BIRKHÄUSER

CORNERSTONES

GEOMETRIC INTEGRATION THEORY

Steven G. Krantz
Harold R. Parks

BIRKHÄUSER

Cornerstones

Steven G. Krantz
Harold R. Parks

Geometric Integration Theory

Birkhäuser
Boston • Basel • Berlin

Steven G. Krantz
American Institute of Mathematics
360 Portage Avenue
Palo Alto, CA 94306-2244
U.S.A.
skrantz@aimath.org

Harold R. Parks
Department of Mathematics
Oregon State University
Corvallis, OR 97331
U.S.A.
parks@math.orst.edu

ISBN: 978-0-8176-4676-9 e-ISBN: 978-0-8176-4679-0
DOI: 10.1007/978-0-8176-4679-0

Library of Congress Control Number: 2008922000

Mathematics Subject Classification (2000): Primary: 49-01; Secondary: 26-01, 26B20, 28-01, 28A05, 28A75, 28C05, 28C10, 49Q05, 49Q20, 58C35, 53C65, 58E12

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com

Dedicated to the memory of Hassler Whitney (1907–1989).

Contents

Preface

Geometric measure theory has roots going back to ancient Greek mathematics, for considerations of the isoperimetric problem (to find the planar domain of given perimeter having greatest area) led naturally to questions about spatial regions and boundaries.

In more modern times, the Plateau problem is considered to be the wellspring of questions in geometric measure theory. Named in honor of the nineteenth century Belgian physicist Joseph Plateau, who studied surface tension phenomena in general, and soap films and soap bubbles in particular, the question (in its original formulation) was to show that a fixed, simple, closed curve in three-space will bound a surface of the type of a disk and having minimal area. Further, one wishes to study uniqueness for this minimal surface, and also to determine its other properties.

Jesse Douglas solved the original Plateau problem by considering the minimal surface to be a harmonic mapping (which one sees by studying the Dirichlet integral). For this work he was awarded the Fields Medal in 1936.

Unfortunately, Douglas's methods do not adapt well to higher dimensions, so it is desirable to find other techniques with broader applicability. Enter the theory of currents. Currents are continuous linear functionals on spaces of differential forms. Brought to fruition by Federer and Fleming in the 1950s, currents turn out to be a natural language in which to formulate the sorts of extremal problems that arise in geometry. One can show that the natural differential operators in the subject are closed when acting on spaces of currents, and one can prove compactness and structure theorems for spaces of currents that satisfy certain natural bounds. These two facts are key to the study of generalized versions of the Plateau problem and related questions of geometric analysis. As a result, Federer and Fleming were able in 1960 to prove the existence of a solution to the general Plateau problem in all dimensions and codimensions.

Today, geometric measure theory, which is properly focused on the study of currents and their geometry, is a burgeoning field in its own right. Furthermore, the techniques of geometric measure theory are finding good use in complex geometry, in partial differential equations, and in many other parts of modern geometry. It is desirable to have a text that introduces the graduate student to key ideas in this subject.

The present book is such a text. Demanding minimal background—only basic courses in calculus and linear algebra and real variables and measure theory—this book treats all the key ideas in the subject. These include the deformation theorem, the area and coarea formulas, the compactness theorem, the slicing theorem, and applications to fundamental questions about minimal surfaces that span given boundaries. In an effort to keep things as fundamental and near-the-surface as possible, we eschew generality and concentrate on the most essential results. As part of our effort to keep the exposition self-contained and accessible, we have limited our treatment of the regularity theory to proving almost-everywhere regularity of mass-minimizing hypersurfaces. We provide a full proof of the Lipschitz space estimate for harmonic functions that underlies the regularity of mass-minimizing hypersurfaces.

The notation in this subject—which is copious and complex—has been carefully considered by these authors and we have made strenuous effort to keep it as streamlined as possible. This is virtually the only graduate-level text in geometric measure theory that has figures and fully develops the subject; we feel that these figures add to the clarity of the exposition.

It should also be stressed that this book provides considerable background to bring the student up to speed. This includes

- measure theory
- lower-dimensional measures and Carathéodory's construction
- Haar measure
- covering theorems and differentiation of measures
- Poincaré inequalities
- differential forms and Stokes's theorem
- a thorough introduction to distributions and currents

Some students will find that they can skip certain of the introductory material; but it is useful to have it all present to establish terminology and notation, as a resource, and for reference. We have also made a special effort to keep this book self-contained. We do not want the reader running off to other sources for key ideas; he or she should be able to read this book while sitting at home.

Geometric measure theory uses techniques from geometry, measure theory, analysis, and partial differential equations. This book showcases all these methodologies, and explains the ways in which they interact. The result is a rich symbiosis that is both rewarding and educational.

The subject of geometric measure theory deserves to be known to a broad audience, and we hope that the present text will facilitate the dissemination of the subject to a new generation of mathematicians. It has been our pleasure to record these topics in a definitive and accessible and, we hope, lively form. We hope that the reader will derive the same satisfaction in studying these ideas in the present text. Of course, we welcome comments and criticisms, so that the book may be kept lively and current and as accurate as possible.

We are happy to thank Randi D. Ruden and Hypatia S. R. Krantz for genealogical help and Susan Parks for continued strength. It is a particular pleasure to thank our teachers and mentors, Frederick J. Almgren and Herbert Federer, for their inspiration and for the model that they set. Geometric measure theory is a different subject because of their work.

American Institute of Mathematics, Palo Alto *Steven G. Krantz*

Oregon State University, Corvallis *Harold R. Parks*

Geometric Integration Theory

1

Basics

Our purpose in this chapter will be to establish notation and terminology. The reader should already be acquainted with most of the concepts discussed and thus might wish to skim the chapter or skip ahead, returning if clarification is needed.

1.1 Smooth Functions

The set of real numbers will be denoted by $\mathbb{R}$. In this book, we will be concerned with questions of geometric analysis in an N-dimensional Euclidean space. That is, we will work in the space $\mathbb{R}^N$ of ordered N-tuples of real numbers. The *inner product* $x \cdot y$ of two elements $x, y \in \mathbb{R}^N$ is defined by setting

$$x \cdot y = \sum_{i=1}^{N} x_i y_i ,$$

where

$$x = (x_1, x_2, \ldots, x_N) \text{ and } y = (y_1, y_2, \ldots, y_N) .$$

Of course, the inner product is a symmetric, bilinear, positive definite function on $\mathbb{R}^N \times \mathbb{R}^N$. The *norm* of the element $x \in \mathbb{R}^N$, denoted by $|x|$, is defined by setting

$$|x| = \sqrt{x \cdot x} , \tag{1.1}$$

as we may since $x \cdot x$ is always nonnegative. The *standard orthonormal basis elements* for $\mathbb{R}^N$ will be denoted by $\mathbf{e}_i$, $i = 1, 2, \ldots, N$. Specifically, $\mathbf{e}_i$ is the vector with N entries, all of which are 0's except the ith entry, which is 1. For computational purposes, elements of $\mathbb{R}^N$ should be considered column vectors. Column vectors can waste space on the page, and so we sometimes take the liberty of using row vector notation, as we did above.

The *open ball of radius* $r > 0$ *centered at* x will be denoted by $\mathbb{B}(x, r)$ and is defined by setting

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_1,

$$\mathbb{B}(x, r) = \{ y \in \mathbb{R}^N : |x - y| < r \} .$$

The *closed ball of radius* $r \geq 0$ *centered at* x will be denoted by $\overline{\mathbb{B}}(x, r)$ and is defined by setting

$$\overline{\mathbb{B}}(x, r) = \{ y \in \mathbb{R}^N : |x - y| \leq r \} .$$

The standard topology on the space $\mathbb{R}^N$ is defined by letting the *open sets* consist of all arbitrary unions of open balls. The *closed sets* are then defined to be the complements of the open sets. For any subset A of $\mathbb{R}^N$ (or of any topological space), there is a largest open set contained in A. That set, denoted by $\mathring{A}$, is called the *interior of* A. Similarly, A is contained in a smallest closed set containing A and that set, denoted by $\overline{A}$, is called the *closure of* A. The *topological boundary of* A, denoted by ∂A, is defined by setting

$$\partial A = \overline{A} \setminus \mathring{A} .$$

Remark 1.1.1.

(1) At this juncture, the only notion of boundary in sight is that of the topological boundary. Since later we shall be led to define another notion of boundary, we are taking care to emphasize that the present definition is the topological one. When it is clear from context that we are discussing the topological boundary, then we will refer simply to the "boundary of A."
(2) The notations $\mathring{A}$ and $\overline{A}$ for the interior and closure, respectively, of the set A are commonly used but are not universal. A variety of notations is used for the topological boundary of A, and ∂A is one of the more popular choices.

Let $U \subseteq \mathbb{R}^N$ be any open set. A function $f : U \to \mathbb{R}^M$ is said to be *continuously differentiable of order* k, or C^k, if f possesses all partial derivatives of order not exceeding k and all of those partial derivatives are continuous; we write $f \in C^k$ or $f \in C^k(U)$ if U is not clear from context. If the range of f is also not clear from context, then we write (for instance) $f \in C^k(U; \mathbb{R}^M)$. When $k = 1$, we simply say that f is *continuously differentiable*. The function f is said to be C^∞, or *infinitely differentiable*, provided that $f \in C^k$ for every positive k. The function f is said to be in C^ω, or *real analytic*, provided that it has a convergent power series expansion about each point of U. We direct the reader to [KPk 02] for matters related to real analytic functions. We also extend the preceding notation by using $f \in C^0$ to indicate that f is continuous.

The order of differentiability of a function is referred to as its *smoothness*. By a *smooth function*, one typically means an $f \in C^\infty$, but sometimes one may mean an $f \in C^k$, where k is an integer as large as turns out to be needed.

The *support* of a continuous function $f : U \to \mathbb{R}^M$, denoted by $\operatorname{supp} f$, is the closure of the set of points where $f \neq 0$. We will use C_c^k to denote the C^k functions with compact support; here k can be a nonnegative integer or ∞.

Let $\mathbb{Z}$ denote the integers, $\mathbb{Z}^+$ the nonnegative integers, and $\mathbb{N}$ the positive integers. A *multi-index* α is an element of $(\mathbb{Z}^+)^N$, the Cartesian product of N copies of $\mathbb{Z}^+$. If $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)$ is a multi-index and $x = (x_1, x_2, \ldots, x_N)$ is a point in $\mathbb{R}^N$, then we introduce the following standard notation:

$$x^\alpha \equiv (x_1)^{\alpha_1}(x_2)^{\alpha_2}\cdots(x_N)^{\alpha_N}\,,$$

$$|\alpha| \equiv \alpha_1+\alpha_2+\cdots+\alpha_N\,,$$

$$\frac{\partial^{|\alpha|}}{\partial x^\alpha} \equiv \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}}\,,$$

$$\alpha! \equiv (\alpha_1!)(\alpha_2!)\cdots(\alpha_N!)\,.$$

With this notation, a function f on U is C^k if $(\partial^{|\alpha|}/\partial x^\alpha)f$ exists and is continuous for all multi-indices α with $|\alpha|\le k$.

We will sometimes find it convenient to use the alternative notations

$$D_{x_i}f=\frac{\partial f}{\partial x_i}\quad\text{and}\quad D_{x_i x_j}f=\frac{\partial^2 f}{\partial x_i\partial x_j}$$

for the partial derivatives of the function f (which may be a real-valued or vector-valued function).

Definition 1.1.2. If f is defined in a neighborhood of $p\in\mathbb{R}^N$, and if f takes values in $\mathbb{R}^M$, then we say that f is *differentiable* at p when there exists a linear function $Df(p):\mathbb{R}^N\to\mathbb{R}^M$ such that

$$\lim_{x\to p}\frac{|f(x)-f(p)-\langle Df(p),\ x-p\rangle|}{|x-p|}=0\,. \tag{1.2}$$

In case f is differentiable at p, we call $Df(p)$ the *differential* of f at p.

Advanced calculus tells us that if f is differentiable as in Definition 1.1.2, then the first partial derivatives of f exist and that we can evaluate the differential applied to the vector v using the equation

$$\langle Df(p),\ v\rangle=\sum_{i=1}^N v_i\,\frac{\partial f}{\partial x_i}(p)=\sum_{i=1}^N(\mathbf{e}_i\cdot v)\,\frac{\partial f}{\partial x_i}(p)\,, \tag{1.3}$$

where $v=\sum_{i=1}^n v_i\mathbf{e}_i$. The *Jacobian matrix*[1] of f at p is denoted by Jac f and is defined by

$$\operatorname{Jac} f \equiv \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(p) & \dfrac{\partial f_1}{\partial x_2}(p) & \cdots & \dfrac{\partial f_1}{\partial x_N}(p)\\ \dfrac{\partial f_2}{\partial x_1}(p) & \dfrac{\partial f_2}{\partial x_2}(p) & \cdots & \dfrac{\partial f_2}{\partial x_N}(p)\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial f_M}{\partial x_1}(p) & \dfrac{\partial f_M}{\partial x_2}(p) & \cdots & \dfrac{\partial f_M}{\partial x_N}(p) \end{pmatrix}.$$

[1] Carl Gustav Jacobi (1804–1851).

For $v \in \mathbb{R}^N$, we have

$$\langle Df(p),\ v\rangle = [\mathrm{Jac}\, f]\, v\,, \tag{1.4}$$

where on the right-hand side of (1.4) the vector v is represented as a column vector and Jac f operates on v by matrix multiplication. Equation (1.4) is simply another way of writing (1.3). We will sometimes find it convenient to use the notation

$$D_v f(p) = \langle Df(p),\ v\rangle\,.$$

We will denote the collection of all M-by-N matrices with real entries by

$$\mathcal{M}_{M,N}\,.$$

The *Hilbert–Schmidt norm*[2] on $\mathcal{M}_{M,N}$ is defined by setting

$$\Big|\,\big(a_{i,j}\big)\,\Big| = \left(\sum_{i=1}^{M}\sum_{j=1}^{N}(a_{i,j})^2\right)^{1/2}$$

for $\big(a_{i,j}\big) \in \mathcal{M}_{M,N}$. The standard topology on $\mathcal{M}_{M,N}$ is that induced by the Hilbert–Schmidt norm. Of course, the mapping

$$\big(a_{i,j}\big) \longmapsto \sum_{i=1}^{M}\sum_{j=1}^{N} a_{i,j}\,\mathbf{e}_{i+(j-1)M}$$

from $\mathcal{M}_{M,N}$ to $\mathbb{R}^{MN}$ is a homeomorphism.

The function sending a point to its differential, when the differential exists, takes its values in the space of linear transformations from $\mathbb{R}^N$ to $\mathbb{R}^M$, a space often denoted by $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$. The space $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$ can be identified with $\mathcal{M}_{M,N}$ by representing each linear transformation by an $M \times N$ matrix. The Jacobian matrix provides that representation for the differential of a function.

The standard topology on $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$ is that induced by the Hilbert–Schmidt norm on $\mathcal{M}_{M,N}$ and the identification of $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$ with $\mathcal{M}_{M,N}$. On a finite-dimensional vector space, all norms induce the same topology, so, in particular, the same topology is given by the *mapping norm* on $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$ defined by

$$\|L\| = \sup\{\,|L(v)|\ :\ v \in \mathbb{R}^N,\ |v| \le 1\,\}\,.$$

We see that $f : U \to \mathbb{R}^M$ is C^1 if and only if

$$p \longmapsto Df(p)$$

is a continuous mapping from U into $\mathrm{Hom}\,(\mathbb{R}^N, \mathbb{R}^M)$.

[2] David Hilbert (1862–1943), Erhard Schmidt (1876–1959).

Definition 1.1.3. If $f \in C^k(U, \mathbb{R}^M)$, $k = 1, 2, \ldots$, we define the kth *differential* of f at p, denoted by $D^k f(p)$, to be the k-linear $\mathbb{R}^M$-valued function given by

$$\langle D^k f(p),\ (v_1, v_2, \ldots, v_k)\rangle = \sum_{i_1, i_2, \ldots, i_k = 1}^{N} \prod_{j=1}^{k} (\mathbf{e}_{i_j} \cdot v_j)\, \frac{\partial^k}{\partial x_{i_1}\, \partial x_{i_2} \cdots \partial x_{i_k}} f(p)\,. \quad (1.5)$$

Note that in the case $k = 1$, equations (1.3) and (1.5) agree. Also note that the equality of mixed partial derivatives guarantees that $D^k f(p)$ is a symmetric function. The interested reader may consult [Fed 69, 1.9, 1.10, 3.1.11] to see the kth differential placed in the context of the symmetric algebra over a vector space.

Finally, note that in case $k > 1$, one can show inductively that (1.5) agrees with the value of the differential at p of the function

$$\langle D^{k-1} f(\cdot),\ (v_1, v_2, \ldots, v_{k-1})\rangle$$

applied to the vector v_k, that is,

$$\langle D^k f(p),\ (v_1, v_2, \ldots, v_k)\rangle = \langle D \, \langle D^{k-1} f(p),\ (v_1, v_2, \ldots, v_{k-1})\rangle,\ v_k\rangle$$

holds.

In case $M = 1$, one often identifies the differential of f with the *gradient vector* of f, denoted by grad f and defined by setting

$$\operatorname{grad} f = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}\, \mathbf{e}_i\,.$$

Similarly, the second differential of f can be identified with the *Hessian matrix*[3] of f, denoted by Hess (f) and defined by

$$\operatorname{Hess}(f) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\, \partial x_N} \\ \dfrac{\partial^2 f}{\partial x_2\, \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\, \partial x_N} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_N\, \partial x_1} & \dfrac{\partial^2 f}{\partial x_N\, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_N^2} \end{pmatrix}.$$

If f is suitably smooth, one has

$$v \cdot \operatorname{grad} f = \langle Df,\ v\rangle$$

and

$$v \cdot ([\operatorname{Hess}(f)]\, w) = \langle D^2 f,\ (v, w)\rangle\,,$$

for vectors v and w represented as columns and where $[\operatorname{Hess}(f)]\, w$ indicates matrix multiplication.

[3] Ludwig Otto Hesse (1811–1874).

1.2 Measures

Standard references for basic measure theory are [Fol 84], [Roy 88], and [Rud 87]. Since there are variations in terminology and notation among authors, we will briefly review measure theory. We shall *not* provide proofs of most statements, but instead refer the reader to [Fol 84], [Roy 88], and [Rud 87] for details.

Definition 1.2.1. Let X be a nonempty set.

(1) By a *measure* on X we mean a function μ defined on all subsets of X satisfying the conditions $\mu(\emptyset) = 0$, $A \subseteq B$ implies $\mu(A) \leq \mu(B)$, and

$$\mu\left(\bigcup_{A\in\mathcal{F}} A\right) \leq \sum_{A\in\mathcal{F}} \mu(A) \quad \text{if } \mathcal{F} \text{ is collection of subsets of } X \text{ with } \operatorname{card}(\mathcal{F}) \leq \aleph_0. \tag{1.6}$$

(2) If a set $A \subseteq X$ satisfies

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A) \quad \text{for all } E \subseteq X, \tag{1.7}$$

then we say that A is *μ-measurable*.

The condition (1.6) is called *countable subadditivity*. Since the empty union is the empty set and the empty sum is zero, countable subadditivity implies $\mu(\emptyset) = 0$. Nonetheless, it is worth emphasizing that $\mu(\emptyset) = 0$ must hold.

Proposition 1.2.2. *Let μ be a measure on the nonempty set X.*

(1) *If $\mu(A) = 0$, then A is μ-measurable.*
(2) *If A is μ-measurable and $B \subseteq X$, then*

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

Definition 1.2.3. Let X be a nonempty set. By a *σ-algebra* on X is meant a family $\mathcal{M}$ of subsets of X such that

(1) $\emptyset \in \mathcal{M}$, $X \in \mathcal{M}$,
(2) $\mathcal{M}$ is closed under countable unions,
(3) $\mathcal{M}$ is closed under countable intersections, and
(4) $\mathcal{M}$ is closed under taking complements in X.

Theorem 1.2.4. *If μ is a measure on the nonempty set X, then the family of μ-measurable sets forms a σ-algebra.*

Theorem 1.2.5. *Let μ be a measure on the nonempty set X.*

(1) *If $\mathcal{F}$ is an at most countable family of pairwise disjoint μ-measurable sets, then*

$$\mu\left(\bigcup_{A\in\mathcal{F}} A\right) = \sum_{A\in\mathcal{F}} \mu(A)\,.$$

(2) *If* $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ *is a nondecreasing family of* μ*-measurable sets, then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i\to\infty} \mu(A_i)\,.$$

(3) *If* $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$ *is a nonincreasing family of* μ*-measurable sets and* $\mu(B_1) < \infty$*, then*

$$\mu\left(\bigcap_{i=1}^{\infty} B_i\right) = \lim_{i\to\infty} \mu(B_i)\,.$$

Remark 1.2.6. The conclusion (1) of Theorem 1.2.5 is called *countable additivity*. Many authors prefer the term *outer measure* for the countably subadditive functions we have called measures. Those authors define a measure to be a countably additive function on a σ-algebra. But if $\mathcal{M}$ is a σ-algebra and

$$m : \mathcal{M} \to \{t : 0 \le t \le \infty\}$$

is a countably additive function, then one can define $\mu(A)$ for any $A \subseteq X$ by setting

$$\mu(A) = \inf\{\, m(E) : A \subseteq E \in \mathcal{M}\,\}\,.$$

With μ so defined, we see that $\mu(A) = m(A)$ holds whenever $A \in \mathcal{M}$ and that every set in $\mathcal{M}$ is μ-measurable. Thus it is no loss of generality to assume from the outset that a measure is defined on all subsets of X. It should be stressed that even though the measure is defined on all subsets of X, some subsets of X will *not* be μ-measurable.

The notion of a regular measure, defined next, gives additional useful structure.

Definition 1.2.7. A measure μ on a nonempty set X is *regular* if for each set $A \subseteq X$ there exists a μ-measurable set B with $A \subseteq B$ and $\mu(A) = \mu(B)$.

One consequence of the additional structure available when one is working with a regular measure is given in the next lemma. The lemma is easily proved using the analogous result for μ-measurable sets, i.e., Theorem 1.2.5(2).

Lemma 1.2.8. *Let* μ *be a regular measure on the nonempty set* X*. If a sequence of subsets* $\{A_j\}$ *of* X *satisfies* $A_1 \subseteq A_2 \subseteq \cdots$*, then*

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} \mu(A_j)\,.$$

Definition 1.2.9. If X is a topological space, then the *Borel sets*[4] are the elements of the smallest σ-algebra containing the open sets.

[4] Émile Borel (1871–1956).

For a measure on a topological space, it is evident that the measurability of all the open sets implies the measurability of all the Borel sets, but it is typical for the Borel sets to be a proper subfamily of the measurable sets. For instance, the sets in $\mathbb{R}^N$ known as Suslin sets[5] or (especially in the descriptive set theory literature) as analytic sets are μ-measurable for measures μ of interest in geometric analysis. Any continuous image of a Borel set is a Suslin set, so every Borel set is ipso facto a Suslin set. Suslin sets are discussed in Section 1.7.

For the study of geometric analysis, the measures of interest always satisfy the following condition of Borel regularity.

Definition 1.2.10. Let μ be a measure on the topological space X. We say that μ is *Borel regular* if every open set is μ-measurable and if for each $A \subseteq X$, there exists a Borel set $B \subseteq X$ with $A \subseteq B$ and $\mu(A) = \mu(B)$.

Often we will be working in the more restrictive class of Radon measures[6] defined next.

Definition 1.2.11. Suppose μ is a measure on a locally compact Hausdorff space[7] X. We say that μ is a *Radon measure* if the following conditions hold:

(1) Every compact set has finite μ measure.
(2) Every open set is μ-measurable, and if $V \subseteq X$ is open, then

$$\mu(V) = \sup\{\, \mu(K) : K \text{ is compact and } K \subseteq V \,\}.$$

(3) For every $A \subseteq X$,

$$\mu(A) = \inf\{\, \mu(V) : V \text{ is open and } A \subseteq V \,\}.$$

Definition 1.2.12. Let X be a metric space with metric ϱ.

(1) For a set $A \subseteq X$, we define the *diameter of* A by setting

$$\operatorname{diam} A = \sup\{\, \varrho(x, y) : x, y \in A \,\}.$$

(2) For sets $A, B \subseteq X$, we define the *distance between* A *and* B by setting

$$\operatorname{dist}(A, B) = \inf\{\, \varrho(a, b) : a \in A,\ b \in B \,\}.$$

If A is the singleton set $\{a_0\}$, then we will abuse the notation by writing $\operatorname{dist}(a_0, B)$ instead of $\operatorname{dist}(\{a_0\}, B)$.

When one is working in a metric space, a convenient tool for verifying the measurability of the open sets is often provided by Carathéordory's criterion,[8] which we now introduce.

---

[5] Mikhail Yakovlevich Suslin (1894–1919).
[6] Johann Radon (1887–1956).
[7] Felix Hausdorff (1869–1942).
[8] Constantin Carathéodory (1873–1950).

Theorem 1.2.13 (Carathéodory's criterion). *Suppose μ is a measure on the metric space X. All open subsets of X are μ-measurable if and only if*

$$\mu(A) + \mu(B) \leq \mu(A \cup B) \tag{1.8}$$

holds whenever $A, B \subseteq X$ with $0 < \operatorname{dist}(A, B)$.

Proof. First, suppose all open subsets of X are μ-measurable and let $A, B \subseteq X$ with $0 < \operatorname{dist}(A, B)$ be given. Setting $d = \operatorname{dist}(A, B)$, we can define the open set

$$V = \{\, x \in X : \operatorname{dist}(x, A) < d/2 \,\}.$$

Since V is open, thus μ-measurable, we have

$$\mu(A \cup B) = \mu[(A \cup B) \cap V] + \mu[(A \cup B) \setminus V] = \mu(A) + \mu(B),$$

so (1.8) holds.

Conversely, let $V \subseteq X$ be open and suppose (1.8) holds whenever $A, B \subseteq X$ with $0 < \operatorname{dist}(A, B)$. Let $E \subseteq X$ be an arbitrary set. Without loss of generality, we may suppose that $\mu(E) < \infty$ holds. Using (1.8) inductively, we see that

$$\mu(E) \geq \sum_{i=1}^{n} \mu(\{x \in E : 1/(2i+1) \leq \operatorname{dist}(x, V) < 1/(2i)\})$$

and likewise,

$$\mu(E) \geq \sum_{i=1}^{n} \mu(\{x \in E : 1/(2i+2) \leq \operatorname{dist}(x, V) < 1/(2i+1)\}).$$

Since n was arbitrary, we conclude that

$$2\mu(E) \geq \sum_{i=1}^{\infty} \mu(\{x \in E : 1/(i+1) \leq \operatorname{dist}(x, V) < 1/i\}),$$

so

$$\begin{aligned} 0 = \in_{n \to \infty} \sum_{i=n}^{\infty} \mu(\{x \in E : 1/(i+1) \leq \operatorname{dist}(x, V) < 1/i\}) \\ \geq \mu(\{x \in E : 0 < \operatorname{dist}(x, V) < 1/n\}). \end{aligned}$$

Again using (1.8), we see that

$$\begin{aligned} \mu(E) &\geq \mu(E \cap V) + \mu(\{x \in E : 1/n \leq \operatorname{dist}(x, V)\}) \\ &\geq \mu(E \cap V) + \mu(E \setminus V) - \mu(\{x \in E : 0 < \operatorname{dist}(x, V) < 1/n\}), \end{aligned}$$

and letting $n \to \infty$, we obtain

$$\mu(E) \geq \mu(E \cap V) + \mu(E \setminus V).$$

Since $E \subseteq X$ was arbitrary, V is μ-measurable. □

1.2.1 Lebesgue Measure

To close out this section, we define Lebesgue measure[9] on $\mathbb{R}$. Other measures will be defined in Chapter 2.

Definition 1.2.14. For $A \subseteq \mathbb{R}$, the (one-dimensional) *Lebesgue measure* of A is denoted by $\mathcal{L}^1(A)$ and is defined by setting $\mathcal{L}^1(A)$ equal to

$$\inf \left\{ \sum_{I \in \mathcal{I}} \text{length}(I) : \mathcal{I} \text{ is a family of bounded open intervals, } A \subseteq \bigcup_{I \in \mathcal{I}} I \right\}. \quad (1.9)$$

Here, of course, if $I = (a, b)$ is an open interval, then $\text{length}(I) = b - a$.

It is easy to see that $\mathcal{L}^1$ is a measure, and it is easy to apply Carathéodory's criterion (by dividing long intervals into short intervals) to see that all open sets in the reals are $\mathcal{L}^1$ measurable. The purpose of the Lebesgue measure is to extend the notion of length to more general sets. It may not be obvious that the result of the construction agrees with the ordinary notion of length, so we confirm that fact next.

Lemma 1.2.15. *If a bounded, closed interval* $[a, b]$ *is contained in the union of finitely many nonempty, bounded, open intervals,* $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$*, then it holds that*

$$b - a \leq \sum_{i=1}^{n} (b_i - a_i) . \quad (1.10)$$

Proof. Noting that the result is obvious when $n = 1$, we argue by induction on n by supposing that the result holds for all bounded, closed intervals and all n less than or equal to the natural number N.

Consider

$$[a, b] \subseteq \bigcup_{i=1}^{N+1} (a_i, b_i) .$$

At least one of the intervals contains a, so by renumbering the intervals if need be, we may suppose $a \in (a_{N+1}, b_{N+1})$. Also, we may suppose $b_{N+1} < b$, because $b \leq b_{N+1}$ would give us $b - a < b_{N+1} - a_{N+1}$.

We have

$$[b_{N+1}, b] \subseteq \bigcup_{i=1}^{N} (a_i, b_i) ,$$

and thus, by the induction hypothesis,

$$b - b_{N+1} \leq \sum_{i=1}^{N} (b_i - a_i) ,$$

so

[9] Henri Léon Lebesgue (1875–1941).

$$b-a \le (b_{N+1}-a_{N+1})+(b-b_{N+1}) \le (b_{N+1}-a_{N+1})+\sum_{i=1}^{N}(b_i-a_i) = \sum_{i=1}^{N+1}(b_i-a_i)\,,$$

as required. □

Corollary 1.2.16. *The Lebesgue measure of the closed, bounded interval* $[a, b]$ *equals* $b - a$.

Proof. Clearly, we have $\mathcal{L}^1([a, b]) \le b - a$. To obtain the reverse inequality, we observe that, if $[a, b]$ is covered by a countable family of open intervals, then by compactness, $[a, b]$ is covered by finitely many of the open intervals. It then follows from the lemma that the sum of the lengths of the covering intervals exceeds $b - a$. □

Lebesgue measure is the unique translation-invariant measure on $\mathbb{R}$ that assigns measure 1 to the unit interval. The next example shows us that not every set is $\mathcal{L}^1$-measurable.

Example 1.2.17. Let $\mathbb{Q}$ denote the rational numbers. Notice that for each $a \in \mathbb{R}$, the set X_a defined by

$$X_a = \{\, a + q : q \in \mathbb{Q} \,\}$$

intersects the unit interval [0, 1]. Of course, if $a_1 - a_2$ is a rational number, then $X_{a_1} = X_{a_2}$, but also the converse is true: if $X_{a_1} = X_{a_2}$, then $a_1 - a_2 \in \mathbb{Q}$.

By the axiom of choice, there exists a set C such that

$$C \cap [0, 1] \cap X_a$$

has exactly one element for every $a \in \mathbb{R}$. By the way C is defined, the sets $C - q = \{\, c - q : c \in C \,\}$, $q \in [0, 1] \cap \mathbb{Q}$, must be pairwise disjoint. Because $\mathcal{L}^1$ is translation-invariant, all the sets $C - q$ have $\mathcal{L}^1$ measure equal to $\mathcal{L}^1(C)$, and if one of those sets is $\mathcal{L}^1$-measurable, then all of them are.

Now, if $t \in [0, 1]$, then there is $c \in [0, 1] \cap X_t$, that is, $c = t + q$ with $q \in \mathbb{Q}$. Equivalently, we can write $q = c - t$, so we see that $-1 \le q \le 1$ and $t \in C - q$. Thus we have

$$[0, 1] \subseteq \bigcup_{q\in[-1,1]\cap\mathbb{Q}} (C - q) \subseteq [-1, 2] \tag{1.11}$$

and the sets in the union are all pairwise disjoint.

If C were $\mathcal{L}^1$-measurable, then the left-hand containment in (1.11) would tell us that $\mathcal{L}^1(C) > 0$, while the right-hand containment would tell us that $\mathcal{L}^1(C) = 0$. Thus we have a contradiction. We conclude that C is not $\mathcal{L}^1$-measurable. □

The construction in the Example 1.2.17 is widely known. Less well known is the general fact that if μ is a Borel regular measure on a complete, separable metric space such that there are sets with positive, finite measure and with the property that no point has positive measure, then there must exist a set that is not μ-measurable (see [Fed 69, 2.2.4]).

The construction of nonmeasurable sets requires the use of the axiom of choice. In fact, Robert Solovay has used the method of forcing (originally developed by Paul Cohen (1934–2007)) to construct a model of set theory in which the axiom of choice is not valid and in which every set of reals is Lebesgue measurable (see [Sov 70]).

1.3 Integration

The definition of the integral in use in the mid 1800s was that given by Augustin-Louis Cauchy (1789–1850). Cauchy's definition is applicable to continuous integrands, and easily extends to piecewise continuous integrands, but does not afford more generality. This lack of generality in the definition of the definite integral compelled Bernhard Riemann (1826–1866) to clarify the notion of an integrable function for his investigation of the representation of functions by trigonometric series.

Recall that Riemann's definition of the integral of a function $f : [a, b] \to \mathbb{R}$ is based on the idea of partitioning the *domain* of the function into sub-intervals. This approach is mandated by the absence of a measure of the size of general subsets of the domain. Measure theory takes away that limitation and allows the definition of the integral to proceed by partitioning the domain via the inverse images of intervals in the *range*. While this change of the partitioning may seem minor, the consequences are far-reaching and have provided a theory that continues to serve us well.

1.3.1 Measurable Functions

Definition 1.3.1. Let μ be a measure on the nonempty set X.

(1) The term *μ-almost* can serve as an adjective or adverb in the following ways:
   (a) Let $\mathcal{P}(x)$ be a statement or formula that contains a free variable $x \in X$. We say that $\mathcal{P}(x)$ holds for *μ-almost every $x \in X$* if

$$\mu\Big(\{\, x \in X : \mathcal{P}(x) \text{ is false} \,\}\Big) = 0\,.$$

   If X is understood from context, then we simply say that $\mathcal{P}(x)$ holds *μ-almost everywhere*.
   (b) Two sets $A, B \subseteq X$ are *μ-almost equal* if their symmetric difference has μ-measure zero, i.e., $\mu\Big[(A \setminus B) \cup (B \setminus A)\Big] = 0$.
   (c) Two functions f and g, each defined for μ-almost every $x \in X$, are said to be *μ-almost equal* if $f(x) = g(x)$ holds for μ-almost every $x \in X$.

(2) Let Y be a topological space. By a *μ-measurable, Y-valued function* we mean a Y-valued function f defined for μ-almost every $x \in X$ such that the inverse image of any open subset U of Y is a μ-measurable subset of X, that is,
   (a) $f : D \subseteq X \to Y$,
   (b) $\mu(X \setminus D) = 0$, and
   (c) $f^{-1}(U)$ is μ-measurable whenever $U \subseteq Y$ is open.

Remark 1.3.2.

(1) For the purposes of measure and integration, two functions that are μ-almost equal are equivalent. This defines an equivalence relation.
(2) It is no loss of generality to assume that a μ-measurable function is defined at every point of X. In fact, suppose f is a μ-measurable, Y-valued function with domain D and let y_0 be any element of Y. We can define the μ-measurable

function $\widetilde{f} : X \to Y$ by setting $\widetilde{f} = f$ on D and $\widetilde{f}(x) = y_0$, for all $x \in X \setminus D$. Then f and $\widetilde{f}$ are μ-almost equal and $\widetilde{f}$ is defined at every point of X.

Next we state two classical theorems concerning measurable functions due to Egorov[10] and Luzin.[11]

Theorem 1.3.3 (Egorov's theorem). *Let μ be a measure on X and let $f_1, f_2, \ldots$ be real-valued, μ-measurable functions. If $A \subseteq X$ with $\mu(A) < \infty$,*

$$\lim_{n\to\infty} f_n(x) = g(x) \ \text{ exists for } \mu\text{-almost every } x \in A,$$

and $\epsilon > 0$, then there exists a μ-measurable set B, with $\mu(A \setminus B) < \epsilon$, such that f_n converges uniformly to g on B.

Theorem 1.3.4 (Luzin's theorem). *Let X be a metric space and let μ be a Borel regular measure on X. If $f : X \to \mathbb{R}$ is μ-measurable, $A \subseteq X$ is μ-measurable with $\mu(A) < \infty$, and $\epsilon > 0$, then there exists a closed set $C \subseteq A$, with $\mu(A \setminus C) < \epsilon$, such that f is continuous on C.*

One reason for the usefulness of the notion of a μ-measurable function is that the set of μ-measurable functions is closed under operations of interest in analysis (including limiting operations). This usefulness is further enhanced by using the extended real numbers, which we define next.

Definition 1.3.5. Often we will allow a function to take the values $+\infty = \infty$ and $-\infty$. To accommodate this generality, we define the *extended real numbers*

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}.$$

The standard ordering on $\overline{\mathbb{R}}$ is defined by requiring

$$x \le y \ \text{ if and only if}$$
$$(x, y) \in \Big(\{-\infty\} \times \overline{\mathbb{R}}\Big) \bigcup \Big(\overline{\mathbb{R}} \times \{\infty\}\Big) \bigcup \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \le y\}.$$

The operation of addition is extended by requiring that it agree with values already defined for the real numbers, by demanding that the operation be commutative, and by assigning the values given in the following table:

$+$	$-\infty$	$x \in \mathbb{R}$	$+\infty$
$+\infty$	undefined	$+\infty$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	undefined

The operation of multiplication is extended by requiring that it agree with values already defined for the real numbers, by demanding that the operation be commutative, and by assigning the values given in the following table:

10 Dmitriĭ Fedorovich Egorov (1869–1931).

11 Nikolai Nikolaevich Luzin (Nicolas Lusin) (1883–1950).

x	$-\infty \le x < 0$	0	$0 < x \le +\infty$
$+\infty$	$-\infty$	undefined	$+\infty$
$-\infty$	$+\infty$	undefined	$-\infty$

The topology on $\overline{\mathbb{R}}$ has as a basis the finite open intervals and the intervals of the form $[-\infty, a)$ and $(a, \infty]$ for $a \in \mathbb{R}$.

The extension of each arithmetic operation given above is maximal subject to the requirement that the operation remain continuous. Nonetheless, when defining integrals, it is convenient to extend the above definitions by adopting the convention that

$$0 \cdot \infty = 0 \cdot (-\infty) = 0\,.$$

Theorem 1.3.6. *Let μ be a measure on the nonempty set X.*

(1) *If f and g are μ-measurable, extended-real-valued functions and if $f + g$ (respectively, fg) is defined μ-almost everywhere, then $f + g$ (respectively, fg) is μ-measurable.*
(2) *If f and g are μ-measurable, extended-real-valued functions, then the functions* $\max\{f, g\}$ *and* $\min\{f, g\}$ *are μ-measurable.*
(3) *If $f_1, f_2, \ldots$ are μ-measurable, extended-real-valued functions, then the functions* $\limsup_{n\to\infty} f_n$ *and* $\liminf_{n\to\infty} f_n$ *are μ-measurable.*

1.3.2 The Integral

Definition 1.3.7. For a function $f : X \to \overline{\mathbb{R}}$ we define the *positive part* of f to be the function $f^+ : X \to [0, \infty]$ defined by setting

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the *negative part* of f is denoted by f^- and is defined by setting

$$f^-(x) = \begin{cases} f(x) & \text{if } f(x) < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 1.3.8.

(1) The *characteristic function of $S \subseteq X$* is the function with domain X defined, for $x \in X$, by setting
$$\chi_S(x) = \begin{cases} 1 \text{ if } x \in S, \\ 0 \text{ if } x \notin S. \end{cases}$$
(2) By a *simple function* is meant a linear combination of characteristic functions of subsets of X; that is, f is a simple function if it can be written in the form
$$f = \sum_{i=1}^{n} a_i\, \chi_{A_i}\,, \tag{1.12}$$
where the numbers a_i can be real or complex, but only finite values are allowed (that is, $a_i \neq \pm\infty$).

The nonnegative, μ-measurable, simple functions are of particular interest for integration theory.

Lemma 1.3.9. *Let μ be a measure on the nonempty set X. If $f : X \to [0, \infty]$ is μ-measurable, then there exists a sequence of μ-measurable, simple functions $h_n : X \to [0, \infty]$, $n = 1, 2, \dots$, such that*

(1) $0 \le h_1 \le h_2 \le \cdots \le f$, *and*
(2) $\lim_{n\to\infty} h_n = f(x)$, *for all* $x \in X$.

Proof. We can set

$$h_n = n\,\chi_{B_n} + \sum_{i=1}^{n2^n-1} i \cdot 2^{-n}\,\chi_{A_i}\,,$$

where $B_n = f^{-1}\Big([n, \infty]\Big)$, and

$$A_i = f^{-1}\Big([i \cdot 2^{-n},\ (i+1) \cdot 2^{-n})\Big), \quad i = 1, 2, \dots, n2^n - 1\,. \qquad \square$$

Definition 1.3.10. Let μ be a measure on the nonempty set X. If $f : X \to \overline{\mathbb{R}}$ is μ-measurable, then the *integral of f with respect to μ* or, more simply, the *μ-integral of f* (or, more simply yet, the *integral of f* when the measure is clear from context) is denoted by

$$\int f\,d\mu = \int_X f(x)\,d\mu(x)$$

and is defined as follows:

(1) In case f is a nonnegative, simple function written as in (1.12) with each A_i μ-measurable, we set

$$\int f\,d\mu = \sum_{i=1}^{n} a_i\,\mu(A_i)\,. \tag{1.13}$$

(2) In case f is a nonnegative function, we set

$$\int f\,d\mu = \sup\left\{ \int h\,d\mu : 0 \le h \le f,\ h \text{ simple, } \mu\text{-measurable} \right\}. \tag{1.14}$$

(3) In case at least one of $\int f^+\,d\mu$ and $\int f^-\,d\mu$ is finite, so that

$$\int f^+\,d\mu - \int f^-\,d\mu$$

is defined, we set

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu\,. \tag{1.15}$$

(4) In case both $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are infinite, the quantity $\int f \, d\mu$ is undefined.

Definition 1.3.11.

(1) To integrate f over a subset A of X, we multiply f by the characteristic function of A, that is,

$$\int_A f \, d\mu = \int f \cdot \chi_A \, d\mu \,.$$

(2) The definition of $\int f \, d\mu$ extends to complex-valued, respectively $\mathbb{R}^N$-valued, functions by separating f into real and imaginary parts, respectively components, and combining the resulting real-valued integrals using linearity.

(3) If $\int |f| \, d\mu$ is finite, then we say that f is *μ-integrable* (or simply *integrable* if the measure μ is clear from context). In particular, f is μ-integrable if and only if $|f|$ is μ-integrable.

Remark 1.3.12.

(1) By a *Lebesgue integrable* function is meant an $\mathcal{L}^1$-integrable function in the terminology of Definition 1.3.11(3).

(2) The theories of Riemann integration and Lebesgue integration are connected by the following theorem:

A bounded, real-valued function on a closed interval is Riemann integrable if and only if the set of points at which the function is discontinuous has Lebesgue measure zero.

We will not prove this result. A proof can be found in [Fol 84, Theorem (2.28)].

(3) The reader should be aware that the terminology in [Fed 69] is different from that which we use: In [Fed 69] a function is said to be "μ integrable" if $\int f \, d\mu$ is defined, the values $+\infty$ and $-\infty$ being allowed, and "μ summable" if $\int |f| \, d\mu$ is finite.

The following basic facts hold for integration of nonnegative functions.

Theorem 1.3.13. *Let μ be a measure on the nonempty set X. Suppose $f, g : X \to [0, \infty]$ are μ-measurable.*

(1) *If $A \subseteq X$ is μ-measurable, and $f(x) = 0$ holds for μ-almost all $x \in A$, then*

$$\int_A f \, d\mu = 0 \,.$$

(2) *If $A \subseteq X$ is μ-measurable and $\mu(A) = 0$, then*

$$\int_A f \, d\mu = 0 \,.$$

(3) *If* $0 \le c < \infty$, *then*

$$\int (c \cdot f)\, d\mu = c \int f\, d\mu\,.$$

(4) *If* $f \le g$, *then*

$$\int f\, d\mu \le \int g\, d\mu\,.$$

(5) *If* $A \subseteq B \subseteq X$ *are* μ*-measurable, then*

$$\int_A f\, d\mu \le \int_B f\, d\mu\,.$$

Proof. Conclusions (1)–(4) are immediate from the definitions, and conclusion (5) follows from (4). □

Of course, it is essential that the equation $\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu$ hold. Unfortunately, this equation is not an immediate consequence of the definition. To prove it we need the next lemma, which is a weak form of Lebesgue's monotone convergence theorem.

Lemma 1.3.14. *Let* μ *be a measure on the nonempty set* X. *If* $f : X \to [0, \infty]$ *is* μ*-measurable and* $0 \le h_1 \le h_2 \le \cdots \le f$ *is a sequence of simple,* μ*-measurable functions with* $\lim\limits_{n\to\infty} h_n = f$, *then*

$$\lim_{n\to\infty} \int h_n\, d\mu = \int f\, d\mu\,.$$

Proof. The inequality $\lim_{n\to\infty} \int h_n\, d\mu \le \int f\, d\mu$ is immediate from the definition of the integral.

To obtain the reverse inequality, let ℓ be an arbitrary simple, μ-measurable function with $0 \le \ell \le f$ and write

$$\ell = \sum_{i=1}^{k} a_i\, \chi_{A_i}\,,$$

where each A_i is μ-measurable. Let $c \in (0, 1)$ also be arbitrary.

For each $m \in \mathbb{N}$, set

$$E_m = \{\, x : c \cdot \ell(x) \le h_m(x)\,\} \text{ and } \ell_m = c \cdot \ell \cdot \chi_{E_m}\,.$$

For $m \le n$, we have $\ell_m \le h_n$, so applying Theorem 1.3.13(4), we obtain

$$\int \ell_m\, d\mu \le \lim_{n\to\infty} \int h_n\, d\mu\,.$$

Finally, we note that for each $i = 1, 2, \ldots, k$, the sets $A_i \cap E_m$ increase to A_i as $m \to \infty$, so $\mu(A_i) = \lim_{m\to\infty} \mu(A_i \cap E_m)$ and thus

$$c \int \ell\, d\mu = \int c \cdot \ell\, d\mu = \lim_{m\to\infty} \int \ell_m\, d\mu \le \lim_{n\to\infty} \int h_n\, d\mu\,.$$

The result follows from the arbitrariness of ℓ and c. □

Theorem 1.3.15. *Let μ be a measure on the nonempty set X. If $f, g : X \to [0,\infty]$ are μ-measurable, then*

$$\int (f+g)\,d\mu = \int f\,d\mu + \int g\,d\mu\,.$$

Proof. The result clearly holds if f and g are simple functions, and the general case then follows from Lemmas 1.3.9 and 1.3.14. □

Corollary 1.3.16. *The μ-integrable functions form a vector space, and the μ-integral is a linear functional on the space of μ-integrable functions.*

The decisive results for integration theory are Fatou's lemma[12] and the monotone and dominated convergence theorems of Lebesgue (see any of [Fol 84], [Roy 88], and [Rud 87]). In the development outlined above, it is easiest first to prove Lebesgue's monotone convergence theorem, arguing as in the proof of Lemma 1.3.14. Then one uses the monotone convergence theorem to prove Fatou's lemma and the dominated convergence theorem. We state these results next.

Theorem 1.3.17. *Let μ be a measure on the nonempty set X.*

(1) **[Fatou's lemma]** *If $f_1, f_2, \ldots$ are nonnegative μ-measurable functions, then*

$$\liminf_{n\to\infty} \int_X f_n\,d\mu \geq \int_X \liminf_{n\to\infty} f_n\ d\mu\,.$$

(2) **[Lebesgue's monotone convergence theorem]** *If $f_1 \leq f_2 \leq \cdots$ are nonnegative μ-measurable functions, then*

$$\lim_{n\to\infty} \int_X f_n\,d\mu = \int_X \lim_{n\to\infty} f_n\ d\mu\,.$$

(3) **[Lebesgue's dominated convergence theorem]** *Let $f_1, f_2, \ldots$ be complex-valued μ-measurable functions that converge μ-almost everywhere to f. If there exists a nonnegative μ-measurable function g such that*

$$\sup_n |f_n(x)| \leq g(x) \ \ \textit{and} \ \ \int_X g\,d\mu < \infty\,,$$

then

$$\lim_{n\to\infty} \int_X |f_n - f|\,d\mu = 0 \ \ \textit{and} \ \ \lim_{n\to\infty} \int_X f_n\,d\mu = \int_X f\,d\mu\,.$$

One of the beauties of measure theory is that we deal in analysis almost exclusively with measurable functions and sets, and the ordinary operations of analysis would never cause us to leave the realm of measurable functions and sets. However, in geometric measure theory it is occasionally necessary to deal with functions that either are nonmeasurable or are not known a priori to be measurable. In such situations, it is convenient to have a notion of upper and lower integral.

[12] Pierre Joseph Louis Fatou (1878–1929).

Definition 1.3.18. Let μ be a measure on the nonempty set X and let $f : X \to [0, \infty]$ be defined μ-almost everywhere. We denote the *upper μ-integral of f* by

$$\overline{\int} f \, d\mu$$

and define it by setting

$$\overline{\int} f \, d\mu = \inf \left\{ \int \psi \, d\mu : 0 \leq f \leq \psi \text{ and } \psi \text{ is } \mu\text{-measurable} \right\} .$$

Similarly, the *lower μ-integral of f* is denoted by

$$\underline{\int} f \, d\mu$$

and defined by setting

$$\underline{\int} f \, d\mu = \sup \left\{ \int \phi \, d\mu : 0 \leq \phi \leq f \text{ and } \phi \text{ is } \mu\text{-measurable} \right\} .$$

Lemma 1.3.19. *If μ is a measure on the nonempty set X and $f, g : X \to [0, \infty]$ are defined μ-almost everywhere, then the following hold:*

(1) $\underline{\int} f \, d\mu \leq \overline{\int} f \, d\mu$,

(2) *if $f \leq g$, then* $\underline{\int} f \, d\mu \leq \underline{\int} g \, d\mu$ *and* $\overline{\int} f \, d\mu \leq \overline{\int} g \, d\mu$,

(3) *if f is μ-measurable, then* $\underline{\int} f \, d\mu = \int f \, d\mu = \overline{\int} f \, d\mu$,

(4) *if* $0 \leq c$, *then* $\underline{\int} cf \, d\mu = c \underline{\int} f \, d\mu$ *and* $\overline{\int} cf \, d\mu = c \overline{\int} f \, d\mu$,

(5) $\underline{\int} f \, d\mu + \underline{\int} g \, d\mu \leq \underline{\int} (f + g) \, d\mu$ *and* $\overline{\int} (f + g) \, d\mu \leq \overline{\int} f \, d\mu + \overline{\int} g \, d\mu$.

The lemma follows easily from the definitions.

Proposition 1.3.20. *Suppose $f : X \to [0, \infty]$ satisfies $\overline{\int} f \, d\mu < \infty$. For such a function,*

$$\underline{\int} f \, d\mu = \overline{\int} f \, d\mu$$

holds if and only if f is μ-measurable.

Proof. Suppose the upper and lower μ-integrals of f are equal. Choose sequences of μ-measurable functions $g_1 \leq g_2 \leq \cdots \leq f$ and $h_1 \geq h_2 \geq \cdots \geq f$ with

$$\lim_{n\to\infty} \int g_n \, d\mu = \underline{\int} f \, d\mu = \overline{\int} f \, d\mu = \lim_{n\to\infty} \int h_n \, d\mu \, .$$

Then $g = \lim_{n\to\infty} g_n$ and $h = \lim_{n\to\infty} h_n$ are μ-measurable with $g \le f \le h$. Since, by Lebesgue's dominated convergence theorem, the μ-integrals of g and h are equal, we see that g and h must be μ-almost equal to each other, and thus μ-almost equal to f. □

1.3.3 Lebesgue Spaces

Definition 1.3.21. Fix $1 \le p \le \infty$. Let μ be a measure on the nonempty set X. The *Lebesgue space* $L^p(\mu)$ (or simply L^p if the choice of the measure is clear from context) is the vector space of μ-measurable, complex-valued functions satisfying

$$\|f\|_p < \infty \, ,$$

where $\|f\|_p$ is defined by setting

$$\|f\|_p = \begin{cases} \left(\displaystyle\int |f|^p \, d\mu \right)^{1/p}, & \text{if } p < \infty, \\ \inf \Big\{ t : \mu\Big(X \cap \{ x : |f(x)| > t \} \Big) = 0 \Big\}, & \text{if } p = \infty. \end{cases}$$

The elements of L^p are called *L^p functions*. Of course, the L^1 functions are just the μ-integrable functions. The L^2 functions are also called *square integrable functions*, and, for $1 \le p < \infty$, the L^p functions are also called *p-integrable functions*.

Remark 1.3.22.

(1) A frequently used tool in analysis is Hölder's inequality[13]

$$\int fg \, d\mu \le \|f\|_p \, \|g\|_q \, ,$$

where f and g are μ-measurable, $1 < p < \infty$, and $1/p + 1/q = 1$. We note that Hölder's inequality is also valid when the integrals are replaced by upper integrals. The proof of this generalization makes use of Lemma 1.3.19(2)5).

(2) The functional $\|\cdot\|_p$ is called the *L^p-norm*. In the cases $p = 1$ and $p = \infty$, it is easy to verify that the L^p-norm is, in fact, a norm, but for the case $1 < p < \infty$, this fact is a consequence of Minkowski's inequality[14]

$$\|f + g\|_p \le \|f\|_p + \|g\|_p \, .$$

(3) Much of the importance of the Lebesgue spaces stems from the discovery that L^p, $1 \le p < \infty$, is a complete metric space. This result is sometimes (for instance in [Roy 88]) called the Riesz–Fischer theorem.[15]

---

[13] Otto Ludwig Hölder (1859–1937).

[14] Hermann Minkowski (1864–1909).

[15] Frigyes Riesz (1880–1956), Ernst Sigismund Fischer (1875–1954).

1.3.4 Product Measures and the Fubini–Tonelli Theorem

Definition 1.3.23. Let μ be a measure on the nonempty set X and let ν be a measure on the nonempty set Y. The *Cartesian product of the measures μ and ν* is denoted $\mu \times \nu$ and is defined by setting

$$(\mu \times \nu)(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) \cdot \nu(B_i) : E \subseteq \bigcup_{i=1}^{\infty} A_i \times B_i, \right. \\ A_i \subseteq X \text{ is } \mu\text{-measurable, for } i = 1, 2, \ldots, \\ \left. B_i \subseteq Y \text{ is } \nu\text{-measurable, for } i = 1, 2, \ldots \right\}. \tag{1.16}$$

It is immediately verified that $\mu \times \nu$ is a measure on $X \times Y$. Clearly the inequality

$$(\mu \times \nu)(A \times B) \leq \mu(A) \cdot \nu(B)$$

holds whenever $A \subseteq X$ is μ-measurable and $B \subseteq Y$ is ν-measurable. The product measure $\mu \times \nu$ is the largest measure satisfying that condition.

One of the main concerns in using product measures is justifying the interchange of the order of integration in a multiple integral. The next example illustrates a situation in which the order of integration in a double integral cannot be interchanged.

Example 1.3.24. The *counting measure on X* is defined by setting

$$\mu(E) = \begin{cases} \operatorname{card}(E) & \text{if } E \text{ is finite,} \\ \infty & \text{otherwise,} \end{cases}$$

for $E \subseteq X$. If ν is another measure on X for which $0 < \nu(X)$ and $\nu(\{x\}) = 0$ for each $x \in X$, and if $f : X \times X \to [0, \infty]$ is the characteristic function of the diagonal, that is,

$$f(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 = x_2, \\ 0 & \text{otherwise,} \end{cases}$$

then

$$\int \left(\int f(x_1, x_2)\, d\mu(x_1) \right) d\nu(x_2) = \int 1\, d\nu = \nu(X) > 0\,,$$

but

$$\int \left(\int f(x_1, x_2)\, d\nu(x_2) \right) d\mu(x_1) = \int 0\, d\mu = 0\,. \qquad \square$$

To avoid the phenomenon in the preceding example we introduce a definition.

Definition 1.3.25. Let μ be a measure on the nonempty set X. We say that μ is *σ-finite* if X can be written as a countable union of μ-measurable sets each having finite μ measure.

The main facts about product measures, which often *do* allow the interchange of the order of integration, are stated in the next theorem. We refer the reader to any of [Fol 84], [Roy 88], and [Rud 87].

Theorem 1.3.26. *Let μ be a σ-finite measure on the nonempty set X and let ν be a σ-finite measure on the nonempty set Y.*

(1) *If $A \subseteq X$ is μ-measurable and $B \subseteq Y$ is ν-measurable, then $A \times B$ is $(\mu \times \nu)$-measurable and*
$$(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B)\,.$$

(2) **(Tonelli's[16] theorem)** *If $f : X \times Y \to [0, \infty]$ is $(\mu \times \nu)$-measurable, then*
$$g(x) = \int f(x, y)\, d\nu(y) \tag{1.17}$$
defines a μ-measurable function on X,
$$h(y) = \int f(x, y)\, d\mu(x) \tag{1.18}$$
defines a ν-measurable function on Y, and
$$\begin{aligned} \int f\, d(\mu \times \nu) &= \int \left(\int f(x, y)\, d\mu(x) \right) d\nu(y) \\ &= \int \left(\int f(x, y)\, d\nu(y) \right) d\mu(x)\,. \end{aligned} \tag{1.19}$$

(3) **(Fubini's[17] theorem)** *If f is $(\mu \times \nu)$-integrable, then*
  (a) *$\phi(x) \equiv f(x, y)$ is μ-integrable for ν-almost every $y \in Y$,*
  (b) *$\psi(y) \equiv f(x, y)$ is ν-integrable for μ-almost every $x \in X$,*
  (c) *$g(x)$ defined by* (1.17) *is a μ-integrable function on X,*
  (d) *$h(y)$ defined by* (1.18) *is a ν-integrable function on Y, and*
  (e) *equation* (1.19) *holds.*

Definition 1.3.27. The *N-dimensional Lebesgue measure on* $\mathbb{R}^N$, denoted by $\mathcal{L}^N$, is defined inductively by setting $\mathcal{L}^N = \mathcal{L}^{N-1} \times \mathcal{L}^1$.

1.4 The Exterior Algebra

In an introductory vector calculus course, a vector is typically described as representing a direction and a magnitude, that is, an oriented line and a length. When later an oriented plane and an area in that plane are to be represented, a direction orthogonal to the plane and a length equal to the desired area are often used. This last device is viable only for $(N-1)$-dimensional oriented planes in N-dimensional space, because the complementary dimension must be 1. For the general case of an

[16] Leonida Tonelli (1885–1946).
[17] Guido Fubini (1879–1943).

oriented m-dimensional plane and an m-dimensional area in $\mathbb{R}^N$, some new idea must be invoked.

The straightforward way to represent an oriented m-dimensional plane in $\mathbb{R}^N$ is to specify an ordered m-tuple of independent vectors parallel to (contained in) the plane. To simultaneously represent an m-dimensional area in that plane, choose the vectors so that the m-dimensional area of the parallelepiped they determine equals that given m-dimensional area. Of course, a given oriented m-dimensional plane and m-dimensional area can be represented equally well by many different ordered m-tuples of vectors, and identifying any two such ordered m-tuples introduces an equivalence relation on the ordered m-tuples of vectors. To facilitate computation and understanding, the equivalence classes of ordered m-tuples are overlaid with a vector space structure. The result is the alternating algebra of m-vectors in $\mathbb{R}^N$. We now proceed to a formal definition.

Definition 1.4.1.

(1) Define an equivalence relation $\sim$ on

$$\left(\mathbb{R}^N\right)^m = \underbrace{\mathbb{R}^N \times \mathbb{R}^N \times \cdots \times \mathbb{R}^N}_{m \text{ factors}}$$

by requiring, for all $\alpha \in \mathbb{R}$ and $1 \le i < j \le m$,
(a)
$(u_1, \ldots, \alpha u_i, \ldots, u_j, \ldots, u_m) \sim (u_1, \ldots, u_i, \ldots, \alpha u_j, \ldots, u_m)$,
(b)
$(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_m) \sim (u_1, \ldots, u_i + \alpha u_j, \ldots, u_j, \ldots, u_m)$,
(c)
$(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_m) \sim (u_1, \ldots, -u_j, \ldots, u_i, \ldots, u_m)$,
and extending the resulting relation to be symmetric and transitive.

(2) The equivalence class of $(u_1, u_2, \ldots, \ldots, u_m)$ under $\sim$ is denoted by $u_1 \wedge u_2 \wedge \cdots \wedge u_m$. We call $u_1 \wedge u_2 \wedge \cdots \wedge u_m$ a *simple m-vector.*

(3) On the vector space of formal linear combinations of simple m-vectors, we define the equivalence relation $\approx$ by extending the relation defined by requiring
(a) $\alpha(u_1 \wedge u_2 \wedge \cdots \wedge u_m) \approx (\alpha u_1) \wedge u_2 \wedge \cdots \wedge u_m$,
(b) $(u_1 \wedge u_2 \wedge \cdots \wedge u_m) + (v_1 \wedge u_2 \wedge \cdots \wedge u_m) \approx (u_1 + v_1) \wedge u_2 \wedge \cdots \wedge u_m$.

(4) The equivalence classes of formal linear combinations of simple m-vectors under the relation $\approx$ are the *m-vectors in* $\mathbb{R}^N$. The vector space of m-vectors in $\mathbb{R}^N$ is denoted by $\bigwedge_m(\mathbb{R}^N)$.

(5) The *exterior algebra of* $\mathbb{R}^N$, denoted by $\bigwedge_*(\mathbb{R}^N)$, is the direct sum of the $\bigwedge_m(\mathbb{R}^N)$ together with the *exterior multiplication* defined by linearly extending the definition

$$(u_1 \wedge u_2 \wedge \cdots \wedge u_\ell) \wedge (v_1 \wedge v_2 \wedge \cdots \wedge v_m) = u_1 \wedge u_2 \wedge \cdots \wedge u_\ell \wedge v_1 \wedge v_2 \wedge \cdots \wedge v_m .$$

Remark 1.4.2.

(1) When $m = 1$, Definition 1.4.1(1) is vacuous, so $\bigwedge_1(\mathbb{R}^N)$ is isomorphic to, and will be identified with, $\mathbb{R}^N$. If the vectors $u_1, u_2, \ldots, u_m$ are linearly dependent, then $u_1 \wedge u_2 \wedge \cdots \wedge u_m$ is the additive identity in $\bigwedge_m(\mathbb{R}^N)$, so we write $u_1 \wedge u_2 \wedge \cdots \wedge u_m = 0$. Consequently, when $N < m$, $\bigwedge_m(\mathbb{R}^N)$ is the trivial vector space containing only 0.

(2) As an exercise, the reader should convince himself that $\mathbf{e}_1 \wedge \mathbf{e}_2 + \mathbf{e}_3 \wedge \mathbf{e}_4 \in \bigwedge_2(\mathbb{R}^4)$ is not a simple 2-vector.

For a nontrivial simple m-vector $u_1 \wedge u_2 \wedge \cdots \wedge u_m$ in $\mathbb{R}^N$, the *associated subspace* is that subspace spanned by the vectors $u_1, u_2, \ldots, u_m$. It is evident from Definition 1.4.1(1) that if $u_1 \wedge u_2 \wedge \cdots \wedge u_m = \pm v_1 \wedge v_2 \wedge \cdots \wedge v_m$, then their associated subspaces are equal. We assert that if $u_1 \wedge u_2 \wedge \cdots \wedge u_m = \pm v_1 \wedge v_2 \wedge \cdots \wedge v_m$, then also the m-dimensional area of the parallelepiped determined by $u_1, u_2, \ldots, u_m$ is equal to the m-dimensional area of the parallelepiped determined by $v_1, v_2, \ldots, v_m$. To see this last fact, we need the next proposition, which gives us a way to compute the m-dimensional areas in question. The proof is based on [Por 96].

Proposition 1.4.3. *Let $u_1, u_2, \ldots, u_m$ be vectors in $\mathbb{R}^N$. Then the parallelepiped determined by those vectors has m-dimensional area*

$$\sqrt{\det\left(U^{\mathrm{t}}\, U\right)}, \tag{1.20}$$

where U is the $N \times m$ matrix with $u_1, u_2, \ldots, u_m$ as its columns.

Proof. If the vectors $u_1, u_2, \ldots, u_m$ are pairwise orthogonal, then the result is immediate. Thus we will reduce the general case to this special case.

Notice that Cavalieri's principle[18] shows us that adding a multiple of u_j to another vector u_i, $i \neq j$, does not change the m-dimensional area of the parallelepiped determined by the vectors. But also notice that such an operation on the vectors u_i is equivalent to multiplying U on the right by an $m \times m$ triangular matrix with 1's on the diagonal. The Gram–Schmidt orthogonalization procedure[19] is effected by a sequence of operations of precisely this type. Thus we see that there is an upper triangular matrix A with 1's on the diagonal such that UA has orthogonal columns and the columns of UA determine a parallelepiped with the same m-dimensional area as the parallelepiped determined by $u_1, u_2, \ldots, u_m$. Since the columns of UA are orthogonal, we know that $\sqrt{\det\left((UA)^{\mathrm{t}}\,(UA)\right)}$ equals the m-dimensional area of the parallelepiped determined by its columns, and thus equals the m-dimensional area of the parallelepiped determined by $u_1, u_2, \ldots, u_m$. Finally, we compute

$$\begin{aligned}
\det\left((UA)^{\mathrm{t}}\,(UA)\right) &= \det\left(A^{\mathrm{t}}\, U^{\mathrm{t}}\, U\, A\right)\\
&= \det\left(A^{\mathrm{t}}\right)\,\det\left(U^{\mathrm{t}}\, U\right)\det(A)\\
&= \det\left(U^{\mathrm{t}}\, U\right).
\end{aligned}$$

□

[18] Bonaventura Francesco Cavalieri (1598–1647).

[19] Jørgen Pedersen Gram (1850–1916).

Corollary 1.4.4. *If* $u_1, u_2, \ldots, u_m$ *and* $v_1, v_2, \ldots, v_m$ *are vectors in* $\mathbb{R}^N$ *with*

$$u_1 \wedge u_2 \wedge \cdots \wedge u_m = \pm v_1 \wedge v_2 \wedge \cdots \wedge v_m \,,$$

then the m*-dimensional area of the parallelepiped determined by the vectors* $u_1, u_2, \ldots, u_m$ *equals the* m*-dimensional area of the parallelepiped determined by the vectors* $v_1, v_2, \ldots, v_m$.

Proof. We consider the m-tuples of vectors on the left-hand and right-hand sides of Definition 1.4.1(1a,b,c). Let U_l be the matrix whose columns are the vectors on the left-hand side and let U_r be the matrix whose columns are the vectors on the right-hand side. For (a), we have $U_r = U_l A$, where A is the $m \times m$ diagonal matrix with $1/\alpha$ in the ith column and α in the jth column. For (b), we have $U_r = U_l A$, where A is an $m \times m$ triangular matrix with 1's on the diagonal. For (c), we have $U_r = U_l A$, where A is an $m \times m$ permutation matrix with one of its 1's replaced by -1. In all three cases, $\det(A) = \pm 1$, and the result follows. □

For computational purposes, it is often convenient to use the basis

$$\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m} \,, \quad 1 \le i_1 < i_2 < \cdots < i_m \le N \,, \tag{1.21}$$

for $\bigwedge_m (\mathbb{R}^N)$. Specifying that the m-vectors in (1.21) are orthonormal induces the *standard inner product on* $\bigwedge_m (\mathbb{R}^N)$. The *exterior product* (sometimes called the *wedge product*)

$$\wedge : \textstyle\bigwedge_\ell (\mathbb{R}^N) \times \bigwedge_m (\mathbb{R}^N) \to \bigwedge_{\ell+m} (\mathbb{R}^N)$$

is an anticommutative, multilinear multiplication. Any linear $F : \mathbb{R}^N \to \mathbb{R}^P$ extends to a linear map $F_m : \bigwedge_m (\mathbb{R}^N) \to \bigwedge_m (\mathbb{R}^P)$ by defining

$$F_m(u_1 \wedge u_2 \wedge \cdots \wedge u_m) = F(u_1) \wedge F(u_2) \wedge \cdots \wedge F(u_m) \,.$$

1.5 The Generalized Pythagorean Theorem

The generalized Pythagorean theorem (Theorem 1.5.2 below) tells us that for a figure Σ lying in an m-dimensional affine subspace of $\mathbb{R}^N$, the square of the m-dimensional area of Σ equals the sum of the squares of the m-dimensional areas of the orthogonal projections of Σ onto all possible coordinate m-planes. For conceptual simplicity, we will restrict our attention to polyhedral figures Σ. We consider a few instances of this theorem:

- If $m = 1$ and Σ is a line segment, then the generalized Pythagorean theorem tells us that the square of the length of the segment is the sum of the squares of the lengths in each of the coordinate directions; that is, we recover the usual Pythagorean theorem.

- Suppose Σ is the parallelepiped generated by the m vectors $u_1, \ldots, u_m$ and U is the matrix whose columns are $u_1, \ldots, u_m$. Then the (signed) m-dimensional area of each projection of Σ onto a coordinate m-plane is given by an m-by-m minor determinant of U. Proposition 1.4.3 tells us that the m-dimensional area of Σ equals $\sqrt{\det(U^{\mathrm{t}}\, U)}$. Thus the generalized Pythagorean theorem implies—and, in fact, is equivalent to—the nontrivial fact that

$$\det(U^{\mathrm{t}}\, U) = \sum_{\lambda} [\det(U_\lambda)]^2 \tag{1.22}$$

  holds, where in (1.22) the summation extends over all $\lambda = \{i_1, \ldots, i_m\} \subseteq \{1, \ldots, N\}$ and where for each such λ, U_λ is the m-by-m submatrix whose rows are the rows numbered $i_1, \ldots, i_m$ in U.
- If Σ is an m-dimensional simplex in $\mathbb{R}^N$, then Σ automatically lies in an m-dimensional affine subspace of $\mathbb{R}^N$, and the generalized Pythagorean theorem applies to Σ. Figure 1.1 illustrates this situation when Σ is a triangle in $\mathbb{R}^3$. We have used A to denote the area of the triangle and A_{ij} to denote the area of the projection of the triangle onto the (x_i, x_j)-coordinate plane.

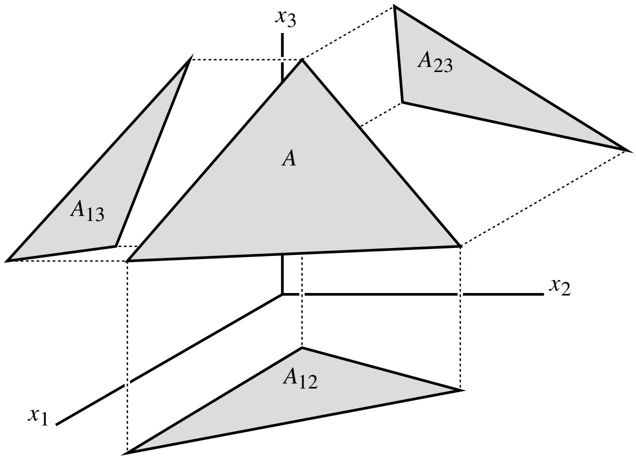

Fig. 1.1. $A^2 = A_{12}^2 + A_{13}^2 + A_{23}^2$.

In this section, we will give a geometrical proof of the generalized Pythagorean theorem. In particular, the proof will make no use of determinants. The main computation in the proof is made by applying the divergence theorem of advanced calculus to a constant vector field, while our other primary tool is the fact that the m-dimensional area of a figure is unchanged when the figure is mapped by an isometry.

Notation 1.5.1.

(1) Any m-dimensional polyhedral figure can be written as the union of m-dimensional simplices that intersect only in their boundaries. Thus, to prove the generalized Pythagorean theorem, it is sufficient to prove it when $\Sigma \subseteq \mathbb{R}^N$ is an m-simplex. Accordingly we will assume throughout the remainder of this section that Σ is the m-dimensional simplex determined by the $m + 1$ points $u_0, \ldots, u_m$.
(2) We will denote the m-dimensional area of Σ by A.
(3) If $\lambda \subseteq \{1, 2, \ldots, N\}$ and $\text{card}(\lambda) = K$, then $\Pi_\lambda : \mathbb{R}^N \to \mathbb{R}^K$ will be the orthogonal projection given by

$$\Pi_\lambda(x_1, x_2, \ldots, x_N) = (x_{i_1}, x_{i_2}, \ldots, x_{i_K}),$$

where $\lambda = \{i_1, i_2, \ldots, i_K\}$ and $i_1 < i_2 < \cdots < i_K$. We will need only the two cases $K = m$ and $K = 2$.
(4) If $\lambda \subseteq \{1, 2, \ldots, N\}$ and $\text{card}(\lambda) = m$, let A_λ denote the m-dimensional area of $\Pi_\lambda(\Sigma)$. We will sometimes abuse this notation (as we did in Figure 1.1) by writing $A_{i_1,i_2,\ldots,i_K}$ instead of the more pedantic $A_{\{i_1,i_2,\ldots,i_K\}}$.
(5) Since a set $\lambda \subseteq \{1, 2, \ldots, m + 1\}$ with $\text{card}(\lambda) = m$ is most easily described by the one element it omits, we will write

$$A_{\widehat{\imath}} = A_{1,\ldots,i-1,i+1,\ldots,m+1}\,.$$

Using the notation given above, we can state our result as follows:

Theorem 1.5.2 (Generalized Pythagorean theorem). *If Σ is an m-dimensional simplex in $\mathbb{R}^N$, then it holds that*

$$A^2 = \sum_{\substack{\lambda \subseteq \{1,\ldots,N\} \\ \text{card}\,(\lambda) = m}} A_\lambda^2\,. \tag{1.23}$$

Note that if $N = m$, the theorem is trivial. We first give a proof of the theorem in the case $N = m + 1$.

The Codimension-One Case, $N = m + 1$

Our proof for the case $N = m + 1$ will be based on an application of the divergence theorem.

Proposition 1.5.3. *Let Σ be an m-simplex in $\mathbb{R}^{m+1}$ with m-dimensional area A. Let $\mathbf{n}_0$ be a unit vector normal to Σ. Then*

$$A\,|\mathbf{n}_0 \cdot \mathbf{e}_i| = A_{\widehat{\imath}}$$

holds for $i = 1, \ldots, m + 1$.

Proof. We may assume for convenience that $i = m + 1$. If $\mathbf{n}_0 \cdot \mathbf{e}_{m+1} = 0$, then the result is trivial, so we also may assume that $\mathbf{n}_0 \cdot \mathbf{e}_{m+1} > 0$, i.e., $\mathbf{n}_0$ points "up."

By translating Σ if necessary, we may assume that all the coordinates of all the points in Σ are positive. Consider the closed polyhedral cylinder $\mathcal{C}$ made up of the line segments connecting each point of Σ with its projection on the $(x_1, \ldots, x_m)$-coordinate hyperplane; that is,

$$\mathcal{C} = \left\{ (1-t)\,x + t\,\Pi_{1,\ldots,m}(x) : x \in \Sigma,\ 0 \leq t \leq 1 \right\}$$

(Figure 1.2 illustrates $\mathcal{C}$ in the case $m = 2$). It will be convenient to call Σ the "top" of $\mathcal{C}$ and to call $B \equiv \Pi_{1,\ldots,m}(\Sigma)$ the "bottom" of $\mathcal{C}$.

Note that except on the top and bottom of $\mathcal{C}$, the outward unit normal to $\partial\mathcal{C}$ is orthogonal to $\mathbf{e}_{m+1}$. On the top of $\mathcal{C}$ the outward unit normal to $\mathcal{C}$ equals $\mathbf{n}_0$, and on the bottom of $\mathcal{C}$ the outward unit normal to $\mathcal{C}$ equals $-\mathbf{e}_{m+1}$ (see Figure 1.2).

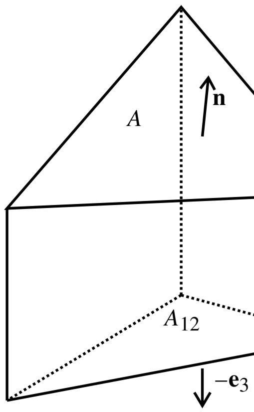

Fig. 1.2. Applying the divergence theorem.

The divergence theorem tells us that if $\mathbf{w}$ is a C^1 vector field on $\mathcal{C}$, then

$$\int_{\partial\mathcal{C}} \mathbf{w} \cdot \mathbf{n}\, d\sigma = \int_{\mathcal{C}} \operatorname{div} \mathbf{w}\, dV$$

holds, where $\mathbf{n}$ is the outward unit normal vector to $\partial\mathcal{C}$, $d\sigma$ is the element of m-dimensional area on $\partial\mathcal{C}$, and dV is the element of $(m+1)$-dimensional volume in $\mathcal{C}$.

Applying the divergence theorem to the constant vector field $\mathbf{w} \equiv \mathbf{e}_{m+1}$ on $\mathcal{C}$, we obtain

$$0 = \int_{\mathcal{C}} \operatorname{div} \mathbf{w}\, dV = \int_{\partial\mathcal{C}} \mathbf{w} \cdot \mathbf{n}\, d\sigma = A\,\mathbf{n}_0 \cdot \mathbf{e}_{m+1} - A_{\widehat{m+1}},$$

and the result follows. □

Corollary 1.5.4. *The generalized Pythagorean theorem holds when $N = m + 1$.*

Proof. Let $\mathbf{n}_0$ be a unit vector normal to $\Sigma \subseteq \mathbb{R}^{m+1}$. Since $\mathbf{n}_0$ is a unit vector, Proposition 1.5.3 gives us

$$A^2 = A^2 \sum_{i=1}^{m+1} (\mathbf{n}_0 \cdot \mathbf{e}_i)^2 = \sum_{i=1}^{m+1} A^2 (\mathbf{n}_0 \cdot \mathbf{e}_i)^2 = \sum_{i=1}^{m+1} A_{\hat{\imath}}^2. \qquad \square$$

The Higher Codimension Case, $N \geq m + 2$

Definition 1.5.5. By a *coordinate-plane rotation of* $\mathbb{R}^N$ we will mean a linear transformation that for some $i < j$, rotates the (x_i, x_j)-plane while leaving the remaining $(N - 2)$ coordinates unchanged. We will call x_i and x_j the *rotated coordinates*.

Our strategy for completing the proof of the generalized Pythagorean theorem is to show that the result holds for Σ if and only if it holds for the image of Σ under a coordinate-plane rotation. We then show that a sequence of coordinate-plane rotations of Σ will move Σ into an m-dimensional plane parallel to a coordinate m-plane—a situation in which the generalized Pythagorean theorem holds trivially.

Notation 1.5.6.

(1) Suppose $F : \mathbb{R}^N \to \mathbb{R}^N$ is a linear transformation. We set

$$\widetilde{\Sigma} = F(\Sigma).$$

For $\lambda \subseteq \{1, 2, \ldots, N\}$ with $\operatorname{card}(\lambda) = m$, $\widetilde{A}_\lambda$ will denote the m-dimensional area of $\Pi_\lambda(\widetilde{\Sigma})$. Similarly, when $N = m + 1$, we will use the notation $\widetilde{A}_{\hat{\imath}}$.

(2) For each positive integer K, we let $\mathbf{I}_{\mathbb{R}^K}$ be the identity map on $\mathbb{R}^K$.

Lemma 1.5.7. *Let* $F = \mathcal{R} \times \mathbf{I}_{\mathbb{R}^{N-2}}$, *where* $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ *is a rotation. Suppose* $\lambda \subseteq \{1, 2, \ldots, N\}$ *with* $\operatorname{card}(\lambda) = m$. *If*

$$\textit{either} \quad \{1, 2\} \cap \lambda = \emptyset \quad \textit{or} \quad \{1, 2\} \cap \lambda = \{1, 2\},$$

then $A_\lambda = \widetilde{A}_\lambda$.

Proof. When $\{1, 2\} \cap \lambda = \emptyset$ holds, we have

$$\Pi_\lambda(\Sigma) = \Pi_\lambda(\widetilde{\Sigma}),$$

so the result is trivial in this case.

Now suppose that $\{1, 2\} \subseteq \lambda$. Then we have

$$\Pi_\lambda \circ F = \Pi_\lambda \circ (\mathcal{R} \times \mathbf{I}_{\mathbb{R}^{N-2}}) = (\mathcal{R} \times \mathbf{I}_{\mathbb{R}^{m-2}}) \circ \Pi_\lambda,$$

and the result follows because $\mathcal{R} \times \mathbf{I}_{\mathbb{R}^{m-2}}$ is an isometry. $\square$

In Lemma 1.5.7, we considered projections Π_λ such that λ either included the indices of both rotated coordinates or omitted the indices of both rotated coordinates. In contrast, the m-dimensional area of the projection is *not preserved* when λ includes *exactly one* of the indices of the rotated coordinates. But we do have the next result.

Lemma 1.5.8. *Let* $F = \mathcal{R} \times \mathbf{I}_{\mathbb{R}^{N-2}}$*, where* $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ *is a rotation. If* $\lambda' \subseteq \{3, 4, \ldots, N\}$ *with* $\operatorname{card}(\lambda') = m - 1$*, then*

$$A^2_{\{1\}\cup\lambda'} + A^2_{\{2\}\cup\lambda'} = \widetilde{A}^2_{\{1\}\cup\lambda'} + \widetilde{A}^2_{\{2\}\cup\lambda'}\,. \tag{1.24}$$

Proof. For notational convenience, suppose that

$$\lambda' = \{3, 4, \ldots, m+1\}\,.$$

Each summand in (1.24) is unchanged if Σ is replaced by its projection into $\mathbb{R}^{m+1}$, so we may and shall assume that $N = m + 1$.

We have already shown that the generalized Pythagorean theorem holds when $N = m + 1$, so we can apply that theorem to $\Sigma \subseteq \mathbb{R}^{m+1}$ and to $\widetilde{\Sigma} \subseteq \mathbb{R}^{m+1}$. Using also the fact that $A = \widetilde{A}$ (which holds because F is an isometry), we obtain

$$\sum_{i=1}^{m+1} A^2_{\widehat{i}} = A^2 = \widetilde{A}^2 = \sum_{i=1}^{m+1} \widetilde{A}^2_{\widehat{i}}.$$

Observe that

$$\sum_{i=1}^{m+1} A^2_{\widehat{i}} = A^2_{\lambda'\cup\{1\}} + A^2_{\lambda'\cup\{2\}} + \sum_{\substack{\lambda''\subseteq\lambda' \\ \operatorname{card}(\lambda'')=m-2}} A^2_{\lambda''\cup\{1,2\}}$$

and, likewise, that

$$\sum_{i=1}^{m+1} \widetilde{A}^2_{\widehat{i}} = \widetilde{A}^2_{\lambda'\cup\{1\}} + \widetilde{A}^2_{\lambda'\cup\{2\}} + \sum_{\substack{\lambda''\subseteq\lambda' \\ \operatorname{card}(\lambda'')=m-2}} \widetilde{A}^2_{\lambda''\cup\{1,2\}}\,.$$

Lemma 1.5.7 tells us that for each $\lambda'' \subseteq \lambda'$ with $\operatorname{card}(\lambda'') = m - 2$,

$$A_{\lambda''\cup\{1,2\}} = \widetilde{A}_{\lambda''\cup\{1,2\}}$$

holds, so the result follows. □

In Lemmas 1.5.7 and 1.5.8, we considered a rotation $\mathcal{R}$ in the (x_1, x_2)-plane merely for convenience of notation. By relabeling coordinates, we see that the following result holds.

Proposition 1.5.9. *Suppose* $F : \mathbb{R}^N \to \mathbb{R}^N$ *rotates the* (x_i, x_j)*-plane while leaving all the other coordinates unchanged (here* $i < j$*).*

(1) *If* $\lambda \subseteq \{1, 2, \ldots, N\}$ *with* $\operatorname{card}(\lambda) = m$ *and if*

$$\textit{either}\quad \{i, j\} \cap \lambda = \emptyset \quad \textit{or} \quad \{i, j\} \cap \lambda = \{i, j\}\,,$$

then $A_\lambda = \widetilde{A}_\lambda$.

(2) *If* $\lambda' \subseteq \{1, 2, \dots, N\}$ *with* $\operatorname{card}(\lambda') = m - 1$ *and if*

$$\{i, j\} \cap \lambda' = \emptyset\,,$$

then

$$A^2_{\{i\}\cup\lambda'} + A^2_{\{j\}\cup\lambda'} = \widetilde{A}^2_{\{i\}\cup\lambda'} + \widetilde{A}^2_{\{j\}\cup\lambda'}\,.$$

In the next result, we show that the generalized Pythagorean theorem holds for Σ if and only if it holds for the image of Σ under a coordinate-plane rotation.

Corollary 1.5.10. *If* $F : \mathbb{R}^N \to \mathbb{R}^N$ *rotates the* (x_i, x_j)*-plane while leaving all the other coordinates unchanged (here* $i < j$*), then we have* $A = \widetilde{A}$ *and*

$$\sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m}} A^2_\lambda = \sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m}} \widetilde{A}^2_\lambda\,.$$

Consequently, the generalized Pythagorean theorem holds for Σ *if and only if it holds for* $\widetilde{\Sigma}$.

Proof. Observe that

$$\sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m}} A^2_\lambda = \sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m,\ \lambda\cap\{i,j\}=\emptyset}} A^2_\lambda + \sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m,\ \lambda\cap\{i,j\}=\{i,j\}}} A^2_\lambda$$
$$+ \sum_{\substack{\lambda'\subseteq\{1,\dots,N\}\\ \text{card }(\lambda')=m-1,\ \lambda'\cap\{i,j\}=\emptyset}} \left(A^2_{\lambda'\cup\{i\}} + A^2_{\lambda'\cup\{j\}}\right)$$

and, likewise, that

$$\sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m}} \widetilde{A}^2_\lambda = \sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m,\ \lambda\cap\{i,j\}=\emptyset}} \widetilde{A}^2_\lambda + \sum_{\substack{\lambda\subseteq\{1,\dots,N\}\\ \text{card }(\lambda)=m,\ \lambda\cap\{i,j\}=\{i,j\}}} \widetilde{A}^2_\lambda$$
$$+ \sum_{\substack{\lambda'\subseteq\{1,\dots,N\}\\ \text{card }(\lambda')=m-1,\ \lambda'\cap\{i,j\}=\emptyset}} \left(\widetilde{A}^2_{\lambda'\cup\{i\}} + A^2_{\lambda'\cup\{j\}}\right)\,.$$

The result now follows from Proposition 1.5.9. □

Proof of the Generalized Pythagorean Theorem. By translating Σ if necessary, we may suppose that u_0 coincides with the origin. Let us also introduce the notation

$$u_i = (u_{i,1},\ u_{i,2}, \dots,\ u_{i,N})\,.$$

By Corollary 1.5.10, it suffices to prove the generalized Pythagorean theorem for the image of Σ after a sequence of coordinate-plane rotations. In fact, we will show that there exists a sequence of coordinate-plane rotations such that the resulting image of Σ is contained in the $(x_1, \ldots, x_m)$-coordinate plane. Since the generalized Pythagorean theorem holds trivially for a simplex lying in an m-dimensional coordinate plane, it follows that the generalized Pythagorean theorem holds for the originally given Σ.

• **The first sequence of coordinate-plane rotations.** We begin with the rotation $\mathcal{R}$ of the (x_1, x_2)-plane that maps $\Pi_{\{1,2\}}(u_1) = (u_{1,1},\ u_{1,2})$ to $(t, 0)$, where $t = (u_{1,1}^2 + u_{1,2}^2)^{1/2}$. When the coordinate-plane rotation $\mathcal{R} \times \mathbf{I}_{\mathbb{R}^{N-2}}$ is applied to Σ and Σ is replaced by its image—without changing notation—we obtain

$$u_1 = (u_{1,1}, 0,\ u_{1,3}, \ldots,\ u_{1,N})\,.$$

The second coordinate-plane rotation will rotate the (x_1, x_3)-plane so that $\Pi_{\{1,3\}}(u_1) = (u_{1,1},\ u_{1,3})$ is mapped to $(t, 0)$, where $t = (u_{1,1}^2 + u_{1,3}^2)^{1/2}$. After again replacing Σ by its image—still without changing notation—we obtain

$$u_1 = (u_{1,1}, 0, 0,\ u_{1,4}, \ldots,\ u_{1,N})\,.$$

After a total of $N-1$ coordinate-plane rotations and replacements, we obtain

$$u_1 = (u_{1,1}, 0, 0, \ldots, 0)\,. \tag{1.25}$$

From now on, x_1 will *not* be one of the rotated coordinates in any of the coordinate-plane rotations we use. Consequently, (1.25) will continue to hold.

• **The $(i+1)$st sequence of coordinate-plane rotations.** Suppose that we have

$$\begin{aligned}
u_1 &= (u_{1,1}, 0, \ \ 0, \ldots, 0, 0, \ldots, 0)\,,\\
u_2 &= (u_{2,1}, u_{2,2}, 0, \ldots, 0, 0, \ldots, 0)\,,\\
&\ \ \vdots\\
u_i &= (u_{i,1}, u_{i,2}, u_{i,3}, \ldots, u_{i,i}, 0, \ldots, 0)\,.
\end{aligned} \tag{1.26}$$

In particular, observe that (1.26) implies that the points $u_1, u_2, \ldots, u_i$ all lie in the $(x_1, x_2, \ldots, x_i)$-coordinate plane.

Arguing inductively, we will show that we can obtain (1.26) with $i = m$. Note that when $i = 1$, (1.26) is the same as (1.25).

Our next coordinate-plane rotation will rotate the (x_{i+1}, x_{i+2})-plane so that $\Pi_{\{i+1,i+2\}}(u_{i+1}) = (u_{i+1,i+1},\ u_{i+1,i+2})$ is mapped to $(t, 0)$, where $t = (u_{i+1,i+1}^2 + u_{i+1,i+2}^2)^{1/2}$. Then we obtain

$$u_{i+1} = (u_{i+1,1},\ u_{i+1,2}, \ldots,\ u_{i+1,i+1}, 0,\ u_{i+1,i+3}, \ldots,\ u_{i+1,N})\,.$$

Continuing in that fashion, we see that after a total of $N-i-1$ coordinate-plane rotations, we obtain

$$u_{i+1} = (u_{i+1,1},\ u_{i+1,2}, \ldots,\ u_{i+1,i+1}, 0, 0, \ldots, 0)\,.$$

Since none of the coordinates $x_1, x_2, \ldots, x_i$ have been rotated coordinates for any of the coordinate-plane rotations we have used, the values of those coordinates will have remained unchanged. Thus we now have (1.26) with i replaced by $i+1$.

Arguing as above for $i = 1, 2, \ldots, m-1$, we see that—including the first sequence of coordinate-plane rotations—after a grand total of $(N-1) + \sum_{i=1}^{m-1}(N-i-1) = (m/2)\,(2N-m-1)$ coordinate-plane rotations, we obtain (1.26) with i replaced by m. Thus we see that the image of Σ lies in the $(x_1, \ldots, x_m)$-coordinate plane, as desired. □

Remark 1.5.11. In [Bar 96], the reader will find a proof of the usual Pythagorean theorem via dimensional analysis. E. Thomann has conjectured (private communication) that the generalized Pythagorean theorem also might be provable via a dimensional analysis argument.

1.6 The Hausdorff Distance and Steiner Symmetrization

Consider the collection $\mathcal{P}(\mathbb{R}^N)$ of all subsets of $\mathbb{R}^N$. It is often useful, especially in geometric applications, to have a metric on $\mathcal{P}(\mathbb{R}^N)$. In this section we address methods for achieving this end. In Definition 1.2.12, we defined $\mathrm{dist}(S, T)$ for subsets S, T of a metric space; unfortunately, this function need not satisfy the triangle inequality. Also, in practice, $\mathcal{P}(\mathbb{R}^N)$ (the entire power set of $\mathbb{R}^N$) is probably too large a collection of objects to have a reasonable and useful metric topology (see [Dug 66, Section IX.9] for several characterizations of metrizability). With these considerations in mind, we shall restrict attention to the collection of nonempty, *bounded* subsets of $\mathbb{R}^N$.

Definition 1.6.1. Let S and T be nonempty, bounded subsets of $\mathbb{R}^N$. We set

$$\mathrm{HD}\,(S, T) = \max\left\{\sup_{s\in S}\mathrm{dist}(s, T),\ \sup_{t\in T}\mathrm{dist}(S, t)\right\}. \tag{1.27}$$

This function is called the *Hausdorff distance*.

Notice that $\mathrm{HD}\,(S, T) = \mathrm{HD}\,(\overline{S}, T) = \mathrm{HD}\,(S, \overline{T}) = \mathrm{HD}\,(\overline{S}, \overline{T})$, so we further restrict our attention to the collection of nonempty sets that are both closed and bounded (i.e., compact) subsets of $\mathbb{R}^N$. For convenience, in this section, we will use $\mathcal{B}$ to denote the collection of nonempty, compact subsets of $\mathbb{R}^N$.

In Figure 1.3, if we let d denote the distance from a point on the left to the line segment on the right, then every point in the line segment is within distance $\sqrt{d^2 + (\epsilon/2)^2}$ of one of the points on the left—and that bound is sharp. Thus we see that $\mathrm{HD}\,(S, T) = \sqrt{d^2 + (\epsilon/2)^2}$.

Lemma 1.6.2. *Let $S, T \in \mathcal{B}$. Then there are points $s \in S$ and $t \in T$ such that* $\mathrm{HD}\,(S, T) = |s - t|$.

We leave the proof as an exercise for the reader (see Figure 1.4).

Proposition 1.6.3. *The function* HD *is a metric on $\mathcal{B}$.*

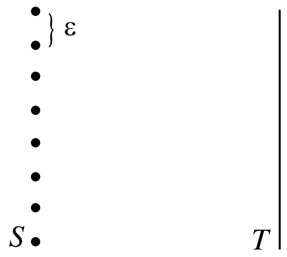

Fig. 1.3. The Hausdorff distance.

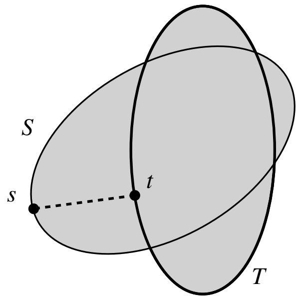

Fig. 1.4. Points that realize the Hausdorff distance.

Proof. Clearly HD ≥ 0, and if $S = T$, then $\text{HD}(S, T) = 0$.

Conversely, if $\text{HD}(S, T) = 0$ then let $s \in S$. By definition, there are points $t_j \in T$ such that $|s - t_j| \to 0$. Since T is compact, we may select a subsequence $\{t_{j_k}\}$ such that $t_{j_k} \to s$. Again, since T is compact, we then conclude that $s \in T$. Hence $S \subseteq T$. Similar reasoning shows that $T \subseteq S$. Hence $S = T$.

Finally, we come to the triangle inequality. Let $S, T, U \in \mathcal{B}$. Let $s \in S$, $t \in T$, $u \in U$. Then we have

$$
\begin{aligned}
|s - u| &\leq |s - t| + |t - u| \\
&\Downarrow \\
\text{dist}(S, u) &\leq |s - t| + |t - u| \\
&\Downarrow \\
\text{dist}(S, u) &\leq \text{dist}(S, t) + |t - u| \\
&\Downarrow \\
\text{dist}(S, u) &\leq \text{HD}(S, T) + |t - u| \\
&\Downarrow \\
\text{dist}(S, u) &\leq \text{HD}(S, T) + \text{dist}(T, u) \\
&\Downarrow \\
\text{dist}(S, u) &\leq \text{HD}(S, T) + \sup_{u \in U} \text{dist}(T, u)
\end{aligned}
$$

$$\Downarrow$$

$$\sup_{u \in U} \operatorname{dist}(S, u) \le \operatorname{HD}(S, T) + \sup_{u \in U} \operatorname{dist}(T, u).$$

By symmetry, we have

$$\sup_{s \in S} \operatorname{dist}(U, s) \le \operatorname{HD}(U, T) + \sup_{s \in S} \operatorname{dist}(T, s)$$

and thus

$$\begin{aligned}\max\{ \sup_{u \in U} \operatorname{dist}(S, u)\ ,\ \sup_{s \in S} \operatorname{dist}(U, s) \} \\ \le \max\Big\{\operatorname{HD}(S, T) + \sup_{u \in U} \operatorname{dist}(T, u)\ ,\ \operatorname{HD}(U, T) + \sup_{s \in S} \operatorname{dist}(T, s)\Big\}.\end{aligned}$$

We conclude that

$$\operatorname{HD}(U, S) \le \operatorname{HD}(U, T) + \operatorname{HD}(T, S). \qquad \square$$

There are fundamental questions concerning completeness, compactness, etc. that we need to ask about any metric space.

Theorem 1.6.4. *The metric space* $(\mathcal{B}, \operatorname{HD})$ *is complete.*

Proof. Let $\{S_j\}$ be a Cauchy sequence in the metric space $(\mathcal{B}, \operatorname{HD})$. We seek an element $S \in \mathcal{B}$ such that $S_j \to S$.

Elementary estimates, as in any metric space, show that the elements S_j are all contained in a common ball $B(0, R)$. We set S equal to

$$\bigcap_{j=1}^{\infty} \left(\overline{\bigcup_{\ell=j}^{\infty} S_\ell} \right).$$

Then S is nonempty, closed, and bounded, so it is an element of $\mathcal{B}$.

To see that $S_j \to S$, select $\epsilon > 0$. Choose J large enough so that if $j, k \ge J$ then $\operatorname{HD}(S_j, S_k) < \epsilon$. For $m > J$ set $T_m = \cup_{\ell=J}^m S_\ell$. Then it follows from the definition, and from Proposition 1.6.3, that $\operatorname{HD}(S_J, T_m) < \epsilon$ for every $m > J$. Therefore, with $U_p = \overline{\cup_{\ell=p}^{\infty} S_\ell}$ for every $p > J$, it follows that $\operatorname{HD}(S_J, U_p) \le \epsilon$.

We conclude that $\operatorname{HD}(S_J, \cap_{p=J+1}^K U_p) \le \epsilon$. Hence, by the continuity of the distance, $\operatorname{HD}(S_J, S) \le \epsilon$. That is what we wished to prove. $\square$

As a corollary of the proof of Theorem 1.6.4 we obtain the following:

Corollary 1.6.5. *Let* $\{S_j\}$ *be a sequence of elements of* $\mathcal{B}$*. Suppose that* $S_j \to S$ *in the Hausdorff metric. Then*

$$\mathcal{L}^n(S) \ge \limsup_{j \to \infty} \mathcal{L}^n(S_j).$$

The next theorem informs us of a seminal fact regarding the Hausdorff distance topology.

Theorem 1.6.6. *The set of nonempty compact subsets of* $\mathbb{R}^N$ *with the Hausdorff distance topology is boundedly compact, i.e., any bounded sequence has a subsequence that converges to a compact set.*

Proof. Let $A_1, A_2, \ldots$ be a bounded sequence in the Hausdorff distance. We may assume without loss of generality that each A_i is a subset of the closed unit N-cube, C_0. For each integer $k \geq 1$, subdivide the unit N-cube into 2^{kN} congruent subcubes of side length 2^{-k}; denote that collection of 2^{kN} subcubes by $\mathcal{S}_k$.

We will use an inductive construction and a diagonalization argument. Let $A_{0,i} = A_i$ for $i = 1, 2, \ldots$. For each $k \geq 1$, the sequence $A_{k,i}$, $i = 1, 2, \ldots$, will be a subsequence of the preceding sequence $A_{k-1,i}$, $i = 1, 2, \ldots$. Also, we will construct sets $C_0 \supseteq C_1 \supseteq \cdots$ inductively. Each C_k will be a union of a set of cubes in $\mathcal{S}_k$. The first set in this sequence is the unit cube C_0 itself. For each $k = 0, 1, \ldots$, the sequence $A_{k,i}$, $i = 1, 2, \ldots$, and the set C_k are to have the properties that

$$\begin{aligned} &D \cap A_{k,i} \neq \emptyset \text{ holds for } i = 1, 2, \ldots \\ &\qquad\qquad \text{whenever } D \in \mathcal{S}_k \text{ is one of the cubes forming } C_k, \end{aligned} \tag{1.28}$$

and

$$A_{k,i} \subseteq C_k \quad \text{holds for all sufficiently large } i. \tag{1.29}$$

It is clear that (1.28) and (1.29) are satisfied when $k = 0$.

Assume $A_{k-1,i}$, $i = 1, 2, \ldots$, and C_{k-1} have been defined so that

$$\begin{aligned} &D \cap A_{k-1,i} \neq \emptyset \text{ holds for } i = 1, 2, \ldots \\ &\qquad\qquad \text{whenever } D \in \mathcal{S}_{k-1} \text{ is one of the cubes forming } C_{k-1}, \end{aligned}$$

and

$$A_{k-1,i} \subseteq C_{k-1} \quad \text{holds for all sufficiently large } i.$$

We let $\mathcal{C}_k$ be the collection of cubes in $\mathcal{S}_k$ that are subsets of C_{k-1} (here we are effectively subdividing the cubes that form C_{k-1}). A subcollection, $\mathcal{C} \subseteq \mathcal{C}_k$, will be called *admissible* if there are infinitely many i for which

$$D \cap A_{k-1,i} \neq \emptyset \quad \text{holds for all } D \in \mathcal{C}. \tag{1.30}$$

Let C_k be the union of a maximal admissible collection of subcubes, which is immediately seen to exist because $\mathcal{C}_k$ is finite. Let $A_{k,1}, A_{k,2}, \ldots$ be the subsequence of $A_{k-1,1}, A_{k-1,2}, \ldots$ consisting of those $A_{k-1,i}$ for which (1.30) is true. Observe that $A_{k,i} \subseteq C_k$ holds for sufficiently large i; otherwise, there is another subcube that could be added to the maximal collection while maintaining admissibility.

We set

$$C = \bigcap_{k=0}^{\infty} C_k$$

and claim that C is the limit in the Hausdorff distance of $A_{k,k}$ as $k \to \infty$. Of course, C is nonempty by the finite intersection property. Let $\epsilon > 0$ be given. Clearly we can find an index k_0 such that

$$C_{k_0} \subseteq \{x : \operatorname{dist}(x, C) < \epsilon\}.$$

There is a number i_0 such that for $i \geq i_0$ we have

$$A_{k_0,i} \subseteq C_{k_0} \subseteq \{x : \operatorname{dist}(x, C) < \epsilon\}.$$

So, for $k \geq k_0 + i_0$, we know that

$$A_{k,k} \subseteq \{x : \operatorname{dist}(x, C) < \epsilon\}$$

holds. We let $k_1 \geq k_0 + i_0$ be such that

$$\sqrt{N}\, 2^{-k_1} < \epsilon.$$

Let $c \in C$ be arbitrary. Then $c \in C_{k_1}$, so there is some cube, D, of side length 2^{-k_1} containing c and for which

$$D \cap A_{k_1,i} \neq \emptyset$$

holds for all i. But then if $k \geq k_1$, we have $D \cap A_{k,k} \neq \emptyset$, so

$$\operatorname{dist}(c, A_{k,k}) \leq \sqrt{N}\, s^{-k} < \epsilon.$$

It follows that $\mathrm{HD}\,(C, A_{k,k}) < \epsilon$ holds for all $k \geq k_1$. □

Next we give two more useful facts about the Hausdorff distance topology.

Definition 1.6.7. A subset C of a vector space is *convex* if for $x, y \in C$ and $0 \leq t \leq 1$ we have

$$(1 - t)\, x + t\, y \in C\,.$$

Proposition 1.6.8. *Let $\mathcal{C}$ be the collection of all closed, bounded, convex sets in $\mathbb{R}^N$. Then $\mathcal{C}$ is a closed subset of the metric space $(\mathcal{B}, \mathrm{HD}\,)$.*

Proof. There are several amusing ways to prove this assertion. One is by contradiction. If $\{S_j\}$ is a convergent sequence in $\mathcal{C}$, then let $S \in B$ be its limit. If S does not lie in $\mathcal{C}$ then S is not convex. Thus there is a segment ℓ with endpoints lying in S but with some interior point p not in S.

Let $\epsilon > 0$ be selected so that the open ball $U(p, \epsilon)$ does not lie in S. Let a, b be the endpoints of ℓ. Choose j so large that $\mathrm{HD}\,(S_j, S) < \epsilon/2$. For such j, there exist points $a_j, b_j \in S_j$ such that $|a_j - a| < \epsilon/3$ and $|b_j - b| < \epsilon/3$. But then each point $c_j(t) \equiv (1 - t)a_j + tb_j$ has distance less than $\epsilon/3$ from $c(t) \equiv (1 - t)a + tb$, $0 \leq t \leq 1$. In particular, there is a point p_j on the line segment ℓ_j connecting a_j to b_j such that $|p_j - p| < \epsilon/3$. Noting that p_j must lie in S_j, we see that we have contradicted our statement about $U(p, \epsilon)$. Therefore S must be convex. □

Proposition 1.6.9. *Let $\{S_j\}$ be a sequence of elements of $\mathcal{B}$, each of which is connected. Suppose that $S_j \to S$ in the Hausdorff metric. Then S must be connected.*

Proof. Suppose not. Then S is disconnected. So we may write $S = A \cup B$ with each of A and B closed and nonempty and $A \cap B = \emptyset$. Then there is a number $\eta > 0$ such that if $a \in A$ and $b \in B$ then $|a - b| > \eta$.

Choose j so large that HD $(S_j, S) < \eta/3$. Define

$$A_j = \{s \in S_j : \mathrm{dist}(s, A) \leq \eta/3\} \quad \text{and} \quad B_j = \{s \in S_j : \mathrm{dist}(s, B) \leq \eta/3\}.$$

Clearly $A_j \cap B_j = \emptyset$ and A_j, B_j are closed and nonempty. Moreover, $A_j \cup B_j = S_j$. That contradicts the connectedness of S_j and completes the proof. □

Remark 1.6.10. It is certainly possible to have totally disconnected sets E_j, $j = 1, 2, \ldots$, such that $E_j \to E$ as $j \to \infty$ and E is connected (exercise).

Now we turn to a new arena in which the Hausdorff distance is applicable.

Definition 1.6.11. Let V be an $(N-1)$-dimensional vector subspace of $\mathbb{R}^N$. *Steiner symmetrization*[20] with respect to V is the operation that associates with each bounded subset T of $\mathbb{R}^N$ the subset $\widetilde{T}$ of $\mathbb{R}^N$ having the property that for each straight line ℓ perpendicular to V, $\ell \cap \widetilde{T}$ is a closed line segment with center in V or is empty and the conditions

$$\mathcal{L}^1(\ell \cap \widetilde{T}) = \mathcal{L}^1(\ell \cap T) \tag{1.31}$$

and

$$\ell \cap \widetilde{T} = \emptyset \quad \text{if and only if} \quad \ell \cap T = \emptyset$$

hold, where in (1.31), $\mathcal{L}^1$ means the Lebesgue measure resulting from isometrically identifying the line ℓ with $\mathbb{R}$.

In Figure 1.5, B is the Steiner symmetrization of A with respect to the line L.

Steiner used symmetrization to give a proof of the isoperimetric theorem that he presented to the Berlin Academy of Science in 1836 (see [Str 36]). The results in the remainder of this section document a number of aspects of the behavior of Steiner symmetrization.

Proposition 1.6.12. *If T is a bounded $\mathcal{L}^N$-measurable subset of $\mathbb{R}^N$ and if S is obtained from T by Steiner symmetrization, then S is $\mathcal{L}^N$-measurable and*

$$\mathcal{L}^N(T) = \mathcal{L}^N(S).$$

Proof. This is a consequence of Fubini's theorem. □

Lemma 1.6.13. *Fix $0 < M < \infty$. If A and $A_1, A_2, \ldots$ are closed subsets of $\mathbb{R}^N \cap \overline{\mathbb{B}}(0, M)$ such that*

[20] Jakob Steiner (1796–1863).

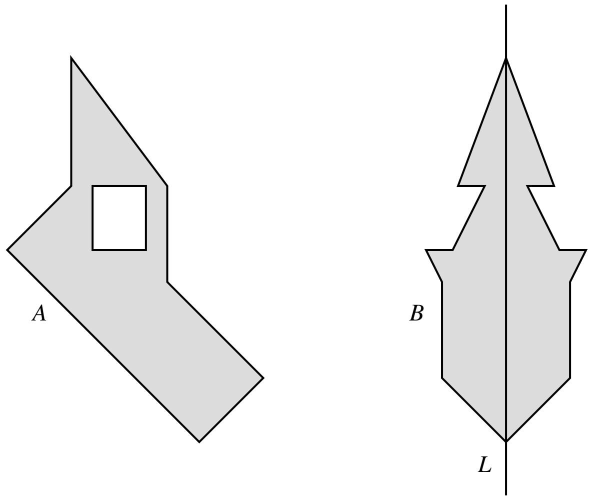

Fig. 1.5. Steiner symmetrization.

$$\bigcap_{i_0=1}^{\infty} \overline{\left[\bigcup_{i=i_0}^{\infty} A_i \right]} \subseteq A,$$

then

$$\limsup_i \mathcal{L}^N(A_i) \leq \mathcal{L}^N(A).$$

Proof. Let $\epsilon > 0$ be arbitrary. Then there exists an open set U with $A \subseteq U$ and

$$\mathcal{L}^N(U) \leq \mathcal{L}^N(A) + \epsilon.$$

A routine argument shows that for all sufficiently large i, $A_i \subseteq U$. It follows that

$$\limsup_i \mathcal{L}^N(A_i) \leq \mathcal{L}^N(U),$$

and the fact that ϵ was arbitrary implies the lemma. □

Proposition 1.6.14. *If T is a compact subset of $\mathbb{R}^N$ and if S is obtained from T by Steiner symmetrization, then S is compact.*

Proof. Let V be an $(N-1)$-dimensional vector subspace of $\mathbb{R}^N$, and suppose that S is the result of Steiner symmetrization of T with respect to V. It is clear that the boundedness of T implies the boundedness of S. To see that S is closed, consider any sequence of points $p_1, p_2, \ldots$ in S that converges to some point p. Each p_i lies in a line ℓ_i perpendicular to V, and we know that

$$\operatorname{dist}(p_i, V) \leq \frac{1}{2}\mathcal{L}^1(\ell_i \cap S) = \frac{1}{2}\mathcal{L}^1(\ell_i \cap T).$$

We also know that the line perpendicular to V and containing p must be the limit of the sequence of lines $\ell_1, \ell_2, \ldots$. Further, we know that

$$\operatorname{dist}(p, V) = \lim_{i\to\infty} \operatorname{dist}(p_i, V).$$

The inequality

$$\limsup_i \mathcal{L}^1(\ell_i \cap T) \le \mathcal{L}^1(\ell \cap T) \tag{1.32}$$

would allow us to conclude that

$$\operatorname{dist}(p, V) = \lim_{i\to\infty} \operatorname{dist}(p_i, V) \le \frac{1}{2} \limsup_{i\to\infty} \mathcal{L}^1(\ell_i \cap T) \le \frac{1}{2}\mathcal{L}^1(\ell \cap T),$$

and thus that $p \in S$.

To obtain the inequality (1.32), we let q_i be the vector parallel to V that translates ℓ_i to ℓ, and we apply Lemma 1.6.13, with N replaced by 1 and with ℓ identified with $\mathbb{R}$, to the sets $A_i = \tau_{q_i}(\ell_i \cap T)$, which are the translates of the sets $\ell_i \cap T$. We can take $A = \ell \cap T$, because T is closed. □

Proposition 1.6.15. *If T is a bounded, convex subset of $\mathbb{R}^N$ and S is obtained from T by Steiner symmetrization, then S is also a convex set.*

Proof. Let V be an $(N-1)$-dimensional vector subspace of $\mathbb{R}^N$, and suppose that S is the result of Steiner symmetrization of T with respect to V. Let x and y be two points of S. We let x' and y' denote the points obtained from x and y by reflection through the hyperplane V. Also, let ℓ_x and ℓ_y denote the lines perpendicular to V and passing through the points x and y, respectively. By the definition of Steiner symmetrization and the convexity of T, we see that $\ell_x \cap T$ must contain a line segment, say from p_x to q_x, of length at least $\operatorname{dist}(x, x')$. Likewise, $\ell_y \cap T$ contains a line segment from p_y to q_y of length at least $\operatorname{dist}(y, y')$. The convex hull of the four points p_x, q_x, p_y, q_y is a trapezoid, Q, which is a subset of T.

We claim that the trapezoid, Q', that is the convex hull of x, x', y, y' must be contained in S. Let x'' be the point of intersection of ℓ_x and V. Similarly, define y'' to be the intersection of ℓ_y and V. For any $0 \le \tau \le 1$, the line ℓ'' perpendicular to V and passing through

$$(1-\tau)x'' + \tau y''$$

intersects the trapezoid $Q \subseteq T$ in a line segment of length

$$d_1 = (1-\tau)\operatorname{dist}(p_x, q_x) + \tau \operatorname{dist}(p_y, q_y)\,,$$

and it intersects the trapezoid Q' in a line segment, centered about V, of length

$$d_2 = (1-\tau)\operatorname{dist}(x, x') + \tau \operatorname{dist}(y, y').$$

But S must contain a closed line segment of ℓ'', centered about V, of length at least d_1. Since d_1 at least as large as d_2,

$$\ell'' \cap Q' \subseteq \ell'' \cap S.$$

Since the choice of $0 \le \tau \le 1$ was arbitrary, we conclude that $Q' \subseteq S$. In particular, the line segment from x to y is contained in Q' and thus in S. □

The power of Steiner symmetrization obtains from the following theorem.

Theorem 1.6.16. *Suppose that $\mathcal{C}$ is a nonempty family of nonempty compact subsets of $\mathbb{R}^N$ that is closed in the Hausdorff distance topology and that is closed under the operation of Steiner symmetrization with respect to any $(N-1)$-dimensional vector subspace of $\mathbb{R}^N$. Then $\mathcal{C}$ contains a closed ball (possibly of radius 0) centered at the origin.*

Proof. Let $\mathcal{C}$ be such a family of compact subsets of $\mathbb{R}^N$ and set

$$r = \inf\{s : \text{ there exists } T \in \mathcal{C} \text{ with } T \subseteq \overline{\mathbb{B}}(0,s)\}.$$

If $r = 0$, we are done, so we may assume $r > 0$. By Theorem 1.6.6, any uniformly bounded family of nonempty compact sets is compact in the Hausdorff distance topology, so we can suppose there exists a $T \in \mathcal{C}$ with $T \subseteq \overline{\mathbb{B}}(0,r)$.

