The Change of Variables Formula

Rich Schwartz

March 17, 2015

1 The Result

Consider the following data.

1. U and V are open subsets in $\boldsymbol{R}^n$.
2. $F : U \to V$ is a diffeomorphism.
3. $f : V \to \boldsymbol{R}$ be a continuous function.
4. $K \subset V$ is a compact set.

The purpose of these notes is to give a self-contained proof of the following result.

$$\int_K f \; dV = \int_{F^{-1}(K)} (f \circ F) \; \det(dF). \tag{1}$$

The result also holds when f is just Lebesgue measurable. But, this result requires some auxiliary results from measure theory, like the monotone convergence theorem. The special case when f is continuous suffices for all applications in the class, because these have to do with integrating smooth differential forms on manifolds.

I'll prove the result through a series of steps, each treating a more general case.

2 Step 1

The case when K is a cube and F is a linear transformation and f is a constant function just boils down to the determinant.

3 Step 2

Let's prove this result when K is a cube and f is a constant function. If $f = 0$ then both integrals are obviously 0. So, we can scale so that $f = 1$. Introduce the function

$$J(K, F) = \frac{\int_{F^{-1}(K)} \det(DF)}{\mu(K)}. \tag{2}$$

Equation 1 is equivalent to the statement that $J(K, F) = 1$.

Suppose that there is some $b > 0$ such that $J(K, F) > 1 + b$. Then, for every $\epsilon > 0$, there is some sub-cube $K' \subset K$ such that the side length of K' is less than ϵ and $J(K', F) > 1 + b$. This comes from the additivity the integral. If the ratio were near 1 on all small scales, it would also be near one on the large scale.

However, once ϵ is sufficiently small, the restriction of F to K' is nearly a linear map, and the ratio $J(K', F)$ must converge to 1. This is a contradiction. The same argument shows that there cannot be any $b > 0$ so that $J(K, F) < 1 - b$.

These two cases combine to show that $J(K, F) = 1$.

4 Step 3

Suppose that K is a cube and f is continuous. This time define

$$J(K, F, f) = \frac{\int_{F^{-1}(K)} (f \circ F) \det(DF)}{\int_K f} \tag{3}$$

The same argument as in Step 2 works here. The point is that the restriction of f to a small cube $K' \subset K$ is nearly constant. So, up to an error which vanishes as $\epsilon > 0$ we are back in the constant function case.

5 Step 4

Say that K is *approximable by cubes* if, for every $\epsilon > 0$, there is some finite collection $Q_1, ..., Q_m$ of almost disjoint cubes (with m depending on ϵ) so that

$$K \subset Q = \bigcup_{i=1}^{m} Q_i, \qquad \mu(K) > \sum_{i=1}^{m} \mu(Q_i) - \epsilon. \tag{4}$$

Here μ denotes Lebesgue measure and *almost disjoint* means that the interiors are disjoint.

Now I'll prove the result assuming that K is approximable by cubes. Once ϵ is sufficiently small, we have $Q \subset V$. By compactness, there is some upper bound C_1 for the restriction of $|f|$ to Q. Hence

$$\Big|\int_Q f - \int_K f\Big| < C_1\epsilon. \tag{5}$$

Since F is a diffeomorphism and Q is compact, there exists some constant C_2' such that the restriction of F^{-1} to Q expands distances by a factor of C_2' and hence volume by at most $C_2 = (C_2')^n$. Hence

$$\mu(F^{-1}(Q-K)) < C_2\epsilon. \tag{6}$$

By compactness again, there is a constant C_3 so that the restriction of $|\det(DF)|$ to $F^{-1}(Q)$ is at most C_3. Hence

$$\Big|\int_{F^{-1}(Q)} (f\circ F)\ \det(DF) - \int_{F^{-1}(K)} (f\circ F)\ \det(DF)\Big| < C_1C_2C_3\epsilon. \tag{7}$$

Since the cubes $Q_1, ..., Q_m$ are almost disjoint, we have

$$\sum_{i=1}^m \int_{Q_i} f = \int_Q f. \tag{8}$$

$$\sum_{i=1}^m \int_{F^{-1}(Q_i)} (f\circ F)\ \det(DF) = \int_{F^{-1}(Q)} (f\circ F)\ \det(DF). \tag{9}$$

Since Equation 1 is true for individual cubes, it is also true for finite sums of cubes, as in the set Q. But then Equations 5 and 7 tell us that Equation 1 holds for K up to an error of $C_4\epsilon$, where $C_4 = C_1 + C_1C_2C_3$. But ϵ is artibrary. Hence Equation 1 holds for K.

6 Step 5

Now we show that every compact $K \subset V$ is approximable by cubes. Without loss of generality, we can assume that $K \subset [0,1]^n$. For notational convenience, set $X = [0,1]^n$. Say that a *dyadic interval* is an interval whose endpoints are rational numbers of the form $k/2^m$ for integers k and m. Say that a *dyadic cube* is the product of dyadic intervals which all have the same length. The set of centers of dyadic cubes is dense in $\boldsymbol{R}^n$ and also the set of possible diameters of such cubes is dense. For this reason, $X - K$ is the countable union of dyadic cubes.

Let $P_1, P_2, P_3, ...$ be this infinite collection. We have

$$\sum_{i=1}^{\infty} \mu(P_i) = \mu(X - K. \tag{10}$$

Setting

$$P^\ell = \bigcup_{i=1}^{\ell} P_i, \tag{11}$$

we have

$$\lim_{\ell \to \infty} \mu(P^\ell) = \mu(X - K), \qquad K \subset X - P^\ell. \tag{12}$$

Given ϵ, we can choose ℓ so that $\mu(X - K) < \mu(P^\ell) + \epsilon$. Using the fact that

$$\mu(K) + \mu(X - K) = 1 = \mu(P^\ell) + \mu(X - P^\ell) \tag{13}$$

we see that

$$\mu(K) > \mu(X - P^\ell) - \epsilon. \tag{14}$$

But $X - P^\ell$ is a finite union of almost disjoint cubes, say $Q_1, ..., Q_m$. The way to see this is that we can scale up the whole picture by some power of 2 so that every cube in sight has integer coordinates. Then the set of interest to us is tiled by integer cubes.