Calculus on Manifolds

Riley Moriss

August 29, 2025

1 Defining Integration

Obviously we can define integration more abstractly on an arbitrary measure space but our interest is in manifolds so we will restrict to the Lebesgue measure on the reals. First we define the volume of a cube in $\mathbb{R}^n$ as one would expect, merely as the product of the lengths of its sides.

$$\mathrm{vol}\big([a_1, a_1+\epsilon] \times \cdots \times [a_n, a_n+\epsilon]\big) = \epsilon^n$$

The Lebesgue (outer) measure is defined for any set $E \subseteq \mathbb{R}^n$ by the infimum over the volume of all coverings of the set by (almost disjoint) cubes. A set is called measurable if for every $\epsilon > 0$ there is an open set such that $\mu(U - E) < \epsilon$.

The other key object is a differential form. Given a manifold then its cotangent bundle is the fiberwise dual of the tangent bundle (linear functions into $\mathbb{R}$). Given local charts on the manifold $(x_1, ..., x_n) : M \to \mathbb{R}^n$ then there is local coordinates on the tangent bundle $\left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}\right) : M \to TM$ and this induces local charts on the cotangent bundle $(dx_1, ..., dx_n) : M \to T^*M$ which are defined as

$$dx_i \frac{\partial}{\partial x_j} = \delta_{ij}$$

A one form is then a smooth section of this cotangent bundle, a general one form can be represented in local coordinates as

$$f_1 dx_1 + \cdots + f_n dx_n$$

where $f_i : M \to \mathbb{R}$ is a smooth function. Given the cotangent bundle then we can construct the bundle of k-forms by taking the fiberwise k-th exterior algebra (defined on any vector space). The exterior algebra is essentially the tensor algebra with some antisymmetry relations, and setting higher powers to zero. In local coordinates there is a basis given then by tensors or tuples of one forms.

1.1 Integration on $\mathbb{R}^n$

To define the integral of a (measurable) function $\mathbb{R}^n \to \mathbb{R}$ we will define it for simple functions and claim that any such function can be approximated by simple functions. A simple function on $\mathbb{R}^n$ is a linear combination of indicator functions on measurable sets.

$$\varphi(x) = \sum_{i=1}^{n} a_i 1(x \in E_i), \quad a_i \in \mathbb{R}$$

We define the integral of such a function as

$$\int_{\mathbb{R}^n} \sum_{i=1}^{n} a_i 1(x \in E_i) = \sum_{i=1}^{n} a_i \mu(E_i)$$

If f is a non-negative function then its integral is defined as the infimum of the integrals of simple functions that cover f.

$$\int f = \inf \left\{ \int \varphi : \varphi \text{ simple and } 0 \leq f \leq \varphi \right\}$$

where the function is leq to the other pointwise. For functions that have a negative part then we just break it into two functions, positive and negative (say by taking the max and min with the zero function) and then subtract the integral of the negative of the negative part

$$\int f = \int f_+ - \int f_-$$

The idea is clear, to find the area of an arbitrary function we cover it with step functions and take the smallest such covering. There are a lot of things to show to make all this work, most importantly that measurable functions can be approximated in this way.

1.2 Integration on Manifolds

On manifolds what we integrate are top forms. We can integrate lower degree forms over submanifolds on which they are top forms. Therefore it is sufficient to define the integration of top forms.

First consider a differential form on $\mathbb{R}^n$, we can take global charts and then a top form will in general be of the form

$$f(x) dx_1 \wedge \cdots \wedge dx_n$$

for a smooth $f : \mathbb{R}^n \to \mathbb{R}$. Then the integral of this form is defined as the integral of f

$$\int_{\mathbb{R}^n} f dx_1 \wedge \cdots \wedge dx_n \quad := \quad \int_{\mathbb{R}^n} f$$

where the right hand side is in the Lebesgue sense. For an abstract manifold M^n then we can define integration in a chart in this way. The final thing is to define it non-locally. We do this by fixing a partition of unity. Then in each local chart we have a well defined (one has to check) integral via the formula above and we define the integral over the total space to be the sum of the integrals over the partition

$$\int_M \omega = \sum_i \int_{U_i, \varphi_i} \varphi^* \omega = \sum_i \int_{\mathbb{R}^n} f_i$$

Remark. The point of the top form machinery is to create another bundle that is locally given by functions into $\mathbb{R}$ in a compatible way, that way we can integrate them. If we were to say take a global function $M \to \mathbb{R}$ and try to integrate it by picking charts and writing it in local coordinates it would not be well defined. See this post for a nice example.

Remark. For this setup to actually work the manifold must be orientable. Otherwise the integral is only defined up to a sign.

