Representations
of a
Physical Universe

Alex Atanasov & Aaron Hillman

July 11, 2016

Contents

3   More Advanced Topics 53

Part 0

Old Notions and New Horizons

Chapter 1

Old Notions

1.1   The Cartesian Coordinate System

One of the most important revolutions in mathematics and physics was ushered in by an idea of the seventeenth century mathematician and philosopher René Descartes. The idea was that any point P in the Euclidean plane could be represented by a pair of numbers (x, y). The numbers themselves represented distances along two perpendicular axes that met at a point $(0, 0)$, called the origin. By introducing this concept, he had done something amazing. He had related the *geometry* of the plane to the *algebra* of variables and equations. Algebra could be *represented* geometrically, and conversely geometric problems could be solved by going into the realm of algebra.

Insert 2D Plane with Coordinates

Coordinates soon became more than just pairs of numbers (x, y). Their use was extended to 3D space, and later to arbitrarily high dimensions. They would subsequently be used to lay the foundations for modern physics and mathematics. Linear algebra, multivariable calculus, and all the connections between algebra and geometry begin with the concept of a coordinate. Since then, the use of coordinate systems has proven indispensable to physicists and mathematicians throughout history. Newton used Descarte's coordinate system to formulate his infinitesimal calculus. Maxwell used it to analyze electromagnetic fields, discovering mathematically that light is a wave in the electromagnetic field. Einstein, going further, made use of coordinate systems to formulate his theory of gravitation. Today, physicists and engineers do their calculations within the frame of coordinate

systems. In mathematics, Descarte's idea planted the roots for what would turn into the modern field of algebraic geometry.

When studying a geometric phenomenon in some n-dimensional space, say $\mathbb{R}^n$, we pick an origin and axes to form our coordinate system. For a ball falling, we could set the origin at some point on the ground, and pick one axis parallel to the ground, and one perpendicular. We can decide to measure the axes in meters, or we could decide to do it in feet (nothing stops us from making bad choices). The physical point P where the ball lies is represented by $(x, y) = (0 \text{ m}, 10 \text{ m})$. The coordinate y is a natural choice of coordinate, as it corresponds to our intuitive notion of height.

Insert Ball Falling

We can now study y, free of geometry, as just a function which we can do arithmetic and calculus on. If we are given an equations of motion, say

$$\frac{d^2 y}{dt^2} = -g, \quad \frac{d^2 x}{dt^2} = 0$$

with initial conditions,

$$\frac{dy}{dt} = 0, \quad \frac{dx}{dt} = 0$$

then we can perform our well-known kinetic calculations for the system, and see how the system evolves in the *time* direction. **A recurring theme will be that dynamics of a system in n-dimensional space can be thought of just a special type of geometry in $n+1$ dimensional space, putting time as an added dimension**.

Because this book will, in large part, be concerned with studying the ways in which geometry, algebra, and physics connect, it is worthwhile to dwell on the *philosophy* behind coordinate systems.

The ball will fall from 10 meters, according to the force of gravity. That is the way the world works. It doesn't matter what coordinate system we set up to do that calculation, we should get the *exact same result*. Plainly: nature doens't *care* what coordinate system we use. This fact, obvious as it may be, is worth thinking about: No matter what coordinate system we use, the equation of motion should give the same dynamics. The laws of physics should be *independent of any coordinate system*.

Newton's law $\mathbf{F} = m\mathbf{a}$ relates the force vector to the acceleration vector. The vector representing the force $\mathbf{F}$ that you apply on a surface is an object independent of coordinate system, and so is the resulting acceleration vector. The *components* of these vectors (F_x, F_y, F_z) and (a_x, a_y, a_z), however, depend on what you have chosen for the x, y, z axes. These components *represent* a real physical vector, but only once we pick a coordinate system. If we were to pick a different coordinate system, they would change. If a physical law ever looked like $F_x dx = dW$ it would not be true in every coordinate system, because it puts emphasis on just one of the components over the others. While in some coordinate system this may be nonzero, in another it would be zero, so the work done, dW, is not an invariant. So such a law would not be universally valid regardless of what coordinate system we pick: it would be wrong. The need for such invariance is why the true formula looks like

$$\mathbf{F} \cdot d\mathbf{r} = F_x dx + F_y dy + F_z dz = dW.$$

Although its not yet obvious that this is invariant under change of coordinates, at the very least it doesn't put one component above any of the others.

1.2   Linear Algebra & Coordinates

The traditional conception of a coordinate system, a series of perpendicular lines that together associate ordered tuples of numbers to each point in space, is not representative of all coordinate systems. For one, we do not need the requirement that the lines be perpendicular (we'll show later that for general spaces, there's not even a good notion for what perpendicular *means*). Our coordinate system could instead look like this:

Graphic of non-perp lines and representing a point like that

In the language of linear algebra, once we choose an origin point, choosing a set of coordinate axes is the same as choosing a basis for the space (a coordinate basis). We can relate the new system of coordinates x_i' in terms of the old system x_i by matrix multiplication: $x_i' = \sum_{j=1}^n \mathbf{A}_{ij} x_j$. This is exactly what's called a change of basis in

linear algebra. Transformations between coordinate bases are exactly the invertible **linear transformations** .

As in linear algebra, we need our coordinate system to both **span** the space so that we can represent any point, and be **linearly independent** so that every point has exactly *one* representation in our coordinate system. That's what a basis *is*: it specifies a good coordinate system.

Definition 1.1. *A set of vectors is said to span a space $\mathbb{R}^n$ if every point P can be represented as $a_1\mathbf{v}_1 + \ldots a_n\mathbf{v}_n$*

Definition 1.2. *A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is called linearly independent if there is only one way to represent the zero vector $\mathbf{0}$ as a combination of them, namely as $\mathbf{0} = 0v_1 + \cdots + 0v_k$.*

This second definition is the same as saying every point has a unique representation. If there were two ways to represent a point P: as

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

and

$$b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$$

then subtracting these two different combinations would give a nonzero way to represent zero. Conversely, if there were a nonzero combination of vectors summing to zero, then we could add that combination to a coordinate representation of any point and get a *different* representation of the same point. So coordinate representations for all vectors are unique as long as long as there is only one representation for zero: the trivial one where each component is zero.

Intuitively, linear independence means that there is no superfluous information in the set of vectors. We cannot linearly combine vectors in some subset to get another vector in the set; each vector is adding its own unique additional piece of information, making the set able to span in an additional direction.

Bases that don't span, or are not linearly independent, would lead to coordinate systems like these:

Show a 2-D basis in a 3-D space, and a basis of 3 vectors in 2-D space

Very often in mathematics, we ask "does a solution exist?", and "if there is a solution, is it unique?". These two questions are dual to one

another. If a set of vectors spans the space, then there *exists* a way to represent any point (at least one way to represent any point). If a set of vectors is linearly independent, then *if* you can represent a point, that representation is *unique* (no more than one way to represent any point).

Now to stress the same idea again: because points in $\mathbb{R}^n$ and vectors are essentially the same thing, the idea that points in space are invariant of a coordinate system of course applies to vectors. If we choose a basis for our vector space $\mathbf{v}_1, \ldots \mathbf{v}_n$, then we can express any vector $\mathbf{u}$ by a unique combination $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$. We then say that in this basis, we can represent $\mathbf{u}$ by a list of numbers:

$$\mathbf{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

In some sense, writing this as an equality is wrong. The vector $\mathbf{u}$ represents something physical: a velocity, a force, the flow of water. It doesn't depend on the coordinate system. On the other hand, the right hand side is just a list of numbers that depend entirely on the coordinate system chosen. If we change coordinate systems, the right hand side changes. Because $\mathbf{u}$ exists (say, in the real world) independently of coordinates used, it does not change.

A vector is *not* a list of numbers. Once we pick a basis, a vector can be *represented by* a list of numbers, but if we change into a different basis, those numbers all have to change as well. This exact same idea will be the reason why a tensor is *not* just a multi-dimensional array (like the ones you see in computer science). It can be *represented by* a multi-dimensional array once we pick a coordinate system, but the numbers in entries will change depending on the coordinate system we pick.

A more careful way to write $\mathbf{u}$ would be:

$$\mathbf{u} = \begin{pmatrix} \mathbf{v}_1 \ldots \mathbf{v}_n \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n. \tag{1.1}$$

Once we pick a basis, that column of coordinates means something. If

we denote our basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ by B, then we will use the notation

$$\mathbf{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_B = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n.$$

Let us do a very simple example to start. In the 2-D plane, say we have our original basis $\mathbf{v}_1, \mathbf{v}_2$ and we rotate it by $\pi/4$ radians to get a new basis. Say we have a point P whose coordinate representation was $\mathbf{v}_1 + \mathbf{v}_2$, or $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ in the original basis.

