Calculus on Manifolds

Talha Ashraf

Abstract

In this paper, I will discuss differential forms as the language of the generalised Stokes' theorem. Through this venture, we will generalise the procedure of integration and differentiation to $\mathbb{R}^N$ as well as to performing these procedures on smooth Manifolds.

1 Introduction

In a Multivariable Calculus class we learn way too many integral theorems. We start out by defining gradient,divergence and curl as

$$\text{grad}\, V = \frac{dV}{dx}\hat{x} + \frac{dV}{dy}\hat{y} + \frac{dV}{dz}\hat{z}$$
$$\text{div}\, V = \frac{dV}{dx} + \frac{dV}{dy} + \frac{dV}{dz}$$
$$\text{curl}\ \langle f, g, h \rangle = \Big(\frac{dh}{dy} - \frac{dg}{dz}\Big)\hat{x} + \Big(\frac{df}{dz} - \frac{dh}{dx}\Big)\hat{y} + \Big(\frac{dg}{dx} - \frac{df}{d}y\Big)\hat{z}.$$

Then we cumbersomely learn several integral theorems:

c $\int_C \nabla f \cdot \mathbf{r}' dt = f(\mathbf{b}) - f(\mathbf{a})$	Fundamental theorem of Calculus
$\int_S \Big(\frac{dQ}{dx} - \frac{dP}{dy}\Big) dS = \int_C P dx + Q dy$	Green's theorem
$\int_S \mathbf{F} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{F} dV$	Divergence theorem
$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_S \nabla \times \mathbf{F} \cdot d\mathbf{S}$	3D Stokes' theorem

These integral theorems are not independent from one another. They are actually special cases of a much more fundamental theorem: the generalised Stokes' Theorem. This is a beautiful result with huge relavance to the problem of integrating in higher dimensions.

The Stokes' theorem, in its most general form is

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$$

and it is written in the language of differential forms. These are mathematical objects that provide us insight into our integration procedures because they can automatically encode information such as orientation and the translation of functions. You may have noticed in the definitions of Green's, Divergence and the 3D Stokes' theorem that gradient divergence and curl are used constantly in the integral theorems. Keeping track of when to do what type of "differentiation method" is annoying but differential forms automatically track this for us. By this I mean that differetnial forms keep track of which procedure to use (gradient, divergence or curl) based on which integral situation we are in automatically. This allows the generalisation of integration theorems beautifully to higher than 3 dimensions.

I will start this paper by defining Manifolds and the procedure of differentiation on them. Then we will love on to defining a differential form and seeing how divergence and curl show up from lower dimensional forms. Finally, we will state the Stokes' theorem and see how the other integral theorems are derived from it.

2 Manifolds

It is well known how to differentiate and integrate in $\mathbb{R}$ and it is fairly straight forward to extend our familiar laws of calculus to $\mathbb{R}^2$, $\mathbb{R}^3$ and even $\mathbb{R}^n$. But we need not stop here. With some effort, it is possible to extend the bounds of Calculus and talk about integration and differentiation on not just Euclidean space but any space that looks locally Euclidean.

A Manifold is any space where if you zoom far enough into any point, it looks like a Euclidean space around your point. We grapple with this idea everyday. If you go to the beach and look around, everything looks flat, almost like you are standing on top of a plane in $\mathbb{R}^2$. So local points on our earth look like a 2 dimensional plane eventhough the Earth itself is a space in $\mathbb{R}^3$. This planar behaviour does not hold up if you zoom out far enough to notice the Earth's curvature, however. In other words, while a Manifold looks Euclidean locally, this does not mean that it will look the same *globally*.

Furthermore, to be classed as a Manifold, every point in the space must look like the same kind of Euclidean space. As an anti-example of a Manifold, consider the shape made by the letter "Y". If you focus on any one of the 3 lines making up the letter "Y", they just look like lines in $\mathbb{R}$. But this is not a manifold because of the point where these 3 lines join. Such intersections are not something that can be seen in $\mathbb{R}$ and hence the space around this intersection point does not locally look like $\mathbb{R}$.