We claim that $T = \overline{\mathbb{B}}(0,r)$. If not, then there exist $p \in \overline{\mathbb{B}}(0,r)$ and $\epsilon > 0$ such that $T \subseteq \overline{\mathbb{B}}(0,r) \setminus \mathbb{B}(p,\epsilon)$. Suppose T_1 is the result of Steiner symmetrization of T with respect to any arbitrarily chosen $(N-1)$-dimensional vector subspace V. Let ℓ be the line perpendicular to V and passing through p. For any line ℓ' parallel to ℓ and at distance less than ϵ from ℓ, the Lebesgue measure of the intersection of ℓ' with T must be strictly less than the length of the intersection of ℓ' with $\overline{\mathbb{B}}(0,r)$, so the intersection of ℓ' with $\partial\overline{\mathbb{B}}(0,r)$ is not in T_1. We conclude that if p_1 is either one of the points of intersection of the sphere of radius r about the origin with the line ℓ, then

$$\mathbb{B}(p_1,\epsilon) \cap \partial\overline{\mathbb{B}}(0,r) \cap T_1 = \emptyset.$$

Choose a finite set of distinct additional points $p_2, p_3, \ldots, p_k$ such that

$$\partial\overline{\mathbb{B}}(0,r) \subseteq \bigcup_{i=1}^{k} \mathbb{B}(p_i,\epsilon).$$

For $i = 1, 2, \ldots, k-1$, let T_{i+1} be the result of Steiner symmetrization of T_i with respect to the $(N-1)$-dimensional vector subspace perpendicular to the line through p_i and p_{i+1}. By the lemma it follows that

$$\mathbb{B}(p_i,\epsilon) \cap \partial\overline{\mathbb{B}}(0,r) \cap T_j = \emptyset$$

holds for $i \le j \le k$. Thus we have

$$T_k \cap \partial\overline{\mathbb{B}}(0,r) = \emptyset,$$

so

$$T_k \subseteq \overline{\mathbb{B}}(0,s)$$

holds for some $s < r$, a contradiction. □

1.7 Borel and Suslin Sets

In this section, we discuss the Borel and Suslin sets. The goal of the section is to show that for all reasonable measures on Euclidean space, the continuous images of Borel sets are measurable sets (Corollary 1.7.19). This result is based on three facts: every Borel set is a Suslin set (Theorem 1.7.9), the continuous image of a Suslin set is a Suslin set (Theorem 1.7.12), and all Suslin sets are measurable (Corollary 1.7.18).

To put it as briefly as possible, the Suslin sets in $\mathbb{R}^N$ are the sets obtained as the orthogonal projections of Borel sets in $\mathbb{R}^{N+M}$. The history of Suslin sets is of some interest. In [Leb 05] (on page 191) Lebesgue had claimed that every projection of a Borel set is again a Borel set—Lebesgue even gave what he believed was a proof. It was Suslin (see [Sus 17]) who showed that, in fact, the Borel sets form a proper subfamily within the Suslin sets, and consequently, there exists a Borel set whose orthogonal projection is *not* a Borel set. While it is clearly of interest to know that there exists a Suslin set that is not a Borel set, we will not prove or use that result. We refer the interested reader to [Fed 69, 2.2.11], [Hau 62, Section 33], or [Jec 78, Section 39].

Construction of the Borel Sets

In Section 1.2 we defined the Borel sets in a topological space to be the members of the smallest σ-algebra that includes all the open sets. The virtue of this definition is its efficiency, but the price we pay for that efficiency is the absence of a mechanism for constructing the Borel sets. In this section, we will provide that constructive definition of the Borel sets.

For definiteness we work on $\mathbb{R}^N$. We will use transfinite induction over the smallest uncountable ordinal ω_1 (see Appendix A.1 for a brief introduction to transfinite induction) to define families of sets Σ^0_α and Π^0_α, for $\alpha < \omega_1$. For us, the superscript 0's are superfluous, but we include them since they are typically used in descriptive set theory.

Definition 1.7.1. Set

$$\Sigma^0_1 = \text{the family of all open sets in } \mathbb{R}^N,$$

$$\Pi^0_1 = \text{the family of all closed sets in } \mathbb{R}^N.$$

If $\alpha < \omega_1$, and Σ^0_β and Π^0_β have been defined for all $\beta < \alpha$, then set

$$\Sigma^0_\alpha = \text{the family of sets of the form}$$

$$A = \bigcup_{i=1}^{\infty} A_i, \quad \text{where each } A_i \in \Pi^0_\beta \text{ for some } \beta < \alpha, \tag{1.33}$$

$$\Pi^0_\alpha = \text{the family of sets of the form } \mathbb{R}^N \setminus A \text{ for } A \in \Sigma^0_\alpha. \tag{1.34}$$

Since the complement of a union is the intersection of the complements, we see that we can also write

Π^0_α = the family of sets of the form

$$A = \bigcap_{i=1}^{\infty} A_i, \quad \text{where each } A_i \in \Sigma^0_\beta \text{ for some } \beta < \alpha. \tag{1.35}$$

By transfinite induction over ω_1, we see that for $\alpha < \omega_1$, all the elements of Σ^0_α and Π^0_α are Borel sets.

Lemma 1.7.2. *If* $1 \le \beta < \alpha < \omega_1$, *then*

$$\Sigma^0_\beta \subseteq \Pi^0_\alpha\,, \qquad \Pi^0_\beta \subseteq \Sigma^0_\alpha\,, \qquad \Sigma^0_\beta \subseteq \Sigma^0_\alpha\,, \qquad \Pi^0_\beta \subseteq \Pi^0_\alpha$$

hold.

Proof. By (1.33) and (1.35), we see that $\Sigma^0_\beta \subseteq \Pi^0_\alpha$ and $\Pi^0_\beta \subseteq \Sigma^0_\alpha$ hold whenever $1 \le \beta < \alpha < \omega_1$.

Every open set in Euclidean space is a countable union of closed sets, so $\Sigma^0_1 \subseteq \Sigma^0_2$ holds. Consequently, we also have $\Pi^0_1 \subseteq \Pi^0_2$. Since $\Sigma^0_1 \subseteq \Pi^0_2 \subseteq \Sigma^0_\alpha$ holds whenever $2 < \alpha$ and since $\Pi^0_1 \subseteq \Sigma^0_2$ holds, we have $\Sigma^0_1 \subseteq \Sigma^0_\alpha$ and $\Pi^0_1 \subseteq \Pi^0_\alpha$ for all $1 < \alpha < \omega_1$.

Fix $1 \le \beta < \alpha < \omega_1$. Suppose $\Sigma^0_\gamma \subseteq \Sigma^0_\alpha$ and $\Pi^0_\gamma \subseteq \Pi^0_\alpha$ hold whenever $\gamma < \beta$. Any set $A \in \Sigma^0_\beta$ must be of the form $A = \cup_{i=1}^{\infty} A_i$ with each $A_i \in \Pi^0_\gamma$ for some $\gamma < \beta$. Then since $\beta < \alpha$, we see that $A \in \Sigma^0_\alpha$. Thus $\Sigma^0_\beta \subseteq \Sigma^0_\alpha$. Similarly, we have $\Pi^0_\beta \subseteq \Pi^0_\alpha$. □

Corollary 1.7.3. *We have*

$$\bigcup_{\alpha<\omega_1} \Sigma^0_\alpha = \bigcup_{\alpha<\omega_1} \Pi^0_\alpha\,. \tag{1.36}$$

Theorem 1.7.4. *The family of sets in* (1.36) *is the* σ*-algebra of Borel subsets of* $\mathbb{R}^N$.

Proof. Let $\mathcal{B}$ denote the family of sets in (1.36). To see that $\mathcal{B}$ is closed under countable unions, suppose we are given $A_1, A_2, \ldots$ in $\mathcal{B}$. Considering the left-hand side of (1.36), we see that for each i, there is $\alpha_i < \omega_1$ such that $A_i \in \Sigma^0_{\alpha_i}$. Since the sequence $\alpha_1, \alpha_2, \ldots$ is countable, but ω_1 is uncountable, there is $\alpha^* < \omega_1$ with $\alpha_i < \alpha^*$ for all i (see Lemma A.1.4). We conclude that $\cup_{i=1}^{\infty} A_i \in \Sigma_{\alpha^*}$. Thus $\mathcal{B}$ is closed under countable unions. We argue similarly to see that $\mathcal{B}$ is closed under countable intersections and complements. □

Because in the definition of Π^0_α, equation (1.34) can be replaced by (1.35), Theorem 1.7.4 has the following corollary.

Corollary 1.7.5. *The family of Borel sets in* $\mathbb{R}^N$ *is the smallest family of sets containing the open sets that is closed under countable unions and countable intersections. Likewise, the family of Borel sets in* $\mathbb{R}^N$ *is the smallest family of sets, containing the closed sets, that is closed under countable unions and countable intersections.*

Suslin Sets

Recall that the positive integers are denoted by $\mathbb{N}$. We let $\widetilde{\mathcal{N}}$ denote the set of all finite sequences of positive integers and we let $\mathcal{N}$ denote the set of all infinite sequences of positive integers, so

$$\widetilde{\mathcal{N}} = \{\, (n_1, n_2, \ldots, n_k) : k \in \mathbb{N},\ n_i \in \mathbb{N} \text{ for } i = 1, 2, \ldots, k \,\},$$

$$\mathcal{N} = \{\, (n_1, n_2, \ldots) : n_i \in \mathbb{N} \text{ for } i = 1, 2, \ldots \,\}.$$

Definition 1.7.6. Let $\mathcal{M}$ be a collection of subsets of a set X. Suppose that there is a set $M_{n_1,n_2,\ldots,n_k} \in \mathcal{M}$ associated with every finite sequence of positive integers. We can represent this relation as a function $\nu : \widetilde{\mathcal{N}} \to \mathcal{M}$ defined by

$$(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} M_{n_1,n_2,\ldots,n_k} .$$

Such a function ν is called a *determining system in* $\mathcal{M}$. Associated with the determining system ν is the set called the *nucleus of* ν denoted by $\mathrm{N}(\nu)$ and defined by

$$\mathrm{N}(\nu) = \bigcup_{\substack{n \in \mathcal{N} \\ n=(n_1,n_2,\ldots)}} \left(M_{n_1} \cap M_{n_1,n_2} \cap \cdots \cap M_{n_1,n_2,\ldots,n_k} \cap \cdots \right) .$$

Suslin's operation (A) is the function that when applied to the argument ν produces the result $\mathrm{N}(\nu)$. We will say that $\mathrm{N}(\nu)$ is a *Suslin set generated by* $\mathcal{M}$. The family of all Suslin sets generated by $\mathcal{M}$ will be denoted by $\mathcal{M}_{(\mathrm{A})}$.

By the *Suslin sets in a topological space* we mean the Suslin sets generated by the family of closed sets.

Since $\mathcal{N}$ has the same cardinality as the real numbers, we see that the nucleus is formed by an *uncountable* union of countable intersections of sets in $\mathcal{M}$. We might expect that operation (A) could be extremely powerful, but at the outset it is not immediately clear what can be done with the operation. The next proposition tells us that operation (A) is at least as powerful as those used to form the Borel sets.

Proposition 1.7.7. *Suppose $A_1, A_2, \ldots \in \mathcal{M}$, then there exist determining systems ν_U and ν_I such that*

$$\mathrm{N}(\nu_U) = \bigcup_{i=1}^{\infty} A_i \quad \textit{and} \quad \mathrm{N}(\nu_I) = \bigcap_{i=1}^{\infty} A_i .$$

Proof. Define ν_U and ν_I by

$$(n_1, n_2, \ldots, n_k) \overset{\nu_U}{\longmapsto} A_{n_1} ,$$

$$(n_1, n_2, \ldots, n_k) \overset{\nu_I}{\longmapsto} A_k .$$

It is easy to see that ν_U and ν_I have the desired properties. □

The next theorem that tells us that repeated applications of operation (A) produce nothing that cannot be produced with only one application of the operation.

Theorem 1.7.8. *If $\mathcal{M}$ is a family of sets, if $\emptyset \in \mathcal{M}$, and if $\mathcal{M}_{(\mathrm{A})}$ is the family of Suslin sets generated by $\mathcal{M}$, then any Suslin set generated by $\mathcal{M}_{(\mathrm{A})}$ is already an element of $\mathcal{M}_{(\mathrm{A})}$. Symbolically, we have*

$$\left(\mathcal{M}_{(\mathrm{A})}\right)_{(\mathrm{A})} = \mathcal{M}_{(\mathrm{A})} .$$

Proof. Let

$$(n_1, n_2, \ldots, n_k) \stackrel{\nu}{\longmapsto} M_{n_1,n_2,\ldots,n_k} \in \mathcal{M}_{(\mathrm{A})}$$

be a determining system in $\mathcal{M}_{(\mathrm{A})}$. For each $n_1, n_2, \ldots, n_k \in \widetilde{\mathcal{N}}$, the set $M_{n_1,n_2,\ldots,n_k}$ must itself be the nucleus of a determining system $\nu_{n_1,n_2,\ldots,n_k}$ in $\mathcal{M}$; that is,

$$(q_1, q_2, \ldots, q_\ell) \xrightarrow{\nu_{n_1,n_2,\ldots,n_k}} M^{q_1,q_2,\ldots,q_\ell}_{n_1,n_2,\ldots,n_k} \in \mathcal{M} ,$$

$$M_{n_1,n_2,\ldots,n_k} = \bigcup_{\substack{q\in\mathcal{N}\\ q=(q_1,q_2,\ldots)}} \left(M^{q_1}_{n_1,n_2,\ldots,n_k} \cap M^{q_1,q_2}_{n_1,n_2,\ldots,n_k} \cap \cdots \cap M^{q_1,q_2,\ldots,q_\ell}_{n_1,n_2,\ldots,n_k} \cap \cdots\right) ,$$

$$\mathrm{N}(\nu) = \bigcup_{\substack{n\in\mathcal{N}\\ n=(n_1,n_2,\ldots)}} \left(M_{n_1} \cap M_{n_1,n_2} \cap \cdots \cap M_{n_1,n_2,\ldots,n_k} \cap \cdots\right) .$$

We can rewrite $\mathrm{N}(\nu)$ as the union of the sets

$$\begin{array}{lllll}
 & \left(M^{q_1^1}_{n_1}\right. & \cap\ M^{q_1^1,q_2^1}_{n_1} & \cap \cdots \cap\ M^{q_1^1,q_2^1,\ldots,q_\ell^1}_{n_1} & \left.\cap \cdots\right) \\
\bigcap & \left(M^{q_1^2}_{n_1,n_2}\right. & \cap\ M^{q_1^2,q_2^2}_{n_1,n_2} & \cap \cdots \cap\ M^{q_1^2,q_2^2,\ldots,q_\ell^2}_{n_1,n_2} & \left.\cap \cdots\right) \\
\vdots & \vdots & \vdots & \vdots & \\
\bigcap & \left(M^{q_1^k}_{n_1,n_2,\ldots,n_k}\right. & \cap\ M^{q_1^k,q_2^k}_{n_1,n_2,\ldots,n_k} & \cap \cdots \cap\ M^{q_1^k,q_2^k,\ldots,q_\ell^k}_{n_1,n_2,\ldots,n_k} & \left.\cap \cdots\right) \\
\vdots & \vdots & \vdots & \vdots & .
\end{array} \tag{1.37}$$

Notice that the set in the kth row and ℓth column of (1.37) is indexed by k subscripts and ℓ superscripts. The choices of the subscripts and superscripts are constrained by the following requirements:

$$\left.\begin{array}{l}
\text{in any row, the list of subscripts is constant,}\\
\text{in any row, the list of superscripts grows by concatenation,}\\
\text{in any column, the list of subscripts grows by concatenation.}
\end{array}\right\} \tag{1.38}$$

Let the prime numbers in increasing numerical order be given by the list

$$p_1,\ p_2,\ p_3,\ldots .$$

We can use the list of primes to encode the information concerning the number of subscripts, the number of superscripts, and their values as follows: Set

$$m = p_1^k \cdot p_2^\ell \cdot p_3^{n_1} \cdot p_4^{n_2} \cdots p_{k+2}^{n_k} \cdot p_{k+3}^{q_1^k} \cdot p_{k+4}^{q_2^k} \cdots p_{\ell+k+2}^{q_\ell^k} . \tag{1.39}$$

Given a positive integer m, the unique factorization of m into prime powers determines whether m is of the form (1.39). Certainly not every positive integer m is of the form (1.39), nor is every sequence of positive integers $m_1, m_2, \ldots$ consistent with the conditions (1.38), even if the individual numbers m_i are of the form (1.39). But it is true that any sequence of sets in (1.37) will give rise to a sequence of positive integers $m_1, m_2, \ldots$ of the form (1.39) that satisfies the conditions (1.38).

We now define the determining system

$$(m_1, m_2, \ldots, m_k) \stackrel{\sigma}{\longmapsto} S_{m_1,m_2,\ldots,m_k} .$$

For each positive integer m, if m is of the form (1.39), then the numbers k, ℓ, $n_1, n_2, \ldots, n_k$, $q_1^k, q_2^k, \ldots, q_\ell^k$ are uniquely determined by (1.39). So we can make the definition

$$T_m = \begin{cases} S_{n_1,n_2,\ldots,n_k}^{q_1^k,q_2^k,\ldots,q_\ell^k} & \text{if } m \text{ is of the form (1.39)}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Then, for the sequence of positive integers $m_1, m_2, \ldots,$ set

$$S_{m_1,m_2,\ldots,m_k} = \begin{cases} T_{m_1} \cap T_{m_2} \cap \cdots \cap T_{m_k} & \text{if (1.38) is not violated,} \\ \emptyset & \text{otherwise.} \end{cases}$$

For $m = (m_1, m_2, \ldots) \in \mathcal{N}$, the set

$$S_{m_1} \cap S_{m_1,m_2} \cap \cdots \cap S_{m_1,m_2,\ldots,m_k} \cap \cdots$$

is either one of the sets in (1.37) or the empty set. By construction, every set in (1.37) gives rise to a sequence $m = (m_1, m_2, \ldots) \in \mathcal{N}$ such that

$$S_{m_1} \cap S_{m_1,m_2} \cap \cdots \cap S_{m_1,m_2,\ldots,m_k} \cap \cdots$$

equals that set in (1.37). Thus we have $\mathrm{N}(\nu) = \mathrm{N}(\sigma)$. □

Theorem 1.7.9. *Every Borel set in* $\mathbb{R}^N$ *is a Suslin set.*

Proof. By Proposition 1.7.7 and Theorem 1.7.8, the collection of Suslin sets is closed under countable unions and countable intersections. Thus by Corollary 1.7.5, the collection of Suslin sets contains all the Borel sets. □

Continuous Images of Suslin Sets

Suppose $f : X \to Y$ is a function from a set X to a set Y. The inverse image of a union of sets equals the union of the inverse images, and likewise the inverse image of an intersection of sets equals the intersection of the inverse images. Images of sets under functions are not as well behaved as inverse images; nonetheless, we do have the following result—which is easily verified.

Proposition 1.7.10. *Let $f : X \to Y$.*

(1) *For $\{A_\alpha\}_{\alpha\in I}$ a collection of subsets of X, $f\left(\bigcup_{\alpha\in I} A_\alpha\right) = \bigcup_{\alpha\in I} f(A_\alpha)$ holds.*
(2) *For $X \supseteq A_1 \supseteq A_2 \supseteq \cdots$, $f\left(\bigcap_{i=1}^\infty A_i\right) \subseteq \bigcap_{i=1}^\infty f(A_i)$ holds and strict inclusion is possible.*

To obtain an equality for images of intersections, we need to look at continuous functions and decreasing sequences of compact sets.

Proposition 1.7.11. *Let X and Y be topological spaces and let $f : X \to Y$ be continuous. If X is sequentially compact, $X \supseteq C_1 \supseteq C_2 \supseteq \cdots$, and if each C_i is a closed subset of X, then $f\left(\bigcap_{i=1}^\infty C_i\right) = \bigcap_{i=1}^\infty f(C_i)$.*

Proof. By Proposition 1.7.10, we need only show that $\bigcap_{i=1}^\infty f(C_i) \subseteq f\left(\bigcap_{i=1}^\infty C_i\right)$, so suppose $y \in \bigcap_{i=1}^\infty f(C_i)$.

For each i, there is $x_i \in C_i$ with $f(x_i) = y$, and because the sets C_i are decreasing, we have $x_j \in C_i$ whenever $j \geq i$.

Set $x_{0,j} = x_j$ for $j = 1, 2, \ldots$. Since C_1 is sequentially compact, there is a convergent subsequence $\{x_{1,j}\}_{j=1}^\infty$ of $\{x_{0,j}\}_{j=1}^\infty$. Arguing inductively, suppose $1 \leq i$ and that we have already constructed a convergent sequence $\{x_{i,j}\}_{j=1}^\infty$ that is a subsequence of $\{x_{h,j}\}_{j=1}^\infty$, for $0 \leq h \leq i-1$, and is such that every $x_{i,j}$ is a point of C_i, for $j = 1, 2, \ldots$. Since $\{x_{i,j}\}_{j=1}^\infty$ is a subsequence of the original sequence $\{x_{0,j}\}_{j=1}^\infty$, there is a j_* such that $x_{i,j} \in C_{i+1}$ holds for all j with $j_* \leq j$. Since C_{i+1} is sequentially compact, we can select a convergent subsequence $\{x_{i+1,j}\}_{j=1}^\infty$ of $\{x_{i,j}\}_{j=j_*}^\infty$, and thus satisfy the induction hypothesis.

By construction, the sequence $\{x_{j,j}\}_{j=1}^\infty$ is convergent. Hence we have $\lim_{j\to\infty} x_{j,j} \in \bigcap_{i=1}^\infty C_i$, $f\left(\lim_{j\to\infty} x_{j,j}\right) = \lim_{j=1}^\infty f\left(x_{j,j}\right) = y$, and thus we have shown that $y \in \bigcap_{i=1}^\infty C_i$. □

Theorem 1.7.12. *If $f : \mathbb{R}^N \to \mathbb{R}^M$ is continuous and $S \subseteq \mathbb{R}^N$ is a Suslin set, then $f(S)$ is a Suslin subset of $\mathbb{R}^M$.*

Proof. Since any closed subset of $\mathbb{R}^N$ is a countable union of compact sets, we see that if $\mathcal{K}$ is the collection of compact subsets of $\mathbb{R}^N$, then $\mathcal{K}_{(\mathrm{A})}$ is the collection of Suslin sets.

Let $S \subseteq \mathbb{R}^N$ be a Suslin set, and let ν be a determining system in $\mathcal{K}$ such that $S = \mathrm{N}(\nu)$. Since any finite intersection of compact sets is compact, we see that the determining system $(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} K_{n_1,n_2,\ldots,n_k}$ has the same nucleus as the determining system $(n_1, n_2, \ldots, n_k) \overset{\widetilde{\nu}}{\longmapsto} H_{n_1,n_2,\ldots,n_k}$ in $\mathcal{K}$ given by

$$H_{n_1,n_2,\ldots,n_k} = K_{n_1} \cap K_{n_1,n_2} \cap \cdots \cap K_{n_1,n_2,\ldots,n_k}.$$

Because the sets $\left\{H_{n_1,n_2,\ldots,n_k}\right\}_{k=1}^\infty$ form a decreasing sequence of compact sets, we can apply Propositions 1.7.10 and 1.7.11 to conclude that

$$
\begin{aligned}
f(S) &= f[\mathrm{N}(\nu)] = f[\mathrm{N}(\widetilde{\nu})] \\
&= f\left[\bigcup_{\substack{n\in\mathcal{N}\\ n=(n_1,n_2,\ldots)}} \left(H_{n_1}\cap H_{n_1,n_2}\cap\cdots\cap H_{n_1,n_2,\ldots,n_k}\cap\cdots\right)\right] \\
&= \bigcup_{\substack{n\in\mathcal{N}\\ n=(n_1,n_2,\ldots)}} \Big(f(H_{n_1})\cap f(H_{n_1,n_2})\cap\cdots\cap f(H_{n_1,n_2,\ldots,n_k})\cap\cdots\Big),
\end{aligned}
$$

and so we see that $f(S)$ is a Suslin set in $\mathbb{R}^M$. □

Measurability of Suslin Sets

In order to prove that the Suslin sets are measurable, we need to introduce some additional structures similar to the nucleus of a determining system.

Definition 1.7.13. Let $(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} A_{n_1,n_2,\ldots,n_k}$ be given. Let $h_1, h_2, \ldots, h_s$ be a finite sequence of positive integers. We define the following sets:

$$
\mathrm{N}^{h_1,h_2,\ldots,h_s}(\nu) = \bigcup_{\substack{(n_1,n_2,\ldots)\in\mathcal{N}\\ n_i\le h_i,\ 1\le i\le s}} A_{n_1}\cap A_{n_1,n_2}\cap\cdots\cap A_{n_1,n_2,\ldots,n_k}\cap\cdots\,, \tag{1.40}
$$

$$
\mathrm{N}_{h_1,h_2,\ldots,h_s}(\nu) = \bigcup_{n_1=1}^{h_1}\bigcup_{n_2=1}^{h_2}\cdots\bigcup_{n_s=1}^{h_s} A_{n_1}\cap A_{n_1,n_2}\cap\cdots\cap A_{n_1,n_2,\ldots,n_s}\,. \tag{1.41}
$$

The next proposition follows immediately from the definition.

Proposition 1.7.14. *Let* $(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} A_{n_1,n_2,\ldots,n_k}$ *be given. We have*

$$
\mathrm{N}^{1}(\nu)\subseteq\cdots\subseteq\mathrm{N}^{h}(\nu)\subseteq\mathrm{N}^{h+1}(\nu)\subseteq\cdots\,,
$$

$$
\mathrm{N}(\nu) = \bigcup_{k=1}^{\infty}\mathrm{N}^{k}(\nu)\,,
$$

$$
\mathrm{N}^{h_1,\ldots,h_s,1}(\nu)\subseteq\cdots\subseteq\mathrm{N}^{h_1,\ldots,h_s,k}(\nu)\subseteq\mathrm{N}^{h_1,\ldots,h_s,k+1}(\nu)\subseteq\cdots\,,
$$

$$
\mathrm{N}^{h_1,\ldots,h_s}(\nu) = \bigcup_{k=1}^{\infty}\mathrm{N}^{h_1,\ldots,h_s,k}(\nu)\,.
$$

Corollary 1.7.15. *If* μ *is a regular measure on the nonempty set* X *and* ν *is a determining system in any family of subsets of* X *and if* E *is any subset of* X*, then*

$$
\lim_{k\to\infty}\mu\Big[E\cap\mathrm{N}^{k}(\nu)\Big] = \mu\Big[E\cap\mathrm{N}(\nu)\Big],
$$

$$
\lim_{k\to\infty}\mu\Big[E\cap\mathrm{N}^{h_1,h_2,\ldots,h_s,k}(\nu)\Big] = \mu\Big[E\cap\mathrm{N}^{h_1,h_2,\ldots,h_s}(\nu)\Big].
$$

Proof. Recall that Lemma 1.2.8 tells us that for a regular measure the measure of the union of an increasing sequence of sets is the limit of the measures of the sets, so the result follows immediately from Proposition 1.7.14. □

We will need the following lemma.

Lemma 1.7.16. *Let* $(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} A_{n_1,n_2,\ldots,n_k}$ *and* $(h_1, h_2, \ldots) \in \mathcal{N}$ *be given. Then we have*

$$\mathrm{N}_{h_1}(\nu) \cap \mathrm{N}_{h_1,h_2}(\nu) \cap \cdots \cap \mathrm{N}_{h_1,h_2,\ldots,h_s}(\nu) \cap \cdots \subseteq \mathrm{N}(\nu)\,. \tag{1.42}$$

Proof. Fix a point x belonging to the left-hand side of (1.42).

First we claim that there exists a positive integer $n_1^0 \le h_1$ such that for every k with $2 \le k$, there exist $n_2, n_3, \ldots, n_k$ with $n_i \le h_i$, for $2 \le i \le k$, and with

$$x \in A_{n_1^0} \cap A_{n_1^0,n_2} \cap \cdots \cap A_{n_1^0,n_2,\ldots,n_k}\,.$$

To verify this, suppose it were not true. Then for each index $n_1 \le h_1$ there would be exist a positive integer $k(n_1)$ such that

$$x \notin A_{n_1} \cap A_{n_1,n_2} \cap \cdots \cap A_{n_1,n_2,\ldots,n_{k(n_1)}}$$

whenever $n_i \le h_i$ for $i = 2, 3, \ldots, k(n_1)$.

Setting $K(1) = \max\{\, k(1),\ k(2), \ldots,\ k(h_1)\,\}$, we see that

$$x \notin \bigcup_{n_1=1}^{h_1} \bigcup_{n_2=1}^{h_2} \cdots \bigcup_{n_{K(1)}=1}^{h_{K(1)}} A_{n_1} \cap A_{n_1,n_2} \cap \cdots \cap A_{n_1,n_2,\ldots,n_{K(1)}}\,,$$

which contradicts our assumption that x is an element of the left-hand side of (1.42).

Arguing inductively, suppose we have selected positive integers $n_1^0, n_2^0, \ldots, n_s^0$ satisfying

$$\left.\begin{array}{l} n_1^0 \le h_1,\, n_2^0 \le h_2, \ldots, n_s^0 \le h_s, \\ \text{for every } k \text{ with } s+1 \le k, \text{ there exist } n_{s+1}, n_{s+2}, \ldots, n_k \\ \qquad \text{with } n_i \le h_i, \text{ for } s+1 \le i \le k, \text{ and with} \\ \qquad x \in A_{n_1^0} \cap A_{n_1^0,n_2^0} \cap \cdots \cap A_{n_1^0,n_2^0,\ldots,n_s^0,n_{s+1},n_{s+2},\ldots,n_k}\,. \end{array}\right\} \tag{1.43}$$

We claim that we can select $n_{s+1}^0 \le h_{s+1}$ so that (1.43) holds with s replaced by $s+1$. If no such n_{s+1}^0 existed, then for each index $n_{s+1} \le h_{s+1}$ there would exist a positive integer $k(n_{s+1})$ such that

$$x \notin A_{n_1^0} \cap A_{n_1^0,n_2^0} \cap \cdots \cap A_{n_1^0,n_2^0,\ldots,n_s^0,n_{s+1},n_{s+2},\ldots,n_{k(n_{s+1})}}$$

whenever $n_i \le h_i$ for $i = s+1, s+2, \ldots, k(n_{s+1})$.

Setting $K(s+1) = \max\{\, k(1),\ k(2), \ldots,\ k(h_{s+1})\,\}$, we see that

$$x \notin \bigcup_{n_1=1}^{h_1} \bigcup_{n_2=1}^{h_2} \cdots \bigcup_{n_{K(s+1)}=1}^{h_{K(s+1)}} A_{n_1} \cap A_{n_1,n_2} \cap \cdots \cap A_{n_1,n_2,\ldots,n_{K(s+1)}} ,$$

which contradicts our assumption that x is an element of the left-hand side of (1.42).

Thus there exists an infinite sequence $n_1^0 \leq h_1, n_2^0 \leq h_2, \ldots$ such that

$$x \in A_{n_1^0} \cap A_{n_1^0,n_2^0} \cap \cdots \cap A_{n_1^0,n_2^0,\ldots,n_k^0} \cap \cdots ;$$

hence $x \in \mathrm{N}(\nu)$. □

Theorem 1.7.17. *Let μ be a regular measure on the nonempty set X, and let $\mathcal{M}$ be the collection of μ-measurable subsets of X. If ν is a determining system in $\mathcal{M}$, then* $\mathrm{N}(\nu)$ *is μ-measurable.*

Proof. Let the determining system ν be $(n_1, n_2, \ldots, n_k) \overset{\nu}{\longmapsto} M_{n_1,n_2,\ldots,n_k}$, and set $A = \mathrm{N}(\nu)$. We need to show that for any set $E \subseteq X$, we have

$$\mu(E \cap A) + \mu(E \setminus A) \leq \mu(E) .$$

We may assume that $\mu(E) < \infty$. Let $\epsilon > 0$ be arbitrary.

Using Corollary 1.7.15, we can inductively define a sequence of positive integers $h_1, h_2, \ldots$ such that

$$\mu\Big[C \cap \mathrm{N}^{h_1}(\nu)\Big] \geq \mu\Big[E \cap \mathrm{N}(\nu)\Big] - \epsilon/2$$

and

$$\mu\Big[C \cap \mathrm{N}^{h_1,h_2,\ldots,h_k}(\nu)\Big] \geq \mu\Big[E \cap \mathrm{N}^{h_1,h_2,\ldots,h_{k-1}}(\nu)\Big] - \epsilon/2^k .$$

We have $\mathrm{N}^{h_1,h_2,\ldots,h_k}(\nu) \subseteq \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)$, so

$$\mu\Big[E \cap \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big] \geq \mu\Big[E \cap \mathrm{N}^{h_1,h_2,\ldots,h_k}(\nu)\Big] \geq \mu(E \cap \mathrm{N}(\nu)) - \epsilon$$

holds, and thus, since $\mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)$ is μ-measurable,

$$\begin{aligned} \mu(E) &= \mu\Big[E \cap \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big] + \mu\Big[E \setminus \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big] \\ &\geq \mu\Big[E \cap \mathrm{N}(\nu)\Big] + \mu\Big[E \setminus \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big] - \epsilon . \end{aligned}$$

Now the sequence of sets $\Big\{\mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big\}_{k=1,2,\ldots}$ is descending, and by Lemma 1.7.16 its limit is a subset of $\mathrm{N}(\nu)$. Consequently the sequence $\Big\{X \setminus \mathrm{N}_{h_1,h_2,\ldots,h_k}\Big\}_{k=1,2,\ldots}$ is ascending and its limit contains the set $X \setminus \mathrm{N}(\nu)$. Hence

$$\lim_{k\to\infty} \mu\Big[E \setminus \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\Big] = \mu\left[E \setminus \bigcup_{k=1}^{\infty} \mathrm{N}_{h_1,h_2,\ldots,h_k}(\nu)\right] \geq \mu\Big[E \setminus \mathrm{N}(\nu)\Big],$$

so

$$\mu(E) \geq \mu\Big[E \cap \mathrm{N}(\nu)\Big] + \mu\Big[E \setminus \mathrm{N}(\nu)\Big] - \epsilon \,,$$

and the result follows since ϵ is an arbitrary positive number. □

Corollary 1.7.18. *If μ is a Borel regular measure on the topological space X, then all the Suslin sets in X are μ-measurable.*

Corollary 1.7.19. *If $f : \mathbb{R}^N \to \mathbb{R}^M$ is continuous, μ is a Borel regular measure on $\mathbb{R}^M$, and $S \subseteq \mathbb{R}^N$ is a Suslin set, then $f(S)$ is μ-measurable.*

Remark 1.7.20. The particular properties of Euclidean space required for Corollary 1.7.19 are that every open set is a countable union of closed sets and that every closed set is a countable union of compact sets.

2

Carathéodory's Construction and Lower-Dimensional Measures

In the study of geometric questions about sets it is useful to have various devices for measuring the size of those sets. Certainly lower-dimensional measures are one such mechanism. The classic construction of Carathéodory provides an umbrella paradigm that generates a great many such measures, suitable for a variety of applications. Our aim in the present chapter is to give a thorough development of this theory and to present a number of examples and applications.

Certainly the ideas that we present here began with Hausdorff [Hau 18] and Carathéodory [Car 14]. In the intervening eighty years they have developed in a number of startling and powerful new directions. We shall endeavor to describe both the history as well as some of the current directions.

2.1 The Basic Definition

Let $\mathcal{F}$ be a collection of sets in $\mathbb{R}^N$. These will be our "test sets" for constructing Hausdorff-type measures. Let $\zeta : \mathcal{F} \to [0, +\infty]$ be a function (called the *gauge* of the measure to be constructed). Then preliminary measures ϕ_δ, $0 < \delta \leq \infty$, are created as follows:

If $A \subseteq \mathbb{R}^N$, then set

$$\begin{aligned} &\phi_\delta(A) \\ &\quad = \inf \left\{ \sum_{S \in \mathcal{G}} \zeta(S) : \mathcal{G} \subseteq \mathcal{F} \cap \{S : \text{diam } S \leq \delta\},\ \text{card}(\mathcal{G}) \leq \aleph_0,\ \text{and}\, A \subseteq \bigcup_{S \in \mathcal{G}} S \right\}. \end{aligned} \tag{2.1}$$

Each number in the set over which we take the infimum in (2.1) is obtained by covering A by sets of diameter not exceeding δ (see Figure 2.1). We can either separately define $\phi_\delta(\emptyset) = 0$; or we can note that the empty sum is 0, so the empty covering of the empty set realizes the infimum in (2.1). Note that ϕ_δ clearly satisfies the monotonicity and subadditivity requirements of Definition 1.2.1 and thus is a measure.

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_2,

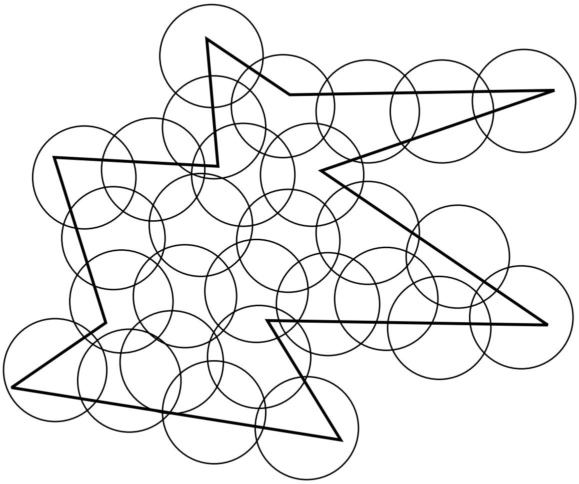

Fig. 2.1. Carathéodory's construction.

If $0 < \delta_1 < \delta_2 \le \infty$ then it is immediate that $\phi_{\delta_1} \ge \phi_{\delta_2}$. Thus we may set

$$\psi(A) = \lim_{\delta \to 0^+} \phi_\delta(A) = \sup_{\delta > 0} \phi_\delta(A) .$$

Certainly ψ is also a measure. This process for constructing the measure ψ is called *Carathéodory's construction.* Once the family of sets $\mathcal{F}$ and the gauge ζ have been selected, the resulting measure ψ is uniquely determined.

By applying Carathéodory's criterion, Theorem 1.2.13, we can immediately see that any open set is ψ-measurable. Indeed, one sees that

$$\phi_\delta(A \cup B) \ge \phi_\delta(A) + \phi_\delta(B)$$

whenever $\operatorname{dist}(A, B) > \delta > 0$. This follows because any set of diameter $\le \delta$ that is part of a covering of $A \cup B$ will either intersect A or intersect B but not both. Thus any collection $\mathcal{G}$ as above will partition naturally into a subcollection that covers A and a subcollection that covers B.

Example 2.1.1. Not every open set is ϕ_δ-measurable, for fixed $\delta > 0$. To see this, let $N = 1$, let $\mathcal{F}$ be the collection of open intervals, and let $\zeta(S) = (\operatorname{diam}(S))^{1/2}$. Define $I_1 = (0, \delta/2)$, $I_2 = (\delta/2, \delta)$, and $I = I_1 \cup I_2$. Then it is easy to see that

$$\phi_\delta(I_1) = (\delta/2)^{1/2} , \quad \phi_\delta(I_2) = (\delta/2)^{1/2} , \quad \phi_\delta(I) = \delta^{1/2} .$$

But then the inequality

$$\phi_\delta(I) \ge \phi_\delta(I_1) + \phi_\delta(I_2)$$

clearly fails. □

It is not difficult to show that if all members of $\mathcal{F}$ are Borel sets, then every subset A of $\mathbb{R}^N$ is contained in a Borel set $\widetilde{A}$ with the same ϕ_δ measure (just take the intersection of the unions of covers). Thus ψ is a Borel regular measure.

We now describe an alternative approach to Carathéodory's construction that is due to Federer [Fed 54]. In fact, $\psi(A)$ can be characterized as the infimum of the set of all numbers t with this property:

For each open covering $\mathcal{U}$ of A there exists a countable subfamily $\mathcal{G}$ of $\mathcal{F}$ such that each member of $\mathcal{G}$ is contained in some member of $\mathcal{U}$, $\mathcal{G}$ covers A, and

$$\sum_{S\in\mathcal{G}} \zeta(S) < t\,. \tag{2.2}$$

One advantage of this new definition—important for us—is that it frees the definition of ψ from any reference to a metric. This is particularly useful if one wants to define Hausdorff measure on a manifold.

2.1.1 Hausdorff Measure and Spherical Measure

Hausdorff measure and spherical measure were introduced by Hausdorff in [Hau 18].

Let m be a nonnegative integer and let Ω_m be the m-dimensional volume of the unit ball in Euclidean m-space, that is,

$$\Omega_m = \frac{2\pi^{m/2}}{m\Gamma(m/2)} = \frac{[\Gamma(1/2)]^m}{\Gamma(m/2+1)}\,. \tag{2.3}$$

Now we specialize to the situation in which $\mathcal{F}$ is the collection of *all sets S* and

$$\zeta_1(S) = \Omega_m\, 2^{-m}(\text{diam } S)^m \tag{2.4}$$

for $S \neq \emptyset$. [Note that this definition makes sense for any $m \geq 0$ with Ω_m defined by (2.3), although the interpretation of Ω_m as the volume of a ball is no longer relevant or valid when m is not an integer.]

We call the resulting measure the *m-dimensional Hausdorff measure* on $\mathbb{R}^N$, denoted by $\mathcal{H}^m$. It is worth noting that the same measure would result if we let $\mathcal{F}$ be the collection of all closed sets or all open sets. In fact, because any set and its convex hull have the same diameter, we could restrict attention to convex sets.

It is immediate that the measure $\mathcal{H}^0$ is counting measure (see Example 1.3.24).

Proposition 2.1.2. *For* $0 \leq s < t < \infty$ *and* $A \subseteq \mathbb{R}^N$, *we have that*

(1) $\mathcal{H}^s(A) < \infty$ *implies that* $\mathcal{H}^t(A) = 0$;
(2) $\mathcal{H}^t(A) > 0$ *implies that* $\mathcal{H}^s(A) = \infty$.

Proof. It will be convenient to use $\mathcal{H}^s_\delta$ (respectively, $\mathcal{H}^t_\delta$) to denote the preliminary measure ϕ_δ constructed using the gauge ζ_1 in (2.4) with $m = s$ (respectively, $m = t$).

For (1), let $A \subseteq \bigcup_i E_i$, with diam $(E_i) \leq \delta$ and

$$\Omega_s\, 2^{-s} \sum_i \operatorname{diam}(E_i)^s \le \mathcal{H}^s_\delta(A) + 1\,.$$

Then

$$\begin{aligned}\mathcal{H}^t_\delta(A) &\le \Omega_t\, 2^{-t} \sum_i \operatorname{diam}(E_i)^t \\ &\le \delta^{t-s}\, \Omega_t\, 2^{-t} \sum_i \operatorname{diam}(E_i)^s \le \delta^{t-s}\, (\Omega_t/\Omega_s)\, 2^{s-t} \left(\mathcal{H}^s_\delta(A) + 1\right).\end{aligned}$$

As $\delta \to 0^+$, this estimate gives (1).

Statement (2) is really just the contrapositive of (1). But it is worth stating separately, since it is the basis for the theory of Hausdorff dimension. □

When $\mathcal{F}$ is the family of all closed balls in $\mathbb{R}^N$, and ζ_1 is as above, then the resulting measure ψ is called the *m-dimensional spherical measure*. We denote this measure by $\mathcal{S}^m$. The same measure results if we use the family of all open balls.

Of course, it is immediate that

$$\mathcal{H}^m \le \mathcal{S}^m \le 2^m \cdot \mathcal{H}^m\,.$$

More precise comparisons are possible, and we shall explore these in due course.

2.1.2 A Measure Based on Parallelepipeds

Let $M > 0$ be an integer and assume that $M \le N$, the dimension of the Euclidean space $\mathbb{R}^N$. Now suppose we use the new gauge function defined by

$$\zeta_2(S) = \Omega_M \cdot 2^{-M} \cdot \sup\left\{ |(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| : a_1, b_1, \ldots, a_M, b_M \in S \right\}. \tag{2.5}$$

See Figure 2.2. We will learn more about this gauge in Lemma 2.1.3. Carathéodory's construction on the family $\mathcal{F}$ of all nonempty subsets of $\mathbb{R}^N$ will be denoted by $\mathcal{T}^M$

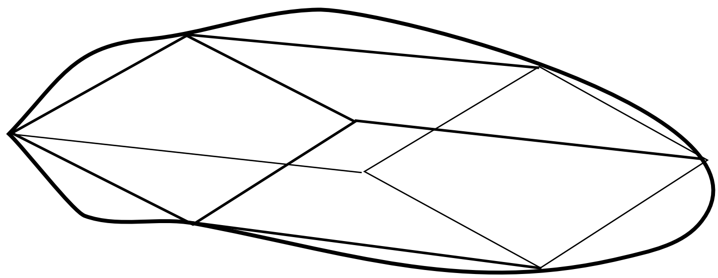

Fig. 2.2. A construction based on exterior algebra.

and will be called *M-dimensional Federer*[1] *measure* on $\mathbb{R}^N$. Of course, we could use all open sets S, or all compact sets S, or all convex sets S; the same measure would result.

Since

$$|(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| \leq \prod_{i=1}^{M} |a_i - b_i| ,$$

we conclude that

$$\zeta_2(S) \leq \Omega_M \cdot 2^{-M} (\text{diam } S)^M$$

and thus that $\mathcal{T}^M \leq \mathcal{H}^M$. Observe that the gauge ζ_2 assigns the same value to any set and to its convex hull. This follows because the map of $(\mathbb{R}^N)^{2M}$ into $\bigwedge_M(\mathbb{R}^N)$ yielding the preceding exterior product is affine with respect to each of the $2M$ variables $a_1, b_1, \ldots, a_M, b_M$.

2.1.3 Projections and Convexity

Continue to assume that $M > 0$ is an integer with $M \leq N$, the dimension of the Euclidean space $\mathbb{R}^N$. We let $\mathbf{O}(N, M)$ denote the collection of orthogonal injections of $\mathbb{R}^M$ into $\mathbb{R}^N$, so each element of $\mathbf{O}(N, M)$ is a linear map from $\mathbb{R}^M$ to $\mathbb{R}^N$ that is represented by an $N \times M$ matrix with orthonormal columns. In case $M = N$, we write $\mathbf{O}(M) = \mathbf{O}(M, M)$, so that $\mathbf{O}(M)$ is the orthogonal group. Furthermore, $\mathbf{O}^*(N, M)$ will be the set of adjoints of elements of $\mathbf{O}(N, M)$ from $\mathbb{R}^N$ onto $\mathbb{R}^M$ (these are of course orthogonal projections). For $S \subseteq \mathbb{R}^N$, we set

$$\zeta_3(S) = \sup \left\{ \mathcal{L}^M[p(S)] : p \in \mathbf{O}^*(N, M) \right\} , \tag{2.6}$$

where $\mathcal{L}^M$ is the M-dimensional Lebesgue measure.

Gross Measure

Let $\mathcal{F}$ be the family of all Borel subsets of $\mathbb{R}^N$. Then Carathéodory's construction, with ζ_3 as in (2.6), gives the *M-dimensional Gross measure*[2] on $\mathbb{R}^N$. It is denoted by $\mathcal{G}^M$.

Carathéodory Measure

Let $\mathcal{F}$ be the family of all *open, convex subsets of* $\mathbb{R}^N$. Then Carathéodory's construction, with ζ_3 as in (2.6), gives the *M-dimensional Carathéodory measure*[3] on $\mathbb{R}^N$. We denote this measure by $\mathcal{C}^M$. The family of all closed, convex subsets gives rise to just the same measure.

It is worth noting that, when $M = 1$, then $\zeta_3(S) = \text{diam}\,(S)$ when S is convex and hence

$$\mathcal{C}^1 = \mathcal{H}^1 .$$

[1] This measure was introduced by H. Federer in [Fed 69].

[2] Introduced in [Gro 18a] and [Gro 18b].

[3] Introduced in [Car 14].

2.1.4 Other Geometric Measures

Fix $\mathbb{R}^N$ as usual and select a positive integer M such that $M \le N$. For $1 \le t \le \infty$, we now proceed to define a gauge function $\zeta_{4,t}$:

For $S \subseteq \mathbb{R}^N$, define $f_S : \mathbf{O}^*(N, M) \to \overline{\mathbb{R}}$ by setting

$$f_S(p) = \mathcal{L}^M[p(S)] \qquad \text{for all } p \in \mathbf{O}^*(N, M)\,.$$

Let $\theta^*_{N,M}$ be the measure on $\mathbf{O}^*(N, M)$, with total measure 1, that is invariant under the action of the orthogonal group. (We will prove the existence of such a measure in Chapter 3, where our arguments are independent of this chapter.) To ensure that the measures resulting from Carathéodory's construction using the gauge $\zeta_{4,t}$ give values that agree with those found for smooth surfaces using calculus, we need to introduce a normalizing factor $\beta_t(N, M)$. For completeness, we give the definition here. For $1 \le t < \infty$, let $\beta_t(N, M)$ be the positive number that satisfies the equation

$$\left(\int |(\wedge_M p)\xi|^t \, d\theta^*_{N,M}(p)\right)^{1/t} = \beta_t(N, M) \cdot |\xi|$$

for any simple M-vector ξ of $\mathbb{R}^N$. Set $\beta_\infty(N, M) = 1$. Finally, set

$$\zeta_{4,t}(S) = \Big(\beta_t(N, M)\Big)^{-1} \left(\int \Big|f_S(p)\Big|^t \, d\theta^*_{N,M}(p)\right)^{1/t} \tag{2.7}$$

whenever $f_S(p) = \mathcal{L}^M[p(S)]$ is $\theta^*_{N,M}$-measurable.

In fact, f_S is $\theta^*_{N,M}$-measurable whenever S is a Borel or Suslin set. This measurability holds because

$$\{\, (x, y, p) : x \in S,\ y = p(x) \,\}$$

is a Suslin set in $\mathbb{R}^N \times \mathbb{R}^M \times \mathbf{O}^*(N, M)$ whenever S is a Borel or Suslin set in $\mathbb{R}^N$.

The map

$$t \longmapsto \beta_t(N, M)\zeta_{4,t}(S) \tag{2.8}$$

sends t to the L^t-norm of a fixed function on a space with total measure 1, so, using Hölder's inequality and Lebesgue's convergence theorems, we see that the map (2.8) is nondecreasing and continuous; thus $\zeta_{4,t}(S)$ is continuous as a function of t.

Integral Geometric Measure

Let $\mathcal{F}$ be the family of all Borel subsets of $\mathbb{R}^N$. Using Carathéodory's construction with gauge $\zeta_{4,t}$, we construct the *M-dimensional integral geometric measure with exponent t* on $\mathbb{R}^N$. This measure is denoted by $\mathcal{I}^M_t$. Roughly speaking, integral geometric measure measures all projections of the given set, and then integrates out (using an invariant measure) over all projections. The M-dimensional integral geometric measure with exponent 1 was introduced by Jean Favard (1902–1965) in [Fav 32] and is sometimes called *Favard measure*.

It is worth noting that $\mathcal{I}^M_t(A) = 0$ if and only if the set A is contained in a Borel set B with $\mathcal{L}^M[p(B)] = 0$ for $\theta^*_{N,M}$-almost every $p \in \mathbf{O}^*_{N,M}$. Thus all the measures $\mathcal{I}^M_t$, $1 \le t \le \infty$, have the same null sets.

Gillespie Measure

Let $\mathcal{F}$ be the family of all open, convex subsets of $\mathbb{R}^N$. The Carathéodory construction with gauge $\zeta_{4,t}$ then gives the measure $\mathcal{Q}_t^M$. We call this measure the *M-dimensional Gillespie*[4] *measure with exponent t* on $\mathbb{R}^N$. The same measure results when we use instead the family of all closed, convex subsets of $\mathbb{R}^N$.

Since the function f_S is continuous for any bounded, open, convex set S, we see that $\mathcal{Q}_\infty^M = \mathcal{C}^M$.

2.1.5 Summary

In Table 2.1, we summarize the measures, and their constructions, that have been described in this section.

To establish the relationships among the measures listed in Table 2.1, we will need to understand ζ_2 better.

Lemma 2.1.3. *If $S \subseteq \mathbb{R}^M$ is a nonempty subset, then*

$$\mathcal{L}^M(S) \leq \Omega_M \cdot 2^{-M} \cdot \sup\{|(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| : a_1, b_1, \ldots, a_M, b_M \in S\}\,.$$

Proof. Let $M = N$ and let $\zeta_2(S)$ be as above. Take $\lambda, \mu > 0$. Define $\mathcal{C}$ to be the collection of all nonempty, compact, convex subsets S of $\mathbb{R}^N$ such that

$$\mathcal{L}^M(S) \geq \lambda \quad \text{and} \quad \zeta_2(S) \leq \mu\,.$$

By the upper semicontinuity of Lebesgue measure with respect to the Hausdorff distance, i.e., Corollary 1.6.5, and by the definition of ζ_2, $\mathcal{C}$ is closed with respect to the Hausdorff metric. We further claim that if the set T is obtained from $S \in \mathcal{C}$ by Steiner symmetrization, then $T \in \mathcal{C}$. To see that this claim holds, recall that Proposition 1.6.12 tells us that Steiner symmetrization preserves Lebesgue measure, while symmetrization also preserves the gauge ζ_2 just by linearity.

Now, in case $\mathcal{C}$ is nonempty, we can conclude from Theorem 1.6.16 that there is some ball $\overline{\mathbb{B}}(0, r)$ in $\mathcal{C}$. Thus

$$\lambda \leq \mathcal{L}^M[\overline{\mathbb{B}}(0, r)] = \Omega_M \cdot r^M = \zeta_2[\overline{\mathbb{B}}(0, r)] \leq \mu\,.$$

That proves our result. □

Corollary 2.1.4. *For $S \subseteq \mathbb{R}^N$, it holds that*

$$\zeta_3(S) \leq \zeta_2(S)\,.$$

[4] David Clinton Gillespie (1879–1935) suggested the measure $\mathcal{Q}_1^M$ to Anthony Perry Morse (1911–1984) and John A. F. Randolph (see [MR 40]).

Table 2.1. Measures resulting from Carathéodory's construction.

Gauges

$\boldsymbol{m \in \mathbb{R}, \quad 0 \leq m < \infty}$

$\zeta_1(S) = \Omega_m\, 2^{-m} (\operatorname{diam}\ S)^m$

$\boldsymbol{M \in \mathbb{Z}, \quad 1 \leq M \leq N,}$

$\zeta_2(S) = \Omega_M \cdot 2^{-M} \cdot \sup\{|(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| : a_1, \ldots, b_M \in S\},$

$\zeta_3(S) = \sup\{\mathcal{L}^M[p(S)] : p \in \mathbf{O}^*(N, M)\},$

$\zeta_{4,t}(S) = (\beta_t(N, M))^{-1}\, \|\mathcal{L}^M[p(S)]\|_t.$

Notation	Name of Measure	Family of Sets $\mathcal{F}$	Gauge
$\mathcal{H}^m$	Hausdorff	all sets	ζ_1
$\mathcal{S}^m$	spherical	balls	ζ_1
$\mathcal{T}^M$	Federer	all sets	ζ_2
$\mathcal{G}^M$	Gross	Borel sets	ζ_3
$\mathcal{C}^M$	Carathéodory	open, convex sets	ζ_3
$\mathcal{I}_1^M$	Favard	Borel sets	$\zeta_{4,1}$
$\mathcal{I}_t^M$	integral geometric with exponent t	Borel sets	$\zeta_{4,t}$
$\mathcal{Q}_t^M$	Gillespie with exponent t	open, convex sets	$\zeta_{4,t}$

Proof. For $p \in \mathbf{O}^*(N, M)$, we have

$$|p(a_1 - b_1) \wedge \cdots \wedge p(a_M - b_M)| \leq |(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)|\,,$$

so, by Lemma 2.1.3,

$$\begin{aligned}
&\mathcal{L}^M[p(S)] \\
&\quad \leq \Omega_M \cdot 2^{-M} \cdot \sup\{|(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| : a_1, b_1, \ldots, a_M, b_M \in p(S)\} \\
&\quad \leq \Omega_M \cdot 2^{-M} \cdot \sup\{|p(a_1 - b_1) \wedge \cdots \wedge p(a_M - b_M)| : a_1, b_1, \ldots, a_M, b_M \in S\} \\
&\quad \leq \Omega_M \cdot 2^{-M} \cdot \sup\{|(a_1 - b_1) \wedge \cdots \wedge (a_M - b_M)| : a_1, b_1, \ldots, a_M, b_M \in S\} \\
&\quad = \zeta_2(S)
\end{aligned}$$

holds. Taking the supremum over $p \in \mathbf{O}^*(N, M)$, we obtain the result. □

The following six facts will allow us to compare the measures we have created using Carathéodory's construction.

(1) making the family of sets $\mathcal{F}$ smaller cannot decrease the measure resulting from Carathéodory's construction,
(2) $\zeta_2 \le \zeta_1$,
(3) $\zeta_3 \le \zeta_2$,
(4) $\beta_t(N, m)\, \zeta_{4,t}(S)$ is a nondecreasing function of t,
(5) $\zeta_{4,\infty} \le \zeta_3$, and
(6) ζ_3 and $\zeta_{4,\infty}$ agree on the open, convex sets.

Proposition 2.1.5. *For M an integer with $1 \le M \le N$ and for $\infty \ge t \ge s \ge 1$, the following relationships hold:*

$$\begin{array}{ccccccccc}
\mathcal{S}^M \ge \mathcal{H}^M \ge & \mathcal{T}^M & & & & & & \\
& \vee\!| & & & & & & \\
& \mathcal{C}^M & = & \mathcal{Q}^M_\infty & \ge & \beta_t(N,M)\cdot\mathcal{Q}^M_t & \ge & \beta_s(N,M)\cdot\mathcal{Q}^M_s \\
& \vee\!| & & \vee\!| & & \vee\!| & & \vee\!| \\
& \mathcal{G}^M & \ge & \mathcal{I}^M_\infty & \ge & \beta_t(N,M)\cdot\mathcal{I}^M_t & \ge & \beta_s(N,M)\cdot\mathcal{I}^M_s\,.
\end{array}$$

Proof. Use the six facts above. □

Noting that $\beta_t(N, N) = 1$ for $1 \le t \le \infty$, we see that when $N = M$, $\mathcal{I}^N_1$ is the smallest of the measures that we have defined in this section. Also note that the inequality

$$\mathcal{I}^N_1(A) \ge \mathcal{L}^N(A), \quad \text{for all } A \subseteq \mathbb{R}^N\,, \tag{2.9}$$

is evident from the definition of $\mathcal{I}^N_1$. Ultimately (see Corollary 4.3.9) we will show that in $\mathbb{R}^N$, the measures $\mathcal{S}^N$, $\mathcal{H}^N$, $\mathcal{T}^N$, $\mathcal{C}^N$, $\mathcal{G}^N$, $\mathcal{Q}^N_t$, and $\mathcal{I}^N_t$ $(1 \le t \le \infty)$ all agree with the N-dimensional Lebesgue measure $\mathcal{L}^N$.

2.2 The Densities of a Measure

At a point p of a smooth m-dimensional surface S in $\mathbb{R}^N$, we know that the m-dimensional area of $S \cap \overline{\mathbb{B}}(p, r)$ approaches 0 like r^m as $r \downarrow 0$. We might hope to generalize that observation to less smooth surfaces and more general measures, or we might wish to show that if some measure behaves in that way on a set S, then that set exhibits some other desirable behavior. The tools for such investigations are the densities of a measure, which we define next.

Definition 2.2.1. Let μ be a measure on $\mathbb{R}^N$. Fix a point $p \in \mathbb{R}^N$ and fix $0 \le m < \infty$ (m need not be an integer).

(1) The *m-dimensional upper density of μ at p* is denoted by $\Theta^{*m}(\mu, p)$ and is defined by setting

$$\Theta^{*m}(\mu, p) = \limsup_{r \downarrow 0} \frac{\mu\left[\overline{\mathbb{B}}(p, r)\right]}{\Omega_m\, r^m} .$$

(2) Similarly, the *m-dimensional lower density of μ at p* is denoted by $\Theta^m_*(\mu, p)$ and is defined by setting

$$\Theta^m_*(\mu, p) = \liminf_{r \downarrow 0} \frac{\mu\left[\overline{\mathbb{B}}(p, r)\right]}{\Omega_m\, r^m} .$$

(3) In case $\Theta^m_*(\mu, p) = \Theta^{*m}(\mu, p)$, we call their common value the *m-dimensional density of μ at p* and denote it by $\Theta^m(\mu, p)$.

Because Hausdorff measure and spherical measure are based on diameters of sets and balls, respectively, a bound on the upper density of a measure μ should imply a relationship between μ and Hausdorff measure and between μ and spherical measure. To obtain such results, we need to require the measure μ to be regular. Recall that Lemma 1.2.8 tells us that for a regular measure, the measure of the union of an increasing sequence of sets equals the limit of their measures.

Proposition 2.2.2. *Let μ be a regular measure on $\mathbb{R}^N$, and let $0 \le t < \infty$ be fixed. If $\mathcal{H}^m(A) < \infty$ and $\Theta^{*m}(\mu, p) \le t$ holds for all $p \in A$, then*

$$\mu(A) \le t \cdot 2^m \cdot \mathcal{H}^m(A) \le t \cdot 2^m \cdot \mathcal{S}^m(A) .$$

Proof. Since $\mathcal{H}^m \le \mathcal{S}^m$, we need only consider the Hausdorff measure.

Let s with $t < s < \infty$ be arbitrary. For each positive integer j, set

$$A_j = A \cap \left\{ p : \mu\left[\overline{\mathbb{B}}(p, r)\right] \le s \cdot \Omega_m\, r^m, \text{ for all } r \le 1/j \right\} .$$

By Lemma 1.2.8, the fact that $\mathcal{H}^m(A_j) < \infty$, and the arbitrariness of s, it suffices to prove that

$$\mu(A_j) \le 2^m \cdot s \cdot \mathcal{H}^m(A_j) \tag{2.10}$$

holds for each j.

Now let δ satisfy $0 < \delta \le 1/j$. Let $S_1, S_2, \ldots$ be a family of sets of diameter not exceeding δ such that $A_j \subseteq \cup_{i=1}^{\infty} S_i$. Without loss of generality, we may assume that each S_i intersects A_j in a point P_i. We conclude that

$$\mu(A_j) \le \sum_{i=1}^{\infty} \mu(S_i) \le \sum_{i=1}^{\infty} \mu\left[\overline{\mathbb{B}}(p_i,\, \operatorname{diam} S_i)\right]$$

$$\le \sum_{i=1}^{\infty} s\, \Omega_m\, (\operatorname{diam} S_i)^m \le 2^m\, s\, \sum_{i=1}^{\infty} \zeta_1(S_i)$$

holds, where $\zeta_1(S)$ is the gauge function

$$\zeta_1(S) = \Omega_m\, 2^{-m}\, (\operatorname{diam} S)^m\,.$$

Since the countable covering $\{S_i\}$ by sets with diameter not exceeding δ was otherwise arbitrary, we conclude that

$$\mu(A_j) \le 2^m \cdot s \cdot \phi_\delta(A_j)\,.$$

Letting $\delta \downarrow 0$, we obtain (2.10). □

Definition 2.2.3. If μ is a measure on the nonempty set X and $A \subseteq X$ is any set, define the measure $\mu \llcorner A$ on X by setting

$$(\mu \llcorner A)(E) = \mu(A \cap E)$$

for each $E \subseteq X$. It is easy to check that $\mu \llcorner A$ is, in fact, a measure, and it is also easy to check that any set that is μ-measurable is also $\mu \llcorner A$-measurable. We call $\mu \llcorner A$ the *restriction of* μ *to* A.

Corollary 2.2.4. *Fix* $0 \le t < 2^{-m}$. *If* $A \subseteq \mathbb{R}^N$ *with* $\mathcal{H}^m(A) < \infty$ *and if* $\Theta^{*m}(\mathcal{H}^m \llcorner A, p) < t$ *holds for each* $p \in A$, *then* $\mathcal{H}^m(A) = 0$.

Proof. Argue by contradiction. Assume $\mathcal{H}^m(A) > 0$ and apply Proposition 2.2.2 to the measure $\mu = \mathcal{H}^m \llcorner A$ on the set A. □

Remark 2.2.5. In fact the conclusion of Corollary 2.2.4 remains true even without the hypothesis $\mathcal{H}^m(A) < \infty$ as long as A is assumed to be a Suslin set. Obtaining this generalization requires the next result, which we shall not prove here.

Theorem 2.2.6 ([Bes 52]). *If* A *is a compact subset of* $\mathbb{R}^N$ *with* $\mathcal{H}^m(A) = \infty$, *then there is a compact set* B *with* $B \subseteq A$ *and* $0 < \mathcal{H}^m(B) < \infty$.

2.3 A One-Dimensional Example

Suppose $g : \mathbb{R} \to \mathbb{R}$ is nondecreasing. Let $\mathcal{F}$ be the family of all nonempty, bounded open subintervals of $\mathbb{R}$. Define the gauge

$$\zeta\left(\{t \in \mathbb{R} : a < t < b\}\right) = g(b) - b(a) \tag{2.11}$$

whenever $-\infty < a < b < \infty$. Now applying Carathéodory's construction produces a measure ψ that we will investigate.

Lemma 2.3.1. *If* g *is continuous at* a *and* b, *then*

$$\psi\left\{t \in \mathbb{R} : a < t < b\right\} = g(b) - g(a)\,.$$

Proof. First we observe that, using the gauge in (2.11), all the measures ϕ_δ, for $0 < \delta < \infty$, in Carathéodory's construction are equal. This is because if g is continuous at points $t_1 < t_2 < \cdots < t_{N+1}$ then

$$g(t_{N+1}) - g(t_1) = \lim_{\epsilon \to 0^+} \sum_{j=1}^{n} [g(t_{j+1} + \epsilon) - g(t_j - \epsilon)] ,$$

which allows us to replace any particular interval by shorter intervals. From the equality of all the approximating measures ϕ_δ, we conclude that $\psi(\{t \in \mathbb{R} : a < t < b\}) \leq g(b) - g(a)$.

To obtain the opposite inequality, notice that if $\mathcal{G}$ is any countable family of open intervals covering the interval (a, b), and if $\epsilon > 0$, then $\{t \in \mathbb{R} : a + \epsilon \leq t \leq b - \epsilon\}$ is covered by some finite subfamily of $\mathcal{G}$. Call this subcovering $(u_1, v_1), (u_2, v_2), \ldots, (u_k, v_k)$. Thus

$$\sum_{j=1}^{k} [g(v_j) - g(u_j)] \geq g(b - \epsilon) - g(a + \epsilon) ,$$

and that proves the result. $\square$

The measure ψ is the measure associated with Riemann–Stieltjes integration[5] with respect to g. See [Rud 76, Chapter 6] or [Fed 69, 2.5.17] for more on the Riemann–Stieltjes integral.

Example 2.3.2. In the special case that $g(x) = x$, the gauge ζ defined in (2.11) agrees with the gauge ζ_1 used to define Hausdorff measure (or spherical measure) on $\mathbb{R}$, so that $\psi = \mathcal{H}^1 = \mathcal{S}^1$. The lemma tells us that $\mathcal{H}^1$ and $\mathcal{S}^1$ assign the same measure to any open interval as does $\mathcal{L}^1$. We conclude that, on $\mathbb{R}$, $\mathcal{L}^1 = \mathcal{H}^1 = \mathcal{S}^1$. $\square$

2.4 Carathéodory's Construction and Mappings

Carathéodory's construction is complicated enough that it is often a challenge to compute values of the resulting measure. For this reason, the next proposition is of considerable utility.

First recall that a *partition* of a set A is a collection $\mathcal{P}$ of pairwise disjoint subsets of A whose union equals A; that is,

$$P_1 \cap P_2 = \emptyset \text{ if } P_1, P_2 \in \mathcal{P} \text{ with } P_1 \neq P_2 ,$$

$$A = \bigcup_{P \in \mathcal{P}} P .$$

---

5 Thomas Jan Stieltjes (1856–1894).

Proposition 2.4.1. *Let ψ be the result of applying Carathéodory's construction to the family $\mathcal{F}$ using a gauge function ζ. Suppose that every element of $\mathcal{F}$ is a Borel set, and suppose that the gauge function satisfies the sub-additivity condition*

$$\zeta(A) \le \sum_{B \in \mathcal{G}} \zeta(B) \tag{2.12}$$

whenever $\mathcal{G}$ is a countable subfamily of $\mathcal{F}$ with $A \subseteq \bigcup_{B \in \mathcal{G}} B$.

If $A \subseteq \mathbb{R}^N$ is any set in $\mathcal{F}$, then we have

$$\psi(A) = \sup \left\{ \sum_{B \in \mathcal{H}} \zeta(B) : \mathcal{H} \text{ is an } \mathcal{F}\text{-partition of } A \right\} .$$

Furthermore, if $\mathcal{H}_1, \mathcal{H}_2, \dots$ are $\mathcal{F}$-partitions of A, then

$$\limsup_{j \to \infty} \{\text{diam } B : B \in \mathcal{H}_j\} = 0 \quad \textit{implies} \quad \lim_{j \to \infty} \sum_{B \in \mathcal{H}_j} \zeta(B) = \psi(A) .$$

Proof. Of course, $\zeta(S) \le \psi(S)$ holds for every set $S \in \mathcal{F}$. Since any $S \in \mathcal{F}$ is a Borel set and any Borel set is ψ-measurable, every $S \in \mathcal{F}$ is ψ-measurable. It follows that

$$\sum_{B \in \mathcal{H}} \zeta(B) \le \sum_{B \in \mathcal{H}} \psi(B) = \psi(A)$$

whenever $\mathcal{H}$ is an $\mathcal{F}$-partition of A.

If the diameters of the members of the partitions $\mathcal{H}_j$ of A approach 0 as $j \to \infty$, then we also have

$$\psi(A) \le \liminf_{j \to \infty} \sum_{B \in \mathcal{H}_j} \zeta(B) \le \liminf_{j \to \infty} \sum_{B \in \mathcal{H}_j} \psi(B) . \qquad \square$$

Proposition 2.4.1 can be applied to the construction of $\mathcal{G}^m$ and $\mathcal{I}_t^m$. One concludes that

$$\mathcal{I}_t^m = \lim_{s \to t^-} \mathcal{I}_s^m \quad \text{for } 1 \le t \le \infty .$$

The theorem cannot be applied to $\mathcal{H}^m$, $\mathcal{S}^m$, $\mathcal{T}^m$, or $\mathcal{Q}_t^m$. For instance, there is no hope of ζ_1 satisfying (2.12), since in general, diam $(A \cup B)$ is in no way bounded by the two numbers diam A and diam B.

Now we introduce the notion of the multiplicity of a mapping.

Definition 2.4.2. Suppose that $f : X \to Y$. We let $N(f, y)$ denote the number of elements of $f^{-1}(\{y\})$. More precisely, for $y \in Y$, we set

$$N(f, y) = \begin{cases} \text{card}\{x \in X : f(x) = y\} & \text{if } \{x \in X : f(x) = y\} \text{ is finite,} \\ \infty & \text{otherwise.} \end{cases}$$

We call $N(f, y)$ the *multiplicity* of f at y.

Proposition 2.4.3. *Let μ be a measure on $\mathbb{R}^N$, let $f : \mathbb{R}^M \to \mathbb{R}^N$, and let $\mathcal{F}$ be the family of Borel subsets of $\mathbb{R}^M$. Assume that $f(A)$ is μ-measurable whenever $A \in \mathcal{F}$. If we set*

$$\zeta(S) = \mu[f(S)] \quad \text{for } S \subseteq X\,,$$

and if ψ is the result of Carathéodory's construction on $\mathbb{R}^M$ using the gauge ζ on the family $\mathcal{F}$, then

$$\psi(A) = \int N(f|A, y)\, d\mu(y) \ \text{ for every } A \in \mathcal{F}.$$

Proof. Let $\mathcal{H}_1, \mathcal{H}_2, \ldots$ be Borel partitions of A such that each member of $\mathcal{H}_j$ is the union of some subfamily of $\mathcal{H}_{j+1}$ and

$$\sup\{\text{diam } S : S \in \mathcal{H}_j\} \to 0 \quad \text{as } j \to \infty\,.$$

Then

$$\sum_{S \in \mathcal{H}_j} \chi_{f(S)}(y) \uparrow N(f|A, y) \quad \text{as } j \uparrow \infty$$

for each $y \in Y$. Thus the last proposition and the Lebesgue monotone convergence theorem imply that

$$\psi(A) = \lim_{j\to\infty} \sum_{S \in \mathcal{H}_j} \mu[f(S)] = \lim_{j\to\infty} \int \sum_{S \in \mathcal{H}_j} \chi_{f(S)}\, d\mu = \int N(F|A, y)\, d\mu(y)\,. \ \square$$

Definition 2.4.4. Let X and Y be metric spaces with metrics dist_X and dist_Y, respectively. A function $f : X \to Y$ is said to be *Lipschitz*[6] *of order* 1, or simply *Lipschitz*, if there exists $M < \infty$ such that

$$\text{dist}_Y[f(x_1), f(x_2)] \leq M\, \text{dist}_X[x, y] \tag{2.13}$$

holds, for all $x_1, x_2 \in X$. The least choice of M that makes (2.13) true is called the *Lipschitz constant for* f and is denoted by $\text{Lip}\, f$.

Corollary 2.4.5. *If f is a Lipschitz mapping of $\mathbb{R}^M$ into $\mathbb{R}^N$, if $0 \leq m < \infty$, and if $A \subseteq \mathbb{R}^M$ is Borel, then*

$$(\text{Lip}\, f)^m \cdot \mathcal{H}^m(A) \geq \int N(f|A, y)\, d\mathcal{H}^m(y)\,.$$

Proof. We apply Proposition 2.4.3 with μ replaced by $\mathcal{H}^m$, so we have $\zeta(S) = \mathcal{H}^m[f(S)]$. It is elementary that

$$\mathcal{H}^m[f(S)] \leq (\text{Lip}\, f)^m \cdot \mathcal{H}^m(S) \qquad \text{for } S \subseteq \mathbb{R}^M\,,$$

and the result follows. $\square$

[6] Rudolf Otto Sigismund Lipschitz (1832–1903).

Now an interesting geometric upshot of this discussion is the following:

Corollary 2.4.6. *If $C \subseteq \mathbb{R}^M$ is connected then*

$$\mathcal{H}^1(C) \geq \text{diam } C\,.$$

Proof. We may of course assume that $\mathcal{H}^1(C) < \infty$. Choose a Borel set $B \supseteq C$ such that $\mathcal{H}^1(B) = \mathcal{H}^1(C)$.

For $a, b \in C$, we define $F : \mathbb{R}^M \to \mathbb{R}$ by setting $F(x) = \text{dist}(a, x)$ for $x \in \mathbb{R}^M$. Then, by the previous corollary and our discussion of Hausdorff measure in one dimension,

$$\mathcal{H}^1(C) = \mathcal{H}^1(B) \geq \int N(F|B, y)\, d\mathcal{H}^1(y) \geq \mathcal{H}^1[F(C)] \geq \text{dist}(a, b)$$

just because $0 = F(a)$ and $F(b)$ belong to the interval $F(C)$. That proves the result. □

In the proof of Corollary 2.4.5 we used the inequality $(\text{Lip } f)^m \cdot \mathcal{H}^m(A) \geq \mathcal{H}^m[f(A)]$, a fact that follows directly from the definition of Hausdorff measure. We now note this fact as a separate proposition.

Proposition 2.4.7. *If f is a Lipschitz mapping of $\mathbb{R}^M$ into $\mathbb{R}^N$, if $0 \leq m < \infty$, and if $A \subseteq \mathbb{R}^M$ is any set, then*

$$(\text{Lip } f)^m \cdot \mathcal{H}^m(A) \geq \mathcal{H}^m[f(A)]\,.$$

2.5 The Concept of Hausdorff Dimension

The concept of Hausdorff dimension relies on the following conclusions of Proposition 2.1.2:

(1) If $\mathcal{H}^m(A) < \infty$ then $\mathcal{H}^k(A) = 0$ for any $m < k < \infty$.
(2) If $\mathcal{H}^m(A) = +\infty$ then $\mathcal{H}^k(A) = +\infty$ for any $0 \leq k < m$.

Definition 2.5.1. The *Hausdorff dimension* of an infinite set $A \subseteq \mathbb{R}^N$ is

$$\begin{aligned}\dim_{\mathcal{H}} A &= \sup\{s : \mathcal{H}^s(A) > 0\} = \sup\{s : \mathcal{H}^s(A) = \infty\} \\ &= \inf\{t : \mathcal{H}^t(A) < \infty\} = \inf\{t : \mathcal{H}^t(A) = 0\}\,.\end{aligned}$$

It is clear that the Hausdorff dimension of a set $A \subseteq \mathbb{R}^N$ is that unique extended real number α with the property that

$$\begin{aligned} s < \alpha &\text{ implies } \mathcal{H}^s(A) = \infty\,, \\ t > \alpha &\text{ implies } \mathcal{H}^t(A) = 0\,. \end{aligned}$$

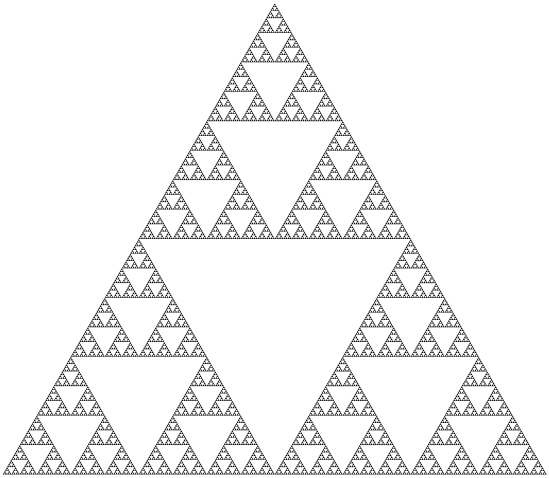

Fig. 2.3. The Sierpiński gasket.

When $\alpha = \dim_{\mathcal{H}} A$, we cannot know anything for sure about $\mathcal{H}^\alpha(A)$. That is to say, the value could be 0 or positive finite or infinity. If, for a given A, we can find an s such that $0 < \mathcal{H}^s(A) < \infty$, then it must be that $s = \dim_{\mathcal{H}} A$. While the Hausdorff dimension of the set A can be an integer, in general this is *not* the case. Figure 2.3 illustrates a classic example [due to Wacław Sierpiński (1882–1969)] of a set with Hausdorff dimension $\log 3/\log 2$.

Clearly the notion of Hausdorff dimension has the properties of monotonicity and stability with respect to countable unions:

$$\dim_{\mathcal{H}} A \le \dim_{\mathcal{H}} B \qquad \text{for } A \subseteq B \subseteq \mathbb{R}^N;$$

$$\dim_{\mathcal{H}} \left(\bigcup_{j=1}^{\infty} A_j \right) = \sup_j \dim_{\mathcal{H}} A_j \qquad \text{for } A_j \subseteq \mathbb{R}^N,\ j = 1, 2, \ldots.$$

It is not difficult to show that $\dim_{\mathcal{H}} \mathbb{R}^N = N$ and that the dimension of a line segment is 1. More generally, the dimension of any compact C^1 curve is 1. For one can use the implicit function theorem to locally flatten the curve, and then the result follows from that for a segment. The dimension of any discrete set is 0.

Sometimes sets have surprising Hausdorff dimensions. Probably the first such surprise was exhibited in [Osg 03] when William Fogg Osgood (1864–1943) published his example of a Jordan arc[7] γ in $\mathbb{R}^2$ that has positive area, hence $\dim_{\mathcal{H}} \gamma = 2$ (see [PS 92] for a generalization to a Jordan arc γ in $\mathbb{R}^N$ with $\dim_{\mathcal{H}} \gamma = N$).

[7] Marie Ennemond Camille Jordan (1838–1922).

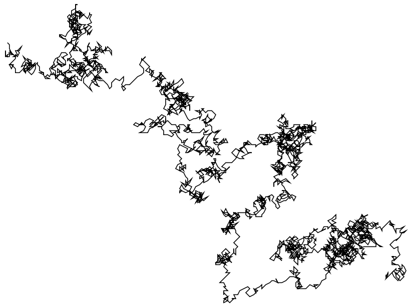

Fig. 2.4. Brownian motion.

A recent result of note is that of Mitsuhiro Shishikura [Shi 98] showing that the boundary of the Mandelbrot set[8] has Hausdorff dimension 2.

We construct the m-dimensional Hausdorff measure by summing mth powers of the diameters of the covering sets. But, in some contexts, it is convenient to apply another function to the diameters. For example, in the study of Brownian motion[9] (see Figure 2.4) it is useful to consider the gauges

$$\zeta(S) = [\text{diam } S]^2 \cdot \log\log[\text{diam } S]^{-1} \qquad \text{in dimension } \geq 3$$

and

$$\zeta(S) = [\text{diam } S]^2 \cdot \log[\text{diam } S]^{-1} \cdot \log\log[\text{diam } S]^{-1} \qquad \text{in dimension } 2\,.$$

It can be shown that the trajectories of Brownian motion have positive and σ-finite measure with respect to the measures that are created from Carathéodory's construction with these gauges ζ.

A planar Brownian path almost surely intersects itself. In the 1980s, Mandelbrot conjectured that the boundary of the set enclosed by the return of a Brownian path is almost surely of Hausdorff dimension 4/3. Recent work of Lawler, Schramm, and Werner in [LSW 01] and [LSW 02] has confirmed the Mandelbrot conjecture.

2.6 Some Cantor Set Examples

In this section, we construct examples of sets of various Hausdorff dimensions. Much of our discussion follows [Mat 95]. Certainly additional examples can be found in Sections 2.10.28, 2.10.29, 3.3.19, and 3.3.20 of [Fed 69].

---

[8] Earlier numerical work by John H. Ewing and Glenn Edward Schober (1938–1991) in [ES 92] had suggested that the boundary of the Mandelbrot set has positive 2-dimensional Lebesgue measure.

[9] Robert Brown (1773–1858).

2.6.1 Basic Examples

Fix a parameter $0 < \lambda < 1/2$. Set $I_0 = [0, 1]$ and let $I_{1,1}$ and $I_{1,2}$ be the intervals $[0, \lambda]$ and $[1 - \lambda, 1]$ respectively. Inductively, if the 2^{k-1} intervals $I_{k-1,1}, I_{k-1,2}, \ldots, I_{k-1,2^{k-1}}$, each having length λ^{k-1}, have been constructed, then we define $I_{k,1}, \ldots, I_{k,2^k}$ by deleting an interval of length $(1 - 2\lambda) \cdot \operatorname{diam}(I_{k-1,j}) = (1 - 2\lambda) \cdot \lambda^{k-1}$ from the middle of each $I_{k-1,j}$. All of the 2^k intervals thus obtained at this kth step have length λ^k, so $\mathcal{H}^1\left[\bigcup_{j=1}^{2^k} I_{k,j}\right] = (2\lambda)^k$.

We may pass to a limit of this construction in the usual "direct limit" or "limsup" manner: We set

$$C(\lambda) = \bigcap_{k=0}^{\infty} \bigcup_{j=1}^{2^k} I_{k,j} \,.$$

See Figure 2.5. Then it is easy to see that $C(\lambda)$ is a compact, nonempty, perfect set and therefore is uncountable. It has no interior and it has Lebesgue measure zero. Every $C(\lambda)$, $0 < \lambda < 1/2$, is a *Cantor set*,[10] and any two are homeomorphic. The most frequently encountered rendition of the sets $C(\lambda)$ is the case $\lambda = 1/3$, which is the *Cantor middle-thirds set*.

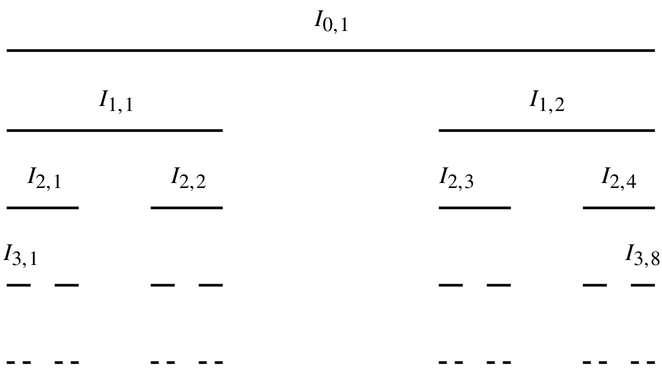

Fig. 2.5. A Cantor set.

It is convenient now to study the Hausdorff measures and dimensions of these Cantor sets. The nature of Carathéodory's construction shows immediately that it is easier to find upper bounds than lower bounds for Hausdorff measure. This is because any particular covering gives an upper bound, but a lower bound requires an estimate over all coverings. Our calculations will bear out this assertion.

We let $\mathcal{H}^m_\delta$ denote the preliminary measure ϕ_δ of (2.1) constructed using the gauge ζ_1 of (2.4); that is,

[10] Georg Ferdinand Ludwig Philipp Cantor (1845–1918).

$$\mathcal{H}^m_\delta(A) = \inf \left\{ \sum_{S \in \mathcal{G}} \Omega_m \, 2^{-m} (\text{diam } S)^m : \right.$$
$$\left. \mathcal{G} \subseteq \{S : \text{diam } S \le \delta\},\ \text{card}(\mathcal{G}) \le \aleph_0,\ \text{and } A \subseteq \bigcup_{S \in \mathcal{G}} S \right\}.$$

To begin, for each $k = 1, 2, \ldots$, we have $C(\lambda) \subseteq \bigcup_j I_{k,j}$, hence

$$\mathcal{H}^m_{\lambda^k}[C(\lambda)] \le \sum_{j=1}^{2^k} \text{diam}\,(I_{k,j})^m = 2^k \lambda^{km} = (2\lambda^m)^k .$$

To make this upper bound truly useful, we would like it to remain uniformly bounded as $k \to +\infty$. Of course, the least value of m for which this occurs is provided by the equation $2\lambda^m = 1$, i.e.,

$$m = \frac{\log 2}{\log(1/\lambda)} .$$

For this choice of m we have

$$\mathcal{H}^m[C(\lambda)] = \lim_{k \to +\infty} \mathcal{H}^m_{\lambda^k}[C(\lambda)] \le 1 .$$

Hence $\dim_{\mathcal{H}} C(\lambda) \le m$.

Our next calculation will show that $\mathcal{H}^m[C(\lambda)] \ge 1/4$. Hence we will be able to conclude that

$$\dim_{\mathcal{H}} C(\lambda) = \frac{\log 2}{\log(1/\lambda)} . \tag{2.14}$$

To prove this new estimate, we need only show that

$$\sum_j \text{diam}\,(I_j)^m \ge \frac{1}{4} \tag{2.15}$$

whenever the I_j are open intervals covering $C(\lambda)$. The set $C(\lambda)$ is compact; hence finitely many of the I_j's cover $C(\lambda)$. Hence we may as well assume from the outset that $C(\lambda)$ is covered by $I_1, \ldots, I_n$.

Since $C(\lambda)$ certainly has no interior, we can suppose (making the I_j slightly larger if necessary) that the endpoints of each I_j lie outside $C(\lambda)$. Then we may select a number $\delta > 0$ such that the Euclidean distance from the set of all endpoints of the I_j to $C(\lambda)$ is at least δ. We select $k > 0$ so large that $\delta > \lambda^k = \text{diam}\,(I_{k,i})$. Thus each interval $I_{k,i}$ is contained in some I_j.

Next we claim that, for any open interval I and any fixed index ℓ, we have the inequality

$$\sum_{I_{\ell,i} \subseteq I} \text{diam}\,(I_{\ell,i})^m \le 4 \cdot \text{diam}\,(I)^m . \tag{2.16}$$

This claim will give (2.15), since

$$4\sum_j \operatorname{diam}(I_j)^m \ge \sum_j \sum_{I_{k,\ell}\subseteq I_j} \operatorname{diam}(I_{k,i})^m \ge \sum_{i=1}^{2^k} \operatorname{diam}(I_{k,i})^m = 1\,.$$

It remains then to prove (2.16).

So suppose that there are some intervals $I_{\ell,i}$ that lie inside I and let n be the least integer for which I contains some $I_{n,i}$. Then $n \le \ell$. Let $I_{n,j_1}, I_{n,j_2}, \ldots, I_{n,j_p}$ be all the nth-generation intervals that have nontrivial intersection with I. Then $p \le 4$, since otherwise, I would contain some $I_{n-1,i}$. Thus

$$\begin{aligned}4\cdot \operatorname{diam}(I)^m &\ge \sum_{s=1}^{p} \operatorname{diam}(I_{n,j_s})^m \\ &= \sum_{s=1}^{p} \sum_{I_{\ell,i}\subseteq I_{n,j_s}} \operatorname{diam}(I_{\ell,i})^m \\ &\ge \sum_{I_{\ell,i}\subseteq I} \operatorname{diam}(I_{\ell,i})^m\,.\end{aligned}$$

That completes the proof. □

It is actually possible, with some refined efforts, to show that

$$\sum \operatorname{diam}(I_j)^m \ge 1\,,$$

which gives the sharper fact that $\mathcal{H}^m[C(\lambda)] = 1$.

It is worth noting the intuitive fact that when λ increases, the size of the deleted holes decreases and therefore the sets $C(\lambda)$ become larger. Corresponding to this intuitive observation, (2.14) shows that $\dim_{\mathcal{H}} C(\lambda)$ increases as λ increases. Also observe that when λ increases from 0 to 1/2, then $\dim_{\mathcal{H}} C(\lambda)$ takes all the values between 0 and 1.

2.6.2 Some Generalized Cantor Sets

In the preceding construction of Cantor sets we always kept constant the ratio of the lengths of intervals in two successive stages of the construction. We are not bound to do so, and we can thus introduce the following variant of the construction.

Let $T = \{\lambda_i\}$ be a sequence of numbers in the interval $(0, 1/2)$. We construct a set $C(T)$ as in the last subsection, but we now take the intervals $I_{k,j}$ to have length $\lambda_k \cdot \operatorname{diam}(I_{k-1,i})$. Then, for each k, we obtain 2^k intervals of length $s_k = \lambda_1 \cdot \lambda_2 \cdots \lambda_k$.

Let $h : [0, \infty) \to [0, \infty)$ be a continuous, increasing function satisfying $h(s_k) = 2^{-k}$. Then, by the argument of the preceding subsection, the measure ψ resulting from Carathéodory's construction using the gauge $\zeta(S) = h(\operatorname{diam} S)$ satisfies

$$\frac{1}{4} \le \psi[C(T)] \le 1\,.$$

We can also run this argument in the converse direction. Beginning with any continuous, increasing function $h : [0, \infty,) \to [0, \infty)$ satisfying $h(0) = 0$ and $h(2r) < 2h(r)$ for $0 < r < \infty$, we inductively select $\lambda_1, \lambda_2, \ldots$ such that $h(s_k) = h(\lambda_1 \cdot \lambda_2 \cdots \lambda_k) = 2^{-k}$ holds. For any such h there is then a compact set $C_h \subseteq \mathbb{R}^1$ such that $0 < \psi_h(C_h) < \infty$, where ψ_h is the measure resulting from Carathéodory's construction using the gauge $\zeta(S) = h(\text{diam } S)$.

Now fix $0 < m \leq 1$. Letting $h(0) = 0$ and $h(r) = r^m \log(1/r)$ for r small, we observe that the condition $h(2r) < 2h(r)$ is satisfied for r small and thus we can find a compact set C_h with $\psi_h(C_h)$ positive and finite. By comparing $r^m \log(1/r)$ to r^m for r small, we conclude that $\mathcal{H}^m(C_h) = 0$, while by comparing $r^m \log(1/r)$ to r^s, $0 \leq s < r$, for r small, we conclude that $\dim_{\mathcal{H}} C_h = m$. On the other hand, choosing $h(r) = r^m / \log(1/r)$ instead (for r small), we see that the condition $h(2r) < 2h(r)$ is again satisfied for r small and we see that C_h has Hausdorff dimension m and is not σ-finite with respect to the measure $\mathcal{H}^m$. In particular, the extreme cases $s = 0$ and $s = 1$ give, respectively, a set of dimension 1 with zero Lebesgue measure and an uncountable set of dimension zero.

2.6.3 Cantor Sets in Higher Dimensions

Of course, Cantor sets can be constructed in dimensions 2 and higher, following the paradigm of the last section. The idea is illustrated in Figure 2.6.

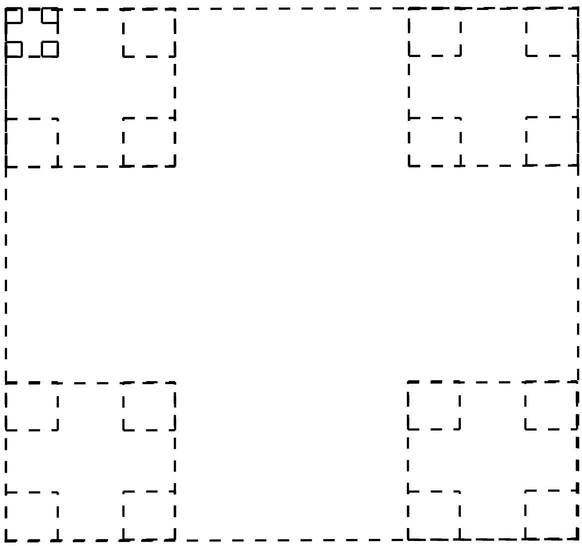

Fig. 2.6. A higher-dimensional Cantor set.

To illustrate the utility of these Cantor sets in constructing examples for Hausdorff dimension, we now describe one result.

Suppose that for $k = 1, 2, \ldots$ we have compact sets $E_{i_1, i_2, \ldots, i_k}$ with $i_j = 1, \ldots, n_j$. assume that

$$E_{i_1, \ldots, i_k, i_{k+1}} \subseteq E_{i_1, \ldots, i_k} \,, \tag{2.17}$$

$$d_k = \max_{i_1, \ldots, i_k} \operatorname{diam}(E_{i_1, \ldots, i_k}) \to 0 \quad \text{as } k \to \infty \,, \tag{2.18}$$

$$\sum_{j=1}^{n_{k+1}} \operatorname{diam}(E_{i_1, \ldots, i_k, j})^m = \operatorname{diam}(E_{i_1, \ldots, i_k})^m \,, \tag{2.19}$$

$$\sum_{B \cap E_{i_1, \ldots, i_k} \neq \emptyset} \operatorname{diam}(E_{i_1, \ldots, i_k})^m \leq c \cdot \operatorname{diam}(B)^m \quad \text{for any ball } B \text{ with } \operatorname{diam}(B) \geq d_k \,, \tag{2.20}$$

where $0 < c < \infty$ is a constant. Define the set

$$\mathcal{E} = \bigcap_{k=1}^{\infty} \bigcup_{i_1, \ldots, i_k} E_{i_1, \ldots, i_k} \,. \tag{2.21}$$

It is immediate from (2.19) that $\mathcal{H}^m(\mathcal{E})$ is finite. To see that $\mathcal{H}^m(\mathcal{E})$ is also positive, suppose that $\mathcal{E}$ is covered by a family of sets of diameter less than δ. We can replace each set in the family by an open ball of slightly more than twice the set's diameter while still covering $\mathcal{E}$. Thus we may suppose that $\mathcal{E}$ is covered by a family of open balls. Since $\mathcal{E}$ is compact, we may suppose the family of open balls is finite. So we have $\mathcal{E} \subseteq \bigcup_{\alpha=1}^{A} U_\alpha$, where each U_α is an open ball. Since, as a function of k, $\bigcup_{i_1, \ldots, i_k} E_{i_1, \ldots, i_k}$ is a decreasing family of compact sets, there is a k_0 such that

$$\bigcup_{i_1, \ldots, i_{k_0}} E_{i_1, \ldots, i_{k_0}} \subseteq \bigcup_{\alpha=1}^{A} U_\alpha \,.$$

Now, using (2.20), we estimate

$$\begin{aligned}
\sum_{\alpha=1}^{A} \operatorname{diam} U_\alpha &\geq c^{-1} \sum_{\alpha=1}^{A} \sum_{U_\alpha \cap E_{i_1, \ldots, i_{k_0}} \neq \emptyset} \operatorname{diam}(E_{i_1, \ldots, i_{k_0}})^m \\
&\geq c^{-1} \sum_{i_1, \ldots, i_{k_0}} \operatorname{diam}(E_{i_1, \ldots, i_{k_0}})^m = c^{-1} \sum_{i_1=1}^{n_1} \operatorname{diam}(E_{i_1})^m \,.
\end{aligned}$$

Thus $\mathcal{H}^m(\mathcal{E})$ is greater than $C \cdot \sum_{i_1=1}^{n_1} \operatorname{diam}(E_{i_1})^m$, where C depends only on c and m.

Example 2.6.1. Let E be the unit ball $\overline{\mathbb{B}}(0, 1) \subseteq \mathbb{R}^2$. Consider the subset $\widetilde{E}$ of E consisting of balls of radius $1/4$ centered at the four points

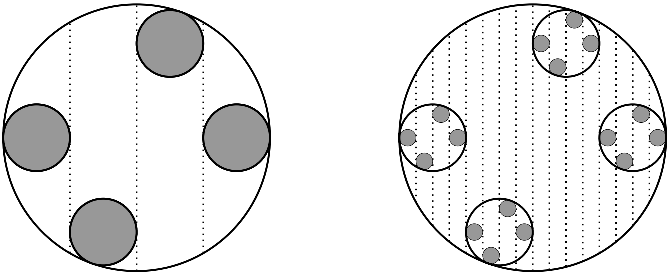

Fig. 2.7. The first two stages in the construction in Example 2.6.1.

$$v_1 = (3/4,\ 0)\,, \quad v_2 = (1/4,\ \sqrt{2}/2)\,,$$
$$v_3 = (-3/4,\ 0)\,,\ v_4 = (-1/4,\ -\sqrt{2}/2)\,.$$

We want to recursively define sets of closed balls by starting with $\widetilde{E}$ and at each stage of the construction replacing each ball with a scaled copy of $\widetilde{E}$ (see Figure 2.7). More precisely, for $k = 1, 2, \ldots$ and $i_j \in \{1, 2, 3, 4\}$, for $j = 1, 2, \ldots, k$, set

$$p_{i_1,i_2,\ldots,i_k} = \sum_{j=1}^{k} (1/4)^{j-1}\, v_{i_j}\,, \qquad E_{i_1,i_2,\ldots,i_k} = \overline{\mathbb{B}}\Big[\, p_{i_1,i_2,\ldots,i_k},\ (1/4)^k \,\Big]\,.$$

These sets satisfy (2.17)–(2.20) with $d_k = 2(1/4)^k$, $m = 1$, and $c = 4$. With $\mathcal{E}$ defined as in (2.21), we conclude that $0 < \mathcal{H}^1(\mathcal{E}) < \infty$, so $\mathcal{E}$ is of Hausdorff dimension 1.

The set $\mathcal{E}$ that we have constructed projects orthogonally onto the full interval $[-1, 1]$ on the x-axis. An interesting property of this set is that there are two mutually orthogonal lines (specifically, with slopes $1/\sqrt{2}$ and $-\sqrt{2}$) onto which $\mathcal{E}$ orthogonally projects to a set of Hausdorff dimension $1/2$. $\square$

There is an extensive literature of self-similar sets and their Hausdorff measures and dimensions. We refer the reader to [Mat 95] and [Rog 98] for further particulars on this topic.

References for additional interesting and instructive sets can be found in Sections 2.10.6 and 3.3.21 of [Fed 69].

3

Invariant Measures and the Construction of Haar Measure

The N-dimensional Lebesgue measure $\mathcal{L}^N$, the most commonly used measure on $\mathbb{R}^N$, has the property that $\mathcal{L}^N(A) = \mathcal{L}^N(b + A)$ for any set A and translation by any element $b \in \mathbb{R}^N$. In fact this translation invariance essentially characterizes Lebesgue measure on $\mathbb{R}^N$. However, consider instead the space $\mathbb{R}^+ \equiv \{x \in \mathbb{R} : x > 0\}$ with the group operation being multiplication (instead of addition). Now what is the invariant measure?

In fact, the reader may verify that the measure dx/x is invariant under the group action. Indeed, if A is a measurable set and $b \in \mathbb{R}^+$, then

$$\int_{\mathbb{R}^+} \chi_A(x \cdot b)\, \frac{dx}{x} = \int_{\mathbb{R}^+} \chi_A(x)\, \frac{dx}{x} .$$

More generally, one may ask, "Is it possible to find an invariant measure on any topological group?" By a topological group we mean a topological space that also comes equipped with a binary operation that induces a group structure on the underlying set. We require that the group operations (product and inverse) be continuous in the given topology. Examples of topological groups are

(1) $(\mathbb{R}^N, +)$, N-dimensional Euclidean space under the operation of vector addition,
(2) $(\mathbb{T}, \cdot)$, the *circle group* consisting of the complex numbers with modulus 1 under the operation of complex multiplication,
(3) $(\mathbf{O}(N), \cdot)$, the *orthogonal group* consisting of the orthogonal transformations of $\mathbb{R}^N$ under the operation of composition or, equivalently, consisting of the $N \times N$ orthogonal matrices under the operation of matrix multiplication,
(4) $(\mathbf{SO}(N), \cdot)$, the *special orthogonal group* consisting of the orientation-preserving orthogonal transformations of $\mathbb{R}^N$ under the operation of composition or, equivalently, consisting of the $N \times N$ orthogonal matrices with determinant 1 under the operation of matrix multiplication.

While an invariant measure, called *Haar measure*,[1] exists on any locally compact group, we shall concentrate our efforts in the present chapter on *compact groups*.

[1] Alfréd Haar (1885–1933).

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_3,

One advantage of compact groups is that the left-invariant Haar measure and the right-invariant Haar measure are identical. For our purposes, the study of compact groups will suffice.

3.1 The Fundamental Theorem

The basic theorem about the existence and uniqueness of Haar measure is as follows. We first enunciate a result about invariant *integrals*. Of course, an integral can be thought of as a linear functional on the *continuous functions*. Then we use a simple limiting argument to extend this functional from continuous functions to characteristic functions (see the corollary). Figure 3.1 illustrates the process of using translates of the graph of a function to approximate the characteristic function of a set.

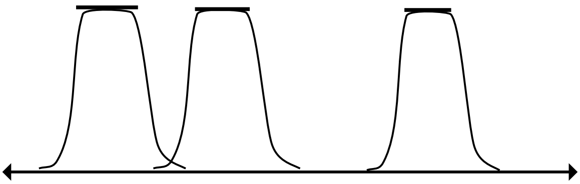

Fig. 3.1. Constructing Haar measure.

Theorem 3.1.1. *Let G be a compact topological group. There is a unique invariant integral λ on G such that $\lambda(1) = 1$.*

Remark 3.1.2. Specifically, the theorem requires that λ be a *monotone* (or *positive*) *Daniell integral*,[2] that is, a linear functional on the continuous functions such that for continuous f, g, and f_n, $n = 1, 2, \ldots$, $f \leq g$ implies $\lambda(f) \leq \lambda(g)$ and $f_n \uparrow f$ implies $\lambda(f_n) \uparrow \lambda(f)$ (see [Fed 69, 2.5] or [Roy 88, Chapter 16]). The invariance of λ means that if φ is a continuous function on G, if $g \in G$, and if $\varphi_g(x) \equiv \varphi(gx)$, then

$$\lambda(\varphi) = \lambda(\varphi_g)\,.$$

Corollary 3.1.3. *Let G be a compact topological group. There is a unique invariant Radon measure μ on G such that $\mu(G) = 1$. The invariance of μ means that for all sets $A \subseteq G$ and $g \in G$,*

$$\mu(A) = \mu\{ga : a \in A\} = \mu\{ag^{-1} : a \in A\}\,.$$

Proof of the Theorem. Let $C(G)$ denote the continuous functions on G, and let $C(G)^+$ denote the nonnegative continuous functions. If $h \in G$ then let A_h denote the operator of left multiplication by h. If $u \in C(G)^+$ and $0 \neq v \in C(G)^+$, then let $W(u, v)$ be the set of all maps

[2] Percy John Daniell (1889–1946).

$$\xi : G \to \{t \in \mathbb{R} : 0 \le t < \infty\}$$

for which

$$\{g \in G : \xi(g) > 0\} \text{ is finite}$$

and

$$u(x) \le \sum_{g \in G} \xi(g) \cdot (v \circ A_g)(x) = \sum_{g \in G} \xi(g) \cdot v(gx) .$$

Now define the *Haar ratio*

$$(u : v) \equiv \inf \left\{ \sum_G \xi \;\middle|\; \xi \in W(u, v) \right\}$$

(here we have altered our set-builder notation to avoid using the colon for two distinct purposes in the same line). Clearly $W(u, v) \ne \emptyset$ and

$$(u : v) \ge [\sup_{x \in G} u(x)]/[\sup_{x \in G} v(x)] .$$

Also we have

$$(u \circ A_h : v) = (u : v) \text{ for } h \in G ;$$

$$(cu : v) = c(u : v) \text{ for } 0 < c < \infty ;$$

$$(u_1 + u_2 : v) \le (u_1 : v) + (u_2 : v) ;$$

$$u_1 \le u_2 \text{ implies } (u_1 : v) \le (u_2 : v) .$$

If $u, v, w \in C(G)^+$ are all nonzero, then

$$(u : w) \le (u : v) \cdot (v : w)$$

just because $\xi \in W(u, v)$ and $\eta \in W(v, w)$ imply

$$u \le \sum_{g \in G} \xi(g) \cdot \sum_{h \in G} \eta(h) \cdot (w \circ A_h \circ A_g) = \sum_{k \in G} (w \circ A_k) ,$$

with $\zeta(k) = \sum_{hg=k} \xi(g) \cdot \eta(h)$ and $\sum_G \zeta = \sum_G \xi \cdot \sum_G \eta$. As a result,

$$\frac{1}{(w : u)} \le \frac{(u : v)}{(w : v)} \le (u : w) . \tag{3.1}$$

Now fix a $0 \ne w \in C(G)^+$ and consider the Cartesian product P of the compact intervals

$$\left\{ t \in \mathbb{R} \;\middle|\; 0 \le t \le (u : w) \right\}$$

corresponding to all $u \in C(G)^+$ (again the set-builder notation has been modified to avoid confusion). Whenever $0 \ne v \in C(G)^+$, we define $p_v \in P$ by

$$p_v(u) = \frac{(u : v)}{(w : v)} \text{ for } u \in C(G)^+ ,$$

as we may by the right-hand inequality in (3.1). Observe that the left-hand inequality in (3.1) tells us that if $u \neq 0$, then

$$0 < \frac{1}{(w : u)} \leq p_v(u) .$$

We let $\mathcal{B}$ be the family of sets of the form

$$\{(x, y) : xy^{-1} \in V\}$$

for V a neighborhood of e, the identity in the group G. (Then $\mathcal{B}$ is the basis for a uniformity on G—see [Kel 55] for the concept of uniformity.) With each $\beta \in \mathcal{B}$ we associate the closed set

$$S(\beta) = \overline{\{p_v : (\operatorname{spt} v) \times (\operatorname{spt} v) \subseteq \beta\}} .$$

If $\beta_1, \beta_2, \beta_3 \in \mathcal{B}$ and $\beta_1 \cap \beta_2 \supseteq \beta_3$ then $S(\beta_1) \cap S(\beta_2) \supseteq S(\beta_3) \neq \emptyset$. Thus, since P is compact (by Tychonoff's theorem), there is a point

$$\lambda \in \bigcap_{\beta \in \mathcal{B}} S(\beta) .$$

This function λ turns out to be a nonzero invariant integral on $C(G)^+$. That is to say, it is a bounded linear functional on $C(G)^+$, and it extends naturally to $C(G)$. The properties that we desire for λ follow immediately from the properties of the approximating functions p_v. The only nontrivial part of the verification is proving that

$$\lambda(u_1 + u_2) \geq \lambda(u_1) + \lambda(u_2) \text{ whenever } u_1, u_2 \in C(G)^+ . \tag{3.2}$$

To prove (3.2), we choose $f \in C(G)^+$ satisfying

$$\operatorname{spt} u_1 \cup \operatorname{spt} u_2 \subseteq \{x \in G : f(x) > 0\} .$$

For any $\epsilon > 0$, we define $s, r_1, r_2 \in C(G)^+$ so that

$$s = u_1 + u_2 + \epsilon f \ , \ \ r_j s = u_j \ , \ \text{ and } \operatorname{spt} r_j = \operatorname{spt} u_j \text{ for } j \in \{1, 2\} .$$

Now we use the uniform continuity of r_1, r_2 to obtain $\beta \in \mathcal{B}$ such that

$$|r_j(x) - f_j(y)| \leq \epsilon \text{ whenever } (x, y) \in \beta,\ j \in \{1, 2\} .$$

For any $v \in S(\beta)$, $a \in \operatorname{spt} v$, $\xi \in W(s, v)$, we define

$$\xi_j(G) = \left[f_j(g^{-1}a) + \epsilon \right] \xi(g) \text{ whenever } g \in G \text{ and } j \in \{1, 2\} .$$

We infer that

$$u_j(x) = r_j(x) \cdot s(x) \le \sum_{g \in G} r_j(x) \cdot \xi(g) \cdot v(gx) \le \sum_{g \in G} \xi_j(g) \cdot v(gx)$$

just because $v(gx) \neq 0$ implies that $(gx, a) \in \beta$ and $(x, g^{-1}a) \in \beta$. Thus $\xi_j \in W(u_j, v)$ and

$$(u_1 : v) + (u_2 : v) \le \sum_G \xi_1 + \sum_G \xi_2 \le (1 + 2\epsilon) \sum_G \xi$$

since $r_1 + r_2 \le 1$.

It follows that

$$p_v(u_1) + p_v(u_2) \le (1 + 2\epsilon) p_v(s) \le (1 + 2\epsilon) \Big[p_v(u_1 + u_2) + \epsilon p_v(f) \Big]$$

whenever $v \in S(\beta)$. Since $\lambda \in \overline{S(\beta)}$, we may now conclude that

$$\lambda(u_1) + \lambda(u_2) \le (1 + 2\epsilon) \Big[\lambda(u_1 + u_2) + \epsilon \lambda(f) \Big] . \qquad \square$$

Proof of Corollary 3.1.3. If $E \subseteq G$ then let us say that a sequence of continuous functions f_j is *adapted to* E if

(a) $0 \le f_1 \le f_2 \le \cdots$,
(b) $1 \le \lim_{j \to \infty} f_j(x) \quad$ whenever $x \in E$.

We define a set function ϕ by

$$\phi(E) = \inf \left\{ \lim_{j \to \infty} \lambda(f_j) : \{f_j\} \text{ is adapted to } E \right\} . \tag{3.3}$$

Of course, λ is monotone, in the sense that $f \le g$ implies $\lambda(f) \le \lambda(g)$. So the limit in (3.3) will always exist.

Claim 1. *The function ϕ is a measure on G.*

To verify this assertion we must show that if $E \subseteq \bigcup_{j=1}^{\infty} B_j$ then $\mu(E) \le \sum_j \mu(B_j)$. This follows because if $\{f_\ell^j\}$ is adapted to B_j then the sequence of functions

$$g_m = \sum_{j=1}^{m} f_m^j$$

is adapted to E. Moreover,

$$\lambda(g_m) = \sum_{j=1}^{m} \lambda(f_m^j) \le \sum_{j=1}^{\infty} \lim_{\ell \to \infty} \lambda(f_\ell^j) .$$

Claim 2. *Suppose that $g \in C(G)^+$, E is a set, $g(x) \le 1$ for $x \in E$, and $g(x) = 0$ for $x \notin E$. Then $\lambda(g) \le \phi(E)$.*

To see this, let $\{f_j\}$ be adapted to E. Then certainly

$$h_m \equiv \inf\{f_m, g\} \uparrow g \quad \text{as } m \uparrow \infty .$$

Thus

$$\lambda(g) = \lim_{m\to\infty} \lambda(h_m) \le \lim_{m\to\infty} \lambda(f_m) .$$

Claim 3. *Every $f \in C(G)^+$ is ϕ-measurable.*

To prove this claim, let $T \subseteq X$ and $-\infty < a < b < \infty$. We shall show that

$$\phi(T) \ge \phi(T \cap \{x : f(x) \le a\}) + \phi(T \cap \{x : f(x) \ge b\}) .$$

The assertion is trivial if $a \le 0$. Thus take $a \ge 0$ and assume that $\{g_j\}$ is adapted to T. Define

$$h = \frac{1}{b-a} \cdot \Big[\inf\{f, b\} - \inf\{f, a\} \Big]$$

and

$$k_m = \inf\{g_m, h\} .$$

Since

(a) $0 \le k_{m+1} - k_m \le g_{m+1} - g_m$,
(b) $h(x) = 1$ whenever $f(x) \ge b$,
(c) $h(x) = 0$ whenever $f(x) \le a$,

we see that the sequence $\{k_j\}$ is adapted to the set

$$B \equiv T \cap \{x : f(x) \ge b\}$$

and the sequence $\{g_j - k_j\}$ is adapted to the set

$$A = T \cap \{x : f(x) \le a\} .$$

In conclusion,

$$\lim_{m\to\infty} \lambda(g_m) = \lim_{m\to\infty} [\lambda(k_m) + \lambda(g_m - k_m)] \ge \phi(B) + \phi(A) .$$

Claim 4. *If $f \in C(G)^+$ then $\lambda(f) = \int f \, d\phi$.*

For this assertion, let $f_t = \inf\{f, t\}$ whenever $t \ge 0$.

Now if $k > 0$ is a positive integer and $\epsilon > 0$, then

(a) $0 \le f_{k\epsilon}(x) - f_{(k-1)\epsilon}(x) \le \epsilon$ for $x \in G$;
(b) $f_{k\epsilon}(x) - f_{(k-1)\epsilon}(x) = \epsilon$ whenever $f(x) \ge k\epsilon$;
(c) $f_{k\epsilon}(x) - f_{(k-1)\epsilon}(x) = 0$ whenever $f(x) \le (k-1)\epsilon$.

As a result,

$$\begin{aligned}
\lambda\left(f_{k\epsilon} - f_{(k-1)\epsilon}\right) &\ge \epsilon\, \phi\{x : f(x) \ge k\epsilon\} \\
&\ge \int (f_{(k+1)\epsilon} - f_{k\epsilon})\, d\phi \\
&\ge \epsilon\, \phi\{x : f(x) \ge (k+1)\epsilon\} \\
&\ge \lambda(f_{(k+2)\epsilon} - f_{(k+1)\epsilon}) .
\end{aligned}$$

Summing on k from 1 to m, we see that

$$\lambda(f_{m\epsilon}) \geq \int (f_{(m+1)\epsilon} - f_{\epsilon})\, d\phi \geq \lambda(f_{(m+2)\epsilon} - f_{2\epsilon})\,.$$

Certainly $f_{m\epsilon} \uparrow f$ as $m \uparrow \infty$ and $\lambda(f) \geq \int (f - f_{\epsilon})\, d\phi \geq \lambda(f - f_{2\epsilon})$. Also $f_{\epsilon} \downarrow 0$. It follows that $\lambda(f) = \int f\, d\phi$.

