Differential Geometry

R.C. Dennis

1 Intro

So far in your physics education you've become quite adept at manipulating scalars and vectors in three dimensions. In fact, you've become capable of manipulating them in any dimension you please and have even learned about how functions can comprise an infinite dimensional vector space themselves (e.g. polynomials, Fourier, Bessel, etc).

You first learned about vectors in the (x, y) plane and were even told you're free to move them around so long as you don't change their lengths or directions. Typically we put all of our vectors at the origin and added/subtracted them at our leisure.

On top of all of that we similarly learned that in some cases it's more practical to work in polar or even spherical coordinates. For example, with gravity it's very pragmatic indeed to write $\hat{r}$ instead of

$$\hat{r} = \sin\theta\cos\phi\hat{x} + \sin\theta\sin\phi\hat{y} + cos\theta\hat{z}$$

writing $\hat{r}$ is indeed much easier in such problems!

You've also learned a couple ways to "multiply" vectors such as

$$\vec{a}\cdot\vec{b} = \sum_i a_i b_i$$

and

$$\vec{a}\times\vec{b} = \det\begin{bmatrix} \hat{x} & \hat{y} & \hat{z} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix}$$

Maybe you've even learned about the Kronecker/tensor product of two vectors

$$\vec{a}\otimes\vec{b}=\begin{bmatrix} a_x b_x & a_x b_y & a_x b_z \\ a_y b_x & a_y b_y & a_y b_z \\ a_z b_x & a_z b_y & a_z b_z \end{bmatrix}$$

Everything we did we these tools was quite straightforward but nobody probably ever told you where these somewhat arbitrary rules came from. For example, why are we free to detach vectors from their home and move them where ever we please? What even is a vector? Why does the dot product give us a real number but the cross product a (pseudo)vector?

To begin our studies of more general things we must completely strip bare everything we thought we knew about vectors and coordinates and start afresh. Our initial route will be horrifically cumbersome and seemingly pointless but the payoff will be legion. Let's begin:

Where do things happen?

When we solve a physical problem we often begin with with the notion of $3D$ space and perhaps 1 dimension of time. We imagine vectors as these things that live in the $3D$ space point in all sorts of different directions representing various concepts like force or velocity. Let's be careful about this stage where everything happens.

1. Let $\mathcal{M}$ be $\mathbb{R}^D$

here $\mathbb{R}^D$ means

$$\mathbb{R}^D=\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\dots$$

This is the Cartesian product of the real number line. A single point in $\mathbb{R}^D$ is the ordered tuple $(x,y,z,\dots)$: a letter for each number line. Usually in physics we work in $\mathbb{R}^3$ using (x,y,z) Cartesian coordinates. Nothing too surprising, just jargon.

Next, we can define a function f such that

$$f:\mathbb{R}^D\to\mathbb{R}$$

This is a function $f(x,y,z,\dots)$ that takes as input D elements and outputs a single number. For example, the topography of some area of the Earth would be a function

$$h:\mathbb{R}^2\to\mathbb{R}$$

that gives the elevation of every point (x, y). Viz, $h(x, y) = H$ where H is the elevation - evidently a real number.

Now, if we can build functions then we certainly can build all sorts of arbitrary black boxes that take some subset of $\mathbb{R}^D$ into some other subset of $\mathbb{R}^D$. For example

$$T : U \in \mathbb{R}^D \to W \in \mathbb{R}^D$$

This is about the most general type of machine/function we can define in $\mathbb{R}^D$. For example, at every point $P \in \mathbb{R}^D$ we might want to know the temperature, mass, energy, and charge at that point (x, y, z). This would be a function from $\mathbb{R}^3$ to $\mathbb{R}^4$. We're certainly allowed to do this.

Next, given that our manifold $\mathcal{M} = \mathbb{R}^D$ and given that we know how to do some multivariable calculus with real numbers let us define a second manifold, related to our original - in fact, we will simply make a copy of our $\mathcal{M}$. Now we have a pair of $\mathbb{R}^D$. Let's call the new one $\mathcal{M}'$.

Because they're identical then for every point in $\mathcal{M}$ I have a corresponding point in $\mathcal{M}'$. Now, I will place them one on top of the other and make a geometric observation: everywhere $\mathcal{M}'$ is exactly tangent to $\mathcal{M}$. For definiteness consider the $\mathbb{R}^2$ plane. If i lay down another on top of it then they're certainly tangent to each other!

Now, for a point $P \in \mathcal{M}$, the original manifold, I will define a tuple $\vec{v} = (v_x, v_y, \cdots)$ in $\mathcal{M}'$. That is, I've established a mapping from one space to the other

$$v : \mathcal{M} \to \mathcal{M}'$$

This mapping takes $(x, y, z, \cdots)$ and produces a tuple of numbers $(v_x, v_y, v_z, \cdots)$.

I shall give $\mathcal{M}'$ a new name: at $P \in \mathcal{M}$ I call $\mathcal{M}' = T_P\mathcal{M}$. We call $T_P\mathcal{M}$ the "tangent space at P" to $\mathcal{M}$. Overkill, I know!

We have two manifolds now: $\mathcal{M}$ and $T_P\mathcal{M}$. We say the tangent space at P is the manifold $T_P\mathcal{M}$. Which is to say that at some $P \in \mathcal{M}$ I can assign a D-tuple from $\mathbb{R}^D$ at that single point P.

But if I can do this for point P then let's do this for EVERY point in our original manifold! We imagine that "above" each point in $\mathcal{M}$ we have an entire universe of possibilities. In fact we can be even more general than having a $\mathbb{R}^D$ at each $x \in \mathcal{M}$ - it could be ANY cartesian product of the real line at each point. Heck, we could even have the complex numbers $\mathbb{C}$ at each and every point or even any cartesian product of those!

For example, we could certainly say for each $x \in \mathcal{M}$ we define the space $\mathbb{R}^M$ where $M \neq D$. A basic function f is cast in a new light - for each $x \in \mathcal{M}$ we attach a real number line. Then a function f is simply a value of that attached number line at each point $x \in \mathcal{M}$ - we will call this a "section" of f.

Think about it this way: suppose you have a string and at each point on the string you attach another string. Haven't you effectively just created a sheet? Sure, but thinking about it this way opens us up to all sorts of new geometries. The quantum wavefunction of an object is a section (a single complex number at every point) of the complex numbers attached to every point in (x, y, z) space.

Or we might take a cylinder and attach a mobius strip at every point - I'm not sure what on earth that would be but it certainly might work from our definitions. This type of construction is called a fiber bundle and is one of the more beautiful ideas in mathematics.

Complexities aside let's go back to our boring $\mathcal{M} = \mathbb{R}^D$ and our shiny new $T_P\mathcal{M}$

Now we need to make a couple definitions

Vector Space

A vector space is any set V such that there exists

1. an "addition" map (e.g. $\vec{v} + \vec{w} = \vec{u}$)

$$\oplus : V \times V \to V$$

2. Scalar Multiplication Map (e.g. $5\vec{v}$)

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

In addition we require ALL of the following axioms to be satisfied

1. Commutativity of vector addition

$$\vec{v} \oplus \vec{w} = \vec{w} \oplus \vec{v}$$

2. Associativity of vector addition

$$\vec{v} \oplus (\vec{w} \oplus \vec{u}) = (\vec{v} \oplus \vec{w}) \oplus \vec{u}$$

3. Existence of additive inverses

$$\vec{v} \oplus \vec{w} = \vec{0}$$

4. Identity for scalar multiplication

$$1 \odot \vec{v} = \vec{v}$$

5. A left identity for addition

$$\vec{0} \oplus \vec{v} = \vec{v}$$

6. associativity of scalar multiplication

$$(\alpha\beta) \odot \vec{v} = \alpha \odot (\beta\vec{v})$$

7. distributivity of scalar multiplication over vector addition

$$\alpha \odot (\vec{v} \oplus \vec{w}) = \alpha \odot \vec{v} \oplus \alpha \odot \vec{w}$$

8. distributivity of scalar multiplication over scalar addition

$$(\alpha + \beta) \odot \vec{v} = \alpha \odot \vec{v} \oplus \beta \odot \vec{v}$$

ANYTHING that satisfies these axioms IS a vector space. This definition has NOTHING to do with numbers per se. The addition and scalar multiplication may also have nothing to do with our everyday plus and multiplication.

We can show that in $\mathbb{R}^D$ then $\oplus = +$ and so forth. But we will be nit-picky in that $\oplus$ acts on vectors which are fundamentally different than scalars.

Now, we ask " is $\mathcal{M}$ and $T_P\mathcal{M}$ vector spaces?"

Now, $\mathcal{M}$, in our case, is identical to $\mathbb{R}^D$ and $\mathbb{R}^D$ is certainly a vector space (check). But suppose we didnt know $\mathcal{M} = \mathbb{R}^D$. Suppose $\mathcal{M}$ was just some general space and we had no idea how it was connected together. Perhaps it was curvy and crazy but still continuous and therefore we might now have a way to assign numbers from $\mathbb{R}^D$ to each point.

We made this HUGE assumption about our everyday (x, y, z) manifold!

Consider a point $P \in \mathcal{M}$ and ask "what shall I use as basis vectors?"

In order to define basis vectors we are fundamentally required to choose coordinates at that point. It seems most natural to choose (x, y, z) coordinates such.

So we ask, for a general $\mathcal{M}$ can we add "points"? No, not at all, I'm not even sure what a point even is in this context. In general, our manifold is some smooth space but I don't know how to add/subtract different parts of it so therefore we should not be so quick to describe it as a vector space!

What about $T_P\mathcal{M}$? Well here we are certainly free to say that $T_P\mathcal{M}$ is $\mathbb{R}^M$ - geometrically it is nothing more than a tangent space to a particular point on $\mathcal{M}$ so we can imagine even if our manifold is curvy at P the tangent space would still be $\mathbb{R}^M$ where M may not be equal to D

Now $T_P\mathcal{M} = \mathbb{R}^M$ does comprise a vector space so we hope we can choose basis vectors with which to expand our vectors which live in $T_P\mathcal{M}$ but it's not clear how.

Let's go back to our manifold $\mathcal{M}$. Remember when we said "Let $\mathcal{M}$ be $\mathbb{R}^D$"? We were too hasty. Let's think about what this means.

Suppose we didn't know that $\mathcal{M}$ could be $\mathbb{R}^D$ and think about what it means to claim "$\mathcal{M}$ IS $\mathbb{R}^D$"

This means we can set up coordinates in just such a way as to cover every single point in our manifold $\mathcal{M}$. It's as though $\mathcal{M}$ is its own universe and suddenly we're teleported into it and we happen to have three rulers which we set up perpendicular to each other and then extend them infinitely in three perpendicular directions.

Is this guaranteed to work? NO! It fundamentally depends on if $\mathcal{M}$ is such that this is possible! If you're plopped down onto a spherical manifold then setting up your rulers this way would almost be meaningless since the manifold curves away from your flat sticks! You couldn't possibly hope to cover the whole sphere with just two rulers!

So, when we claimed "$\mathcal{M}$ IS $\mathbb{R}^D$" we were imposing WAY too much upon our manifold. What we should have done instead is say "I shall choose coordinates for this manifold such that every part of the manifold has an address".

There is an infinity of ways we could choose coordinates - just like you learned in intro physics - but if we make a choice then we have an address for each point. And if we have that address for each point then we have at our disposal a tangent space that IS a vector space at each point.

Let us define coordinates for our manifold as

$$\mathbb{X} : \mathcal{M} \to \mathbb{R}^D$$

$\mathbb{X}$ takes a generic point, say ξ, and spits out an address using real numbers.

For our everyday manifold universe we use in intro physics this map is

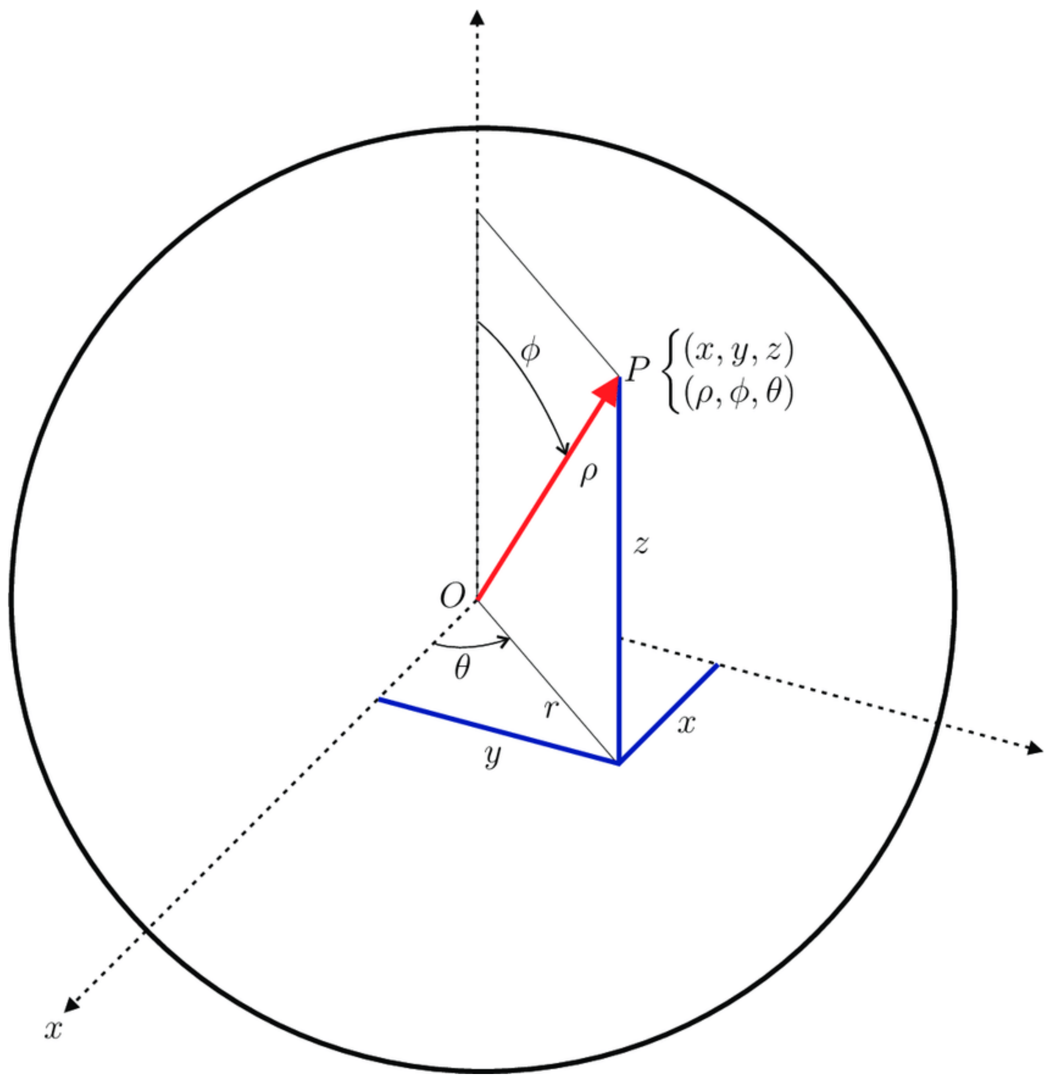
$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

This takes a point in $\mathcal{M}$ and spits out an address that WE CHOSE to be $\mathbb{R}^3$. Massive overkill dear reader but this idea allows us to work on really weird manifolds with ease.

Now, please keep in mind, we could've chosen $\mathbb{X}$ to be anything else we wanted. Sometimes we choose spherical coordinates for our "intro physics" manifold such that

$$\mathbb{X}(\xi) = \begin{bmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{bmatrix}$$

So we can impose almost any kind of $\mathbb{X}(\xi)$ we wish so long as our manifold allows is.



So what would it mean for our manifold to "allow" these choices? Well, we know in the manifold we usually play around in in intro physics Pythagoras must hold so between any two points $(\xi, \zeta) \in \mathcal{M}$ we require

$$\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$$

where s is the straight line length between (ξ, ζ). We could have an equivalent expression in spherical coordinates or any other coordinates we choose. Pythagoras is ultimately what allows us to lay down three perpendicular rulers and extend them to infinity - other manifolds may not let us (for example the 2-sphere!).

Okay, but despite all of this $\mathcal{M}$ is still technically not a vector space. Now let us study the vector space $T_P\mathcal{M}$.

Why might $T_P\mathcal{M}$ be a vector space exactly? Let's think about what this vector space is.

We choose a point in $\mathcal{M}$ and just like we drew a tangent LINE to a curve in intro calculus we "draw" a space tangent to a point $P \in \mathcal{M}$. Potentially, if we move to a different point $P' \in \mathcal{M}$ this tangent space might tip and turn depending on the manifold. We say the tangent line to a curve is a "linear approximation at P to the curve/function $\gamma(x)$".

Similarly $T_P\mathcal{M}$ is a linear approximation to $\mathcal{M}$ at P. Since it's a linear approximation is must be inherently $\mathbb{R}^M$ which is a bonafide vector space.

It's the linearity of $T_P\mathcal{M}$ that gives us a clean vector space to work with. $\mathcal{M}$ could be wacky as all hell but each linear approximation to a point will be a vector space.

