Mat 1501 lecture notes, October 7-17, 2013

Area and coarea

Throughout the following discussion we assume that n, N are positive integers and that $n \le N$. We will generally be interested in the case $n < N$.

Our goal is to obtain

- a formula of the form
$$\mathcal{H}^n(f(A)) = \int_A (\cdots)\, dx$$
for $f : \mathbb{R}^n \to \mathbb{R}^N$ Lipschitz and injective and $A \subset \mathbb{R}^n$ Lebesgue measurable.
- A dual formula for Lipschitz maps $f : \mathbb{R}^N \to \mathbb{R}^n$.

These will be called the "area formula" and "coarea formula" respectively. We will also prove more general results, where for example we drop the asssumption of injectivity (in the "area" case)

In fact, the integrand on the right-hand side of the area formula will be essentially the "volume element" or "area element" for an n-dimensional manifold embedded in $\mathbb{R}^N$, which may be familiar from some calculus class. So the significance of the area formula is

- two ways that we have of measuring the size of smooth n-dimensional submanifolds – using Hausdorff measure or using calculus – are completely consistent with each other; and
- this continues to be true when we consider "submanifolds" that are parametrized by Lipschitz maps rather than smooth maps.

1. linear maps

1.1. some linear algebra.

Definition 1. A linear map $O : \mathbb{R}^n \to \mathbb{R}^N$ is *orthogonal* if
$$Ox \cdot O\tilde{x} = x \cdot \tilde{x} \qquad \text{for all } x, \tilde{x} \in \mathbb{R}^n.$$

We recall that if $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear (for any n, m) then $L^* : \mathbb{R}^m \to \mathbb{R}^n$ is the linear map defined by requiring that
$$Lx \cdot y = x \cdot L^* y \qquad \text{for all } x \in \mathbb{R}^n, y \in \mathbb{R}^m.$$
If we write L as an $m \times n$ matrix with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, then the matrix corresponsing to L^* is the transpose of L.

Lemma 1. *Let $O : \mathbb{R}^n \to \mathbb{R}^N$ be a linear map. The following are equivalent.*

(1) *O is orthogonal.*
(2) *$O^* \circ O = Id_n$ (that is, the $n \times n$ identity matrix).*
(3) *O is an isometric embedding of $\mathbb{R}^n$ into $\mathbb{R}^N$, ie*
$$|Ox - O\tilde{x}|_{\mathbb{R}^N} = |x - \tilde{x}|_{\mathbb{R}^n} \qquad \textit{for all } x, \tilde{x} \in \mathbb{R}^n.$$
(4) *if we write O as a $N \times n$ matrix, then the columns of O form an orthonormal set.*
(5) *If $v_1, \ldots, v_n$ is any orthonornal basis for $\mathbb{R}^n$, then $\{Ov_i\}_{i=1}^n$ is an orthonormal subset of $\mathbb{R}^N$.*

(6) *O can be written in the form*

$$O = Q \circ i_{\mathbb{R}^n \to \mathbb{R}^N}$$

where ${}_{\mathbb{R}^n \to \mathbb{R}^N} i : \mathbb{R}^n \to \mathbb{R}^N$ is defined by $i_{\mathbb{R}^n \to \mathbb{R}^N} x = (x, 0, \ldots, 0)$ (with $N - n$ zeros) and $Q : \mathbb{R}^N \to \mathbb{R}^N$ is orthogonal, i.e. $Q^* \circ Q = Id_N$

lem.orth

Exercise 1. Prove some or all of the above Lemma. In the final conclusion, note that Q is not unique as long as $N > n$; that is, there is more than one possible choice of Q such that $O = Q \circ i_{\mathbb{R}^n \to \mathbb{R}^N}$.

Lemma 2. *Let $L : \mathbb{R}^n \to \mathbb{R}^N$ be a linear map.*

Then there exists $S : \mathbb{R}^n \to \mathbb{R}^n$, symmetric and nonnegative definite, and $O : \mathbb{R}^n \to \mathbb{R}^N$ orthogonal, such that

$$L = O \circ S.$$

polard

The lemma implies that $O = L \circ S^{-1}$ if S is invertible. If S is not invertible, as the proof will show, then O is in general not unique.

Proof. **Step 1**. Let $A = L^* \circ L$, Then $Ax \cdot x = |Lx|^2$ for all $x \in \mathbb{R}^n$, and x is clearly symmetric, so there exist nonnegative numbers $\lambda_1, \ldots \lambda_n$ and vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ such that

$$Av_i = \lambda_i v_i, \quad i = 1, \ldots, n, \qquad \text{and} \qquad v_i \cdot v_j = \delta_{ij}.$$

Thus $\{v_i\}_{i=1}^n$ form an orthonormal basis for $\mathbb{R}^n$.

To define any linear map, it therefore suffices to specify how is acts on $v_1, \ldots, v_n$.

We define

$$Sv_i = \sqrt{\lambda_i} v_i$$

and

$$Ov_i = \begin{cases} \lambda_i^{-1/2} L v_i & \text{if } \lambda_i \neq 0 \\ w_i \in \mathbb{R}^N, \text{ to be determined later} & \text{if } \lambda_i = 0. \end{cases}$$

Step 2. We first claim that $L = O \circ S$ (regardless of the choice of vectors w_i in the definition of O.)

It suffices to check that $Lv_i = O(Sv_i)$ for every i.

This is obvious if $\lambda_i \neq 0$.

If $\lambda_i = 0$, then $O(Sv_i) = 0$. In addition, $|Lv_i|^2 = Av_i \cdot v_i = 0$, so $Lv_i = 0$ and the desired identity holds.

Step 3. We now note that

$$\{Ov_i : \lambda_i \neq 0\}$$

forms an orthonormal set. Indeed, if $\lambda_i, \lambda_j \neq 0$, then

$$Ov_i \cdot Ov_j \;=\; (\lambda_i \lambda_j)^{-1/2} L v_i \cdot L v_j \;=\; (\lambda_i \lambda_j)^{-1/2} A v_i \cdot v_j \;=\; \frac{\lambda_i}{(\lambda_i \lambda_j)^{1/2}} v_i \cdot v_j = \delta_{ij}.$$

Step 4. For i such that $\lambda_i = 0$, we may thus choose w_i in the definition of O in such a way that $\{Ov_i\}_{i=1}^n$ is an orthonormal set, and then (according to the previous Lemma) O is an orthogonal map. $\square$

Lemma 3. *If $L : \mathbb{R}^N \to \mathbb{R}^n$ is a linear map, then L can be written*

$$L = S \circ P_{\mathbb{R}^N \to \mathbb{R}^n} \circ Q$$

where $Q : \mathbb{R}^N \to \mathbb{R}^N$ is orthogonal, $P_{\mathbb{R}^N \to \mathbb{R}^n}$ is the canonical projection of $\mathbb{R}^N$ onto $\mathbb{R}^n$ (that is, $P(x_1, \ldots, x_N) = (x_1, \ldots, x_n)$ and $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric.