Now we use linearity to extend our assertion to all of $C(G)$. Let f be any continuous function on G. Write $f = f^+ - f^-$, where $f^+ \geq 0$ and $f^- \geq 0$. Then

$$\lambda(f) = \lambda(f^+) - \lambda(f^-) = \int f^+\, d\phi - \int f^-\, d\phi = \int f\, d\phi\,.$$

Finally, if U is any open subset of G, then let $f_1 \leq f_2 \leq \cdots$ be continuous functions such that $f_j(x)$ converges to the characteristic function χ_U of U. Then it follows that μ is translation invariant on U. This assertion may then be extended to Borel sets in an obvious way. Finally, one deduces the invariance of μ for measurable sets. This establishes the corollary. □

If G is a compact topological group and also happens to be a metric space (such as the orthogonal group—see below), then we say that the metric d is *invariant* if

$$d(gh, gk) = d(hg, kg) = d(h, k)$$

for any g, h, k in the group. It follows, for such a metric, that $g[\mathbb{B}(h, r)] = \mathbb{B}(gh, r)$ for any (open) metric ball. Since the Haar measure is invariant, we conclude that the Haar measure of all balls with the same radii are the same. In fact this property characterizes Haar measure, as we shall now see.

Definition 3.1.4. A Borel regular measure μ on a metric space X is called *uniformly distributed* if the measures of all nontrivial balls are positive and, in addition,

$$\mu(\mathbb{B}(x, r)) = \mu(\mathbb{B}(y, r)) \quad \text{for all } x, y \in X, 0 < r < \infty\,.$$

Proposition 3.1.5. *Let μ and ν be uniformly distributed Borel regular measures on a separable metric space X. Then there is a positive constant c such that $\mu = c \cdot \nu$.*

Proof. Define

$$g(r) = \mu(\mathbb{B}(x, r)) \quad \text{and} \quad h(r) = \nu(\mathbb{B}(x, r))\,,$$

where our hypothesis guarantees that these definitions are unambiguous (i.e., do not depend on $x \in X$). Suppose that $U \subseteq X$ is any nonempty, open, bounded subset of X. Then

$$\lim_{r \downarrow 0} \frac{\nu(U \cap \mathbb{B}(x, r))}{h(r)}$$

clearly exists and equals 1 for any $x \in U$. Now we have

$$\begin{aligned}
\mu(U) &= \int_U \lim_{r\downarrow 0} \frac{\nu(U\cap\mathbb{B}(x,r))}{h(r)}\,d\mu(x) \\
&\overset{\text{(Fatou)}}{\leq} \liminf_{r\downarrow 0}\left[\frac{1}{h(r)}\int_U \nu(U\cap\mathbb{B}(x,r))\,d\mu(x)\right] \\
&\overset{\text{(Fubini)}}{=} \liminf_{r\downarrow 0}\left[\frac{1}{h(r)}\int_U \mu(\mathbb{B}(y,r))\,d\nu(y)\right] \\
&= \left[\liminf_{r\downarrow 0}\frac{g(r)}{h(r)}\right]\nu(U)\,.
\end{aligned}$$

A symmetric argument shows that

$$\nu(U) \leq \left[\liminf_{r\downarrow 0}\frac{h(r)}{g(r)}\right]\mu(U)\,.$$

It follows immediately that

$$c \equiv \lim_{r\downarrow 0}\frac{g(r)}{h(r)}$$

exists. Furthermore, $\mu(U) = c\cdot\nu(U)$ for any bounded, open set $U \subseteq X$. Now the full equality follows by Borel regularity. □

It is a matter of some interest to determine the Haar measure on some specific groups and symmetric spaces. We have already noted that Haar measure on $\mathbb{R}^N$ is Lebesgue measure (or any constant multiple thereof). Since this group is noncompact, we must forgo the stipulation that the total mass of the measure be 1.

In this book we are particularly interested in groups that bear on the geometry of Euclidean space. We have already noted the Haar measure on the multiplicative reals, which corresponds to the dilation group. And the preceding paragraph treats the Haar measure of the group of translations. The next section treats the other fundamental group acting on space, which is the group of rotations.

3.2 Haar Measure for the Orthogonal Group and the Grassmannian

Let S^{N-1} be the standard unit sphere in $\mathbb{R}^N$,

$$S^{N-1} = \left\{x\in\mathbb{R}^N : |x| = \sum_{j=1}^N x_j^2 = 1\right\}.$$

Of course, S^{N-1} bounds $\mathbb{B}(0,1)$, which is the open unit ball in $\mathbb{R}^N$. Then S^{N-1} is an $(N-1)$-dimensional manifold, and is naturally equipped with the Hausdorff measure $\mathcal{H}^{N-1}$.

An equivalent method for defining an invariant measure on S^{N-1} is as follows: If $A \subseteq S^{N-1}$ we define

$$\widetilde{A} = \{ta : 0 \le t \le 1, a \in A\}.$$

Then set

$$\sigma_{N-1}(A) = \mathcal{H}^{N-1}(S^{N-1}) \cdot \frac{\mathcal{L}^N(\widetilde{A})}{\mathcal{L}^N[\mathbb{B}(0,1)]}.$$

It may be verified—by first checking on spherical caps in S^{N-1} and then using Vitali's theorem and outer regularity of the measure—that $\mathcal{H}^{N-1}$ and σ_{N-1} are equal measures on S^{N-1}. Of course, we may normalize either measure to have total mass 1 by dividing out by the surface area of the sphere, and we will assume this normalization in what follows. That is, we redefine σ_{N-1} by setting

$$\sigma_{N-1}(A) = \frac{\mathcal{L}^N(\widetilde{A})}{\mathcal{L}^N[\mathbb{B}(0,1)]}.$$

The orthogonal group $\mathbf{O}(N)$ consists of those linear transformations L with the property that

$$L^{-1} = L^{\mathtt{t}}. \tag{3.4}$$

This is the standard, if not the most enlightening, definition. Because of the identity

$$Lx \cdot Ly = x \cdot (L^{\mathtt{t}}Ly), \tag{3.5}$$

one can easily see that L is orthogonal if and only if

$$Lx \cdot Ly = x \cdot y$$

for all $x, y \in \mathbb{R}^N$.

A useful interpretation of (3.5) is that L will take any orthonormal basis for $\mathbb{R}^N$ to another orthonormal basis. Conversely, if $u_1, \dots, u_N$ and $v_1, \dots, v_N$ are orthonormal bases for $\mathbb{R}^N$ and if we set $L(u_j) = v_j$ for every j and extend by linearity, then the result is an orthogonal transformation of $\mathbb{R}^N$.

Recall that the special orthogonal group $\mathbf{SO}(N)$ consists of those orthogonal transformations having determinant 1. These will be just the rotations.

In $\mathbb{R}^2$ the condition of orthogonality has a particularly simple formulation: if u_1, u_2 is an orthonormal basis for $\mathbb{R}^2$ then any orthogonal transformation will either preserve the orientation (i.e., the order) of the pair, or it will not. In the first instance the transformation is a rotation. In the second it is a reflection in some line through the origin. In $\mathbb{R}^N$ we may say analogously that a linear transformation is orthogonal if and only if it is (i) a rotation, (ii) a reflection in some hyperplane through the origin, or (iii) a composition of these.

We know that the orthogonal group is compact. Indeed, the row entries of the matrix representation of an element of $\mathbf{O}(N)$ will just be an orthonormal basis of $\mathbb{R}^N$; so the set is closed and bounded. It is convenient to describe Haar measure θ_N on the orthogonal group $\mathbf{O}(N)$ by letting the measure be induced by the action of the group on the sphere.

Proposition 3.2.1. *Fix a point $s \in S^{N-1}$. Let $A \subseteq \mathbf{O}(N)$. Then it holds that*

$$\theta_N(A) = \sigma_{N-1}(\{gs : g \in A\}) .$$

Proof. Define $f : \mathbf{O}(N) \to S^{N-1}$ by $f(g) = gs$. We define the *pushforward measure* $[f_*\theta_N]$ on S^{N-1} by

$$[f_*\theta_n](B) = \theta_N(f^{-1}(B)) \quad \text{for } B \subseteq S^{N-1} .$$

We observe that, with $f^{-1}(B) = A$,

$$[f_*\theta_N](B) = \theta_N(A) = \theta_N(\{g \in \mathbf{O}(N) : gs \in B\} .$$

It is our job, then, to show that $[f_*\theta_N] = \sigma_{N-1}$. Since both these measures have total mass 1 on S^{N-1}, it suffices by Proposition 3.1.5 to show that $f_*\theta_N$ is uniformly distributed.

Now let $a, b \in S^{N-1}$. There is a (not necessarily unique) element $\widetilde{g} \in \mathbf{O}(N)$ such that $\widetilde{g}a = b$. In order to discuss the concept of "uniformly distributed" on S^{N-1}, we need a metric; we simply take that metric induced on the sphere by the standard metric on Euclidean space.[3] Let $\overline{\mathbb{B}}(x, r)$ denote the closed metric ball with center $x \in S^{N-1}$ and radius r. Then it is clear that $g(\mathbb{B}(a, r)) = \mathbb{B}(b, r)$ for any $r > 0$. But then the invariance of θ_N (since it is Haar measure) gives

$$\begin{aligned}
[f_*\theta_N](\mathbb{B}(b, r)) &= \theta_N(\{g \in \mathbf{O}(N) : |gs - b| \le r\}) \\
&= \theta_N(\{g \in \mathbf{O}(N) : |gs - \widetilde{g}a| \le r\}) \\
&= \theta_N(\{g \in \mathbf{O}(N) : |\widetilde{g}^{-1}gs - a| \le r\} \\
&= \theta_N(\{h \in \mathbf{O}(N) : |hs - a| \le r\} = [f_*\theta_N](\mathbb{B}(a, r)) .
\end{aligned}$$

Thus $[f_*\theta_N]$ is uniformly distributed and we are done. □

Now fix $0 < M < N$. The *Grassmannian*[4] $G(N, M)$ is the collection of all M-dimensional linear subspaces of $\mathbb{R}^N$. In fact it is possible to equip $G(N, M)$ with a manifold structure, and we shall say more about this point later. For the moment, we wish to consider a natural measure on $G(N, M)$.

In case $M = 1$ the task is fairly simple. When $N = 2$, each line is uniquely determined by the angle it subtends with the positive x-axis. Thus we may measure subsets of $G(N, M)$ by measuring the cognate set in the interval $[0, \pi)$ using Lebesgue measure. Similarly, a line in $\mathbb{R}^N$, $N \ge 2$, is determined by its two points of intersection

[3] It is worth noting that $\mathbf{O}(N)$ is also a metric space: If $g, h \in \mathbf{O}(N)$ then we define $d(g, h)$ as usual by

$$d(g, h) = \|g - h\| = \sup_{x \in S^{N-1}} |g(x) - h(x)| .$$

[4] Hermann Grassmann (1809–1877).

with the unit sphere S^{N-1}. So we may measure a set in $G(N, M)$ by measuring the cognate set in the sphere. When $N > M > 1$ then things are more complicated.

To develop a general framework for defining a measure on $G(N, M)$, we make use of Euclidean orthogonal projections. Let $0 < M < N$ and let $E \in G(N, M)$. Define

$$\mathcal{P}_E : \mathbb{R}^N \to \mathbb{R}^N$$

to be the Euclidean orthogonal projection onto E. If $E, F \in G(N, M)$ then we define a metric

$$d(E, F) = \|\mathcal{P}_E - \mathcal{P}_F\| \, ;$$

here, as usual, $\| \ \|$ denotes the standard operator norm. This metric makes $G(N, M)$ compact. To see that this compactness claim holds, we argue as follows: Any set of M orthonormal vectors determines an element of $G(N, M)$. Thus we have a (many-to-one) map from a closed subset of

$$\underbrace{S^{N-1} \times S^{N-1} \times \cdots \times S^{N-1}}_{M \text{ factors}}$$

to $G(N, M)$, and the metric makes the map continuous.

We see immediately that the action of $\mathbf{O}(N)$ on $G(N, M)$ is distance-preserving. Namely, the action of an orthogonal transformation on space will evidently preserve the relative positions of two M-planes. Alternatively, such a transformation preserves inner products, so it will preserve the set of vectors to which each of $E, F \in G(N, M)$ is orthogonal and hence will preserve $d(E, F)$. More specifically, if $g \in \mathbf{O}(N)$, then

$$d(gE, gF) = d(E, F) \, .$$

We further verify that $\mathbf{O}(N)$ acts transitively on $G(N, M)$. This means that if $E, F \in G(N, M)$, then there is an element $g \in \mathbf{O}(N)$ such that $gE = F$. To see this, let $u_1, \ldots, u_M$ be an orthonormal basis for E and $v_1, \ldots, v_M$ be an orthonormal basis for F. Complete the first basis to an orthonormal basis $u_1, \ldots, u_N$ for $\mathbb{R}^N$ and likewise complete the second basis to an orthonormal basis $v_1, \ldots, v_N$ for $\mathbb{R}^N$. Then the map $u_j \leftrightarrow v_j$, $j = 1, \ldots, N$, extends by linearity to an element of $\mathbf{O}(N)$, and it takes E to F.

Now fix an element $H \in G(N, M)$. Define the map

$$\begin{aligned} f_H : \mathbf{O}(N) &\to G(N, M) \\ g &\mapsto gH \, . \end{aligned}$$

Now we define a measure on $G(N, M)$ by

$$\gamma_{N,M} = [f_H]_* \theta_N \, .$$

More explicitly, if $A \subseteq G(N, M)$ then

$$\gamma_{N,M}(A) = \theta_N \{g \in G(N, M) : gH \in A\} \, .$$

Now, since θ_N is an invariant measure on $\mathbf{O}(N)$, we may immediately deduce that the measure $\gamma_{N,M}$ is invariant on $G(N, M)$ under the action of $\mathbf{O}(N)$. This means that, for $g \in \mathbf{O}(N)$ and $A \subseteq G(N, M)$,

$$\gamma_{N,M}(gA) = \gamma_{N,M}(A) .$$

Since $\mathbf{O}(N)$ acts transitively on $G(N, M)$, and in a distance-preserving manner, it is immediate that each $\mathbf{O}(N)$-invariant Radon measure on $G(N, M)$ is uniformly distributed. As a result, by Proposition 3.1.5, the measure is unique up to multiplication by a constant. One important consequence of this discussion is that the measure $\gamma_{N,M}$ is independent of the choice of H.

We may also note that, for any $A \subseteq G(N, M)$,

$$\gamma_{N,M}(A) = \gamma_{N,N-M}(\{E^\perp : E \in A\} . \tag{3.6}$$

Here $E^\perp$ is the usual Euclidean orthogonal complement of E in $\mathbb{R}^N$. One may check this assertion by showing that the right-hand side of (3.6) is $\mathbf{O}(N)$-invariant (just because $[gE]^\perp = g(E^\perp)$ for $g \in \mathbf{O}(N)$, $E \in G(N, M)$).

Again, the uniqueness of uniformly distributed measures allows us to relate $\gamma_{N,M}$ to the surface measure σ_{N-1} on the sphere. To wit, for $A \subseteq G(N, 1)$,

$$\gamma_{N,1}(A) = \sigma_{N-1}\left(\bigcup_{E \in A} E \cap S^{N-1}\right)$$

and

$$\gamma_{N,N-1}(A) = \sigma_{N-1}\left(\bigcup_{E \in A} E^\perp \cap S^{N-1}\right) .$$

We leave the details of these identities to the interested reader.

Similarly we can construct the invariant measure $\theta^*_{N,M}$ on $\mathbf{O}^*(N, M)$, the collection of orthogonal projections from $\mathbb{R}^N$ onto $\mathbb{R}^M$. Fix $p \in \mathbf{O}^*(N, M)$ and define $f_p : \mathbf{O} \to \mathbf{O}^*(N, M)$ by $f_p(g) = p \circ g$. Then we define $\theta^*_{N,M} = [f_p]_*\theta_N$.

3.2.1 Remarks on the Manifold Structure of $G(N, M)$

Fix $0 < M < N < \infty$ and consider $G(N, M)$. We will now sketch two methods for giving $G(N, M)$ a manifold structure.

Method 1. Let E be an M-dimensional subspace of $\mathbb{R}^N$. Then there is a natural bijection Φ between $\mathrm{Hom}\,(E, E^\perp)$ and a subset $U_E \subseteq G(N, M)$. Specifically, Φ sends a linear map $\mathcal{L}$ from E to $E^\perp$ to its graph $\Gamma_{\mathcal{L}} \subseteq E \oplus E^\perp$. An element of the graph is of course an ordered pair $(x, \mathcal{L}(x))$, with $x \in \mathbb{R}^M$ and $\mathcal{L}(x) \in \mathbb{R}^{N-M}$. The graph is thus a linear subspace of $\mathbb{R}^N$ of dimension M; it is therefore an element of $G(N, M)$.

We use the inverse mappings $\Phi^{-1} : U_E \to \mathrm{Hom}\,(E, E^\perp)$ as the coordinate charts for our manifold structure. □

Method 2. Let E be an M-dimensional subspace of $\mathbb{R}^N$, and let $\mathcal{P}_E : \mathbb{R}^N \to \mathbb{R}^N$ be orthogonal projection onto E. If $T = T_E$ is the $N \times N$ matrix representation of $\mathcal{P}_E$ then T is symmetric (since a projection must be self-adjoint), has rank M, and is idempotent (i.e., $T^2 = T$). Conversely, if $\widetilde{T}$ is any symmetric $N \times N$ matrix that has rank M and is idempotent then there is an M-dimensional subspace $\widetilde{E} \subseteq \mathbb{R}^N$ for which $\widetilde{T}$ is the matrix representation of the orthogonal projection onto $\widetilde{E}$. The reference [Hal 51] contains an incisive discussion of these ideas. Because of these considerations, we may identify $G(N, M)$ with the set of symmetric, idempotent, $N \times N$ matrices of rank M.

Now we take T to have the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A_{M\times M} & B_{M\times(N-M)} \\ C_{(N-M)\times M} & D_{(N-M)\times(N-M)} \end{pmatrix}, \tag{3.7}$$

where we take A to be an $M \times M$ matrix and thus the sizes of B, C, D are as indicated.

If A is nonsingular, then we can compute

$$\begin{pmatrix} I & 0 \\ -C & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix},$$

so we see that T has rank M if and only if $D = CA^{-1}B$. If we further assume that T is *symmetric* of rank M, then A is nonsingular and symmetric, $C = B^{\mathrm{t}}$, and so it must be that $D = B^{\mathrm{t}}A^{-1}B$. It follows that T is idempotent if and only if $A^2 + BB^{\mathrm{t}} = A$.

From the last paragraph, we see that $G(N, M)$ can be identified with the set of $N \times N$ matrices of the form (3.7) satisfying

(1) A is nonsingular and symmetric;
(2) $C = B^{\mathrm{t}}$;
(3) $D = B^{\mathrm{t}}A^{-1}B$;
(4) $A^2 + BB^{\mathrm{t}} = A$.

We observe that (1) is equivalent to $(M^2 - M)/2$ scalar conditions, (2) is equivalent to $M(N - M)$ scalar conditions, (3) is equivalent to $(N - M)^2$ scalar conditions, and (4) is equivalent to $(M^2 + M)/2$ scalar conditions. It then follows from the implicit function theorem that $G(N, M)$ is a manifold of dimension $M(N - M)$.

4

Covering Theorems and the Differentiation of Integrals

A number of fundamental problems in geometric analysis—ranging from decompositions of measures to density of sets to approximate continuity of functions—depend on the theory of differentiation of integrals. These results, in turn, depend on a variety of so-called covering theorems for families of balls (and other geometric objects). Thus we come upon the remarkable, and profound, fact that deep analytic facts reduce to rather elementary (but often difficult) facts about Euclidean geometry.

The technique of covering lemmas has become an entire area of mathematical analysis (see, for example, [DGu 75] and [Ste 93]). It is intimately connected with problems of differentiation of integrals, with certain maximal operators (such as the Hardy–Littlewood maximal operator), with the boundedness of multiplier operators in harmonic analysis, and (concomitantly) with questions of summation of Fourier series.

The purpose of the present chapter is to introduce some of these ideas. We do not strive for any sort of comprehensive treatment, but rather to touch upon the key concepts and to introduce some of the most pervasive techniques and applications.

4.1 Wiener's Covering Lemma and Its Variants

Let $S \subseteq \mathbb{R}^N$ be a set. A *covering* of S will be a collection $\mathcal{U} = \{U_\alpha\}_{\alpha \in \mathcal{A}}$ of sets such that $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \supseteq S$. If all the sets of $\mathcal{U}$ are open, then we call $\mathcal{U}$ an *open covering* of S. A *subcovering* of the covering $\mathcal{U}$ is a covering $\mathcal{V} = \{V_\beta\}_{\beta \in \mathcal{B}}$ such that each V_β is one of the U_α. A *refinement* of the covering $\mathcal{U}$ is a collection $\mathcal{W} = \{W_\gamma\}_{\gamma \in \mathcal{G}}$ of sets such that each W_γ is a subset of some U_α. If $\mathcal{U}$ is a covering of a set S, then the *valence* of $\mathcal{U}$ is the least positive integer M such that no point of S lies in more than M of the sets in $\mathcal{U}$.

It is elementary to see that any open covering of a set $S \subseteq \mathbb{R}^N$ has a countable subcover. We also know, thanks to Lebesgue, that any open covering of S has a refinement with valence at most $N + 1$ (see [HW 41, Theorem V 1]).

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_4,

Wiener's covering lemma[1] concerns a covering of a set by a collection of balls. When applying the lemma, one must be willing to replace any particular ball by a ball with the same center but triple its radius—see Figure 4.1.

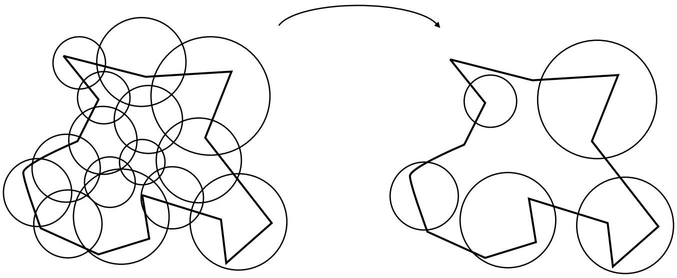

Fig. 4.1. Wiener's covering lemma.

Lemma 4.1.1 (Wiener). *Let $K \subseteq \mathbb{R}^N$ be a compact set with a covering $\mathcal{U} = \{B_\alpha\}_{\alpha \in A}$, $B_\alpha = \mathbb{B}(c_\alpha, r_\alpha)$, by open balls. Then there is a subcollection $B_{\alpha_1}, B_{\alpha_2}, \ldots, B_{\alpha_m}$, consisting of pairwise disjoint balls, such that*

$$\bigcup_{j=1}^{m} \mathbb{B}(c_{\alpha_j}, 3r_{\alpha_j}) \supseteq K.$$

Proof. Since K is compact, we may immediately assume that there are only finitely many B_α. Let B_{α_1} be the ball in this collection that has the greatest radius (this ball may not be unique). Let B_{α_2} be the ball that is disjoint from B_{α_1} and has greatest radius among those balls that are disjoint from B_{α_1} (again, this ball may not be unique). At the jth step choose the (not necessarily unique) ball disjoint from $B_{\alpha_1}, \ldots, B_{\alpha_{j-1}}$ that has greatest radius among those balls that are disjoint from $B_{\alpha_1}, \ldots, B_{\alpha_{j-1}}$. Continue. The process ends in finitely many steps. We claim that the B_{α_j} chosen in this fashion do the job.

For each j, we will write $B_{\alpha_j} = \mathbb{B}(c_{\alpha_j}, r_{\alpha_j})$. It is enough to show that $B_\alpha \subseteq \bigcup_j \mathbb{B}(c_{\alpha_j}, 3r_{\alpha_j})$ for every α. Fix an α. If $\alpha = \alpha_j$ for some j then we are done. If $\alpha \notin \{\alpha_j\}$, let j_0 be the first index j with $B_{\alpha_j} \cap B_\alpha \neq \emptyset$ (there must be one; otherwise, the process would not have stopped). Then $r_{\alpha_{j_0}} \geq r_\alpha$; otherwise, we selected $B_{\alpha_{j_0}}$ incorrectly. But then (by the triangle inequality) $\mathbb{B}(c_{\alpha_{j_0}}, 3r_{\alpha_{j_0}}) \supseteq \mathbb{B}(c_\alpha, r_\alpha)$ as desired. □

For completeness, and because it is such an integral part of the classical theory of measures, we now present the venerable covering theorem of Vitali.[2]

---

[1] Norbert Wiener (1894–1964).

[2] Giuseppe Vitali (1875–1932).

Proposition 4.1.2. *Let $A \subseteq \mathbb{R}^N$ and let $\mathcal{B}$ be a family of open balls. Suppose that each point of A is contained in arbitrarily small balls belonging to $\mathcal{B}$. Then there exist pairwise disjoint balls $B_j \in \mathcal{B}$ such that*

$$\mathcal{L}^N\left(A \setminus \bigcup_j B_j\right) = 0\,.$$

Furthermore, for any $\epsilon > 0$, we may choose the balls B_j in such a way that

$$\sum_j \mathcal{L}^N(B_j) \le \mathcal{L}^N(A) + \epsilon\,.$$

Proof. The last statement will follow from the substance of the proof. For the first statement, let us begin by making the additional assumption (which we shall remove at the end) that the set $A \equiv A_0$ is bounded. We may select a bounded open set U_0 that contains $\overline{A_0}$ and that is such that $\mathcal{L}^N(U_0)$ exceeds $\mathcal{L}^N(A_0)$ by as small a quantity as we may wish. In fact, we demand that

$$\mathcal{L}^N(U_0) \le (1 + 5^{-N})\mathcal{L}^N(A_0)\,.$$

Now focus attention on those balls that lie in U_0. By Lemma 4.1.1, we may select a finite, pairwise disjoint collection $B_j = \mathbb{B}(x_j, r_j) \in \mathcal{B}$, $j = 1, \ldots, k_1$, such that $B_j \subseteq U_0$ for each j and

$$\overline{A}_0 \subseteq \bigcup_{j=1}^{k_1} \mathbb{B}(x_j, 3r_j)\,.$$

Now we may calculate that

$$3^{-N}\mathcal{L}^N(A_0) \le 3^{-N}\sum_j \mathcal{L}^N[\mathbb{B}(x_j, 3r_j)] = 3^{-N}\sum_j 3^N\mathcal{L}^N(B_j) = \sum_j \mathcal{L}^N(B_j)\,.$$

Let

$$A_1 = A_0 \setminus \bigcup_{j=1}^{k_1} \overline{B_j}\,.$$

Then

$$\begin{aligned}
\mathcal{L}^N(A_1) &\le \mathcal{L}^N\left(U_0 \setminus \bigcup_{j=1}^{k_1} \overline{B_j}\right) \\
&= \mathcal{L}^N\left(U_0 \setminus \bigcup_{j=1}^{k_1} B_j\right) = \mathcal{L}^N(U_0) - \sum_{j=1}^{k_1} \mathcal{L}^N(B_j) \\
&\le (1 + 5^{-N} - 3^{-N})\,\mathcal{L}^N(A_0) \equiv u \cdot \mathcal{L}^N(A_0)\,,
\end{aligned}$$

where $u \equiv 1 + 5^{-N} - 3^{-N} < 1$. Now A_1 is a bounded subset of $\mathbb{R}^N \setminus \bigcup_{j=1}^{k_1} \overline{B_j}$. Hence we may find a bounded, open set $U_1 \subseteq U_0$ such that

$$A_1 \subseteq U_1 \subseteq \mathbb{R}^N \setminus \bigcup_{j=1}^{k_1} \overline{B_j}$$

and

$$\mathcal{L}^N(U_1) \leq (1 + 5^{-N})\, \mathcal{L}^N(A_1)\,.$$

Just as in the first iteration of this construction, we may now find disjoint balls B_j, $j = k_1 + 1, \ldots, k_2$, for which $B_j \subseteq U_1$ and

$$\mathcal{L}^N(A_2) \leq u \cdot \mathcal{L}^N(A_1) \leq u^2\, \mathcal{L}^N(A_0)\,;$$

here

$$A_2 = A_1 \setminus \bigcup_{j=k_1+1}^{k_2} \overline{B_j} = A_0 \setminus \bigcup_{j=1}^{k_2} \overline{B_j}\,.$$

By our construction, all the balls $B_1, \ldots, B_{k_2}$ are disjoint.

After m repetitions of this procedure, we find that we have balls $B_1, B_2, \ldots, B_{k_m}$ such that

$$\mathcal{L}^N\left(A_0 \setminus \bigcup_{j=1}^{k_m} B_j\right) \leq u^m\, \mathcal{L}^N(A_0)\,.$$

Since $u < 1$, the result follows.

For the general case, we simply decompose $\mathbb{R}^N$ into closed unit cubes Q_ℓ with disjoint interiors and sides parallel to the axes and apply the result just proved to each $A_0 \cap Q_\ell$. □

The Maximal Function

A classical construct, due to Hardy and Littlewood,[3] is the so-called maximal function. It is used to control other operators, and also to study questions of differentiation of integrals.

Definition 4.1.3. If f is a locally integrable function on $\mathbb{R}^N$, we let

$$Mf(x) = \sup_{R>0} \frac{1}{\mathcal{L}^N[\mathbb{B}(x,R)]} \int_{\mathbb{B}(x,R)} |f(t)|\, d\mathcal{L}^N(t)\,.$$

The operator M is called the *Hardy–Littlewood maximal operator*. The functions to which M is applied may be real-valued or complex-valued. A few facts are immediately obvious about M:

(1) M is not linear, but it is *sublinear* in the sense that

$$M[f+g](x) \leq Mf(x) + Mg(x)\,.$$

[3] Godfrey Harold Hardy (1877–1947), John Edensor Littlewood (1885–1977).

(2) Mf is always nonnegative, and it *could* be identically equal to infinity.
(3) Mf makes sense for any locally integrable f.

We will in fact prove that Mf is finite $\mathcal{L}^N$-almost everywhere, for any $f \in L^p$. In order to do so, it is convenient to formulate a weak notion of boundedness for operators. To begin, we say that a measurable function f is *weak type* p, $1 \le p < \infty$, if there exists a $C = C(f)$ with $0 < C < \infty$ such that, for any $\lambda > 0$,

$$\mathcal{L}^N(\{x \in \mathbb{R}^N : |f(x)| > \lambda\}) \le \frac{C}{\lambda^p}.$$

An operator T on L^p taking values in the collection of measurable functions is said to be of *weak type* (p, p) if there exists a $C = C(T)$ with $0 < C < \infty$ such that, for any $f \in L^p$ and for any $\lambda > 0$,

$$\mathcal{L}^N(\{x \in \mathbb{R}^N : |Tf(x)| > \lambda\}) \le C \cdot \left(\frac{\|f\|_{L^p}}{\lambda}\right)^p .$$

A function is defined to be *weak type* ∞ when it is L^∞. For $1 \le p < \infty$, an L^p function is certainly weak type p, but the converse is not true. In fact, we note that the function $f(x) = |x|^{-1/p}$ on $\mathbb{R}^1$ is weak type p, but not in L^p, for $1 \le p < \infty$. The Hilbert transform (see [Kra 99]) is an important operator that is not bounded on L^1 but is in fact weak type (1, 1).

Proposition 4.1.4. *The Hardy–Littlewood maximal operator M is weak type* (1, 1).

Proof. Let $\lambda > 0$. Set $S_\lambda = \{x : |Mf(x)| > \lambda\}$. Because Mf is the supremum of a set of continuous functions, Mf is lower semicontinuous, and consequently, S_λ is open.

Since S_λ is open, we may let $K \subseteq S_\lambda$ be a compact subset with $2\,\mathcal{L}^N(K) \ge \mathcal{L}^N(S_\lambda)$. For each $x \in K$, there is a ball $B_x = \mathbb{B}(x, r_x)$ with

$$\lambda < \frac{1}{\mathcal{L}^N(B_x)} \int_{B_x} |f(t)| d\mathcal{L}^N(t) .$$

Then $\{B_x\}_{x \in K}$ is an open cover of K by balls. By Lemma 4.1.1, there is a subcollection $\{B_{x_j}\}_{j=1}^M$ that is pairwise disjoint but such that the threefold dilates of these selected balls still cover K. Then

$$\begin{aligned}
\mathcal{L}^N(S_\lambda) \le 2\,\mathcal{L}^N(K) \le 2\,\mathcal{L}^N\left(\bigcup_{j=1}^M \mathbb{B}(x_j, 3r_j)\right) &\le 2\sum_{j=1}^M \mathcal{L}^N[\mathbb{B}(x_j, 3r_j)] \\
&\le \sum_{j=1}^M 2 \cdot 3^N\,\mathcal{L}^N(B_{x_j}) \\
&\le \sum_{j=1}^M \frac{2 \cdot 3^N}{\lambda} \int_{B_{x_j}} |f(t)| d\mathcal{L}^N(t) \\
&\le \frac{2 \cdot 3^N}{\lambda} \|f\|_{L^1}.
\end{aligned}$$

□

One of the venerable applications of the Hardy–Littlewood operator is the Lebesgue differentiation theorem:

Theorem 4.1.5. *Let f be a locally Lebesgue integrable function on $\mathbb{R}^N$. Then, for $\mathcal{L}^N$-almost every $x \in \mathbb{R}^N$, it holds that*

$$\lim_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) = f(x)\,.$$

Proof. Multiplying f by a compactly supported C^∞ function that is identically 1 on a ball, we may as well suppose that $f \in L^1$. We may also assume, by linearity, that f is real-valued. We begin by proving that

$$\lim_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)$$

exists.

Let $\epsilon > 0$. Select a function φ, continuous with compact support, and real-valued, such that $\|f - \varphi\|_{L^1} < \epsilon^2$. Then

$$\begin{aligned}
&\mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)\Bigg| > \epsilon\Bigg\} \\
&\le \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |f(t) - \varphi(t)|\, d\mathcal{L}^N(t) \\
&\qquad + \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t)\Bigg| \\
&\qquad + \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |\varphi(t) - f(t)|\, d\mathcal{L}^N(t) > \epsilon\Bigg\} \\
&\le \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |f(t) - \varphi(t)|\, d\mathcal{L}^N(t) > \frac{\epsilon}{3}\Bigg\} \\
&\quad + \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \Bigg|\limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t) \\
&\qquad - \liminf_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} \varphi(t)\, d\mathcal{L}^N(t)\Bigg| > \frac{\epsilon}{3}\Bigg\} \\
&\quad + \mathcal{L}^N\Bigg\{x \in \mathbb{R}^N : \limsup_{R \to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x, R)]} \int_{\mathbb{B}(x,R)} |\varphi(t) - f(t)|\, d\mathcal{L}^N(t) > \frac{\epsilon}{3}\Bigg\} \\
&\equiv I + II + III\,.
\end{aligned}$$

Now $II = 0$ because the set being measured is empty (since φ is continuous). Each of I and III may be estimated by

$$\mathcal{L}^N \left\{ x \in \mathbb{R}^N : M(f - \varphi)(x) > \epsilon/3 \right\},$$

and this, by Proposition 4.1.4, is majorized by

$$C \cdot \frac{\epsilon^2}{\epsilon/3} = c \cdot \epsilon\,.$$

In sum, we have proved the estimate

$$\mathcal{L}^N \left\{ x \in \mathbb{R}^N : \left| \limsup_{R\to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x,R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \right.\right.$$
$$\left.\left. - \liminf_{R\to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x,R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t) \right| > \epsilon \right\} \le c \cdot \epsilon\,.$$

It follows immediately that

$$\lim_{R\to 0^+} \frac{1}{\mathcal{L}^N[\mathbb{B}(x,R)]} \int_{\mathbb{B}(x,R)} f(t)\, d\mathcal{L}^N(t)$$

exists for $\mathcal{L}^N$-almost every $x \in \mathbb{R}^N$.

The proof that the limit actually equals $f(x)$ at $\mathcal{L}^N$-almost every point follows exactly the same lines. We shall omit the details. □

Corollary 4.1.6. *If $A \subseteq \mathbb{R}^N$ is Lebesgue measurable, then for almost every $x \in \mathbb{R}^N$, it holds that*

$$\chi_A(x) = \lim_{r\to 0^+} \frac{\mathcal{L}^N(A \cap \mathbb{B}(x,r))}{\mathcal{L}^N(\mathbb{B}(x,r))}.$$

Proof. Set $f = \chi_A$. Then

$$\int_{\mathbb{B}(x,r)} f(t)\, d\mathcal{L}^N(t) = \mathcal{L}^N(A \cap \mathbb{B}(x,r)),$$

and the corollary follows from Theorem 4.1.5. □

Definition 4.1.7. A function $f : \mathbb{R}^N \to \mathbb{R}$ is said to be *approximately continuous* if, for $\mathcal{L}^N$-almost every $x_0 \in \mathbb{R}^N$ and for each $\epsilon > 0$, the set

$$\{x : |f(x) - f(x_0)| > \epsilon\}$$

has density 0 at x_0, that is,

$$0 = \lim_{r\to 0^+} \frac{\mathcal{L}^N(\{x : |f(x) - f(x_0)| > \epsilon\} \cap \mathbb{B}(x_0,r))}{\mathcal{L}^N(\mathbb{B}(x_0,r))}.$$

Corollary 4.1.8. *If a function $f : \mathbb{R}^N \to \mathbb{R}$ is Lebesgue measurable, then it is approximately continuous.*

Proof. Suppose that f is Lebesgue measurable. Let $q_1, q_2, \ldots$ be an enumeration of the rational numbers. For each positive integer i, let E_i be the set of points $x \notin \{z : f(z) < q_i\}$ for which

$$0 < \limsup_{r \to 0^+} \frac{\mathcal{L}^N(\{z : f(z) < q_i\} \cap \mathbb{B}(x,r))}{\mathcal{L}^N(\mathbb{B}(x,r))}$$

and let E^i be the set of points $x \notin \{z : q_i < f(z)\}$ for which

$$0 < \limsup_{r \to 0^+} \frac{\mathcal{L}^N(\{z : q_i < f(z)\} \cap \mathbb{B}(x,r))}{\mathcal{L}^N(\mathbb{B}(x,r))}.$$

By Corollary 4.1.6 and the Lebesgue measurability of f, we know that

$$\mathcal{L}^N(E_i) = 0 \text{ and } \mathcal{L}^N(E^i) = 0 .$$

Thus we see that

$$E \equiv \bigcup_{i=1}^{\infty} (E_i \cup E^i)$$

is also a set of Lebesgue measure zero.

Consider any point $x_0 \notin E$ and any $\epsilon > 0$. There exist rational numbers q_i and q_j such that

$$f(x_0) - \epsilon < q_i < f(x_0) < q_j < f(x_0) + \epsilon.$$

We have $\{x : |f(x) - f(x_0)| > \epsilon\} \subseteq \{z : f(z) < q_i\} \cup \{z : q_j < f(z)\}$. By the definition of E_i and E^j we have

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{z : f(z) < q_i\} \cap \mathbb{B}(x_0,r))}{\mathcal{L}^N(\mathbb{B}(x_0,r))}$$

and

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{z : q_j < f(z)\} \cap \mathbb{B}(x_0,r))}{\mathcal{L}^N(\mathbb{B}(x_0,r))}.$$

It follows that

$$0 = \lim_{r \to 0^+} \frac{\mathcal{L}^N(\{x : |f(x) - f(x_0)| > \epsilon\} \cap \mathbb{B}(x_0,r))}{\mathcal{L}^N(\mathbb{B}(x_0,r))}.$$

Since $x_0 \notin E$ and $\epsilon > 0$ were arbitrary, we conclude that f is approximately continuous. $\square$

4.2 The Besicovitch Covering Theorem

Preliminary Remarks

The Besicovitch covering theorem,[4] which we shall treat in the present section, is of particular interest to geometric analysis because its statement and proof do not depend on a measure. This is a result about the geometry of balls in space.

The Besicovitch Covering Theorem

Theorem 4.2.1. *Let N be a positive integer. There is a constant $K = K(N)$ with the following property. Let $\mathcal{B} = \{B_j\}_{j=1}^{M}$, where $M \in \mathbb{N} \cup \{\infty\}$, be any finite or countable collection of balls in $\mathbb{R}^N$ with the property that the interior of no ball contains the center of any other. Then we may write*

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_K$$

so that each $\mathcal{B}_j$, $j = 1, \ldots, K$, is a collection of balls with pairwise disjoint closures.

Here by a ball we mean a set B satisfying $\mathbb{B}(x, r) \subseteq B \subseteq \overline{\mathbb{B}}(x, r)$, for some $x \in \mathbb{R}^N$ and some $r > 0$.

It is a matter of some interest to determine what the best possible K is for any given dimension N. Significant progress on this problem has been made in [Sul 94]. See also [Loe 93]. Certainly our proof below will give little indication of the best K.

We shall see that the heart of this theorem is the following lemma about balls.

Lemma 4.2.2. *There is a constant $\widetilde{K} = \widetilde{K}(N)$, depending only on the dimension of the space $\mathbb{R}^N$, with the following property: Let $B_0 = \overline{\mathbb{B}}(x_0, r_0)$ be a ball of fixed positive radius. Let $B_1 = \overline{\mathbb{B}}(x_1, r_1)$, $B_2 = \overline{\mathbb{B}}(x_2, r_2)$, $\ldots$, $B_p = \overline{\mathbb{B}}(x_p, r_p)$ be balls such that*

(1) *Each B_j has nonempty intersection with B_0, $j = 1, \ldots, p$;*
(2) *The radii r_j satisfy $r_j \geq r_0$ for all $j = 1, \ldots, p$;*
(3) *The interior of no ball B_j contains the center of any other B_k for $j, k \in \{0, \ldots, p\}$ with $j \neq k$.*

Then $p \leq \widetilde{K}$.

Here is what the lemma says in simple terms: Fix the ball B_0. Then at most $\widetilde{K}$ pairwise disjoint balls of (at least) the same size can touch B_0. Note here that being "pairwise disjoint" and "intersecting but not containing the center of the other ball" are essentially equivalent: if the second condition holds then shrinking each ball by a factor of one-half makes the balls pairwise disjoint; if the balls are already pairwise disjoint, have equal radii, and are close together, then doubling their size arranges for the first condition to hold.

Our proof of Lemma 4.2.2 is based on the next two lemmas—which in essence rely only on two-dimensional Euclidean geometry (trigonometry)—and on the fact

[4] Abram Samoilovitch Besicovitch (1891–1970).

that we can choose a set of unit vectors in $\mathbb{R}^N$ such that every direction is within a small angle of one of our chosen unit vectors (where the measure of an angle between two vectors is defined to be in the interval $[0, \pi]$).

Lemma 4.2.3. *Suppose the ball $\overline{\mathbb{B}}(q, r)$, with $r \geq 1$, intersects the closed unit ball and does not contain the origin in its interior, i.e., $r \leq |q|$. If u is a unit vector making an angle $\phi \leq \pi/6$ with q, then $\sqrt{3}\, u \in \overline{\mathbb{B}}(q, r)$.*

Proof. Because $\overline{\mathbb{B}}(q, r)$ intersects the closed unit ball and does not contain the origin in its interior, we can write $|q| = x + r$ with $0 \leq x \leq 1 \leq r$. By the law of cosines we have

$$\begin{aligned} |q - \sqrt{3}\, u|^2 &= |q|^2 + 3 - 2\sqrt{3}\, |q| \cos\phi \\ &\leq |q|^2 + 3 - 2\sqrt{3}\, |q| \cos\tfrac{\pi}{6} \\ &= (x + r)^2 + 3 - 3\,(x + r)\,. \end{aligned}$$

Thus it will suffice to show that

$$(x + r)^2 + 3 - 3\,(x + r) \leq r^2$$

or, equivalently,

$$f(x, r) = x^2 + 2xr + 3 - 3x - 3r \leq 0\,.$$

Since for each fixed r, $f(x, r)$ is quadratic in x with positive second derivative and since we are concerned only with the range $0 \leq x \leq 1$, it will suffice to consider only the endpoints $x = 0$ and $x = 1$. But we have

$$f(0, r) = 3 - 3r \leq 0 \quad \text{and} \quad f(1, r) = 1 + 2r + 3 - 3 - 3r = 1 - r \leq 0\,,$$

as required. □

Lemma 4.2.4. *Suppose neither of the balls $\overline{\mathbb{B}}(q_1, r_1)$ and $\overline{\mathbb{B}}(q_2, r_2)$ contains the center of the other ball in its interior. If the point p is in both balls, then the angle between $q_1 - p$ and $q_2 - p$ is at least $\pi/3$.*

Proof. To see this, we denote the angle in question by θ and use the law of cosines to compute

$$|q_1 - q_2|^2 = |q_1 - p|^2 + |q_2 - p|^2 - 2\,|q_1 - p|\,|q_2 - p| \cos\theta\,.$$

So we have

$$\cos\theta \leq \frac{|q_1 - p|^2 + |q_2 - p|^2 - |q_1 - q_2|^2}{2\,|q_1 - p|\,|q_2 - p|}\,.$$

Since neither ball contains the center of the other ball in its interior, we know that $|q_1 - q_2|$ is at least as large as the radius of either ball. So we have both $|q_1 - p| \leq$

$r_1 \leq |q_1 - q_2|$ and $|q_2 - p| \leq r_2 \leq |q_1 - q_2|$. Suppose without loss of generality that $|q_1 - p| \leq |q_2 - p|$. Then we estimate

$$\begin{aligned}\cos\theta &\leq \frac{|q_1 - p|^2 + |q_2 - p|^2 - |q_1 - q_2|^2}{2\,|q_1 - p|\,|q_2 - p|} \\ &\leq \frac{|q_1 - p|^2}{2\,|q_1 - p|\,|q_2 - p|} \\ &= \frac{1}{2} \cdot \frac{|q_1 - p|}{|q_2 - p|} \leq \frac{1}{2}\,,\end{aligned}$$

as required. □

Proof of Lemma 4.2.2. Suppose for the moment (we confirm this construction later) that we have chosen a set of unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ in $\mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there is a j such that the angle between u and u_j is strictly less than $\pi/6$ (picture points sufficiently dense on the unit sphere—see the discussion below). The number, $\kappa(N)$, of vectors u_j will be used below.

Consider balls $B_0, B_1, \ldots, B_p$ as in the statement of Lemma 4.2.2 and suppose that $p \geq \kappa(N)^2 + 1$. Without loss of generality, we may assume that $B_0 = \overline{\mathbb{B}}(0, 1)$. The direction to the center of each ball is within an angle strictly less than $\pi/6$ of one of the unit vectors u_j and so, by Lemma 4.2.3, must contain the point $\sqrt{3}\,u_j$. Since there are at least $\kappa(N)^2 + 1$ balls and only $\kappa(N)$ possible u_j's, there must be (at least) one j^* such that $\kappa(N) + 1$ of the balls contain the point $\sqrt{3}\,u_{j^*}$.

Now consider those $\kappa(N) + 1$ balls. The direction from $\sqrt{3}\,u_{j^*}$ to each center is within an angle strictly less than $\pi/6$ of one of the unit vectors u_k. But since there are $\kappa(N) + 1$ balls and only $\kappa(N)$ possible u_k's, there must be two centers within angle less than $\pi/6$ of the same direction and thus within an angle less than $\pi/3$ of each other, contradicting Lemma 4.2.4. We conclude that $p \leq \kappa(N)^2$.

Finally, we show that there exists a set of unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ in $\mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there is a j such that the angle between u and u_j is strictly less than $\pi/6$. Let

$$\mathcal{F} = \{\, \mathbb{B}(u_j, 1/4) : j = 1, 2, \ldots, \kappa(N) \,\}$$

be a maximal pairwise disjoint family of balls with centers in the unit sphere. All of the balls $\mathbb{B}(u_j, 1/4)$ are contained in $\mathbb{B}(0, 5/4)$, so, by comparing volumes, we see that

$$\kappa(N) \leq \frac{\Omega_N\,(5/4)^N}{\Omega_N\,(1/4)^N} = 5^N\,.$$

[In Remark 4.2.5, we give an alternative construction for the u_j that avoids any use of volume in $\mathbb{R}^N$ or $(N-1)$-dimensional area in the unit sphere.]

To see that the unit vectors $u_1, u_2, \ldots, u_{\kappa(N)}$ have the requisite property, let u be an arbitrary unit vector. There must exist a j such that $|u - u_j| < 1/2$; otherwise, we could add the ball $\mathbb{B}(u, 1/4)$ to the family $\mathcal{F}$, contradicting the maximality of $\mathcal{F}$. Fix such a j and let θ denote the angle between u_j and u. Using the law of cosines we estimate

$$\cos\theta = \frac{|u_j|^2 + |u|^2 - |u_j - u|^2}{2\,|u_j|\,|u|} = 1 - \tfrac{1}{2}|u_j - u|^2$$

$$\geq 7/8 > \sqrt{3}/2 = \cos\tfrac{\pi}{6}\,,$$

so the angle θ is strictly less than $\pi/6$. □

Remark 4.2.5. We now give another, more explicit, construction of a set of unit vectors $\mathcal{U} \subseteq \mathbb{R}^N$ with the property that for any unit vector $u \in \mathbb{R}^N$, there exists $u^* \in \mathcal{U}$ such that the angle between u and u^* is strictly less than $\pi/6$.

The vectors in $\mathcal{U}$ are formed by choosing $\theta_1, \theta_2, \ldots, \theta_{N-1}$ from the set

$$\left\{0,\ \tfrac{\pi}{m},\ \tfrac{2\pi}{m},\ \ldots,\ \tfrac{(m-1)\pi}{m},\ \pi\right\} \tag{4.1}$$

and choosing a sign $\tau \in \{-1, +1\}$. We then set

$$u_{\theta_1,\ldots,\theta_{N-1},\tau} = \Big(\cos\theta_1,\ \cos\theta_2\,\sin\theta_1,\ \ldots,\ \cos\theta_{N-1}\,\textstyle\prod_{i=1}^{N-2}\sin\theta_i,\ \tau\cdot\prod_{i=1}^{N-1}\sin\theta_i\Big).$$

Given a unit vector $u \in \mathbb{R}^N$, there exist $0 \leq \phi_i \leq \pi$, $i = 1, 2, \ldots, N-1$, and $\tau' \in \{-1, +1\}$ such that

$$u = \Big(\cos\phi_1,\ \cos\phi_2\,\sin\phi_1,\ \ldots,\ \cos\phi_{N-1}\,\textstyle\prod_{i=1}^{N-2}\sin\phi_i,\ \tau'\cdot\prod_{i=1}^{N-1}\sin\phi_i\Big).$$

The sign τ' represents a hemisphere containing u.

The main fact needed to verify that u is within $\pi/6$ of some $u_{\theta_1,\ldots,\theta_{N-1},\tau}$ is that if $\tau = \tau'$, then

$$u\cdot u_{\theta_1,\ldots,\theta_{N-1},\tau} = \cos(\theta_1-\phi_1) - \sum_{k=1}^{N-1}\left(\Big[1-\cos(\theta_k-\phi_k)\Big]\prod_{\ell=1}^{k-1}\sin\theta_\ell\,\sin\phi_\ell\right). \tag{4.2}$$

Equation (4.2) is proved by induction on N.

One completes the construction by choosing a sufficiently large value for m in (4.1). □

H. Federer's concept of a *directionally limited metric space*—see [Fed 69, 2.8.9]—abstracts and formalizes the geometry that goes into the proof of Lemma 4.2.2. More precisely, it generalizes to abstract contexts the notion that a cone in a given direction can contain only a certain number of points with distance $\eta > 0$ from the vertex and distance η from each other. The interested reader is advised to study that source.

Now we can present the proof of Besicovitch's covering theorem.

Proof of Theorem 4.2.1. First consider the case $M < \infty$ (recall that M was the number of balls in $\mathcal{B}$, the given collection of balls).

We have an iterative procedure for selecting balls.

Select B_1^1 to be a ball of maximum radius (this ball may not be unique). Then select B_2^1 to be a ball of maximum radius such that $\overline{B_2^1}$ is disjoint from $\overline{B_1^1}$ (again, this ball may not be unique). Continue until this selection procedure is no longer possible (remember that there are only finitely many balls in total). Set $\mathcal{B}_1 = \left\{B_j^1\right\}$.

Now work with the remaining balls. Let B_1^2 be the ball with greatest radius. Then select B_2^2 to be the remaining ball with greatest radius such that $\overline{B_2^2}$ is disjoint from $\overline{B_1^2}$. Continue in this fashion until no further selection is possible. Set $\mathcal{B}_2 = \left\{B_j^2\right\}$.

Working with the remaining balls, we now produce the family $\mathcal{B}_3$, and so forth (see Figure 4.2). Clearly, since in total there are only finitely many balls, this procedure must stop. We will have produced finitely many—say q—nonempty families of balls, each family consisting of balls having pairwise disjoint closures: $\mathcal{B}_1, \ldots, \mathcal{B}_q$. It remains to say how large q can be.

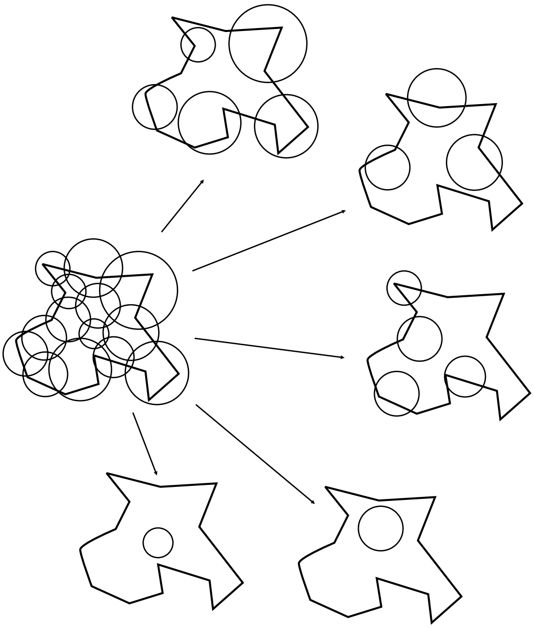

Fig. 4.2. Besicovitch's covering theorem.

Suppose that $q > \widetilde{K}(N) + 1$, where $\widetilde{K}(N)$ is as in the lemma. Let B_1^q be the first ball in the family $\mathcal{B}_q$. The closure of that ball must have intersected the closure of a ball in each of the preceding families (in case there are several such balls in a family, we consider the ball chosen earliest); by our selection procedure, each of those balls must have been at least as large in radius as B_1^q. Thus B_1^q is a ball with at least $\widetilde{K}(N)+1$ "neighbors" as in the lemma. But the lemma says that a ball can have only $\widetilde{K}(N)$ such neighbors. That is a contradiction.

We conclude that $q \le \widetilde{K}(N)+1$. That proves the theorem when M is finite.

When $M = \infty$, recursive application of the above iterative procedure completes the proof of the theorem. We argue as follows:

Suppose that for each $M = 1, 2, \ldots,$ the iterative procedure above is carried out for the set of balls $\{B_j\}_{j=1}^M$ resulting in the families of balls $\mathcal{B}_{M,i_1}$, $1 \le i_1 \le \widetilde{K}(N)+1$.

There must be a particular i_1 with $1 \le i_1 \le \widetilde{K}(N)+1$ such that the ball B_1 is assigned to $\mathcal{B}_{M,i_1}$ for infinitely many values of M. We assign B_1 to a family that we label $\mathcal{B}_{i_1}$.

Let $M_{1,1}$ be the smallest value of M for which B_1 is assigned to $\mathcal{B}_{M,i_1}$. Proceeding inductively, we assume that $M_{1,1} < M_{1,2} < \cdots < M_{1,\ell}$ have been defined. Let $M_{1,\ell+1}$ be the smallest value of M that is greater than $M_{1,\ell}$ and is such that B_1 is assigned to $\mathcal{B}_{M,i_1}$. Thus we define the increasing sequence $M_{1,\ell}$, $\ell = 1, 2, \ldots,$ with the property that B_1 is assigned to $\mathcal{B}_{M,i_1}$ when our procedure is carried out with $M = M_{1,\ell}$.

There must be a particular i_2 with $1 \le i_2 \le \widetilde{K}(N)+1$ such that the ball B_2 is assigned to $\mathcal{B}_{M,i_2}$ for infinitely many $M \in \{M_{1,1}, M_{1,2}, \ldots\}$. If $i_2 = i_1$ holds, then we assign B_2 to the family $\mathcal{B}_{i_1}$ that already contains B_1. In this case, we see that the closures of B_1 and B_2 do not intersect because there is an $M = M_{1,\ell}$ for which $B_1, B_2 \in \mathcal{B}_{M,i_1} = \mathcal{B}_{M,i_2}$ (in fact, there are infinitely many such M's). On the other hand, if $i_2 \ne i_1$, then we assign B_2 to a new family that we label $\mathcal{B}_{i_2}$.

Let $M_{2,1}$ be the smallest $M \in \{M_{1,1}, M_{1,2}, \ldots\}$ for which B_2 is assigned to $\mathcal{B}_{M,i_2}$. Proceeding inductively, we assume that $M_{2,1} < M_{2,2} < \cdots < M_{2,\ell}$ have been defined. Let $M_{2,\ell+1}$ be the smallest $M \in \{M_{1,1}, M_{1,2}, \ldots\}$ that is greater than $M_{2,\ell}$ and is such that B_2 is assigned to $\mathcal{B}_{M,i_2}$. Thus we define the increasing sequence $M_{2,\ell}$, $\ell = 1, 2, \ldots,$ that is a subsequence of $\{M_{1,p}\}_{p=1}^\infty$ and has the property that B_2 is assigned to $\mathcal{B}_{M,i_2}$ when our procedure is carried out with $M = M_{2,\ell}$.

Continuing in this way we assign each ball B_p to one of the families $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_{\widetilde{K}(N)+1}$. □

Remark 4.2.6. Note that there do not exist uncountable families of balls none of which contains the center of any of the other balls. That is because shrinking each ball by a factor of one-half—while keeping the same centers—makes the balls pairwise disjoint.

The next lemma show us one situation in which we can construct a covering of a set by a family of open balls with the property that no ball contains the center of any other ball.

Lemma 4.2.7. *Let $\mathcal{B}$ be a family of open balls centered at points of a compact set A. Suppose $\mathcal{B}$ is such that*

(1) *every point of A is the center of at least one ball in $\mathcal{B}$,*
(2) $\sup\{r : \mathbb{B}(x, r) \in \mathcal{B}\} < \infty$,

(3) $\{\mathbb{B}(x_i, r_i)\}_{i=1}^{\infty} \subseteq \mathcal{B}$ *with* $x_i \to x$ *and* $r_i \to r > 0$ *implies* $\mathbb{B}(x, r) \in \mathcal{B}$.

Then there are finitely many balls $\mathbb{B}(x_i, r_i) \in \mathcal{B}$, $i = 1, 2, \ldots, n$, *such that* $x_i \notin \mathbb{B}(x_j, r_j)$ *whenever* $i \neq j$ *and* $A \subseteq \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$.

Proof. Let $\mathbb{B}(x_1, r_1) \in \mathcal{B}$ be such that r_1 is maximal. Inductively we define $\mathbb{B}(x_{n+1}, r_{n+1})$ to be such that $x_{n+1} \in A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$ and r_{n+1} is maximal. If $A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i) = \emptyset$, the construction terminates and we do not define x_{n+1}.

Our construction ensures that we have $x_i \notin \mathbb{B}(x_j, r_j)$ whenever $i \neq j$. We claim that the construction terminates after finitely many steps. To see this fact, we argue by contradiction. Thus we suppose that $\mathbb{B}(x_i, r_i)$ has been defined for $i = 1, 2, \ldots$. Since the balls $\mathbb{B}(x_i, r_i/2)$ are disjoint and all lie in a bounded set, we see that $r_i \downarrow 0$, as $i \to \infty$.

Because A is compact and $\emptyset \neq A \setminus \bigcup_{i=1}^{n} \mathbb{B}(x_i, r_i)$ holds for each n, we see that there is $x \in A \setminus \bigcup_{i=1}^{\infty} \mathbb{B}(x_i, r_i)$. Let $\mathbb{B}(x, r) \in \mathcal{B}$. Since r_i is a nonincreasing sequence with limit 0, there must be an i such that $r_{i+1} < r \leq r_i$, but then we see that $\mathbb{B}(x_{i+1}, r_{i+1})$ was incorrectly chosen. □

Sometimes the requirement that no ball can contain the center of any other ball is too restrictive. In that case the condition we give next may be useful.

Definition 4.2.8. By a *controlled family of balls* we mean a family $\mathcal{B}$ of closed balls with positive radii such that if $\overline{\mathbb{B}}(a, r) \in \mathcal{B}$, $\overline{\mathbb{B}}(b, s) \in \mathcal{B}$, and $\overline{\mathbb{B}}(a, r) \neq \overline{\mathbb{B}}(b, s)$, then

$$\text{either} \qquad |a - b| > r > 4s/5 \qquad \text{or} \qquad |a - b| > s > 4r/5.$$

The next lemma tells us that if we shrink the balls in a controlled family by a factor of one-third, the balls become disjoint. Of course, that also implies that there are no uncountable controlled families.

Lemma 4.2.9. *If* $\overline{\mathbb{B}}(a, r)$ *and* $\overline{\mathbb{B}}(b, s)$ *are members of a controlled family, then* $\overline{\mathbb{B}}(a, r/3) \cap \overline{\mathbb{B}}(b, s/3) = \emptyset$.

Proof. We may assume without loss of generality that

$$|a - b| > r > 4s/5\,.$$

Suppose $p \in \overline{\mathbb{B}}(a, r/3) \cap \overline{\mathbb{B}}(b, s/3)$. Then we have

$$|a - b| \leq |a - p| + |p - b| \leq r/3 + s/3 \leq r/3 + (5/4) \cdot s/3 = 3r/4\,,$$

a contradiction. □

The geometric lemma applicable to balls in a controlled family is given next.

Lemma 4.2.10. *If* $\overline{\mathbb{B}}(a, r)$ *and* $\overline{\mathbb{B}}(b, s)$ *are members of a controlled family and if additionally*

$$4 \leq r \leq |a| \leq r + 1\,,$$

$$4 \leq s \leq |b| \leq s + 1\,,$$

then the angle between $a/|a|$ *and* $b/|b|$ *is at least* $\cos^{-1}(7/8)$.

Proof. Let θ denote the angle between $a/|a|$ and $b/|b|$. Since the balls are members of a controlled family, we may suppose without loss of generality that

$$|a-b| > r > 4s/5\,.$$

Using the law of cosines, we see that

$$\cos\theta = \frac{|a|^2+|b|^2-|a-b|^2}{2\,|a|\,|b|} = \frac{|a|}{2\,|b|}+\frac{|b|}{2\,|a|}-\frac{|a-b|^2}{2\,|a|\,|b|}$$

$$\leq \frac{r+1}{2s}+\frac{s+1}{2r}-\frac{r^2}{2rs} = \frac{1}{2s}+\frac{s}{2r}+\frac{1}{2r} \leq \frac{1}{8}+\frac{5}{8}+\frac{1}{8}\,. \qquad \square$$

As before, we have a bound, depending only on the dimension, for how many balls in a controlled family can intersect one particular ball.

Lemma 4.2.11. *There is a constant $K = K(N)$, depending only on the dimension of our space $\mathbb{R}^N$, with the following property: Let $B_0 = \overline{\mathbb{B}}(x_0, r_0)$ be a ball of fixed positive radius. Let $B_1 = \overline{\mathbb{B}}(x_1, r_1)$, $B_2 = \overline{\mathbb{B}}(x_2, r_2)$, $\ldots$, $B_p = \overline{\mathbb{B}}(x_p, r_p)$ be balls such that*

(1) *Each B_j has nonempty intersection with B_0, $j = 1, \ldots, p$;*
(2) *The radii $r_j \geq r_0$ for all $j = 1, \ldots, p$;*
(3) *The balls $\{B_j\}_{j=0}^p$ are members of a controlled family.*

Then $p \leq K$.

Proof. Without loss of generality we may suppose that $x_0 = 0$ and $r_0 = 1$. Divide the balls $B_1, B_2, \ldots, B_p$ into two collections:

$$\mathcal{B}_1 = \{\, B_j : 4 \leq r_j \leq |x_j| \leq r_j + 1 \,\}$$

and

$$\mathcal{B}_2 = \{B_j\}_{j=0}^p \Big\backslash\, \mathcal{B}_1\,.$$

By Lemma 4.2.10, the number of balls in $\mathcal{B}_1$ can be bounded by a number depending only on N. So our task is to bound the number of balls in $\mathcal{B}_2$.

We claim that

$$\bigcup_{B \in \mathcal{B}_2} B \subseteq \overline{\mathbb{B}}(0, 9)\,.$$

Observe that $|x_j| \leq r_j + 1$ holds for every j because $B_0 \cap B_j \neq \emptyset$. Thus

$$\mathcal{B}_2 = \{\, B_j : r_j < 4 \ \text{ or } \ |x_j| < r_j \,\}\,.$$

In case $r_j < 4$ holds, we have $|x_j| + r_j \leq 2r_j + 1 < 9$. Also, if $|x_j| < r_j$ and $j \neq 0$, then, because the balls are members of a controlled family, we have $|x_j| > 1 > 4r_j/5$, which yields $|x_j| + r_j < 2r_j < 5/2$.

Since the balls in $\left\{\overline{\mathbb{B}}(x_j, r_j/3)\right\}_{j=0}^p$ are pairwise disjoint (by Lemma 4.2.9) and since $r_j \geq 1$ holds for all the balls in $\mathcal{B}_2$, we see that $\mathcal{B}_2$ contains no more than $9^N/(1/3)^N = 3^{3N}$ balls. $\square$

Theorem 4.2.12. *Let N be a positive integer. There is a constant $K = K(N)$ with the following property. Given a set $A \subseteq \mathbb{R}^N$, a positive finite number R, and a family $\mathcal{B}$ of closed balls of positive radius not exceeding R, if every point of A is the center of at least one ball in $\mathcal{B}$, then there exist $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_K$ such that*

$$A \subseteq \bigcup_{j=1}^{K} \bigcup_{B \in \mathcal{B}_j} B \,,$$

and for each j, the balls in $\mathcal{B}_j$ are pairwise disjoint.

Proof. Enlarge A, if necessary, so that it contains *all* centers of balls in $\mathcal{B}$. It will certainly suffice to prove the result for this possibly larger set, which we will continue to denote by A.

If we construct a controlled family $\mathcal{B}' \subseteq \mathcal{B}$ with

$$A \subseteq \bigcup_{B \in \mathcal{B}'} B \,, \tag{4.3}$$

then we can obtain the desired conclusion by applying the argument used in the proof of Theorem 4.2.1, but with the role of Lemma 4.2.2 filled by Lemma 4.2.11.

We proceed to construct such a controlled family. To this end, we consider the class Ξ of all controlled subfamilies $\mathcal{B}'$ of $\mathcal{B}$ that also satisfy the condition that for any $\overline{\mathbb{B}}(y, s) \in \mathcal{B}$,

$$\left.\begin{aligned} \text{either} \quad & |x - y| \le r \quad \text{holds for some } \overline{\mathbb{B}}(x, r) \in \mathcal{B}', \\ \text{or} \quad & |x - y| > r > 4s/5 \quad \text{holds for every } \overline{\mathbb{B}}(x, r) \in \mathcal{B}'. \end{aligned}\right\} \tag{4.4}$$

We note that $\emptyset \in \Xi$, and we partially order Ξ using the relation $\subseteq$. It is easy to see that the union of any subclass of Ξ that is linearly ordered by $\subseteq$ is itself an element of Ξ. Therefore Zorn's lemma[5] tells us that Ξ has a maximal element $\mathcal{B}'$. It remains to verify that $\mathcal{B}'$ satisfies (4.3).

If $\mathcal{B}'$ does not satisfy (4.3), then

$$Y = \{y \in A : |y - x| > r \ \text{ holds for all } \overline{\mathbb{B}}(x, r) \in \mathcal{B}'\} \neq \emptyset \,.$$

Select $\overline{\mathbb{B}}(y^*, s^*)$ such that $y^* \in Y$ and

$$s^* > (4/5) \cdot \sup\{s : \exists y \in Y \ \text{ such that } \overline{\mathbb{B}}(y, s) \in \mathcal{B}\} \tag{4.5}$$

(this is where we use the fact that the radii of the balls are bounded by $R < \infty$). We will now show that $\mathcal{B}'' = \mathcal{B}' \cup \{\overline{\mathbb{B}}(y^*, s^*)\}$ is controlled and satisfies the condition (4.4).

To see that $\mathcal{B}''$ is controlled, we need only consider $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$ and $\overline{\mathbb{B}}(y^*, s^*)$. Since $y^* \in Y$, (4.4) tells us that $|x - y^*| > r > 4s^*/5$, verifying that $\mathcal{B}''$ is controlled.

[5] Max August Zorn (1906–1993).

To check that $\mathcal{B}''$ satisfies (4.4), we consider an arbitrary $\overline{\mathbb{B}}(y, s) \in \mathcal{B}$. If there already exists a $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$ for which $|x - y| \leq r$ holds, then (4.4) is satisfied. On the other hand, if $|x - y| > r$ holds for every $\overline{\mathbb{B}}(x, r) \in \mathcal{B}'$, then $y \in Y$. We consider $\overline{\mathbb{B}}(y^*, s^*)$. If $|y - y^*| \leq s^*$, then again (4.4) holds. Finally, we have the case in which $|y - y^*| > s^*$ holds. But now we also have $s^* > 4s/5$ by (4.5) and again (4.4) holds.

We have shown that $\mathcal{B}'' \in \Xi$ and we know that $\mathcal{B}'$ is a proper subset of $\mathcal{B}''$. This contradicts the maximality of $\mathcal{B}'$, so we conclude that in fact (4.3) is satisfied. □

Recall the notion of a Radon measure from Definition 1.2.11 in Section 1.2. Using the Besicovitch covering theorem instead of Wiener's covering lemma, we can prove a result like Vitali's (Proposition 4.1.2) for more general Radon measures:

Proposition 4.2.13. *Let μ be a Radon measure on $\mathbb{R}^N$. Let $A \subseteq \mathbb{R}^N$ and let $\mathcal{B}$ be a family of closed balls, with positive radius, such that each point of A is the center of arbitrarily small balls in $\mathcal{B}$. Then there are disjoint balls $B_j \in \mathcal{B}$ such that*

$$\mu\left(A \setminus \bigcup_j B_j\right) = 0\,.$$

Proof. We shall follow the same proof strategy as for Proposition 4.1.2. We may as well suppose that $\mu(A) > 0$; otherwise, there is nothing to prove. We also suppose (as we have done in the past) that A is bounded. Let K be as in Theorem 4.2.1.

Let U be a bounded open set with $A \subseteq U$ and choose a compact set C such that $C \subseteq U$ and $\mu(A \cap C) \geq (1/2)\,\mu(A)$. We define $\widetilde{\mathcal{B}}$ to be the family of balls in $\mathcal{B}$ that are centered in $A \cap C$ and contained in U.

By Theorem 4.2.1, we obtain subfamilies $\widetilde{\mathcal{B}}_1, \widetilde{\mathcal{B}}_2, \ldots, \widetilde{\mathcal{B}}_K$ such that each $\widetilde{\mathcal{B}}_j$ is a collection of balls that are pairwise disjoint. We have

$$A \cap C \subseteq \bigcup_{j=1}^{K} \bigcup_{B \in \widetilde{\mathcal{B}}_j} B.$$

Now it is clear that

$$\mu(A \cap C) \leq \sum_{j=1}^{K} \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_j} (A \cap B)\right).$$

Hence there is a particular index j_0 such that

$$\mu(A \cap C) \leq K \cdot \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_{j_0}} (A \cap B)\right).$$

We have

$$\mu(A) \le 2\,\mu(A \cap C) \le 2K \cdot \mu\left(\bigcup_{B \in \widetilde{\mathcal{B}}_{j_0}} (A \cap B)\right).$$

We can choose a finite subfamily $\widehat{\mathcal{B}} \subseteq \widetilde{\mathcal{B}}_{j_0}$ such that

$$\mu(A) \le 3K \cdot \mu\left(\bigcup_{B \in \widehat{\mathcal{B}}} (A \cap B)\right).$$

So setting

$$A_1 = A \setminus \bigcup_{B \in \widehat{\mathcal{B}}} B,$$

we conclude that

$$\mu(A_1) \le \mu(A)\,[1 - 1/(3K)]$$

and that A_1 is contained in the bounded open set $U_1 = U \setminus \bigcup_{B \in \widehat{\mathcal{B}}} B$. Now we simply iterate the construction, just as in the proof of Proposition 4.1.2.

We may dispense with the hypothesis that A is bounded just as in the proof of Proposition 4.1.2—making the additional observation that, since the Radon measure μ is σ-finite, it can measure at most countably many hyperplanes parallel to the axes with positive measure (so that we can avoid them when we chop up space into cubes). □

4.3 Decomposition and Differentiation of Measures

Next we turn to differentiation theorems for measures. These are useful in geometric measure theory and also in the theory of singularities for partial differential equations.

Suppose that μ and λ are Radon measures on $\mathbb{R}^N$. We define the *upper derivate* of μ with respect to λ at a point $x \in \mathbb{R}^N$ to be

$$\overline{D}_\lambda(\mu, x) \equiv \limsup_{r \downarrow 0} \frac{\mu[\mathbb{B}(x, r)]}{\lambda[\mathbb{B}(x, r)]}$$

and the *lower derivate* of μ with respect to λ at a point $x \in \mathbb{R}^N$ to be

$$\underline{D}_\lambda(\mu, x) \equiv \liminf_{r \downarrow 0} \frac{\mu[\mathbb{B}(x, r)]}{\lambda[\mathbb{B}(x, r)]}.$$

At a point x where the upper and lower derivates are equal, we define the *derivative of μ by λ* to be

$$D_\lambda(\mu, x) = \overline{D}_\lambda(\mu, x) = \underline{D}_\lambda(\mu, x).$$

Remark 4.3.1. It is convenient when calculating these derivates to declare $0/0 = 0$ (this is analogous to other customs in measure theory). The derivates that we have defined are Borel functions. To see this, first observe that $x \mapsto \mu[\mathbb{B}(x, r)]$ is continuous. This is in fact immediate from Lebesgue's dominated convergence theorem. Next notice that our definition of the three derivates does not change if we restrict r to lie in the positive rationals. Since, for each fixed r, the function

$$x \longmapsto \frac{\mu[\mathbb{B}(x, r)]}{\lambda[\mathbb{B}(x, r)]}$$

is continuous, and since the supremum and infimum of a countable family of Borel functions is Borel, we are done.

Definition 4.3.2. Let μ and λ be measures on $\mathbb{R}^N$. We say that μ is *absolutely continuous* with respect to λ if, for $A \subseteq \mathbb{R}^N$,

$$\lambda(A) = 0 \quad \text{implies} \quad \mu(A) = 0\,.$$

It is common to denote this relation by $\mu \ll \lambda$.

Our next result will require the following lemma:

Lemma 4.3.3. *Let μ and λ be Radon measures on $\mathbb{R}^N$. Let $0 < t < \infty$ and suppose that $A \subseteq \mathbb{R}^N$.*

(1) *If $\underline{D}_\lambda(\mu, x) \leq t$ for all $x \in A$ then $\mu(A) \leq t\lambda(A)$.*
(2) *If $\overline{D}_\lambda(\mu, x) \geq t$ for all $x \in A$ then $\mu(A) \geq t\lambda(A)$.*

Proof. If $\epsilon > 0$ then the Radon property gives us an open set U such that $A \subseteq U$ and $\lambda(U) \leq \lambda(A) + \epsilon$. Then the Vitali theorem for Radon measures (Proposition 4.2.13) gives disjoint closed balls $B_j \subseteq U$ such that

$$\mu(B_j) \leq (t + \epsilon)\lambda(B_j) \quad \text{(provided the balls are sufficiently small)}$$

and

$$\mu\left(A \setminus \bigcup_j B_j\right) = 0\,.$$

We conclude that

$$\mu(A) \leq \sum_j \mu(B_j) \leq (t + \epsilon) \sum_j \lambda(B_j)$$

$$\leq (t + \epsilon)\lambda(U) \leq (t + \epsilon)(\lambda(A) + \epsilon)\,.$$

Letting $\epsilon \to 0$ yields $\mu(A) \leq t \cdot \lambda(A)$. This is assertion (1). Assertion (2) may be established in just the same way. □

Theorem 4.3.4. *Suppose that μ and λ are Radon measures on $\mathbb{R}^N$.*

(1) *The derivative* $D_\lambda(\mu, x)$ *exists and is finite* λ*-almost everywhere.*
(2) *For any Borel set* $B \subseteq \mathbb{R}^N$,

$$\int_B D_\lambda(\mu, x)\, d\lambda(x) \leq \mu(B)\,,$$

with equality if $\mu \ll \lambda$.
(3) *The relation* $\mu \ll \lambda$ *holds if and only if* $\underline{D}_\lambda(\mu, x) < \infty$ *for* μ*-almost all* $x \in \mathbb{R}^N$.

Proof.
(1) Let $0 < r < \infty$ and $0 < s < t < \infty$. Define

$$A_{s,t}(r) = \{x \in \mathbb{B}(0, r) : \underline{D}_\lambda(\mu, x) \leq s < t \leq \overline{D}_\lambda(\mu, x)\}$$

and

$$A_t(r) = \{x \in \mathbb{B}(0, r) : \overline{D}_\lambda(\mu, x) \geq t\}\,.$$

Now Lemma 4.3.3 implies that

$$t \cdot \lambda(A_{s,t}(r)) \leq \mu(A_{s,t}(r)) \leq s \cdot \lambda(A_{s,t}(r)) < \infty$$

and, for $u > 0$,

$$u \cdot \lambda(A_u(r)) \leq \mu(A_u(r)) \leq \mu[\mathbb{B}(0, r)] < \infty\,.$$

Since $s < t$, these inequalities imply that $\lambda(A_{s,t}(r)) = 0$ and $\lambda(\bigcap_{u>0} A_u(r)) = \lim_{u\to\infty} \lambda(A_u(r)) = 0$. But

$$\begin{aligned}
&\mathbb{R}^N \setminus \{x \in \mathbb{R}^N : D_\lambda(\mu, x) \text{ exists and is finite}\}\\
&\quad = \bigcup_{r\in\mathbb{N}} \bigcup_{\substack{0<s<t\\ s,t\in\mathbb{Q}}} A_{s,t}(r) \;\cup\; \bigcup_{r\in\mathbb{N}} \bigcap_{u>0} A_u(r)\,. \qquad (4.6)
\end{aligned}$$

We see then that the set in (4.6) has λ-measure 0, and this proves assertion (1).

(2) For $1 < t < \infty$ and $p = 0, \pm 1, \pm 2, \ldots$, we define

$$B_p = \{x \in B : t^p \leq D_\lambda(\mu, x) < t^{p+1}\}\,.$$

Then part (1) above and Lemma 4.3.3(2) yield that

$$\begin{aligned}
\int_B D_\lambda(\mu, x)\, d\lambda(x) &= \sum_{k=-\infty}^{\infty} \int_{B_k} D_\lambda(\mu, x)\, d\lambda(x)\\
&\leq \sum_{k=-\infty}^{\infty} t^{k+1}\, \lambda(B_k)\\
&\leq t \cdot \sum_{k=-\infty}^{\infty} \mu(B_k)\\
&\leq t \cdot \mu(B)\,.
\end{aligned}$$

Letting $t \downarrow 1$ yields then $\int_B D_\lambda(\mu, x)\, d\lambda(x) \le \mu(B)$.

Suppose now that $\mu \ll \lambda$. Then the sets of λ-measure 0 are of course also sets of μ-measure zero. Part (1) tells us that $D_\lambda(\mu, x) = 1/D_\mu(\lambda, x) > 0$ for μ-almost every x. We conclude that $\mu(B) = \sum_{k=-\infty}^{\infty} \mu(B_k)$, and an argument similar to the one just given (using Lemma 4.3.3(2)) gives the inequality $\int_B D_\lambda(\mu, x)\, d\lambda(x) \ge \mu(B)$.
(3) By (1), we know that $\underline{D}_\lambda(\mu, x) < \infty$ at λ-almost every x; if $\mu \ll \lambda$ then this also holds at μ-almost every x.

For the reverse direction in (3), assume that $\underline{D}_\lambda(\mu, x) < \infty$ for μ-almost all $x \in \mathbb{R}^N$. Take $A \subseteq \mathbb{R}^N$ with $\lambda(A) = 0$. For $u = 1, 2, \ldots$, Lemma 4.3.3(2) gives

$$\mu\big(\{x \in A : \underline{D}_\lambda(\mu, x) \le u\} \le u \cdot \lambda(A) = 0\,.$$

We conclude that $\mu(A) = 0$. □

Now we reach our first goal, which is a density theorem and a theorem on the differentiation of integrals for Radon measures.

Theorem 4.3.5. *Let λ be a Radon measure on $\mathbb{R}^N$.*

(1) *If $A \subseteq \mathbb{R}^N$ is λ-measurable then the limit*

$$\lim_{r \downarrow 0} \frac{\lambda(A \cap \mathbb{B}(x, r))}{\lambda[\mathbb{B}(x, r)]}$$

exists and equals 1 for λ-almost every $x \in A$ and equals 0 for λ-almost every $x \in \mathbb{R}^N \setminus A$.
(2) *If $f : \mathbb{R}^N \to \overline{\mathbb{R}}$ is locally λ-integrable, then*

$$\lim_{r \downarrow 0} \frac{1}{\lambda[\mathbb{B}(x, r)]} \int_{\mathbb{B}(x,r)} f(x)\, d\lambda(x) = f(x)$$

for λ-almost every $x \in \mathbb{R}^N$.

Proof. Part (1) follows from part (2) by setting $f = \chi_A$. To prove (2), we may take $f \ge 0$. Define $\mu(A) = \int_A f(x)\, d\lambda(x)$. Then μ is a Radon measure and $\mu \ll \lambda$. Theorem 4.3.4(2) now yields that

$$\int_E D_\lambda(\mu, x)\, d\lambda(x) = \mu(E) = \int_E f\, d\lambda$$

for all Borel sets E. This clearly entails $f(x) = D_\lambda(\mu, x)$ for λ-almost all $x \in \mathbb{R}^N$. That proves (2). □

We say that two Radon measures μ and λ are *mutually singular* if there is a set $A \subseteq \mathbb{R}^N$ such that $\lambda(A) = 0 = \mu(\mathbb{R}^N \setminus A)$. Now we have a version of the Radon–Nikodym theorem combined with the Lebesgue decomposition.

Theorem 4.3.6. *Suppose that λ and μ are finite Radon measures on $\mathbb{R}^N$. Then there is a Borel function f and a Radon measure ν such that λ and ν are mutually singular and*

$$\mu(E) = \int_E f\, d\lambda + \nu(E)$$

for any Borel set $E \subseteq \mathbb{R}^N$. Furthermore, $\mu \ll \lambda$ if and only if $\nu = 0$.

Proof. Define

$$A = \{x \in \mathbb{R}^N : \underline{D}_\lambda(\mu, x) < \infty\}.$$

Recalling that $\llcorner$ denotes the restriction of a measure, we set

$$\mu_1 = \mu \llcorner A \quad \text{and} \quad \nu = \mu \llcorner (\mathbb{R}^N \setminus A).$$

Then obviously $\mu = \mu_1 + \nu$, and λ and ν are mutually singular by Theorem 4.3.4(1). Now Lemma 4.3.3(1) gives $\mu_1 \ll \lambda$; hence μ_1 has the required representation by Theorem 4.3.4(2) with $f(x) = D_\lambda(\mu, x)$. The last statement of the theorem is now obvious. □

We conclude this section with some results concerning densities of measures (see Definition 2.2.1).

Theorem 4.3.7. *Fix $0 < t$. If μ is a Borel regular measure on $\mathbb{R}^N$ and $A \subseteq C \subseteq \mathbb{R}^N$, then*

$$t \le \Theta^{*M}(\mu \llcorner C, x)\,, \text{ for all } x \in A, \quad \text{implies} \quad t \cdot \mathcal{S}^M(A) \le \mu(C)\,.$$

Remark 4.3.8. Since spherical measure is always at least as large as Hausdorff measure, we also have the conclusion

$$t \le \Theta^{*M}(\mu \llcorner C, x)\,, \text{ for all } x \in A, \quad \text{implies} \quad t \cdot \mathcal{H}^M(A) \le \mu(C)\,.$$

Proof. Without loss of generality, we may assume that $\mu(C) < \infty$. It will also be sufficient to prove that $t < \Theta^{*M}(\mu \llcorner C, x)$, for all $x \in A$, implies $t \cdot \mathcal{S}^M(A) \le \mu(C)$.

Fix $0 < \delta$. We will estimate the approximating measure $\mathcal{S}^M_{6\delta}(A)$. This estimation will require a special type of covering, which we construct next.

Set

$$\mathcal{B} = \{\overline{\mathbb{B}}(x, r) : x \in A, \quad 0 < r \le \delta, \quad t \cdot \Omega_M \cdot r^M \le (\mu \llcorner C)\overline{\mathbb{B}}(x, r)\},$$

$$\mathcal{B}_1 = \{\overline{\mathbb{B}}(x, r) \in \mathcal{B} : \quad 2^{-1}\delta < r \le \delta\},$$

and let $\mathcal{B}_1'$ be a maximal pairwise disjoint subfamily of $\mathcal{B}_1$.

Assuming that $\mathcal{B}_1', \mathcal{B}_2', \ldots, \mathcal{B}_k'$ have already been defined, set

$$\mathcal{B}_{j+1} = \left\{\overline{\mathbb{B}}(x, r) \in \mathcal{B} : \ 2^{-(j+1)}\delta < r \le 2^{-j}\delta, \ \ \emptyset = \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=1}^{j} \bigcup_{B \in \mathcal{B}_i'} B \right\},$$

and let $\mathcal{B}'_{j+1}$ be a maximal pairwise disjoint subfamily of $\mathcal{B}_{j+1}$.

Note that the assumption $\mu(C) < \infty$ ensures that each $\mathcal{B}'_i$ is finite. Also note that, by construction, any two closed balls in the family $\bigcup_{i=1}^{\infty} \mathcal{B}'_i$ are disjoint, so we have

$$\sum_{i=1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C)(B) = (\mu \llcorner C) \left(\bigcup_{i=1}^{\infty} \bigcup_{B \in \mathcal{B}'_i} B \right) \leq \mu(C) < \infty . \tag{4.7}$$

Claim. For each n,

$$A \subseteq \left(\bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}'_i} B \right) \cup \left(\bigcup_{i=n+1}^{\infty} \bigcup_{B \in \mathcal{B}'_i} \widehat{B} \right) \tag{4.8}$$

holds, where, for each ball $B = \overline{\mathbb{B}}(x, r)$, we set $\widehat{B} = \overline{\mathbb{B}}(x, 3r)$.

To verify the claim, consider $x \notin \bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}_i} B$. Since $\bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}_i} B$ is closed, there is $\overline{\mathbb{B}}(x, r) \in \mathcal{B}$ such that

$$\emptyset = \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=1}^{n} \bigcup_{B \in \mathcal{B}'_i} B .$$

Letting k be such that $2^{-k} < r \leq 2^{-(k-1)}$, we see that if $k > n$ and $\overline{\mathbb{B}}(x, r) \notin \mathcal{B}'_k$, then

$$\emptyset \neq \overline{\mathbb{B}}(x, r) \bigcap \bigcup_{i=n+1}^{k} \bigcup_{B \in \mathcal{B}'_i} B .$$

Thus there is $\overline{\mathbb{B}}(y, t) \in \mathcal{B}'_i$, where $n + 1 \leq i \leq k$, such that $\emptyset \neq \overline{\mathbb{B}}(x, r) \cap \overline{\mathbb{B}}(y, t)$. Since $r \leq 2^{-(k-1)}$ and $2^{-k} < t$, we have $x \in \overline{\mathbb{B}}(y, r + t) \subseteq \overline{\mathbb{B}}(y, 3t)$. The claim is proved.

Let $\epsilon > 0$ be arbitrary. By (4.7) (see also (4.8)), we choose n such that

$$\sum_{i=1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C)(B) < \epsilon .$$

Using the claim and letting rad B denote the radius of the ball B, we estimate

$$\begin{aligned}
\mathcal{S}^M_{6\delta}(A) &\leq \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) + \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} \widehat{B})^M \right) \\
&= \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) + 3^M \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} \Omega_M \, (\operatorname{rad} B)^M \right) \\
&\leq t^{-1} \left(\sum_{i=1}^{n} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C) B \right) + 3^M \, t^{-1} \left(\sum_{i=n+1}^{\infty} \sum_{B \in \mathcal{B}'_i} (\mu \llcorner C) B \right) \\
&\leq t^{-1} \, [\, \mu(C) + 3^M \, \epsilon \,] .
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we conclude that $\mathcal{S}^M_{6\delta}(A) \le t^{-1}\,\mu(C)$. The result follows, since $\delta > 0$ was also arbitrary. □

Corollary 4.3.9. *In $\mathbb{R}^N$, the measures $\mathcal{S}^N$, $\mathcal{H}^N$, $\mathcal{T}^N$, $\mathcal{C}^N$, $\mathcal{G}^N$, $\mathcal{Q}^N_t$, and $\mathcal{I}^N_t$ ($1 \le t \le \infty$) all agree with the N-dimensional Lebesgue measure $\mathcal{L}^N$.*

Proof. Noting that $\beta_t(N, N) = 1$, for $1 \le t \le \infty$, and using Proposition 2.1.5, we see that $\mathcal{S}^N$ is the largest of the measures $\mathcal{S}^N$, $\mathcal{H}^N$, $\mathcal{T}^N$, $\mathcal{C}^N$, $\mathcal{G}^N$, $\mathcal{Q}^N_t$, and $\mathcal{I}^N_t$, while $\mathcal{I}^N_1$ is the smallest. Theorem 4.3.7 implies $\mathcal{S}^N \le \mathcal{L}^N$ and (2.9) gives us $\mathcal{I}^N_1 \ge \mathcal{L}^N$, so the result follows. □

Corollary 4.3.10. *If μ is a Borel regular measure on $\mathbb{R}^N$, $A \subseteq \mathbb{R}^N$ is μ-measurable, and $\mu(A) < \infty$, then*

$$\Theta^{*M}(\mu \llcorner A, x) = 0$$

holds for $\mathcal{S}^M$-almost every $x \in \mathbb{R}^N \setminus A$.

Proof. Let j be a positive integer and set

$$C_j = \left\{ x \in (\mathbb{R}^N \setminus A) : j^{-1} \le \Theta^{*M}(\mu \llcorner A, x) \right\}.$$

Arguing by contradiction, suppose that $\mathcal{S}^M(C_j)$ is positive. Then, by the Borel regularity of μ, we can find a closed set $E \subseteq A$ such that

$$\mu(A \setminus E) < j^{-1} \cdot \mathcal{S}^M(C_j).$$

For $x \in C_j$, since E is closed and $x \notin E$, we have

$$\begin{aligned} j^{-1} \le \Theta^{*M}(\mu \llcorner A, x) &= \Theta^{*M}[\,\mu \llcorner (A \setminus E)\,, x] \\ &= \Theta^{*M}[\,(\mu \llcorner A) \llcorner (\mathbb{R}^N \setminus E)\,, x]\,. \end{aligned}$$

So we can apply Theorem 4.3.7 (with the roles of μ, A, and B played by $\mu \llcorner A$, $\mathbb{R}^N \setminus E$, and C_j, respectively), to conclude that

$$t \cdot \mathcal{S}^M(C_j) \le (\mu \llcorner A)(\mathbb{R}^N \setminus E) = \mu(A \setminus E)\,,$$

a contradiction.

Thus we have $\mathcal{S}^M(C_j) = 0$ and the result follows. □

4.4 The Riesz Representation Theorem

In this section, we prove a version of the Riesz representation theorem for linear functionals. Anticipating that our main application of this theorem will be to currents with finite mass, we have taken our linear functionals to be defined on the space of real-valued, infinitely differentiable, compactly supported functions on $\mathbb{R}^N$. Standard versions of the theorem apply to linear functionals on the space of *continuous*, compactly supported functions (see, for example, [Fol 84], [Roy 88], or [Rud 87]). In [EG 92], Evans and Gariepy prove a version of the theorem for linear functionals on the space of *vector-valued*, continuous, compactly supported functions.

Theorem 4.4.1 (Riesz Representation Theorem). *Let $\mathcal{D}$ denote the set of real-valued, infinitely differentiable, compactly supported functions on $\mathbb{R}^N$. If $L : \mathcal{D} \to \mathbb{R}$ is a linear functional satisfying*

$$M = \sup \left\{ |L(\phi)| : \phi \in \mathcal{D}, \sup_{x \in \mathbb{R}^N} |\phi| \le 1 \right\} < \infty , \tag{4.9}$$

then there exists a Radon measure λ on $\mathbb{R}^N$ and a λ-measurable function $g : \mathbb{R}^N \to \mathbb{R}$ such that

(1) $\lambda \left(\mathbb{R}^N \right) = M$,

(2) $L(\phi) = \int_{\mathbb{R}^N} \phi \, g \, d\lambda$, *for all $\phi \in \mathcal{D}$.*

Proof. First, we note that it follows immediately from (4.9) that

$$|L(\phi)| \le M \cdot \sup_x |\phi(x)| , \text{ for } \phi \in \mathcal{D} . \tag{4.10}$$

Step 1: Definition of the measure λ. We define the function λ on subsets of $\mathbb{R}^N$ by setting $\lambda(\emptyset) = 0$, setting

$$\lambda(U) = \sup \left\{ |L(\phi)| : \phi \in \mathcal{D}, \sup_x |\phi(x)| \le 1, \text{ supp}\,\phi \subseteq U \right\} \tag{4.11}$$

when U is a nonempty open set, and setting

$$\lambda(E) = \inf \{ \lambda(U) : U \text{ is open, } E \subseteq U \} \tag{4.12}$$

when E is not an open set.

Ultimately we will show that λ is a measure. It follows immediately that

$$\lambda(\mathbb{R}^N) = M , \tag{4.13}$$

$$A \subseteq B \text{ implies } \lambda(A) \le \lambda(B) . \tag{4.14}$$

To show that μ is a measure, we first show that λ is countably subadditive on the family of open sets. To see this, let U_i, $i = 1, 2, \ldots$, be a sequence of open sets. We need to show that

$$\lambda\left(\bigcup_i U_i \right) \le \sum_i \lambda(U_i) \tag{4.15}$$

holds. It is no loss of generality to assume that $\sum_i \lambda(U_i) < \infty$.

Suppose that $\phi \in \mathcal{D}$, $\sup_{x \in \mathbb{R}^N} |\phi| \le 1$, and $\text{supp}\,\phi \subseteq \cup_i U_i$. Let α_i be a smooth partition of unity for the set $\text{supp}\,\phi$, subordinate to the cover $\{U_i\}_{i=1}^{\infty}$ (see [KPk 99]).

We estimate

$$\left|L\left(\sum_{i=m}^{n}\phi\cdot\alpha_i\right)\right| = \left|\sum_{i=m}^{n}L\left(\phi\cdot\alpha_i\right)\right| \le \sum_{i=m}^{n}\left|L\left(\phi\cdot\alpha_i\right)\right| \le \sum_{i=m}^{\infty}\lambda(U_i)\,.$$

Thus $L(\sum_i \phi\cdot\alpha_i)$ and $\sum_i |L(\phi\cdot\alpha_i)|$ are convergent. We then have

$$|L(\phi)| = \left|L\left(\phi\sum_i\alpha_i\right)\right| = \left|L\left(\sum_i\phi\cdot\alpha_i\right)\right| \le \sum_i\left|L\left(\phi\cdot\alpha_i\right)\right| \le \sum_i\lambda(U_i)\,,$$

and (4.15) follows.

To complete the proof that λ is a measure, we show that λ is countably subadditive on the family of all subsets of $\mathbb{R}^N$. To see this, we let E_i, $i = 1, 2, \ldots$, be a sequence of sets. We need to show that $\lambda(\bigcup_i E_i) \le \sum_i \lambda(E_i)$. We may suppose without loss of generality that $\sum_i \lambda(E_i) < \infty$.

Let $\epsilon > 0$ be arbitrary. For each i, let U_i be an open set with $\lambda(U_i) \le \lambda(E_i)+2^{-i}\epsilon$. Then, by (4.15), we have

$$\lambda(\textstyle\bigcup_i E_i) \le \lambda(\textstyle\bigcup_i U_i) \le \displaystyle\sum_i\lambda(U_i) \le \epsilon + \sum_i\lambda(E_i)\,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

Step 2: A bound on L. We claim that

$$|L(\phi)| \le \sup_x|\phi(x)|\cdot\lambda\Big(\{x : \phi(x)\ne 0\}\Big)\,,\ \text{for } \phi\in\mathcal{D}\,. \tag{4.16}$$

To see this, fix a nonzero $\phi\in\mathcal{D}$, set $\kappa = \sup_x|\phi(x)|$, and set

$$U = \{x : \phi(x)\ne 0\}\,.$$

Let $\alpha_\ell : \mathbb{R}\to\mathbb{R}$, $\ell = 1, 2, \ldots$, be a sequence of infinitely differentiable functions such that

$$\begin{aligned} \alpha_\ell(t) = 0 \quad &\text{if} \quad |t| \le 1/(2\ell),\\ |\alpha_\ell(t)| \le 1/\ell\ &\text{if}\ 1/(2\ell) < |t| < 1/\ell,\\ \alpha_\ell(t) = t \quad &\text{if} \quad 1/\ell \le |t|\,. \end{aligned}$$

For ℓ such that $1/\ell \le \sup_x|\phi(x)|$, we have $\kappa = \sup_x \alpha_\ell\circ\phi(x)$ and

$$\operatorname{supp}\alpha_\ell\circ\phi \subseteq U\,,$$

so

$$|L(\alpha_\ell\circ\phi)| \le \kappa\,\lambda(U)\,.$$

Since $\sup_x|\phi - \alpha_\ell\circ\phi| \le 1/\ell$ holds, we conclude from (4.10) that

$$|L(\phi) - L(\alpha_\ell\circ\phi)| = |L(\phi - \alpha_\ell\circ\phi)| \le M/\ell$$

holds. Letting $\ell\to\infty$, we obtain the claim.

Step 3: Showing that λ is a Radon measure. First, we claim that λ is finitely additive on the family of open sets. To see this, let U and V be disjoint open sets. Let $\epsilon > 0$ be arbitrary. Let $\phi_U\in\mathcal{D}$ satisfy

- $\sup_x |\phi_U(x)| \le 1$,
- $\operatorname{supp} \phi_U \subseteq U$,
- $\lambda(U) \le |L(\phi_U)| + \epsilon$.

Replacing ϕ_U by $-\phi_U$ if necessary, we may assume that $L(\phi_U) = |L(\phi_U)|$. Choose $\phi_V \in \mathcal{D}$ similarly. Then we have

$$\begin{aligned} \lambda(U) + \lambda(V) &\le |L(\phi_U)| + |L(\phi_V)| + 2\epsilon \\ &= L(\phi_U) + L(\phi_V) + 2\epsilon \\ &= L(\phi_U + \phi_V) + 2\epsilon \\ &\le |L(\phi_U + \phi_V)| + 2\epsilon \le \lambda(U \cup V) + 2\epsilon\,, \end{aligned}$$

and since $\epsilon > 0$ was arbitrary, the claim follows.

Next, we claim that λ satisfies Carathéodory's criterion. To see this, let A and B be sets that are separated by a positive distance.

Let $\epsilon > 0$ be arbitrary. We can find an open set U with $A \cup B \subseteq U$ and $\lambda(U) \le \lambda(A \cup B) + \epsilon$. Since A and B are at a positive distance from each other, we may assume without loss of generality that $U = U_A \cup U_B$, where U_A and U_B are disjoint open sets containing A and B, respectively. Then we have

$$\lambda(A) + \lambda(B) \le \lambda(U_A) + \lambda(U_B) = \lambda(U_A \cup U_B) \le \lambda(A \cup B) + \epsilon\,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

Since λ satisfies Carathéodory's criterion, we know that all open sets are λ-measurable. The fact that λ is a Radon measure follows from (4.12) and the fact that $\lambda(\mathbb{R}^N) < \infty$.

Step 4: Extension of L. Let $\overline{\mathcal{D}}$ denote the set of functions $f : \mathbb{R}^N \to \mathbb{R}$ such that f is bounded and f is the pointwise limit of a sequence of functions in $\mathcal{D}$. We observe that

- $\overline{\mathcal{D}}$ contains the characteristic function of any open subset of $\mathbb{R}^N$,
- $\overline{\mathcal{D}}$ is a vector space,
- $\overline{\mathcal{D}}$ is closed under multiplication.

We will define the extension of L from $\mathcal{D}$ to $\overline{\mathcal{D}}$.

Let $f \in \overline{\mathcal{D}}$. Let ϕ_i be a sequence of functions in $\mathcal{D}$ with $f = \lim_i \phi_i$. We may assume without loss of generality that the functions ϕ_i are uniformly bounded.

Set

$$\kappa \equiv \sup_i \sup_x \phi_i(x) < \infty\,.$$

Fix $\epsilon > 0$. For each n, set

$$A_n = \{x : \exists\, i, j \ge n \;\text{ such that }\; |\phi_i(x) - \phi_j(x)| \ge \epsilon\,\}\,.$$

Then we have $A_1 \supseteq A_2 \supseteq \cdots$ and $\cap_n A_n = \emptyset$. So $\lambda(A_n) \downarrow 0$ as $n \to \infty$. Fix an n such that $\lambda(A_n) < \epsilon$.

Let $\beta : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function satisfying

- β takes its values in $[0, 1]$,
- $\beta(t) = 1$ if $|t| \geq 2\epsilon$,
- $\beta(t) = 0$ if $|t| < \epsilon$.

For $i, j \geq n$, we have

$$\begin{aligned} |L(\phi_i - \phi_j)| &\leq \Big| L[\,\beta \circ (\phi_i - \phi_j) \cdot (\phi_i - \phi_j)\,] \Big| \\ &\qquad + \Big| L[\,(1 - \beta \circ (\phi_i - \phi_j)) \cdot (\phi_i - \phi_j)\,) \Big| \\ &\leq 2\,(\kappa + M)\,\epsilon\,. \end{aligned}$$

Thus we see that $L(\phi_i)$ forms a Cauchy sequence. We define

$$L\Big(\lim_{i\to\infty} \phi_i\Big) = \lim_{i\to\infty} L(\phi_i)\,.$$

It is easy to see that the extension of L is well-defined and linear.

The extension of L satisfies an estimate like (4.16); specifically, we claim that if $f \in \overline{\mathcal{D}}$, then it holds that

$$|L(f)| \leq \sup_x |f(x)| \cdot \lambda\Big(\{x : f(x) \neq 0\}\Big)\,. \tag{4.17}$$

To see this, fix the function $f \in \overline{\mathcal{D}}$ and fix a uniformly bounded sequence $\phi_i \in \mathcal{D}$ that converges pointwise to f. It is no loss of generality to assume that

$$\sup_x |f(x)| = \lim_{i\to\infty} \Big(\sup_x |\phi_i(x)| \Big)\,.$$

Set $W = \{x : f(x) \neq 0\}$.

Let $\epsilon > 0$ be arbitrary. Then we can find an open set U with $W \subseteq U$ and $\lambda(U) \leq \lambda(W) + \epsilon$.

Let $\alpha_\ell : \mathbb{R}^N \to \mathbb{R}$ be a sequence of infinitely differentiable functions with values in $[0, 1]$ such that $\{x : \alpha_\ell(x) = 1\}$ increases to χ_U. Then $\phi_i \cdot \alpha_i$ is a uniformly bounded sequence that converges to f.

We have

$$\begin{aligned} |L(\phi_i \cdot \alpha_i)| &\leq \sup_x |(\phi_i \cdot \alpha_i)(x)| \cdot \lambda\Big(\{x : (\phi_i \cdot \alpha_i)(x) \neq 0\}\Big) \\ &\leq \sup_x |f(x)| \cdot \lambda\Big(\{x : \alpha_i(x) \neq 0\}\Big) \\ &\leq \sup_x |f(x)| \cdot \lambda(U) \\ &\leq \sup_x |f(x)| \cdot (\lambda(W) + \epsilon)\,, \end{aligned}$$

and the claim follows.

Step 5: A family of subsets of $\mathbb{R}^N$. Let $\mathcal{O}$ denote the family of subsets A of $\mathbb{R}^N$ for which $\chi_A \in \overline{\mathcal{D}}$. Since

$$\chi_{A\cap B} = \chi_A\,\chi_B\,,$$

$$\chi_{A\cup B} = \chi_A + \chi_B - \chi_A\,\chi_B\,,$$

$$\chi_{A\setminus B} = (1-\chi_B)\,\chi_A\,,$$

we see that $\mathcal{O}$ is closed under finite unions, finite intersections, and complements. Also every element of $\mathcal{O}$ is a Borel set. Note that

$$L(\chi_U) + \lambda(U) \ge 0$$

holds, for any $U \in \mathcal{O}$.

Step 6: Definition of the measure μ. We define the function μ on subsets of $\mathbb{R}^N$ by setting

$$\mu(U) = L(\chi_U) + \lambda(U)\,, \tag{4.18}$$

when U is open, and setting

$$\mu(E) = \inf\{\,\mu(U) : U \text{ is open},\ E \subseteq U\,\}\,, \tag{4.19}$$

when E is not open.

For sets $U, V \in \mathcal{O}$ with $U \subseteq V$, we have

$$\begin{aligned} L(\chi_V) + \lambda(V) &= L(\chi_U + \chi_{V\setminus U}) + \lambda(\,U \cup (V\setminus U)\,) \\ &= L(\chi_U) + L(\chi_{V\setminus U}) + \lambda(U) + \lambda(V\setminus U) \\ &\ge L(\chi_U) + \lambda(U)\,. \end{aligned}$$

If U and V are open with $U \subseteq V$, then we conclude that $\mu(U) \le \mu(V)$. Then by (4.19), μ is monotone on all sets.

We claim that

$$\mu(E) = L(\chi_E) + \lambda(E)\,,\quad \text{for}\ \ E \in \mathcal{O}\,. \tag{4.20}$$

The argument above also shows that if U is open, $E \in \mathcal{O}$, and $E \subseteq U$, then

$$L(\chi_E) + \lambda(E) \le \mu(U)\,.$$

Let $\epsilon > 0$ be arbitrary. Then we can find an open U with $E \subseteq U$ and

$$\lambda(U) \le \lambda(E) + \epsilon\,.$$

Since

$$\lambda(U) = \lambda(U\setminus E) + \lambda(E)\,,$$

we have

$$\lambda(U \setminus E) \leq \epsilon\,.$$

By (4.17), we have

$$L(\chi_{U \setminus E}) \leq \epsilon\,,$$

so

$$L(\chi_U) + \lambda(U) = L(\chi_E) + \lambda(E) + L(\chi_{U \setminus E}) + \lambda(U \setminus E) \leq L(\chi_E) + \lambda(E) + 2\epsilon$$

holds. Thus we have

$$\mu(E) \leq L(\chi_E) + \lambda(E) + 2\epsilon\,,$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

By (4.20), we see that we obtain the same function μ on subsets of $\mathbb{R}^N$ if we define μ by setting

$$\mu(U) = L(\chi_U) + \lambda(U)\,, \tag{4.21}$$

when $U \in \mathcal{O}$, and setting

$$\mu(E) = \inf \{\, \mu(U) : U \in \mathcal{O},\ E \subseteq U \,\}\,, \tag{4.22}$$

when $E \notin \mathcal{O}$. We shall use this alternative definition. Ultimately we will show that μ is a measure. We note that the original definition of μ is useful for verifying that μ is a Radon measure.

By (4.17), we see that

$$0 \leq \mu(E) \leq 2\lambda(E)$$

holds, for every set E. In particular, μ is absolutely continuous with respect to λ. We also note that if $U, V \in \mathcal{O}$, then

$$\begin{aligned}
\mu(V) &= L(\chi_V) + \lambda(V) \\
&= L(\chi_U + \chi_{V \setminus U}) + \lambda(\,U \cup (V \setminus U)\,) \\
&= L(\chi_U) + L(\chi_{V \setminus U}) + \lambda(U) + \lambda(V \setminus U) \\
&\geq L(\chi_U) + \lambda(U) \\
&= \mu(U)
\end{aligned}$$

and

$$\begin{aligned}
\mu(U \cup V) &= L(\chi_{U \cup V}) + \lambda(U \cup V) \\
&= L(\chi_U) + L(\chi_V) - L(\chi_{U \cap V}) + \lambda(U \cup V) \\
&= L(\chi_U) + L(\chi_V) - L(\chi_{U \cap V}) + \lambda(U) + \lambda(V) - \lambda(U \cap V) \\
&= \mu(U) + \mu(V) - \mu(U \cap V) \leq \mu(U) + \mu(V)\,,
\end{aligned}$$

so λ is finitely additive and finitely subadditive on $\mathcal{O}$.

Step 7: Showing that μ is a Radon measure. First, we claim that μ is countably subadditive on $\mathcal{O}$. To see this, let a sequence $\{U_i\} \subseteq \mathcal{O}$ be given. We need to show that

$$\mu\Big(\bigcup_i U_i\Big) \leq \sum_i \mu(U_i) \tag{4.23}$$

holds.

Let $\epsilon > 0$ be arbitrary. Set

$$A_n = \big(\bigcup_{i=1}^{\infty} U_i\big) \setminus \big(\bigcup_{i=1}^{n} U_i\big) .$$

Then $\lambda(A_n) \to 0$ as $n \to \infty$. Choose n such that $\lambda(A_n) < \epsilon$. We have

$$\begin{aligned}\mu(\bigcup_{i=1}^{\infty} U_i) &= \mu(\bigcup_{i=1}^{n} U_i) + L(\chi_{A_n}) + \lambda(A_n) \\ &\leq \mu(\bigcup_{i=1}^{n} U_i) + 2\epsilon \leq 2\epsilon + \sum_{i=1}^{\infty} \mu(U_i) ,\end{aligned}$$

and the claim follows from the fact that $\epsilon > 0$ was arbitrary.

We see that μ is countably subadditive by using the same argument that showed that λ is subadditive. We can also see that Carathéodory's criterion holds for μ in the same way that we saw that it holds for λ, and we similarly conclude that λ is a Radon measure.

Step 8: Obtaining the function g. By Theorem 4.3.6, there exists a Borel function f such that

$$\mu(E) = \int_E f \, d\lambda$$

holds, for any Borel set E. Set $g = f - 1$. For $U \in \mathcal{O}$, we have

$$L(\chi_U) = \mu(U) - \lambda(U) = \int_U (f-1)\, d\lambda = \int_U g \, d\lambda .$$

For $\phi \in \mathcal{D}$, we obtain

$$L(\phi) = \int \phi \, g \, d\lambda$$

by uniformly approximating ϕ by simple functions of the form $\sum_i \alpha_i \chi_{E_i}$, with $E_i \in \mathcal{O}$, and applying (4.17). $\square$

4.5 Maximal Functions Redux

It is possible to construe the Hardy–Littlewood maximal function in the more general context of measures.

Definition 4.5.1. Let μ be a Radon measure on $\mathbb{R}^N$. If f is a μ-measurable function and $x \in \mathbb{R}^N$ then we define

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu[\mathbb{B}(x,r)]} \int_{\mathbb{B}(x,r)} |f(t)| \, d\mu(t) \, .$$

Further, and more generally, if ν is a Radon measure on $\mathbb{R}^N$ then we define

$$M_\mu \nu(x) = \sup_{r>0} \frac{\nu[\mathbb{B}(x,r)]}{\mu[\mathbb{B}(x,r)]} \, .$$

Finally, it is sometimes useful to have the noncentered maximal operator $\widetilde{M}_\mu$ defined by

$$\widetilde{M}_\mu f(x) = \sup_{\mathbb{B}(z,r) \ni x} \frac{1}{\mu[\mathbb{B}(z,r)]} \int_{\mathbb{B}(z,r)} |f(t)| \, d\mu(t) \, .$$

A similar definition may be given for the maximal function of a Radon measure.

The principal result about these maximal functions is the following:

Theorem 4.5.2. *The operator M_μ is weak type $(1, 1)$ in the sense that*

$$\mu \left\{ x \in \mathbb{R}^N : M_\mu \nu(x) > s \right\} \leq C \cdot \frac{\nu(\mathbb{R}^N)}{s} \, .$$

In particular, if $f \in L^1(\mu)$ then

$$\mu \left\{ x \in \mathbb{R}^N : M_\mu f(x) > s \right\} \leq C \cdot \frac{\|f\|_{L^1}}{s} \, .$$

In case the measure μ satisfies the enlargement condition $\mu[\mathbb{B}(x, 3r)] \leq c \cdot \mu[\mathbb{B}(x,r)]$, then we have

$$\mu \left\{ x \in \mathbb{R}^N : \widetilde{M}_\mu \nu(x) > s \right\} \leq c \cdot s^{-1} \cdot \nu \left\{ x \in \mathbb{R}^N : \widetilde{M}_\mu \nu(x) > s \right\} \, .$$

The proof of this result follows the same lines as the development of Proposition 4.1.4, and we omit the details. A full account may be found in [Mat 95].

5

Analytical Tools: The Area Formula, the Coarea Formula, and Poincaré Inequalities

5.1 The Area Formula

The main result of this section is the following theorem.

Theorem 5.1.1 (Area Formula). *If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function and $M \le N$, then*

$$\int_A J_M f(x)\, d\mathcal{L}^M(x) = \int_{\mathbb{R}^N} \operatorname{card}(A \cap f^{-1}(y))\, d\mathcal{H}^M(y) \tag{5.1}$$

holds for each Lebesgue measurable subset A of $\mathbb{R}^M$.

See Figure 5.1. Here $J_M f$ denotes the M-dimensional Jacobian of f, which will be defined below in Definition 5.1.3. In case $M = N$, the M-dimensional Jacobian agrees with the usual Jacobian $|\det(Df)|$.

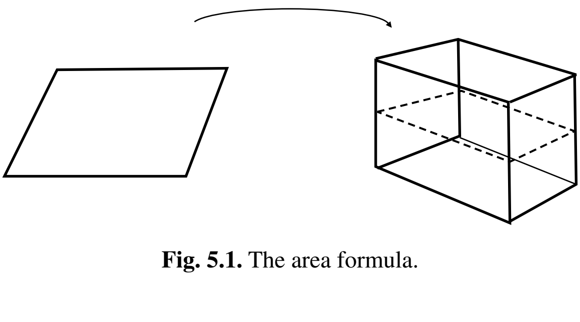

Fig. 5.1. The area formula.

