2.6. To define the usual Cantor set $C \subset \mathbf{R}^1$, let $C_1 = [0,1]$; construct C_{j+1} by removing the open middle third of each interval of C_j and put

$$C = \bigcap\{C_j : j \in \mathbf{Z}^+\}.$$

Let $m = \dfrac{\ln 2}{\ln 3}$.

(a) Prove that $\mathscr{H}^m(C) \le \dfrac{\alpha_m}{2^m}$, and hence, $\dim C \le m$.

(b) Try to prove that $\mathscr{H}^m(C) = \dfrac{\alpha_m}{2^m}$ or at least that $\mathscr{H}^m(C) > 0$ and hence that the Hausdorff dimension of C is m.

2.7. Give a function $f: \mathbf{R}^2 \to \mathbf{R}$ which is approximately continuous at $\mathbf{0}$, but for which $\mathbf{0}$ is not a Lebesgue point.

2.8. Prove that if $f: \mathbf{R}^m \to \mathbf{R}$ has $\mathbf{0}$ as a Lebesgue point, then f is approximately continuous at $\mathbf{0}$.

2.9. Deduce Corollary 2.10 from Corollary 2.9.
Hint: Let $\{q_i\}$ be a countable dense subset of $\mathbf{R}$, $A_i = \{x : f(x) > q_i\}$, $E_i = \{x : \Theta(A_i, x) = \chi_{A_i}\}$, and show that f is approximately continuous at each point in $\bigcap E_i$.

"Frank Morgan "Geometric Measure Theory: A Beginner's Guide"

CHAPTER 3

Lipschitz Functions and Rectifiable Sets

This chapter introduces the m-dimensional surfaces of geometric measure theory, called rectifiable sets. These sets have folds, corners, and more general singularities. The relevant functions are not smooth functions as in differential geometry, but *Lipschitz* functions.

3.1. Lipschitz Functions. A function $f: \mathbf{R}^m \to \mathbf{R}^n$ is *Lipschitz* if there is a constant C such that

$$|f(x) - f(y)| \le C|x - y|.$$

The least such constant is called the *Lipschitz constant* and denoted by Lip f.

Figure 3.1.1 gives the graphs of two typical Lipschitz functions.

Theorems 3.2 and 3.3 show that a Lipschitz function comes very close to being differentiable.

3.2. Rademacher's Theorem [GMT 3.1.6]. *A Lipschitz function $f: \mathbf{R}^m \to \mathbf{R}^n$ is differentiable almost everywhere.*

The *Proof* has five steps:

(1) A monotonic function $f: \mathbf{R} \to \mathbf{R}$ is differentiable almost everywhere.
(2) Every function $f: \mathbf{R} \to \mathbf{R}$ which is locally of bounded variation (and hence every Lipschitz function) is differentiable almost everywhere.

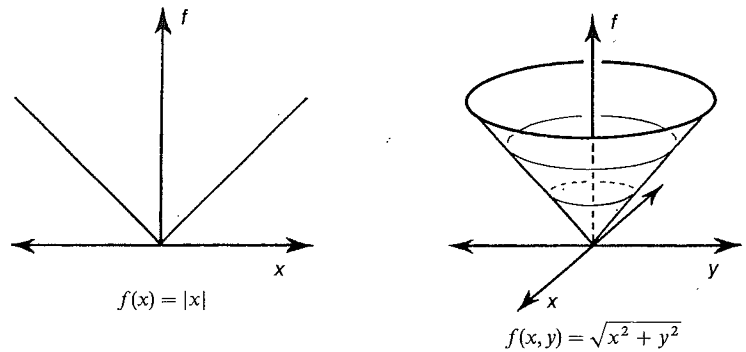

Figure 3.1.1. Examples of Lipschitz functions.

(3) A Lipschitz function $f:\mathbf{R}^m \to \mathbf{R}^n$ has partial derivatives almost everywhere.
(4) A Lipschitz function $f:\mathbf{R}^m \to \mathbf{R}^n$ is approximately differentiable almost everywhere.
(5) A Lipschitz function $f:\mathbf{R}^m \to \mathbf{R}^n$ is differentiable almost everywhere.

Step (1) is a standard result of real analysis, proved by differentiation of measures. Step (2) follows by decomposing a function of bounded variation as the difference of two monotonic functions. Step (3) follows (Fubini) immediately from Step (2). The deduction of (4) from (3) relies on a technical measure-theoretic argument which I do not find sufficiently edifying to include here. If (3) holds everywhere, it does not follow that (4) holds everywhere.