`adjointpd`

Proof. By applying Lemma 2 to $L^* : \mathbb{R}^n \to \mathbb{R}^N$, we find that

$$L^* = O \circ S.$$

Thus $L = S \circ O^*$. Moreover, by the final conclusion of Lemma 1, we know that $O = Q \circ i_{\mathbb{R}^n \to \mathbb{R}^N}$, so that $O^* = (i_{\mathbb{R}^n \to \mathbb{R}^N})^* \circ Q^*$. Thus

$$L = S \circ (i_{\mathbb{R}^n \to \mathbb{R}^N})^* \circ Q^*.$$

Since $Q^* : \mathbb{R}^N \to \mathbb{R}^N$ is orthogonal whenever Q is, it suffices to check that $(i_{\mathbb{R}^n \to \mathbb{R}^N})^* = P_{\mathbb{R}^N \to \mathbb{R}^n}$, and this follows directly from the definitions. □

1.2. area formula for linear maps. We now prove the area and coarea formulas for linear maps. We will use the notation

$$\text{for linear } L : \mathbb{R}^n \to \mathbb{R}^N, \qquad JL := \sqrt{\det(L^* \circ L)}.$$

$$\text{for linear } L : \mathbb{R}^N \to \mathbb{R}^n, \qquad JL := \sqrt{\det(L \circ L^*)}.$$

(Here "det" means the determinant of the associated matrix.) Thus, in both cases, $JL = \det S$, where S is the symmetric matrix appearing in the polar decomposition.

The notation JL will be explained a bit later.

First, the area formula:

Proposition 1. *Assume that $L : \mathbb{R}^n \to \mathbb{R}^N$ is linear, and assume (as always) that $N \geq n$. Then for every measurable $A \subset \mathbb{R}^n$,*

$$\mathcal{H}^n(L(A)) \;=\; JL \cdot \mathcal{L}^n(A) \;=\; \int_A JL \; dx.$$

`area.linear`

Exercise 2. The proposition relies on the fact that if $O : \mathbb{R}^n \to \mathbb{R}^N$ is orthogonal and A is any subset of $\mathbb{R}^n$, then $\mathcal{H}^n(O(A)) = \mathcal{H}^n(A)$.

Prove that this holds. Note that it is not necessary to assume that A is measurable.

`ex.orthH`

Exercise 3. Here is a two-sentence proof that if A is any Lebesgue measurable subset of $\mathbb{R}^n$, then $O(A)$ is a $\mathcal{H}^n$ measurable subset of $\mathbb{R}^N$:

If A is Lebesgue measurable, then it is a countable union of compact sets, together with a set of $\mathcal{L}^n$ measure zero. Thus $O(A)$ is a countable union of compact sets, together with a set of $\mathcal{H}^n$ measure zero (by the previous exercise) and is hence $\mathcal{H}^n$ measurable.

Is there a one-sentence proof? I can't think of one at the moment, but perhaps it should exist. ("This is obvious" is not allowed.)

`ex.meas`

Proof of Proposition 1. Fix L linear and A measurable, and write $L = O \circ S$ as in Lemma 2. Then by exercise 2 and the fact that $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$,

$$\mathcal{H}^n(L(A)) = \mathcal{H}^n(O(S(A)) = \mathcal{H}^n(S(A)) = \mathcal{L}^n(S(A)).$$

Since $JL = \det S$, it therefore suffices to show that

cvbasic (1) $$\mathcal{L}^n(S(A)) = \det S\ \mathcal{L}^n(A).$$

Let us assume that we already know, from earlier in our lives, that this holds if S is diagonal (or see Exercise 4 below). In the general case it follows by writing $S = Q^* D Q$, where $Q : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal and $D : \mathbb{R}^n \to \mathbb{R}^n$ is diagonal, and then using Exercise 2 again (or more simply, in this case, the invariance of $\mathcal{L}^n$ with respect to rotations of $\mathbb{R}^n$). □

Exercise 4. Prove (1) if S is diagonal and nonnegative definite.
hint 1: It is clear if $\det S = 0$.
hint 2: If S is both diagonal and invertible and A is any subset of $\mathbb{R}^n$, then $|S(R)| = \det S \cdot |R|$ for every rectangle R, and

$$A \subset \cup_i R_i \qquad \text{if and only if } S(A) \subset \cup_i S(R_i).$$

As usual, we follow the convention that a "rectangle" is one whose sides are parallel to the coordinate axes, or in other words, a product of intervals.

ex.cvbasic

1.3. coarea formula for linear maps. Now we prove

Proposition 2. *Assume that $L : \mathbb{R}^N \to \mathbb{R}^n$ is linear, with $N \geq n$. Then for every measurable $A \subset \mathbb{R}^N$,*

coarealinear (2) $$JL \cdot \mathcal{L}^N(A) = \int_A JL\ dx = \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap L^{-1}\{y\})\, dy$$

The proof will use

Lemma 4. *Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is integrable and that $S : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear map with $\det S > 0$.*
Then

$$\int f(y)\, \mathcal{L}^n(dy) \;\; = \;\; \det S \int f \circ S(z)\ \mathcal{L}^n(dz).$$

cvlinear

(TERSE) PROOF. If f is the characteristic function of a measurable set, then this just reduces to the $N = n$ case of Proposition 1, or equivalently identity (1), established above.

It then follows for finite linear combinations of characteristic functions of measurable sets, and then by a standard approximation argument for general integrable functions.

□

The proposition just follows from Fubini's Theorem and a change of variables.

PROOF OF PROPOSITION 2. Let us write $L = S \circ P \circ Q$ as in Lemma 3, but writing P instead of $P_{\mathbb{R}^N \to \mathbb{R}^n}$ for simplicity.

Step 1. First, note that $JL = 0$ if and only if $\det S = 0$, and if this holds, the image of L is contained in a subspace of $\mathbb{R}^n$ of dimension at most $n-1$. Thus $A \cap L^{-1}\{y\}$ is empty at $\mathcal{L}^n$ a.e. y. It follows that (2) holds in this case, since both sides equal zero.

Step 2 We henceforth assume that $JL > 0$, so that S is invertible.

We may then use Lemma 4 to change variables in the integral on the right-hand side of (2). Since $\det S = JL > 0$, this yields

lca.int1

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap L^{-1}\{y\})\, dy = JL \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap L^{-1}\{Sz\})\, dz. \tag{3}$$

For every $y \in \mathbb{R}^n$, since $L^{-1} = Q^{-1} \circ P^{-1} \circ S^{-1}$,

$$A \cap L^{-1}\{Sz\} = A \cap Q^{-1}(P^{-1}\{z\}) = Q^{-1}\Big(Q(A) \cap P^{-1}\{z\}\Big).$$

It follows from the rotational invariance of Hausdorff measure that

$$\mathcal{H}^{N-n}(A \cap L^{-1}\{Sz\}) = \mathcal{H}^{N-n}\Big(Q(A) \cap P^{-1}\{z\}\Big)$$

for every $z \in \mathbb{R}^n$.