So, given our vector space at P we should be able to define a set of basis vectors at that point P - but how?

Obviously we chose

$$\hat{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\hat{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\hat{z} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

But why these particular vectors? It always "felt" obvious but what's going on behind the scenes? our manifold isn't necessarily a vector space and these vectors "live in" $\mathbb{R}^3$ in our everyday example.

Let's remind ourselves what we did. Ye holy gods plopped us down into a $3D$ manifold and we set up our rulers/coordinates by applying

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(x, y, z) gives an address from $(0, 0, 0)$ which we take to be our "origin" - where the gods placed us. Next we said "at P there is a tangent space which comprises a vector space". In our coordinates P is located at, say (x, y, z) and at (x, y, z) we decided to use basis vectors for $T_P\mathcal{M}$ called $\hat{x}$ and so on.

How are these basis vectors related to the coordinates on our manifold?

well,

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\frac{\partial \mathbb{X}}{\partial x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \hat{x}$$

$$\frac{\partial \mathbb{X}}{\partial y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \hat{y}$$

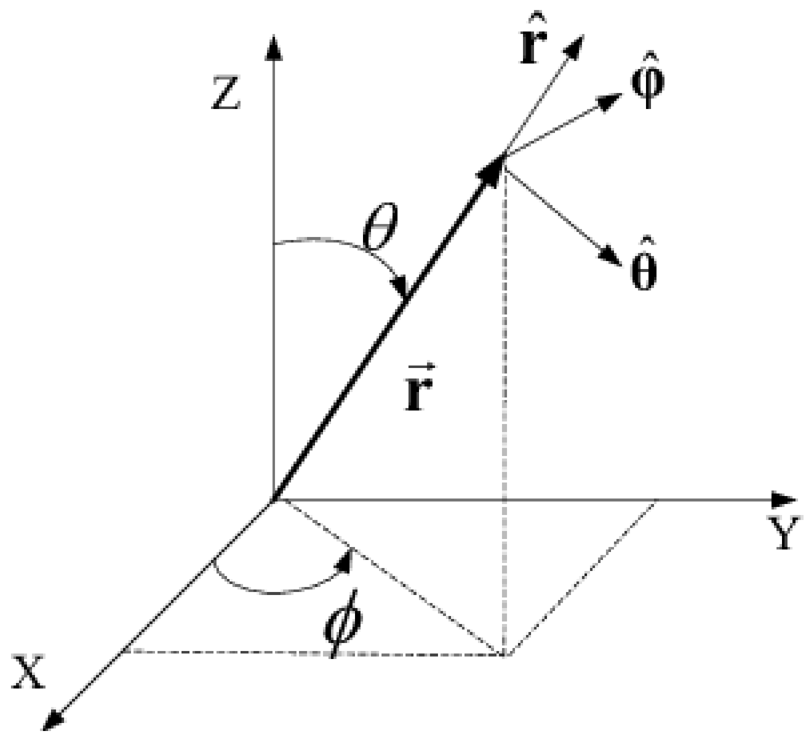
$$\frac{\partial \mathbb{X}}{\partial z} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \hat{z}$$

The basis vectors we CHOSE for $T_P\mathcal{M}$ depend on our choice of coordinates $\mathbb{X}$. If instead we used our spherical coordinates then the basis vectors at P would look much different.

$$\mathbb{X}(\xi) = \begin{bmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{bmatrix}$$

$$\hat{r} = \frac{\partial \mathbb{X}}{\partial r} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}$$

and similarly for $\hat{\theta}, \hat{\phi}$



Your whole life you've been defining you basis vectors in this manner without realizing that these basis vectors live and breath in the tangent space rather than your universe's manifold.

Now, that we have an honest to goodness vector space then we can sprinkle vectors all around our manifold: the vectors are inside their tangents spaces. Suppose we have a point in $P \in \mathcal{M}$ and a different point $P' \in \mathcal{M}$.

How can we compare vectors in $T_P\mathcal{M}$ and in $T'_P\mathcal{M}$? These are two different vector spaces. Well in our (x, y, z) you simply moved the vector from P' and put it at P and added/subtracted and did whatever you had to do.

Is this justified? It's only justified if we can drag the vector from P' to P and be absolutely sure its components don't go all crazy. Which is to say the basis vectors at P and P' MUST be identical. Is that the case in our (x, y, z) world?

well, at P we have

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and at P' we have

$$\mathbb{X}(\zeta) = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

Next, what are the basis vectors at each point
At P we have

$$\frac{\partial \mathbb{X}}{\partial x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \hat{x}$$

at P' we have

$$\frac{\partial \mathbb{X}}{\partial x'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \hat{x'}$$

but $\hat{x'} = \hat{x}$ so that we are welcome to slide vectors around and compare them anywhere and everywhere without fear that their components change!
But this is only true for our

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

is it also true for

$$\mathbb{X}(\xi) = \begin{bmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{bmatrix}$$

Check whether we can slide vectors around in spherical coordinates in this $\mathcal{M} = \mathbb{R}^3$ world.

2 Recap

Now, let's recap what we've done:

1. There exists a space called a manifold $\mathcal{M}$ that "locally" looks like $\mathbb{R}^D$
2. We impose a set of coordinates on $\mathcal{M}$
3. At every point $x \in \mathcal{M}$ in this space we define a tangent space called $T_x\mathcal{M} = \mathbb{R}^D$ which is a linear approximation to $\mathcal{M}$ at x.
4. The basis vectors in $T_x\mathcal{M} = \mathbb{R}^D$ are given by derivatives of our chosen coordinates on $\mathcal{M}$
5. "Vectors" live in $T_x\mathcal{M}$. A vector at x is different than a vector at y - they live in different tangent spaces.

As a fiber bundle $\mathcal{M} \times T_x\mathcal{M}$ a vector at some particular point in $\mathcal{M}$ is a tuple $(x, y, z; v^x, v^y, v^z)$ specifying the location and components of a vector.

$$\vec{v} = v^x\hat{x} + v^y\hat{y} + v^z\hat{z} = \begin{bmatrix} v^x \\ v^y \\ v^z \end{bmatrix}$$

Where we have expanded $\vec{v}$ according to our coordinate map $\mathbb{X}$

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Since our vectors are always written in terms of a basis regardless of the form of $\mathbb{X}$ we say that at some point P

$$\frac{\partial}{\partial x^i} = e_i$$

This statement is independent of the coordinate choice. When our manifold is our everyday $\mathcal{M} = \mathbb{R}^D$ then making the choice

$$\mathbb{X}(\xi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

yields the possibility that at every point $\xi \in \mathcal{M}$ the tangent spaces are identical.

$$e_x(\xi) = e_x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore if we have a vector at (x, y, z) then we can compare it with a vector at (x', y', z') since the two separate tangent spaces are identical in everyway.

$$e_i \in T_x\mathcal{M} = \mathbb{R}^D$$

and

$$e_i \in T_{x'}\mathcal{M} = \mathbb{R}^D$$

fundamentally $T_x\mathcal{M} = T_{x'}\mathcal{M}$

This fact needn't be true at all - it just so happens this is true for what we usually call our (x, y, z) world. We say this manifold is "flat" or Euclidean.

If it weren't true then $T_x\mathcal{M} \neq T_{x'}\mathcal{M}$ so that vectors at x could not be compared with vectors at x'. We would have to do some heavy work in order to find some well defined way to compare vectors at two different points. The only reason we've gotten away with sliding vectors around willy-nilly was a happy accident. From our more abstract and principled point of view there would be no obvious reason why vectors could be dragged from tangent space to tangent space!

So given a manifold $\mathcal{M}$ we discovered infinitely many tangent spaces $T_x\mathcal{M}$.

A vector field on $\mathcal{M}$ is defined as a specification of $(x, y, \cdots; v^x, v^y, \cdots)$ from $\mathcal{M} \times T_x\mathcal{M}$ for all points of the manifold.

There's another vector space we can build from $\mathcal{M}$. It's a vector space you've always used without any thought to it. In your physics classes you've had column vectors (those guys that live in the tangent space) and "row vectors". The row vectors were easy-peasy, you just transpose your column vector to get your row vector.

But now, since we're being enormously picky we realize rows and columns are two different beasts! Column vectors are easy to write. given our standard coordinate choice $\mathbb{X}$

$$\vec{v} = v^x\hat{x} + v^y\hat{y} + v^z\hat{z} = \begin{bmatrix} v^x \\ v^y \\ v^z \end{bmatrix}$$

if we transpose $\vec{v}$ we get

$$\vec{v}^T = \begin{bmatrix} v^{xT} & v^{xT} \\ v^{xT} & \end{bmatrix}$$

But what is this thing? it doesn't have the basis vectors from $T_x\mathcal{M}$ - if it did it would be written as a column of elements

What could

$$v^{Tx}\hat{x}^T + v^{Ty}\hat{y}^T + v^{Tz}\hat{z}^T$$

possibly be? What is the purpose of transposing a vector anyhow?

Well, one reason we transpose a vector is so we can form the standard dot product

$$\vec{a} \cdot \vec{b} = \vec{a}^T\vec{b}$$

$$\begin{bmatrix} a^{Tx} & a^{Ty} \\ a^{Tz} & \end{bmatrix} \begin{bmatrix} b^x \\ b^y \\ b^z \end{bmatrix} = a^{Tx}b^x + a^{Ty}b^y + a^{Tz}b^z$$

But we know we dont have to write a^{Tx} everywhere - let's write it $a^{Tx} = a_x$: the subscript means we've "done something" to the component of a vector with an upper index.

Let's abstract what this row vector does.

A row vector eats a column vector and produces a real number. Lets call this row vector ω_P to be definite. It is a machine that eats vectors and spits out numbers. Therefore

$$\omega : T_x\mathcal{M} \to \mathbb{R}$$

such that

$$\omega(\vec{v}) \in \mathbb{R}$$

Okay, this is weird but it's certainly how a row vector behaves. Let's consider all the gory details written out with the tangent space's bases vectors.

$$\vec{b} = b^x \frac{\partial \mathbb{X}}{\partial x} + b^y \frac{\partial \mathbb{X}}{\partial y} + b^z \frac{\partial \mathbb{X}}{\partial z}$$

That's just how we defined our unit vectors - but remember, the basis vectors are this way for ANY $\mathbb{X}$, so let's write

$$\vec{b} = b^x \frac{\partial}{\partial x} + b^y \frac{\partial}{\partial y} + b^z \frac{\partial}{\partial z}$$

Kind of weird but so be it: you give me a coordinate choice then this would be the vector. Now, what should we write for the row vector?

$$\omega_a = a^{Tx}\hat{x}^T + a^{Ty}\hat{y}^T + a^{Tz}\hat{z}^T = a_x\hat{x}^T + a_y\hat{y}^T + a_z\hat{z}^T$$

where we made use of our lower index notation. Now, we need the "bases" $\hat{x}^T$. Since we know what $\hat{x}$ is then evidently

$$\hat{x}^T = \left(\frac{\partial}{\partial x}\right)^T$$

Lets put it all together!

$$\vec{a} \cdot \vec{b} = a_x b^x \left(\frac{\partial}{\partial x}\right)^T \frac{\partial}{\partial x} + a_y b^y \left(\frac{\partial}{\partial y}\right)^T \frac{\partial}{\partial y} + a_z b^z \left(\frac{\partial}{\partial z}\right)^T \frac{\partial}{\partial z}$$

and this MUST somehow be

$$= a_x b^y + a_x b^y + a_z b^z$$

So we demand that

$$\left(\frac{\partial}{\partial x}\right)^T \frac{\partial}{\partial x} = 1$$

This looks a LOT like the transposed basis vector is the "inverse" of the basis vector. Let us write

$$\left(\frac{\partial}{\partial x}\right)^T = dx$$

This is apt as the differential dx feels like the inverse of the derivative. Then

$$dx(\frac{\partial}{\partial x}) = 1$$

or, using our notation for the basis vectors,

$$e^i e_i = 1$$

$$e^i = dx^i$$

$$e_i = \frac{\partial}{\partial x}$$

e^i is one of these weird objects that eats a vector and spits our a number. A row vector can be written as

$$\omega = \omega_x dx + \omega_y dy + \omega_z dz = \begin{bmatrix} \omega_x & \omega_y \\ \omega_z & \end{bmatrix}$$

Where the heck do these things live? Certainly not in $T_x\mathcal{M}$. Let us suppose they live in a "dual" space we shall call the cotangent space $T_x^*\mathcal{M}$. In our case these two spaces are related by the transpose operation but it could potentially be more general than that.

Notice that these co-vectors ω eat vectors to make numbers. We should write the co-vectors first. Also notice that the way we set things up we have

$$e^i e_j = \delta^i_j$$

We've even, inadvertently, discovered a way to measure the lengths of our $T_x\mathcal{M}$. We know the right thing to do is dot the vector with itself and take the square root.

$$|\vec{v}|^2 = \vec{v} \cdot \vec{v} = v_x v^x + v_y v^y + v_z v^z$$

$$= \sum_j v_j v^j = \sum_{kj} \delta_{jk} v^k v^j$$

Where in the last line we put in by hand something fundamentally new: the ability to "lower an index"

$$v_j = \sum_k \delta_{jk} v^k$$

In our everyday space the tangent and cotangent spaces are related in this fashion. We could've written something more general like

$$v_j = \sum_k g_{jk} v^k$$

But then by studying the distance between two points in our (x, y, z) space we would quickly discover that it MUST be δ_{ij}. Pythagoras says

$$\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 = \Delta \vec{r} \cdot \Delta \vec{r}$$

If we had some more general g_{ij} then we wouldn't have our familiar triangles and circles we're so accustomed to!

Incidentally, by being even more careful we quickly find that

$$e_i \cdot e_j = g_{ij} = \delta_{ij}$$

which is to say our basis vectors of the tangent space allow us to measure lengths. We call this $g_{ij} = \delta_{ij}$ the "Metric" on $\mathcal{M}$

Similarly we could use the co-vector basis

$$e^i \cdot e^j = g^{ij} = \delta^{ij}$$

My goodness, our everyday (x, y, z) is MUCH more rich than we could've ever imagined. Given a manifold we can get two other spaces out: the tangent space at a point and the co-tangent space at a point. We can also have scalar fields, vector fields, and now co-vector fields as well as the ability to measure lengths and drag vectors around from one place to another!

So, to recap, in our boring everyday manifold we call "Euclidean space" we have at some point P

1. A system of coordinates $\mathbb{X}$ we choose
2. $T_P\mathcal{M}$ - the tangent space
3. $T_P^*\mathcal{M}$ - the cotangent space
4. basis vectors $e_i = \frac{\partial}{\partial x^i}$ from $T_P\mathcal{M}$
5. basis co-vectors $e^i = dx^i$ from $T_P^*\mathcal{M}$

6. The ability to measure lengths using a "metric" g_{ij}

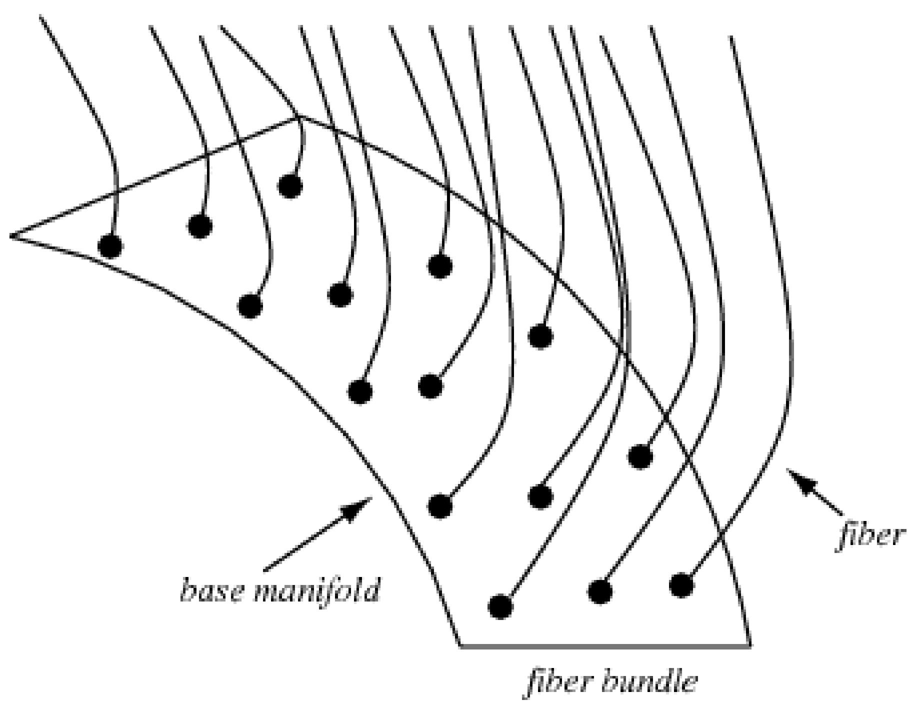
7. general scalars, vectors, and co-vectors

8. The ability to "connect" (or add/subtract vectors) from point P to point P' without any issue or consequence.

From analyzing our own boring (x, y, z) manifold we've inadvertently discovered a wealth of possible generalizations.

Our manifold could be anything and we would still be able to build these concepts - although the last item: comparing vectors, might be very challenging to develop.