Now our new basis is the old one rotated by $\pi/4$ so

$$\mathbf{v}_1' = \frac{\sqrt{2}}{2}\mathbf{v}_1 + \frac{\sqrt{2}}{2}\mathbf{v}_2$$
$$\mathbf{v}_2' = -\frac{\sqrt{2}}{2}\mathbf{v}_1 + \frac{\sqrt{2}}{2}\mathbf{v}_2.$$

PUT GRAPHIC HERE

As a matrix transform, we can write this as[1]:

$$\begin{pmatrix} \mathbf{v}_1' & \mathbf{v}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

This relates the actual basis vectors themselves. On the other hand, if we wanted to see the *coordinates* representing $\mathbf{v}_1'$ and $\mathbf{v}_1'$, then in the new basis they would simply be represented in coordinates as:

$$\mathbf{v}_1' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\text{new}}, \quad \mathbf{v}_2' = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\text{new}}$$

and in the old basis they'd be represented as:

---

[1]We're going to put the basis vectors in row-tuples instead of column tuples. The reason is due to Equation (1.1), where we've decided to have the coordinates be represented in a column. Because of the way we do matrix multiplication, we then want the basis vectors to be rows for the multiplication to make sense. It's really just an issue of styling and indexing, and not a physical issue. If we were to write the basis vector transform in terms of columns, we'd get a matrix that's the transpose of the one above, and matrix transposes would appear in subsequent equations, making them less clean.

$$\mathbf{v}_1' = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}_{\text{old}} , \quad \mathbf{v}_2' = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}_{\text{old}}$$

If we know that we can describe a point through coordinates as $\begin{pmatrix} x \\ y \end{pmatrix}$ in the new basis, then we can easily get its description in the old basis as:

$$
\begin{aligned}
\begin{pmatrix} x \\ y \end{pmatrix}_{\text{new}} &= x\mathbf{v}_1' + y\mathbf{v}_2' \\
&= x \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}_{\text{old}} + y \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}_{\text{old}} \\
&= \left(\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}_{\text{old}\leftarrow\text{new}} \begin{pmatrix} x \\ y \end{pmatrix}_{\text{new}} \right)_{\text{old}}
\end{aligned}
\tag{1.2}
$$

This is the same matrix that related the basis vectors. We'll call it $\mathbf{A}$. $\mathbf{A}$ takes the new coordinate representations (x, y) and tells us how they'd look like in the *old* basis.

So then it is the *inverse* $\mathbf{A}^{-1}$ that tells us how our old coordinate representations of a point P will look like in our new basis.

For our point P, represented as $(1, 1)$ in our original basis, in the new basis, we would have:

$$
P = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\text{old}} = \left(\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}^{-1}_{\text{new}\leftarrow\text{old}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\text{old}} \right)_{\text{new}} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}_{\text{new}}
$$

Indeed, $\mathbf{v}_1 + \mathbf{v}_2 = \sqrt{2}\mathbf{v}_1' + 0\mathbf{v}_2'$, as we wanted.

This is the idea. If we *vary* the $\mathbf{v}_i$ to a different basis, $\mathbf{v}_i'$, then the coordinates a_i' will vary the *other* way, so that

$$\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = a_1'\mathbf{v}_1' + \cdots + a_n'\mathbf{v}_n'$$

is *invariant* regardless of coordinate choice.

Let's make this precise in the general case. If we start with a set of basis vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and we make the linear transformation to a new basis $\{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$ so that, as before:

$$(\mathbf{v}_1', \ldots, \mathbf{v}_n') = (\mathbf{v}_1, \ldots, \mathbf{v}_n)\mathbf{A} \tag{1.3}$$

then since the vector **u** should not change when we change our basis, we must have:

$$\begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \tag{1.4}$$

so that

$$(\mathbf{v}'_1, \ldots, \mathbf{v}'_n) \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix} = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \mathbf{A} \mathbf{A}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

as desired. **THIS WOULD BE A GOOD EXERCISE: PROVE IT HAS TO BE A INVERSE**

We say that the basis vectors $\mathbf{v}_i$ **co-vary** and the coordinates a_i **contra-vary** with the change of basis. The idea, although it sounds simple, is rather hard to get the feel of. It's worth thinking a good bit about how coordinates and bases need to vary in opposite ways so that the physical object represented by the coordinates stays the same regardless of how we look at it.

This will be a caption for a sketch of a 3-D rotation

When you rotate your character in a video game (or in real life too...), the world rotates *contrary* to the direction you've rotated in. The coordinates of what you've seen have *contra-varied* while your basis vectors have *co-varied*. The end result is that despite changing your coordinate system, physics stays the same: invariant. This extends beyond just rotations to *all* linear transformations point.

As one further example, consider the following transformation:

$$\mathbf{v}'_1 = 2\mathbf{v}_1 + \mathbf{v}_2$$
$$\mathbf{v}'_2 = \mathbf{v}_1 + 2\mathbf{v}_2.$$

It's easy to compute the inverse of this matrix and see how the new coordinates should work, but it worthwhile looking at how this looks geometrically. It is not a rotation, but more of a "stretching". Notice that while in our original perspective, $\mathbf{v}_1$ and $\mathbf{v}_2$ were orthogonal vectors, in the *new* perspective, they are *no longer* orthogonal. This is very important: linear transformations in general do not preserve orthogonality.

Include graphic here

If we were to take the "dot product" $\mathbf{v}'_1 \cdot \mathbf{v}'_2$ just by multiplying corresponding coordinates and summing them up, then in the new basis we'd get zero, but in the original basis we would *not*. This dot product actually changes depending on the coordinate system that we use! In some sense, this is expected: all we're doing is multiplying contravariant coordinates together and summing them up. The result should be contravariant as well (in fact doubly contravariant).

But then does that mean that the state of being orthogonal is coordinate-dependent? Two vectors can be orthogonal in one coordinate system and not in another? Immediately, you should say "No, the state of being perpendicular is a real-world observation that doesn't depend on what coordinate system you use." So what is going on? The only possibility is that we're not doing the dot product right. We can't just multiply the contravariant coordinates: we need to make use of both the coordinates and the vectors.

In the succeeding chapter, we will study how to define these *inner products* on vectors. To obtain the invariant scalar from the contravariant coordinates, we will need to introduce a doubly-covariant quantity called the *metric tensor*.

As soon as we define a product operation between vectors $\mathbf{v} \cdot \mathbf{w}$, we have added structure to our vector space. An inner product makes it so that not all bases of vectors are equal. There are bases where the basis vectors satisfy $v_i \cdot v_j = 0$ unless $i = j$, meaning the basis vectors are all orthogonal. The orthogonal bases are special among the set of all bases. Not all linear transformations, therefore, will preserve orthogonal bases: only orthogonal ones will.

1.3 Curvilinear Coordinates: Polar & Beyond

Perhaps you may be wondering why we've spent so much time on changing between coordinate systems represented by basis vectors centered at a fixed origin. You've seen change between cartesian and polar coordinates, and that has little to do with the linear change of basis that we've just discussed, right?

We could use something like a polar system of (r, θ) or a spherical system (r, ϕ, θ). These are now not representable in terms of three

axes, but instead look like this:

Graphic of polar coordinate system/spherical

This is an example of a non-linear coordinate transformation. They are more commonly referred to as **curvilinear**. Whereas linear ones map lines to lines, curvilinear ones more generally map lines to curves. The idea for making sure that the equations of physics still stay true for non-linear coordinate transformations is to note that just like a curve locally looks like a line, a *non-linear* transformation locally looks like a *linear* one. The linear transformation that it locally looks like is called the **Jacobian** J. If the laws of physics are invariant under linear transformations locally at each point, then *globally*, they will be invariant under non-linear ones as well. That is why we cared about studying covariance and contravariance for linear transformations: more complicated cases can be reduced to their local linear behavior.

As an example, consider going from a cartesian to a polar coordinate system. We have $x = r \cos \theta$ and $y = r \sin \theta$. Certainly, this is not a linear transformation of coordinates. There is sinusoidal dependence on θ in this transformation. Physics and geometry, however, do not have laws in terms of absolute coordinates (it doesn't make sense to say "That object is located at 50 meters") but only in terms of relative distances (you'd instead say "That object is located 50 meters *relative to* me"). It is the changes over relative distances between points that we care about, and these are obtained by integrating the *infinitesimal* changes at each point.