Step 3. By Fubini's Theorem $z \mapsto \mathcal{H}^{N-n}\big(Q(A) \cap P^{-1}\{z\}\big)$ is $\mathcal{L}^n$ measurable, and

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}\Big(Q(A) \cap P^{-1}\{z\}\Big) = \mathcal{L}^N(Q(A)) = \mathcal{L}^N(A).$$

(We have used again the rotational invariance of Lebesgue measure.) Combining this with (3), we obtain the conclusion of the Proposition. □

2. the area formula

In this section we will state without proof the area formula and some corollaries. The proof is rather similar in spirit to that of the coarea formula, which we provide in detail in the next section.

Theorem 1 (Area Formula)**.** *Assume that $n \le N$, and let $f : \mathbb{R}^n \to \mathbb{R}^N$ be a Lipschitz map and $A \subset \mathbb{R}^n$ a Lenesgue measurable set.*

Let

$$Jf(x) := \sqrt{\det(\nabla f^*(x)\nabla f(x))}$$

If f is injective, then

$$\int_A Jf(x) d\mathcal{H}^n(x) \;=\; \mathcal{H}^n(f(A)).$$

More generally,

$$\int_A Jf(x) d\mathcal{H}^n(x) \;=\; \int_{\mathbb{R}^N} \mathcal{H}^0\big(A \cap f^{-1}\{y\}\big)\; d\mathcal{H}^n(y)$$

We will sometimes use the notation

$$N(f, A, y) := \mathcal{H}^0\big(A \cap f^{-1}\{y\}\big)$$

for the number of point in the intersection of A with the preimage of f. Note that $N(f, A, y) \neq 0$ if and only if $y \in f(A)$.

As a corollary, we have a change of variables formula. We state it in the general case (ie, when f is not injective). The formulas become much more transparent if f is injective.

Theorem 2. *Assume that $n \le N$, and let $f : \mathbb{R}^n \to \mathbb{R}^N$ be a Lipschitz map and $A \subset \mathbb{R}^n$ a Lenesgue measurable set, and defie the Jacobian Jf as above.*

If $u : \mathbb{R}^n \to [0, \infty]$ is Lebesgue measurable, then

$$y \in \mathbb{R}^N \mapsto \sum_{x \in f^{-1}(y)} u(x)$$

is $\mathcal{H}^n$ measurable, and

$$\int_{\mathbb{R}^n} u(x)\, Jf(x)\, d\mathcal{H}^n(x) \;=\; \int_{\mathbb{R}^N} \left(\sum_{x\in f^{-1}(y)} u(x) \right) d\mathcal{H}^n(y).$$

As a result, if $v:\mathbb{R}^N\to[0,\infty]$ and $A\subset\mathbb{R}^n$ are $\mathcal{H}^n$ measurable, then

$$\int_A v\circ f(x)\, Jf(x)\, d\mathcal{H}^n(x) \;=\; \int_{\mathbb{R}^N} v(y) N(f,A,y) d\mathcal{H}^n(y).$$

Example 1. If $n = 1$ and $f:\mathbb{R}\to\mathbb{R}^N$ is a Lipschitz curve, then $Jf(x) = |f'(x)|$ a.e., and the area formula implies that, for an interval I,

$$\mathcal{H}^1(f(I)) = \int_I |f'(x)|\, dx \qquad \text{if } f \text{ is injective,}$$

The case when f is not injective is illustrated by the example $f(x) = (\cos x, \sin x)$. If we take I to be the interval $[0, 3\pi]$, then

$$3\pi \;=\; \int_I |f'(x)| dx \;=\; \int_{\mathbb{R}^2} N(f,I,y)\; d\mathcal{H}^1(y)$$

where the "multiplicity function $N(f,I,y)$ is

$$N(f,I,y) = \begin{cases} 0 & \text{if } y \text{ is not in the unit circle} \\ 1 & \text{if } y=(y_1,y_2) \text{ belongs to the unit circle, and } y_2<0 \\ 2 & \text{if } y=(y_1,y_2) \text{ belongs to the unit circle, and } y_2\ge 0 \end{cases}$$

Example 2. Given a Lipschitz function $f:\mathbb{R}^n\to\mathbb{R}$, let us define $F(x) = (x, f(x))$. Then $F:\mathbb{R}^n\to\mathbb{R}^{n+1}$ is Lipschitz, and one can check that $JF(x) = (1+|\nabla f|^2)^{1/2}$. Thus for any measurable $A\subset\mathbb{R}^n$,

$$\mathcal{H}^n(\{(x,f(x)) : x\in A\}) = \int_A (1+|\nabla f|^2)^{1/2} dx.$$

area.ex2 Some version of this formula is probably familiar for smooth functions.

We will see more examples later.

Exercise 5. If one actually has to compute a Jacobian $Jf(x)$, by hand, computations can often be simplified by working in a coordinate system in which $\nabla f(x)$ has a simple form.

Use this observation to check quickly that in the situation described in Example 2, it is in fact the case that $JF(x) = (1+|\nabla f(x)|^2)^{1/2}$, as asserted above.

For a bit more of a challenge, assume that $f:\mathbb{R}^3\to\mathbb{R}^2$ is Lipschitz, define $F:\mathbb{R}^3\to\mathbb{R}^5$ by $F(x) = (x, f(x))$, and check that

$$JF(x)^2 = 1 + |\nabla f^1|^2 + |\nabla f^2|^2 + |\nabla f^1\times\nabla f^2|^2$$

where (f^1, f^2) are the components of f. This can be done in a few minutes by choosing a good basis for $\mathbb{R}^3$ at the point x.

3. the coarea formula

Theorem 3 (Coarea formula)**.** *Assume that $N \geq n$ and let $f : \mathbb{R}^N \to \mathbb{R}^n$ be a Lipschitz map.*

Define

$$Jf(x) := \sqrt{\det(\nabla f(x) \nabla f^*(x))}.$$

Then for every measurable $A \subset \mathbb{R}^N$,

coarea1
$$\int_A Jf(x)\, dx \;=\; \int_{\mathbb{R}^n} \mathcal{H}^{N-n}\big(A \cap f^{-1}\{y\}\big) d\mathcal{H}^n(y). \tag{4}$$

And for every measurable $g : \mathbb{R}^N \to [0, \infty]$,

coarea2
$$\int_{\mathbb{R}^N} g(x)\, Jf(x)\, dx \;=\; \int_{\mathbb{R}^n} \left(\int_{f^{-1}\{y\}} g(x) d\mathcal{H}^{N-n}(x) \right) d\mathcal{H}^n(y). \tag{5}$$

We will omit the proof of (5). The idea is that (4) shows that it holds if g is the characteristic function of a measurable set, and it follows for finite linear combinations of characteristic functions. It is then deduced for arbitrary measurable g by an approximation argument.