Furthermore, given some manifold we've suddenly unlocked infinite possibilities using the notion of a fiber bundle.



At each point in our manifold we can have entire vector spaces at that point and they could potentially twist around in unexpected ways!

The mobius strip is an example: we have a circle and at each point of the circle we attach a real number line - but this number line fundamentally twists as you traverse the circle.

A fiber bundle is the following map (called a projection)

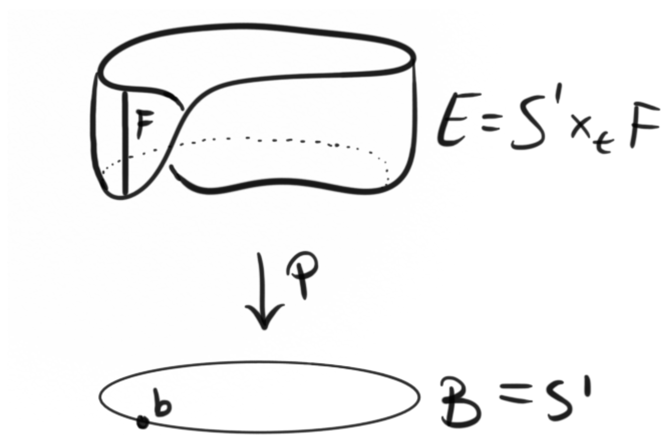
$$\pi : E \to B$$

Where E is called the "total space, B is called the "base space/manifold, and F is the "fiber".

This map has to be such that at all points of the base space a neighborhood around a subset of the base space has to be such that

$$E = F \times B$$

which is to say the bundle is "locally" a cartesian product of the base space and the fiber space.



This is a really elegant route towards differential geometry that we may decide to pursue after our less abstract considerations.

At this point we've just about run out of things we can say about our flat Euclidean manifold. Our universe, on local as well as cosmic scales, is measured to be very nearly flat. Why this is so is a mystery as Einstein discovered that gravity is a manifestation of curvature in a $4D$ manifold we call space-time. At this point it is very hard to understand what we mean by "the space-time 4-manifold has curvature" but as we putter along in our studies it will slowly appear more and more reasonable.

We are widely familiar with our spatial world, once we develop the machinery, we can combine our spatial and temporal manifolds together to form a single 4 dimensional space-time manifold whose points we call "events" labeled by (t, x, y, z). There is much to discuss and study.

Our strategy we will employ to study arbitrary manifolds will hinge on the requirement that "geometry is independent of coordinate choices". For example, a sphere is a sphere whether we describe it with (x, y, z) coordinates or (r, θ, ϕ) coordinates or anything complicated in between!

By studying general coordinate transformations we will make definitions for objects such as vectors, scalars, and tensors. These are geometric objects that should also be invariant under coordinate transformations.

Next we shall study how the basis vectors and co-vectors vary from point to point on our manifold. This will allow us to "connect" the tangent spaces together such that we can then compare vectors from different places.

Finally the way the "connection" varies from place to place will allow us to completely describe the curvature at any point on our manifold. This was the dream of Gauss and Euler among others that was fully realized by the great Riemann and put to use by the great Einstein. It is a straightforward yet enormously difficult journey. We will develop notational tools to allow us to study these concepts more clearly under a choice of coordinates.

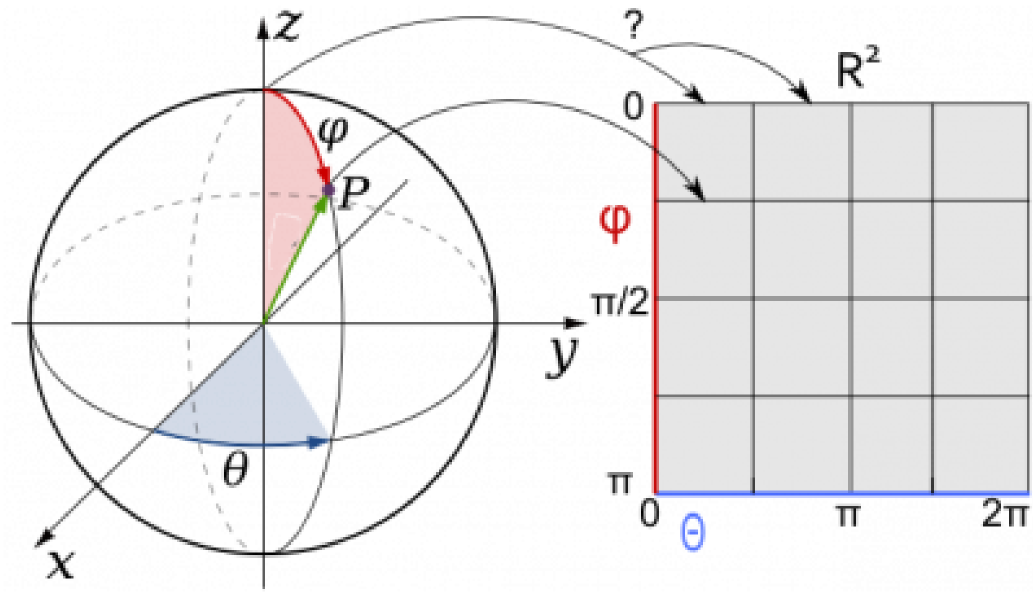
Later, the coordinates will be stripped away and we will have the full, yet abstract, mathematical structure of the geometry of curved manifolds. With that in hand, we can build almost any complicated geometry we desire and push the limits of what is and isnt known about geometry.

---

3 Building Our World

Suppose we have an arbitrary smooth manifold $\mathcal{M}$.

A manifold is a mathematical structure with certain specific properties. For us we will imagine that it's some continuous space such that if we zoom into a small enough region it looks like $\mathbb{R}^D$ where D is the dimension of the manifold. That is to say, in small enough regions it is Euclidean. This is similar in spirit to our own Earth: locally it appears flat and we can have parallel lines and follow Euclid's axioms but if we go out far enough we soon discover our otherwise parallel lines will intersect and triangles will no longer sum to π! The surface of the Earth is manifestly two dimensional and we label points on our surface with θ and ϕ: latitude and longitude.



Smooth essentially means we can take as many derivatives as we please. We call a smooth manifold C^∞ meaning "infinitely differentiable". There are ways for relaxing this restriction but this happens to be convenient for our purposes and isn't too severe of a restriction.

A n-sphere S^n is an example of a smooth manifold as is a -n-torus T^n and much much more.

Now, we will label each and every point of our manifold using ξ. This is a generic assignment meant only to distinguish between a point ξ of $\mathcal{M}$ and a coordinate value x^i. Points are geometric objects whereas coordinates are a "choice".

Now, let us choose coordinates for $\mathcal{M}$ which we call x^i - that is, some coordinates (like θ and ϕ for the sphere or (x, y, z) and so forth). We mentally imagine these coordinates as being a set of rulers erected at each and every point $\xi \in \mathcal{M}$.

These coordinates can be essentially anything; maybe they're really wacky and wavy lines all over the place: any choice that covers the manifold and doesn't do anything terrible like loop around a point will usually be valid choices. What's more is we want our coordinates to cover every single point of $\xi \in \mathcal{M}$.

Now, the manifold is MUCH more fundamental than the rulers/coordinates we lay down. We expect, indeed demand, that when describing functions, and vectors, and other things on our manifold that different coordinates wont make any difference to our geometric or inherent properties.

For example, we could define a physical temperature field giving us a temperature at every point ξ on $\mathcal{M}$. If we have a single number at every point in $\mathcal{M}$ then we can write

$$T(\xi) \in \mathbb{R}$$

Temperature is a real number with specific units, say "Kelvin". Therefore we have a mapping

$$T : \mathcal{M} \to \mathbb{R}$$

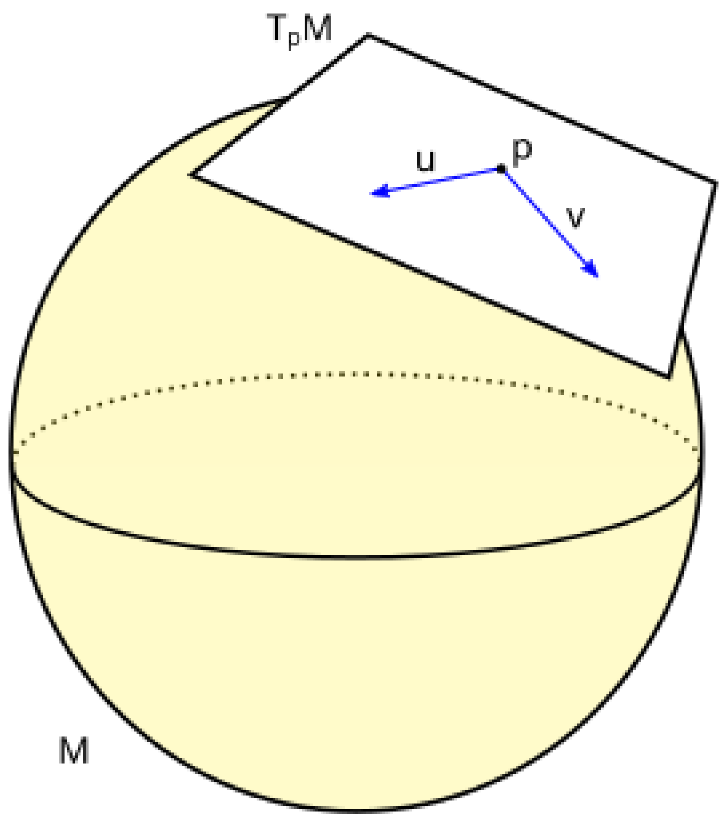
Now, if we choose specific coordinates then we do not expect this temperature field to change one bit. Suppose we choose coordinates x^μ where μ ranges from 0 to the dimension of $\mathcal{M}$. E.g. if $\mathcal{M}$ is a $5D$ manifold then our coordinates would be x^0, x^1, x^2, x^3, x^4. You could also write these coordinates as x, y, z, w, v but we will generally label them with the same letters indexed by natural numbers.

Within these coordinates we have chosen then

$$T = T(\vec{r})$$

where $\vec{r}$ is the position vector in the specific coordinates we've chosen.

Also note that given our manifold we also have $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$ such that we can define vectors and co-vectors at a point p.



Now suppose we chose our x^μ coordinates and then change our minds and wish to transform into a different coordinate system. We do this all the time in physics - you've defined a system in the (x, y, z) Cartesian coordinates and then said "nevermind, I want to study this in spherical coordinates (r, θ, ϕ)" and you simply make a transformation.

Let's say that for this particular manifold we chose the usual (x, y, z) coordinates near a point p and then wish to change into spherical coordinates around that point. (Let me be clear; I am saying "near a point $p \in \mathcal{M}$" since we've required our manifold to be locally like $\mathbb{R}^{\mathbb{D}}$. It wont be these coordinates EVERYWHERE...just in a small patch around p - we could define this properly as a "topological manifold" but I think the general idea is apt enough).

We wish to transform our everyday $(x, y, z) \in \mathbb{R}^3$ into spherical coordinates so we perform the mapping

$$\mathbb{X} : (x, y, z) \rightarrow (r, \theta, \phi)$$

$$\mathbb{X}(r, \theta, \phi) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r \sin\theta \cos\phi \\ r \sin\theta \sin\phi \\ r \cos\theta \end{bmatrix}$$

$\mathbb{X}$ is called a coordinate map. Notice that $\mathbb{X}$ is a map from the tangent space $T_p\mathcal{M}$ into the tangent space. That's where vectors live!

In both sets of coordinates we have basis vectors such as $\hat{x}, \hat{y}, \hat{z}$ and $\hat{r}, \hat{\theta}, \hat{\phi}$. We shall also distinguish between previous coordinates x^μ and new coordinates x'^μ (with a prime symbol) rather than always coming up with new letters and symbols anytime we wish to change! In general we shall write for our basis vectors

$$e_i \in T_p\mathcal{M}$$

For example

$$e_2 = \hat{y}$$

$$e'_3 = \hat{\phi}$$

and so forth.

If the coordinates change in this manner then how will our (x, y, z) derivative operators transform into these new coordinates?

Notice that we're asking two questions in one.

1. How do elements of $T_p\mathcal{M}$ transform?

2. How do elements of $T_p^*\mathcal{M}$ transform?

Vectors live in $T_p\mathcal{M}$ and co-vectors in $T_p^*\mathcal{M}$. This is in the same spirit as we studied earlier where we have basis elements dx^i and $\frac{\partial}{\partial x^i}$

However, let's pretend we don't know any f that and simply follow our standard procedure to find our transformed derivatives.

Well, to do so, we studied the differentials of this coordinate transformation and found

$$dx = (\sin\theta\cos\phi)dr + (r\cos\theta\cos\phi)d\theta - r\sin\theta\sin\phi d\phi$$

$$dy = (\sin\theta\sin\phi)dr + (r\cos\theta\sin\phi)d\theta + r\sin\theta\cos\phi d\phi$$

$$dz = \cos\theta dr - r\sin\theta d\theta$$

We then inverted these three equations to find

$$\begin{aligned} dr &= \dots \\ d\theta &= \dots \\ d\phi &= \dots \end{aligned}$$

Which is to say, compactly

$$dx'^i = \sum_j \frac{\partial x'^i}{\partial x^j} dx^j = \frac{\partial x'^i}{\partial x^j} dx^j$$

Once we had this we found we could form derivative operators in our new coordinates as

$$\frac{\partial}{\partial x'^i} = \sum_j \frac{\partial x^j}{\partial x'^i}\frac{\partial}{\partial x^j} = \frac{\partial x^j}{\partial x'^i}\frac{\partial}{\partial x^j}$$

Here, I have intentionally removed the sums over j following Einstein's advice that "If you see an up index and a down index repeated then you always sum over that index" . If you have an "up" index in the bottom half of a fraction then we consider that to be a "down" index.

E.g.

$$\frac{\partial}{\partial x^j} = \partial_j$$

Notice that $\frac{\partial x^i}{\partial x'^j}$ occurs in all of these expressions as though it were a matrix. Let us therefore write

$$\frac{\partial x'^i}{\partial x^j} = S^i_j$$

$$\frac{\partial x^j}{\partial x'^i} = (S^{-1})^j_i$$

These expressions comprise what we call a coordinate transformation. In the expressions we will study we will investigate how various objects transform under a coordinate transformation.

S governs how the basis elements of $T_p^*\mathcal{M}$ transform and S^{-1} governs how the basis elements of $T_p\mathcal{M}$ transform. We say they "transform oppositely".

Now that we have these transformation laws of our tangent and cotangent spaces then we are done (at least around our point p): the rest is physics!

We expect the components of a particular vector $\vec{v}$ to be different in different coordinate systems. However, their length should be completely invariant upon a legitimate coordinate transformation if we are to have a notion of "vector" that we grew up with. We would have HUGE problems if, when we rotate our coordinate system, the length of the vector changed!

Let's revisit rotations in the plane. Our coordinate transformations (the map $\mathbb{X}$) are given by

$$x' = \cos(\theta)x + \sin(\theta)y$$

$$y' = -\sin(\theta)x + \cos(\theta)y$$

Then, from our definition of S^i_j we have

$$S^i_j = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

But this is our standard rotation matrix $R(\theta)$

We have seen before that if under a rotation we have

$$\vec{v}' = R(\theta)\vec{v}$$

or, in components, we have

$$v'^i = \sum_j S^i_j v^j = S^i_j v^j$$

then

$$\vec{v}' \cdot \vec{v}' = \vec{v} \cdot \vec{v}$$

Therefore, for a legitimate rotation we have

$$R^T R = \mathbb{1}$$

In our general notation then

$$S^i_k (S^{-1})^k_j = \delta^i_j$$

Notice we have a "transpose operation" and we recently learned that vectors are columns and co-vectors are rows. For now, even though we're familiar with $R^T R$ and basic linear algebra, in our new language of tangent and co-tangent spaces we're unsure how to interpret this. Let's leave that subtlety aside for now. The point is the rotations leave our vector's length invariant.

This must be true regardless of what our coordinate transformation functions are. S^i_j is the primal connection between one and another coordinate change. It's not "oh this just worked because of our particular transformation law of rotations" - we could've chosen anything and the length of our vector would have HAD to be invariant.

What's more is that the lengths of any curves had also be invariant under these coordinate changes.

We found that under a coordinate change the length of a vector remained the same if its "components" (emphasis on this) changed as

$$v'^i = S^i_j v^j$$

This is a linear transformation regardless of the non-linear functions that transformed our individual coordinates.

Let us use this revelation to update our definition of what a vector is. This new definition will apply to ANY manifold no matter how curvy or how flat.

Definition: A vector is anything that transforms under coordinate transformation like dx^i

$$dx'^i = \sum_j \frac{\partial x'^i}{\partial x^j} dx^j = S^i_j dx^j$$

Notice, we are defining a vector's components in terms of how the differential elements transform. A vector would be anything like

$$v'^i(x') = S^i_j v^j(x)$$

Often, when working with vectors we only care about the values of the "components" and we tend to ignore the basis vectors. The basis vectors have a down index and the components have an up index such that the vector components transform just like dx^i. Yes, it is a bit confusing.