With this intuition for Manifolds, we can construct a definition for these spaces. Since we want small regions on our Manifold to look like Euclidean space, we can set this on firm mathematical foundations by considering mappings from that Euclidean space to our Manifold. Consider a small open subset on our Manifold encompassing any point on it. If this locale of the point really looks like a Euclidean space $\mathbb{R}^n$, then we should be able to find a clean mapping from an open set of $\mathbb{R}^n$ onto this subset of our manifold. We would like this mapping to be one to one and onto. Additionally, since we soon plan to talk about performing differentiation on our Manifolds, the map had better be smooth. These ideas are summed up in definitions 1 and 2.

DEFINITION 1. A map between two subsets X, and Y of two Euclidian spaces $f : X \to Y$ is diffeomorphic if it is one-to-one and onto and if both $f : X \to Y$ and its inverse $f^{-1} : Y \to X$ are smooth.

DEFINITION 2. Suppose X is a subset of a larger Euclidean space R^n. X is said to be a m-dimensional Manifold if it is locally diffeommorphic to $\mathbb{R}^m$, meaning every point $x \in X$ possesses a neighbourhood U which is diffeomorphic to an open subset $V \in \mathbb{R}^m$.

3 Differentiation on Manifolds

We can define the differentiation of a single-valued function $f : \mathbb{R} \to \mathbb{R}$ as

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

For a vector valued function $f : \mathbb{R}^n \to \mathbb{R}^m$, the derivative can have different values depending on the direction we move on along it. Hence, instead of using a scalar "h" in the definition, we can generalise by considering a vector $\mathbf{h} \in \mathbb{R}^n$ in the direction we are differentiating in to get

$$df_x(\mathbf{h}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x})}{t}. \tag{1}$$

There are two important things to note here. Firstly, the derivative has a dependence on the vector $\mathbf{h}$, and secondly we know that derivatives are linear operations. Hence differentiation of a function f can really be categorised as a linear transformation $df_x : \mathbf{h} \in \mathbb{R}^n \to \mathbb{R}^m$.

Now, we get to the problem of defining a notion of differentiation on a Manifold. For this, we imagine we have a parameterization $\phi : U \to X$ where X is an m-dimensional manifold in $\mathbb{R}^n$ and U is an open subset of $\mathbb{R}^m$. We know that derivatives of a surface in 3 dimensions lie on the tangent plane: a plane defined by the linear combination of vectors tangent to the plane at a point x_0, but with the vectors displaced from the origin by $f(x_0)$.

A surface is really just a 2 dimensional Manifold embedded in 3 dimensions and this idea of a tangent plane extends seamlessly to higher dimensions. From this we can conclude that differentiation on a Manifold just corresponds to

$$d_X(\mathbf{h}) = d\,\phi(\mathbf{h}).$$

4 Basic Properties of Differential Forms

In elementary Calculus, we talk about objects like dx, dy, dz. These "differentials" are a cornerstone of the notation of integral calculus. But what are these objects really and what properties do they possess? To answer these questions we define differential forms. These are objects that transform in such a way that they encode a remarkable amount of information used in the integration process. Hence differential forms provide a pathway to devising a unified general approach to integrating on manifolds.

To be more specific, differential forms can encode orientation and any change of coordinates. Perhaps even more surprisingly, a general d operator may be defined on forms which does the exact right thing in the right time. By this, I mean the result of applying the d operator to a form can produce the operation of gradient, curl, divergence and so on exactly when each operation is appropriate. Hence forms can be used to greatly simplify and provide a general pathway to integration on Manifolds. Furthermore, differential forms are the language of the generalized Stokes theorem.

For now, we define abstract symbols that are written as dx^i. To be able to do anything interesting with these symbols, we must define a product on them denoted by the symbol $\wedge$. This product will have the usual associativity and distributive properties

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$$
$$(\alpha + \beta) \wedge \gamma = \alpha \wedge \gamma + \beta \wedge \gamma$$
$$(c\alpha) \wedge \beta = \alpha \wedge (c\beta).$$

However, it will also have the more unique property of anti-symmetry:

$$\alpha \wedge \beta = -\beta \wedge \alpha \tag{2}$$

This property actually encodes information about the orientation we will follow during integration. I will stall the explanation of orientation for when I define the different types of forms. For now, we must note that this anti-symmetry will be our main guiding principle in defining the actual operation of wedge products.