Before giving the full proof, we give an

OUTLINE OF THE PROOF OF THE COAREA FORMULA. **Step 1**. $f(A)$ is a measurable subset of $\mathbb{R}^n$. This will be left as an exercise.

Step 2. For any measure μ on $\mathbb{R}^n$ and function $g : \mathbb{R}^n \to \mathbb{R}$, not necessarily μ-measurable, define

intstar
$$\int_{\mathbb{R}^n}^{*} g(x) d\mu(y) := \inf \left\{ \int_{\mathbb{R}^n} h(y) d\mu(y) \;:\; h \text{ is } \mu\text{-meas., } g \leq h \ \mu \text{ a.e.} \right\}. \tag{6}$$

Then

$$\int_{\mathbb{R}^n}^{*} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n \;\leq\; \frac{\omega_{N-n}\omega_n}{\omega_N} (\mathrm{Lip}(f))^n \mathcal{H}^N(A).$$

Step 3. $y \in \mathbb{R}^n \mapsto \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$ measurable, and thus

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n \;\leq\; \frac{\omega_{N-n}\omega_n}{\omega_N} (\mathrm{Lip}(f))^n \mathcal{H}^N(A).$$

Step 4. reduction to the case when f is C^1.

Step 5. proof of the coarea formula if f is C^1 and $Jf(x) > 0$ in A.

Step 6. proof of the coarea formula if f is C^1 and $Jf(x) = 0$ in A. □

Here are some remarks about how this is done, working backwards through the argument.

- The idea of the final two steps is to approximate f by linear functions, for which we already know the formula. These rely, among other things, on the measurability proved in Step 3 (which is not any easier for $f \in C^1$ than for Lipschitz f.

- The case $Jf = 0$ is a little harder in some ways, and is handled by, in effect, perturbing f to get a function whose Jacobian is everywhere positive, using the previous case, and then taking limits as the perturbation tends to zero.
- The reduction to the case of C^1 maps relies on a Lusin-type theorem for Lipschitz maps, which says that given a Lipschitz map f, one can find a C^1 map that agrees with f (and such that the gradients also agree) outside a set of arbitrarily small measure.

  In order to control the approximations, however, we need the preliminary upper bound of $\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})$ proved in Steps 2 and 3.
- Measurability is proved by first considering the case when A is compact. For general A, it is proved by an approximation argument. The preliminary estimate of $\int_{\mathbb{R}^n}^* \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})$ is needed to control the errors in the approximation.
- Step 2 is proved by a smart argument that directly uses the definition of Hausdorff measure. The idea is that a cover of A, suitable for estimating the $\mathcal{H}^N(A)$, also yields covers of $A \cap f^{-1}\{y\}$ that can be used to estimate $\mathcal{H}^{N-n}(A \cap f^{-1}\{y\})$ for every y.

We remark that one often sees arguments with the above structure: in order to prove some inequality, we first prove a weaker version of it, which can be used to show, for example, that sets of measure zero are harmless. In fact, we have already seen this in the proof that $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$. Here the crucial point is that subsets of $\mathbb{R}^N$ of $\mathcal{L}^N$ measure zero have a negligible impact on the function $y \mapsto \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})$.

With those preliminaries, here is the complete proof:

Proof. **Step 1**.

Exercise 6. Prove that under the assumptions of the theorem, $f(A)$ is a $\mathcal{L}^n$-measurable of $\mathbb{R}^n$. (Consult the discussion in exercise 3 above.)

Step 2. We will use the following fact.

Lemma 5. *If $A \subset \mathbb{R}^n$, then*

$$\mathcal{H}^n(A) = \mathcal{H}^n_S(A),$$

where $\mathcal{H}^s_S(A)$ ("spherical Hausdorff measure") is defined by

$$\mathcal{H}^s_S(A) = \lim_{\delta \searrow 0} \mathcal{H}^s_{S,\delta}(A),$$

with

$$\mathcal{H}^s_{S,\delta}(A) = \inf\{\sum \omega_s (\operatorname{diam} B_i)^s : A \subset \cup_i B_i, \textit{every } B_i \text{ a ball } \textit{of diameter less than } \delta\}$$

Exercise 7. Sketch a proof that the lemma holds. I *do not* recommend trying to do this from scratch, so to speak – it's much easier and more sensible to cite some theorem that will take care of much of the proof.

It is not in general true for $s < n$ that $\mathcal{H}^s = \mathcal{H}^s_S$ in $\mathbb{R}^n$.

In view of the lemma, for every $j > 0$ we can find a collection of balls $\{B_i^j\}_{i=1}^\infty$ such that

$$A \subset \cup_i B_i^j, \qquad \sum_i \omega_N (\operatorname{diam} B_i^j)^N \le \mathcal{H}^N(A) + \frac{1}{j}, \qquad \operatorname{diam} B_i^j < \frac{1}{j} \text{ for all } i.$$

These balls clearly cover $A \cap f^{-1}\{y\}$ for every y, so that for every y,

$$\mathcal{H}^{N-n}_{1/j}(A \cap f^{-1}\{y\}) \le \sum_{i=1}^\infty \left[\omega_{N-n} (\operatorname{diam} B_i^j)^{N-n} \mathbf{1}_{y \in f(B_i^j)}\right]$$

We define $g_i^j(y)$ to be the term in square brackets $[\cdots]$ on the right-hand side above. Then

$$\mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) \le \liminf_{j\to\infty} \sum_i g_i^j(y) =: g(y) \qquad \text{for every } y.$$

Note also that g_i^j is $\mathcal{H}^n$-measurable for every i, j, so that $g = \liminf_{j\to\infty} \sum_i g_i^j$ is also $\mathcal{H}^n$-measurable. Thus

$$\int^*_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n \;\le\; \int^*_{\mathbb{R}^n} g(y) d\mathcal{H}^n(y) \;=\; \int_{\mathbb{R}^n} g(y)\, d\mathcal{H}^n(y).$$

since it is easy to see that $\int^* g d\mathcal{H}^n = \int g \; d\mathcal{H}^n$ is g is measurable. And by Fatou's Lemma,

$$\begin{aligned}
\int_{\mathbb{R}^n} g(y) \; d\mathcal{H}^n(y) &= \int_{\mathbb{R}^n} \liminf_{j\to\infty} \sum_i g_i^j \; d\mathcal{H}^n(y) \\
&\le \liminf_{j\to\infty} \int_{\mathbb{R}^n} \sum_i g_i^j \; d\mathcal{H}^n(y) \\
&= \liminf_{j\to\infty} \sum_i \omega_{N-n} (\operatorname{diam} B_i^j)^{N-n} \mathcal{H}^n(f(B_i^j)).
\end{aligned}$$