What about the other transformation law?

Definition: a co-vector is anything that transforms under coordinate transformation like

$$\frac{\partial}{\partial x'^i} = \sum_j \frac{\partial x^j}{\partial x'^i}\frac{\partial}{\partial x^j} = (S^{-1})^j_i \frac{\partial}{\partial x^j}$$

where

$$\frac{\partial x^i}{\partial x'^j} = (S^{-1})^i_j$$

An example of a co-vector would be and object that transforms as

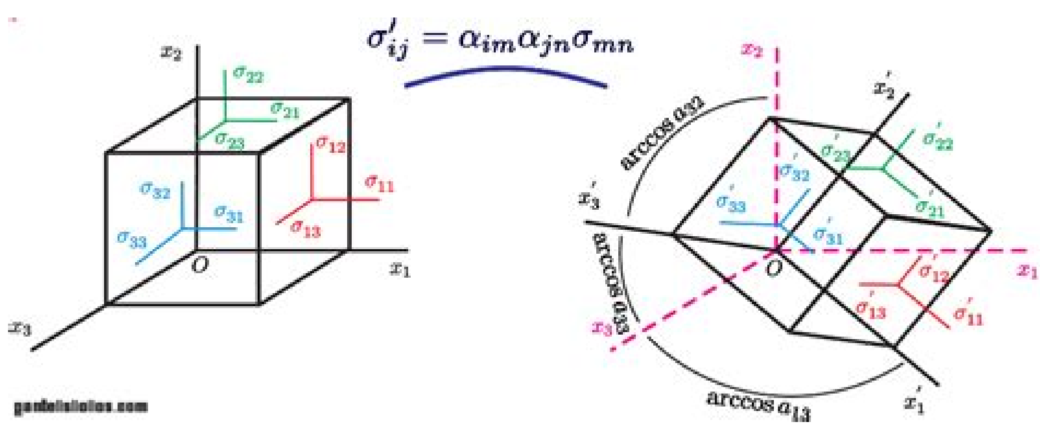
$$v'_i(x') = (S^{-1})^j_i v_j(x)$$

This is how the elements of a row-vector transform: what we previously called

$$\omega = \omega_x dx + \omega_y dy + \omega_z dz$$

Next we have a new type of object called a tensor

Definition: a Tensor is anything that transforms under coordinate transformation like S for each upper index and S^{-1} for each lower index



These will be beasts that we write like

$$T^{ij}_{klm} = S^i_a S^j_b (S^{-1})^c_k (S^{-1})^d_l (S^{-1})^f_m T^{ab}_{cdf}$$

They have multiple indices and straddle both $T_p\mathcal{M}$ and $T^*_p\mathcal{M}$.

and so forth. The above example is called a rank $(2, 3)$ tensor. In principle we can have as many indices as we want but usually in physics we deal with just a few. In general we have a rank (m, n) tensor writeen as

$$T^{abc...m}_{pqr...n}$$

A general tensor can be created from elements of $T_p\mathcal{M}$ and $T^*_p\mathcal{M}$ by invoking the outer or tensor product of vectors and co-vectors. For example

$$\omega_a e^a \otimes \sigma_b e^b = \begin{bmatrix} \omega_1\sigma_1 & \omega_1\sigma_2 & \omega_1\sigma_3 \\ \omega_2\sigma_1 & \omega_2\sigma_2 & \omega_2\sigma_3 \\ \omega_3\sigma_1 & \omega_3\sigma_2 & \omega_3\sigma_3 \end{bmatrix} = T_{ij}(e^i \otimes e^j)$$

would be a rank $(0, 2)$ composed from a pair of co-vectors

We could build anything we'd like in this manner.

$$\omega_a e^a \otimes w^j e_j \otimes \sigma_b e^b \otimes v^i e_i = \Lambda^{ij}_{ab}(e^a \otimes e_j \otimes e^b \otimes e_i)$$

For now the details of tensors aren't too important. We will discuss more on them in time. For now we just want to build up our machinery.

Now, let's take stock of what we've found by considering coordinate transformations. We have two primary ways things transform and the differential operators are especially important

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j$$

$$\frac{\partial}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i}\frac{\partial}{\partial x^j}$$

Here we take a major conceptual leap. Let's abstract what basis vectors even are. Suppose I have a vector $\vec{X}$ on a manifold $\mathcal{M}$ at some point p; then I can define the derivative of a function $f(x^1, ...x^D)$ with respect to my vector $\vec{X}$ as

$$D_X(f) = \sum_j \frac{\partial f}{\partial x^j}(p) X^j$$

where the derivative is evaluated at p. Clearly this derivative depends on our coordinates so that we hope that for an arbitrary function f the value it returns will be independent of our coordinate choice.

Your task: apply a coordinate transformation to this expression and show that it is invariant.

Now, for all practical purposes our vector $\vec{X}$ at p acts on functions as

$$\vec{X} = \sum_j X^j \frac{\partial}{\partial x^j} = X^j \frac{\partial}{\partial x^j}$$

This looks an awful lot like an expansion of a vector in a basis given by

$$e_j = \frac{\partial}{\partial x^j}$$

Indeed this is the case: consider that at some local point p we can choose coordinates and then use those coordinates to define basis vectors at that point to be the tangent vectors. The tangent vector to that point in such and such direction is just what the derivative means. They comprise our basis of $T_p\mathcal{M}$ - the tangent space!

In our everyday Euclidean (x, y) coordinates this never showed up because at a point p the tangent vector to that point would be

$$e_j = \frac{\partial}{\partial x^j} \begin{bmatrix} x \\ y \end{bmatrix}$$

which yields

$$\hat{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\hat{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

But for curved spaces these vectors can potentially be more complicated - even in polar coordinates now the basis vectors depend on θ and as you travel around the basis vectors tilt and twist - this is merely due to a particular choice of coordinates.

Also consider that given this definition for our basis vectors at point p then on an arbitrary manifold we generally have basis vectors vary from one point to the next generally regardless of the coordinates we choose!

We say that vectors live in the "tangent space" at $T_p\mathcal{M}$ and any vector can be written in a chosen set of coordinates as

$$\vec{v} = v^j \frac{\partial}{\partial x^j} = v^j e_j$$

Conversely, the differentials dx^i define a dual basis for the "cotangent" space $T_p^*\mathcal{M}$ via

$$\vec{v} = v_j dx^j = v_j e^j$$

This means for each point of the manifold we have two fundamental spaces: the tangent space and the cotangent space; and any vector can be expanded in either one of the bases.

Importantly, in general $v^i \neq v_i$. We can write the same vector in either space but they will in general have different values for their components. We haven't seen this happen before. Every time you switched a column vector to a row vector the components stayed the same: we took it for granted. It certainly does matter in quantum mechanics since when you change a ket into a bra you have to complex conjugate! The same thing as we discuss here happens in quantum theory!

v_i are called "covariant components" and v^i are called "contravariant components". Sometimes we say things like $v_i dx^i$ is a 1-form or differential form.

Now, if we choose a coordinate system then we can ask how the components of $\vec{v}$ transform and we found earlier that

$$v'^i = S^i_j v^j$$

$$v'_i = (S^{-1})^j_i v_j$$

where we said "anything that transforms like a derivative is a vector" and "anything that transforms like a differential is a co-vector".

Given all of this then we have at some point $p \in \mathcal{M}$ the tangent and cotangent space bases are given by

$$e_i = \frac{\partial}{\partial x^i}$$

$$e^i = dx^i$$

They are such that

$$e_i \cdot e^j = e^j \cdot e_i = \delta_i^j$$

We demand that these basis elements within their respective spaces transform under arbitrary coordinate changes as

$$e'^i = dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j = S^i_j dx^j$$

$$e'_i = \frac{\partial}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i} \frac{\partial}{\partial x^j} = (S^{-1})^j_i \frac{\partial}{\partial x^j}$$

Due to this relationship we can write an arbitrary vector in terms of either the tangent or cotangent space as

$$\vec{v} = v^1 \hat{e}_1 + v^2 \hat{e}_2 + v^3 \hat{e}_3$$

$$\vec{v} = v_1 \hat{e}^1 + v_2 \hat{e}^2 + v_3 \hat{e}^3$$

These v^i,v_i can be numbers or even functions themselves: in which case we will call it a vector field or a co-vector field $v(x)$ for $x \in \mathcal{M}$

It is not yet clear how to switch back and forth between upper and lower indices but in our analysis of $\mathbb{R}^3$ it will wind up being the metric which connects our rows and columns.

4 The 2-Sphere

Now, after all this long winded abstraction and formalism, let's see a specific example of a manifold that isn't our everyday $\mathbb{R}^N$. Everything we have done has been aspects of the same boring flat Euclidean $(x, y, z, \cdots)$ space. The 2-sphere is inherently and definitely a curved manifold. How can we proceed?

The key insight is that we studied a coordinate transformation from Cartesian to spherical coordinates. What would happen if we restricted our world to be the surface of the sphere? That should be the same as imposing that $dr = 0$ and the radius is some constant R.

In this manner of thinking we are really asking "Can we embed the 2-sphere in our flat Euclidean $3D$ space? Then can we study functions and vectors existing only on that surface?

Clearly dear reader you know verifiably that you can embed a 2-sphere in $3D$ space. We play with them all the time: some folks even get paid billions to embed and manipulate 2-spheres in $3D$ space!

Suppose we have a 2-sphere of radius $R = 1$. Our coordinate map, restricted to $R = 1$ will then define a surface parameterized by (θ, ϕ); our everyday latitude and longitude on the sphere.

Our generalization will be via the following steps we've applied many times already:

1. Define a coordinate map $\mathbb{X}(\theta, \phi)$
   This map defines a surface in (x, y, z) with (θ, ϕ) coordinates.

2. Define basis vectors in the tangent space (e_θ, e_ϕ) and co-tangent space (e^θ, e^ϕ). The basis vectors and co-vectors will naturally depend on the parameterized coordinates of the surface.

Let's apply this to the case of a 2-sphere. We already have our $\mathbb{X}(\theta, \phi)$ from earlier; now we just sub in a constant radius.

$$\mathbb{X}(\theta, \phi) = \begin{bmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \end{bmatrix}$$

Next, $d\mathbb{X}$ is given by the regular old chain rule

$$d\mathbb{X}^i = \sum_j \frac{\partial x^i}{\partial u^j} du^j$$

Where, for the 2-sphere, we have $u^1 = \theta$ and $u^2 = \phi$ and $(x, y, z) = (x^1, x^2, x^3)$ in this generalized notation.

So we have our transformation matrix

$$S^i_j = \frac{\partial x^i}{\partial u^j}$$

This is how the coordinates change from one set to the other.
Next, we define the spherical basis vectors as

$$\hat{e}_i = \frac{\partial \mathbb{X}}{\partial u^i}$$

In our spherical example $(R = 1)$ this yields

$$e_\theta = \frac{\partial}{\partial \theta} \mathbb{X}$$

$$= \frac{\partial}{\partial \theta} \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}$$

$$e_\theta = \begin{bmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{bmatrix}$$

Similarly,

$$e_\phi = \begin{bmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{bmatrix}$$

which is indeed the correct form of the standard $\hat{\theta}$ and $\hat{\phi}$ vectors we know so well.

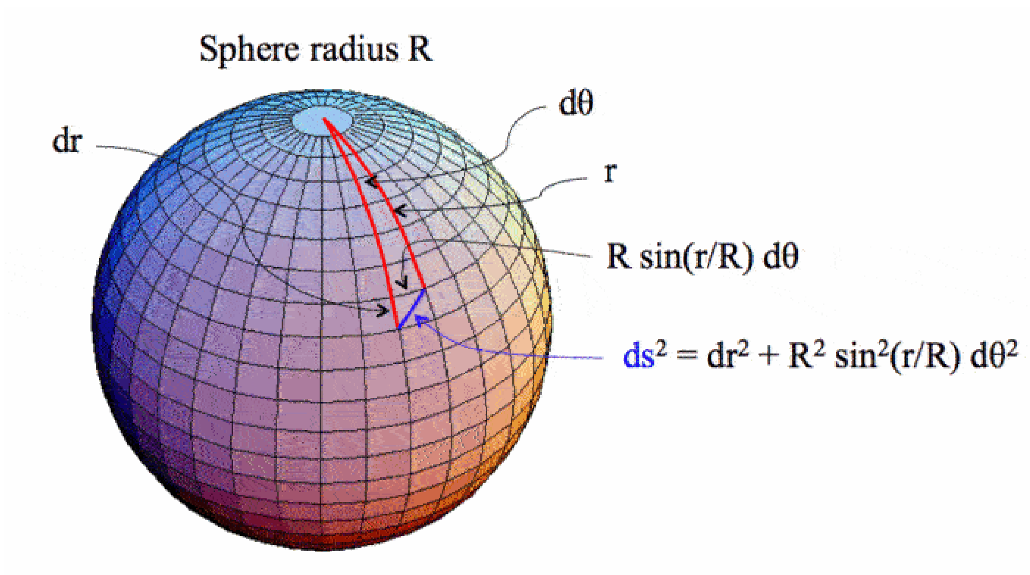
e_θ and e_ϕ fundamentally live in the tangent space T_pS^2 but we can see they "stick out" in the (x, y, z) directions: this is due to our choice of proceeding via an embedding into a $3D$ space.

In our 2-sphere case we have here a general mapping from (θ, ϕ) to (x, y, z)

That is

$$\mathbb{X} : \mathbb{R}^2 \to \mathbb{E}^3$$

where $\mathbb{E}^3$ means a "Euclidean 3-space" like our everyday geometric world.



Notice what we've done: once we specified our coordinates (θ, ϕ) we had no more information about any kind of third dimension. Our 2-sphere is manifestly two dimensional and in these new coordinates everything we do will be constrained to that surface. For all practical purposes there is no "third" dimension when we speak only of θ and ϕ coordinates. We almighty $3D$ beings, looking down from $3D$ to the $2D$ S^2 certainly see the curvature and all the tangent space vectors "sticking out" from the surface. However, a bug constrained to this surface sees none of that - it may even appear locally like $\mathbb{E}^2$ to them!

What could a bug study in these coordinates? Well anything we could study in $\mathbb{R}^2$ and potentially more since S^2 is fundamentally different globally from the real plane.

θ and ϕ act like an (x, y) plane with the stipulation of periodicity. A natural question to ask might be "what is the radius of this sphere?" You and I in our $3D$ space know that the radius is 1 in some suitable units of distance. However, is it possible to figure this out given only the (θ, ϕ) coordinates? Can our civilization of bugs on the surface discover the curvature of their world?

Maybe they could ask "how long do I have to walk straight before I get back to where I started?" They travel for thousands of miles and make it back and then possibly assume they live on a sphere and have secretly walked a great circle. Then they could infer the radius from the circumference.

But notice - to do this they have to a priori ASSUME it's a sphere. In general that's not guaranteed. Perhaps its secretly an ellipsoid or some other periodic space. Maybe there are mountains and trenches. How then could they possibly find out?

The ancients had the same problem and couldn't exactly send out satellites into a third dimension to find out "from above"!

One thing you could do if restricted to some surface is the following experiment

1. walk 100 miles east

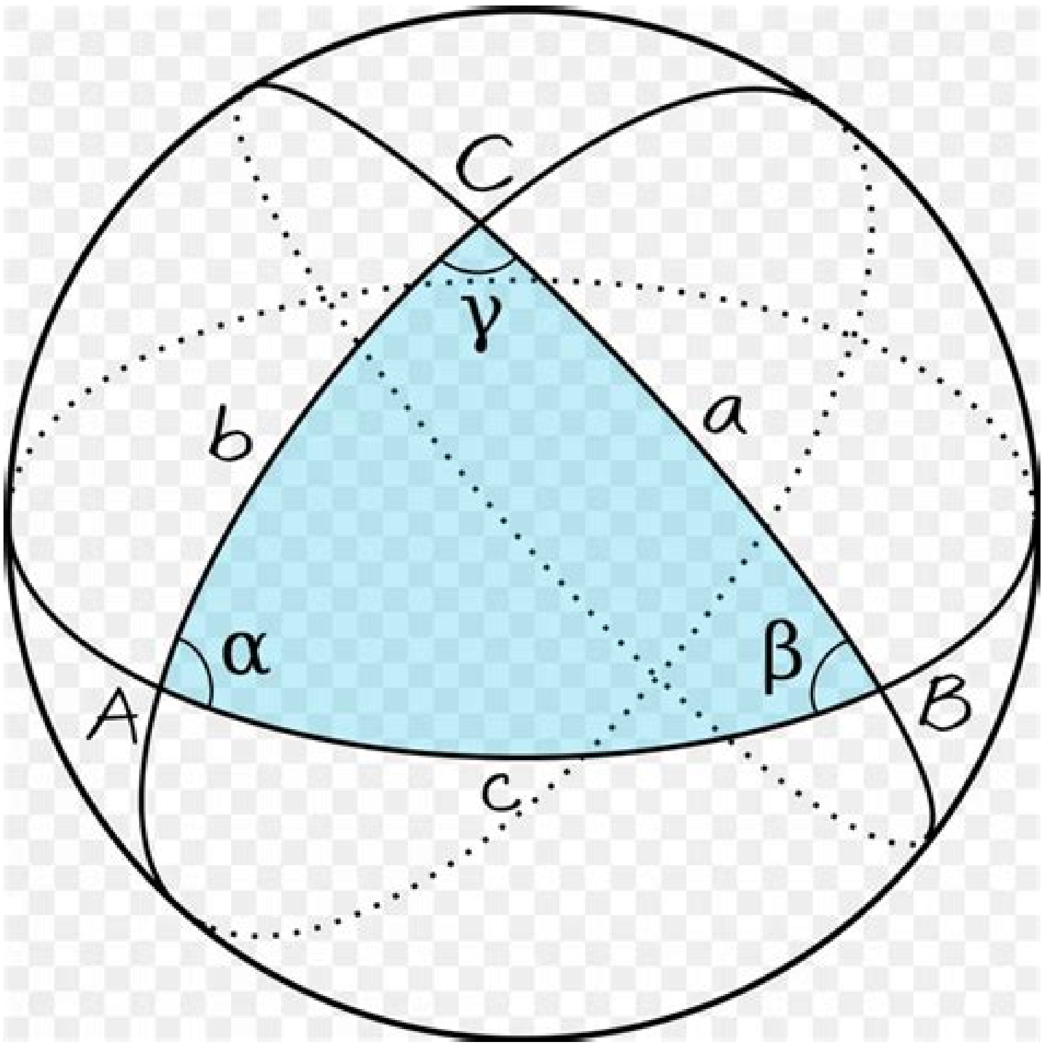
2. walk 100 miles north

3. walk 100 miles west

4. walk 100 miles south

Then you check if you've made it back home or somewhere else. If you live in a perfectly flat manifold then you will wind up back home. Otherwise, like on Earth, you can infer that there is some "local" curvature.