Remark. In the case of manifolds we integrate top forms which at least in local charts are given by a single smooth function. Notice that at least in a single chart this is measure theoretically a very strong condition to be smooth.

Remark. So far we have exclusively seen how to integrate functions whose codomain is the real numbers or in the case of manifolds the cotangent bundle, which at least locally is given by a function into the real numbers. Given a curve in a manifold $[0,1] \to M$ then we can find its so called arc length. Note here that the integral we take is also of a real valued function *constructed from the curve.* Thus this is not anthing different. The same is true for line integrals.

Remark. If we have a function $\mathbb{R}^n \to \mathbb{R}^m$ then it is given by m functions $\mathbb{R}^n \to \mathbb{R}$. We may then take the integral of the vector valued function by just integrating each of the functions and collecting them into a vector. Thus the integral of a vector valued function in this sense is another vector. A special case of this is complex valued functions which are essentially $\mathbb{R}^2$ and we treat them as such by integrating the imaginary and real parts seperately.

Remark. In the context of automorphic forms mostly we integrate $\mathbb{C}$ valued functions. On occasion we have need of integrals over functions valued in infinite dimensional vector spaces. The Gelfand–Pettis integral is used in such cases (or Bochner integral). However even in this case the it is merely a construction to give a real valued function (via a pairing) that can then be Lebesgue integrated.

2 Total and Partial Derivatives

2.1 Derivatives on $\mathbb{R}^n$

The idea of the derivative is to find "the best linear approximation" to a function. Here we will consider all functions to be smooth. First consider a function $\mathbb{R} \to \mathbb{R}$ then we know that its derivative is given by the function

$$\frac{d}{dx}f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This derivative function defines the slope of the tangent line at any given point, and so the *equation* of this tangent line at any given point x_0 is given by (using slope intercept form)

$$df(x) = \left(\frac{d}{dx}f(x_0)\right)(x - x_0) + f(x_0)$$

Which is the linear approximation of f at the point x_0. Alternitively by centering the function at the origin we can approximate it not just be a line but by a *linear function* $\mathbb{R} \to \mathbb{R}$ given by $\frac{d}{dx}f(x_0)x$. Thus there are two senses in which we can think of this as a linear approximation, as a tangent line or as a linear function.

If we have a function $\mathbb{R} \to \mathbb{R}^n$ then its derivative is just the derivative of the component functions, as it can be expressed as n functions $\mathbb{R} \to \mathbb{R}$. Using the vector addition and multiplication by scalars this is seen to be exactly the same formula above where $f(x)$ is now some vector. This gives us a line in $\mathbb{R}^n$ by taking the exact same equation as above too.

Functions $\mathbb{R}^n \to \mathbb{R}$ again define functions from $\mathbb{R} \to \mathbb{R}$ in a less canonical way, by fixing $n-1$ of the variables. So at a point $(x_0, ..., x_n)$ the derivative in the x_i direction is denoted

$$\frac{\partial}{\partial x_i} f(x_0, ..., x_n) = \lim_{h \to 0} \frac{f((x_0, ..., x_n) + he_i) - f(x_0, ..., x_n)}{h}$$

where $e_i = (0, ..., 0, 1, 0, ..., 0)$ is the $i-th$ standard basis vector. We again have a function into $\mathbb{R}$ and so our linear approximation should be a line in $\mathbb{R}$. It will be the tangent to the hypersurface in the x_i direction. Our equation for df above still gives the equation for this line by fixing $n-1$ of the entries it is the equation of a line $\mathbb{R} \to \mathbb{R}$. Alternitively by taking the span of these lines we get a subspace of $\mathbb{R}^{n+1}$ that can be considered as a hyperplane to the *graph* of the function.

Finally functions $\mathbb{R}^n \to \mathbb{R}^m$ combine the last two generalisations. First we consider it as a colletion of m functions $\mathbb{R}^n \to \mathbb{R}$ and then we consider the partial derivatives of these functions in all directions. There is no clear way of interpreting this as a tangent line because there are multiple parametrers, it is more like the approximation of the m dimensional surface by an n dimensional hyperplane; taking the span of the n partial derivatives gives a plane for each function into $\mathbb{R}$ and there are m such planes approximating the function in each direction. We could have considered the function as n functions $\mathbb{R} \to \mathbb{R}^m$ howevever we claim that this should give the same matrix. In the end we have lots of derivatives and the only *clean* way of expressing it now is as df that is the linear (in the sense of vector spaces) approximation, if $f = (f_1, ..., f_m$ then

$$df = \left(\frac{\partial}{\partial x_i} f_j \right)_{ij}$$

This is a matrix that summarises all the partial derivatives, however it also defines a linear function $\mathbb{R}^n \to \mathbb{R}^m$ which approximates the original in the same way as the other linear functions did. If we were to use it as the gradient of a hyper plane and then translate that plane to the relevant point we would get the required tangent plane thing. Note this is sometimes called the "total derivative" of the function.