By the isodiamteric inequality,

$$\mathcal{H}^n(f(B_i^j)) \le \omega_n (\operatorname{diam} f(B_i^j))^n \le \omega_n \operatorname{Lip}(f)^n \operatorname{diam}(B_i^j)^n.$$

Inserting this into the above, we find that

$$\int^*_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n \le \liminf_{j\to\infty} \omega_{N-n}\omega_n \operatorname{Lip}(f)^n \sum_i \operatorname{diam}(B_i^j)^N.$$

So by the choice of the balls B_i^j we conclude that

step2 (7)
$$\int^*_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n \le \frac{\omega_{N-n}\omega_n}{\omega_N} \operatorname{Lip}(f)^n \mathcal{H}^N(A).$$

Step 3. We now prove that

step3 (8)
$$y \in \mathbb{R}^n \mapsto \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) \text{ is } \mathcal{H}^n \text{ measurable.}$$

Since $\int^* g = \int^g$ when g is measurable, this will allow us to replace $\int^*$ by $\int$ in (7) above.

We consider several cases.

Case 1: A is compact.

Since

$$\mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) = \lim_{\delta\to 0} \mathcal{H}^{N-n}_\delta(A \cap f^{-1}\{y\})$$

it suffices to prove that $y \mapsto \mathcal{H}^{N-n}_\delta(A \cap f^{-1}\{y\})$ is measurable for every $\delta > 0$. To do this, we will show that for every δ, this function is upper semicontinuous. Thus, we fix some $y \in \mathbb{R}^n$, and some sequence $y_k \to y$, and we will show that

step3a

$$\limsup_{k\to\infty} \mathcal{H}^{N-n}_\delta(A \cap f^{-1}\{y_k\}) \le \mathcal{H}^{N-n}_\delta(A \cap f^{-1}\{y\}). \tag{9}$$

To do this, fix $\varepsilon > 0$, and fix *open* sets C_i such that $\operatorname{diam} C_i < \delta$ for all i,

$$A \cap f^{-1}\{y\} \subset \cup C_i, \qquad \text{and} \quad \sum \omega_{N-n}(\operatorname{diam} C_i)^{N-n} \le \mathcal{H}^{N-n}_\delta(A \cap f^{-1}\{y\}) + \varepsilon.$$

Then the compactness of A implies that $f^{-1}\{y_k\} \cap A \subset \cup C_i$ for all sufficiently large k. To see this, suppose toward a contradiction that there exists some subsequence (still labelled (x_k)) such that

$$x_k \in f^{-1}\{y_k\} \cap A, \qquad x_k \notin \cup_i C_i$$

for every k. Since A is compact, we can pass to a further subsequence (which however we still label (x_k)) that converges to a limit $x \in A$. It is then clear that $x \in f^{-1}\{y\} \cap A$, and hence that $x \in C_i$ for some i. Thus all sufficiently large x_k also belong to C_i, which is a contradiction.

Since $f^{-1}\{y_k\} \cap A \subset \cup C_i$ for all sufficiently large k, it follows from the choice of (C_i) that

$$\limsup_{k\to\infty} \mathcal{H}^{N-n}_\delta\big(A \cap f^{-1}\{y_k\}\big) \le \mathcal{H}^{N-n}\big(A \cap f^{-1}\{y\}\big) + \varepsilon$$

for every $\varepsilon > 0$, which clearly implies (9).

Exercise 8. In the lecture the claim was made, and then retracted, that $y \mapsto \mathcal{H}^{N-n}(f^{-1}\{y\} \cap A)$ is upper semicontinuous if A is compact.

Construct an example to show that this claim is in fact *false* for $\mathcal{H}^{N-n}$ (although as proved above it is true for $\mathcal{H}^{N-n}_\delta$ for every positive δ.)

That is, explicitly define a function $f : \mathbb{R}^N \to \mathbb{R}^n$ for some $N > n$, a compact set $A \subset \mathbb{R}^N$ and some $y \in \mathbb{R}^n$ such that

$$\limsup_{z\to y} \mathcal{H}^{N-n}(A \cap f^{-1}\{z\}) > \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}).$$

Recall also that we have drawn a picture to give a argument (in retrospect, a misleading argument) that the opposite inequality is plausible. Presumably the example you construct should look rather unlike that picture.

returning to the proof.....

Case 2: A is measurable.

Then there is an increasing sequence $K_1 \subset K_2 \subset \dots$ of compact sets such that $A \setminus \cup_{i=1}^\infty K_i$ has Lebesgue measure zero. For simplicity, we will write $E := \cup_{i=1}^\infty K_i$. Then for every y,

step3b

$$\mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) = \mathcal{H}^{N-n}(E \cap f^{-1}\{y\}) + \mathcal{H}^{N-n}((A \setminus E) \cap f^{-1}\{y\}). \tag{10}$$

Since $\mathcal{H}^{N-n}(E \cap f^{-1}\{y\}) = \lim_{i\to\infty} \mathcal{H}^{N-n}(K_i \cap f^{-1}\{y\})$, it follows from Case 1 above that it is Borel measurable as a function of $y \in \mathbb{R}^n$.

And since $\mathcal{L}^n(A \setminus E) = 0$, we see from (7) that

$$\int^*_{\mathbb{R}^n} \mathcal{H}^{N-n}((A \setminus E) \cap f^{-1}\{y\})\, dy = 0$$

which implies that $\mathcal{H}^{N-n}((A \setminus E) \cap f^{-1}\{y\}) = 0$ a.e.. Then the measurability of $\mathcal{H}^{N-n}(A \cap f^{-1}\{\cdot\})$ follows from (10).

Step 4. Reduction to the case when f is C^1.

Say we are given $f : \mathbb{R}^N \to \mathbb{R}^n$ Lipschitz and measurable $A \subset \mathbb{R}^N$.