The final conclusion (5) rests on the interesting fact that *if a Lipschitz function is approximately differentiable at a, it is differentiable at a.* We conclude this discussion with a proof of that fact.

Suppose that the Lipschitz function $f:\mathbf{R}^m \to \mathbf{R}^n$ is approximately differentiable at a but not differentiable at a. We may assume $a = \mathbf{0}$ and $\operatorname{ap} Df(\mathbf{0}) = \mathbf{0}$.

For some $0 < \varepsilon < 1$, there is a sequence of points $a_i \to \mathbf{0}$ such that

$$|f(a_i)| \geq \varepsilon |a_i|.$$

Let $C = \max\{\operatorname{Lip} f, 1\}$. Then for x in the closed ball $\mathbf{B}(a_i, \varepsilon|a_i|/3C)$,

$$|f(x)| \geq \varepsilon|a_i| - \varepsilon|a_i|/3 \geq \varepsilon|x|/3.$$

Thus for $x \in E = \bigcup_{i=1}^{\infty} \mathbf{B}(a_i, \varepsilon|a_i|/3C)$,

$$|f(x)| \geq \varepsilon|x|/3.$$

But E does not have density 0 at $\mathbf{0}$, because

$$\frac{\mathscr{L}^n \mathbf{B}(a_i, \varepsilon|a_i|/3C)}{\alpha_n(|a_i| + \varepsilon|a_i|/3C)^n} \geq \frac{(\varepsilon|a_i|/3C)^n}{(2|a_i|)^n} = \frac{\varepsilon^n}{6^n C^n} > 0.$$

This contradiction of the approximate differentiability of f at $\mathbf{0}$ completes the proof.

3.3. Approximation of a Lipschitz Function by a $\mathbf{C^1}$ Function [GMT 3.1.15]. *Suppose $A \subset \mathbf{R}^m$, and $f: A \to \mathbf{R}^n$ is Lipschitz. Given $\varepsilon > 0$ there is a C^1 function $g:\mathbf{R}^m \to \mathbf{R}^n$ such that $\mathscr{L}^m\{x \in A: f(x) \neq g(x)\} \leq \varepsilon$.*

Note that the approximation is in the strongest sense: the functions *coincide* except on a set of measure ε. The proof of 3.3 depends on Whitney's Extension Theorem, which gives the coherence conditions on prescribed values for a desired C^1 function.

3.4. Lemma (Whitney's Extention Theorem) [GMT 3.1.14]. *Let A be a closed set of points a in $\mathbf{R}^m$ at which the values and derivatives of a desired C^1 function are prescribed by linear polynomials $P_a:\mathbf{R}^m \to \mathbf{R}$. For each compact subset C of A and $\delta > 0$, let $\rho(C, \delta)$ be the supremum of the numbers $|P_a(b) - P_b(b)|/|a - b|$, $\|DP_a(b) - DP_b(b)\|$, over all $a, b \in C$ with $0 < |a - b| \leq \delta$. If the prescribed data satisfy the coherence condition that $\lim_{\delta \to 0} \rho(C, \delta) = 0$ for each compact subset C of A, then there exists a C^1 function g satisfying*

$$g(a) = P_a(a), \; Dg(a) = DP_a(a)$$

for all $a \in A$.

Remarks. A more general version of Whitney's Extension Theorem gives the analogous conditions to obtain a C^k function with values and derivatives prescribed by polynomials P_a of degree k. In the proof, the value $g(x)$ assigned at a point $x \notin A$ is a smoothly weighted average of the values prescribed at nearby points of A. The averaging uses a partition of unity subordinate to a cover of A^C which becomes finer and finer as one approaches A.

***Sketch of proof of* 3.3.** First extend f to a Lipschitz function on all of $\mathbf{R}^m$ (see [GMT 2.10.43]). Second, by Rademacher's Theorem 3.2, f is differentiable almost everywhere. Third, by Lusin's Theorem [GMT 2.3.5], there is a closed subset E of A such that Df is continuous on E and $\mathscr{L}^m(A - E) < \varepsilon$. Fourth, for any $a \in E$, $\delta > 0$, define

$$\eta_\delta(a) = \sup_{\substack{0<|x-a|<\delta \\ x \in C}} \frac{|f(x) - f(a) - Df(a)(x-a)|}{|x-a|}$$

Since as $\delta \to 0$, $\eta_\delta \to 0$ pointwise, then by Egoroff's Theorem [GMT 2.3.7] there is a closed subset F of E such that $\mathscr{L}^m(A - F) < \varepsilon$ and $\eta_\delta \to 0$ uniformly on compact subsets of F. This condition implies the hypotheses of Whitney's Extension Theorem 3.4, with $P_a(x) = f(a) + Df(a)(x)$. Consequently there is a C^1 function $g : \mathbf{R}^m \to \mathbf{R}^n$ which coincides with f on F.

