$$

then $f \in BV(U)$.

REMARK Compare this with Proposition 2 in Section 3.4.4. ▮

PROOF Let $\varphi \in C_c^1(U;\mathbb{R}^n)$, $|\varphi| \le 1$.

1. Claim #1: $\int_U f \operatorname{div}\varphi\, dx = \int_{-\infty}^{\infty}\left(\int_{E_t}\operatorname{div}\varphi\, dx\right) dt.$

Proof of Claim #1: First suppose $f \ge 0$, so that

$$f(x) = \int_0^\infty \chi_{E_t}(x)\, dt \qquad (\text{a.e. } x \in U).$$

Thus

$$\begin{aligned}\int_U f \operatorname{div}\varphi\, dx &= \int_U \left(\int_0^\infty \chi_{E_t}(x)\, dt\right)\operatorname{div}\varphi(x)\, dx\\ &= \int_0^\infty \left(\int_U \chi_{E_t}(x)\operatorname{div}\varphi(x)\, dx\right) dt\\ &= \int_0^\infty \left(\int_{E_t}\operatorname{div}\varphi\, dx\right) dt.\end{aligned}$$

Similarly, if $f \le 0$,

$$f(x) = \int_{-\infty}^0 (\chi_{E_t}(x) - 1)\, dt,$$

whence

$$\begin{aligned}\int_U f \operatorname{div}\varphi\, dx &= \int_U \left(\int_{-\infty}^0 (\chi_{E_t}(x) - 1)\, dt\right)\operatorname{div}\varphi(x)\, dx\\ &= \int_{-\infty}^0 \left(\int_U (\chi_{E_t}(x) - 1)\operatorname{div}\varphi(x)\, dx\right) dt\\ &= \int_{-\infty}^0 \left(\int_{E_t}\operatorname{div}\varphi\, dx\right) dt.\end{aligned}$$

For the general case, write $f = f^+ + (-f^-)$.

2. From Claim #1 we see that for all φ as above,

$$\int_U f \operatorname{div}\varphi\, dx \le \int_{-\infty}^\infty ||\partial E_t||(U)\, dt.$$

Hence

$$||Df||(U) \le \int_{-\infty}^\infty ||\partial E_t||(U)\, dt. \qquad (\star)$$

3. Claim #2: Assertion (ii) holds for all $f \in BV(U) \cap C^\infty(U)$.

Proof of Claim #2: Let

$$m(t) \equiv \int_{U-E_t} |Df|\,dx = \int_{\{f\le t\}} |Df|\,dx.$$

Then the function m is nondecreasing, and thus m' exists $\mathcal{L}^1$ a.e., with

$$\int_{-\infty}^{\infty} m'(t)\,dt \le \int_U |Df|\,dx. \qquad (\star\star)$$

Now fix any $-\infty < t < \infty$, $r > 0$, and define $\eta : \mathbb{R} \to \mathbb{R}$ this way:

$$\eta(s) \equiv \begin{cases} 0 & \text{if } s \le t \\ \dfrac{s-t}{r} & \text{if } t \le s \le t+r \\ 1 & \text{if } s \ge t+r. \end{cases}$$

Then

$$\eta'(s) = \begin{cases} \dfrac{1}{r} & \text{if } t < s < t+r \\ 0 & \text{if } s < t \text{ or } s > t+r. \end{cases}$$

Hence, for all $\varphi \in C_c^1(U;\mathbb{R}^n)$,

$$\begin{aligned} -\int_U \eta(f(x))\operatorname{div}\varphi\,dx &= \int_U \eta'(f(x))Df\cdot\varphi\,dx \\ &= \frac{1}{r}\int_{E_t-E_{t+r}} Df\cdot\varphi\,dx. \qquad (\star\star\star) \end{aligned}$$

Now

$$\begin{aligned} \frac{m(t+r)-m(t)}{r} &= \frac{1}{r}\left[\int_{U-E_{t+r}} |Df|\,dx - \int_{U-E_t} |Df|\,dx\right] \\ &= \frac{1}{r}\int_{E_t-E_{t+r}} |Df|\,dx \\ &\ge \frac{1}{r}\int_{E_t-E_{t+r}} Df\cdot\varphi\,dx \\ &= -\int_U \eta(f(x))\operatorname{div}\varphi\,dx \qquad \text{by } (\star\star\star). \end{aligned}$$

For those t such that $m'(t)$ exists, we then let $r \to 0$:

$$m'(t) \ge -\int_{E_t} \operatorname{div}\varphi\,dx \qquad \mathcal{L}^n \text{ a.e. } t.$$

Take the supremum over all φ as above:

$$||\partial E_t||(U) \le m'(t),$$

and recall $(\star\star)$ to find

$$\int_{-\infty}^{\infty} ||\partial E_t||(U)\,dt \le \int_U |Df|\,dx = ||Df||(U).$$

This estimate and $(\star)$ complete the proof.

4. *Claim #3*: Assertion (ii) holds for each function $f \in BV(U)$.

Proof of Claim #3: Fix $f \in BV(U)$ and choose $\{f_k\}_{k=1}^{\infty}$ as in Theorem 2 in Section 5.2.2. Then

$$f_k \to f \qquad \text{in } L^1(U) \text{ as } k \to \infty.$$

Define

$$E_t^k \equiv \{x \in U \mid f_k(x) > t\}.$$

Now

$$\int_{-\infty}^{\infty} |\chi_{E_t^k}(x) - \chi_{E_t}(x)|\,dt = \int_{\min\{f(x),f_k(x)\}}^{\max\{f(x),f_k(x)\}} dt = |f_k(x) - f(x)|;$$

consequently,

$$\int_U |f_k(x) - f(x)|\,dx = \int_{-\infty}^{\infty} \left(\int_U \left|\chi_{E_t^k}(x) - \chi_{E_t}(x)\right|\,dx\right)\,dt.$$

Since $f_k \to f$ in $L^1(U)$, there exists a subsequence which, upon reindexing by k if needs be, satisfies

$$\chi_{E_t^k} \to \chi_{E_t} \text{ in } L^1(U), \qquad \text{for } \mathcal{L}^1 \text{ a.e. } t.$$

Then, by the Lower Semicontinuity Theorem,

$$||\partial E_t||(U) \le \liminf_{k\to\infty} ||\partial E_t^k||(U).$$

Thus Fatou's Lemma implies

$$\begin{aligned}\int_{-\infty}^{\infty} ||\partial E_t||(U)\,dt &\le \liminf_{k\to\infty} \int_{-\infty}^{\infty} ||\partial E_t^k||(U)\,dt \\ &= \lim_{k\to\infty} ||Df_k||(U) \\ &= ||Df||(U).\end{aligned}$$

This calculation and $(\star)$ complete the proof. ∎

5.6 Isoperimetric Inequalities

We now develop certain inequalities relating the $\mathcal{L}^n$-measure of a set and its perimeter.

5.6.1 Sobolev's and Poincaré's inequalities for BV

THEOREM 1

(i) *There exists a constant C_1 such that*

$$||f||_{L^{n/n-1}(\mathbb{R}^n)} \leq C_1||Df||(\mathbb{R}^n)$$

for all $f \in BV(\mathbb{R}^n)$.

(ii) *There exists a constant C_2 such that*

$$||f - (f)_{x,r}||_{L^{n/n-1}(B(x,r))} \leq C_2||Df||(U(x,r))$$

for all $B(x,r) \subset \mathbb{R}^n$, $f \in BV_{\mathrm{loc}}(\mathbb{R}^n)$, where $(f)_{x,r} = ⨍_{B(x,r)} f\, dy$.

(iii) *For each $0 < \alpha \leq 1$, there exists a constant $C_3(\alpha)$ such that*

$$||f||_{L^{n/n-1}(B(x,r))} \leq C_3(\alpha)||Df||(U(x,r))$$

for all $B(x,r) \subset \mathbb{R}^n$ and all $f \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ satisfying

$$\frac{\mathcal{L}^n(B(x,r) \cap \{f = 0\})}{\mathcal{L}^n(B(x,r))} \geq \alpha.$$

PROOF

1. Choose $f_k \in C_c^\infty(\mathbb{R}^n)$ $(k = 1, \ldots)$ so that

$$\begin{cases} f_k \to f \text{ in } L^1(\mathbb{R}^n), \qquad f_k \to f \qquad \mathcal{L}^n \text{ a.e.} \\ ||Df_k||(\mathbb{R}^n) \to ||Df||(\mathbb{R}^n). \end{cases}$$

Then by Fatou's Lemma and the Gagliardo–Nirenberg–Sobolev inequality,

$$\begin{aligned} ||f||_{L^{n/n-1}(\mathbb{R}^n)} &\leq \liminf_{k\to\infty} ||f_k||_{L^{n/n-1}(\mathbb{R}^n)} \\ &\leq \lim_{k\to\infty} C_1||Df_k||_{L^1(\mathbb{R}^n)} \\ &= C_1||Df||(\mathbb{R}^n). \end{aligned}$$

This proves (i).

2. Statement (ii) follows similarly from Poincaré's inequality, Section 4.5.2.
3. Suppose

$$\frac{\mathcal{L}^n(B(x,r)\cap\{f=0\})}{\mathcal{L}^n(B(x,r))}\geq\alpha>0. \tag{$\star$}$$

Then

$$\begin{aligned}||f||_{L^{n/n-1}(B(x,r))} &\leq ||f-(f)_{x,r}||_{L^{n/n-1}(B(x,r))}+||(f)_{x,r}||_{L^{n/n-1}(B(x,r))}\\ &\leq C_2||Df||(U(x,r))+|(f)_{x,r}|(\mathcal{L}^n(B(x,r)))^{1-1/n}. \qquad (\star\star)\end{aligned}$$

But

$$\begin{aligned}&|(f)_{x,r}|(\mathcal{L}^n(B(x,r)))^{1-1/n}\\ &\quad\leq\frac{1}{\mathcal{L}^n(B(x,r))^{1/n}}\int_{B(x,r)\cap\{f\neq0\}}|f|\,dy\\ &\quad\leq\left(\int_{B(x,r)}|f|^{n/n-1}\,dy\right)^{1-1/n}\left(\frac{\mathcal{L}^n(B(x,r)\cap\{f\neq0\})}{\mathcal{L}^n(B(x,r))}\right)^{1/n}\\ &\quad\leq||f||_{L^{n/n-1}(B(x,r))}(1-\alpha)^{1/n},\end{aligned}$$

by ($\star$). We employ this estimate in ($\star\star$) to compute

$$||f||_{L^{n/n-1}(B(x,r))}\leq\frac{C_2}{(1-(1-\alpha)^{1/n})}||Df||(U(x,r)). \quad \blacksquare$$

5.6.2 Isoperimetric Inequalities

THEOREM 2
Let E be a bounded set of finite perimeter in $\mathbb{R}^n$. Then

(i) $\mathcal{L}^n(E)^{1-1/n}\leq C_1||\partial E||(\mathbb{R}^n)$, *and*
(ii) for each ball $B(x,r)\subset\mathbb{R}^n$,

$$\min\{\mathcal{L}^n(B(x,r)\cap E),\mathcal{L}^n(B(x,r)-E)\}^{1-\frac{1}{n}}\leq 2C_2||\partial E||(U(x,r)).$$

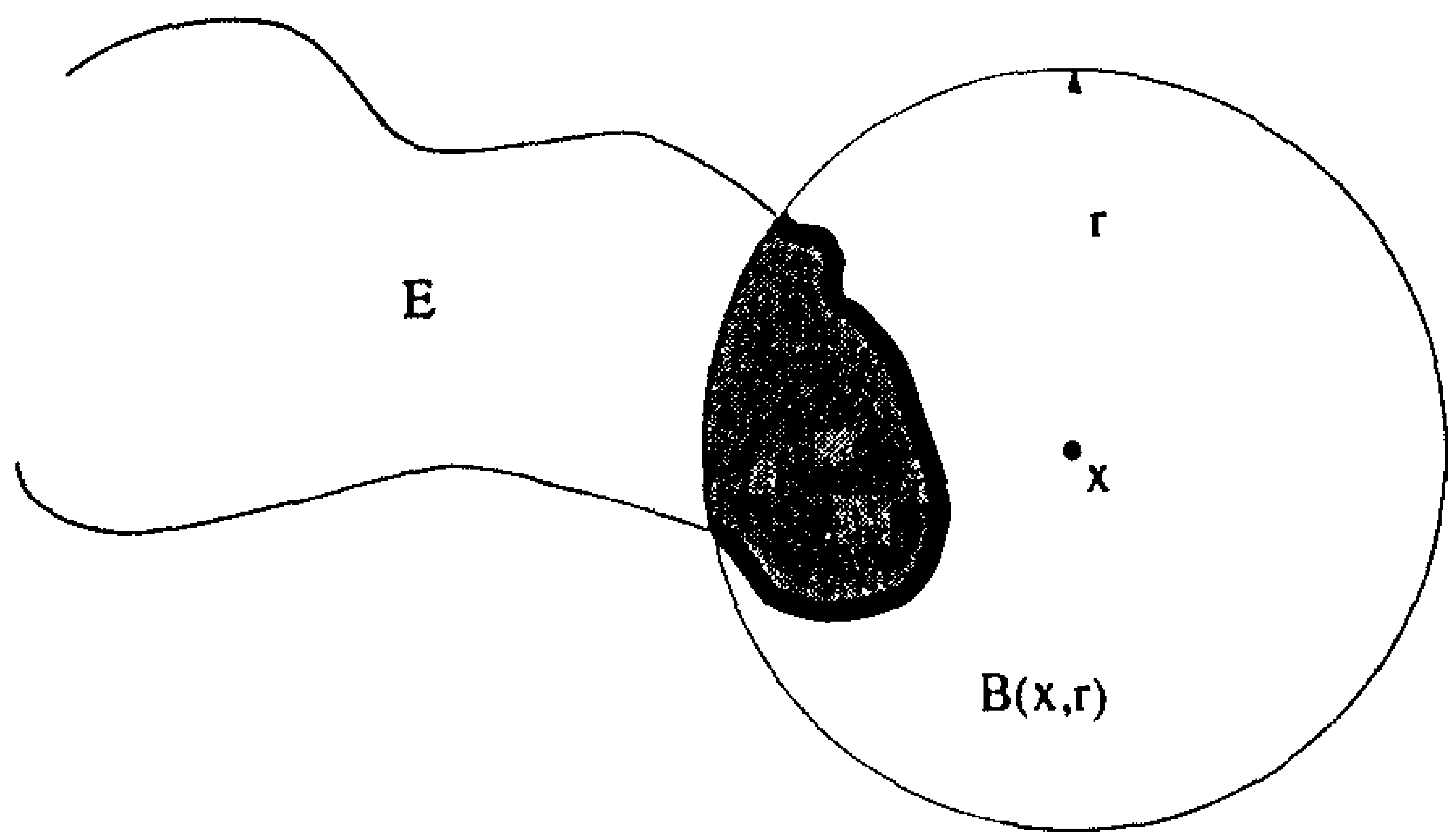

FIGURE 5.3
Relative Isoperimetric Inequality.

REMARK Statement (i) is the *Isoperimetric Inequality* and (ii) is the *Relative Isoperimetric Inequality*. The constants C_1 and C_2 are those from Theorems 1 and 2 in Section 4.5. ∎

PROOF

1. Let $f = \chi_E$ in assertion (i) of Theorem 1 to prove (i).

2. Let $f = \chi_{B(x,r)\cap E}$ in assertion (ii) of Theorem 1, in which case

$$(f)_{x,r} = \frac{\mathcal{L}^n(B(x,r)\cap E)}{\mathcal{L}^n(B(x,r))} .$$

Thus

$$\int_{B(x,r)} |f - (f)_{x,r}|^{n/n-1}\, dy = \left(\frac{\mathcal{L}^n(B(x,r)-E)}{\mathcal{L}^n(B(x,r))}\right)^{n/n-1} \mathcal{L}^n(B(x,r)\cap E)$$
$$+\left(\frac{\mathcal{L}^n(B(x,r)\cap E)}{\mathcal{L}^n(B(x,r))}\right)^{n/n-1} \mathcal{L}^n(B(x,r)-E).$$

Now if $\mathcal{L}^n(B(x,r)\cap E) \le \mathcal{L}^n(B(x,r)-E)$, then

$$\begin{aligned}\left(\int_{B(x,r)} |f-(f)_{x,r}|^{n/n-1}\,dy\right)^{1-1/n} &\\ &\ge \left[\frac{\mathcal{L}^n(B(x,r)-E)}{\mathcal{L}^n(B(x,r))}\right]\mathcal{L}^n(B(x,r)\cap E)^{1-1/n}\\ &\ge \frac{1}{2}\min\{\mathcal{L}^n(B(x,r)\cap E),\mathcal{L}^n(B(x,r)-E)\}^{1-1/n}.\end{aligned}$$

The other case is similar. ∎

REMARK We have shown that the Gagliardo–Nirenberg–Sobolev Inequality implies the Isoperimetric Inequality. In fact, *the converse is true as well.*

To see this, assume $f \in C_c^1(\mathbb{R}^n)$, $f \ge 0$. We calculate

$$\begin{aligned}\int_{\mathbb{R}^n} |Df|\,dx = \|Df\|(\mathbb{R}^n) &= \int_{-\infty}^{\infty} \|\partial E_t\|(\mathbb{R}^n)\,dt\\ &= \int_0^{\infty} \|\partial E_t\|(\mathbb{R}^n)\,dt\\ &\ge \frac{1}{C_1}\int_0^{\infty} \mathcal{L}^n(E_t)^{1-1/n}\,dt.\end{aligned}$$

Now let

$$f_t \equiv \min\{t,f\}, \qquad \chi(t) \equiv \left(\int_{\mathbb{R}^n} f_t^{n/n-1}\,dx\right)^{1-1/n} \qquad (t \in \mathbb{R}).$$

Then χ is nondecreasing on $(0,\infty)$ and

$$\lim_{t\to\infty}\chi(t) = \left(\int_{\mathbb{R}^n} |f|^{n/n-1}\,dx\right)^{1-1/n}$$

Also, for $h > 0$,

$$\begin{aligned}0 &\le \chi(t+h)-\chi(t)\\ &\le \left(\int_{\mathbb{R}^n} |f_{t+h}-f_t|^{n/n-1}\,dx\right)^{1-1/n}\\ &\le h\mathcal{L}^n(E_t)^{1-1/n}.\end{aligned}$$

Thus χ is locally Lipschitz, and

$$\chi'(t) \le \mathcal{L}^n(E_t)^{1-1/n} \qquad \mathcal{L}^1 \text{ a.e. } t.$$

Integrate from 0 to ∞:

$$\begin{aligned}\left(\int_{\mathbb{R}^n} |f|^{n/n-1}\,dx\right)^{1-1/n} &= \int_0^\infty \chi'(t)\,dt \\ &\le \int_0^\infty \mathcal{L}^n(E_t)^{n/n-1}\,dt \\ &\le C_1 \int_{\mathbb{R}^n} |Df|\,dx. \quad \blacksquare\end{aligned}$$

5.6.3 $\mathcal{H}^{n-1}$ and Cap_1

As a first application of the Isoperimetric Inequalities, we establish this refinement of Theorem 4 in Section 4.7.2:

THEOREM 3
Assume $A \subset \mathbb{R}^n$ is compact. Then $\mathrm{Cap}_1(A) = 0$ *if and only if* $\mathcal{H}^{n-1}(A) = 0$.

PROOF According to Theorem 2 in Section 4.7.1, $\mathrm{Cap}_1(A)=0$ if $\mathcal{H}^{n-1}(A)=0$.
Now suppose $\mathrm{Cap}_1(A) = 0$. If $f \in K^1$ and $A \subset \{f \ge 1\}^o$, then by Theorem 1 in Section 5.5,

$$\int_0^1 \|\partial E_t\|(\mathbb{R}^n)\,dt \le \int_{\mathbb{R}^n} |Df|\,dx$$

where $E_t \equiv \{f > t\}$. Thus for some $t \in (0,1)$,

$$\|\partial E_t\|(\mathbb{R}^n) \le \int_{\mathbb{R}^n} |Df|\,dx.$$

Clearly $A \subset E_t{}^o$, and by the Isoperimetric Inequality, $\mathcal{L}^n(E_t) < \infty$. Thus for each $x \in A$, there exists a $r > 0$ such that

$$\frac{\mathcal{L}^n(E_t \cap B(x,r))}{\alpha(n) r^n} = \frac{1}{4}\,.$$

In light of the Relative Isoperimetric Inequality, we have for each such $B(x,r)$,

$$\left[\frac{1}{4}\alpha(n) r^n\right]^{\frac{n-1}{n}} = \left(\mathcal{L}^n(E_t \cap B(x,r))\right)^{\frac{n-1}{n}} \le C\|\partial E_t\|(B(x,r));$$

that is,

$$r^{n-1} \le C\|\partial E_t\|(B(x,r)).$$

By Vitali's Covering Theorem there exists a disjoint collection of balls $\{B(x_j, r_j)\}_{j=1}^{\infty}$ as above, with $x_j \in A$ and

$$A \subset \bigcup_{j=1}^{\infty} B(x_j, 5r_j).$$

Thus

$$\sum_{j=1}^{\infty}(5r_j)^{n-1} \leq C||\partial E_t||(\mathbb{R}^n) \leq C \int_{\mathbb{R}^n} |Df| \, dx.$$

Since $\mathrm{Cap}_1(A) = 0$, given $\epsilon > 0$, the function f can be chosen so that

$$\int_{\mathbb{R}^n} |Df| \, dx \leq \epsilon,$$

and hence for each j,

$$r_j \leq (C||\partial E_t||(\mathbb{R}^n))^{\frac{1}{n-1}} \leq C\epsilon^{\frac{1}{n-1}}.$$

This implies $\mathcal{H}^{n-1}(A) = 0$. ▮

5.7 The reduced boundary

In this and the next section we study the detailed structure of sets of locally finite perimeter. Our goal is to verify that such a set has "a C^1 boundary measure theoretically."

5.7.1 Estimates

We hereafter assume

E is a set of locally finite perimeter in $\mathbb{R}^n$.

Recall the definitions of ν_E, $||\partial E||$, etc., from Section 5.1.

DEFINITION *Let $x \in \mathbb{R}^n$. We say $x \in \partial^* E$, the* ***reduced boundary*** *of E, if*

(i) $||\partial E||(B(x, r)) > 0$ *for all* $r > 0$,

(ii) $\lim_{r\to 0} ⨍_{B(x,r)} \nu_E \, d||\partial E|| = \nu_E(x)$, *and*

(iii) $|\nu_E(x)| = 1$.