So although x, y do not depend linearly on theta, through the use of the chain rule, we have:

$$dx = \cos \theta dr - r \sin \theta d\theta$$
$$dy = \sin \theta dr + r \cos \theta d\theta$$

This is a linear relationship now:

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} \begin{pmatrix} dr \\ d\theta \end{pmatrix}$$

So every nonlinear transformation from some coordinate system $x_1 \ldots x_n$

to $x' \ldots x'_n$ has the local linear transformation law:

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}_{\text{old}} = \begin{pmatrix} \frac{\partial x_1}{\partial x'_1} & \cdots & \frac{\partial x_1}{\partial x'_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial x'_1} & \cdots & \frac{\partial x_n}{\partial x'_n} \end{pmatrix}_{\text{old}\leftarrow\text{new}} \begin{pmatrix} dx'_1 \\ \cdots \\ dx'_n \end{pmatrix}_{\text{new}}$$

$$\Rightarrow \begin{pmatrix} dx'_1 \\ \cdots \\ dx'_n \end{pmatrix}_{\text{new}} = \begin{pmatrix} \frac{\partial x_1}{\partial x'_1} & \cdots & \frac{\partial x_1}{\partial x'_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial x'_1} & \cdots & \frac{\partial x_n}{\partial x'_n} \end{pmatrix}_{\text{new}\leftarrow\text{old}}^{-1} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}_{\text{old}}$$

This is exactly the analogue of Equation (1.4), so indeed the changes in coordinates dx_i can be called *contravariant*, just like the coordinates themselves were for the linear case. This is all an extension of the principle of local linearity learned in calculus.

Similarly, we can express the new derivative operators in terms of the old ones by using the chain rule. For polar coordinates we have

$$\frac{\partial f}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial f}{\partial y}$$

$$\frac{\partial f}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial f}{\partial y}$$

or more compactly we can relate just the differential operators themselves:

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$$

so that more generally:

$$\begin{pmatrix} \frac{\partial}{\partial x'_1} \\ \vdots \\ \frac{\partial}{\partial x'_n} \end{pmatrix}_{\text{new}} = \begin{pmatrix} \frac{\partial x_1}{\partial x'_1} & \cdots & \frac{\partial x_1}{\partial x'_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial x'_1} & \cdots & \frac{\partial x_n}{\partial x'_n} \end{pmatrix}_{\text{new}\leftarrow\text{old}}^{T} \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}_{\text{old}}$$

If you look carefully, you will see that this accordingly mirrors the covariant change of vectors in Equation (1.3). So while the infinitesimal changes in the coordinates themselves are contravariant, just like the linear coordinates themseles, the *differential operators* corresponding

to changes in these coordinates become *covariant*, just like the vectors $\mathbf{v}_i$ in the linear case. This is the first correspondence that will hint that our basis vectors $\mathbf{v}_i$ actually *correspond* to the differential operators $\frac{\partial}{\partial x_i}$.

Indeed, as we begin to move away from 3-D and n-D Euclidean space, we will see why the old notions of unit vectors $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ are better viewed as the operators $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$, and $\frac{\partial}{\partial z}$, and in general $\mathbf{v}_i \to \frac{\partial}{\partial x_i}$. This change of notation will allow us to easily pass onto far more general spaces than the Euclidean ones we've gotten used to.

Chapter 2

New Horizons

2.1   The Manifold

It is impossible to conceive of a complete description of physics that need not make use of a geometric space in which states exists. And even those physical degrees of freedom not tied to positions or times, such as spin, naturally lend themselves to having their states described in a geometric space. As such, a natural starting point for understanding physical systems is first understanding the space in which the system lives, the *stage* on which physics takes place. For many years it was merely posited that physical space had the mathematical character of $\mathbb{R}^3$ (Euclidean space) and that time was merely another orthogonal direction, making spacetime $\mathbb{R}^4$.

Everything we've done so far has been in $\mathbb{R}^n$. We have used different bases to represent it. Perhaps when we were first born, we expect that the universe looks like $\mathbb{R}^3$, and it goes out infinitely far in every direction. Perhaps the first humans, too, expected that the surface of the earth looks like $\mathbb{R}^2$, stretching out infinitely. But we know now that the surface of the earth is *not* the plane $\mathbb{R}^2$. We don't know so much about the universe, but there is no reason to expect it to be $\mathbb{R}^3$.

We exist in the universe like 2-D people would exist on a sphere, or on some other object with rich geometry that by no means needs to be flat. Because the object is so large, we have no idea how it *globally* looks, but *locally*, just like 2-D humans would see a flat plane, we see $\mathbb{R}^3$. These geometric objects we are talking about are called *manifolds*. What is a manifold? It is a set of points that looks like Euclidean space around each point.

A line is a one-dimensional manifold (in fact it *is* a Euclidean space) a circle is a one-dimensional manifold (when you zoom in on a point, it looks like a line), so is an ellipse, parabola, hyperbola, and the graph of any smooth function. A sphere is a two dimensional manifold (note that the sphere is just the 2-D surface of a 3-D ball). It locally looks like the euclidean plane, just as the world looks flat to us. The Mobius strip is also a two-dimensional manifold. Although globally it is a twisted band, locally it, too looks like flat two-dimensional space.

This will be our intuitive basis for what we mean by a manifold: a manifold is anything that *locally* looks like Euclidean space. This concept shall have such fundamental weight that it is especially important to try to make it formal. In order to do that, we need to define what we mean by locally.

In order to formally define a manifold, we must first define a **topological space**. A topological space X is a set with a family of open subsets T satisfying

1) $\emptyset$ (empty set) and X are in T

2) A finite union of members of T is in T

3) A finite intersection of members of T is in T

and T is called the **topology** on X. We can talk about points x in the set X and we say that a collection of open sets $\{O_\alpha\}$ **covers** X if X is contained in their union. Intuitively, the information contained in such a collection of sets should be enough to talk about all of X. The last thing we need is a **chart**, which is an open set O_a paired with a continuous function $\varphi_\alpha : O_\alpha \to \mathbb{R}^n$ with continuous inverse. This map is identifying each point in the subset O_α with a corresponding point in $\mathbb{R}^n$. We can now define a manifold

DRAW A PICTURE OF A SPHERE COVERED BY CHARTS

We need to be able to use **Talk about the sphere here (not the metric part)**

Just like in $\mathbb{R}^n$ where had different coordinate systems around an origin, on a manifold M, we will *locally* at each point have coordinate systems that look exactly like the ones we used for $\mathbb{R}^n$ in section 1.2.

Just because we have a coordinate system to describe the manifold doens't mean we have everything. It may seem strange, but up until

now we have missed talking about a vital part of geometry: the notion of *distance*. Even though we've talked in all the way points can be represented by coordinates, none of these numbers representing coordinates have any *inherent* notion of distance to them.

Concept 2.1. *A smooth manifold is a continuous collection of points which can be invertibly mapped into Euclidean space in a neighborhood of any point, allowing us to locally talk about our manifold by working in Euclidean space.*

Now that we prepared the stage, we are ready to study the fields which can exist on it.

2.2   The Field

One of the most important aspects of physics is studying the *fields* that live on manifolds. Just like in multivariable calculus, this means the study of scalar fields ϕ that associate a number to every point P. Examples are voltage, potential energy, mass/charge density, etc. This also includes the study of vector fields **v** associating a vector to each point. These can be wind speeds, electric fields.

While scalar fields associate a number $\phi(P)$ to each point P, which looks the same regardless of the local coordinate system used at P, vector fields will associate a specific vector **u** to a point P whose coordinate representation will, of course, change depending on the local coordinate system at P.

[I want to somehow motivate why the vectors should looks like $\partial/\partial q_i$]

Show a graph of the curves for coordinates q_i

2.3   What Follows

The rest of this book will expand both on the geometry of fields and manifolds, and also on the larger ideas of groups, homogenous spaces, and representations.

In chapter 1, we will continue studying the fields that live on manifolds. We'll prove the General Stokes' theorem, an elegant generalization of the divergence, curl, and line integral theorems that have

been taught in multivariable calculus. From there, we will study more thoroughly the concept of a metric, and how this relates vector fields to differential forms. The notion of a derivative will be extended to manifolds, and will take the form of a "Lie Derivative".

In chapter 2, we will introduce Fourier Analysis as a powerful tool for studying functions on the real line an Euclidean space. Then we will shift to looking at the representation theory of *finite* groups and illustrate the parallels. We will then return to the study of continuous group actions on especially symmetric "homogenous" spaces, and show how Fourier analysis is related to their representation theory. Towards the end, we will expand on the idea behind a group actions on manifolds and look at the representation theory, giving a small glimpse into harmonic analysis: the Fourier transform on manifolds. Just as in the first chapter, we'll recognize the importance of the underlying differential geometry of the group action. The underlying differential structure is known as the "Lie Algebra" of the group, and we will discuss that.

In chapter 3, we introduce some background behind Lie Algebras. We put almost all of our focus on one special case: the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. The relationship between this algebra and the symmetries of the sphere are explored, as well as its applications in quantum physics for studying angular momentum. The representation theory of a variant of this algebra gives rise to the concept of spin.

