5. Next choose any set E_k, which is of the form $E(c,T,i)$ for some $c \in C$, $T \in \mathbf{S}$, $i = 1,2,\ldots$. Let $T_k = T$. According to $(\star\star\star)$,

$$\frac{1}{t}|T_k(a-b)| \le |f(a)-f(b)| \le t|T_k(a-b)|$$

for all $b \in E_k$, $a \in B(b,2/i)$. As $E_k \subset B(c,1/i) \subset B(b,2/i)$, we thus have

$$\frac{1}{t}|T_k(a-b)| \le |f(a)-f(b)| \le t|T_k(a-b)| \qquad (\star\star\star\star)$$

for all $a,\ b \in E_k$; hence $f|_{E_k}$ is one-to-one.

6. Finally, notice $(\star\star\star\star)$ implies

$$\mathrm{Lip}\,((f|_{E_k}) \circ T_k^{-1}) \le t, \qquad \mathrm{Lip}\,(T_k \circ (f|_{E_k})^{-1}) \le t,$$

whereas the claim provides the estimate

$$t^{-n}|\det T_k| \le Jf\,|_{E_k} \le t^n|\det T_k|.$$

Assertion (iii) is proved. ▌

3.3.2 Proof of the Area Formula

THEOREM 1 AREA FORMULA

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \le m$. Then for each $\mathcal{L}^n$-measurable subset $A \subset \mathbb{R}^n$,

$$\int_A Jf\,dx = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\,d\mathcal{H}^n(y).$$

PROOF

1. In view of Rademacher's Theorem, we may as well assume $Df(x)$ and $Jf(x)$ exist for all $x \in A$. We may also suppose $\mathcal{L}^n(A) < \infty$.

2. Case 1. $A \subset \{Jf > 0\}$. Fix $t > 1$ and choose Borel sets $\{E_k\}_{k=1}^\infty$ as in Lemma 3. We may assume the sets $\{E_k\}_{k=1}^\infty$ are disjoint. Define $\mathcal{B}_k$ as in the proof of Lemma 2. Set

$$F_j^i = E_j \cap Q_i \cap A \qquad (Q_i \in \mathcal{B}_k, j = 1,2,\ldots).$$

Then the sets F_j^i are disjoint and $A = \cup_{i,j=1}^\infty F_j^i$.

3. Claim #1:

$$\lim_{k\to\infty} \sum_{i,j=1}^\infty \mathcal{H}^n(f(F_j^i)) = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\,d\mathcal{H}^n.$$

Proof of Claim #1: Let

$$g_k \equiv \sum_{i,j=1}^\infty \chi_{f(F_j^i)}$$

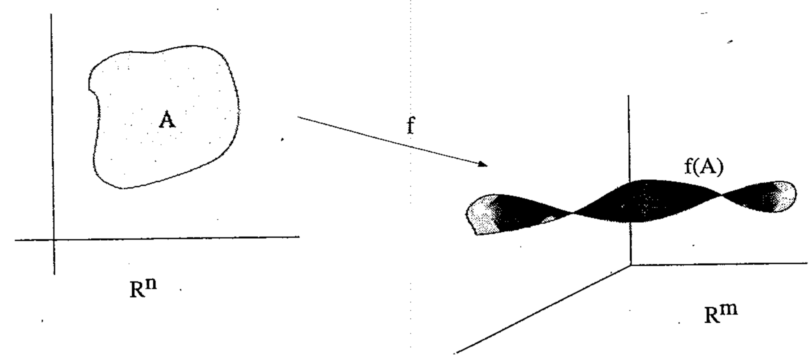

FIGURE 3.2
The Area Formula.

so that $g_k(y)$ is the number of the sets $\{F_j^i\}$ such that $F_j^i \cap f^{-1}\{y\} \ne \emptyset$. Then $g_k(y) \uparrow \mathcal{H}^0(A \cap f^{-1}\{y\})$ as $k \to \infty$. Apply the Monotone Convergence Theorem.

4. Note

$$\mathcal{H}^n(f(F_j^i)) = \mathcal{H}^n(f|_{E_j} \circ T_j^{-1} \circ T_j(F_j^i)) \le t^n\mathcal{L}^n(T_j(F_j^i))$$

and

$$\mathcal{L}^n(T_j(F_j^i)) = \mathcal{H}^n(T_j \circ (f|_{E_j})^{-1} \circ f(F_j^i)) \le t^n\mathcal{H}^n(f(F_j^i))$$

by Lemma 3. Thus

$$\begin{aligned} t^{-2n}\mathcal{H}^n(f(F_j^i)) &\le t^{-n}\mathcal{L}^n(T_j(F_j^i)) \\ &= t^{-n}|\det T_j|\mathcal{L}^n(F_j^i) \\ &\le \int_{F_i^j} Jf\,dx \\ &\le t^n|\det T_j|\mathcal{L}^n(F_j^i) \\ &= t^n\mathcal{L}^n(T_j(F_j^i)) \\ &\le t^{2n}\mathcal{H}^n(f(F_j^i)), \end{aligned}$$

where we repeatedly used Lemmas 1 and 3. Now sum on i and j:

$$t^{-2n}\sum_{i,j=1}^\infty \mathcal{H}^n(f(F_j^i)) \le \int_A Jf\,dx \le t^{2n}\sum_{i,j=1}^\infty \mathcal{H}^n(f(F_j^i)).$$