From antisymmetry we have that the wedge product of any form with itself must be zero since

$$\alpha \wedge \alpha = -\alpha \wedge \alpha$$

which is a property satisfied only by 0.

5 Wedge Products

The wedge product is more generally a product between a certain group of tensors. Tensors are real valued functions which take in vectors as outputs. Tensors are multilinear, as explained in definition 3. Why are we suddenly shifting focus from the objects we just labeled as differential forms to the whole new object of tensors? Well, we will soon see that a differential form may be described as a type of tensor. This will involve some geometric insight which we will discuss in the following section. Talking about tensors first allows us to come up with a rigorous definition of the wedge product which will be used constantly as we continue our discussion of forms.

DEFINITION 3. A p-tensor is a function $T : v_1, v_2, \ldots v_p \to \mathbb{R}$ on a group of p vectors $v_1, v_2 \ldots v_p \in V$ where V is some vector space. Tensors are multilinear maps, meaning that they are individually linear in each variable:

$$T(v_1, v_2 \ldots v_k + \alpha v_k', \ldots v_p) = T(v_1, v_2 \ldots v_k, \ldots v_p) + \alpha T(v_1, v_2 \ldots v_k', \ldots v_p).$$

If we are given two tensors, we can define a tensor product between them that creates a larger tensor. More specifically, given a p-tensor and a q-tensor, we can create a $(p+q)$-tensor by defining the tensor product $\otimes$ as in definition .

DEFINITION 4. Given a p-tensor T_p and a q-tensor T_q, we define their tensor product applied on a set of $p+q$ vectors $(v_1, v_2 \dots v_p, v_{p+1}, v_{p+2}, \dots v_{p+q})$ as

$$T_p \otimes T_q = T_p(v_1, v_2, \dots v_p) \cdot T_q(v_{p+1}, v_{p+2}, \dots v_{p+q}).$$

This tensor product is then a tensor itself that maps $V^{\times p+q}$ to $\mathbb{R}$. Anti-symmetry of a tensor is reached when interchanging the positions of two of its parameters produces a change of sign:

$$T(v_1, v_2, \dots v_i, \dots v_j, \dots v_n) = -T(v_1, v_2, \dots v_j, \dots v_i, \dots v_n).$$

Such a tensor is called an alternating tensor.

A key detail to notice is that if T_p and T_q are two tensors, then their tensor product is normally not antisymmetric. This means that a general tensor cannot be a candidate for an object to rigorously define forms, as the product of two forms should be a third anti-symmetric form. However, there exists an operation that can take a non-alternating tensor and convert it into an alternating one.

In order for a tensor to be antisymmetric, everytime we swap one of its variables, there must be a flip of signs. Let's consider the set of all permutations of these parameters S_π. Elements $\pi \in S_\pi$ are permutations of the parameters. Let us also define $\mathrm{sgn}(\pi)$ as a function that returns 1 if going from the original permutation to π requires an even number of swaps and -1 for an odd number of swaps.

Then, consider we have an alternating tensor T and a tensor T^π which is the result of swapping its variables until they are in the permutation π. We should see that

$$T^\pi(v_{\pi_1}, v_{\pi_2}, \dots v_{\pi_p}) = (-1)^{\mathrm{sgn}(\pi)} T(v_1, v_2, \dots v_p). \tag{3}$$

Alternately, if T is a non-alternating tensor, then we can define an operation to produce an alternating tensor from it:

$$\mathrm{Alt}(T) = \frac{1}{p!} \sum_{\pi \in S_\pi} (-1)^{\mathrm{sgn}(\pi)} T^\pi. \tag{4}$$

This tensor is alternating because every element of the sum is alternating. The reason to divide by $p!$ is that, when T is already alternating, we want that $\text{Alt}(T) = T$ and not $\text{Alt}(T) = p!\,T$.

With all of this in mind, we can finally define a product that results in an alternating tensor.

DEFINITION 5. The wedge product between two tensors T and T' is defined as

$$T \wedge T' = \text{Alt}(T \otimes T'). \tag{5}$$

6 Differential Forms

With all this out of the way, we can finally develop the powerful and elegant mathematical concept of forms.