This is very much akin to drawing triangles and checking if the angles sum to π. On the Earth at some point, for large enough triangles, you will find angles summing to greater than π. You can try this on a globe or balloon.



Therefore, whatever is going on we might conjecture that this effect of local "curvature" is proportional to the area of some loop or path that you circulate around. The amount that your starting and ending points deviate would then yield some expression for the radius of curvature within that area. Or, conversely, if your path is a closed polygon, then the angular deviation from Euclidean space would be proportional to the curvature within that area.

However, determining that your global manifold is globally a sphere seems more challenging. Everything idea we've come up with has been related to "local" curvature.

You might send out signals in multiple directions and wait for them to all return and if they all arrive at exactly the same time then you could claim that the radius of curvature is independent of direction which would then imply a spherical world.

All of this curvature business aside what we would ultimately love to do is figure out how to integrate along a parameterized curve $\gamma(s)$ to find its length or some other property by utilizing some generalization of

$$\int_\gamma ds$$

In our everyday Euclidean world we would integrate

$$\int_\gamma ds = \int_\gamma \sqrt{dx^2 + dy^2 + dz^2}$$

based upon Pythagoras. However, at this point it is unclear how and what to integrate on a curved manifold. We will revisit this idea shortly.

---

5 Polar Coordinates

Now let's study the flat plane in polar coordinates.

For the manifold $\mathbb{R}^2$ define the coordinate transformation from $(x^1 = x, x^2 = y)$ to $(x'^1 = r, x'^2 = \theta)$

$$x^1 = x = r\cos\theta$$

$$x^2 = y = r\sin\theta$$

and

$$x'^1 = r = \sqrt{(x^1)^2 + (x^2)^2}$$

$$x'^2 = \theta = tan^{-1}\frac{x^2}{x^1}$$

These transformations are certainly non-linear functions but the point is that $e^i = dx^i$ and $e_i = \frac{\partial}{\partial x^i}$ transform linearly with a matrix.

Now, given that we're transforming from unprimed to primed coordinates and we have our coordinate map then we can compute the S matrix. It is (take all the derivatives)

$$S^i_j = \begin{bmatrix} \frac{x_1}{\sqrt{x_1^2+x_2^2}} & \frac{x_2}{\sqrt{x_1^2+x_2^2}} \\ \frac{-x_2}{x_1^2+x_2^2} & \frac{x_1}{x_1^2+x_2^2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \frac{-\sin\theta}{r} & \frac{\cos\theta}{r} \end{bmatrix}$$

and its inverse

$$(S^{-1})^i_j = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

Complicated as it may appear each of the elements of S are simple derivatives of the respective coordinate transformations (it's just (x, y) or (r, θ) - you're used to this just not written in this manner).

Importantly, clearly we hope this to be a bijective transformation: if we go one way and then transform back we expect what we started with (so that $SS^{-1} = \mathbb{1}$)

Our basis vectors and basis co-vectors are

$$dx^1 = dx = e^x = \frac{\partial x^1}{\partial x'^j} dx'^j = \cos\theta dr - r\sin\theta d\theta$$

$$dx^2 = dy = e^y = \frac{\partial x^2}{\partial x'^j} dx'^j = \sin\theta dr + r\cos\theta d\theta$$

$$dx'^1 = dr = e^r = \frac{\partial x'^2}{\partial x^j} dx^j = \frac{x_1}{\sqrt{x_1^2+x_2^2}} dx + \frac{x_2}{\sqrt{x_1^2+x_2^2}} dy$$

$$dx'^2 = d\theta = e^\theta = \frac{\partial x'^2}{\partial x^j} dx^j = -\frac{x_2}{x_1^2+x_2^2} dx + \frac{x_1}{x_1^2+x_2^2} dy$$

$$e_x = \frac{\partial}{\partial x} = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}\frac{\partial}{\partial r} + \frac{x_2}{\sqrt{x_1^2 + x_2^2}}\frac{\partial}{\partial \theta}$$

$$e_y = \frac{\partial}{\partial y} = -\frac{x_2}{x_1^2 + x_2^2}\frac{\partial}{\partial r} + \frac{x_1}{x_1^2 + x_2^2}\frac{\partial}{\partial \theta}$$

$$e_r = \frac{\partial}{\partial r} = \cos\theta\frac{\partial}{\partial x} - r\sin\theta\frac{\partial}{\partial y}$$

$$e_\theta = \frac{\partial}{\partial \theta} = \sin\theta\frac{\partial}{\partial x} + r\cos\theta\frac{\partial}{\partial y}$$

Complete the last step: write each of these in terms of the same variables on each side of the equation. It's a mess but abstracting the fussy details away serves to clarify what's really happening here.

Now, for fun notice what happens if we compute

$$e_i \cdot e_j = g_{ij}$$

computing all the dot products we delightfully find

$$g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}$$

This appears to be shockingly simple (compared to the mess above) and is ultimately the primary object of study for differential geometry!

6 The Metric

In introductory calculus you studied the Pythagorean line element

$$ds^2 = dx^2 + dy^2 + dz^2$$

The way we worked with this thing is that we chose a parameterization of some curve $\gamma(t) \in \mathbb{R}^D$ such that we could integrate along that parameter t.

Often the parameter is time but it could also be the total distance one travels along the curve (i.e. you start at $s = 0$ and arrive at $s = 1$ that you measure during your trip along the curve. In this manner the total length of the curve is $L = 1$).

Under any parameterization then

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt$$

No matter how we choose to parameterize our curve this line element will be fundamentally true and is due to the fundamental nature of the underlying manifold $\mathcal{M}$: in this case $\mathbb{E}^3$.

We can integrate over our parameter to find the length of the curve as

$$L = \int \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$$

Thus all is well in the world. Given some curve $\gamma(t)$ we merrily compute derivatives and integrate to study arbitrary parameterized curves.

However, this analysis has been with respect to our purely Euclidean space - what else could we have envisioned? Perhaps you've heard the parallel line axiom of Euclid fails for exotic geometries such as curves on spheres and hyperbolic spaces.

How can we generalize these ideas to cover geometries that are not intrinsically flat like we're accustomed to? Pythagoras, as amazing as it is, only applies to Euclidean geometries after all.

We've spoken a bit about this type of generalization before. A key to figuring out how to proceed to the most general case is to remind ourselves that coordinates don't matter!

For example, we know that $ds^2 = dx^2 + dy^2 + dz^2$ defines our Euclidean geometry and we also know that not a damn thing is different if we choose some other coordinates. ds will surely be the same (you and I agree how long a curve is right?) The other side of the equation, however, will be written in different coordinates and this expression could potentially be wild.

As an example, suppose we have regular old (x, y, z) Euclidean flat space. We're perfectly comfortable with using spherical coordinates so let's figure out what's ds will be in these coordinates. This transformation doesn't change ANYTHING about the manifold - only how we describe it. Curves are curves and will have their respective lengths no matter how we choose coordinates: Lengths are invariant!

Applying our change of coordinates then, we find

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 = dx^2 + dy^2 + dz^2$$

This looks completely different than our original choice but we're guaranteed that a curve by any other name will have the same length under a change of coordinates. It looks nothing like Pythagoras and if you handed this to a person on the street and asked them if it was the Pythagorean theorem they would surely scoff (lest they were more learned than most!).

Now, we might stare at these two forms of the same thing and based upon what we've previously learned about our basis elements might notice that both of these forms look like a matrix times a basis element squared - what's known as a quadratic form. If you studied a bit of mechanics then you're used to seeing such things like kinetic energy written as a matrix equation. So, let's see where this leads: we have, in general

$$ds^2 = g_{ij} dx^i dx^j$$

where, in our Cartesian (x, y, z) case

$$g_{ij} = \delta_{ij}$$

and in our spherical coordinate case

$$g'_{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{bmatrix}$$

One of them looks like the identity matrix and the other a mildly complicated diagonal matrix.

We might speculate that these two matrices are equivalent for ds^2 due to the fact that they're both diagonal. Would all diagonal matrices describe equivalent geometries?

This is a really good guess, however, if you choose your favorite crazy coordinate functions you'll find that the matrix g_{ij} you wind up with will generally have weird off-diagonal elements depending on how weird of a choice of coordinates you make. So no, flat space isn't described solely by diagonal metrics.

But we do expect an entire class of matrices to be equivalent under coordinate changes: but how? what class of matrices are Euclidean?

Well, in its most general form we have a relationship between all the possible matrices we could consider for our everyday three dimensional Euclidean world. We have the fact that lengths are invariant and therefore

$$ds^2 = g_{ij} dx^i dx^j = g'_{ij} dx'^i dx'^j$$

In one case we have unprimed coordinates (e.g. (x, y, z)) and in the other case we have primed coordinates (e.g. (r, θ, ϕ)).

Okay, so we stare a bit and think back to what we know and it dawns on us that we've studied how dx^i transforms under coordinate changes: that's part of the reason we worked so hard earlier! We can substitute these equations in for dx'^i and dx'^j to get

$$ds^2 = g_{ij} dx^i dx^j = g'_{ij} dx'^i dx'^j$$

$$= g'_{ij} \sum_{\nu} \sum_{\mu} \frac{\partial x'^i}{\partial x^\mu} \frac{\partial x'^j}{\partial x^\nu} dx^\mu dx^\nu$$

where we've explicitly put in the $\sum$'s but didn't have to according to our double-index convention.

Notice: ALL of the indices are being summed over including i and j and whenever you sum over an index you are totally free to give it ANY name you'd like. Let's then switch ij with $\mu\nu$ and compare our two equations (and also remove the summation symbols since we are invoking Einstein's notation here anyways). We have

$$ds^2 = g_{ij} dx^i dx^j = g'_{\mu\nu} \frac{\partial x'^\mu}{\partial x^i} \frac{\partial x'^\nu}{\partial x^j} dx^i dx^j$$

This says that since the differentials are now written exactly the same on both sides then it must be the case that the matrix itself must change as

$$g_{ij} = g'_{\mu\nu} \frac{\partial x'^\mu}{\partial x^i} \frac{\partial x'^\nu}{\partial x^j}$$

or, we could go the other way around (watch the primes like a hawk!)

$$g'_{ij} = g_{\mu\nu} \frac{\partial x^\mu}{\partial x'^i} \frac{\partial x^\nu}{\partial x'^j}$$

We can also write this using our definition of S from earlier for these partials

$$g'_{ij} = g_{\mu\nu} (S^{-1})^\mu_i (S^{-1})^\nu_j$$

This says that two different metrics describe the same manifold if and only if they're related in this exact way. We perform a coordinate transformation and invariance of lengths requires this. If we find two different matrices that are NOT related in this manner then we say "we have two completely different geometries".

Recall what we learned in our introduction from considering our everyday (x, y, z) Euclidean space in gory abstract detail.

The metric is defined in some chosen coordinates as

$$e_i \cdot e_j = g_{ij}$$

the metric itself gives us a way of converting a^i to a_i via

$$a_i = g_{ij}a^j$$

$$a^i = g^{ij}a_j$$

such that

$$g_{ik}g^{kj} = \delta^i_j$$

This means that the metric with upper indices is the inverse of that with lower indices. The metric serves to connect the tangent and cotangent spaces of a manifold - once again, the metric is king.

This raising and lowering is so natural in the formalism that one becomes adept at doing it by eye. This is also the reason we never had to distinguish between upper and lower indices in Euclidean space: we could raise and lower willy-nilly since the metric was the identity! In more complicated manifolds we MUST be very careful about raising and lowering indices (which is a casual way of saying "moving from the tangent to cotangent spaces").

The beauty of this subject is that you can largely get away with simply taking what you already know, try to find a way to re-write it using the Kronecker delta, and then replace the delta with the general metric. In fact that's basically entirely how quantum field theory in curved spacetime is done!

Interestingly if we confine ourselves to the surface of a sphere like we did earlier we get $dr = 0$ and so the length of curves on a sphere of some radius R can be found by studying

$$ds^2 = R^2 d\theta^2 + R^2 \sin^2\theta d\phi^2$$

and suddenly we have a manifestly different geometry than

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$$

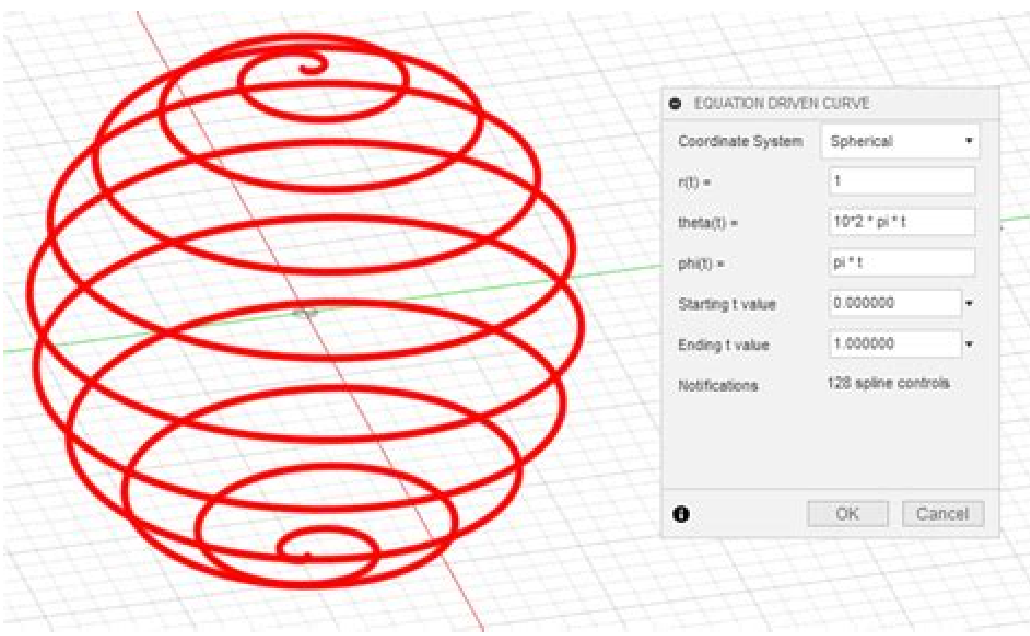
The sphere is curved and the spherical coordinate metric for Euclidean space is not. There is no way that we can perform a coordinate change to this spherical metric such that our two metrics are related. Zilch!

The sphere's metric describes a fundamentally different geometry than $\mathbb{E}^3$ even though we arrived at the sphere from $\mathbb{E}^3$.