Now let $k \to \infty$ and recall Claim #1:

$$t^{-2n} \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n \le \int_A Jf\, dx \le t^{2n} \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n.$$

Finally, send $t \to 1^+$.

5. *Case 2.* $A \subset \{Jf = 0\}$. Fix $0 < \epsilon \le 1$. We factor $f = p \circ g$, where

$$g : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n, \qquad g(x) \equiv (f(x), \epsilon x) \text{ for } x \in \mathbb{R}^n,$$

and

$$p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m, \qquad p(y,z) = y \text{ for } y \in \mathbb{R}^m,\ z \in \mathbb{R}^n.$$

6. *Claim #2*: There exists a constant C such that

$$0 < Jg(x) \le C\epsilon$$

for $x \in A$.

Proof of Claim #2: Write $g = (f^1, \dots, f^m, \epsilon x_1, \dots, \epsilon x_n)$; then

$$Dg(x) = \begin{pmatrix} Df(x) \\ \epsilon I \end{pmatrix}_{(n+m)\times n}.$$

Since $Jf(x)^2$ equals the sum of the squares of the $(n \times n)$-subdeterminants of $Df(x)$ according to the Binet–Cauchy formula, we see

$$Jg(x)^2 = \text{ sum of squares of } (n \times n)\text{-subdeterminants of } Dg(x) \ge \epsilon^{2n} > 0.$$

Furthermore, since $|Df| \le \operatorname{Lip}(f) < \infty$, we may employ the Binet–Cauchy Formula to compute

$$Jg(x)^2 = Jf(x)^2 + \left\{\begin{matrix}\text{sum of squares of terms each} \\ \text{involving at least one } \epsilon\end{matrix}\right\} \le C\epsilon^2$$

for each $x \in A$.

7. Since $p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ is a projection, we can compute, using Case 1 above,

$$\begin{aligned}\mathcal{H}^n(f(A)) &\le \mathcal{H}^n(g(A)) \\ &\le \int_{\mathbb{R}^{n+m}} \mathcal{H}^0(A \cap g^{-1}\{y,z\})\, d\mathcal{H}^n(y,z) \\ &= \int_A Jg(x)\, dx \\ &\le \epsilon C \mathcal{L}^n(A).\end{aligned}$$

Let $\epsilon \to 0$ to conclude $\mathcal{H}^n(f(A)) = 0$, and thus

$$\int_{\mathbb{R}^n} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n = 0,$$

since spt $\mathcal{H}^0(A \cap f^{-1}\{y\}) \subset f(A)$. But then

$$\int_{\mathbb{R}^n} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n = 0 = \int_A Jf\, dx.$$

8. In the general case, write $A = A_1 \cup A_2$ with $A_1 \subset \{Jf > 0\}$, $A_2 \subset \{Jf = 0\}$, and apply Cases 1 and 2 above. ∎

3.3.3 Change of variables formula

THEOREM 2
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \le m$. Then for each $\mathcal{L}^n$-summable function $g : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathbb{R}^n} g(x) Jf(x)\, dx = \int_{\mathbb{R}^m} \left[\sum_{x \in f^{-1}\{y\}} g(x) \right] d\mathcal{H}^n(y).$$

REMARK Using the Area Formula, we see $f^{-1}\{y\}$ is at most countable for $\mathcal{H}^n$ a.e. $y \in \mathbb{R}^m$. ∎

PROOF

1. *Case 1.* $g \ge 0$. According to Theorem 7 in Section 1.1.2 we can write

$$g = \sum_{i=1}^{\infty} \frac{1}{i} \chi_{A_i}$$

for appropriate $\mathcal{L}^n$-measurable sets $\{A_i\}_{i=1}^{\infty}$. Then the Monotone Convergence Theorem implies

$$\begin{aligned}\int_{\mathbb{R}^n} gJf\, dx &= \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^n} \chi_{A_i} Jf\, dx \\ &= \sum_{i=1}^{\infty} \frac{1}{i} \int_{A_i} Jf\, dx \\ &= \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^m} \mathcal{H}^0(A_i \cap f^{-1}\{y\})\, d\mathcal{H}^n(y)\end{aligned}$$