For each $\varepsilon > 0$, by a Lusin-type theorem[1] for Lipschitz functions, there exists a C^1 function f_ε and a measurable set G_ε such that

$$f = f_\varepsilon \text{ and } \nabla f = \nabla f_\varepsilon \text{ in } G_\varepsilon, \qquad \mathcal{L}^N(\mathbb{R}^N \setminus G_\varepsilon) < \varepsilon.$$

Let $A_\varepsilon := A \cap G_\varepsilon$. It is straightforward to verify from the definition of the Jacobian that $0 \leq Jf(x) \leq C \operatorname{Lip}(f)$ for $\mathcal{L}^N$ a.e. x, for some constant C depending on N, n and $\operatorname{Lip}(f)$, so

reduction1

$$\left| \int_A Jf(x)\,dx - \int_{A_\varepsilon} Jf(x)\,dx \right| = \int_{A \setminus A_\varepsilon} Jf(x)\;dx \leq \varepsilon C \tag{11}$$

And by Step 2,

reduction2

$$\begin{aligned} \left| \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})\,dy - \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A_\varepsilon \cap f^{-1}\{y\})\,dy \right| \\ = \int_{\mathbb{R}^n} \mathcal{H}^{N-n}((A \setminus A_\varepsilon) \cap f^{-1}\{y\})\,dy \\ \leq C\mathcal{L}^n(A \setminus A_\varepsilon) = C\varepsilon \end{aligned} \tag{12}$$

for a (possibly different) constant C depending on N, n and $\operatorname{Lip}(f)$. If the coarea formula holds for C^1 functions, then

$$\int_{A_\varepsilon} Jf(x)\,dx \;=\; \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A_\varepsilon \cap f^{-1}\{y\})\,dy$$

for every ε, and the coarea formula for Lipschitz functions then follows from (11) and (12).

Step 5. We now prove the coarea formula under the assumptions that $f \in C^1(\mathbb{R}^N, \mathbb{R}^n)$ and $Jf(x) > 0$ in A.

Step 5.1 Fix $t > 1$. (We will eventually let $t \searrow 1$.)

We claim that for each $\xi \in A$, there exists an open set D_ξ, an invertible map $h_\xi \in C^1(D_\xi, \mathbb{R}^N)$ and a linear map $L_\xi : \mathbb{R}^N \to \mathbb{R}^n$ such that

step5.1

$$f = L_\xi \circ h_\xi \quad \text{on } D_\xi, \qquad \operatorname{Lip}(h_\xi) \leq t \text{ in } D_\xi, \qquad \operatorname{Lip}(h_\xi^{-1}) \leq t \text{ in } h(D_\xi). \tag{13}$$

Thus, on each set, f is the composition of a linear map (to which we can apply the coarea formula) and an approximate isometry.

To see this, let

$$I(N, N-n) := \{\alpha \in \mathbb{Z}^{N-n} : 1 \leq \alpha_1 < \ldots < \alpha_{N-n} \leq N\},$$

and for $\alpha \in I(N, N-n)$, define $p_\alpha : \mathbb{R}^N \to \mathbb{R}^{N-n}$ by

$$p_\alpha(x) = (x_{\alpha_1}, \ldots, x_{\alpha_{N-n}}).$$

---

[1] We stated this result, and briefly discussed its proof, some time ago. We recall that the basic point was to combine Rademacher's Theorem with Lusin's Theorem (applied to ∇f) and the Whitney Extension Theorem.

Then define $F_\alpha : \mathbb{R}^N \to \mathbb{R}^N$ by

$$F_\alpha(x) = (f(x), p_\alpha(x)).$$

For every α, note that $f = \pi \circ F_\alpha$, where $\pi(x_1, \ldots, x_N) = (x_1, \ldots, x_n)$.

Fix an arbitrary $\xi \in A$. Since $Jf(\xi) > 0$, the matrix $\nabla f(\xi)$ must have rank n, and it follows that $\nabla F_\alpha(\xi)$ is invertible for some $\alpha = \alpha(\xi)$. For such an index α (which may not be unique), let $h_\xi := (\nabla F_\alpha(\xi))^{-1} \circ F_\alpha$, and $L_\xi = \pi \circ \nabla F_\alpha(\xi)$. Then

$$L_\xi \circ h_\xi = \pi \circ \nabla F_\alpha(\xi) \circ (\nabla F_\alpha(\xi))^{-1} \circ F_\alpha = \pi \circ F_\alpha = f$$

as desired. Moreover, $\nabla h_\xi(\xi) = Id_N$ by construction. So the inverse function theorem implies that h_α is invertible in a neighborhood of ξ, and that $\nabla(h_\xi^{-1})(h(\xi)) = Id_N$. Then a Taylor series expansions of both h_ξ and h_ξ^{-1} implies that one can find an open neighborhood D_ξ of ξ so small that the remaining conclusions of (13) hold.

Step 5.2.

We next claim that there exist Borel sets D_k, invertible maps $h_k \in C^1(D_k, \mathbb{R}^N)$, and linear maps $L_k : \mathbb{R}^N \to \mathbb{R}^n$ such that

step5.2a

$$A = \sqcup_k D_k \qquad \text{ie, a } \textit{disjoint} \text{ union} \tag{14}$$

and

step5.2

$$f = L_k \circ h_k \quad \text{on } D_k, \qquad \mathrm{Lip}(h_k) \le t \text{ in } D_k, \qquad \mathrm{Lip}(h_k^{-1}) \le t \text{ in } h_k(D_k). \tag{15}$$

To see this, note that the sets $\{D_\xi\}_{\xi \in A}$ found above form an open cover of A, and hence there is a countable subcover, associated to a countable sequence $\{\xi_1, \xi_2, \ldots\}$ of points in A. Then (14) and (15) hold if we define $L_k = L_{\xi_k}, h_k = h_{\xi_k}$, and

$$D_1 := D_{\xi_1}, \quad D_2 := D_{\xi_2} \setminus D_1, \quad \ldots, \quad D_k := D_{\xi_{k+1}} \setminus (\cup_{j=1}^k D_j).$$

Step 5.3. Next we claim that for every k,

step5.3

$$t^{-n} JL_k \le Jf(x) \le t^n JL_k \quad \text{ in } D_k. \tag{16}$$

Fix $x \in D_k$, and note that

nabhkbd

$$\|\nabla h_k(x)\| = \|\nabla h_k^*(x)\| \le t \text{ in } D_k, \tag{17}$$

where $\|M\|$ denotes the operator norm of a matrix M, where we recall

$$\begin{aligned}\|M\| := \sup\{|Mv| \ : \ |v| \le 1\} &= \sup\{Mv \cdot w \ : \ |v| \le 1, |w| \le 1\} \\ &= \sup\{v \cdot M^* w \ : \ |v| \le 1, |w| \le 1\} = \|M^*\|.\end{aligned}$$

Estimate (17) follows directly from the fact that $\mathrm{Lip}(h_k) \le t$ and the characterization of $\nabla h_k(x) v$ as a limit of difference quotients.

By definition,

$$Jf(x)^2 = \det(\nabla f(x) \circ \nabla f^*(x)) = \det(L_k \circ \nabla h_k(x) \circ \nabla h_k^*(x) \circ L_k^*).$$

Clearly $\nabla h_k(x) \circ \nabla h_k^*(x)$ is a symmetric, positive definite $N \times N$ matrix. As such it can be written in the form $\nabla h_k(x) \circ \nabla h_k^*(x) = Q^* \circ D \circ Q$, where $Q^* \circ Q = Id_N$ and D is diagonal, with all eigenvalues of D bounded by t^2. (This follows from (17).)