REMARK According to Theorem 1 in Section 1.7.1,

$$||\partial E||(\mathbb{R}^n - \partial^* E) = 0. \quad ▮$$

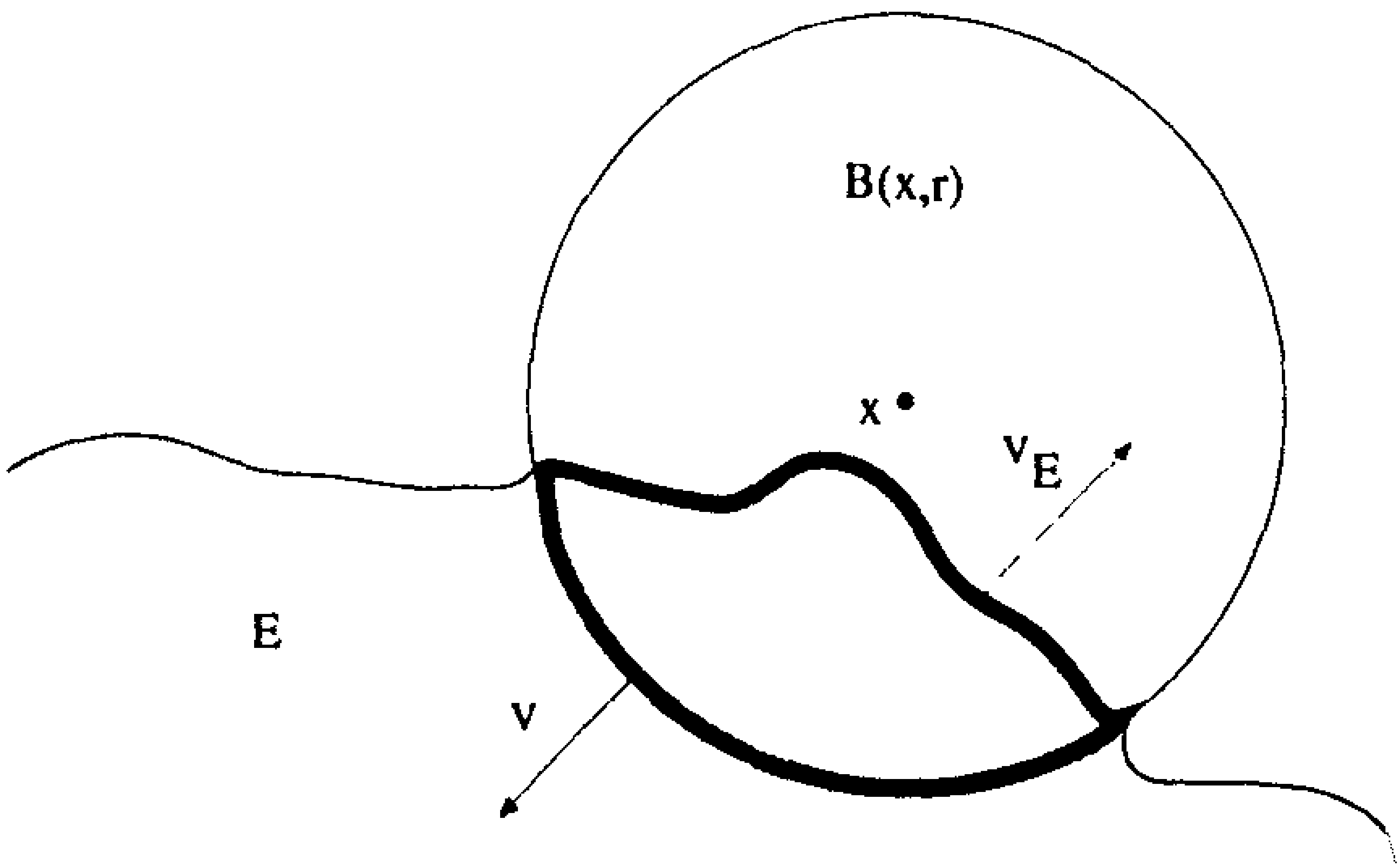

FIGURE 5.4
Normals to E and to $B(x,r)$.

LEMMA 1
Let $\varphi \in C^1_c(\mathbb{R}^n;\mathbb{R}^n)$. Then for each $x \in \mathbb{R}^n$,

$$\int_{E\cap B(x,r)} \operatorname{div}\varphi\, dy = \int_{B(x,r)} \varphi\cdot\nu_E\, d\|\partial E\| + \int_{E\cap\partial B(x,r)} \varphi\cdot\nu\, d\mathcal{H}^{n-1}$$

for $\mathcal{L}^1$ a.e. $r>0$, ν denoting the outward unit normal to $\partial B(x,r)$.

PROOF Assume $h:\mathbb{R}^n\to\mathbb{R}$ is smooth, then

$$\int_E \operatorname{div}(h\varphi)\, dy = \int_E h\operatorname{div}\varphi\, dy + \int_E Dh\cdot\varphi\, dy.$$

Thus

$$\int_{\mathbb{R}^n} h\varphi\cdot\nu_E\, d\|\partial E\| = \int_E h\operatorname{div}\varphi\, dy + \int_E Dh\cdot\varphi\, dy. \qquad (\star)$$

By approximation, $(\star)$ holds also for

$$h_\epsilon(y) \equiv g_\epsilon(|y-x|),$$

where

$$g_\epsilon(s) \equiv \begin{cases} 1 & \text{if } 0 \le s \le r \\ \dfrac{r-s+\epsilon}{\epsilon} & \text{if } r \le s \le r+\epsilon \\ 0 & \text{if } s \ge r+\epsilon. \end{cases}$$

Notice

$$g_\epsilon'(s) = \begin{cases} 0 & \text{if } 0 \le s < r \text{ or } s > r+\epsilon \\ -\dfrac{1}{\epsilon} & \text{if } r < s < r+\epsilon\,, \end{cases}$$

and therefore

$$Dh_\epsilon(y) = \begin{cases} 0 & \text{if } |y-x| < r \text{ or } |y-x| > r+\epsilon \\ -\dfrac{1}{\epsilon}\dfrac{y-x}{|y-x|} & \text{if } r < |y-x| < r+\epsilon\,. \end{cases}$$

Set $h = h_\epsilon$ in $(\star)$:

$$\int_{\mathbb{R}^n} h_\epsilon \varphi \cdot \nu_E \, d\|\partial E\| = \int_E h_\epsilon \operatorname{div} \varphi \, dy - \frac{1}{\epsilon} \int_{E\cap\{y \mid r<|y-x|<r+\epsilon\}} \varphi \cdot \frac{y-x}{|y-x|} \, dy.$$

Let $\epsilon \to 0$ and recall Proposition 1 in Section 3.4.4 to find

$$\int_{B(x,r)} \varphi \cdot \nu_E \, d\|\partial E\| = \int_{E\cap B(x,r)} \operatorname{div} \varphi \, dy - \int_{E\cap\partial B(x,r)} \varphi \cdot \nu \, d\mathcal{H}^{n-1}$$

for $\mathcal{L}^1$ a.e. $r > 0$. ∎

LEMMA 2

There exist positive constants $A_1, \dots, A_5$, depending only on n, such that for each $x \in \partial^* E$,

(i) $\liminf_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap E)}{r^n} > A_1 > 0$,

(ii) $\liminf_{r\to 0} \frac{\mathcal{L}^n(B(x,r)-E)}{r^n} > A_2 > 0$,

(iii) $\liminf_{r\to 0} \frac{\|\partial E\|(B(x,r))}{r^{n-1}} > A_3 > 0$,

(iv) $\limsup_{r\to 0} \frac{\|\partial E\|(B(x,r))}{r^{n-1}} \le A_4$,

(v) $\limsup_{r\to 0} \frac{\|\partial(E\cap B(x,r))\|(\mathbb{R}^n)}{r^{n-1}} \le A_5$.

PROOF

1. Fix $x \in \partial^* E$. According to Lemma 1, for $\mathcal{L}^1$ a.e. $r > 0$

$$\|\partial(E\cap B(x,r))\|(\mathbb{R}^n) \le \|\partial E\|(B(x,r)) + \mathcal{H}^{n-1}(E \cap \partial B(x,r)). \qquad (\star)$$

On the other hand, choose $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$ such that

$$\varphi \equiv \nu_E(x) \text{ on } B(x,r).$$

Then the formula from Lemma 1 reads

$$\int_{B(x,r)} \nu_E(x) \cdot \nu_E \, d\|\partial E\| = -\int_{E\cap\partial B(x,r)} \nu_E(x) \cdot \nu \, d\mathcal{H}^{n-1}. \qquad (\star\star)$$

Since $x \in \partial^* E$,

$$\lim_{r\to 0} \nu_E(x) \cdot \fint_{B(x,r)} \nu_E \, d\|\partial E\| = |\nu_E(x)|^2 = 1;$$

thus for $\mathcal{L}^1$ a.e. sufficiently small $r > 0$, say $0 < r < r_0 = r_0(x)$, $(\star\star)$ implies

$$\frac{1}{2}\|\partial E\|(B(x,r)) \leq \mathcal{H}^{n-1}(E \cap \partial B(x,r)). \qquad (\star\star\star)$$

This and $(\star)$ give

$$\|\partial(E \cap B(x,r))\|(\mathbb{R}^n) \leq 3\mathcal{H}^{n-1}(E \cap \partial B(x,r)) \qquad (\star\star\star\star)$$

for a.e. $0 < r < r_0$.

2. Write $g(r) \equiv \mathcal{L}^n(B(x,r) \cap E)$. Then

$$g(r) = \int_0^r \mathcal{H}^{n-1}(\partial B(x,s) \cap E) \, ds,$$

whence g is absolutely continuous, and

$$g'(r) = \mathcal{H}^{n-1}(\partial B(x,r) \cap E) \qquad \text{for a.e. } r > 0.$$

Using now the Isoperimetric Inequality and $(\star\star\star\star)$, we compute

$$\begin{aligned} g(r)^{1-1/n} &= \mathcal{L}^n(B(x,r) \cap E)^{1-1/n} \leq C\|\partial(B(x,r) \cap E)\|(\mathbb{R}^n) \\ &\leq C\mathcal{H}^{n-1}(B(x,r) \cap E) \\ &= C_1 g'(r) \qquad \text{for a.e. } r \in (0, r_0). \end{aligned}$$

Thus

$$\frac{1}{C_1} \leq g(r)^{(1/n)-1} g'(r) = n(g^{1/n}(r))',$$

and so

$$g^{\frac{1}{n}}(r) \geq \frac{r}{C_1 n}$$

and

$$g(r) \geq \frac{r^n}{(C_1 n)^n}$$

for $0 < r \leq r_0$. This proves assertion (i).

3. Since for all $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$

$$\int_E \operatorname{div} \varphi \, dx + \int_{\mathbb{R}^n - E} \operatorname{div} \varphi \, dx = \int_{\mathbb{R}^n} \operatorname{div} \varphi \, dx = 0,$$

it is easy to check

$$||\partial E|| = ||\partial(\mathbb{R}^n - E)||, \qquad \nu_E = -\nu_{\mathbb{R}^n - E}.$$

Consequently, statement (ii) follows from (i).

4. According to the Relative Isoperimetric Inequality,

$$\frac{||\partial E||(B(x,r))}{r^{n-1}} \geq C \min\left\{\frac{\mathcal{L}^n(B(x,r) \cap E)}{r^n}, \frac{\mathcal{L}^n(B(x,r) - E)}{r^n}\right\}^{\frac{n-1}{n}},$$

and thus assertion (iii) follows from (i), (ii).

5. By $(\star\star\star)$,

$$||\partial E||(B(x,r)) \leq 2\mathcal{H}^{n-1}(E \cap \partial B(x,r)) \leq Cr^{n-1} \qquad (0 < r < r_0);$$

this is (iv).

6. Statement (v) is a consequence of $(\star)$ and (iv). ∎

5.7.2 Blow-up

DEFINITION *For each $x \in \partial^* E$, define the hyperplane*

$$H(x) \equiv \{y \in \mathbb{R}^n \mid \nu_E(x) \cdot (y - x) = 0\}$$

and the half-spaces

$$H^+(x) \equiv \{y \in \mathbb{R}^n \mid \nu_E(x) \cdot (y - x) \geq 0\},$$

$$H^-(x) \equiv \{y \in \mathbb{R}^n \mid \nu_E(x) \cdot (y - x) \leq 0\}.$$

NOTATION Fix $x \in \partial^* E$, $r > 0$, and set

$$E_r \equiv \{y \in \mathbb{R}^n \mid r(y - x) + x \in E\}.$$

REMARK Observe $y \in E \cap B(x,r)$ if and only if $g_r(y) \in E_r \cap B(x,1)$, where $g_r(y) \equiv ((y - x)/r) + x$. ∎

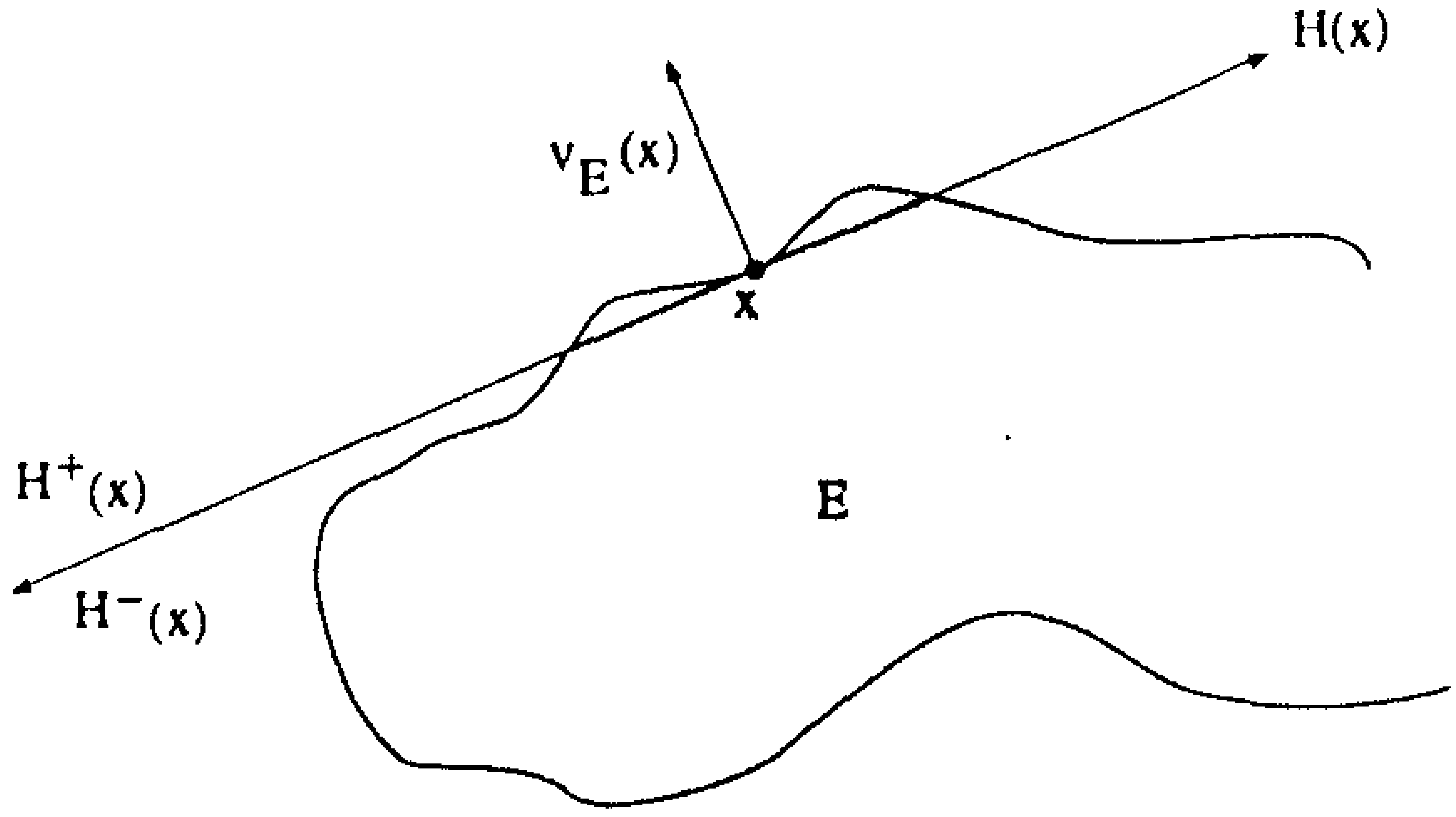

FIGURE 5.5
Approximate tangent plane.

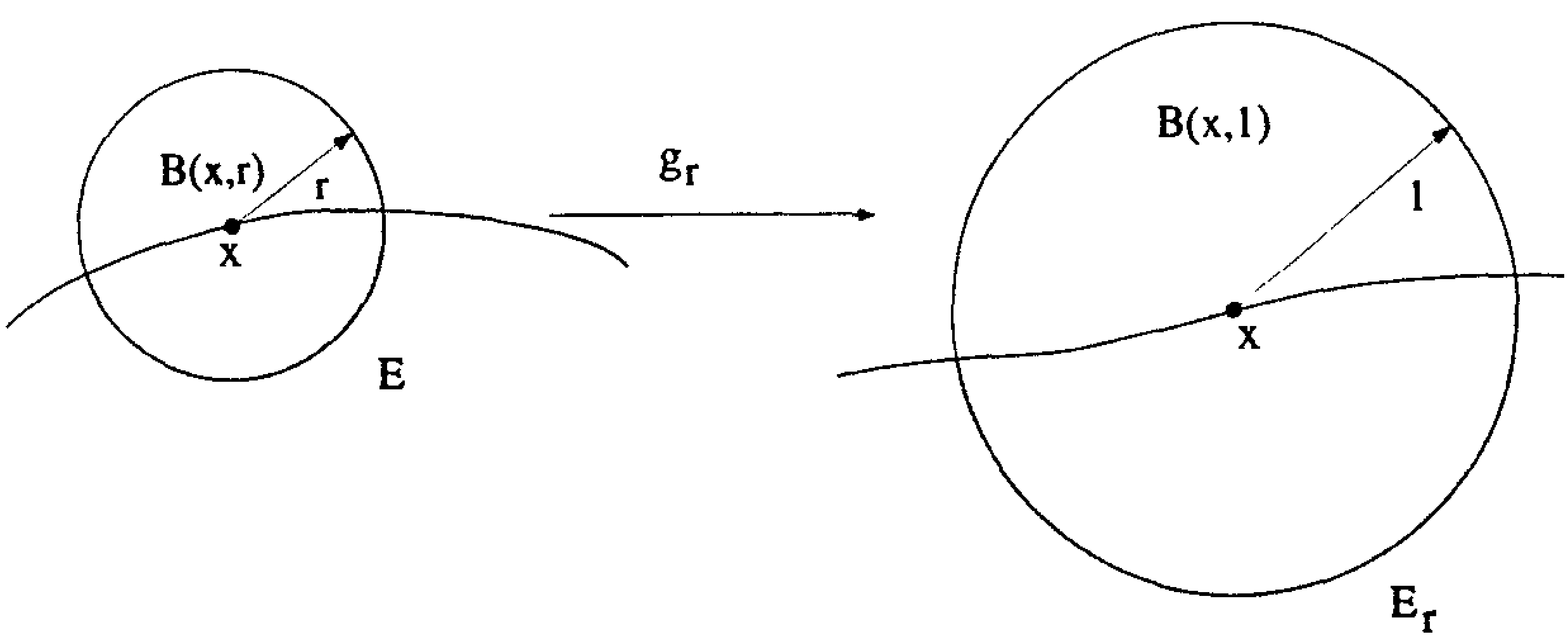

FIGURE 5.6
Blow-up.

THEOREM 1 BLOW-UP OF REDUCED BOUNDARY
Assume $x \in \partial^* E$. Then

$$\chi_{E_r} \to \chi_{H^-(x)} \text{ in } L^1_{\mathrm{loc}}(\mathbb{R}^n)$$

as $r \to 0$.

Thus for small enough $r > 0$, $E \cap B(x,r)$ approximately equals the half ball $H^-(x) \cap B(x,r)$.

PROOF

1. First of all, we may as well assume:

$$\begin{cases} x = 0, \ \nu_E(0) = e_n = (0, \ldots, 0, 1), \\ H(0) = \{y \in \mathbb{R}^n \mid y_n = 0\}, \\ H^+(0) = \{y \in \mathbb{R}^n \mid y_n \geq 0\}, \\ H^-(0) = \{y \in \mathbb{R}^n \mid y_n \leq 0\}. \end{cases}$$

2. Choose any sequence $r_k \to 0$. It will be enough to show there exists a subsequence $\{s_j\}_{j=1}^\infty \subset \{r_k\}_{k=1}^\infty$ for which

$$\chi_{E_{s_j}} \to \chi_{H^-(0)} \text{ in } L^1_{\text{loc}}(\mathbb{R}^n).$$

3. Fix $L > 0$ and let

$$D_r \equiv E_r \cap B(0,L), \qquad g_r(y) = \frac{y}{r}.$$

Then for any $\varphi \in C^1_c(\mathbb{R}^n;\mathbb{R}^n)$, $|\varphi| \leq 1$, we have

$$\begin{aligned} \int_{D_r} \operatorname{div} \varphi \, dz &= \frac{1}{r^{n-1}} \int_{E \cap B(0,rL)} \operatorname{div}(\varphi \circ g_r) \, dy \\ &= \frac{1}{r^{n-1}} \int_{\mathbb{R}^n} (\varphi \circ g_r) \cdot \nu_{E \cap B(0,rL)} \, d\|\partial(E \cap B(0,rL))\| \\ &\leq \frac{\|\partial(E \cap B(0,rL))\|(\mathbb{R}^n)}{r^{n-1}} \\ &\leq C < \infty \end{aligned}$$

for all $r \in (0,1]$, according to Lemma 2(v). Consequently,

$$\|\partial D_r\|(\mathbb{R}^n) \leq C < \infty \qquad (0 < r \leq 1),$$

and furthermore,

$$\|\chi_{D_r}\|_{L^1(\mathbb{R}^n)} = \mathcal{L}^n(D_r) \leq \mathcal{L}^n(B(0,L)) < \infty \qquad (r > 0).$$

Hence

$$||\chi_{D_r}||_{BV(\mathbb{R}^n)} \leq C < \infty$$

for all $0 < r \leq 1$.

In view of this estimate and the Compactness Theorem from Section 5.2.3, there exists a subsequence $\{s_j\}_{j=1}^{\infty} \subset \{r_k\}_{k=1}^{\infty}$ and a function $f \in BV_{\text{loc}}(\mathbb{R}^n)$ such that, writing $E_j \equiv E_{s_j}$, we have

$$\chi_{E_j} \to f \text{ in } L^1_{\text{loc}}(\mathbb{R}^n).$$

We may assume also $\chi_{E_j} \to f$ $\mathcal{L}^n$ a.e.; hence $f(x) \in \{0, 1\}$ for $\mathcal{L}^n$ a.e. x and so

$$f = \chi_F \qquad \mathcal{L}^n \text{ a.e.},$$

where

$$F \subset \mathbb{R}^n \qquad \text{has locally finite perimeter.}$$

Hence if $\varphi \in C^1_c(\mathbb{R}^n; \mathbb{R}^n)$,

$$\int_F \operatorname{div} \varphi \, dy = \int_{\mathbb{R}^n} \varphi \cdot \nu_F \, d||\partial F||, \tag{$\star$}$$

for some $||\partial F||$-measurable function ν_F, with $|\nu_F| = 1$ $||\partial F||$ a.e.

We must prove $F = H^-(0)$.

4. *Claim #1*: $\nu_F = e_n$ $||\partial F||$ a.e.