The proof of the area formula separates into three fundamental parts. The first is understanding the situation for linear maps. The second is extending our understanding to the behavior of maps that are well approximated by linear maps. This second part of the proof is essentially multivariable calculus, and the area formula for C^1 maps follows readily. The third part of the proof brings in the measure theory

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_5,

that allows us to reduce the behavior of Lipschitz maps to that of maps that are well approximated by linear maps.

In the next section we will treat the coarea formula that applies to a Lipschitz map $f : \mathbb{R}^M \to \mathbb{R}^N$, but with $M \geq N$ instead of $M \leq N$. The proof of the coarea formula is similar to the proof of the area formula in that the same three steps of understanding linear maps, understanding maps well approximated by linear maps, and applying measure theory are fundamental. The discussion of linear maps in the next subsection will be applicable to both the area formula and the coarea formula.

5.1.1 Linear Maps

A key ingredient in the area formula is the K-dimensional Jacobian, which is a measure of how K-dimensional area transforms under the differential of a mapping. Since a linear map sends one parallelepiped into another, the fundamental question is, "What is the K-dimensional area of the parallelepiped determined by a set of K vectors in $\mathbb{R}^N$?" Of course, the answer is known, and G. J. Porter gave a particularly lucid derivation in [Por 96]. We follow Porter's approach in the argument given below (this argument also appeared earlier, in Section 1.4).

Since we will often need to divide by the K-dimensional area of a parallelepiped, when we say that P is a K-dimensional parallelepiped, we will assume that P *is not* contained in any $(K-1)$-dimensional subspace. That is, when P is a K-dimensional parallelepiped we mean that there are *linearly independent* vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K$ such that

$$P = \left\{ \sum_{i=1}^{K} \lambda_i \, \mathbf{v}_i : 0 \leq \lambda_i \leq 1, \text{ for } i = 1, 2, \ldots, K \right\} .$$

Proposition 5.1.2. *If*

$$\mathbf{v}_i = \begin{pmatrix} v_{1\,i} \\ v_{2\,i} \\ \vdots \\ v_{N\,i} \end{pmatrix}, \textit{ for } i = 1, 2, \ldots, K \,, \tag{5.2}$$

are vectors in $\mathbb{R}^N$, *then the parallelepiped determined by those vectors has* K*-dimensional area*

$$\sqrt{\det\left(V^{\mathrm{t}}\, V\right)}\,, \tag{5.3}$$

where V *is the* $N \times K$ *matrix having* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K$ *as its columns.*

Proof. If the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K$ are orthogonal, then the result is immediate. Thus we will reduce the general case to this special case.

Notice that Cavalieri's principle shows us that adding a multiple of $\mathbf{v}_i$ to another vector $\mathbf{v}_j$, $j \neq i$, does not change the K-dimensional area of the parallelepiped determined by the vectors. But also notice that such an operation on the vectors $\mathbf{v}_i$ is equivalent to multiplying V on the right by a $K \times K$ triangular matrix with 1's on the diagonal (upper triangular if $i < j$ and lower triangular if $i > j$). The Gram–Schmidt orthogonalization procedure is effected by a sequence of operations of precisely this type. Thus we see that there is an upper triangular matrix A with 1's on the diagonal such that VA has orthogonal columns and the columns of VA determine a parallelepiped with the same K-dimensional area as the parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$. Since the columns of VA are orthogonal, we know that $\sqrt{\det\left((VA)^{\mathrm{t}}\,(VA)\right)}$ equals the K-dimensional area of the parallelepiped determined by its columns, and thus equals the K-dimensional area of the parallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K$. Finally, we compute

$$\begin{aligned}\det\left((VA)^{\mathrm{t}}\,(VA)\right) &= \det\left(A^{\mathrm{t}}\,V^{\mathrm{t}}\,V\,A\right)\\ &= \det\left(A^{\mathrm{t}}\right)\,\det\left(V^{\mathrm{t}}\,V\right)\det(A)\\ &= \det\left(V^{\mathrm{t}}\,V\right).\end{aligned}$$ □

Definition 5.1.3. Suppose that $U \subseteq \mathbb{R}^M$, $f : U \to \mathbb{R}^N$, f is differentiable at a, and $K \le M$. We define the *K-dimensional Jacobian of f at a*, denoted by $J_K f(a)$, by setting

$$J_K f(a) = \sup\left\{\frac{\mathcal{H}^K[Df(a)(P)]}{\mathcal{H}^K[P]} : \right.$$
$$\left. P \text{ is a } K\text{-dimensional parallelepiped contained in } \mathbb{R}^M\right\}. \tag{5.4}$$

The conventional situation considered in elementary multivariable calculus is that in which $K = M = N$. In that case, it is easily seen from Proposition 5.1.2 that one may choose P to be the unit M-dimensional cube and that $J_M f(a) = J_N f(a) = |\det(Df(a))|$.

Two other special cases are of interest: They are $K = M < N$ and $M > N = K$. When $K = M < N$, again one can choose P to be the unit M-dimensional cube in $\mathbb{R}^M$. The image of P under $Df(a)$ is the parallelepiped determined by the columns of the matrix representing $Df(a)$. It follows from Proposition 5.1.2 that $J_M f(a) = \sqrt{\det\left[(Df(a))^{\mathrm{t}}\,(Df(a))\right]}$.

When $M > N = K$, then P should be chosen to lie in the orthogonal complement of the kernel of $Df(a)$. This follows because if P is any parallelepiped in $\mathbb{R}^M$, then the image under $Df(a)$ of the orthogonal projection of P onto the orthogonal complement of the kernel of $Df(a)$ is the same as the image of P under $Df(a)$, while N-dimensional area of the orthogonal projection is no larger than the N-dimensional area of P.

It is plain to see that the orthogonal complement of the kernel of $Df(a)$ is the span of the columns of $(Df(a))^{\mathrm{t}}$. If we begin with the parallelepiped determined by the columns of $(Df(a))^{\mathrm{t}}$, then that parallelepiped maps onto the parallelepiped determined by the columns of $(Df(a))\,(Df(a))^{\mathrm{t}}$. By Proposition 5.1.2, the N-dimensional area of the first parallelepiped is

$$\sqrt{\det\left[(Df(a))\,(Df(a))^{\mathrm{t}}\right]},$$

and the N-dimensional area of the second parallelepiped is

$$\sqrt{\det\left[\left((Df(a))\,(Df(a))^{\mathrm{t}}\right)^{\mathrm{t}}\,\left((Df(a))\,(Df(a))^{\mathrm{t}}\right)\right]}$$

$$= \det\left[(Df(a))\,(Df(a))^{\mathrm{t}}\right],$$

so the ratio is $J_N f(a) = \sqrt{\det\left[(Df(a))\,(Df(a))^{\mathrm{t}}\right]}$. (The preceding discussion could also have been phrased in terms of the effect of the adjoint of Df on the area of a parallelepiped in $\mathbb{R}^N$.)

We summarize the above facts in the following lemma.

Lemma 5.1.4. *Suppose that $f : \mathbb{R}^M \to \mathbb{R}^N$ is differentiable at a.*

(1) *If $M = N$, then*
$$J_M f(a) = J_N f(a) = |\det(Df(a))|. \tag{5.5}$$

(2) *If $M \le N$, then*
$$J_M f(a) = \sqrt{\det\left[(Df(a))^{\mathrm{t}}\,(Df(a))\right]}. \tag{5.6}$$

(3) *If $M \ge N$, then*
$$J_N f(a) = \sqrt{\det\left[(Df(a))\,(Df(a))^{\mathrm{t}}\right]}. \tag{5.7}$$

Remark 5.1.5. The generalized Pythagorean theorem (see Section 1.5) allows one to see that the right-hand side of either (5.6) or (5.7) is equal to the square root of the sum of the squares of the $K \times K$ minors of $Df(a)$, where $K = \min\{M, N\}$. This is the form one is naturally led to if one develops the K-dimensional Jacobian via the alternating algebra over $\mathbb{R}^M$ and $\mathbb{R}^N$ as in [Fed 69].

We will also need to make use of the polar decomposition of linear maps.

Theorem 5.1.6 (Polar Decomposition).

(1) *If $M \le N$ and $T : \mathbb{R}^M \to \mathbb{R}^N$ is linear, then there exist a symmetric linear map $S : \mathbb{R}^M \to \mathbb{R}^M$ and an orthogonal injection $U : \mathbb{R}^M \to \mathbb{R}^N$ such that $T = U \circ S$.*

(2) *If $M \ge N$ and $T : \mathbb{R}^M \to \mathbb{R}^N$ is linear, then there exist a symmetric linear map $S : \mathbb{R}^N \to \mathbb{R}^N$ and an orthogonal injection $U : \mathbb{R}^N \to \mathbb{R}^M$ such that $T = S \circ U^{\mathrm{t}}$.*

Proof.
(1) For convenience, let us first suppose that T is of full rank. The $M \times M$ matrix $T^{\mathrm{t}}\, T$ is symmetric and positive definite. So $T^{\mathrm{t}}\, T$ has a complete set of M orthonormal eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$ associated with the positive eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$.

We define $S : \mathbb{R}^M \to \mathbb{R}^M$ by setting

$$S(\mathbf{v}_i) = \sqrt{\lambda_i}\, \mathbf{v}_i \, .$$

Using the orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$, we see that S is represented by a diagonal matrix; thus S is symmetric.

We define $U : \mathbb{R}^M \to \mathbb{R}^N$ by setting

$$U(\mathbf{v}_i) = \frac{1}{\sqrt{\lambda_i}}\, T(\mathbf{v}_i) \, .$$

We calculate

$$\begin{aligned} U(\mathbf{v}_i) \cdot U(\mathbf{v}_j) &= \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}}\, T(\mathbf{v}_i) \cdot T(\mathbf{v}_j) \\ &= \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}}\, \mathbf{v}_i \cdot (T^{\mathrm{t}}\, T)(\mathbf{v}_j) \\ &= \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}}\, \lambda_j\, \mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} \, . \end{aligned}$$

Thus U is an orthogonal injection.

In case T is not of full rank, it follows that some of the λ_i's may be zero. For such an index i, we may choose $U(\mathbf{v}_i)$ arbitrarily, subject only to the requirement that $U(\mathbf{v}_1), U(\mathbf{v}_2), \ldots, U(\mathbf{v}_n)$ be an orthonormal set.

(2) We apply (1) to the mapping T^{t} to obtain a symmetric map S and an orthogonal injection U such that $T^{\mathrm{t}} = U \circ S$, but then $T = (U \circ S)^{\mathrm{t}} = S \circ U^{\mathrm{t}}$. □

The first application of the Jacobian is in the following basic lemma concerning the behavior of Lebesgue measure under a linear map.

Lemma 5.1.7. *If $A \subseteq \mathbb{R}^M$ is Lebesgue measurable and $T : \mathbb{R}^M \to \mathbb{R}^M$ is linear, then*

$$\mathcal{L}^M\big(T(A)\big) = |\det(T)|\, \mathcal{L}^M(A) \, .$$

Proof. By countable additivity, it will suffice to prove the result for bounded sets A. Given $\epsilon > 0$, we can find an open U with $A \subseteq U$ and $\mathcal{L}^M(U \setminus A) < \epsilon$. We write U as an increasing union of sets C_n such that each C_n is a union of cubes that intersect only on their faces. Then we have

$$\mathcal{L}^M(T(U)) = \lim_{n\to\infty} \mathcal{L}^M(T(C_n)) = \lim_{n\to\infty} |\det(T)| \mathcal{L}^M(C_n) = |\det(T)| \mathcal{L}^M(U) \, .$$

So we conclude that

$$\mathcal{L}^M(T(A)) \le \mathcal{L}^M(T(U)) \le |\det(T)|\, \mathcal{L}^M(U) \le |\det(T)|\, [\epsilon + \mathcal{L}^M(A)]\,.$$

Letting $\epsilon \downarrow 0$, we see that

$$\mathcal{L}^M(T(A)) \le |\det(T)|\, \mathcal{L}^M(A)\,.$$

Now we need to prove the reverse inequality. Note that if $\det(T) = 0$, then we are done. Assuming $\det(T) \neq 0$, we apply the case already proved to $T(A)$ and T^{-1} to see that

$$\mathcal{L}^M(A) = \mathcal{L}^M(T^{-1}(T(A))) \le |\det(T^{-1})|\, \mathcal{L}^M(T(A))\,.$$

The result follows since $\det(T^{-1}) = (\det(T))^{-1}$. □

Lemma 5.1.8 (Main Estimates for the Area Formula). *Suppose that $M \le N$, $T : \mathbb{R}^M \to \mathbb{R}^N$ is linear and of full rank, and that $0 < \epsilon < \frac{1}{2}$. Let Π be orthogonal projection onto the image of T. Set*

$$\lambda = \inf \left\{ |\langle T, v\rangle| : |v| = 1 \right\}. \tag{5.8}$$

If the Lebesgue measurable set $A \subseteq \mathbb{R}^M$ is such that

(1) *$Df(a)$ exists for $a \in A$,*

(2) *$\|Df(a) - T\| < \epsilon$ holds for $a \in A$,*

(3) *$|f(y) - f(a) - \langle Df(a), y - a\rangle| < \epsilon\, |y - a|$ holds for $y, a \in A$,*

(4) *$\Pi|_{f(A)}$ is one-to-one,*

then

$$\begin{aligned}\left(1 - 3\epsilon\lambda^{-1}\right)^M \cdot J_M T \cdot \mathcal{L}^M(A) &\le \mathcal{H}^M\big(f(A)\big) \\ &\le \left(1 + 2\epsilon\lambda^{-1}\right)^M \cdot J_M T \cdot \mathcal{L}^M(A)\,. \end{aligned} \tag{5.9}$$

Proof. First we bound $\mathcal{H}^M\big(f(A)\big)$ from above. We use the polar decomposition to write $T = U \circ S$, where $S : \mathbb{R}^M \to \mathbb{R}^M$ is a symmetric map and $U : \mathbb{R}^M \to \mathbb{R}^N$ is an orthogonal injection, and we note that S is nonsingular with $J_M S = J_M T$ and with $\lambda^{-1} = \|S^{-1}\|$.

Set $B = S(A)$ and $g = f \circ S^{-1}$. We know that

$$\mathcal{L}^M(B) = J_M S \cdot \mathcal{L}^M(A) = J_M T \cdot \mathcal{L}^M(A)\,.$$

We claim that

$$\mathrm{Lip}\,(g|_B) \le 1 + 2\epsilon\lambda^{-1}\,.$$

To see this, suppose $z, b \in B$. Then with $a = S^{-1}(b)$, $y = S^{-1}(z)$, it follows that $|y - a| \le \lambda^{-1}|z - b|$. Therefore we have

$$
\begin{aligned}
&|g(z) - g(b)| \\
&\le |g(z) - g(b) - \langle Dg(b), z-b\rangle| + |\langle Dg(b) - U, z-b\rangle| + |\langle U, z-b\rangle| \\
&= |f(y) - f(a) - \langle Df(a), y-a\rangle| \\
&\qquad\qquad + |\langle (Df(a) - T)\circ S^{-1}, z-b\rangle| + |z-b| \\
&\le \epsilon\,|y-a| + \|Df(a) - T\| \cdot \|S^{-1}\| \cdot |z-b| + |z-b| \\
&\le \left(1 + 2\epsilon\lambda^{-1}\right)|z-b|\,. \qquad (5.10)
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\mathcal{H}^M\big(f(A)\big) &= \mathcal{H}^M\big(g(B)\big) \\
&\le \left(1 + 2\epsilon\lambda^{-1}\right)^M \cdot \mathcal{L}^M(B) \\
&= \left(1 + 2\epsilon\lambda^{-1}\right)^M \cdot J_M T \cdot \mathcal{L}^M(A)\,.
\end{aligned}
$$

Next we bound $\mathcal{H}^M\big(f(A)\big)$ from below. We continue to use the same notation for the polar decomposition and we will continue to write $g = f \circ S^{-1}$. Set $C = \Pi\big(f(A)\big) = \Pi\big(g(B)\big)$ and $h = \big(\Pi \circ g|_B\big)^{-1}$. We claim that

$$
\operatorname{Lip}(h|_C) \le \left(1 - 3\epsilon\lambda^{-1}\right)^{-1}.
$$

To see this, suppose $w, c \in C$. Let $b \in B$ be such that $\Pi \circ g(b) = c$ and $z \in B$ be such that $\Pi \circ g(z) = w$. Arguing as we did to obtain the upper bound (5.10), but with some obvious changes, we see that

$$
|g(z) - g(b)| \ge \left(1 - 2\epsilon\lambda^{-1}\right)|z-b|\,.
$$

Also we have

$$
\begin{aligned}
\epsilon\lambda^{-1}|z-b| &\ge |g(z) - g(b) - \langle Dg(b), z-b\rangle| \\
&= |\Pi\big(g(z) - g(b) - \langle Dg(b), z-b\rangle\big) \\
&\qquad\qquad + \Pi^{\perp}\big(g(z) - g(b) - \langle Dg(b), z-b\rangle\big)| \\
&\ge |\Pi^{\perp}\big(g(z) - g(b) - \langle Dg(b), z-b\rangle\big)| \\
&= |\Pi^{\perp}\big(g(z) - g(b)\big)|\,.
\end{aligned}
$$

Thus we conclude that

$$
\begin{aligned}
|\Pi(g(z)) - \Pi(g(b))| &\ge |g(z) - g(b)| - |\Pi^{\perp}\big(g(z) - g(b)\big)| \\
&\ge \left(1 - 2\epsilon\lambda^{-1}\right)|z-b| - \epsilon\lambda^{-1}\,|z-b|\,.
\end{aligned}
$$

Finally, we calculate that

$$\begin{aligned} J_M T \cdot \mathcal{H}^M(A) &= \mathcal{L}^M(B) \\ &\le \left(1 - 3\epsilon\lambda^{-1}\right)^M \cdot \mathcal{L}^M(C) \\ &\le \left(1 - 3\epsilon\lambda^{-1}\right)^M \cdot \mathcal{H}^M\left(f(A)\right). \end{aligned}$$ □

5.1.2 C^1 Functions

Now we can prove the area formula for C^1 functions.

Theorem 5.1.9. *Suppose that $M \le N$. If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a C^1 function, then*

$$\int_A J_M f(x)\, d\mathcal{L}^M(x) = \int_{\mathbb{R}^N} \operatorname{card}(A \cap f^{-1}(y))\, d\mathcal{H}^M(y)$$

holds for each Lebesgue measurable subset A of $\mathbb{R}^M$.

Proof. By countable additivity, it will suffice to prove the result for bounded sets A. We first prove the result under the additional assumptions that f is one-to-one and that $J_M f(a) > 0$ holds at every point of A.

It is plain that, for any $\epsilon > 0$, every subset of A with sufficiently small diameter satisfies conditions (1)–(3) of Lemma 5.1.8 for some full-rank linear $T : \mathbb{R}^M \to \mathbb{R}^N$—namely, we can choose T to be Df at any point in such a sufficiently small set. Since Df on A is the restriction of a continuous function, we can find a positive lower bound for λ in (5.8). To see that condition (4) of Lemma 5.1.8 is also satisfied on a subset of A of small enough diameter, we suppose, to the contrary, that $\Pi \circ f(y) = \Pi \circ f(z)$; we show that in this case, $\epsilon > 0$ can be chosen small enough compared to λ that conditions (1)–(3) lead to a contradiction. Using (1)–(3), we estimate

$$\begin{aligned} |\langle T, y - z\rangle| &= |\Pi\, \langle T, y - z\rangle)| \\ &\le |\Pi\, \langle T - Df(a), y - z\rangle| + |\Pi\, \langle Df(a) - Df(z), y - z\rangle| \\ &\qquad + |\Pi\, \langle Df(z), y - z\rangle| \\ &\le \|T - Df(a)\|\, |y - z| + \|Df(a) - Df(z)\|\, |y - z| \\ &\qquad + |\Pi\, \langle Df(z), y - z\rangle| \\ &= \|T - Df(a)\|\, |y - z| + \|Df(a) - Df(z)\|\, |y - z| \\ &\qquad + |\Pi(f(y) - f(z) - \langle Df(z), y - z\rangle)| \\ &\le \|T - Df(a)\|\, |y - z| + \|Df(a) - Df(z)\|\, |y - z| \\ &\qquad + |f(y) - f(z) - \langle Df(a), y - z\rangle|\,. \end{aligned}$$

By choosing a, y, z in a small enough set we can bound the right-hand side of the preceding inequality above by $3\,\epsilon\,|y-z|$, while the left-hand side is bounded below by $\lambda\,|y-z|$. Choosing ϵ smaller than $\frac{1}{3}\,\lambda$ gives a contradiction. Thus (4) also must hold on subsets of small enough diameter, and the result follows by decomposing A into such sufficiently small sets.

In case f is not necessarily one-to-one, but still assuming $J_M f(a) > 0$ holds at every point of A, there is $\sigma > 0$ such that f is one-to-one in any ball of radius σ about any point in A. Write

$$A = \bigcup_j A_j\,,$$

where the sets A_j, $j = 1, 2, \ldots$, are pairwise disjoint $\mathcal{H}^M$-measurable sets all having diameter less than σ. Then we have

$$\sum_j \chi_{f(A_{i,j})}(y) = \operatorname{card}(A \cap f^{-1}(y)) \;\text{ for each } y \in \mathbb{R}^N\,.$$

We conclude that

$$\begin{aligned}
\int_A J_M f(x)\, d\mathcal{L}^M(x) &= \sum_j \int_{A_j} J_M f(x)\, d\mathcal{L}^M(x)\\
&= \sum_j \mathcal{H}^M[f(A_{i,j})]\\
&= \int_{\mathbb{R}^N} \sum_j \chi_{f(A_{i,j})}\, d\mathcal{H}^M\\
&= \int_{\mathbb{R}^N} \operatorname{card}(A \cap f^{-1}(y))\, d\mathcal{H}^M\,.
\end{aligned}$$

To complete the proof, we need to show that the image of a set on which $J_M f = 0$ has measure zero. That fact follows by defining $f_\epsilon : \mathbb{R}^M \to \mathbb{R}^{M+N}$ by

$$x \longmapsto \Big(\epsilon x,\, f(x)\Big)\,.$$

This definition of f_ϵ gives us the full-rank hypothesis, but increases the Jacobian only by a bounded multiple of ϵ. The image of f is the orthogonal projection of the image of f_ϵ, and thus its Hausdorff measure is no larger than the Hausdorff measure of the image of f_ϵ. By letting ϵ decrease to 0, we conclude that the Hausdorff measure of the image of f is 0. □

The last part of the preceding proof gives us the next corollary, which is known as Sard's theorem.[1] The sharp version of Sard's theorem, the Morse–Sard–Federer theorem, can be found in [Fed 69, 3.4.3].

Corollary 5.1.10. *Suppose that $M \le N$. If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a C^1 function and $A = \{x : J_M f(x) = 0\}$, then $\mathcal{H}^M[\,f(A)\,] = 0$.*

[1] Arthur Sard (1909–1980).

5.1.3 Rademacher's Theorem

Theorem 5.1.11 (Rademacher's Theorem[2]). *If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function, then f is differentiable $\mathcal{L}^M$-almost everywhere and the differential of f is a measurable function.*

Proof. We may assume $N = 1$. We use induction on M. In case $M = 1$, the result follows from the classical theorem stating that an absolutely continuous function from $\mathbb{R}$ to $\mathbb{R}$ is differentiable $\mathcal{L}^1$-almost everywhere.

We consider the inductive step $M > 1$. Note that if $M - 1$ variables are held constant, then, as a function of the one remaining variable, f is absolutely continuous. By Fubini's theorem, we see that all M partial derivatives of f are defined $\mathcal{L}^M$-almost everywhere and are measurable functions. The goal is to show that these partial derivatives actually represent the differential at almost every point.

Let us write $\mathbb{R}^M = \mathbb{R}^{M-1} \times \mathbb{R}$ and denote points $p \in \mathbb{R}^{M-1} \times \mathbb{R}$ by $p = (x, y)$, $x \in \mathbb{R}^{M-1}$, $y \in \mathbb{R}$. We consider a point $p_0 = (x_0, y_0)$ at which the following two conditions are satisfied:

(1) As a function of the first $M - 1$ variables, f is differentiable.
(2) All M partial derivatives of f exist and are approximately continuous (see Definition 4.1.7).

For convenience of notation, we assume that $f(p_0) = 0$, that $p_0 = (0, 0)$, and that all the partial derivatives of f at p_0 vanish.

Fix an ϵ with $1 > \epsilon > 0$. By (1), we can choose $r_0 > 0$ such that $|x| < r_0$ implies that

$$|f(x, 0)| \le \epsilon |x|$$

holds. By (2), the M-dimensional density at $(0, 0)$ of

$$\left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon \right\}$$

is zero. Thus, by choosing a smaller value for r_0 if necessary, we may assume that for $0 < r < r_0$,

$$\mathcal{L}^M \left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon,\ |x'| < 2r,\ -2r < y' < 2r \right\} \le \tfrac{1}{2} \Omega_{M-1} \cdot \epsilon^M r^M \tag{5.11}$$

holds.

Now consider $(0, 0) \neq (x, y) \in \mathbb{R}^{M-1} \times \mathbb{R}$ with $|x| < r_0$ and $|y| < r_0$. Set $r = \max\{\, |x|,\ |y| \,\}$ If for every $x' \in \mathbb{R}^{M-1}$ with $|x' - x| < \epsilon r$, we have

$$\mathcal{L}^1 \left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon,\ -2r < y' < 2r \right\} \ge \epsilon r\,,$$

then we can estimate

[2] Hans Rademacher (1892–1969).

$$
\begin{aligned}
&\mathcal{L}^M \left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon,\ |x'| < 2r,\ -2r < y' < 2r \right\} \\
&\quad \geq \mathcal{L}^M \left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon,\ |x' - x| < \epsilon r,\ -2r < y' < 2r \right\} \\
&\quad \geq \epsilon r \cdot \mathcal{L}^{M-1}\{ x' \in \mathbb{R}^{M-1} : |x' - x| < r \} \\
&\quad \geq \Omega_{M-1} \cdot \epsilon^M r^M ,
\end{aligned}
$$

contradicting (5.11).

By the last paragraph, there exists $x' \in \mathbb{R}^{M-1}$, with $|x' - x| < \epsilon r$, such that

$$
\mathcal{L}^1 \left\{ (x', y') : \left| \frac{\partial f}{\partial y}(x', y') \right| > \epsilon,\ -2r < y' < 2r \right\} < \epsilon r
$$

holds; select and fix such an x'. We have

$$
\begin{aligned}
|f(x', y) - f(x', 0)| &= \left| \int_0^y \frac{\partial f}{\partial y}(x', \eta)\, d\mathcal{L}^1(\eta) \right| \\
&\leq \epsilon |y| + M\epsilon r \\
&< (M + 1)\epsilon r , \qquad (5.12)
\end{aligned}
$$

where we have used the fact that $\left| \frac{\partial f}{\partial y}(x', \eta) \right| \leq M$ holds for $\mathcal{L}^1$-almost all η. Also, we have

$$
|f(x, y) - f(x', y)| \leq M|x - x'| < M\epsilon r , \qquad (5.13)
$$

$$
|f(x, 0) - f(x', 0)| \leq M|x - x'| < M\epsilon r , \qquad (5.14)
$$

$$
|f(x, 0)| \leq \epsilon |x| < \epsilon r. \qquad (5.15)
$$

Combining (5.12), (5.13), (5.14), and (5.15), we obtain

$$
|f(x, y)| \leq (3M + 2)\epsilon r ,
$$

from which it follows that $Df(0, 0) = 0$. □

As a consequence of Rademacher's theorem and the Whitney extension theorem[3] (see [Fed 69] or [KPk 99]), we have the following approximation theorem for Lipschitz functions.

Theorem 5.1.12. *If $f : \mathbb{R}^N \to \mathbb{R}^\nu$ is Lipschitz and if $\epsilon > 0$, then there exists a C^1 function $g : \mathbb{R}^N \to \mathbb{R}^\nu$ for which*

$$
\mathcal{L}^N\{x : f(x) \neq g(x)\} \leq \epsilon ,
$$

$$
\mathcal{L}^N\{x : Df(x) \neq Dg(x)\} \leq \epsilon .
$$

[3] Hassler Whitney (1907–1989).

Proof. It will suffice to prove the result when $\nu = 1$.

Recall that the Whitney extension theorem for C^1 functions tells us the following:

Let $A \subseteq \mathbb{R}^N$ be closed. Suppose that $f : A \to \mathbb{R}$ and $v : A \to \mathbb{R}^N$ are continuous. If the limit of

$$\frac{f(y) - f(x) - v(x) \cdot (y - x)}{|y - x|}$$

is zero as $x, y \in A$, with $x \neq y$, approach any point of A, then there exists a C^1 function $g : \mathbb{R}^N \to \mathbb{R}$ with $g(a) = f(a)$ and grad $g(a) = v(a)$ *for all $a \in A$.*

By Rademacher's theorem applied to f and Lusin's theorem (i.e., Theorem 1.3.4) applied to grad f (for $\mathcal{L}^N$ on $\mathbb{R}^N$, Lusin's theorem is easily seen to be applicable to sets with infinite measure), there is a closed set $B \subseteq \mathbb{R}^N$ with $\mathcal{L}^N(\mathbb{R}^N \setminus B) < \epsilon/2$ such that grad f exists and is continuous on B. We set $v(x) = \operatorname{grad} f(x)$ and

$$h_k(x) = \sup \left\{ \frac{f(y) - f(x) - v(x) \cdot (y - x)}{|y - x|} : y \in B,\ 0 < |y - x| < 1/k \right\},$$

for $x \in B$, $k = 1, 2, \ldots$. Since f is differentiable on B, $h_k(x) \to 0$ for each $x \in B$. By Egorov's theorem (i.e., Theorem 1.3.3), there exists a closed set $A \subseteq B$ with $\mathcal{L}^N(B \setminus A) \leq \epsilon/2$ such that h_k converges to 0 uniformly on compact subsets of A. Thus we can apply Whitney's extension theorem to f and v on A to obtain the desired function g. □

Proof of the Area Formula. As usual, it will suffice to consider the case in which A is bounded. Use Theorem 5.1.12 to replace f by the C^1 function g when A is replaced by a set B with $\mathcal{L}^M(A \setminus B) < \epsilon$. Theorem 5.1.9 applies to g on B.

To complete the proof, observe that for any $A_j \subseteq A$, it holds that $\mathcal{H}^M[f(A_j)] \leq (\operatorname{Lip} f)^M\, \mathcal{L}^M(A_j)$. In particular, by decomposing $A \setminus B$ into pairwise disjoint sets A_j on which f is one-to-one, we obtain

$$\int_{\mathbb{R}^N} \operatorname{card}((A \setminus B) \cap f^{-1}(y))\, d\mathcal{H}^M(y) \leq (\operatorname{Lip} f)^M\, \epsilon\,. \qquad \square$$

Corollary 5.1.13. *If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function and $M \leq N$, then*

$$\int_A g(x)\, J_M f(x)\, d\mathcal{L}^M(x) = \int_{\mathbb{R}^N} \sum_{x \in A \cap f^{-1}(y)} g(x)\, d\mathcal{H}^M(y) \tag{5.16}$$

holds for each Lebesgue measurable subset A of $\mathbb{R}^M$ and each nonnegative $\mathcal{L}^M$-measurable function $g : A \to \mathbb{R}$.

Proof. Approximate g by simple functions. □

5.2 The Coarea Formula

The main result of this section is the following theorem.

Theorem 5.2.1 (Coarea Formula). *If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function and $M \geq N$, then*

$$\int_A J_N f(x)\, d\mathcal{L}^M(x) = \int_{\mathbb{R}^N} \mathcal{H}^{M-N}(A \cap f^{-1}(y))\, d\mathcal{L}^N(y) \tag{5.17}$$

holds for each Lebesgue measurable subset A of $\mathbb{R}^M$.

See Figure 5.2. Here $J_N f$ denotes the N-dimensional Jacobian of f, which was defined in the previous section in Definition 5.1.3 and which was seen in (5.7) to be given by

$$J_N f(a) = \sqrt{\det\left[(Df(a)) \cdot (Df(a))^{\mathrm{t}}\right]}.$$

In case $M = N$, the N-dimensional Jacobian agrees with the usual Jacobian $|\det(Df)|$, and the area and coarea formulas coincide. In case $M > N$, and $f : \mathbb{R}^M = \mathbb{R}^N \times \mathbb{R}^{M-N} \to \mathbb{R}^N$ is orthogonal projection onto the first factor, then the coarea formula simplifies to Fubini's theorem; thus one can think of the coarea formula as a generalization of Fubini's theorem to functions more complicated than orthogonal projection. The coarea formula was first proved in [Fed 59].

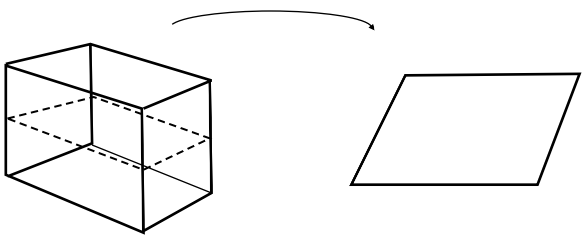

Fig. 5.2. The coarea formula.

As in the proof of the area formula, the proof of the coarea formula separates into three fundamental parts. The first is to understand the situation for linear maps. This was done in the previous section. The second part is to extend our understanding to the behavior of maps that are well approximated by linear maps. The third part of the proof brings in the measure theory that allows us to reduce the behavior of Lipschitz maps to that of maps that are well approximated by linear maps.

Main Estimates for the Coarea Formula

Lemma 5.2.2. *Suppose that $M > N$, $f : \mathbb{R}^M \to \mathbb{R}^N$, $U : \mathbb{R}^N \to \mathbb{R}^M$ is orthogonal, and $0 < \epsilon < 1/2$. If the Lebesgue measurable set $A \subseteq \mathbb{R}^M$ is such that*

(1) $Df(a)$ *exists for* $a \in A$,

(2) $\|Df(a) - U^{\mathtt{t}}\| < \epsilon$ *holds for* $a \in A$,

(3) $|f(y) - f(a) - \langle Df(a), y - a\rangle| < \epsilon\, |y - a|$ *holds for* $y, a \in A$,

then

$$
\begin{aligned}
(1-2\epsilon)^M \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^N(y) &\le \int_A J_M f(a)\, d\mathcal{L}^M(a) \\
&\le \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^N(y)\,. \qquad (5.18)
\end{aligned}
$$

Proof. The N columns of the matrix representing U form an orthonormal set, but not a full basis. By completing that set of vectors to an orthonormal basis, we can obtain a complementary set of $M - N$ orthonormal vectors. Those $M - N$ vectors can be used as the columns of a matrix representing an orthogonal map $V : \mathbb{R}^{M-N} \to \mathbb{R}^M$ such that $\ker(U^{\mathtt{t}})$ and $\ker(V^{\mathtt{t}})$ are orthogonal complements.

Define $F : \mathbb{R}^M \to \mathbb{R}^N \times \mathbb{R}^{M-N}$ by setting

$$
F(x) = \big(f(x), V^{\mathtt{t}}(x)\big),
$$

and let $\Pi : \mathbb{R}^N \times \mathbb{R}^{M-N} \to \mathbb{R}^N$ be projection on the first factor. It is easy to see that

$$
J_M F = J_N f.
$$

Subsequently we will show that $F|_A$ is one-to-one, so that by the area formula,

$$
\mathcal{L}^M[F(A)] = \int_A J_M F\, d\mathcal{L}^M = \int_A J_N f\, d\mathcal{L}^M\,.
$$

Thus, using Fubini's theorem, we have

$$
\begin{aligned}
\int_A J_N f\, d\mathcal{L}^M &= \mathcal{L}^M\big[F(A)\big] \\
&= \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big[F(A) \cap \Pi^{-1}(z)\big]\, d\mathcal{L}^N(z) \\
&= \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big[F(A \cap f^{-1}(z))\big]\, d\mathcal{L}^N(z)\,.
\end{aligned}
$$

To complete the proof, we show $F|_A$ to be one-to-one and estimate the Lipschitz constant of F on $A \cap f^{-1}(z)$ and the Lipschitz constant of F^{-1} on $F(A \cap f^{-1}(z))$. Suppose that $a, y \in A \cap f^{-1}(z)$. Then

$$
F(a) = (f(a), V^{\mathtt{t}}(a)) = (z, V^{\mathtt{t}}(a)) \quad \text{and} \quad F(y) = (f(y), V^{\mathtt{t}}(y)) = (z, V^{\mathtt{t}}(y))\,.
$$

We should like to compare $|a - y|$ and $|F(a) - F(y)|$. But the first components are the same, so

$$|F(a) - F(y)| = |V^{\mathrm{t}}(a) - V^{\mathrm{t}}(y)|.$$

On the one hand, V^{t} is distance-decreasing, so

$$|F(a) - F(y)| \le |a - y|.$$

On the other hand,

$$\begin{aligned} |\langle U^{\mathrm{t}}, y - a\rangle| &\le |\langle Df(a), y - a\rangle| + \|Df(a) - U^{\mathrm{t}}\|\,|y - a| \\ &= |f(y) - f(a) - \langle Df(a), y - a\rangle| + \|Df(a) - U^{\mathrm{t}}\|\,|y - a| \\ &< 2\epsilon|y - a|, \end{aligned}$$

and

$$|y - a|^2 = |V^{\mathrm{t}}(a) - V^{\mathrm{t}}(y)|^2 + |\langle U^{\mathrm{t}}, y - a\rangle|^2,$$

so

$$|V^{\mathrm{t}}(a) - V^{\mathrm{t}}(y)|^2 \ge |y - a|^2\,(1 - 4\epsilon^2).$$

Thus we have

$$\sqrt{1 - 4\epsilon^2}\;|y - a| \le |F(y) - F(a)| \le |y - a|\,,$$

so we see that F is one-to-one and we have obtained bounds on the Lipschitz constants of both F and F^{-1}. Finally, we note that $1 - 2\epsilon < \sqrt{1 - 4\epsilon^2}$. □

Corollary 5.2.3. *Suppose that $M > N$, $f : \mathbb{R}^M \to \mathbb{R}^N$, $T : \mathbb{R}^M \to \mathbb{R}^N$ is of rank N, and $0 < \epsilon < 1/2$. If the Lebesgue measurable set $A \subseteq \mathbb{R}^M$ is such that*

(1) *$Df(a)$ exists for $a \in A$,*

(2) *$\|Df(a) - T\| < \epsilon$ holds for $a \in A$,*

(3) *$|f(y) - f(a) - \langle Df(a), y - a\rangle| < \epsilon\,|y - a|$ holds for $y, a \in A$,*

then

$$\begin{aligned} (1 - 2\epsilon)^M \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^N(y) &\le \int_A J_M f(a)\, d\mathcal{L}^M(a) \\ &\le \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^N(y)\,. \end{aligned} \tag{5.19}$$

Proof. By the polar decomposition (Theorem 5.1.6), there exists a symmetric linear map $S : \mathbb{R}^N \to \mathbb{R}^N$ and an orthogonal map $U : \mathbb{R}^N \to \mathbb{R}^M$ such that $T = S \circ U^{\mathrm{t}}$. Set $g = S^{-1} \circ f$. Then we apply the lemma to g and U to obtain

$$\begin{aligned} (1 - 2\epsilon)^M \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap g^{-1}(z)\big)\, d\mathcal{L}^N(z) &\le \int_A J_M g(a)\, d\mathcal{L}^M(a) \\ &\le \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap g^{-1}(z)\big)\, d\mathcal{L}^N(z). \end{aligned} \tag{5.20}$$

Notice that if $y = S(z)$, then

$$A \cap g^{-1}(z) = A \cap f^{-1}(y),$$

so by the change of variables formula in $\mathbb{R}^N$ applied to the mapping S, we have

$$\int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap g^{-1}(z)\big)\, J_N S \, d\mathcal{L}^N(z) = \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^N(y).$$

Also we have $J_N S \, J_M g = J_M f$, so

$$\int_A J_N g \, J_M g(a)\, d\mathcal{L}^M(a) = \int_A J_M f(a)\, d\mathcal{L}^M(a)$$

holds. By multiplying all three terms in (5.20) by $J_N S$, we obtain (5.19). □

5.2.1 Measure Theory of Lipschitz Maps

We need to verify that the integrand on the right-hand side of (5.17) is measurable. (The measurability of the integrand on the left-hand side of (5.17) is given by Rademacher's theorem, i.e., Theorem 5.1.11.) First we obtain a useful preliminary estimate that generalizes a result originally proved in [EH 43].

Lemma 5.2.4. *Suppose* $0 \le N \le M < \infty$. *There exists a constant* $C(M,N)$ *such that the following statement is true: If* $f : \mathbb{R}^M \to \mathbb{R}^N$ *is a Lipschitz function and* $A \subseteq \mathbb{R}^M$, *then*

$$\overline{\int}_{\mathbb{R}^N} \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)\, d\mathcal{H}^N y \le C(M,N)\,\big[\mathrm{Lip}\,(f)\big]^N \,\mathcal{H}^M(A) \tag{5.21}$$

holds.

Proof. We may assume that the right-hand side of (5.21) is finite.

Fix $\sigma > 0$. By the definition of Hausdorff measure, there exists a cover of A by closed sets $S_1, S_2, \ldots$, all having diameter less than σ, such that

$$\sum_i \Omega_M \left(\frac{\mathrm{diam}\,(S_i)}{2}\right)^M \le \mathcal{H}^M(A) + \sigma.$$

For $y \in \mathbb{R}^N$ we observe that

$$\begin{aligned}
\mathcal{H}^{M-N}_\sigma(A \cap f^{-1}(y)) &\le \sum_{\{i : S_i \cap f^{-1}(y) \neq \emptyset\}} \Omega_{M-N} \left(\frac{\mathrm{diam}\,(S_i)}{2}\right)^{M-N} \\
&= 2^{N-M}\, \Omega_{M-N} \sum_i \big(\mathrm{diam}\,(S_i)\big)^{M-N} \chi_{f(S_i)}(y).
\end{aligned}$$

Note also that if $p \in S_i$, then

$$f(S_i) \subseteq \overline{\mathbb{B}}\Big(f(p), [\mathrm{Lip}\,(f)]\,\mathrm{diam}\,(S_i)\Big),$$

so

$$\int_{\mathbb{R}^N} \chi_{f(S_i)}\, d\mathcal{H}^N \leq [\mathrm{Lip}\,(f)]^N\, \Omega_N \Big(\mathrm{diam}\,(S_i)\Big)^N .$$

Thus we have

$$\begin{aligned}
&\overline{\int}_{\mathbb{R}^N} \mathcal{H}^{M-N}_\sigma\big(A \cap f^{-1}(y)\big)\, d\mathcal{H}^N y \\
&\leq 2^{N-M}\, \Omega_{M-N} \sum_i \Big(\mathrm{diam}\,(S_i)\Big)^{M-N} \int_{\mathbb{R}^N} \chi_{f(S_i)}\, d\mathcal{H}^N \\
&\leq 2^{N-M}\, \Omega_{M-N}\, \Omega_N\, [\mathrm{Lip}\,(f)]^N \sum_i \Big(\mathrm{diam}\,(S_i)\Big)^N \\
&\leq 2^N\, \frac{\Omega_{M-N}\, \Omega_N}{\Omega_M} \Big(\mathcal{H}^M(A) + \sigma\Big).
\end{aligned}$$

The result follows by letting σ decrease to 0. □

Lemma 5.2.5. *Suppose that $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function. If $A \subseteq \mathbb{R}^M$ is $\mathcal{H}^M$-measurable, then the mapping*

$$\mathbb{R}^N \ni y \longmapsto \mathcal{H}^{M-N}\big(A \cap f^{-1}(y)\big)$$

is $\mathcal{H}^N$-measurable.

Proof. By the previous lemma, we can ignore sets of arbitrarily small measure; hence we may and shall assume that A is compact.

Observe that, for $U \subseteq \mathbb{R}^M$,

$$f(A) \bigcap \{\, y : f^{-1}(y) \cap A \subseteq U \,\} = f(A) \Big\backslash f(A \setminus U)\,. \tag{5.22}$$

Additionally, note that if $U \subseteq \mathbb{R}^M$ is open, then $f(A)$ and $f(A \setminus U)$ are compact, and thus the set in (5.22) is a Borel subset of $\mathbb{R}^N$.

Let $\mathcal{U}$ denote the family of open subsets of $\mathbb{R}^N$ that are finite unions of open balls with rational radii and centers in $\mathbb{Q}^N$.

We will show that for $t \in \mathbb{R}$, $\{y : \mathcal{H}^{M-N}(A \cap f^{-1}(y)) \leq t\}$ is a Borel subset of $\mathbb{R}^N$. For $t < 0$, we have $\{y : \mathcal{H}^{M-N}(A \cap f^{-1}(y)) \leq t\} = \emptyset$, so we may assume that $t \geq 0$.

For each $i = 1, 2, \ldots$, let $\mathcal{F}_i$ denote the collection of finite subfamilies of $\mathcal{U}$ such that $\{\, U_{i,1}, U_{i,2}, \ldots, U_{i,k_j} \,\} \in \mathcal{F}_i$ if and only if

$$\mathrm{diam}\,(U_{i,j}) < 1/i, \ \text{ for } j = 1, 2, \ldots, k_j,$$

$$\sum_{j=1}^{k_i} \Omega_{M-N} \left(\frac{\mathrm{diam}\,(U_{i,j})}{2}\right)^{M-N} \leq t + \frac{1}{i}\,.$$

Since $\mathcal{F}_i$ is at most countable, we see that

$$B_i = \bigcup_{\{U_{i,1},\ldots,U_{i,k_i}\}\in\mathcal{F}_i} f(A) \Big\backslash f(A \setminus \cup_{j=1}^{k_i} U_{i,j}) \tag{5.23}$$

is a Borel subset of $\mathbb{R}^N$. Finally, we observe that

$$\begin{aligned} &\{\, y : \mathcal{H}^{M-N}(A \cap f^{-1}(y)) \le t \,\} \\ &= \Big[\mathbb{R}^N \setminus f(A)\Big] \bigcup \Big[f(A) \bigcap \{\, y : \mathcal{H}^{M-N}(A \cap f^{-1}(y)) \le t \,\}\Big], \end{aligned}$$

and that $f(A) \bigcap \{\, y : \mathcal{H}^{M-N}(A \cap f^{-1}(y)) \le t \,\}$ is the intersection of the sets B_i in (5.23). □

5.2.2 Proof of the Coarea Formula

By Theorem 5.1.11 and (5.21), we may assume that $Df(a)$ exists at every point $a \in A$. We first prove the result under the additional assumption that $J_N f(a) > 0$ at every point of A. By Lusin's theorem (i.e., Theorem 1.3.4), we may assume that $Df(a)$ is the restriction to A of a continuous function. By Egorov's theorem (i.e., Theorem 1.3.3) we may suppose that

$$\frac{|f(y) - f(a) - \langle Df(a), y - a\rangle|}{|y-a|}$$

converges uniformly to 0 as $y \in A$ approaches $a \in A$. It is plain that, for any $\epsilon > 0$, conditions (1)–(3) of Corollary 5.2.3 are satisfied in any subset of A that has small enough diameter.

Finally, to complete the proof, we need to consider the case in which $J_N f = 0$ holds on all of A. In that case, the left-hand side of (5.17) is 0. We need to show that the right-hand side of (5.17) also equals 0. To this end, consider $f_\epsilon : \mathbb{R}^{M+N} \to \mathbb{R}^N$ defined by

$$(x, y) \longmapsto f(x) + \epsilon y.$$

We can apply what has already been proved to the set

$$A \times [-1, 1]^N \subseteq \mathbb{R}^M \times \mathbb{R}^N.$$

We have $\mathcal{L}^{M+N}(A \times [-1,1]^N) = 2^N \mathcal{L}^M(A)$, $J_N f_\epsilon \le \epsilon\,[\epsilon + \mathrm{Lip}\,(f)]^{N-1}$, and

$$\int_{A\times[-1,1]^N} J_N f_\epsilon \, d\mathcal{L}^{M+N} = \int_{\mathbb{R}^N} \mathcal{H}^M\Big[\big(A \times [-1,1]^N\big) \cap f_\epsilon^{-1}(z)\Big]\, d\mathcal{L}^N(z)\,.$$

By (5.21) we observe that

$$\begin{aligned} &C(M,N)\,\mathcal{H}^M\Big[\big(A \times [-1,1]^N\big) \cap f_\epsilon^{-1}(z)\Big] \\ &\ge \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\Big[\big(A \times [-1,1]^N\big) \cap f_\epsilon^{-1}(z) \cap \Pi^{-1}(y)\Big]\, d\mathcal{L}^N(y) \\ &= \int_{[-1,1]^N} \mathcal{H}^{M-N}\big[A \cap f^{-1}(z - \epsilon y)\big]\, d\mathcal{L}^N(y)\,. \end{aligned}$$

Thus

$$
\begin{aligned}
&2^N \mathcal{L}^M(A)\,\epsilon\,[\epsilon + \mathrm{Lip}\,(f)]^{N-1} \\
&\geq \int_{A\times[-1,1]^N} J_N f_\epsilon \, d\mathcal{L}^{M+N} \\
&\geq \frac{1}{C(M,N)} \int_{\mathbb{R}^N} \int_{[-1,1]^N} \mathcal{H}^{M-N}\big[A \cap f^{-1}(z-\epsilon y)\big]\, d\mathcal{L}^N(y)\, d\mathcal{L}^N(z) \\
&= \frac{1}{C(M,N)} \int_{[-1,1]^N} \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big[A \cap f^{-1}(z-\epsilon y)\big]\, d\mathcal{L}^N(z)\, d\mathcal{L}^N(y) \\
&= \frac{2^N}{C(M,N)} \int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big[A \cap f^{-1}(z)\big]\, d\mathcal{L}^N(z)
\end{aligned}
$$

holds, where the last equation holds by translation invariance. Letting $\epsilon \downarrow 0$, we see that

$$
\int_{\mathbb{R}^N} \mathcal{H}^{M-N}\big[A \cap f^{-1}(z)\big]\, d\mathcal{L}^N(z) = 0\,. \qquad \square
$$

Corollary 5.2.6. *If $f : \mathbb{R}^M \to \mathbb{R}^N$ is a Lipschitz function and $M \geq N$, then*

$$
\int_A g(x)\, J_N f(x)\, d\mathcal{L}^M(x) = \int_{\mathbb{R}^N} \int_{A\cap f^{-1}(y)} g\, d\mathcal{H}^{M-N}\, d\mathcal{L}^N(y) \tag{5.24}
$$

holds for each Lebesgue measurable subset A of $\mathbb{R}^M$ and each nonnegative $\mathcal{L}^M$-measurable function $g : A \to \mathbb{R}$.

Remark 5.2.7. Observe that when $M = \nu$ and $g \equiv 1$, the integral with respect to 0-dimensional Hausdorff measure over $A \cap f^{-1}(y)$ gives the cardinality of $A \cap f^{-1}(y)$.

Proof. Approximate g by simple functions. $\square$

5.3 The Area and Coarea Formulas for C^1 Submanifolds

Definition 5.3.1. By an *M-dimensional C^1 submanifold of* $\mathbb{R}^N$ we will mean a set $S \subseteq \mathbb{R}^N$ for which each point has an open neighborhood V in $\mathbb{R}^N$ such that there exists a one-to-one C^1 map $\phi : U \to \mathbb{R}^N$, where $U \subseteq \mathbb{R}^M$ is open, with

(1) $D\phi$ of rank M at all points of U,
(2) $\phi(U) = V \bigcap S$.

Remark 5.3.2. The object defined in Definition 5.3.1 is sometimes called a *regularly embedded C^1* submanifold.

Definition 5.3.3. Suppose that S is an M-dimensional C^1 submanifold of $\mathbb{R}^N$. Let x be a point of S and let ϕ be as in Definition 5.3.1.

(1) The range of $D\phi(u)$, $u \in U$, will be called the *tangent space to S at $x = \phi(u)$* and will be denoted by $\mathbf{T}_x S$.
(2) Now suppose $x \in S$ and $f : W \to \mathbb{R}^\nu$, where W contains a neighborhood of x in S. We say that f is *differentiable relative to S at x* if there is $\widetilde{f} : \widetilde{W} \to \mathbb{R}^\nu$ such that
   (a) $\widetilde{W}$ is a neighborhood of x in $\mathbb{R}^N$,
   (b) $f\big|_{S\cap\widetilde{W}} = \widetilde{f}\big|_{S\cap\widetilde{W}}$,
   (c) $\widetilde{f}$ is differentiable at x.

   In case f is differentiable relative to S at x, we will call the restriction of $D\widetilde{f}(x)$ to $\mathbf{T}_x S$ the *differential of f relative to S at x* and we will denote $D\widetilde{f}(x)\big|_{\mathbf{T}_x S}$ by $D_S f(x)$.
(3) For $K \leq M$, we define the *K-dimensional Jacobian of f relative to S at x*, denoted by $J_K^S f(x)$, by setting

$$J_K^S f(x) = \sup\left\{\frac{\mathcal{H}^K[D_S f(P)]}{\mathcal{H}^K[P]} : \right.$$
$$\left. P \text{ is a } K\text{-dimensional parallelepiped contained in } \mathbf{T}_x S\right\}. \tag{5.25}$$

Remark 5.3.4. In case $\nu = 1$, we define the *gradient of f relative to S* to be that vector $\nabla^S f(x) \in \mathbf{T}_x S$ for which

$$\langle D_S f, v \rangle = \nabla^S f(x) \cdot v$$

holds for all $v \in \mathbf{T}_x S$. If fact, $\nabla^S f(x)$ is simply the orthogonal projection of grad $\widetilde{f}(x)$ on $\mathbf{T}_x S$, where $\widetilde{f}$ is as in (2) of the preceding definition.

Lemma 5.3.5. *Suppose S is an M-dimensional C^1 submanifold of $\mathbb{R}^N$. Suppose the $\mathbb{R}^\nu$-valued function f is differentiable relative to S at x.*

(1) *If $M \leq \nu$, then*
$$J_M^S f \cdot \mathcal{H}^M[P] = \mathcal{H}^M[D_S f(P)]$$
holds for any M-dimensional parallelepiped P contained in $\mathbf{T}_x S$.
(2) *If $\nu \leq M$, then*
$$J_\nu^S f \cdot \mathcal{H}^\nu[P] = \mathcal{H}^\nu[D_S f(P)]$$
holds for any ν-dimensional parallelepiped P contained in the orthogonal complement of ker $D_S f$ *in* $\mathbf{T}_x S$.

Proof. (1) Choose the orthonormal coordinate system in $\mathbb{R}^N$ so that $\mathbf{T}_x S$ is the span of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_M$. With this choice of coordinate system, $D_S f$ can be represented by an $\nu \times M$ matrix T.

Consider two M-dimensional parallelepipeds P_1 and P_2 contained in $\mathbf{T}_x S$. For $i = 1, 2$, let V_i be the $M \times M$ matrix whose columns are the vectors that determine P_i. There is a nonsingular $M \times M$ matrix A such that V_2 equals the matrix product

$A\,V_1$ (recall that we assume that our M-dimensional parallelepipeds are determined by M linearly independent vectors).

Using Proposition 5.1.2, we compute

$$\mathcal{H}^M[P_1] = \sqrt{\det(V_1^{\mathtt{t}}\,V_1)} = |\det(V_1)|\,,$$

$$\mathcal{H}^M[P_2] = \sqrt{V_2^{\mathtt{t}}\,V_2} = \sqrt{V_1^{\mathtt{t}}\,A^{\mathtt{t}}\,V_1\,A} = |\det(A)|\,|\det(V_1)|\,,$$

$$\mathcal{H}^M[D_S f(P_1)] = \sqrt{\det(V_1^{\mathtt{t}}\,T^{\mathtt{t}}\,T\,V_1)} = \sqrt{\det(T^{\mathtt{t}}\,T)}\,|\det(V_1)|\,,$$

$$\begin{aligned}\mathcal{H}^M[D_S f(P_2)] &= \sqrt{\det(V_2^{\mathtt{t}}\,T^{\mathtt{t}}\,T\,V_2)}\,,\\ &= \sqrt{\det(V_1^{\mathtt{t}}\,A^{\mathtt{t}}\,T^{\mathtt{t}}\,T\,A\,V_1)} = \sqrt{\det(T^{\mathtt{t}}\,T)}\,|\det(A)|\,|\det(V_1)|\,,\end{aligned}$$

and the result follows.

(2) If P is a ν-dimensional parallelepiped and $\widetilde{P}$ is its orthogonal projection on the orthogonal complement of the kernel of $D_S f$, then we have $D_S f(P) = D_S f(\widetilde{P})$ and $\mathcal{H}^\nu(P) \geq \mathcal{H}^\nu(\widetilde{P})$. Thus the supremum in (5.25) will be realized by a parallelepiped contained in the orthogonal projection on the orthogonal complement of the kernel of $D_S f$.

Choosing the orthonormal coordinate system in $\mathbb{R}^N$ so that the orthogonal complement of the kernel of $D_S f$ is the span of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_\nu$, and arguing as in the proof of (1), we see that the supremum is realized by any such parallelepiped. □

Lemma 5.3.6. *Suppose that $M \leq \nu$, S is an M-dimensional C^1 submanifold of $\mathbb{R}^N$, and ϕ is as in Definition* 5.3.1 *above. If the $\mathbb{R}^\nu$-valued function f is C^1 in a neighborhood of x in S and if $x = \phi(u)$, then*

$$J_M^S\,f[\phi(u)]\,J_M\phi(u) = J_M(f \circ \phi)(u)\,.$$

Proof. Let P be an M-dimensional parallelepiped contained in $\mathbb{R}^M$. By Definition 5.1.3 and Lemma 5.1.4, we have $\mathcal{H}^M[D\phi(P)] = J_M\phi(u)\,\mathcal{H}^M[P]$ and $\mathcal{H}^M[D(f \circ \phi)(P)] = J_M(f \circ \phi)(u)\,\mathcal{H}^M[P]$. By Lemma 5.3.5, we have $\mathcal{H}^M[D_S(\phi(P))] = J_M^S\,f\,\mathcal{H}^M[D\phi(P)]$. Since $D_S(\phi(P)) = D(f \circ \phi)(P)$, we conclude that

$$\begin{aligned}J_M^S\,f\,J_M\phi(u)\,\mathcal{H}^M[P] &= J_M^S\,f\,\mathcal{H}^M[D\phi(P)]\\ &= \mathcal{H}^M[D_S(\phi(P))]\\ &= \mathcal{H}^M[D(f \circ \phi)(P)]\\ &= J_M(f \circ \phi)(u)\,\mathcal{H}^M[P]\,,\end{aligned}$$

from which the result follows. □

We now can prove the following version of the area formula for C^1 submanifolds.

Theorem 5.3.7. *Suppose $M \le \nu$ and $f : \mathbb{R}^N \to \mathbb{R}^\nu$ is Lipschitz. If $S \subseteq \mathbb{R}^N$ is an M-dimensional C^1 submanifold, then*

$$\int_S g\, J_M^S f \, d\mathcal{H}^M = \int_{\mathbb{R}^\nu} g(y)\, \mathrm{card}(S \cap f^{-1}(y))\, d\mathcal{H}^M(y)$$

for every $\mathcal{H}^M$-measurable function g.

Proof. It suffices to consider $g \equiv 1$ and $S = \phi(U)$, where $\phi : U \to \mathbb{R}^N$. By part (1) of Lemma 5.3.5 and Corollary 5.1.13, we have

$$\begin{aligned}
\int_S J_M^S f \, d\mathcal{H}^M &= \int_U J_M^S f[\phi(u)]\, J_M\phi(u)\, d\mathcal{L}^M(u) \\
&= \int_U J_M(f \circ \phi)(u)\, d\mathcal{L}^M(u) \\
&= \int_{\mathbb{R}^\nu} \mathrm{card}(U \cap (f \circ \phi)^{-1}(y))\, d\mathcal{H}^M(y) \\
&= \int_{\mathbb{R}^\nu} \mathrm{card}(S \cap f^{-1}(y))\, d\mathcal{H}^M(y)\,.
\end{aligned}$$

□

Lemma 5.3.8. *Suppose that $\nu < M$, S is an M-dimensional C^1 submanifold of $\mathbb{R}^N$, and ϕ is as in Definition* 5.3.1 *above. If the $\mathbb{R}^\nu$-valued function f is C^1 in a neighborhood of x in S and if $z = f(x)$, then*

$$J_\nu(f \circ \phi) \cdot J_{M-\nu}^{(f\circ\phi)^{-1}(z)}\, \phi = J_M\,\phi \cdot J_\nu^S\, f\,. \tag{5.26}$$

Proof. The two linear functions $D(f \circ \phi)$ and $D_S f$ clearly have the same rank. If that common rank is less than ν, then both sides of (5.26) are zero. Thus we may assume that both functions have rank ν.

Let $\Pi : \mathbf{T}_x S \to \mathbf{T}_x S$ be orthogonal projection onto the orthogonal complement of $\ker D_S f$. Choose an $(M-\nu)$-dimensional parallelepiped P_1 in $\ker D(f \circ \phi)$ and a ν-dimensional parallelepiped P_2 in the orthogonal complement of $\ker D(f \circ \phi)$. Since $D\phi$ maps $\ker D(f \circ \phi)$ onto $\ker D_S f$, we have

$$\mathcal{H}^M[(D\phi(P_1)) \times (\Pi \circ D\phi(P_2))] = \mathcal{H}^M[(D\phi(P_1)) \times (D\phi(P_2))]\,. \tag{5.27}$$

Since $\Pi \circ D\phi(P_2)$ is a ν-dimensional parallelepiped in the orthogonal complement of $\ker D_S f$ and P_2 is a ν-dimensional parallelepiped in the orthogonal complement of $\ker D(f \circ \phi)$, Lemma 5.3.5 gives us

$$\begin{aligned}
J_\nu^S f \cdot \mathcal{H}^\nu[\Pi \circ D\phi(P_2)] &= \mathcal{H}^\nu[D_S f(\Pi \circ D\phi(P_2))] \\
&= \mathcal{H}^\nu[D_S f \circ D\phi(P_2)] \\
&= \mathcal{H}^\nu[D(f \circ \phi)(P_2)] \\
&= J_\nu(f \circ \phi) \cdot \mathcal{H}^\nu[P_2]\,.
\end{aligned} \tag{5.28}$$

We also have

$$J_{M-\nu}^{(f\circ\phi)^{-1}(z)}\,\phi\cdot\mathcal{H}^{M-\nu}[P_1]=\mathcal{H}^{M-\nu}[D\phi(P_1)]\,. \tag{5.29}$$

Combining (5.28) and (5.29), using (5.27), and applying Lemma 5.3.5 again, we obtain

$$\begin{aligned}
&J_\nu\,(f\circ\phi)\cdot J_{M-\nu}^{(f\circ\phi)^{-1}(z)}\,\phi\cdot\mathcal{H}^{M-\nu}[P_1]\cdot\mathcal{H}^\nu[P_2]\\
&=J_\nu^S\,f\cdot\mathcal{H}^{M-\nu}[D\phi(P_1)]\cdot\mathcal{H}^\nu[\Pi\circ D\phi(P_2)]\\
&=J_\nu^S\,f\cdot\mathcal{H}^{M}[(D\phi(P_1))\times(\Pi\circ D\phi(P_2))]\\
&=J_\nu^S\,f\cdot\mathcal{H}^{M}[(D\phi(P_1))\times(D\phi(P_2))]\\
&=J_\nu^S\,f\cdot\mathcal{H}^{M}[D\phi(P_1\times P_2)]\\
&=J_\nu^S\,f\cdot J_M\,\phi\cdot\mathcal{H}^{M}[P_1\times P_2]\\
&=J_\nu^S\,f\cdot J_M\,\phi\cdot\mathcal{H}^{M-\nu}[P_1]\cdot\mathcal{H}^\nu[P_2]\,,
\end{aligned}$$

and the result follows. □

To end this section, we prove the coarea formula for C^1 submanifolds. As we shall see in the next section, the condition that f be C^1 is not essential; it suffices to assume only that f is Lipschitz.

Theorem 5.3.9. *Suppose $M\ge\nu$ and $f:\mathbb{R}^N\to\mathbb{R}^\nu$ is C^1. If $S\subseteq\mathbb{R}^N$ is an M-dimensional C^1 submanifold, then*

$$\int_S g\,J_\nu^S\,f\,d\mathcal{H}^M=\int_{\mathbb{R}^\nu}\int_{S\cap f^{-1}(y)}g\,d\mathcal{H}^{M-\nu}\,d\mathcal{H}^\nu(y)$$

for every $\mathcal{H}^M$-measurable function g.

Proof. It suffices to consider $g\equiv 1$ and $S=\phi(U)$, where $\phi:U\to\mathbb{R}^N$. By Lemma 5.3.5 and Theorem 5.3.7, we have

$$\begin{aligned}
\int_S J_\nu^S\,f\,d\mathcal{H}^M&=\int_U J_\nu^S\,f(x)J_M\,\phi(u)\,d\mathcal{L}^M\\
&=\int_{\mathbb{R}^\nu}J_\nu\,(f\circ\phi)\,J_{M-\nu}^{(f\circ\phi)^{-1}(z)}\,\phi\,d\mathcal{H}^{M-\nu}\,d\mathcal{H}^\nu(y)\\
&=\int_{\mathbb{R}^\nu}\int_{U\cap(f\circ\phi)^{-1}(y)}J_{M-\nu}^{(f\circ\phi)^{-1}(z)}\,\phi\,d\mathcal{H}^{M-\nu}\,d\mathcal{H}^\nu(y)\\
&=\int_{\mathbb{R}^\nu}\int_{S\cap f^{-1}(y)}d\mathcal{H}^{M-\nu}\,d\mathcal{H}^\nu(y)\,.
\end{aligned}$$

□

5.4 Rectifiable Sets

Definition 5.4.1. Let M be an integer with $1 \le M \le N$. A set $S \subseteq \mathbb{R}^N$ is said to be *countably M-rectifiable* if

$$S \subseteq S_0 \bigcup \left(\bigcup_{j=1}^{\infty} F_j(\mathbb{R}^M) \right) ,$$

where

(1) $\mathcal{H}^M(S_0) = 0$;
(2) $F_j : \mathbb{R}^M \to \mathbb{R}^N$ are Lipschitz functions, $j = 1, 2, \ldots$.

We will usually use countably M-rectifiable sets in conjunction with the hypothesis of $\mathcal{H}^M$-measurability and the assumption that the intersection with any compact set has finite Hausdorff measure.

Our terminology follows that of [Sim 83] rather than that of [Fed 69]. The distinction here is that we are allowing the set S_0 with $\mathcal{H}^M(S_0) = 0$, but that set is excluded in [Fed 69].

It is easy to see that a Lipschitz function $f : A \to \mathbb{R}^N$ can be extended to a Lipschitz function $F : \mathbb{R}^M \to \mathbb{R}$ with $\operatorname{Lip}(F)$ bounded by a constant multiple[4] of $\operatorname{Lip}(f)$. Thus condition (2) in Definition 5.4.1 is equivalent to mandating that

$$S = S_0 \bigcup \left(\bigcup_{j=1}^{\infty} F_j(S_j) \right) ,$$

where $\mathcal{H}^M(S_0) = 0$, $S_j \subseteq \mathbb{R}^M$, and $F_j : S_j \to \mathbb{R}^N$ is Lipschitz. In practice this is the way that we think of an M-rectifiable set.

Lemma 5.4.2. *The set S is countably M-rectifiable ($1 \le M$) if and only if $S \subseteq \bigcup_{j=0}^{\infty} T_j$, where $\mathcal{H}^M(T_0) = 0$ and where each T_j for $j \ge 1$ is an M-dimensional, embedded C^1 submanifold of $\mathbb{R}^N$.*

Proof. The "if" direction of the result is trivial. For the "only if" part, we use Theorem 5.1.12. Specifically, we select C^1 functions $h_1^{(j)}, h_2^{(j)}, \ldots$ such that if F_j are Lipschitz functions as in Definition 5.4.1, then

$$F_j(\mathbb{R}^M) \subseteq E_j \bigcup \left(\bigcup_{\ell=1}^{\infty} h_\ell^{(j)}(\mathbb{R}^M) \right) , \quad j = 1, 2, \ldots ,$$

where $\mathcal{H}^M(E_j) = 0$. Then set

[4] The deeper result that an $\mathbb{R}^N$-valued function on a subset of $\mathbb{R}^M$ can be extended without increasing the Lipschitz constant is Kirszbraun's theorem (see [Fed 69] or [KPk 99]).

$$C_{\ell j} = \left\{ x \in \mathbb{R}^M : J_M \, h_\ell^{(j)}(x) = 0 \right\} ,$$

where $J_M \, h_\ell^{(j)}(x)$ denotes the M-dimensional Jacobian of $h_\ell^{(j)}$ at x (see Definition 5.1.3), and define

$$T_0 = \left(\bigcup_{j=1}^{\infty} E_j \right) \bigcup \left(\bigcup_{\ell, j=1}^{\infty} h_\ell^{(j)}(C_{\ell j}) \right) .$$

Theorem 5.1.1, the area formula, now tells us that $\mathcal{H}^M \left(\bigcup_{\ell,j=1}^{\infty} h_\ell^{(j)}(C_{\ell j}) \right) = 0$ and hence $\mathcal{H}^M(T_0) = 0$.

Because the open set $\mathbb{R}^M \setminus C_{\ell j}$ consists only of points at which $J_M \, h_\ell^{(j)}$ is non-vanishing, $\mathbb{R}^M \setminus C_{\ell j}$ can be written as the union of countably many open sets $U_{\ell jk}$ that may be chosen small enough that each $T_{\ell jk} = h_\ell^{(j)}(U_{\ell jk})$ is an M-dimensional, embedded C^1 submanifold of $\mathbb{R}^N$. Then we have

$$S \subseteq T_0 \cup \bigcup_{\ell, j, k=1}^{\infty} T_{\ell jk} ,$$

as required. □

Proposition 5.4.3. *Suppose $M \geq 1$. If the set S is $\mathcal{H}^M$-measurable and countably M-rectifiable, then $S = \bigcup_{j=0}^{\infty} S_j$, where*

(1) $\mathcal{H}^M(S_0) = 0$,
(2) $S_i \bigcap S_j = \emptyset$ *if* $i \neq j$,
(3) *for* $j \geq 1$, $S_j \subseteq T_j$, *and* T_j *is an M-dimensional, embedded C^1 submanifold of* $\mathbb{R}^N$.

Proof. Let the T_j be as in Lemma 5.4.2. Define the S_j inductively by setting $S_0 = S \bigcap T_0$ and $S_{j+1} = (S \bigcap T_{j+1}) \setminus \bigcup_{i=0}^{j} S_i$. □

Definition 5.4.4. Let $S \subseteq \mathbb{R}^N$ be $\mathcal{H}^M$-measurable with $\mathcal{H}^M(S \cap K) < \infty$ for every compact K. We say that an M-dimensional linear subspace W of $\mathbb{R}^N$ is the *approximate tangent space* to S at $x \in \mathbb{R}^N$ if

$$\lim_{\lambda \to 0^+} \int_{\lambda^{-1}(S-x)} f(y) \, d\mathcal{H}^M(y) = \int_W f(y) \, d\mathcal{H}^M(y)$$

for all compactly supported continuous functions f. Here

$$y \in \lambda^{-1}(S - x) \iff \lambda y + x \in S \iff y = \lambda^{-1}(z - x) \text{ for some } z \in S .$$

Of course, if S is an M-dimensional C^1 submanifold of $\mathbb{R}^N$, then the approximate tangent space coincides with the usual tangent space arising from the smooth structure. When S is not a C^1 submanifold, there may exist various exceptional points x of S for which there is a set W that is not an M-dimensional linear subspace, but nonetheless ought to be considered a tangent object for S at x—for example, at a vertex of a simplex. Even so, our definition will be justified by the fact that in the case that S is countably M-rectifiable, the set of such exceptional points x has $\mathcal{H}^M$ measure zero.

Notation 5.4.5. When the approximate tangent space to S at x exists, we will denote it by $\mathbf{T}_x S$. Here the dimension M should always be understood to be the Hausdorff dimension of S. This notation extends that introduced in Definition 5.3.3, (1).

Theorem 5.4.6. *If S is $\mathcal{H}^M$-measurable and countably M-rectifiable and if $\mathcal{H}^M(S \cap K) < \infty$ holds for every compact $K \subseteq \mathbb{R}^N$, then $\mathbf{T}_x S$ exists for $\mathcal{H}^M$-almost every $x \in S$.*

Proof. Write S as in Proposition 5.4.3 and consider $j \geq 1$. By Corollary 4.3.10, we have (using the notation of Proposition 5.4.3)

$$\Theta^{*M}[\mathcal{H}^M \llcorner (S \setminus S_j), x] = 0$$

for $\mathcal{H}^M$-almost every $x \in S_j$. By Theorem 4.3.5, we have

$$\lim_{r \downarrow 0} \frac{\mathcal{H}^M[S_j \bigcap \overline{\mathbb{B}}(x,r)]}{\mathcal{H}^M[T_j \bigcap \overline{\mathbb{B}}(x,r)]} = 1$$

for $\mathcal{H}^M$-almost every $x \in S_j$. Since T_j is an M-dimensional C^1 submanifold of $\mathbb{R}^N$, the result follows with $\mathbf{T}_x S = \mathbf{T}_x T_j$. □

Definition 5.4.7. Suppose that the set S is $\mathcal{H}^M$-measurable and countably M-rectifiable and suppose that $\mathcal{H}^M(S \cap K) < \infty$ holds for every compact $K \subseteq \mathbb{R}^N$. Let $f : S \to \mathbb{R}^\nu$. We define $D_S f$ and J_K^S, $K \leq M$, by writing S as in Proposition 5.4.3 and setting

$$D_S f(x) = D_{T_j} f(x)\,,$$

$$J_K^S f(x) = J_K^{T_j} f(x)$$

whenever $j \geq 1$ and the respective right-hand side exists. We call $D_S f$ the *approximate differential* of f and $J_K^S f$ the *approximate K-dimensional Jacobian* of f. In case $\nu = 1$, we similarly define the *approximate gradient* of f, denoted by $\nabla^S f$.

Now that the requisite definitions have been made, the area and coarea formulas for countably M-rectifiable sets follow readily from the corresponding results for C^1 submanifolds.

Theorem 5.4.8. *Suppose that $M \leq \nu$ and $f : \mathbb{R}^N \to \mathbb{R}^\nu$ is Lipschitz. If $S \subseteq \mathbb{R}^N$ is $\mathcal{H}^M$-measurable and countably M-rectifiable and if $\mathcal{H}^M(S \cap K) < \infty$ holds for every compact $K \subseteq \mathbb{R}^N$, then $J_M^S f$ exists $\mathcal{H}^M$-almost everywhere in S and*

$$\int_S g\, J_M^S f \, d\mathcal{H}^M = \int_{\mathbb{R}^\nu} g(y)\, \mathrm{card}(S \cap f^{-1}(y))\, d\mathcal{H}^M(y)$$

holds for every $\mathcal{H}^M$-measurable function g.

Proof. Write S as in Proposition 5.4.3 and apply Theorem 5.3.7. □

Theorem 5.4.9. *Suppose $M \geq \nu$ and $f : \mathbb{R}^N \to \mathbb{R}^\nu$ is Lipschitz. If $S \subseteq \mathbb{R}^N$ is $\mathcal{H}^M$-measurable and countably M-rectifiable and if $\mathcal{H}^M(S \cap K) < \infty$ holds for every compact $K \subseteq \mathbb{R}^N$, then $J_\nu^S f$ exists $\mathcal{H}^M$-almost everywhere in S and*

$$\int_S g \, J_\nu^S f \, d\mathcal{H}^M = \int_{\mathbb{R}^\nu} \int_{S \cap f^{-1}(y)} g \, d\mathcal{H}^{M-\nu} \, d\mathcal{H}^\nu(y)$$

holds for every $\mathcal{H}^M$-measurable function g.

Proof. Write S as in Proposition 5.4.3 and, using Theorem 5.1.12 to approximate the Lipschitz map f by C^1 maps, apply Theorem 5.3.7. □

5.5 Poincaré Inequalities

The Poincaré inequalities[5] are like a weak version of the Sobolev inequalities[6] (see [Zie 89, Section 2.4] for an introduction to Sobolev inequalities). They are of a priori interest, but they also are adequate for many of our applications in geometric measure theory.

We shall require a bit of preliminary machinery in order to formulate and prove the results that follow. In most partial differential equations texts, the Poincaré inequalities are formulated for smooth testing functions. Here we must have such inequalities for functions of bounded variation. So some extra effort is required.

A function u on a domain $U \subseteq \mathbb{R}^N$ is said to be of *local bounded variation* on U, written $u \in BV_{\text{loc}}(U)$, if for each $W \subset\subset U$ there is a constant $c = c(W) < \infty$ such that

$$\int_W u(x) \operatorname{div} g(x) \, d\mathcal{L}^N(x) \leq c(W) \cdot \sup |g| \tag{5.30}$$

holds for all compactly supported, vector-valued functions $g = (g^1, \ldots, g^N)$ with each $g^j \in C^\infty(W)$. For convenience we denote the space of such g by $\mathcal{K}_W(U, \mathbb{R}^N)$. Then we see from (5.30) that the linear functional

$$\mathcal{K}_W(U, \mathbb{R}^N) \ni g \longmapsto \int_W u(x) \operatorname{div} g(x) \, d\mathcal{L}^N(x)$$

is bounded in the supremum norm. Thus the Riesz representation theorem, i.e., Theorem 4.4.1, tells us that there is a Radon measure μ on U and a μ-measurable function $\nu = (\nu^1, \ldots, \nu^N)$, with $|\nu| = 1$ almost everywhere, such that[7]

[5] Jules Henri Poincaré (1854–1912).

[6] Sergei Lvovich Sobolev (1908–1989).

[7] Of course, the usual formulation of the Riesz theorem does not include the vector-valued function ν. That function is necessitated by the fact that g is vector-valued. The extension of Riesz's theorem to the vector-valued case is routine.

$$\int_U u(x)\,\mathrm{div}\,g(x)\,d\mathcal{L}^N(x) = \int_U g(x)\cdot\nu(x)\,d\mu(x)\,.$$

In the language of distribution theory, the weak derivatives $D_j u$ of u are represented by the signed measures $\nu_j d\mu$, $j = 1, \ldots, N$. It is thus convenient to denote the total variation measure[8] μ by $|Du|$.

We will find it useful in our discussions to use Friedrichs mollifiers[9] to smooth our bounded variation functions.

Definition 5.5.1. We call φ a *mollifier* if (see Figure 5.3)

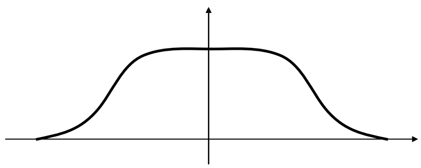

Fig. 5.3. The graph of a mollifier.

- $\varphi \in C^\infty(\mathbb{R}^N)$;
- $\varphi \geq 0$;
- $\mathrm{supp}\,\varphi \subseteq \mathbb{B}(0, 1)$;
- $\displaystyle\int_{\mathbb{R}^N} \varphi(x)\,d\mathcal{L}^N(x) = 1$;
- $\varphi(x) = \varphi(-x)$.

For $\sigma > 0$ we set $\varphi_\sigma(x) = \sigma^{-N}\varphi(x/\sigma)$. We call $\{\varphi_\sigma\}_{\sigma>0}$ a *family of mollifiers* or an *approximation to the identity*.

In case $f \in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ and $\sigma > 0$, we define

$$f_\sigma(x) = f * \varphi_\sigma(x) = \int_{\mathbb{R}^N} f(z)\,\varphi_\sigma(x-z)\,d\mathcal{L}^N(z) = \int_{\mathbb{R}^N} f(x-z)\,\varphi_\sigma(z)\,d\mathcal{L}^N(z)\,. \tag{5.31}$$

Then $f_\sigma \in C^\infty$ and f_σ converges back to f in a variety of senses. In particular, $f_\sigma \to f$ pointwise almost everywhere and $f_\sigma \to f$ in the L^1_{loc} topology. In case f is continuous then f_σ converges uniformly on compact sets to f. The reference [SW 71] contains details of these assertions.

We begin with a version of the Poincaré inequality for smooth functions. If f is a Lebesgue measurable function and U is a subset of positive Lebesgue measure of the domain of f then we let

[8] Indeed, if $u \in W^{1,1}_{\mathrm{loc}}(U)$ then $d\mu = |Du|d\mathcal{L}^N$ and $\nu_j = \frac{D_j u}{|Du|}$ provided $|Du| \neq 0$.

[9] Kurt Otto Friedrichs (1901–1982).

$$f_U = \frac{1}{\mathcal{L}^N(U)} \int_U f(t)\, d\mathcal{L}^N(t) \tag{5.32}$$

be the average of f over U.

Lemma 5.5.2. *Let U be a bounded, convex, open subset of $\mathbb{R}^N$. Let f be a continuously differentiable function on U. Then there is a constant $c = c(U)$ such that*

$$\int_U |f - f_U|\, d\mathcal{L}^N \le c \cdot \int_U |Df|\, d\mathcal{L}^N .$$

Proof. We will use the notation $|U| = \mathcal{L}^N(U)$. We calculate that

$$\int_U |f - f_U|\, d\mathcal{L}^N = \int_U \left| f(x) - \frac{1}{|U|} \int_U f(t)\, d\mathcal{L}^N(t) \right| d\mathcal{L}^N(x)$$

$$= \int_U \left| \frac{1}{|U|} \int_U [\, f(x) - f(t)\,]\, d\mathcal{L}^N(t) \right| d\mathcal{L}^N(x)$$

$$\le \frac{1}{|U|} \int_U \int_U |f(x) - f(t)|\, d\mathcal{L}^N(x)\, d\mathcal{L}^N((t)$$

$$= \frac{1}{|U|} \int_U \int_U \left| \int_0^1 \frac{d}{ds} f((1-s)t + sx)\, d\mathcal{L}^1(s) \right| d\mathcal{L}^N(x)\, d\mathcal{L}^N(t)$$

$$\le \frac{1}{|U|} \int_U \int_U \int_0^1 |Df((1-s)t + sx)| \cdot |x - t|\, d\mathcal{L}^1(s)\, d\mathcal{L}^N(x)\, d\mathcal{L}^N(t)$$

$$\le \operatorname{diam}(U) \cdot \frac{1}{|U|} \int_U \int_U \int_0^1 |Df((1-s)t + sx)|\, d\mathcal{L}^1(s)\, d\mathcal{L}^N(x)\, d\mathcal{L}^N(t)$$

$$= \operatorname{diam}(U) \cdot \frac{1}{|U|} \int_U \int_0^{1/2} \left(\int_U |Df((1-s)t + sx)|\, d\mathcal{L}^N(t) \right) d\mathcal{L}^1(s)\, d\mathcal{L}^N(x)$$

$$+ \operatorname{diam}(U) \cdot \frac{1}{|U|} \int_U \int_{1/2}^1 \left(\int_U |Df((1-s)t + sx)|\, d\mathcal{L}^N(x) \right) d\mathcal{L}^1(s)\, d\mathcal{L}^N(t) .$$

For $1/2 \le s \le 1$, by making the change of variable $\widehat{x} = (1-s)t + sx$, we see that

$$\int_U |Df((1-s)t + sx)|\, d\mathcal{L}^N(x) = \int_{\widehat{U}} |Df(\widehat{x})|\, s^{-N}\, d\mathcal{L}^N(\widehat{x}) ,$$

where

$$\widehat{U} = \{(1-s)t + sx : x \in U\} .$$

Observing that $\widehat{U} \subseteq U$, we obtain

$$\int_{\widehat{U}} |Df(\widehat{x})|\, s^{-N}\, d\mathcal{L}^N(\widehat{x}) \le s^{-N}\, \|Df\|_{L^1(U)} \le 2^N\, \|Df\|_{L^1(U)} .$$

Similarly, for $0 \le s \le 1/2$ we have

$$\int_U |Df((1-s)t + sx)| \, d\mathcal{L}^N(t) \le 2^N \, \|Df\|_{L^1(U)} \, .$$

We conclude that

$$\int_U |f - f_U| \, d\mathcal{L}^N \le \operatorname{diam}(U) \cdot \frac{1}{|U|} \cdot 2^N \, \|Df\|_{L^1(U)} \int_U d\mathcal{L}^N$$

$$= 2^N \operatorname{diam}(U) \, \|Df\|_{L^1(U)} \, . \qquad \square$$

Remark 5.5.3. Observe that we used the convexity property of U in order to invoke the fundamental theorem of calculus in line 4 of the calculation. In fact, with extra effort, a result may be proved on a smoothly bounded domain. One then instead uses a piecewise linear curve with the fundamental theorem.

Next we wish to replace the average f_U in the statement of the lemma with a more arbitrary constant.

Lemma 5.5.4. *Let $\beta \in \mathbb{R}$ and $0 < \theta < 1$ be constants. Let f and U be as in Lemma* 5.5.2*, and let f_U be as in* (5.32)*. Assume that*

$$\mathcal{L}^N \{x \in U : f(x) \ge \beta\} \ge \theta \, \mathcal{L}^N(U)$$

and

$$\mathcal{L}^N \{x \in U : f(x) \le \beta\}| \ge \theta \, \mathcal{L}^N(U) \, .$$

Then there is a constant $C = C(\theta)$ such that

$$\int_U |f(x) - \beta| \, d\mathcal{L}^N(x) \le \theta^{-1}(1+\theta) \cdot \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) \, .$$

Proof. We write

$$U_+ = \{x \in U : f(x) \ge \beta\} \, , \qquad U_- = \{x \in U : f(x) \le \beta\} \, .$$

First we shall prove that

$$\int_U |f_U - \beta| \, d\mathcal{L}^N \le C \cdot \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) \, .$$

We consider two cases:

(1) First we treat the case $\beta > f_U$. Then we have

$$\begin{aligned}
\int_U |f_U - \beta| \, d\mathcal{L}^N &= \int_U (\beta - f_U) \, d\mathcal{L}^N \\
&= \mathcal{L}^N(U) \cdot (\beta - f_U) \\
&\le \mathcal{L}^N(U) \cdot \left[\left(\frac{1}{\mathcal{L}^N(U_+)} \int_{U_+} f(x) \, d\mathcal{L}^N(x) \right) - f_U \right] \\
&= \mathcal{L}^N(U) \cdot \left(\frac{1}{\mathcal{L}^N(U_+)} \int_{U_+} (f(x) - f_U) \, d\mathcal{L}^N(x) \right) .
\end{aligned}$$

Now, on the set where $f \geq \beta$ we certainly have, since $\beta > f_U$, that $f > f_U$. Therefore the last line is (by our hypotheses about θ and β)

$$\leq C \cdot \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) .$$

Thus

$$\int_U |f_U - \beta| \, d\mathcal{L}^N \leq C \cdot \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) .$$

(2) Now we treat the case $\beta \leq f_U$. Then we have

$$\begin{aligned}\int_U |f_U - \beta| \, d\mathcal{L}^N &= \int_U (f_U - \beta) \, d\mathcal{L}^N \\ &\leq \mathcal{L}^N(U) \cdot \left(f_U - \frac{1}{\mathcal{L}^N(U_-)} \int_{U_-} f(x) \, d\mathcal{L}^N(x) \right) \\ &= \mathcal{L}^N(U) \cdot \left(\frac{1}{\mathcal{L}^N(U_-)} \int_{U_-} (f_U - f(x)) \, d\mathcal{L}^N(x) \right) .\end{aligned}$$

Now clearly $f \leq \beta \leq f_U$ on U_-. So we may estimate the last line, in view of our hypotheses about θ and β, by

$$C \cdot \int_U |f_U - f(x)| \, d\mathcal{L}^N(x) .$$

We have the simple estimates

$$\begin{aligned}\int_U |f(x) - \beta| \, d\mathcal{L}^N(x) &\leq \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) + \int_U |f_U - \beta| \, d\mathcal{L}^N(x) \\ &\leq \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) + C \cdot \int_U |f(x) - f_U| \, d\mathcal{L}^N(x) .\end{aligned}$$

That is the desired result. □

Theorem 5.5.5. *Let U be a bounded, convex, open subset of $\mathbb{R}^N$. Let β, θ be as in Lemma* 5.5.4. *Let f be a continuously differentiable function on U. Then*

$$\int_U |f - \beta| \, d\mathcal{L}^N \leq c \cdot \int_U |Df| \, d\mathcal{L}^N .$$

Proof. Combine the two lemmas. □

Theorem 5.5.6. *Let U be a bounded, convex, open subset of $\mathbb{R}^N$. Let β, θ be as in Lemma* 5.5.4. *Let u be a function of bounded variation on U. Then*

$$\int_U |u - \beta| \, d\mathcal{L}^N \leq c \cdot \int_U |Du| .$$

Proof. Use a standard approximation argument to reduce the result to the preceding theorem. □

Our next Poincaré inequality mediates between the variation of a function on $\mathbb{R}^N$ and its variation on the natural domain of support U. Of course, the boundary of U will play a key role in the result.

Theorem 5.5.7. *Let $U \subseteq \mathbb{R}^N$ be a bounded, open, and convex domain. If $u \in BV_{\mathrm{loc}}(\mathbb{R}^N)$ with* $\operatorname{supp} u \subseteq \overline{U}$*, then there is a constant $c = c(U)$ such that*

$$\int_{\mathbb{R}^N} |Du|\, d\mathcal{L}^N \leq c \cdot \left(\int_U |Du| + \int_U |u|\, d\mathcal{L}^N \right) .$$

Proof. For $\delta > 0$ small, set $U_\delta = \{x \in U : \operatorname{dist}(x, \partial U) > \delta\}$. Let ϕ_δ be a compactly supported C^∞ function satisfying

(1) $\phi_\delta = 1$ in U_δ;
(2) $\phi_\delta = 0$ in $\mathbb{R}^N \setminus U_{\delta/2}$;
(3) $0 \leq \phi_\delta \leq 1$ in $\mathbb{R}^N$;
(4) for some point $a \in U$ and some $c = c(U, a) > 0$,

$$|D\phi_\delta(x)| \leq -c \cdot (x - a) \cdot D\phi_\delta(x) \quad \text{for all } x \in U .$$

Condition (4) is perhaps unfamiliar, and merits some discussion. The point a should be thought of as lying in the "middle" of U, and its existence as mandated in (4) is simply a manifestation of the starlike quality of U (see Figure 5.4). The effect of the boundary of U will be expressed via the value of $c(U, a)$ in condition (4).

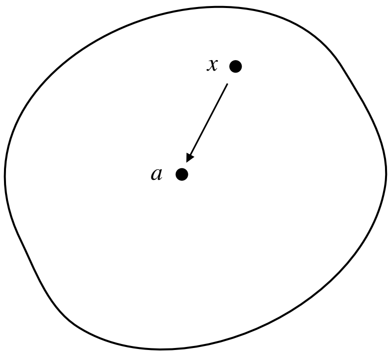

Fig. 5.4. The point a representing the middle of the set U.

We now apply the definition of $|Dw|$ with $w = \phi_\delta \cdot u$ to obtain

$$\int_{\mathbb{R}^N} |D(\phi_\delta \cdot u)|\, d\mathcal{L}^N \leq \int_{\mathbb{R}^N} |D\phi_\delta| \cdot |u|\, d\mathcal{L}^N + \int_{\mathbb{R}^N} \phi_\delta \cdot |Du| . \tag{5.33}$$

Property (4) of the function ϕ_δ tells us that

$$\int_{\mathbb{R}^N} |D\phi_\delta| \cdot |u| \, d\mathcal{L}^N \le -c \int_{\mathbb{R}^N} [(x-a) \cdot D\phi_\delta] \cdot |u| \, d\mathcal{L}^N(x) \, .$$

Notice that

$$-\int_{\mathbb{R}^N} \operatorname{div} [(x-a) \cdot \phi_\delta] \cdot |u| \, d\mathcal{L}^N = -\int_{\mathbb{R}^N} N \cdot \phi_\delta \cdot |u| + (x-a) \cdot D\phi_\delta \cdot |u| \, d\mathcal{L}^N \, .$$

Here we have used the fact that $\operatorname{div}(x-a) = N$, the dimension of the ambient space. Thus we see that

$$\int_{\mathbb{R}^N} -\operatorname{div} [(x-a) \cdot \phi_\delta] \cdot |u| + N\phi_\delta |u| \, d\mathcal{L}^N = \int_{\mathbb{R}^N} (x-a) \cdot D\phi_\delta \cdot |u| \, d\mathcal{L}^N \, .$$

In conclusion,

$$\int_{\mathbb{R}^N} |D\phi_\delta| \cdot |u| \, d\mathcal{L}^N \le c \cdot \int_{\mathbb{R}^N} (-|u| \cdot \operatorname{div}((x-a)\phi_\delta) + N|u|\phi_\delta) \, d\mathcal{L}^N(x) \, .$$

This last is majorized by

$$c \left(\int_U \big| D|u| \big| + \int_{\mathbb{R}^N} |u| \, d\mathcal{L}^N \right) \le c \left(\int_U |Du| \, d\mathcal{L}^N + \int_{\mathbb{R}^N} |u| \, d\mathcal{L}^N \right) . \tag{5.34}$$

Here we have used the definition of $\big| D|u| \big|$ and the fact that $\big| D|u| \big| \le |Du|$ as seen by a standard approximation argument.

Now it is not difficult to verify that

$$\int_{\mathbb{R}^N} |Du| \, d\mathcal{L}^N \le \liminf_{\delta \to 0^+} \int_{\mathbb{R}^N} |D(\phi_\delta u)| \, . \tag{5.35}$$

The result follows by combining (5.33), (5.34), and (5.35). □

6

The Calculus of Differential Forms and Stokes's Theorem

In this chapter, we give a brief treatment of the classical theory of differential forms and Stokes's theorem. These topics provide motivation for the more abstract theory of currents.

6.1 Differential Forms and Exterior Differentiation

Multilinear Functions and m-Covectors

The dual space of $\mathbb{R}^N$ is very useful in the formulation of line integrals (see Appendices A.2 and A.3), but to define surface integrals we need to go beyond the dual space to consider functions defined on ordered m-tuples of vectors.

Definition 6.1.1. Let $(\mathbb{R}^N)^m$ be the Cartesian product of m copies of $\mathbb{R}^N$.

(1) A function $\phi : (\mathbb{R}^N)^m \to \mathbb{R}$ is *m-linear* if it is linear as a function of each of its m arguments; that is, for each $1 \le \ell \le m$, it holds that

$$\begin{aligned}\phi(u_1, \dots, u_{\ell-1},\ \alpha u + \beta v,\ u_{\ell+1}, \dots, u_m)\\ = \alpha\, \phi(u_1, \dots, u_{\ell-1},\ u,\ u_{\ell+1}, \dots, u_m)\\ + \beta\, \phi(u_1, \dots, u_{\ell-1},\ v,\ u_{\ell+1}, \dots, u_m)\,,\end{aligned}$$

where $\alpha, \beta \in \mathbb{R}$ and $u, v, u_1, \dots, u_{\ell-1}, u_{\ell+1}, \dots, u_m \in \mathbb{R}^N$. The more inclusive term *multilinear* means m-linear for an appropriate m.

(2) A function $\phi : (\mathbb{R}^N)^m \to \mathbb{R}$ is *alternating* if interchanging two arguments results in a sign change for the value of the function; that is, for $1 \le i < \ell \le m$, it holds that

$$\begin{aligned}&\phi(u_1, \dots, u_{i-1},\ u_i,\ u_{i+1}, \dots, u_{\ell-1},\ u_\ell,\ u_{\ell+1}, \dots, u_m)\\ &\qquad = -\,\phi(u_1, \dots, u_{i-1},\ u_\ell,\ u_{i+1}, \dots, u_{\ell-1},\ u_i,\ u_{\ell+1}, \dots, u_m)\,,\end{aligned}$$

where $u_1, \dots, u_m \in \mathbb{R}^N$.

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_6,

(3) We denote by $\bigwedge^m(\mathbb{R}^N)$ the set of m-linear, alternating functions from $(\mathbb{R}^N)^m$ to $\mathbb{R}$. We endow $\bigwedge^m(\mathbb{R}^N)$ with the usual vector space operations of addition and scalar multiplication, namely,

$$(\phi + \psi)(u_1, u_2, \ldots, u_m) = \phi(u_1, u_2, \ldots, u_m) + \psi(u_1, u_2, \ldots, u_m)$$

and

$$(\alpha\,\phi)(u_1, u_2, \ldots, u_m) = \alpha \cdot \phi(u_1, u_2, \ldots, u_m)\,,$$

so $\bigwedge^m(\mathbb{R}^N)$ is itself a vector space. The elements of $\bigwedge^m(\mathbb{R}^N)$ are called *m-covectors of* $\mathbb{R}^N$.

Remark 6.1.2.

(1) In case $m = 1$, requiring a map to be alternating imposes no restriction; also, 1-linear is the same as linear. Consequently, we see that $\bigwedge^1(\mathbb{R}^N)$ is the dual space of $\mathbb{R}^N$; that is, $\bigwedge^1(\mathbb{R}^N) = (\mathbb{R}^N)^*$.

(2) Recalling that the standard basis for $\mathbb{R}^N$ is written $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$, we let $\mathbf{e}_i^*$ denote the dual of $\mathbf{e}_i$ defined by

$$\langle\, \mathbf{e}_i^*,\ \mathbf{e}_j\,\rangle = \begin{cases} 1 \text{ if } j = i, \\ 0 \text{ if } j \neq i. \end{cases}$$

Then $\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_N^*$ form the *standard dual basis* for $(\mathbb{R}^N)^*$.

(3) If $x_1, x_2, \ldots, x_N$ are the coordinates in $\mathbb{R}^N$, then it is traditional to use the alternative notation dx_i to denote the dual of $\mathbf{e}_i$; that is,

$$dx_i = \mathbf{e}_i^*\,, \quad \text{for } i = 1, 2, \ldots, N\,.$$

Example 6.1.3. The archetypical multilinear, alternating function is the determinant. As a function of its columns (or rows), the determinant of an N-by-N matrix is N-linear and alternating. It is elementary to verify that every element of $\bigwedge^N(\mathbb{R}^N)$ is a real multiple of the determinant function. □

The next definition shows how we can extend the use of determinants to define examples of m-linear, alternating functions when m is strictly smaller than N.

Definition 6.1.4. Let $a_1, a_2, \ldots, a_m \in \bigwedge^1(\mathbb{R}^N)$ be given. Each a_i can be written

$$a_i = a_{i\,1}\, dx_1 + a_{i\,2}\, dx_2 + \cdots + a_{i\,N}\, dx_N\,.$$

We define $a_1 \wedge a_2 \wedge \cdots \wedge a_m \in \bigwedge^m(\mathbb{R}^N)$, called the *exterior product of* $a_1, a_2, \ldots, a_m$, by setting

$$\begin{aligned}
&(a_1 \wedge a_2 \wedge \cdots \wedge a_m)(u_1, u_2, \ldots, u_m) \\
&\quad = \det\left[\begin{pmatrix} a_{1\,1} & a_{1\,2} & \ldots & a_{1\,N} \\ a_{2\,1} & a_{2\,2} & \ldots & a_{2\,N} \\ \vdots & \vdots & & \vdots \\ a_{m\,1} & a_{m\,2} & \ldots & a_{m\,N} \end{pmatrix} \begin{pmatrix} u_{1\,1} & u_{1\,2} & \ldots & u_{1\,m} \\ u_{2\,1} & u_{2\,2} & \ldots & u_{2\,m} \\ \vdots & \vdots & & \vdots \\ u_{N\,1} & u_{N\,2} & \ldots & u_{N\,m} \end{pmatrix}\right], \qquad (6.1)
\end{aligned}$$

where the u_{ij} are the components of the vectors $u_1, u_2, \ldots, u_m \in \mathbb{R}^N$; that is, each u_j is given by

$$u_j = u_{1j}\,\mathbf{e}_1 + u_{2j}\,\mathbf{e}_2 + \cdots + u_{Nj}\,\mathbf{e}_N \,.$$

To see that the function in (6.1) is m-linear and alternating, rewrite it in the form

$$\begin{aligned}
&(a_1 \wedge a_2 \wedge \cdots \wedge a_m)(u_1, u_2, \ldots, u_m) \\
&\quad = \det \begin{pmatrix} \langle a_1,\, u_1\rangle & \langle a_1,\, u_2\rangle & \ldots & \langle a_1,\, u_m\rangle \\ \langle a_2,\, u_1\rangle & \langle a_2,\, u_2\rangle & \ldots & \langle a_2,\, u_m\rangle \\ \vdots & \vdots & & \vdots \\ \langle a_m,\, u_1\rangle & \langle a_m,\, u_2\rangle & \ldots & \langle a_m,\, u_m\rangle \end{pmatrix}, \qquad (6.2)
\end{aligned}$$

where $\langle a_i,\, u_j\rangle$ is the dual pairing of a_i and u_j (see Section A.2).

Elements of $\bigwedge^m \mathbb{R}^N$ that can be written in the form $a_1 \wedge a_2 \wedge \cdots \wedge a_m$ are called *simple* m-covectors.

Recall that $\bigwedge_m (\mathbb{R}^N)$ is the space of m-vectors in $\mathbb{R}^N$ defined in Section 1.4. It is easy to see that any element of $\bigwedge^m (\mathbb{R}^N)$ is well-defined on $\bigwedge_m (\mathbb{R}^N)$ (just consider the equivalence relation in Definition 1.4.1). Thus $\bigwedge^m (\mathbb{R}^N)$ can be considered the dual space of $\bigwedge_m (\mathbb{R}^N)$. Evidently

$$dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}\,, \quad 1 \le i_1 < i_2 < \cdots < i_m \le N\,, \qquad (6.3)$$

is the dual basis to the basis

$$\mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m}\,, \quad 1 \le i_1 < i_2 < \cdots < i_m \le N\,,$$

for $\bigwedge_m (\mathbb{R}^N)$.

Differential Forms

Definition 6.1.5. Let $W \subset \mathbb{R}^N$ be open. A *differential m-form* on W is a function $\phi : W \to \bigwedge^m (\mathbb{R}^N)$. We call m the *degree* of the form. We say that the differential m-form ϕ is C^k if for each set of (constant) vectors $v_1, v_2, \ldots, v_m$, the real-valued function $\langle \phi(p), v_1 \wedge v_2 \wedge \cdots \wedge v_m\rangle$ is a C^k function of $p \in W$.

The differential form can be rewritten in terms of a basis and component functions as follows: For each m-tuple $1 \le i_1 < i_2 < \cdots < i_m \le N$, define the real-valued function

$$\phi_{i_1,i_2,\ldots,i_m}(p) = \langle \phi(p), \mathbf{e}_{i_1} \wedge \mathbf{e}_{i_2} \wedge \cdots \wedge \mathbf{e}_{i_m}\rangle\,.$$

Then we have

$$\phi = \sum_{1 \le i_1 < i_2 < \cdots < i_m \le N} \phi_{i_1,i_2,\ldots,i_m}\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}\,.$$

The natural role for a differential m-form is to serve as the integrand in an integral over an m-dimensional surface. This is consistent with and generalizes integration of a 1-form along a curve.

Definition 6.1.6. Suppose

(1) the m-dimensional surface $S \subseteq \mathbb{R}^N$ is parametrized by the function $F : U \to \mathbb{R}^N$, where U is an open subset of $\mathbb{R}^m$; that is, F is a one-to-one C^k ($k \geq 1$) function, DF is of rank m, and $S = F(U)$,
(2) $W \subseteq \mathbb{R}^N$ is open with $F(U) \subseteq W$, and
(3) ϕ is a differential m-form on W.

Then the *integral* of ϕ over S is defined by

$$\int_S \phi = \int_U \left\langle \phi \circ F(t), \frac{\partial F}{\partial t_1} \wedge \frac{\partial F}{\partial t_2} \wedge \cdots \wedge \frac{\partial F}{\partial t_m} \right\rangle d\mathcal{L}^m(t) \tag{6.4}$$

whenever the right-hand side of (6.4) is defined.

The surface S in Definition 6.1.6 is an oriented surface for which the orientation is induced by the orientation on $\mathbb{R}^m$ and the parametrization F. The value of the integral is unaffected by a reparametrization as long as the reparametrization is orientation-preserving.

Exterior Differentiation

In Appendix A.3 one can see how the exterior derivative of a function allows the fundamental theorem of calculus to be applied to the integrals of 1-forms along curves. The exterior derivative of a differential form, which we discuss next, is the mechanism that allows the fundamental theorem of calculus to be extended to higher-dimensional settings.

Definition 6.1.7. Suppose that $U \subset \mathbb{R}^N$ is open and $f : U \to \mathbb{R}$ is a C^k function, $k \geq 1$.

(1) The *exterior derivative* of f is the 1-form df on U defined by setting

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_N} dx_N \,. \tag{6.5}$$

Note that (6.5) is equivalent to

$$\langle df(p),\ v \rangle = \langle Df(p),\ v \rangle \,, \tag{6.6}$$

for $p \in U$ and $v \in \mathbb{R}^N$.

(2) The *exterior derivative* of the m-form $\phi = f\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}$, $m \geq 1$, is the $(m+1)$-form $d\phi$ given by setting

$$d\phi = (df) \wedge dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m} \,.$$

(3) The definition of exterior differentiation in (2) is extended by linearity to all C^k m-forms, $m \geq 1$.

The rules analogous to those for ordinary derivatives of sums and products of functions are given in the next lemma.

Lemma 6.1.8. *Let ϕ and ψ be C^1 m-forms and let θ be a C^1 ℓ-form. It holds that*

(1) $d(\phi + \psi) = (d\phi) + (d\psi)$,
(2) $d(\phi \wedge \theta) = (d\phi) \wedge \theta + (-1)^m \phi \wedge (d\theta)$.

Proof.
(1) Equation (1) follows immediately from Definition 6.1.7(3).

(2) Note that in case $m = 0$, equation (2) reduces to Definition 6.1.7(2) and the usual product rule. Now suppose that $m \geq 1$, $\phi = f\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}$, and $\theta = g\, dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_\ell}$. Using Definition 6.1.7(2), we compute

$$\begin{aligned}
&d(\phi \wedge \theta) \\
&\quad = d(fg)\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_\ell} \\
&\quad = [(df)\, g + f\, (dg)]\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m} \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_\ell} \\
&\quad = [(df) \wedge dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}] \wedge [g\, dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_\ell}] \\
&\qquad + (-1)^m [f\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}] \wedge [(dg) \wedge dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_\ell}] \\
&\quad = (d\phi) \wedge \theta + (-1)^m \phi \wedge (d\theta)\,.
\end{aligned}$$

□

In contrast to the situation for ordinary derivatives of functions, repeated exterior differentiation results in a trivial form.

Theorem 6.1.9. *If the differential m-form $\phi : U \to \bigwedge^m (\mathbb{R}^N)$ is C^k, $k \geq 2$, then $d\, d\phi = 0$ holds.*

Proof. For $m = 0$, ϕ is a real-valued function, so we have

$$\begin{aligned}
d\, d\phi &= \sum_{j \neq i} \sum_{i} \frac{\partial}{\partial x_j} \left(\frac{\partial \phi}{\partial x_i} \right) dx_j \wedge dx_i \\
&= \sum_{i<j} \left[\frac{\partial}{\partial x_i} \left(\frac{\partial \phi}{\partial x_j} \right) - \frac{\partial}{\partial x_j} \left(\frac{\partial \phi}{\partial x_i} \right) \right] dx_i \wedge dx_j = 0\,.
\end{aligned}$$

For $m \geq 1$ and $\phi = f\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_m}$, we have

$$\begin{aligned}
d\, d\phi &= \sum_{\substack{j \neq i \\ j \notin \{i_1, i_2, \ldots, i_m\}}} \sum_{i \notin \{i_1, i_2, \ldots, i_m\}} \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) dx_j \wedge dx_i \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_m} \\
&= \sum_{\substack{i<j \\ i,j \notin \{i_1, i_2, \ldots, i_m\}}} \left[\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) - \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) \right] dx_i \wedge dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_m} \\
&= 0\,.
\end{aligned}$$

The result now follows from the linearity of exterior differentiation. □

Definition 6.1.10.

(1) An m-form ϕ is said to be *closed* if $d\phi = 0$.
(2) An m-form ϕ is said to be *exact* if there exists an $(m-1)$-form ψ such that $d\psi = \phi$.

Remark 6.1.11. Theorem 6.1.9 tells us that every exact form is closed. It is *not* the case that every closed form is exact. In fact, the distinction between closed forms and exact forms underlies the celebrated theorem of Georges de Rham (1903–1990) relating the geometrically defined singular cohomology of a smooth manifold to the cohomology defined by differential forms (see [DRh 31] or Theorem 29A in Chapter IV of [Whn 57]).

6.2 Stokes's Theorem

Motivation

Stokes's theorem[1] expresses the equality of the integral of a differential form over the boundary of a surface and the integral of the exterior derivative of the form over the surface itself. The simplest instance of this equality is found in the part of the fundamental theorem of calculus that assures us that the difference between the values of a (continuously differentiable) function at the endpoints of an interval is equal to the integral of the derivative of the function over that interval—here the interval plays the role of the surface and the endpoints form the boundary of that surface. In fact, Stokes's theorem can be considered the higher-dimensional generalization of the fundamental theorem of calculus.

Oriented Rectangular Solids in $\mathbb{R}^N$

In order to state Stokes's theorem, one needs to define the oriented geometric boundary of an m-dimensional surface. In fact, the general definitions are designed so that the proof of Stokes's theorem can be reduced to the special case of a nicely bounded region in $\mathbb{R}^N$, indeed, to the even more special case of a rectangular solid that has its faces parallel to the coordinate hyperplanes.

The space $\mathbb{R}^N$ itself is oriented by the unit N-vector $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$. The orientation of a Lebesgue measurable subset of $\mathbb{R}^N$ will be induced by the orientation of $\mathbb{R}^N$ as described in the next definition.

Definition 6.2.1. Let $U \subseteq \mathbb{R}^N$ be $\mathcal{L}^N$-measurable, and let ω be a continuous differential N-form defined on U.

(1) The integral of ω over U is defined by setting

$$\int_U \omega = \int_U \langle \omega(x), \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \rangle \, d\mathcal{L}^N(x) \,. \tag{6.7}$$

Note that on the left-hand side of (6.7), U denotes the *oriented* set, while on the right-hand side, U denotes the set of points. On the left-hand side of (6.7), U is

[1] George Gabriel Stokes (1819–1903).

deemed to have the *positive orientation* given by the unit N-vector $\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N$. One must recognize from the context which meaning of U is being used. In Chapter 7, we will introduce a notation that allows us to explicitly indicate when U is to be considered an oriented set.

(2) If U is to be given the opposite, or negative, orientation, the resulting oriented set will be denoted by $-U$. We define

$$\int_{-U} \omega = \int_U -\langle \omega(x), \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \rangle \, d\mathcal{L}^N(x) \,. \tag{6.8}$$

Definition 6.2.1 gives us a broadly applicable definition of the integral for an oriented set of top dimension. The matter is much more difficult for lower-dimensional sets.

A lower-dimensional case that is straightforward is that of a singleton set consisting of the point $p \in \mathbb{R}^N$. The point itself will be considered to be positively oriented. A 0-form is simply a function, and the "integral" over p is evaluation at p. Traditionally, evaluation at a point is called a *Dirac delta function*,[2] so we will use the notation

$$\boldsymbol{\delta}_p(f) = f(p)$$

for any real-valued function whose domain includes p.

The next definition will specify a choice of orientation for an $(N-1)$-dimensional rectangular solid in $\mathbb{R}^N$ that is parallel to a coordinate hyperplane.

Definition 6.2.2. Suppose that $N \geq 2$.

(1) An $(N-1)$-dimensional rectangular solid, parallel to a coordinate hyperplane in $\mathbb{R}^N$, is a set of the form

$$\mathcal{F} = [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times \{c\} \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_N, b_N] \,,$$

where $a_i < b_i$ for $i = 1, \ldots, i-1, i+1, \ldots, N$.

(2) The $(N-1)$-dimensional rectangular solid $\mathcal{F} \subseteq \mathbb{R}^N$ will be oriented by the $(N-1)$-vector

$$\widehat{\mathbf{e}}_i = \bigwedge_{j \neq i} \mathbf{e}_j = \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{i-1} \wedge \mathbf{e}_{i+1} \wedge \cdots \wedge \mathbf{e}_N \,.$$

(3) Let ω be a continuous $(N-1)$-form defined on $\mathcal{F}$. The *integral of ω over $\mathcal{F}$* is defined by

$$\int_{\mathcal{F}} \omega = \int_{\mathcal{F}} \langle \omega(x), \widehat{\mathbf{e}}_i \rangle \, d\mathcal{H}^{N-1}(x) \,.$$

Similarly, the *integral of ω over $-\mathcal{F}$* is defined by

$$\int_{-\mathcal{F}} \omega = \int_{\mathcal{F}} -\langle \omega, \widehat{\mathbf{e}}_i \rangle \, d\mathcal{H}^{N-1} \,.$$

Note that $\int_{-\mathcal{F}} \omega = -\int_{\mathcal{F}} \omega$ holds.

[2] Paul Adrien Maurice Dirac (1902–1984).

(4) For a formal linear combination of $(N-1)$-dimensional rectangular solids as described in (1),

$$\sum \alpha_\ell \mathcal{F}_\ell \,, \tag{6.9}$$

we define

$$\int_{\sum \alpha_\ell \mathcal{F}_\ell} \omega = \sum \alpha_\ell \int_{\mathcal{F}_\ell} \omega \,. \tag{6.10}$$

We can now define the oriented boundary of the rectangular solid in $\mathbb{R}^N$ that has its faces parallel to the coordinate hyperplanes.

Definition 6.2.3. Let

$$\mathcal{R} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_N, b_N] \,,$$

where $a_i < b_i$, for $i = 1, 2, \ldots, N$.

(1) If $N \geq 2$, then for $i = 1, 2, \ldots, N$, set

$$\begin{aligned} \mathcal{R}_i^+ &= [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times \{b_i\} \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_N, b_N] \,, \\ \mathcal{R}_i^- &= [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times \{a_i\} \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_N, b_N] \,. \end{aligned}$$

In case $N = 1$, set $\mathcal{R}_1^+ = \boldsymbol{\delta}_{b_1}$ and $\mathcal{R}_1^- = \boldsymbol{\delta}_{a_1}$.

(2) The oriented boundary of $\mathcal{R}$, denoted by $\partial_{\text{o}} \mathcal{R}$ to distinguish it from the topological boundary, is the formal sum

$$\partial_{\text{o}} \mathcal{R} = \begin{cases} \boldsymbol{\delta}_{b_1} - \boldsymbol{\delta}_{a_1} & \text{if } N \geq 1 \,, \\ \displaystyle\sum_{i=1}^{N} (-1)^{i-1} \left(\mathcal{R}_i^+ - \mathcal{R}_i^- \right) & \text{if } N \geq 2 \,. \end{cases}$$

Stokes's Theorem on a Rectangular Solid

We now state and prove the basic form of Stokes's theorem.