We can also write $L_k = S \circ O^*$, with $O : \mathbb{R}^n \to \mathbb{R}^N$ orthogonal and $S : \mathbb{R}^n \to \mathbb{R}^n$ symmetric. Thus

$$Jf(x)^2 = \det(S \circ \tilde{O}^* \circ D \circ \tilde{O} \circ S^*), \qquad \tilde{O} := Q \circ O \ \text{ orthogonal}.$$

Moreover, $O^* \circ D \circ \tilde{O}$ is a $n \times n$ matrix with all eigenvalues less than t^2, so

$$Jf(x)^2 = \det S \det(O^* \circ D \circ \tilde{O}) \det S^* = t^{2n} \det S^2 = t^{2n}(JL_k)^2.$$

This proves one of the inequalities in (16). The proof used only the fact that $\nabla f(x) = L_k \circ \nabla h_k(x)$, and the bounds (17), so the opposite inequality follows by the same argument from the fact that $L_k = \nabla f(x) \circ \nabla h_k(x)^{-1}$, with $\nabla h_k(x)$ satisfying (17).

Step 5.4 To complete the proof we will use the following lemma, whose proof is an easy execise.

Lemma 6. *Assume that $h : \mathbb{R}^N \to \mathbb{R}^N$ is a Lipschitz function. Then for any $A \subset \mathbb{R}^N$ and any $s > 0$,*

step5.4

$$\mathcal{H}^s(h(A)) \le (Lip(h))^s \mathcal{H}^s(A). \tag{18}$$

Exercise 9. prove the lemma.

Applying the Lemma to both h_k and h_k^{-1}, we deduce that

step5.4a

$$t^{-N}\mathcal{H}^N(h_k(D_k)) \le \mathcal{H}^N(D_k) \le t^N \mathcal{H}^N(h_k(D_k)). \tag{19}$$

Moreover, since $h_k(D_k \cap f^{-1}\{y\}) = h_k(D_k) \cap L_k^{-1}\{y\}$, the Lemma also implies that

step5.4b

$$\begin{aligned} t^{n-N}\mathcal{H}^{N-n}(h_k(D_k) \cap L_k^{-1}\{y\}) &\le \mathcal{H}^{N-n}(D_k \cap f^{-1}\{y\}) \\ &\le t^{N-n}\mathcal{H}^{N-n}(h_k(D_k) \cap L_k^{-1}\{y\}) \end{aligned} \tag{20}$$

Step 5.5. We now use the above information to compute:

$$\begin{aligned} \int_{D_k} Jf(x)\, d\mathcal{H}^N(x) &\overset{(16)}{\le} t^n \ JL_k \ \mathcal{H}^N(D_k) \\ &\overset{(19)}{\le} t^{n+N} \ JL_k \ \mathcal{H}^N(h_k(D_k)) \\ &\overset{(2)}{=} t^{n+N} \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(h_k(D_k) \cap L_k^{-1}\{y\}) \ d\mathcal{H}^n(y) \\ &\overset{(20)}{\le} t^{2N} \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(D_k \cap f^{-1}\{y\}) \ d\mathcal{H}^n(y). \end{aligned}$$

And using the opposite inequalities,

$$\begin{aligned} \int_{D_k} Jf(x)\, d\mathcal{H}^N(x) &\ge t^n \ JL_k \ \mathcal{H}^N(D_k) \\ &\ge t^{n+N} \ JL_k \ \mathcal{H}^N(h_k(D_k)) \\ &= t^{n+N} \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(h_k(D_k) \cap L_k^{-1}\{y\}) \ d\mathcal{H}^n(y) \\ &\ge t^{2N} \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(D_k \cap f^{-1}\{y\}) \ d\mathcal{H}^n(y). \end{aligned}$$

We sum over k to find that

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

Finally, we obtain the coarea formula (in $Jf > 0$ in A) by letting $t \searrow 1$.

Step 6. Finally, assume that f is C^1 and $Jf(x) = 0$ in A, and that $\mathcal{L}^n(A) < \infty$, and we prove the coarea formula. Clearly, we have to show that the right-hand side of the formula vanishes.

This will complete the proof, since for $f \in C^1(\mathbb{R}^N, \mathbb{R}^n)$ and measurable $A \subset \mathbb{R}^N$, we can write A as the disjoint union of $\{x \in A : Jf(x) > 0\}$ and sets of finite measure on which $Jf = 0$.

We will argue by reducing this case to the previous case. To do this, fix $\varepsilon > 0$ and define $g : \mathbb{R}^N \times \mathbb{R}^n \to \mathbb{R}^n$ and $\hat{\pi} : \mathbb{R}^N \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$g(x, y) = f(x) + \varepsilon y, \qquad \hat{\pi}(x, y) = y.$$

Then one computes that

$$\nabla g(x, y) = (\nabla f(x), \varepsilon \operatorname{Id}_n), \qquad \nabla g(x, y) \nabla^* g(x, y) = \nabla f(x) \nabla^* f(x) + \operatorname{Id}_n.$$

By writing $\nabla g \nabla^* g$ in a basis in which $\nabla f \nabla^* f$ is diagonal (with at least one zero eigenvalue and all eigenvalues bounded by $\operatorname{Lip}(f)^2$) we see that

$$\varepsilon^n \le Jg(x, y) \le C\varepsilon, \qquad \text{for } C = \operatorname{Lip}(f)^{n-1}$$

for all $x \in A$ and $y \in \mathbb{R}^n$. (Here and below we are assuming that $\varepsilon \ll 1$, so that $\varepsilon^n \ll \varepsilon$.)