A p-dimensional differential form, or just a "p-form", is an alternating p-tensor defined on the structure a manifold. More specifically, if p is a point on a manifold then a k-form is any alternating tensor acting on vectors in the tangent space at p.

DEFINITION 6. Let Λ^k be the set of all alternating k-tensors. Let $\mathbb{R}^n_p$ be the tangent space of a point p. Then a k-form is any tensor ω that follows $\omega \in \Lambda^k(\mathbb{R}^n_p)$.

With all this defined, we are ready to start going over the various types of forms and see how the "d" operator on forms returns the right "derivative" at the right time. We will start with the zero form.

6.1 0-Forms and Orientation

A 0-form is a simple 1-dimensional function (with the antisymmetric property described before). If f is a 0-form, then at a point p on the manifold we may define

$$\int_p f = f(p). \tag{6}$$

This notation is different from what we ordinarily see in Calculus where we may right

$$\int_a^b df = f(b) - f(a).$$

This different in notation exists as it will better allow us to generalise to higher forms.

Rather than just evaluating the integral in equation (6) at a single point, we may instead want to evaluate it between two boundary points that bound a curve in space. Let's denote a and b as the boundary points of a 1-dimensional manifold in $\mathbb{R}^n$. We know from experience with the fundamental theorem of calculus that

$$\int_a^b df = f(b) - f(a).$$

For our theory of differential forms to work, it must replicate this result. This can be achieved by defining an orientation for forms. For 0 forms, this means that for a positively oriented point P, we have $\int_P f = f(P)$, while for a negatively oriented point Q, we have $\int_Q f = -f(Q)$. The orientation really just describes what direction we are integrating along. For a line integral on a curve from point a to point b on a 1-dimensional manifold, integrating a function on this manifold from a to b should give the negative result of integrating from b to a:

$$\int_a^b df = -\int_b^a df.$$

To express this in the language of differential forms, we can say that the boundary of our 1-dimensional manifold M is $\partial M = b - a$ where the minus sign next to a just denotes that we will evaluate it with a negative orientation. Then we can define

$$\int_{\partial M} f = \int_{b-a} f = f(b) - f(a).$$

The idea of orientation generalises to higher dimensions. For a 1-dimensional manifold, the boundary was 0-dimensional, just two points in space. We could go from one boundary point to the other or vice versa and the integrals of these two paths were the negatives of eachother. Similarly, for a 2-dimensional manifold, the boundary is itself a curve in space. We can either integrate on the boundary clockwise or anti-clockwise, getting opposite results. Hence we may define integrating clockwise as the positive orienttation and anti-clockwise as the negative. On a 3-dimensional manifold, The orientation is based on whether you define the normal vector as pointing "inwards" or "outwards".

6.2 1-Forms and the d Operator

So far for 0-forms, we have defined some notation that connects forms to our everyday knowledge in integration. For 1-forms, we will define yet more notation, but then for 2-forms, we will see the notational machinery from 0 and 1 forms to come together to give us a beautiful result.

Let us say that ω is some 1-form. We define 1-forms as expressions of the form

$$\omega = f(x, y, z)dx + g(x, y, z)dy + h(x, y, z)dz. \tag{7}$$

We will also now define the differentiation operator d on forms. This operator takes a 0-form and transforms it intp a 1-form in the following way:

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \dots \frac{\partial f}{\partial x_n}dx_n \tag{8}$$

This is reminiscent of the chain rule. We will also define that

$$d(f dx) = df \wedge dx.$$

You should also much notice the similarity between equation (8) and the expression for gradient

$$\frac{\partial f}{\partial x_1}\hat{x}_1 + \frac{\partial f}{\partial x_2}\hat{x_2} + \dots \frac{\partial f}{\partial x_n}\hat{x}_n.$$

If we try to integrate on a 1-form, we are simply met with a familiar line integral:

$$\int \omega = \int \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \dots \frac{\partial f}{\partial x_n}dx_n.$$