Proof of Claim #1: Let us write $\nu_j \equiv \nu_{E_j}$. Then if $\varphi \in C^1_c(\mathbb{R}^n; \mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \varphi \cdot \nu_j \, d||\partial E_j|| = \int_{E_j} \operatorname{div} \varphi \, dy \qquad (j = 1, 2, \ldots).$$

Since

$$\chi_{E_j} \to \chi_F \text{ in } L^1_{\text{loc}},$$

we see from the above and ($\star$) that

$$\int_{\mathbb{R}^n} \varphi \cdot \nu_j \, d||\partial E_j|| \to \int_{\mathbb{R}^n} \varphi \cdot \nu_F \, d||\partial F|| \text{ as } j \to \infty.$$

Thus

$$\nu_j ||\partial E_j|| \rightharpoonup \nu_F ||\partial F||$$

weakly in the sense of Radon measures. Consequently, for each $L > 0$ for which $||\partial F||(\partial B(0,L)) = 0$, and hence for all but at most countably many $L > 0$,

$$\int_{B(0,L)} \nu_j \, d||\partial E_j|| \to \int_{B(0,L)} \nu_F \, d||\partial F||. \tag{$\star\star$}$$

On the other hand, for all φ as above,

$$\int_{\mathbb{R}^n} \varphi \cdot \nu_j \, d||\partial E_j|| = \frac{1}{s_j^{n-1}} \int_{\mathbb{R}^n} (\varphi \circ g_{s_j}) \cdot \nu_E \, d||\partial E||,$$

whence

$$\left\{ \begin{array}{l} ||\partial E_j||(U(0,L)) = \dfrac{1}{s_j^{n-1}} ||\partial E||(B(0,s_jL)) \\ \displaystyle\int_{B(0,L)} \nu_j \, d||\partial E_j|| = \frac{1}{s_j^{n-1}} \int_{B(0,s_jL)} \nu_E \, d||\partial E||. \end{array} \right. \tag{$\star\star\star$}$$

Therefore

$$\lim_{j\to\infty} \fint_{B(0,L)} \nu_j \, d||\partial E_j|| = \lim_{j\to\infty} \fint_{B(0,s_jL)} \nu_E \, d||\partial E|| = \nu_E(0) = e_n,$$

since $0 \in \partial^* E$. If $||\partial F||(\partial B(0,L)) = 0$, the Lower Semicontinuity Theorem implies

$$\begin{aligned} ||\partial F||(B(0,L)) &\le \liminf_{j\to\infty} ||\partial E_j||(B(0,L)) \\ &= \lim_{j\to\infty} \int_{B(0,L)} e_n \cdot \nu_j \, d||\partial E_j|| \\ &= \int_{B(0,L)} e_n \cdot \nu_F \, d||\partial F||, \qquad \text{by } (\star\star). \end{aligned}$$

Since $|\nu_F| = 1$ $||\partial F||$ a.e., the above inequality forces

$$\nu_F = e_n \qquad ||\partial F|| \text{ a.e.}$$

It also follows from the above inequality that

$$||\partial F||(B(0,L)) = \lim_{j\to\infty} ||\partial E_j||(B(0,L))$$

whenever $||\partial F||(\partial B(0,L)) = 0$.

5. *Claim #2.* F is a half space.

Proof of Claim #2: By Claim #1, for all $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$,

$$\int_F \operatorname{div} \varphi \, dz = \int_{\mathbb{R}^n} \varphi \cdot e_n \, d||\partial F||.$$

Fix $\epsilon > 0$ and let $f^\epsilon \equiv \eta_\epsilon * \chi_F$, where η_ϵ is the usual mollifier. Then $f^\epsilon \in C^\infty(\mathbb{R}^n)$, and so

$$\begin{aligned}\int_{\mathbb{R}^n} f^\epsilon \operatorname{div}\varphi\, dz &= \int_F \operatorname{div}(\eta_\epsilon * \varphi)\, dz\\ &= \int_{\mathbb{R}^n} \eta_\epsilon * (\varphi \cdot e_n)\, d\|\partial F\|.\end{aligned}$$

But also

$$\int_{\mathbb{R}^n} f^\epsilon \operatorname{div}\varphi\, dz = -\int_{\mathbb{R}^n} Df^\epsilon \cdot \varphi\, dz.$$

Thus

$$\frac{\partial f^\epsilon}{\partial z_i} = 0 \qquad (i = 1, \dots, n-1), \qquad \frac{\partial f^\epsilon}{\partial z_n} \leq 0.$$

As $f_\epsilon \to \chi_F$ $\mathcal{L}^n$ a.e. as $\epsilon \to 0$, we conclude that — up to a set of $\mathcal{L}^n$-measure zero —

$$F = \{y \in \mathbb{R}^n \mid y_n \leq \gamma\} \qquad \text{for some } \gamma \in \mathbb{R}.$$

6. *Claim #3*: $F = H^-(0)$.

Proof of Claim #3: We must show $\gamma = 0$ above. Assume instead $\gamma > 0$. Since $\chi_{E_j} \to \chi_F$ in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$,

$$\begin{aligned}\alpha(n)\gamma^n = \mathcal{L}^n(B(0,\gamma) \cap F) &= \lim_{j\to\infty} \mathcal{L}^n(B(0,\gamma) \cap E_j)\\ &= \lim_{j\to\infty} \frac{\mathcal{L}^n(B(0,\gamma s_j) \cap E)}{s_j^n},\end{aligned}$$

a contradiction to Lemma 2(ii).

Similarly, the case $\gamma < 0$ leads to a contradiction to Lemma 2(i). ∎

We at once read off more detailed information concerning the blow-up of E around a point $x \in \partial^* E$:

COROLLARY 1
Assume $x \in \partial^* E$. Then

(i) $$\lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r) \cap E \cap H^+(x))}{r^n} = 0$$

$$\lim_{r\to 0} \frac{\mathcal{L}^n((B(x,r) - E) \cap H^-(x))}{r^n} = 0, \textit{ and}$$

(ii) $$\lim_{r\to 0} \frac{\|\partial E\|(B(x,r))}{\alpha(n-1)r^{n-1}} = 1.$$

DEFINITION *A unit vector $\nu_E(x)$ for which (i) holds (with $H^\pm(x)$ as defined above) is called the* ***measure theoretic unit outer normal*** *to E at x.*

PROOF

1. We have

$$\frac{\mathcal{L}^n(B(x,r)\cap E\cap H^+(x))}{r^n} = \mathcal{L}^n(B(x,1)\cap E_r\cap H^+(x))$$
$$\to \mathcal{L}^n(B(x,1)\cap H^-(x)\cap H^+(x)) = 0 \text{ as } r\to 0.$$

The other limit in (i) has a similar proof.

2. Assume $x=0$. By $(\star\star\star)$ in the proof of Theorem 1,

$$\frac{||\partial E||(B(0,r))}{r^{n-1}} = ||\partial E_r||(B(0,1)).$$

Since $||\partial H^-(0)||(\partial B(0,1)) = \mathcal{H}^{n-1}(\partial B(0,1)\cap H(0)) = 0$, part 2 of the proof of Theorem 1 implies

$$\begin{aligned}\lim_{r\to 0}\frac{||\partial E||(B(0,r))}{r^{n-1}} &= ||\partial H^-(0)||(B(0,1))\\ &= \mathcal{H}^{n-1}(B(0,1)\cap H(0))\\ &= \alpha(n-1). \quad \blacksquare\end{aligned}$$

5.7.3 Structure Theorem for sets of finite perimeter

LEMMA 3

There exists a constant C, depending only on n, such that

$$\mathcal{H}^{n-1}(B) \le C||\partial E||(B)$$

for all $B\subset\partial^* E$.

PROOF Let $\epsilon,\delta>0$, $B\subset\partial^*E$. Since $||\partial E||$ is a Radon measure, there exists an open set $U\supset B$ such that

$$||\partial E||(U)\le ||\partial E||(B)+\epsilon.$$

According to Lemma 2, if $x\in\partial^*E$, then

$$\liminf_{r\to 0}\frac{||\partial E||(B(x,r))}{r^{n-1}} > A_3 > 0.$$

Let

$$\mathcal{F}\equiv\left\{B(x,r)\mid x\in B, B(x,r)\subset U, r<\frac{\delta}{10}, ||\partial E||(B(x,r))>A_3r^{n-1}\right\}.$$

According to Vitali's Covering Theorem, there exist disjoint balls $\{B(x_i,r_i)\}_{i=1}^{\infty}\subset\mathcal{F}$ such that

$$B\subset\bigcup_{i=1}^{\infty}B(x_i,5r_i).$$

Since diam $B(x_i, 5r_i) \le \delta \quad (i = 1, \dots)$,

$$\begin{aligned}\mathcal{H}^{n-1}_\delta(B) \le \sum_{i=1}^\infty \alpha(n-1)(5r_i)^{n-1} &\le C\sum_{i=1}^\infty ||\partial E||(B(x_i, r_i)) \\ &\le C||\partial E||(U) \\ &\le C(||\partial E||(B) + \epsilon).\end{aligned}$$

Let $\epsilon \to 0$ and then $\delta \to 0$. ∎

Now we show that a set of locally finite perimeter has "measure theoretically a C^1 boundary."

THEOREM 2 STRUCTURE THEOREM FOR SETS OF FINITE PERIMETER *Assume E has locally finite perimeter in $\mathbb{R}^n$.*

(i) Then

$$\partial^* E = \bigcup_{k=1}^\infty K_k \cup N,$$

where

$$||\partial E||(N) = 0$$

and K_k is a compact subset of a C^1-hypersurface S_k $(k = 1, 2, \dots)$.

(ii) Furthermore, $\nu_E \,|_{S_k}$ is normal to S_k $(k = 1, \dots)$, and

(iii) $||\partial E|| = \mathcal{H}^{n-1} \llcorner \partial^* E$.

PROOF

1. For each $x \in \partial^* E$, we have according to Corollary 1

$$\begin{cases} \lim_{r\to 0} \dfrac{\mathcal{L}^n(B(x,r) \cap E \cap H^+(x))}{r^n} = 0, \\ \lim_{r\to 0} \dfrac{\mathcal{L}^n((B(x,r) - E) \cap H^-(x))}{r^n} = 0. \end{cases} \qquad (\star)$$

Using Egoroff's Theorem, we see that there exist disjoint $||\partial E||$-measurable sets $\{F_i\}_{i=1}^\infty \subset \partial^* E$ such that

$$\begin{cases} ||\partial E||\left(\partial^* E - \bigcup_{i=1}^\infty F_i\right) = 0, \ ||\partial E||(F_i) < \infty, \text{ and} \\ \text{the convergence in } (\star) \text{ is uniform for } x \in F_i \ \ (i = 1, \dots). \end{cases}$$

Then, by Lusin's Theorem, for each i there exist disjoint compact sets $\{E_i^j\}_{j=1}^\infty \subset F_i$ such that

$$\begin{cases} \|\partial E\| \left(F_i - \bigcup_{j=1}^\infty E_i^j \right) = 0 \text{ and} \\ \nu_E \,|_{E_i^j} \quad \text{is continuous.} \end{cases}$$

Reindex the sets $\{E_i^j\}_{i,j=1}^\infty$ and call them $\{K_k\}_{k=1}^\infty$. Then

$$\begin{cases} \partial^* E = \bigcup_{k=1}^\infty K_k \cup N, \ \|\partial E\|(N) = 0, \\ \text{the convergence in } (\star) \text{ is uniform on } \ K_k, \text{ and} \\ \nu_E \,|_{K_k} \quad \text{is continuous} \quad (k = 1, 2, \ldots). \end{cases} \tag{$\star\star$}$$

2. Define for $\delta > 0$

$$\rho_k(\delta) \equiv \sup \left\{ \frac{|\nu_E(x) \cdot (y - x)|}{|y - x|} \ \middle| \ 0 < |x - y| \le \delta, x, y \in K_k \right\}.$$

3. *Claim*: For each $k = 1, 2, \ldots$, $\rho_k(\delta) \to 0$ as $\delta \to 0$.

Proof of Claim: We may as well assume $k = 1$. Fix $0 < \epsilon < 1$. By $(\star)$, $(\star\star)$ there exists $0 < \delta < 1$ such that if $z \in K_1$ and $r < 2\delta$, then

$$\begin{cases} \mathcal{L}^n(E \cap B(z, r) \cap H^+(z)) < \dfrac{\epsilon^n}{2^{n+2}} \alpha(n) r^n \\ \mathcal{L}^n(E \cap B(z, r) \cap H^-(z)) > \alpha(n) \left(\dfrac{1}{2} - \dfrac{\epsilon^n}{2^{n+2}} \right) r^n. \end{cases} \tag{$\star\star\star$}$$

Assume now $x, y \in K_1$, $0 < |x - y| \le \delta$.
Case 1. $\nu_E(x) \cdot (y - x) > \epsilon |x - y|$.
Then, since $\epsilon < 1$,

$$B(y, \epsilon|x - y|) \subset H^+(x) \cap B(x, 2|x - y|). \tag{$\star\star\star\star$}$$

To see this, observe that if $z \in B(y, \epsilon|x - y|)$, then $z = y + w$, where $|w| \le \epsilon|x - y|$, whence

$$\nu_E(x) \cdot (z - x) = \nu_E(x) \cdot (y - x) + \nu_E(x) \cdot w > \epsilon|x - y| - |w| \ge 0.$$

On the other hand, $(\star\star\star)$ with $z = x$ implies

$$\begin{aligned} \mathcal{L}^n(E \cap B(x, 2|x - y|) \cap H^+(x)) &< \frac{\epsilon^n}{2^{n+2}} \alpha(n) (2|x - y|)^n \\ &= \frac{\epsilon^n \alpha(n)}{4} |x - y|^n, \end{aligned}$$

and $(\star\star\star)$ with $z = y$ implies

$$\begin{aligned}\mathcal{L}^n(E \cap B(y, \epsilon|x-y|)) &\geq \mathcal{L}^n(E \cap B(y, \epsilon|x-y|) \cap H^-(y)) \\ &\geq \frac{\epsilon^n \alpha(n)|x-y|^n}{2}\left(1 - \frac{\epsilon^n}{2^{n+1}}\right) \\ &> \frac{\epsilon^n \alpha(n)}{4}|x-y|^n.\end{aligned}$$

However, our applying $\mathcal{L}^n \llcorner E$ to both sides of $(\star\star\star\star)$ yields an estimate contradicting the above inequalities.

Case 2. $\nu_E(x) \cdot (y - x) \leq -\epsilon|x-y|$.

This similarly leads to a contradiction.

4. Now apply Whitney's Extension Theorem (found in Section 6.5) with

$$f = 0 \quad \text{and} \quad d = \nu_E \quad \text{on} \quad K_k.$$

We conclude that there exist C^1-functions $\bar{f}_k : \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{cases} \bar{f}_k = 0 & \text{on } K_k \\ D\bar{f}_k = \nu_E & \text{on } K_k. \end{cases}$$

Let

$$S_k \equiv \left\{x \in \mathbb{R}^n \mid \bar{f}_k = 0,\ |D\bar{f}_k| > \frac{1}{2}\right\} \qquad (k = 1, 2, \ldots).$$

By the Implicit Function Theorem, S_k is a C^1, $(n-1)$-dimensional submanifold of $\mathbb{R}^n$. Clearly $K_k \subset S_k$. This proves (i) and (ii).

5. Choose a Borel set $B \subset \partial^* E$. According to Lemma 3,

$$\mathcal{H}^{n-1}(B \cap N) \leq C||\partial E||(B \cap N) = 0.$$

Thus we may as well assume $B \subset \bigcup_{k=1}^{\infty} K_k$, and in fact $B \subset K_1$. By (ii) there exists a C^1-hypersurface $S_1 \supset K_1$. Let

$$\nu \equiv \mathcal{H}^{n-1} \llcorner S_1.$$

Since S_1 is C^1,

$$\lim_{r \to 0} \frac{\nu(B(x,r))}{\alpha(n-1)r^{n-1}} = 1 \qquad (x \in B).$$

Thus Corollary 1(ii) implies

$$\lim_{r \to 0} \frac{\nu(B(x,r))}{||\partial E||(B(s,r))} = 1 \qquad (x \in B).$$

Since ν and $||\partial E||$ are Radon measures, Theorem 2 in Section 1.6.2 implies

$$||\partial E||(B) = \nu(B) = \mathcal{H}^{n-1}(B). \quad \blacksquare$$

5.8 The measure theoretic boundary; Gauss–Green Theorem

As above, we continue to assume E is a set of locally finite perimeter in $\mathbb{R}^n$. We next refine Corollary 3 in Section 1.7.1.

DEFINITION *Let $x \in \mathbb{R}^n$. We say $x \in \partial_* E$, the* ***measure theoretic boundary*** *of E, if*

$$\limsup_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{r^n} > 0$$

and

$$\limsup_{r \to 0} \frac{\mathcal{L}^n(B(x,r) - E)}{r^n} > 0.$$

LEMMA 1

(i) $\partial^* E \subset \partial_* E$.

(ii) $\mathcal{H}^{n-1}(\partial_* E - \partial^* E) = 0$.

PROOF

1. Assertion (i) follows from Lemma 2 in Section 5.7.

2. Since the mapping

$$r \mapsto \frac{\mathcal{L}^n(B(x,r) \cap E)}{r^n}$$

is continuous, if $x \in \partial_* E$, there exists $0 < \alpha < 1$ and $r_j \to 0$ such that

$$\frac{\mathcal{L}^n(B(x,r_j) \cap E)}{\alpha(n) r_j^n} = \alpha.$$

Thus

$$\min\{\mathcal{L}^n(B(x,r_j) \cap E), \mathcal{L}^n(B(x,r_j) - E)\} = \min\{\alpha, 1-\alpha\}\alpha(n) r_j^n,$$

and so the Relative Isoperimetric Inequality implies

$$\limsup_{r \to 0} \frac{||\partial E||(B(x,r))}{r^{n-1}} > 0.$$

Since $||\partial E||(\mathbb{R}^n - \partial^* E) = 0$, standard covering arguments imply

$$\mathcal{H}^{n-1}(\partial_* E - \partial^* E) = 0. \quad \blacksquare$$

Now we prove that if E has locally finite perimeter, then the usual Gauss–Green formula holds, provided we consider the measure theoretic boundary of E.

THEOREM 1 GENERALIZED GAUSS–GREEN THEOREM
Let $E \subset \mathbb{R}^n$ have locally finite perimeter.

(i) *Then $\mathcal{H}^{n-1}(\partial_* E \cap K) < \infty$ for each compact set $K \subset \mathbb{R}^n$.*

(ii) *Furthermore, for $\mathcal{H}^{n-1}$ a.e. $x \in \partial_* E$, there is a unique measure theoretic unit outer normal $\nu_E(x)$ such that*

$$\int_E \operatorname{div} \varphi \, dx = \int_{\partial_* E} \varphi \cdot \nu_E \, d\mathcal{H}^{n-1} \tag{$\star$}$$

for all $\varphi \in C_c^1(\mathbb{R}^n; \mathbb{R}^n)$.

PROOF By the foregoing theory,

$$\int_E \operatorname{div} \varphi \, dx = \int_{\mathbb{R}^n} \varphi \cdot \nu_E \, d\|\partial E\|.$$

But

$$\|\partial E\|(\mathbb{R}^n - \partial^* E) = 0$$

and, by Theorem 2 in Section 5.7.3 and Lemma 1,

$$\|\partial E\| = \mathcal{H}^{n-1} \llcorner \partial_* E.$$

Thus ($\star$) follows from Lemma 1. ∎

REMARK We will see in Section 5.11 below that if $E \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable and $\mathcal{H}^{n-1}(\partial_* E \cap K) < \infty$ for all compact $K \subset \mathbb{R}^n$, then E has locally finite perimeter. In particular, we see that *the Gauss–Green Theorem is valid for $E = U$, an open set with Lipschitz boundary.* ∎

5.9 Pointwise properties of BV functions

We next extend our analysis of sets of finite perimeter to general BV functions. The goal will be to demonstrate that a BV function is "measure theoretically piecewise continuous," with "jumps along a measure theoretically C^1 surface."

We now assume $f \in BV(\mathbb{R}^n)$ and investigate the approximate limits of $f(y)$ as y approaches a typical point $x \in \mathbb{R}^n$.

DEFINITIONS

(i) $\mu(x) \equiv \operatorname{ap}\limsup_{y \to x} f(y) = \inf \left\{ t \mid \lim_{r \to 0} \dfrac{\mathcal{L}^n(B(x,r) \cap \{f > t\})}{r^n} = 0 \right\}.$

(ii) $\lambda(x) \equiv \operatorname{ap}\liminf_{y \to x} f(y) = \sup \left\{ t \mid \lim_{r \to 0} \dfrac{\mathcal{L}^n(B(x,r) \cap \{f < t\})}{r^n} = 0 \right\}.$

REMARK Clearly $-\infty \leq \lambda(x) \leq \mu(x) \leq \infty$ for all $x \in \mathbb{R}^n$. ∎

LEMMA 1
The functions $x \mapsto \lambda(x)$, $\mu(x)$ are Borel measurable.

PROOF For each $t \in \mathbb{R}$, the set $E_t \equiv \{x \in \mathbb{R}^n \mid f(x) > t\}$ is $\mathcal{L}^n$-measurable, and so for each $r > 0$, $t \in \mathbb{R}$, the mapping

$$x \mapsto \frac{\mathcal{L}^n(B(x,r) \cap E_t)}{r^n}$$

is continuous. This implies

$$\mu_t(x) \equiv \limsup_{\substack{r \to 0 \\ r \text{ rational}}} \frac{\mathcal{L}^n(B(x,r) \cap E_t)}{r^n}$$

is a Borel measurable function of x for each $t \in \mathbb{R}$.

Now, for each $s \in R$,

$$\{x \in \mathbb{R}^n \mid \mu(x) \leq s\} = \bigcap_{k=1}^{\infty} \{x \in \mathbb{R}^n \mid \mu_{s+\frac{1}{k}}(x) = 0\},$$

and so μ is a Borel measurable function.

The proof that λ is Borel measurable is similar. ∎

DEFINITION *Let* J *denote $\{x \in \mathbb{R}^n \mid \lambda(x) < \mu(x)\}$, the set of points at which the approximate limit of f does **not** exist.*

According to Theorem 2 in Section 1.7.2,

$$\mathcal{L}^n(\mathrm{J}) = 0.$$

We will see below that for $\mathcal{H}^{n-1}$ a.e. point $x \in \mathrm{J}$, f has a "measure theoretic jump" across a hyperplane through x.

THEOREM 1
There exist countably many C^1-hypersurfaces $\{S_k\}_{k=1}^{\infty}$ such that

$$\mathcal{H}^{n-1}\left(\mathrm{J} - \bigcup_{k=1}^{\infty} S_k\right) = 0.$$

PROOF Define, as in Section 5.5,

$$E_t \equiv \{x \in \mathbb{R}^n \mid f(x) > t\} \qquad (t \in \mathbb{R}).$$

According to the Coarea Formula for BV functions, E_t is a set of finite perimeter in $\mathbb{R}^n$ for $\mathcal{L}^1$ a.e. t. Furthermore, observe that if $x \in \mathrm{J}$ and $\lambda(x) < t < \mu(x)$,

then

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap\{f>t\})}{r^n} > 0$$

and

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap\{f<t\})}{r^n} > 0.$$

Thus

$$\{x\in \mathbf{J} \mid \lambda(x) < t < \mu(x)\} \subset \partial_* E_t. \tag{$\star$}$$

Choose $D \subset \mathbb{R}^1$ to be a countable, dense set such that E_t is of finite perimeter for each $t\in D$. For each $t\in D$, $\mathcal{H}^{n-1}$ almost all of $\partial_* E_t$ is contained in a countable union of C^1-hypersurfaces: this is a consequence of the Structure Theorem in Section 5.7.

Now, according to ($\star$),

$$\mathbf{J} \subset \bigcup_{t\in D} \partial_* E_t,$$

and the theorem follows. ▮

THEOREM 2
$-\infty < \lambda(x) \le \mu(x) < +\infty$ *for* $\mathcal{H}^{n-1}$ *a.e.* $x\in\mathbb{R}^n$.

PROOF

1. Claim #1: $\mathcal{H}^{n-1}(\{x \mid \lambda(x) = +\infty\}) = 0$, $\mathcal{H}^{n-1}(\{x \mid \mu(x) = -\infty\}) = 0$.