Note that for every $z \in \mathbb{R}^n$, a simple change of variables implies that

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n(y) = \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y - \varepsilon z\}) d\mathcal{H}^n(y).$$

If we fix some measurable $E \subset \mathbb{R}^n$ with $\mathcal{L}^n(E) = 1$, it follows that

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\}) d\mathcal{H}^n(y) \\ = \int_E \left(\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y - \varepsilon z\}) d\mathcal{H}^n(y) \right) d\mathcal{H}^n(z).$$

Now let $\mathbf{A} := A \times E \subset \mathbb{R}^{N+n}$. Straightforward manipulations of the definitions show that

$$\mathbf{A} \cap g^{-1}\{y\} \cap \hat{\pi}^{-1}\{z\} = \begin{cases} (A \cap f^{-1}\{y + \varepsilon z\}) \times \{z\} & \text{if } w \in E \\ \emptyset & \text{if not.} \end{cases}$$

So

$$\mathcal{H}^{N-n}(A \cap f^{-1}\{y - \varepsilon z\}) \mathbf{1}_{w \in E} \;=\; \mathcal{H}^{N-n}(\mathbf{A} \cap g^{-1}\{y\} \cap \hat{\pi}^{-1}\{z\}).$$

Thus

$$\int_E \left(\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y - \varepsilon z\}) d\mathcal{H}^n(y) \right) d\mathcal{H}^n(z) \\ = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{H}^{N-n}(\mathbf{A} \cap g^{-1}\{y\} \cap \hat{\pi}^{-1}\{z\}) d\mathcal{H}^n(z)\, d\mathcal{H}^n(y)$$

We will argue below that for every Borel $\mathbf{B} \subset \mathbb{R}^{N+n}$ (and in particular for $\mathbf{B} = \mathbf{A} \cap g^{-1}\{y\}$)

`Fubini?` (21)
$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(\mathbf{B} \cap \hat{\pi}^{-1}\{z\}) d\mathcal{H}^n(z) \le C\mathcal{H}^N(\mathbf{B})$$

for some constant C. We accept this for the time being and continue. Since $Jg > 0$ everywhere, we can combine the previous few statements and apply the previous case of the coarea formula to conclude that

$$\begin{aligned}\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})\, d\mathcal{H}^n(y) &\leq \int_{\mathbb{R}^n} \mathcal{H}^N(\mathbf{A} \cap g^{-1}\{y\}) d\mathcal{H}^n(y) \\ &= \int_{\mathbf{A}} Jg\, d\mathcal{H}^{N+n} \\ &\leq \varepsilon C \mathcal{H}^{N+n}(\mathbf{A}) = \varepsilon C \mathcal{H}^N(A).\end{aligned}$$

Since ε is arbitrary, we conclude that

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

This concludes the proof, modulo estimate (21) above. This can be established by the argument that we used to prove (7) above. This yields the inequality

step6z

$$\int_{\mathbb{R}^n}^{*} \mathcal{H}^{N-n}(\mathbf{B} \cap \pi^{-1}\{z\})\ d\mathcal{H}^n(z) \ \leq\ \frac{\omega_{N-n}\omega_n}{\omega_N} \mathrm{Lip(f)}^n \mathcal{H}^N(\mathbf{B}), \tag{22}$$

which is the same formula as above, with the difference that this version concerns an N-dimesional subset of $\mathbb{R}^{N+n}$, whereas in the previous version we had an N-dimensional subset of $\mathbb{R}^N$.

Exercise 10. Check that the proof of (7) indeed shows that (22) holds.

In view of the fact that $\int^*$ rather than $\int$ appears in (22), to be really correct we should strictly replace $\int$ by $\int^*$ in several places at the very end of the proof.

□

Exercise 11. The book *Cartesian Currents in the Calculus of Variations, vol. 1* by Giaquinta, Modica and Souček states (in Step 6 of its discussion of the coarea formula, in Section 2.1.3) that for every Borel $\mathbf{B} \subset \mathbb{R}^{N+n}$ (and in particular for $\mathbf{B} = \mathbf{A} \cap g^{-1}\{y\}$)

Fubini-wrong

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(\mathbf{B} \cap \hat{\pi}^{-1}\{z\}) d\mathcal{H}^n(z) = \mathcal{H}^N(\mathbf{B}) \tag{23}$$

by Fubini's Theorem, where $\mathbf{B}$ is a subset of $\mathbb{R}^{N+n}$.

Given an example to show that (23) is not always correct.

Giaquinta *et al* deduce from the incorrect statement (23) that, using notation from Step 6 of the above proof,

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(A \cap f^{-1}\{y\})\, d\mathcal{H}^n(y) = \int_{\mathbf{A}} Jg\, d\mathcal{H}^{N+n}.$$

Since the right-hand side of this alleged identity depends on ε — recalling the way in which g depends on ε, we see that it lies in the interval $[\varepsilon^n \mathcal{H}^N(A), C\varepsilon\mathcal{H}^N(A)]$ — and the left-hand side does not, this is clearly false.

Now we state some consequences of the coarea formula.

Corollary 1. *Assume that $g : \mathbb{R}^N \to \mathbb{R}$ is integrable. Then*

$$\int_{\mathbb{R}^N} g(x)\, d\mathcal{H}^N(x) = \int_0^\infty \left(\int_{\{x \in \mathbb{R}^N : |x| = r\}} g\ d\mathcal{H}^{N-1} \right) dr.$$

Proof. This is a special case of the change-of-variables formula (5), with $f(x) = |x|$. For this choice of f one can check that $Jf(x) = 1$ almost everywhere. $\square$

We remark that in general, the definition more or less immediately implies that

$$\boxed{Jf(x) = |\nabla f(x)| \quad \text{for } f : \mathbb{R}^N \to \mathbb{R} \text{ (ie, in the case when } n = 1),}$$

Thus for a Lipchitz function,

$$\int_A |\nabla f| \, d\mathcal{H}^N \ = \ \int_{-\infty}^{\infty} \mathcal{H}^{N-1}(\{x \in A : f(x) = y\}) \, dy.$$

Another corollary is

Corollary 2 ("C^1 Sard-type Theorem")**.** *Suppose $f : \mathbb{R}^N \to \mathbb{R}^n$ is C^1 and $A \subset \mathbb{R}^N$ is open.*

Then for $\mathcal{H}^n$ a.e. $y \in \mathbb{R}^n$, $f^{-1}\{y\}$ is the union of a $(N-n)$-dimensional C^1 submanifold of $\mathbb{R}^N$ and a set of $\mathcal{H}^{N-n}$ measure zero. Precisely, if we define

$$C := \{x \in A : Jf(x) = 0\} = \{x \in A : \mathit{rank}(\nabla f(x)) < n\}$$

then $\mathcal{H}^{N-n}(f^{-1}\{y\} \cap C) = 0$ for a.e. y, and $f^{-1}\{y\} \setminus C$ is a $(N-n)$-dimensional C^1 submanifold of $\mathbb{R}^N$.

Recall that Sard's Theorem states that if f is C^{N-n+1}, then $f^{-1}\{y\} \cap C$ is in fact empty for $\mathcal{H}^n$ a.e. y. Here we get weaker, but still nontrivial, conclusions with weaker hypotheses.

Proof. The fact that $\mathcal{H}^{N-n}(f^{-1}\{y\} \cap C) = 0$ for a.e. y follows directly from the coarea formula:

$$\int_{\mathbb{R}^n} \mathcal{H}^{N-n}(f^{-1}\{y\} \cap C) \, d\mathcal{H}^n(y) \ = \int_C Jf(x) \, d\mathcal{H}^N(x) = 0.$$

The fact that $f^{-1}\{y\} \setminus C$ is a C^1 submanifold is a consequence of the Implicit Function Theorem. $\square$