In chapter 4, we move further into physics, going over classical Lagrangian and Hamiltonian Mechanics. We discuss Noether's theorem in both the Lagrangian and Hamiltonian Pictures, and then we move to study Hamiltonian mechanics using the language of differential geometry that we have developed. This will give rise to *symplectic geometry*. In chapter 7, combining this with representation theory gives rise to *quantum mechanics*.

In chapter 5, we apply differential geometry first to the study of electromagnetism, and then to gravitation. We shall arrive at Einstein's theory of gravity. Along the way, we study in even greater detail the notion of a metric, a connection, and curvature.

In chapter 7, we use the representation theory and differential geometry that we have developed so far to study how quantum mechanics can arise from quantizing a symplectic manifold.

Finally, chapter 8 studies Lie algebras in greater detail, working towards the *classification of complex semisimple Lie algebras*. Along the way, we will look at the relationship between representation theory of Lie algebras and modern physics.

Part 1

A Better Language

Chapter 3

Differential Geometry

"In mathematics you don't understand things. You just get used to them."

John Von Neumann

"There is no royal road to geometry"

Euclid

Setting the Stage

It is impossible to conceive of a complete description of physics that need not make use of a geometric space in which states exists. And even those physical degrees of freedom not tied to positions or times, such as spin, naturally lend themselves to having their states described in a geometric space. As such, a natural starting point for understanding physical systems is first understanding the space in which the system lives, the *stage* on which physics takes place. For many years it was merely posited that physical space had the mathematical character of $\mathbb{R}^3$ (Euclidean space) and that time was merely another orthogonal direction making spacetime $\mathbb{R}^4$. Such assumptions were poor for two reasons. For one, a physicist should strive to make as few assumptions as possible in the pursuit of a general theory and

presuming that physical space is $\mathbb{R}^3$ indirectly assumes many things about how physical systems behave. Second, and much more egregious, in many cases these assumptions are simply false; that is, they are inconsistent with what we experimentally observe about the universe. But what is the way forward, then? First, we must explore in abstract the geometries that spaces may have and how objects behave on them. This naturally leads us to what is at present the best mathematical conception of the stage on which physics takes place, the **manifold**.

The term manifold describes a very general class of mathematical objects. Moreover, without certain restrictions on their properties, they may take certain pathological and non-physical forms. As such, we will restrict our discussion to smooth manifolds, since spacetime (and other geometric spaces we care about) appears to be such a thing. It is worth noting that perhaps, as physics progresses, we may learn that this assumption was also a mere approximation to a more fundamental description of reality. This is a fair point, but nonetheless, this assumption is completely compatible with the physics of interest here.

In order to formally define a manifold, we must first define a **topological space**. A topological space X is a set with a family of open subsets T satisfying

1) $\emptyset$ (empty set) and X are in T

2) A finite union of members of T is in T

3) A finite intersection of members of T is in T

and T is called the **topology** on X. We can talk about points x in the set X and we say that a collection of open sets $\{O_\alpha\}$ **covers** X if X is contained in their union. Intuitively, the information contained in such a collection of sets should be enough to talk about all of X. The last thing we need is a **chart**, which is an open set O_a paired with a continuous function $\varphi_\alpha : O_\alpha \to \mathbb{R}^n$ with continuous inverse. This map is identifying each point in the subset O_α with a corresponding point in $\mathbb{R}^n$. We can now define a manifold

An n-dimensional **smooth manifold** is a topological space M equipped with a cover $\{O_\alpha\}$ and charts $\{\varphi_\alpha\}$ such that when O_α and

O_β have nontrivial intersection (they overlap), $\varphi_\alpha \circ \varphi_\beta^{-1}$ is **smooth**. Smooth here means infinitely differentiable. This composition is called a **transition function** and its smoothness guarantees that our collection of charts (called an **atlas**) will allow us to define smooth functions on our manifold with no trouble.

It is also important to stress that the reader should not at the moment think of charts as mapping identifying each point p in a manifold M with n coordinates in $\mathbb{R}^n$. The point of this construction is to have coordinate independent notions of geometry. When working in a chart, we are identifying each point in the manifold with a corresponding point in $\mathbb{R}^n$, which exists independent of a set of numbers we choose to identify it with.

This definition of a manifold may still seem too abstract. Intuitively, this definition is saying that an n-dimensional Manifold is a set which *locally* looks like $\mathbb{R}^n$. In the context of physics, by considering general manifolds, we are laxing our stipulation that space is Euclidean in its totality to a requirement that space *seems* Euclidean in a small patch around any given point. For example, we know by various means, that the surface of the earth is roughly like the surface of a spherical ball. But, ignoring irregularities in the surface like hills and mountains, at any given point on earth the surface appears flat in some patch around you. We could, in principle, draw lines determining a collection of relatively flat patches of earth which collectively cover all of earth and then map each of these onto part of the plane, $\mathbb{R}^2$. Therefore the surface of the earth is to good approximation a perfectly fine 2-dimensional smooth manifold, but we note that as it is different from the plane. In a similar manner, its conceivable that although in 3-dimensional patches around us physical space appears to be Euclidean, the entirety of physical space may be described by some much more interesting geometry. It is on these non-Euclidean manifolds that much interesting and often unintuitive physics arises.

As we endow our geometric spaces with more structure it is important to keep in mind at each stage the limits of what we can say about our space. At this point, we only have information about the smoothness and *topology* of our space, but we have no notion of how distances behave on our manifold. A concrete example of this limit is

that we have not provided enough structure to our notion of manifold to differentiate between an egg and a ping-pong ball. The surfaces of these objects are topologically the same, and differentiating between the two requires a notion of how long different curves are on their surfaces; such differentiation is facilitated by the presence of Riemann structure, which our manifolds have not yet been equipped with.

Moreover, the reader should not presume that they are able to visualize the totality of a manifold. The capacity to do so for an n-dimensional manifold is contingent upon the ability to *embed* the manifold in $\mathbb{R}^{n+1}$, as is commonly done with the 2-sphere (picturing the two dimensional surface of the sphere as a curved space in three dimensional space. We are not generally able to do this, and so the reader should not rely on intuitions rooted in embeddings, as they are a geometric crutch.

Concept 3.1. *A smooth manifold is a continuous collection of points which can be invertibly mapped into Euclidean space in a neighborhood of any point, allowing us to locally talk about our manifold by working in Euclidean space.*

Now that we prepared the stage, we are ready to study the fields which may exist on it.

Vectors Reinvented

Surely the reader is familiar with the concept of a vector field. The geometric intuition behind a vector field $v : \mathbb{R}^3 \to \mathbb{R}^3$ in Euclidean space is that we input a position in space and get out an object at said point with a magnitude and direction. This may describe the velocity of a particle at the point, or a force acting at such point. But as we generalize the geometry of the spaces in which physics is taking place, it becomes natural and necessary to utilize differently defined mathematical objects analogous to (or even in special cases reducing to) those previously used. In order to build up our new geometric machinery, we will start simple and make an effort to intuitively motivate the need for this new formalism, often by pointing out the defunctness of old ways. Nonetheless, this modern view of geometry may never feel as intuitive as arrows in three dimensional space, and so

the reader should take seriously the epigraph at the beginning of the chapter.

Perhaps the simplest object to conceive of existing on a manifold is a smooth function

$$f : M \to \mathbb{R}$$

This definition should not seem surprising to the reader. At each point in the manifold the function assigns a value and in order for it to be smooth we require that

$$f \circ \varphi_\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}$$

be smooth. The collection of such functions is denoted by $C^\infty(M)$. Perhaps the most fundamental functions are coordinates themselves! For example, each Cartesian coordinate in $\mathbb{R}^3$ (x, y, and z) returns a value at each point in $\mathbb{R}^3$ corresponding to what we call the x, y, or z coordinate. Clearly these are infinitely differentiable, defining a smooth function. We see that, working in a chart, we can then use the local coordinate functions $\{x^\mu\}$ on $\mathbb{R}^n$ as arguments of our function by relating

$$f(p) = f(x^1 \circ \varphi_a, \ldots, x^n \circ \varphi_a)$$

We will proceed writing functions in the way on the right, and omitting the composition with the chart (i.e. just writing $f(x^1, \ldots, x^n)$).

Slightly less intuitive is our modern definition of a tangent vector. In order to motivate this, we must first introduce a **curve** in M, which is a map $\gamma : T \to M$ where T is the compact interval $[0, T] \subset \mathbb{R}$. For each value of some parameter in $\mathbb{R}$, we get a point in M, and so γ is describing a path traced out in the manifold. The reader should be familiar with curves in $\mathbb{R}^3$ as described by the methods of basic multivariable calculus. In this case we were unknowingly spoiled, as $\mathbb{R}^3$ can be viewed both as a manifold and a vector space. We could sloppily interchange these viewpoints of $\mathbb{R}^3$, representing the points on the curve as endpoints of vectors starting from the origin. We could then acquire the tangent vector to a curve by differentiating the position vector component-wise. When one conceives of such a tangent vector, they cannot help but conceive of it existing as an arrow in three dimensional Euclidean space. And so one way to explain why our old Euclidean ways fall short is this: our manifold is not

in general a vector space. *We cannot linearly combine two distinct points on the manifold to get something meaningful.***(probably worthy of stressing as a concept)** We have not equipped our space with a structure enabling comparison of information at two distinct points and so we must not subconsciously presume one.