Proof of Claim #1: We may assume spt (f) is compact. Let

$$F_t \equiv \{x\in\mathbb{R}^n \mid \lambda(x) > t\}.$$

Since $\mu(x) = \lambda(x) = f(x)$ $\mathcal{L}^n$ a.e., E_t and F_t differ at most by a set of $\mathcal{L}^n$-measure zero, whence

$$\|\partial E_t\| = \|\partial F_t\|.$$

Consequently, the Coarea Formula for BV functions implies

$$\int_{-\infty}^{\infty} \|\partial F_t\|(\mathbb{R}^n)\,dt = \|Df\|(\mathbb{R}^n) < \infty,$$

and so

$$\liminf_{t\to\infty} \|\partial F_t\|(\mathbb{R}^n) = 0. \tag{$\star$}$$

Since spt (f) is compact, there exists $d>0$ such that

$$\mathcal{L}^n(\text{spt }(f)\cap B(x,r)) \le \frac{1}{8}\alpha(n) r^n \text{ for all } x\in \text{spt }(f) \text{ and } r\ge d. \tag{$\star\star$}$$

Fix $t > 0$. By the definitions of λ and F_t,

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap F_t)}{\alpha(n) r^n} = 1 \text{ for } x \in F_t.$$

Thus for each $x \in F_t$, there exists $r > 0$ such that

$$\frac{\mathcal{L}^n(B(x,r) \cap F_t)}{\alpha(n) r^n} = \frac{1}{4}. \qquad (\star\star\star)$$

According to $(\star\star)$, $r \leq d$.

We apply Vitali's Covering Theorem to find a countable disjoint collection $\{B(x_i, r_i)\}_{i=1}^{\infty}$ of balls satisfying $(\star\star\star)$ for $x = x_i$, $r = r_i \leq d$, such that

$$F_t \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_i).$$

Now $(\star\star\star)$ and the Relative Isoperimetric Inequality imply

$$\left(\frac{\alpha(n)}{4}\right)^{\frac{n-1}{n}} \leq \frac{C\|\partial F_t\|(B(x_i, r_i))}{r_i^{n-1}};$$

that is,

$$r_i^{n-1} \leq C\|\partial F_t\|(B(x_i, r_i)) \qquad (i = 1, 2, \ldots).$$

Thus we may calculate

$$\begin{aligned} \mathcal{H}_{10d}^{n-1}(F_t) &\leq \sum_{i=1}^{\infty} \alpha(n-1)(5r_i)^{n-1} \\ &\leq C \sum_{i=1}^{\infty} \|\partial F_t\|(B(x_i, r_i)) \\ &\leq C\|\partial F_t\|(\mathbb{R}^n). \end{aligned}$$

In view of $(\star)$,

$$\mathcal{H}_{10d}^{n-1}(\{x \mid \lambda(x) = +\infty\}) = 0,$$

and so

$$\mathcal{H}^{n-1}(\{x \mid \lambda(x) = +\infty\}) = 0.$$

The proof that $\mathcal{H}^{n-1}(\{x \mid \mu(x) = -\infty\}) = 0$ is similar.

2. *Claim #2*: $\mathcal{H}^{n-1}(\{x \mid \mu(x) - \lambda(x) = \infty\}) = 0$.

Proof of Claim #2: By Theorem 1, $\mathbf{J}$ is σ-finite with respect to $\mathcal{H}^{n-1}$ in $\mathbb{R}^n$, and thus $\{(x,t) \mid x \in \mathbf{J},\ \lambda(x) < t < \mu(x)\}$ is σ-finite with respect to $\mathcal{H}^{n-1} \times \mathcal{L}^1$

in $\mathbb{R}^{n+1}$. Consequently, Fubini's Theorem implies

$$\int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\{\lambda(x) < t < \mu(x)\})\, dt = \int_{\mathbb{R}^n} \mu(x) - \lambda(x)\, d\mathcal{H}^{n-1}.$$

But by statement ($\star$) in the proof of Theorem 1, and the theory developed in Section 5.7,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\{\lambda(x) < t < \mu(x)\})\, dt &\le \int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\partial_* E_t)\, dt \\ &= \int_{-\infty}^{\infty} ||\partial E_t||(\mathbb{R}^n)\, dt \\ &= ||Df||(\mathbb{R}^n) < \infty.\end{aligned}$$

Consequently, $\mathcal{H}^{n-1}(\{x \mid \mu(x) - \lambda(x) = \infty\}) = 0$. ∎

NOTATION $F(x) \equiv (\lambda(x) + \mu(x))/2$.

DEFINITIONS *Let ν be a unit vector in $\mathbb{R}^n$, $x \in \mathbb{R}^n$. We define the hyperplane*

$$H_\nu \equiv \{y \in \mathbb{R}^n \mid \nu \cdot (y - x) = 0\}$$

and the half-spaces

$$H_\nu^+ \equiv \{y \in \mathbb{R}^n \mid \nu \cdot (y - x) \ge 0\},$$

$$H_\nu^- \equiv \{y \in \mathbb{R}^n \mid \nu \cdot (y - x) \le 0\}.$$

THEOREM 3 FINE PROPERTIES OF BV FUNCTIONS
Assume $f \in BV(\mathbb{R}^n)$. Then

(i) $\lim_{r\to 0} ⨍_{B(x,r)} |f - F(x)|^{n/n-1}\, dy = 0$ *for $\mathcal{H}^{n-1}$ a.e. $x \in \mathbb{R}^n - \mathbf{J}$,*

and

(ii) *for $\mathcal{H}^{n-1}$ a.e. $x \in \mathbf{J}$, there exists a unit vector $\nu = \nu(x)$ such that*

$$\lim_{r\to 0} ⨍_{B(x,r)\cap H_\nu^-} |f - \mu(x)|^{n/n-1}\, dy = 0$$

and

$$\lim_{r\to 0} ⨍_{B(x,r)\cap H_\nu^+} |f - \lambda(x)|^{n/n-1}\, dy = 0.$$

In particular,

$$\mu(x) = \operatorname{ap}\lim_{\substack{y\to x\\ y\in H_\nu^+}} f(y), \qquad \lambda(x) = \operatorname{ap}\lim_{\substack{y\to x\\ y\in H_\nu^-}} f(y).$$

REMARK Thus we see that for $\mathcal{H}^{n-1}$ a.e. $x \in J$, f has a "measure theoretic jump" across the hyperplane $H_{\nu(x)}$. ▮

PROOF We will prove only the second part of assertion (ii), as the other statements follow similarly.

1. For $\mathcal{H}^{n-1}$ a.e. $x \in \mathbf{J}$, there exists a unit vector ν such that ν is the measure theoretic exterior unit normal to $E_t = \{f > t\}$ at x for $\lambda(x) < t < \mu(x)$. Thus for each $\epsilon > 0$,

$$\begin{cases} \dfrac{\mathcal{L}^n(B(x,r) \cap \{f > \lambda(x) + \epsilon\} \cap H_\nu^+)}{r^n} = 0, \\[2ex] \dfrac{\mathcal{L}^n(B(x,r) \cap \{f < \lambda(x) - \epsilon\})}{r^n} = 0. \end{cases} \tag{$\star$}$$

Hence if $0 < \epsilon < 1$,

$$\begin{aligned} \frac{1}{r^n} \int_{B(x,r)\cap H_\nu^+} & |f - \lambda(x)|^{n/n-1}\, dy \\ & \leq \frac{1}{2}\alpha(n)\epsilon^{n/n-1} \\ & \quad + \frac{1}{r^n} \int_{B(x,r)\cap H_\nu^+ \cap \{f > \lambda(x)+\epsilon\}} |f - \lambda(x)|^{n/n-1}\, dy \\ & \quad + \frac{1}{r^n} \int_{B(x,r)\cap H_\nu^+ \cap \{f < \lambda(x)-\epsilon\}} |f - \lambda(x)|^{n/n-1}\, dy. \end{aligned} \tag{$\star\star$}$$

Now fix $M > \lambda(x) + \epsilon$. Then

$$\begin{aligned} \frac{1}{r^n} \int_{B(x,r)\cap H_\nu^+ \cap \{f > \lambda(x)+\epsilon\}} & |f - \lambda(x)|^{n/n-1}\, dy \\ & \leq (M - \lambda(x))^{n/n-1} \frac{\mathcal{L}^n(B(x,r) \cap H_\nu^+ \cap \{f > \lambda(x) + \epsilon\})}{r^n} \\ & \quad + \frac{1}{r^n} \int_{B(x,r)\cap \{f > M\}} |f - \lambda(x)|^{n/n-1}\, dy. \end{aligned}$$

Similarly, if $-M < \lambda(x) - \epsilon$,

$$\begin{aligned} \frac{1}{r^n} \int_{B(x,r)\cap \{f < \lambda(x)-\epsilon\}} & |f - \lambda(x)|^{n/n-1}\, dy \\ & \leq (M + \lambda(x))^{n/n-1} \frac{\mathcal{L}^n(B(x,r) \cap \{f < \lambda(x) - \epsilon\})}{r^n} \\ & \quad + \frac{1}{r^n} \int_{B(x,r)\cap \{f < -M\}} |f - \lambda(x)|^{n/n-1}\, dy. \end{aligned}$$

We employ the two previous calculations in ($\star\star$) and then recall ($\star$) to compute

$$\begin{aligned}\limsup_{r\to 0}\frac{1}{r^n}\int_{B(x,r)\cap H_\nu^+}|f-\lambda(x)|^{n/n-1}\,dy\\ \le \limsup_{r\to 0}\frac{1}{r^n}\int_{B(x,r)\cap\{|f|>M\}}|f-\lambda(x)|^{n/n-1}\,dy \qquad (\star\star\star)\end{aligned}$$

for all sufficiently large $M>0$.

2. Now

$$\begin{aligned}\frac{1}{r^n}\int_{B(x,r)\cap\{f>M\}}|f-\lambda(x)|^{n/n-1}\,dy \le \frac{C}{r^n}\int_{B(x,r)}(f-M)^{+^{n/n-1}}\,dy\\ +(M-\lambda(x))^{n/n-1}\frac{\mathcal{L}^n(B(x,r)\cap\{f>M\})}{r^n}.\end{aligned}$$

If $M>\mu(x)$, the second term on the right-hand side of this inequality goes to zero as $r\to 0$. Furthermore, for sufficiently small $r>0$,

$$\frac{\mathcal{L}^n(B(x,r)\cap\{f>M\})}{\mathcal{L}^n(B(x,r))}\le\frac{1}{2}$$

and hence by Theorem 1(iii) in Section 5.6.1 we have

$$\left(⨍_{B(x,r)}(f-M)^{+^{n/n-1}}\,dy\right)^{\frac{n-1}{n}} \le \frac{C}{r^{n-1}}\|D(f-M)^+\|(B(x,r)).$$

This estimate and the analogous one over the set $\{f<-M\}$ combine with ($\star\star\star$) to prove

$$\begin{aligned}\limsup_{r\to 0}\left(⨍_{B(x,r)\cap H_\nu^+}|f-\lambda(x)|^{n/n-1}\,dy\right)^{\frac{n-1}{n}}\\ \le C\limsup_{r\to 0}\frac{\|D(f-M)^+\|(B(x,r))}{r^{n-1}}\\ +C\limsup_{r\to 0}\frac{\|D(-M-f)^+\|(B(x,r))}{r^{n-1}} \qquad (\star\star\star\star)\end{aligned}$$

for all sufficiently large $M>0$.

3. Fix $\epsilon>0$, $N>0$, and define

$$A_\epsilon^N\equiv\left\{x\in\mathbb{R}^n \mid \limsup_{r\to 0}\frac{\|D(f-M)^+\|(B(x,r))}{r^{n-1}}>\epsilon \text{ for all } M\ge N\right\}.$$

Then

$$\epsilon\mathcal{H}^{n-1}(A_\epsilon^N)\le C\|D(f-M)^+\|(\mathbb{R}^n)=C\int_M^\infty\|\partial E_t\|(\mathbb{R}^n)\,dt$$

for all $M \geq N$. Thus

$$\mathcal{H}^{n-1}(A_\epsilon^N) = 0,$$

and so

$$\lim_{M\to\infty} \limsup_{r\to 0} \frac{||D(f-M)^+||(B(x,r))}{r^{n-1}} = 0$$

for $\mathcal{H}^{n-1}$ a.e. $x \in J$. Similarly,

$$\lim_{M\to\infty} \limsup_{r\to 0} \frac{||D(-M-f)^+||(B(x,r))}{r^{n-1}} = 0.$$

These estimates and $(\star\star\star\star)$ prove

$$\lim_{r\to 0} \fint_{B(x,r)\cap H_\nu^+} |f-\lambda(x)|^{n/n-1}\, dy = 0. \quad \blacksquare$$

COROLLARY 1

(i) *If $f \in BV(\mathbb{R}^n)$, then*

$$f^\star(x) \equiv \lim_{r\to 0} (f)_{x,r} = F(x)$$

exists for $\mathcal{H}^{n-1}$ a.e. $x \in \mathbb{R}^n$.

(ii) *Furthermore, if η_ϵ is the standard mollifier and $f^\epsilon \equiv \eta_\epsilon * f$, then*

$$f^\star(x) = \lim_{\epsilon\to 0} f^\epsilon(x)$$

for $\mathcal{H}^{n-1}$ a.e. $x \in \mathbb{R}^n$.

5.10 Essential variation on lines

We now ascertain the behavior of a BV function on lines.

5.10.1 BV functions of one variable

We first study BV functions of one variable.

Suppose $f : \mathbb{R} \to \mathbb{R}$ is $\mathcal{L}^1$-measurable, $-\infty \leq a < b \leq \infty$.

DEFINITION *The **essential variation** of f on the interval (a,b) is*

$$\text{ess } V_a^b f \equiv \sup\left\{\sum_{j=1}^m |f(t_{j+1}) - f(t_j)|\right\},$$

the supremum taken over all finite partitions $\{a < t_1 < \cdots < t_{m+1} < b\}$ such that each t_j is a point of approximate continuity of f.

REMARK The *variation* of f on (a, b) is similarly defined, but without the proviso that each partition point t_j be a point of approximate continuity. Since we demand that a function remain BV even after being redefined on a set of $\mathcal{L}^1$ measure zero, we see that essential variation is the proper notion here.

In particular, if $f = g$ $\mathcal{L}^1$ a.e. on (a, b), then

$$\text{ess } V_a^b f = \text{ess } V_a^b g. \quad \blacksquare$$

THEOREM 1

Suppose $f \in L^1(a, b)$. Then $\|Df\|(a, b) = \text{ess } V_a^b f$. Thus $f \in BV(a, b)$ if and only if ess $V_a^b f < \infty$.

PROOF

1. Consider first ess $V_a^b f$. Fix $\epsilon > 0$ and let $f^\epsilon \equiv \eta_\epsilon * f$ denote the usual smoothing of f. Choose any $a + \epsilon < t_1 < \cdots < t_{m+1} < b - \epsilon$. Since $\mathcal{L}^1$ a.e. point is a point of approximate continuity of f, $t_j - s$ is a point of approximate continuity of f for $\mathcal{L}^1$ a.e. s. Hence

$$\begin{aligned}\sum_{j=1}^{m} |f^\epsilon(t_{j+1}) - f^\epsilon(t_j)| &= \sum_{j=1}^{m} \left| \int_{-\epsilon}^{\epsilon} \eta_\epsilon(s)(f(t_{j+1} - s) - f(t_j - s))\, ds \right| \\ &\le \int_{-\epsilon}^{\epsilon} \eta_\epsilon(s) \sum_{j=1}^{m} |f(t_{j+1} - s) - f(t_j - s)|\, ds \\ &\le \text{ess } V_a^b f.\end{aligned}$$

It follows that

$$\int_{a+\epsilon}^{b-\epsilon} |(f^\epsilon)'|\, dx = \sup \left\{ \sum_{j=1}^{m} |f^\epsilon(t_{j+1}) - f^\epsilon(t_j)| \right\} \le \text{ess } V_a^b f.$$

Thus if $\varphi \in C_c^1(a, b)$, $|\varphi| \le 1$, we have

$$\int_a^b f^\epsilon \varphi'\, dx = -\int_a^b (f^\epsilon)'\varphi\, dx \le \int_{a+\epsilon}^{b-\epsilon} |(f^\epsilon)'|\, dx \le \text{ess } V_a^b f$$

for ϵ sufficiently small. Let $\epsilon \to 0$ to find

$$\int_a^b f\varphi'\, dx \le \text{ess } V_a^b f.$$

Hence

$$||Df||(a,b) = \sup\left\{\int_a^b f\varphi'\,dx \mid \varphi \in C_c^1(a,b), |\varphi| \le 1\right\}$$
$$\le \text{ess } V_a^b f \le \infty.$$

In particular, if $f \notin BV(a,b)$,

$$||Df||(a,b) = \text{ess } V_a^b f = +\infty.$$

2. Now suppose $f \in BV(a,b)$ and choose $a < c < d < b$. Then for each $\varphi \in C_c^1(c,d)$, with $|\varphi| \le 1$, and each small $\epsilon > 0$, we calculate

$$\begin{aligned}\int_c^d (f^\epsilon)'\varphi\,dx &= -\int_c^d f^\epsilon\varphi'\,dx \\ &= -\int_c^d (\eta_\epsilon * f)\varphi'\,dx \\ &= -\int_a^b f(\eta_\epsilon * \varphi)'\,dx \\ &\le ||Df||(a,b).\end{aligned}$$

Thus $\int_c^d |(f^\epsilon)'|\,dx \le ||Df||(a,b)$.

3. *Claim*: $f \in L^\infty(a,b)$.

Proof of Claim: Choose $\{f_j\}_{j=1}^\infty \subset BV(a,b) \cap C^\infty(a,b)$ so that

$$f_j \to f \quad \text{in} \quad L^1(a,b), \qquad f_j \to f \ \ \mathcal{L}^n \text{ a.e.}$$

and

$$\int_a^b |f_j'|\,dx \to ||Df||(a,b).$$

For each $y, z \in (a,b)$,

$$f_j(z) = f_j(y) + \int_y^z f_j'\,dx.$$

Averaging with respect to $y \in (a,b)$, we obtain

$$|f_j(z)| \le \fint_a^b |f_j|\,dy + \int_a^b |f_j'|\,dx,$$

and so

$$\sup_j ||f_j||_{L^\infty(a,b)} < \infty.$$

Since $f_j \to f$ $\mathcal{L}^n$ a.e., $||f||_{L^\infty(a,b)} < \infty$.

4. It follows from the claim that each point of approximate continuity of f is a Lebesgue point and hence

$$f^\epsilon(t) \to f(t) \tag{$\star$}$$

as $\epsilon \to 0$ for each point of approximate continuity of f. Consequently, for each partition $\{a < t_1 < \cdots < t_{m+1} < b\}$, with each t_j a point of approximate continuity of f,

$$\begin{aligned}\sum_{j=1}^{m} |f(t_{j+1}) - f(t_j)| &= \lim_{\epsilon \to 0} \sum_{j=1}^{m} |f^\epsilon(t_{j+1}) - f^\epsilon(t_j)| \\ &\leq \limsup_{\epsilon \to 0} \int_a^b |(f^\epsilon)'|\, dx \\ &\leq \|Df\|(a,b).\end{aligned}$$

Thus

$$\operatorname{ess} V_a^b f \leq \|Df\|(a,b) < \infty. \quad \blacksquare$$

5.10.2 Essential variation on a.e. line

We next extend our analysis to BV functions on $\mathbb{R}^n$.

NOTATION Suppose $f : \mathbb{R}^n \to \mathbb{R}$. Then for $k = 1, \ldots, n$, set $x' = (x_1, \ldots x_{k-1}, x_{k+1}, \ldots x_n) \in \mathbb{R}^{n-1}$, and $t \in \mathbb{R}$, write

$$f_k(x', t) \equiv f(\ldots, x_{k-1}, t, x_{k+1}, \ldots).$$

Thus $\operatorname{ess} V_a^b f_k$ means the essential variation of f_k as a function of $t \in (a,b)$, for each fixed x'.

LEMMA 1
Assume $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, $k \in \{1, \ldots, n\}$, $-\infty \leq a < b \leq \infty$. Then the mapping

$$x' \mapsto \operatorname{ess} V_a^b f_k$$

is $\mathcal{L}^{n-1}$-measurable.

PROOF According to Theorem 1, for $\mathcal{L}^{n-1}$ a.e. $x' \in \mathbb{R}^{n-1}$,

$$\begin{aligned}\operatorname{ess} V_a^b f_k &= \|Df_k\|(a,b) \\ &= \sup\left\{ \int_a^b f_k(x', t)\varphi'(t)\, dt \mid \varphi \in C_c^1(a,b),\ |\varphi| \leq 1 \right\}.\end{aligned}$$

Let $\{\varphi_j\}_{j=1}^{\infty}$ be a countable, dense subset of $C_c^1(a,b) \cap \{|\varphi| \leq 1\}$. Then

$$x' \mapsto \int_a^b f_k(x',t)\varphi_j'(t)\,dt$$

is $\mathcal{L}^{n-1}$-measurable for $j = 1, \dots$ and so

$$x' \mapsto \sup_j \left\{ \int_a^b f_k(x',t)\varphi_j'(t)\,dt \right\} = \text{ess } V_a^b f_k$$

is $\mathcal{L}^{n-1}$-measurable. ∎

THEOREM 2
Assume $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. Then $f \in BV_{\text{loc}}(\mathbb{R}^n)$ if and only if

$$\int_K \text{ess } V_a^b f_k \, dx' < \infty$$

for each $k = 1, \dots, n$, $a < b$, and compact set $K \subset \mathbb{R}^{n-1}$.

PROOF

1. First suppose $f \in BV_{\text{loc}}(\mathbb{R}^n)$. Choose k, a, b, K as above. Se

$$C = \{x \mid a \leq x_k \leq b, (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \in K\}$$

Let $f^\epsilon \equiv \eta_\epsilon * f$, as before. Then

$$\begin{cases} \lim_{\epsilon \to 0} \int_C |f^\epsilon - f|\,dx = 0, \\ \limsup_{\epsilon \to 0} \int_C |Df^\epsilon|\,dx < \infty. \end{cases}$$

Thus for $\mathcal{H}^{n-1}$ a.e. $x' \in K$,

$$f_k^\epsilon \to f_k \text{ in } L^1(a,b),$$

where

$$f_k^\epsilon(x',t) \equiv f^\epsilon(\dots x_{k-1}, t, x_{k+1} \dots).$$

Hence

$$\text{ess } V_a^b f_k \leq \liminf_{\epsilon \to 0} \text{ ess } V_a^b f_k^\epsilon \qquad \text{for } \mathcal{H}^{n-1} \text{ a.e. } x' \in K$$

Thus Fatou's Lemma implies

$$\begin{aligned}\int_K \text{ess } V_a^b f_k \, dx' &\leq \liminf_{\epsilon \to 0} \int_K \text{ess } V_a^b f_k^\epsilon \, dx' \\ &= \liminf_{\epsilon \to 0} \int_C \left| \frac{\partial f^\epsilon}{\partial x_k} \right| dx \\ &\leq \limsup_{\epsilon \to 0} \int_C |Df^\epsilon| \, dx < \infty.\end{aligned}$$

2. Now suppose $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and

$$\int_K \text{ess } V_a^b f_k \, dx' < \infty$$

for all $k = 1, \ldots, n$, $a < b$, and compact $K \subset \mathbb{R}^{n-1}$. Fix $\varphi \in C_c^\infty(\mathbb{R}^n)$, $|\varphi| \leq 1$, and choose a, b, and k such that

$$\text{spt}(\varphi) \subset \{x \mid a < x_k < b\}.$$

Then Theorem 1 implies

$$\int_{\mathbb{R}^n} f \frac{\partial \varphi}{\partial x_k} \, dx \leq \int_K \text{ess } V_a^b f_k \, dx' < \infty,$$

for

$$K \equiv \{x' \in \mathbb{R}^{n-1} \mid (\ldots x_{k-1}, t, x_{k+1}, \ldots) \in \text{spt}(\varphi) \text{ for some } t \in \mathbb{R}\}.$$

As this estimate holds for $k = 1, \ldots, n$, $f \in BV_{\text{loc}}(\mathbb{R}^n)$. ∎

5.11 A criterion for finite perimeter

We conclude this chapter by establishing a relatively simple criterion for a set E to have locally finite perimeter.

