

# Chapter 0

# Philosophy



# Chapter 1

## Differential Geometry

In calculus class you were taught the fundamental theorem, that the total difference of a function's value at the end of an interval from its value at the beginning is the sum of the infinitesimal changes in the function over the points of the interval:

$$\int_a^b f'(x)dx = f|_a^b \quad (1.1)$$

And later, in multivariable calculus, you encountered more elaborate integral formulae, such as the divergence theorem of Gauss:

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, dV = \int_S \mathbf{F} \cdot d\mathbf{S} \quad (1.2)$$

where  $\Omega$  is the volume of a 3D region we are integrating over, with infinitesimal volume element  $dV$  and  $S$  is the surface that forms the boundary of  $\Omega$ .  $dS$  then represents an infinitesimal parallelogram through which  $\mathbf{F}$  is flowing out, giving the flux integral on the right. Read in english, Gauss' divergence theorem says "Summing up the infinitesimal flux over every volume element of the region is the same as calculating the total flux coming out of the region". The total flux coming out of a region is the sum of its parts over the region. You might see that in english, this reads very similar to the description of the fundamental theorem of calculus.

Alongside this, there is Stokes' theorem for a 2D region. In english: summing up the infinitesimal amount of circulation of a vector field  $\mathbf{F}$  over every infinitesimal area is equal to calculating the total circulation of  $\mathbf{F}$  around the boundary of the region. In mathematical language:

$$\int_R \nabla \times \mathbf{F} \, dA = \int_C \mathbf{F} \cdot d\mathbf{r} \quad (1.3)$$

where  $R$  is our region and  $C$  is its boundary.

Perhaps now, the pattern is more evident. In all the above cases, summing up some *differential* of the function on the interior of some region is the same as summing up the function itself at the *boundary* of the region. All these theorems, that on their own look so strange to a first-year calculus student, are part of a much more general statement, the **General Stokes' Theorem**:

$$\int_{\Omega} d\omega = \int_{\partial\Omega} \omega. \quad (1.4)$$

Above,  $\omega$  is an object that will generalize both the “functions” and “vector fields” that you’ve seen in multivariable calculus, and  $d$  will generalize all the differential operators (gradient, divergence, curl) that you’ve dealt with. Lastly, when  $\Omega$  is the region in question  $\partial\Omega$  represents the *boundary* of the region  $\Omega$ . The fact that it looks like a derivative symbol is no coincidence, as we’ll see that the natural way to define the “derivative” of a region is as its boundary.

Through abstraction, we can reach results like this that not only look elegant and beautiful, but also provide us with insight into the natural way to view the objects that we’ve been working with for centuries. This gives us not only understanding of what language to use when studying mathematics, but also what is the natural language in which to describe the natural world. The general Stokes’ theorem is one of the first examples of this beautiful phenomenon, and this book will work to illustrate many more.

For the first half of this chapter, we will work towards giving the intuition behind this result. On our way, we will begin to slowly move into a much more general setting, beyond the 3-dimensional world in which most of multivariable calculus was taught. That doesn’t just mean we’ll be going into  $n$ -dimensional space. We’ll move outside of euclidean spaces that look like  $\mathbb{R}^n$ , into non-euclidean geometries. This will put into question what we really mean by the familiar concepts of “vector”, “derivative”, and “distance” as the bias towards Euclidean geometry no longer remains central in our minds. At its worst, the introduction of new concepts and notation will seem confusing and even unnecessary. At its best, it will open your mind away from the biases you’ve gained from growing up in a euclidean-looking world, and give you a glimpse of how modern mathematics *actually* looks.

Modern mathematics is learning that the earth isn’t flat. To someone who’s never had those thoughts, it is difficult to get used to, tiring, and sometimes even rage inducing, but to someone who has spent months

thinking and reflecting on it, it quickly becomes second nature. Far from being the study of numbers or circles, it is a systematic climb towards abstraction. It is a struggle towards creating one language, free from all-encompassing human bias, in order to try and describe a world that all other human languages, for so many centuries, have failed to grasp. It is humbling, and in the strangest of ways, it is profoundly beautiful.

## 1.1 The Derivative and the Boundary

Let's start working towards understanding Equation (1.4). First, let's work with what we've already seen to try and explore the relation between integrating within a region and integrating on the boundary.

If we are in one dimension, we have a function  $f$  defined on the interval  $[a, b]$ . Proving Equation (1.1) is much easier than you'd think. Let's take a bunch of steps:  $x_i = a + (b - a)i/N$ , so that  $x_0 = a, x_N = b$ . Then all we need is to form the telescoping sum:

$$\begin{aligned} f|_a^b &= f(x_N) - f(x_0) \\ &= \sum_{i=1}^N f(x_i) - f(x_{i-1}). \end{aligned}$$

If we make the number of steps  $N$  large enough, the stepsize shrinks so that in the limit, we get

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{i=1}^N f(x_i) - f(x_{i-1}) &= \lim_{N \rightarrow \infty} \sum_{i=1}^N \Delta f \\ &= \int_I df. \end{aligned}$$

Of course, the way its written more often is:

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N \frac{\Delta f}{\Delta x} \Delta x = \int_a^b \frac{df}{dx} dx.$$

What is the idea of what we've done? At each point we've taken a difference of  $f$  at that point with  $f$  at the preceding one. Because we're summing over all points, the sum of differences between neighboring points will lead to cancellation everywhere *except* at the boundary, where there

will not be further neighbors to cancel out the  $f(b)$  and  $f(a)$ . From this, we get Equation (1.1).