Moreover, we need to talk about geometric objects only making reference to the *intrinsic* properties of the manifold, and without embedding our n-dimensional manifold in $n+1$ - dimensional Euclidean space. How then can we view our tangent vector as an element of the vector space $\mathbb{R}^m$? We cannot in general, but I ask the reader, what good are tangent vectors? That is, what do we want from a tangent vector? A tangent vector to a curve provides us with the information necessary to determine the rate of change of a function as its argument moves along the curve. In the old way mentioned above, this is done by taking a directional derivative of a function along an arrow in $\mathbb{R}^3$, but on our arbitrary manifold, the output of γ is not a vector, but a point! Let us proceed demanding the desired *result* from our tangent vector and see if we can construct a consistent geometric viewpoint. We simply demand that the tangent vector to γ at point $p = \gamma(t)$ be the map

$$\gamma'(t) : C^\infty(M) \to C^\infty(M)$$

sending the smooth function f to its derivative

$$\frac{d}{dt}(f(\gamma(t)))$$

the collection of such maps for all possible curves γ at a point $\gamma(t)$ is called the Tangent Space at $\gamma(t)$, denoted $T_{\gamma(t)}M$. We will formally define this space later. Before seeing how this lends itself to a generalized notion of vector fields on manifolds, we must first discuss how mappings between manifolds give rise to mappings between the objects that exist on them.

Covariance, Contravariance, and Duality

Covariance, contravariance, and duality are very general concepts utilized in category theory, but in order to remain geometrically grounded we will restrict ourselves to a discussion of their application to manifolds. Consider a smooth map $\varphi : M \to N$ between manifolds M

and N. Such a map need not be invertible in general. If we have a smooth function f defined on N, φ can allow us to compute this function on M by considering the composition $\varphi^* f = f \circ \varphi : M \to \mathbb{R}$. This operation $\varphi^* : f \to f \circ \varphi$ is called a **pullback** because it brought the function defined on N *back* and defined it on M. Notice that since φ is not in general invertible, we are not in general able to send functions from M forward to N by composing with φ^{-1} (which may not exist). Because the map that goes forward from M to N gives rise to a backward map $\varphi* : C^\infty(N) \to C^\infty(M)$ we say that smooth functions are **contravariant** objects. Now, there exist other objects which, unlike functions, give rise to maps that would go the same way as φ. Consider a map $v : C^\infty(M) \to C^\infty(M)$. This v takes a smooth function on M and sends yields another smooth function on M. We may define $(\varphi_* v)(f) = v(f \circ \varphi)$ where f is a smooth function on N. Therefore, φ_* took a v defined on M and yielded a map $\varphi_* v : C^\infty(N) \to C^\infty(N)$ defined on N. Since φ_* maps in the same direction as φ we call it a **pushforward**. In practice, when we actually perform analysis on manifolds, we will be manipulating coordinates $\{x^\mu\}$ on $\mathbb{R}^n$, but technically speaking we are working with the *pullbacks* of these coordinate functions by our charts, since *those* are what are defined on the manifold. There exists a *duality* between contravariant and covariant objects, which can be summarized abstractly.

If we have a vector space V, we may consider the space of linear maps $v^* : V \to \mathbb{R}$. This too comprises a vector space V^* which we call the **dual space**. If we have a map $\varphi : V \to W$ we can see that, being careful about composing maps properly, this gives rise to $\varphi^* : W^* \to V^*$. Maps between vector spaces produce pullbacks between the dual spaces. With this in mind, we are prepared to discuss vector fields and differential forms.
(excercise on basis inducing a corresponding dual basis and $(V^*)^* = V$)

Fields

Vector Fields

Given how organically the concept emerged of a tangent vector being a map differentiating functions, it is natural to generalize the concept to vector *fields* on a manifold. A vector field will be a map which sends smooth functions to a directional derivative. Since we define our smooth functions using local coordinate functions on a chart, we can compute this derivative as

$$vf = \sum_\mu v^\mu \frac{\partial}{\partial x^\mu} f(x^1, ..., x^n)$$

where we will call the v^μ the components of the vector v written with respect to a basis $\{\partial_\mu\}$ of partial derivative operators with respect to coordinates on $\mathbb{R}^n$. The $\{\partial_\mu\}$ serve as a basis for vector fields in $\mathbb{R}^n$ in the sense that any function f on $\mathbb{R}^n$ can be taken to any of its possible directional derivatives by taking a unique linear combination

$$v^\mu \partial_\mu = \sum_\mu v^\mu \frac{\partial}{\partial x^\mu}$$

where the left side is short hand meant to represent the right side in accordance with Einstein summation notation. But in the spirit of generality, we must provide a coordinate independent definition of a vector field, which will naturally be more abstract. A **vector field** is a map $v : C^\infty(M) \to C^\infty(M)$ satisfying

1. $v(\alpha f + \beta g) = \alpha v(f) + \beta v(g)$

2. $v(fg) = v(f)g + v(g)f$

for $f, g \in C^\infty(M)$ and a scalar α.

The essence of a vector field as a derivative operator is captured by the second property, the so called Liebnitz property (essentially the chain rule). In this way, the presence of vector fields which differentiate functions is not contingent upon the details of a coordinate system. Given this construction, we can now generally define a tangent vector.

The vector v_p is a map taking f to its derivative at p, a number which would be acquired by feeding in p to the resultant function yielded by vf. We say v_p is a tangent vector at p and the set of all such vectors is the **Tangent Space** at p, T_pM. If we have an embedding of our n-dimensional manifold, we can identify the components v_p^μ of the vector field at a point p as the components of traditional vector in $\mathbb{R}^n$ and as such place a copy of $\mathbb{R}^n$ tangent to M at p and draw the corresponding arrow. In this way, we would get a field of arrows on our manifold. But to indulge this practice is to betray the essence of our new concept of a vector field; our definition provides a useful form of a vector that does not need an embedding of the manifold in higher dimensional Euclidean space. Nonetheless, it is nice to see that this advanced machinery is consistent with our old intuitions in special cases.

Form Fields

In order to motivate forms, it is valuable to begin by asking how we might try to perform integration on a manifold or a submanifold (which is merely a manifold whose points are contained in a greater manifold). The simplest example we might consider is integrating over a curve $\gamma : [0,1] \to M$. We want our integration to correspond to a summation of tiny scalar values, computed along a series of infinitesimal displacements conceptually given by the tangent vectors to γ. But we cannot merely sum vector values to get a scalar, so we see already the need for a map ω from tangent vectors to real numbers. This map ω assigns at each point in the manifold a map

$$\omega_p : T_pM \to \mathbb{R}$$

which lives in T_p^*M, the **cotangent space**, which is dual to the tangent space as its elements are linear maps from tangent vectors to scalars. These are also called **one-forms**, and just as we have vector fields, we can have form fields, which assign a covector in the cotangent space at each point in the manifold.

More concretely, we can consider the one-form df where f is a smooth function on M. Though there is delicate formalism which we

need to develop, we still intuitively identify df with the differential of f. We have that df is defined by

$$df(v) = vf$$

and so we can consider integration

$$\int_{\gamma} df = \int_{[0,1]} df(\gamma'(t))dt = \int_{[0,1]} \frac{d}{dt} f(\gamma(t))dt$$

where we use the notation on the right to indicate that we are performing a typical Riemann sum. So what occurs here is that we define our form df and it eats a vector, yielding a real number. Intuitively this is the infinitesimal change in f for an infinitesimal change in γ. That is to say, df is *not* the scalar value of small change itself, but a map which gives us this small change if we feed it a vector. By making these identifications, we see that forms are the things which we integrate over manifolds, or more generally, manifolds with boundary which we will define later. As such, forms merit a great deal of study and so we must proceed unpacking their machinery.

If we recall that a basis $\{e_i\}$ for a vector space V directly induces a basis $\{e^j\}$ for the dual space V^* then we recognize that their should be such a basis induced on the space of forms. Indeed, given the basis $\{\partial_\mu\}$ for vector fields on $\mathbb{R}^n$, we have a dual basis of forms $\{dx^\nu\}$ satisfying

$$dx^\nu(\partial_\mu) = \delta^\nu_\mu$$

where the symbol on the right is the Kronecker delta, which is 1 when $\mu = \nu$ and 0 otherwise. Being a basis, we can express any one-form in a chart as a linear combination

$$\omega = \omega_\mu dx^\mu$$

But we will not stop at mere one-forms, we must generalize in order to integrate over higher dimensional manifolds.