So what do we get? Well now we can integrate along curves drawn on the surface of a sphere by parameterizing our curve and integrating ds. For example, I might have

$$\gamma(s) = \begin{bmatrix} \frac{\pi}{4}(1 + \frac{s}{2\pi}) \\ s \end{bmatrix} = \begin{bmatrix} \theta(s) \\ \phi(s) \end{bmatrix}$$

which describes a spiral on the sphere. Then if i want to study its length I integrate



$$ds^2 = R^2 \left(\frac{d\theta}{ds}\right)^2 + R^2 \sin^2\theta \left(\frac{d\phi}{ds}\right)^2 ds^2$$

but

$$\frac{d\theta}{ds} = \frac{1}{8}$$

$$\frac{d\phi}{ds} = 1$$

then

$$ds = R\sqrt{\frac{1}{64} + \sin^2\theta} ds$$

or

$$L = \int_0^{4\pi} \sqrt{\frac{1}{64} + \sin^2(\frac{\pi}{4} + \frac{s}{8}))} ds = 11.42$$

This is a curve that wraps around the "z-axis" twice (e.g. $\phi = [0, 4\pi]$)

Now let's try to generalize these ideas

For example, our everyday dot product says

$$\vec{a} \cdot \vec{b} = \sum_i a_i b^i = a_i b^i = a_1 b^1 + a_2 b^2 + \cdots$$

In this way of writing it the dot product is a row vector times a column vector. But we studied this in gory detail earlier: we use the metric to write this as

Using the identity matrix we also can write it as a vector times a vector

$$\vec{a} \cdot \vec{b} = \sum_k g_{ki} a^k b^i = a_i b^i = a_1 b^1 + a_2 b^2 + \cdots$$

Where a^k and b^i both transform under coordinate change as vectors. Previously our dot product was defined as a co-vector (row vector) times a column vector but now we have a slick way of realizing it as a type of multiplication between two bonafide tangent space vectors. We can write each vector under some chosen coordinates as

$$\vec{a} = a^i e_i$$

$$\vec{b} = b^j e_j$$

Then we have from our definition of the metric

$$a \cdot b = a^i e_i \cdot b^j e_j = a^i b^j e_i \cdot e_j$$

$$= a^i b^j g_{ij} = a^i b_i = a_j b^j$$

since $g_{ij} = e_i \cdot e_j$

Everything nice and cute and tidy! Secretly you've been doing differential geometry this whole time - it was always just a Euclidean space with its boring metric! The metric is king.

Suppose we have two vectors on the sphere at the same point

$$\vec{a} = \begin{bmatrix} 4\pi \\ 17 \end{bmatrix}$$

$$\vec{b} = \begin{bmatrix} 12 \\ \frac{1}{\pi} \end{bmatrix}$$

Those are certainly two vectors at some point $p = (\pi, \pi/3)$ on the sphere. How do we form their dot product? well it certainly is NOT $12 \times 4\pi + \frac{17}{\pi}$!! We MUST use the metric

$$ds^2 = R^2 d\theta^2 + R^2 \sin^2 \theta d\phi^2$$

then

$$g_{ij} = \begin{bmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{bmatrix}$$

Then

$$\vec{a} \cdot \vec{b} = 12 \times 4\pi R^2 + R^2 \sin^2 \theta \frac{17}{\pi}$$

at our chosen point $p = (\pi, \pi/3)$ then this dot product evaluates to $(\sin\pi = 0)$

$$\vec{a} \cdot \vec{b} = 48\pi R^2$$

everything is con trolled by the metric!

6.1 Riemannian Manifolds

Now we mention an important aspect of all of this that I've been intentionally putting off. The type of manifold we've been working with is called "Riemannian". We have been dealing with a special manifold such that locally we can plop down Cartesian coordinates. We say our manifold is "locally flat" in the same way that when we stand in Kansas it appears flat as flat can be. On large scales we see the curvature of the Earth but locally it always looks flat.

What's more we could build a really accurate map of the whole earth by taking pictures of 1 meter by 1 meter areas of the surface from a high tech satellite above and then stacking all of those flat images together in a book we call an "atlas"

Every single page, called a "chart", will be so flat that if we drew a triangle on it then the angles would sum to almost exactly π - as close to π as we want. The smaller we make our charts to more accurate we get to π. So for all practical purposes EVERY page in our atlas can be given regular old Cartesian $(x, y) \in \mathbb{R}^2$ coordinates!

Now, each page/chart is an entirely different coordinate system and if we want to cover the entire globe then each chart will overlap by some tiny amount. For example, a chart of Kansas and a chart of South Dakota will both contain the border between the two states. It's the same idea with our local coordinates.

So every chart in our atlas is flat as flat can be. But we're all familiar with the fact that if you try and map the surface of the earth using only a $2D$ flat page then you are guaranteed to have wild distortions. Common maps of the Earth make Antarctica and Greenland look HUGE - much bigger than they really are. But our atlas doesn't distort squat - everything is the correct size since we have billions of flat local charts. Need more accuracy? use a google of charts! A giant atlas indeed but we can do this in principle.

Okay, given this subtlety we recognize that as we move from point to point or chart to chart our basis vectors and co-vectors will also generally change. The basis vectors for one page of our atlas won't work for any other page of the atlas! Therefore we must generally write for all points $x \in \mathcal{M}$

$$e_i(x)$$

and

$$e^i(x)$$

Even more importantly our coordinate transformation laws will also depend on what chart we're in so that

$$S^i_j(x)$$

Given this we could ask fruitful questions such as "How do our basis vectors change from chart to chart?" We would want to compute something like

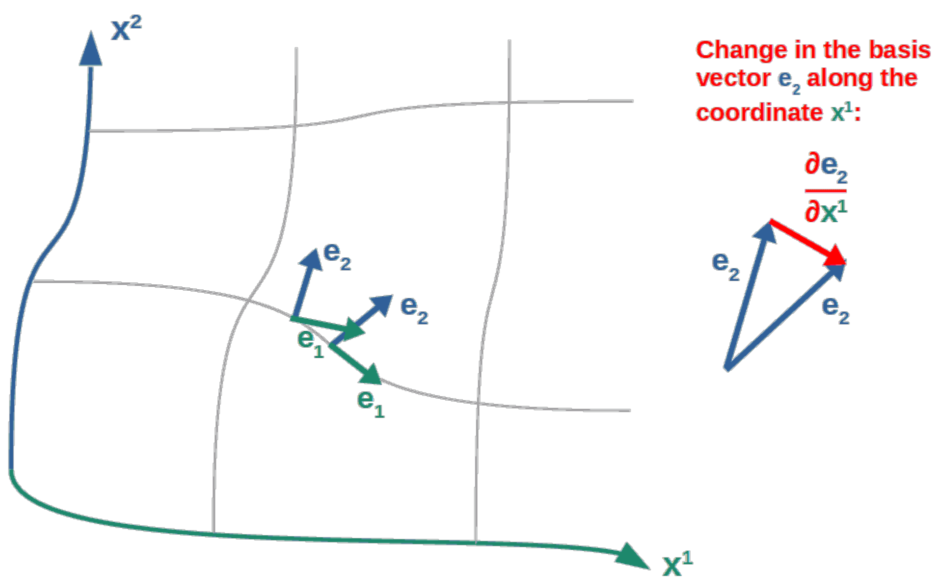
$$\frac{\partial e_i}{\partial x^\mu} = ?$$

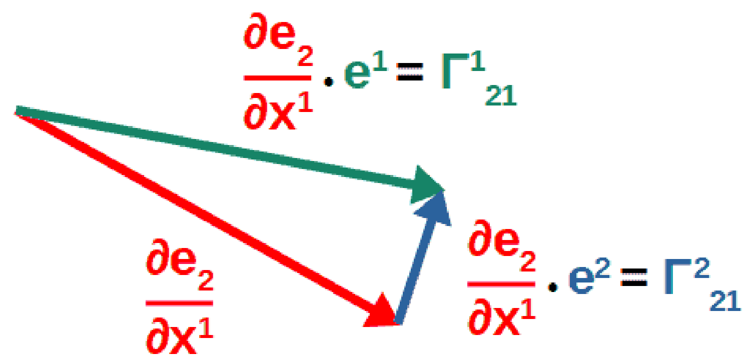
or even

$$\frac{\partial g_{ij}}{\partial x^\mu} = ?$$

since if the basis vary from chart to chart then so too does the metric itself!

Now, importantly; if we're in a chart labeled by x, our point in $\mathcal{M}$ with basis vectors $e_i(x)$ then whatever we get upon differentiation should not only be a vector or co-vector but it should be expressible in terms of some linear combination of our $e_i(x)$ basis vectors at that point!





Therefore, we expect the derivative of our basis vectors to be some linear combination of $e_k(x)$ like

$$\frac{\partial e_i}{\partial x^\mu} = \Gamma^k_{i\mu} e_k$$

(notice the sum over k)

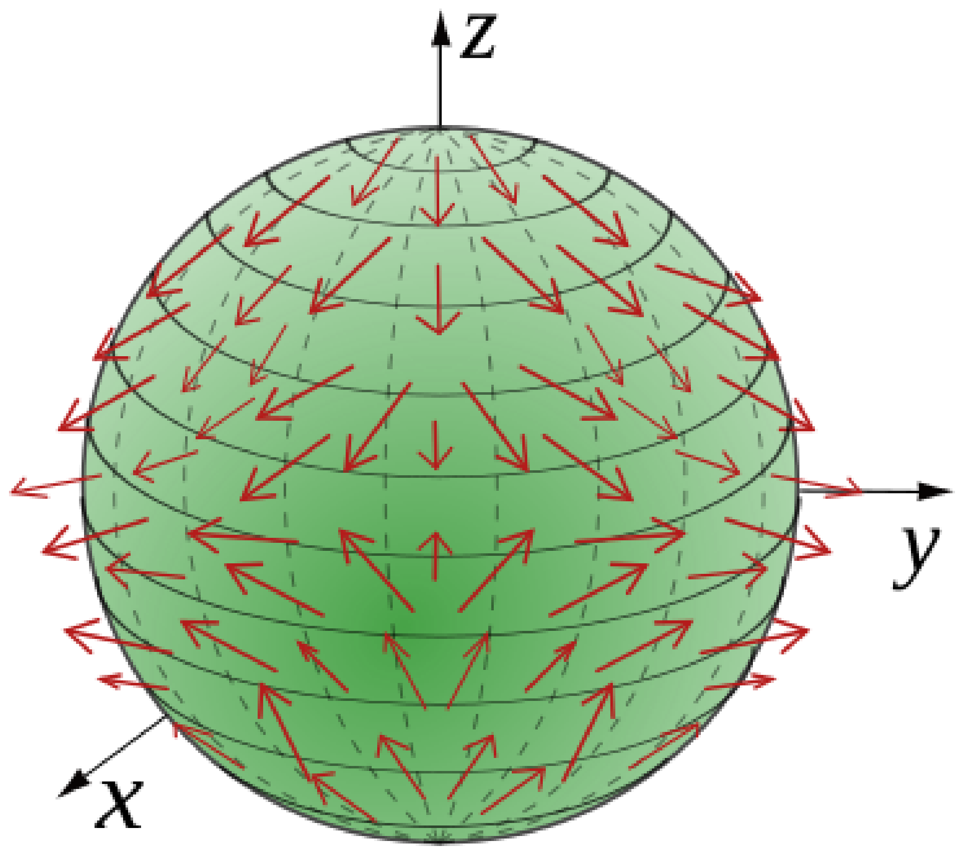
Whatever this thing is would be fundamentally connected to how basis vectors change smoothly from point to point and chart to chart. We don't want wild changes in our basis vectors as we flip from page to page in our atlas. They will assuredly change, the must, but we wish to set things up such that the change in basis vector from page to page will be expressible using this $\Gamma^k_{i\mu}$. We ONLY want LINEAR changes between charts/pages.

We might even say this $\Gamma^k_{i\mu}$, whatever it is, "connects" the pages/charts of our atlas together and it too, in principle, may even change as we flip through our atlas. It only tells us how to get to nearby basis vectors. We call it a "connection".

In addition it may even allow up to connect vectors at one point to other vectors far away in some other tangent space: the tangent spaces at x and x' are entirely different places and we cant compare vectors from different pages of our atlas unless we "connect" them somehow.

But first let us ask "can we handle vector fields on an arbitrary manifold $\mathcal{M}$?"

6.2 Vector Fields



Following our rule that a vector is anything which transforms in the same way as dx^i under coordinate transformation. A vector field is defined to be a vector at each point (or even just a subset I guess) on our manifold. Once we choose coordinates x^i then a general vector field MUST transform as

$$W'^{\mu}(x') = S^{\mu}_{\nu}(x)W^{\nu}(x)$$

Check this

We furthermore can define a co-vector field the same way: just lower the index and make sure it transforms with S^{-1}

$$W'_\mu(x') = (S^{-1})^\nu_\mu(x) W_\nu(x)$$

We could do this for tensor fields as well: more indices, more transformation

$$T'^{\mu\nu}(x') = S^\mu_\rho S^\nu_\sigma T^{\rho\sigma}(x)$$

and similarly for the tensor fields $T_{\mu\nu}$ and T^μ_ν with their appropriate transformation laws.

However, we will encounter objects that "look" like a tensor and which have indices like this but they will NOT transform as above. Don't blindly claim something is a vector or a tensor unless you've done an explicit coordinate change!

Question: if

$$\frac{\partial e_i}{\partial x^\mu} = \Gamma^k_{i\mu} e_k$$

then is $\Gamma^k_{i\mu}$ a $(1,2)$-tensor?

Let's check. We know how e_i and e^μ change under coordinate transformation. First things first: we ought to isolate this $\Gamma^k_{i\mu}$

$$\Gamma^k_{i\mu} = \Gamma^k_{i\mu} e_k \cdot e^k = e^k \frac{\partial e_i}{\partial x^\mu}$$

Next, transform to primed coordinates

$$\Gamma'^k_{i\mu} = e'^k \frac{\partial e'_i}{\partial x'^\mu}$$

but

$$e'^k = \frac{\partial x'_k}{\partial x^n} e^n$$

$$e'_i = \frac{\partial x^l}{\partial x'^i} e_l$$

then, plugging in

$$\Gamma'^k_{i\mu} = \frac{\partial x'_k}{\partial x^n} e^n \frac{\partial}{\partial x'^\mu} \left(\frac{\partial x^l}{\partial x'^i} e_l \right)$$

Boy howdy, we've got a product rule. Let's proceed

$$\Gamma'^{k}_{i\mu} = -\frac{\partial x'_k}{\partial x^n} e^n \left(\frac{\partial^2 x^l}{\partial x'^{\mu} \partial x'^{i}} e_l + \frac{\partial x^l}{\partial x'^{i}} \frac{\partial e_l}{\partial x'^{\mu}} \right)$$

Then this beast has to be begrudgingly distributed through to yield

$$\Gamma'^{k}_{i\mu} = -\frac{\partial x'_k}{\partial x^n} \frac{\partial^2 x^l}{\partial x'^{\mu} \partial x'^{i}} e^n \cdot e_l + \frac{\partial x'_k}{\partial x^n} \frac{\partial x^l}{\partial x'^{i}} \frac{\partial x^m}{\partial x'^{\mu}} e^n \cdot \frac{\partial e_l}{\partial x^m}$$

but $e^n \cdot e_l = \delta^n_l$ and

$$e^n \cdot \frac{\partial e_l}{\partial x^m}$$

is just our definition of $\Gamma^n_{l\mu}$ - our old one. So all together we have

$$\Gamma'^{k}_{i\mu} = -\frac{\partial x'_k}{\partial x^n} \frac{\partial^2 x^n}{\partial x'^{\mu} \partial x'^{i}} + \frac{\partial x'_k}{\partial x^n} \frac{\partial x^l}{\partial x'^{i}} \frac{\partial x^m}{\partial x'^{\mu}} \Gamma^n_{l\mu}$$

What a mess! Now, dear reader, is this how we defined how a tensor transforms? Heck no! a Tensor transforms as

$$T'^{\mu}_{ij}(x') = S^{\mu}_a (S^{-1})^b_i (S^{-1})^c_j T^a_{bc}(x)$$

Whatever Γ^k_{ij} is, it definitely ain't a tensor!

Moving on, scalar fields simply transform as

$$\psi'(x') = \psi(x)$$

A scalar field is a number defined at various points in $\mathcal{M}$ which generally can vary from place to place (e.g. Temperature, energy, etc). it is a map

$$\psi : \mathcal{M} \to \mathbb{R}$$

Next, given the metric and a D-dimensional manifold we can also find the volume element in a given coordinate system as

$$dV = \sqrt{\det g} dx^1 dx^2 dx^3 = \sqrt{g} dx^1 dx^2 dx^3$$

or

$$dV = \sqrt{g} d^D x$$

where g will henceforth represent the determinant of the matrix given by g_{ij}. Show that this is true.