6.3 2-Forms

For this subsection, I will restrict our analysis to a 2-dimensional manifold in $\mathbb{R}^3$. You should soon see why I'm doing this in the following results. We define a 2-form in 3 dimensions as

$$\omega = f\, dx \wedge dy + g\, dy \wedge dz + h\, dz \wedge dx.$$

The same d operator defined earlier takes a 1-form and transforms it into a 2-form.

$$\begin{aligned} d(fdx+gdy+hdz) =& df\wedge dx + dg\wedge dy + dh\wedge dz \\ =& \Big(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\Big)\wedge dx + \Big(\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy + \frac{\partial g}{\partial z}dz\Big)\wedge dy \\ &+ \Big(\frac{\partial h}{\partial x}dx + \frac{\partial h}{\partial y}dy + \frac{\partial h}{\partial z}dz\Big)\wedge dz \\ =& \Big(\frac{dh}{dy} - \frac{dg}{dz}\Big)dy\wedge dz + \Big(\frac{df}{dz} - \frac{dh}{dx}\Big)dz\wedge dx + \Big(\frac{dg}{dx} - \frac{df}{d}y\Big)dy\wedge dx \end{aligned}$$

But this expression is just completely analogous to the curl of a vector field (f, g, h) with the unit vectors replaced by the differential wedge products. Moreover, we can take the wedge product of two 1-forms to get

$$\begin{aligned} (Adx+Bdy+Cdz)\wedge(Ddx+Edy+Fdz) =& (BF-CE)dy\wedge dz + (CD-AF)dz\wedge dx \\ &+ (AE-BD)dx\wedge dy. \end{aligned}$$

which is analogous to the cross product between two vectors in $\mathbb{R}^3$.

What is going on? Why are operations on forms producing the results of these familiar operations? As promised in the introduction, forms keep track of which "differentiaion" procedure to perform when. Finally defining Stokes' Theorem and seeings its cases in lower dimensions should show this.

7 Generalised Stokes' Theorem

DEFINITION 7. let M be a Manifold with boundary ∂M and let ω be a k-dimensional form. Then

$$\int_{\partial M}\omega = \int_M d\omega. \tag{9}$$

This is a beautiful theorem that applies to any finite dimensional space. As a end to this paper, however, let's restrict to 3 or fewer dimensions so that we can see how the other integral theorems fall out of Stokes' theorem as special cases.

We already showed how the fundamental theorem of Calculus follows from the definitions on forms.

Green's theorem can be derived by considering a 1-form $\omega = Pdx + Qdy$ on a 2 dimensional Manifold. Then

$$d\omega = \frac{\partial P}{dy} dy \wedge dx + \frac{\partial Q}{\partial x} dx \wedge dy = \Big(\frac{\partial Q}{dx} - \frac{\partial P}{dy}\Big) dx \wedge dy$$
$$\int_C P dx + Q dy = \int_S \Big(\frac{\partial Q}{dx} - \frac{\partial P}{dy}\Big) dS.$$

The form transforms under d like the 2 dimensional curl as required.

The divergence theorem can be similarly proved by considering a 2-form $\omega = f\,dx \wedge dy + g\,dy \wedge dz + h\,dz \wedge dx$ so that

$$d\omega = \Big(\frac{\partial f}{dx} + \frac{\partial g}{dy} + \frac{\partial h}{dz}\Big) dx \wedge dy \wedge dz$$
$$\int_S \langle f, g, h\rangle \cdot d\mathbf{S} = \int_V \nabla \cdot \langle f, g, h\rangle \, dV.$$

The form transforms under d like the divergence as required.

Finally, let us now assume $\omega = f dx + g dy + h dz$ is a 1-form to prove Stoke's theorem in 3 dimensions. Then as we saw, $d\omega$ is analogous to $\nabla \times \omega$.

$$\int_S d\omega = \int \int_S \nabla \times \langle f, g, h\rangle \cdot d\mathbf{S}$$
$$\int_{\partial S} \omega = \int_{\partial S} \langle f, g, h\rangle \cdot d\mathbf{r}$$
$$\therefore \int_C \mathbf{F} \cdot d\mathbf{r} = \int_S \nabla \times \mathbf{F} \cdot d\mathbf{S}$$

The form transforms under d like the curl as required.