2.2 Differentials on Manifolds

Over $\mathbb{R}$ we saw that the derivative gives us the linear approximation in many senses, and we can go between them using basic point line geometry. The easiest to write down however was always the linear approximation in the sense of linear algebra, that is a linear function $\mathbb{R}^n \to \mathbb{R}^m$. In the context of differential geometry it is no different, by talking about embeddings of submanifolds one can regain the explit equations of the tangent lines as linear subspaces, however this is cumbersome. Much easier to work with is often the embedding free definition.

Thus if we have a map between two manifolds $f : N \to M$ then we define its derivative as the best linear approximation in the sense of linear algebra as the smooth map on the only vector spaces around $df : TN \to TM$. It has a natural definition for whatever definition of the tangent bundle you choose, however the key point is that in local charts f can be written as a function $\mathbb{R}^n \to \mathbb{R}^m$ and its derivative is the matrix given above.

Remark. Another notion of derivative is given by distributions. One can in this setting differentiate more than just smooth functions but only "in the sense of distributions". Note that functions often embed into the relevant space of distributions and so we can consider this as a generalisation.

3 Fundamental Results of Calculus

3.1 Stokes Theorem

Consider A manifold with boundary M^n and a differential $n-1$ form on M call it ω. Then $d\omega$ the exterior derivative is a top form on M and we have

$$\int_M d\omega = \int_{\partial M} \omega$$

A special case of this is if we consider a differential top form on $\mathbb{R}$, $f(x)dx$ then it restricts to a top form on a submanifolds with boundary given by closed intervals $[a, b]$ Stokes theorem then implies that

$$\int_{[a,b]} f(x)dx = \int_{\{a,b\}} f(x) = f(b) - f(a)$$

where we have used the fact that f is a zero form and that $d(f) = f dx$.

3.2 Laplacians

Given a function $\mathbb{R}^n \to \mathbb{R}$ then it has several partial derivatives. These can be summarised by the ∇ operator given by

$$\nabla = \left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n} \right)$$

By considering this as the degenerate $m = 1$ case of a map $\mathbb{R}^n \to \mathbb{R}^m$ then we see that

$$\nabla f = df$$

as we have defined above. That is applying ∇ is exactly getting its total derivative. Note that this is now a vector field $\mathbb{R}^n \to \mathbb{R}^n$!

Taking $\Delta := \nabla^2 = \nabla \cdot \nabla$ we get the Laplacian of a scalar function. This operator is informing how much nearby points are influcing the behaviour at a given point. An important point of view is via the divergence of a vector field, that is a function $F : \mathbb{R}^n \to \mathbb{R}^n$,

$$\text{div}(F) = \nabla \cdot F$$

and so we see that $\Delta(f) = \text{div}(\nabla f)$.

This definition generalises to function between Riemannian manifolds (manifolds with a nice metric) via the Riemannian divergence and gradient operations. The idea is to reformulate the definition of the divergence and grad operator in terms of the inner product on $\mathbb{R}^n$, this then generalises by just taking the inner product given by the metric.

4 PDE's

According to [Eva10] a PDE is an expression of the form

$$F\big(D^k u(x), ..., D^1 u(x), u(x), x\big) = 0, \quad x \in U$$

for $U \subseteq \mathbb{R}^n$ an open subset and where F is a function

$$F : \mathbb{R}^{n^k} \times \cdots \mathbb{R}^{n^1} \times \mathbb{R} \times U \to \mathbb{R}$$

The D^k are the k-th partial derivatives of u, from our discussion above given $u : \mathbb{R}^n \to \mathbb{R}$ there are n first partial derivatives, and as we are taking k-th partial derivatives we can have permutations, possibly n^k many derivatives, if they dont commute with one another (you need to pick one of n derivatives k times). *Solving* the PDE is finding all $u : U \to \mathbb{R}$ that satisfy the equation above.

If we want to consider functions with codomain $\mathbb{R}^m$ then we need a setup of the same form but now u and the 0 are vectors. F will have to therefore intake more data as there are even more partial derivatives than before

$$F : \mathbb{R}^{mn^k} \times \cdots \mathbb{R}^{mn^1} \times \mathbb{R}^m \times U \to \mathbb{R}^m$$

and we are looking for solutions $u : U \to \mathbb{R}^m = (u_1, ..., u_m)$. Notice that we can consider each component of u seperately and therefore we have m PDE's in the first sense of functions $U \to \mathbb{R}$ and we look for compatible solutions. Therefore solving higher dimensional PDE's immediately reduces to considering just those for functions into $\mathbb{R}$.