NOTATION Write $x \in \mathbb{R}^n$ as $x = (x', t)$, for $x' = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$, $t = x_n \in \mathbb{R}$. The *projection* $P : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is

$$P(x) = x' \qquad (x = (x', x_n) \in \mathbb{R}^n).$$

DEFINITION *Set $N(P \mid A, x') = \mathcal{H}^0(A \cap P^{-1}\{x'\})$ for Borel sets $A \subset \mathbb{R}^n$ and $x' \in \mathbb{R}^{n-1}$.*

LEMMA 1

(i) *The mapping* $x' \mapsto N(P \mid A, x')$ *is* $\mathcal{L}^{n-1}$*-measurable.*

(ii) $\int_{\mathbb{R}^{n-1}} N(P \mid A, x')\, dx' \leq \mathcal{H}^{n-1}(A)$.

PROOF Assertions (i) and (ii) follow as in the proof of Lemma 2, Section 3.4.1; see also the remark in Section 3.4.1.

DEFINITIONS *Let* $E \subset \mathbb{R}^n$ *be* $\mathcal{L}^n$*-measurable. We define*

$$I \equiv \left\{ x \in \mathbb{R}^n \mid \lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r) - E)}{r^n} = 0 \right\}$$

*to be the **measure theoretic interior** of E and*

$$O \equiv \left\{ x \in \mathbb{R}^n \mid \lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{r^n} = 0 \right\}$$

*to be the **measure theoretic exterior** of E.*

REMARK Note $\partial_* E = \mathbb{R}^n - (I \cup O)$. Think of I as denoting the "inside" and O as denoting the "outside" of E. ∎

LEMMA 2

(i) I, O, *and* $\partial_* E$ *are Borel measurable sets.*

(ii) $\mathcal{L}^n((I - E) \cup (E - I)) = 0$.

PROOF

1. There exists a Borel set $C \subset \mathbb{R}^n - E$ such that $\mathcal{L}^n(C \cap T) = \mathcal{L}^n(T - E)$ for all $\mathcal{L}^n$-measurable sets T. Thus

$$I = \left\{ x \mid \lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r) \cap C)}{r^n} = 0 \right\},$$

and so is Borel measurable. The proof for O is similar.

2. Assertion (ii) follows from Corollary 3 in Section 1.7.1. ∎

THEOREM 1 CRITERION FOR FINITE PERIMETER

Let $E \subset \mathbb{R}^n$ *be* $\mathcal{L}^n$*-measurable. Then E has locally finite perimeter if, and only if,*

$$\mathcal{H}^{n-1}(K \cap \partial_* E) < \infty \qquad (\star)$$

for each compact set $K \subset \mathbb{R}^n$.

PROOF

1. Assume first $(\star)$ holds, fix $a > 0$, and set

$$U \equiv (-a, a)^n \subset \mathbb{R}^n.$$

To simplify notation slightly, let us write $z = x' \in \mathbb{R}^{n-1}$, $t = x_n \in \mathbb{R}$. Note from Lemma 1 and hypothesis $(\star)$

$$\int_{\mathbb{R}^{n-1}} N(P \mid U \cap \partial_\star E, z)\, dz \le \mathcal{H}^{n-1}(U \cap \partial_\star E) < \infty. \tag{$\star\star$}$$

Define for each $z \in \mathbb{R}^{n-1}$

$$f^z(t) \equiv \chi_I(z,t) \qquad (t \in \mathbb{R}).$$

Assume $\varphi \in C^1_c(U)$, $|\varphi| \le 1$, and then compute

$$\begin{aligned} \int_E \operatorname{div}(\varphi e_n)\, dx = \int_I \operatorname{div}(\varphi e_n)\, dx &= \int_I \frac{\partial \varphi}{\partial x_n}\, dx \\ &= \int_{\mathbb{R}^{n-1}} \left[\int_{\mathbb{R}} f^z(t) \frac{\partial \varphi}{\partial x_n}(z,t)\, dt \right] dz \\ &\le \int_V \operatorname{ess} V^a_{-a} f^z\, dz \end{aligned} \tag{$\star\star\star$}$$

where

$$V \equiv (-a,a)^{n-1} \subset \mathbb{R}^{n-1}.$$

2. For positive integers k and m, define these sets:

$$\begin{cases} G(k) \equiv \left\{ x \in \mathbb{R}^n \mid \mathcal{L}^n(B(x,r) \cap O) \le \dfrac{\alpha(n-1)}{3^{n+1}} r^n \text{ for } 0 < r < \dfrac{3}{k} \right\}, \\ H(k) \equiv \left\{ x \in \mathbb{R}^n \mid \mathcal{L}^n(B(x,r) \cap I) \le \dfrac{\alpha(n-1)}{3^{n+1}} r^n \text{ for } 0 < r < \dfrac{3}{k} \right\}, \end{cases}$$

$$\begin{cases} G^+(k,m) \equiv G(k) \cap \{x \mid x + se_n \in O \text{ for } 0 < s < 3/m\}, \\ G^-(k,m) \equiv G(k) \cap \{x \mid x - se_n \in O \text{ for } 0 < s < 3/m\}, \\ H^+(k,m) \equiv H(k) \cap \{x \mid x + se_n \in I \text{ for } 0 < s < 3/m\}, \\ H^-(k,m) \equiv H(k) \cap \{x \mid x - se_n \in I \text{ for } 0 < s < 3/m\}. \end{cases}$$

3. Claim #1:

$$\mathcal{L}^{n-1}(P(G^\pm(k,m))) = \mathcal{L}^{n-1}(P(H^\pm(k,m))) = 0 \qquad (k,m = 1,2,\ldots).$$

Proof of Claim #1: For fixed k, m, write

$$G^+(k,m) = \bigcup_{j=-\infty}^{\infty} G_j,$$

where

$$G_j \equiv G^+(k,m) \cap \left\{ x \mid \frac{j-1}{m} \leq x_n < \frac{j}{m} \right\}.$$

Assume $z \in \mathbb{R}^{n-1}, 0 < r < \min\{1/k, 1/m\}$, and $B(z,r) \cap P(G_j) \neq \emptyset$. Then there exists a point $b \in G_j \cap P^{-1}(B(z,r)) \subset G(k)$ such that

$$b_n + \frac{r}{2} > \sup\{x_n \mid x \in G_j \cap P^{-1}(B(x,r))\}.$$

Thus, by the definition of $G^+(k,m)$, we have

$$\left\{ y \mid b_n + \frac{r}{2} \leq y_n \leq b_n + r \right\} \cap P^{-1}(P(G_j) \cap B(z,r)) \subset O \cap B(b, 3r).$$

Take the $\mathcal{L}^n$-measure of each side above to calculate

$$\frac{r}{2}\mathcal{L}^{n-1}(P(G_j) \cap B(z,r)) \leq \mathcal{L}^n(O \cap B(b,3r)) \leq \frac{\alpha(n-1)}{3^{n+1}}(3r)^n,$$

since $b \in G(k)$. Then

$$\limsup_{r \to 0} \frac{\mathcal{L}^{n-1}(P(G_j) \cap B(z,r))}{\alpha(n-1)r^{n-1}} \leq \frac{2}{3}$$

for all $z \in \mathbb{R}^{n-1}$. This implies

$$\mathcal{L}^{n-1}(P(G_j)) = 0 \qquad (j = 0 \pm 1, \pm 2, \ldots),$$

and consequently

$$\mathcal{L}^{n-1}(P(G^+(k,m))) = 0.$$

Similar arguments imply

$$\mathcal{L}^{n-1}(P(G^-(k,m))) = \mathcal{L}^{n-1}(P(H^\pm(k,m))) = 0$$

for all k, m.

4. Now suppose

$$z \in V - \bigcup_{k,m=1}^{\infty} P[G^+(k,m) \cup G^-(k,m) \cup H^+(k,m) \cup H^-(k,m)] \qquad (\star\star\star\star)$$

and

$$N(P \mid U \cap \partial_* E, z) < \infty.$$

Assume $-a < t_1 < \cdots < t_{m+1} < a$ are points of approximate continuity of f^z. Notice that $|f^z(t_{j+1}) - f^z(t_j)| \neq 0$ if and only if $|f^z(t_{j+1}) - f^z(t_j)| = 1$. In the latter case we may, for definiteness, suppose $(z, t_j) \in I$, $(z, t_{j+1}) \notin I$. Since t_{j+1} is a point of approximate continuity of f^z and since $\mathbb{R}^n - (O \cup I) = \partial_* E$, it follows from the finiteness of $N(P \mid U \cap \partial_* E, z)$ that every neighborhood of t_{j+1} must contain points s such that $(z, s) \in O$ and f^z is approximately continuous at s. Consequently,

$$\text{ess } V_{-a}^{a} f^z = \sup \left\{ \sum_{j=1}^{m} |f^z(t_{j+1}) - f^z(t_j)| \right\},$$

the supremum taken over points $-a < t_1 < \cdots < t_{m+1} < a$ such that $(z, t_j) \in O \cup I$ and f^z is approximately continuous at each t_j.

5. Claim #2: If $(z, u) \in I$ and $(z, v) \in O$, with $u < v$, there exists $u < t < v$ such that $(z, t) \in \partial_* E$.

Proof of Claim #2: Suppose not; then $(z, t) \in O \cup I$ for all $u < t < v$. We observe that

$$I \subset \bigcup_{k=1}^{\infty} G(k), \qquad O \subset \bigcup_{k=1}^{\infty} H(k),$$

and that the $G(k)$, $H(k)$ are increasing and closed. Hence there exists k_0 such that $(z, u) \in G(k_0)$, $(z, v) \in H(k_0)$. Now $H(k_0) \cap G(k_0) = \emptyset$, and so

$$u_0 \equiv \sup\{t \mid (z, t) \in G(k_0),\ t < v\} < v.$$

Set

$$v_0 \equiv \inf\{t \mid (z, t) \in H(k_0),\ t > u_0\}.$$

Then

$$(z, u_0) \in G(k_0), \qquad (z, v_0) \in H(k_0),$$

$$u \le u_0 < v_0 \le v,$$

and

$$\{(z, t) \mid u_0 < t < v_0\} \cap [H(k_0) \cup G(k_0)] = \emptyset.$$

Next, there exist

$$u_0 < s_1 < t_1 < v_0$$

with $(z, s_1) \in I$, $(z, t_1) \in O$; this is a consequence of $(\star\star\star\star)$. Arguing as above, we find $k_1 > k_0$ and numbers u_1, v_1 such that

$$u_0 < u_1 < v_1 < v_0, \qquad (z, u_1) \in G(k_1), \qquad (z, v_1) \in H(k_1),$$

and $(z,t) \notin H(k_1) \cup G(k_1)$ if $u_1 < t < v_1$. Continuing, there exist $k_j \to \infty$ and sequences $\{u_j\}_{j=1}^{\infty}$, $\{v_j\}_{j=1}^{\infty}$ such that

$$\begin{cases} u_0 < u_1 < \ldots, \quad v_0 > v_1 > v_2 \ldots, \\ u_j < v_j \text{ for all } j = 1, 2, \ldots, \\ (z, u_j) \in G(k_j), \quad (z, v_j) \in H(k_j), \\ (z,t) \notin G(k_j) \cup H(k_j) \text{ if } u_j < t < v_j. \end{cases}$$

Choose

$$\lim_{j\to\infty} u_j \le t \le \lim_{j\to\infty} v_j.$$

Then

$$y \equiv (z,t) \notin \bigcup_{j=1}^{\infty} [G(k_j) \cup H(k_j)],$$

whence

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(y,r)\cap E)}{r^n} \ge \frac{\alpha(n-1)}{3^{n+1}}$$

and

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(y,r) - E)}{r^n} \ge \frac{\alpha(n-1)}{3^{n+1}} \ .$$

Thus $y \in \partial_* E$.

6. Now, by Claim #2, if z satisfies $(\star\star\star\star)$,

$$\begin{aligned} \text{ess } V_{-a}^{a} f^z &\le \text{ Card } \{t \mid -a < t < a, (z,t) \in \partial_* E\} \\ &= N(P \mid U \cap \partial_* E, z). \end{aligned}$$

Thus $(\star\star\star)$ implies

$$\int_V \text{ess } V_{-a}^{a} f^z \, dz \le \int_V N(P \mid U \cap \partial_* E, z)\, dz \le \mathcal{H}^{n-1}(U \cap \partial_* E) < \infty$$

and analogous inequalities hold for the other coordinate directions. According to Theorem 2 in Section 5.10, E has locally finite perimeter.

7. The necessity of $(\star)$ was established in Theorem 1 in Section 5.8. ∎

6

Differentiability and Approximation by C^1 Functions

In this final chapter we examine more carefully the differentiability properties of BV, Sobolev, and Lipschitz functions. We will see that such functions are differentiable in various senses for $\mathcal{L}^n$ a.e. point in $\mathbb{R}^n$, and as a consequence are equal to C^1 functions except on small sets.

Section 6.1 investigates differentiability $\mathcal{L}^n$ a.e. in certain L^p-senses, and Section 6.2 extends these ideas to show functions in $W^{1,p}$ for $p > n$ are in fact $\mathcal{L}^n$ a.e. differentiable in the classical sense. Section 6.3 recounts the elementary properties of convex functions. In Section 6.4 we prove Aleksandrov's Theorem, asserting a convex function is twice differentiable $\mathcal{L}^n$ a.e. Whitney's Extension Theorem, ensuring the existence of C^1 extensions, is proved in Section 6.5 and is utilized in Section 6.6 to show approximation by C^1 functions.

6.1 L^p differentiability; Approximate differentiability

6.1.1 L^{1^*} differentiability a.e. for BV

Assume $f \in BV_{\mathrm{loc}}(\mathbb{R}^n)$.

NOTATION We recall from Section 5.1 the notation

$$[Df] = [Df]_{\mathrm{ac}} + [Df]_{\mathrm{s}} = \mathcal{L}^n \llcorner Df + [Df]_{\mathrm{s}},$$

where $Df \in L^1_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^n)$ is the density of the absolutely continuous part $[Df]_{\mathrm{ac}}$ of $[Df]$, and $[Df]_{\mathrm{s}}$ is the singular part.

We first demonstrate that near $\mathcal{L}^n$ a.e. point x, f can be approximated in an integral norm by a linear tangent mapping.

THEOREM 1
Assume $f \in BV_{\text{loc}}(\mathbb{R}^n)$. Then for $\mathcal{L}^n$ a.e. $x \in \mathbb{R}^n$,

$$\left(\fint_{B(x,r)} |f(y) - f(x) - Df(x) \cdot (x - y)|^{1^*} \, dy \right)^{\frac{1}{1^*}} = o(r) \qquad \text{as } r \to 0.$$

PROOF
1. $\mathcal{L}^n$ a.e. point $x \in \mathbb{R}^n$ satisfies these conditions:

(a) $\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)| \, dy = 0.$

(b) $\lim_{r\to 0} \fint_{B(x,r)} |Df(y) - Df(x)| \, dy = 0.$

(c) $\lim_{r\to 0} |[Df]_s|(B(x,r))/r^n = 0.$

2. Fix such a point x; we may as well assume $x = 0$. Choose $r > 0$ and let $f^\epsilon \equiv \eta_\epsilon * f$. Select $y \in B(r)$ and write $g(t) \equiv f^\epsilon(ty)$. Then

$$g(1) = g(0) + \int_0^1 g'(s) \, ds,$$

that is,

$$\begin{aligned} f^\epsilon(y) &= f^\epsilon(0) + \int_0^1 Df^\epsilon(sy) \cdot y \, ds \\ &= f^\epsilon(0) + Df(0) \cdot y + \int_0^1 [Df^\epsilon(sy) - Df(0)] \cdot y \, ds. \end{aligned}$$

3. Choose any function $\varphi \in C_c^1(B(r))$ with $|\varphi| \leq 1$, multiply by φ, and average over $B(r)$:

$$\begin{aligned} &\fint_{B(r)} \varphi(y)(f^\epsilon(y) - f^\epsilon(0) - Df(0) \cdot y) \, dy \\ &\qquad = \int_0^1 \left(\fint_{B(r)} \varphi(y)[Df^\epsilon(sy) - Df(0)] \cdot y \, dy \right) ds \\ &\qquad = \int_0^1 \frac{1}{s} \left(\fint_{B(rs)} \varphi\left(\frac{z}{s}\right) [Df^\epsilon(z) - Df(0)] \cdot z \, dz \right) ds. \qquad (\star) \end{aligned}$$

Now

$$\begin{aligned}
g_\epsilon(s) &\equiv \int_{B(rs)} \varphi\left(\frac{z}{s}\right) Df^\epsilon(z)\cdot z\,dz \\
&= -\int_{B(rs)} f^\epsilon(z)\,\mathrm{div}\left(\varphi\left(\frac{z}{s}\right)z\right)\,dz \\
&\to -\int_{B(rs)} f(z)\,\mathrm{div}\left(\varphi\left(\frac{z}{s}\right)z\right)\,dz \qquad \text{as } \epsilon\to 0 \\
&= \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z\cdot\, d[Df] \\
&= \int_{B(rs)} \varphi\left(\frac{z}{s}\right) Df(z)\cdot z\,dz + \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z\cdot\, d[Df]_s.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\frac{|g_\epsilon(s)|}{s^{n+1}} &\le \frac{r}{s^n}\int_{B(rs)} |Df^\epsilon(z)|\,dz \\
&= \frac{r}{s^n}\int_{B(rs)} \left|\int_{\mathbb{R}^n} D\eta_\epsilon(z-y)f(y)\,dy\right|\,dz \\
&= \frac{r}{s^n}\int_{B(rs)} \left|\int_{\mathbb{R}^n} \eta_\epsilon(z-y)\,d[Df]\right|\,dz \\
&\le \frac{r}{s^n}\int_{B(rs)}\int_{\mathbb{R}^n} \eta_\epsilon(z-y)\,d\|Df\|\,dz \\
&= \frac{r}{s^n}\int_{\mathbb{R}^n}\int_{B(rs)} \eta_\epsilon(z-y)\,dz\,d\|Df\| \\
&\le \frac{C}{s^n\epsilon^n}\int_{B(rs+\epsilon)}\int_{B(rs)\cap B(y,\epsilon)} dz\,d\|Df\| \\
&\le C\frac{\min((rs)^n,\epsilon^n)}{s^n\epsilon^n}\|Df\|(B(rs+\epsilon)) \\
&\le C\frac{\min((rs)^n,\epsilon^n)(rs+\epsilon)^n}{s^n\epsilon^n} \\
&\le C \qquad \text{for } 0<\epsilon, s\le 1.
\end{aligned}$$

4. Therefore, applying the Dominated Convergence Theorem to ($\star$), we find

$$\begin{aligned}\fint_{B(r)} &\varphi(y)(f(y) - f(0) - Df(0)\cdot y))\,dy \\ &\leq Cr\int_0^1 \fint_{B(rs)} |Df(z) - Df(0)|\,dz\,ds + Cr\int_0^1 \frac{\|[Df]_s\|(B(rs))}{(rs)^n}\,ds \\ &= o(r) \text{ as } r \to 0.\end{aligned}$$

Take the supremum over all φ as above to find

$$\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|\,dy = o(r) \text{ as } r \to 0. \qquad (\star\star)$$

5. Finally, observe from Theorem 1(ii) in Section 5.6.1 that

$$\begin{aligned}\left(\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|^{\frac{n}{n-1}}\,dy\right)^{\frac{n-1}{n}} &\leq C\frac{\|D(f - f(0) - Df(0)\cdot y)\|(B(r))}{r^{n-1}} \\ &\quad + C\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|\,dy \\ &= o(r) \text{ as } r \to 0,\end{aligned}$$

according to ($\star\star$), (b), and (c). ∎

6.1.2 L^{p^*} differentiability a.e. for $W^{1,p}$ $(1 \leq p < n)$

We can improve the local approximation by tangent planes if f is a Sobolev function.

THEOREM 2
Assume $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ for $1 \leq p < n$. Then for $\mathcal{L}^n$ a.e. $x \in \mathbb{R}^n$,

$$\left(\fint_{B(x,r)} |f(y) - f(x) - Df(x)\cdot(y - x)|^{p^*}\,dy\right)^{1/p^*} = o(r) \quad \text{as} \quad r \to 0.$$

PROOF

1. $\mathcal{L}^n$ a.e. point $x \in \mathbb{R}^n$ satisfies

(a) $\lim_{r\to 0} \fint_{B(x,r)} |f(x) - f(y)|^p\,dy = 0$.

(b) $\lim_{r\to 0} \fint_{B(x,r)} |Df(x) - Df(y)|^p\,dy = 0$.

2. Fix such a point x; we may as well assume $x = 0$. Select $\varphi \in C^1_c(B(r))$ with $\|\varphi\|_{L^q(B(r))} \leq 1$, where $1/p + 1/q = 1$. Then, as in the previous proof, we calculate

$$\begin{aligned}
&\fint_{B(r)} \varphi(y)(f(y) - f(0) - Df(0)\cdot y)\, dy \\
&\quad = \int_0^1 \frac{1}{s} \fint_{B(rs)} \varphi\left(\frac{z}{s}\right) [Df(z) - Df(0)]\cdot z \, dz\, ds \\
&\quad \le r \int_0^1 \left(\fint_{B(rs)} \left|\varphi\left(\frac{z}{s}\right)\right|^q dz\right)^{1/q} \left(\fint_{B(rs)} |Df(z) - Df(0)|^p \, dz\right)^{1/p} ds.
\end{aligned}$$

Since

$$\fint_{B(rs)} \left|\varphi\left(\frac{z}{s}\right)\right|^q dz = \fint_{B(r)} |\varphi(y)|^q \, dy \le \frac{1}{\alpha(n) r^n},$$

we obtain

$$\fint_{B(r)} \varphi(y)(f(y) - f(0) - Df(0)\cdot y)\, dy = o(r^{1-n/q}) \text{ as } r \to 0.$$

Taking the supremum over all functions φ as above gives

$$\frac{1}{r^n}\left(\int_{B(r)} |f(y) - f(0) - Df(0)\cdot y|^p \, dy\right)^{1/p} = o(r^{1-n/q}),$$

and so

$$\left(\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|^p \, dy\right)^{1/p} = o(r) \text{ as } r \to 0. \qquad (\star)$$

3. Thus Theorem 2(ii) in Section 4.5.1 implies

$$\begin{aligned}
&\left(\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|^{p^*} \, dy\right)^{1/p^*} \\
&\qquad \le Cr \left(\fint_{B(r)} |Df(y) - Df(0)|^p \, dy\right)^{1/p} \\
&\qquad\quad + C \left(\fint_{B(r)} |f(y) - f(0) - Df(0)\cdot y|^p \, dy\right)^{1/p} \\
&\qquad = o(r) \text{ as } r \to 0,
\end{aligned}$$

according to $(\star)$ and (b). ∎

6.1.3 Approximate differentiability

DEFINITION *Let $f : \mathbb{R}^n \to \mathbb{R}^m$. We say f is **approximately differentiable** at $x \in \mathbb{R}^n$ if there exists a linear mapping*

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$\text{ap}\lim_{y \to x} \frac{|f(y) - f(x) - L(y - x)|}{|y - x|} = 0.$$

(See Section 1.7.2 for the definition of the approximate limit.)