Notice something: It doesn't matter how we partition the interval, so long as that partition becomes infinitesimal as  $N \rightarrow \infty$ . The end result doesn't depend on the  $x_i$  at all! We are summing up the change of  $f$  over some interval, but it doesn't matter what coordinate system we use to describe this interval. It is *coordinate independent*. We chose to use  $x$  as our coordinate, going from  $a$  to  $b$ . This makes perfect physical sense. For example, if we had a temperature at each point in space, the temperature difference between two fixed points shouldn't depend on whether we use meters or feet to measure their distance apart.

Written mathematically:

$$\int_I df = \int_I \frac{df}{dx} dx = \int_I \frac{df}{du} du$$

This means that at a particular point of the integration, we get the familiar chain rule that we use in  $u$ -substitution: **DO WE WANT THIS HERE, OR LATER?**

$$\frac{df}{dx} dx = \frac{df}{du} du \Rightarrow \frac{df}{dx} = \frac{df}{du} \frac{du}{dx}$$

Now what if  $f$  was a function defined not on the real line  $\mathbb{R}$ , but on 2-dimensional space  $\mathbb{R}^2$ , or more generally  $n$ -dimensional space  $\mathbb{R}^n$ . To each point  $p = (p_1, \dots, p_n)$  we associate  $f(p)$ . Now again, consider  $f(p_f) - f(p_i)$  for two points in this space.

For any curve  $C$  going between  $p_f$  and  $p_i$ , say defined by  $\mathbf{r}(t)$  for  $t$  a real number going from  $a$  to  $b$ , we can make the same partition  $t_i = a + (b - a)i/N$  and let  $N$  get large. Again, it becomes a telescoping sum:

$$\begin{aligned} f(p_f) - f(p_i) &= f(\mathbf{r}(b)) - f(\mathbf{r}(a)) \\ &= \sum_{i=1}^N f(\mathbf{r}(t_i)) - f(\mathbf{r}(t_{i-1})) \\ &= \sum_{i=1}^N \Delta f_i \rightarrow \int_C df. \end{aligned}$$

Now if we cared about coordinates, we could ask "how can we write  $df$  in terms of  $dt$  or  $dx_i$ ?"

We know from the multivariable chain rule that the infinitesimal change of  $f$  is the sum of the change in  $f$  due to every individual variable, so:

$$df = \sum_i \frac{df}{dx_i} dx_i \quad (1.5)$$

We know that  $dx_i$  together must lie along  $C$ . In terms of  $t$  since  $x_i = r_i(t)$ , we have  $dx_i = \frac{dr_i}{dt} dt$  giving:

$$f|_{p_i}^{p_f} = \int_C \sum_i \frac{df}{dx_i} \frac{dr_i}{dt} dt = \int_C \nabla f \cdot \frac{d\mathbf{r}}{dt} dt = \int_C \nabla f \cdot d\mathbf{r} \quad (1.6)$$

This is the fundamental theorem of line integrals. The proof of it was no different from the 1-D case.

Let's go further. Consider a region in three dimensions. We want to relate the total flux coming out of the region to the infinitesimal flux at each point inside the region. To do this, as before, we will subdivide the region. This time, it will not be into a series of intervals, but instead into a mesh of increasingly small *cubes*, as below.

#### PUT A GRAPHIC HERE

See that the flux out a side of each cube is cancelled out by the corresponding side on its neighboring cube. That means that the only sides that do not cancel are for the cubes at the boundary<sup>11</sup>, giving us the desired flux out.

So if we sum the fluxes over all infinitesimal cubes, we will get the total flux out of the boundary. For a single cube of sides  $dx, dy, dz$ , drawn below, the total flux will be the sum over each side.

$$\begin{aligned} \text{Flux} = & \mathbf{F}(x, y, z + dz/2) dx dy - \mathbf{F}(x, y, z - dz/2) dx dy \\ & + \mathbf{F}(x, y + dy/2, z) dx dz - \mathbf{F}(x, y - dy/2, z) dx dz \\ & + \mathbf{F}(x + dx/2, y, z) dy dz - \mathbf{F}(x - dx/2, y, z) dy dz \end{aligned}$$

#### SHOW GRAPHIC HERE

But we can write this as:

$$\left( \frac{\partial \mathbf{F}(x, y, z)}{\partial x} + \frac{\partial \mathbf{F}(x, y, z)}{\partial y} + \frac{\partial \mathbf{F}(x, y, z)}{\partial z} \right) dx dy dz = \nabla \cdot \mathbf{F} dV$$

---

<sup>1</sup>You may be worried that the cubes do not perfectly fit into the boundary when it is not rectangular. As the mesh gets smaller and smaller, this does not pose a problem. This can be made more rigorous (c.f. **GIVE A REFERENCE HERE**)

So the total flux will be the sum over all these cubes of each of their total fluxes. But then this becomes exactly the divergence theorem:

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, dV = \int_{\partial\Omega} \mathbf{F} \cdot d\mathbf{S}$$

It is an easy **exercise** to show that this exact same argument holds for an  $n$ -cube.

