NOTES ON DIFFERENTIAL FORMS. PART 4: INTEGRATION

1. The whole is the sum of the parts

Before we go about making sense of integrating forms over manifolds, we need to understand what integrating functions over $\mathbb{R}^n$ actually means. When somebody writes

$$\int_0^3 e^x dx$$

or

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx\, dy$$

or

$$\int_R f(x) d^n x,$$

what is actually being computed?

The simplest case is in $\mathbb{R}$. When we write $\displaystyle\int_a^b f(x)dx$, we have a quantity with density $f(x)$ spread out over the interval $[a, b]$. We imagine breaking that interval into small sub-intervals $[x_0, x_1]$, $[x_1, x_2]$, up to $[x_{N-1}, x_N]$, where $a = x_0$ and $b = x_N$. We then have

$$\begin{aligned} \int_a^b f(x)dx &= \text{Amount of stuff in } [a,b] \\ &= \sum_{k=1}^N \text{Amount of stuff in } [x_{k-1}, x_k] \\ &\approx \sum_{k=1}^N f(x_k^*)\Delta_k x, \qquad (1) \end{aligned}$$

where $\Delta_k x = x_k - x_{k-1}$ is the length of the kth interval, and x_k^* is an arbitrarily chosen point in the kth interval. As long as f is continuous and each interval is small, all values of $f(x)$ in the kth interval are close to $f(x_k^*)$, so $f(x_k^*)\Delta_k x$ is a good approximation to the amount of stuff in the kth interval. As $N \to \infty$ and the intervals are chosen smaller and smaller, the errors go to zero, and we have

$$\int_a^b f(x)dx = \lim_{N\to\infty} \sum_{k=1}^N f(x_k^*)\Delta_k x.$$

Note that I have not required that all of the intervals $[x_{k-1}, x_k]$ be the same size! While that's convenient, it's not actually necessary. All we need for convergence is for all of the sizes to go to zero in the $N \to \infty$ limit.

Date: April 15, 2016.

The same idea goes in higher dimensions, when we want to integrate any continuous bounded function over any bounded region. We break the region into tiny pieces, estimate the contribution of each piece, and add up the contributions. As the pieces are chosen smaller and smaller, the errors in our estimates go to zero, and the limit of our sum is our exact integral.

If we want to integrate an unbounded function, or integrate over an unbounded region, we break things up into bounded pieces and add up the integrals over the (infinitely many) pieces. A function is (absolutely) integrable if the pieces add up to a finite sum, no matter how we slice up the pieces. Calculus books sometimes distinguish between "Type I" improper integrals like $\int_1^\infty x^{-3/2}dx$ and "Type II" improper integrals like $\int_0^1 y^{-1/2}dy$, but they are really the same. Just apply the change of variables $y = 1/x$:

$$\begin{aligned} \int_1^\infty x^{-3/2}dx &= \sum_{k=1}^\infty \int_k^{k+1} x^{-3/2}dx \\ &= \sum_{k=1}^\infty \int_{1/(k+1)}^{1/k} y^{-1/2}dy \\ &= \int_0^1 y^{-1/2}dy. \end{aligned} \tag{2}$$

When doing such a change of variables, the width of the intervals can change drastically. Δy is not Δx, and x-intervals of size 1 turn into y-intervals of size $\frac{1}{k(k+1)}$. Likewise, the integrand is not the same. However, the contribution of the interval, whether written as $x^{-3/2}\Delta x$ or $y^{-1/2}\Delta y$, *is* the same (at least in the limit of small intervals).

In other words, we need to stop thinking about $f(x)$ and dx separately, and think instead of the combination $f(x)dx$, which is a machine for extracting the contribution of each small interval.

But that's exactly what the differential form $f(x)dx$ is for! In one dimension, the covector dx just gives the value of a vector in $\mathbb{R}^1$. If we evaluate $f(x)dx$ at a sample point x_k^* and apply it to the vector $x_k - x_{k-1}$, we get

$$f(x)dx(\vec{x}_k - \vec{x}_{k-1}) = f(x_k^*)\Delta_k x.$$

2. Integrals in 2 or More Dimensions

Likewise, let's try to interpret the integral of $f(x, y)dxdy$ over a rectangle $R = [a, b] \times [c, d]$ in $\mathbb{R}^2$. The usual approach is to break the interval $[a, b]$ into N pieces and the interval $[c, d]$ into M pieces, and hence the rectangle R into NM little rectangles with vertices at (x_{i-1}, y_{j-1}), (x_i, y_{j-1}), (x_{i-1}, y_j) and (x_i, y_j), where $i = 1, \ldots, N$ and $j = 1, \ldots, M$.