The following theorem implies for example that Lipschitz images of sets of Hausdorff measure 0 have masure 0.

3.5. Proposition [GMT 2.10.11]. *Suppose $f : \mathbf{R}^l \to \mathbf{R}^n$ is Lipschitz and A is a Borel subset of $\mathbf{R}^l$. Then*

$$\int N(f \mid A, y)\, d\mathscr{H}^m y \le (\operatorname{Lip} f)^m \mathscr{H}^m(A).$$

Here $N(f \mid A, y) \equiv \operatorname{card}\{x \in A : f(x) = y\}$.

Proof. Any covering of A by sets S_i of diameter d_i yields a covering of $f(A)$ by the sets $f(S_i)$, of diameter at most $(\operatorname{Lip} f) d_i$. Since the approximating sum $\Sigma\, \alpha_m(\text{diam}/2)^m$ for Hausdorff measure contains $(\text{diam})^m$,

$$\mathscr{H}^m(f(A)) \le (\operatorname{Lip} f)^m \mathscr{H}^m(A).$$

Notice that this formula gives the proposition in the case that f is injective. In the general case, chop A up into little pieces A_i and add up the formulas for each piece to obtain

$$\int_{f(A)} (\text{the number of } A_i \text{ intersecting } f^{-1}\{y\})\, d\mathscr{H}^m y \le (\operatorname{Lip} f)^m \mathscr{H}^m(A).$$

As the pieces subdivide, the integrand increases monotonically to the multiplicity function $N(f \mid A, y)$, and the proposition is proved.

The beginning of this proof illustrates the virtue of allowing coverings by arbitrary sets rather than just balls in the definition of Hausdorff measure. If $\{S_i\}$ covers A, then $\{f(S_i)\}$ is an admissible covering of $f(A)$.

3.6. Jacobians. Jacobians are the corrective factors relating the elements of areas of the domains and images of functions. If $f : \mathbf{R}^m \to \mathbf{R}^n$ is differentiable at a, we define the *k-dimensional Jacobian of f at a, $J_k f(a)$,* as the maximum k-dimensional volume of the image under $Df(a)$ of a unit k-dimensional cube.

If rank $Df(a) < k$, $J_k f(a) = 0$. If rank $Df(a) \le k$, as holds in most applications, then $J_k f(a)^2$ equals the sum of the squares of the determinants of the $k \times k$ submatrices of $Df(a)$. If $k = m$ or n, then $J_k f(a)^2$ equals the determinant of the $k \times k$ product of $Df(a)$ with its transpose. If $k = m = n$, then $J_k f(a)$ is just the absolute value of the determinant of $Df(a)$. In general, computations are sometimes simplified by viewing $Df(a)$ as a map from the orthogonal complement of its kernel onto its image.

If $L : \mathbf{R}^m \to \mathbf{R}^m$ is linear, then $\mathscr{L}^m(L(A)) = J_m L \cdot \mathscr{L}^m(A)$.

3.7. The Area Formula [GMT 3.2.3]. *Consider a Lipschitz function $f : \mathbf{R}^m \to \mathbf{R}^n$ for $m \le n$.*

(1) *If A is an $\mathscr{L}^m$ measurable set, then*

$$\int_A J_m f(x)\, d\mathscr{L}^m x = \int_{\mathbf{R}^n} N(f \mid A, y)\, d\mathscr{H}^m y.$$

(2) *If u is an $\mathscr{L}^m$ integrable function, then*

$$\int_{\mathbf{R}^m} u(x) J_m f(x)\, d\mathscr{L}^m x = \int_{\mathbf{R}^n} \sum_{x \in f^{-1}\{y\}} u(x)\, d\mathscr{H}^m y.$$

Remark. If f is a smooth embedding, then (1) equates the classical area of the parameterized surface $f(A)$ with the Hausdorff measure of $f(A)$. Therefore for all smooth surfaces, the Hausdorff measure coincides with the classical area.