Next, under a coordinate change we expect the volume element to be invariant but have a different form. Show that

$$dV' = \sqrt{g}' d^D x' = \sqrt{g} d^D x = dV$$

6.2.1 Some Interesting Metrics

In a flat Euclidean space we have a metric tensor given by δ_{ij} since we measure our line element using Pythagoras

$$ds^2 = dx^2 + dy^2 + dz^2$$

In special relativity we combine space and time into a "spacetime" manifold such that the line element given by

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 = -c^2 dt^2 + d\vec{x}^2$$

We say that spacetime is Minkowski flat when this metric is manifest. It is not Euclidean but it evidently is not curved either. It's a very strange version of Pythagoras as though the time direction were an imaginary coordinate. Surprising things can happen in this geometry physically. For example, electro-magnetic radiation travels under the condition that

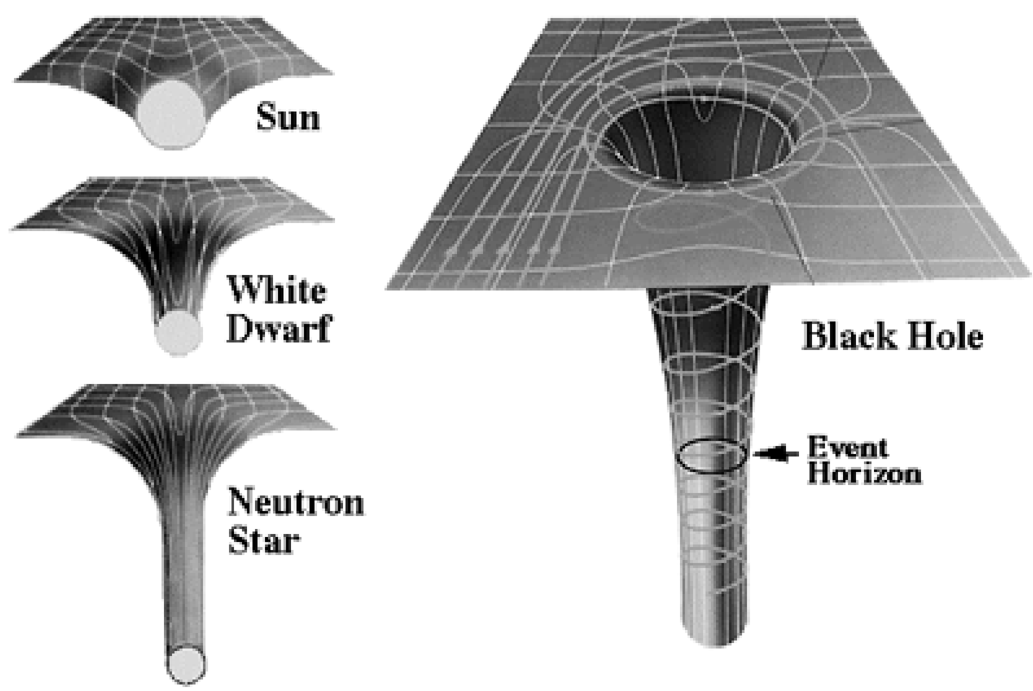
$$ds = 0 \implies dx^2 + dy^2 + dz^2 = c^2 dt^2$$

That's not too surprising. What is surprising in this view is that length and time are fundamentally connected. This geometry leads directly to the notion of length contraction and time dilation and all of the strange relativistic effects that show up in our verifiable reality!

Another common geometry is that of space-time near a massive object and can be found to be

$$ds^2 = ds^2 = \left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2.$$

Notice that as $r \to \infty$ this reduces to the Minkowski metric. Einstein's general relativity field equation allows this solution to be found. Not only does it describe spherically symmetric stars it can also describe a black hole when $r \to r_s$. Something horrific happens to this metric at $r = r_s$! It can be shown however, via a coordinate change, that nothing too crazy happens when $r = r_s$. You could cross this point and not even know it according to this geometry. It would be locally Minkowski flat. However, the physics says that nothing can move across this special boundary from the inside to the outside. It is called the "event horizon".

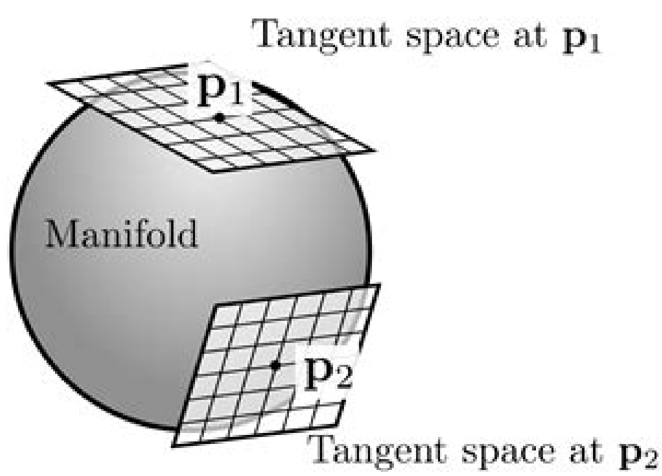


7 Calculus on Arbitrary Manifolds

Now, let's study the consequence of building our atlas of charts and the fact that our basis vectors and co-vectors can vary from point to point on our manifold $\mathcal{M}$

Imagine some multi-dimensional curved manifold (I envision some kind of 2 dimensional wobbly sheet). We lay down a bunch of small Cartesian coordinate patches (since manifolds are locally flat) which generally overlap in potentially some complicated way. What we previously called our charts.

At every point on this manifold we may define both a tangent and cotangent space (again, I envision a sheet of paper kissing my manifold at exactly one point and as I move from point to point my sheet of paper undulates around at different angles).



Therefore we have basis vectors at a point $p \in \mathcal{M}$ given as (picking two dimensions for clarity)

$$e_1(x_p) = \frac{\partial}{\partial x^1}$$

$$e_2(x_p) = \frac{\partial}{\partial x^2}$$

Notice that these basis vectors are unique to each point on the manifold and generally change from one place to another which we will write generally as $e_i(x)$.

The vectors at p (which live within the tangent space we call T_pM) are not the same as the vectors at p' (which live within the tangent space $T_p'M$). We hope to develop some way to "connect" them in a way such that we could compare them between different points which by now you know depends crucially on that Γ symbol.

Imagining a point on the 2-sphere embedded in our $\mathbb{E}^3$ space we see that any vectors at that point don't even "live" in our manifold (think of an arrow's tail at P - the arrow sort of pokes "outside" the surface of our sphere into the 3^{rd} dimension).

So it's not like we can form a derivative like

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

unless we can "connect" our two vector spaces. Let's try

$$\vec{v} = v^i(x)e_i(x)$$

$$\frac{\partial \vec{v}(x)}{\partial x^j}$$

is the partial in the x^j direction. So, let's plug in our definition of $\vec{v}$

$$\frac{\partial v^i(x)e_i(x)}{\partial x^j} = \frac{\partial v^i(x)}{\partial x^j} e_i(x) + v^i(x)\frac{\partial e_i(x)}{\partial x^j}$$

Now, harken back to what we learned about the derivatives of bases vectors - we formed a "connection" called Γ^a_{ij} such that

$$\frac{\partial \vec{v}(x)}{\partial x^j} = \frac{\partial v^i(x)}{\partial x^j} e_i(x) + v^i(x)\Gamma^k_{ij} e_k(x)$$

Now because we are free to choose the names of any indices that are summed we exchange k and i to yield

$$\frac{\partial \vec{v}(x)}{\partial x^j} = \frac{\partial v^i(x)}{\partial x^j} e_i(x) + v^k(x)\Gamma^i_{kj} e_i(x)$$

Now we can "factor out" e_i and we get

$$\frac{\partial \vec{v}(x)}{\partial x^j} = \left[\frac{\partial v^i(x)}{\partial x^j} + v^k(x)\Gamma^i_{kj}\right] e_i(x)$$

or, equivalently

$$\frac{\partial v^i(x)}{\partial x^j} = \left[\frac{\partial v^i(x)}{\partial x^j} + v^k(x)\Gamma^i_{kj}\right]$$

Which, admittedly looks suspicious since we have $\frac{\partial v^i(x)}{\partial x^j}$ on both sides! Let's fix this with some notations and call the left side "the covariant derivative" and then explain why we can do this below

$$\frac{Dv^i(x)}{Dx^j} = v^i_{;k} = \left[v^i_{,j} + v^k\Gamma^i_{kj}\right]$$

Where we introduced the semi-colon notation to mean "perform a covariant derivative on v^i" and $v^i_{,j} = \partial_j v^i$.

Further still we could abstractly write something like

$$\frac{D}{Dx^j} = \left[\partial_j + \Gamma^i_{kj}\right]$$

Where this operator would act on vector components v^i in the appropriate way.

In an entirely analogous way would could covariantly differentiatie covectors. Show that

$$\frac{D\omega_i(x)}{Dx^j} = \omega_{i;k} = \left[\omega_{i,j} - \omega_k\Gamma^k_{ij}\right]$$

Be very careful with the placement of your indices!

There we have it, a way to differentiate on an arbitrary manifold. But what is geometrically going on here?

7.1 Another Route

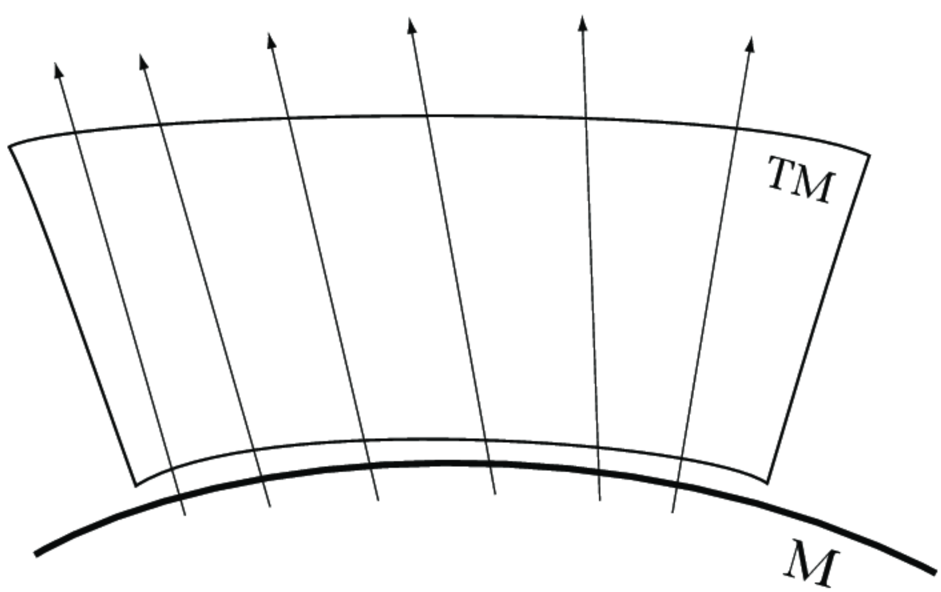
Suppose now you didn't know what we just learned in establishing the covariant derivative: you got amnesia and have to redefine it.

Since we have no way to compare vectors at different points on our manifold then we don't know how to add/subtract vectors from different spaces. They have to be in the same space to add/subtract!

Is all hope lost then?

7.1.1 Tangent Bundle Idea

One idea might be to gather up ALL the tangent spaces T_xM at every point $x \in \mathcal{M}$ and their vectors (labeled by which point they came from) and toss them into a big bag in hopes we might be able to compare them in some "higher" space.



This might be written as

$$TM = \bigsqcup_x T_xM = \{(x, v_y) \mid x \in M, v_y \in T_xM\}$$

This vector space is called the "tangent bundle" of the manifold $\mathcal{M}$ and the squared cup symbol means we take the "disjoint union". This is just like the regular union only we keep track of where the vectors came from.

For example, in the disjoint union we might have two identical vectors where one came from x and the other from x'. In the regular union we would only include ONE of those two vectors since they're identical. It's best to think of TM as a bag full of ALL vectors such that when you pull one out of the bag it has a label telling you which x it came from!

This is a VERY fruitful path to pursue (and is one of the most elegant in all of mathematics) but for our purposes not intuitively clear how a comparison could be made in this manner. We will save this for much later when we rework everything entirely abstractly without coordinates anywhere.

7.1.2 Dragging Vectors Idea

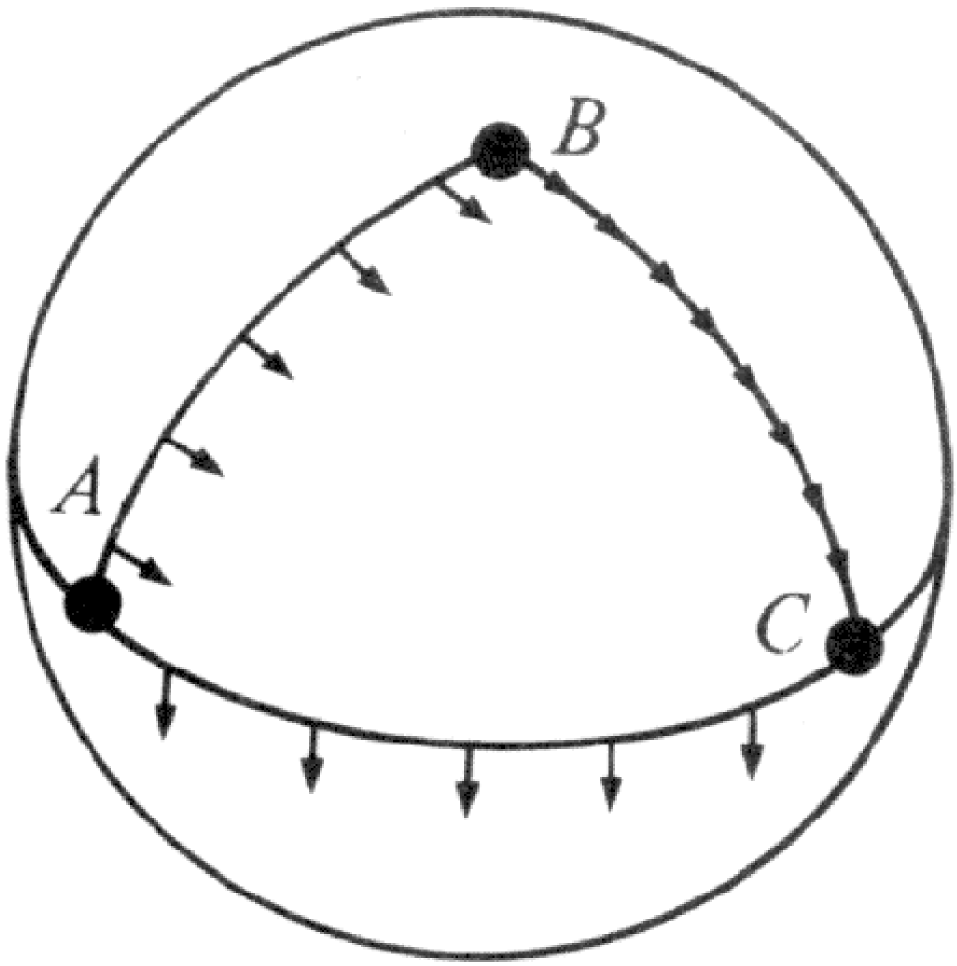
Another idea would be to pick a curve $\gamma(t)$ in our manifold and define the tangent vector in the direction of that curve (velocity!) and then ever so carefully drag our candidate vector along the curve at all times making sure its direction relative to tangent/velocity vector of the curve doesn't change.

Once you arrive at your target point where you wish to compare then you can place that vector you dragged into the tangent space and compare. This is almost like what you've done ion intro physics where you were told "move the vectors around; just don't change their direction or length". When we do this on an arbitrary manifold, however, the direction of the vector may very well appear to change - not because it itself is changing but because the basis vectors are changing from point to point!

In this manner, for an arbitrary surface, we will have our original vector from T_xM and our velocity vector from T'_xM in the same space. At this point since we dragged along our vector from T_xM to T'_xM we can add/subtract and compute anything else since we have brought the vector to x'.

7.1.3 A Spherical Example

Let's imagine a specific example. Consider the 2-sphere and you're at the north pole. Suppose you have two gyroscopes with angular momentum pointing in some identical direction - say, perfectly horizontal (which is to say tangent to the surface) at the north pole. Now you hop into a car (or plane) with one of the gyros and drive directly towards the equator always forcing your gyroscope to be tangent to the direction of your path (you indeed must manually force it since gyros want to stay pointing in the same direction - notice we always want the vector to be parallel to the path we take - so, on a sphere, this direction is changing).



By the time you get to the equator you notice your vector has not changed one bit - it is still tangent to the surface in your local frame of reference. A person on the moon would observe your vector starting in the $+y$ direction at the north pole and then be in the $-z$ direction at the equator. You and your friend will disagree, but your friend has the benefit of living in three dimensions whereas you are stuck on a two dimensional sphere - in fact, you might not even KNOW the Earth is spherical. Based upon what you know about gyroscopes you find it strange that you have to apply a torque to keep it in the same direction (but maybe you haven't learned about conservation of angular momentum so it doesn't concern you very much!)

Next, you turn (your vector doesn't!) and drive along the equator until you have traveled across 90° longitude - again, keeping your vector tangent in the same direction according to you. Your friend on the moon sees your vector stay in the same $-z$ direction. You, however, notice nothing other than requiring to physically force the vector to not change. You can change direction, but not your vector!

Finally, again keeping your vector in the same orientation in your local frame you travel all the way back up to the north pole to meet up with your other gyro. Your moon friend now sees your vector pointing in his $+x$ direction. You, however, meet up with your static gyro to discover they aren't in the same direction at all! They're 90° perpendicular to each other! You are now extremely confused. You kept your vector in the same direction ALWAYS along your path yet somehow the gyroscopes disagree! Conundrum indeed!

You decide to take a similar trip only this time you don't got nearly as far: you start at the north pole travel for several miles and make your way back (not along your original path). When you return you find your gyros pointing in different directions but the angle between them is less than 90° this time. Next you try again, only instead of a loop, you simply back-track your path and find the gyros agree completely. It appears to only be different if you traverse a loop.

You might try a very small experiment and simply walk around for a few feet and return. Here you barely notice any difference between the gyros at all. To you, on small scales, it looks like you're free to move vectors around so long as their direction and length doesn't change. That's what you've been doing since you were a kid and your teacher told you "you're free to move your vectors around in the plane so long as they don't change direction and magnitude ". You studied Newton and learned to copy/paste all your vectors to be at your origin so you can decompose them and solve problems.