Theorem 6.2.4. *Let*

$$\mathcal{R} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_N, b_N] \,,$$

where $a_i < b_i$, for $i = 1, 2, \ldots, N$. If ϕ is a C^k, $k \geq 1$, $(N-1)$-form on an open set containing $\mathcal{R}$, then it holds that

$$\int_{\partial_{\text{o}} \mathcal{R}} \phi = \int_{\mathcal{R}} d\phi \,.$$

Proof. For $N = 1$, the result is simply the fundamental theorem of calculus, so we will suppose that $N \geq 2$.

Write

$$\phi = \sum_{i=1}^{N} \phi_i \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N \, .$$

It suffices to prove that

$$\int_{\mathcal{R}} d(\phi_i \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N)$$
$$= \int_{\partial_{\mathrm{o}} \mathcal{R}} (\phi_i \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N)$$

holds for each $1 \leq i \leq N$.

Fix an i between 1 and N. We compute

$$d(\phi_i \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N)$$
$$= (d\phi_i) \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$
$$= \sum_{j=1}^{N} \frac{\partial \phi_i}{\partial x_j} \; dx_j \wedge \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$
$$= \frac{\partial \phi_i}{\partial x_i} \; dx_i \wedge dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$
$$= \frac{\partial \phi_i}{\partial x_i} \; (-1)^{i-1} \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_i \wedge dx_{i+1} \wedge \cdots \wedge dx_N \, ,$$

so we have

$$\int_{\mathcal{R}} d(\phi_i \; dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N)$$
$$= \int_{\mathcal{R}} (-1)^{i-1} \frac{\partial \phi_i}{\partial x_i} \; \langle dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N, \; \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \mathbf{e}_N \rangle \, d\mathcal{L}^N$$
$$= (-1)^{i-1} \int_{\mathcal{R}} \frac{\partial \phi_i}{\partial x_i} \, d\mathcal{L}^N \, .$$

By applying Fubini's theorem to evaluate $\int_{\mathcal{R}} (\partial \phi_i / \partial x_i) \, d\mathcal{L}^N$, we obtain

$$\int_{\mathcal{R}} \frac{\partial \phi_i}{\partial x_i} \, d\mathcal{L}^N$$
$$= \int_{[a_1,b_1] \times \cdots \times [a_{i-1},b_{i-1}] \times [a_{i+1},b_{i+1}] \times \cdots \times [a_N,b_N]} \left(\int_{a_i}^{b_i} \frac{\partial \phi_i}{\partial x_i} \, d\mathcal{L}^1(x_i) \right) d\mathcal{L}^{N-1}$$
$$= \int_{[a_1,b_1] \times \cdots \times [a_{i-1},b_{i-1}] \times [a_{i+1},b_{i+1}] \times \cdots \times [a_N,b_N]} \phi_i \big|_{x_i = b_i} \; d\mathcal{L}^{N-1}$$

$$- \int_{[a_1,b_1]\times\cdots\times[a_{i-1},b_{i-1}]\times[a_{i+1},b_{i+1}]\times\cdots\times[a_N,b_N]} \phi_i|_{x_i=a_i} \, d\mathcal{L}^{N-1}$$

$$= \int_{\mathcal{R}_i^+} \phi_i \, d\mathcal{H}^{N-1} - \int_{\mathcal{R}_i^-} \phi_i \, d\mathcal{H}^{N-1} .$$

We conclude that

$$\int_{\mathcal{R}} d(\phi_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N)$$

$$= (-1)^{i-1} \left(\int_{\mathcal{R}_i^+} \phi_i \, d\mathcal{H}^{N-1} - \int_{\mathcal{R}_i^-} \phi_i \, d\mathcal{H}^{N-1} \right) . \tag{6.11}$$

On the other hand, we compute

$$\int_{\partial_\mathrm{o} \mathcal{R}} \phi_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$

$$= \sum_{j=1}^{N} (-1)^{j-1} \int_{\mathcal{R}_j^+} \phi_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$

$$- \sum_{j=1}^{N} (-1)^{j-1} \int_{\mathcal{R}_j^-} \phi_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N$$

$$= \sum_{j=1}^{N} (-1)^{j-1} \int_{\mathcal{R}_j^+} \phi_i \, \langle dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N, \widehat{\mathbf{e}}_j \rangle \, d\mathcal{H}^{N-1}$$

$$- \sum_{j=1}^{N} (-1)^{j-1} \int_{\mathcal{R}_j^-} \phi_i \, \langle dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N, \widehat{\mathbf{e}}_j \rangle \, d\mathcal{H}^{N-1}$$

$$= (-1)^{i-1} \left(\int_{\mathcal{R}_i^+} \phi_i \, d\mathcal{H}^{N-1} - \int_{\mathcal{R}_i^-} \phi_i \, d\mathcal{H}^{N-1} \right) . \tag{6.12}$$

Since (6.11) and (6.12) agree, we have the result. □

The Gauss–Green Theorem

A vector field on an open set $U \subseteq \mathbb{R}^N$ is a function $V : U \to \mathbb{R}^N$. The component functions V_i, $i = 1, 2, \ldots, N$, are defined by setting

$$V_i(x) = V(x) \cdot \mathbf{e}_i ,$$

so we have $V = \sum_{i=1}^N V_i \, \mathbf{e}_i$. We say that V is C^k if the component functions are C^k. The *divergence* of V, denoted by div V, is the real-valued function

$$\operatorname{div} V = \sum_{i=1}^{N} \frac{\partial V_i}{\partial x_i} .$$

Given an $(N-1)$-form ϕ in $\mathbb{R}^N$ we can associate with it a vector field V by the following means: if ϕ is written

$$\phi = \sum_{i=1}^{N} \phi_i \, dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_N ,$$

then set

$$V = \sum_{i=1}^{N} (-1)^{i-1} \phi_i \, \mathbf{e}_i .$$

Direct calculation shows that

$$d\phi = (\operatorname{div} V) \, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N$$

holds. One can also verify that

$$\int_{\partial_o \mathcal{R}} \phi = \int_{\partial \mathcal{R}} V \cdot \mathbf{n} \, d\mathcal{H}^{N-1}$$

holds, where $\mathbf{n}$ is the outward-pointing unit vector orthogonal to the topological boundary $\partial\mathcal{R}$. We call $\mathbf{n}$ the *outward unit normal vector*.

By converting the statement of Theorem 6.2.4 about integrals of forms into the corresponding statement about vector fields, one obtains the following result, called the Gauss–Green theorem[3] or the divergence theorem:

Corollary 6.2.5. *If V is a C^1 vector field on an open set containing $\mathcal{R}$, then*

$$\int_{\mathcal{R}} \operatorname{div} V \, d\mathcal{L}^N = \int_{\partial \mathcal{R}} V \cdot \mathbf{n} \, d\mathcal{H}^{N-1} .$$

By piecing together rectangular solids and estimating the error at the boundary, one can prove a more general version of Theorem 6.2.4 or of Corollary 6.2.5. Thus we have the following result.

Theorem 6.2.6. *Let $A \subseteq \mathbb{R}^N$ be a bounded open set with C^1 boundary, and let $\mathbf{n}(x)$ denote the outward unit normal to ∂A at x. If V is a C^1 vector field defined on $\overline{A}$, then*

$$\int_{A} \operatorname{div} V \, d\mathcal{L}^N = \int_{\partial A} V \cdot \mathbf{n} \, d\mathcal{H}^{N-1} .$$

Theorem 6.2.6 is by no means the most general result available. The reader should see [Fed 69, 4.5.6] for an optimal version of the Gauss–Green theorem.

[3] Johann Carl Friedrich Gauss (1777–1855), George Green (1793–1841).

The Pullback of a Form

Definition 6.2.7. Suppose that $U \subseteq \mathbb{R}^N$ is open and $F : U \to \mathbb{R}^M$ is C^k, $k \geq 1$. Fix a point $p \in U$. If the differential m-form ϕ is defined at $F(p)$, then the *pullback* of ϕ is the m-form, defined at p, denoted by $F^\# \phi$ and evaluated on $v_1, v_2, \ldots, v_m$ by setting

$$\langle F^\# \phi(p), v_1 \wedge v_2 \wedge \cdots \wedge v_m \rangle = \langle \phi[F(p)], D_{v_1} F \wedge D_{v_2} F \wedge \cdots \wedge D_{v_m} F \rangle \,, \quad (6.13)$$

where we use the notation

$$D_{v_i} F = \langle DF,\ v_i \rangle \,,$$

for $i = 1, 2, \ldots, m$. In case $m = 0$, (6.13) reduces to $F^\# \phi = \phi \circ F$.

Remark 6.2.8. We now have three similar notations in use: $D_{v_i} F$ as above; $D_\lambda(\mu, x)$ for differentiation of measures, which was introduced in Section 4.3; and $D_S f(x)$ for the differential of f relative to the surface S, which was introduced in Section 5.3 for smooth surfaces and extended to rectifiable sets in Section 5.4. The notation that is meant should always be clear from context.

The next theorem tells us that the operations of pullback and exterior differentiation commute. This seems like an insignificant observation, but in fact, it is key to generalizing Stokes's theorem, i.e., Theorem 6.2.4.

Theorem 6.2.9. *Suppose that $U \subseteq \mathbb{R}^N$ is open and $F : U \to \mathbb{R}^M$ is C^k, $k \geq 2$. Fix a point $p \in U$. If the differential m-form ϕ is defined and C^k, $k \geq 2$, in a neighborhood of $F(p)$, then*

$$d(F^\# \phi) = F^\# (d\phi) \quad (6.14)$$

holds at p.

Proof. First we consider the case $m = 0$ in which $F^\# \phi = \phi \circ F$. Fix $v \in \mathbb{R}^N$. Using the chain rule and (6.6), we compute

$$\begin{aligned} \langle dF^\# \phi, v \rangle &= \langle d[\phi \circ F], v \rangle = \langle D[\phi \circ F],\ v \rangle \\ &= \langle D\phi[F(p)],\ \langle DF,\ v \rangle \rangle = \langle d\phi[F(p)],\ \langle DF,\ v \rangle \rangle \,. \end{aligned}$$

The most efficient argument to deal with the case $m \geq 1$ is to first consider a 1-form ϕ that can be written as an exterior derivative; that is, $\phi = d\psi$ for a 0-form ψ. Then we have

$$d(F^\# \phi) = d(F^\# d\psi) = d(dF^\# \psi) = 0 = F^\# (d\, d\psi) = F^\# (d\phi) \,.$$

Lemma 6.1.8 allows us to see that the set of forms satisfying (6.14) is closed under addition and exterior multiplication. The general case then follows by addition and exterior multiplication of 0-forms and exterior derivatives of 0-forms. □

In Appendix A.4, the reader can see an alternative argument that is less elegant, but which reveals the inner workings of interchanging a pullback and an exterior differentiation.

Stokes's Theorem

Let $\mathcal{R}$ be a rectangular solid in $\mathbb{R}^N$. If U is open with $\mathcal{R} \subseteq U \subseteq \mathbb{R}^N$ and $F : U \to \mathbb{R}^M$ is one-to-one and C^k, $k \geq 1$, then the F-image of $\mathcal{R}$ is an N-dimensional C^k surface parametrized by F. We denote this surface by

$$F_{\#}\mathcal{R}\,.$$

This definition extends to formal sums by setting $F_{\#}\Big[\sum_\alpha \mathcal{R}_\alpha\Big] = \sum_\alpha F_{\#}\mathcal{R}_\alpha$.

In Definition 6.1.6, we gave a definition for the integral of a differential form over a surface. The next lemma gives us another way of looking at that definition.

Lemma 6.2.10. *If ω is a continuous N-form defined in a neighborhood of $F(\mathcal{R})$, then*

$$\int_{F_{\#}\mathcal{R}} \omega = \int_{\mathcal{R}} F^{\#}\omega\,.$$

Proof. By Definition 6.1.6, we have

$$\int_{F_{\#}\mathcal{R}} \omega = \int_{\mathcal{R}} \left\langle \omega \circ F(t),\, \frac{\partial F}{\partial t_1} \wedge \frac{\partial F}{\partial t_2} \wedge \cdots \wedge \frac{\partial F}{\partial t_N} \right\rangle d\mathcal{L}^N(t)\,.$$

Observing that

$$\frac{\partial F}{\partial t_i} = \langle DF,\, \mathbf{e}_i \rangle\,,$$

for $i = 1, 2, \ldots, N$, we see that

$$\begin{aligned}
&\left\langle \omega \circ F(t),\, \frac{\partial F}{\partial t_1} \wedge \frac{\partial F}{\partial t_2} \wedge \cdots \wedge \frac{\partial F}{\partial t_N} \right\rangle \\
&\qquad = \langle \omega \circ F(t), \langle DF,\, \mathbf{e}_1 \rangle \wedge \langle DF,\, \mathbf{e}_2 \rangle \wedge \cdots \wedge \langle DF,\, \mathbf{e}_N \rangle \rangle \\
&\qquad = \left\langle F^{\#}\omega,\, \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \right\rangle,
\end{aligned}$$

and the result follows. □

The boundary of a smooth surface is usually defined by referring back to the space of parameters. That is our motivation for the next definition.

Definition 6.2.11. The *oriented boundary* of $F_{\#}\mathcal{R}$ will be denoted by $\partial_{\mathrm{o}} F_{\#}\mathcal{R}$ and is defined by

$$\partial_{\mathrm{o}} F_{\#}\mathcal{R} = \sum_{i=1}^{N} (-1)^{i-1} \left(F_{\#}\mathcal{R}_i^+ - F_{\#}\mathcal{R}_i^- \right) = F_{\#}\partial_{\mathrm{o}}\mathcal{R}\,.$$

Some explanation of this definition is called for because $F_{\#}\mathcal{R}_i^+$ and $F_{\#}\mathcal{R}_i^-$ do not quite fit our earlier discussion. Recall that $\mathcal{R}_i^+$ and $\mathcal{R}_i^-$ lie in planes parallel to the coordinate hyperplanes, so F restricted to either $\mathcal{R}_i^+$ or $\mathcal{R}_i^-$ can be thought of as a function on $\mathbb{R}^{N-1}$. Note that both $\mathcal{R}_i^+$ and $\mathcal{R}_i^-$ are oriented in a manner consistent with this interpretation.

We are now in a position to state and prove a general version of Stokes's theorem.

Theorem 6.2.12 (Stokes's Theorem). *Let $\mathcal{R}$ be a rectangular solid in $\mathbb{R}^N$. Suppose that U is open with $\mathcal{R} \subseteq U \subseteq \mathbb{R}^N$ and that $F : U \to \mathbb{R}^M$ is one-to-one and C^k, $k \geq 1$, with DF of rank N at every point of U. If ω is a C^k, $k \geq 2$, $(N-1)$-form defined on $F(\mathcal{R})$, then*

$$\int_{F_{\#}\mathcal{R}} d\omega = \int_{\partial_{\circ} F_{\#}\mathcal{R}} \omega\,.$$

Proof. We compute

$$\begin{aligned}\int_{F_{\#}\mathcal{R}} d\omega &\overset{\text{(Lemma 6.2.10)}}{=} \int_{\mathcal{R}} F^{\#}(d\omega) \overset{\text{(Thm. 6.2.9)}}{=} \int_{\mathcal{R}} d(F^{\#}\omega) \\ &\overset{\text{(Thm. 6.2.4)}}{=} \int_{\partial_{\circ}\mathcal{R}} F^{\#}\omega \overset{\text{(Lemma 6.2.10)}}{=} \int_{F_{\#}\partial_{\circ}\mathcal{R}} \omega \overset{\text{(Def. 6.2.11)}}{=} \int_{\partial_{\circ} F_{\#}\mathcal{R}} \omega\,.\end{aligned}$$ □

As was true for the earlier version of Stokes's theorem (Theorem 6.2.4) and for the Gauss–Green theorem (Corollary 6.2.5), a more general version of Theorem 6.2.12 may be obtained by piecing together patches of surface. Since the theory of currents gives a still more general expression to Stokes's theorem, we will defer further discussion of Stokes's theorem until we have introduced the language of currents.

7

Introduction to Currents

In the traditional setup (see our Chapter 6), a differential form is a smooth function that assigns to each point of space a covector. For the purposes of integration on smooth surfaces, de Rham cohomology, and other standard applications of geometric analysis, differential forms with smooth coefficients are the perfect device. But for applications in geometric measure theory and certain areas of partial differential equations, something more general is needed. In particular, differential forms in the raw (as just described) are *not* convenient for limit processes. Thus was born the theory of currents. The earliest provenance of currents occurs in [Sch 51] and [DRh 55], but the theory came into its own in [FF 60] and later works. See [Fed 69] for a complete bibliography as of that writing.

Intuitively, a current is a differential form with coefficients that are *distributions*. [The rigorous definition of current is more technical; this intuitive definition will suffice for our introductory remarks.] It will turn out, for example, that integration over a rectifiable set, with suitable orientation, can be thought of as a current. However, it cannot be thought of as a traditional differential form.

The main advantage of the space of currents is that it possesses useful compactness properties. Just as it is useful to extend the domain of an elliptic differential operator to L^2, with the definition of differentiation taken in the distribution sense, so that the operator becomes closed, so it is useful to study the Plateau problem, and questions of minimal surface theory, and a variety of variational problems, in the context of currents. For it turns out that a collection of currents that is bounded in a rather weak sense will have a convergent subsequence or subnet. Frequently, the limit of that sequence or net will be the solution of the variational problem that we seek. It generally requires considerable extra effort to verify in practice that that limiting current can actually be represented by integration over a regular surface; but it can be done. This has become the standard approach to a variety of extremal problems in geometric measure theory.

Currents may also be used to construct representation theorems for measures and other linear operators of geometric analysis, to produce approximation theorems, to solve partial differential equations, and to prove isoperimetric inequalities. They have become a fundamental device of geometric analysis.

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_7,

Our purpose in the present chapter is to give a rigorous but very basic introduction to the theory of currents and to indicate some of their applications. Our exposition in this chapter owes a debt to [Fed 69], [Sim 83], and [Whn 57]. For further reading, we recommend [Fed 69], [FF 60], [LY 02], and [Mat 95]. Some modern treatments of currents may also be found in [Blo 98], [Kli 91], [Lel 69], [LG 86]. We note particularly the extensive monograph [GMS 98] treating the subject of Cartesian currents—heuristically Cartesian currents can be thought of as being weak limits of graphs of smooth maps—and their application to variational problems.

7.1 A Few Words about Distributions

The theory of currents is built on the framework of distributions. We will quickly cover those portions of distribution theory for which we have immediate use. For the reader familiar with the basic theory of distributions, the main purpose will be to fix some notation. The reader who wishes to pursue some background reading should see [Hor 69], [Kra 92], [Tre 80].

Fix $M, N \in \mathbb{N}$. Let $U \subseteq \mathbb{R}^N$ be open and let V be an M-dimensional vector space. By choosing a basis, we can identify V with $\mathbb{R}^M$ and thus apply all the usual constructions of calculus. We let $\mathcal{E}(U, V)$ denote the C^∞ mappings of U into V. Now, as is customary in the theory of distributions, we define a family of seminorms. If $i \in \mathbb{Z}$, $i \geq 0$, and $K \subseteq U$ is compact, then we let, for $\phi \in \mathcal{E}(U, V)$,

$$\nu_K^i(\phi) = \sup\{\|D^j\phi(x)\| : 0 \leq j \leq i \text{ and } x \in K\}.$$

Here D^j is, of course, the jth differential and $\|D^j\phi(x)\|$ is its operator norm (also called the mapping norm—see Definition 1.1.3). Equivalently, one could use the seminorms $\widetilde{\nu}_K^i$ defined by taking the supremum over K of the partial derivatives up to and including order i of all M component functions.

The family of all the seminorms ν_K^i induces a locally convex, translation-invariant Hausdorff topology on $\mathcal{E}(U, V)$. A subbasis for the topology consists of sets of the form

$$\mathcal{O}(\psi, i, K, r) = \{\phi \in \mathcal{E}(U, V) : \nu_K^i(\phi - \psi) < r\}$$

for $\psi \in \mathcal{E}(U, V)$ fixed and $r > 0$. Then $\mathcal{E}(U, V)$ is a topological vector space.

We define $\mathcal{E}'(U, V)$ to be the set of all continuous, real-valued linear functionals on $\mathcal{E}(U, V)$. We endow $\mathcal{E}'(U, V)$ with the weak topology generated by the subbasis consisting of sets of the form

$$\{T \in \mathcal{E}'(U, V) : a < T(\phi) < b\}$$

for $\phi \in \mathcal{E}(U, V)$ and $a < b \in \mathbb{R}$. This topology is also referred to as the *weak-$*$ topology*.

Now, for $\phi \in \mathcal{E}(U, V)$, recall that $\operatorname{supp}\phi$, the support of ϕ, is defined by

$$\operatorname{supp}\phi = U \setminus \bigcup\{W : W \text{ is open, } \phi(x) = 0 \text{ whenever } x \in W\}.$$

We would like to give a similar definition of the support of an element of $\mathcal{E}'(U, V)$. To match the standard notation for currents, we will denote the support of $T \in \mathcal{E}'(U, V)$ by $\operatorname{spt} T$ instead of $\operatorname{supp} T$. For $T \in \mathcal{E}'(U, V)$, we define

$$\operatorname{spt} T =$$
$$U \setminus \bigcup\{W : W \text{ is open, } T(\phi) = 0 \text{ whenever } \phi \in \mathcal{E}(U, V), \operatorname{supp}\phi \subseteq W\}.$$

This is the *support* of T. Then each element of $\mathcal{E}'(U, V)$ is compactly supported just because, given $T \in \mathcal{E}'(U, V)$, there exist $0 < M < \infty$, $i \in \mathbb{Z}^+$, and $K \subset\subset \mathbb{R}^N$ such that

$$|T(\phi)| \leq M \cdot \nu_K^i(\phi)$$

holds, for all $\phi \in \mathcal{E}(U, V)$,[1] and this inequality implies $\operatorname{spt} T \subseteq K$. In conclusion, we see that $\mathcal{E}'(U, V)$ is the union of its closed subsets

$$\mathcal{E}'_K(U, V) \equiv \{T \in \mathcal{E}'(U, V) : \operatorname{spt} T \subseteq K\}$$

corresponding to all compact subsets K of U. In fact, one may see (and this is important in practice) that all the members of any convergent sequence in $\mathcal{E}'(U, V)$ belong to some single set $\mathcal{E}'_K(U, V)$.

For each compact $K \subseteq U$ we let

$$\mathcal{D}_K(U, V) = \{\phi \in \mathcal{E}(U, V) : \operatorname{supp}\phi \subseteq K\}.$$

We notice that $\mathcal{D}_K(U, V)$ is closed in $\mathcal{E}(U, V)$. Now we define the vector space

$$\mathcal{D}(U, V) = \bigcup\{\mathcal{D}_K(U, V) : K \text{ is a compact subset of } U\}.$$

We endow $\mathcal{D}(U, V)$ with the largest topology such that the inclusion maps $\mathcal{D}_K(U, V) \hookrightarrow \mathcal{D}(U, V)$ are all continuous. It follows that a subset W of $\mathcal{D}(U, V)$ is open if and only if $W \cap \mathcal{D}_K(U, V)$ belongs to the relative topology of $\mathcal{D}_K(U, V)$ in $\mathcal{E}(U, V)$. Thus the inclusion map $\mathcal{D}(U, V) \hookrightarrow \mathcal{E}(U, V)$ is continuous. This map is *not* a homeomorphism unless $U = \emptyset$ or $M = 0$. But it should be noted that the topologies of $\mathcal{E}(U, V)$ and $\mathcal{D}(U, V)$ induce the same relative topology on each $\mathcal{D}_K(U, V)$.

Now we define the dual space $\mathcal{D}'(U, V)$ to be the vector space of all continuous, real-valued linear functionals on $\mathcal{D}(U, V)$. We endow $\mathcal{D}'(U, V)$ with the weak topology generated by the sets

$$\{T \in \mathcal{D}'(U, V) : a < T(\phi) < b\}$$

corresponding to $\phi \in \mathcal{D}(U, V)$ and $a < b \in \mathbb{R}$. Again, this topology is sometimes referred to as the *weak-$*$ topology*.

[1] To see this, note that $T^{-1}(-1, 1)$ must be open in $\mathcal{E}(U, V)$ and consider a neighborhood basis of $0 \in \mathcal{E}(U, V)$.

Each member of $\mathcal{D}(U, V)$ has compact support. However, the support of a member of $\mathcal{D}'(U, V)$ need not be compact. For example, if $U = V = \mathbb{R}$ and $\boldsymbol{\delta}_p$ is the Dirac delta mass at p [i.e., the functional defined by $\boldsymbol{\delta}_p(\phi) = \phi(p)$] then

$$\eta \equiv \sum_{j=1}^{\infty} 2^{-j} \boldsymbol{\delta}_j$$

is an element of $\mathcal{D}'(U, V)$ that certainly does not have compact support. In point of fact, a real-valued linear functional T on $\mathcal{D}(U, V)$ belongs to $\mathcal{D}'(U, V)$ if and only if, for each compact subset $K \subseteq U$, there are nonnegative integers i and M such that

$$T(\phi) \leq M \cdot \nu_K^i(\phi) \text{ whenever } \phi \in \mathcal{D}_K(U, V)\,.$$

An element of $\mathcal{D}'(U, V)$ is called a *distribution in U with values in V*. Since $\mathcal{D}(U, V) \subseteq \mathcal{E}(U, V)$, it follows that $\mathcal{E}'(U, V) \subseteq \mathcal{D}'(U, V)$. We sometimes refer to the elements of $\mathcal{E}'(U, V)$ as the *compactly supported distributions*.

In case $U = V = \mathbb{R}$, we see that any L^1 function f defines a distribution $T_f \in \mathcal{D}'(\mathbb{R}, \mathbb{R})$ by setting

$$T_f(\phi) = \int_{-\infty}^{\infty} f(t)\phi(t)\, d\mathcal{L}^1(t)\,,$$

for each $\phi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$. If f is continuously differentiable, then integration by parts gives us

$$T_{f'}(\phi) = \int_{-\infty}^{\infty} f'(t)\phi(t)\, d\mathcal{L}^1(t) = -\int_{-\infty}^{\infty} f(t)\phi'(t)\, d\mathcal{L}^1(t) = -T_f(\phi')\,.$$

This last equation motivates the general definition for differentiation of distributions.

Definition 7.1.1. For $T \in \mathcal{D}'(U, V)$, the *partial derivative* of T with respect to the ith variable, $1 \leq i \leq N$, is the element $D_{x_i}T$ of $\mathcal{D}'(U, V)$ defined by setting

$$(D_{x_i}T)(\phi) = -T(\partial\phi/\partial x_i)\,.$$

A similar definition is applicable to the currents with compact support.

We use the notation $D_{x_i}T$ (instead of $\partial T/\partial x_i$) for the partial derivative of the distribution T to avert possible confusion later with the boundary operator on currents.

The distributions in $\mathcal{D}'(U, \mathbb{R})$ are sometimes called *generalized functions*. The next result generalizes the fact that if the derivative of a function vanishes, then the function is constant.

Proposition 7.1.2. *If $T \in \mathcal{D}'(\mathbb{R}, \mathbb{R})$ and $D_x T = 0$, i.e., $T(\phi') = 0$, for all $\phi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$, then there is $c \in \mathbb{R}$ such that $T = c$, i.e., $T(\phi) = c \int_{\mathbb{R}} \phi\, d\mathcal{L}^1$, for all $\phi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$.*

Proof. Fix $\psi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ with $\int_{\mathbb{R}} \psi \, d\mathcal{L}^1 \neq 0$. Given $\phi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$, set

$$f(t) = \int_{-\infty}^{t} [\phi(\tau) + q\,\psi(\tau)]\, d\mathcal{L}^1(\tau) \text{ where } q = -\int_{\mathbb{R}} \phi \, d\mathcal{L}^1 \Big/ \int_{\mathbb{R}} \psi \, d\mathcal{L}^1 \; .$$

Then $f \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ and $f' = \phi + q\,\psi$. Thus we have

$$0 = -D_x T(f) = T(f') = T(\phi) + q\, T(\psi) \, ,$$

and we see that the result holds with

$$c = \left(T(\psi) \Big/ \int_{\mathbb{R}} \psi \, d\mathcal{L}^1 \right) . \qquad \square$$

Proposition 7.1.2 is the simplest case of a more general result that tells us that if all the partial derivatives of a distribution on $\mathbb{R}^N$ vanish, then the distribution is just a constant. Another form of that theorem in the context of currents is called the constancy theorem, and that result will be particularly important to us later. We treat it in detail below.

7.2 The Definition of a Current

With notation as in the last section, we define

$$\mathcal{E}^M(U) = \mathcal{E}\left(U, \bigwedge\nolimits^M \mathbb{R}^N\right) , \qquad \mathcal{E}_M(U) = \mathcal{E}'\left(U, \bigwedge\nolimits^M \mathbb{R}^N\right) ,$$
$$\mathcal{D}^M(U) = \mathcal{D}\left(U, \bigwedge\nolimits^M \mathbb{R}^N\right) , \qquad \mathcal{D}_M(U) = \mathcal{D}'\left(U, \bigwedge\nolimits^M \mathbb{R}^N\right) .$$

Thus, in brief, $\mathcal{E}^M(U)$ is the space of differential forms on U with degree M and having C^∞ coefficients. Also $\mathcal{D}^M(U)$ is the subspace of $\mathcal{E}^M(U)$ having coefficients of compact support in U. The spaces $\mathcal{D}_M(U)$ and $\mathcal{E}_M(U)$ are duals of these spaces. The members of $\mathcal{D}_M(U)$ are called the *M-dimensional currents* on U, and the image of $\mathcal{E}_M(U)$ in $\mathcal{D}_M(U)$ consists of all M-dimensional currents with compact support in U. To summarize, we have $\mathcal{D}^M(U) \subseteq \mathcal{E}^M(U)$ and $\mathcal{E}_M(U) \subseteq \mathcal{D}_M(U)$.

A simple example of an M-dimensional current on U is provided by considering an $\mathcal{L}^N$-measurable function $\xi : U \to \bigwedge_M (\mathbb{R}^N)$ with the property that its operator norm $\|\xi\|$ has finite integral over U, i.e., $\|\xi\| \in L^1(U)$. Then define $T \in \mathcal{D}_M(U)$ by setting

$$T(\phi) = \int_U \langle\, \phi(x),\ \xi(x)\, \rangle \, d\mathcal{L}^N(x)$$

for each $\phi \in \mathcal{D}^M(U)$. Certainly this example can be generalized by considering measures μ different from $\mathcal{L}^N$. The function ξ will then need to be μ-measurable and satisfy $\int_U \|\xi\| \, d\mu < \infty$ or, to generalize further, $\int_K \|\xi\| \, d\mu < \infty$ for each compact $K \subseteq U$. As will become clear, such examples in $\mathcal{D}_M(U)$ are particularly useful when the measure μ is concentrated on a set of dimension M.

Now we define some operations on currents that are dual to those on differential forms. Those who know some algebraic topology will recognize some of the classical cohomology operations lurking in the background (see [BT 82] or [Spa 66]).

Let $T \in \mathcal{D}_M(U)$. If $\phi \in \mathcal{E}^k(U)$ and $k \le M$ then we define

$$T \llcorner \phi \in \mathcal{D}_{M-k}(U)$$

according to the identity

$$(T \llcorner \phi)(\psi) \equiv T(\phi \wedge \psi) \quad \text{for all } \psi \in \mathcal{D}^{M-k}(U)\,.$$

Now let ξ be a p-vector field with C^∞ coefficients on U (that is, a smooth map into $\bigwedge_p \mathbb{R}^N$). We let

$$T \wedge \xi \in \mathcal{D}_{M+p}(U)$$

be specified by the identity

$$(T \wedge \xi)(\psi) \equiv T(\xi \rfloor \psi) \quad \text{for all } \psi \in \mathcal{D}^{M+p}(U)\,,$$

where $\xi \rfloor \psi$ is the *interior product* characterized by $\langle \xi \rfloor \psi, \alpha \rangle = \langle \psi, \alpha \wedge \xi \rangle$ for $\alpha \in \bigwedge_M \mathbb{R}^N$. (This last definition is consistent with [Fed 69, 1.5] despite the fact that in the dual pairing $\langle \cdot, \cdot \rangle$, we are placing M-covectors on the left and M-vectors on the right.)

Since the interior product used above may not be familiar, we will say more about it here. Suppose that

$$1 \le i_1 < \cdots < i_p \le N \ \text{ and } \ 1 \le j_1 < \cdots < j_{M+p} \le N\,.$$

If $\{ i_1, \ldots, i_p \} \not\subseteq \{ j_1, \ldots, j_{M+p} \}$, then

$$(\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}) \rfloor (dx_{j_1} \wedge \cdots \wedge dx_{j_{M+p}}) = 0\,.$$

On the other hand, if $\{ i_1, \ldots, i_p \} \subseteq \{ j_1, \ldots, j_{M+p} \}$, then we write

$$\{ k_1, \ldots, k_M \} = \{ j_1, \ldots, j_{M+p} \} \setminus \{ i_1, \ldots, i_p \}\,,$$

where $1 \le k_1 < \cdots < k_M \le N$. In this case, we have

$$(\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_p}) \rfloor (dx_{j_1} \wedge \cdots \wedge dx_{j_{M+p}}) = \sigma\, dx_{k_1} \wedge \cdots \wedge dx_{k_M}\,,$$

where $\sigma \in \{ -1, \ +1 \}$ is the sign of the permutation

$$(j_1, \ldots, j_{M+p}) \longmapsto (k_1, \ldots, k_M, i_1, \ldots, i_p)\,.$$

In practice, it is often not necessary to require that ϕ and ξ have C^∞ coefficients. It is only necessary to be able to make sense of the expressions that we use. Thus, in the special case that T is given by an integral, we need only require that ϕ and ξ be measurable and that their norms have finite integral over every compact set in U. In particular, we may let

$$T \llcorner A = T \llcorner \chi_A \in \mathcal{E}_M(U)$$

for each set A that is measurable with respect to the measure used to define T.

One of the features that makes currents important is that there is an associated homology theory. For this we need a boundary operator. If $M \geq 1$ and $T \in \mathcal{D}_M(U)$, then we let the *boundary* of T,

$$\partial T \in \mathcal{D}_{M-1}(U) ,$$

be defined by setting

$$(\partial T)(\psi) = T(d\psi) \tag{7.1}$$

whenever $\psi \in \mathcal{D}^{M-1}(U)$. This definition is motivated by and consistent with Stokes's theorem, as we shall see later. It is also convenient to define $\partial T = 0$ for $T \in \mathcal{D}_0(U)$.

The reader should keep in mind that, for a current $T \in \mathcal{D}_M(U)$, there is a significant distinction between the boundary of the current, $\partial T \in \mathcal{D}_{M-1}(U)$, defined in (7.1) and a partial derivative of the current, $D_{x_\ell} T \in \mathcal{D}_M(U)$, $1 \leq \ell \leq N$. Definition 7.1.1 tells us that, for any C^∞ real-valued function with compact support in U and any choice of $1 \leq j_1 < \cdots < j_M \leq N$,

$$D_{x_\ell} T(\phi \, dx_{j_1} \wedge \cdots \wedge dx_{j_M}) = -T \left[\, (D_{x_\ell}\phi) \, dx_{j_1} \wedge \cdots \wedge dx_{j_M} \, \right]$$

holds, where

$$D_{x_\ell}\phi = \frac{\partial \phi}{\partial x_\ell}$$

is the ordinary partial derivative of the real-valued function ϕ.

Proposition 7.2.1. *Suppose that ϕ and ξ have C^∞ coefficients on U, where ϕ is a form of degree k and ξ is a p-vector field. Then*

(1) $\partial(\partial T) = 0 \quad \textit{if} \quad \dim T \geq 2\,;$
(2) $(\partial T) \llcorner \phi = T \llcorner d\phi + (-1)^k \partial(T \llcorner \phi)\,;$
(3) $\partial T = -\sum_{j=1}^{N} (D_{x_j} T) \llcorner dx_j \quad \textit{if} \ \dim T \geq 1\,;$
(4) $T = \sum_{1 \leq j_1 < \cdots < j_M \leq N} \left[T \llcorner dx_{j_1} \wedge \cdots \wedge dx_{j_M} \right] \wedge \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_M}\,;$
(5) $D_{x_j}(T \llcorner \phi) = (D_{x_j} T) \llcorner \phi + T \llcorner (\partial\phi/\partial x_j)\,;$
(6) $D_{x_j}(T \wedge \xi) = (D_{x_j} T) \wedge \xi + T \wedge (\partial\xi/\partial x_j)\,;$
(7) $(T \wedge \xi) \llcorner \phi = T \wedge (\xi \llcorner \phi) \quad \textit{if} \ \dim T = 0 \ \textit{and} \ k \leq p\,;$
(8) $\partial(T \wedge \xi) = -T \wedge \mathit{div}\,\xi - \sum_{j=1}^{N} (D_{x_j} T) \wedge (\xi \llcorner dx_j) \quad \textit{if} \ \dim T = 0 \leq p\,.$

In the above, the partial derivatives $\partial\phi/\partial x_j$ of the form ϕ and $\partial\xi/\partial x_j$ of the vector field ξ are obtained by differentiating the coefficient functions. The reader may easily verify the identities given in the proposition.

Currents Representable by Integration

If $U \subseteq \mathbb{R}^N$ is an open set and μ is a Radon measure on U (see Definition 1.2.11), then the functional

$$\varphi \longmapsto \int_U \varphi \, d\mu$$

is positive (i.e., $\int_U \varphi \, d\mu \geq 0$ whenever $\varphi \geq 0$), $\mathbb{R}$-linear, and continuous on the space of compactly supported continuous functions on U. The topology on the compactly supported continuous functions can be characterized by defining φ_0 to be the limit of the sequence $\{\varphi_j\}$ if and only if $\varphi_j \to \varphi_0$ uniformly and, in addition, $\bigcup_j \operatorname{supp} \varphi_j$ is a compact subset of U.

The Riesz representation theorem, i.e., Theorem 4.4.1, tells us that every positive, $\mathbb{R}$-linear, continuous functional on the space of compactly supported continuous functions arises in this way. Similarly, each $\mathbb{R}$-linear, continuous functional T on the space of compactly supported continuous functions gives rise to a pair of mutually singular Radon measures μ_1 and μ_2 such that

$$T(\varphi) = \int_U \varphi \, d\mu_1 - \int_U \varphi \, d\mu_2 .$$

For our purposes, it is more convenient to form the total variation measure μ by setting $\mu = \mu_1 + \mu_2$, to define a Borel function f that equals $+1$ at μ_1-almost every point and equals -1 at μ_2-almost every point, and to write then

$$T(\varphi) = \int_U f \, \varphi \, d\mu \tag{7.2}$$

(see Figure 7.1).

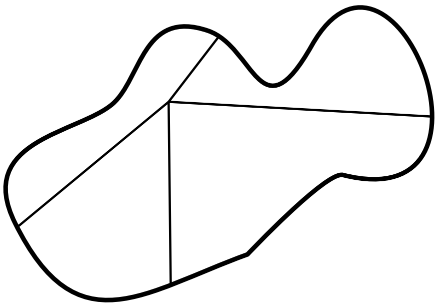

Fig. 7.1. A current representable by integration.

We would like to know which 0-dimensional currents $T \in \mathcal{D}'(U, \mathbb{R})$ can be represented by integrals of Radon measures. Not every 0-dimensional current can be so written (consider for instance derivatives of the Dirac delta $\boldsymbol{\delta}_p$). The characterizing property is that for each open $W \subset\subset U$ there exists an $M < \infty$ such that

$$|T(\phi)| \leq M \sup\{\, |\phi(x)| : x \in U \,\} \tag{7.3}$$

holds for all $\phi \in \mathcal{D}(U, \mathbb{R})$. In fact, when (7.3) holds, T extends to all compactly supported continuous functions on U, thereby defining an $\mathbb{R}$-linear, continuous functional.

Now suppose that $T \in \mathcal{D}_M(U)$. We define the *mass* of T on the open set U by

$$\mathbf{M}(T) = \sup_{\substack{|\omega| \leq 1 \\ \omega \in \mathcal{D}^M(U)}} T(\omega)\,.$$

If $W \subseteq U$ is an open subset then we have the refined notion of mass given by

$$\mathbf{M}_W(T) = \sup_{\substack{|\omega| \leq 1,\, \omega \in \mathcal{D}^M(U) \\ \operatorname{supp} \omega \subseteq W}} T(\omega)\,.$$

Notice that if $\mathbf{M}_W(T) < \infty$ for all open $W \subset\subset U$, then, for each sequence $1 \leq j_1 < j_2 < \cdots < j_M \leq N$, the 0-dimensional current

$$T \llcorner (dx_{j_1} \wedge dx_{j_2} \wedge \cdots \wedge dx_{j_M})$$

satisfies the condition (7.3) and thus defines a total variation measure $\mu_{j_1,\ldots,j_M}$ and function $f_{j_1,\ldots,j_M}$ as in (7.2). Using the identity

$$T = \sum_{1 \leq j_1 < \cdots < j_M \leq N} \left[T \llcorner dx_{j_1} \wedge \cdots \wedge dx_{j_M} \right] \wedge \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_M}\,,$$

we see that we can add together the total variation measures $\mu_{j_1,\ldots,j_M}$ and functions $f_{j_1,\ldots,j_M}\, \mathbf{e}_{j_1} \wedge \cdots \wedge \mathbf{e}_{j_M}$ and normalize the resulting function to obtain a Radon measure μ_T on U and a μ_T-measurable *orientation function* $\overrightarrow{T}$ with values in $\bigwedge_M(\mathbb{R}^N)$ such that $|\overrightarrow{T}| = 1$ μ_T-almost everywhere and

$$T(\omega) = \int_U \langle\, \omega(x),\ \overrightarrow{T}(x) \,\rangle\, d\mu_T(x)\,. \tag{7.4}$$

In consequence of (7.4), the current T is said to be *representable by integration*. The measure μ_T—which we call the *total variation measure associated with T*—is characterized by the identity

$$\mu_T(W) = \sup_{\substack{|\omega| = 1,\omega \in \mathcal{D}^M(U) \\ \operatorname{supp} \omega \subseteq W}} T(\omega)\,,$$

and this last equals $\mathbf{M}_W(T)$ for any open $W \subseteq U$. We have in particular that $\mu_T(U) = \mathbf{M}(T)$.

The total variation measure μ_T will also be denoted by $\|T\|$. This alternative notation $\|T\|$ is the only one used in [Fed 69].

If E is a μ_T-measurable set and $\mu_T(\mathbb{R}^N \setminus E) = 0$, then we have $T = T \llcorner E$ and we say that T is *carried by E*. Certainly T is carried by spt T, but since spt T is by definition a closed set, T can be carried on a much smaller set than spt T.

It is worth noting that mass $\mathbf{M}$ is lower semicontinuous in the sense that if $T_j \to T$ in U in the topology of weak convergence then

$$\mathbf{M}_W(T) \le \liminf_{j\to\infty} \mathbf{M}_W(T_j) \quad \text{for all open } W \subset U\,. \tag{7.5}$$

Currents Associated to Oriented Submanifolds

A particularly important type of current representable by integration is that associated with an oriented submanifold of $\mathbb{R}^N$. Suppose that S is a C^1 oriented M-dimensional submanifold. By saying that S is oriented we mean that at each point $x \in S$ there is a set of M orthonormal tangent vectors $\xi_1(x), \xi_2(x), \ldots, \xi_M(x)$ such that

$$\overrightarrow{S}(x) = \xi_1(x) \wedge \xi_2(x) \wedge \cdots \wedge \xi_M(x)$$

defines a continuous function $\overrightarrow{S} : S \to \bigwedge_M(\mathbb{R}^N)$. We define the current $[\![S]\!] \in \mathcal{D}_M(\mathbb{R}^N)$ by setting

$$[\![S]\!](\omega) = \int_S \langle\, \omega,\ \overrightarrow{S}\,\rangle\, d\mathcal{H}^M\,.$$

As a special case of this definition, we can take S to be a Lebesgue measurable subset of $\mathbb{R}^N$ and define

$$[\![S]\!](\omega) = \int_S \langle\, \omega,\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N\,\rangle\, d\mathcal{L}^N\,, \tag{7.6}$$

for $\omega \in \mathcal{D}^N(\mathbb{R}^N)$.

In case S is an oriented submanifold with oriented boundary, the classical Stokes's theorem tells us that

$$[\![S]\!](d\omega) = [\![\partial_{\mathrm{o}} S]\!](\omega)\,, \tag{7.7}$$

where $\partial_{\mathrm{o}} S$ is the oriented boundary of S (see Definition 6.2.11 and Theorem 6.2.12). By the definition of the boundary of a current we have

$$[\![S]\!](d\omega) = (\,\partial [\![S]\!]\,)(\omega)\,. \tag{7.8}$$

Equations (7.7) and (7.8) show that the definition of the boundary of a current is consistent with the classical definition of the oriented boundary.

We also observe that

$$\mathbf{M}(\,[\![S]\!]\,) = \mathcal{H}^M(S)\,,$$

which shows that the mass generalizes the area of a submanifold.

In case $M = N - 1$, one can identify[2] $\overrightarrow{S}$ with a unit vector normal to S. Figure 7.2 uses this identification to illustrate a current associated with a 2-dimensional submanifold of $\mathbb{R}^3$.

[2] This identification is effected by the Hodge star operator, which is discussed in Section 7.5.

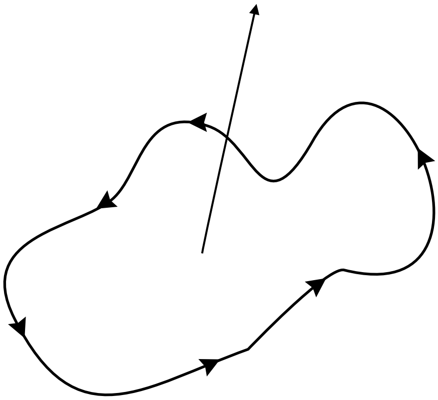

Fig. 7.2. A current associated with a 2-dimensional submanifold.

7.3 Constructions Using Currents and the Constancy Theorem

We can think of $\mathcal{L}^N$ as the 0-dimensional current that gives the value $\int_U \phi \, d\mathcal{L}^N$ when applied to $\phi \in \mathcal{D}^0(\mathbb{R}^N)$. If ξ is an M-vector field with $\mathcal{L}^N$-measurable coefficients, satisfying

$$\int_K \|\xi\| \, d\mathcal{L}^N < \infty$$

for each compact subset $K \subseteq \mathbb{R}^N$, then there is a corresponding current $\mathcal{L}^N \wedge \xi \in \mathcal{D}_M(\mathbb{R}^N)$ given by

$$(\mathcal{L}^N \wedge \xi)(\psi) = \int \langle \psi, \xi \rangle \, d\mathcal{L}^N \quad \text{for } \psi \in \mathcal{D}^M(\mathbb{R}^N).$$

Recalling the definitions in the last section, we see that for $\phi \in \mathcal{E}^k(U)$, with $k \le M$, $(\mathcal{L}^N \wedge \xi) \llcorner \phi \in \mathcal{D}_{M-k}(U)$ is given by

$$\left[(\mathcal{L}^N \wedge \xi) \llcorner \phi \right](\psi) = \int \langle \phi \wedge \psi, \xi \rangle \, d\mathcal{L}^N$$

for $\psi \in \mathcal{D}^{M-k}(\mathbb{R}^N)$. We can also write this as $(\mathcal{L}^N \wedge \xi) \llcorner \phi = \mathcal{L}^N \wedge (\xi \llcorner \phi)$, where we define the *interior product* $\xi \llcorner \phi$ by requiring that $\langle \psi, \xi \llcorner \phi \rangle = \langle \phi \wedge \psi, \xi \rangle$.

As we did for the interior product defined in the preceding section, we can examine the effect of the interior product $\xi \llcorner \phi$ on the basis vectors for $\bigwedge_M (\mathbb{R}^N)$ and $\bigwedge^{M-k} (\mathbb{R}^N)$. Suppose that

$$1 \le i_1 < \cdots < i_M \le N \ \text{ and } \ 1 \le j_1 < \cdots < j_{M-k} \le N.$$

If $\{ i_1, \ldots, i_M \} \not\supseteq \{ j_1, \ldots, j_{M-k} \}$, then

$$(\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_M}) \llcorner (dx_{j_1} \wedge \cdots \wedge dx_{j_{M-k}}) = 0.$$

On the other hand, if $\{ i_1, \dots, i_M \} \supseteq \{ j_1, \dots, j_{M-k} \}$, then we write

$$\{ \ell_1, \dots, \ell_k \} = \{ i_1, \dots, i_M \} \setminus \{ j_1, \dots, j_{M-k} \},$$

where $1 \le \ell_1 < \cdots < \ell_k \le N$. In this case, we have

$$(\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_M}) \llcorner (dx_{j_1} \wedge \cdots \wedge dx_{j_{M-k}}) = \sigma\, dx_{\ell_1} \wedge \cdots \wedge dx_{\ell_k} ,$$

where $\sigma \in \{ -1, +1 \}$ is the sign of the permutation

$$(i_1, \dots, i_M) \longmapsto (j_1, \dots, j_{M-k}, \ell_1, \dots, \ell_k) .$$

If it happens that ξ has C^1 coefficients, then (using the fact that when $\mathcal{L}^N$ is treated as a current, all its partial derivatives vanish) we have

$$D_{x_j}(\mathcal{L}^N \wedge \xi) = \mathcal{L}^N \wedge (\partial\xi/\partial x_j)$$

and

$$\partial(\mathcal{L}^N \wedge \xi) = -\sum_{j=1}^{N} [\, D_{x_j}(\mathcal{L}^N \wedge \xi)\,] \llcorner dx_j = -\mathcal{L}^N \wedge \left(\sum_{j=1}^{N} (\partial\xi/\partial x_j) \llcorner dx_j \right) .$$

In case $M = 1$, in which case ξ is an ordinary vector field, we see that

$$\sum_{j=1}^{N} (\partial\xi/\partial x_j) \llcorner dx_j = \operatorname{div} \xi . \tag{7.9}$$

Letting (7.9) define the *divergence of an M-vector field* for all $1 \le M \le N$, we have

$$\partial(\mathcal{L}^N \wedge \xi) = -\mathcal{L}^N \wedge \operatorname{div} \xi .$$

Let ξ be an M-vector field on U. We define the differential form $\mathbf{D}_M \xi$ by setting

$$\mathbf{D}_M \xi = \xi \lrcorner (dx_1 \wedge \cdots \wedge dx_N) .$$

Of course, $\mathbf{D}_M \xi$ has degree $N - M$. Also, with each differential form ϕ of degree M on U we associate the $(N - M)$-vector field

$$\mathbf{D}^M \phi = (\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N) \llcorner \phi .$$

If $\phi \in \mathcal{D}^{N-M}$ and $\psi \in \mathcal{D}^M$, then we see that

$$\begin{aligned}(\mathcal{L}^N \wedge \mathbf{D}^{N-M}\phi)(\psi) &= \int \langle\, \psi,\ \mathbf{D}^{N-M}\phi \,\rangle \, d\mathcal{L}^N \\ &= \int \langle\, dx_1 \wedge \cdots \wedge dx_N,\ \phi \wedge \psi \,\rangle \, d\mathcal{L}^N .\end{aligned}$$

The following commutative diagram helps to clarify the roles of the different spaces and their interaction with the various boundary and coboundary operators:

$$
\begin{array}{ccccc}
\mathcal{E}^{N-M}(\mathbb{R}^N) & \xrightarrow{\mathbf{D}^{N-M}} & \mathcal{E}(\mathbb{R}^N, \bigwedge_M \mathbb{R}^N) & \xrightarrow{\mathcal{L}^N \wedge} & \mathcal{D}_M(\mathbb{R}^N) \\
\Big\downarrow {\scriptstyle (-1)^{N-M} d} & & \Big\downarrow {\scriptstyle \mathrm{div}} & & \Big\downarrow {\scriptstyle -\partial} \\
\mathcal{E}^{N-M+1}(\mathbb{R}^N) & \xrightarrow{\mathbf{D}^{N-M+1}} & \mathcal{E}(\mathbb{R}^N, \bigwedge_{M-1} \mathbb{R}^N) & \xrightarrow{\mathcal{L}^N \wedge} & \mathcal{D}_{M-1}(\mathbb{R}^N)
\end{array}
$$

The special notation

$$\mathbf{E}^N = \mathcal{L}^N \wedge \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_N \in \mathcal{D}_N(\mathbb{R}^N)$$

is often used. Of course, this means that if $\phi \in \mathcal{D}^N(\mathbb{R}^N)$, then

$$\mathbf{E}^N(\phi) = \int \langle\, \phi(x),\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \,\rangle \, d\mathcal{L}^N(x)\,.$$

We see that

$$D_{x_j}\mathbf{E}^N = 0 \ \text{ for each } j = 1, \ldots, N \quad \text{and} \quad \partial \mathbf{E}^N = 0\,.$$

Comparing with (7.6), we see that for any Lebesgue measurable set $A \subseteq \mathbb{R}^N$,

$$\mathbf{E}^N \llcorner A = [\![\, A \,]\!]\,.$$

If $T \in \mathcal{D}_N(U)$ and $j \in \{1, \ldots, N\}$, then, using the formula

$$\partial T = -\sum_{\ell=1}^{N} (D_{x_\ell} T) \llcorner dx_\ell$$

and the fact that $\bigwedge^{N+1} \mathbb{R}^N = 0$, we can calculate that

$$(\partial T) \wedge \mathbf{e}_j = (-1)^N D_{x_j} T\,. \tag{7.10}$$

Thus the vanishing of the boundary of an N-dimensional current is equivalent to the vanishing of its partial derivatives. Accordingly we expect that an N-dimensional current with vanishing boundary should be given by integration. That intuition is confirmed by the next proposition.

Proposition 7.3.1 (Constancy Theorem). *If $T \in \mathcal{D}_N(U)$ with $\partial T = 0$ and if U is a connected open set, then there is a real number c such that*

$$T = c\,(\mathbf{E}^N \llcorner U) = c\,[\![\, U \,]\!]\,.$$

In order to prove the constancy theorem, we will need to introduce the notion of smoothing currents. In what follows, we will use mollifiers in a standard manner. Mollifiers were introduced in Section 5.5. Recall from Definition 5.5.1 that φ is a mollifier if

- $\varphi \in C^\infty(\mathbb{R}^N)$;
- $\varphi \geq 0$;
- $\operatorname{supp} \varphi \subseteq \mathbb{B}(0, 1)$;
- $\displaystyle\int_{\mathbb{R}^N} \varphi(x)\, d\mathcal{L}^N(x) = 1$;
- $\varphi(x) = \varphi(-x)$.

For $\sigma > 0$ we set $\varphi_\sigma(x) = \sigma^{-N}\varphi(x/\sigma)$. Also recall that in case $f \in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ and $\sigma > 0$, equation (5.31) defined

$$f_\sigma(x) = f * \varphi_\sigma(x) = \int_{\mathbb{R}^N} f(z)\, \varphi_\sigma(x - z)\, d\mathcal{L}^N(z) = \int_{\mathbb{R}^N} f(x - z)\, \varphi_\sigma(z)\, d\mathcal{L}^N(z)\,.$$

Definition 7.3.2. Given a current $T \in \mathcal{D}_M(\mathbb{R}^N)$, we define a new current $T_\sigma \in \mathcal{D}_M(\mathbb{R}^N)$ by

$$T_\sigma(\omega) = T(\varphi_\sigma * \omega)\,. \tag{7.11}$$

[Note here that we convolve φ_σ with a form by convolving with each of the coefficient functions.] The process of forming T_σ from T is called *smoothing* the current T.

The crucial facts are collected in the next lemma.

Lemma 7.3.3.

(1) *T_σ converges to T in $\mathcal{D}_M(\mathbb{R}^N)$ as $\sigma \downarrow 0$,*
(2) *$D_{x_j} T_\sigma = (D_{x_j} T)_\sigma$, for $j = 1, 2, \ldots, N$,*
(3) *for each $\sigma > 0$, T_σ corresponds to a function in $\mathcal{E}(\mathbb{R}^N, \bigwedge_M \mathbb{R}^N)$.*

Proof.
(1) This is immediate from the fact that, for $\omega \in \mathcal{D}^M(\mathbb{R}^N)$, $\varphi_\sigma * \omega$ converges to ω in the topology of $\mathcal{D}^M(\mathbb{R}^N)$.
(2) Fix $j \in \{1, \ldots, N\}$ and $\omega \in \mathcal{D}^M(\mathbb{R}^N)$. We have

$$\varphi_\sigma * (\partial\omega/\partial x_j) = \partial(\varphi_\sigma * \omega)/\partial x_j\,,$$

so we compute

$$\begin{aligned}(D_{x_j} T_\sigma)(\omega) &= -T_\sigma(\partial\omega/\partial x_j) = -T[\varphi_\sigma * (\partial\omega/\partial x_j)] \\ &= -T[\partial(\varphi_\sigma * \omega)/\partial x_j] = D_{x_j} T(\varphi_\sigma * \omega) = (D_{x_j} T)_\sigma(\omega)\,.\end{aligned}$$

(3) In order to focus on the essential ideas, we will consider just the case $M = N$. Let $\boldsymbol{t}_z : \mathbb{R}^N \to \mathbb{R}^N$ denote translation by $z \in \mathbb{R}^N$, so that

$$\boldsymbol{t}_z(x) = x + z\,.$$

We define the real-valued function F_σ by setting

$$F_\sigma(z) = T[\,(\varphi_\sigma \circ \boldsymbol{t}_{-z})\, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\,]\,. \tag{7.12}$$

Another way to write (7.12) is as

$$F_\sigma(z) = T_x[\,\varphi_\sigma(x-z)\,dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\,]\,, \tag{7.13}$$

where the subscript x on T indicates that we are considering x as the operant variable for the current, while z is treated as a parameter. It is routine to verify that F_σ is C^∞ using the fact that φ_σ is C^∞.

We claim that T_σ corresponds to the function $F_\sigma\,\mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \in \mathcal{E}_N(\mathbb{R}^N)$; that is,

$$T_\sigma(\omega) = \int_{\mathbb{R}^N} F_\sigma \cdot \langle\, \omega,\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N\,\rangle\, d\mathcal{L}^N \tag{7.14}$$

holds, for each $\omega \in \mathcal{D}^M(\mathbb{R}^N)$.

To verify the claim, fix $\omega \in \mathcal{D}^M(\mathbb{R}^N)$ and write

$$\omega = g\, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\,,$$

where g is scalar-valued and C^∞. By definition, the left-hand side of (7.14) equals

$$T_x\left[\left(\int_{\mathbb{R}^N} g(z)\,\varphi_\sigma(x-z)\,d\mathcal{L}^N(z)\right) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\right].$$

We can approximate

$$\int_{\mathbb{R}^N} g(z)\,\varphi_\sigma(x-z)\,d\mathcal{L}^N(z)$$

(in the topology of $\mathcal{D}(\mathbb{R}^N, \mathbb{R})$) by finite sums

$$\sum_{k=1}^{p} g(z_k)\,\varphi_\sigma(x-z_k)\,\mathcal{L}^N(A_k)\,,$$

where $z_k \in A_k$ and where the A_k are Borel subsets of the support of g. Thus

$$T_x\left[\sum_{k=1}^{p} g(z_k)\,\varphi_\sigma(x-z_k)\,\mathcal{L}^N(A_k)\,dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\right]$$

will approximate $T_\sigma(\omega)$.

By the linearity of T and using (7.13), we have

$$\begin{aligned}
&T_x\left[\sum_{k=1}^{p} g(z_k)\,\varphi_\sigma(x-z_k)\,\mathcal{L}^N(A_k)\,dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N\right]\\
&\quad= \sum_{k=1}^{p} T_x\,[\,\varphi_\sigma(x-z_k)\,dx_1 \wedge dx_2 \wedge \cdots \wedge dx_N]\,g(z_k)\,\mathcal{L}^N(A_k)\\
&\quad= \sum_{k=1}^{p} F_\sigma(z_k) \cdot \langle\, \omega(z_k),\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N\,\rangle\, \mathcal{L}^N(A_k)\,.
\end{aligned}$$

But as the diameters of the A_k approach 0,

$$\sum_{k=1}^{p} F_\sigma(z_k) \cdot \langle\, \omega(z_k),\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \,\rangle\, \mathcal{L}^N(A_k)$$

approaches

$$\int_{\mathbb{R}^N} F_\sigma \cdot \langle\, \omega,\ \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_N \,\rangle\, d\mathcal{L}^N\,,$$

verifying the claim. □

Smoothing is defined in a general open set $U \subseteq \mathbb{R}^N$ by introducing functions $w_j \in \mathcal{D}(U, \mathbb{R})$ such that the sets $K_j = \{x : w_j(x) = 1\}$ are increasing and exhaust U. For $T \in \mathcal{D}_M(U)$, one then considers $(T \llcorner w_j)_\sigma$—as one may, since $T \llcorner w_j \in \mathcal{D}_M(\mathbb{R}^N)$.

Proposition 7.3.4. *If $T \in \mathcal{D}_M(U)$, where $U \subseteq \mathbb{R}^M$, and if $\mathbf{M}(T) < \infty$ and $\mathbf{M}(\partial T) < \infty$ hold, then $T = [\![U]\!] \llcorner F$ with $F \in BV(U)$.*

Proof. Referring to Lemma 7.3.3(3), we observe that $T_\sigma = [\![U]\!] \llcorner F_\sigma$ and that the L^1-norm of F_σ equals $\mathbf{M}(T_\sigma)$, which is bounded by $\mathbf{M}(T)$. Also, $\int |DF_\sigma|$ equals $\mathbf{M}(\partial T_\sigma)$, which is bounded by $\mathbf{M}(\partial T)$. By the compactness theorem for functions of bounded variation (see [KPk 99, Corollary 3.6.14]), we can select a sequence $\sigma_i \downarrow 0$ such that F_{σ_i} converges to a BV-function F and conclude from Lemma 7.3.3(1) that $T = [\![U]\!] \llcorner F$. □

Now we return to the constancy theorem.

Proof of the Constancy Theorem. For convenience of exposition we suppose that $U = \mathbb{R}^N$. By (7.10), the hypothesis $\partial T = 0$ tells us that all the partial derivatives of T vanish. Then, for any $\sigma > 0$, the partial derivatives of T_σ must vanish. We know that T_σ corresponds to a function in $\mathcal{E}(\mathbb{R}^N, \bigwedge_N \mathbb{R}^N)$ and that function must be constant since its partial derivatives vanish. Letting $\sigma \downarrow 0$, we obtain the result. □

We end this section with the following variant of the constancy theorem.

Proposition 7.3.5. *If $T \in \mathcal{D}_M(\mathbb{R}^N)$ with $\partial T = 0$ and* spt $T \subseteq V$*, where V is an M-dimensional plane, then there is a real number c such that*

$$T = c\, [\![V]\!]\,,$$

that is, $T = c\, \left(\mathcal{H}^M \llcorner V\right) \wedge v_1 \wedge v_2 \wedge \cdots \wedge v_M$, where $v_1, v_2, \ldots, v_M$ is an orthonormal family of vectors parallel to V.

Proof. Without loss of generality, we may suppose that

$$V = \{\, (x_1, x_2, \ldots, x_N) : x_{M+1} = x_{M+2} = \cdots = x_N = 0 \,\}\,.$$

Fix $\sigma : \mathbb{R} \to \mathbb{R}$, a compactly supported, C^∞ function satisfying $\sigma(t) = t$, for $|t| < 1$.

Consider $1 \leq i_1 < i_2 < \cdots < i_M \leq N$ and suppose that $M < i_M$. Let ϕ be an arbitrary compactly supported, real-valued C^∞ function on $\mathbb{R}^N$. Setting

$$\omega = \sigma\left(x_{i_M}\right) \cdot \phi(x)\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_{M-1}},$$

we see that, on V, $d\omega = \phi(x)\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_M}$, so that

$$0 = (\partial T)(\omega) = T(d\omega) = T\Big(\phi(x)\, dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_M}\Big)$$

holds. Thus we have

$$T \llcorner dx_{i_1} \wedge \cdots \wedge dx_{i_M} = 0\,.$$

Using the preceding paragraph, we conclude that

$$\begin{aligned} T &= \sum_{1 \leq i_1 < \cdots < i_M \leq N} \Big[T \llcorner dx_{i_1} \wedge \cdots \wedge dx_{i_M}\Big] \wedge \mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_M} \\ &= \Big[T \llcorner dx_1 \wedge \cdots \wedge dx_M\Big] \wedge \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_M\,. \end{aligned}$$

Thus we can identify T with an element of $\mathcal{D}_N(\mathbb{R}^N)$ and apply the constancy theorem. □

7.4 Further Constructions with Currents

7.4.1 Products of Currents

Next we need the notion of a Cartesian product of currents.

Definition 7.4.1. Suppose $U_1 \subseteq \mathbb{R}^{N_1}$, $T_1 \in \mathcal{D}_{M_1}(U_1)$ and $U_2 \subseteq \mathbb{R}^{N_2}$, $T_2 \in \mathcal{D}_{M_2}(U_2)$. We define $T_1 \times T_2 \in \mathcal{D}_{M_1+M_2}(U_1 \times U_2)$, the *Cartesian product of* T_1 *and* T_2 as follows:

(1) We will denote the basis covectors in $\mathbb{R}^{N_1}$ by dx_α and the basis covectors in $\mathbb{R}^{N_2}$ by dy_β.

(2) If $1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_{M_1} \leq N_1$, $1 \leq \beta_1 < \beta_2 < \cdots < \beta_{M_2} \leq N_2$, and $g \in \mathcal{D}(U_1 \times U_2, \mathbb{R})$, then set

$$\begin{aligned} &[T_1 \times T_2](g\, dx_{\alpha_1} \wedge \cdots \wedge dx_{\alpha_{M_1}} \wedge dy_{\beta_1} \wedge \cdots \wedge dy_{\beta_{M_2}}) \\ &\quad = T_1\left(T_2[\, g(x,y)\, dy_{\beta_1} \wedge \cdots \wedge dy_{\beta_{M_2}}\,]\, dx_{\alpha_1} \wedge \cdots \wedge dx_{\alpha_{M_1}} \right). \end{aligned}$$

(3) If $\omega_1 \in \mathcal{D}^{M_1'}(U_1)$, $\omega_2 \in \mathcal{D}^{M_2'}(U_2)$ with $M_1' + M_2' = M_1 + M_2$ but $M_1' \neq M_1$ and $M_2' \neq M_2$, then $[T_1 \times T_2](\omega_1 \wedge \omega_2) = 0$.

(4) Extend $T_1 \times T_2$ to $\mathcal{D}^{M_1+M_2}(U_1 \times U_2)$ by linearity.

Now it is immediate that

$$\partial(T_1 \times T_2) = (\partial T_1) \times T_2 + (-1)^{M_1} T_1 \times \partial T_2 \,. \tag{7.15}$$

In case either $M_1 = 0$ or $M_2 = 0$, then the last formula is still valid, provided the corresponding terms are interpreted to be zero.

In the special case that $T \in \mathcal{D}_M(U)$ with $U \subseteq \mathbb{R}^N$ and $[\![(0,1)]\!]$ is the 1-current in $\mathbb{R}^1$ given by integration over the oriented unit interval, then (7.15) becomes

$$\begin{aligned}\partial([\![(0,1)]\!] \times T) &= (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_0) \times T - [\![(0,1)]\!] \times \partial T \\ &= \boldsymbol{\delta}_1 \times T - \boldsymbol{\delta}_0 \times T - [\![(0,1)]\!] \times \partial T \,.\end{aligned}$$

Of course, $\boldsymbol{\delta}_p$ denotes the 0-current that is given by a point mass at p.

7.4.2 The Pushforward

Now we shall define the notion of the pushforward of a current. Some of the most important and profound properties of currents will be formulated in terms of the preservation of certain structures under the pushforward. The setup is this. We are given open sets $U \subseteq \mathbb{R}^{N_1}$ and $V \subseteq \mathbb{R}^{N_2}$ and a smooth mapping $f : U \to V$. If $\omega \in \mathcal{D}^M(V)$ then let $f^\# \omega$ be the standard pullback of the form ω (see Definition 6.2.7). Now the current T is given on U, and we must suppose that $f\big|_{\operatorname{spt} T}$ is proper: this means that the inverse image under f of any compact set, intersected with spt T, is compact. We define the *pushforward* $f_\# T$ under f (see Figure 7.3) of the current T by

$$f_\# T(\omega) = T(\zeta \cdot f^\# \omega) \qquad \forall \omega \in \mathcal{D}^M(V) \,, \tag{7.16}$$

where ζ is any compactly supported $C^\infty(U)$ function that equals 1 in a neighborhood of spt $T \cap$ supp $f^\# \omega$. The definition of $f_\# T$ given in (7.16) is independent of ζ. Notice that

$$\partial f_\# T = f_\# \partial T \tag{7.17}$$

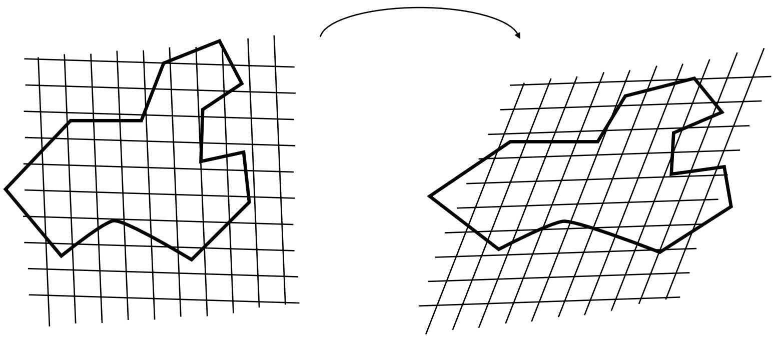

Fig. 7.3. The pushforward of a current.

holds for f, T as above. In fact, (7.17) holds because one can interchange the exterior differentiation and pullback operations on forms (see Theorem 6.2.9).

If $\mathbf{M}_W(T) < \infty$ for every $W \subseteq U$ then f is representable by integration and $f_\# T$ is given explicitly by

$$\begin{aligned} f_\# T(\omega) &= \int \langle \overrightarrow{T}, f^\# \omega \rangle \, d\mu_T \\ &= \int \langle \langle \textstyle\bigwedge_M Df, \overrightarrow{T}(x) \rangle, \omega(f(x)) \rangle \, d\mu_T(x) . \end{aligned}$$

This formula gives a way to make sense of $f_\# T$ even when f is only continuously differentiable and proper.

The next result is about vanishing of currents on sets that project to measure 0 in all coordinate directions. For notation, if $\alpha = (i_1, \dots, i_M)$ is a multi-index with $1 \le i_1 < i_2 < \cdots < i_M \le N$ then we let $\mathbf{p}_\alpha$ denote the orthogonal projection of $\mathbb{R}^N$ onto $\mathbb{R}^M$ given by

$$(x^1, \dots, x^P) \longmapsto (x^{i_1}, \dots, x^{i_M}) .$$

Lemma 7.4.2. *Let $U \subseteq \mathbb{R}^N$ be open as usual. Let $E \subseteq U$ be closed. Assume that $\mathcal{L}^M(\mathbf{p}_\alpha(E)) = 0$ for each multi-index $\alpha = (i_1, \dots, i_M)$, $1 \le i_1 < i_2 < \cdots < I_M \le N$. Then $T \llcorner E = 0$ whenever $T \in \mathcal{D}_M(U)$ with $\mathbf{M}_W(T)$ and $\mathbf{M}_W(\partial T)$ finite for every $W \subset\subset U$.*

Proof. Let $\omega \in \mathcal{D}^M(U)$. We write

$$\omega = \sum_{\alpha \in \Lambda(N,M)} \omega_\alpha dx^\alpha$$

with $\omega_\alpha \in C^\infty(U)$ and compactly supported. Thus

$$\begin{aligned} T(\omega) &= \sum_\alpha T(\omega_\alpha dx^\alpha) \\ &= \sum_\alpha (T \llcorner \omega_\alpha) \, dx^\alpha \\ &= \sum_\alpha (T \llcorner \omega_\alpha) \mathbf{p}_\alpha^\# \, dy . \end{aligned}$$

Here $dy \equiv dy^1 \wedge \cdots \wedge dy^M$ in the standard coordinates on $\mathbb{R}^M$.

So we have

$$T(\omega) = \sum_\alpha \mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha)(dy) . \tag{7.18}$$

This last makes sense just because $\operatorname{spt} T \llcorner \omega_\alpha \subseteq \operatorname{supp} \omega_\alpha$, which is a compact subset of U.

On the other hand, we know for any $\tau \in \mathcal{D}^{N-1}(U)$ that

$$
\begin{aligned}
\partial(T \llcorner \omega_\alpha)(\tau) &= (T \llcorner \omega_\alpha)(d\tau) \\
&= T(\omega_\alpha d\tau) \\
&= T(d(\omega_\alpha \tau)) - T(d\omega_\alpha \wedge \tau) \\
&= \partial T(\omega_\alpha \tau) - T(d\omega_\alpha \wedge \tau)\,,
\end{aligned}
$$

and so

$$
\mathbf{M}_W(\partial(T \llcorner \omega_\omega)) \le \mathbf{M}_W(\partial T) \cdot \sup\, |\omega_\alpha| + \mathbf{M}_W(T) \cdot \sup\, |d\omega_\alpha|\,.
$$

From this we conclude that

$$
\mathbf{M}(\partial \mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha)) = \mathbf{M}(\mathbf{p}_{\alpha\#}\partial(T \llcorner \omega_\alpha)) \le \mathbf{M}(\partial(T \llcorner \omega_\alpha)) < \infty\,.
$$

Now we apply Proposition 7.3.4 to see that there is a $\theta_\alpha \in BV(\mathbf{p}_\alpha(U))$ such that

$$
\mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha) = [\![\, \mathbf{p}_\alpha(U)]\!] \llcorner \theta_\alpha\,.
$$

It follows that $\mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha) \llcorner \mathbf{p}_\alpha(E) = 0$ because $\mathcal{L}^M(\mathbf{p}_\alpha(E)) = 0$. Assuming without loss of generality that E is closed, we now see that

$$
\begin{aligned}
\mathbf{M}(\mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha)) &\le \mathbf{M}(\mathbf{p}_{\alpha\#}(T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^M \setminus \mathbf{p}_\alpha(E))) \\
&= \mathbf{M}(\,\mathbf{p}_{\alpha\#}[\,(T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^N \setminus \mathbf{p}_\alpha^{-1}\mathbf{p}_\alpha(E))\,]\,) \\
&\le \mathbf{M}((T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^N \setminus \mathbf{p}_\alpha^{-1}\mathbf{p}_\alpha E)) \qquad (7.19) \\
&\le \mathbf{M}_W(T \llcorner (\mathbb{R}^N \setminus \mathbf{p}_\alpha^{-1}\mathbf{p}_\alpha E)) \cdot |\omega_\alpha| \\
&\le \mathbf{M}_W(T \llcorner (\mathbb{R}^N \setminus E)) \cdot |\omega_\alpha| \qquad (7.20)
\end{aligned}
$$

for any open set W such that $\operatorname{supp} \omega \subseteq W \subseteq U$.

Now we combine (7.18) and (7.20) to obtain

$$
\mathbf{M}_W(T) \le c\,\mathbf{M}_W(T \llcorner (\mathbb{R}^N \setminus E))\,.
$$

In particular, we see that

$$
\mathbf{M}_W(T \llcorner E) \le c\,\mathbf{M}_W(T \llcorner (\mathbb{R}^N \setminus E))\,. \qquad (7.21)
$$

If K is any compact subset of E, then we can choose sets $\{W_q\}$ such that

- $W_q \subset\subset U$;
- $W_{q+1} \subseteq W_q$;
- $\bigcap_{q=1}^\infty W_q = K$.

By (7.21), with $W = W_q$, we conclude that $\mathbf{M}(T \llcorner K) = 0$. Since K was arbitrary, we see that $\mathbf{M}(T \llcorner E) = 0$. $\square$

7.4.3 The Homotopy Formula

Next we have the homotopy formula for currents. Let $f, g : U \to V$ be smooth mappings, with $U \subseteq \mathbb{R}^{N_1}$ and $V \subseteq \mathbb{R}^{N_2}$. Let h be a smooth homotopy of f to g; that is, $h : [0,1] \times U \to V$, $h(0,x) = f(x)$, and $h(1,x) = g(x)$. If $T \in \mathcal{D}_M(U)$ and if the restriction of h to $[0,1] \times \operatorname{spt} T$ is proper, then $h_\#(\llbracket\, (0,1)\,\rrbracket \times T)$ is well-defined and

$$\begin{aligned} \partial h_\#(\llbracket\, (0,1)\,\rrbracket \times T) &= h_\# \partial(\llbracket\, (0,1)\,\rrbracket \times T) \\ &= h_\#(\boldsymbol{\delta}_1 \times T - \boldsymbol{\delta}_0 \times T - \llbracket\, (0,1)\,\rrbracket \times \partial T) \\ &= g_\# T - f_\# T - h_\#(\llbracket\, (0,1)\,\rrbracket \times \partial T)\,. \end{aligned}$$

The *homotopy formula* is then a simple rearrangement of this last equality:

$$g_\# T - f_\# T = \partial h_\#(\llbracket\, (0,1)\,\rrbracket \times T) + h_\#(\llbracket\, (0,1)\,\rrbracket \times \partial T)\,. \tag{7.22}$$

An important instance of the homotopy formula occurs when

$$h(t,x) = tg(x) + (1-t)f(x) = f(x) + t(g(x) - f(x))\,;$$

we call this an *affine homotopy* of f to g. Then we can obtain that

$$\mathbf{M}[\, h_\#(\llbracket\, (0,1)\,\rrbracket \times T)\,] \le \sup_{\operatorname{spt} T} |f - g| \cdot \sup_{x \in \operatorname{spt} T} (\|Df(x)\| + \|Dg(x)\|)^M \mathbf{M}(T)\,. \tag{7.23}$$

In fact, this inequality follows immediately once we notice that

$$\begin{aligned} &h_\#(\llbracket\, (0,1)\,\rrbracket \times T)(\omega) \\ &= \int_0^1 \int \Big\langle \langle \textstyle\bigwedge_{M+1} Dh(t,x),\ e_1 \wedge \overrightarrow{T}(x) \rangle,\ \omega(h(t,x)) \Big\rangle d\mu_T(x)\, d\mathcal{L}^1(t) \\ &= \int_0^1 \int \Big\langle (g(x) - f(x)) \wedge \\ &\quad \langle t \textstyle\bigwedge_M Df(x) + (1-t) \bigwedge_M Dg(x),\ \overrightarrow{T}(x) \rangle,\ \omega(h(t,x)) \Big\rangle d\mu_T(x)\, d\mathcal{L}^1(t)\,. \end{aligned} \tag{7.24}$$

Figure 7.4 illustrates the homotopy formula. In this figure, T is the 1-dimensional current associated with the oriented line segment on the left, f is the identity, and g maps the line segment on the left to the polygonal path on its right. The six-sided polygonal region then corresponds to $h_\#(\llbracket\, (0,1)\,\rrbracket \times T)$ with h the affine homotopy of f to g.

7.4.4 Applications of the Homotopy Formula

The next lemma shows us how the homotopy formula can be used to define $f_\# T$ in case f is only Lipschitz—provided that the restriction of f to the support of T is proper and both $\mathbf{M}_W(T)$, $\mathbf{M}_W(\partial T)$ are finite for all $W \subset\subset U$. We will use smoothing of currents as described in Definition 7.3.2.

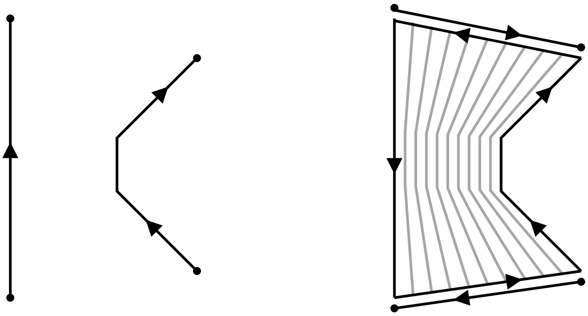

Fig. 7.4. The homotopy formula.

Lemma 7.4.3. *Let T be a current, $T \in \mathcal{D}_M(U)$, and suppose that $\mathbf{M}_W(T)$, $\mathbf{M}_W(\partial T)$ are finite for each $W \subset\subset U$. Let $f : U \to V$ be a Lipschitz mapping, and assume that the restriction of f to the support of T is proper. Then we may define*

$$f_\#(T) \equiv \lim_{\sigma \to 0^+} f_{\sigma\#} T(\omega)$$

because the limit on the right-hand side exists for each $\omega \in D^M(V)$. We also may conclude that

$$\operatorname{spt} f_\# T \subseteq f(\operatorname{spt} T) \qquad \textit{and} \qquad \mathbf{M}_W(f_\# T) \le \left(ess \sup_{f^{-1}(W)} |Df| \right)^M \mathbf{M}_{f^{-1}(W)}(T)$$

for all $W \subset\subset V$.

Proof. If $\sigma, \tau > 0$ are small then the homotopy formula gives us that

$$f_{\sigma\#} T(\omega) - f_{\tau\#} T(\omega) = h_\#([\![[0, 1)]\!] \times T)(d\omega) + h_\#([\![(0, 1)]\!] \times \partial T)(\omega),$$

where h is the usual affine homotopy of f_τ to f_σ. Now (7.23) tells us, for small σ, τ, that

$$|f_{\sigma\#} T(\omega) - f_{\tau\#} T(\omega)| \le c \sup_{f^{-1}(K) \cap \operatorname{spt} T} |f_\sigma - f_\tau| \cdot \|f\|_{\mathrm{Lip}}.$$

Here K is a compact subset of V with $\operatorname{supp} \omega \subseteq \operatorname{interior}(K)$. Since $f_\sigma \to f$ uniformly on compact subsets of U, the result clearly follows. □

Now we need the notion of a cone over a current $T \in \mathcal{D}_M(U)$. Any definition that we give should have the property that in the special case that $T = [\![S]\!]$, where S is a submanifold of the sphere $\mathcal{S}^{N-1} \subseteq \mathbb{R}^N$, the cone over T is $[\![C_S]\!]$, where

$$C_S = \{\lambda x : x \in S, 0 \le \lambda \le 1\}.$$

We define the cone using ideas and terminology that we have introduced thus far. We let

- $T \in \mathcal{D}_M$;
- U be star-shaped with respect to the point 0 (i.e., $t \in U$, for each $x \in U$ and each $0 \le t \le 1$);
- spt T be compact;
- $h : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ be defined by $h(t, x) = tx$.

Then the *cone over* T, denoted by $\boldsymbol{\delta}_0 ⨳ T$, is given by

$$\boldsymbol{\delta}_0 ⨳ T = h_\#([\![(0,1)]\!] \times T) . \tag{7.25}$$

It follows that $\boldsymbol{\delta}_0 ⨳ T \in \mathcal{D}_{M+1}(U)$ and, by the homotopy formula,

$$\partial(\boldsymbol{\delta}_0 ⨳ T) = T - \boldsymbol{\delta}_0 ⨳ \partial T .$$

Also, if spt $T \subseteq \{ x : |x| = r \}$ holds, then we can estimate

$$\mathbf{M}(\boldsymbol{\delta}_0 ⨳ T) \le \frac{r}{M+1} \mathbf{M}(T) .$$

This last estimate follows from observing that

$$\begin{aligned} &h_\#([\![(0,1)]\!] \times T)(\omega) \\ &= \int_0^1 \int \Big\langle \langle \textstyle\bigwedge_{M+1} Dh(t,x),\ e_1 \wedge \overrightarrow{T}(x) \rangle,\ \omega(h(t,x)) \Big\rangle d\mu_T(x)\, d\mathcal{L}^1(t) \\ &= \int_0^1 \int t^M \langle x \wedge \overrightarrow{T}(x),\ \omega(tx) \rangle \, d\mu_T(x)\, d\mathcal{L}^1(t) . \end{aligned}$$

By making the obvious modifications, we can define the *cone over* T *with vertex* p, which we denote by $\boldsymbol{\delta}_p ⨳ T$. In this case, we have

$$\partial(\boldsymbol{\delta}_p ⨳ T) = T - \boldsymbol{\delta}_p ⨳ \partial T \tag{7.26}$$

and, if spt $T \subseteq \{ x : |x - p| = r \}$ holds,

$$\mathbf{M}(\boldsymbol{\delta}_p ⨳ T) \le \frac{r}{M+1} \mathbf{M}(T) . \tag{7.27}$$

7.5 Rectifiable Currents with Integer Multiplicity

Now we consider integer-multiplicity currents $T \in \mathcal{D}_N(U)$, which are similar to, but more general than, the currents associated with smooth surfaces. These new currents will be based on the notion of a countably M-rectifiable set that was introduced in Section 5.4.

Definition 7.5.1. Let M be an integer with $1 \le M \le N$. Let $T \in \mathcal{D}_M(U)$ for $U \subseteq \mathbb{R}^N$ an open set. We say that T is an *integer-multiplicity rectifiable* M*-current* (or, more succinctly, an *integer-multiplicity current*) if there are S, θ, and ξ such that

(1) S is an $\mathcal{H}^M$-measurable, countably M-rectifiable subset of U with $\mathcal{H}^M(S \bigcap K) < \infty$ for each compact $K \subseteq U$;
(2) θ is a locally $\mathcal{H}^M$-integrable, nonnegative, integer-valued function;
(3) $\xi : S \to \bigwedge_M(\mathbb{R}^N)$ is an $\mathcal{H}^M$-measurable function such that, for $\mathcal{H}^M$-almost every point $x \in S$, $\xi(x)$ is a simple unit M-vector in $\mathbf{T}_x S$;
(4) the current T is given by

$$T(\omega) = \int_S \langle\, \omega(x),\ \xi(x)\,\rangle\, \theta(x)\, d\mathcal{H}^M(x)$$

for $\omega \in \mathcal{D}^M(U)$.

For (3), recall that $\xi(x)$ is simple if $\xi(x) = \tau_1 \wedge \cdots \wedge \tau_M$; in this situation it is usually most convenient to choose $\{\tau_j\}$ to be an orthonormal basis for $\mathbf{T}_x S$.

In the preceding definition, we call θ the *multiplicity* of T and ξ the *orientation* of T. It will be convenient for us to write $T = \boldsymbol{\tau}(S, \theta, \xi)$. In terms of the notation for currents representable by integration introduced in (7.4) we have

$$\overrightarrow{S} = \xi, \quad \mu_S = \|S\| = \mathcal{H}^M \llcorner (\theta\, \chi_S)\,.$$

Figure 7.5 illustrates a current that fails to be integer-multiplicity rectifiable because the orientation does not lie in the tangent space.

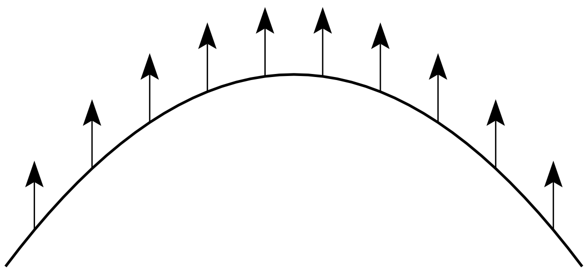

Fig. 7.5. A current that is not integer-multiplicity rectifiable.

Let $T \in \mathcal{D}_0(U)$ for $U \subseteq \mathbb{R}^N$ an open set. We say that T is an *integer-multiplicity rectifiable* 0*-current* if there are $S \subseteq U$ and $\theta : S \to \mathbb{Z}$ such that

$$S \bigcap K \ \text{ is finite if } \ K \subseteq U \ \text{ is compact,}$$

$$T(\omega) = \sum_{x \in S \cap \operatorname{supp} \omega} \theta(x)\, \omega(x) \ \text{ for } \ \omega \in \mathcal{D}^0(U). \tag{7.28}$$

In this case, the multiplicity function of T is the absolute value of θ and the orientation function of T is the sign of θ, so we may write

$$T = \boldsymbol{\tau}(S, |\theta|, \operatorname{sgn}(\theta))\,.$$

Some Properties of Integer-Multiplicity Currents

(1) If $T_1, T_2 \in \mathcal{D}_M(U)$ are integer-multiplicity currents, then so is $p_1 T_1 + p_2 T_2$ for any $p_1, p_2 \in \mathbb{Z}$.

(2) If $T_1 = \boldsymbol{\tau}(V_1, \theta_1, \xi_1) \in \mathcal{D}_M(U)$ and $T_2 = \boldsymbol{\tau}(V_2, \theta_2, \xi_2) \in \mathcal{D}_N(V)$ then $T_1 \times T_2 \in \mathcal{D}_{M+N}(U \times V)$ is also integer-multiplicity and

$$T_1 \times T_2 = \boldsymbol{\tau}(V_1 \times V_2,\ \theta_1\theta_2,\ \xi_1 \wedge \xi_2)\,.$$

(3) If $F : U \to V$ is Lipschitz, $S \subseteq U$, and $T = \boldsymbol{\tau}(S, \theta, \xi) \in \mathcal{D}_M(U)$, and if $f\big|_{\operatorname{spt} T}$ is proper, then $F_\# T \in \mathcal{D}_M(V)$ is integer-multiplicity and

$$F_\# T = \boldsymbol{\tau}(\,F(S),\ \phi,\ \eta)\,,$$

where $\phi \in \bigwedge_M \mathbb{R}^N$ and $\eta \in \mathbb{Z}^+$ are characterized, $\mathcal{H}^M$-almost everywhere in $F(S)$, by

$$\sum_{x \in F^{-1}(y) \cap S_+} \theta(x) \cdot \frac{\langle \bigwedge_M D_S F(x), \xi(x)\rangle}{|\langle \bigwedge_M D_S F(x), \xi(x)\rangle|} = \phi(y)\,\eta(y)\,. \tag{7.29}$$

Here S_+ is the set of $x \in S$ for which $\mathbf{T}_x S$ and $D_S F(x)$ both exist and $D_S F(x)$ is of rank M on $\mathbf{T}_x S$.

Statements (1) and (2) are immediate. To see statement (3) we reason as follows: By definition,

$$F_\# T(\omega) = \int_V \langle\, \omega(f(x)),\ \langle \textstyle\bigwedge_M D_S F(x), \xi(x)\rangle\,\rangle\, \theta(x)\, d\mathcal{H}^M(x)\,.$$

Corollary 5.1.13 of the area formula allows us to rewrite the last equation as

$$F_\# T(\omega) = \int_{F(S)} \left\langle \omega(y),\ \sum_{x \in F^{-1}(y) \cap S_+} \theta(x) \cdot \frac{\langle \bigwedge_M D_S F, \xi(x)\rangle}{|\langle \bigwedge_M D_S F, \xi(x)\rangle|} \right\rangle d\mathcal{H}^M(y)\,. \tag{7.30}$$

For $\mathcal{H}^M$-almost every y the approximate tangent space $\mathbf{T}_y(F(S))$ exists and $\mathbf{T}_x S$ and $D_S F$ exist for all $x \in F^{-1}(y) \cap S_+$. Hence

$$\frac{\langle \bigwedge_M D_S F, \xi(x)\rangle}{|\langle \bigwedge_M D_S F, \xi(x)\rangle|} = \pm\tau_1 \wedge \cdots \wedge \tau_M\,, \tag{7.31}$$

where $\tau_1, \ldots, \tau_M$ is an orthonormal basis for $\mathbf{T}_y(F(S))$. Thus we obtain (7.29).

Considering a y such that the approximate tangent space $\mathbf{T}_y(F(S))$ exists and $\mathbf{T}_x S$ and $D_S F$ exist for all $x \in F^{-1}(y) \cap S_+$ and replacing τ_1 by $-\tau_1$ if necessary, we may suppose that $\tau_1 \wedge \cdots \wedge \tau_M = \eta(y)$. Then we have

$$\phi(y) = \sum_{A_1} \theta(x) - \sum_{A_2} \theta(x)\,,$$

where A_1 is the set of $x \in F^{-1}(y) \cap S_+$ for which

$$\eta = \frac{\langle \bigwedge_M D_S F(x), \xi(x) \rangle}{|\langle \bigwedge_M D_S F(x), \xi(x) \rangle|}$$

and A_2 is the set of $x \in F^{-1}(y) \cap S_+$ for which

$$-\eta = \frac{\langle \bigwedge_M D_S F(x), \xi(x) \rangle}{|\langle \bigwedge_M D_S F(x), \xi(x) \rangle|} .$$

Thus, for $\mathcal{H}^M$-almost every $y \in F(W)$, we have

$$\eta(y) = \sum_{x \in F^{-1}(y) \cap W_+} \theta(x) - 2 \sum_{A_2} \theta(x) \leq \sum_{x \in F^{-1}(y) \cap W_+} \theta(x) .$$

We also note that, for $\mathcal{H}^M$-almost every $y \in F(W)$, $\eta(y)$ is congruent modulo 2 to $\sum_{x \in F^{-1}(y) \cap W_+} \theta(x)$.

One of the main things that we do in this subject is to extract "convergent" subsequences from collections of currents. This is, for instance, how we prove an existence theorem for the solution of the Plateau problem.[3] The next compactness theorem is an instance of this point of view.

Theorem 7.5.2 (Compactness for Integer-Multiplicity Currents). *Let $\{T_j\} \subseteq \mathcal{D}_M(U)$ be a sequence of integer-multiplicity currents such that*

$$\sup_{j \geq 1} \left[\mathbf{M}_W(T_j) + \mathbf{M}_W(\partial T_j) \right] < \infty \quad \textit{for all } W \subset\subset U .$$

Then there is an integer-multiplicity current $T \in \mathcal{D}_M(U)$ and a subsequence $\{T_{j'}\}$ such that $T_{j'} \to T$ weakly in U.

The compactness theorem was first proved by Federer and Fleming in [FF 60]. Their proof had the drawback of relying on the structure theorem for sets of finite Hausdorff measure (for the structure theorem, see [Whe 98] and the references therein). An alternative proof was developed by Bruce Solomon (see [Som 84]). Solomon's proof used facts about multivalued functions, which led Brian White to give a third proof that avoided both the structure theory and multivalued functions (see [Whe 89]). Later in this book we will give a proof of the compactness theorem using metric-space-valued functions of bounded variation in a manner similar to that in [LY 02]. It should be noted that the work of Ambrosio and Kirchheim [AK 00] puts this metric space approach into a more general and natural context. Yet another extension of this theory appears in [Whe 99].

Remark 7.5.3. It is important to realize that the existence of the subsequence $\{T_{j'}\}$ and the limit current T in Theorem 7.5.2 is an immediate consequence of the Banach–Alaoglu theorem.[4] What is nontrivial is the fact that T is an *integer-multiplicity*

[3] Joseph Antoine Ferdinand Plateau (1801–1883).

[4] Stefan Banach (1892–1945), Leonidas Alaoglu (1914–1981).

current. In the codimension-1 case, that is, when the ambient space has dimension $N = M+1$, Theorem 7.5.2 can be proved using Proposition 7.3.4 and the compactness theorem for functions of bounded variation. In case $M = 0$, because of (7.28), Theorem 7.5.2 is a consequence of the Bolzano–Weierstrass theorem.[5]

To end this section we will prove a decomposition theorem for integer-multiplicity currents of codimension 1. The statement of this theorem invokes the notion of a set of locally finite perimeter. We recall the relevant definitions here (see [KPk 99, Section 3.7]; the original definition is due to De Giorgi):

Definition 7.5.4.

(1) If A is a Borel set and $U \subseteq \mathbb{R}^N$ is open, then the *perimeter of A in U* is denoted by $P(A, U)$ and is defined by

$$\begin{aligned} P(A,U) &= \int_U |D\chi_A| \\ &= \sup\left\{ \int_A \operatorname{div}(g)\, d\mathcal{L}^N : g \in C^1(U;\mathbb{R}^N),\ \operatorname{supp} g \subset\subset U,\ |g| \le 1 \right\}. \end{aligned}$$

(2) We say that A is of *locally finite perimeter* if

$$P(A, U) < \infty$$

holds for every bounded open set U. Sets of locally finite perimeter are also called *Caccioppoli sets*.[6]

(3) If A is of locally finite perimeter, then there is a positive Radon measure μ and a μ-measurable $\mathbb{R}^N$-valued function σ, with $|\sigma(x)| = 1$ for μ-almost every x, such that the distribution derivative of χ_A is given by $D\chi_A = \sigma\mu$. It is customary to use the notation $|D\chi_A|$ for the Radon measure μ and to write $\mathbf{n}_A = -\sigma$, so that

$$D\chi_A = -\mathbf{n}_A\, |D\chi_A|$$

and

$$P(A,U) = \int_U |D\chi_A| .$$

We have defined $\mathbf{n}_A$ to be the *negative* of σ so that $\mathbf{n}_A$ will be the outward unit normal to A.

(4) In case A has locally finite perimeter in U, the *reduced boundary* of A, denoted by $\partial^* A$, is the set of $x \in U$ such that

(a) $|D\chi_A|\big(\mathbb{B}(x,r)\big) > 0$ holds for $r > 0$,

---

[5] Bernard Placidus Johann Nepomuk Bolzano (1781–1848), Karl Theodor Wilhelm Weierstrass (1815–1897).

[6] Renato Caccioppoli (1904–1959).

(b) $\mathbf{n}_A(x) = \lim_{r\downarrow 0} \dfrac{\int_{\mathbb{B}(x,r)} \mathbf{n}_A \, d|D\chi_A|}{|D\chi_A|\big(\mathbb{B}(x,r)\big)}$,

(c) $|\mathbf{n}_A| = 1$.

The structure theorem for sets of finite perimeter tells us that

$$|D\chi_A| = \mathcal{H}^{N-1} \llcorner \partial^* A \,. \tag{7.32}$$

Theorem 7.5.5. *Let U be an open set in $\mathbb{R}^{M+1}$ and let R be an integer-multiplicity current in $\mathcal{D}_{M+1}(U)$ with $\mathbf{M}_W(\partial R) < \infty$ for all $W \subset\subset U$. Then $T = \partial R$ is of integer multiplicity, and we can find a decreasing sequence of $(M+1)$-dimensional Lebesgue measurable sets $\{U_j\}_{j=-\infty}^{\infty}$ of locally finite perimeter in U such that*

$$R = \sum_{j=1}^{\infty} [\![U_j]\!] - \sum_{j=-\infty}^{0} [\![U \setminus U_j]\!] \,,$$

$$T = \sum_{j=-\infty}^{\infty} \partial [\![U_j]\!] \,,$$

$$\mu_T = \sum_{j=-\infty}^{\infty} \mu_{\partial [\![U_j]\!]} \,.$$

In particular,

$$\mathbf{M}_W(T) = \sum_{j=-\infty}^{\infty} \mathbf{M}_W(\partial [\![U_j]\!]) \quad \textit{for all} \quad W \subset\subset U \,.$$

Remark 7.5.6. Lebesgue measurable sets whose boundaries as currents have locally finite mass correspond to domains with locally finite perimeter. Here we describe that correspondence.

Let $\star : \mathcal{D}(U, \mathbb{R}^{M+1}) \to \mathcal{D}^M(U)$ be the version of the Hodge star operator[7] given by

$$\star g = \sum_{j=1}^{M+1} (-1)^{j-1} g_j \, dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_{N+1} \,.$$

Thus $d \star g = (\operatorname{div} g)\, dx_1 \wedge \cdots \wedge dx_{M+1}$. Then, for any $(M+1)$-dimensional Lebesgue measurable set $A \subseteq U$, we see that

$$\partial [\![A]\!](\star g) = [\![A]\!](d \star g) = \int_U \chi_A \operatorname{div} g \, d\mathcal{L}^{M+1} \,.$$

Thus, by definition of $|D\chi_A|$ and $\mathbf{M}(T)$, we find that *for any $(M+1)$-dimensional Lebesgue measurable $A \subseteq U$,*

[7] William Vallance Douglas Hodge (1903–1975).

(1) *A has locally finite perimeter in U if and only if* $\mathbf{M}_W(\partial[\![A]\!]) < \infty$ *holds for all* $W \subset\subset U$,
(2) *in case A has locally finite perimeter in U, then*

$$\mathbf{M}_W(\partial[\![A]\!]) = \int_W |D\chi_A|\,, \text{ for all } W \subset\subset U\,,$$

$$\overrightarrow{\partial[\![A]\!]} = \star\mathbf{n}_A\,, \text{ at } |D\chi_A|\text{-almost every point of } U\,.$$

Proof of Theorem 7.5.5. Now R must have the form

$$R = \boldsymbol{\tau}(S, \theta, \xi)\,,$$

where S is an $(M+1)$-dimensional Lebesgue measurable subset of U. We may suppose that $\xi(x) = \pm\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{M+1}$ and $\theta \in \mathbb{Z}^+$ for all $x \in U$ and that $\theta(x) = 0$ holds for $x \in U \setminus S$.

Set

$$\theta_+(x) = \begin{cases} \theta(x) & \text{if } \xi(x) = \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{M+1}, \\ 0 & \text{otherwise,} \end{cases}$$

$$\theta_-(x) = \begin{cases} \theta(x) & \text{if } \xi(x) = -\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{M+1}, \\ 0 & \text{otherwise,} \end{cases}$$

$$\widetilde{\theta} = \theta_+ - \theta_-\,.$$

We have

$$R(\omega) = \int_S a\,\widetilde{\theta}\,d\mathcal{L}^{M+1}(x)\,,$$

where $\omega = a\,dx_1 \wedge \cdots \wedge dx_{M+1} \in \mathcal{D}^{M+1}(U)$ and

$$\mathbf{M}_W(R) = \int_W |\widetilde{\theta}|\,d\mathcal{L}^{M+1}(x) \tag{7.33}$$

for all $W \subset\subset U$. Also we have

$$\mathbf{M}_W(T) = \int_W |D\widetilde{\theta}| \tag{7.34}$$

for all $W \subset\subset U$, because we can convert between the left-hand and right-hand sides of (7.34) using the operation $\star$. Thus we see that $\widetilde{\theta} \in BV_{\text{loc}}(U)$.

Now let

$$U_j = \{x \in U : \theta_+(x) \geq j\}\,,$$

$$W_j = \{x \in U : \theta_-(x) \geq j\}\,,$$

for $j = 1, 2, \ldots,$ so that

$$\widetilde{\theta} = \theta_+ - \theta_- = \sum_{j=1}^{\infty} \chi_{U_j} - \sum_{j=1}^{\infty} \chi_{W_j} \,.$$

Since

$$\begin{aligned} W_j &= \{\, x : \widetilde{\theta}(x) \le -j \,\} \\ &= U \setminus \{\, x : \widetilde{\theta}(x) > -j \,\} = U \setminus \{\, x : \widetilde{\theta}(x) \ge -j + 1 \,\} \,, \end{aligned}$$

we can set

$$U_j = \{x \in U : \theta(x) \ge -j\} \,,$$

for $j = 0, -1, -2, \ldots$, and conclude that

$$\widetilde{\theta} = \sum_{j=1}^{\infty} \chi_{U_j} - \sum_{j=-\infty}^{0} \chi_{U \setminus U_j}$$

and that

$$R = \sum_{j=1}^{\infty} [\![\, U_j \,]\!] - \sum_{j=-\infty}^{0} [\![\, U \setminus U_j \,]\!]$$

in U.

Since $T(\omega) = \partial R(\omega) = R(d\omega)$, $\omega \in \mathcal{D}^M(U)$, we have

$$\begin{aligned} T &= \partial R \\ &= \sum_{j=1}^{\infty} \partial [\![\, U_j \,]\!] - \sum_{j=0}^{\infty} \partial [\![\, V_j \,]\!] \\ &= \sum_{j=-\infty}^{\infty} \partial [\![\, U_j \,]\!] \,. \end{aligned} \tag{7.35}$$

Hence we have the necessary decomposition of T; it remains only to prove that each U_j has locally finite perimeter in U and that the corresponding measures sum up.

To this end, we will use a smoothing argument. Choose $0 < \epsilon < 1/6$ and let $\psi_j \in C^1(\mathbb{R})$, $j \in \mathbb{Z}$, satisfy

- $\psi_j(t) = 0$ for $t \le j - 1 + \epsilon$;
- $\psi_j(t) = 1$ for $t \ge j - \epsilon$;
- $0 \le \psi_j \le 1$;
- $\sup |\psi_j'| \le 1 + 3\epsilon$.

Then, because $\widetilde{\theta}$ is integer-valued, we have $\chi_{U_j} = \psi_j \circ \widetilde{\theta}$ for all $j \in \mathbb{Z}$.

Suppose that a is a nonnegative, compactly supported, continuous function on U and that $g = (g^1, \ldots, g^{M+1})$, where each component g^j is a compactly supported, continuous function on U. Suppose that $|g| \le a$ holds. For any choices of $k, \ell \in \mathbb{Z}$ with $k \le \ell$, we have

$$\int_U (\operatorname{div} g)\left(\sum_{j=k}^{\ell} \chi_{U_j}\right) d\mathcal{L}^{M+1} = \int_U (\operatorname{div} g)\left(\sum_{j=k}^{\ell} \psi_j \circ \widetilde{\theta}\right) d\mathcal{L}^{M+1}$$

$$= \lim_{\sigma \to 0^+} \int_U (\operatorname{div} g)\left(\sum_{j=k}^{\ell} \psi_j \circ \widetilde{\theta}_\sigma\right) d\mathcal{L}^{M+1}$$

$$= -\lim_{\sigma \to 0^+} \int_U g \cdot \left(\sum_{j=k}^{\ell} [\operatorname{grad} \widetilde{\theta}_\sigma]\, [\psi_j' \circ \widetilde{\theta}_\sigma]\right) d\mathcal{L}^{M+1}$$

$$\leq (1+3\epsilon) \lim_{\sigma \to 0^+} \int_U a\, |\operatorname{grad} \widetilde{\theta}_\sigma|\, d\mathcal{L}^{M+1}$$

$$= (1+3\epsilon) \int_U a\, |D\widetilde{\theta}|$$

$$= (1+3\epsilon) \int_U a\, d\mu_T\,.$$

Here $\widetilde{\theta}_\sigma$ are the mollified functions formed in our usual way (see Definition 5.5.1); we have used the fact that the mollification of a bounded variation function converges back to that function in a suitable topology (see [KPk 99, Section 3.6]), and we have also used (7.34).

By taking $k = \ell$, we see that each U_j has locally finite perimeter in U. If instead we take $k = -\ell$ and set $R_\ell = \sum_{j=1}^{\ell} [\![\, U_j \,]\!] - \sum_{j=0}^{\ell} [\![\, V_j \,]\!]$, we see that (with g as in Remark 7.5.6) the last display implies that

$$|R_\ell(d \star g)| \leq (1+3\epsilon) \int_U a\, d\mu_T\,.$$

Thus, with $T_\ell = \partial R_\ell$, we have that

$$\int_U a\, d\mu_{T_\ell} \leq \int_U a\, d\mu_T$$

holds for all $1 \leq \ell$ and all compactly supported $0 \leq a \in C^\infty(U)$.

Using (7.32), we also know that

$$R_\ell(d \star g) = \sum_{j=-\ell}^{\ell} \int_U \operatorname{div} g \cdot \chi_{U_j}\, d\mathcal{L}^{M+1}(x)$$

$$= \sum_{j=-\ell}^{\ell} \int_{\partial^* U_j} \mathbf{n}_j \cdot g\, d\mathcal{H}^M\,.$$

Here $\mathbf{n}_j$ is the outward unit normal for U_j and $\partial^* U_j$ is the reduced boundary for U_j. Since $U_{j+1} \subseteq U_j$, we have $\mathbf{n}_j = \mathbf{n}_k$ on $\partial^* U_j \cap \partial^* U_k$. Thus the last line may be rewritten as

$$T_\ell(\star g) = \int_U \mathbf{n} \cdot g\, h_\ell \, d\mathcal{H}^M \,. \tag{7.36}$$

In (7.36) we have let $h_\ell = \sum_{j=-\ell}^{\ell} \chi_{\partial^* U_j}$ and let $\mathbf{n}$ be defined on $\bigcup_{j=-\infty}^{\infty} \partial^* U_j$ by $\mathbf{n} = \mathbf{n}_j$ on $\partial^* U_j$.

Since $|\mathbf{n}| = 1$ on $\bigcup_{j=-\infty}^{\infty} \partial^* U_j$, we may thus conclude that

$$\begin{aligned}\int a \, d\mu_{T_\ell} &= \int a\, h_\ell \, d\mathcal{H}^M \\ &= \sum_{j=-\ell}^{\ell} \int_{\partial^* U_j} a \, d\mathcal{H}^M \\ &= \sum_{j=-\ell}^{\ell} \int a \, d\mu_{\partial [\![U_j]\!]} \,.\end{aligned}$$

Letting $\ell \to +\infty$, we can now conclude that

$$\mu_T \geq \sum_{j=-\infty}^{\infty} \mu_{\partial [\![U_j]\!]} \,.$$

The reverse inequality of course follows directly from (7.35). Hence the proof is complete. □

7.6 Slicing

Our first goal in this section is to define the concept of the "slice" of an integer-multiplicity current. Roughly speaking, we slice a current by intersecting it with the level set of a Lipschitz function. The process is closely related to the content of the coarea formula. First recall from Theorem 5.4.9 that if S is an $\mathcal{H}^M$-measurable, countably M-rectifiable set, then for $\mathcal{H}^M$-almost every $x \in S$, the approximate tangent plane $\mathbf{T}_x S$ exists. If, additionally, $f : \mathbb{R}^{M+K} \to \mathbb{R}$ is Lipschitz, then for $\mathcal{H}^M$-almost every $x \in S$, the approximate gradient $\nabla^S f(x) : \mathbf{T}_x S \to \mathbb{R}$ also exists.

The following lemma is a special case of Theorem 5.4.9.

Lemma 7.6.1. *Let S be an $\mathcal{H}^M$-measurable, countably M-rectifiable set and let $f : \mathbb{R}^{M+K} \to \mathbb{R}$ be Lipschitz. If we define S_+ to be the set of $x \in S$ for which $\mathbf{T}_x S$ and $\nabla^S f(x)$ exist and for which $\nabla^S f(x) \neq 0$, then for $\mathcal{L}^1$-almost all $t \in \mathbb{R}$, the following statements hold:*

(1) *$S_t = f^{-1}(t) \cap S_+$ is countably $\mathcal{H}^{M-1}$-rectifiable.*
(2) *For $\mathcal{H}^{M-1}$-almost every $x \in S_t$, the tangent spaces $\mathbf{T}_x S$ and $\mathbf{T}_x S_t$ both exist. In fact, $\mathbf{T}_x S_t$ is an $(M-1)$-dimensional subspace of $\mathbf{T}_x S$ and*

$$\mathbf{T}_x S = \{ y + \lambda \, \nabla^S f(x) : y \in \mathbf{T}_x S_t ,\ \lambda \in \mathbb{R} \} \,.$$

Finally, for any nonnegative $\mathcal{H}^M$-measurable function g on S we have

$$(3)\quad \int_{-\infty}^{\infty}\left(\int_{S_t} g\, d\mathcal{H}^{M-1}\right) d\mathcal{L}^1(t) = \int_S |\nabla^S f|\, g\, d\mathcal{H}^M .$$

Now we apply the lemma. We replace g in statement (3) by $g \cdot \chi_{\{x: f(x)<t\}}$. We thus obtain the identity

$$\int_{S\cap\{x:f(x)<t\}} |\nabla^S f|\, g\, d\mathcal{H}^M = \int_{-\infty}^{t}\int_{S_u} g\, d\mathcal{H}^{M-1} d\mathcal{L}^1(u) .$$

Hence the left-hand side is an absolutely continuous function of t and we may write

$$\frac{d}{dt}\int_{S\cap\{x:f(x)<t\}} |\nabla^S f|\, g\, d\mathcal{H}^M = \int_{S_t} g\, d\mathcal{H}^{M-1} \quad \text{for all } t \in \mathbb{R} .$$

We let $T = \boldsymbol{\tau}(S, \theta, \xi)$ be an integer-multiplicity current in U, with U an open set in $\mathbb{R}^{M+K}$. Let f be a Lipschitz function on U and let

$$\theta_+(x) = \begin{cases} 0 & \text{if } \nabla^S f(x) = 0 , \\ \theta(x) & \text{if } \nabla^S f(x) \neq 0 . \end{cases}$$

For $\mathcal{L}^1$-almost every $t \in \mathbb{R}$ with $\mathbf{T}_x S$, $\mathbf{T}_x S_t$ existing for $\mathcal{H}^{M-1}$-almost every $x \in S_t$, and such that the identity (3) of Lemma 7.6.1 holds, we define $\xi_t(x)$ by

$$\xi_t(x) = \xi(x) \llcorner \left(\frac{\nabla^S f(x)}{|\nabla^S f(x)|}\right) \tag{7.37}$$

and we note that $\xi_t(x)$ has the following properties

- $\xi_t(x)$ is simple;
- $\xi_t(x)$ lies in $\bigwedge_{M-1}(\mathbf{T}_x S_t) \subseteq \bigwedge_{M-1}(\mathbf{T}_x S)$;
- $\xi_t(x)$ has unit length for $\mathcal{H}^{M-1}$-almost every $x \in S_t$.

Continuing to assume that $T \in \mathcal{D}_M(U)$ is given by $T = \boldsymbol{\tau}(S, \theta, \xi)$, we define the slice of T by the Lipschitz mapping f as follows.

Definition 7.6.2. For $\mathcal{L}^1$-almost every $t \in \mathbb{R}$, we know that $\mathbf{T}_x S$, $\mathbf{T}_x S_t$ exist and (3) of Lemma 7.6.1 holds for $\mathcal{H}^{M-1}$-almost every $x \in S_t$. We now define the integer-multiplicity current $\langle T, f, t\rangle \in \mathcal{D}_{M-1}$ by

$$\langle T, f, t\rangle = \boldsymbol{\tau}(S_t, \theta_t, \xi_t) ,$$

where $\xi_t(x)$ is as in (7.37) and

$$\theta_t = \theta_+\big|_{S_t} .$$

We call $\langle T, f, t\rangle$ the *slice* of the current T by the function f at t. See Figure 7.6.

The next lemma records some of the main properties of slices.

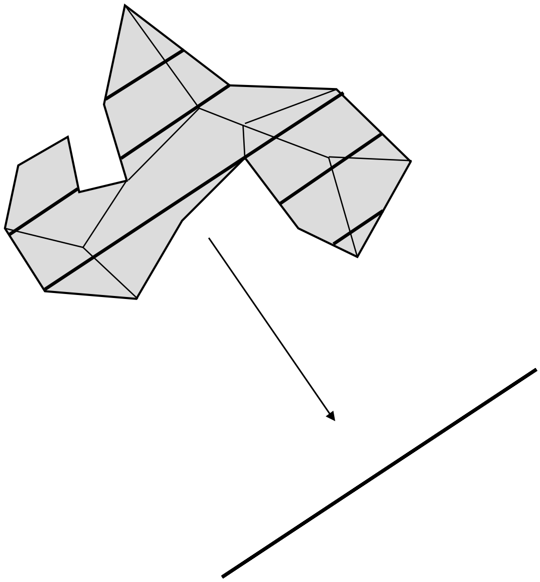

Fig. 7.6. Slicing.