***Sketch of the Proof of the Area Formula* 3.7(1).** We will split A up into two cases, according to the rank of Df. In either case, by Rademacher's Theorem 3.2 and 3.5, we may assume that f is differentiable.

CASE 1. *Df has rank m.* Let $\{s_i\}$ be a countable dense set of affine maps of $\mathbf{R}^m$ onto m-dimensional planes in $\mathbf{R}^n$. Let E_i be a piece of A such that for each $a \in E_i$ the affine functions $f(a) + Df(a)(x)$ and $s_i(x)$ are approximately equal. It follows that

(1) $\det s_i \approx J_m f$ on E_i,
(2) f is injective on E_i, and
(3) the associated map from $s_i(E_i)$ to $f(E_i)$ and its inverse both have Lipschitz constant ≈ 1.

Since f is differentiable, the E_i cover A. Refine $\{E_i\}$ into a countable disjoint covering of A by tiny pieces. On each piece E, by (3) and 3.5,

$$\begin{aligned}\mathscr{H}^m(f(E)) &\approx \mathscr{H}^m(s_i(E))\\ &= \mathscr{L}^m(s_i(E))\\ &= \int_E \det s_i \, d\mathscr{L}^m\\ &\approx \int_E J_m f \, d\mathscr{L}^m.\end{aligned}$$

Summing over all the sets E yields

$$\int (\text{number of sets } E \text{ intersecting } f^{-1}\{y\}) \, d\mathscr{H}^m y \approx \int_A J_m f \, d\mathscr{L}^m.$$

Taking a limit yields

$$\int N(f \mid A, y) \, d\mathscr{H}^m y = \int_A J_m f \, d\mathscr{L}^m,$$

and completes the proof of case 1.

We remark that it does not suffice in the proof just to cut A up into tiny pieces without using the s_i. Without the requirement that for $a, b \in E$, $Df(a) \approx Df(b)$, f need not even be injective on E, no matter how small E is.

CASE 2. *Df has rank $< m$.* In this case the left-hand side $\int_A J_m f$ is zero. Define a function

$$g : \mathbf{R}^m \to \mathbf{R}^{n+m}$$
$$x \to (f(x), \varepsilon x).$$

Then $J_m(g) \le \varepsilon(\operatorname{Lip} f + \varepsilon)^{m-1}$. Now by Case 1,

g(A)

$$\begin{aligned}\mathscr{H}^m(f(A)) &\le \mathscr{H}^m(g(A))\\ &= \int_A J_m g\\ &\le \varepsilon(\operatorname{Lip} f + \varepsilon)^{m-1}\mathscr{L}^m(A).\end{aligned}$$

Therefore the right-hand side also must vanish.

Finally we remark that 3.7(2) follows from 3.7(1) by approximating u by simple functions.

The following useful formula relates integrals of a function f over a set A to the areas of the level sets $A \cap f^{-1}\{y\}$ of the function.

3.8. The Coarea Formula [GMT 3.2.11]. *Consider a Lipschitz function $f : \mathbf{R}^m \to \mathbf{R}^n$ with $m > n$. If A is a $\mathscr{L}^m$ measurable set, then*

$$\int_A J_n f(x) \, d\mathscr{L}^m x = \int_{\mathbf{R}^n} \mathscr{H}^{m-n}(A \cap f^{-1}\{y\}) \, d\mathscr{L}^n y.$$

Proof.
CASE 1. *f is orthogonal projection.* If f is orthogonal projection, then $J_n f = 1$, and the coarea formula is reduced to Fubini's Theorem.

GENERAL CASE. We treat just the main case $J_n f \neq 0$. By subdividing A as in the proof of the area formula, we may assume that f is linear. Then $f = L \circ P$, where P denotes projection onto the n-dimensional orthogonal complement V of the kernel of f and L is a nonsingular linear map from V to $\mathbf{R}^n$. Now

$$\begin{aligned}\int_A J_n f \, d\mathscr{L}^m &= |\det L| \mathscr{H}^m(A)\\ &= |\det L| \int_{P(A)} \mathscr{H}^{m-n}(P^{-1}\{y\}) \, d\mathscr{L}^n y\\ &= \int_{L \circ P(A)} \mathscr{H}^{m-n}((L \circ P)^{-1}\{y\}) \, d\mathscr{L}^n y\end{aligned}$$

as desired.