The reader should recall from multivariable calculus that when integrating over surfaces in $\mathbb{R}^3$ one would parameterize the surface

using a map $\varphi : [0,1]^2 \to \mathbb{R}^3$ and the derivatives with respect to each parameter would each define tangent vector. We would identify with the norm of the cross product of these tangent vectors a small area element on the surface, and we could then integrate functions over these infinitesimal areas. We can intuit that since a k-manifold will be parameterized by k-parameters, we need something that can eat k tangent vectors and give us scalars to sum over our manifold. Out of necessity we therefore develop k-forms.
(maybe note that mapping from unit square implies compact surface)

We need an operation which can allow us to build k-forms from our humble one-forms, and we will denote it by $\wedge$. But how can we get a unique operation to combine them? For one, there is a requirement we would like satisfied by these forms. Given our new formalism, our tangent vectors merely get eaten by the forms, so some of the work that was done by interesting operations between vectors must now be done by forms. In particular, 3-vectors had an algebra where the operation was the cross-product. We still want to have an analog of a cross-product in our new formalism. Therefore, like the cross product, we demand antisymmetry from our $\wedge$. That is, for any two $v, w \in V$ where V is the vector space of one-forms, we demand

$$v \wedge w = -w \wedge v$$

we call this operation the **wedge product** and with it we construct something called the **exterior algebra**, ΛV. We can build k-forms by wedging any two n and m forms together (satisfying $n + m = k$) and taking linear combinations of k-forms with functions as coefficients. This will now allow us to integrate over k-manifolds, but before we attempt to do so, we need to discuss another operation on forms and some additional structure on our manifold. **(rigorous definition with rules)** **(Prove $\omega \wedge \omega = 0$ as excercise)**
Remaining

We saw before that we wanted df to represent a one-form which emerged from a smooth function. Therefore we want d to be an operation which sends smooth functions to one forms. But really, we can identify functions as 0-forms, so we are motivated to make d a general

operation from k forms to $k+1$ forms. We will call this the **exterior derivative** and as a derivative we would like it to obey

1. $d(\alpha\omega + \beta\mu) = \alpha d\omega + \beta d\mu$ for forms ω, μ and scalars α, β

2. $d(\omega \wedge \mu) = d\omega \wedge \mu + (-1)^p \omega \wedge d\mu$ for p-form ω and one-form μ.

3. $d(d\omega) = 0$ for all forms ω

These properties uniquely define the directional derivative. We can write a simple algorithm for taking a general exterior derivative as a result of these properties. Consider a multi-index I which is a list of p indices, where each index corresponds to those labeling the coordinate functions in a chart (and hence the basis of forms and vector fields). We can write a p-form as

$$\omega = \omega_I dx^I$$

where the summation is over all possible *combinations* of p indices. We chose to say combinations because the dx^I for different orderings of indices in I can only differ by a sign, so we just combine the functions for different permutations into one ω_I coefficient. We then have that

$$d\omega = (\partial_\mu \omega_I) dx^\mu \wedge dx^I$$

Now, this generalization of the d operation may at first appear contrived or useless, but we again call upon the intuition that in our new framework of geometry, we are transferring some of what was physically expressed by operations on vectors, to operations on *forms*. Also, despite this rather involved generalization to k-forms, we can still preserve our old notion of d as representing small change.

(form wedge parallelipiped somewhere)

Consider a chart on an n-dimensional manifold M. We introduce a *fixed* index, which is just a collection $(i_1, \ldots, i_p)$ where the i's are ordered and correspond to the indexation of the basis of coordinate functions and hence one-forms in our chart. We now consider a p-form of the form

$$\omega = \omega_{(i_1,\ldots,i_p)} dx^{(i_1,\ldots,i_p)}$$

where

$$dx^{(i_1,\ldots,i_p)} = dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

We can consider

$$d\omega = (\partial_\mu \omega_{(i_1,\ldots,i_p)}) dx^\mu \wedge dx^{(i_1,\ldots,i_p)}$$

where the nonzero terms result in a sum over basis forms dx^μ not in the p-form dx^I (otherwise we have a repeated wedge of basis forms which equals 0). Now, consider a vector

$$v = v^\nu \partial_\nu$$

where the v^ν are only nonzero for ν not in the index. Then we see that we get

$$d\omega(v) = v^\mu (\partial_\mu \omega_{(i_1,\ldots,i_p)}) dx^\mu(\partial_\mu) \wedge dx^{(i_1,\ldots,i_p)} = (v\omega_{(i_1,\ldots,i_p)}) dx^{(i_1,\ldots,i_p)}$$

so the effect of this was to essentially get a directional derivative, the change, of the form's coefficient $\omega_{(i_1,\ldots,i_p)}$ in the directions excluded from the index.

(Diagram of this with parallelogram moving)

This example is clearly not a general result, but it is nonetheless comforting to see that a notion of exterior derivative as small changed can be recovered, and that we are *generalizing* our old intuitions, not completely trashing them.

Stoke's and Integration

We have now built up the geometric intuition and mathematical sophistication necessary to discuss a very powerful and general theorem in differential geometry: Stoke's Theorem. The statement concerns a form ω defined on a manifold with boundary, M and is a generalization of the many integration theorems taught to students in multivariable calculus courses. As such, we will motivate the theorem by reframing each of those old calculus theorems in our modern geometric language.

Surely the reader recalls the fundamental theorem of calculus

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

but how is this written in our modern language? First we recognize that the integrand

$$f'(x)dx = \partial_x f dx = df$$

and that the closed interval $[a, b]$, though not a manifold, does qualify as a manifold with boundary.

A manifold with boundary differs from a manifold in that we allow the charts to map to $H^n := \{(x^1, \ldots, x^n) \mid x^n \geq 0\}$. The requirement that transition functions be smooth becomes slippery around the boundary
(define this better)

So, denoting the manifold with boundary as M and the boundary, which is the endpoints of the interval, as ∂M, we have the fundamental theorem of calculus is

$$\int_M df = \sum_{\partial M} f$$

well this does not seem to have gotten us much, but lets see how the divergence theorem is reformulated in this modern language. Recall that for a vector field F on $\mathbb{R}^3$ that the divergence theorem states

$$\int_V \nabla \cdot \mathbf{F} dV = \int_S \mathbf{F} \cdot d\mathbf{S}$$

for some 3-dimensional volume V with boundary surface S. Again, V is a suitable manifold with boundary whose boundary is the surface S. We will again reframe this statement in accordance with the new formalism. Since we are taking some piece of information from F and integrating it over a two-dimensional surface, it is natural to look for a way to express F as a two-form. We will make the identifications

$$F_x dS_x \to F_x dy \wedge dz$$
$$F_y dS_y \to F_y dz \wedge dx$$
$$F_z dS_z \to F_z dx \wedge dy$$
$$S \to \partial M$$

and we see that the right side becomes

$$\int_{\partial M} F$$

This should make sense to the reader, as the wedge product conceptually emerged as a generalization of the cross product, and here the F_i component is the coefficient of the parallelogram normal to the ith direction. **(maybe diagram)** Now we reformulate the right side. Consider

$$dF = \partial_\mu dx^I \wedge dx^\mu = (\partial_x F_x + \partial_y F_y + \partial_z F_z)dx \wedge dy \wedge dz$$

and we see the coefficient is the traditional divergence! So the equality becomes

$$\int_M dF = \int_{\partial M} F$$

perhaps the reader notices a pattern emerging.

(Rewrite curl theorem in differential forms language excercise)

The generalization of this pattern is the elegant equality

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

in English, we are saying that the integral of the form over the boundary is equal to the integral of its exterior derivative over the manifold. Even in plain english, this may not be entirely intuitive yet. In order to understand why this equality should make sense, we should remain grounded in the concept that as we evaluate the integral of a form, we are summing small changes of the form, as convoluted as they may be in general. In this sense what Stoke's theorem is saying is the following: the sum of the small changes of a form *throughout* a manifold is equal to the sum of the form's values over the boundary of the manifold. That is, when we sum all the small changes throughout

the manifold, the result is what's left on the boundary, just the *net* change. We have packaged the intuition and mathematical machinery behind the various integration theorems in calculus, and distilled it into an impressively general and elegant statement.

Coord Transforms

Perhaps the greatest impetus for the use of this geometric formalism in physics is the notion that we need coordinate independent laws of physics. That is, when we apply a coordinate transformation on our manifold, we want the mathematical statements or equalities which we call physical laws, to say the same thing they did before. In the context of relativity, coordinate transformations correspond to changes of reference frame, and this condition says that we do not want the laws of physics to be different for two different observers. It should be no surprise then, that the study of coordinate transformations on manifolds merits a bit of attention.