Lemma 7.6.3. *Slices enjoy these features:*

(1) *For each open $W \subseteq U$,*

$$\int_{-\infty}^{\infty} \mathbf{M}_W(\langle T, f, t\rangle)\, d\mathcal{L}^1(t) = \int_{S\cap W} |\nabla^S f|\, \theta\, d\mathcal{H}^M$$

$$\leq \left(\operatorname*{ess\,sup}_{S\cap W} |\nabla^S f|\right) \mathbf{M}_W(T)\,.$$

(2) *If $\mathbf{M}_W(\partial T) < \infty$ for all $W \subset\subset U$, then for $\mathcal{L}^1$-almost every $t \in \mathbb{R}$, we have*

$$\langle T, f, t\rangle = \partial[\, T \llcorner \{x : f(x) < t\}\,] - (\partial T) \llcorner \{x : f(x) < t\}\,.$$

(3) *If ∂T is of integer multiplicity in $\mathcal{D}_{M-1}(U)$, then for $\mathcal{L}^1$-almost every $t \in \mathbb{R}$, we have*

$$\langle \partial T, f, t\rangle = -\partial\langle T, f, t\rangle\,.$$

Remark 7.6.4. The equation in (2) is often the most intuitively helpful way to think of a slice of T.

Proof.
(1) To prove (1), take $g = \theta_+$ in formula (3) of Lemma 7.6.1.

(2) Recall that the countable M-rectifiability of S allows us to write

$$S = \bigcup_{j=0}^{\infty} S_j\,,$$

where $S_i \cap S_j = \emptyset$ when $i \neq j$, $\mathcal{H}^M(S_0) = 0$, and each $S_j \subseteq V_j$, $j \geq 1$, with V_j an embedded C^1 submanifold of $\mathbb{R}^{M+K}$. This decomposition, together with the definition of ∇^S, shows that if h is Lipschitz on $\mathbb{R}^{M+K}$ and if h_σ are the mollifications of h (formed in the usual way—see (5.31)) then as $\sigma \to 0$,

$$v \cdot \nabla^S h_\sigma \text{ converges to } v \cdot \nabla^S h \tag{7.38}$$

for any fixed, bounded $\mathcal{H}^M$-measurable function v with values in $\mathbb{R}^{M+K}$; that is, $\nabla^S h_\sigma$ converges to $\nabla^S h$ in the weak topology of $L^2(\mu_T)$. To verify this assertion, one need only check that (2) holds with the C^1 submanifolds V_j replacing S_j and with v vanishing on $\mathbb{R}^{M+K} \setminus S_j$; one approximates v by a smooth function and exploits the fact that the h_σ converge uniformly to h.

Now let $\epsilon > 0$ and let γ be the unique piecewise linear, continuous function satisfying

$$\gamma(s) = \begin{cases} 1 & \text{if } s < t - \epsilon\,, \\ 0 & \text{if } s > t\,. \end{cases}$$

Then γ is Lipschitz, and we apply the reasoning of the preceding paragraph to $h = \gamma \circ f$. Letting $\omega \in \mathcal{D}^M(U)$, we have

$$\begin{aligned} \partial T(h_\sigma \omega) &= T(d(h_\sigma \omega)) \\ &= T(dh_\sigma \wedge \omega) + T(h_\sigma d\omega)\,. \end{aligned}$$

Now, applying the integral representation (7.4) to ∂T, we see that

$$(\partial T \llcorner h)(\omega) = \lim_{\sigma \to 0^+} T(dh_\sigma \wedge \omega) + (T \llcorner h)(d\omega)\,. \tag{7.39}$$

Since $\xi(x)$ orients $\mathbf{T}_x S$, we have

$$\begin{aligned} \langle\, dh_\sigma \wedge \omega,\ \xi(x)\,\rangle &= \langle\, (dh_\sigma(x))^T \wedge \omega^T,\ \xi(x)\,\rangle \\ &= \langle\, (dh_\sigma(x))^T \wedge \omega,\ \xi(x)\,\rangle\,. \end{aligned}$$

Here λ^T denotes the orthogonal projection of $\Lambda^q(\mathbb{R}^{M+K})$ onto $\Lambda^q(\mathbf{T}_x S))$. We conclude that

$$\begin{aligned} T(dh_\sigma \wedge \omega) &= \int_S \langle\, (dh_\sigma(x))^T \wedge \omega,\ \xi(x)\,\rangle\, \theta\, d\mathcal{H}^M \\ &= \int_S \langle\, \omega,\ \xi(x) \llcorner \nabla^S h_\sigma(x)\,\rangle\, \theta\, d\mathcal{H}^M\,. \end{aligned}$$

Thus we may use (7.38) to write

$$\lim_{\sigma \to 0^+} T(dh_\sigma \wedge \omega) = \int_S \langle\, \omega,\ \xi(x) \llcorner \nabla^S h(x)\,\rangle\, \theta\, d\mathcal{H}^M\,. \tag{7.40}$$

By definition of $\nabla^S h$, and by the chain rule for Lipschitz functions, we have

$$\nabla^S h = \gamma'(f)\, \nabla^S f \quad \text{for } \mathcal{H}^M\text{-almost every point of } S\,. \tag{7.41}$$

Here we have used the convention that $\gamma'(f) = 0$ when f takes one of the values t or $t - \epsilon$ for which γ is not differentiable. Notice also that

$$\nabla^S h(x) = \nabla^S f(x) = 0$$

for $\mathcal{H}^M$-almost every point in $\{x \in S : f(x) = c\}$, c a constant.

Now (7.39), (7.40), and (7.41) tell us that

$$(\partial T \llcorner h)(\omega) = -\frac{1}{\epsilon}\int_{S\cap\{t-\epsilon<f<t\}} \langle\, \omega,\ \xi \llcorner \nabla^S f\,\rangle\, \theta\, d\mathcal{H}^M + (T \llcorner h)(d\omega)\,.$$

We conclude by letting $\epsilon \to 0$ and exploiting the remark following the proof of Lemma 7.6.1 with $g = \theta\,\langle\, \omega,\ \xi \llcorner \nabla^S f/|\nabla^S f|\,\rangle$. In fact, by considering a countable dense set of $\omega \in \mathcal{D}^M(U)$, we can show that the aforementioned remark is applicable with this choice of g except on a set F of points t having measure 0, with F independent of ω. That completes the proof of (2).

(3) To prove part (3) of the theorem, we begin by applying part (2) with ∂T replacing T. Since $\partial^2 = 0$, we find that

$$\langle \partial T, f, t\rangle = \partial[(\partial T) \llcorner \{f < t\}]\,.$$

If we instead apply ∂ to the identity in (2) we obtain

$$\partial[\,(\partial T) \llcorner \{x : f(x) < t\}\,] = -\partial\langle T, f, t\rangle\,.$$

Therefore part (3) is proved. □

The right-hand side of the equation in part (2) of Lemma 7.6.3 makes sense when T and ∂T are representable by integration, without the necessity of assuming that T is an integer-multiplicity current. Thus we may consider slicing for an arbitrary current $T \in \mathcal{D}_M(U)$ that together with its boundary has locally finite mass in U. So suppose that $\mathbf{M}_W(T) + \mathbf{M}_W(\partial T) < \infty$ for all $W \subset\subset U$. Initially, we define two types of slices by

$$\langle T, f, t_-\rangle = \partial[\,T \llcorner \{x : f(x) < t\}\,] - (\partial T) \llcorner \{x : f(x) < t\} \tag{7.42}$$

and

$$\langle T, f, t_+\rangle = -\partial[\,T \llcorner \{x : f(x) > t\}\,] + (\partial T) \llcorner \{x : f(x) > t\}\,. \tag{7.43}$$

For only countably many values of t does it hold that

$$\mathbf{M}[\,T \llcorner \{x : f(x) = t\}\,] + \mathbf{M}[\,(\partial T) \llcorner \{x : f(x) = t\}\,] > 0\,.$$

For all other values of t, we have

$$\langle T, f, t_- \rangle - \langle T, f, t_+ \rangle = \partial[\, T \llcorner \{x : f(x) \neq t\}\,] - (\partial T) \llcorner \{x : f(x) \neq t\} = 0\,,$$

and we denote the common value of $\langle T, f, t_+ \rangle$ and $\langle T, f, t_- \rangle$ by $\langle T, f, t \rangle$.

The important facts about these slices are that if f is Lipschitz on U, then

$$\operatorname{spt} \langle T, f, t_\pm \rangle \subset \operatorname{spt} T \cap \{x : f(x) = t\} \tag{7.44}$$

and, for all open $W \subset U$,

$$\mathbf{M}_W\Big(\langle T, f, t_+ \rangle\Big) \\ \leq \operatorname*{ess\,sup}_W |Df| \cdot \liminf_{h \to 0^+} \frac{1}{h} \mathbf{M}_W\Big(T \llcorner \{t < f < t+h\}\Big)\,, \tag{7.45}$$

$$\mathbf{M}_W\Big(\langle T, f, t_- \rangle\Big) \\ \leq \operatorname*{ess\,sup}_W |Df| \cdot \liminf_{h \to 0^+} \frac{1}{h} \mathbf{M}_W\Big(T \llcorner \{t-h < f < t\}\Big)\,. \tag{7.46}$$

Certainly $\mathbf{M}_W(T \llcorner \{f < t\})$ is increasing in t; thus the function is differentiable for $\mathcal{L}^1$-almost every $t \in \mathbb{R}$ and

$$\int_a^b \frac{d}{dt} \mathbf{M}_W(T \llcorner \{f < t\})\, d\mathcal{L}^1(t) \leq \mathbf{M}_W(T \llcorner \{a < f < b\})$$

for any $a < b$. Thus (7.46) yields the following bound on the upper integral of the mass of the slices:

$$\overline{\int_a^b} \mathbf{M}_W(\langle T, f, t_\pm \rangle)\, d\mathcal{L}^1(t) \leq \operatorname*{ess\,sup}_W |Df| \cdot \mathbf{M}_W(T \llcorner \{a < f < b\}) \tag{7.47}$$

for every open $W \subset U$.

Now we prove (7.44), (7.45), and (7.46). First consider the case that f is C^1 and let γ be any smooth, increasing function from $\mathbb{R}$ to $\mathbb{R}^+$. We have

$$\begin{aligned} \partial(T \llcorner \gamma \circ f)(\omega) - ((\partial T) \llcorner \gamma \circ f)(\omega) &= (T \llcorner \gamma \circ f)(d\omega) - ((\partial T) \llcorner \gamma \circ f)(\omega) \\ &= T(\gamma \circ f\, d\omega) - T(d(\gamma \circ f \omega)) \\ &= -T(\gamma'(f) df \wedge \omega)\,. \end{aligned} \tag{7.48}$$

Now let $\epsilon > 0$ be arbitrary and select γ piecewise linear such that

$$\gamma(t) = \begin{cases} 0 & \text{for } t < a\,, \\ 1 & \text{for } t > b\,. \end{cases}$$

We also suppose that $0 \leq \gamma'(t) \leq [1+\epsilon]/[b-a]$ for $a < t < b$. Then the left side of (7.48) converges to $\langle T, f, a_+ \rangle$ if we let b decrease to a. Hence (7.44) now follows because $\operatorname{spt} \gamma' \subset [a, b]$.

Furthermore, the right-hand side R of (7.48) is majorized by

$$|R| \le (\sup_W |Df|) \cdot \left(\frac{1+\epsilon}{b-a}\right) \cdot \mathbf{M}_W(T \llcorner \{a < f < b\}) \cdot (\sup_W |\omega|)$$

for all ω with support in W. Hence we have (7.45) for $f \in C^1$. Equation (7.46) for $f \in C^1$ is proved similarly.

To handle the more general Lipschitz f, we simply examine f_σ in place of f in (7.42), (7.43), and in the preceding argument, and let $\sigma \to 0^+$ to obtain the conclusion.

We conclude this section with a discussion of slicing a current $T \in \mathcal{D}_M$ by a Lipschitz function $F : \mathbb{R}^{M+K} \to \mathbb{R}^L$, where $2 \le L \le M$. The most straightforward approach is to formulate the definition iteratively. For example, if T is integer-multiplicity, then define

$$\langle\, T,\ F,\ (t_1, \ldots, t_L)\,\rangle = \langle\, \langle \cdots \langle\, \langle\, T,\ F_1,\ t_1\,\rangle,\ F_2,\ t_2\,\rangle, \cdots \rangle, F_L,\ t_L \rangle\,,$$

where $F_1, F_2, \ldots, F_L$ are the components of F.

Of particular interest to us will be slicing the integer-multiplicity current $T = \boldsymbol{\tau}(S, \theta, \xi)$ by the orthogonal projection onto a coordinate M-plane. Let $\mathbf{p} : \mathbb{R}^{M+K} \to \mathbb{R}^M$ map $(x_1, x_2, \ldots, x_{M+K})$ to $(x_1, x_2, \ldots, x_M)$. Proceeding in a manner similar to Lemma 7.6.1, we define S_+ to be the set of $x \in S$ for which $\mathbf{T}_x S$ and $D_S\mathbf{p}(x)$ exist and for which rank $D_S\mathbf{p}(x) = M$. Then for $\mathcal{L}^M$-almost every $t = (t_1, \ldots, t_M)$, we have

$$\langle\, T,\ \mathbf{p},\ t\,\rangle = \sum_{x \in \mathbf{p}^{-1}(t) \cap S_+} \sigma(x)\, \theta(x)\, \boldsymbol{\delta}_x\,, \tag{7.49}$$

where $\sigma(x) = \operatorname{sgn}(a)$ when $\langle \bigwedge_M \mathbf{p},\ \xi(x)\,\rangle = a\, dx_1 \wedge \cdots \wedge dx_M$.

The next proposition is then evident from the definition in (7.49).

Proposition 7.6.5. *Let* $\mathbf{p} : \mathbb{R}^{M+K} \to \mathbb{R}^M$ *be projection onto the coordinate plane as above.*

(1) *If* $h : \mathbb{R}^M \to \mathbb{R}^K$, $A \subseteq \mathbb{R}^M$ *is* $\mathcal{L}^M$*-measurable, and* $H : \mathbb{R}^M \to \mathbb{R}^{M+K}$ *is given by* $H(t) = (t, h(t))$*, then*

$$\langle\, H_\# [\![A]\!],\ \mathbf{p},\ t\,\rangle = \boldsymbol{\delta}_{H(t)}\,.$$

(2) *For continuous* $\phi : \mathbb{R}^{M+K} \to \mathbb{R}$ *and* $\psi : \mathbb{R}^M \to \mathbb{R}$*, if at least one of the two functions is compactly supported, then*

$$\int \psi(x) \langle T, p, x\rangle\, (\phi)\, d\mathcal{L}^M(x) = [T \llcorner (\psi \circ \mathbf{p})\, dx_1 \wedge \cdots \wedge dx_M](\phi)\,.$$

The interested reader will find an extremely thorough treatment of slicing in a very general context in [Fed 69, Section 4.3].

7.7 The Deformation Theorem

One of the cornerstones of geometric measure theory, and more particularly of the theory of currents, is the deformation theorem. There are both scaled and unscaled versions of this result. The scaled version of the theorem is obtained by applying homotheties to the unscaled version, so we will concentrate on the unscaled version. First we need some notation that will be particular to this treatment:

- $1 \le M, K \in \mathbb{Z}$ (we will be considering M-dimensional currents in $\mathbb{R}^{M+K}$);
- $C = \underbrace{[0, 1] \times [0, 1] \times \cdots \times [0, 1]}_{M+K \text{ factors}}$ (the standard unit cube in $\mathbb{R}^{M+K}$);
- $\mathbb{Z}^{M+K} = \{(z_1, z_2, \ldots, z_{M+K}) : z_j \in \mathbb{Z}\}$ (the integer lattice in $\mathbb{R}^{M+K}$);
- for $j = 0, 1, \ldots, M + K$, we will use $\mathcal{L}_j$ to denote the collection of all the j-dimensional faces occurring in the cubes

$$\boldsymbol{t}_z(C) = [z_1,\, z_1 + 1] \times [z_2,\, z_2 + 1] \times \cdots \times [z_{M+K},\, z_{M+K} + 1]$$

as $z = (z_1, z_2, \ldots, z_{M+K}) \in \mathbb{Z}^{M+K}$ ranges over the integer lattice.

Each M-dimensional face $F \in \mathcal{L}_M$ corresponds (once we make a choice of orientation) to an integer-multiplicity current $[\![F]\!]$. For currents having finite mass and having boundaries of finite mass, the deformation theorem tells us how such a current can be approximated by a linear combination of the $[\![F]\!]$, $F \in \mathcal{L}_M$. The name "deformation theorem" arises from the proof of the theorem. The precise statement is as follows.

Theorem 7.7.1 (Deformation Theorem, Unscaled Version). *Suppose that T is an M-dimensional current in $\mathbb{R}^{M+K}$ with*

$$\mathbf{M}(T) + \mathbf{M}(\partial T) < \infty\,.$$

Then we may write

$$T - P = \partial R + S\,,$$

where $P \in \mathcal{D}_M(\mathbb{R}^{M+K})$, $R \in \mathcal{D}_{M+1}(\mathbb{R}^{M+K})$, and $S \in \mathcal{D}_M(\mathbb{R}^{M+K})$ satisfy

$$P = \sum_{F \in \mathcal{L}_M} p_F\, [\![F]\!]\,, \quad \text{where } p_F \in \mathbb{R}, \text{for } F \in \mathcal{L}_M, \tag{7.50}$$

$$\mathbf{M}(P) \le c\,\mathbf{M}(T)\,, \qquad \mathbf{M}(\partial P) \le c\,\mathbf{M}(\partial T)\,, \tag{7.51}$$

$$\mathbf{M}(R) \le c\,\mathbf{M}(T)\,, \qquad \mathbf{M}(S) \le c\,\mathbf{M}(\partial T)\,. \tag{7.52}$$

The constant c depends on M and K. Further,

$$\operatorname{spt} P \cup \operatorname{spt} R \subset \left\{ x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{M+K} \right\},$$

$$\operatorname{spt} \partial P \cup \operatorname{spt} S \subset \left\{ x : \operatorname{dist}(x, \operatorname{spt} \partial T) < 2\sqrt{M+K} \right\}.$$

Moreover, if T is an integer-multiplicity current, then P and R can be chosen to be integer-multiplicity currents. Also, in this case, the numbers p_F in (7.50) *are integers. If in addition ∂T is integer-multiplicity, then S can be chosen to be integer-multiplicity.* [*We shall see later that in case T is integer-multiplicity and $\mathbf{M}(\partial T) < \infty$, then ∂T is automatically integer-multiplicity.*]

See Figure 7.7.

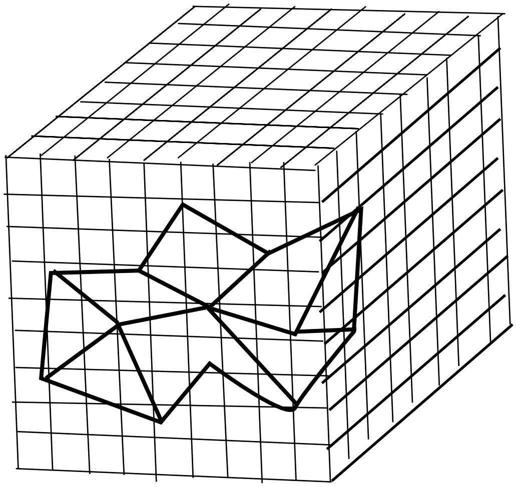

Fig. 7.7. The deformation theorem.

A few remarks about the unscaled deformation theorem are now in order. First, since $\partial S = \partial T - \partial P$ and $\mathbf{M}(\partial P) \le c\,\mathbf{M}(\partial T)$, it is an immediate corollary that $\mathbf{M}(\partial S) \le c\,\mathbf{M}(\partial T)$. Also, the inequalities $\mathbf{M}(\partial P) \le c\,\mathbf{M}(\partial T)$ and $\mathbf{M}(S) \le c\,\mathbf{M}(\partial T)$ yield immediately that when $\partial T = 0$ then $\partial P = 0$ and $S = 0$.

The estimate for the mass of P in (7.51) depends only on the mass of T and not on the mass of ∂T. This estimate, due to Leon Simon, is an improvement over the original estimate given by Federer and Fleming.

For the record now, we shall also state the scaled version of the deformation theorem. In the statement, we will use the notation $\boldsymbol{\eta}_t : \mathbb{R}^{M+K} \to \mathbb{R}^{M+K}$ for the homothety defined by

$$\boldsymbol{\eta}_t(x) = tx\,.$$

Theorem 7.7.2 (Deformation Theorem, Scaled Version). *Fix $\rho > 0$. Suppose that T is an M-dimensional current in $\mathbb{R}^{M+K}$ with*

$$\mathbf{M}(T) + \mathbf{M}(\partial T) < \infty\,.$$

Then we may write

$$T - P = \partial R + S\,,$$

where $P \in \mathcal{D}_M(\mathbb{R}^{M+K})$, $R \in \mathcal{D}_{M+1}(\mathbb{R}^{M+K})$, and $S \in \mathcal{D}_M(\mathbb{R}^{M+K})$. We have

$$P = \sum_{F \in \mathcal{L}_M} p_F \, \boldsymbol{\eta}_{\rho\#}[\![F]\!] , \tag{7.53}$$

where $p_F \in \mathbb{R}$, for $F \in \mathcal{L}_M$, and

$$\mathbf{M}(P) \le c\,\mathbf{M}(T) , \qquad \mathbf{M}(\partial P) \le c\,\mathbf{M}(\partial T) , \tag{7.54}$$

$$\mathbf{M}(R) \le c\,\rho\,\mathbf{M}(T) , \qquad \mathbf{M}(S) \le c\,\rho\,\mathbf{M}(\partial T) . \tag{7.55}$$

The constant c depends only on M and K. Further,

$$\operatorname{spt} P \cup \operatorname{spt} R \subset \left\{ x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{M+K}\rho \right\} ,$$

$$\operatorname{spt} \partial P \cup \operatorname{spt} S \subset \left\{ x : \operatorname{dist}(x, \operatorname{spt} \partial T) < 2\sqrt{M+K}\rho \right\} .$$

In the case that T is integer-multiplicity then so are P and R. If ∂T is integer-multiplicity then so is S.

The scaled deformation theorem is an immediate consequence of applying the unscaled theorem to $\boldsymbol{\eta}_{1/\rho\#}T$ and then applying $\boldsymbol{\eta}_{\rho\#}$ to the P, R, and S so obtained. The two factors of ρ in (7.55) occur because the dimension of R is 1 more than the dimension of T and the dimension of S is 1 more than the dimension of ∂T. Thus it will suffice to prove the unscaled deformation theorem.

The essence of the proof of the unscaled theorem consists in pushing forward by a retraction to deform the current T onto the M-skeleton L_M. The first step in our presentation of the proof will therefore be the construction of the retraction. For this construction, we introduce additional notation.

- For $j = 0, 1, \ldots, M+K$, set

  $$L_j = \bigcup_{F \in \mathcal{L}_j} F ,$$

  so that L_j is the j-skeleton of the cubical decomposition

  $$\bigcup_{z \in \mathbb{Z}^{M+K}} (z + C)$$

  of $\mathbb{R}^{M+K}$;
- for $j = 0, 1, \ldots, M+K$, set

  $$\widetilde{L}_j = (\tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}) + L_j .$$

Clearly we have

$$\mathbb{R}^{M+K} = L_{M+K} \supseteq L_{M+K-1} \supseteq L_{M+K-2} \supseteq \cdots \supseteq L_0 \,,$$

and similar containments hold for the $\widetilde{L}_j$.

Observe that

$$\begin{aligned} \widetilde{L}_0 \cap L_{M+K-1} &= \emptyset, \\ \widetilde{L}_1 \cap L_{M+K-2} &= \emptyset, \\ &\vdots \\ \widetilde{L}_{K-1} \cap L_M &= \emptyset\,; \end{aligned}$$

these equations hold because

- a point in $L_{M+K-j-1}$ must have $j+1$ integral coordinate values,

whereas

- a point in $\widetilde{L}_j$ must have $M+K-j$ coordinate values that are multiples of $1/2$.

Similarly we see that, for any face $F \in \mathcal{L}_{M+K-j}$, the center of F is the point of intersection of F and $\widetilde{L}_j$. Thus the nearest-point-retraction

$$\xi_j : L_{M+K-j} \setminus L_{M+K-j-1} \to \widetilde{L}_j$$

is well-defined. We define the retraction

$$\psi_j : L_{M+K-j} \setminus \widetilde{L}_j \to L_{M+K-j-1}$$

by requiring that

- $\psi_j(x) = x$, if $x \in L_{M+K-j-1}$;
- the line segment connecting $\psi_j(x)$ and $\xi_j(x)$ contains x if $x \in L_{M+K-j} \setminus [\,\widetilde{L}_j \bigcup L_{M+K-j-1}\,]$.

In effect, ψ_j radially projects the points in $F \in \mathcal{L}_{M+K-j}$ from the center of F onto the relative boundary of F, so of course ψ_j cannot be defined at the center of F and still be continuous.

We define

$$\psi : \mathbb{R}^{M+K} \setminus \widetilde{L}_{K-1} \to L_M$$

by

$$\psi = \psi_{K-1} \circ \psi_{K-2} \circ \cdots \circ \psi_0 \,.$$

Figure 7.8 illustrates the mapping ψ (for $M = 1$ and $K = 2$) by showing how ψ_0 maps a curve in the unit cube onto the faces of the cube by radially projecting from the center of the cube. Then ψ_1 maps that projected curve onto the edges of the cube by radially projecting from the centers of the faces.

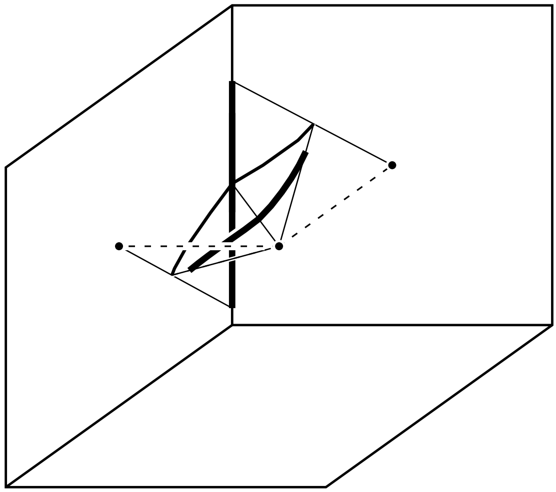

Fig. 7.8. The mapping ψ.

It is crucial to estimate the norm of the differential of ψ. Because ψ is the composition of radial projections, one can bound $|D\psi|$ from below by

$$1 \le |D\psi| .$$

One also expects to be able to bound $|D\psi|$ from above by a constant divided by the minimum distance to any of the centers of projection. We will prove such an upper bound, but in fact our proof will be more analytic than geometric. We will need the next elementary lemma.

Lemma 7.7.3. *If* $0 \le a_0 \le a_1 \le \cdots \le a_{j+1} < 1/2$*, then*

$$\prod_{i=0}^{j} (1 + 2a_i - 2a_{i+1})^{-1} \le \frac{1}{1 - 2a_{j+1}} .$$

Proof. We argue by induction. The result is obvious if $j = 0$ and easily verified if $j = 1$.

Now, assuming that the result holds for j, we see that

$$\begin{aligned} \prod_{i=0}^{j+1} (1 + 2a_i - 2a_{i+1})^{-1} &\le (1 - 2a_{j+1})^{-1} \, (1 + 2a_{j+1} - 2a_{j+2})^{-1} \\ &\le \frac{1}{1 - 2a_{j+2}} , \end{aligned}$$

where the first inequality follows from the induction hypothesis and the second inequality follows from the case $j = 1$. $\square$

Lemma 7.7.4. *There is a constant $c = c = c(M, K)$ such that*

$$|D\psi(x)| \le \frac{c}{\rho}$$

holds for $\mathcal{L}^{M+K}$-almost every $x \in \mathbb{R}^{M+K} \setminus \widetilde{L}_{K-1}$, where $\rho = \mathrm{dist}(x, \widetilde{L}_{K-1})$.

Proof. First note that if θ is the composition of reflections through planes of the form $\mathbf{e}_j \cdot x = k/2$, $k \in \mathbb{Z}$, translations of the form $\boldsymbol{t}_z$, $z \in \mathbb{Z}^{M+K}$, and permutations of coordinates, then $\theta \circ \psi \circ \theta^{-1} = \psi$. Thus it suffices to consider points $x = (x_1, x_2, \dots, x_{M+K})$ of the form

$$0 < x_1 < x_2 < \cdots < x_{M+K} < 1/2\,.$$

Since no coordinate of x equals $1/2$, we have $x \notin \widetilde{L}_{M+K}$. One computes $\psi_0(x)$ by finding the smallest value of $t \in \mathbb{R}$ for which

$$(1-t)\left(\tfrac{1}{2}, \tfrac{1}{2}, \dots, \tfrac{1}{2}\right) + t\left(x_1, x_2, \dots, x_{M+K}\right)$$

has a coordinate equal to 0. In fact, that smallest value of t is $1/(1-2x_1)$, and we see that

$$\psi_0(x) = \frac{1}{1-2x_1}\,(0, x_2 - x_1, \dots, x_{M+K} - x_1)\,.$$

Proceeding in this way, we see that

$$\psi_1 \circ \psi_0(x) = \frac{1}{1-2x_1}\,\frac{1}{1-2(x_2-x_1)}\,(0, 0, x_3 - x_2, \dots, x_{M+K} - x_2)$$

and, ultimately, that

$$\begin{aligned}
\psi(x) &= \psi_{K-1} \circ \psi_{K-2} \circ \cdots \circ \psi_0(x) \\
&= (1-2x_1)^{-1} \prod_{j=1}^{K-1} [\,1 - 2(x_{j+1} - x_j)\,]^{-1} \\
&\qquad (0, 0, \dots, 0, x_{K+1} - x_K, \dots, x_{M+K} - x_K) \\
&= \prod_{j=0}^{K-1} (1 + 2x_j - 2x_{j+1})^{-1} \\
&\qquad (0, 0, \dots, 0, x_{K+1} - x_K, \dots, x_{M+K} - x_K) \in L_M\,,
\end{aligned} \tag{7.56}$$

where $x_0 = 0$.

By computing the partial derivatives of

$$(x_I - x_K) \prod_{j=0}^{K-1} (1 + 2x_j - 2x_{j+1})^{-1}\,, \quad \text{for } 1 + K \le I \le M + K,$$

and using the estimate in Lemma 7.7.3, we see that each

$$\left|\frac{\partial(\mathbf{e}_I \cdot \psi)}{\partial x_J}\right|$$

can be bounded by a constant multiple of $(1 - 2x_{M+K})^{-1}$. Since the point of $\widetilde{L}_{K-1}$ nearest to x is $\left(x_1, x_2, \ldots, x_{K-1}, \frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\right)$, we have

$$\rho = 2^{-1} \left(\sum_{j=K}^{M+K} (1 - 2x_j)^2\right)^{1/2} \geq 2^{-1} (1 - 2x_{M+K}),$$

so the desired bound holds. □

7.8 Proof of the Unscaled Deformation Theorem

We divide the proof into four steps.

Step 1. *We claim that*

$$\int_{\widetilde{C}} |D\psi(x)|^M \, d\mathcal{L}^{M+K}(x) < \infty,$$

where $\widetilde{C} = [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}] \times \cdots \times [-\frac{1}{2}, \frac{1}{2}]$.

Using the estimate in Lemma 7.7.4, we see that it will suffice to bound $\int_{\widetilde{C}} (\widetilde{\rho})^{-M} \, d\mathcal{L}^{M+K}$, where $\widetilde{\rho}$ is the distance from a point in $\mathbb{R}^{M+K}$ to the union of the $(K-1)$-dimensional coordinate planes. Since the distance from a point to the union of the $(K-1)$-dimensional coordinate planes is the minimum of the distances to each of the individual $(K-1)$-dimensional coordinate planes, if we write $x = (x', x'') \in \mathbb{R}^{M+K}$, where $x' \in \mathbb{R}^{M+1}$ and $x'' \in \mathbb{R}^{K-1}$, then it will suffice to bound $\int_{\widetilde{C}} |x'|^{-M} \, d\mathcal{L}^{M+K}(x)$. We may also replace $\widetilde{C}$ by the larger set $B_1 \times B_2$, where

$$B_1 = \{x' \in \mathbb{R}^{M+1} : |x'| \leq 2^{-1}(M+1)^{1/2}\},$$

$$B_2 = \{x'' \in \mathbb{R}^{K-1} : |x''| \leq 2^{-1}(K-1)^{1/2}\}.$$

We have

$$\int_{\widetilde{C}} |x'|^{-M} \, d\mathcal{L}^{M+K}(x) \leq \int_{B_1} \int_{B_2} |x'|^{-M} \, d\mathcal{L}^{M+1}(x') \, d\mathcal{L}^{K-1}(x'')$$

$$= \mathcal{L}^{K-1}(B_2) \cdot \int_0^{2^{-1}(M+1)^{1/2}} \int_{\mathbb{R}^{M+1} \cap \{\xi : |\xi| = r\}} r^{-M} \, d\mathcal{H}^M(\xi) \, d\mathcal{L}^1(r)$$

$$= \mathcal{L}^{K-1}(B_2) \cdot \mathcal{H}^M\Big(\mathbb{R}^{M+1} \cap \{\xi : |\xi| = 1\}\Big) \cdot 2^{-1}(M+1)^{1/2} < \infty.$$

Step 2. *There exists a point $a \in \widetilde{C}$ such that*

$$\int |D\psi(x)|^M \, d\|\boldsymbol{t}_{a\#}T\|(x) \le c\,\mathbf{M}(T)\,,$$

$$\int |D\psi(x)|^M \, d\|\boldsymbol{t}_{a\#}\partial T\|(x) \le c\,\mathbf{M}(\partial T)$$

hold, where c depends only on M and K. (*Recall that $\|W\|$ denotes the total variation measure of the current W.*)

Set

$$c = 4 \int_{\widetilde{C}} |D\psi(x)|^M \, d\mathcal{L}^{M+K}(x)\,.$$

By the symmetry in the construction of ψ we have

$$\int_{\widetilde{C}} |D\psi(x+a)|^M \, d\mathcal{L}^{M+K}(a) = \int_{\widetilde{C}} |D\psi(a)|^M \, d\mathcal{L}^{M+K}(a) = c/4\,.$$

By Fubini's theorem, we have

$$\begin{aligned}(c/4)\,\mathbf{M}(T) &= \int \int_{\widetilde{C}} |D\psi(x+a)|^M \, d\mathcal{L}^{M+K}(a)\, d\|T\|(x) \\ &= \int_{\widetilde{C}} \int |D\psi(x+a)|^M \, d\|T\|(x)\, d\mathcal{L}^{M+K}(a)\,.\end{aligned}$$

Set

$$G_1 = \left\{ a \in \widetilde{C} : \int |D\psi(x+a)|^M \, d\|T\|(x) \le c\,\mathbf{M}(T) \right\}\,,$$

$$H_1 = \widetilde{C} \setminus G_1 = \left\{ a \in \widetilde{C} : \int |D\psi(x+a)|^M \, d\|T\|(x) > c\,\mathbf{M}(T) \right\}\,.$$

We have

$$\int_{\widetilde{C}} \int |D\psi(x+a)|^M \, d\|T\|(x)\, d\mathcal{L}^{M+K}(a) \ge c\,\mathbf{M}(T)\,\mathcal{L}^{M+K}(H_1)\,,$$

so if $\mathcal{L}^{M+K}(H_1) \ge 1/3$ held, then we would have $(c/4)\,\mathbf{M}(T) \ge (c/3)\,\mathbf{M}(T)$. That is a contradiction. So we have $\mathcal{L}^{M+K}(H_1) < 1/3$ and $\mathcal{L}^{M+K}(G_1) \ge 2/3$.

A similar argument shows that

$$G_2 = \left\{ a \in \widetilde{C} : \int |D\psi(x+a)|^M \, d\|\partial T\|(x) \le c\,\mathbf{M}(\partial T) \right\}$$

satisfies $\mathcal{L}^{M+K}(G_2) \ge 2/3$.

We have

$$\begin{aligned}\mathcal{L}^{M+K}(G_1 \bigcap G_2) &= \mathcal{L}^{M+K}(G_1) + \mathcal{L}^{M+K}(G_2) - \mathcal{L}^{M+K}(G_1 \bigcup G_2) \\ &\geq \mathcal{L}^{M+K}(G_1) + \mathcal{L}^{M+K}(G_2) - \mathcal{L}^{M+K}(\widetilde{C}) \geq 1/3\,.\end{aligned}$$

Thus there exists $a \in G_1 \bigcap G_2$. Finally, we observe that

$$\int |D\psi(x)|^M \, d\|\boldsymbol{t}_{a\,\#}T\|(x) = \int |D\psi(x+a)|^M \, d\|T\|(x)$$

and

$$\int |D\psi(x)|^M \, d\|\partial \boldsymbol{t}_{a\,\#}T\|(x) = \int |D\psi(x+a)|^M \, d\|\partial T\|(x)$$

hold.

Step 3. Now we fix an $a \in \widetilde{C}$ as in Step 2 above and write $\widetilde{T} = \boldsymbol{t}_{a\,\#}T$. Applying the homotopy formula (see (7.22) in Section 7.4), we have

$$T = \widetilde{T} + \partial h_\#([\![\, (0,1) \,]\!] \times T) + h_\#([\![\, (0,1) \,]\!] \times \partial T)\,, \tag{7.57}$$

where h is the affine homotopy

$$h(t,x) = t\,x + (1-t)\psi(x)$$

between the identity map and $\boldsymbol{t}_a$. We have the estimates

$$\mathbf{M}[\, h_\#([\![\, (0,1) \,]\!] \times T)\,] \leq |a|\,\mathbf{M}(T)\,,$$

$$\mathbf{M}[\, h_\#([\![\, (0,1) \,]\!] \times \partial T)\,] \leq |a|\,\mathbf{M}(\partial T)\,.$$

We also have

$$\widetilde{T} = \psi_\#\,\widetilde{T} + \partial k_\#([\![\, (0,1) \,]\!] \times \widetilde{T}) + k_\#([\![\, (0,1) \,]\!] \times \partial \widetilde{T})\,, \tag{7.58}$$

where k is the affine homotopy

$$k(t,x) = t\,x + (1-t)\psi(x)$$

between the identity map and ψ. We note the estimates

$$\begin{aligned}\mathbf{M}[\, k_\#([\![\, (0,1) \,]\!] \times \widetilde{T})\,] &\leq 2^{-1}\,(M+K)^{1/2} \int |D\psi(x)|^M \, d\|\widetilde{T}\|(x) \\ &\leq 2^{-1}\,(M+K)^{1/2}\, c\,\mathbf{M}(T)\,, \\ \mathbf{M}[\, k_\#([\![\, (0,1) \,]\!] \times \partial\widetilde{T})\,] &\leq 2^{-1}\,(M+K)^{1/2} \int |D\psi(x)|^{M-1} \, d\|\partial\widetilde{T}\|(x) \\ &\leq 2^{-1}\,(M+K)^{1/2} \int |D\psi(x)|^{M} \, d\|\partial\widetilde{T}\|(x)\end{aligned}$$

$$\leq 2^{-1}\,(M+K)^{1/2}\,c\,\mathbf{M}(\partial T)\,,$$

$$\mathbf{M}\Big(\psi_{\#}\,\widetilde{T}\Big) \leq \int |D\psi(x)|^{M}\,d\|\widetilde{T}\|(x) \leq c\,\mathbf{M}(T)\,,$$

$$\mathbf{M}\Big(\psi_{\#}\,\partial\widetilde{T}\Big) \leq \int |D\psi(x)|^{M-1}\,d\|\partial\widetilde{T}\|(x)$$

$$\leq \int |D\psi(x)|^{M}\,d\|\partial\widetilde{T}\|(x) \leq c\,\mathbf{M}(\partial T)\,.$$

Combining (7.57) and (7.58), we have

$$T - \psi_{\#}\widetilde{T} = \partial\Big[\,h_{\#}([\![\ (0,1)\ ]\!] \times T) + k_{\#}([\![\ (0,1)\ ]\!] \times \widetilde{T})\,\Big]$$

$$+\,h_{\#}([\![\ (0,1)\ ]\!] \times \partial T) + k_{\#}([\![\ (0,1)\ ]\!] \times \partial\widetilde{T})\,.$$

We set

$$R = h_{\#}([\![\ (0,1)\ ]\!] \times T) + k_{\#}([\![\ (0,1)\ ]\!] \times \widetilde{T})$$

and

$$S_1 = h_{\#}([\![\ (0,1)\ ]\!] \times \partial T) + k_{\#}([\![\ (0,1)\ ]\!] \times \partial\widetilde{T})\,.$$

Note that R is integer-multiplicity if T is, and S_1 is integer-multiplicity if ∂T is. Also we have

$$\operatorname{spt} R \subset \left\{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{M+K}\right\},$$

$$\operatorname{spt} S_1 \subset \left\{x : \operatorname{dist}(x, \operatorname{spt} \partial T) < 2\sqrt{M+K}\right\}.$$

Step 4. While $\psi_{\#}\widetilde{T}$ is supported in L_M, it need not have the form

$$\sum_{F\in\mathcal{L}_M} p_F\,[\![\,F\,]\!]$$

required by (7.50). We will now show how $\psi_{\#}\widetilde{T}$ can be used to construct P as in (7.50).

Write $Q = \psi_{\#}\widetilde{T}$. We have

$$\operatorname{spt} Q \subset L_M\,. \tag{7.59}$$

Let F be one of the faces in L_M (that is to say, $F \in \mathcal{L}_M$) and let $\mathring{F}$ be the interior of F. Suppose that $F \subset \mathbb{R}^M \times \{0\} \subset \mathbb{R}^{M+K}$ and let p be orthogonal projection onto $\mathbb{R}^M \times \{0\}$. The construction of ψ tells us that $p \circ \psi = \psi$ in a neighborhood of any point $y \in \mathring{F}$. Thus we have that

$$p_{\#}(Q \llcorner \mathring{F}) = Q \llcorner \mathring{F}\,.$$

Identifying $\mathbb{R}^M \times \{0\}$ with $\mathbb{R}^M$ and applying Proposition 7.3.4, we obtain a function of bounded variation θ_F such that

$$\mathbf{M}(Q \llcorner \mathring{F}) = \int_{\mathring{F}} |\theta_F| \, d\mathcal{L}^M(x) \tag{7.60}$$

and

$$\mathbf{M}((\partial Q) \llcorner \mathring{F}) = \int_{\mathring{F}} |D\theta_F| \tag{7.61}$$

hold and such that

$$(Q \llcorner \mathring{F})(\omega) = \int_{\mathring{F}} \langle \, \omega(x), \ e_1 \wedge e_2 \wedge \cdots \wedge e_n \, \rangle \, \theta_F(x) \, d\mathcal{L}^M(x) \tag{7.62}$$

holds for all $\omega \in \mathcal{D}^M(\mathbb{R}^M)$.

In addition, by (7.62),

$$(Q \llcorner \mathring{F} - \beta[\![F]\!])(\omega) = \int_{\mathring{F}} (\theta_F - \beta) \, \langle \, \omega(x), \ e_1 \wedge \cdots \wedge e_M \, \rangle \, d\mathcal{L}^M(x) \, .$$

Thus, we have

$$\mathbf{M}(Q \llcorner \mathring{F} - \beta[\![F]\!]) = \int_{\mathring{F}} |\theta_F - \beta| \, d\mathcal{L}^M(x) \, , \tag{7.63}$$

$$\mathbf{M}(\partial(Q \llcorner \mathring{F} - \beta[\![F]\!])) = \int_{\mathbb{R}^M} |D(\chi_{\mathring{F}} \, (\theta_F - \beta))| \, . \tag{7.64}$$

Now let us take $\beta = \beta_F$ such that

$$\min\left\{ \mathcal{L}^M\{x \in \mathring{F} : \theta_F(x) \geq \beta\}, \ \mathcal{L}^M\{x \in \mathring{F} : \theta_F(x) \leq \beta\} \right\} \geq \frac{1}{2} \, .$$

Note that we can do this because $\mathcal{L}^M(\mathring{F}) = 1$. Also we may take $\beta_F \in \mathbb{Z}$ whenever θ_F is integer-valued.

We have now, by Theorem 5.5.6, Theorem 5.5.7, (7.60), (7.61), (7.63), and (7.64), that

$$\mathbf{M}(Q \llcorner \mathring{F} - \beta[\![F]\!]) \leq c \int_{\mathring{F}} |D\theta_F| = c \, \mathbf{M}(\partial Q \llcorner \mathring{F}) \, , \tag{7.65}$$

$$\mathbf{M}(\partial(Q \llcorner \mathring{F} - \beta[\![F]\!])) \leq c \int_{\mathring{F}} |D\theta_F| = c \, \mathbf{M}(\partial Q \llcorner \mathring{F}) \, . \tag{7.66}$$

It is also the case that

$$Q \llcorner \partial F = 0 \, . \tag{7.67}$$

Now, summing over $F \in \mathcal{L}_M$ and using (7.65), (7.66), and (7.67), with $P = \sum_{F \in \mathcal{L}_M} \beta_F [\![F]\!]$, we see that

$$\mathbf{M}(Q - P) \leq c\mathbf{M}(\partial Q)\,, \tag{7.68}$$

$$\mathbf{M}(\partial Q - \partial P) \leq c\mathbf{M}(\partial Q)\,. \tag{7.69}$$

Actually our choice of β_F tells us that

$$|\beta_F| \leq 2 \int_{\mathring{F}} |\theta_F|\, d\mathcal{L}^M(x)\,.$$

Thus, again using (7.63), and since $\mathbf{M}(P) = \sum_F |\beta_F|$, we see that

$$\mathbf{M}(P) \leq c\,\mathbf{M}(Q)\,. \tag{7.70}$$

We also know, from (7.69) above (and the triangle inequality), that

$$\mathbf{M}(\partial P) \leq c\,\mathbf{M}(\partial Q)\,. \tag{7.71}$$

Finally, we write

$$T - P = \partial R + S\,, \tag{7.72}$$

where $S = S_1 + (Q - P)$, and the deformation theorem follows. □

7.9 Applications of the Deformation Theorem

There are some immediate applications of the deformation theorem that amply illustrate the power of the theorem. These are:

- the isoperimetric theorem;
- the weak polyhedral approximation theorem;
- the boundary rectifiability theorem.

Theorem 7.9.1 (Isoperimetric Inequality). *Let $M \geq 2$. Suppose that $T \in \mathcal{D}_{M-1}(\mathbb{R}^{M+K})$ is integer-multiplicity. Assume that* spt T *is compact and that $\partial T = 0$. Then there is a compactly supported, integer-multiplicity current $R \in \mathcal{D}_M(\mathbb{R}^{M+K})$ such that $\partial R = T$ and*

$$[\,\mathbf{M}(R)\,]^{(M-1)/M} \leq c\,\mathbf{M}(T)\,.$$

Here the constant c depends on M and K.

The theorem deserves some commentary. In its most classical formulation, the current T is a current of integration on a simple, closed curve γ in $\mathbb{R}^2$. Of course, the mass of T is then its length. The current R is then a 2-dimensional current (i.e., a *region* in the plane) whose boundary is T. And the conclusion of the theorem is then that the square root of the area of R is majorized by a constant times the mass of T. We know, both intuitively and because of the classical isoperimetric theorem, that the extremal curve T—that is, the curve that encloses the largest area for a given perimeter (see Figure 7.9)—is the circle. Let us say that that extremal curve is a circle of radius r. Its mass is $2\pi r$. The region inside this curve is a disk of radius r, and its mass is πr^2. In this situation the asserted inequality is obvious with constant $c = 1/(2\sqrt{\pi})$.

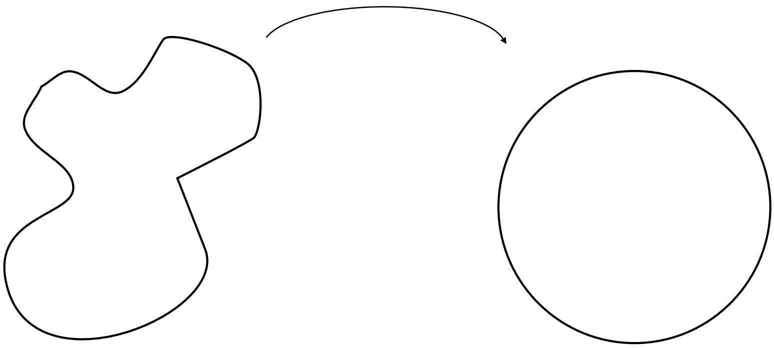

Fig. 7.9. The isoperimetric inequality.

Proof of the Theorem. The case $T = 0$ is trivial, so let us assume that $T \neq 0$. Let P, R, S be integer-multiplicity currents as in Theorem 7.7.2, the scaled version of the deformation theorem, applied with M replaced by $M-1$ and with K replaced by $K+1$. For the moment, $\rho > 0$ is arbitrary; observe also that $S = 0$ because $\partial T = 0$.

Clearly, because

$$\mathbf{M}(\boldsymbol{\eta}_{\rho\#}[\![F]\!]) = \mathcal{H}^{M-1}[\boldsymbol{\eta}(F)] = \rho^{M-1}$$

for all $F \in \mathcal{F}_{M-1}$, we know that

$$\mathbf{M}(P) = N(\rho)\,\rho^{M-1}$$

for some nonnegative integer $N(\rho)$. Theorem 7.7.2 tells us that $\mathbf{M}(P) \leq c\,\mathbf{M}(T)$. If we take

$$\rho = [\,2\,c\,\mathbf{M}(T)\,]^{1/(M-1)}\,, \tag{7.73}$$

then we have

$$N(\rho)\,2\,c\,\mathbf{M}(T) = N(\rho)\,\rho^{M-1} = \mathbf{M}(P) \leq c\,\mathbf{M}(T)\,,$$

so $2\,N(\rho) \leq 1$, implying that $N(\rho) = 0$.

Choosing ρ as in (7.73), we have $P = 0$. Theorem 7.7.1 now tells us that $T = \partial R$ for the compactly supported, integer-multiplicity current R and we have

$$\mathbf{M}(R) \leq c\,\rho\,\mathbf{M}(T) = 2^{1/(M-1)}\,c^{M/(M-1)}\,[\,\mathbf{M}(T)\,]^{M/(M-1)}\,. \qquad \square$$

Theorem 7.9.2 (Weak Polyhedral Approximation). *Let $T \in \mathcal{D}_M(U)$ be any integer-multiplicity current with $\mathbf{M}_W(\partial T) < \infty$ for all $W \subset\subset U$. Then there is a sequence $\{P_\ell\}$ of currents of the form*

$$P_\ell = \sum_{F \in \mathcal{F}_M} p_F^{(\ell)}\,\boldsymbol{\eta}_{\rho_\ell\#}[\![F]\!]\,, \tag{7.74}$$

with $p_F^{(\ell)} \in \mathbb{Z}$, $\rho_\ell \downarrow 0$ and with P_ℓ converging weakly to T (so ∂P_ℓ also converges weakly to ∂T) in U.

Proof. First consider the case $U = \mathbb{R}^{M+K}$ and $\mathbf{M}(T) < \infty$, $\mathbf{M}(\partial T) < \infty$. Now we just use the deformation theorem directly: For any sequence $\rho_\ell \downarrow 0$, Theorem 7.7.1, the scaled version of the deformation theorem, applied with $\rho = \rho_\ell$, yields P_ℓ as in (7.74) such that

$$T - P_\ell = \partial R_\ell + S_\ell$$

for some R_ℓ, S_ℓ such that

$$\mathbf{M}(R_\ell) \leq c\,\rho_\ell\,\mathbf{M}(T) \to 0\,,$$

$$\mathbf{M}(S_\ell) \leq c\,\rho_\ell\,\mathbf{M}(\partial T) \to 0\,,$$

and

$$\mathbf{M}(P_\ell) \leq c\,\mathbf{M}(T) \qquad \text{and} \qquad \mathbf{M}(\partial P_\ell) \leq c\,\mathbf{M}(\partial T)\,.$$

Clearly the last three lines give $P_\ell(\omega) \to T_\ell(\omega)$ for all $\omega \in \mathcal{D}^M(\mathbb{R}^{M+K})$. Also $\partial P_\ell = 0$ if $\partial T = 0$. Hence the theorem is established if $U = \mathbb{R}^{M+K}$ and T, ∂T are of finite mass.

For the general case, let us take any Lipschitz function ϕ on $\mathbb{R}^{M+K}$ such that $\phi > 0$ in U and $\phi = 0$ on $\mathbb{R}^{M+K} \setminus U$. We further assume that $\{x = \phi(x) > \lambda\} \subset\subset U$ for all $\lambda > 0$. For $\mathcal{L}^1$-almost every $\lambda > 0$, Lemma 7.6.3 implies that $T_\lambda \equiv T \llcorner \{x : \phi(x) > \lambda\}$ is such that $\mathbf{M}(\partial T_\lambda) < \infty$. Since $\operatorname{spt} T_\lambda \subset\subset U$, we can use the above argument to approximate T_λ for any such λ. Then, for a suitable sequence $\lambda_j \downarrow 0$, the required approximation is an immediate consequence. □

Theorem 7.9.3 (Boundary Rectifiability). *Let the integer-multiplicity current $T \in \mathcal{D}_M$ be such that $\mathbf{M}_W(\partial T) < \infty$ for all $W \subset\subset U$. Then ∂T, which is an element of $\mathcal{D}_{M-1}(U)$, is also an integer-multiplicity current.*

Proof. This is a direct consequence of the last theorem and of the compactness theorem, Theorem 7.5.2, applied to integer-multiplicity currents of dimension $M - 1$. □

Remark 7.9.4. The compactness theorem is not proved until Section 8.1. We will see there that the proof of the compactness theorem for integer-multiplicity currents of dimension M uses the boundary rectifiability theorem for currents of dimension $M - 1$. So logically the compactness theorem and boundary rectifiability theorem are proved together in an induction that begins with the compactness theorem for integer-multiplicity currents of dimension 0.

8

Currents and the Calculus of Variations

8.1 Proof of the Compactness Theorem

First let us recall the statement of the compactness theorem, Theorem 7.5.2:

The Compactness Theorem for Integer-Multiplicity Currents. *Let* $\{T_j\} \subseteq \mathcal{D}_M(U)$ *be a sequence of integer-multiplicity currents such that*

$$\sup_{j \geq 1} \Big[\mathbf{M}_W(T_j) + \mathbf{M}_W(\partial T_j) \Big] < \infty \quad \textit{for all } W \subset\subset U \, .$$

Then there is an integer-multiplicity current $T \in \mathcal{D}_M(U)$ *and a subsequence* $\{T_{j'}\}$ *such that* $T_{j'} \to T$ *weakly in* U.

Logically the compactness theorem and boundary rectifiability theorem are proved in tandem by induction on M, the dimension of the currents. The induction begins with the straightforward proof of the compactness theorem in the case $M = 0$. That proof is given in the next subsection.

The inductive step is then in two parts. First it is shown that the boundary rectifiability theorem is valid. Note that the boundary rectifiability theorem is vacuous when $M = 0$. In Section 7.9, we showed that when $M \geq 1$, the boundary rectifiability theorem is an easy consequence of the compactness theorem for currents of dimension $M - 1$.

The second part of the induction step is to prove the compactness theorem for dimension M assuming the boundary rectifiability theorem for dimension M and the compactness theorem for dimension $M - 1$. The strategy for this part of the proof is to use slicing to convert a sequence of weakly convergent M-dimensional integer-multiplicity currents into a sequence of functions that take their values in the space of 0-dimensional integer-multiplicity currents. These functions are of bounded variation in an appropriate sense. We then analyze the behavior of the graphs of such functions of bounded variation to understand the structure of the limit M-dimensional current.

To carry out this program we must study the 0-dimensional integer-multiplicity currents in some detail and we must define and investigate the appropriate space of functions of bounded variation.

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_8,

8.1.1 Integer-Multiplicity 0-Currents

Notation 8.1.1.

(1) Let $\mathcal{R}_0(\mathbb{R}^{M+K})$ denote the space of finite-mass integer-multiplicity 0-currents in $\mathbb{R}^{M+K}$.
(2) By (7.28), a nonzero current T in $\mathcal{R}_0(\mathbb{R}^{M+K})$ can be written

$$T = \sum_{j=1}^{\alpha} c_j\, \boldsymbol{\delta}_{p_j}\,, \tag{8.1}$$

where α is a positive integer, $p_j \in \mathbb{R}^{M+K}$ for each $1 \le j \le \alpha$, $p_i \ne p_j$ for $1 \le i \ne j \le \alpha$, $\boldsymbol{\delta}_{p_j}$ is the Dirac mass at p_j, and $c_j \in \mathbb{Z} \setminus \{0\}$ for each $1 \le j \le \alpha$.

Proof of the Compactness Theorem for Integer-Multiplicity Currents of Dimension 0. Suppose that $T_j \in \mathcal{R}_0(\mathbb{R}^{M+K})$, $j = 1, 2, \ldots$, and that

$$L = \sup_{j \ge 1} \mathbf{M}(T_j) < \infty\,.$$

By the Banach–Alaoglu theorem there is a $T \in \mathcal{D}_0(\mathbb{R}^{M+k})$ such that a subsequence of the T_j converges weakly to T. For simplicity, we will not change notation. Instead we will suppose that the original sequence T_j converges weakly to T. What we must prove is that $T \in \mathcal{R}_0(\mathbb{R}^{M+K})$.

Consider $0 < m < \infty$ chosen large enough that $T \llcorner \mathbb{B}(0, m) \ne 0$. We can write each $T_j \llcorner \overline{\mathbb{B}}(0, m) \in \mathcal{R}_0(\mathbb{R}^{M+K})$ as

$$T_j \llcorner \overline{\mathbb{B}}(0, m) = \sum_{i=1}^{L} c_i^{(j)}\, \boldsymbol{\delta}_{p_i^{(j)}}\,,$$

where

$$c_i^{(j)} \in \mathbb{Z}\,, \qquad -L \le c_i^{(j)} \le L\,, \qquad p_i^{(j)} \in \overline{\mathbb{B}}(0, m)\,.$$

We now allow $c_i^{(j)} = 0$ because it is possible that $\mathbf{M}[\, T_j \llcorner \overline{\mathbb{B}}(0, m)\,] < L$ holds.

By the Bolzano–Weierstrass theorem, we can pass to a subsequence—but again we will not change notation—so that for $j = 1, 2, \ldots, L$, $c_i^{(j)} \to c_i \in \mathbb{Z}$ and $p_i^{(j)} \to p_i \in \overline{\mathbb{B}}(0, m)$ as $j \to \infty$.

If $\phi \in \mathcal{D}^0(\mathbb{R}^{M+K})$ with $\operatorname{supp} \phi \subseteq \mathbb{B}(0, m)$, then we have

$$T_j(\phi) = T_j \llcorner \overline{\mathbb{B}}(0, m)(\phi) \to \sum_{i=1}^{L} c_i\, \phi(p_i)$$

and we have $T_j(\phi) \to T(\phi)$ because T_j converges weakly to T. Thus we can write

$$T \llcorner \mathbb{B}(0, m) = \sum_{i=1}^{\alpha} c_i\, \boldsymbol{\delta}_{p_i}\,,$$

where (by renaming if necessary) we can suppose that $\alpha \le L$ is a positive integer, $p_i \in \mathbb{B}(0, m)$ for each $1 \le i \le \alpha$, $p_h \ne p_i$ for $1 \le h \ne i \le \alpha$, and $c_i \in \mathbb{Z} \setminus \{0\}$ for each $1 \le i \le \alpha$. Since $\mathbf{M}(T) \le L < \infty$, we see that in fact we can choose m large enough that $T = T \llcorner \mathbb{B}(0, m)$. □

Notation 8.1.2.

(1) Equation (8.1) tells us that, for $\phi \in \mathcal{D}^0(\mathbb{R}^{M+K})$,

$$T(\phi) = \sum_{j=1}^{\alpha} c_j \, \phi(p_j) \,. \tag{8.2}$$

We can also use (8.2) to define $T(\phi)$ when ϕ is merely continuous.

(2) We will use the metric d_0 on $\mathcal{R}_0(\mathbb{R}^{M+K})$ defined by

$$\begin{aligned} &\mathrm{d}_0(T_1, T_2) \\ &\quad = \sup\{\, (T_1 - T_2)(\phi) : \ \phi \text{ is Lipschitz, } \|\phi\|_\infty \le 1, \ \|d\phi\|_\infty \le 1 \,\} \,. \end{aligned}$$

(3) We let $\mathcal{F}^{M+K}$ denote the space of nonempty finite subsets of $\mathbb{R}^{M+K}$ metrized by the Hausdorff distance. The Hausdorff distance is defined in Section 1.6. The Hausdorff distance between A and B is denoted by $\mathrm{HD}\,(A, B)$.

(4) Define

$$\varrho : \mathcal{R}_0(\mathbb{R}^{M+K}) \to \overline{\mathbb{R}}$$

by

$$\varrho(T) = \inf\{\, |p - q| : p, q \in \operatorname{spt}(T), \ p \ne q \,\} \,.$$

Note that if either $T = 0$ or $\operatorname{card}[\operatorname{spt}(T)] = 1$, then $\varrho(T) = +\infty$.

Lemma 8.1.3. *If $T_j \in \mathcal{R}_0(\mathbb{R}^{M+K})$ and $T_j \to T \in \mathcal{R}_0(\mathbb{R}^{M+K})$ weakly as $j \to \infty$, then*

$$\operatorname{card}[\operatorname{spt}(T)] \le \liminf_{j\to\infty} \operatorname{card}[\operatorname{spt}(T_j)] \,.$$

If additionally

$$\operatorname{card}[\operatorname{spt}(T)] = \operatorname{card}[\operatorname{spt}(T_j)], \quad j = 1, 2, \ldots ,$$

then

$$\varrho(T) = \lim_{j\to\infty} \varrho(T_j) \,.$$

Proof. For each $p \in \operatorname{spt}(T)$ we can find $\phi_p \in \mathcal{D}^0(\mathbb{R}^{M+K})$ for which $\phi_p(p) = 1$, $\phi_p(x) < 1$ for $x \ne p$, and $\phi_p(q) = 0$ for $q \in \operatorname{spt}(T)$ with $q \ne p$. The existence of such a function ϕ_p implies that p is a limit point of any set of the form $\bigcup_{i \ge I} \operatorname{spt}[\, T_{j_i} \,]$, and the result follows. □

The proof of the next lemma is elementary, but we treat it in detail because the result is so essential to proving the compactness theorem.

Lemma 8.1.4. *If* $T, \widetilde{T} \in \mathcal{R}_0(\mathbb{R}^{M+K})$ *satisfy* $0 < \mathbf{M}(T) = \mathbf{M}(\widetilde{T})$*, then it holds that*

$$\min\left\{1,\ (1/3)\,\varrho(T),\ \mathrm{HD}[\,\mathrm{spt}\,(T),\ \mathrm{spt}\,(\widetilde{T})\,]\right\} \le \mathrm{d}_0(T, \widetilde{T})\,.$$

Proof. Write $T = \sum_{j=1}^{\alpha} c_j\, \boldsymbol{\delta}_{p_j}$ as in (8.1), and write $\widetilde{T} = \sum_{q \in \mathrm{spt}\,(\widetilde{T})} \gamma_q\, \boldsymbol{\delta}_q$. Set

$$r = \min\left\{1,\ (1/3)\,\varrho(T)\right\}.$$

We may assume that $\mathrm{d}_0(T, \widetilde{T}) < r$.

Because $\mathbf{M}(T) = \mathbf{M}(\widetilde{T})$ holds, we have

$$\sum_{j=1}^{\alpha} |c_j| = \sum_{q \in \mathrm{spt}\,(\widetilde{T})} |\gamma_q|\,. \tag{8.3}$$

For $j = 1, 2, \ldots, \alpha$, define ϕ_j by setting

$$\phi_j(x) = \begin{cases} \mathrm{sgn}(c_j) \cdot [\,r - |x - p_j|\,] & \text{if } |x - p_j| < r, \\ 0 & \text{if } |x - p_j| \ge r. \end{cases}$$

Since $|\phi_j| \le r_T \le 1$ and $|d\phi_j| \le 1$ hold, we have $(T - \widetilde{T})(\phi_j) \le \mathrm{d}_0(T, \widetilde{T})$.

If there were $1 \le j \le \alpha$ for which $\mathrm{spt}\,(\widetilde{T}) \bigcap \mathbb{B}(p_j, r) = \emptyset$ held, then we would have

$$\mathrm{d}_0(T, \widetilde{T}) \ge (T - \widetilde{T})(\phi_j) = T(\phi_j) = r\,|c_j| \ge r\,,$$

contradicting the assumption that $\mathrm{d}_0(T, \widetilde{T}) < r$ holds. We conclude that

$$\mathrm{spt}\,(\widetilde{T}) \bigcap \mathbb{B}(p_j, r) \ne \emptyset\,, \quad \text{for } j = 1, 2, \ldots, \alpha. \tag{8.4}$$

Now define $\phi = \sum_{j=1}^{\alpha} \phi_j$. Since the ϕ_j have disjoint supports, we see that $|\phi| \le r_T \le 1$ and $|d\phi| \le 1$ hold. Setting

$$A_j = \mathrm{spt}\,(\widetilde{T}) \bigcap \mathbb{B}(p_j, r)\,, \qquad B = \mathrm{spt}\,(\widetilde{T}) \setminus \bigcup_{j=1}^{\alpha} A_j\,,$$

and using (8.3), we have

$$\begin{aligned} \mathrm{d}_0(T, \widetilde{T}) &\ge (T - \widetilde{T})(\phi) = T(\phi) - \widetilde{T}(\phi) \\ &= r \sum_{j=1}^{\alpha} |c_j| - \sum_{j=1}^{\alpha} \sum_{q \in A_j} \mathrm{sgn}(c_j)\,[\,r - |q - p_j|\,]\,\gamma_q \\ &= r \sum_{q \in \mathrm{spt}\,(\widetilde{T})} |\gamma_q| - \sum_{j=1}^{\alpha} \sum_{q \in A_j} \mathrm{sgn}(c_j)\,[\,r - |q - p_j|\,]\,\gamma_q \\ &= \sum_{q \in B} r\,|\gamma_q| + \sum_{j=1}^{\alpha} \sum_{q \in A_j} \Big(r\,|\gamma_q| - \mathrm{sgn}(c_j)\,[\,r - |q - p_j|\,]\,\gamma_q \Big). \end{aligned} \tag{8.5}$$

Note that every summand in (8.5) is nonnegative.

If there existed $q \in B$, then we would have

$$\mathrm{d}_0(T, \widetilde{T}) \geq r\,|\gamma_q| \geq r\,,$$

contradicting the assumption that $\mathrm{d}_0(T, \widetilde{T}) < r$ holds. We conclude that

$$\mathrm{spt}\,(\widetilde{T}) \subseteq \textstyle\bigcup_{j=1}^{\alpha} \mathbb{B}(p_j, r)\,. \tag{8.6}$$

Now we consider $q_* \in \mathrm{spt}\,(\widetilde{T})$ and $1 \leq j_* \leq \alpha$ such that $q_* \in A_{j_*}$. Looking only at the summand in (8.5) that corresponds to j_* and q_*, we see that

$$\mathrm{d}_0(T, \widetilde{T}) \geq r\,|\gamma_{q_*}| - \mathrm{sgn}(c_{j_*})\,[\,r - |q_* - p_{j_*}|\,]\,\gamma_{q_*} \tag{8.7}$$

holds.

In assessing the significance of (8.7) there are two cases to be considered according to the sign of $c_{j_*}\gamma_{q_*}$.

Case 1. In case $\mathrm{sgn}(c_{j_*}\,\gamma_{q_*}) = -1$ holds, we have

$$\mathrm{sgn}(c_{j_*})\,\gamma_{q_*} = \mathrm{sgn}(c_{j_*})\,\mathrm{sgn}(\gamma_{q_*})\,|\gamma_{q_*}| = \mathrm{sgn}(c_{j_*}\,\gamma_{q_*})\,|\gamma_{q_*}| = -|\gamma_{q_*}|\,.$$

The fact that $\mathrm{sgn}(c_{j_*})\,\gamma_{q_*} = -|\gamma_{q_*}|$ holds implies

$$\begin{aligned}\mathrm{d}_0(T, \widetilde{T}) &\geq r\,|\gamma_q| - \mathrm{sgn}(c_j)\,[\,r - |q - p_j|\,]\,\gamma_q \\ &= (r + r - |q_* - p_{j_*}|)\,|\gamma_{q_*}| \geq r\,,\end{aligned}$$

and this last inequality contradicts the assumption that $\mathrm{d}_0(T, \widetilde{T}) < r$.

Case 2. Because of the contradiction obtained in the last paragraph, we see that $\mathrm{sgn}(c_{j_*}\,\gamma_{q_*}) = +1$ must hold. Consequently we have $\mathrm{sgn}(c_{j_*})\,\gamma_{q_*} = |\gamma_{q_*}|$, which implies that

$$\mathrm{d}_0(T, \widetilde{T}) \geq (r - r + |q - p_{j_*}|)\,|\gamma_{q_*}| \geq |q_* - p_{j_*}|\,.$$

By (8.6), for $q_* \in \mathrm{spt}\,(\widetilde{T})$, there exists j_* such that $q_* \in A_{j_*}$. Similarly, by (8.4), for $1 \leq j_* \leq \alpha$, there exists $q_* \in \mathrm{spt}\,(\widetilde{T})$ such that $q_* \in A_{j_*}$. Thus we conclude that $\mathrm{d}_0(T, \widetilde{T}) \geq \mathrm{HD}\,[\,\mathrm{spt}\,(T),\ \mathrm{spt}\,(\widetilde{T})\,]$. □

Theorem 8.1.5.

(1) *If* $A \subseteq \mathbb{R}^M$ *and* $f : A \to \mathcal{F}^{M+K}$ *is a Lipschitz function, then*

$$\bigcup_{x \in A} f(x) \tag{8.8}$$

is a countably M*-rectifiable subset of* $\mathbb{R}^{M+K}$.

(2) *If $A \subseteq \mathbb{R}^M$ and $g : A \to \mathcal{R}_0(\mathbb{R}^{M+K})$ is a Lipschitz function, then*

$$\bigcup_{x \in A} \operatorname{spt} [\, g(x) \,] \tag{8.9}$$

is a countably M-rectifiable subset of $\mathbb{R}^{M+K}$.

Proof.
(1) Let m be a Lipschitz bound for f. Then 1 will be a Lipschitz bound for $f(x/m)$. Thus, without loss of generality, we may suppose that 1 is a Lipschitz bound for f.

In this proof, we will need to consider open balls in both $\mathbb{R}^M$ and in $\mathbb{R}^{M+K}$. Accordingly, we will use the notation $\mathbb{B}^M(x, r)$ for the open ball in $\mathbb{R}^M$ and $\mathbb{B}^{M+K}(x, r)$ for the open ball in $\mathbb{R}^{M+K}$.

For $\ell = 1, 2, \ldots$, set $A_\ell = \{ x \in A : \operatorname{card}[\, f(x) \,] = \ell \}$. Note that $\bigcup_{x \in A_1} f(x)$ is the image of the Lipschitz function $u : A_1 \to \mathbb{R}^{M+K}$ defined by requiring $f(x) = \{u(x)\}$.

Now consider $\ell \geq 2$ and $x \in A_\ell$. Write $f(x) = \{ p_1, p_2, \ldots, p_\ell \}$ and set $r(x) = \min_{i \neq j} |p_i - p_j|$.

If $z \in A_\ell \bigcap \mathbb{B}^M(x, r(x)/4)$, then for each $i = 1, 2, \ldots, \ell$ there is a unique $q \in f(z) \bigcap \mathbb{B}^{M+K}(p_i, r(x)/4)$ and we define $u_i(z) = q$.

The functions $u_1, u_2, \ldots, u_\ell$ are Lipschitz because, for

$$z_1, z_2 \in A_\ell \bigcap \mathbb{B}^M(x, r(x)/4)\,,$$

we have

$$\operatorname{HD}[\, f(z_1),\ f(z_2) \,] = \max\{\, |u_i(z_1) - u_i(z_2)| : i = 1, 2, \ldots, \ell \,\}\,.$$

Since

$$\bigcup_{z \in A_\ell \cap \mathbb{B}^M(x, r(x)/4)} f(z) = \bigcup_{i=1}^{\ell} \left\{ u_i(z) : z \in A_\ell \bigcap \mathbb{B}^M(x, r(x)/4) \right\},$$

we see that $\bigcup_{z \in A_\ell \cap \mathbb{B}^M(x, r(x)/4)} f(z)$ is a countably M-rectifiable subset of $\mathbb{R}^{M+K}$.

As a subspace of a second countable space, A_ℓ is second countable, so it has the Lindelöf[1] property; that is, every open cover has a countable subcover. Thus there is a countable cover of A_ℓ by sets of the form $A_\ell \bigcap \mathbb{B}^M(x, r(x)/4)$, $x \in A_\ell$. We conclude that $\bigcup_{z \in A_\ell} f(z)$ is a countably M-rectifiable subset of $\mathbb{R}^{M+K}$ and hence $\bigcup_{\ell=1}^{\infty} \bigcup_{z \in A_\ell} f(z)$ is also countably M-rectifiable.

(2) Without loss of generality, suppose that 1 is a Lipschitz bound for g. For i and j positive integers, set

$$A_{i,j} = \{\, x \in A : \mathbf{M}[\, g(x) \,] = j \ \text{ and } \ 2^{-i} < r_{g(x)} \,\}\,,$$

[1] Ernst Leonard Lindelöf (1870–1946).

where

$$r_{g(x)} = \min\left\{ 1,\ (1/3)\,\rho[\,g(x)\,] \right\}.$$

Fix $x \in A_{i,j}$. For $z_1, z_2 \in A_{i,j} \bigcap \mathbb{B}(x, 2^{-i-1})$, we have

$$\mathbf{M}[\,g(z_1)\,] = \mathbf{M}[\,g(z_2)\,] = j \ \text{ and } \ \mathrm{d}_0[\,g(z_1),\ g(z_2)\,] < 2^{-i} < r_{g(z_1)}\,.$$

So, by Lemma 8.1.4, HD [spt $(g(z_1))$, spt $(g(z_2))$] $\le \mathrm{d}_0[\,g(z_1),\ g(z_2)\,]$ holds. Thus

$$f : A_{i,j} \bigcap \mathbb{B}(x, 2^{-i-1}) \to \mathcal{F}^{M+K}$$

defined by $f(z) = \operatorname{spt}[\,g(z)\,]$ is Lipschitz. By part (1) we conclude that

$$\bigcup_{z \in A_{i,j} \cap \mathbb{B}(x, 2^{-i-1})} \operatorname{spt}[\,g(z)\,] \tag{8.10}$$

is a countably M-rectifiable subset of $\mathbb{R}^{M+K}$. As in the proof of (1), we observe that $A_{i,j}$ has the Lindelöf property, and so the result follows. □

8.1.2 A Rectifiability Criterion for Currents

The next theorem provides a criterion for guaranteeing that a current is an integer-multiplicity rectifiable current. Later we shall use this criterion to complete the proof of the compactness theorem.

Theorem 8.1.6 (Rectifiability Criterion). *If $T \in \mathcal{D}_M(\mathbb{R}^{M+K})$ satisfies the following conditions:*

(1) $\mathbf{M}(T) + \mathbf{M}(\partial T) < \infty$,
(2) $\|T\| = \mathcal{H}^M \llcorner \theta$, *where θ is integer-valued and nonnegative,*
(3) $\{\,x : \theta(x) > 0\,\}$ *is a countably M-rectifiable set,*

then T is an integer-multiplicity rectifiable current.

Proof. Set $S = \{\,x : \theta(x) > 0\,\}$. We need to show that for $\mathcal{H}^M$-almost every point in S, $\overrightarrow{T}(x) = v_1 \wedge \cdots \wedge v_M$, where $v_1, \ldots, v_M$ is an orthonormal system parallel to $\mathbf{T}_x S$.

Of course, $\mathcal{H}^M$-almost every point x of S is a Lebesgue point of θ and is a point where $\overrightarrow{T}(x)$ and $\mathbf{T}_x S$ both exist. Also, by Theorem 4.3.7, we see that $\Theta^{*M}(\|\partial T\|, x) < \infty$ holds for $\mathcal{H}^M$-almost every $x \in S$. Hence $\Theta^{M-1}(\|\partial T\|, x) = 0$ also holds for $\mathcal{H}^M$-almost every $x \in S$. Let us consider such a point and, for convenience of notation, suppose that $x = 0$. Consider a sequence $r_i \downarrow 0$. Passing to a subsequence if necessary, but without changing notation, we may suppose that $\boldsymbol{\eta}_{r_i\,\#}T$ and $\boldsymbol{\eta}_{r_i\,\#}\partial T$ converge weakly to R and ∂R, respectively. Here $\boldsymbol{\eta}_r : \mathbb{R}^{M+K} \to \mathbb{R}^{M+K}$ is given by $\boldsymbol{\eta}_r(z) = r^{-1} z$. Then we have $\overrightarrow{R}(0) = \overrightarrow{T}(0)$, $\partial R = 0$, and $\operatorname{spt} R \subseteq \mathbf{T}_0 S$. By Proposition 7.3.5 (a variant of the constancy theorem), we have $\overrightarrow{R}(x) = v_1 \wedge \cdots \wedge v_M$, where $v_1, \ldots, v_M$ is an orthonormal system parallel to $\mathbf{T}_0 S$. □

8.1.3 MBV Functions

In this subsection, we introduce a class of metric-space-valued functions of bounded variation. The notion of metric-space-valued functions of bounded variation was introduced in [Amb 90] and applied to currents in [AK 00].

Definition 8.1.7.

(1) A function $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ can be written

$$u(y) = \sum_{i=1}^{\infty} c_i(y)\, \boldsymbol{\delta}_{p_i(y)}\,, \tag{8.11}$$

where only finitely many $c_i(y)$ are nonzero, for any $y \in \mathbb{R}^M$.

(2) If u is as in (8.11) and and $\phi : \mathbb{R}^{M+K} \to \mathbb{R}$, then we define $u \diamond \phi : \mathbb{R}^M \to \mathbb{R}$ by setting

$$(u \diamond \phi)(y) = \sum_{i=1}^{\infty} c_i(y)\, \phi\Big[\, p_i(y)\,\Big], \tag{8.12}$$

for $y \in \mathbb{R}^M$; thus the value of $(u \diamond \phi)(y)$ is the result of applying the 0-current $u(y)$ to the function ϕ. We use the notation $\diamond$ in analogy with the notation $\circ$ for composition.

(3) A Borel function $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ is a *metric-space-valued function of bounded variation* if for every bounded Lipschitz function $\phi : \mathbb{R}^{M+K} \to \mathbb{R}$, the function $u \diamond \phi$ is locally BV in the traditional sense (see for instance [KPk 99, Section 3.6]). We will abbreviate "u is a metric-space-valued function of bounded variation" to simply "u is *MBV*."

(4) If $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ is MBV, then we denote the *total variation measure* of u by V_u and define it by

$$
\begin{aligned}
(\mathrm{V}_u)(A) &= \sup\left\{ \int_A |D(u \diamond \phi)| : \phi : \mathbb{R}^{M+K} \to \mathbb{R},\ |\phi| \le 1,\ |d\phi| \le 1 \right\} \\
&= \sup\left\{ \int (u \diamond \phi)\, \mathrm{div} g\, d\mathcal{L}^M : \mathrm{supp}\, g \subseteq A,\ |g| \le 1,\ |\phi| \le 1,\ |d\phi| \le 1 \right\}
\end{aligned}
$$

for $A \subseteq \mathbb{R}^M$ open.

For us the most important example of an MBV function will be provided by slicing a current. That is the content of the next proposition.

Proposition 8.1.8. *Let* $\mathbf{p} : \mathbb{R}^{M+K} = \mathbb{R}^M \times \mathbb{R}^K \to \mathbb{R}^M$ *be projection onto the first factor. If* $T \in \mathcal{D}_M(\mathbb{R}^{M+K})$ *is an integer-multiplicity current with* $\mathbf{M}(T) + \mathbf{M}(\partial T) < \infty$*, then* $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ *defined by*

$$u(x) = \langle T, \mathbf{p}, x \rangle$$

is MBV and

$$\mathrm{V}_u(A) \le M\Big[\, \|\partial T\|(A) + \|T\|(A) \Big]$$

holds, for each open set $A \subseteq \mathbb{R}^M$.

Proof. Fix an open set $A \subseteq \mathbb{R}^M$. Let $g \in C^1(\mathbb{R}^M, \mathbb{R}^M)$ satisfy $|g| \le 1$ and $\operatorname{supp} g \subseteq A$. Let $\phi : \mathbb{R}^{M+K} \to \mathbb{R}$ be such that $|\phi| \le 1$ and $|d\phi| \le 1$.

Pick i with $1 \le i \le M$ and set

$$\psi = g_i\,, \qquad dx_{\widehat{\imath}} = dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_M\,.$$

Using Proposition 7.6.5(2), we estimate

$$\begin{aligned}
&\left| \int D_{x_i}\psi \,\langle T, \mathbf{p}, x\rangle\,(\phi)\, d\mathcal{L}^M(x) \right| \\
&\qquad = \left| \Big(T \llcorner \big[(D_{x_i}\psi)\circ \mathbf{p}\big]\, dx_1 \wedge \cdots \wedge dx_M \Big)\,(\phi) \right| \\
&\qquad = \left| T\Big(\phi\, \big[(D_{x_i}\psi)\circ \mathbf{p}\big]\, dx_1 \wedge \cdots \wedge dx_M \Big) \right| \\
&\qquad = \left| T\,[\,\phi\, d(\psi \circ \mathbf{p}) \wedge dx_{\widehat{\imath}}\,] \right| \\
&\qquad = \left| (\partial T)\,[\phi\,(\psi \circ \mathbf{p})\, dx_{\widehat{\imath}}] - T\,[(\psi \circ \mathbf{p})\, d\phi \wedge dx_{\widehat{\imath}}] \right| \\
&\qquad \le \|\partial T\|(A) + \|T\|(A)\,,
\end{aligned}$$

so

$$\left| \int \langle T, \mathbf{p}, x\rangle\, \phi\, \operatorname{div}(g)\, d\mathcal{L}^n(x) \right| \le M\,[\, \|\partial T\|(A) + \|T\|(A)\,]\,. \qquad \square$$

In fact, we have the following more general result.

Theorem 8.1.9. *Let* $\mathbf{p} : \mathbb{R}^{M+K} = \mathbb{R}^M \times \mathbb{R}^K \to \mathbb{R}^M$ *be projection onto the first factor and fix* $0 < L < \infty$. *If for* $\ell = 1, 2, \ldots,$ *we have that* $T_\ell \in \mathcal{D}_M(\mathbb{R}^{M+K})$ *is an integer-multiplicity current with* $\mathbf{M}(T_\ell) + \mathbf{M}(\partial T_\ell) \le L$ *and if* $T_\ell \to T$ *weakly, then for* $\mathcal{L}^M$*-almost every* $x \in \mathbb{R}^M$, *it holds that* $\langle T, \mathbf{p}, x\rangle$ *is an integer-multiplicity current. Furthermore, the function* $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ *defined by*

$$u(x) = \langle T, \mathbf{p}, x\rangle$$

is MBV, and

$$\mathrm{V}_u(A) \le M\,L$$

holds for each open set $A \subseteq \mathbb{R}^M$.

Proof. Since $\langle T_\ell, \mathbf{p}, x\rangle \to \langle T, \mathbf{p}, x\rangle$ weakly for $\mathcal{L}^M$-almost every $x \in \mathbb{R}^M$, we see that $\langle T, \mathbf{p}, x\rangle$ is an integer-multiplicity current by the compactness theorem for

0-dimensional currents. Then, using the same notation as in the proof of Proposition 8.1.8, we estimate

$$\begin{aligned}
&\left| \int \psi_{x_i} \langle T, \mathbf{p}, x \rangle (\phi) \, d\mathcal{L}^M(x) \right| \\
&\qquad = \left| \left(T \llcorner \left(\psi_{x_i} \circ \mathbf{p} \right) dx_1 \wedge \cdots \wedge dx_M \right) (\phi) \right| \\
&\qquad = \left| T (\phi \left(\psi_{x_i} \circ \mathbf{p} \right) dx_1 \wedge \cdots \wedge dx_M) \right| \\
&\qquad = \left| T \left[\phi \, d(\psi \circ \mathbf{p}) \wedge dx_{\hat{\imath}} \right] \right| \\
&\qquad = \left| \lim_{\ell \to \infty} T_\ell \left[\phi \, d(\psi \circ \mathbf{p}) \wedge dx_{\hat{\imath}} \right] \right| \\
&\qquad = \lim_{\ell \to \infty} \left| (\partial T_\ell) \left[\phi \, (\psi \circ \mathbf{p}) \, dx_{\hat{\imath}} \right] - T_\ell \left[(\psi \circ \mathbf{p}) \, d\phi \wedge dx_{\hat{\imath}} \right] \right| \\
&\qquad \leq \lim_{\ell \to \infty} \Big[\, \|\partial T_\ell\|(A) + \|T_\ell\|(A) \Big],
\end{aligned}$$

and the result follows. □

Definition 8.1.10. For a measure μ on $\mathbb{R}^M$, we define the *maximal function for* μ, denoted by $\mathcal{M}_\mu$, by

$$\mathcal{M}_\mu(x) = \sup_{r>0} \frac{1}{\Omega_M r^M} \, \mu \left[\overline{\mathbb{B}}(x, r) \right].$$

(This definition is a variation on the definition given in Section 4.5.)

Lemma 8.1.11. *If v is a real-valued BV function and* 0 *is a Lebesgue point for f, then it holds that*

$$\begin{aligned}
&\frac{1}{\Omega_M r^M} \int_{\mathbb{B}(0,r)} \frac{|v(x) - v(0)|}{|x|} \, d\mathcal{L}^M(x) \\
&\qquad \leq \int_0^1 \frac{1}{\Omega_M (\tau r)^M} \int_{\mathbb{B}(0,\tau r)} |Dv(x)| \, d\mathcal{L}^M(x) \, d\mathcal{L}^1(\tau) \;\leq\; \mathcal{M}_{|Dv|}(0) \, .
\end{aligned}$$

Proof. For a C^1 function $v : \mathbb{R}^M \to \mathbb{R}$, we have

$$\begin{aligned}
|v(x) - v(0)| &= \left| \int_0^1 \frac{d}{d\tau} v(\tau x) \, d\mathcal{L}^1(\tau) \right| \\
&= \left| \int_0^1 \langle \, Dv(\tau x) \, , \; x \, \rangle \, d\mathcal{L}^1(\tau) \right| \leq \int_0^1 |Dv(\tau x)| \, |x| \, d\mathcal{L}^1(\tau) \, .
\end{aligned}$$

So

$$\begin{aligned}
&\frac{1}{\Omega_M r^M}\int_{\mathbb{B}(0,r)} \frac{|v(x)-v(0)|}{|x|}\, d\mathcal{L}^M(x)\\
&\qquad \le \int_{\mathbb{B}(0,r)}\int_0^1 \frac{1}{\Omega_M r^M}|Dv(\tau x)|\, d\mathcal{L}^1(\tau)\, d\mathcal{L}^M(x)\\
&\qquad = \int_0^1\int_{\mathbb{B}(0,r)} \frac{1}{\Omega_M r^M}|Dv(\tau x)|\, d\mathcal{L}^M(x)\, d\mathcal{L}^1(\tau)\\
&\qquad = \int_0^1 \frac{1}{\Omega_M (\tau r)^M}\int_{\mathbb{B}(0,\tau r)} |Dv(x)|\, d\mathcal{L}^M(x)\, d\mathcal{L}^1(\tau)\,.
\end{aligned}$$

The result follows by smoothing (see [KPk 99, Theorem 3.6.12]). □

Theorem 8.1.12. *If $v:\mathbb{R}^M\to\mathbb{R}$ is a BV function and y and z are Lebesgue points for v, then*

$$|v(y)-v(z)| \le \Big[\,\mathcal{M}_{|Dv|}(y)+\mathcal{M}_{|Dv|}(z)\,\Big]\,|y-z|\,.$$

Proof. Suppose that $y\ne z$. Let p be the midpoint of the segment connecting y and z and set $r=|y-z|$.

For $x\in\overline{\mathbb{B}}(p,r/2)$ we have

$$\frac{|v(y)-v(z)|}{|y-z|} \le \frac{|v(y)-v(x)|}{|y-z|}+\frac{|v(x)-v(z)|}{|y-z|}\,,$$

$$|x-y| \le |x-p|+|p-y| \le r/2+r/2=|y-z|\,,$$

$$|x-z| \le |x-p|+|p-z| \le r/2+r/2=|y-z|\,,$$

so

$$\begin{aligned}
\frac{|v(y)-v(z)|}{|y-z|} &\le \frac{|v(y)-v(x)|}{|y-z|}+\frac{|v(x)-v(z)|}{|y-z|}\\
&\le \frac{|v(y)-v(x)|}{|y-x|}+\frac{|v(x)-v(z)|}{|x-z|}\,.
\end{aligned}$$

As a result,

$$\begin{aligned}
\frac{|v(y)-v(z)|}{|y-z|} &= \frac{1}{\Omega_M r^M}\int_{\overline{\mathbb{B}}(p,r/2)} \frac{|v(y)-v(z)|}{|y-z|}\, d\mathcal{L}^M\\
&\le \frac{1}{\Omega_M r^M}\int_{\overline{\mathbb{B}}(p,r/2)} \frac{|v(y)-v(x)|}{|y-x|}\, d\mathcal{L}^M\\
&\qquad + \frac{1}{\Omega_M r^M}\int_{\overline{\mathbb{B}}(p,r/2)} \frac{|v(x)-v(z)|}{|x-z|}\, d\mathcal{L}^M
\end{aligned}$$

$$\leq \frac{1}{\Omega_M r^M} \int_{\overline{\mathbb{B}}(y,r)} \frac{|v(y) - v(x)|}{|y - x|} \, d\mathcal{L}^M$$

$$+ \frac{1}{\Omega_M r^M} \int_{\overline{\mathbb{B}}(z,r)} \frac{|v(x) - v(z)|}{|x - z|} \, d\mathcal{L}^M$$

$$\leq \mathcal{M}_{|Dv|}(y) + \mathcal{M}_{|Dv|}(z) \, . \qquad \square$$

Corollary 8.1.13. *If $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ is an MBV function, then there is a set E with $\mathcal{L}^M(E) = 0$ such that, for $y, z \in \mathbb{R}^M \setminus E$, it holds that*

$$\mathrm{d}_0[\, u(y), u(z) \,] \leq \Big[\, \mathcal{M}\mathrm{V}_u(y) + \mathcal{M}\mathrm{V}_u(z) \,\Big] \, |y - z| \, .$$

Proof. Let ϕ_i, $i = 1, 2, \ldots$, be a dense set in $\mathcal{D}^0(\mathbb{R}^M)$ and let E_i be the set of non-Lebesgue points for $u \diamond \phi_i$. Then we set $E = \bigcup_{i=1}^{\infty} E_i$ and the result follows from Theorem 8.1.12. $\square$

The preceding corollary tells us that an MBV function u is Lipschitz on any set where the maximal function for V_u is bounded. As we saw in Chapter 4, we can use covering-theorem methods to show that maximal functions are well behaved. We do so again in the next lemma.

Lemma 8.1.14. *For each $\lambda > 0$, it holds that*

$$\mathcal{L}^M\{x : \mathcal{M}_\mu(x) > \lambda\} \leq \frac{B_M}{\lambda} \, \mu(\mathbb{R}^M) \, ,$$

where B_M is the constant for $\mathbb{R}^M$ from the Besicovitch covering theorem.

Proof. Set

$$L = \{x : \mathcal{M}_\mu(x) > \lambda\} \, .$$

For each $x \in L$, choose a ball $\overline{\mathbb{B}}(x, r_x)$ such that

$$\frac{1}{\Omega_M r^M} \, \mu[\, \overline{\mathbb{B}}(x, r_x) \,] > \lambda \, .$$

Since $L \subseteq \bigcup_{x \in L} \overline{\mathbb{B}}(x, r_x)$, we can apply the Besicovitch covering theorem to find families $F_1, F_2, \ldots, F_{B_M}$ of pairwise disjoint balls $\overline{\mathbb{B}}(x, r_x)$, $x \in L$, such that $L \subseteq \bigcup_{i=1}^{B_M} \bigcup_{B \in F_i} B$. Then we have

$$\mathcal{L}^M(L) \leq \mathcal{L}^M\Big(\bigcup_{i=1}^{B_M} \bigcup_{B \in F_i} B\Big) \leq \sum_{i=1}^{B_M} \sum_{B \in F_i} 2^{-M} \, \Omega_M \, \mathrm{diam}\,(B)$$

$$< \frac{1}{\lambda} \sum_{i=1}^{B_M} \sum_{B \in F_i} \mu(B) \leq \frac{B_M}{\lambda} \, \mu(\mathbb{R}^M) \, . \qquad \square$$

Theorem 8.1.15. *If* $u : \mathbb{R}^M \to \mathcal{R}_0(\mathbb{R}^{M+K})$ *is an MBV function, then there is a set* E *with* $\mathcal{L}^M(E) = 0$ *such that*

$$\bigcup_{x \in \mathbb{R}^M \setminus E} \operatorname{spt}[\, u(x)\,]$$

is a countably M*-rectifiable subset of* $\mathbb{R}^{M+K}$.

Proof. We apply Lemma 8.1.14 to write $\mathbb{R}^M$ as the union of sets A_i on which the maximal function for V_u is bounded. By Corollary 8.1.13, there is a set $E_i \subseteq A_i$ of measure zero such that u is Lipschitz on $A_i \setminus E_i$. So we can apply Theorem 8.1.5 to see that $\bigcup_{x \in A_i \setminus E_i} \operatorname{spt}[\, u(x)\,]$ is countably M-rectifiable. □

8.1.4 The Slicing Lemma

Lemma 8.1.16. *Suppose that* $f : U \to \mathbb{R}$ *is Lipschitz.*

If T_i *converges weakly to* T *and*

$$\sup \Big(\mathbf{M}_W(T_i) + \mathbf{M}_W(\partial T_i) \Big) < \infty$$

for every $W \subset\subset U$*, then, for* $\mathcal{L}^1$*-almost every* r*, there is a subsequence* i_j *such that*

$$\langle T_{i_j}, f, r \rangle \ \text{converges weakly to} \ \langle T, f, r \rangle \tag{8.13}$$

and

$$\sup \Big(\mathbf{M}_W[\, \langle T_{i_j}, f, r \rangle \,] + \mathbf{M}_W[\, \partial \langle T_{i_j}, f, r \rangle \,] \Big) < \infty$$

holds for $W \subset\subset U$.

If additionally $W_0 \subset\subset U$ *is such that*

$$\lim_{i \to \infty} \Big(\mathbf{M}_{W_0}(T_i) + \mathbf{M}_{W_0}(\partial T_i) \Big) = 0\,,$$

then the subsequence can be chosen so that

$$\lim_{i \to \infty} \Big(\mathbf{M}_{W_0}[\, \langle T_{i_j}, f, r \rangle \,] + \mathbf{M}_{W_0}[\, \partial \langle T_{i_j}, f, r \rangle \,] \Big) = 0\,.$$

Proof. Passing to a subsequence for which $\|T_{i_j}\| + \|\partial T_{i_j}\|$ converges weakly to a Radon measure μ, we see that (8.13) holds, except possibly for the at most countably many r for which $\mu\{x : f(x) = r\}$ has positive measure.

The remaining conclusions follow by passing to additional subsequences and using (7.47) and the fact that $\partial \langle\, T_i,\ f,\ r \,\rangle = \langle\, \partial T_i,\ f,\ r \,\rangle$. □

8.1.5 The Density Lemma

Lemma 8.1.17. *Suppose that $T \in \mathcal{D}_M(U)$. For $\mathbb{B}(x,r) \subseteq U$, set*

$$\lambda(x,r) = \inf\{\mathbf{M}(S) : \partial S = \partial[\, T \llcorner \mathbb{B}(x,r)\,],\ S \in \mathcal{D}_M(U)\}\,.$$

(1) *If $\mathbf{M}_W(T) + \mathbf{M}_W(\partial T) < \infty$ holds for every $W \subset\subset U$, then*

$$\lim_{r \downarrow 0} \frac{\lambda(x,r)}{\|T\|\,(\mathbb{B}(x,r)\,)} = 1 \tag{8.14}$$

holds for $\|T\|$-almost every $x \in U$.

(2) *If*

(a) $\partial T = 0$,

(b) $\partial[\, T \llcorner \mathbb{B}(x,r)\,]$ *is integer-multiplicity for every $x \in U$ and almost every $0 < r$,*

(c) $\mathbf{M}_W(T) + \mathbf{M}_W(\partial T) < \infty$ *holds for every $W \subset\subset U$,*

then there exists $\delta > 0$ such that

$$\Theta_*^M(\|T\|, x) > \delta$$

holds for $\|T\|$-almost every $x \in U$.

Proof.
(1) We argue by contradiction. Since $\lambda(x,r) \le \|T\|\,(\mathbb{B}(x,r)\,)$ is true by definition, we suppose that there is an $\epsilon > 0$ and $E \subseteq U$ with $\|T\|(E) > 0$ such that for each $x \in E$ there exist arbitrarily small $r > 0$ such that

$$\lambda(x,r) < (1-\epsilon)\,\|T\|\,(\mathbb{B}(x,r)\,)\,.$$

We may assume that $E \subseteq W$ for an open $W \subset\subset U$.

Consider $\rho > 0$. Cover $\|T\|$-almost all of E by disjoint balls $B_i = \mathbb{B}(x_i, r_i)$, where $x_i \in E$ and $r_i < \rho$. For each i, let $S_i \in \mathcal{D}_M(U)$ satisfy

$$\partial S_i = [\, T \llcorner \mathbb{B}(x_i, r_i)\,], \qquad \mathbf{M}(S_i) < (1-\epsilon)\,\mathbf{M}[\, T \llcorner \mathbb{B}(x_i, r_i)\,]\,.$$

Set

$$T_\rho = T - \sum_i T \llcorner B_i + \sum_i S_i\,.$$

For any $\omega \in \mathcal{D}^M(U)$ we have

$$\begin{aligned}
(T - T_\rho)(\omega) &= \sum_i (T \llcorner B_i - S_i)(\omega) \\
&= \sum_i [\,\partial(\,\boldsymbol{\delta}_{x_i} \times\!\!\!\!\times (T \llcorner B_i - S_i)\,)\,](\omega) \\
&= \sum_i (\,\boldsymbol{\delta}_{x_i} \times\!\!\!\!\times (T \llcorner B_i - S_i)\,)(d\omega)
\end{aligned}$$

$$\leq \sum_i \mathbf{M}(\,\delta_{x_i} \mathbin{⨳} (T \llcorner B_i - S_i)\,) \cdot \sup |d\omega|$$

$$\leq \rho \sum_i \mathbf{M}(T \llcorner B_i - S_i) \cdot \sup |d\omega|$$

$$\leq 2\rho \sum_i \mathbf{M}(T \llcorner B_i) \cdot \sup |d\omega|$$

$$\leq 2\rho\, \mathbf{M}(T) \cdot \sup |d\omega|\,.$$

Thus we see that T_ρ converges weakly to T as ρ decreases to zero. By the lower semicontinuity of mass, we have

$$\mathbf{M}_W(T) \leq \liminf_{\rho \downarrow 0} \mathbf{M}_W(T_\rho)\,.$$

On the other hand, we have

$$\begin{aligned}
\mathbf{M}_W(T_\rho) &\leq \mathbf{M}_W\Big(T - \sum_i T \llcorner B_i\Big) + \sum_i \mathbf{M}_W(S_i) \\
&\leq \mathbf{M}_W\Big(T - \sum_i T \llcorner B_i\Big) + (1-\epsilon) \sum_i \mathbf{M}_W(T \llcorner B_i) \\
&\leq \mathbf{M}_W(T) - \epsilon \sum_i \mathbf{M}_W(T \llcorner B_i) \\
&\leq \mathbf{M}_W(T) - \epsilon \|T\|(E)\,,
\end{aligned}$$

a contradiction.

(2) Let x be a point at which (8.14) holds. Set $f(r) = \mathbf{M}(\,T \llcorner \mathbb{B}(x,r)\,)$. For sufficiently small r we have

$$f(r) < 2\lambda(x,r)\,. \tag{8.15}$$

To be specific, let us suppose that (8.15) holds for $0 < r < R$.

For $\mathcal{L}^1$-almost every r, we have

$$\mathbf{M}[\,\partial(\,T \llcorner \mathbb{B}(x,r)\,)\,] \leq f'(r)\,.$$

Applying the isoperimetric inequality, we have

$$\lambda(x,r)^{(M-1)/M} \leq c_0\, f'(r)\,,$$

where c_0 is a constant depending only on the dimensions M and K. So, by (8.15), we have

$$[f(r)]^{(M-1)/M} \leq c_1\, f'(r) \qquad (0 < r < R)\,,$$

where c_1 is another constant. Thus we have

$$\frac{d}{dr}\Big[f(r)\Big]^{1/M} = (1/M)\, f'(r)\Big[f(r)\Big]^{(1-M)/M} \geq 1/c_1\,.$$

Since f is a nondecreasing function, we have

$$\Big[f(\rho)\Big]^{1/M} \geq \int_0^\rho \frac{d}{d\mathcal{L}^1(r)}\Big[f(r)\Big]^{1/M}\, d\mathcal{L}^1(r) \geq \int_0^\rho 1/c_1\, d\mathcal{L}^1(r) = \rho/c_1\,.$$

We conclude that $f(r) \geq (r/c_1)^M$ holds for $0 < r < R$. □

8.1.6 Completion of the Proof of the Compactness Theorem

Now that we have all the requisite tools at hand, we can complete the proof of the compactness theorem. Recall that by hypothesis we have a sequence $\{T_j\} \subseteq \mathcal{D}_M(U)$ of integer-multiplicity currents such that

$$\sup_{j\geq 1}\Big[\mathbf{M}_W(T_j) + \mathbf{M}_W(\partial T_j)\Big] < \infty \quad \text{for all } W \subset\subset U\,.$$

By applying the Banach–Alaoglu theorem and passing to a subsequence if necessary, but without changing notation, we may assume that there is a current $T \in \mathcal{D}_M(U)$ such that $T_j \to T$ weakly in U. Our task is to show that T is an integer-multiplicity rectifiable current.

By the slicing lemma applied with $f(x) = |x - a|$ ($a \in U$), we see that it suffices to consider the case in which $U = \mathbb{R}^{M+K}$ and all the T_j are supported in a fixed compact set.

By the boundary rectifiability theorem, each ∂T_j is integer-multiplicity. By the compactness theorem for currents of dimension $M - 1$, ∂T is integer-multiplicity (since ∂T_j converges weakly to ∂T). We know then that $\boldsymbol{\delta}_0 \mathbin{⨳} (\partial T_j)$ and $\boldsymbol{\delta}_0 \mathbin{⨳} (\partial T)$ are integer-multiplicity. By subtracting those currents from T_j and T, we may suppose that $\partial T_j = 0$, for all j (and, of course, $\partial T = 0$).

By Lemma 8.1.17, we know that $\|T\| = \mathcal{H}^M \llcorner \theta$, where θ is real-valued and nonnegative. In fact, θ is bounded below by a positive number, so we see that

$$A = \{x \in \mathbb{R}^{M+K} : \theta(x) > 0\}$$

has finite $\mathcal{H}^M$ measure.

Consider α a multi-index with

$$1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_M \leq M + K\,. \tag{8.16}$$

Let $\mathbf{p}_\alpha : \mathbb{R}^{M+K} \to \mathbb{R}^M$ be the orthogonal projection mapping

$$\mathbb{R}^{M+K} \ni x \longmapsto \sum_{i=1}^{M} (\mathbf{e}_{\alpha_i} \cdot x)\, \mathbf{e}_i\,.$$

By Theorem 8.1.9, we see that $\langle T, \mathbf{p}_\alpha, x\rangle$ is an MBV function of x with total variation measure bounded by ML. By Theorem 8.1.15, we see that there is a set $E_\alpha \subseteq \mathbb{R}^M$ with $\mathcal{L}^M(E_\alpha) = 0$ such that

$$S_\alpha = \bigcup_{x \in \mathbb{R}^M \setminus E_\alpha} \operatorname{spt}\,[\,\langle T, \mathbf{p}_\alpha, x\rangle\,]$$

is a countably M-rectifiable subset of $\mathbb{R}^{M+K}$. Also set

$$B_\alpha = A \bigcap \mathbf{p}_\alpha^{-1}(E_\alpha)\,.$$

We have $A \subseteq S_\alpha \bigcup B_\alpha$.

Letting I denote the set of all the multi-indices as in (8.16), we see that

$$A \subseteq \bigcap_{\alpha \in I} \Big[\, S_\alpha \bigcup B_\alpha \,\Big] \subseteq S \bigcup B\,,$$

where

$$S = \bigcup_{\alpha \in I} S_\alpha\,, \qquad B = \bigcap_{\alpha \in I} B_\alpha\,.$$

By Lemma 7.4.2, $T \llcorner B = 0$, so $T = T \llcorner S$.

We may suppose that $A \subseteq S$. By Theorem 8.1.9 we know that, for each $\alpha \in I$ and for $\mathcal{L}^M$-almost every $x \in \mathbb{R}^M$, $\langle T, \mathbf{p}_\alpha, x\rangle$ is integer-valued. So we conclude that θ is in fact integer-valued.

Finally, Theorem 8.1.6 tells us that T is an integer-multiplicity rectifiable current. □

8.2 The Flat Metric

Here we introduce a new topology given by the so-called flat metric. Our main result is that, for a sequence of integer-multiplicity currents $\{T_j\} \subset \mathcal{D}_M(U)$ with $\sup_{j \geq 1}[\,\mathbf{M}_W(T_j) + \mathbf{M}_W(\partial T_j)\,] < \infty$, for all $W \subset\subset U$, this new topology is equivalent to that given by weak convergence. There is some confusion in the literature because readers assume that the word "flat" has some geometric connotation of a lack of curvature. In point of fact the use of this word is an allusion to Hassler Whitney's use of the musical notation $\flat$ to denote the metric.

Let U denote an arbitrary open set in $\mathbb{R}^{M+K}$. Set

$$\mathcal{I}(U) = \{T \in \mathcal{D}_M(U) : T \text{ is integer-multiplicity}, \mathbf{M}_W(\partial T) < \infty \text{ if } W \subset\subset U\}\,.$$

Also set, for any $L > 0$ and $W \subset\subset U$,

$$\mathcal{I}_{L,W}(U) = \{T \in \mathcal{I} : \operatorname{spt} T \subset \overline{W},\ \mathbf{M}(T) + \mathbf{M}(\partial T) \leq L\}\,.$$

When the open set U is clear from context, as it usually is, we will simply write $\mathcal{I}$ and $\mathcal{I}_{L,W}$ for $\mathcal{I}(U)$ and $\mathcal{I}_{L,W}(U)$, respectively.

On $\mathcal{I}$ we define a family of pseudometrics $\{\mathrm{d}_W\}_{W \subset\subset U}$ by

$$\mathrm{d}_W(T_1, T_2) = \inf \Big\{ \mathbf{M}_W(S) + \mathbf{M}_W(R) : T_1 - T_2 = \partial R + S,$$
$$R \in \mathcal{D}_{M+1}(U),\ S \in \mathcal{D}_M(U) \text{ are of integer multiplicity} \Big\}\,.$$

It is worth explicitly noting that if $\omega \in \mathcal{D}^M(U)$ with $\operatorname{spt}\omega \subset W$, then

$$|(T_1 - T_2)(\omega)| \le \mathrm{d}_W(T_1, T_2) \cdot \max\left\{ \sup_{x\in W} |\omega(x)|,\ \sup_{x\in W} |d\omega(x)| \right\}. \tag{8.17}$$

In what follows we shall assume that $\mathcal{I}$ is equipped with the topology given by the family $\{\mathrm{d}_W\}_{W\subset\subset U}$ of pseudometrics. This topology is the *flat metric topology* for $\mathcal{I}$. Obviously there is a countable topological base of neighborhoods at each point, and $T_j \to T$ in this topology if and only if $\mathrm{d}_W(T_j, T) \to 0$ for all $W \subset\subset U$.

Theorem 8.2.1. *Let T, $\{T_j\}$ in $\mathcal{D}_M(U)$ be integer-multiplicity currents with $\sup_{j\ge 1}\{\mathbf{M}_W(T_j) + \mathbf{M}_W(\partial T_j)\} < \infty$ for all $W \subset\subset U$. Then T_j converges weakly to T if and only if*

$$\mathrm{d}_W(T_j, T) \to 0 \ \textit{for each}\ \ W \subset\subset U\,. \tag{8.18}$$

Remark 8.2.2. The statement of this last theorem in no way invokes the compactness theorem (Theorem 7.5.2), but we must note that if we combine the result with the compactness theorem then we can see that, for any family of positive (finite) constants $\{c(W)\}_{W\subset\subset U}$, the set

$$\{\, T \in \mathcal{I} : \mathbf{M}_W(T) + \mathbf{M}_W(\partial T) \le c(W) \ \text{ for all }\ W \subset\subset U \,\}$$

is sequentially compact when equipped with the flat metric topology.

Proof of the Theorem. First observe that if (8.18) holds, then (8.17) implies that T_j converges weakly to T.

In proving the converse, that weak convergence implies flat metric convergence, the main point is demonstrating the appropriate total boundedness property. More particularly, we shall show that for any given $\epsilon > 0$ and $W \subset\subset \widetilde{W} \subset\subset U$, we can find a number $n = n(\epsilon, W, \widetilde{W}, L)$ and integer-multiplicity currents $P_1, P_2, \ldots, P_n \in \mathcal{D}_M(U)$ such that

$$\mathcal{I}_{L,W} \subset \bigcup_{j=1}^{n} \{\, S \in \mathcal{I} : \mathrm{d}_{\widetilde{W}}(S, P_j) < \epsilon \,\}\,; \tag{8.19}$$

that is, each element of $\mathcal{I}_{L,W}$ is within ϵ of one of the currents $P_1, P_2, \ldots, P_n$, as measured by the pseudometric $\mathrm{d}_{\widetilde{W}}$. This fact follows immediately from the deformation theorem. To wit, for any $\rho > 0$, Theorem 7.7.2 shows that for $T \in \mathcal{I}_{L,W}$ we can find integer-multiplicity currents P, R, S such that

(1) $T - P = \partial R + S$;
(2) $P = \sum_{F\in\mathcal{L}_M} p_F\, \boldsymbol{\eta}_{\rho\#}[\![F]\!]\,, \quad p_F \in \mathbb{Z}$;
(3) $\operatorname{spt} P \subset \{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{M+K}\,\rho\}$;
(4) $\mathbf{M}(P) = \sum_{F\in\mathcal{L}_M(\rho)} |p_F|\, \rho^M$ and $\mathbf{M}(P) \le c\,\mathbf{M}(T) \le c\,L$;

(5) $\operatorname{spt} R \bigcup \operatorname{spt} S \subset \{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{M+K}\,\rho\}$

and $\mathbf{M}(R) + \mathbf{M}(S) \leq c\,\rho\,\mathbf{M}(T) \leq c\,\rho\,L$.

It follows that for ρ small enough to ensure $2\sqrt{M+K} < \operatorname{dist}(W, \partial\widetilde{W})$, the estimates (1) and (5) imply that

$$\mathrm{d}_{\widetilde{W}}(T, P) \leq c\,\rho\,L\,.$$

Since there are only finitely many currents P as in (2), (3), (4), they may be indexed $P_1, \ldots, P_n$ as in (8.19), where the number n depends only on L, W, M, K, and ρ.

Next we choose an increasing family of sets $W_i \subset\subset U$ such that the boundaries of the W_i cut the T_j in a controlled way. Specifically, we notice that by (1) and (2) of Lemma 7.6.3 and Sard's theorem (i.e., Corollary 5.1.10), we can find a subsequence $\{T_{j'}\} \subset \{T_j\}$ and a sequence $\{W_i\}$ with $W_i \subset\subset W_{i+1} \subset\subset U$ and $\cup_{i=1}^{\infty} W_i = U$ such that $\sup_{j' \geq 1} \mathbf{M}[\,\partial(T_{j'} \llcorner W_i)\,] < \infty$ for all i. It follows that we may henceforth assume without loss of generality that $W \subset\subset U$ and

$$\operatorname{spt} T_j \subset \overline{W} \quad \text{for all } j\,.$$

Now we take any $\widetilde{W}$ such that $W \subset\subset \widetilde{W} \subset\subset U$. We apply (8.19) with $\epsilon = 2^{-r}$, $r = 1, 2, \ldots$, so that we may extract a subsequence $\{T_{j_r}\}_{r=1}^{\infty}$ from $\{T_j\}$ such that

$$\mathrm{d}_{\widetilde{W}}(T_{j_{r+1}}, T_{j_r}) < 2^{-r}$$

and so

$$T_{j_{r+1}} - T_{j_r} = \partial R_r + S_r\,. \tag{8.20}$$

Here R_r, S_r are integer-multiplicity,

$$\operatorname{spt} R_r \bigcup \operatorname{spt} S_r \subset \widetilde{W}\,,$$

and

$$\mathbf{M}(R_r) + \mathbf{M}(S_r) \leq 2^{-r}\,.$$

Thus, by the compactness theorem, Theorem 7.5.2, we can define integer-multiplicity currents $R^{(\ell)}$, $S^{(\ell)}$ via series

$$R^{(\ell)} = \sum_{r=\ell}^{\infty} R_r$$

and

$$S^{(\ell)} = \sum_{r=\ell}^{\infty} S_r\,,$$

which converge in the mass topology. It follows then that

$$\mathbf{M}[\,R^{(\ell)}\,] + \mathbf{M}[\,S^{(\ell)}\,] \leq 2^{-\ell+1}$$

and, from (8.20),

$$T - T_{j_\ell} = \partial R^{(\ell)} + S^{(\ell)} .$$

Hence we have a subsequence $\{T_{j_\ell}\}$ of $\{T_j\}$ such that $\mathrm{d}_{\widetilde{W}}(T, T_{j_\ell}) \to 0$. Since we can in this manner extract a subsequence converging relative to $\mathrm{d}_{\widetilde{W}}$ from any given subsequence of $\{T_j\}$, we have $\mathrm{d}_{\widetilde{W}}(T, T_j) \to 0$. Since this process can be repeated with $W = W_i$, $\widetilde{W} = W_{i+1}$ for all i, the desired result follows. □

8.3 Existence of Currents Minimizing Variational Integrals

8.3.1 Minimizing Mass

One of the problems that motivated the development of the theory of integer-multiplicity currents is the problem of finding an area-minimizing surface having a prescribed boundary. The study of area-minimizing surfaces is quite old, dating back to Euler's discovery[2] of the area-minimizing property of the catenoid in the 1740s and to Lagrange's discovery[3] of the minimal surface equation in the 1760s. But despite the many advances since the time of Euler and Lagrange, many interesting questions and avenues of research remain.