What did we do? In the fundamental theorem of calculus/line integrals, we had a function  $f$  evaluated on the 1-D boundary, and we chopped the curve into little pieces that cancelled on their neighboring boundaries, making a telescoping sum. Then we evaluated the contribution at each individual piece, and found that it was  $df = f'(x_i)dx$ , meaning that the evaluation on the boundary could be expressed as an integral of this differential quantity over the curve. That is Equation (1.1).

For the divergence theorem, we had a vector field  $\mathbf{F}$ , again *evaluated on the boundary*, this time in the form of a surface integral. We chopped the region into little pieces (cubes now) that cancelled on their neighboring boundaries, making a telescoping sum. Then we evaluated the contribution at each individual piece and found that it was  $\nabla \cdot \mathbf{F} \, dV$ , meaning that the integration on the boundary could be expressed as an integral of this differential quantity over the region. That is Equation (1.2).

Through abstraction, we see that there is really no difference. Perhaps now Equation (1.4) does not look so mysterious and far-off.

For Equation (1.3), we have a vector field  $\mathbf{F}$  evaluated on the boundary in the form of a contour integral around a region. This is the total circulation of  $\mathbf{F}$  around the region. Let us chop the region into little pieces.

### INSERT GRAPHIC HERE

On an infinitesimal square, we get that the circulation is:

$$\begin{aligned} \text{Circulation} = & \mathbf{F}(x + dx/2, y)dy - \mathbf{F}(x - dx/2, y)dy \\ & + \mathbf{F}(x, y - dy/2)dx - \mathbf{F}(x, y + dy/2)dx \end{aligned}$$

This can be written as:

$$\left( \frac{\partial \mathbf{F}}{\partial x} - \frac{\partial \mathbf{F}}{\partial y} \right) dxdy = \nabla \times \mathbf{F} \, dA$$

Exercise (**MAKE AN EXERCISE**) generalizes this to a surface in 3D, to get the more general version of Stokes' theorem:



$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S} \quad (1.7)$$

The philosophy behind these proofs is always the same. It is the manipulation of the differentials that seems wildly different every time. The curl looks nothing like a divergence, and a divergence is distinct from a gradient. Moreover, its not clear in what way each one generalizes the one dimensional derivative  $df = f'(x)dx$ . This is the problem that the symbol ‘d’ in Equation (1.4) was made to solve.

We need to stop thinking of the 1-d derivative, the gradient, the divergence, and the curl, as unrelated operations. They are in fact, the same operation, applied in different circumstances. Infinitesimal change, flux, and circulation are all the same type of derivative, acting on different types of objects.

Perhaps part of this was clear from multivariable calculus: the gradient is nothing more than a generalization of the derivative to functions of multiple variables. But then why are there seemingly two different, unrelated types of “derivative” on vector fields? Instead of a regular, gradient-like object, we have two: the divergence and the curl.

It will turn out that the reason that there are two is this: the vector fields that we take curls of are a different type of object from the vector fields we take the divergence of. To see this more clearly, we need to stop thinking of functions and vector fields as totally separate objects. Every object that we’ve encountered when integrating: from functions in 1-D or 3-D, to vector fields in  $n$ -D, have been examples of **forms** .

## 1.2 The Notion of a Form

**Talk about coordinate independence of the FTOC and now how we get it for the proof in the divergence theorem**

**Transition from hat vectors to forms**

### **1.3 Stokes' Theorem**

### **1.4 To Manifolds, Coordinate Freedom**

### **1.5 Vectors, Forms, and Tensors**

### **1.6 Distance, a Metric**

### **1.7 Movement, Lie's Ideas**

First, something cool. Euler's identity  $\rightarrow e^{a\frac{\partial}{\partial x}}$

### **1.8 Exercises**

## Chapter 2

# Beyond Harmonics: Representation Theory



## Chapter 3

# The group $\mathrm{SL}_2(\mathbb{C})$ and friends



Chapter 4

**Classical Mechanics and  
Symplectic Geometry**





# Chapter 5

## Einstein's Gravity



# **Chapter 6**

## **An Introduction to Quantization**



Chapter 7

Classification of Simple Lie  
Algebras over  $\mathbb{C}$

# Index

Differential Forms, 9  
Stokes' Theorem  
    General, 4