Recall that over a manifold M we have its tangent bundle, TM. The fibers of this bundle are given by spaces of derivations, that is maps $\mathbb{C}^\infty(M) \to \mathbb{R}$ satisfying the Liebniz rule at the point over which they fiber

$$T_xM = \{\nu : \mathbb{C}^\infty(M) \to \mathbb{R} | \nu(fg)(x) = \nu(f)g(x) + f(x)\nu(g)\}$$

clearly then if we define derivation as a function satisfying this Leiniz rule at every point then the tangent bundle is just the space of (global) derivations. It is clear then that a vector field is a section of this bundle of derivations and so itself can be considered as a derivation by just permuting variables or currying; if $X : M \to TM$ then it defines a function $X : C^\infty(M) \to C^\infty(M)$ given by

$$X(f)(x) = X(x)(f)$$

Vector fields are therefore the global analogues of derivations, and this defines an isomorphism $TM \cong \mathrm{Der}(M)$. In local coordinates they are indeed simply

$$X = \sum_i a_i \frac{\partial}{\partial x_i}$$

Thus to get a (linear) PDE for a function $f \in \mathbb{C}^\infty(M)$ we can take an equation of the form

$$X_k \cdots X_1 f(x) = 0$$

More generally we just consider functions of these types of things as above, note that it can be more complicated as there are more vector fields than differential operators, so the F needs to be a function taking in some vector fields and combining them in some way; if we denote $\mathrm{Der}^k(M)$ the $\mathbb{R}$ span of the space of k-fold compositions of derivations (which notice themselves will almost never be derivations) then we want some function

$$F : \mathrm{Der}^k(M) \times \cdots \mathrm{Der}^1(M) \times C^\infty(M) \to C^\infty(M)$$

and we want to solve an equation of the form

$$F(X_1, ..., X_p, f) = 0$$

where 0 is the constant map $M \to \mathbb{R}$.

The next step of considering systems of these equations as equations for functions in $C^\infty(M)^n$ or as functions into $\mathbb{R}^n$ goes as one expects, instead of applying derivations in TM we look at $(TM)^n$ and just do everything component wise. In differential geometry we can usually replace $\mathbb{R}^n$ with a vector bundle over a space, here the relevant realisation is that $C^\infty(M) \cong \Gamma_M(\underline{\mathbb{R}})$, that is the space of (smooth global sections of the trivial line bundle over M. Then the multidimensional case is also clearly given by $\Gamma_M(\underline{\mathbb{R}^n})$. Our operators were defined on $C^\infty(M) \to C^\infty(M)$, that is from one space of sections to another and so the direction we will generalise is to consider two vector bundles $E, F \to M$ and we will want to look at linear operators

$$\Gamma(E) \to \Gamma(F)$$

Previously our operators were simply vector fields or combinations of them (our F above if we fixed some vector fields). Note that these spaces of sections are actually $C^\infty(M)$ modules by pointwise multiplication, we will define *differential operators* as linear operators that behave in a certain way with this module structure, following a Leibniz rule type thing. Precisely we call the set of k-th order partial differential opperators as the kernel of the adjoint representation composed $k+1$ times,

$$PDO^k = \ker \mathrm{ad} \;:=\; \{T \text{ linear operator } : [f, T] = 0 \quad \forall f \in C^\infty(M)\}$$

Locally every vector bundle is trivial so even in this setup we have that we are locally looking at maps from $M \to \mathbb{R}^n$, and this definition ensures that the normal partial derivatives are contained in the class of functions [Nic, Cor 10.1.5].

Remark. This setup is of course not realistic, we would need to put restrictions on both the form of the PDE (for example making it linear) and the functions we are looking (regularity conditions as well as boundary conditions) for to make the problem actually tractable.

Remark. If we have a function $\mathbb{R}^n \to \mathbb{R}$ then it has a vector of (first) partial derivatives. It has a matrix of k-th partial derivatives. If we have a function $\mathbb{R}^n \to \mathbb{R}^m$ then it has a matrix of (first) partial derivatives, but has a cube of k-th partial derivatives.

Remark. All of this can be done in a "weak sense", that is using distributions instead of functions as above.

References

[Eva10] Lawrence C. Evans. *Partial Differential Equations*. Amer Mathematical Society, Providence, R.I, 2nd edition edition, March 2010.

[Nic] Liviu I Nicolaescu. Lectures on the Geometry of Manifolds.