In the context of integer-multiplicity currents, it is appropriate to investigate the problem of minimizing the *mass* of the current, as the mass accounts for both the area of the corresponding surface and the multiplicity attached to the surface. The next definition applies in very general situations to make precise the notion of a current being mass-minimizing in comparison with currents having the same boundary.

Definition 8.3.1. Suppose that $U \subseteq \mathbb{R}^N$ is open and $T \in \mathcal{D}_M(\mathbb{R}^N)$ is an integer-multiplicity current. For a subset $B \subseteq U$, we say that T is *mass-minimizing in B* if

$$\mathbf{M}_W[T] \leq \mathbf{M}_W[S] \tag{8.21}$$

holds whenever S is an integer-multiplicity current and

$$\begin{aligned} &W \subset\subset U , \\ &\partial S = \partial T , \\ &\operatorname{spt}[S - T] \ \text{ is a compact subset of } B \cap W. \end{aligned}$$

Remark 8.3.2. In case $B = \mathbb{R}^N$, we say simply that T is *mass-minimizing*. If, additionally, T has compact support, then Definition 8.3.1 reduces to the requirement that

$$\mathbf{M}[T] \leq \mathbf{M}[S]$$

hold whenever $\partial S = \partial T$.

[2] Leonhard Euler (1707–1783).

[3] Joseph-Louis Lagrange (1736–1813).

If R is a nontrivial $(M-1)$-dimensional current that is the boundary of *some* integer-multiplicity current, then it makes sense to ask whether there exists a mass-minimizing integer-multiplicity current with R as its boundary. The next theorem tells us that indeed, such a mass-minimizing current does exist.

Theorem 8.3.3. *Suppose that* $1 \leq M \leq N$. *If* $R \in \mathcal{D}_{M-1}(\mathbb{R}^N)$ *has compact support and if there exists an integer-multiplicity current* $Q \in \mathcal{D}_M(\mathbb{R}^N)$ *with* $R = \partial Q$, *then there exists a mass-minimizing integer-multiplicity current* T *with* $\partial T = R$.

Proof. Let $\{T_i\}_{i=1}^{\infty}$ be a sequence of integer-multiplicity currents with $\partial T_i = R$, for $i = 1, 2, \ldots$, and with

$$\lim_{i\to\infty} \mathbf{M}[T_i] = \inf\{\, \mathbf{M}[S] \;:\; \partial S = R,\ S \text{ is integer-multiplicity}\,\}\,.$$

Set $m = \operatorname{dist}(\operatorname{spt} R,\ 0)$ and let $f : \mathbb{R}^N \to \overline{\mathbb{B}}(0,\ m)$ be the nearest-point retraction. Because the boundary operator and the pushforward operator commute, we have

$$\partial(f_\# T_i) = f_\#(\partial T_i) = f_\# R = R$$

for $i = 1, 2, \ldots$. Noting that $\operatorname{Lip}(f) = 1$, we conclude that

$$\mathbf{M}[f_\# T_i] \leq \mathbf{M}[T_i]$$

holds, for $i = 1, 2, \ldots$. Thus, by replacing T_i with $f_\# T_i$ if need be, we may suppose that $\operatorname{spt} T_i \subseteq \overline{\mathbb{B}}(0,\ m)$ holds for $i = 1, 2, \ldots$.

Now consider the sequence of integer-multiplicity currents $\{S_i\}_{i=1}^{\infty}$ defined by setting $S_i = T_i - Q$, for each $i = 1, 2, \ldots$. Noting that $\partial S_i = 0$ for each i, we see that the sequence $\{S_i\}_{i=1}^{\infty}$ satisfies the conditions of the compactness theorem (Theorem 7.5.2). We conclude that there exist a subsequence $\{S_{i_k}\}_{k=1}^{\infty}$ of $\{S_i\}_{i=1}^{\infty}$ and an integer-multiplicity current S^* such that $S_{i_k} \to S^*$ as $k \to \infty$. We conclude also that $\partial S^* = 0$.

Setting $T = S^* + Q$, we see that $T_{i_k} = S_{i_k} + Q \to S^* + Q = T$ as $k \to \infty$ and that $\partial T = \partial(S^* + Q) = \partial S^* + \partial Q = \partial Q = R$. By the lower semicontinuity of the mass, we have

$$\mathbf{M}[T] = \inf\{\, \mathbf{M}[S] \;:\; \partial S = R,\quad S \text{ is integer-multiplicity}\,\}\,. \qquad \square$$

8.3.2 Other Integrands and Integrals

Minimizing the mass of a current is only one of many possible variational problems that can be considered in the space of integer-multiplicity currents. To introduce more general problems, we first define an appropriate class of integrands.

Definition 8.3.4. Let $U \subseteq \mathbb{R}^N$ be open and suppose that $1 \leq M \leq N$.

(1) By an *M-dimensional parametric integrand on U* we mean a continuous function $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ satisfying the homogeneity condition

$$F(x, r\omega) = r\,F(x, \omega)\,, \quad \text{for } r \geq 0,\ x \in U,\ \omega \in \bigwedge\nolimits_M \mathbb{R}^N\,.$$

The integrand is *positive* if

$$F(x, \omega) > 0$$

holds whenever $\omega \neq 0$. We will limit our attention to positive integrands (see Remark 8.3.5).

(2) If F is an M-dimensional parametric integrand on U and $T = \boldsymbol{\tau}(V, \theta, \xi)$ is an M-dimensional integer-multiplicity current supported in U, then the *integral of F over T*, denoted by $\int_T F$, is defined by setting

$$\int_T F = \int_V F(x, \theta(x)\,\xi(x))\,d\mathcal{H}^M(x) = \int_U F(x, \overrightarrow{T}(x))\,d\|T\|(x)\,.$$

(3) We say that the parametric integrand F is a *constant-coefficient integrand* if $F(x_1, \omega) = F(x_2, \omega)$ holds for $x_1, x_2 \in U$ and $\omega \in \bigwedge_M \mathbb{R}^N$. If F is a constant-coefficient integrand, then it is no loss of generality to assume that $U = \mathbb{R}^N$.

(4) Given any $x_0 \in U$, we define the constant-coefficient parametric integrand F_{x_0} by setting

$$F_{x_0}(x, \omega) = F(x_0, \omega)\,, \quad \text{for } x \in \mathbb{R}^N,\ \omega \in \bigwedge\nolimits_M \mathbb{R}^N\,.$$

Remark 8.3.5. The limitation to considering a positive integrand is convenient when one seeks a current that minimizes the integral of the integrand, because one knows immediately that zero is a lower bound for the possible values of the integral.

Example 8.3.6.

(1) The *M-dimensional area integrand* is the constant-coefficient parametric integrand A given by

$$A(x, \omega) = |\omega|\,, \quad \text{for } x \in U,\ \omega \in \bigwedge\nolimits_M \mathbb{R}^N\,.$$

We see that

$$\int_T A = \mathbf{M}[T]\,.$$

(2) Let F be an $(N-1)$-dimensional parametric integrand on $\mathbb{R}^N$. If W is a bounded open subset of $\mathbb{R}^N$ and T is the $(N-1)$-dimensional integer-multiplicity current associated with the graph of a function $g : W \to \mathbb{R}$, then

$$\int_T F = \int_W F\left[\,(x, g(x)),\ \mathbf{e}^N + \textstyle\sum_{i=1}^{N-1} D_i g(x)\,\mathbf{e}_{\hat{\imath}}\right] d\mathcal{L}^{N-1}(x)\,.$$

Comparing with [Mor 66, p. 2] for instance, we see that integrating the parametric integrand F over a surface defined by the graph of a function g gives the same result as evaluating the classical nonparametric functional

$$\int_W \mathcal{F}[x, g(x), Dg(x)]\,d\mathcal{L}^{N-1}(x)$$

over the region W, where the integrand $\mathcal{F}$ is given by

$$\mathcal{F}[x, z, p] = F\Big[\,(x, z), \mathbf{e}^N + \textstyle\sum_{i=1}^{N-1} p_i\, \mathbf{e}_{\hat{i}}\Big], \tag{8.22}$$

for $x \in \mathbb{R}^{N-1}$, $z \in \mathbb{R}$, and $p = (p_1, p_2, \ldots, p_{N-1}) \in \mathbb{R}^{N-1}$.

A similar comparison can be made in higher codimensions, but the notation becomes increasingly unwieldy. □

The notion of minimizing a parametric integrand is defined analogously to Definition 8.3.1, but with the appropriate modification of (8.21); more precisely, we have the following definition.

Definition 8.3.7. Let $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ be an M-dimensional parametric integrand on U. Suppose that $T \in \mathcal{D}_M(\mathbb{R}^N)$ is an integer-multiplicity current. For a subset $B \subseteq U$, we say that T is *F-minimizing in B* if

$$\int_{T \llcorner W} F \le \int_{S \llcorner W} F \tag{8.23}$$

holds whenever S is an integer-multiplicity current and

$$\begin{aligned} &W \subset\subset U\,, \\ &\partial S = \partial T\,, \\ &\operatorname{spt}[S - T] \ \text{ is a compact subset of } B \cap W. \end{aligned}$$

The existence of mass-minimizing currents was guaranteed by Theorem 8.3.3. The proof of that theorem, as given above, is an instance of the "direct method" in the calculus of variations. In the direct method, a minimizing sequence is chosen (always possible as long as the infimum of the values of the functional is finite), a convergent subsequence is extracted (a compactness theorem is needed—in our case Theorem 7.5.2), and a lower-semicontinuity result is applied (lower semicontinuity is immediate for the mass functional). Thus the question naturally arises whether the integral of a parametric integrand is lower semicontinuous.

Definition 8.3.8. Let $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ be an M-dimensional positive parametric integrand on U. We say that F is *semielliptic* if for each $x_0 \in U$, the integer-multiplicity current associated with any oriented M-dimensional plane is F_{x_0}-minimizing.

Remark 8.3.9. What Definition 8.3.8 tells us is that F is semielliptic if and only if for every $x_0 \in U$, the conditions

(1) $v_1, v_2, \ldots, v_M \in \mathbb{R}^N$ are linearly independent,
(2) V is a bounded, relatively open subset of $\operatorname{span}\{v_1, v_2, \ldots, v_M\}$,
(3) $\xi = v_1 \wedge v_2 \wedge \cdots \wedge v_M / |v_1 \wedge v_2 \wedge \cdots \wedge v_M|$,
(4) $T = \boldsymbol{\tau}(V, 1, \xi)$,
(5) R is a compactly supported integer-multiplicity current,
(6) $\partial R = \partial T$,

imply that

$$\int_T F_{x_0} \le \int_R F_{x_0} \,. \tag{8.24}$$

The hypothesis of semiellipticity for the integrand F is sufficient to guarantee the lower semicontinuity of the integral of F as a functional on integer-multiplicity currents. We state the result here without proof.

Theorem 8.3.10. *Suppose that* $1 \le M \le N$. *Let* $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ *be an* M*-dimensional positive parametric integrand on* U*. If* F *is semielliptic, then the functional* $T \longmapsto \int_T F$ *is lower semicontinuous. That is, if* $K \subset U$ *is compact,* $T_i \to T$ *in the flat metric, and* $\operatorname{spt} T_i \subseteq K$ *for* $i = 1, 2, \ldots,$ *then it holds that*

$$\int_T F \le \liminf_{i\to\infty} \int_{T_i} F \,.$$

The heuristic of the proof is that, for $\|T\|$-almost every x_0, T can be approximated by an M-dimensional plane and F can be approximated by F_{x_0}. The details can be found in [Fed 69, 5.1.5].

Corollary 8.3.11. *Suppose that* $1 \le M \le N$. *Let* $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ *be an* M*-dimensional semielliptic positive parametric integrand. Let* K *be a compact subset of* U*. If* $R \in \mathcal{D}_{M-1}(\mathbb{R}^N)$ *and if there exists an integer-multiplicity current* $Q \in \mathcal{D}_M(\mathbb{R}^N)$ *with* $R = \partial Q$ *and with* $\operatorname{spt} Q \subseteq K$*, then there exists an integer-multiplicity current* T *with* $\partial T = R$ *and with* $\operatorname{spt} T \subseteq K$ *that is* F*-minimizing in* K.

Proof. Proceeding as in the proof of Theorem 8.3.3, we let $\{T_i\}_{i=1}^{\infty}$ be a sequence of integer-multiplicity currents with $\partial T_i = R$ and with $\operatorname{spt} T_i \subseteq K$, for $i = 1, 2, \ldots,$ chosen so that

$$\lim_{i\to\infty} \int_{T_i} F$$
$$= \inf \left\{ \int_S F \;:\; \partial S = R, \quad \operatorname{spt} S \subseteq K, \quad S \text{ is integer-multiplicity} \right\}.$$

By the compactness theorem, we can extract a convergent subsequence, and then the result follows from Theorem 8.3.10. □

As regards being convenient for guaranteeing lower semicontinuity, the condition of semiellipticity is hardly satisfactory, since it may be difficult to verify that currents associated with M-dimensional planes are F_{x_0}-minimizing. A more practical condition is that each F_{x_0} be convex.

Definition 8.3.12. Let $F : U \times \bigwedge_M \mathbb{R}^N \to \mathbb{R}$ be an M-dimensional parametric integrand on U. We say that F is *convex* if for each $x_0 \in U$, F_{x_0} is a convex function on $\bigwedge_M \mathbb{R}^N$, that is, if

$$F\Big(x_0, \lambda\omega_1 + (1-\lambda)\omega_2\Big) \le \lambda\, F(x_0, \omega_1) + (1-\lambda)\, F(x_0, \omega_2)$$

holds for $\omega_1, \omega_2 \in \bigwedge_M \mathbb{R}^N$ and $0 \le \lambda \le 1$.

Theorem 8.3.13. *If the M-dimensional parametric integrand F is convex, then it is semielliptic.*

Proof. Let F be convex and fix $x_0 \in U$. Suppose that the conditions of Remark 8.3.9(1)–(6) hold.

First we claim that

$$\int \overrightarrow{T}\, d\|T\| = \int \overrightarrow{R}\, d\|R\|\,. \tag{8.25}$$

Both sides of (8.25) are elements of $\bigwedge_M \mathbb{R}^N$. If (8.25) were false, then we could find $\omega \in \bigwedge^M \mathbb{R}^N$ such that

$$\left\langle \omega,\ \int \overrightarrow{T}\, d\|T\| - \int \overrightarrow{R}\, d\|R\| \right\rangle \neq 0\,.$$

But choosing $W \in \mathcal{D}_{M+1}(\mathbb{R}^N)$ such that $\partial W = T - R$, as we may because $\partial(T-R) = 0$, and thinking of ω as a differential form having a constant value (so that $d\omega = 0$ holds), we see that

$$\begin{aligned} 0 = W[d\omega] = (\partial W)[\omega] &= \int \langle \omega, \overrightarrow{T}\,\rangle\, d\|T\| - \int \langle \omega, \overrightarrow{R}\,\rangle\, d\|R\| \\ &= \left\langle \omega,\ \int \overrightarrow{T}\, d\|T\| - \int \overrightarrow{R}\, d\|R\| \right\rangle , \end{aligned}$$

a contradiction.

Now by the homogeneity of F_{x_0}, the fact that $\overrightarrow{T}$ is constant, equation (8.25), and using Jensen's inequality,[4] we obtain

$$\begin{aligned} \int_T F_{x_0} &= \int F\left(x_0, \overrightarrow{T}\right) d\|T\| = F\left(x_0, \overrightarrow{T}\right) \|T\|[\mathbb{R}^N] \\ &= F\left(x_0, \overrightarrow{T}\ \|T\|[\mathbb{R}^N]\right) = F\left(x_0, \int \overrightarrow{T}\, d\|T\|\right) \\ &= F\left(x_0, \int \overrightarrow{R}\, d\|R\|\right) \leq \int F\left(x_0, \overrightarrow{R}\right) d\|R\| = \int_R F_{x_0}\,. \end{aligned}$$

□

Finally, we illustrate the subtle difference between the notion of a convex parametric integrand and the notion of convexity of integrands in the nonparametric setting.

Example 8.3.14. The 2-dimensional parametric area integrand on $\mathbb{R}^4$ is convex, but the integrand that gives the 2-dimensional area of the graph of a function g over a region in $\mathbb{R}^2$ is not a convex function of Dg. In fact, if $g = (g_1, g_2)$ is a function of (x_1, x_2), then the area of the graph of g is found by integrating

[4] Johan Ludwig William Valdemar Jensen (1859–1925).

$$\mathcal{F}(p_{1,1}, p_{1,2}, p_{2,1}, p_{2,2}) = \sqrt{1 + \sum_{i,j=1}^{2} p_{i,j}^2 + \big(p_{1,1}\, p_{2,2} - p_{1,2}\, p_{2,1}\big)^2}\,, \tag{8.26}$$

where we set

$$p_{i,j} = \frac{\partial g_i}{\partial x_j}\,.$$

We see that the function in (8.26) is not convex by comparing

$$\frac{\mathcal{F}(T, T, 0, 0) + \mathcal{F}(0, 0, -T, T)}{2} = \sqrt{1 + 2T^2} \tag{8.27}$$

with

$$\mathcal{F}\Big(\tfrac{1}{2}T, \tfrac{1}{2}T, -\tfrac{1}{2}T, \tfrac{1}{2}T\Big) = \sqrt{1 + T^2 + \tfrac{1}{4}T^4}\,, \tag{8.28}$$

and noting that, for large $|T|$, the value in (8.28) is larger than the value in (8.27). □

8.4 Density Estimates for Minimizing Currents

One gains information about a current that minimizes a variational integral by using comparison surfaces. A comparison surface can be any surface having the same boundary as the minimizer. To be useful, a comparison surface should be one that you construct in such a way that the variational integral on the comparison surface can be estimated. Since the variational integral for the minimizer must be less than or equal to the integral for the comparison surface, some information can thereby be gleaned from the estimate for the variational integral on the comparison surface. The next lemma illustrates this idea.

Lemma 8.4.1. *If* $T \in \mathcal{D}_M(\mathbb{R}^N)$ *is a mass-minimizing, integer-multiplicity current,* $p \in \operatorname{spt} T$, *and* $\mathbb{B}(p, r) \cap \operatorname{spt} \partial T = \emptyset$, *where* $0 < r$, *then*

$$\mathbf{M}[\, T \llcorner \mathbb{B}(p, r)\,] \le \frac{r}{M}\, \mathbf{M}[\, \partial(T \llcorner \mathbb{B}(p, r))\,]\,. \tag{8.29}$$

Proof. The comparison surface C that we use is the cone over $\partial(T \llcorner \mathbb{B}(p, r))$ with vertex p; see Figure 8.1. That is, we set

$$C = \boldsymbol{\delta}_p \mathbin{⋇} \partial(T \llcorner \mathbb{B}(p, r))$$

using the cone construction in (7.25) with 0 replaced by p and M replaced by $M - 1$. Then by (7.26) we have

$$\partial C = \partial(T \llcorner \mathbb{B}(p, r)) \tag{8.30}$$

and by (7.27) we have

$$\mathbf{M}[C] \le \frac{r}{M}\, \mathbf{M}[\, \partial(T \llcorner \mathbb{B}(p, r))\,]\,. \tag{8.31}$$

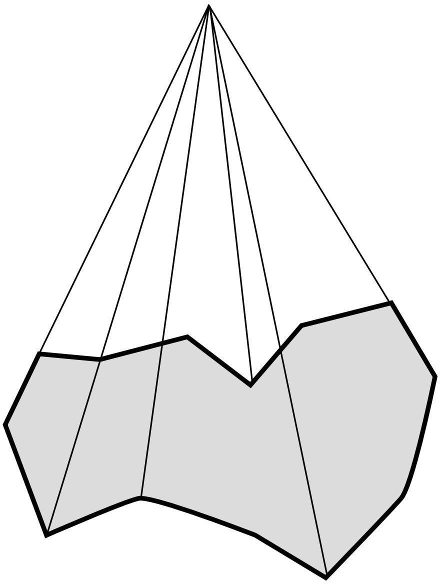

Fig. 8.1. The conical comparison surface.

By (8.30), we see that

$$\partial\Big(T + C - T \llcorner \mathbb{B}(p,r)\Big) = \partial T\,,$$

so, because T is mass-minimizing, we have

$$\mathbf{M}[T] \le \mathbf{M}[\,T + C - T \llcorner \mathbb{B}(p,r)\,]$$

and we conclude that

$$\mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,] \le \mathbf{M}[C] \le \frac{r}{M}\,\mathbf{M}[\,\partial(T \llcorner \mathbb{B}(p,r))\,]$$

holds. □

The upper bound (8.29) for the mass of a mass-minimizer inside a ball is interesting, but the reader may have noticed the absence of a bound for the quantity on the right-hand side of (8.29). The next lemma, which follows readily from Lemma 7.6.3, provides that missing bound.

Lemma 8.4.2. *If $T \in \mathcal{D}_M(\mathbb{R}^N)$ is an integer-multiplicity current, $p \in \operatorname{spt} T$, and $\mathbb{B}(p,R) \cap \operatorname{spt} \partial T = \emptyset$, where $0 < R$, then for $\mathcal{L}^1$-almost every $0 < r < R$, it holds that*

$$\mathbf{M}[\,\partial(T \llcorner \mathbb{B}(p,r)\,] \le \frac{d}{dr}\,\mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,]\,. \tag{8.32}$$

The remarkable fact is that by combining Lemma 8.4.1 and Lemma 8.4.2, we can obtain the lower bound on the density of a mass-minimizing current given in the next theorem.

Theorem 8.4.3. *If* $T \in \mathcal{D}_M(\mathbb{R}^N)$ *is a mass-minimizing integer-multiplicity current,* $p \in \operatorname{spt} T$*, and* $\mathbb{B}(p, R) \cap \operatorname{spt} \partial T = \emptyset$*, where* $0 < R$*, then*

$$\Omega_M \, r^M \leq \mathbf{M}[\, T \llcorner \mathbb{B}(p, r)\,] \tag{8.33}$$

holds, for $0 < r < R$.

Proof. Define $\phi : (0, R) \to \mathbb{R}$ by setting

$$\phi(r) = \mathbf{M}[\, T \llcorner \mathbb{B}(p, r)\,] \,.$$

Then ϕ is a nondecreasing function and (8.29) and (8.32) tell us that

$$\phi(r) \leq \frac{r}{M}\, \phi'(r)$$

holds, for $\mathcal{L}^1$-almost every $0 < r < R$.

Now choose $0 < r_0 < r < R$. Since

$$\begin{aligned}\log r^M - \log r_0^M &= \int_{r_0}^{r} \frac{M}{\rho}\, d\mathcal{L}^1(\rho) \leq \int_{r_0}^{r} \Big(\log \circ \phi\Big)'(\rho)\, d\mathcal{L}^1(\rho) \\ &\leq \Big(\log \circ \phi\Big)(r) - \Big(\log \circ \phi\Big)(r_0)\,,\end{aligned}$$

we conclude that

$$\frac{\mathbf{M}[\, T \llcorner \mathbb{B}(p, r_0)\,]}{r_0^M} \leq \frac{\mathbf{M}[\, T \llcorner \mathbb{B}(p, r)\,]}{r^M}\,. \tag{8.34}$$

Fixing $0 < r < R$ and letting $r_0 \downarrow 0$ in (8.34), we see that

$$\Theta^*_M(\|T\|, p)\; \Omega_M \, r^M \leq \mathbf{M}[\, T \llcorner \mathbb{B}(p, r)\,] \tag{8.35}$$

holds. Replacing p in (8.35) by a nearby $q \in \operatorname{spt} T$ for which $1 \leq \Theta_M(\|T\|, q)$ is true, we obtain

$$\Omega_M \,(r - |p - q|)^M \leq \mathbf{M}[\, T \llcorner \mathbb{B}(p, r - |p - q|)\,]\,. \tag{8.36}$$

Finally, letting $q \to p$ in (8.36), we obtain (8.33). □

The inequality (8.34) expresses the monotonicity of the density of an M-dimensional area-minimizing surface. In fact, the monotonicity property holds very generally for surfaces that are extremal with respect to the area integrand (see for instance [All 72, 5.1(1)]). Allard has also shown in [All 74] that the methods used to prove monotonicity for surfaces that are extremal for the area integrand will not extend to more general integrands.

The preceding paragraph notwithstanding, a lower bound on density does hold for surfaces that minimize more general variational integrals. In the general case, the comparison surface used is not the cone, but rather the surface whose existence is guaranteed by the isoperimetric inequality.

Lemma 8.4.4. *Fix* $0 < \lambda < 1$. *Let* F *be an* M*-dimensional parametric integrand on* $\mathbb{R}^N$ *satisfying the bounds*

$$\lambda|\omega| \le F(x,\omega) \le \lambda^{-1}|\omega|\,, \tag{8.37}$$

for $x \in \mathbb{R}^N$ *and* $\omega \in \bigwedge_M(\mathbb{R}^N)$.

If $T \in \mathcal{D}_M(\mathbb{R}^N)$ *is an* F*-minimizing integer-multiplicity current,* $p \in \operatorname{spt} T$, *and* $\mathbb{B}(p,r) \cap \operatorname{spt} \partial T = \emptyset$, *where* $0 < r$, *then*

$$\mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,] \le \lambda^{-2}\, C_{M,N} \left(\mathbf{M}[\,\partial(T \llcorner \mathbb{B}(p,r)\,]\right)^{M/(M-1)}. \tag{8.38}$$

Here $C_{M,N}$ *is the constant in the isoperimetric inequality for* $(M-1)$*-dimensional boundaries and* M*-dimensional surfaces in* $\mathbb{R}^N$.

Proof. By the isoperimetric inequality, there is an integer-multiplicity current Q with $\partial Q = \partial(T \llcorner \mathbb{B}(p,r)$ and

$$\mathbf{M}[Q] \le C_{M,N} \left(\mathbf{M}[\,\partial(T \llcorner \mathbb{B}(p,r)\,]\right)^{M/(M-1)}.$$

Using (8.37), we obtain

$$\begin{aligned}
\mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,] &\le \lambda^{-1} \int_{T \llcorner \mathbb{B}(p,r)} F \\
&\le \lambda^{-1} \int_Q F \\
&\le \lambda^{-2}\, \mathbf{M}[Q] \le \lambda^{-2}\, C_{M,N} \left(\mathbf{M}[\,\partial(T \llcorner \mathbb{B}(p,r)\,]\right)^{M/(M-1)}.
\end{aligned}$$ □

By combining Lemma 8.4.2 and Lemma 8.4.4, we obtain the next theorem.

Theorem 8.4.5. *Fix* $0 < \lambda < 1$. *Let* F *be an* M*-dimensional parametric integrand on* $\mathbb{R}^N$ *satisfying the bounds*

$$\lambda|\omega| \le F(x,\omega) \le \lambda^{-1}|\omega|\,,$$

for $x \in \mathbb{R}^N$ *and* $\omega \in \bigwedge_M(\mathbb{R}^N)$.

If $T \in \mathcal{D}_M(\mathbb{R}^N)$ *is an* F*-minimizing integer-multiplicity current,* $p \in \operatorname{spt} T$, *and* $\mathbb{B}(p,R) \cap \operatorname{spt} \partial T = \emptyset$, *where* $0 < R$, *then*

$$M^{-M}\, \lambda^{2(M-1)}\, C_{M,N}^{(1-M)}\, r^M \le \mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,] \tag{8.39}$$

holds, for $0 < r < R$.

Proof. As in the proof of Theorem 8.4.3, we define $\phi : (0,R) \to \mathbb{R}$ by setting

$$\phi(r) = \mathbf{M}[\,T \llcorner \mathbb{B}(p,r)\,]\,.$$

Then ϕ is a nondecreasing function and (8.38) and (8.32) tell us that

$$\phi(r) \leq \lambda^{-2}\, C_{M,N}\, \big[\phi'(r)\big]^{M/(M-1)}$$

or, equivalently,

$$\lambda^{2(M-1)/M}\, C_{M,N}^{(1-M)/M} \leq \Big[\phi(r)\Big]^{(1-M)/M} \phi'(r) = M\, \frac{d}{dr} \Big[\phi(r)\Big]^{1/M}$$

holds, for $\mathcal{L}^1$-almost every $0 < r < R$.

Now fix $0 < r < R$. Since we have

$$\begin{aligned} M^{-1}\, \lambda^{2(M-1)/M}\, C_{M,N}^{(1-M)/M}\, r &= \int_0^r M^{-1}\, \lambda^{2(M-1)/M}\, C_{M,N}^{(1-M)/M}\, d\rho \\ &\leq \int_0^r M^{-1} \frac{d}{d\rho} \Big[\phi(\rho)\Big]^{1/M} d\rho \\ &\leq \Big[\phi(r)\Big]^{1/M}, \end{aligned}$$

(8.39) follows. □

Theorem 8.4.5 applies to an integer-multiplicity current that minimizes an elliptic integrand. The theorem gives us a lower bound on the mass of the minimizing current T in any ball that is centered in the support of T and that does not intersect the support of ∂T. Remarkable as Theorem 8.4.5 is, Theorem 8.4.3, which applies to mass-minimizing currents, gives an even larger, and in fact optimal, lower bound for the mass in a ball.

9

Regularity of Mass-Minimizing Currents

In the last chapter we proved the existence of solutions to certain variational problems in the context of integer-multiplicity rectifiable currents. In this chapter, we address the question of whether such solutions are in fact smooth surfaces. Such a question is quite natural: Indeed, Hilbert's 19th problem asked [Hil 02], "Are the solutions of regular problems in the calculus of variations always necessarily analytic?''

While Hilbert proposed his famous problems in 1900, the earliest precursors of currents as a tool for solving variational problems are the generalized curves of Laurence Chisholm Young (1905–2000) [You 37]. So of course, Hilbert could not have been been referring to variational problems in the context of integer-multiplicity currents.

Sets of finite perimeter are essentially equivalent to codimension-one integer-multiplicity rectifiable currents. It was Ennio de Giorgi (1928–1996) [DGi 61a], [DGi 61b] who first proved the existence and almost-everywhere regularity of area-minimizing sets of finite perimeter. Subsequently, Ernst Robert Reifenberg (1928–1964) [Rei 64a], [Rei 64b] proved the almost-everywhere regularity of area-minimizing surfaces in higher codimensions.

Later work of W. Fleming [Fle 62], E. De Giorgi [DGi 65], Frederick Justin Almgren, Jr. (1933–1997) [Alm 66], J. Simons [Sis 68], E. Bombieri, E. De Giorgi, and E. Giusti [BDG 69], and H. Federer [Fed 70], led to the definitive result that states that, in $\mathbb{R}^N$, an $(N-1)$-dimensional mass-minimizing integer-multiplicity current is a smooth, embedded manifold in its interior, except for a singular set of Hausdorff dimension at most $N-8$.

The extension of the regularity theory to general elliptic integrands was made by Almgren [Alm 68]. His result is that an integer-multiplicity current that minimizes the integral of an elliptic integrand is regular on an open dense set. Later work of Almgren, R. Schoen, and L. Simon [SSA 77] gave a stronger result in codimension one.

In our exposition, we will limit the scope of what we prove in favor of including more detail. Specifically, we will limit our attention to the area integrand and to codimension-one surfaces. An advantage of this approach is that we can include a complete derivation of the needed a priori estimates. Our exposition is based on the direct argument of R. Schoen and L. Simon [SS 82].

G.S. Krantz, R.H. Parks (eds.), *Geometric Integration Theory,* doi: 10.1007/978-0-8176-4679-0_9,

9.1 Preliminaries

Notation 9.1.1.

(1) We let M be a positive integer, $M \geq 2$.
(2) We identify $\mathbb{R}^{M+1}$ with $\mathbb{R}^M \times \mathbb{R}$ and let $\mathbf{p}$ be the projection onto $\mathbb{R}^M$ and $\mathbf{q}$ be the projection onto $\mathbb{R}$.
(3) We let $\mathbb{B}^M(y, \rho)$ denote the open ball in $\mathbb{R}^M$ of radius ρ, centered at y. The closed ball of radius ρ, centered at y, will be denoted by $\overline{\mathbb{B}}^M(y, \rho)$.
(4) The cylinder $\mathbb{B}^M(y, \rho) \times \mathbb{R}$ will be denoted by $\mathrm{C}(y, \rho)$ and its closure by $\overline{\mathrm{C}}(y, \rho)$.
(5) Recall that $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{M+1}$ is the standard basis for $\mathbb{R}^{M+1}$, and $dx_1, dx_2, \ldots, dx_{M+1}$ is the dual basis in $\bigwedge^1 \mathbb{R}^{M+1}$.
(6) As basis elements for $\bigwedge_M \mathbb{R}^{M+1}$ we will use

$$\mathbf{e}_{\widehat{1}}, \mathbf{e}_{\widehat{2}}, \ldots, \mathbf{e}_{\widehat{M+1}}, \tag{9.1}$$

where

$$\mathbf{e}_{\widehat{i}} = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_{i-1} \wedge \mathbf{e}_{i+1} \wedge \cdots \wedge \mathbf{e}_{M+1} .$$

Since the M-dimensional subspace associated with $\mathbf{e}_{\widehat{M+1}}$ will play a special role in what follows, we will also use the notation

$$\mathbf{e}^M = \mathbf{e}_{\widehat{M+1}} = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \cdots \wedge \mathbf{e}_M .$$

(7) We will identify $\bigwedge^M \mathbb{R}^{M+1}$ and the dual space of $\bigwedge_M \mathbb{R}^{M+1}$ using the standard isomorphism. Thus we will write $\langle \phi, \eta \rangle$ and $\phi(\eta)$ interchangeably when $\eta \in \bigwedge_M \mathbb{R}^{M+1}$ and $\phi \in \bigwedge^M \mathbb{R}^{M+1} \simeq \left[\bigwedge_M \mathbb{R}^{M+1}\right]'$.
(8) We set

$$dx_{\widehat{i}} = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_{M+1} \tag{9.2}$$

for $i = 1, 2, \ldots, M+1$. We will also use the notation

$$dx^M = dx_{\widehat{M+1}} = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_M . \tag{9.3}$$

Definition 9.1.2.

(1) According to the definition given in Example 8.3.6(1), the M-dimensional area integrand on $\mathbb{R}^{M+1}$ is a function on $\mathbb{R}^{M+1} \times \bigwedge_M \mathbb{R}^{M+1}$, but a function that is in fact independent of the first component of the argument. For simplicity of notation, we will consider the M-dimensional area integrand to be a function only on $\bigwedge_M \mathbb{R}^{M+1}$, so that

$$A : \bigwedge\nolimits_M \mathbb{R}^{M+1} \to \mathbb{R}$$

is given by

$$A(\xi) = |\xi|$$

for $\xi \in \bigwedge_M \mathbb{R}^{M+1}$.

(2) The M-dimensional area functional $\mathbf{A}$ is defined by setting

$$\mathbf{A}(S) = \int A\left(\overrightarrow{S}(x)\right) d\|S\|(x)$$

whenever S is an M-dimensional current representable by integration. We also have $\mathbf{A}(S) = \mathbf{M}(S) = \|S\|(\mathbb{R}^{M+1})$. Of course, the area integrand is called that because, when S is the current associated with a classical M-dimensional surface, then $\mathbf{A}(S)$ equals the area of that surface.

Next we will calculate the first and second derivatives of the area integrand and note some important identities.

Using the basis (9.1), we find that if $\xi = \sum_{i=1}^{M} \xi_i \mathbf{e}_{\hat{\imath}}$, then

$$A(\xi) = \sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_{M+1}^2}\,; \tag{9.4}$$

so the derivative of the area integrand, DA, is represented by the 0-by-$(M+1)$ matrix

$$DA(\xi) = \Big(\xi_1/|\xi|,\ \xi_2/|\xi|, \ldots,\ \xi_{M+1}/|\xi|\Big)\,. \tag{9.5}$$

That is,

$$\langle\, DA(\xi),\ \eta\,\rangle = (\xi \cdot \eta)/|\xi| \tag{9.6}$$

holds for $\xi, \eta \in \bigwedge_M \mathbb{R}^{M+1}$, or equivalently, we have

$$DA(\xi) = |\xi|^{-1} \sum_{i=1}^{M+1} \xi_i \, dx_{\hat{\imath}}\,. \tag{9.7}$$

In particular, we have

$$DA(\mathbf{e}_{\hat{\imath}}) = dx_{\hat{\imath}}\,. \tag{9.8}$$

We see that the second derivative of the area integrand, D^2A, is represented by the Hessian matrix

$$\begin{aligned} D^2A(\xi) = {} & |\xi|^{-1} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \\ & - |\xi|^{-3} \begin{pmatrix} \xi_1^2 & \xi_1\xi_2 & \ldots & \xi_1\xi_{M+1} \\ \xi_2\xi_1 & \xi_2^2 & \ldots & \xi_2\xi_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{M+1}\xi_1 & \xi_{M+1}\xi_2 & \ldots & \xi_{M+1}^2 \end{pmatrix}. \end{aligned} \tag{9.9}$$

Equivalently, for the partial derivatives $\partial^2 A/\partial\xi_i\partial\xi_j = D_{\xi_i\,\xi_j}A$, we have

$$D_{\xi_i \, \xi_j} A(\xi) = |\xi|^{-3} \, (|\xi|^2 \, \delta_{i\,j} - \xi_i \, \xi_j) \, , \tag{9.10}$$

where $\delta_{i\,j}$ is the Kronecker delta.[1]

Using (9.10), we can compute the Hilbert–Schmidt norm of $D^2 A$ as follows:

$$\begin{aligned} |D^2 A(\xi)|^2 &= \sum_{i,j=1}^{M+1} [D_{\xi_i \, \xi_j} A(\xi)]^2 \\ &= |\xi|^{-6} \sum_{i,j=1}^{M+1} \Big[|\xi|^2 \, \delta_{i\,j} - \xi_i \, \xi_j \Big]^2 \\ &= |\xi|^{-6} \sum_{i,j=1}^{M+1} \Big[|\xi|^4 \, \delta_{i\,j} - 2 \, |\xi|^2 \, \xi_i \, \xi_j \, \delta_{i\,j} + \xi_i^2 \, \xi_j^2 \Big] \\ &= |\xi|^{-6} \Big[(M+1) \, |\xi|^4 - 2 \, |\xi|^4 + |\xi|^4 \Big] \\ &= M \, |\xi|^{-2} \, . \end{aligned}$$

So we have

$$|D^2 A| = \sqrt{M} / |\xi| \, . \tag{9.11}$$

We note that

$$\frac{1}{2} \, |\xi - \eta|^2 = A(\eta) - \langle \, DA(\xi), \; \eta \, \rangle, \; \text{ for } |\xi| = |\eta| = 1 \, . \tag{9.12}$$

Equation (9.12) follows because

$$\begin{aligned} \frac{1}{2} |\xi - \eta|^2 &= \frac{1}{2} \Big(|\xi|^2 - 2\xi \cdot \eta + |\eta|^2 \Big) \\ &= 1 - \xi \cdot \eta \\ &= |\eta| - (\xi \cdot \eta)/|\xi| \\ &= A(\eta) - \langle \, DA(\xi), \; \eta \, \rangle \, , \end{aligned}$$

where the last equality follows from (9.6).

Equation (9.12) will play an important role in the regularity theory, but it is the inequality

$$\frac{1}{2} \, |\xi - \eta|^2 \leq A(\eta) - \langle \, DA(\xi), \; \eta \, \rangle, \; \text{ for } |\xi| = |\eta| = 1, \tag{9.13}$$

[1] Leopold Kronecker (1823–1891).

that is essential. Any inequality of the form (9.13) (but with $\frac{1}{2}$ possibly replaced by another positive constant) is called a *Weierstrass condition*. Ellipticity of an integrand is equivalent to the integrand satisfying a Weierstrass condition (see [Fed 75, Section 3]).

Definition 9.1.3. We say that the M-dimensional integer-multiplicity current T is *mass-minimizing* if

$$\mathbf{A}(T) \leq \mathbf{A}(S) \tag{9.14}$$

holds whenever $S \in \mathcal{D}_M(\mathbb{R}^{M+1})$ is integer-multiplicity with $\partial S = \partial T$.

When a current is projected into a plane, the mass of the projection is less than or equal to the mass of the original current. The difference between the two masses is the "excess" (see Figure 9.1). The fundamental quantity used in the regularity theory is the "cylindrical excess," which is the excess of the part of a current in a cylinder, normalized to account for the radius of the cylinder. We give the precise definition next.

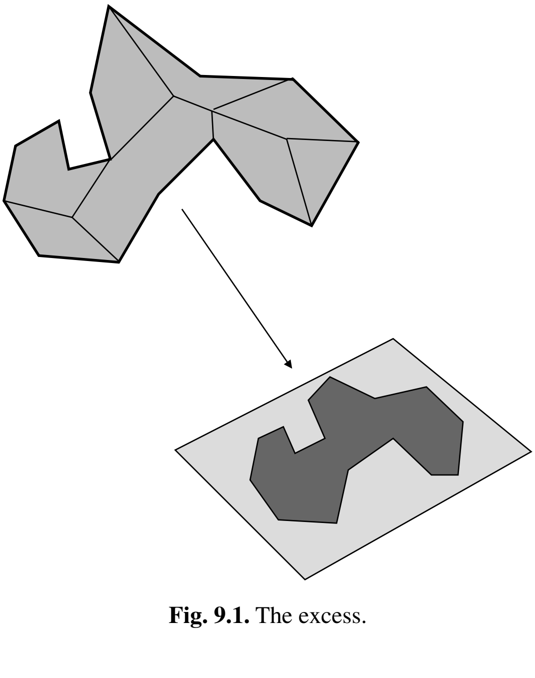

Fig. 9.1. The excess.

Definition 9.1.4. For an integer-multiplicity $T \in \mathcal{D}_M(\mathbb{R}^{M+1})$, $y \in \mathbb{R}^M$, and $\rho > 0$, the *cylindrical excess* $E(T, y, \rho)$ is defined by

$$E(T, y, \rho) = \frac{1}{2}\,\rho^{-M} \int_{C(y,\rho)} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| \,, \tag{9.15}$$

where we recall that

$$T = \|T\| \wedge \overrightarrow{T} .$$

The next lemma shows the connection between equation (9.15), which defines the excess, and the more heuristic description of the excess given before the definition.

Lemma 9.1.5. *Suppose that $T \in \mathcal{D}_M(\mathbb{R}^{M+1})$ is integer-multiplicity, $y \in \mathbb{R}^M$, ℓ is a positive integer, and $\rho > 0$. If*

$$\mathbf{p}_\#(T \llcorner \mathrm{C}(y, \rho)) = \ell\, \mathbf{E}^M \llcorner \mathbb{B}^M(y, \rho)$$

and $\operatorname{spt} \partial T \subseteq \mathbb{R}^{M+1} \setminus \mathrm{C}(y, \rho)$*, then it holds that*

$$\begin{aligned} E(T, y, \rho) &= \rho^{-M} \Big(\|T\|(\mathrm{C}(y, \rho)) - \|\mathbf{p}_\# T\|(\mathbb{B}^M(y, \rho)) \Big) \\ &= \rho^{-M} \left(\|T\|(\mathrm{C}(y, \rho)) - \ell\, \Omega_M\, \rho^M \right) . \end{aligned} \tag{9.16}$$

Proof. Since $|\overrightarrow{T}| = |\mathbf{e}^M| = 1$, we have

$$\begin{aligned} |\overrightarrow{T} - \mathbf{e}^M|^2 &= |\overrightarrow{T}|^2 + |\mathbf{e}^M|^2 - 2 \left(\overrightarrow{T} \cdot \mathbf{e}^M \right) \\ &= 2 - 2 \left(\overrightarrow{T} \cdot \mathbf{e}^M \right) . \end{aligned}$$

So we have

$$\begin{aligned} \frac{1}{2} \int_{\mathrm{C}(y,\rho)} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| &= \int_{\mathrm{C}(y,\rho)} 1 - \left(\overrightarrow{T} \cdot \mathbf{e}^M \right) d\|T\| \\ &= \|T\|(\mathrm{C}(y, \rho)) - \|\mathbf{p}_\# T\|(\mathbb{B}^M(y, \rho)) \\ &= \|T\|(\mathrm{C}(y, \rho)) - \ell\, \Omega_M\, \rho^M . \end{aligned}$$ □

We now give two corollaries of the lemma. The first is an immediate consequence of the proof of Lemma 9.1.5 and the second shows us the effect of an isometry on the excess.

Corollary 9.1.6. *Suppose that $T \in \mathcal{D}_M(\mathbb{R}^{M+1})$ is integer-multiplicity, $y \in \mathbb{R}^M$, ℓ is a positive integer, and $\rho > 0$. If*

$$\mathbf{p}_\#(T \llcorner \mathrm{C}(y, \rho)) = \ell\, \mathbf{E}^M \llcorner \mathbb{B}^M(y, \rho)$$

and $\operatorname{spt} \partial T \subseteq \mathbb{R}^{M+1} \setminus \mathrm{C}(y, \rho)$*, then for any $\mathcal{L}^M$-measurable $B \subseteq \mathbb{B}^M(y, \rho)$, it holds that*

$$\|T\|(B \times \mathbb{R}) \le \frac{1}{2} \int_{B \times \mathbb{R}} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| + \ell\, \mathcal{L}^M(B) . \tag{9.17}$$

Proof. The corollary is an immediate consequence of the proof of the lemma. □

Corollary 9.1.7. *Suppose that $T \in \mathcal{D}_M(\mathbb{R}^{M+1})$ is integer-multiplicity, $\rho > 0$,*

$$\mathbf{p}_\#(T \llcorner \mathrm{C}(0,\rho)) = \ell\, \mathbf{E}^M \llcorner \mathbb{B}^M(0,\rho)\,,$$

and $\operatorname{spt} \partial T \subseteq \mathbb{R}^{M+1} \setminus \mathrm{C}(0,\rho)$.

If $1 < \lambda < \infty$, $\mathbf{j} : \mathbb{R}^{M+1} \to \mathbb{R}^{M+1}$ *is an isometry,* $0 < \rho' < \rho$, *and*

$$\operatorname{spt} \mathbf{j}_\# T \llcorner \mathrm{C}(0,\rho') \subseteq \mathbf{j}\Big(\operatorname{spt} T \llcorner \mathrm{C}(0,\rho)\Big)\,,$$

then

$$\begin{aligned} E(\mathbf{j}_\# T,\ 0,\ \rho') \le{}& \lambda\,(\rho/\rho')^M\, E(T,\ 0,\ \rho) \\ &+ \frac{\lambda}{2(\lambda-1)} \cdot (\rho/\rho')^M \cdot \ell \cdot \|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\|^{2M} \cdot E(T,\ 0,\ \rho) \\ &+ \frac{\lambda\,\ell\,\Omega_M}{2(\lambda-1)} \cdot (\rho/\rho')^M \cdot \|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\|^{2M}\,. \end{aligned}$$

Proof. Using

$$\left| \textstyle\bigwedge_M \mathbf{j}\left(\overrightarrow{T}\right) - \mathbf{e}^M \right| \le \left| \textstyle\bigwedge_M \mathbf{j}\left(\overrightarrow{T}\right) - \textstyle\bigwedge_M \mathbf{j}\left(\mathbf{e}^M\right) \right| + \left| \textstyle\bigwedge_M \mathbf{j}\left(\mathbf{e}^M\right) - \mathbf{e}^M \right|$$

and

$$\begin{aligned} (\,|\alpha| + |\beta|\,)^2 &= \lambda\,\alpha^2 + \frac{\lambda}{\lambda-1}\,\beta^2 - \left(\sqrt{\lambda-1}\,|\alpha| - |\beta|/\sqrt{\lambda-1}\right)^2 \\ &\le \lambda\,\alpha^2 + \frac{\lambda}{\lambda-1}\,\beta^2\,, \end{aligned}$$

we obtain

$$\begin{aligned} E(\mathbf{j}_\# T,\ 0,\ \rho') &\le \frac{1}{2}\,(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \textstyle\bigwedge_M \mathbf{j}\left(\overrightarrow{T}\right) - \mathbf{e}^M \right|^2 d\|T\| \\ &\le \frac{\lambda}{2}(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \textstyle\bigwedge_M \mathbf{j}\left(\overrightarrow{T}\right) - \textstyle\bigwedge_M \mathbf{j}\left(\mathbf{e}^M\right) \right|^2 d\|T\| \\ &\qquad + \frac{\lambda}{2(\lambda-1)}(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \textstyle\bigwedge_M \mathbf{j}\left(\mathbf{e}^M\right) - \mathbf{e}^M \right|^2 d\|T\| \\ &= \frac{\lambda}{2}(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \overrightarrow{T} - \mathbf{e}^M \right|^2 d\|T\| \\ &\qquad + \frac{\lambda}{2(\lambda-1)}(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \textstyle\bigwedge_M \mathbf{j}\left(\mathbf{e}^M\right) - \mathbf{e}^M \right|^2 d\|T\| \end{aligned}$$

$$\leq \frac{\lambda}{2}(\rho')^{-M} \int_{\mathrm{C}(0,\rho)} \left| \overrightarrow{T} - \mathbf{e}^M \right|^2 d\|T\|$$

$$+ \frac{\lambda}{2(\lambda-1)}(\rho')^{-M} \| \mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}} \|^{2M} \|T\|\mathrm{C}(0,\rho)\,,$$

and the result follows from Lemma 9.1.5. □

Notation 9.1.8. Certain hypotheses will occur frequently in what follows, so we collect them here (with labels) for easy reference:

(H1) $\operatorname{spt} \partial T \subseteq \mathbb{R}^{M+1} \setminus \mathrm{C}(y,\rho)$,
(H2) $\mathbf{p}_{\#}[T \llcorner \mathrm{C}(y,\rho)] = \mathbf{E}^M \llcorner \mathbb{B}^M(y,\rho)$,
(H3) $\Omega_M\, r^M \leq \|T\|\{X \in \mathbb{R}^{M+1} : |X - Y| < r\}$ holds whenever $Y \in \operatorname{spt} T$ and $\{X \in \mathbb{R}^{M+1} : |X - Y| < r\} \cap \operatorname{spt} \partial T = \emptyset$,
(H4) $E(T, y, \rho) < \epsilon$,
(H5) T is mass-minimizing.

Here ρ and ϵ are positive and $y \in \mathbb{R}^M$.

Note that the constancy theorem, i.e., Proposition 7.3.1, implies that if $\operatorname{spt} T \subseteq \mathbb{R}^{M+1} \setminus \mathrm{C}(y,\rho)$, then, because $\partial \mathbf{p}_{\#} T = \mathbf{p}_{\#} \partial T$, we have

$$\mathbf{p}_{\#}(T \llcorner \mathrm{C}(y,\rho)) = \ell\, \mathbf{E}^M \llcorner \mathbb{B}^M(y,\rho)\,, \tag{9.18}$$

where ℓ is an integer. So in (H2) we are making the simplifying assumption that $\ell = 1$.

Note that (H5) allows us to apply Theorem 8.4.3 to obtain (H3), so (H3) is, in fact, a consequence of (H5).

9.2 The Height Bound and Lipschitz Approximation

We begin this section with the height bound lemma. The proof we give is simplified by using hypothesis (H3). While the height bound lemma remains true for currents minimizing the integral of an integrand other than area, the proof is more difficult because the lower bound on mass that they satisfy (see Theorem 8.4.5) is weaker than that in (H3).

Lemma 9.2.1 (Height bound). *For each σ with $0 < \sigma < 1$, there are $\epsilon_0 = \epsilon_0(M,\sigma)$ and $c_1 = c_1(M,\sigma)$ such that the hypotheses* (H1–H4), *with $\epsilon = \epsilon_0$ in* (H4), *imply*

$$\sup\Big\{ |\mathbf{q}(X_1) - \mathbf{q}(X_2)| : X_1, X_2 \in \operatorname{spt} T \cap \mathrm{C}(y,\sigma\rho) \Big\}$$

$$\leq c_1\, \rho \left(E(T,y,\rho) \right)^{\frac{1}{2M}}\,.$$

Proof. By using a translation and homothety if need be, we may assume that $y = 0$ and $\rho = 1$. We write

$$E = E(T, 0, 1)\,.$$

Set

$$r_0 = \tfrac{1}{2}(1 - \sigma) \tag{9.19}$$

and

$$\epsilon_0 = 2^{-M}\,\Omega_M\,(1 - \sigma)^M\,. \tag{9.20}$$

First we consider points whose projections onto $\mathbb{B}^M(0, 1)$ are separated by a distance less than $2\,r_0$. So suppose that $X_1, X_2 \in \operatorname{spt} T \cap \mathrm{C}(0, \sigma)$ are such that

$$\tfrac{1}{2}\Big|\mathbf{p}\,(X_1) - \mathbf{p}\,(X_2)\Big| < r_0\,.$$

We set

$$r = \tfrac{1}{2}\Big|\mathbf{p}\,(X_1) - \mathbf{p}\,(X_2)\Big|, \qquad h = \tfrac{1}{2}\Big|\mathbf{q}\,(X_1) - \mathbf{q}\,(X_2)\Big|\,.$$

Then we have

$$\Big|X_1 - X_2\Big| = 2\sqrt{r^2 + h^2}\,.$$

We set

$$s = \min\{\sqrt{r^2 + h^2} - r\,,\ r_0\,\}\,.$$

Then we have

$$\mathbb{B}(X_1, r + s) \bigcap \mathbb{B}(X_2, r + s) = \emptyset$$

and

$$\mathbb{B}(X_1, r + s) \bigcup \mathbb{B}(X_2, r + s) \subseteq \mathrm{C}(0, 1)\,.$$

Setting

$$x^* = \tfrac{1}{2}\,(\mathbf{p}\,(X_1) + \mathbf{p}\,(X_2))\ ,$$

so that

$$\big|\mathbf{p}\,(X_1) - x^*\big| = \big|\mathbf{p}\,(X_2) - x^*\big| = r\,,$$

we see (Figure 9.2) that

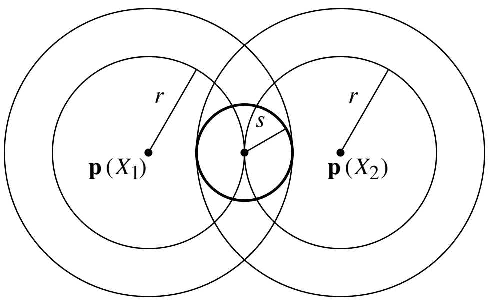

Fig. 9.2. The projections of the balls.

$$\mathbb{B}^M(x^*, s) \subseteq \mathbf{p}\,(\mathbb{B}(X_1, r+s)) \bigcap \mathbf{p}\,(\mathbb{B}(X_2, r+s))$$

and thus that

$$\mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s)) \bigcap \mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big] \geq \Omega_M\, s^M\,.$$

By (H3) we have

$$\begin{aligned}\|T\|\,\mathbb{B}(X_1, r+s) + \|T\|\,\mathbb{B}(X_2, r+s) &\geq 2\,\Omega_M\,(r+s)^M \\ &= \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s))\Big] + \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big].\end{aligned}$$

Thus we have

$$\begin{aligned}E &\geq \|T\|\Big[\mathbb{B}(X_1, r+s) \bigcup \mathbb{B}(X_2, r+s)\Big] \\ &\qquad - \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s)) \bigcup \mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big] \\ &\geq \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s))\Big] + \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big] \\ &\qquad - \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s)) \bigcup \mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big] \\ &= \mathcal{L}^M\Big[\mathbf{p}\,(\mathbb{B}(X_1, r+s)) \bigcap \mathbf{p}\,(\mathbb{B}(X_2, r+s))\Big] \geq \Omega_M\, s^M\,.\end{aligned}$$

We now consider two possibilities.

Case 1. $s = r_0$,
Case 2. $s = \sqrt{r^2 + h^2} - r < r_0$.

In Case 1, by the definition of r_0, i.e., (9.19), the definition of ϵ_0, i.e., (9.20), and (H4), we have

$$E \geq \Omega_M\, s^M = \Omega_M\, r_0^M = 2^{-M}\,\Omega_M\,(1-\sigma)^M = \epsilon_0 > E\,,$$

a contradiction. Thus we may assume that Case 2 holds.

In Case 2, we note that

$$\begin{aligned}h &\leq \sqrt{r^2 + h^2} \\ &\leq (\sqrt{r^2 + h^2} - r) + r_0 \\ &\leq 2\, r_0\,.\end{aligned}$$

Then it follows that

$$
\begin{aligned}
E &\geq \Omega_M\, s^M \\
&= \Omega_M\, (\sqrt{r^2+h^2}-r)^M \\
&= \Omega_M\, \left(\frac{(r^2+h^2)-r^2}{\sqrt{r^2+h^2}+r}\right)^M \\
&\geq \Omega_M\, \left(\frac{h^2}{\sqrt{r_0^2+4r_0^2}+r_0}\right)^M \\
&\geq \Omega_M\, 2^{-M}\,(1-\sigma)^{-M}\, h^{2M}\,,
\end{aligned}
$$

where we obtain the last inequality by using the definition of r_0, i.e., (9.19), and, for simplicity, we have replaced $\sqrt{5}+1$ by the larger number 4.

We have shown that any two points in spt $T \cap \mathrm{C}(0,\sigma)$ whose projections onto $\mathbb{B}^M(0,1)$ are separated by a distance less than $2\,r_0$ will have their projections by $\mathbf{q}$ separated by less than

$$
2^{1/2}\,\Omega_M^{-1/(2M)}\,(1-\sigma)^{1/2}\,E^{1/(2M)}\,.
$$

But any two points x_1 and x_2 in $\mathbb{B}^M(0,\sigma)$ are separated by a distance less than $2\,\sigma$, so if the two points are separated by more than $2\,r_0 = (1-\sigma)$, then we can form a sequence of points $z_1 = x_1, z_2, \ldots, z_M = x_2$ such that $|z_{i+1}-z_i| \leq (1-\sigma) = 2r_0$. We can take L to be the smallest integer exceeding $2\,\sigma/(1-\sigma)$. Thus we have

$$
L \leq 1 + \frac{2\,\sigma}{1-\sigma} = \frac{1+\sigma}{1-\sigma} < \frac{2}{1-\sigma}\,.
$$

Hence we may set

$$
\begin{aligned}
c_1(M,\sigma) &= L \cdot 2^{1/2}\,\Omega_M^{-1/(2M)}\,(1-\sigma)^{1/2} \\
&\leq 2^{3/2}\,\Omega_M^{-1/(2M)}\,(1-\sigma)^{-1/2}\,.
\end{aligned}
$$

□

Lemma 9.2.2 (Lipschitz approximation). *Let γ with $0 < \gamma \leq 1$ be given. There exist constants c_2, c_3, and c_4 such that the following holds:*

If the hypotheses (H1–H4) *are satisfied with $\epsilon = \epsilon_0(M, 2/3)$ in* (H4), *where $\epsilon_0(M, 2/3)$ is as in Lemma* 9.2.1, *then there is a Lipschitz function $g : \mathbb{B}^M(y,\rho/4) \to \mathbb{R}$ satisfying the following conditions:*

$$
\operatorname{Lip} g \leq \gamma, \tag{9.21}
$$

$$
\sup\Big\{|g(z)-g(y)| : z \in \mathbb{B}^M(y,\rho/4)\Big\} \leq c_2\,\rho\,\Big(E(T,y,\rho)\Big)^{\frac{1}{2M}}, \tag{9.22}
$$

$$\mathcal{L}^M\Big[\,\mathbb{B}^M(y,\rho/4)\setminus\Big\{z\in\mathbb{B}^M(y,\rho/4):\mathbf{p}^{-1}(z)\cap\operatorname{spt}T=\{(z,g(z))\}\Big\}\Big]$$

$$\leq\ \rho^M\,c_3\,\gamma^{-2M}\,E(T,y,\rho), \tag{9.23}$$

$$\|T-T^g\|\mathrm{C}(y,\rho/4)\leq\rho^M\,c_4\,\gamma^{-2M}\,E(T,y,\rho)\,, \tag{9.24}$$

where

$$T^g=G_\#\Big(\mathbf{E}^M\llcorner\mathbb{B}^M(y,\rho/4)\Big)\,, \tag{9.25}$$

with $G:\mathbb{B}^M(y,\rho/4)\to\mathrm{C}(y,\rho/4)$ *defined by*

$$G(x)=(x,g(x))\,,\qquad \textit{for } x\in\mathbb{B}^M(y,\rho/4)\,.$$

Proof. Fix the choice of $0<\gamma\leq 1$ and specify a value of ϵ_0 for which the conclusion of Lemma 9.2.1 holds with σ chosen to equal $2/3$. That is, if the hypotheses (H1–H4) hold with $\epsilon=\epsilon_0$ and with z and δ in place of y and ρ, respectively, then

$$\sup\Big\{|\mathbf{q}\,(X_1)-\mathbf{q}\,(X_2)|:X_1,\ X_2\in\operatorname{spt}T\cap\mathrm{C}(z,2\delta/3)\Big\}$$

$$\leq\ c_1\,\delta\,\Big(E(T,z,\delta)\Big)^{\frac{1}{2M}}\,. \tag{9.26}$$

Consider η with

$$0<\eta<\epsilon_0\,. \tag{9.27}$$

Set

$$A=\Big\{z\in\mathbb{B}^M(y,\rho/4)\ :\ E(T,z,\delta)\leq\eta\text{ for all }\delta\text{ with }0<\delta<3\rho/4\Big\}\,, \tag{9.28}$$

and set

$$B=\mathbb{B}^M(0,\rho/4)\setminus A\,.$$

For each $b\in B$ there exists $\delta(b)$ with $0<\delta(b)<3\rho/4$ such that the excess $E(T,b,\delta(b))$ is greater than η, that is,

$$\frac{1}{2}\int_{\mathrm{C}(b,\delta(b))}|\overrightarrow{T}-\mathbf{e}^M|^2\,d\|T\|=\delta(b)^M\cdot E(T,b,\delta(b))>\eta\cdot\delta(b)^M\,. \tag{9.29}$$

Applying the Besicovitch covering theorem (i.e., Theorem 4.2.12) to the family of closed balls

$$\mathcal{B}=\Big\{\overline{\mathbb{B}}^M(b,\delta(b))\ :\ b\in B\Big\}\,,$$

we obtain the subfamilies $\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_K$ of $\mathcal{B}$ such that each $\mathcal{B}_i$ consists of pairwise disjoint balls and

$$B\subseteq\bigcup_{i=1}^K B_i\,,$$

where

$$B_i = \bigcup_{\overline{\mathbb{B}}^M(b,\delta(b))\in\mathcal{B}_i} \overline{\mathbb{B}}^M(b, \delta(b)) .$$

Here K is a number that depends only on the dimension M. Using (9.29), we see that, for each $i = 1, 2, \ldots, K$, we have

$$\begin{aligned}
\eta\,\mathcal{L}^M(B_i) &= \eta \sum_{\overline{\mathbb{B}}^M(b,\delta(b))\in\mathcal{B}_i} \Omega_M \Big[\, \delta(b) \,\Big]^M \\
&< \Omega_M \sum_{\overline{\mathbb{B}}^M(b,\delta(b))\in\mathcal{B}_i} \delta(b)^M \; E(T, b, \delta(b)) \\
&= \frac{1}{2}\,\Omega_M \int_{B_i} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| \\
&\le \frac{1}{2}\,\Omega_M \int_{\mathrm{C}(y,\rho)} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| .
\end{aligned}$$

We conclude that

$$\begin{aligned}
\eta\,\mathcal{L}^M(B) &\le \sum_{i=1}^{K} \eta\,\mathcal{L}^M\left(\bigcup_i B_i\right) \\
&\le \frac{K}{2}\,\Omega_M \int_{\mathrm{C}(y,\rho)} |\overrightarrow{T} - \mathbf{e}^M|^2 \, d\|T\| \\
&= c_5\,\rho^M \; E(T, y, \rho) . \qquad (9.30)
\end{aligned}$$

If $x_1, x_2 \in \mathbb{B}^M(0, \rho/4) \cap A$, and if X_1, X_2 are points with

$$X_i \in \operatorname{spt} T \cap \mathbf{p}^{-1}(x_i), \;\; i = 1, 2,$$

then

$$|x_1 - x_2| < \rho/2 ,$$

so we can apply (9.26) with $z = x_1$ and with δ chosen to satisfy

$$3\,|x_1 - x_2|/2 < \delta < 3\rho/4 . \qquad (9.31)$$

Letting δ in (9.31) decrease to $3\,|x_1 - x_2|/2$, we conclude that

$$|\mathbf{q}\,(X_1) - \mathbf{q}\,(X_2)| \le c_6\,\eta^{1/(2M)}\,|x_1 - x_2| , \qquad (9.32)$$

where we set

$$c_6 = \max\{\, 3/2,\;(3/2)\,c_1,\; \epsilon_0^{-1} \,\} . \qquad (9.33)$$

Thus we may choose

$$\eta = \gamma^{2M}\, c_6^{-2M} \le c_6^{-2M} < c_6^{-1} \le \epsilon_0\,, \tag{9.34}$$

so that $c_6\, \eta^{1/(2M)} = \gamma$ holds, and consequently we have

$$|\mathbf{q}\,(X_1) - \mathbf{q}\,(X_2)| \le \gamma\, |x_1 - x_2| \tag{9.35}$$

for any points

$$x_1, x_2 \in \mathbb{B}^M(0, \rho/4) \bigcap A\,,$$

where

$$X_1 \in \operatorname{spt} T \bigcap \mathbf{p}^{-1}(x_1) \ \text{ and } \ X_2 \in \operatorname{spt} T \bigcap \mathbf{p}^{-1}(x_2)\,.$$

In particular, (9.35) shows that, for any $x \in A \cap \mathbb{B}^M(0, \rho/4)$, there is exactly one $X \in \mathbf{p}^{-1}(x) \bigcap \operatorname{spt} T$. Thus, we can define $g^* : A \bigcap \mathbb{B}^M(0, \rho/4) \to \mathbb{R}$ by requiring

$$\Big\{ (x, g^*(x)) \Big\} = \mathbf{p}^{-1}(x) \bigcap \operatorname{spt} T, \quad \text{whenever } \ x \in A \bigcap \mathbb{B}^M(0, \rho/4)\,.$$

Inequality (9.35) tells us that $\operatorname{Lip}(g^*) \le \gamma$ holds on $A \bigcap \mathbb{B}^M(y, \rho/4)$, so by Kirszbraun's extension theorem (see [KPk 99, Theorem 5.2.2]) g^* extends to $g^{**} : \mathbb{B}^M(y, \rho/4) \to \mathbb{R}$ with the same Lipschitz constant.

By Lemma 9.2.1, if we set

$$g = \min \Big\{ \alpha,\ \max\{ \beta,\ g^{**} \} \Big\}\,,$$

where

$$\alpha = g(y) - c_1\, E^{1/(2M)}(T, y, \rho)\, \rho\,, \quad \beta = g(y) + c_1\, E^{1/(2M)}(T, y, \rho)\, \rho\,,$$

then

$$\Big\{ (x, g(x)) \Big\} = \mathbf{p}^{-1}(x) \bigcap \operatorname{spt} T \ \text{ whenever } \ x \in A \bigcap \mathbb{B}^M(0, \rho/4)$$

and

$$\sup \Big\{ |g(x) - g(y)| : \mathbb{B}^M(y, \rho/4) \Big\} \le c_1\, E^{1/(2M)}(T, y, \rho)\, \rho$$

will both hold.

Using (9.17), (9.30), and (9.34), we see that

$$\begin{aligned}
&\|T\| \Big[(\mathbb{B}^M(y, \rho/4) \setminus A) \times \mathbb{R} \Big] \\
&= \mathcal{L}^M \Big[\mathbb{B}^M(y, \rho/4) \setminus A \Big] + \frac{1}{2} \int_{(\mathbb{B}^M(y,\rho/4)\setminus A)\times\mathbb{R}} |\overrightarrow{T} - \mathbf{e}^M|^2\, d\|T\| \\
&\le \mathcal{L}^M[B] + \frac{1}{2} \int_{\mathrm{C}(y,\rho)} |\overrightarrow{T} - \mathbf{e}^M|^2\, d\|T\| \\
&\le (\eta^{-1} c_5 + 1)\, \rho^M\, E(T, y, \rho) \\
&= (c_5\, c_6^{2M}\, \gamma^{-2M} + 1)\, \rho^M\, E(T, y, \rho) \\
&\le (c_5\, c_6^{2M} + 1)\, \gamma^{-2M}\, \rho^M\, E(T, y, \rho)\,.
\end{aligned}$$

So we conclude that (9.23) holds with $c_3 = c_5\, c_6^{2M} + 1$.

Finally, we have

$$\begin{aligned}\|T - T^g\|\mathrm{C}(y,\rho/4) &\le \|T\|\Big[\, (\mathbb{B}^M(y,\rho/4)\setminus A)\times\mathbb{R}\,\Big] \\ &\qquad + \|T^g\|\Big[\, (\mathbb{B}^M(y,\rho/4)\setminus A)\times\mathbb{R}\,\Big] \\ &\le \|T\|\big[\, (\mathbb{B}^M(y,\rho/4)\setminus A)\times\mathbb{R}\,\Big] + \gamma\,\mathcal{L}^M[B] \\ &\le 2\,(c_5\, c_6^{2M} + 1)\,\gamma^{-2M}\,\rho^M\, E(T,y,\rho)\,,\end{aligned}$$

so we see that (9.24) holds with $c_4 = 2\,(c_5\, c_6^{2M} + 1)$. □

9.3 Currents Defined by Integrating over Graphs

Currents obtained by integration over the graph of a function are particularly nice and are helpful to our intuitive understanding of the concepts being developed here. We will show how the cylindrical excess of such a current relates to a familiar quantity from analysis, namely the Dirichlet integral (see Corollary 9.3.7).

Notation 9.3.1. Let $f : \mathbb{B}^M(0,\sigma) \to \mathbb{R}$ be Lipschitz.

(1) We use the notation F for the function from $\mathbb{B}^M(0,\sigma)$ to $\mathbb{R}^{M+1}$ given by $F(x) = (x, f(x))$.

(2) We use the notation G_F for the M-dimensional current that is defined by integration over the graph of f, that is,

$$G_F = F_\#(\mathbf{E}^M \llcorner \mathbb{B}^M(0,\sigma))\,.$$

Writing

$$J_F(x) = \langle\, \textstyle\bigwedge_M (DF(x)),\ \mathbf{e}^M\,\rangle\,,$$

we have

$$G_F[\psi] = \int_{\mathbb{B}^M(0,\sigma)} \langle\, \psi(x, f(x)),\ J_F(x)\,\rangle\, d\mathcal{L}^M(x) \tag{9.36}$$

for any differential M-form ψ defined on $\mathrm{C}(0,\sigma)$.

Lemma 9.3.2. *If $f : \mathbb{B}^M(0,\sigma) \to \mathbb{R}$ is Lipschitz, then we have*

$$\overrightarrow{G}_F(F(x)) = (1 + |Df|^2)^{-1/2}\left(\mathbf{e}^M + \sum_{i=i}^M \frac{\partial f}{\partial x_i}\,\mathbf{e}_{\hat{\imath}}\right), \tag{9.37}$$

$$DA(\overrightarrow{G}_F) = (1 + |Df|^2)^{-1/2}\left(dx^M + \sum_{i=1}^M \left(\frac{\partial f}{\partial x_i}\right) dx_{\hat{\imath}}\right), \tag{9.38}$$

$$DA(\overrightarrow{G}_F) - DA(\mathbf{e}^M) =$$

$$(1 + |Df|^2)^{-1/2} \left(dx^M + \sum_{i=1}^{M} \left(\frac{\partial f}{\partial x_i} \right) dx_{\hat{\imath}} \right) - dx^M . \tag{9.39}$$

Proof. By definition, we have

$$\langle \bigwedge\nolimits_M (DF(x)), \, \mathbf{e}^M \rangle = \bigwedge_{i=1}^{M} \left(\mathbf{e}_i + \frac{\partial f}{\partial x_i} \mathbf{e}_{M+1} \right) .$$

So

$$J_F = \mathbf{e}^M + \sum_{i=i}^{M} \frac{\partial f}{\partial x_i} \, \mathbf{e}_{\hat{\imath}} . \tag{9.40}$$

We obtain (9.37) from (9.40) by dividing by the norm of J_F. Equation (9.38) follows from (9.37) and (9.7). Equation (9.39) follows from (9.38) and (9.8). □

For the record, we note that the coefficient of dx^M in (9.39) is

$$(1 + |Df|^2)^{-1/2} - 1 .$$

Lemma 9.3.3. *Define a map from* $\mathbb{R}^M$ *to* $\mathbb{R}^{M+1}$ *by*

$$x = (x_1, x_2, \ldots, x_M) \longmapsto X = (1 + |x|^2)^{-1/2} \, (1, x_1, x_2, \ldots, x_M) .$$

If A and B are the images of a and b under this map then

(1) $|A - B| \le |a - b|$;
(2) *for each* $0 < c < \infty$, *it holds that*
$|a|, |b| \le c$ *implies* $|a - b| \le (1 + c^2)^2 \, |A - B|$.

Proof. The mapping $x \mapsto X$ is the composition of two mappings: the distance-preserving map

$$x = (x_1, x_2, \ldots, x_k) \longmapsto (1, x_1, x_2, \ldots, x_k)$$

followed by the radial projection onto the unit sphere

$$y = (y_1, y_2, \ldots, y_{k+1}) \longmapsto |y|^{-1} \, (y_1, y_2, \ldots, y_{k+1}) .$$

Part (1) follows from the fact that the radial projection does not increase the distance between points that are outside of the open unit ball.

To prove (2), we note that

$$|1 + a \cdot b| \le (1 + |a|^2)^{1/2} \, (1 + |b|^2)^{1/2}$$

holds, with equality if and only if $a = b$. Thus

$$0 < (1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2} + (1 + a \cdot b)$$

always holds, so we may compute

$$\begin{aligned}
&(1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2}\,|A - B|^2 \\
&\qquad = 2\Big[\,(1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2} - (1 + a \cdot b)\,\Big] \\
&\qquad = 2\Big[\,(1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2} + (1 + a \cdot b)\,\Big]^{-1} \\
&\qquad\qquad\qquad \cdot\Big[\,(1 + |a|^2)\,(1 + |b|^2) - (1 + a \cdot b)^2\,\Big] \\
&\qquad = 2\Big[\,(1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2} + (1 + a \cdot b)\,\Big]^{-1} \\
&\qquad\qquad\qquad \cdot\Big[\,|a - b|^2 + |a|^2\,|b|^2 - (a \cdot b)^2\,\Big] \\
&\qquad \geq 2\Big[\,(1 + |a|^2)^{1/2}\,(1 + |b|^2)^{1/2} + (1 + a \cdot b)\,\Big]^{-1}\,|a - b|^2\,.
\end{aligned}$$

The estimate in (2) now follows readily. □

Proposition 9.3.4. *We have*

$$\left|\overrightarrow{G}_F(F(x)) - \overrightarrow{G}_F(F(y))\right| \leq |Df(x) - Df(y)| \tag{9.41}$$

and, provided $|Df(x)|, |Df(y)| \leq c$, *we have*

$$|Df(x) - Df(y)| \leq (1 + c^2)^2\,\left|\overrightarrow{G}_F(F(x)) - \overrightarrow{G}_F(F(y))\right|\,. \tag{9.42}$$

Proof. This result follows immediately from Lemma 9.3.3 and (9.37). □

We leave the easy proof of the next lemma to the reader.

Lemma 9.3.5. *For* $t \in \mathbb{R}$ *we have*

$$0 \leq 1 - (1 + t^2)^{-1/2} \leq \min\{\tfrac{1}{2}t^2\,,\ |t|\}\,. \tag{9.43}$$

If additionally $|t| \leq C < \infty$ *holds, then we have*

$$\frac{t^2}{2(1 + C^2)} \leq 1 - (1 + t^2)^{-1/2}\,. \tag{9.44}$$

Proposition 9.3.6. *It holds that*

$$[1 + \mathrm{Lip}\,(f)]^{-2}\ |Df|^2 \leq \left|\overrightarrow{G}_F - \mathbf{e}^M\right|^2 \leq \min\Big\{\,|Df|^2,\ 2|Df|\,\Big\}\,. \tag{9.45}$$

Proof. By (9.37) we have

$$\overrightarrow{G}_F - \mathbf{e}^M = (1+|Df|^2)^{-1/2}\left[(1-(1+|Df|^2)^{1/2})\mathbf{e}^M + \sum_{i=1}^M \frac{\partial f}{\partial x_i}\,\mathbf{e}_{\widehat{\imath}} \right],$$

so

$$\begin{aligned}
|\overrightarrow{G}_F - \mathbf{e}^M|^2 &= (1+|Df|^2)^{-1}\Big[\,1-2(1+|Df|^2)^{1/2} + (1+|Df|^2) + |Df|^2\,\Big] \\
&= (1+|Df|^2)^{-1}\Big[\,2(1+|Df|^2) - 2(1+|Df|^2)^{1/2}\,\Big] \\
&= 2\Big[\,1-(1+|Df|^2)^{-1/2}\,\Big].
\end{aligned}$$

The upper bound follows from (9.43), while the lower bound follows from (9.44). □

Corollary 9.3.7. *It holds that*

$$\begin{aligned}
2^{-1}\ [1+\mathrm{Lip}\,(f)]^{-2}\ \sigma^{-M} \int_{\mathbb{B}^M(0,\sigma)} |Df|^2\, d\mathcal{L}^M &\le E(G_F, 0, \sigma) \\
&\le\ 2^{-1}\,\sigma^{-M} \int_{\mathbb{B}^M(0,\sigma)} |Df|^2\, d\mathcal{L}^M\,.
\end{aligned}$$

Proof. The corollary is an immediate consequence of Proposition 9.3.6 and the definition of the cylindrical excess, i.e., Definition 9.1.4. □

Proposition 9.3.8. *We have*

$$\left| DA(\overrightarrow{G}_F) - DA(\mathbf{e}^M) \right| \le \min\Big\{\,|Df|^2,\ 2\,|Df|\,\Big\}. \tag{9.46}$$

Proof. By (9.39), we have

$$\begin{aligned}
&DA(\overrightarrow{G}_F) - DA(\mathbf{e}^M) \\
&\quad = (1+|Df|^2)^{-1/2}\left[(1-(1+|Df|^2)^{1/2} dx^M + \sum_{i=1}^M \left(\frac{\partial f}{\partial x_i}\right) dx_{\widehat{\imath}} \right],
\end{aligned}$$

so we can proceed as in the proof of Proposition 9.3.6 and apply (9.43). □

9.4 Estimates for Harmonic Functions

The heuristic behind the regularity theory for area-minimizing surfaces is that, at a point where an area-minimizing surface is horizontal, the closer you look at the surface, the more it looks like the graph of a harmonic function. This is made plausible by the fact that an area-minimizing graph is given by a function u that minimizes the integral of the area integrand

$$\sqrt{1+|Du|^2}\,,$$

while a harmonic function u minimizes the integral of

$$\frac{1}{2}|Du|^2\,.$$

Since the area integrand $\sqrt{1+|Du|^2}$ has the expansion

$$1+\frac{1}{2}|Du|^2+\sum_{k=2}^{\infty}\binom{1/2}{k}|Du|^{2k}\,,$$

we see that, at a point where the graph is horizontal, minimizing $\frac{1}{2}|Du|^2$ must be nearly the same as minimizing $\sqrt{1+|Du|^2}$.

To turn the heuristic discussion above into a useful estimate, we will need to investigate the boundary regularity of solutions for the Dirichlet problem[2] for Laplace's equation[3] on the unit ball. To obtain a sharp result we must use the Lipschitz spaces that we introduce next.

Notation 9.4.1. Let B denote the open unit ball in $\mathbb{R}^M$ and let Σ denote the unit sphere.

(1) For $g : \Sigma \to \mathbb{R}$, we say that g is *differentiable at* $x \in \Sigma$ if G defined by

$$G(z) = g(z/|z|) \qquad (z \neq 0)$$

is differentiable at x. This definition exploits the special structure of Σ, but it is easily seen to be equivalent to the usual definition of differentiability for a function defined on a surface (for example, see [Hir 76, pp. 15ff.]).

(2) If $g : \Sigma \to \mathbb{R}$ is differentiable at $x \in \Sigma$ and if v a unit vector, then the *directional derivative of g at x in the direction v* is defined by

$$\frac{\partial g}{\partial v}(x) = \langle\, DG(x),\ v\,\rangle\,. \tag{9.47}$$

We will also use (9.47) as the definition of $\partial g/\partial v$ when v is not a unit vector.

(3) For δ with $1 < \delta < 2$, we say that $g : \Sigma \to \mathbb{R}$ is *Lipschitz of order* δ, written $g \in \Lambda_\delta(\Sigma)$, if g is differentiable at every point of Σ, $\frac{\partial g}{\partial v}(x)$ is a continuous function of x for each unit vector v, and there exists $C < \infty$ such that for each unit vector v,

$$\left|\frac{\partial g}{\partial v}(x_1) - \frac{\partial g}{\partial v}(x_0)\right| \le C\,|x_1 - x_0|^{\delta-1}$$

holds for $x_0, x_1 \in \Sigma$.

[2] Johann Peter Gustav Lejeune Dirichlet (1805–1859).

[3] Pierre-Simon Laplace (1749–1827).

(4) If $g : \Sigma \to \mathbb{R}$ is Lipschitz of order δ on Σ ($1 < \delta < 2$), then we set

$$\|g\|_{\Lambda_\delta} = \sup_{\substack{x\in\Sigma \\ |v|=1}} \left| \frac{\partial g}{\partial v}(x) \right| + \sup_{\substack{x_0, x_1 \in \Sigma,\, x_0 \neq x_1 \\ |v|=1}} |x_1 - x_0|^{1-\delta} \left| \frac{\partial g}{\partial v}(x_1) - \frac{\partial g}{\partial v}(x_0) \right| . \tag{9.48}$$

The number $\|g\|_{\Lambda_\delta}$ defines a seminorm on $\Lambda_\delta(\Sigma)$. Had we wished to define a norm, we could have done so by including the term $\sup_{x\in\Sigma} |g(x)|$ as an additional summand on the right-hand side of (9.48).

We have defined the Lipschitz spaces $\Lambda_\delta(\Sigma)$ for δ in the limited range $1 < \delta < 2$ because those are the only spaces we will need in this section. For a comprehensive study of Lipschitz spaces, the reader should see [Kra 83].

Lemma 9.4.2. *For δ with $1 < \delta < 2$ there exists a constant $c_7 = c_7(\delta)$ with the following property:*

If $g \in \Lambda_\delta(\Sigma)$ and if $u \in C^0(\overline{B}) \bigcap C^2(B)$ satisfies

$$\begin{aligned} \triangle u &= 0 \quad on \quad B\,, \\ u &= g \quad on \quad \Sigma\,, \end{aligned} \tag{9.49}$$

then the Hilbert–Schmidt norm of the Hessian matrix of u (i.e., the square root of the sum of the squares of the entries in the matrix) is bounded by

$$\Big| \operatorname{Hess}[u(x)] \Big| \le c_7 \cdot \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2}\,. \tag{9.50}$$

Here, of course, $\triangle$ denotes the Laplacian $\sum_{i=1}^M \partial^2/\partial x_i^2$.

Proof. Our proof will be based on the fact that the function u solving (9.49) is given by the Poisson integral formula.[4] Recall (see [CH 62, pp. 264ff.], [Kra 99, p. 186], or [Kra 05, p. 143]) that the Poisson kernel for the unit ball in $\mathbb{R}^M$ is given by

$$P(x, y) = \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \frac{1 - |x|^2}{|x-y|^M} \tag{9.51}$$

$$= \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \frac{\varrho(x)\,(2-\varrho(x))}{|x-y|^M}\,, \tag{9.52}$$

where

$$\varrho(x) = 1 - |x|$$

[4] Siméon Denis Poisson (1781–1840).

is the distance from $x \in B$ to Σ. The solution to the Dirichlet problem (9.49) is given by

$$u(x) = \int_\Sigma P(x, y)\, g(y)\, d\mathcal{H}^{M-1}(y)\,. \tag{9.53}$$

Interior estimate. Observe that if $x \in B$ stays at least a fixed positive distance away from Σ, then each $|\partial P/\partial x_i|$ (and all higher derivatives of P as well) will be bounded above. Thus we can obtain estimates for the derivatives of u by differentiating the right-hand side of (9.53) under the integral and estimating the resulting integral. Thus we have (9.50) for $x \in \mathbb{B}^M(0, 1/2)$.

Notation. For $v \in \mathbb{R}^M$ a unit vector, $\partial f/\partial v$ will denote the *directional derivative* of the function f in the direction v. Here f may be real-valued or vector-valued.

Of particular interest are the directional derivatives of the Poisson kernel $P(x, y)$. Since P depends on the two arguments $x \in \mathbb{R}^M$ and $y \in \mathbb{R}^M$, we will augment our notation for directional derivatives to indicate the variable with respect to which the differentiation is to be performed. The notation $\partial P/\partial_x v$ will mean that the directional derivative of $P(x, y)$ in the direction v is to be computed by differentiating with respect to x while treating y as a parameter. We have

$$\frac{\partial P}{\partial_x v} = \sum_{i=1}^{M} v_i \, \frac{\partial P}{\partial x_i}\,. \tag{9.54}$$

On the other hand, when we wish to differentiate $P(x, y)$ as a function of y while treating x as a parameter, we will write $\partial P/\partial_y v$. We have

$$\frac{\partial P}{\partial_y v} = \sum_{i=1}^{M} v_i \, \frac{\partial P}{\partial y_i}\,. \tag{9.55}$$

Equations (9.54) and (9.55) remain meaningful when v is not a unit vector, and later we will have occasion to apply (9.55) in such a circumstance.

Estimates for derivatives of *P*. Fix a point $x \in B \setminus \{0\}$. Let y be a point on Σ. Using (9.51), we compute the derivatives of $P(x, y)$ as follows: Let v be a unit vector. Since

$$\frac{\partial x}{\partial v} = v$$

(that is, the directional derivative, in the direction v, of the map $x \mapsto x$ is v itself), we have

$$\frac{\partial P}{\partial_x v}(x, y) = \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \left(-\frac{2\,x \cdot v}{|x-y|^M} - \frac{M\,(1-|x|^2)\,(x-y)\cdot v}{|x-y|^{M+2}} \right).$$

Similarly, we find that

$$\frac{\partial P}{\partial_y v}(x, y) = \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \frac{M\,(1-|x|^2)\,(x-y)\cdot v}{|x-y|^{M+2}} = M\,\frac{(x-y)\cdot v}{|x-y|^2}\,P(x, y)\,.$$

If we consider $v = \tau$, where τ is a unit vector tangent at x to the sphere of radius $|x|$ centered at the origin, then we have $x \cdot \tau = 0$. We conclude that

$$\frac{\partial P}{\partial_x \tau}(x, y) = \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \frac{-M\,(1 - |x|^2)\,(x - y) \cdot \tau}{|x - y|^{M+2}} = -M\,\frac{(x - y) \cdot \tau}{|x - y|^2}\,P(x, y) \tag{9.56}$$

and that

$$\frac{\partial P}{\partial_x \tau}(x, y) = -\frac{\partial P}{\partial_y \tau}(x, y)\,. \tag{9.57}$$

(Note that the vector τ is the same vector on both sides of (9.57). The subscript y in the notation $\frac{\partial P}{\partial_y \tau}(x, y)$ on the right-hand side of (9.57) merely tells us to differentiate with respect to y while treating x as a constant; the subscript in no way implies that τ is tangent to Σ at y.) From (9.56), we also obtain the estimate

$$\left|\frac{\partial P}{\partial_x \tau}(x, y)\right| \le M\,|x - y|^{-1}\,P(x, y)\,. \tag{9.58}$$

Similarly, if $\widehat{\tau}$ is also a unit vector tangent at x to the sphere of radius $|x|$ centered at the origin, we have

$$\frac{\partial^2 P}{\partial_x \tau\,\partial_x \widehat{\tau}}(x, y) = -\frac{\partial^2 P}{\partial_y \tau\,\partial_x \widehat{\tau}}(x, y)\,. \tag{9.59}$$

For the vector v, which here need not be a unit vector, we find that

$$\begin{aligned}\frac{\partial^2 P}{\partial_y v\,\partial_x \tau}(x, y) = {}& M\,\frac{v \cdot \tau}{|x - y|^2}\,P(x, y)\\ &-(2M + M^2)\,\frac{[\,(x - y) \cdot \tau\,]\,[\,(x - y) \cdot v\,]}{|x - y|^4}\,P(x, y)\,,\end{aligned}$$

and we obtain the estimate

$$\left|\frac{\partial^2 P}{\partial_y v\,\partial_x \tau}(x, y)\right| \le (3M + M^2)\,|v|\,|x - y|^{-2}\,P(x, y)\,. \tag{9.60}$$

Suppose $x \in B \setminus \{0\}$ and let $\nu = x/|x|$ be the outward unit normal vector at x to the sphere of radius $|x|$ centered at the origin. We compute

$$\frac{\partial P}{\partial_x \nu}(x, y) = \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \left(-\frac{2\,x \cdot \nu}{|x - y|^M} - \frac{M\,(1 - |x|^2)\,(x - y) \cdot \nu}{|x - y|^{M+2}}\right).$$

We obtain the estimate

$$\left|\frac{\partial P}{\partial_x \nu}(x, y)\right| \le \frac{\Gamma(M/2)}{2\pi^{M/2}} \cdot \frac{1 - |x|^2}{|x - y|^M}\left(\frac{2\,|x \cdot \nu|}{1 - |x|^2} + M\,\frac{|(x - y) \cdot \nu|}{|x - y|^2}\right)$$

$$\le \frac{\Gamma(M/2)}{2\,\pi^{M/2}} \cdot \frac{1-|x|^2}{|x-y|^M} \left(\frac{2\,|x|}{\varrho(x)\,(2-\varrho(x))} + M\,\frac{|x-y|}{|x-y|^2} \right)$$

$$\le P(x,y)\,(2\,\varrho(x)^{-1} + M\,|x-y|^{-1})$$

$$\le P(x,y) \cdot (M+2) \cdot \varrho(x)^{-1}\,, \tag{9.61}$$

where we have used the fact that $\varrho(x) \le |x-y|$ (which holds because $y \in \Sigma$), thus implying

$$\frac{1}{|x-y|} \le \varrho(x)^{-1}\,. \tag{9.62}$$

In the remainder of the proof, we will use the identity (9.59) for tangential derivatives and the estimates for the derivatives of P to obtain estimates for the second derivatives of u.

Estimates for tangential second derivatives of u. Fix a point $x \in B \setminus \{0\}$. Let τ and $\widehat{\tau}$ be unit vectors tangent at x to the sphere of radius $|x|$ centered at the origin.

Since Hess $[\,u(x)\,]$ is unaffected by adding a constant to g, we may suppose for convenience that

$$g(\zeta(x)) = 0\,, \tag{9.63}$$

where $\zeta(x) = x/|x|$ is the radial projection of x into Σ. It also will be convenient to use "C" to denote a generic constant, the specific value of which may vary from line to line.

We compute

$$\begin{aligned}
\left| \frac{\partial^2 u}{\partial \tau\, \partial \widehat{\tau}} \right| &= \left| \int_\Sigma \frac{\partial^2 P}{\partial_x \tau\, \partial_x \widehat{\tau}}(x,y)\, g(y)\, d\mathcal{H}^{M-1}(y) \right| \\
&= \left| \int_\Sigma -\frac{\partial^2 P}{\partial_y \tau\, \partial_x \widehat{\tau}}(x,y)\, g(y)\, d\mathcal{H}^{M-1}(y) \right| \\
&= \left| \int_\Sigma \frac{\partial P}{\partial_x \widehat{\tau}}(x,y)\, \frac{\partial g}{\partial_y \tau}(y)\, d\mathcal{H}^{M-1}(y) \right. \\
&\qquad \left. - \int_\Sigma \frac{\partial}{\partial_y \tau} \left(\frac{\partial P}{\partial_x \widehat{\tau}}(x,y)\, g(y) \right) d\mathcal{H}^{M-1}(y) \right| \\
&\le \left| \int_\Sigma \frac{\partial P}{\partial_x \widehat{\tau}}(x,y) \left[\frac{\partial g}{\partial_y \tau}(y) - \frac{\partial g}{\partial_y \tau}(\zeta(x)) \right] d\mathcal{H}^{M-1}(y) \right| \\
&\qquad + \left| \int_\Sigma \frac{\partial}{\partial_y \tau} \left(\frac{\partial P}{\partial_x \widehat{\tau}}(x,y)\, g(y) \right) d\mathcal{H}^{M-1}(y) \right| \\
&= I + II\,.
\end{aligned}$$

Here we have also used the fact that

$$\int_{\Sigma} \frac{\partial P}{\partial_x \widehat{\tau}}(x, y)\, d\mathcal{H}^{M-1}(y) = 0\,. \tag{9.64}$$

Equation (9.64) holds because

$$\int_{\Sigma} P(x, y)\, d\mathcal{H}^{M-1}(y) \equiv 1 \tag{9.65}$$

implies

$$0 = \frac{\partial}{\partial \widehat{\tau}} \int_{\Sigma} P(x, y)\, d\mathcal{H}^{M-1}(y) = \int_{\Sigma} \frac{\partial P}{\partial_x \widehat{\tau}}(x, y)\, d\mathcal{H}^{M-1}(y)\,.$$

Set

$$S_1 = \Big\{ y \in \Sigma \;:\; |y - \zeta(x)| \le \varrho(x) \Big\}, \tag{9.66}$$

$$S_2 = \Big\{ y \in \Sigma \;:\; |y - \zeta(x)| > \varrho(x) \Big\} \tag{9.67}$$

(see Figure 9.3).

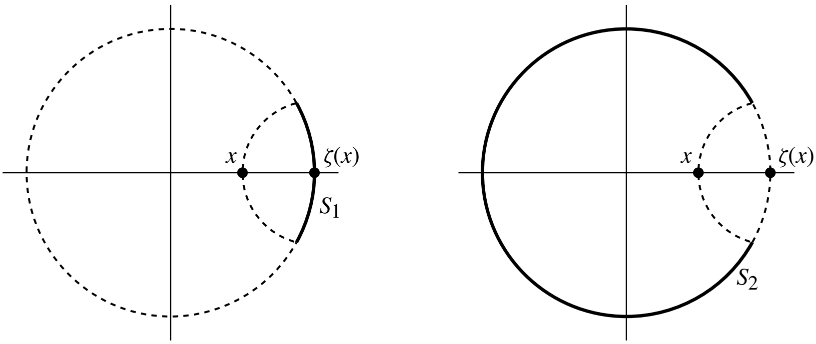

Fig. 9.3. The regions S_1 and S_2 in Σ.

Using (9.58), we can estimate that I is bounded by

$$\begin{aligned}
M & \int_{\Sigma} \frac{1}{|x-y|}\; P(x, y)\; \|g\|_{\Lambda_\delta}\; |y - \zeta(x)|^{\delta-1}\, d\mathcal{H}^{M-1}(y) \\
&= M \int_{S_1} \frac{1}{|x-y|}\; P(x, y)\; \|g\|_{\Lambda_\delta}\; |y - \zeta(x)|^{\delta-1}\, d\mathcal{H}^{M-1}(y) \\
&\quad + M \int_{S_2} \frac{1}{|x-y|}\; P(x, y)\; \|g\|_{\Lambda_\delta}\; |y - \zeta(x)|^{\delta-1}\, d\mathcal{H}^{M-1}(y) \\
&= I_1 + I_2\,.
\end{aligned}$$

We estimate I_1 by using (9.62), (9.65), the nonnegativity of P, and the fact that on S_1, it holds that

$$|y - \zeta(x)|^{\delta-1} \leq \varrho(x)^{\delta-1}$$

because $\delta - 1 > 0$. We have

$$\begin{aligned} I_1 &\leq \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{-1} \int_{S_1} P(x, y)\, |y - \zeta(x)|^{\delta-1}\, d\mathcal{H}^{M-1}(y) \\ &\leq \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{-1} \int_{S_1} P(x, y)\, \varrho(x)^{\delta-1}\, d\mathcal{H}^{M-1}(y) \\ &= \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2} \int_{S_1} P(x, y)\, d\mathcal{H}^{M-1}(y) \\ &\leq \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2} \int_{\Sigma} P(x, y)\, d\mathcal{H}^{M-1}(y) = \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2}\,. \end{aligned}$$

To estimate I_2, we first note that

$$|y - \zeta(x)| \leq |y - x| + |\zeta(x) - x| = |y - x| + \varrho(x) \leq 2|y - x|\,, \tag{9.68}$$

which implies that

$$\frac{1}{|x - y|} \leq 2|y - \zeta(x)|^{-1}\,.$$

Also we note that on S_2, it holds that

$$|y - \zeta(x)|^{\delta-2} \leq \varrho(x)^{\delta-2}$$

because $\delta - 2 < 0$. We estimate

$$\begin{aligned} I_2 &\leq 2\, \|g\|_{\Lambda_\delta} \int_{S_2} P(x, y)\, |y - \zeta(x)|^{\delta-2}\, d\mathcal{H}^{M-1}(y) \\ &\leq 2\, \|g\|_{\Lambda_\delta} \int_{S_2} P(x, y)\, \varrho(x)^{\delta-2}\, d\mathcal{H}^{M-1}(y) \\ &= 2\, \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2} \int_{S_2} P(x, y)\, d\mathcal{H}^{M-1}(y) \\ &\leq 2\, \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2} \int_{\Sigma} P(x, y)\, d\mathcal{H}^{M-1}(y) = 2\, \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2}\,. \end{aligned}$$

To obtain an estimate for II, suppose without loss of generality that $\zeta(x) = \mathbf{e}_1$ and $\tau = \mathbf{e}_2$. Setting

$$T = T(y) = (y_1^2 + y_2^2)^{-1/2}\, (-y_2\, \mathbf{e}_1 + y_1\, \mathbf{e}_2)\,,$$

for each $y = (y_1, y_2, \ldots, y_M) \in \Sigma$, with $(y_1, y_2) \neq (0, 0)$, and applying the fundamental theorem of calculus, we see that

$$\int_\Sigma \frac{\partial}{\partial_y T} \left(\frac{\partial P}{\partial_x \widehat{\tau}}(x, y)\, g(y) \right) d\mathcal{H}^{M-1}(y) = 0\,;$$

more specifically, we parametrize the sphere by

$$\left(r\cos\theta, r\sin\theta, y', \pm\sqrt{1 - r^2 - |y'|^2} \right),$$

where $0 < r < 1$, $0 < \theta < 2\pi$, $y' \in \mathbb{R}^{M-3}$, with $0 < |y'| < \sqrt{1 - r^2}$, and integrate first with respect to θ.

Setting $v = v(y) = \tau - T(y)$ and using (9.63), we have

$$\begin{aligned}
II &= \left| \int_\Sigma \left(\frac{\partial}{\partial_y \tau} - \frac{\partial}{\partial_y T} \right) \left(\frac{\partial P}{\partial_x \widehat{\tau}}(x, y)\, g(y) \right) d\mathcal{H}^{M-1}(y) \right| \\
&= \left| \int_\Sigma \frac{\partial}{\partial_y v} \left(\frac{\partial P}{\partial_x \widehat{\tau}}(x, y)\, g(y) \right) d\mathcal{H}^{M-1}(y) \right| \\
&\leq \left| \int_\Sigma \frac{\partial^2 P}{\partial_y v\, \partial_x \widehat{\tau}}(x, y) \left[g(y) - g(\zeta(x)) \right] d\mathcal{H}^{M-1}(y) \right| \\
&\quad + \left| \int_\Sigma \frac{\partial P}{\partial_x \widehat{\tau}}(x, y) \frac{\partial g}{\partial_y v}(y)\, d\mathcal{H}^{M-1}(y) \right| \\
&= II_1 + II_2\,,
\end{aligned}$$

where we have used the assumption that $g(\zeta(x)) = 0$.

Consider $y = (y_1, y_2, \ldots, y_M) \in \Sigma$ and write $(y_1, y_2) = (r\cos\theta, r\sin\theta)$, where $0 \leq r \leq 1$. It is easy to check that $1 - \cos\theta \leq 2(1 - r\cos\theta)$ holds for $0 \leq r \leq 1$. The law of cosines tells us that $|\tau - T(y)| = \sqrt{2(1 - \cos\theta)}$ and that $|(y_1, y_2) - (1, 0)| = \sqrt{2(1 - r\cos\theta)}$, so we have

$$|\tau - T(y)| \leq \sqrt{2}\, |y - \zeta(x)| \overset{(9.68)}{\leq} 2\sqrt{2}\, |y - x|\,. \tag{9.69}$$

Observe that $|g(y) - g(\zeta(x))|$ is bounded by $\|g\|_{\Lambda_\delta}$ multiplied by the distance from y to $\zeta(x)$ *measured along the sphere.* Thus we have

$$|g(y) - g(\zeta(x))| \leq C \cdot \|g\|_{\Lambda_\delta} \cdot |y - \zeta(x)| \leq 2C \cdot \|g\|_{\Lambda_\delta} \cdot |y - x|\,.$$

Using (9.60) and (9.69), we may estimate

$$\begin{aligned}
II_1 &\leq C \int_\Sigma \frac{|\tau - T|}{|x - y|^2} P(x, y) \cdot \|g\|_{\Lambda_\delta} \cdot |y - x|\, d\mathcal{H}^{M-1}(y) \\
&\leq C \cdot \|g\|_{\Lambda_\delta}\,.
\end{aligned}$$

Next, observe that

$$\left|\frac{\partial g}{\partial_y v}(y)\right| \le |v| \cdot \|g\|_{\Lambda_\delta} ,$$

so, by (9.58) and (9.69), we see that

$$\begin{aligned} II_2 &\le C \int_\Sigma |x-y|^{-1} P(x,y) \cdot \|g\|_{\Lambda_\delta} \cdot |\tau - T| \, d\mathcal{H}^{M-1}(y) \\ &\le C \cdot \|g\|_{\Lambda_\delta} . \end{aligned}$$

Thus we have

$$\left|\frac{\partial^2 u}{\partial \tau \, \partial \widehat{\tau}}\right| \le C \cdot \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta-2} , \tag{9.70}$$

for $x \in B \setminus \{0\}$ and unit vectors $\tau, \widehat{\tau}$ with $\tau \cdot x = \widehat{\tau} \cdot x = 0$.

Mixed normal and tangential second derivatives. Fix a point $x \in B \setminus \{0\}$, let τ be a unit vector tangent at x to the sphere of radius $|x|$ centered at the origin, and let $\nu = x/|x|$ be the outward unit normal vector at x to the sphere of radius $|x|$.

We have

$$\begin{aligned} \frac{\partial^2 u}{\partial \nu \, \partial \tau} &= \int_\Sigma \frac{\partial^2 P}{\partial \nu \, \partial \tau}(x,y) \, g(y) \, d\mathcal{H}^{M-1}(y) \\ &= \int_\Sigma \frac{\partial P}{\partial_x \nu}(x,y) \, \frac{\partial g}{\partial_y \tau}(y) \, d\mathcal{H}^{M-1}(y) \\ &= \int_\Sigma \frac{\partial P}{\partial_x \nu}(x,y) \left[\frac{\partial g}{\partial_y \tau}(y) - \frac{\partial \, (g \circ \zeta)}{\partial_y \tau}(g \circ \zeta)(x)\right] d\mathcal{H}^{M-1}(y) . \end{aligned} \tag{9.71}$$

We can proceed as before, with S_1 and S_2 defined as in (9.66) and (9.67), to estimate

$$\begin{aligned} \left|\frac{\partial^2 u}{\partial \nu \, \partial \tau}\right| &\le \|g\|_{\Lambda_\delta} \int_\Sigma \left|\frac{\partial P}{\partial_x \nu}(x,y)\right| \, |y - \zeta(x)|^{\delta-1} \, d\mathcal{H}^{M-1}(y) \\ &= \|g\|_{\Lambda_\delta} \int_{S_1} \left|\frac{\partial P}{\partial_x \nu}(x,y)\right| \, |y - \zeta(x)|^{\delta-1} \, d\mathcal{H}^{M-1}(y) \\ &\quad + \|g\|_{\Lambda_\delta} \int_{S_2} \left|\frac{\partial P}{\partial_x \nu}(x,y)\right| \, |y - \zeta(x)|^{\delta-1} \, d\mathcal{H}^{M-1}(y) \\ &= III + IV . \end{aligned}$$

We use (9.61) to estimate

$$III \le \|g\|_{\Lambda_\delta} \cdot (M+2) \cdot \varrho(x)^{\delta-2} .$$

Estimating IV is more complicated. We use the estimate (9.61) to see that

$$\begin{aligned}\left|\frac{\partial P}{\partial_x \nu}(x,y)\right| &\le (M+2)\cdot \varrho(x)^{-1}\cdot P(x,y)\\ &= (M+2)\cdot\varrho(x)^{-1}\cdot\frac{\Gamma(M/2)}{2\,\pi^{M/2}}\cdot\frac{\varrho(x)\,(2-\varrho(x))}{|x-y|^M}\\ &= (M+2)\cdot\frac{\Gamma(M/2)}{2\,\pi^{M/2}}\cdot\frac{2-\varrho(x)}{|x-y|^M}\\ &\le \frac{(M+2)\,\Gamma(M/2)}{\pi^{M/2}}\cdot\frac{1}{|x-y|^M}\,.\end{aligned}$$

Then, using the estimate $|y-x|^{-1}\le 2|y-\zeta(x)|^{-1}$, we obtain

$$IV \le C\cdot\ \|g\|_{\Lambda_\delta}\int_{S_2}|y-\zeta(x)|^{\delta-1-M}\,d\mathcal{H}^{M-1}(y)\,.$$

To estimate this last integral, we suppose without loss of generality that $\zeta(x) = (1,0,\ldots,0)$. We write

$$(y_1,y_2,\ldots,y_M)=(y',y'',\eta)\ \text{ with }\ y'=y_1,\ y''=(y_2,y_3,\ldots,y_{M-1}),\ \eta=y_M\,,$$

so that Σ can be parametrized by

$$\eta=\pm(1-y'^2-|y''|^2)^{1/2}$$

with

$$d\mathcal{H}^{M-1}(y)=(1-y'^2-|y''|^2)^{-1/2}\,d\mathcal{L}^{M-1}(y',y'')\,.$$

We have $|y-\zeta(x)|=(2-2y')^{1/2}$, so

$$IV\le C\,\|g\|_{\Lambda_\delta}\int_{-1}^{1-\varrho(x)^2/2}\int_{|y''|=\sqrt{1-y'^2}}\frac{(2-2y')^{(\delta-1-M)/2}}{(1-y'^2-|y''|^2)^{1/2}}\,d\mathcal{L}^{M-2}(y'')\,d\mathcal{L}(y')\,.$$

We note that the integral

$$\int_{|y''|=\sqrt{1-y'^2}}(1-y'^2-|y''|^2)^{-1/2}\,d\mathcal{L}^{M-2}(y'')$$

equals the $(M-2)$-dimensional area of the upper hemisphere of radius $\sqrt{1-y'^2}$ in $\mathbb{R}^{M-1}$. Thus we have

$$\begin{aligned}IV &= C\,\|g\|_{\Lambda_\delta}\int_{-1}^{1-\varrho(x)^2/2}(2-2y')^{(\delta-1-M)/2}\,(1-y'^2)^{(M-2)/2}\,d\mathcal{L}(y')\\ &\le C\,\|g\|_{\Lambda_\delta}\int_{-1}^{1-\varrho(x)^2/2}(1-y')^{(\delta-3)/2}\,d\mathcal{L}(y')\\ &\le C\,\|g\|_{\Lambda_\delta}\,2^{(M+\delta-1)/2}/(\delta-1)\,,\end{aligned}$$

and we conclude that

$$\left|\frac{\partial^2 u}{\partial \nu\, \partial \tau}\right| \le C \cdot \frac{1}{\delta - 1} \cdot \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta - 2} . \tag{9.72}$$

The second normal derivative. Fix a point $x \in B \setminus \{0\}$ and let $\nu = x/|x|$ be the outward unit normal vector to the sphere of radius $|x|$ centered at the origin.

If $\tau_1, \tau_2, \ldots, \tau_{M-1}$ are pairwise orthogonal unit vectors, all tangent at x to the sphere of radius $|x|$, then

$$\frac{\partial^2 u}{\partial \nu^2} = -\sum_{i=1}^{M-1} \frac{\partial^2 u}{\partial \tau_i^2} ,$$

so that

$$\left|\frac{\partial^2 u}{\partial \nu^2}\right| \le C \cdot \|g\|_{\Lambda_\delta} \cdot \varrho(x)^{\delta - 2} . \tag{9.73}$$

Summary. For $x \in B \setminus \{0\}$, we can make an orthogonal change of basis such that $x/|x|$ coincides with one of the standard basis vectors. Then (9.70), (9.72), and (9.73) give us the required bound for the Hilbert–Schmidt norm of the Hessian matrix for u at x. □

Lemma 9.4.3. *Fix* $0 < \delta < 1$ *and* $1 < \hat{\sigma} < 2$*. There is a constant* $c_8 = c_8(\delta)$ *such that if*

$$g : \mathbb{B}^M(0, \hat{\sigma}) \to \mathbb{R}$$

is smooth and $u \in C^0(\overline{B}) \bigcap C^2(B)$ *satisfies*

$$\triangle u = 0 \ \ on \ \ B ,$$

$$u = g \ \ on \ \ \Sigma ,$$

then

(1) $$\sup\left\{ |x - z|^{-\delta}\, |Du(x) - Du(z)| : x, z \in B,\ x \ne z \right\} + \sup_B |Du|$$
$$\le c_8 \cdot \left(\sup\left\{ |x - z|^{-\delta}\, |Dg(x) - Dg(z)| : x, z \in \mathbb{B}^M(0, \hat{\sigma}),\ x \ne z \right\} + \sup_{\mathbb{B}^M(0,\hat{\sigma})} |Dg| \right),$$

(2) $$\sup_{\mathbb{B}^M(0,1/2)} \Big|\mathrm{Hess}\,[u(x)]\Big| \le c_8 \left(\int_B \Big|\mathrm{Hess}\,[u(x)]\Big|^2 d\mathcal{L}^M \right)^{1/2} ,$$

(3) $$\sup_{x \in \mathbb{B}^M(0,\hat{\eta})} |Du(x) - Du(0)|^2 \le c_8\, \hat{\eta}^2 \int_B \Big|\mathrm{Hess}\,[u(x)]\Big|^2 d\mathcal{L}^M ,$$
for each $0 < \hat{\eta} < 1/2$.

Proof.
(1) Since

$$\sup_B |Du| \leq \sup_\Sigma |Dg|$$

holds by the maximum principle, it suffices to estimate

$$\sup \left\{ |x - z|^{-\delta} \, |Du(x) - Du(z)| : x, z \in B, \ x \neq z \right\}.$$

We do so by comparing

$$|Du(x_1) - Du(x_0)|$$

to h^δ, where $x_0, x_1 \in B$ and $h = |x_1 - x_0|$. We need only consider h small, and again by the maximum principle, we need to consider only x_0 near Σ.

Set $\hat{\delta} = 1 + \delta$. We will apply Lemma 9.4.2 with δ replaced by $\hat{\delta}$. By that lemma, we have

$$\Big| \mathrm{Hess}\,[u(x)] \Big| \leq c_7 \cdot \|g\|_{\Lambda_{\hat{\delta}}} \cdot \varrho(x)^{\hat{\delta}-2}$$

for $x \in B$, where $\varrho(x) = 1 - |x|$. Note that

$$\|g\|_{\Lambda_{\hat{\delta}}} \leq \sup \left\{ |x - z|^{-\delta} \, |Dg(x) - Dg(z)| : x, z \in \mathbb{B}^M(0, \hat{\sigma}), \ x \neq z \right\} + \sup_{\mathbb{B}^M(0,\hat{\sigma})} |Dg|$$

holds. In what follows, C will denote a generic positive, finite constant incorporating the value of c_7.

We need to estimate $|Du(x_1) - Du(x_0)|$. The proximity of the boundary Σ makes it difficult to obtain the needed estimate. Rather than proceeding directly, we replace each point x_i by a point $\widetilde{x_i}$ that is at distance h farther away from Σ (see Figure 9.4). Remarkably, it is then feasible to estimate the individual terms $|Du(\widetilde{x_0}) - Du(x_0)|$, $|Du(\widetilde{x_1}) - Du(x_1)|$, and $|Du(\widetilde{x_0}) - Du(\widetilde{x_1})|$.

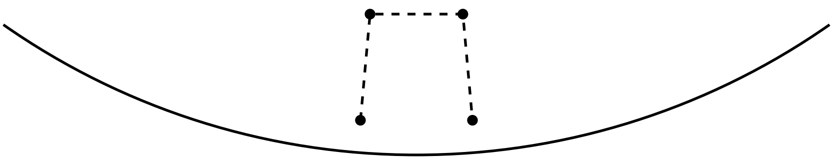

Fig. 9.4. Moving the points away from the boundary.

Let $\widetilde{x_i}$ be such that

$$\zeta(\widetilde{x_i}) = \zeta(x_i)\,,$$

$$|\widetilde{x_i}| = |x_i| - h\,;$$

then we have

$$\begin{aligned}|Du(x_1) - Du(x_0)| &\leq |Du(x_1) - D(\widetilde{x_1})| \\ &\quad + |Du(\widetilde{x_1}) - Du(\widetilde{x_0})| \\ &\quad + |Du(\widetilde{x_0}) - Du(x_0)| \\ &= I + II + III\,.\end{aligned}$$

Set $\nu = x_0/|x_0|$. We have

$$\begin{aligned}III &\leq \int_0^h \left| \frac{\partial(Du)}{\partial \nu}(x_0 - t\nu) \right| d\mathcal{L}^1(t) \\ &\leq \int_0^h \Big| \mathrm{Hess}\,[u(x_0 - t\nu)] \Big| \, d\mathcal{L}^1(t) \\ &\leq C\,\|g\|_{\Lambda_{\hat{\delta}}} \int_0^h \varrho(x_0 - t\nu)^{\hat{\delta}-2}\, d\mathcal{L}^1(t) \\ &\leq C\,\|g\|_{\Lambda_{\hat{\delta}}} \int_0^h [\varrho(x_0) + t]^{\hat{\delta}-2}\, d\mathcal{L}^1(t) \\ &= C\,\|g\|_{\Lambda_{\hat{\delta}}} \left([\varrho(x_0) + h]^{\hat{\delta}-1} - \varrho(x_0)^{\hat{\delta}-1} \right) \\ &\leq C\, h^{\hat{\delta}-1} = C\, h^{\delta}\,,\end{aligned}$$

if $\varrho(x_0)$ is small. (Note that $\hat{\delta} - 1 > 0$.)