NOTATION As proved below, such an L, if it exists, is unique. We write

$$\text{ap } Df(x)$$

for L and call ap $Df(x)$ the *approximate derivative* of f at x.

THEOREM 3
An approximate derivative is unique and, in particular, ap $Df = 0$ $\mathcal{L}^n$ a.e. on $\{f = 0\}$.

PROOF Suppose

$$\text{ap}\lim_{y \to x} \frac{|f(y) - f(x) - L(y - x)|}{|y - x|} = 0$$

and

$$\text{ap}\lim_{y \to x} \frac{|f(y) - f(x) - L'(y - x)|}{|y - x|} = 0.$$

Then for each $\epsilon > 0$,

$$\lim_{r \to 0} \frac{\mathcal{L}^n\left(B(x,r) \cap \left\{y \mid \frac{|f(y)-f(x)-L(y-x)|}{|y-x|} > \epsilon\right\}\right)}{\mathcal{L}^n(B(x,r))} = 0 \qquad (\star)$$

and

$$\lim_{r \to 0} \frac{\mathcal{L}^n\left(B(x,r) \cap \left\{y \mid \frac{|f(y)-f(x)-L'(y-x)|}{|y-x|} > \epsilon\right\}\right)}{\mathcal{L}^n(B(x,r))} = 0. \qquad (\star\star)$$

If $L \neq L'$, set

$$6\epsilon \equiv ||L - L'|| \equiv \max_{|z|=1} |(L - L')(z)| > 0$$

and consider then the sector

$$S \equiv \left\{y \mid \ |(L - L')(y - x)| \geq \frac{||L - L'|| \, |y - x|}{2}\right\}.$$

Note

$$\frac{\mathcal{L}^n(B(x,r)\cap S)}{\mathcal{L}^n(B(x,r))} \equiv a > 0 \qquad (\star\star\star)$$

for all $r > 0$.

But if $y \in S$,

$$\begin{aligned} 3\epsilon|y-x| &= \frac{\|L-L'\||y-x|}{2} \\ &\leq |(L-L')(y-x)| \\ &\leq |f(y)-f(x)-L(y-x)| + |f(y)-f(x)-L'(y-x)| \end{aligned}$$

so that

$$S \subset \left\{ y \mid \frac{|f(y)-f(x)-L(y-x)|}{|y-x|} > \epsilon \right\} \cup \left\{ y \mid \frac{|f(y)-f(x)-L'(y-x)|}{|y-x|} > \epsilon \right\}.$$

Thus $(\star)$ and $(\star\star)$ imply

$$\lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap S)}{\mathcal{L}^n(B(x,r))} = 0,$$

a contradiction to $(\star\star\star)$. ∎

THEOREM 4

Assume $f \in BV_{\text{loc}}(\mathbb{R}^n)$. Then f is approximately differentiable $\mathcal{L}^n$ a.e.

REMARK

(i) We show in addition that

$$\text{ap } Df = Df \qquad \mathcal{L}^n \text{ a.e.},$$

the right-hand function defined in Section 5.1.

(ii) Since $W^{1,p}_{\text{loc}}(\mathbb{R}^n) \subset BV_{\text{loc}}(\mathbb{R}^n)$ $(1 \leq p \leq \infty)$, we see that *each Sobolev function is approximately differentiable $\mathcal{L}^n$ a.e. and its approximate derivative equals its weak derivative $\mathcal{L}^n$ a.e.* ∎

PROOF Choose a point $x \in \mathbb{R}^n$ such that

$$⨍_{B(x,r)} |f(y)-f(x)-Df(x)\cdot(y-x)|\, dy = o(r) \text{ as } r \to 0; \qquad (\star)$$

$\mathcal{L}^n$ a.e. x will do according to Theorem 1.

Suppose

$$\text{ap } \limsup_{y\to x} \frac{|f(y)-f(x)-Df(x)\cdot(y-x)|}{|y-x|} > \theta > 0.$$

Then there exist $r_j \to 0$ and $\gamma > 0$ such that

$$\frac{\mathcal{L}^n(\{y \in B(x, r_j) \mid |f(y) - f(x) - Df(x) \cdot (y - x)| > \theta |y - x|\})}{\alpha(n) r_j^n} \geq \gamma > 0.$$

Hence there exists $\sigma > 0$ such that

$$\frac{\mathcal{L}^n(\{y \in B(x, r_j) - B(x, \sigma r_j) \mid |f(y) - f(x) - Df(x) \cdot (y - x)| > \theta |y - x|\})}{\alpha(n) r_j^n} \geq \frac{\gamma}{2}$$

for $j = 1, 2, \ldots$. Since $|y - x| > \sigma r_j$ for $y \in B(x, r_j) - B(x, \sigma r_j)$,

$$\frac{\mathcal{L}^n(\{y \in B(x, r_j) \mid |f(y) - f(x) - Df(x) \cdot (y - x)| > \theta \sigma r_j\})}{\alpha(n) r_j^n} \geq \frac{\gamma}{2} \qquad (\star\star)$$

for $j = 1, \ldots$. But by ($\star$), the expression on the left-hand side of ($\star\star$) is less than or equal to

$$\frac{o(r_j)}{\theta \sigma r_j} = o(1) \text{ as } r_j \to 0,$$

a contradiction to ($\star\star$).

Thus

$$\operatorname{ap} \limsup_{y \to x} \frac{|f(y) - f(x) - Df(x) \cdot (y - x)|}{|y - x|} = 0,$$

and so

$$\operatorname{ap} Df(x) = Df(x). \quad \blacksquare$$

6.2 Differentiability a.e. for $W^{1,p}$ ($p > n$)

Recall from Section 3.1 the following definition:

DEFINITION *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at $x \in \mathbb{R}^n$ if there exists a linear mapping*

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that

$$\lim_{y \to x} \frac{|f(y) - f(x) - L(x - y)|}{|x - y|} = 0.$$

NOTATION If such a linear mapping L exists at x, it is clearly unique, and we write

$$Df(x)$$

for L. We call $Df(x)$ the *derivative* of f at x.

THEOREM 1
Let $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ for some $n < p \le \infty$. Then f is differentiable $\mathcal{L}^n$ a.e., and its derivative equals its weak derivative $\mathcal{L}^n$ a.e.

PROOF Since $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n) \subset W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, we may as well assume $n < p < \infty$.
For $\mathcal{L}^n$ a.e. $x \in \mathbb{R}^n$, we have

$$\lim_{r\to 0} ⨍_{B(x,r)} |Df(z) - Df(x)|^p \, dz = 0. \tag{$\star$}$$

Choose such a point x, and write

$$g(y) \equiv f(y) - f(x) - Df(x)\cdot(y-x) \qquad (y \in B(x,r)).$$

Employing Morrey's estimate from Section 4.5.3, we deduce

$$|g(y) - g(x)| \le Cr\left(⨍_{B(x,r)} |Dg|^p \, dz\right)^{1/p}$$

for $r \equiv |x-y|$. Since $g(x) = 0$ and $Dg = Df - Df(x)$, this reads

$$\begin{aligned}\frac{|f(y) - f(x) - Df(x)\cdot(y-x)|}{|y-x|} &\le C\left(⨍_{B(x,r)} |Df(z) - Df(x)|^p \, dz\right)^{1/p} \\ &= o(1) \qquad \text{as } y \to x\end{aligned}$$

according to $(\star)$. ∎

As an application we have a new proof of Rademacher's Theorem, Section 3.1.2:

THEOREM 2
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function. Then f is differentiable $\mathcal{L}^n$ a.e

PROOF According to Theorem 5 in Section 4.2.3, $f \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n)$. ∎

6.3 Convex functions

DEFINITION *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* **convex** *if*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$$

for all $0 \le \lambda \le 1$, $x, y \in \mathbb{R}^n$.

THEOREM 1
Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex.

(i) *Then f is locally Lipschitz on $\mathbb{R}^n$, and there exists a constant C, depending only on n, such that*

$$\sup_{B(x,\frac{r}{2})} |f| \le C \fint_{B(x,r)} |f| \, dy$$

and

$$\operatorname*{ess\,sup}_{B(x,\frac{r}{2})} |Df| \le \frac{C}{r} \fint_{B(x,r)} |f| \, dy$$

for each ball $B(x,r) \subset \mathbb{R}^n$.

(ii) *If, in addition, $f \in C^2(\mathbb{R}^n)$, then*

$$D^2 f \ge 0 \text{ on } \mathbb{R}^n,$$

that is, $D^2 f$ is a nonnegative definite symmetric matrix on $\mathbb{R}^n$.

PROOF
1. Suppose first that $f \in C^2(\mathbb{R}^n)$ and is convex. Fix $x \in \mathbb{R}^n$. Then for each $y \in \mathbb{R}^n$ and $\lambda \in (0,1)$,

$$f(x + \lambda(y-x)) \le f(x) + \lambda(f(y) - f(x)).$$

Thus

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \le f(y) - f(x).$$

Let $\lambda \to 0$ to obtain

$$f(y) \ge f(x) + Df(x) \cdot (y-x) \tag{$\star$}$$

for all $x, y \in \mathbb{R}^n$.

2. Given now $B(x,r) \subset \mathbb{R}^n$, we fix a point $z \in B(x, r/2)$. Then ($\star$) implies

$$f(y) \ge f(z) + Df(z) \cdot (y - z).$$

We integrate this inequality with respect to y over $B(z,r/2)$ to find

$$f(z) \leq ⨍_{B(z,\frac{r}{2})} f(y)\,dy \leq C ⨍_{B(x,r)} |f|\,dy. \tag{$\star\star$}$$

Next choose a smooth cutoff function $\zeta \in C_c^\infty(\mathbb{R}^n)$ satisfying

$$\begin{cases} 0 \leq \zeta \leq 1, & |D\zeta| \leq \frac{C}{r}, \\ \zeta \equiv 1 \text{ on } B(x,\frac{r}{2}), & \zeta \equiv 0 \text{ on } \mathbb{R}^n - B(x,r). \end{cases}$$

Now ($\star$) implies

$$f(z) \geq f(y) + Df(y) \cdot (z-y).$$

Multiply this inequality by $\zeta(y)$ and integrate with respect to y over $B(x,r)$:

$$\begin{aligned} f(z) \int_{B(x,r)} \zeta(y)\,dy &\geq \int_{B(x,r)} f(y)\zeta(y)\,dy + \int_{B(x,r)} \zeta(y) Df(y) \cdot (z-y)\,dy \\ &= \int_{B(x,r)} f(y)[\zeta(y) - \operatorname{div}(\zeta(y)(z-y))]\,dy \\ &\geq -C \int_{B(x,r)} |f|\,dy. \end{aligned}$$

This inequality implies

$$f(z) \geq -C ⨍_{B(x,r)} |f|\,dy,$$

which estimate together with ($\star\star$) proves

$$|f(z)| \leq C ⨍_{B(x,r)} |f|\,dy. \tag{$\star\star\star$}$$

3. For z as above, define

$$S_z = \left\{ y \mid \frac{r}{4} \leq |y-x| \leq \frac{r}{2}, Df(z) \cdot (y-z) \geq \frac{1}{2}|Df(z)|\,|y-z| \right\},$$

and observe

$$\mathcal{L}^n(S_z) \geq Cr^n$$

where C depends only on n. Use ($\star$) to write

$$f(y) \geq f(z) + \frac{r}{8}|Df(z)|$$

for all $y \in S_z$. Integrating over S_z gives

$$|Df(z)| \leq \frac{C}{r} \fint_{B(x,\frac{r}{2})} |f(y) - f(z)|\, dy.$$

This inequality and ($\star\star\star$) complete the proof of assertion (i) for C^2 convex functions f.

4. If f is merely convex, define $f^\epsilon \equiv \eta_\epsilon * f$, where $\epsilon > 0$ and η_ϵ is the standard mollifier.

Claim: f^ϵ is convex.

Proof of Claim: Fix $x, y \in \mathbb{R}^n$, $0 \leq \lambda \leq 1$. Then for each $z \in \mathbb{R}^n$,

$$\begin{aligned} f(z - (\lambda x + (1-\lambda)y)) &= f(\lambda(z-x) + (1-\lambda)(z-y)) \\ &\leq \lambda f(z-x) + (1-\lambda) f(z-y). \end{aligned}$$

Multiply this estimate by $\eta_\epsilon(z) \geq 0$ and integrate over $\mathbb{R}^n$:

$$\begin{aligned} f^\epsilon(\lambda x + (1-\lambda)y) &= \int_{\mathbb{R}^n} f(z - (\lambda x + (1-\lambda)y))\eta_\epsilon(z)\, dz \\ &\leq \lambda \int_{\mathbb{R}^n} f(z-x)\eta_\epsilon(z)\, dz \\ &\quad + (1-\lambda)\int_{\mathbb{R}^n} f(z-y)\eta_\epsilon(z)\, dz \\ &= \lambda f^\epsilon(x) + (1-\lambda) f^\epsilon(y). \end{aligned}$$

5. According to the estimate proved above for smooth convex functions, we have

$$\sup_{B(x,\frac{r}{2})} \left(|f^\epsilon| + r|Df^\epsilon|\right) \leq C \fint_{B(x,r)} |f^\epsilon|\, dy$$

for each ball $B(x,r) \subset \mathbb{R}^n$. Letting $\epsilon \to 0$, we obtain in the limit the same estimates for f. This proves assertion (i).

6. To prove assertion (ii), recall from Taylor's Theorem

$$f(y) = f(x) + Df(x)\cdot(y-x) + (y-x)^T \cdot \int_0^1 (1-s)D^2 f(x+s(y-x))\, ds \cdot (y-x).$$

This equality and ($\star$) yield

$$(y-x)^T \cdot \int_0^1 (1-s)D^2 f(x+s(y-x))\, ds \cdot (y-x) \geq 0 \qquad (\star\star\star\star)$$

for all $x, y \in \mathbb{R}^n$. Thus, given any vector ξ, set $y = x + t\xi$ in $(\star\star\star\star)$ for $t > 0$ to compute:

$$\xi^T \cdot \int_0^1 (1-s) D^2 f(x + st\xi)\, ds \cdot \xi \geq 0.$$

Send $t \to 0$ to prove

$$\xi^T \cdot D^2 f(x) \cdot \xi \geq 0. \quad \blacksquare$$

THEOREM 2
Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then there exist signed Radon measures $\mu^{ij} = \mu^{ji}$ such that

$$\int_{\mathbb{R}^n} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j}\, dx = \int_{\mathbb{R}^n} \varphi\, d\mu^{ij} \qquad (i, j = 1, \dots, n)$$

for all $\varphi \in C_c^2(\mathbb{R}^n)$. Furthermore, the measures μ^{ii} are nonnegative ($i = 1, \dots, n$).

PROOF
1. Fix any vector $\xi \in \mathbb{R}^n$, $|\xi| = 1$, $\xi = (\xi_1, \dots, \xi_n)$. Let η_ϵ be the standard mollifier. Write $f^\epsilon \equiv \eta_\epsilon * f$. Then f^ϵ is smooth and convex, whence

$$D^2 f^\epsilon \geq 0.$$

Thus for all $\varphi \in C_c^2(\mathbb{R}^n)$ with $\varphi \geq 0$,

$$\sum_{i,j=1}^n \int_{\mathbb{R}^n} f^\epsilon \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \xi_i \xi_j\, dx = \int_{\mathbb{R}^n} \varphi \sum_{i,j=1}^n \frac{\partial^2 f^\epsilon}{\partial x_i \partial x_j} \xi_i \xi_j\, dx \geq 0.$$

Let $\epsilon \to 0$ to conclude

$$L(\varphi) \equiv \sum_{i,j=1}^n \int_{\mathbb{R}^n} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \xi_i \xi_j\, dx \geq 0.$$

Then Corollary 1 in Section 1.8 implies the existence of a Radon measure μ^ξ such that

$$L(\varphi) = \int_{\mathbb{R}^n} \varphi\, d\mu^\xi$$

for all $\varphi \in C_c^2(\mathbb{R}^n)$.

2. Let $\mu^{ii} = \mu^{e_i}$ for $i = 1, \dots, n$. If $i \neq j$, set $\xi \equiv (e_i + e_j)/\sqrt{2}$. Note that in this case

$$\sum_{k,l=1}^n \frac{\partial^2 \varphi}{\partial x_k \partial x_l} \xi_k \xi_l = \frac{1}{2} \left[\frac{\partial^2 \varphi}{\partial x_i \partial x_i} + 2 \frac{\partial^2 \varphi}{\partial x_i \partial x_j} + \frac{\partial^2 \varphi}{\partial x_j \partial x_j} \right].$$

Thus

$$
\begin{aligned}
\int_{\mathbb{R}^n} f \frac{\partial^2 \varphi}{\partial x_i \partial x_j}\, dx &= \int_{\mathbb{R}^n} f \sum_{k,l=1}^{n} \frac{\partial^2 \varphi}{\partial x_k \partial x_l} \xi_k \xi_l \, dx \\
&\quad - \frac{1}{2}\left[\int_{\mathbb{R}^n} f \frac{\partial^2 \varphi}{\partial x_i \partial x_i}\, dx + \int_{\mathbb{R}^n} f \frac{\partial^2 \varphi}{\partial x_j \partial x_j}\, dx\right] \\
&= \int_{\mathbb{R}^n} \varphi \, d\mu^{\xi} - \frac{1}{2}\int_{\mathbb{R}^n} \varphi \, d\mu^{ii} - \frac{1}{2}\int_{\mathbb{R}^n} \varphi \, d\mu^{jj} \\
&= \int_{\mathbb{R}^n} \varphi \, d\mu^{ij},
\end{aligned}
$$

where

$$\mu^{ij} \equiv \mu^{\xi} - \frac{1}{2}\mu^{ii} - \frac{1}{2}\mu^{jj}. \quad \blacksquare$$

THEOREM 3
Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then

$$\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \in BV_{\mathrm{loc}}(\mathbb{R}^n).$$

PROOF Let $V \subset\subset \mathbb{R}^n$, $\varphi \in C_c^2(V, \mathbb{R}^n)$, $|\varphi| \le 1$. Then for $k = 1, \ldots, n$,

$$
\begin{aligned}
\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_k} \operatorname{div} \varphi \, dx &= - \int_{\mathbb{R}^n} f \sum_{i=1}^{n} \frac{\partial^2 \varphi^i}{\partial x_i \partial x_k}\, dx \\
&= \sum_{i=1}^{n} \int_{\mathbb{R}^n} \varphi^i \, d\mu^{ik} \\
&\le \sum_{i=1}^{n} \mu^{ik}(V) < \infty. \quad \blacksquare
\end{aligned}
$$

NOTATION In analogy with the notation introduced in Section 5.1, let us write for a convex function f:

$$[D^2 f] \equiv \begin{pmatrix} \mu^{11} & \cdots & \mu^{1n} \\ \vdots & & \vdots \\ \mu^{n1} & \cdots & \mu^{nn} \end{pmatrix} = \|D^2 f\| \llcorner \Sigma,$$

where $\Sigma : \mathbb{R}^n \to M^{n \times n}$ is $\|D^2 f\|$-measurable, with $|\Sigma| = 1$ $\|D^2 f\|$ a.e. We also write

$$\left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] = \mu^{ij} \qquad (i, j = 1, \ldots. n).$$

By Lebesgue's Decomposition Theorem, we may further set

$$\mu^{ij} = \mu^{ij}_{\text{ac}} + \mu^{ij}_{s},$$

where

$$\mu^{ij}_{\text{ac}} << \mathcal{L}^n, \qquad \mu^{ij}_{s} \perp \mathcal{L}^n.$$

But then

$$\mu^{ij}_{\text{ac}} = \mathcal{L}^n \llcorner f_{ij}$$

for some $f_{ij} \in L^1_{\text{loc}}(\mathbb{R}^n)$. Set

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \equiv f_{ij} \qquad (i, j = 1, \dots, n),$$

$$D^2 f \equiv \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix},$$

$$\left[D^2 f\right]_{\text{ac}} \equiv \begin{pmatrix} \mu^{11}_{\text{ac}} & \cdots & \mu^{1n}_{\text{ac}} \\ \vdots & & \vdots \\ \mu^{n1}_{\text{ac}} & \cdots & \mu^{nn}_{\text{ac}} \end{pmatrix} = \mathcal{L}^n \llcorner D^2 f,$$

$$\left[D^2 f\right]_{s} \equiv \begin{pmatrix} \mu^{11}_{s} & \cdots & \mu^{1n}_{s} \\ \vdots & & \vdots \\ \mu^{n1}_{s} & \cdots & \mu^{nn}_{s} \end{pmatrix}.$$

Thus $[D^2 f] = \left[D^2 f\right]_{\text{ac}} + \left[D^2 f\right]_{s} = \mathcal{L}^n \llcorner D^2 f + \left[D^2 f\right]_{s}$, so that $D^2 f \in L^1_{\text{loc}}(\mathbb{R}^n; M^{n\times n})$ is the density of the absolutely continuous part $[D^2 f]_{\text{ac}}$ of $[D^2 f]$

6.4 Second derivatives a.e. for convex functions

Next we show that a convex function is twice differentiable a.e. This assertion is in the same spirit as Rademacher's Theorem, but is perhaps even more remarkable in that we have only "one-sided control" on the second derivatives.