So what is the contribution of the (i, j)-th sub-rectangle R_{ij}? We evaluate $f(x, y)$ at a sample point (x_i^*, y_j^*) and multiply by the area of R_{ij}. However, that area is exactly what you get from applying $dx \wedge dy$ to the vectors $\vec{v}_1 = (x_i - x_{i-1}, 0)$ and $\vec{v}_2 = (0, y_j - y_{j-1})$ that span the sides of the rectangle. In other words, $f(x_i^*, y_j^*)\Delta_i x \Delta_j y$ is exactly what you get when you apply the 2-form $f(x,y)dx \wedge dy$ to the vectors $(\vec{v}_1, \vec{v}_2)$ at the point (x_i^*, y_j^*). [Note that this interpretation requires the normalization $C_{k,\ell} = \dfrac{(k+\ell)!}{k!\ell!}$ for wedge products. If we had used $C_{k,\ell} = 1$, as in Guillemin and Pollack, then $dx \wedge dy(\vec{v}_1, \vec{v}_2)$ would only be half the area of the rectangle.]

The same process works for integrals over any bounded domain R in $\mathbb{R}^n$. To compute $\int_R f(x) d^n x$:

(1) Break R into a large number of small pieces $\{R_I\}$, which we'll call "boxes", each of which is approximately a parallelpiped spanned by vectors $\vec{v}_1, \ldots, \vec{v}_n$, where the vectors don't have to be the same for different pieces.
(2) To get the contribution of a box R_I, pick a point $x_I^* \in R_I$, evaluate the n-form $f(x)dx^1 \wedge \cdots \wedge dx^n$ at x_I^*, and apply it to the vectors $\vec{v}_1, \ldots, \vec{v}_k$. Your answer will depend on the choice of x_I^*, but all choices will give approximately the same answer.
(3) Add up the contributions of all of the different boxes.
(4) Take a limit as the sizes of the boxes go to zero uniformly. Integrability means that this limit does not depend on the choices of the sample points x_I^*, or on the way that we defined the boxes. When f is continuous and bounded, this always works. When f is unbounded or discontinuous, or when R is unbounded, work is required to show that the limit is well-defined.

For instance, to integrate $e^{-(x^2+y^2)}dxdy$ over the unit disk, we need to break the disk into pieces. One way is to use Cartesian coordinates, where the boxes are rectangles aligned with the coordinate axes and of size $\Delta x \times \Delta y$. Another way is to use polar coordinates, where the boxes have r and θ ranging over small intervals.

Exercise 1: Let R_I be a "polar rectangle" whose vertices p_1, p_2, p_3 and p_4 have polar coordinates (r_0, θ_0), $(r_0 + \Delta r, \theta_0)$, $(r_0, \theta_0 + \Delta\theta)$ and $r_0 + \Delta r, \theta_0 + \Delta\theta)$, respectively, where we assume that Δr is much smaller than r_0 and that $\Delta\theta$ is small in absolute terms. Let $\vec{v}_1$ be the vector from p_1 to p_2 and $\vec{v}_2$ is the vector from p_1 to p_3.
(a) Compute $dx \wedge dy(\vec{v}_1, \vec{v}_2)$.
(b) If our sample point x_I^* has polar coordinates (r^*, θ^*), evaluate the approximate contribution of this box.
(c) Express the limit of the sum over all boxes as a double integral over r and θ.
(d) Evaluate this integral.

3. Integration Over Manifolds

Now let X be an oriented n-manifold (say, embedded in $\mathbb{R}^N$), and let α be an n-form. The integral $\int_X \alpha$ is the result of the following process.

(1) Break X into a number of boxes X_I, where each box can be approximated as a parallelpiped containing a point p_I^*, with the *oriented* collection of vectors $\vec{v}_1, \ldots, \vec{v}_n$ representing the edges.
(2) Evaluate α at p_I^* and apply it to the vectors $\vec{v}_1, \ldots, \vec{v}_n$.
(3) Add up the contributions of all the boxes.
(4) Take a limit as the size of the boxes goes to zero uniformly.