Likewise, we estimate

$$I \leq C\,\|g\|_{\Lambda_{\hat{\delta}}}\, h^{\hat{\delta}-1}\,.$$

To estimate II, we note that

$$II \leq \int_0^h h \Big| \mathrm{Hess}\,[u(\widetilde{x_0} + \xi)] \Big| \, d\mathcal{L}^1(t)\,, \tag{9.74}$$

where $\widetilde{x_0} + \xi$ is a point on the segment between $\widetilde{x_0}$ and $\widetilde{x_1}$. The right-hand side of (9.74) is bounded above by

$$\begin{aligned}C\,\|g\|_{\Lambda_{\hat{\delta}}}\, h \int_0^h \varrho(\widetilde{x_0} + \xi)^{\hat{\delta}-2}\, d\mathcal{L}^1(t) &\leq C\,\|g\|_{\Lambda_{\hat{\delta}}}\, h \int_0^h h^{\hat{\delta}-2}\, d\mathcal{L}^1(t) \\ &\leq C\,\|g\|_{\Lambda_{\hat{\delta}}}\, h^{\hat{\delta}}\,.\end{aligned}$$

(2) Fix $i, j \in \{1, 2, \ldots, M\}$ and $x \in \mathbb{B}^M(0, 1/2)$. For $0 < r < 1/2$, by the mean value property of harmonic functions, we have

$$\frac{\partial^2 u}{\partial x_i\, \partial x_j}(x) = C \cdot r^{1-M} \int_{\{y:|y|=r\}} \frac{\partial^2 u}{\partial x_i\, \partial x_j}(x + y)\, d\mathcal{H}^{M-1}(y)\,.$$

But then

$$\begin{aligned}\left|\frac{\partial^2 u}{\partial x_i\,\partial x_j}(x)\right| &= C\left|\int_{1/4}^{1/2} r^{1-M}\int_{\{y:|y|=r\}}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(x+y)\,d\mathcal{H}^{M-1}(y)\,d\mathcal{L}^1(r)\right|\\ &\le C\left|\int_{\mathbb{B}^M(x,1/2)}\frac{\partial^2 u}{\partial x_i\,\partial x_j}(z)\,d\mathcal{L}^M(z)\right|\\ &\le C\left(\int_B\left|\frac{\partial^2 u}{\partial x_i\,\partial x_j}\right|^2 d\mathcal{L}^M\right)^{1/2}\end{aligned}$$

holds and the result follows.

(3) Fix $i \in \{1, 2, \ldots, M\}$ and $x \in \mathbb{B}^M(0, 1/2) \setminus \{0\}$. Set $\nu = x/|x|$ and

$$\psi(t) = \frac{\partial u}{\partial x_i}(t\nu)$$

for $-1 < t < 1$. Thus $\psi'(t)$ is the directional derivative of $\partial u/\partial x_i$ in the direction ν at the point $t\nu$. It follows that $|\psi'(t)|$ is bounded by the operator norm of the Hessian matrix for u at $t\nu$. Hence $|\psi'(t)|$ is bounded by a multiple of $\Big|\operatorname{Hess}[u(t\nu)]\Big|$.

Using the fundamental theorem of calculus, we estimate

$$\begin{aligned}\left|\frac{\partial u}{\partial x_i}(x) - \frac{\partial u}{\partial x_i}(0)\right|^2 &= \left|\int_0^{|x|}\psi'(t)\,d\mathcal{L}^1(t)\right|^2\\ &\le |x|^2\cdot\sup\left\{|\psi'(t)|^2 : 0 \le t \le |x|\right\}\\ &\le |x|^2\cdot\sup_{y\in\mathbb{B}^M(0,1/2)}\Big|\operatorname{Hess}[u(y)]\Big|^2,\end{aligned}$$

so we see that conclusion (3) follows from conclusion (2). □

9.5 The Main Estimate

The next lemma is the main tool in the regularity theory. The lemma tells us that once the cylindrical excess (see Definition 9.1.4) of an area-minimizing surface is small enough, then the excess on a smaller cylinder can be made even smaller by appropriately rotating the surface.

Lemma 9.5.1. *There exist constants*

$$0 < \theta < 1/8\,, \quad 0 < \epsilon_* \le (\theta/4)^{2M}\,, \tag{9.75}$$

depending only on M, with the following property:

If $0 \in \operatorname{spt} T$, *if* $T_0 = T \llcorner \mathrm{C}(0, \rho/2)$, *and if the hypotheses* (H1–H5) (*see page* 262) *hold with*

$$y = 0\,, \quad \epsilon = \epsilon_*\,,$$

then

$$\sup_{X \in \operatorname{spt} T_0} |\mathbf{q}(X)| \le \rho/8 \tag{9.76}$$

holds and there exists a linear isometry $\mathbf{j} : \mathbb{R}^{M+1} \to \mathbb{R}^{M+1}$ *with*

$$\theta^{-2M} E(T, 0, \rho) \le 1/64\,, \tag{9.77}$$

$$\|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\|^2 \le \theta^{-2M} E(T, 0, \rho)\,, \tag{9.78}$$

$$E(\mathbf{j}_\# T_0, 0, \theta\rho) \le \theta\, E(T, 0, \rho)\,. \tag{9.79}$$

Here $\mathbf{I}_{\mathbb{R}^{M+1}}$ *is the identity map on* $\mathbb{R}^{M+1}$.

Proof. Since we may change scale if need be, it will be sufficient to prove the lemma with $\rho = 1$. We ultimately will choose

$$\epsilon_* < \epsilon_0\,, \tag{9.80}$$

where ϵ_0 is as in Lemmas 9.2.1 and 9.2.2 (in particular, Lemma 9.2.1 is invoked with $\sigma = 2/3$), so we will assume that $0 \in \operatorname{spt} T$ and that the hypotheses (H1–H5) hold with $y = 0$, $\rho = 1$, and with $\epsilon = \epsilon_0$, where ϵ_0 is as in Lemma 9.2.1.

We set

$$\delta = \frac{1}{9M^2}\,,$$

$$E = E(T, 0, 1)\,.$$

Lipschitz approximations. We can apply Lemma 9.2.2 to obtain a Lipschitz function whose graph approximates $\operatorname{spt} T$. In fact, there are two such approximating functions that will be of interest:

- We let $g_\delta : \mathbb{B}^M(0, 1/4) \to \mathbb{R}$ be a Lipschitz function as in Lemma 9.2.2 corresponding to the choice
$$\gamma = E^{2\delta}\,.$$
- We let $h : \mathbb{B}^M(0, 1/4) \to \mathbb{R}$ be a Lipschitz function as in Lemma 9.2.2 corresponding to the choice $\gamma = 1$.

Smoothing g_δ. Let $\varphi \in C^\infty(\mathbb{R}^M)$ be a mollifier as in Definition 5.5.1 with N replaced by M. As usual, for $0 < \nu$,

- set
$$\varphi_\nu(z) = \nu^{-M} \varphi(\nu^{-1} z);$$

- let $f * \varphi_\nu$ denote convolution of f with φ_ν.

Let $0 < c_9 < \infty$ satisfy

$$\sup|\varphi| \le c_9\,,$$

$$\sup|D\varphi| \le c_9\,,$$

$$\sup_{x\neq z}|x-z|^{-\delta}\,|D\varphi(x)-D\varphi(z)| \le c_9\,.$$

Defining

$$\widetilde{g}_\delta = g_\delta * \varphi_E\,, \tag{9.81}$$

we obtain the following standard estimates:

$$\sup_{\mathbb{B}^M(0,1/8)}|D\widetilde{g}_\delta| \le \sup_{\mathbb{B}^M(0,1/4)}|Dg_\delta| \le E^{2\delta} \le E^\delta, \tag{9.82}$$

$$\sup_{\mathbb{B}^M(0,1/8)}|\widetilde{g}_\delta - g_\delta| \le E\sup_{\mathbb{B}^M(0,1/4)}|Dg_\delta| \le E^{1+\delta}, \tag{9.83}$$

$$\begin{aligned}
&\sup\{\,|x-z|^{-\delta}\,|D\widetilde{g}_\delta(x)-D\widetilde{g}_\delta(z)|\ :\ x,z\in\mathbb{B}^M(0,1/8),\ x\neq z\,\}\\
&\qquad\le \sup_{\mathbb{B}^M(0,1/4)}|Dg_\delta|\cdot\sup_{x\neq z}|x-z|^{-\delta}\,|\phi(E^{-1}x)-\phi(E^{-1}z)|\\
&\qquad\le\ E^{2\delta}\cdot E^{-\delta}\cdot\sup_{x\neq z}|x-z|^{-\delta}\,|\phi(x)-\phi(z)|\\
&\qquad\le c_9\,E^\delta\,.
\end{aligned} \tag{9.84}$$

The graph of $\widetilde{g}_\delta$. We next define

$$\widetilde{S} = \widetilde{G}_\#(\mathbf{E}^M \llcorner \mathbb{B}^M(0,1/8))\,, \tag{9.85}$$

where $\widetilde{G} : \mathbb{B}^M(0,1/8) \to \mathrm{C}(0,1/8)$ is defined by

$$\widetilde{G}(x) = (x, \widetilde{g}_\delta(x))\,.$$

Choosing σ. For each $0 < \sigma < 1/8$ we let

$$T_\sigma = T \llcorner \mathrm{C}(0,\sigma), \qquad \widetilde{S}_\sigma = \widetilde{S} \llcorner \mathrm{C}(0,\sigma)\,.$$

We wish to show that there is a finite positive constant c_{10} such that there are infinitely many choices of $1/16 < \sigma < 1/8$ for which the following inequalities all hold:

$$\mathcal{H}^{M-1}\Big\{x\in\partial\mathbb{B}^M(0,\sigma) : g_\delta(x)\neq h(x)\Big\} \le c_{10}\,E^{1-4M\delta}\,, \tag{9.86}$$

$$\|\partial T_\sigma\|(\mathbb{R}^{M+1}) \le c_{10}, \tag{9.87}$$

$$\|\partial T_\sigma\|\Big\{X : |P(X)-X| > E^{1+\delta}\Big\} \le c_{10}\,E^{1-4M\delta}\,, \tag{9.88}$$

where P is the "vertical retraction" of $\mathrm{C}(0, 1/8)$ onto the graph of $\widetilde{g}_\delta$. That is, for $X \in \mathrm{C}(0, 1/8)$ we have

$$P(X) = (\mathbf{p}\,(X), \widetilde{g}_\delta(\mathbf{p}\,(X)))\,.$$

Notice that $P_\# T_\sigma = \widetilde{S}_\sigma$ by (9.18) and the definition of $\widetilde{S}$.

- First, by (9.23) and by Theorem 5.2.1, i.e., the coarea formula, we have

$$\int_{1/16}^{1/8} \mathcal{H}^{M-1}\Big\{ x \in \partial\mathbb{B}^M(0,\sigma) : g_\delta(x) \neq h(x) \Big\}\, d\mathcal{L}^1(\sigma)$$

$$\leq \mathcal{L}^M\Big(\mathbb{B}^M(y, 1/4) \setminus \Big\{ z \in \mathbb{B}^M(y, 1/4) : \mathbf{p}^{-1}(z) \bigcap \operatorname{spt} T = \{(x, h(x))\} \Big\} \Big)$$

$$+ \mathcal{L}^M\Big(\mathbb{B}^M(y, 1/4) \setminus \Big\{ z \in \mathbb{B}^M(y, 1/4) : \mathbf{p}^{-1}(z) \bigcap \operatorname{spt} T = \{(x, g_\delta(x))\} \Big\} \Big)$$

$$\leq c_3\,(1 + E^{-4\delta})\, E \leq 2\, c_3\, E^{1-4\delta}\,.$$

- Because ∂T has its support outside the cylinder of radius 1, we can identify ∂T_σ with the slice $\langle T, r, \sigma+\rangle$, where r is the distance from the axis of the cylinder. We conclude that

$$\int_{1/16}^{1/8} \|\partial T_\sigma\|(\mathbb{R}^{M+1})\, d\mathcal{L}^1(\sigma) \leq \int_{\mathrm{C}(0,1/8)} d\|T\|$$

holds.

- Third, by (9.83), if $X = (x, g_\delta(x))$ coincides with the point $\mathbf{p}^{-1}(x) \bigcap \operatorname{spt} T$, then X and $P(X)$ are separated by a distance not exceeding $E^{1+\delta}$. So we use (9.24) to estimate

$$\int_{1/16}^{1/8} \|\partial T_\sigma\|\{\, X \;:\; |P(X) - X| > E^{1+\delta}\,\}\, d\mathcal{L}^1(\sigma)$$

$$= \int_{1/16}^{1/8} \|\langle T, r, \sigma+\rangle\|\{\, X \;:\; |P(X) - X| > E^{1+\delta}\,\}\, d\mathcal{L}^1(\sigma)$$

$$= \int_{1/16}^{1/8} \|\langle T - \widetilde{S}, r, \sigma+\rangle\|\mathrm{C}(y, 1/4)\, d\mathcal{L}^1(\sigma)$$

$$\leq \|T - \widetilde{S}\|\mathrm{C}(y, 1/4) \leq c_4\, E^{-4M\delta}\, E\,,$$

where we note that, in the notation of Lemma 9.2.2, $\widetilde{S}$ corresponds to T^{g_δ}.

The homotopy between T_σ and $\widetilde{S}_\sigma$. Let $H : [0, 1] \times \mathrm{C}(0, 1/8) \to \mathbb{R}^{M+1}$ be defined by $H(t, x) = t P(X) + (1 - t)X$. By the homotopy formula (7.22), we have

$$\partial V = \partial T_\sigma - \partial \widetilde{S}_\sigma\,, \tag{9.89}$$

where

$$V = H_\#([\![0,1]\!] \times \partial T_\sigma)\,.$$

By (7.23) and Lemma 9.2.2 applied with $\gamma = E^{2\delta}$ (in particular, using (9.21) and (9.23)), and by (9.83), (9.86), and (9.88), we have

$$\begin{aligned}
&\|V\|(\mathbb{R}^{M+1}) \\
&\quad \le 2 \int |P(X) - X|\, d\|T_\sigma\| \\
&\quad \le 2 \Big(\sup_{X \in \operatorname{spt} \partial T_\sigma} |P(X) - X| \Big) \cdot \|\partial T_\sigma\| \Big\{ X : |X - P(X)| > E^{1+\delta} \Big\} \\
&\qquad + c_{10}\, E^{1+\delta} \\
&\quad \le c_{11}\, E^{1+1/(2M)-4M\delta} + c_{10}\, E^{1+\delta} \\
&\quad \le c_{12}\, E^{1+\delta}\,, \qquad (9.90)
\end{aligned}$$

where we have made use of the fact that $\delta = (9M^2)^{-1}$.

The approximating harmonic function. The aim is to show that with $1/16 < \sigma < 1/8$ chosen such that (9.86), (9.87), and (9.88) hold, $T \llcorner \mathrm{C}(0,\sigma)$ can be very closely approximated by the graph of a harmonic function.

Let $1/16 < \sigma < 1/8$ be such that (9.86), (9.87), and (9.88) (and consequently (9.90)) hold. Let $u : \overline{\mathbb{B}}^M(0,\sigma) \to \mathbb{R}$ be continuous and satisfy

$$\left.\begin{aligned} \triangle u = 0 \ \text{ on } \mathbb{B}^M(0,\sigma), \\ u = \widetilde{g}_\delta \ \text{ on } \partial\mathbb{B}^M(0,\sigma), \end{aligned}\right\} \qquad (9.91)$$

where $\widetilde{g}_\delta$ is as in (9.81), so (9.82) and (9.84) will hold.

Recall that (9.82) and (9.84) are the estimates

$$\sup_{\mathbb{B}^M(0,1/8)} |D\widetilde{g}_\delta| \le E^\delta$$

and

$$\sup\{\, |x - z|^{-\delta}\, |D\widetilde{g}_\delta(x) - D\widetilde{g}_\delta(z)| \ : \ x, z \in \mathbb{B}^M(0,1/8),\ x \ne z\} \le c_9\, E^\delta\,.$$

By applying Lemma 9.4.3 with $\hat{\sigma} = 1/(8\sigma)$, $g(x) = \widetilde{g}_\delta(x/\sigma)$, and $\hat{\eta} = \eta/\sigma$, we see that there exist constants c_{13} and c_{14} such that if u is as in (9.91), then the following estimates hold:

$$\begin{aligned}
&\sup\{\, |x - z|^{-\delta}\, |Du(x) - Du(z)| \ : \ x, z \in \mathbb{B}^M(0,\sigma),\ x \ne z\} \\
&\qquad + \sup_{\mathbb{B}^M(0,\sigma)} |Du| \le c_{13}\, E^\delta\,, \qquad (9.92)
\end{aligned}$$

$$\sup_{x \in \mathbb{B}^M(0,\eta)} |Du(x) - Du(0)|^2 \le c_{14}\, \eta^2 \int_{\mathbb{B}^M(0,\sigma)} |Du|^2\, d\mathcal{L}^M\,, \qquad (9.93)$$

for each $0 < \eta < \sigma/2$.

The comparison surface and the first use of the minimality of T. Define $G : \mathbb{B}^M(0,\sigma) \to \mathrm{C}(0,\sigma)$ by setting $G(x) = (x, u(x))$ and set

$$S = G_\#(\mathbf{E}^M \llcorner \mathbb{B}^M(0,\sigma))\,.$$

We have $\partial S = \partial \widetilde{S}_\sigma$, where we recall that $\widetilde{S}_\sigma = \widetilde{S} \llcorner \mathrm{C}(0,\sigma)$ and that $\widetilde{S}$ is defined in (9.85). Consequently, we have

$$\partial(V + S - T_\sigma) = 0\,, \tag{9.94}$$

by (9.89). This last equation tells us that

$$\partial(V + S) = \partial T_\sigma\,,$$

so we can use $V + S$ as a comparison surface for the area-minimizing surface T_σ. Since it is true for any V and S that

$$\mathbf{A}[V] + \mathbf{A}[S] \geq \mathbf{A}[V + S]\,,$$

we have

$$\mathbf{A}[V] + \mathbf{A}[S] \geq \mathbf{A}[V + S] \geq \mathbf{A}[T_\sigma]\,, \tag{9.95}$$

because T_σ is area-minimizing.

The first calculation of the difference between T_σ and S. We extend $\overrightarrow{S}$ to all of $\mathrm{C}(0,\sigma)$ by setting

$$\overrightarrow{S}(X) = \overrightarrow{S}\Big(\mathbf{p}(X), u(\mathbf{p}(X))\Big)\,. \tag{9.96}$$

Using the extension of $\overrightarrow{S}$ in (9.96) and noting that $\overrightarrow{T_\sigma} = \overrightarrow{T}$ holds $\|T_\sigma\|$-almost everywhere, we get

$$\begin{aligned}
\mathbf{A}[T_\sigma] - \mathbf{A}[S] &= \int A(\overrightarrow{T})\,d\|T_\sigma\| - \int A(\overrightarrow{S})\,d\|S\| \\
&= \int \Big(A(\overrightarrow{T}) - \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{T}\Big\rangle\Big)\,d\|T_\sigma\| \\
&\quad + \int \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{T}\Big\rangle d\|T_\sigma\| - \int A(\overrightarrow{S})\,d\|S\| \\
&= \int \Big(A(\overrightarrow{T}) - \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{T}\Big\rangle\Big)\,d\|T_\sigma\| \\
&\quad + \int \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{T}\Big\rangle d\|T_\sigma\| - \int \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{S}\Big\rangle d\|S\|\,,
\end{aligned} \tag{9.97}$$

where we have also used (9.6) to conclude that $A(\overrightarrow{S}) = \Big\langle DA(\overrightarrow{S}),\ \overrightarrow{S}\Big\rangle$.

By (9.12) we have

$$A(\overrightarrow{T}) - \left\langle DA(\overrightarrow{S}),\ \overrightarrow{T} \right\rangle = \frac{1}{2} \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 . \tag{9.98}$$

For integrands other than area, a Weierstrass condition would be used here instead of (9.12). Recalling from (9.7) that we may also treat $DA(\overrightarrow{S})$ as a differential M-form, we have

$$\int \left\langle DA(\overrightarrow{S}),\ \overrightarrow{T} \right\rangle d\|T_\sigma\| - \int \left\langle DA(\overrightarrow{S}),\ \overrightarrow{S} \right\rangle d\|S\| = [T_\sigma - S]\Big(DA(\overrightarrow{S}) \Big) . \tag{9.99}$$

Using (9.97), (9.98), and (9.99), we see that

$$\mathbf{A}[T_\sigma] - \mathbf{A}[S] = \frac{1}{2} \int \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T_\sigma\| + [T_\sigma - S]\Big(DA(\overrightarrow{S}) \Big) . \tag{9.100}$$

Use of the comparison surface and the second use of the minimality of *T*. Since (9.94) tells us that $\partial(V + S - T_\sigma) = 0$, we have

$$V + S - T_\sigma = \partial R$$

for some $(M + 1)$-dimensional current R, so (see (9.3) for notation)

$$(V + S - T_\sigma)\left(dx^M\right) = (\partial R)\left(dx^M\right) = R\left(d\,dx^M\right) = 0 .$$

Since (9.7) tells us that $DA(\mathbf{e}^M) = dx^M$, we conclude that

$$(V + S - T_\sigma)\left(DA(\mathbf{e}^M)\right) = 0 .$$

Thus we have

$$\begin{aligned} \mathbf{A}[T_\sigma] - \mathbf{A}[S] = {} & \frac{1}{2} \int \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T_\sigma\| \\ & + (T_\sigma - S)\left(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \right) \\ & + V\left(DA(\mathbf{e}^M) \right) . \end{aligned} \tag{9.101}$$

From(9.95), (9.100), and (9.101) we obtain

$$\begin{aligned} \mathbf{A}[V] & \geq \mathbf{A}[T_\sigma] - \mathbf{A}[S] \\ & \geq \frac{1}{2} \int \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T_\sigma\| \\ & \quad + (T_\sigma - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big) \\ & \quad + V(DA(\mathbf{e}^M)) . \end{aligned} \tag{9.102}$$

By (9.90), we have $\mathbf{A}[V] = \|V\|(\mathbb{R}^{M+1}) \le c_{12}\, E^{1+\delta}$ and consequently also

$$\left| V(\, DA(\mathbf{e}^M)\,) \right| \le c_{12}\, E^{1+\delta}\,.$$

Thus we have

$$2c_{12}\, E^{1+\delta} \ge \frac{1}{2} \int \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T_\sigma\| \\ + (T_\sigma - S)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\,\Big)\,. \qquad (9.103)$$

Estimating the second term on the right in (9.103). We wish to estimate the second term on the right in (9.103) by an expression similar to the first term on the right. The argument to obtain the desired estimate is sufficiently complicated that we state the result as a separate claim.

Claim. *There exist constants c_{15} and c_{16} such that*

$$\left| (T_\sigma - S)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\,\Big) \right| \\ \le c_{15}\, E^{1+\delta} + 2\, c_{16}\, E^\delta \int \left| \overrightarrow{S} - \overrightarrow{T} \right|^2 d\|T_\sigma\|\,. \qquad (9.104)$$

Proof of the Claim. We recall that h is as in Lemma 9.2.2 with $\gamma = 1$, and we introduce

$$T_\sigma^0 = G_\#^0(\mathbf{E}^M \llcorner \mathbb{B}^M(0,\sigma))\,,$$

where $G_0(x) = (x, h(x))$. By (9.24) of the Lipschitz approximation lemma, we have

$$\|T_\sigma^0 - T_\sigma\|\mathrm{C}(0,\sigma) \le c_4\, E\,, \qquad (9.105)$$

because $\gamma = 1$, $\rho = 1$, and $\sigma < 1/8$.

The estimate (9.92) gives us the bound $|Du| \le c_{13}\, E^\delta$. Then, using (9.46), we obtain

$$\left| DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M) \right| \le 2\, c_{13}\, E^\delta\,. \qquad (9.106)$$

By (9.105) and (9.106) we have

$$\left| (T_\sigma - S)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\Big) \right| \\ \le \left| (T_\sigma^0 - S)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\Big) \right| + \left| (T_\sigma - T_\sigma^0)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\,\Big) \right| \\ \le \left| (T_\sigma^0 - S)\Big(\, DA(\,\overrightarrow{S}\,) - DA(\mathbf{e}^M)\Big) \right| + c_4\, E \cdot 2\, c_{13}\, E^\delta\,. \qquad (9.107)$$

Because S is the current defined by integrating over the graph of u, we apply (9.39) with $f = u$ to obtain

$$\begin{aligned} &DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \\ &\qquad = (1+|Du|^2)^{-1/2}\left(dx^M + \sum_{i=1}^{M}(D_{x_i}u)\,dx_{\widehat{\imath}}\right) - dx^M . \end{aligned} \tag{9.108}$$

Because T_σ^0 is the current defined by integration over the graph of h, we may apply (9.36), (9.40), and (9.37), with $f = h$, and use (9.108) to find that

$$\begin{aligned} &T_\sigma^0\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M)\Big) \\ &\qquad = \int_{\mathbb{B}^M(0,\sigma)}\left[(1+|Du|^2)^{-1/2}\left(1+\sum_{i=1}^{M} D_{x_i}u\,D_{x_i}h\right) - 1\right] d\mathcal{L}^M . \end{aligned} \tag{9.109}$$

Similarly, taking $f = u$, we obtain

$$\begin{aligned} &S\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M)\Big) \\ &\qquad = \int_{\mathbb{B}^M(0,\sigma)}\left[(1+|Du|^2)^{-1/2}\left(1+\sum_{i=1}^{M} D_{x_i}u\,D_{x_i}u\right) - 1\right] d\mathcal{L}^M . \end{aligned} \tag{9.110}$$

Combining (9.109) and (9.110), we find that

$$\begin{aligned} &(T_\sigma^0 - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M)\Big) \\ &\qquad = \int_{\mathbb{B}^M(0,\sigma)}\left[(1+|Du|^2)^{-1/2}\sum_{i=1}^{M} D_{x_i}u\,D_{x_i}(h-u)\right] d\mathcal{L}^M . \end{aligned} \tag{9.111}$$

We will simplify the integrand in (9.111) so that we can use the fact that u is a harmonic function. To this end we use (9.43) to bound

$$\begin{aligned} \Bigg|\int_{\mathbb{B}^M(0,\sigma)}&\left[(1+|Du|^2)^{-1/2}\sum_{i=1}^{M} D_{x_i}u\,D_{x_i}(h-u)\right] d\mathcal{L}^M \\ &\qquad - \int_{\mathbb{B}^M(0,\sigma)}\left[\sum_{i=1}^{M} D_{x_i}u\,D_{x_i}(h-u)\right] d\mathcal{L}^M\Bigg| \end{aligned}$$

above by

$$\begin{aligned}
&\int_{\mathbb{B}^M(0,\sigma)} |Du| \left| \sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u) \right| d\mathcal{L}^M \\
&\qquad \leq \int_{\mathbb{B}^M(0,\sigma)} |Du|\,|Du|\,|D(h-u)|\, d\mathcal{L}^M \\
&\qquad \leq \int_{\mathbb{B}^M(0,\sigma)} |Du|\,|Du| \Big(|Dh| + |Du| \Big)\, d\mathcal{L}^M \\
&\qquad \leq \int_{\mathbb{B}^M(0,\sigma)} |Du|^3\, d\mathcal{L}^M + \int_{\mathbb{B}^M(0,\sigma)} |Du|\,|Du|\,|Dh|\, d\mathcal{L}^M \\
&\qquad \leq \int_{\mathbb{B}^M(0,\sigma)} |Du|^3\, d\mathcal{L}^M + \frac{1}{2} \int_{\mathbb{B}^M(0,\sigma)} |Du| \Big(|Du|^2 + |Dh|^2 \Big)\, d\mathcal{L}^M \\
&\qquad \leq \frac{3}{2} \int_{\mathbb{B}^M(0,\sigma)} |Du| \Big(|Du|^2 + |Dh|^2 \Big)\, d\mathcal{L}^M .
\end{aligned}$$

So, using the bound $|Du| \leq c_{13}\, E^{\delta}$ from (9.92), we can write

$$(T^0_\sigma - S)(DA(\overrightarrow{S}) - DA(\mathbf{e}^M)) = \int_{\mathbb{B}^M(0,\sigma)} \left[\sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u) \right] d\mathcal{L}^M + R\,, \tag{9.112}$$

where

$$|R| \leq (3/2)\, c_{13}\, E^{\delta} \int_{\mathbb{B}^M(0,\sigma)} \Big(|Du|^2 + |Dh|^2 \Big)\, d\mathcal{L}^M . \tag{9.113}$$

The fact that u is harmonic will allow us to express the integrand

$$\sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u)$$

in (9.112) as the divergence of a vector field, and thereby allow us to use the Gauss–Green theorem to replace the integral over the disk by an integral over the boundary of the disk.

Set

$$\boldsymbol{w} = (h-u) \sum_{i=1}^{M} D_{x_i} u \, \mathbf{e}_i \, .$$

We compute

$$\operatorname{div} \boldsymbol{w} = \sum_{i=1}^{M} \frac{\partial}{\partial x_i} [(h-u) D_{x_i} u]$$

$$= \sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u) + (h-u) \sum_{i=1}^{M} \frac{\partial^2 u}{\partial x_i^2}$$

$$= \sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u) \, .$$

Applying the Gauss–Green theorem (Theorem 6.2.6), we obtain

$$\int_{\mathbb{B}^M(0,\sigma)} \operatorname{div} \boldsymbol{w} \, d\mathcal{L}^M = \int_{\partial \mathbb{B}^M(0,\sigma)} \boldsymbol{w} \cdot \boldsymbol{\eta} \, d\mathcal{H}^{M-1} \, ,$$

where $\boldsymbol{\eta}$ is the outward unit normal to $\partial \mathbb{B}^M(0,\sigma)$. Hence we conclude that

$$\int_{\mathbb{B}^M(0,\sigma)} \left[\sum_{i=1}^{M} D_{x_i} u \, D_{x_i}(h-u) \right] d\mathcal{L}^M$$

$$= \int_{\partial \mathbb{B}^M(0,\sigma)} (h-u) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1}$$

$$= \int_{\partial \mathbb{B}^M(0,\sigma)} (h-\widetilde{g}_\delta) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1} \, ,$$

where we use the boundary condition in (9.91) to replace u by $\widetilde{g}_\delta$ in the last term. Thus we have

$$(T^0_\sigma - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big)$$

$$= \int_{\partial \mathbb{B}^M(0,\sigma)} (h-\widetilde{g}_\delta) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1} + R \, .$$

Now, using (9.92) to estimate $|Du| \leq c_{13}\, E^\delta$, (9.22) to estimate $|h - g_\delta| \leq 2\, c_2\, E^{1/(2M)}$, (9.83) to estimate $|g_\delta - \widetilde{g}_\delta| \leq E^{1+\delta}$, and (9.86) to estimate

$$\mathcal{H}^{M-1}\Big\{ x \in \partial \mathbb{B}^M(0,\sigma) : g_\delta(x) \neq h(x) \Big\} \leq c_{10}\, E^{1-4M\delta} \, ,$$

and recalling that $\delta = 1/(9M^2)$, we obtain the estimate

$$\left| \int_{\partial \mathbb{B}^M(0,\sigma)} (h-\widetilde{g}_\delta) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1} \right|$$

$$\leq \left| \int_{\partial \mathbb{B}^M(0,\sigma)} (h-g_\delta) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1} \right|$$

$$+ \left| \int_{\partial \mathbb{B}^M(0,\sigma)} (g_\delta - \widetilde{g}_\delta) \sum_{i=1}^{M} D_{x_i} u \, \boldsymbol{\eta}_i \, d\mathcal{H}^{M-1} \right|$$

$$\leq c_{13}\, E^\delta \, \left(\int_{\partial \mathbb{B}^M(0,\sigma)} |h - g_\delta| \, d\mathcal{H}^{M-1} \right.$$

$$\left. + \int_{\partial \mathbb{B}^M(0,\sigma)} |g_\delta - \widetilde{g}_\delta| \, d\mathcal{H}^{M-1} \right)$$

$$\leq c_{13}\, E^\delta \, \left(2\, c_2 \, E^{1/(2M)} \;\; c_{10}\, E^{1-4M\delta} + E^{1+\delta}\, M\, \Omega_M \right)$$

$$= \; c_{13} \left(2\, c_2\, c_{10}\, E^{6^{-1}\delta^{1/2}} + M\, \Omega_M \, E^\delta \right) \; E^{1+\delta} \, . \tag{9.114}$$

Combining equation (9.112) with the estimates (9.113) and (9.114), we obtain the estimate

$$\left| (T_\sigma^0 - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big) \right|$$

$$\leq c_{17}\, E^{1+\delta} + (3/2)\, c_{13}\, E^\delta \int_{\mathbb{B}^M((,0),\sigma)} (|Du|^2 + |Dh|^2)\, d\mathcal{L}^M \; ,$$

where we set $c_{17} = c_{13}\, (\, 2\, c_2\, c_{10} + M\, \Omega_M \,)$, as we may since $E \,<\, 1$.

Next, noting that we have $\operatorname{Lip} u \,\leq\, 1$ and $\operatorname{Lip} h \,\leq\, 1$, we apply Proposition 9.3.6 to conclude that

$$|Du|^2 + |Dh|^2 \leq 4 \left(|\, \overrightarrow{S} - \mathbf{e}^M|^2 + |\, \overrightarrow{T_\sigma^0} - \mathbf{e}^M|^2 \right) .$$

Assume now that the function $\overrightarrow{T_\sigma^0}$ has been extended (as has $\overrightarrow{S}$) to all of $\mathrm{C}(0, \sigma)$ by defining $\overrightarrow{T_\sigma^0}(X) \,=\, \overrightarrow{T_\sigma^0}[\mathbf{p}\,(X), h(\mathbf{p}\,(X))]$ at points where the right-hand side is defined and $\overrightarrow{T_\sigma^0}(X) \,=\, \mathbf{e}^M$ otherwise. Using also the fact that the measure $\|T_\sigma\|$ is larger than the measure $\mathcal{L}^M$, we obtain

$$\left| (T_\sigma^0 - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big) \right|$$

$$\leq c_{17}\, E^{1+\delta} + c_{16}\, E^\delta \int \left(|\, \overrightarrow{S} - \mathbf{e}^M|^2 + |\, \overrightarrow{T_\sigma^0} - \mathbf{e}^M|^2 \right) d\|T_\sigma\| \, ,$$

with $c_{16} = 4 \cdot (3/2)\, c_{13}$.

Since

$$\left| \overrightarrow{S} - \mathbf{e}^M \right|^2 \leq \left(\left| \overrightarrow{S} - \overrightarrow{T} \right| + \left| \overrightarrow{T} - \mathbf{e}^M \right| \right)^2 \leq 2 \Big(\left| \overrightarrow{S} - \overrightarrow{T} \right|^2 + \left| \overrightarrow{T} - \mathbf{e}^M \right|^2 \Big),$$

we deduce that

$$
\begin{aligned}
&\Big| (T_\sigma^0 - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big) \Big| \\
&\qquad \le c_{17}\, E^{1+\delta} \\
&\qquad\quad + c_{16}\, E^\delta \int \Big(2\Big| \overrightarrow{S} - \overrightarrow{T} \Big|^2 + 2\Big| \overrightarrow{T} - \mathbf{e}^M \Big|^2 + \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \Big|^2 \Big)\, d\|T_\sigma\| \\
&\qquad = c_{17}\, E^{1+\delta} + 2\, c_{16}\, E^\delta \int \Big| \overrightarrow{S} - \overrightarrow{T} \Big|^2 d\|T_\sigma\| \\
&\qquad\quad + 2\, c_{16}\, E^\delta \int \Big| \overrightarrow{T} - \mathbf{e}^M \Big|^2 d\|T_\sigma\| \\
&\qquad\quad + c_{16}\, E^\delta \int \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \Big|^2 d\|T_\sigma\| \\
&\qquad \le c_{17}\, E^{1+\delta} + 2\, c_{16}\, E^\delta \int \Big| \overrightarrow{S} - \overrightarrow{T} \Big|^2 d\|T_\sigma\| \\
&\qquad\quad + 4\, c_{16}\, E^\delta \cdot E + c_{16}\, E^\delta \int \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \Big|^2 d\|T_\sigma\| \,. \qquad (9.115)
\end{aligned}
$$

Using the fact that $\overrightarrow{T_\sigma^0}$ and $\overrightarrow{T}$ are $\mathcal{H}^M$-almost always simple unit M-vectors, we note that

$$
\begin{aligned}
&\int \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \Big|^2 d\|T_\sigma\| \\
&\qquad \le \int \Big| \overrightarrow{T} - \mathbf{e}^M \Big|^2 d\|T_\sigma\| + \int \Big| \, \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M |^2 - \Big| \overrightarrow{T} - \mathbf{e}^M \Big|^2 \, \Big| \, d\|T_\sigma\| \\
&\qquad \le \ 2\, E + \int \Big| \, \Big| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \Big|^2 - \Big| \overrightarrow{T} - \mathbf{e}^M \Big|^2 \, \Big| \, d\|T_\sigma\| \\
&\qquad \le 2\, E + 2 \int \Big| (\overrightarrow{T_\sigma^0} - \overrightarrow{T}) \cdot \mathbf{e}^M \Big| \, d\|T_\sigma\| \\
&\qquad \le 2\, E + 2 \int \Big| \overrightarrow{T_\sigma^0} - \overrightarrow{T} \Big| \, d\|T_\sigma\| \,.
\end{aligned}
$$

By (9.24), we have

$$
\|T_\sigma^0 - T_\sigma\| \mathrm{C}(0, \sigma) \le c_4\, E \,,
$$

so

$$
\int \Big| \overrightarrow{T_\sigma^0} - \overrightarrow{T} \Big| \, d\|T_\sigma\| \le c_4\, E \,,
$$

and we conclude that

$$\int \left| \overrightarrow{T_\sigma^0} - \mathbf{e}^M \right|^2 d\|T_\sigma\| \le 2\,(1 + c_4)\, E\,. \tag{9.116}$$

Combining (9.107), (9.115), and (9.116), we obtain the estimate

$$\begin{aligned} &\left| (T_\sigma - S)\Big(DA(\overrightarrow{S}) - DA(\mathbf{e}^M) \Big) \right| \\ &\qquad \le c_{15}\, E^{1+\delta} + 2\, c_{16}\, E^{\delta} \int \left| \overrightarrow{S} - \overrightarrow{T} \right|^2 d\|T_\sigma\|\,, \end{aligned}$$

with

$$c_{15} = c_4 \cdot 2\, c_{13} + c_{17} + 4\, c_{16} + c_{16} \cdot 2\,(1 + c_4)\,.$$

Thus the claim has been proved.

Combining the estimates. Combining (9.101) and (9.104), we obtain the estimate

$$\Big(1/2 - 2\, c_{16}\, E^{\delta}\Big) \int \left| \overrightarrow{S} - \overrightarrow{T} \right|^2 d\|T_\sigma\| \le 2\, c_{12}\, E^{1+\delta} + c_{15}\, E^{1+\delta}\,.$$

So we have

$$\int \left| \overrightarrow{S} - \overrightarrow{T} \right|^2 d\|T_\sigma\| \le c_{18}\, E^{1+\delta}\,, \tag{9.117}$$

where $c_{18} = 4\,(2\, c_{12} + c_{15})$, provided that

$$c_{16}\, E^{\delta} \le 1/8 \tag{9.118}$$

holds.

Considering candidates for θ. Consider an arbitrary $0 < \theta < \sigma/4$. We have

$$\begin{aligned} &\int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S}(0) \right|^2 d\|T\| \\ &\quad \le 2 \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T\| + 2 \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{S} - \overrightarrow{S}(0) \right|^2 d\|T\| \\ &\quad \le 2 \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T\| + 2 \Big(\sup_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{S} - \overrightarrow{S}(0) \right|^2 \Big) \cdot \|T\|\mathrm{C}(0, 2\theta)\,. \end{aligned}$$

Now

$$\|T\|\mathrm{C}(0, 2\theta) - \Omega_M\,(2\theta)^M = \frac{1}{2} \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \mathbf{e}^M \right|^2 d\|T\| \le E$$

(see (9.16)), so that

$$\|T\|\mathrm{C}(0, 2\theta) \le \Omega_M\,(2\theta)^M + E \le (1 + \Omega_M 2^M)\, \theta^M\,, \tag{9.119}$$

provided that

$$E \le \theta^M \tag{9.120}$$

holds. Successively applying (9.41), (9.93), and Proposition 9.3.6, we see that

$$\begin{aligned}
\sup_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{S} - \overrightarrow{S}(0) \right|^2 &\le \sup_{\mathrm{C}(0,2\theta)} |Du - Du(0)|^2 \\
&\le c_{14}\,\theta^2 \int_{\mathbb{B}^M(0,\sigma)} |Du|^2 \, d\mathcal{L}^M \\
&\le 4\,c_{14}\,\theta^2 \int \left| \overrightarrow{S} - \mathbf{e}^M \right|^2 d\|T_\sigma\| \,.
\end{aligned} \tag{9.121}$$

Using (9.119) and (9.121), we then deduce, subject to (9.120), that

$$\begin{aligned}
&\frac{1}{2} \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S}(0) \right|^2 d\|T\| \\
&\quad\le \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T\| \\
&\qquad\quad + c_{19}\,\theta^{M+2} \int \left| \overrightarrow{S} - \mathbf{e}^M \right|^2 d\|T_\sigma\| \\
&\quad\le \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T\| \\
&\qquad\quad + 2\,c_{19}\,\theta^{M+2} \int \left(\left| \overrightarrow{S} - \overrightarrow{T} \right|^2 + \left| \overrightarrow{T} - \mathbf{e}^M \right|^2 \right) d\|T_\sigma\| \\
&\quad\le (1 + 2\,c_{19}) \int \left| \overrightarrow{T} - \overrightarrow{S} \right|^2 d\|T_\sigma\| + 4\,c_{19}\,\theta^{M+2}\,E \,,
\end{aligned} \tag{9.122}$$

where $c_{19} = 4\,c_{14} \cdot (1 + \Omega_M 2^M)$. Combining (9.122) and (9.117), we deduce that

$$\frac{1}{2} \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S}(0) \right|^2 d\|T\| \le (1 + 2\,c_{19}) \cdot 2\,c_{18}\,E^{1+\delta} + 4\,c_{19}\,\theta^{M+2}\,E \,,$$

so

$$\frac{1}{2}\,\theta^{-M} \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S}(0) \right|^2 d\|T\| \le (1 + 4\,c_{19})\,\theta^2\,E \tag{9.123}$$

holds, provided that

$$c_{16}\,E^\delta \le 1/8, \qquad E \le \theta^M, \qquad (1 + 2\,c_{19})\,c_{18}\,E^\delta \le \theta^2 \,. \tag{9.124}$$

Note that (9.124) includes conditions (9.118) and (9.120).

Bounding the slope of the harmonic function at 0. By definition we have

$$\frac{1}{2}\,\theta^{-M}\int_{\mathrm{C}(0,2\theta)}\left|\overrightarrow{T}-\mathbf{e}^M\right|^2 d\|T\|\le\theta^{-M}\,E\,. \tag{9.125}$$

Using $\Omega_M(2\,\theta)^M\le\|T\|[\mathrm{C}(0,2\,\theta)]$, we can estimate

$$\begin{aligned}
&\left|\overrightarrow{S}(0)-\mathbf{e}^M\right|^2\\
&\quad=\frac{1}{\|T\|\mathrm{C}(0,2\,\theta)}\int_{\mathrm{C}(0,2\,\theta)}\left|\overrightarrow{S}(0)-\mathbf{e}^M\right|^2 d\|T\|\\
&\quad\le\frac{1}{\Omega_M\,(2\,\theta)^M}\int_{\mathrm{C}(0,2\,\theta)}\left|\overrightarrow{S}(0)-\mathbf{e}^M\right|^2 d\|T\|\\
&\quad\le\frac{2}{\Omega_M\,(2\,\theta)^M}\int_{\mathrm{C}(0,2\,\theta)}\left(\left|\overrightarrow{S}(0)-\overrightarrow{T}\right|^2+\left|\overrightarrow{T}-\mathbf{e}^M\right|^2\right)d\|T\|\\
&\quad\le\frac{1}{\Omega_M\,2^{M-2}}\,\frac{1}{2}\,\theta^{-M}\int_{\mathrm{C}(0,2\,\theta)}\left|\overrightarrow{S}(0)-\overrightarrow{T}\right|^2 d\|T\|\\
&\qquad\quad+\frac{1}{\Omega_M\,2^{M-2}}\,\frac{1}{2}\,\theta^{-M}\int_{\mathrm{C}(0,2\,\theta)}\left|\overrightarrow{T}-\mathbf{e}^M\right|^2 d\|T\|\,.
\end{aligned}$$

By (9.123) and (9.125), we have

$$\left|\overrightarrow{S}(0)-\mathbf{e}^M\right|^2\le c_{20}\,\theta^{-M}\,E\,, \tag{9.126}$$

provided that (9.124) holds, where we may set $c_{20}=2^{3-M}\,\Omega_M^{-1}\,(1+2\,c_{19})$.

Defining the isometry. It is easy to see that there exists a constant c_{21} such that (9.126) implies the existence of a linear isometry $\mathbf{j}$ of $\mathbb{R}^{M+1}$ with

$$\left\langle\textstyle\bigwedge_M\mathbf{j},\ \overrightarrow{S}(0)\right\rangle=\mathbf{e}^M\quad\text{and}\quad\|\mathbf{j}-\mathbf{I}_{\mathbb{R}^{M+1}}\|^2\le c_{21}\,\theta^{-M}\,E\,. \tag{9.127}$$

One way to construct such a $\mathbf{j}$ is to set $v_i=\langle\,Du(0),\ \mathbf{e}_i\,\rangle$ for $i=1,2,\ldots,M$. Then apply the Gram–Schmidt orthogonalization procedure to the set

$$\{v_1,v_2,\ldots,v_M,\mathbf{e}_{M+1}\}$$

to obtain the orthonormal basis $\{w_1,w_2,\ldots,w_{M+1}\}$. Finally, let $\mathbf{j}$ be the inverse of the isometry represented by the matrix having the vectors w_i as its columns.

Recall that $T_0=T\llcorner\mathrm{C}(0,1/2)$. By (H1) (see page 262), we have

$$\operatorname{spt}\partial T\subseteq\mathbb{R}^{M+1}\setminus\mathrm{C}(0,1)\,.$$

So we see that

$$\operatorname{dist}(\operatorname{spt}\partial T_0,\ \mathrm{C}(0,1/4))=1/4\,.$$

By Lemma 9.2.1 and the assumption that $0 \in \operatorname{spt} T$, we have

$$\sup_{X \in \mathrm{C}(0,1/2) \cap \operatorname{spt} T} |\mathbf{q}(X)| \le c_4 \, E^{1/(2M)} \,, \tag{9.128}$$

so $\operatorname{spt} \partial T_0 \subseteq \overline{\mathbb{B}}(0, 1/2 + c_4 \, E^{1/(2M)})$. By (9.127), we have

$$|x - \mathbf{j}(x)| \le (c_{21} \, \theta^{-M} \, E)^{1/2} \cdot (1/2 + c_4 \, E^{1/(2M)})$$

for $x \in \operatorname{spt} \partial T_0$. Thus if

$$(c_{21} \, \theta^{-M} \, E)^{1/2} \cdot (1/2 + c_4 \, E^{1/(2M)}) < 1/4 \tag{9.129}$$

holds, then we have

$$\operatorname{spt} \partial \mathbf{j}_{\#} T_0 \subseteq \mathbb{R}^N \setminus \mathrm{C}(0, 1/4) \,.$$

A similar argument shows that if

$$(c_{21} \, \theta^{-M} \, E)^{1/2} \cdot (\theta + c_4 \, E^{1/(2M)}) < \theta \tag{9.130}$$

holds, then we have

$$\operatorname{spt} T_0 \bigcap \mathbf{j}^{-1} \mathrm{C}(0, \theta) \subseteq \mathrm{C}(0, 2\theta) \,.$$

Selecting θ and ϵ_* to complete the proof of the lemma. If we satisfy the conditions (9.124), (9.129), and (9.130), then we obtain the estimates (9.123), (9.127), and (9.128). Those estimates are

$$\frac{1}{2} \, \theta^{-M} \int_{\mathrm{C}(0,2\theta)} \left| \overrightarrow{T} - \overrightarrow{S}(0) \right|^2 d\|T\| \overset{(9.123)}{\le} (1 + 4 \, c_{19}) \, \theta^2 \, E \,,$$

$$\|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\|^2 \overset{(9.127)}{\le} c_{21} \, \theta^{-M} \, E \,,$$

$$\sup_{X \in \mathrm{C}(0,1/2) \cap \operatorname{spt} T} |\mathbf{q}(X)| \overset{(9.128)}{\le} c_4 \, E^{1/(2M)} \,.$$

We must choose θ and ϵ_* so that the estimates (9.123), (9.127), and (9.128) will imply that (9.76), (9.78), and (9.79) hold. Finally, we need to meet the conditions (9.75) in the statement of the lemma and the condition (9.80) that allowed the use of Lemmas 9.2.1 and 9.2.2. Thus a full set of conditions that, if satisfied, complete the proof of the lemma is the following (of course, θ and ϵ_* must be positive):

$$\begin{aligned} \theta \quad & \overset{(9.75)}{<} \quad 1/8 \,, \\ \epsilon_* \quad & \overset{(9.75)}{\le} \quad (\theta/4)^{2M} \,, \\ \epsilon_* \quad & \overset{(9.80)}{<} \quad \epsilon_0 \,, \end{aligned} \tag{9.131}$$

$$c_{16}\,E^{\delta} \overset{(9.124)}{\leq} 1/8\,,$$

$$E \overset{(9.124)}{\leq} \theta^{M}\,,$$

$$(1+2\,c_{19})\,c_{18}\,E^{\delta} \overset{(9.124)}{\leq} \theta^{2}\,,$$

$$(c_{21}\,\theta^{-M}\,E)^{1/2}\cdot(1/2+c_4\,E^{1/(2M)}) \overset{(9.129)}{<} 1/4\,,$$

$$(c_{21}\,\theta^{-M}\,E)^{1/2}\cdot(\theta+c_4\,E^{1/(2M)}) \overset{(9.130)}{<} \theta\,,$$

$$c_4\,E^{1/(2M)} \overset{\text{so (9.128)}\Rightarrow\text{(9.76)}}{\leq} 1/8\,,$$

$$c_{21}\,\theta^{-M}\,E \overset{\text{so (9.127)}\Rightarrow\text{(9.78)}}{\leq} \theta^{-2M}\,E\,, \tag{9.132}$$

$$\theta^{-2M}\,E \overset{(9.77)}{\leq} 1/64\,,$$

$$(1+4\,c_{19})\,\theta^{2}\,E \overset{\text{so (9.123)}\Rightarrow\text{(9.79)}}{\leq} \theta\,E\,. \tag{9.133}$$

We first choose and fix $0 < \theta$ such that (9.131), (9.132), and (9.133) hold. This choice is clearly independent of the value of E and the choice of ϵ_*. Then we select $0 < \epsilon_*$ such that, assuming that $E < \epsilon_*$ holds, the remaining conditions are satisfied. □

9.6 The Regularity Theorem

The next theorem gives us a flexible tool that we can use in proving regularity; the proof of the theorem is based on iteratively applying Lemma 9.5.1.

Theorem 9.6.1. *Let θ and ϵ_* be as in Lemma* 9.5.1. *There exist constants c_{22} and c_{23}, depending only on M, with the following property:*

If $0 \in \operatorname{spt} T$, if $T_0 = T \llcorner \mathrm{C}(0,\rho/2)$, and if the hypotheses (H1–H5) (*see page* 262) *hold with*

$$y = 0\,,\quad \epsilon = \epsilon_*\,,$$

then

$$E(T,\,0,\,r) \leq c_{22}\,E(T,\,0,\,\rho)\,,\quad \textit{for}\quad 0 < r \leq \rho\,, \tag{9.134}$$

and there exists a linear isometry $\mathbf{j}$ of $\mathbb{R}^{M+1}$ such that

$$\operatorname{spt}\partial\mathbf{j}_{\#}T_0 \cap \mathrm{C}(0,\rho/4) = \emptyset\,,$$

$$\|\mathbf{j}-\mathbf{I}_{\mathbb{R}^{M+1}}\| \leq 4\,\theta^{-2M}\,E(T,\,0,\,\rho) \leq 4^{-2}\,, \tag{9.135}$$

$$E(\mathbf{j}_{\#}T_0,\,0,\,r) \leq c_{23}\cdot\frac{r}{\rho}\cdot E(T,\,0,\,\rho)\quad \textit{for}\quad 0 < r \leq \rho/4\,. \tag{9.136}$$

Proof. Set $\mathbf{j}_0 = \mathbf{I}_{\mathbb{R}^{M+1}}$. We will show inductively that, for $q = 1, 2, \ldots$, there are linear isometries $\mathbf{j}_q$ of $\mathbb{R}^{M+1}$ such that, writing

$$T_q = \mathbf{j}_{q\#} T_0 \,,$$

we have

$$\sup_{X \in \operatorname{spt} T_{q-1} \cap \mathrm{C}(0, \theta^{q-1}\rho/4)} |\mathbf{q}(X)| \leq \theta^{q-1}\rho/2 \quad \text{for} \quad q \geq 2 \,, \tag{9.137}$$

$$E(T_q, 0, \theta^q \rho) \leq \theta\, E(T_{q-1}, 0, \theta^{q-1}\rho) \quad \text{for} \quad q \geq 2 \,, \tag{9.138}$$

$$\|\mathbf{j}_q - \mathbf{j}_{q-1}\| \leq \theta^{-M}\, \theta^{(q-1)/2}\, E(T, 0, \rho)^{1/2} \,, \tag{9.139}$$

$$E(T_q, 0, \theta^q \rho) \leq \theta^q\, E(T, 0, \rho) \,. \tag{9.140}$$

Note that for $q = 2, 3, \ldots$, (9.140) follows from (9.138) and from the instance of (9.140) in which q is replaced by $q - 1$. Thus we need only verify (9.140) for the specific value $q = 1$.

Start of induction on q to prove (9.137)–(9.140). For $q = 1$, conditions (9.137) and (9.138) are vacuous, so we need only verify (9.139) and (9.140). Let $\mathbf{j}_1$ be the isometry whose existence is guaranteed by Lemma 9.5.1. Then the inequality (9.78) gives us (9.139), and the inequality (9.79) gives us (9.140).

Inductive step. Now suppose that (9.137)–(9.140) hold for q. We apply Lemma 9.5.1 to T_q with ρ replaced by $\theta^q \rho$. We may do so because $T_q = \mathbf{j}_{q\#} T_0$ is mass-minimizing. Inequality (9.76) of Lemma 9.5.1 gives us (9.137) with q replaced by $q + 1$.

The isometry $\mathbf{j}$ whose existence is guaranteed by Lemma 9.5.1 satisfies

$$\|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\| \leq \theta^M\, E(T_q, 0, \theta^q \rho)^{1/2} \,, \tag{9.141}$$

$$E\Big(\mathbf{j}_\#\Big(T_q \llcorner \mathrm{C}(0, \theta^q \rho/2)\Big), 0, \theta^{q+1}\rho\Big) \leq \theta\, E(T_q, 0, \theta^q \rho) \,. \tag{9.142}$$

By (9.140) and (9.141), we have

$$\|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\| \leq \theta^{-M}\, \theta^{q/2}\, E(T, 0, \rho)^{1/2} \,.$$

Setting $\mathbf{j}_{q+1} = \mathbf{j} \circ \mathbf{j}_q$, we obtain

$$\|\mathbf{j}_{q+1} - \mathbf{j}_q\| = \|(\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}) \circ \mathbf{j}_q\| = \|\mathbf{j} - \mathbf{I}_{\mathbb{R}^{M+1}}\| \leq \theta^{-M}\, \theta^{q/2}\, E(T, 0, \rho)^{1/2} \,,$$

which gives us (9.139) with q replaced by $q + 1$.

Since

$$\mathbf{j}_\#\Big(T_q \llcorner \mathrm{C}(0, \theta^q \rho/2)\Big) \llcorner \mathrm{C}(0, \theta^{q+1}\rho) = (\mathbf{j}_\# T_q) \llcorner \mathrm{C}(0, \theta^{q+1}\rho) \,,$$

we have

$$
\begin{aligned}
&E(\,T_{q+1},\ 0,\ \theta^{q+1}\rho\,) \\
&\quad = E\Big(\mathbf{j}_{\#}\Big(\,T_q \llcorner \mathrm{C}(0, \theta^q \rho/2)\Big),\ 0,\ \theta^{q+1}\rho\Big) \le \theta\ E(\,T_q,\ 0,\ \theta^q\rho\,)\,,
\end{aligned}
$$

which gives us (9.138) with q replaced by $q+1$. The inductive step has been completed.

Next we show that $\mathbf{j}_q$ has a well-defined limit as $q \to \infty$. For $Q > q \ge 0$, we estimate

$$
\begin{aligned}
\|\mathbf{j}_Q - \mathbf{j}_q\| &\le \sum_{s=q}^{Q+1} \|\mathbf{j}_{s+1} - \mathbf{j}_s\| \le \theta^{-M} \sum_{s=q}^{\infty} \theta^{s/2}\ E(T_0, 0, \rho)^{1/2} \\
&= \theta^{(q/2)-M}\ E(T_0, 0, \rho)^{1/2} \cdot \frac{1}{1-\sqrt{\theta}} \le 2\,\theta^{(q/2)-M}\ E(T_0, 0, \rho)^{1/2}\,.
\end{aligned}
$$

Thus the $\mathbf{j}_q$ form a Cauchy sequence in the mapping-norm topology. We set

$$
\mathbf{j} = \lim_{q\to\infty} \mathbf{j}_q
$$

and conclude that

$$
\|\mathbf{j} - \mathbf{j}_q\|^2 \le 4\,\theta^{q-2M}\ E(T_0, 0, \rho) \le 1/16 \tag{9.143}
$$

holds for $0 \le q$.

Recall Corollary 9.1.7, which tells us how the excess is affected by an isometry. Using (9.143) together with (9.137), (9.139), and (9.140), we see that with an appropriate choice of c_{24},

$$
E(\mathbf{j}_{\#}T_0, 0, \theta^q\rho) \le c_{24}\,\theta^q\ E(T_0, 0, \rho) \tag{9.144}
$$

holds for each $q \ge 1$. Using (9.144) together with (9.76) and (9.143) with $q = 0$, we see that, with an appropriate choice of c_{25},

$$
E(\mathbf{j}_{\#}T_0, 0, r) \le c_{25}\ (r/\rho)\ E(T_0, 0, \rho)
$$

holds for $0 < r < \rho/4$, proving (9.136). Finally, we see that (9.134) follows from (9.76), (9.136), (9.137), and (9.143), again with $q = 0$. □

We are now ready to state and prove the regularity theorem.

Theorem 9.6.2 (Regularity). *There exist constants*

$$
0 < \epsilon_1\,, \qquad 0 < c_{26} < \infty\,,
$$

depending only on M, with the following property:

If the hypotheses (H1–H5) (*see page* 262) *hold with*

$$
\epsilon = \epsilon_1\,,
$$

then spt $T \cap \mathrm{C}(y, \rho/4)$ *is the graph of a* C^1 *function* u*. Moreover,* u *satisfies the following Hölder condition with exponent* $1/2$*:*

$$\sup_{\mathbb{B}^M(y,\rho/4)} \|Du\| + \rho^{1/2} \sup_{x,z\in\mathbb{B}^M(y,\rho/4),x\neq z} |x-z|^{-1/2}\,\|Du(x)-Du(z)\|$$

$$\leq c_{26}\Big(E(T,\,y,\,\rho)\Big)^{1/2}. \tag{9.145}$$

Remark 9.6.3.

(1) Once (9.145) is established, the higher regularity theory applies to show that u is in fact real analytic. The treatise [Mor 66] is the standard reference for the higher regularity theory including the results for systems of equations needed when surfaces of higher codimension are considered.

(2) By the constancy theorem, the regularity theorem implies immediately that $T \llcorner \mathrm{C}(y, \rho/4) = G_\#\Big(\mathbf{E}^M \llcorner \mathbb{B}^M(y, \rho/4)\Big)$, where G is the mapping $x \longmapsto (x, u(x))$.

Proof. We set

$$\epsilon_1 = \min\{\,\theta^{2M}\,\epsilon_*,\; 2^{-M}\,c_6^{-2M}\,c_{22}^{-1}\,\},$$

where θ and ϵ_* are as in Lemma 9.5.1, c_{22} is as in (9.134) in Theorem 9.6.1, and c_6 is as in (9.32) in the proof of Lemma 9.2.2.

In (9.75) in the statement of Lemma 9.5.1, we required that $0 < \theta < 1/8$ and that $0 < \epsilon_* < (\theta/4)^{2M}$. Thus we have $\epsilon_1 < \epsilon_*/2^M$, so $E(T,\,y,\,\rho) < \epsilon_1$ implies that $E(T,\,z,\,\rho/2) < \epsilon_*$ for each $z \in \mathbb{B}^M(y, \rho/2)$. Therefore, after translating the origin and replacing ρ by $\rho/2$, we can apply Theorem 9.6.1 to conclude that

$$E(T,\,z,\,r) \leq c_{22}\,E(T,\,z,\,\rho/2) \leq 2^M\,c_{22}\,E(T,\,y,\,\rho) \tag{9.146}$$

holds for $0 < r \leq \rho/2$ and $z \in \mathbb{B}^M(y, \rho/2)$. Theorem 9.6.1 also tells us that

$$E(\mathbf{j}_{z\#}\,T_z,\,z,\,r) \leq c_{23}\cdot\frac{r}{\rho/2}\cdot E(T,\,z,\,\rho/2)$$

$$\leq 2^{M+1}\,c_{23}\,E(T,\,y,\,\rho) \tag{9.147}$$

holds for $0 < r \leq \rho/8$, where $T_z = T \llcorner \mathrm{C}(y, \rho/4)$. It also says that $\mathbf{j}_z$ is an isometry of $\mathbb{R}^{M+1}$ with spt $\partial\mathbf{j}_{z\#}T_z \cap \mathrm{C}(z, \rho/8) = \emptyset$, $\mathbf{j}_z(z, w) = (z, w)$ for some point $(z, w) \in$ spt T, and

$$\|D\mathbf{j}_z - \mathbf{I}_{\mathbb{R}^{M+1}}\| \leq 4\,\theta^{-2M}\,E(T,\,z,\,\rho/2) \leq 4^{-2}. \tag{9.148}$$

In (9.80) of the proof of Lemma 9.5.1 we required that $\epsilon_* < \epsilon_0$, where ϵ_0 is as in Lemma 9.2.1. Thus we also have $\epsilon_1 < \epsilon_0$. Now we look in detail at the construction in the proof of Lemma 9.2.2 with $\gamma = 1$. In particular, when the choice

$$\eta = c_6^{-2M}$$

is made in (9.34), we guarantee that $\eta = c_6^{-2M}$ is strictly less than ϵ_0. Since $\epsilon_1 \le 2^{-M}\, c_6^{-2M}$ holds, (9.146) implies that

$$E(\, T,\ z,\ r\,) \le c_6^{-2M} = \eta$$

holds for $0 < r \le \rho/2$ and $z \in \mathbb{B}^M(y, \rho/2)$. Thus the set A defined in (9.28) contains all of $\mathbb{B}^M(y, \rho/2)$. We conclude that there exists a Lipschitz function $g : \mathbb{B}^M(y, \rho/4) \to \mathbb{R}$ such that

$$\operatorname{Lip}\ g \le 1\,, \tag{9.149}$$

$$T \llcorner \mathrm{C}(y, \rho/4) = G_{\#}\Big(\mathbf{E}^M \llcorner \mathbb{B}^M(y, \rho/4)\Big)\,, \tag{9.150}$$

with $G : \mathbb{B}^M(y, \rho/4) \to \mathrm{C}(y, \rho/4)$ defined by $G(x) = (x, g(x))$.

If $L_z : \mathbb{R}^M \to \mathbb{R}$ denotes the linear map whose graph is mapped to $\mathbb{R}^M \times \{0\}$ by $D\mathbf{j}_z$, then estimates (9.147), (9.148), (9.149) and equation (9.150) imply that

$$r^{-M} \int_{\mathbb{B}^M(z,r)} \|Dg - L_z\|^2\, d\mathcal{L}^M \le c_{27}\ (r/\rho)\ E(\, T,\ y,\ \rho\,) \tag{9.151}$$

holds for $0 < r \le \rho/8$ and $z \in \mathbb{B}^M(y, \rho/4)$, where c_{27} is an appropriate constant.

We will apply (9.151) with $z_1, z_2 \in \mathbb{B}^M(y, \rho/4)$ and with $r = |z_1 - z_2| < \rho/8$. Setting $z_* = (z_1 + z_2)/2$ and $B = \mathbb{B}^M(z_1, r) \bigcap \mathbb{B}^M(z_2, r)$, we estimate

$$\begin{aligned}
\Omega_M\ (r/2)^M\ \|L_{z_1} - L_{z_2}\|^2 &\le \int_B \|L_{z_1} - L_{z_2}\|^2\, d\mathcal{L}^M \\
&\le 2 \int_B \Big(\|DL_{z_1} - Dg\|^2 + \|Dg - L_{z_2}\|^2\Big)\ d\mathcal{L}^M \\
&\le 2 \int_{\mathbb{B}^M(z_1,r)} \|DL_{z_1} - Dg\|^2\, d\mathcal{L}^M \\
&\quad + 2 \int_{\mathbb{B}^M(z_2,r)} \|Dg - L_{z_2}\|^2\, d\mathcal{L}^M \\
&\le 2\, r^M\ c_{27}\ (r/\rho)\ E(\, T,\ y,\ \rho\,)\,.
\end{aligned}$$

Thus we have

$$\|L_{z_1} - L_{z_2}\|^2 \le 2^{M+1}\ \Omega_M^{-1}\ c_{27}\ (|z_1 - z_2|/\rho)\ E(\, T,\ y,\ \rho\,)\,.$$

Since (9.151) also implies that

$$Dg(z) = L_z$$

holds for $\mathcal{L}^M$-almost all $z \in \mathbb{B}^M(y, \rho/4)$, we conclude that

$$\|Dg(z_1) - Dg(z_2)\| \le c_{28}\,(|z_1 - z_2|/\rho)^{1/2}\, E(T,\ y,\ \rho)^{1/2} \tag{9.152}$$

holds for $\mathcal{L}^M$-almost all $z_1, z_2 \in \mathbb{B}^M(y, \rho/4)$, where we set

$$c_{28} = 2^{(M+1)/2}\,\Omega_M^{-1/2}\,c_{27}^{1/2}\,.$$

Since g is Lipschitz, we conclude that g is C^1 in $\mathbb{B}^M(y, \rho/4)$, that (9.152) holds for all $z_1, z_2 \in \mathbb{B}^M(y, \rho/4)$, and that (9.145) follows from (9.148) and (9.152) when we set $u = g$. □

9.7 Epilogue

In our exposition of the regularity results, we made the simplifying assumptions that the current being studied was of *codimension one* and that it minimized the integral of the *area integrand*. Relaxing these assumptions introduces notational and technical complexity and requires deeper results to obtain bounds for solutions of the appropriate partial differential equation or system of partial differential equations. Nonetheless the proof of the regularity theorem goes through—as Schoen and Simon showed.

What is affected fundamentally by relaxing the assumptions is the applicability of the regularity theorem and the further results that can be proved. It is the hypothesis (H3) that causes the most difficulty in applying Theorem 9.6.2.

Because we have limited our attention to the codimension-one case, we have Theorem 7.5.5 available to decompose a mass-minimizing current into a sum of mass-minimizing currents each of which is the boundary of the current associated with a set of locally finite perimeter. Thus we have proved the following theorem.

Theorem 9.7.1. *If T is a mass-minimizing, integer-multiplicity current of dimension M in $\mathbb{R}^{M+1}$, then, for $\mathcal{H}^M$-almost every $a \in \operatorname{spt} T \setminus \operatorname{spt} \partial T$, there is $r > 0$ such that $\mathbb{B}(a, r) \bigcap \operatorname{spt} T$ is the graph of a C^1 function.*

The more general form of the regularity theorem in [SS 82] extends Theorem 9.7.1 to currents minimizing the integral of smooth elliptic integrands and, in higher codimensions, yields a set of regular points that is dense, though not necessarily of full measure.

Suppose that T is an M-dimensional, integer-multiplicity current in $\mathbb{R}^N$, and suppose that T minimizes the integral of a smooth M-dimensional elliptic integrand F. Let us denote the set of regular points of the current T by $\operatorname{reg} T$ and the set of singular points of T by $\operatorname{sing} T$. More precisely, $\operatorname{reg} T$ is defined by

$$\begin{aligned}\operatorname{reg} T = {} & (\operatorname{spt} T \setminus \operatorname{spt} \partial T) \\ & \cap \left\{a : \exists r > 0 \text{ such that } \mathbb{B}(a, r) \bigcap \operatorname{spt} T \text{ is the graph of a } C^1 \text{ function}\right\}\end{aligned}$$

and

$$\operatorname{sing} T = \operatorname{spt} T \setminus \left(\operatorname{spt} \partial T \bigcup \operatorname{reg} T\right)\,.$$

Table 9.1. Interior regularity of minimizing currents.

	$\boldsymbol{F = A}$	$\boldsymbol{F \neq A}$
$\boldsymbol{N - M = 1}$	$\dim_{\mathcal{H}}(\operatorname{sing} T) \leq M - 7$ [Fed 70]	$\mathcal{H}^{M-2}(\operatorname{sing} T) = 0$ [SSA 77]
$\boldsymbol{N - M \geq 2}$	$\dim_{\mathcal{H}}(\operatorname{sing} T) \leq M - 2$ [Alm 00]	$\operatorname{reg} T$ is dense in $\operatorname{spt} T \setminus \operatorname{spt} \partial T$ [Alm 68]

Table 9.1 summarizes what is known about $\operatorname{reg} T$ and $\operatorname{sing} T$ (and gives a reference for each result). In the table, A denotes the M-dimensional area integrand.

One can also consider the question of what happens near points of $\operatorname{spt} \partial T$, that is, *boundary regularity* as opposed to the *interior regularity* considered above. The earliest results in the context of geometric measure theory are in William K. Allard's work [All 68], [All 75]. Allard's results focus on the area integrand. Robert M. Hardt considered more general integrands in [Har 77]. For area-minimizing hypersurfaces, the definitive result is that of Hardt and Simon [HS 79], which tells us that if ∂T is associated with a C^2 submanifold, then, near every point of $\operatorname{spt} \partial T$, the set $\operatorname{spt} T$ is a C^1 embedded submanifold-with-boundary. More recently, Frank Duzaar and Klaus Steffen (see [DS 02]) have given a unified argument applicable to the interior and boundary regularity of currents that "almost" locally minimize the integral of a general elliptic integrand.

Regularity theory is not a finished subject. The finer structure of the singular set is not generally known (2-dimensional area-minimizing currents are an important exception—see [Cha 88]), so understanding the singular set remains a challenge. Also, techniques created to answer questions about surfaces that minimize integrals of elliptic integrands have found applicability in other areas, for instance, to systems of partial differential equations (e.g., [Eva 86]), mean curvature flows (e.g., [Whe 05]), and harmonic maps (e.g., [Whe 97]). The future will surely see more progress.

Appendix

A.1 Transfinite Induction

We provide a sketch of transfinite induction over the smallest uncountable ordinal. Since we use transfinite induction only for the specific purpose of constructing the Borel sets, we have kept the discussion here minimal. The reader interested in a more complete discussion should see [Hal 74, Sections 17–19].

Definition A.1.1. A relation $<$ on a set $\mathcal{Z}$ is a *well-ordering* if

(1) for $x, y \in \mathcal{Z}$ exactly one of $x < y$, $y < x$, and $x = y$ holds,
(2) for $x, y, z \in \mathcal{Z}$, $x < y$ and $y < z$ imply $x < z$,
(3) if $A \subseteq \mathcal{Z}$ is nonempty, then there exists $a \in A$ such that $a < x$ holds for all $x \in A$ with $x \neq a$; in this case, we call a the *least element of* A and write $a = \min A$.

Recall the well-ordering principle (see for instance [Fol 84] or [Roy 88]).

Theorem A.1.2 (Well-Ordering Principle). *Every set can be well-ordered.*

Now choose any uncountable set $\mathcal{Z}$, and let it be well-ordered by the relation $<$. Every nonempty set has a least element. In particular, the entire well-ordered set will have a least element: Let 1 denote that least element of $\mathcal{Z}$, so $1 = \min \mathcal{Z}$. Now that 1 has been defined, we can write $2 = \min(\mathcal{Z} \setminus \{1\})$. Of course, this process can be continued by using induction over the positive integers. Below we will describe induction over an ordered set of cardinality strictly larger than the cardinality of the integers.

The set of *predecessors of* $\alpha \in \mathcal{Z}$ is $\{ z \in \mathcal{Z} : z < \alpha \}$. We would like to consider the minimum of the set

$$A = \left\{ x \in \mathcal{Z} : \{z \in \mathcal{Z} : z < x\} \text{ is uncountable} \right\}.$$

At the moment we cannot guarantee that this set is nonempty. If A happens to be empty, change the name of $\mathcal{Z}$ to $\widetilde{\mathcal{Z}}$ and let the new $\mathcal{Z}$ be $\mathcal{Z} = \widetilde{\mathcal{Z}} \cup \{x^*\}$, where $x^* \notin \widetilde{\mathcal{Z}}$.

Extend the ordering $<$ from $\widetilde{\mathcal{Z}}$ to $\mathcal{Z}$ by requiring that $x < x^*$ hold for every $x \in \widetilde{\mathcal{Z}}$. With these changes, $\mathcal{Z}$ will still be well-ordered and A will be nonempty.

Let ω_1 be the least element of $\mathcal{Z}$ for which the set of predecessors is uncountable; that is,

$$\omega_1 = \min\left\{ x \in \mathcal{Z} : \{z \in \mathcal{Z} : z < x\} \text{ is uncountable} \right\}.$$

By Definition A.1.1(3), we have

$$\{z \in \mathcal{Z} : z < \omega_1\} \text{ is uncountable.} \tag{A.1}$$

The next lemma describes induction over ω_1. This is an instance of *transfinite induction*.

Lemma A.1.3 (Transfinite Induction over ω_1). *Suppose that $\mathbf{P}(\alpha)$ is a statement that is either true or false depending on the choice of the parameter $\alpha < \omega_1$. If*

(1) $\mathbf{P}(1)$ *is true and*
(2) *for $\alpha < \omega_1$, $\mathbf{P}(\alpha)$ is true whenever $\mathbf{P}(\beta)$ is true for all $\beta < \alpha$,*

then $\mathbf{P}(\alpha)$ is true for all $\alpha < \omega_1$.

Proof. If $A = \{\, \alpha : \alpha < \omega_1,\ \mathbf{P}(\alpha) \text{ is false} \,\}$ were nonempty, then $\widetilde{\alpha} = \min A$ would exist. Note that by (1), $\widetilde{\alpha}$ cannot equal 1. Then by (2), $\widetilde{\alpha}$ cannot be any other element of $\{z \in \mathcal{Z} : z < \omega_1\}$, and we have reached a contradiction. □

The next lemma tells us that we cannot traverse ω_1 in countably many steps. Thus there is an essential difference between induction over the positive integers and induction over ω_1. In the construction of the Borel sets, this lemma allows us to conclude that induction over ω_1 is sufficient to construct all the Borel sets; that is, no new sets would be constructed if we continued the inductive construction beyond ω_1.

Lemma A.1.4. *If $\alpha_1, \alpha_2, \ldots$ is a sequence in $\mathcal{Z}$ and if $\alpha_i < \omega_1$ holds for each $i = 1, 2, \ldots$, then there is α^* with $\alpha^* < \omega_1$ and $\alpha_i < \alpha^*$ for all i.*

Proof. Since $\alpha_i < \omega_1$, the set of predecessors of α_i is countable. Thus the set

$$A = \{\alpha_i : i = 1, 2, \ldots\} \cup \bigcup_{i=1}^{\infty} \{x \in \mathcal{Z} : x < \alpha_i\}$$

is a countable union of countable sets and hence is countable.

By (A.1), $\{z \in \mathcal{Z} : z < \omega_1\}$ is uncountable, while A is merely countable, so there exists

$$\alpha^* \in \{z : z < \omega_1\} \setminus A\,.$$

For each i, α^* is unequal to α_i and is not a predecessor of α_i, so $\alpha_i < \alpha^*$ must hold. Thus α^* is as required. □

A.2 Dual Spaces

Throughout this section we let V be a vector space over the real numbers.

Definition A.2.1. The *dual space* of V, denoted by V^*, is the set of real-valued linear functions on V together with the operations of scalar multiplication and vector addition defined, for $\alpha \in \mathbb{R}$ and $\xi, \eta \in V^*$, by setting

$$(\alpha\xi)(v) = \alpha(\xi(v)), \qquad \text{for } v \in V,$$
$$(\xi + \eta)(v) = (\xi(v)) + (\eta(v)), \text{ for } v \in V.$$

With these operations, V^* forms a vector space in its own right.

Remark A.2.2.

(1) The elements of the dual space V^* are often called *functionals*, providing a briefer way to say "real-valued linear functions." Elements of V^* are also called *dual vectors* or *covectors*.
(2) Our emphasis in this section will be algebraic. On the other hand, if the vector space V is endowed with a topology and if the vector space operations are continuous with respect to that topology, then V is called a *topological vector space*. It then makes sense to consider the *continuous linear functionals* on V. The set of continuous linear functionals is denoted by V' and it forms a subspace of V^*. The continuous linear functionals are the object of study in *functional analysis*.

Notation A.2.3. Because of the way the vector space operations are defined in V^*, the expression

$$\xi(v)\,,$$

where $\xi \in V^*,\ v \in V$, is linear in *both* ξ and v. The symmetry of this situation is better emphasized by writing

$$\langle \xi,\ v\rangle = \xi(\eta)\,.$$

The bilinear function $\langle \xi,\ v\rangle$ is called the *dual pairing*.

Example A.2.4. When $\mathbb{R}^N$ is viewed as a vector space, its elements are typically represented by column vectors:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}\,.$$

Elements of the dual space $(\mathbb{R}^N)^*$ are represented by row vectors:

$$\xi = (\xi_1 \ \ \xi_2 \ \ \ldots \ \ \xi_N)\,.$$

With these notational conventions the dual pairing is expressed as

$$\langle \xi,\, x\rangle = (\xi_1 \;\; \xi_2 \;\; \ldots \;\; \xi_N) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}, \tag{A.2}$$

where the operation on the right-hand side of (A.2) is *ordinary matrix multiplication*. Equation (A.2) justifies our convention of writing the element of the dual space on the left in the dual pairing. This convention is not followed universally, since some authors put the dual space element on the right. □

Definition A.2.5. Suppose a basis for V has been selected:

$$\mathcal{B} = \{b_a\}_{a\in A}\,,$$

where A is some index set. For each b_a we define $b_a^* \in V^*$ by setting

$$\langle b_a^*,\, b_{a'}\rangle = \begin{cases} 1, \text{ if } a' = a, \\ 0, \text{ if } a' \neq a, \end{cases}$$

for basis elements $b_{a'}$ and extending by linearity to all of V. The mapping

$$b_a \longmapsto b_a^*$$

can in turn be extended from $\mathcal{B}$ to all of V by linearity, thus defining a mapping $i_{\mathcal{B}} : V \to V^*$.

Remark A.2.6. We will see in Corollary A.2.9 that when V is finite-dimensional, the set of $\{b_a^*\}_{a\in A}$ forms a basis for V^* called the "dual basis."

Lemma A.2.7. *The map $i_{\mathcal{B}} : V \to V^*$ is one-to-one.*

Proof. Suppose $i_{\mathcal{B}}(v) = 0$. Write $v = \sum_{j=1}^n \alpha_j b_{a_j}$ as we may since $\mathcal{B}$ is a basis for V. By linearity

$$i_{\mathcal{B}}(v) = \sum_{j=1}^n \alpha_j\, i_{\mathcal{B}}(b_{a_j})$$

holds, so for any $j_0 \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned} 0 &= \langle i_{\mathcal{B}}(v),\, b_{a_{j_0}}\rangle \\ &= \sum_{j=1}^n \alpha_j\, \langle b_{a_j}^*,\, b_{a_{j_0}}\rangle \\ &= \alpha_{j_0}\,. \end{aligned}$$

Thus we have $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ and consequently $v = 0$. □

Lemma A.2.8. *The map $i_{\mathcal{B}} : V \to V^*$ is an isomorphism if and only if V is finite-dimensional.*

Proof. By Lemma A.2.7, we need to show that $i_{\mathcal{B}}$ is surjective if and only if V is finite-dimensional.

First suppose V is infinite-dimensional. We define $X \in V^*$ by setting

$$\left\langle X, \sum_{j=1}^{n} \alpha_j b_{a_j} \right\rangle = \sum_{j=1}^{n} \alpha_j .$$

We cannot express X as a finite linear combination of the functionals b_a^*, so X is not in the range of $i_{\mathcal{B}}$ (one can write X *formally* as an infinite linear combination of b_a^*, namely as $X = \sum_{a \in \mathcal{B}} b_a^*$, because whenever X is evaluated on $v \in V$ only finitely many of the summands will be nonzero).

Now suppose that V is finite-dimensional. We can write

$$\mathcal{B} = \{b_1, b_2, \ldots, b_N\} .$$

Letting $\xi \in V^*$ be arbitrary, we see by linearity that

$$\xi = \sum_{i=1}^{N} \langle \xi, b_i \rangle b_i^* . \qquad \square$$

From the proof of Lemma A.2.8 we obtain the following corollary.

Corollary A.2.9. *If V is finite-dimensional with basis $\mathcal{B} = \{b_1, b_2, \ldots, b_N\}$, then $\mathcal{B}^* = \{b_1^*, b_2^*, \ldots, b_N^*\}$ is a basis for V^* called the* dual basis.

Remark A.2.10. As was noted in Section 6.1, for the special case of $\mathbb{R}^N$ with coordinates $x_1, x_2, \ldots, x_N$ and standard basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$, it is traditional to write dx_i to denote the dual of $\mathbf{e}_i$; that is,

$$dx_i = \mathbf{e}_i^* , \quad \text{for } i = 1, 2, \ldots, N . \tag{A.3}$$

The reason for this notation is made clear in Section A.3.

Remark A.2.11. One can consider the dual space of V^*, denoted by V^{**}. It is always possible to embed V into V^{**} using the mapping $\mathcal{I} : V \to V^{**}$ defined by setting

$$\langle \mathcal{I}(v), \xi \rangle = \langle \xi, v \rangle ,$$

for $v \in V$ and $\xi \in V^*$. If V is finite-dimensional with basis $\mathcal{B}$ and dual basis $\mathcal{B}^*$, then one checks that $\mathcal{I} = i_{\mathcal{B}^*} \circ i_{\mathcal{B}}$. Thus by Lemma A.2.8, we see that *if V is finite-dimensional, then $\mathcal{I}$ is an isomorphism.* Because the natural embedding $\mathcal{I}$ is an isomorphism when V is finite-dimensional, it is common in the finite-dimensional case to identify V and V^{**}.

The Dual of an Inner Product Space

In this subsection, we assume that V also has the structure of an inner product space and let the inner product of $x, y \in V$ be denoted by $x \cdot y$. In this case, every element $x \in V$ defines a corresponding element $\xi_x \in V^*$ by setting

$$\langle \xi_x, y \rangle = x \cdot y\,.$$

The mapping $x \longmapsto \xi_x$ is one-to-one because $\langle \xi_x, x \rangle = x \cdot x = 0$ if and only if $x = 0$.

Remark A.2.12. If V has the orthonormal basis $\mathcal{B}$, then the mapping $i_{\mathcal{B}}$ is the same as the mapping $x \longmapsto \xi_x$

Lemma A.2.13. *If V is a finite-dimensional inner product space, then the mapping $x \longmapsto \xi_x$ is an isomorphism of V onto V^*.*

Proof. If V is finite-dimensional, then V^* is also finite-dimensional and $\dim V = \dim V^*$. Since the mapping $x \longmapsto \xi_x$ is one-to-one, its image must have the same dimension as its domain, thus it maps onto V^*. □

Lemma A.2.13 gives us a natural way to define an inner product on the dual of a finite-dimensional inner product space, which we do in the next definition.

Definition A.2.14. If V is a finite-dimensional inner product space, then the *dual inner product* on V^* is defined by requiring the mapping $x \longmapsto \xi_x$ to be an isometry. Equivalently, if $\mathcal{B}$ is an orthonormal basis for V, then we decree $\mathcal{B}^*$ to be an orthonormal basis for V^*.

Remark A.2.15. Even with the extra structure of an inner product on V, if V is infinite-dimensional, then V and V^* are not isomorphic. What is true is that if V is given the metric topology derived from the inner product and if V is complete in that metric, then V is isomorphic to the vector space V' of continuous linear functionals (a result known as the Riesz representation theorem). Such a space V, namely, an inner product space that forms a complete metric space when endowed with the metric derived from the inner product, is called a *Hilbert space*. A pertinent reference is [Con 90].

A.3 Line Integrals

In a course on vector calculus, a student will learn about line integrals along a curve in Euclidean space, first in $\mathbb{R}^2$ and then more generally in $\mathbb{R}^3$, or perhaps even in $\mathbb{R}^N$. Such an introduction typically will involve two types of line integral, one being the integral with respect to arc length

$$\int_C f\,ds$$

and the second being the the integral of a differential form

$$\int_C f\,dx + g\,dy + h\,dz\,.$$

The vector calculus definition of a line integral is operational.

Definition A.3.1. If the curve C is parametrized by the smooth function $\gamma : [a, b] \to \mathbb{R}^N$, then the *integral with respect to arc length* of the function f over the curve C is given by

$$\int_C f\,ds := \int_a^b f[\gamma(t)]\,|\gamma'(t)|\,dt\,.$$

If we suppose the component functions of C are $\gamma_1, \gamma_2, \ldots,$ and γ_N, then the *integral of the differential form* $f_1\,dx_1 + f_2\,dx_2 + \cdots + f_N\,dx_N$ over the curve C is given by

$$\int_C f_1\,dx_1 + f_2\,dx_2 + \cdots + f_N\,dx_N := \int_a^b \left(\textstyle\sum_{i=1}^N f_i[\gamma(t)]\,\gamma_i'(t)\right) dt\,. \tag{A.4}$$

The mnemonic for the latter definition is that the component functions *could* be written

$$x_1(t),\ x_2(t),\ \ldots,\ x_N(t)\,,$$

inspiring the mechanical calculations

$$dx_1 = x_1'(t)\,dt,\ dx_2 = x_2'(t)\,dt,\ \ldots,\ dx_N = x_N'(t)\,dt\,.$$

The operational definition of the line integral of a differential form leaves unanswered the question of what a differential form *is*. To answer that question, recall that if $\mathbb{R}^N$ has coordinates $x_1, x_2, \ldots, x_N$ and has the standard basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$, then $dx_1, dx_2, \ldots, dx_N$ are dual to $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N$, respectively. So for any fixed point in the domain of $f_1, f_2, \ldots, f_N$,

$$f_1\,dx_1 + f_2\,dx_2 + \cdots + f_N\,dx_N$$

is an element of the dual space of $\mathbb{R}^N$. In light of this interpretation of the differential form, the integrand on the right-hand side of (A.4),

$$\textstyle\sum_{i=1}^N f_i[\gamma(t)]\,\gamma_i'(t)\,,$$

is the dual pairing of

$$f_1\,dx_1 + f_2\,dx_2 + \cdots + f_N\,dx_N$$

against the velocity vector of the curve

$$\gamma_1'(t)\,\mathbf{e}_1 + \gamma_2'(t)\,\mathbf{e}_2 + \cdots + \gamma_N'(t)\,\mathbf{e}_N\,.$$

Exterior Differentiation

The fundamental theorem of calculus tells us that integration and differentiation of functions can be thought of as inverse operations. We might wonder whether the line integral is also inverse to some type of differentiation. Indeed, "exterior differentiation," which we define next, plays that role.

Definition A.3.2. Suppose that $U \subseteq \mathbb{R}^N$ is open. If $F : U \to \mathbb{R}$ is differentiable, then the *exterior derivative* of F, denoted by dF, is the differential form defined by

$$dF = \frac{\partial F}{\partial x_1}\,dx_1 + \frac{\partial F}{\partial x_2}\,dx_2 + \cdots + \frac{\partial F}{\partial x_N}\,dx_N\,. \tag{A.5}$$

Example A.3.3. Fix $i \in \{1, 2, \ldots, N\}$. Suppose $F : \mathbb{R}^N \to \mathbb{R}$ is defined by setting

$$F(x_1, x_2, \ldots, x_N) = x_i\,. \tag{A.6}$$

We compute

$$dF = dx_i\,. \tag{A.7}$$

The function F defined by (A.6) is often denoted by x_i. If we were to use that notation then (A.7) would become the tautology "$dx_i = dx_i$." □

The next theorem shows us that the line integral is indeed the inverse operation to exterior differentiation, justifying the use of the notation "dF."

Theorem A.3.4. *If $U \subseteq \mathbb{R}^N$ is open, $F : U \to \mathbb{R}$ is continuously differentiable, and $C \subseteq U$ is a curve with initial point p_0 and terminal point p_1, then*

$$\int_C dF = F(p_1) - F(p_0)\,.$$

Proof. Suppose C is parametrized by the smooth function $\gamma : [a, b] \to \mathbb{R}^N$. Then the initial point of the curve is $p_0 = \gamma(a)$ and the terminal point of the curve is $p_1 = \gamma(b)$.

Consider the function $\phi : \mathbb{R} \to \mathbb{R}$ defined by $\phi(t) = F[\gamma(t)]$. The fundamental theorem of calculus tells us that

$$\int_a^b \phi'(t)\,dt = \phi(b) - \phi(a) = F(p_1) - F(p_0)\,.$$

On the other hand, the chain rule and (A.4) tell us that

$$\int_a^b \phi'(t)\,dt = \int_a^b \left(\frac{\partial F}{\partial x_1}\,\gamma_1'(t) + \frac{\partial F}{\partial x_2}\,\gamma_2'(t) + \cdots + \frac{\partial F}{\partial x_N}\,\gamma_N'(t) \right) dt = \int_C dF\,. \qquad \square$$

A.4 Pullbacks and Exterior Derivatives

Theorem 6.2.9 tells us that for differential forms, the operations of pullback and exterior differentiation commute. In this section, we give an alternative proof of that theorem. The proof given here hinges on the fact that the order of differentiation does not matter in a second derivative of a C^2 function.

We will need to develop a new expression for the exterior derivative.

Definition A.4.1. Suppose that the differential m-form $\phi : U \to \bigwedge^m (\mathbb{R}^N)$ is given and is C^1. For any vector $v \in \mathbb{R}^N$, the *directional derivative* of ϕ in the direction v is the m-form, denoted by $D_v\phi$, that when applied to the m vectors $v_1, v_2, \ldots, v_m \in \mathbb{R}^N$ at the point p is defined by setting

$$\begin{aligned} &\Big\langle \Big(D_v\phi(p) \Big),\ v_1 \wedge v_2 \wedge \cdots \wedge v_m \Big\rangle \\ &\quad = \lim_{t \to 0} \frac{\langle \phi(p + tv),\ v_1 \wedge v_2 \wedge \cdots \wedge v_m \rangle - \langle \phi(p),\ v_1 \wedge v_2 \wedge \cdots \wedge v_m \rangle}{t}\,. \qquad \text{(A.8)} \end{aligned}$$

To obtain an $(m+1)$-form by differentiating ϕ, we need to modify the directional derivative so as to make it an alternating function of $m+1$ vectors. The standard way to convert a multilinear function into an *alternating* multilinear function is to average the alternating sum over all permutations of the arguments. Since the underlying m-form ϕ is already alternating in its m arguments, the required alternating sum simplifies to the following:

$$\frac{1}{m+1} \sum_{i=1}^{m+1} (-1)^{i+1} \Big\langle \Big(D_{v_i}\phi(p) \Big),\ v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_{m+1} \Big\rangle\,. \qquad \text{(A.9)}$$

Expressions such as $v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} \wedge \cdots \wedge v_{m+1}$ occur with enough frequency that it is useful to have a special notation for them.

Notation A.4.2. Given vectors $v_1, v_2, \ldots, v_\ell$, we set

$$\begin{aligned} v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_\ell &= v_1 \wedge \cdots \wedge v_{j-1} \wedge v_{j+1} \wedge \cdots \wedge v_\ell\,, \qquad &\text{(A.10)} \\ v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_\ell & \\ = v_1 \wedge \cdots \wedge v_{i-1} \wedge v_{i+1} &\wedge \cdots \wedge v_{j-1} \wedge v_{j+1} \wedge \cdots \wedge v_\ell\,. \qquad &\text{(A.11)} \end{aligned}$$

Of course, we would like to see that the expression in (A.9) agrees with the exterior derivative as previously defined. That is the content of Proposition A.4.3.

Proposition A.4.3. *Suppose that the differential m-form $\phi : U \to \bigwedge^m (\mathbb{R}^N)$ is given and is C^1. Then, for any set of $m+1$ vectors $v_1, v_2, \ldots, v_{m+1} \in \mathbb{R}^N$, we have*

$$\langle d\phi(p),\ v_1 \wedge v_2 \wedge \cdots \wedge v_m \wedge v_{m+1}\rangle$$

$$= \frac{1}{m+1} \sum_{i=1}^{m+1} (-1)^{i+1} \Big\langle \Big(D_{v_i}\phi(p) \Big),\ v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_{m+1} \Big\rangle . \tag{A.12}$$

Proof. The result is easily verified by checking that it is true when the v_i are all chosen to be standard basis vectors. □

Theorem A.4.4. *Suppose that $U \subseteq \mathbb{R}^N$ is open and $F : U \to \mathbb{R}^M$ is C^2. Fix a point $p \in U$. If the differential m-form ϕ is defined and C^1 in a neighborhood of $F(p)$, then $d\left(F^{\#}\phi\right) = F^{\#}\Big(d\phi\Big)$ holds at p.*

Proof. Fix vectors $u, v_1, v_2, \ldots, v_m \in \mathbb{R}^N$. We do a preliminary calculation of the directional derivative in the direction u of $\phi^{\#}F$ applied to the m-vector $v_1 \wedge v_2 \wedge \cdots \wedge v_m$. Writing $w = D_u F$, we obtain

$$\begin{aligned}
&\langle D_u(F^{\#}\phi), v_1 \wedge v_2 \wedge \cdots \wedge v_m\rangle \\
&\quad = \lim_{t\to 0} \frac{\langle (F^{\#}\phi)(p+tu),\ v_1 \wedge \cdots \wedge v_m\rangle - \langle (F^{\#}\phi)(p),\ v_1 \wedge \cdots \wedge v_m\rangle}{t} \\
&\quad = \lim_{t\to 0} \Big[\langle \phi \circ F(p+tu),\ D_{v_1}F(p+tu) \wedge \cdots \wedge D_{v_m}F(p+tu)\rangle \\
&\qquad\qquad - \langle\ \phi \circ F(p),\ D_{v_1}F(p) \wedge \cdots \wedge D_{v_m}F(p)\ \rangle \Big] \Big/ t \\
&\quad = \langle D_w\phi[F(p)],\ D_{v_1}F \wedge D_{v_2}F \wedge \cdots \wedge D_{v_m}F\rangle \\
&\qquad + \langle \phi \circ F, D_u D_{v_1}F \wedge D_{v_2}F \wedge \cdots \wedge D_{v_m}F\rangle \\
&\qquad + \langle \phi \circ F, D_{v_1}F \wedge D_u D_{v_2}F \wedge \cdots \wedge D_{v_m}F\rangle \\
&\qquad + \cdots + \langle \phi \circ F, D_{v_1}F \wedge D_{v_2}F \wedge \cdots \wedge D_u D_{v_m}F\rangle .
\end{aligned}$$

Now fix vectors $v_1, v_2, \ldots, v_{m+1} \in \mathbb{R}^N$. Writing $w_i = D_{v_i}F$ and using (A.12), we see that

$$\begin{aligned}
&(m+1)\Big\langle d\Big(F^{\#}\phi\Big), v_1 \wedge v_2 \wedge \cdots \wedge v_{m+1}\Big\rangle \\
&= \sum_{i=1}^{m+1} (-1)^{i+1} \langle D_{v_i}(F^{\#}\phi), v_1 \wedge \cdots \wedge \widehat{v_i} \wedge \cdots \wedge v_{m+1}\rangle \\
&= \sum_{i=1}^{m+1} (-1)^{i+1} \langle D_{w_i}\phi[F(p)], D_{v_1}F \wedge \cdots \wedge \widehat{D_{v_i}F} \wedge \cdots \wedge D_{v_{m+1}}F\rangle
\end{aligned}$$

$$+\sum_{i=1}^{m+1}(-1)^{i+1}\Bigg(\sum_{j=1}^{i-1}\Big\langle\phi\circ F,$$

$$D_{v_1}F\wedge\cdots\wedge D_{v_i}D_{v_j}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_{m+1}}F\Big\rangle$$

$$+\sum_{j=i+1}^{m+1}\langle\phi\circ F, D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_i}D_{v_j}F\wedge\cdots\wedge D_{v_{m+1}}F\rangle\Bigg).$$

Again using (A.12), we have

$$\sum_{i=1}^{m+1}(-1)^{i+1}\langle D_{w_i}\phi[F(p)], D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_{m+1}}F\rangle$$

$$= (m+1)\,\langle d\phi[F(p)], D_{v_1}F\wedge\cdots\wedge D_{v_{m+1}}F\rangle$$

$$= (m+1)\Big\langle F^{\#}\Big(d\phi\Big), v_1\wedge v_2\wedge\cdots\wedge v_{m+1}\Big\rangle.$$

Also we have

$$\sum_{i=1}^{m+1}(-1)^{i+1}\Bigg(\sum_{j=1}^{i-1}\langle\phi\circ F, D_{v_1}F\wedge\cdots\wedge D_{v_i}D_{v_j}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_{m+1}}F\rangle$$

$$+\sum_{j=i+1}^{m+1}\langle\phi\circ F, D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_i}D_{v_j}F\wedge\cdots\wedge D_{v_{m+1}}F\rangle\Bigg)$$

$$= \sum_{i=1}^{m+1}\sum_{j=1}^{i-1}(-1)^{i+j}\Big\langle\phi\circ F,$$

$$D_{v_i}D_{v_j}F\wedge D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_j}F}\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_{m+1}}F\Big\rangle$$

$$+\sum_{i=1}^{m+1}\sum_{j=i+1}^{m+1}(-1)^{i+j-1}\Big\langle\phi\circ F,$$

$$D_{v_j}D_{v_i}F\wedge D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge\widehat{D_{v_j}F}\wedge\cdots\wedge D_{v_{m+1}}F\Big\rangle$$

$$= \sum_{1\le j<i\le m+1}(-1)^{i+j}\Big\langle\phi\circ F,$$

$$D_{v_i}D_{v_j}F\wedge D_{v_1}F\wedge\cdots\wedge\widehat{D_{v_j}F}\wedge\cdots\wedge\widehat{D_{v_i}F}\wedge\cdots\wedge D_{v_{m+1}}F\Big\rangle$$

$$+\sum_{1\le i<j\le m+1}(-1)^{i+j-1}\Big\langle\phi\circ F,$$

$$D_{v_j} D_{v_i} F \wedge D_{v_1} F \wedge \cdots \wedge \widehat{D_{v_i} F} \wedge \cdots \wedge \widehat{D_{v_j} F} \wedge \cdots \wedge D_{v_{m+1}} F \Big\rangle$$

$$= 0\,,$$

where the last equality follows from the fact that $D_{v_j} D_{v_i} F = D_{v_i} D_{v_j} F$, that is, from the fact that the order of differentiation can be interchanged. $\square$

References

[All 68] W. K. Allard, *On Boundary Regularity for the Plateau Problem*, Ph.D. thesis, Brown University, Providence, RI, 1968.

[All 72] W. K. Allard, On the first variation of a varifold, *Ann. Math.* (2), **95** (1972), 417–491.

[All 74] W. K. Allard, A characterization of the area integrand, in *Convegno di Teoria Geometrica dell'Integrazione e Varietà Minimali, INDAM, Rome*, 1973, Symposia Mathematica, Vol. XIV, Academic Press, London, 1974, 429–444.

[All 75] W. K. Allard, On the first variation of a varifold: Boundary behavior, *Ann. Math.* (2), **101** (1975), 418–446.

[Alm 66] F. J. Almgren, Jr., Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem, *Ann. Math.* (2), **84** (1966), 277–292.

[Alm 68] F. J. Almgren, Jr., Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure, *Ann. Math.* (2), **87** (1968), 321–391.

[Alm 00] F. J. Almgren, Jr., *Almgren's Big Regularity Paper: Q-Valued Functions Minimizing Dirichlet's Integral and the Regularity of Area-Minimizing Rectifiable Currents Up to Codimension* 2 (with a preface by J. E. Taylor and V. Scheffer), World Scientific Publishing, River Edge, NJ, 2000.

[Amb 90] L. Ambrosio, Metric space valued functions of bounded variation, *Ann. Sc. Norm. Sup. Pisa* (4), **17** (1990), 439–478.

[AK 00] L. Ambrosio and B. Kirchheim, Currents in metric spaces, *Acta Math.*, **185** (2000), 1–80.

[Bar 96] G. I. Barenblatt, *Scaling, Self-Similarity, and Intermediate Asymptotics*, Cambridge University Press, Cambridge, UK, 1996.

[Bes 52] A. S. Besicovitch, On existence of subsets of finite measure of sets of infinite measure, *Koninklijke Nederlandse Akad. Wetenschappen Indag. Math.*, **14** (1952), 339–344.

[Blo 98] Z. Błocki, Regularity of the fundamental solution for the Monge–Ampère operator, in *Progress in Partial Differential Equations, Vol.* 1 (*Pont-à-Mousson*, 1997), Pitman Research Notes in Mathematics Series, Longman, Harlow, UK, 1998, 40–45.

[BDG 69] E. Bombieri, E. de Giorgi, and E. Giusti, Minimal cones and the Bernstein problem, *Invent. Math.*, **7** (1969), 243–268.

[BT 82] R. H. Bott and L. W. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag, New York, 1982.

[CZ 52] A. P. Caldéron and A. S. Zygmund, On the existence of certain singular integrals, *Acta Math.*, **88** (1952), 85–139.

[Car 14] C. Carathéodory, Über das lineare Maß von Punktmengen: Eine Verallgemeinerung des Längenbegriffs, *Nachr. Gesellsch. Wiss. Göttingen Math.-Phys. Klasse*, **1914** (1914), 404–426.

[Cha 88] S. X.-D. Chang, Two-dimensional area minimizing integral currents are classical minimal surfaces, *J. Amer. Math. Soc.*, **1** (1988), 699–778.

[Con 90] J. B. Conway, *A Course in Functional Analysis*, 2nd ed., Springer-Verlag, New York, 1990.

[CH 62] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2, Interscience, New York, 1962.

[DGi 61a] E. de Giorgi, *Complementi alla teoria della misura* $(n-1)$*-dimensionale in uno spazio* n*-dimensionale*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960–1961, Editrice Tecnico Scientifica, Pisa, 1961.

[DGi 61b] E. de Giorgi, *Frontiere orientate di misura minima*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960–1961, Editrice Tecnico Scientifica, Pisa, 1961.

[DGi 65] E. de Giorgi, Una estensione del teorema di Bernstein, *Ann. Sc. Norm. Sup. Pisa* (3), **19** (1965), 79–85.

[DGu 75] M. de Guzmán, *Differentiation of Integrals in* $\mathbb{R}^n$, Lecture Notes in Mathematics, Vol. 481, Springer-Verlag, Berlin, New York, 1975.

[DRh 31] G. de Rham, Sur l'analyse situs des variétés à n dimensions, *J. Math. Pures Appl.* (9), **10** (1931), 115–200.

[DRh 55] G. de Rham, *Variétés différentiables: Formes, courants, formes harmoniques*, Hermann, Paris, 1955.

[Dug 66] J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.

[DS 02] F. Duzaar and K. Steffen, Optimal interior and boundary regularity for almost minimizers to elliptic variational integrals, *J. Reine Angew. Math.*, **546** (2002), 73–138.

[EH 43] S. Eilenberg and O. G. Harrold, Jr., Continua of finite linear measure I, *Amer. J. Math.*, **65** (1943), 137–146.

[Eva 86] L. C. Evans, Quasiconvexity and partial regularity in the calculus of variations, *Arch. Rational Mech. Anal.*, **95** (1986), 227–252.

[EG 92] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, FL, 1992.

[ES 92] J. H. Ewing and G. E. Schober, The area of the Mandelbrot set, *Numer. Math.*, **61** (1992), 59–72.

[Fav 32] J. Favard, Une définition de la longueur et de l'aire, *C. R. Séances Acad. Sci. Paris*, **194** (1932), 344–346.

[Fed 54] H. Federer, An analytic characterization of distributions whose partial derivatives are representable by measures, *Bull. Amer. Math. Soc.*, **60** (1954), 339.

[Fed 59] H. Federer, Curvature measures, *Trans. Amer. Math. Soc.*, **93** (1959), 418–491.

[Fed 69] H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin, 1969.

[Fed 75] H. Federer, A minimizing property of extremal submanifolds, *Arch. Rational Mech. Anal.*, **59** (1975), 207–217.

[Fed 70] H. Federer, The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension, *Bull. Amer. Math. Soc.*, **76** (1970), 767–771.

[FF 60] H. Federer and W. H. Fleming, Normal and integral currents, *Ann. Math.* (2), **72** (1960), 458–520.

[Fle 62] W. H. Fleming, On the oriented Plateau problem, *Rend. Circ. Mat. Palermo* (2), **11** (1962), 69–90.

[Fol 84] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Wiley, New York, 1984.

[GMS 98] M. Giaquinta, G. Modica, and J. Souček, *Cartesian Currents in the Calculus of Variations* I: *Cartesian Currents,* II: *Variational Integrals*, Springer-Verlag, Berlin, 1998.

[Gro 18a] W. Gross, Über das Flächenmaß von Punktmengen, *Monat. Math. Phys.*, **29** (1918), 145–176.

[Gro 18b] W. Gross, Über das lineare Maß von Punktmengen, *Monat. Math. Phys.*, **29** (1918), 177–193.

[Hal 51] P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, New York, 1951.

[Hal 74] P. R. Halmos, *Naive Set Theory*, Springer-Verlag, New York, 1974.

[Har 77] R. M. Hardt, On boundary regularity for integral currents or flat chains modulo two minimizing the integral of an elliptic integrand, *Comm. Partial Differential Equations*, **2** (1977), 1163–1232.

[HS 79] R. M. Hardt and L. M. Simon, Boundary regularity and embedded solutions for the oriented Plateau problem, *Ann. Math.* (2), **110** (1979), 439–486.

[Hau 18] F. Hausdorff, Dimension and äußeres Maß, *Math. Ann.*, **79** (1918), 157–179.

[Hau 62] F. Hausdorff, *Set Theory*, 2nd English ed., translated by J. R. Aumann et al., Chelsea, New York, 1962.

[Hil 02] D. Hilbert, Mathematical problems, translated by M. F. Winston Newson, *Bull. Amer. Math. Soc.*, **8** (1902), 437–479; reprinted in *Bull. Amer. Math. Soc.* (*N.S.*), **37** (2000), 407–436.

[Hir 76] M. W. Hirsch, *Differential Topology*, Springer-Verlag, New York, 1976.

[Hor 69] L. Hörmander, *Linear Partial Differential Operators*, Springer-Verlag, New York, 1969.

[HW 41] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton University Press, Princeton, NJ, 1941.

[Jec 78] T. J. Jech, *Set Theory*, Academic Press, New York, 1978.

[Kel 55] J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955.

[Kli 91] M. Klimek, *Pluripotential Theory*, London Mathematical Society Monographs, New Series, Vol. 6, Oxford University Press, New York, 1991.

[Kra 83] S. G. Krantz, Lipschitz spaces, smoothness of functions, and approximation theory, *Expos. Math.*, **1** (1983), 193–260.

[Kra 92] S. G. Krantz, *Partial Differential Equations and Complex Analysis*, CRC Press, Boca Raton, FL, 1992.

[Kra 99] S. G. Krantz, *A Panorama of Harmonic Analysis*, Carus Mathematical Monographs, Vol. 27, Mathematical Association of America, Washington, DC, 1999.

[Kra 05] S. G. Krantz, Calculation and estimation of the Poisson kernel, *J. Math. Anal. Appl.*, **302** (2005), 143–148.

[KPk 99] S. G. Krantz and H. R. Parks, *The Geometry of Domains in Space*, Birkhäuser Boston, Cambridge, MA, 1999.

[KPk 02] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, 2nd ed., Birkhäuser Boston, Cambridge, MA, 2002.

[KPn 86] S. G. Krantz and T. D. Parsons, Antisocial subcovers of self-centered coverings, *Amer. Math. Monthly*, **93** (1986), 45–48.

[LSW 01] G. F. Lawler, O. Schramm, and W. Werner, Values of Brownian intersection exponents I: Half-plane exponents, II: Plane exponents, *Acta Math.*, **187** (2001), 237–273, 275–308.

[LSW 02] G. F. Lawler, O. Schramm, and W. Werner, Analyticity of intersection exponents for planar Brownian motion, *Acta Math.*, **189** (2002), 179–201.

[Leb 05] H. L. Lebesgue, Sur les fonctions représentables analytiquement, *J. Math. Pures Appl.* (6), **1** (1905), 139–216.

[Lel 69] P. Lelong, *Plurisubharmonic Functions and Positive Differential Forms*, translated by M. A. Dostál, Gordon and Breach, New York, 1969.

[LG 86] P. Lelong and L. Gruman, *Entire Functions of Several Complex Variables*, Springer-Verlag, Berlin, 1986.

[LY 02] F. Lin and X. Yang, *Geometric Measure Theory: An Introduction*, International Press, Boston, 2002.

[Loe 93] P. A. Loeb, An optimization of the Besicovitch covering, *Proc. Amer. Math. Soc.*, **118** (1993), 715–716.

[Mat 95] P. Mattila, *Geometry of Sets and Measures in Euclidean Space*, Cambridge University Press, Cambridge, UK, 1995.

[Mor 66] C. B. Morrey, Jr., *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, New York, 1966.

[MR 40] A. P. Morse and J. A. F. Randolph, Gillespie measure, *Duke Math. J.*, **6** (1940), 408–419.

[Osg 03] W. F. Osgood, A Jordan curve of positive area, *Trans. Amer. Math. Soc.*, **4** (1903), 107–112.

[PS 92] H. R. Parks and R. M. Schori, A Jordan arc in $\mathbb{R}^m$ with positive m-dimensional Lebesgue measure, in *The Problem of Plateau: A Tribute to Jesse Douglas and Tibor Rado*, T. M. Rassias, ed., World Scientific, River Edge, NJ, 1992, 229–236.

[Por 96] G. J. Porter, k-volume in $\mathbb{R}^n$ and the generalized Pythagorean theorem, *Amer. Math. Monthly*, **103** (1996), 252–256.

[Rei 64a] E. R. Reifenberg, An epiperimetric inequality related to the analyticity of minimal surfaces, *Ann. Math.* (2), **80** (1964), 1–14.

[Rei 64b] E. R. Reifenberg, On the analyticity of minimal surfaces, *Ann. Math.* (2), **80** (1964), 15–21.

[Rog 98] C. A. Rogers, *Hausdorff Measures*, Cambridge University Press, Cambridge, UK, 1998.

[Roy 88] H. L. Royden, Jr., *Real Analysis*, 3rd ed., Macmillan, New York, 1988.

[Rud 76] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw–Hill, New York, 1976.

[Rud 87] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw–Hill, New York, 1987.

[SS 82] R. M. Schoen and L. M. Simon, A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals, *Indiana Univ. Math. J.*, **31** (1982), 415–434.

[SSA 77] R. M. Schoen, L. M. Simon, and F. J. Almgren, Jr., Regularity and singularity estimates on hypersurfaces minimizing parametric elliptic variational integrals I, II, *Acta Math.*, **139** (1977), 217–265.

[Sch 51] L. Schwartz, *Théorie des distributions*, Tome 1, 2, Hermann, Paris, 1950, 1951.

[Shi 98] M. Shishikura, The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets, *Ann. Math.* (2), **147** (1998), 225–267.

[Sim 83] L. M. Simon, *Lectures on Geometric Measure Theory*, Proceedings of the Centre for Mathematical Analysis, Vol. 3, Australian National University, Canberra, ACT, Australia, 1983.

[Sis 68] J. H. Simons, Minimal varietes in riemannian manifolds, *Ann. Math.* (2), **88** (1968), 62–105.

[Som 84] B. M. Solomon, A new proof of the closure theorem for integral currents, *Indiana Univ. Math. J.*, **33** (1984), 393–418.

[Sov 70] R. M. Solovay, A model of set-theory in which every set of reals is Lebesgue measurable, *Ann. Math.* (2), **92** (1970), 1–56.

[Spa 66] E. H. Spanier, *Algebraic Topology*, McGraw–Hill, New York, 1966.

[Ste 93] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton, NJ, 1993.

[SW 71] E. M. Stein and G. L. Weiss, *Introduction to Fourier Analysis on Euclidean Space*, Princeton University Press, Princeton, NJ, 1971.

[Str 36] J. Steiner, Einfache Beweise der isoperimetrischen Hauptsätze, *J. Reine Angew. Math.*, **18** (1838), 281–296; reprinted in *Jacob Steiner's Gesammelte Werke*, Vol. 2, K. Weierstrass, ed., G. Reimer, Berlin, 1882, 75–91.

[Sul 94] J. M. Sullivan, Sphere packings give an explicit bound for the Besicovitch covering theorem, *J. Geom. Anal.*, **4** (1994), 219–231.

[Sus 17] M. Y. Suslin, Sur une définition des ensembles mesurables B sans nombres transfinis, *C. R. Séances Acad. Sci. Paris*, **164** (1917), 88–91.

[Tre 80] F. Trèves, *Introduction to Pseudodifferential and Fourier Integral Operators*, Plenum Press, New York, 1980.

[Whe 89] B. C. White, A new proof of the compactness theorem for integral currents, *Comm. Math. Helvetici*, **64** (1989), 207–220.

[Whe 97] B. C. White, Stratification of minimal surfaces, mean curvature flows, and harmonic maps, *J. Reine Angew. Math.*, **488** (1997), 1–35.

[Whe 98] B. C. White, A new proof of Federer's structure theorem for k-dimensional subsets of $\mathbb{R}^N$, *J. Amer. Math. Soc.*, **11** (1998), 693–701.

[Whe 99] B. C. White, Rectifiability of flat chains, *Ann. Math.* (2), **150** (1999), 165–184.

[Whe 05] B. C. White, A local regularity theorem for mean curvature flow, *Ann. Math.* (2), **161** (2005), 1487–1519.

[Whn 57] H. Whitney, *Geometric Integration Theory*, Princeton University Press, Princeton, NJ, 1957.

[You 37] L. C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *C. R. Soc. Sci. Lett. Varsovie* III, **30** (1937), 212–234.

[Zie 89] W. P. Ziemer, *Weakly Differentiable Functions*, Springer-Verlag, New York, 1989.

Index of Notation

Index