THEOREM 1 ALEKSANDROV'S THEOREM
Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Then f has a second derivative $\mathcal{L}^n$ a.e. More precisely, for $\mathcal{L}^n$ a.e. x,

$$\left| f(y) - f(x) - Df(x) \cdot (y - x) - \frac{1}{2}(y - x)^T \cdot D^2 f(x) \cdot (y - x) \right|$$
$$= o(|y - x|^2) \text{ as } y \to x. \qquad (\star)$$

PROOF
1. $\mathcal{L}^n$ a.e. point x satisfies these conditions:

$$(a)\ Df(x) \text{ exists and } \lim_{r \to 0} ⨍_{B(x,r)} |Df(y) - Df(x)|\, dy = 0.$$

$$(b)\ \lim_{r \to 0} ⨍_{B(x,r)} |D^2 f(y) - D^2 f(x)|\, dy = 0. \qquad (\star\star)$$

$$(c)\ \lim_{r \to 0} |[D^2 f]_s|\, (B(x,r))/r^n = 0.$$

2. Fix such a point x; we may as well assume $x = 0$. Choose $r > 0$ and let $f^\epsilon \equiv \eta_\epsilon * f$. Fix $y \in B(r)$. By Taylor's Theorem,

$$f^\epsilon(y) = f^\epsilon(0) + Df^\epsilon(0) \cdot y + \int_0^1 (1 - s) y^T \cdot D^2 f^\epsilon(sy) \cdot y\, ds.$$

Add and subtract $(1/2) y^T \cdot D^2 f(0) \cdot y$:

$$f^\epsilon(y) = f^\epsilon(0) + Df^\epsilon(0) \cdot y + \frac{1}{2} y^T \cdot D^2 f(0) \cdot y$$
$$+ \int_0^1 (1 - s) y^T \cdot [D^2 f^\epsilon(sy) - D^2 f(0)] \cdot y\, ds.$$

3. Fix any function $\varphi \in C_c^2(B(r))$ with $|\varphi| \le 1$, multiply the equation above by φ, and average over $B(r)$:

$$⨍_{B(r)} \varphi(y)(f^\epsilon(y) - f^\epsilon(0) - Df^\epsilon(0) \cdot y - \frac{1}{2} y^T \cdot D^2 f(0) \cdot y)\, dy$$
$$= \int_0^1 (1 - s) \left(⨍_{B(r)} \varphi(y) y^T \cdot [D^2 f^\epsilon(sy) - D^2 f(0)] \cdot y\, dy \right) ds$$
$$= \int_0^1 \frac{(1 - s)}{s^2} \left(⨍_{B(rs)} \varphi(\frac{z}{s}) z^T \cdot [D^2 f^\epsilon(z) - D^2 f(0)] \cdot z\, dz \right) ds. \ (\star\star\star)$$

Now

$$
\begin{aligned}
g_\epsilon(s) &\equiv \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z^T \cdot D^2 f^\epsilon(z) \cdot z\, dz \\
&= \int_{B(rs)} f^\epsilon(z) \sum_{i,j=1}^n \frac{\partial^2}{\partial z_i \partial z_j}\left(\varphi\left(\frac{z}{s}\right) z_i z_j\right) dz \\
&\to \int_{B(rs)} f(z) \sum_{i,j=1}^n \frac{\partial^2}{\partial z_i \partial z_j}\left(\varphi\left(\frac{z}{s}\right) z_i z_j\right) dz \qquad \text{as } \epsilon \to 0 \\
&= \sum_{i,j=1}^n \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z_i z_j \, d\mu^{ij} \\
&= \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z^T \cdot D^2 f(z) \cdot z\, dz + \sum_{i,j=1}^n \int_{B(rs)} \varphi\left(\frac{z}{s}\right) z_i z_j \, d\mu_s^{ij}.
\end{aligned}
$$

Furthermore, as in Section 6.1.1, we may calculate

$$
\begin{aligned}
\frac{|g_\epsilon(s)|}{s^{n+2}} &\le \frac{r^2}{s^n} \int_{B(rs)} |D^2 f^\epsilon(z)|\, dz \\
&= \frac{r^2}{s^n} \int_{B(rs)} \left| \int_{\mathbb{R}^n} D^2 \eta_\epsilon(z-y) f(y)\, dy \right| dz \\
&\le \frac{r^2}{s^n} \int_{B(rs)} \left| \int_{\mathbb{R}^n} \eta_\epsilon(z-y)\, d[D^2 f] \right| dz \\
&\le \frac{C}{s^n \epsilon^n} \int_{B(rs+\epsilon)} \left(\int_{B(rs)\cap B(y,\epsilon)} dz \right) d\|D^2 f\| \\
&\le C \frac{\min((rs)^n, \epsilon^n)}{s^n \epsilon^n} \|D^2 f\|(B(rs+\epsilon)) \\
&\le C \frac{\min((rs)^n, \epsilon^n)(rs+\epsilon)^n}{s^n \epsilon^n} \\
&\le C \qquad \text{for } 0 < \epsilon, s \le 1 \text{ by } (\star\star).
\end{aligned}
$$

4. Hence we may apply the Dominated Convergence Theorem to let $\epsilon \to 0$ in $(\star\star\star)$:

$$
\begin{aligned}
&⨍_{B(r)} \varphi(y) \left[f(y) - f(0) - Df(0)\cdot y - \frac{1}{2} y^T \cdot D^2 f(0) \cdot y \right] dy \\
&\quad \le Cr^2 \int_0^1 ⨍_{B(rs)} |D^2 f(z) - D^2 f(0)|\, dz\, ds + Cr^2 \int_0^1 \frac{|[D^2 f]_s|(B(rs))}{(sr)^n}\, ds \\
&\quad = o(r^2) \text{ as } r \to 0, \qquad \text{by } (\star\star) \text{ with } x = 0.
\end{aligned}
$$

Take the supremum over all φ as above to obtain

$$\fint_{B(r)} |h(y)|\, dy = o(r^2) \text{ as } r \to 0 \qquad (\star\star\star\star)$$

for

$$h(y) \equiv f(y) - f(0) - Df(0)\cdot y - \frac{1}{2}y^T \cdot D^2 f(0)\cdot y.$$

5. *Claim #1*: There exists a constant C such that

$$\sup_{B(r/2)} |Dh| \le \frac{C}{r}\fint_{B(r)} |h|\, dy + Cr \qquad (r > 0).$$

Proof of Claim #1: Let $\Lambda \equiv |D^2 f(0)|$. Then $g \equiv h + (\Lambda/2)|y|^2$ is convex: apply Theorem 1 from Section 6.3.

6. *Claim #2:* $\sup_{B(r/2)} |h| = o(r^2)$ as $r \to 0$.

Proof of Claim #2: Fix $0 < \epsilon, \eta < 1$, $\eta^{1/n} \le 1/2$. Then

$$\begin{aligned}
\mathcal{L}^n\{z \in B(r) \mid |h(z)| \ge \epsilon r^2\} &\le \frac{1}{\epsilon r^2}\int_{B(r)} |h|\, dz \\
&= o(r^n) \text{ as } r \to 0, \text{ by } (\star\star\star\star) \\
&< \eta \mathcal{L}^n(B(r)) \qquad \text{for } 0 < r < r_0 \equiv r_0(\epsilon, \eta).
\end{aligned}$$

Thus for each $y \in B(r/2)$ there exists $z \in B(r)$ such that

$$|h(z)| \le \epsilon r^2$$

and

$$|y - z| \le \sigma \equiv \eta^{1/n} r,$$

for if not,

$$\mathcal{L}^n\{z \in B(r) \mid |h(z)| \ge \epsilon r^2\} \ge \mathcal{L}^n(B(y, \sigma)) = \alpha(n)\eta r^n = \eta \mathcal{L}^n(B(r)).$$

Consequently,

$$\begin{aligned}
|h(y)| &\le |h(z)| + |h(y) - h(z)| \\
&\le \epsilon r^2 + \sigma \sup_{B(r)} |Dh| \\
&\le \epsilon r^2 + C\eta^{1/n} r^2 \qquad \text{by Claim \#1 and } (\star\star\star\star) \\
&= 2\epsilon r^2,
\end{aligned}$$

provided we fix η such that $C\eta^{1/n} = \epsilon$ and then choose $0 < r < r_0$.

7. According to Claim #2,

$$\sup_{B(r/2)} \left| f(y) - f(0) - Df(0) \cdot y - \frac{1}{2} y^T \cdot D^2 f(0) \cdot y \right| = o(r^2) \qquad \text{as } r \to 0.$$

This proves ($\star$) for $x = 0$. ∎

6.5 Whitney's Extension Theorem

We next identify conditions ensuring the existence of a C^1 extension $\bar{f}$ of a given function f defined on a closed subset C of $\mathbb{R}^n$.

Let $C \subset \mathbb{R}^n$ be a closed set and assume $f : C \to \mathbb{R}$, $d : C \to \mathbb{R}^n$ are given functions.

NOTATION

(i) $R(y, x) \equiv \frac{f(y) - f(x) - d(x) \cdot (y - x)}{|x - y|} \qquad (x, y \in C, x \neq y)$.

(ii) Let $K \subset C$ be compact, and set

$$\rho_K(\delta) \equiv \sup\{|R(y, x)| \mid 0 < |x - y| \leq \delta,\ x, y \in K\}.$$

THEOREM 1 WHITNEY'S EXTENSION THEOREM

Assume f, d are continuous, and for each compact set $K \subset C$,

$$\rho_K(\delta) \to 0 \text{ as } \delta \to 0. \tag{$\star$}$$

Then there exists a function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

(i) $\bar{f}$ is C^1.

(ii) $\bar{f} = f$, $D\bar{f} = d$ on C.

PROOF

1. The proof will be a kind of "C^1-version" of the proof of the Extension Theorem presented in Section 1.2.

Let $U \equiv \mathbb{R}^n - C$; U is open. Define

$$r(x) \equiv \frac{1}{20} \min\{1, \operatorname{dist}(x, C)\}.$$

By Vitali's Covering Theorem, there exists a countable set $\{x_j\}_{j=1}^{\infty} \subset U$ such that

$$U = \bigcup_{j=1}^{\infty} B(x_j, 5r(x_j))$$

and the balls $\{B(x_j, r(x_j))\}_{j=1}^{\infty}$ are disjoint. For each $x \in U$, define

$$S_x \equiv \{x_j \mid B(x, 10r(x)) \cap B(x_j, 10r(x_j)) \neq \emptyset\}.$$

2. *Claim #1*: Card $(S_x) \leq (129)^n$ and $1/3 \leq r(x)/r(x_j) \leq 3$ if $x_j \in S_x$.

Proof of Claim #1: If $x_j \in S_x$, then

$$\begin{aligned} |r(x) - r(x_j)| \leq \frac{1}{20}|x - x_j| &\leq \frac{1}{20}(10(r(x) + r(x_j))) \\ &= \frac{1}{2}(r(x) + r(x_j)). \end{aligned}$$

Hence

$$r(x) \leq 3r(x_j), \qquad r(x_j) \leq 3r(x).$$

In addition, we have

$$\begin{aligned} |x - x_j| + r(x_j) &\leq 10(r(x) + r(x_j)) + r(x_j) \\ &= 10r(x) + 11r(x_j) \\ &\leq 43r(x); \end{aligned}$$

consequently,

$$B(x_j, r(x_j)) \subset B(x, 43r(x)).$$

As the balls $\{B(x_j, r(x_j))\}_{j=1}^{\infty}$ are disjoint and $r(x_j) \geq r(x)/3$,

$$\text{Card } (S_x)\alpha(n)\left(\frac{r(x)}{3}\right)^n \leq \alpha(n)(43r(x))^n$$

whence

$$\text{Card } (S_x) \leq (129)^n.$$

3. Now choose $\mu : \mathbb{R} \to \mathbb{R}$ such that

$$\mu \in C^\infty, \qquad 0 \leq \mu \leq 1, \qquad \mu(t) \equiv 1 \text{ for } t \leq 1. \qquad \mu(t) \equiv 0 \text{ for } t \geq 2.$$

For each $j = 1, \ldots,$ define

$$u_j(x) \equiv \mu\left(\frac{|x - x_j|}{5r(x_j)}\right) \qquad (x \in \mathbb{R}^n).$$

Then

$$\begin{cases} u_j \in C^\infty, \qquad 0 \leq u_j \leq 1. \\ u_j = 1 \text{ on } B(x_j, 5r(x_j)), \\ u_j = 0 \text{ on } \mathbb{R}^n - B(x_j, 10r(x_j)). \end{cases}$$

Also

$$|Du_j(x)| \leq \frac{C}{r(x_j)} \leq \frac{C_1}{r(x)} \text{ if } x_j \in S_x \qquad (\star\star)$$

and

$$u_j = 0 \text{ on } B(x, 10r(x)) \text{ if } x_j \notin S_x.$$

Define

$$\sigma(x) \equiv \sum_{j=1}^{\infty} u_j(x) \qquad (x \in \mathbb{R}^n).$$

Since $u_j = 0$ on $B(x, 10r(x))$ if $x_j \notin S_x$,

$$\sigma(y) = \sum_{x_j \in S_x} u_j(y) \text{ if } y \in B(x, 10r(x)).$$

By Claim #1, Card $(S_x) \leq (129)^n$; this fact and $(\star\star)$ imply

$$\begin{cases} \sigma \in C^\infty(U), & \sigma \geq 1 \text{ on U} \\ |D\sigma(x)| \leq \dfrac{C_2}{r(x)} & (x \in U). \end{cases}$$

Now for each $j = 1, \dots,$ define

$$v_j(x) \equiv \frac{u_j(x)}{\sigma(x)} \qquad (x \in U).$$

Notice

$$Dv_j = \frac{Du_j}{\sigma} - \frac{u_j D\sigma}{\sigma^2}.$$

Thus

$$\begin{cases} \sum_{j=1}^{\infty} v_j(x) = 1 \\ \sum_{j=1}^{\infty} Dv_j(x) = 0 & (x \in U) \\ |Dv_j(x)| \leq \dfrac{C_3}{r(x)}. \end{cases}$$

The functions $\{v_j\}_{j=1}^{\infty}$ are thus a smooth partition of unity in U.

4. Now for each $j = 1, \dots,$ choose any point $s_j \in C$ such that

$$|x_j - s_j| = \text{dist}(x_j, C).$$

Finally, define $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ this way:

$$\bar{f}(x) \equiv \begin{cases} f(x) & \text{if } x \in C \\ \displaystyle\sum_{j=1}^{\infty} v_j(x)[f(s_j) + d(s_j) \cdot (x - s_j)] & \text{if } x \in U. \end{cases}$$

Observe

$$\bar{f} \in C^\infty(U)$$

and

$$D\bar{f}(x) = \sum_{x_j \in S_x} \{[f(s_j) + d(s_j) \cdot (x - s_j)]Dv_j(x) + v_j(x)d(s_j)\} \qquad (x \in U)$$

5. Claim #2: $D\bar{f}(a) = d(a)$ for all $a \in C$.

Proof of Claim #2: Fix $a \in C$ and let $K \equiv C \cap B(a, 1)$; K is compact. Define

$$\begin{aligned} \varphi(\delta) \equiv\ & \sup\{|R(x,y)| \mid x, y \in K, 0 < |x - y| \leq \delta\} \\ & + \sup\{|d(x) - d(y)| \mid x, y \in K, |x - y| \leq \delta\}. \end{aligned}$$

Since $d : C \to \mathbb{R}^n$ is continuous and $(\star)$ holds,

$$\varphi(\delta) \to 0 \text{ as } \delta \to 0. \qquad (\star\star\star)$$

If $x \in C$ and $|x - a| \leq 1$, then

$$\begin{aligned} |\bar{f}(x) - \bar{f}(a) - d(a) \cdot (x - a)| &= |f(x) - f(a) - d(a) \cdot (x - a)| \\ &= |R(x,a)|\,|x - a| \\ &\leq \varphi(|x - a|)|x - a| \end{aligned}$$

and

$$|d(x) - d(a)| \leq \varphi(|x - a|).$$

Now suppose $x \in U$, $|x-a| \le 1/6$. We calculate

$$\begin{aligned}|\bar{f}(x) - \bar{f}(a) - d(a)\cdot(x-a)| &= |\bar{f}(x) - f(a) - d(a)\cdot(x-a)|\\ &\le \sum_{x_j \in S_x} |v_j(x)[f(s_j) - f(a) + d(s_j)\cdot(x-s_j) - d(a)\cdot(x-a)]|\\ &\le \sum_{x_j \in S_x} v_j(x)|f(s_j) - f(a) + d(s_j)\cdot(a-s_j)|\\ &\quad + \sum_{x_j \in S_x} v_j(x)|(d(s_j)-d(a))\cdot(x-a)|.\end{aligned}$$

Now $|x-a| \le 1/6$ implies $r(x) \le (1/20)|x-a|$. Thus for $x_j \in S_x$,

$$\begin{aligned}|a-s_j| &\le |a-x_j| + |x_j - s_j|\\ &\le 2|a-x_j|\\ &\le 2(|x-a| + |x-x_j|)\\ &\le 2(|x-a| + 10(r(x)+r(x_j)))\\ &\le 2(|x-a| + 40r(x))\\ &\le 6|x-a|.\end{aligned}$$

Hence the calculation above and Claim #1 show

$$|\bar{f}(x) - \bar{f}(a) - d(a)\cdot(x-a)| < C\varphi(6|x-a|)|x-a|.$$

In view of $(\star\star\star)$, the calculations above imply that for each $a \in C$,

$$|\bar{f}(x) - \bar{f}(a) - d(a)\cdot(x-a)| = o(|x-a|) \text{ as } x \to a.$$

Thus $D\bar{f}(a)$ exists and equals $d(a)$.

6. *Claim #3*: $\bar{f} \in C^1(\mathbb{R}^n)$.

Proof of Claim #3: Fix $a \in C$, $x \in \mathbb{R}^n$, $|x-a| \le 1/6$. If $x \in C$, then

$$|D\bar{f}(x) - D\bar{f}(a)| = |d(x) - d(a)| \le \varphi(|x-a|).$$

If $x \in U$, choose $b \in C$ such that

$$|x-b| = \operatorname{dist}(x, C).$$

Thus

$$|D\bar{f}(x) - D\bar{f}(a)| = |D\bar{f}(x) - d(a)| \le |D\bar{f}(x) - d(b)| + |d(b) - d(a)|.$$

Since

$$|b - a| \le |b - x| + |x - a| \le 2|x - a|,$$

we have

$$|d(b) - d(a)| \le \varphi(2|x - a|).$$

We thus must estimate:

$$\begin{aligned}
\left|D\bar{f}(x) - d(b)\right| &= \Bigg|\sum_{x_j \in S_x} [f(s_j) + d(s_j)\cdot(x - s_j)]Dv_j(x) \\
&\qquad + v_j(x)[d(s_j) - d(b)]\Bigg| \\
&\le \left|\sum_{x_j \in S_x} [-f(b) + f(s_j) + d(s_j)\cdot(b - s_j)]Dv_j(x)\right| \\
&\qquad + \left|\sum_{x_j \in S_x} [(d(s_j) - d(b))\cdot(x - b)]Dv_j(x)\right| \\
&\qquad + \left|\sum_{x_j \in S_x} v_j(x)[d(s_j) - d(b)]\right| \\
&\le \frac{C}{r(x)} \sum_{x_j \in S_x} \varphi(|b - s_j|)|b - s_j| \\
&\qquad + \frac{C}{r(x)} \sum_{x_j \in S_x} \varphi(|b - s_j|)|x - b| \\
&\qquad + \sum_{x_j \in S_x} \varphi(|b - s_j|). \qquad (\star\star\star\star)
\end{aligned}$$

Now

$$|x - b| \le |x - a| \le \frac{1}{6}\,.$$

and therefore

$$r(x) = \frac{1}{20}|x - b| \le \frac{1}{120}\,.$$

If $x_j \in S_x$,

$$r(x_j) \le 3r(x) \le \frac{1}{40} < \frac{1}{20}\,.$$

Hence

$$r(x_j) = \frac{1}{20}|x_j - s_j| \qquad (x_j \in S_x).$$

Accordingly, if $x_j \in S_x$,

$$\begin{aligned}|b - s_j| &\le |b - x| + |x - x_j| + |x_j - s_j| \\ &\le 20r(x) + 10(r(x) + r(x_j)) + 20r(x_j) \\ &\le 120r(x) = 6|x - b| \le 6|x - a|.\end{aligned}$$

Consequently $(\star\star\star\star)$ implies

$$|D\bar{f}(x) - d(b)| \le C\varphi(6|x - a|).$$

This estimate and the calculations before show

$$|D\bar{f}(x) - D\bar{f}(a)| \le C\varphi(6|x - a|). \quad \blacksquare$$

6.6 Approximation by C^1 functions

We now make use of Whitney's Extension Theorem to show that if f is a Lipschitz, BV, or Sobolev function, then f actually equals a C^1 function $\bar{f}$, except on a small set. In addition, $Df = D\bar{f}$, except on a small set.

6.6.1 Approximation of Lipschitz functions

THEOREM 1
Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous. Then for each $\epsilon > 0$, there exists a C^1 function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n\{x \mid \bar{f}(x) \ne f(x) \text{ or } D\bar{f}(x) \ne Df(x)\} \le \epsilon.$$

In addition,

$$\sup_{\mathbb{R}^n} |D\bar{f}| \le C\,\mathrm{Lip}(f)$$

for some constant C depending only on n.

PROOF By Rademacher's Theorem, f is differentiable on a set $A \subset \mathbb{R}^n$, with $\mathcal{L}^n(\mathbb{R}^n - A) = 0$. Using Lusin's Theorem, we see there exists a closed set $B \subset A$ such that $Df\,|_B$ is continuous and $\mathcal{L}^n(\mathbb{R}^n - B) < \epsilon/2$. Set

$$d(x) \equiv Df(x)$$

and

$$R(y,x) \equiv \frac{f(y) - f(x) - d(x) \cdot (y - x)}{|x - y|} \qquad (x \neq y).$$

Define also

$$\eta_k(x) \equiv \sup\{|R(y,x)| \mid y \in B,\ 0 < |x - y| \leq 1/k\}.$$

Then

$$\eta_k(x) \to 0 \text{ as } k \to \infty, \text{ for all } x \in B.$$

By Egoroff's Theorem, there exists a closed set $C \subset B$ such that

$$\eta_k \to 0 \qquad \text{uniformly on compact subsets of C,}$$

and

$$\mathcal{L}^n(B - C) \leq \frac{\epsilon}{2}.$$

This implies hypothesis ($\star$) of Whitney's Extension Theorem.

The stated estimate on $\sup_{\mathbb{R}^n} |D\bar{f}|$ follows from the construction of $\bar{f}$ in the proof in Section 6.5, since $\sup_C |d| \leq \text{Lip}\,(f)$ and thus

$$|R|, |\varphi| \leq C \text{ Lip } (f). \quad \blacksquare$$

6.6.2 Approximation of BV functions

THEOREM 2

Let $f \in BV(\mathbb{R}^n)$. Then for each $\epsilon > 0$, there exists a Lipschitz function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n\{x \mid \bar{f}(x) \neq f(x)\} \leq \epsilon.$$

PROOF

1. Define for $\lambda > 0$

$$R^\lambda \equiv \left\{x \in \mathbb{R}^n \mid \frac{||Df||(B(x,r))}{r^n} \leq \lambda \text{ for all } r > 0\right\}.$$

2. *Claim #1*: $\mathcal{L}^n(\mathbb{R}^n - R^\lambda) \leq \frac{\alpha(n)5^n}{\lambda}||Df||(\mathbb{R}^n)$.

Proof of Claim #1: According to Vitali's Covering Theorem, there exist disjoint balls $\{B(x_i, r_i)\}_{i=1}^\infty$ such that

$$\mathbb{R}^n - R^\lambda \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_i)$$

and

$$\frac{\|Df\|(B(x_i, r_i))}{r_i^n} > \lambda.$$

Thus

$$\mathcal{L}^n(\mathbb{R}^n - R^\lambda) \leq 5^n \alpha(n) \sum_{i=1}^{\infty} r_i^n \leq \frac{5^n \alpha(n)}{\lambda} \|Df\|(\mathbb{R}^n).$$

3. *Claim #2*: There exists a constant C, depending only on n, such that

$$|f(x) - f(y)| \leq C\lambda |x - y|$$

for $\mathcal{L}^n$ a.e. $x, y \in R^\lambda$.