There are two ways to apply coordinate transformations on a manifold, actively and passively. We will focus on the active way. Active coordinate transformations are represented by **diffeomorphisms**, which are isomorphisms of manifolds. If we have a map $\varphi : M \to N$, the condition for it to be a diffeomorphism is that φ is smooth with smooth inverse; this means φ can essentially pair each point in M with a point in N smoothly. And so, if two manifolds M and N are diffeomorphic, they are really the same manifold being represented in two different ways. Perhaps the reader is not convinced of this, but we will show that diffeomorphisms are equivalent to our old notion of coordinate transformation.

To make the equivalence of diffeomorphic manifolds more concrete we consider a conceptual example. Suppose I have a 2-sphere on which I draw a great circle to create two hemispheres and a natural choice of two poles which I call Northern and Southern. If I have a map switching each point p in the Northern hemisphere with the closest point q in the Southern hemisphere that has the same distance from the South pole as p has from the North pole (a sort of "reflecting" about the equator) **(figure)**, all we have done is exchanged the

two hemispheres. We have completely preserved the structure of the manifold, just changed the way the coordinate functions choose to talk about it. It should then be clear that the things of physical interest on our manifold are those unaffected by diffeomorphisms. First, we will consider how a vector v changes under a diffeomorphism.

We introduce manifolds M and M' which are diffeomorphic, so two different representations of the same thing, and denote objects defined on each space as primed and unprimed, respectively. As such, there exists a diffeomorphism φ which allows us to pushforward and pullback objects. We say the pushforward of v will be sent to v' and we know it obeys

$$v'(f') = (\varphi_* v)(f') = v(f' \circ \varphi)$$

but since this is a diffeomorphism, we are free to use φ^{-1} and say that f' is $f \circ \varphi^{-1}$, effectively the pushforward of the function f on M. And so

$$v'(f') = v(f \circ \varphi^{-1} \circ \varphi) = v(f)$$

and so the action of the vector field on functions is invariant under the diffeomorphism, and hence invariant under coordinate transformations. It is sensible that the action of vectors on functions is invariant, but *something* must be changing, right? What changes are the the coordinate vector fields and the vectors components; they do so in essentially opposite ways so that their changing cancels out to leave the *whole* object invariant. When we are doing actual physics later on, we are going to want to know how these components change. To do this, we will develop the tensor transformation law, but first we must define tensors.

Tensors

The starting point for defining tensors, is defining the **tensor product** of vector spaces. We begin with an abstract, coordinate independent definition. Given vector spaces $V^1, \ldots, V^n$ we can consider their cartesian product, $V^1 \times \cdots \times V^n$, merely the set of ordered pairs $(v^1, \ldots, v^n)$. We introduce a space $V^1 \otimes \cdots \otimes V^n$ whose elements are of the form

$$v^1 \otimes \cdots \otimes v^n$$

We define the tensor product by its universal property. That is, when given a multilinear map $\phi : V^1 \times \cdots \times V^n \to Z$ (a map linear in each "slot")

$$\phi(\alpha v^1 + \beta w^1, \ldots, \gamma v^n + \delta w^n) =$$
$$\alpha\gamma\phi(v^1, \ldots, v^n) + \alpha\delta f(v^1, \ldots, w^n) + \beta\gamma f(w^1, \ldots, v^n) + \beta\delta f(w^1, \ldots, w^n)$$

we can find a multilinear map $\varphi : V^1 \times \cdots \times V^n \to V^1 \otimes \cdots \otimes V^n$ and find that there exists a unique *linear* map $F : V^1 \otimes \cdots \otimes V^n \to Z$ such that

$$F(v^1 \otimes \cdots \otimes v^n) = f(v^1, \ldots, v^n)$$

We say that F is a map over the tensor product of these vector spaces. We can then see that a given bases $\{e_i^1\}, \ldots, \{e_i^n\}$ for these vector spaces, we have that $\{e_i^1 \otimes \cdots \otimes e_i^n\}$ must be a basis for the tensor product space. This definition appears devoid of intuitive motivation, and so we now try to make this construction seem more natural.

One perspective to take on what we have done in constructing our tensor product space is to interpret tensoring as introducing *coupling* between elements of the different vector spaces. Indeed this is a useful notion when applying tensor products to quantum mechanics. What we mean by coupling is that two tensored vectors $v \otimes v'$ behave now as a collective object, which is dependent on how each sub-object (each vector in the tensor product) changes. To illustrate this notion more explicitly, we will compare $V \otimes W$ and $V \times W$ for vector spaces V and W. If I have

$$(v_1 + v_2, w_1 + w_2) \in V \times W$$

(maybe even talk about how this is a direct sum)
I can decompose this as

$$(v_1, w_1) + (v_2, w_2)$$

so we can independently separate each component. I can separate vectors in V and in W separately without worrying about any coupling or interaction between vectors in the two spaces. This is why we used Cartesian products to describe velocity and acceleration vectors; we have the physical intuition that forces applied in the x direction of

motion should not affect my y direction of motion. Now consider

$$(v_1 + v_2) \otimes (w_1 + w_2) = v_1 \otimes \omega_1 + v_2 \otimes w_2 + v_1 \otimes w_2 + v_2 \otimes w_1$$
$$\neq v_1 \otimes w_1 + v_2 \otimes w_2$$

There are cross terms here emerging from the definition of our tensor product. These cross terms are a manifestation of the coupling, which is the essential property of the tensor product.

Concept 3.2. *The tensor product of vector spaces can be viewed as the space of coupled elements of each vector space. A map which is multilinear over the Cartesian product extends to a linear map over the tensor product because the tensoring the vectors packages them into a single, pre-coupled object.*

This perspective will prove useful in our use of tensor products in quantum mechanics, but in order to consider tensors in relativity, we will need to consider the tensor products of linear maps. Given a set of linear maps

$$L^i : V^i \to W^i$$

We have tensored maps

$$L^1 \otimes \cdots \otimes L^n : V^1 \otimes \cdots \otimes V^n \to W^1 \otimes \cdots \otimes W^n$$

which are linear over the tensor product space and obey

$$L^1 \otimes \cdots \otimes L^n(v^1 \otimes \cdots \otimes v^n) = L^1(v^1) \otimes \cdots \otimes L^n(v^n)$$

These maps can also be viewed as multilinear maps over the Cartesian product of the vector spaces.. We will now use this concept and the duality of vector spaces to construct tensor fields. Recall that given a vector space V, the dual space V^* is the space of linear maps from V to $\mathbb{R}$. As such, we can identify an element of

$$V^1 \otimes \cdots \otimes V^r \otimes V^{*^1} \otimes V^{*^s}$$

as a linear map

$$T : V^{*^1} \otimes \cdots \otimes V^{*^r} \otimes V^1 \otimes V^s \to \mathbb{R}$$

where we note that n tensor products of $\mathbb{R}$ is just $\mathbb{R}$. Such an object is called a rank (r, s) **tensor** . This construction straightforwardly extends to smooth tensor fields on manifolds.

Given a manifold M, we may consider the tangent and cotangent spaces at p, T_pM and T_p^*M respectively. These are vector spaces, and so we may consider objects in the tensor product of r copies of the tangent space and s copies of the cotangent space. A tensor field will assign an element in this tensor product space at each point p in the manifold M. Given our coordinate bases for the tangent and cotanget spaces in a manifold, we can expand a tensor as

$$T = T^{\mu_1...\mu_r}_{\nu_1...\nu_s} \partial_{\mu_1} \otimes \ldots \partial_{\mu_r} \otimes dx^{\nu_1} \otimes \cdots \otimes dx^{\nu_s}$$

where the smooth functions in front are the components of the tensor. Now that we have tensor fields on our manifold, we can formulate the tensor transformation law.

In order to formulate the tensor transformation law it is now more straightforward to work with passive coordinate transformations. In this case, we merely change the coordinate functions on $\mathbb{R}^n$ that we use in a chart. As we transform from the

(tensor transformation law)

(contraction etc.?)

(exterior algebra as anti-symmetric subspace of tensor product space)

Metric title

In order to put manifolds to use in physics, in particular general relativity, we need to define something called a metric. In the sense of stripping away old Euclidean assumptions and generalizing, the metric allows us to generalize the notion of a scalar or dot product. In particular, it gives us a way to talk about the magnitudes of vectors.

In Euclidean space we would multiply the components of two vectors in order to obtain their dot product, and not think much about it. As such, the magnitude of a vector was the square-root of the sum of the squares of its components. But implicit in this operation was a choice of metric, we are free to choose *how* we want infinitesimal distances to vary across points in our manifold. We will motivate the need for the distance function but making an effort to compute the length of a curve.