In practice, Step 1 is usually done via a parametrization ψ, and letting the box X_I be the image under ψ of an actual $\Delta x_1 \times \cdots \times \Delta x_n$ rectangle in $\mathbb{R}^n$, and setting $\vec{v}_i = d\psi_a(\Delta x_i \vec{e}_i)$, where $p_I^* = \psi(a)$. Note that p_I^* is not necessarily a vertex. It's just an arbitrary point in the box.

If the box is constructed in this way, then Step 2 is *exactly* the same as applying $\psi^*\alpha(a)$ to the vectors $\{\Delta x_i \vec{e}_i\}$. But that makes integrating α over X the same as integrating $\psi^*\alpha$ over $\mathbb{R}^n$! This shows directly that different choices of coordinates give the same integrals, as long as the coordinate patches are oriented correctly.

When a manifold consists of more than one coordinate patch, there are several things we can do. One is to break X into several large pieces, each within a coordinate patch, and then break each large piece into small coordinate-based boxes, exactly as described above. Another is to use a partition of unity to write $\alpha = \sum \rho_i \alpha$ as a sum of pieces supported in a single coordinate chart, and then integrate each α_i separately.

This allows for a number of natural constructions where forms are defined intrinsically rather than via coordinates.

Let X be an oriented $(n-1)$-manifold in $\mathbb{R}^n$, and let $\vec{n}(x)$ be the unit normal to X at x whose sign is chosen such that, for any oriented basis $\vec{v}_1, \ldots, \vec{v}_{n-1}$ of T_xX, the basis $(\vec{n}, \vec{v}_1, \ldots, \vec{v}_{n-1})$ of $T_x\mathbb{R}^n$ is positively oriented. (E.g, if $X = \partial Y$, then n is the normal pointing out from Y). Let $dV = dx^1 \wedge \cdots \wedge dx^n$ be the volume form on $\mathbb{R}^n$. Define a form ω on X by

$$\omega(\vec{v}_1, \ldots, \vec{v}_{n-1}) = dV(\vec{n}, \vec{v}_1, \ldots, \vec{v}_{n-1}).$$

Exercise 2: Show that $\int_X \omega$ is the $(n-1)$-dimensional volume of X.

More generally, let α be any k-form on a manifold X, and let $\vec{w}(x)$ be any vector field. We define a new $(k-1)$-form $i_w\alpha$ by

$$(i_w\alpha)(\vec{v}_1, \ldots, \vec{v}_{k-1}) = \alpha(\vec{w}, \vec{v}_1, \ldots, \vec{v}_{k-1}).$$

Exercise 3: Let S be a surface in $\mathbb{R}^3$ and let $\vec{v}(x)$ be a vector field. Show *directly* that $\int_S i_v(dx \wedge dy \wedge dz)$ is the flux of $\vec{v}$ through S. That is, show that $i_v(dx \wedge dy \wedge dz)$ applied to a pair of (small) vectors gives (approximately) the flux of $\vec{v}$ through a parallelogram spanned by those vectors.

Exercise 4: In $\mathbb{R}^3$ we have already seen $i_v(dx \wedge dy \wedge dz)$. What did we call it?

Exercise 5: Let $\vec{v}$ be any vector field in $\mathbb{R}^n$. Compute $d(i_v(dx^1 \wedge \cdots \wedge dx^n))$.

Exercise 6: Let $\alpha = \sum \alpha_I(x)dx^I$ be a k-form on $\mathbb{R}^n$ and let $\vec{v}(x) = \vec{e}_i$, the i-th standard basis vector for $\mathbb{R}^n$. Compute $d(i_v\alpha) + i_v(d\alpha)$. Generalize to the case where $\vec{v}$ is an arbitrary *constant* vector field.

When $\vec{v}$ is not constant, the expression $d(i_v\alpha) + i_v(d\alpha)$ is more complicated, and depends both on derivatives of v and derivatives of α_I, as we saw in the last two exercises. This quantity is called the *Lie derivative* of α with respect to $\vec{v}$.

It is certainly possible to feed more than one vector field to a k-form, thereby reducing its degree by more than 1. It immediately follows that $i_v i_w = -i_w i_v$ as a map $\Omega^k(X) \to \Omega^{k-2}(X)$.