Proof of Claim #2: Let $x \in R^\lambda$, $r > 0$. By Poincaré's inequality, Theorem 1(ii) in Section 5.6.1,

$$\fint_{B(x,r)} |f - (f)_{x,r}| \, dy \leq \frac{C\|Df\|(B(x,r))}{r^{n-1}} \leq C\lambda r.$$

Thus, in particular,

$$\begin{aligned} |(f)_{x,r/2^{k+1}} - (f)_{x,r/2^k}| &\leq \fint_{B(x,r/2^{k+1})} |f - (f)_{x,r/2^k}| \, dy \\ &\leq 2^n \fint_{B(x,r/2^k)} |f - (f)_{x,r/2^k}| \, dy \\ &\leq \frac{C\lambda r}{2^k} \, . \end{aligned}$$

Since

$$f(x) = \lim_{r \to 0} (f)_{x,r}$$

for $\mathcal{L}^n$ a.e. $x \in R^\lambda$,

$$|f(x) - (f)_{x,r}| \leq \sum_{k=1}^{\infty} |(f)_{x,r/2^{k+1}} - (f)_{x,r/2^k}| \leq C\lambda r.$$

Now for $x, y \in R^\lambda$, $x \neq y$, set $r = |x - y|$. Then

$$\begin{aligned} |(f)_{x,r} - (f)_{y,r}| &\leq \fint_{B(x,r)\cap B(y,r)} |(f)_{x,r} - f(z)| + |f(z) - (f)_{y,r}| \, dz \\ &\leq C \left(\fint_{B(x,r)} |f(z) - (f)_{x,r}| \, dz + \fint_{B(y,r)} |f(z) - (f)_{y,r}| \, dz \right) \\ &\leq C\lambda r. \end{aligned}$$

We combine the inequalities above to estimate

$$|f(x) - f(y)| \le C\lambda r = C\lambda|x - y|$$

for $\mathcal{L}^n$ a.e. $x, y \in R^\lambda$.

4. In view of Claim #2, there exists a Lipschitz mapping $\bar{f} : R^\lambda \to \mathbb{R}$ such that $\bar{f} = f$ $\mathcal{L}^n$ a.e. on R^λ. Now recall Theorem 1 in Section 3.1 and extend $\bar{f}$ to a Lipschitz mapping $\bar{f} : \mathbb{R}^n \to \mathbb{R}$. ∎

COROLLARY 1

Let $f \in BV(\mathbb{R}^n)$. Then for each $\epsilon > 0$ there exists a C^1-function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n\{x \mid f(x) \neq \bar{f}(x) \text{ or } Df(x) \neq D\bar{f}(x)\} \le \epsilon.$$

PROOF According to Theorems 1 and 2, there exists $\bar{f} \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n(\{\bar{f} \neq f\}) < \epsilon.$$

Furthermore,

$$D\bar{f}(x) = Df(x)$$

$\mathcal{L}^n$ a.e. on $\{f = \bar{f}\}$, according to Theorem 4 in Section 6.1. ∎

6.6.3 Approximation of Sobolev functions

THEOREM 3

Let $f \in W^{1,p}(\mathbb{R}^n)$ for some $1 \le p < \infty$. Then for each $\epsilon > 0$ there exists a Lipschitz function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n\{x \mid f(x) \neq \bar{f}(x)\} \le \epsilon$$

and

$$||f - \bar{f}||_{W^{1,p}(\mathbb{R}^n)} \le \epsilon.$$

PROOF

1. Write $g \equiv |f| + |Df|$, and define for $\lambda > 0$

$$R^\lambda \equiv \left\{x \in \mathbb{R}^n \mid ⨍_{B(x,r)} g\, dy \le \lambda \text{ for all } r > 0\right\}.$$

2. Claim #1: $\mathcal{L}^n(\mathbb{R}^n - R^\lambda) = o(1/\lambda^p)$ as $\lambda \to \infty$.

Proof of Claim #1: By Vitali's Covering Theorem, there exist disjoint balls $\{B(x_i, r_i)\}_{i=1}^{\infty}$ such that

$$\mathbb{R}^n - R^\lambda \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_i) \tag{$\star$}$$

and

$$⨍_{B(x_i,r_i)} g\, dy > \lambda \qquad (i = 1, \ldots).$$

Hence

$$\lambda \leq \frac{1}{\mathcal{L}^n(B(x_i,r_i))} \int_{B(x_i,r_i)\cap\{g>\frac{\lambda}{2}\}} g\, dy + \frac{1}{\mathcal{L}^n(B(x_i,r_i))} \int_{B(x_i,r_i)\cap\{g\leq\lambda/2\}} g\, dy$$

$$\leq \frac{1}{\mathcal{L}^n(B(x_i,r_i))} \int_{B(x_i,r_i)\cap\{g>\lambda/2\}} g\, dy + \frac{\lambda}{2}$$

and so

$$\alpha(n) r_i^n \leq \frac{2}{\lambda} \int_{B(x_i,r_i)\cap\{g>\lambda/2\}} g\, dy \qquad (i = 1, \ldots).$$

Using ($\star$) therefore, we see

$$\begin{aligned}
\mathcal{L}^n(\mathbb{R}^n - R^\lambda) &\leq 5^n \alpha(n) \sum_{i=1}^{\infty} r_i^n \\
&\leq \frac{2 \cdot 5^n}{\lambda} \int_{\{g>\lambda/2\}} g\, dy \\
&\leq \frac{2 \cdot 5^n}{\lambda} \left(\int_{\{g>\lambda/2\}} g^p\, dy \right)^{\frac{1}{p}} (\mathcal{L}^n(\{g > \lambda/2\}))^{1-\frac{1}{p}} \\
&\leq \frac{C}{\lambda^p} \int_{\{|f|+|Df|>\lambda/2\}} |Df|^p + |f|^p\, dy \\
&= o\left(\frac{1}{\lambda^p}\right) \quad \text{as } \lambda \to \infty.
\end{aligned}$$

3. *Claim #2*: There exists a constant C, depending only on n, such that

$$|f(x)| \leq \lambda, \qquad |f(x) - f(y)| \leq C\lambda|x - y|$$

for $\mathcal{L}^n$ a.e. $x, y \in R^\lambda$.

Proof of Claim #2: This is almost exactly like the proof of Claim #2 in the proof of Theorem 2.

4. In view of Claim #2 we may extend f using Theorem 1 in Section 3.1 to a Lipschitz mapping $\bar{f} : \mathbb{R}^n \to \mathbb{R}$, with

$$|\bar{f}| \leq \lambda, \qquad \text{Lip}(\bar{f}) \leq C\lambda, \qquad \bar{f} = f \ \ \mathcal{L}^n \text{ a.e. on } R^\lambda.$$

5. Claim #3: $||f - \bar{f}||_{W^{1,p}(\mathbb{R}^n)} = o(1)$ as $\lambda \to \infty$.

Proof of Claim #3: Since $f = \bar{f}$ on R^λ, we have

$$\begin{aligned}\int_{\mathbb{R}^n} |f - \bar{f}|^p \, dx &= \int_{\mathbb{R}^n - R^\lambda} |f - \bar{f}|^p \, dx \\ &\leq C \int_{\mathbb{R}^n - R^\lambda} |f|^p \, dx + C\lambda^p \mathcal{L}^n(\mathbb{R}^n - R^\lambda) \\ &= o(1) \text{ as } \lambda \to \infty,\end{aligned}$$

according to Claim #1. Similarly, $Df = D\bar{f}$ $\mathcal{L}^n$ a.e. on R^λ, and so

$$\begin{aligned}\int_{\mathbb{R}^n} |Df - D\bar{f}|^p \, dx &\leq C \int_{\mathbb{R}^n - R^\lambda} |Df|^p \, dx + C\lambda^p \mathcal{L}^n(\mathbb{R}^n - R^\lambda) \\ &= o(1) \text{ as } \lambda \to \infty. \quad \blacksquare\end{aligned}$$

COROLLARY 2

Let $f \in W^{1,p}(\mathbb{R}^n)$ for some $1 \leq p < \infty$. Then for each $\epsilon > 0$, there exists a C^1-function $\bar{f} : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n\{x \mid f(x) \neq \bar{f}(x) \ \ \textit{or} \ \ Df(x) \neq D\bar{f}(x)\} \leq \epsilon$$

and

$$||f - \bar{f}||_{W^{1,p}(\mathbb{R}^n)} \leq \epsilon.$$

PROOF The assertion follows from Theorems 1 and 3. $\blacksquare$

Bibliography

[FA] K. Falconer, *Fractal Geometry*. Wiley, New York, 1990.

[F] H. Federer, *Geometric Measure Theory*. Springer-Verlag, New York, 1969.

[F-Z] H. Federer and W. Ziemer, "The Lebesgue set of a function whose distribution derivatives are p-th power summable," *Indiana U. Math. J.*, **22**, (1972), 139-158.

[F-R] W. Fleming and R. Rishel, An integral formula for the total gradient variation, *Arch. Math.* **11**, 218–222 (1960).

[GA] F. R. Gantmacher, *The Theory of Matrices*, Vol. I. Chelsea Publishing Co., New York, 1977.

[G-T] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, (2nd edn.). Springer-Verlag, New York, 1983.

[G] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, Boston, 1984.

[H] R. Hardt, *An Introduction to Geometric Measure Theory*. Lecture notes, Melbourne University, 1979.

[L] F.-C. Liu, "A Lusin property of Sobolev functions," *Indiana U. Math. J.* **26**, 645–651 (1977).

[M] V. Maz'ja, *Sobolev Spaces*. Springer-Verlag, New York, 1985.

[MO] F. Morgan, *Geometric Measure Theory: A Beginners Guide*. Academic Press, Boston, 1988.

[MY] C. B. Morrey, Jr., *Multiple Integrals in the Calculus of Variations*. Springer-Verlag, New York, 1966.

[R] Ju. G. Reshetnjak, "Generalized derivatives and differentiability almost everywhere," *Math USSR-Sbornik* **4**, 293–302 (1968).

[S] L. Simon, *Lectures on Geometric Measure Theory*. Centre for Mathematical Analysis, Australian National University, 1984.

[ST] E. Stein, *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, Princeton, NJ, 1970.

[Z] W. Ziemer, *Weakly Differentiable Functions*. Springer-Verlag, New York, 1989.

Chapters 1 through 3 are based primarily on Federer [F], as well as Simon [S] and Hardt [H].

Chapter 1: Essentially this entire chapter closely follows [F]. In Section 1.1, Theorems 1 and 2 are [F, Sections 2.1.3, 2.1.5]. Lemma 1 is [F, Section 2.2.2] and Theorem 4 is [F, Section 2.2.5]. Theorems 5 and 6 are from [F, Section 2.3.2] and Theorem 7 is [F, Section 2.3.3]. See also [S, Chapter 1] and [H].

In Section 1.2, Theorem 1 is a simplified version of [F, Sections 3.1.13, 3.1.14]. Lusin's and Egoroff's theorems are in [F, Sections 2.3.5, 2.3.7]. The general theory of integration may be found in [F, Section 2.4]: see in particular [F, Sections 2.4.6, 2.4.7, 2.4.9] for Fatou's Lemma, the Monotone Convergence Theorem, and the Dominated Convergence Theorem. Our treatment of product measure and Fubini's Theorem is taken directly from [F, Section 2.6].

We relied heavily on [S] for Vitali's Covering Theorem and [H] for Besicovitch's Covering Theorem: see also [F, Sections 2.9.12, 2.9.13]. The differentiation thoery in Section 1.6 is based on [H], [S], and [F, Section 2.9]. Approximate limits and approximate continuity are in [F, Sections 2.9.12, 2.9.13]. We took the proof of the Riesz Representation Theorem from [S, Sections 4.1, 4.2] (cf. [F, Section 2.5]). A. Damlamian showed us the proof of Corollary 1 in Section 1.8. See [G, Appendix A] for Theorem 2 in Section 1.9.

Chapter 2: Again, our primary source is [F], especially [F, Section 2.10]. Steiner symmetrization may be found in [F, Sections 2.10.30, 2.10.31]. We closely follow [H] for the proof of the Isodiametric Inequality, but incorporated a simplification due to Tam (communicated to us by R. Hardt). The proof $\mathcal{H}^n = \mathcal{L}^n$ is from [H] and [S, Sections 2.3–2.6]. We used [S, Section 3] for the density theorems in Section 2.3. Falconer [FA] and Morgan [MO] provide nice introductions to Hausdorff measure.

Chapter 3: The primary reference is [F, Chapters 1 and 3]. Theorem 1 in Section 3.1 is [S, Section 5.1]. The proof of Rademacher's Theorem, which we took from [S, Section 5.2], is due to Morrey (cf. [MY, p. 65]). Corollary 1 in Section 3.1 is [F, Section 3.2.8].

The discussion of linear maps and Jacobians in Section 3.2 is strongly based on [H]. S. Antman helped us with the proof of the Polar Decomposition Theorem, and A. Damlamian provided the calculations for the Binet–Cauchy formula: see also Gantmacher [GA, pp. 9–12, 276–278].

The proof of the Area Formula in Section 3.3, originating with [F, Sections 3.2.2–3.2.5], follows Hardt's exposition in [H]. Our proof in Section 3.4 of the Coarea Formula also closely follows [H], and is in turn based on [F, Sections 3.2.8–3.2.13].

Chapter 4: Our main sources for Sobolev functions were Gilbarg and Trudinger [G-T, Chapter 7] and Federer and Ziemer [F-Z].

See [G-T, Sections 7.2 and 7.3] for mollification and local approximation by smooth functions. Theorem 2 in Section 4.2 is from [G-T, Section 7.6] and Theorem 3 is based on [G-T, proof of Theorem 7.25]. The product and chain rules are in [G-T, Section 7.4]. See also [G-T, Section 7.12] for extensions. Sobolev inequalities of various sorts are in [G-T, Section 7.7]. Lemma 1 in Section 4.5 is a variant of [G-T, Lemma 7.16]. Compactness assertions are in [G-T, Section 7.10]. We follow [F-Z] (cf. also Ziemer [Z] and Maz'ja [M]) in our treatment of capacity. Theorems 1–4 in Section 4.7 are from [F-Z], as are all the results in Section 4.8. Much more information is available in [Z] and [M].

Chapter 5: We principally used Giusti [G] and Federer [F, Section 4.5] for BV theory (cf. also [S, Section 6]). The Structure Theorem is stated, for instance, in [S, Section 6.1]. The Lower Semicontinuity Theorem in Section 5.2 is [G, Section 1.9], and the Local Approximation Theorem is [G, Theorem 1.17]. (This result is due to Anzellotti and Giaquinta). The compactness assertion in Theorem 4 of Section 5.2 follows [G, Theorem 1.19]. The discussion of traces in Section 5.3 follows [G, Chapter 2]. Our treatment of extensions in Section 5.4 is an elaboration of [G, Remark 2.13]. The Coarea Formula for BV functions, due to Fleming and Rishel [F-R], is proved as in [G, Theorem 1.23]. For the Isoperimetric Inequalities, consult [G, Theorem 1.28 and Corollary 1.29]. The Remark in Section 5.6 is related to [F, Section 4.5.9(18)]. Theorem 3 in Section 5.6 is due to Fleming; we followed [F-Z]. The results in Sections 5.7 and 5.8 on the reduced and measure-theoretic boundaries are from [G, Chapters 3 and 4]: these assertions were originally established by De Giorgi. [F, Section 4.5.9] presents a long list of properties of BV functions, from which we extracted the theory set forth in Section 5.9. Essential variation occurs in [F, Section 4.5.10] and the criterion for finite perimeter described in Section 5.11 is [F, Section 4.5.11].

Chapter 6: The principal sources for this chapter are Federer [F], Liu [L], Reshetnjak [R], and Stein [S]. Our treatment of L^p-differentiability utilizes ideas from [ST, Section 8.1]. Approximate differentiability is discussed in [F, Sections 3.1.2–3.1.5]. D. Adams showed us the proof of Theorem 1 in Section 6.2. We followed [R] for the proof of Aleksandrov's Theorem in Section 6.4, and we took Whitney's Extension Theorem from [F, Sections 3.1.13–3.1.14]. The approximation of Lipschitz functions by C^1 functions is from [S, Section 5.3]: see also [F, Section 3.1.15]. We used [L] for the approximation of Sobolev functions.

We refer the reader to the sources listed in the Bibliography for complete citations of original papers, historical comments, etc.

Notation

A Vector and set notation

$\mathbb{R}^n$	n-dimensional real Euclidean space
e_i	$(0,\dots,1,\dots,0)$, with 1 in the ith slot
$x = (x_1, x_2, \dots, x_n)$	a typical point in $\mathbb{R}^n$
$\lvert x\rvert$	$(x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}$
$x \cdot y$	$x_1y_1 + x_2y_2 + \cdots + x_ny_n$
$x^T \cdot A \cdot y$	bilinear form $\sum_{i,j=1}^n a_{ij}x_iy_j$, where $x, y \in \mathbb{R}^n$ and $A = ((a_{ij}))$ is an $n \times n$ matrix
$B(x,r)$	$\{y \in \mathbb{R}^n \mid \lvert x-y\rvert \leq r\}$ = closed ball with center x, radius r
$B(r)$	$B(0,r)$
$U(x,r)$	$\{y \in \mathbb{R}^n \mid \lvert x-y\rvert < r\}$ = open ball with center x, radius r
$C(x,r,h)$	$\{y \in \mathbb{R}^n \mid \lvert y' - x'\rvert < r, \lvert y_n - x_n\rvert < h\}$ = open cylinder with center x, radius r, height $2h$
$\alpha(s)$	$\dfrac{\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2}+1)}$ $(0 \leq s < \infty)$
$\alpha(n)$	volume of the unit ball in $\mathbb{R}^n$
$Q(x,r)$	$\{y \in \mathbb{R}^n \mid \lvert x_i - y_i\rvert < r,\ i = 1,\dots,n\}$ = open cube with center x, side length $2r$
U, V, W	open sets, usually in $\mathbb{R}^n$
$V \subset\subset U$	V is compactly contained in U; i.e., $\overline{V}$ is compact and $\overline{V} \subset U$
K	compact set, usually in $\mathbb{R}^n$

χ_E	indicator function of the set E
$\overline{E}$	closure of E
E^o	interior of E
$S_a(E)$	Steiner symmetrization of a set E; Section 2.3
∂E	topological boundary of E
$\partial^* E$	reduced boundary of E; Section 5.7.1
$\partial_* E$	measure theoretic boundary of E; Section 5.8
$\|\|\partial E\|\|$	perimeter measure of E; Section 5.1

B Functional notation

$⨍_E f\, d\mu$ or $(f)_E$	$\frac{1}{\mu(E)} \int_E f\, d\mu =$ average of f on E with respect to the measure μ
$(f)_{x,r}$	$⨍_{B(x,r)} f\, dx$
$\mathrm{spt}(f)$	support of f
f^+, f^-	$\max(f, 0)$, $\max(-f, 0)$
f^*	precise representative of f; Section 1.7.1
$f\,\lfloor_E$	f restricted to the set E
$\bar{f}$ or Ef	an extension of f; cf. Sections 1.2, 3.1.1, 4.4, 5.4, 6.5
Tf	trace of f; Sections 4.3, 5.3
Df	derivative of f
$[Df]$	(vector-valued) measure for gradient of $f \in BV$; Section 5.1
$[Df]_{\mathrm{ac}}$, $[Df]_{\mathrm{s}}$	absolutely continuous, singular parts of $[Df]$; Section 5.1
ap Df	approximate derivative; Section 6.1.3
Jf	Jacobian of f; Section 3.2.2
$\mathrm{Lip}(f)$	Lipschitz constant of f; Sections 2.4.1, 3.1.1
$D^2 f$	Hessian matrix of f

$[D^2 f]$	(matrix-valued) measure for Hessian of convex f; Section 6.3
$[D^2 f]_{ac}$, $[D^2 f]_s$	absolutely continuous, singular parts of $[D^2 f]$; Section 6.3
$G(f, A)$	graph of f over the set A; Section 2.4.2

C Function spaces

Let $U \subset \mathbb{R}^n$ be an open set.

$C(U)$	$\{f : U \to \mathbb{R} \mid f$ continuous $\}$
$C(\overline{U})$	$\{f \in C(U) \mid f$ uniformly continuous$\}$
$C^k(U)$	$\{f : U \to \mathbb{R} \mid f$ is k-times continuously differentiable$\}$
$C^k(\overline{U})$	$\{f \in C^k(U) \mid D^\alpha f$ is uniformly continuous on U for $\lvert\alpha\rvert \leq k\}$
$C_c(U)$, $C_c(\overline{U})$, etc.	functions in $C(U)$, $C(\overline{U})$, etc. with compact support
$C(U; \mathbb{R}^m)$, $C(\overline{U}; \mathbb{R}^m)$, etc.	functions $f : U \to \mathbb{R}^m$, $f = (f^1, f^2, \ldots, f^m)$, with $f^i \in C(U)$, $C(\overline{U})$, etc. for $i = 1, \ldots, m$
$L^p(U)$	$\{f : U \to \mathbb{R} \mid \left(\int_U \lvert f\rvert^p \, dx\right)^{\frac{1}{p}} < \infty$, f Lebesgue measurable $\}$ $(1 \leq p < \infty)$
$L^\infty(U)$	$\{f : U \to \mathbb{R} \mid \operatorname{ess\,sup}_U \lvert f\rvert < \infty$, f Lebesgue measurable $\}$
$L^p_{\text{loc}}(U)$	$\{f : U \to \mathbb{R} \mid f \in L^p(V)$ for each open $V \subset\subset U\}$
$L^p(U; \mu)$	$\{f : U \to \mathbb{R} \mid \left(\int_U \lvert f\rvert^p \, d\mu\right)^{\frac{1}{p}} < \infty$, f μ-measurable $\}$ $(1 \leq p < \infty)$
$L^\infty(U; \mu)$	$\{f : U \to \mathbb{R} \mid f$ is μ-measurable, μ- $\operatorname{ess\,sup}_U \lvert f\rvert < \infty\}$
$W^{1,p}(U)$	Sobolev space; Section 4.1

K^p	$\{f : \mathbb{R}^n \to \mathbb{R} \mid f \geq 0,\ f \in L^{p^*},\ Df \in L^p\}$; Section 4.7
$BV(U)$	space of functions of bounded variation; Section 5.1

D Measures and capacity

$\mathcal{L}^n$	n-dimensional Lebesgue measure
$\mathcal{H}^s_\delta$	approximate s-dimensional Hausdorff measure; Section 2.1
$\mathcal{H}^s$	s-dimensional Hausdorff measures; Section 2.1
$\mathcal{H}_{\text{dim}}$	Hausdorff dimension; Section 2.1
Cap_p	p-capacity; Section 4.7.1

E Other notation

$\mu \llcorner A$	μ restricted to the set A; Section 1.1.1
$\mu \llcorner f$	(signed) measure with density f with respect to μ; §1.3
$D_\mu \nu$	derivative of ν with respect to μ; Section 1.6.1
$\nu << \mu$	ν is absolutely continuous with respect to μ; §1.6.2
$\nu \perp \mu$	ν and μ are mutually singular: Section 1.6.2
$\text{ap} \lim\limits_{y \to x} f$	approximate limit: Section 1.7.2
$\text{ap} \limsup\limits_{y \to x} f$, $\text{ap} \liminf\limits_{y \to x} f$	approximate lim sup, approximate lim inf; Section 1.7.2
$\rightharpoonup$	weak convergence; Section 1.9
S	symmetric linear mapping; Section 3.2.1
O	orthogonal linear mapping; Section 3.2.1
L^*	adjoint of L; Section 3.2.1
$[\![L]\!]$	Jacobian of linear mapping L; Section 3.2.1
$\Lambda(m, n)$	$\{\lambda : \{1, \ldots, n\} \to \{1 \ldots, m\} \mid \lambda \text{ increasing}\}$; Section 3.2.1

P_λ	projection associated with $\lambda \in \Lambda(m,n)$; Section 3.2.1
η, η_ϵ	mollifiers; Section 4.2.1
p^*	$\frac{np}{n-p}$ = Sobolev conjugate of p; Section 4.5.1
H, H^+, H^-	hyperplane, half spaces; Section 5.7.2
μ, λ	approximate lim sup, lim inf for BV function; Section 5.9
J	set of "measure theoretic jumps" for BV function; Section 5.9
$\operatorname{ess} V_a^b f$	essential variation; Section 5.10

Index