Consider two points p and q in a manifold. Call the function yielding the distance between them for a given path s. We then identify the exterior derivative ds with a tiny distance change along the curve. We would like to integrate the form ds in order to get a length for the curve $\gamma : [0,1] \to M$. In accordance with the integration formalism we have developed so far, we will write this as

$$\int_\gamma ds = \int_{[0,1]} ds(\gamma'(t))dt = \int_{[0,1]} \frac{d}{dt}s(\gamma(t))dt$$

So, in order to compute this integral, we need a way for ds to take a tangent vector and give us an infinitesimal change in distance along the curve, or a way to get the instantaneous rate of change of the distance function at each point along the curve. There is no *right* way to define the tiny changes in distance on an arbitrary manifold given only the structure that we have built so far. In fact, the way that small distances vary on a manifold is how we differentiate between manifolds which are topologically identical. As such, we need to define on our own how s or ds will determine these distances; we are actually going to define the one-form ds by running two copies of γ through a two-form g. We write

$$ds(\gamma'(t)) = \sqrt{g(\gamma'(t), \gamma'(t))}$$

and we call g the **metric tensor** on the manifold. It is the object which tells us how to measure tiny distances. Really, g is the metric tensor *field*, defined over the entirety of M. At each point p in M, g assigns a metric tensor $g_p : T_pM \times T_pM \to \mathbb{R}$, taking two vectors and returning a scalar. Before proceeding, we will now rigorously define a semi-Riemannian metric.

A semi-Riemannian metric is a bilinear map

$$g : V \times V \to \mathbb{R}$$

obeying

- $g(v, w) = g(w, v)$

- $g(v, w) = 0$ for all $w \in V$ if and only if $v = 0$

these conditions are called symmetry and non-degeneracy, respectively. Indeed, as g is bilinear in the Cartesian product, it naturally extends to being defined as a rank $(0, 2)$ tensor. We write

$$ds \otimes ds = g_{\mu\nu} dx^\mu \otimes dx^\nu$$

for the metric tensor in a basis of coordinate one-forms. As is the case with a tensor field, the metric components can vary from point to point on the manifold. It is this added Riemann structure on the manifold that lets us differentiate between the egg and the ping-pong ball! On the ping-pong ball, from a given point, the change in small distances is not uniform over directions, whereas they are on the sphere. Aside from its use as implied its motivation, the metric is a useful algebraic tool for index manipulation. We can get a dual vector from a vector and vice versa by "raising" and "lowering" with the metric.

(excercise showing $v^a g_{ab} = v_b$)
(Notion of tangent bundle?)

Derivatives Section

Up to now we have built up a great deal of structure on our manifold, but there still remain some physically useful notions that we do not have the capacity to discuss with the tools at hand. For example, we know how functions change as they move along curves, but what about a vector or even a tensor? We have no notion of how to propagate a vector along a curve because we have no notion of how to take a directional derivative of a vector. The mathematical apparatus which allows us to talk about these things is called a **connection** and with

it comes a **covariant derivative**.

Even given how many old Euclidean preconceptions we have blasted thus far, the reader might be tempted to think they can see *exactly* how you would move a vector along a curve, from one tangent space to another. If so, I would wager that the reader is visualizing their tangent vector, which so eagerly awaits propagation, in higher dimensional Euclidean space. Euclid too thought there was an obvious to move vectors around, but he was working in his space. And again, just as with the metric tensor, these beliefs implicitly restricted the mathematician to special cases of a more general geometric apparatus.

Euclid's fifth postulate stated that parallel lines never intersect. Working in $\mathbb{R}^n$ this seems clear, we just infinitely extended two arrows that point in the same direction, and they will never cross paths. But this is just a Euclidean intuition; how would we *carefully* address Euclid's claim? The claim is first trying to saying that we have two objects called lines, and at a point on each line, there is some form of information which can be compared telling us whether or not the lines are something called *parallel*. We can already see a breaking down of the question itself. How do we pick which points on each line to compare? Does it matter? What is the information and how are we comparing it? The desire for answers to these questions motivates the need for a connection; after all, the spirit of the name is that it allows us to *connect* different tangent spaces, cotangent spaces, or tensor products thereof. Holding these questions in the back of our minds, we will now proceed to define the object capable of answering them.

The desire to for a way to connect different tangent spaces and the desire for a way to determine how vectors infinitesimally change in some direction are dual. To see this, consider one perspective on what derivatives of single variable functions are: the limiting value of the difference of a function evaluated at two points (i.e. $f(x+\epsilon) - f(x)$) over the interval length. Our notion of small change of functions was and is contingent upon a way of connecting or comparing evaluations of f at separate points. There is no *right* way to pick an operator that yields a form of small change of a vector field, but we can sensibly construct one from the ground up.

First, we should think about sensible properties are derivative operator should have. For one, it will need to tell us the derivative in the direction of some vector field, in particular mapping a vector field to a vector field. We will want it reduce to the typical directional derivative when acting on a function and, as any good derivative operator, have the Liebniz property. Properties of linearity are of course desirable and necessary as well. Therefore we have a map

$$D_v : \text{Vect}(M) \to \text{Vect}(M)$$

which obeys

1. $D_{fv}(\alpha g(u \otimes t)) = f\alpha\,(v(g) + u \otimes D_v(t) + t \otimes D_v(s))$

2. $D_{v+w}(u) = D_v(u) + D_w(u)$

where α is a scalar, f is a function, and v, w, s, and t are vector fields. A final property we require is that the operator is natural, that is, it commutes with diffeomorphisms. This stipulation says that the operation of taking a covariant derivative should be coordinate independent.

With these requirements in mind, we now work with our covariant derivative in a coordinate chart with the standard basis of coordinate vector fields. We have

$$D_v w = v^\mu(w^\nu D_{\partial_\mu}\partial_\nu + D_{\partial_\mu}(w^\nu)\partial_\nu) = v^\nu((\partial_\mu w^\nu)\partial_\nu + D_\mu\partial_\nu)$$

where D_μ is short hand for D_μ and we used the derivative rules since the vector components are functions and we can employ the Liebniz rule.
(tensor product of function and vector is just multiplication)

The remaining part of this equation is the derivative at the end

$$D_\mu\partial_\nu$$

given the linear nature of this map, we can say

$$D_\mu : \partial_\nu \to \Gamma^\lambda_{\mu\nu}\partial_\lambda$$

this should seem completely fair to the reader, as the result of mapping ∂_ν to another vector field is some linear combination of basis vector fields. The set of functions denoted by $\Gamma^\lambda_{\mu\nu}$ are the λth expansion coefficient of the result of differentiating the νth basis vector field in the μth direction. These coefficients also provide another essential property: they make sure the covariant derivative operator is natural. The coefficients might also be interpreted as corrections to the normal derivative operator for generalized manifolds, and as such they vanish for flat space. Returning to our original expression, we then have

$$D_v w = v^\mu \left((\partial_\mu \omega^\nu)\partial_\nu + \omega^\nu \Gamma^\lambda_{\mu\nu}\partial_\lambda) \right)$$

now, we are going to rearrange this into a way common for the use of physicists, careful to note the manipulations we are making and why. It should be clear that the above implies

$$D_\mu w = (\partial_\mu \omega^\nu)\partial_\nu + \omega^\nu \Gamma^\lambda_{\mu\nu}\partial_\lambda$$

and we can now focus on the σth component of the result of the above, which will be the coefficient of ∂_σ in the expansion. As such, we get

$$(D_\mu w)^\sigma = \partial_\mu w^\sigma + \Gamma^\sigma_{\mu\nu}\omega^\nu$$

this is how physicists usually write the operation of a covariant derivative. They in fact go even further, and take the formalism of saying that

$$D_\mu : w \to \left(\partial_\mu w^\sigma + \Gamma^\sigma_{\mu\nu}\omega^\nu \right)\partial_\sigma \otimes dx^\mu$$

and in general say that the covariant derivative sends (k, l) tensors to $(k, l+1)$ tensors. The convenience of using the right side of the above mapping is that we can now feed this object a vector field v and get $D_v w$, we don't have to feed D_μ the components. **transition**

This object Γ currently has an uncomfortable amount of mysteriousness given its importance to taking covariant derivatives. Indeed these symbols are the deciders of *how* we take derivatives on our manifold. In gauge theory this object is identified with the vector potential and its components are in the gauge group. In general relativity these are called Christoffel symbols and are consequence of choosing the Levi-Civita connection. In order to discuss parallel transport. **(parallel propagation? Or save for GR)**
(Lie derivative-¿ what's the difference between the two?)

Chapter 4

Beyond Harmonics: Representation Theory

Part 2

Physics

Chapter 5

Symmetries of the sphere: $SU(2)$ and friends

Chapter 6

Classical Mechanics and Symplectic Geometry

Chapter 7

Einstein's Gravity

Part 3

More Advanced Topics

Chapter 8

An Introduction to Quantization

Chapter 9

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

Index