Yet here you tried to do exactly that: you took a vector and kept its length and direction constant yet magically the vector changed direction when you brought it back! Mysterious indeed. You've discovered that vectors absolutely cant be copy/pasted to the origin - it matters what path you take.

Your friend on the moon knows what's REALLY happening since she has the benefit of three dimensional space. The vector is trying its damndest to stay in the same direction but YOU are changing it! Your friend knows about conservation of angular momentum - you don't! The gyro would prefer to always point in the same direction as it goes around the Earth.

What's more is your friend also realizes you live on a sphere. As you travel your paths curves and changes vector orientations. You friend knows that your vector at the north pole lives in the tangent space at the north pole which is entirely different than the tangent space anywhere else. You do not live in a flat, Euclidean, space: it is manifestly curved!

In our 2D flat Euclidean Cartesian plane we could slide vectors around and compare vectors at two different points with no problems. This is because the tangent space of the plane at one point is precisely the same tangent space at any other. There was no path dependence to our vectors and we could define derivatives with ease. The basis vectors were constant everywhere.

In our general manifold case we cant compare or add/subtract vectors naively. But, our experiment with our choosing a path and carrying our vector from one place to another DOES indeed allow us to compare vectors: that's what you did when you brought your gyro back to the north pole: you compared and found they were different by exactly $90°$ when you went by the route to the equator! You've discovered that vectors will change in proportion to the area of your loop. curious: let's put that fact in our pocket for now.

Now clearly, had you gone straight down to the equator and straight back up your vectors wouldn't be different at all - just like if you do a karate chop with your hand and bring it back up.

Clearly there's something funny about traveling in a loop - try this experiment with your hand laying flat: start by pointing it up, come down with your elbow, rotate your arm by ninety degrees and come back up. You'll find you hand has twisted.

There appears to be some deep relationship between traveling in a loop and the fundamental geometry of the space.

What this experiment has done is the notion of "parallel transport" and it allows us to define the "covariant derivative" of a vector.

In flat Euclidean space we expect our definition of derivative will reduce to our everyday derivative. if the basis vectors don't vary from place to place then their derivatives are identically zero. A non-Euclidean geometry however, introduces more complexity due to the parallel transport across the surface.

Let's define our derivative as follows. Given a vector W^μ on $\mathcal{M}$ at x then we have, via Taylor, to first order in δx^ν

$$W^\mu(x + \delta x) = W^\mu(x) + \partial_\nu W^\mu \delta x^\nu$$

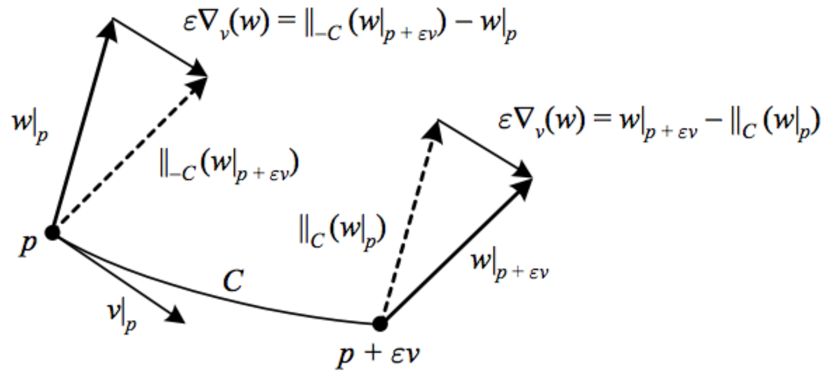
This is just the definition of our usual derivative

$$\partial_\nu W^\mu = \frac{W^\mu(x + \delta x) - W^\mu(x)}{\delta x^\nu}$$

Now, from our discussion above we wish to compare a vector at $x + \delta x$ with the parallel transported vector along $\gamma(t)$. Therefore let us write this parallel transported vector as

$$W^\mu_\parallel(x) = W^\mu(x) - \bar{\delta} W^\mu(x)$$

to first order, for small δx^ν we anticipate this $\bar{\delta} W^\mu(x)$ term (the piece caused by dragging our vector around) will be linear in δx^ν. For a curved surface we anticipate our vector W^μ "loses" a bit of itself as it travels.



We shall write this as a linear combination of our original vector

$$\bar{\delta} W^\mu(x) = \Gamma^\mu_{\rho\nu} W^\rho(x) \delta x^\nu$$

That is to say that the vector at x^μ generally loses part of itself when parallel transported to $x^\mu + \delta x^\mu$ and this will generally be some complicated linear combination of the original $W^\mu(x)$ but fundamentally dependent on the metric. Note that this assumption has two repeated indices we must sum over. $\Gamma^\mu_{\rho\nu}$ is some box of numbers which we suspect will depend on the geometry of our manifold but it has not yet been defined.

Now we shall define our "covariant" derivative in the manner that allows us to compare vectors from two different tangent spaces

$$\frac{DW^\mu(x)}{dx^\nu} = \lim_{\delta x^\nu \to 0} \frac{W^\mu(x+\delta x) - W^\mu_\parallel(x)}{\delta x^\nu}$$

We often write this in a variety of different ways:

$$\frac{DW^\mu(x)}{dx^\nu} = D_\nu W^\mu(x) = \nabla_\nu W^\mu$$

We will almost always use $D_\nu W^\mu(x)$.

Now, we have an expressions for $W_{\parallel}^{\mu}(x)$ so plugging in we have

$$D_\nu W^\mu(x) = \lim_{\delta x^\nu \to 0} \frac{W^\mu(x+\delta x) - (W^\mu(x) - \bar{\delta} W^\mu(x))}{\delta x^\nu}$$

But look, if we separate this into two terms then the first term looks EXACTLY like the normal derivative of a vector fields we know and love. i.e.

$$\begin{aligned} D_\nu W^\mu(x) &= \lim_{\delta x^\nu \to 0} \frac{W^\mu(x+\delta x) - W^\mu(x)}{\delta x^\nu} + \frac{\bar{\delta} W^\mu(x)}{\delta x^\nu} \\ &= \partial_\nu W^\mu + \lim_{\delta x^\nu \to 0} \frac{\bar{\delta} W^\mu(x)}{\delta x^\nu} \end{aligned}$$

Now, we also have a first order assumption about the form of $\bar{\delta} W^\mu(x)$ from above. All together then we have

$$D_\nu W^\mu(x) = \partial_\nu W^\mu + \lim_{\delta x^\nu \to 0} \frac{\Gamma^\mu_{\rho\nu} W^\rho(x) \delta x^\nu}{\delta x^\nu}$$

And thus we see the limit drops out leaving us with the complete definition of our brand new general version of the derivative on a curved manifold as

$$D_\nu W^\mu(x) = \partial_\nu W^\mu + \Gamma^\mu_{\rho\nu} W^\rho(x)$$

Just like we found more easily earlier!

But now we have to figure out what this $\Gamma^\mu_{\rho\nu}$ could possibly be. We hope, beyond hope, that for our everyday Euclidean world the complicated object is precisely zero. In the case of a flat manifold then the covariant derivative is exactly the derivative we grew up with!

Clearly this $\Gamma^\mu_{\rho\nu}$ must depend, in some complicated way, on the metric $g_{\mu\nu}$ and since the whole thing is a derivative then we expect that $\Gamma^\mu_{\rho\nu}$ might be something like $\partial_\rho g_{\mu\nu}$ which is three distinct indices reminiscent of this Γ guy.

Furthermore, we also know that as we move from place to place on our manifold that since the tangent spaces are wobbling about then our basis vectors themselves wobble about and so we might be interested in studying how these basis vectors change from place to place. i.e.

$$\partial_\nu e_\mu(x)$$

Remember, basis VECTORS have a down index; components of a vector have an up index.

7.1.4 Polar Coordinates Again!

As an example to guide us let's consider polar coordinates of the plane. Here the basis vectors depend on where we are in the plane (r, θ). We have

$$\hat{r} = e_r = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$\hat{\theta} = e_\theta = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

now, clearly, these do not depend on r so that $\partial_r e_i = 0$. Then we compute $\partial_\theta e_i$.

$$\partial_\theta e_r = \partial_\theta \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$\partial_\theta e_\theta = \partial_\theta \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} -\cos\theta \\ -\sin\theta \end{bmatrix}$$

Now we ask "can we write these derivatives of the basis vectors as linear combinations of our original $\hat{r}, \hat{\theta}$ basis vectors?" Sure, we can always do that: the derivative doesnt take us "out" of our vector space. The derivatives change the input vectors into something else but we could always express that new vector in terms of the old. Doing that here we have

$$\partial_\theta e_r = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = e_\theta$$

$$\partial_\theta e_\theta = \begin{bmatrix} -\cos\theta \\ -\sin\theta \end{bmatrix} = -e_r$$

Here we have found the linear combinations. In general this linear combination might be written

$$\partial_i e_j = \sum_a \Gamma^a_{ij} e_a$$

Usually we will not explicitly write the summation. The Γ are the coefficients of the linear combination of basis vectors

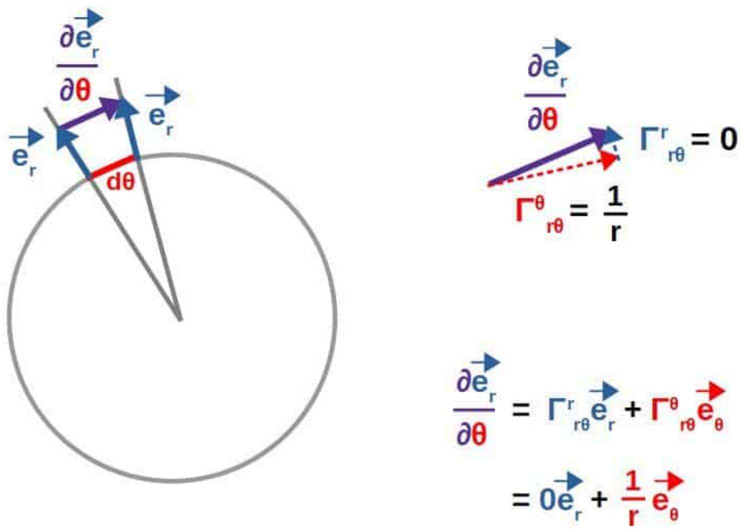
They happen to be simple in our polar flat case. For the four possibilities of $(i, j) = (r, r), (\theta, \theta), (r, \theta), (\theta, r)$ we found the coefficients of the linear combination to be 0 if $i = r$ and The other two non-zero ones are ± 1. Writing them out completely we have

$$\partial_\theta e_r = e_\theta = \Gamma^r_{\theta r} e_r + \Gamma^\theta_{\theta r} e_\theta$$

and

$$\partial_\theta e_\theta = -e_r = \Gamma^r_{\theta\theta} e_r + \Gamma^\theta_{\theta\theta} e_\theta$$

In these two equations above all of the Γ terms are individual numbers. The only way the first equation could be true is if $\Gamma^r_{\theta r} = 0$ and $\Gamma^\theta_{\theta r} = 1$. Similarly for the second equation $\Gamma^r_{\theta\theta} = -1$ and $\Gamma^\theta_{\theta\theta} = 0$.



Then we can write out the full set of Γ's as

$$\Gamma^r_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\Gamma^\theta_{ij} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Note that these three index objects called Γ are like a "cube" of numbers - in this case it is a 2x2x2 box and so I wrote out both the r and θ parts. It's relatively easy to visualize these things in the 2D plane but in higher dimensions it becomes much harder. You can always use the upper index to write out each individual matrix for specific coordinates of the upper index.

Now, in our example, even though we know know know that the $2D$ plane is flat we found non-zero Γ's. Our Covariant derivative says

$$D_\nu W^\mu(x) = \partial_\nu W^\mu + \Gamma^\mu_{\rho\nu} W^\rho(x)$$

What if we had a constant radial vector field in the plane? For example

$$W = \begin{bmatrix} 6 \\ 0 \end{bmatrix} = 6e_r + 0e_\theta$$

Then our derivative should wind up being zero. Is that what we get? Well the first term, the ordinary derivative is certainly zero. What about the second term? write it out

$$D_i W^j = \Gamma^j_{\rho i} W^\rho = \Gamma^j_{ri} W^r + \Gamma^j_{\theta i} W^\theta$$

Now, we haven't explicitly written down Γ^j_{ri} or $\Gamma^j_{\theta i}$ - notice we had written down the matrices if the upper indices were known such as Γ^r_{ai}

However, if we choose Γ^r_{ri} and Γ^θ_{ri} then these symbols would be the columns of the Γ's we found. They have only one undetermined symbol which means they should look like a row of numbers.

$$\Gamma^r_{ri} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$\Gamma^\theta_{ri} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

Which mean our covariant derivative $D_i W^j = 0$! Whew! That was the single most difficult easy derivative you've ever done in your life!

Of course the derivative of a constant vector in the flat plane is zero! We would've been in trouble if it weren't! The whole point is that this new type of derivative can handle even the curviest of spaces!

These Γ are called the "Christoffel symbols". They encode how the basis vectors tilt around as you move from point to point on your manifold. We have a common notation that is used in this tensor calculus

$$\partial_\nu e_\mu(x) = \frac{\partial e_\mu}{\partial x^\nu} = e_{\mu,\nu}$$

If we use this comma notation to represent partial derivatives then our covariant derivative becomes

$$D_\nu W^\mu(x) = W^\mu_{\ ,\nu} + \Gamma^\mu_{\rho\nu} W^\rho(x) = W^\mu_{\ ;\nu}(x)$$

Where in the last line we have defined the semi-colon notation for the covariant derivative. We will make ample use of the comma but I will tend to write the capital D for the covariant derivative for clarity. However, be advised that the comma/semi-colon notation is fairly standard in the literature.

Now, geometrically, what is this covariant derivative? It looks like it would transform like a co-vector due to its downstairs index D_ν but we're skeptical since we previously found Γ certainly does not transform as a tensor: it has a VERY wacky transformation law!

If we know how it transforms and we know how vectors transform then we'll know how something like $D_\nu W^\mu$ transforms. No way around it, we HAVE to check.

$$D_{\nu'} W^{\mu'} = \frac{\partial}{\partial x^{\nu'}} W^{\mu'} + \Gamma'^{\mu'}_{\nu'\lambda'} W'^{\lambda'}$$

First let's consider

$$\frac{\partial}{\partial x^{\nu'}} W^{\mu'} = \frac{\partial}{\partial x^{\nu'}} \left(\frac{\partial x^{\mu'}}{\partial x^{\alpha}} W^{\alpha} \right)$$

Invoke the product rule to find

$$\frac{\partial}{\partial x^{\nu'}} W^{\mu'} = \frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} + \frac{\partial x^{\mu'}}{\partial x^{\alpha}} \frac{\partial W^{\alpha}}{\partial x^{\nu'}}$$

but

$$\frac{\partial W^{\alpha}}{\partial x^{\nu'}} = \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial W^{\alpha}}{\partial x^{\sigma}}$$

or

$$\frac{\partial}{\partial x^{\nu'}} W^{\mu'} = \frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} + \frac{\partial x^{\mu'}}{\partial x^{\alpha}} \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial W^{\alpha}}{\partial x^{\sigma}}$$

That's our first piece and it's definitely not a vector or tensor or anything else geometrically nice! Now for the second piece, the Γ - but we've transformed this guy before - we found

$$\Gamma^{\mu'}_{\nu'\lambda'} = \frac{\partial x^{\mu'}}{\partial x^{l}} \frac{\partial x^{m}}{\partial x^{\nu'}} \frac{\partial x^{n}}{\partial x^{\lambda'}} \Gamma^{l}_{mn} - \frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\lambda'}}$$

Then this piece is multiplied by $W^{\mu'}$

$$W^{\lambda'} = \frac{\partial x^{\lambda'}}{\partial x^{\beta}} W^{\beta}$$

Now put everything together

$$D_{\nu'} W^{\mu'} = \left[\frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} + \frac{\partial x^{\mu'}}{\partial x^{\alpha}} \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial W^{\alpha}}{\partial x^{\sigma}} \right]$$

$$+ \left[\frac{\partial x^{\mu'}}{\partial x^{l}} \frac{\partial x^{m}}{\partial x^{\nu'}} \frac{\partial x^{n}}{\partial x^{\lambda'}} \Gamma^{l}_{\ mn} \frac{\partial x^{\lambda'}}{\partial x^{\beta}} W^{\beta} - \frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\lambda'}} \frac{\partial x^{\lambda'}}{\partial x^{\beta}} W^{\beta} \right]$$

Dear reader, this is a horrific but necessary computation! However, we do have some cancellations, namely the derivative orders can be switched (since we have no torsion or anything too crazy).

The first piece of the second term becomes

$$\frac{\partial x^{\mu'}}{\partial x^{l}} \frac{\partial x^{m}}{\partial x^{\nu'}} \frac{\partial x^{n}}{\partial x^{\beta}} \Gamma^{l}_{\ mn} W^{\beta}$$

and the second piece of the second term simplifies to

$$-\frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\lambda'}} \frac{\partial x^{\lambda'}}{\partial x^{\alpha}} W^{\alpha} = -\frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha}$$

all together we have

$$\left[\frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} + \frac{\partial x^{\mu'}}{\partial x^{\alpha}} \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial W^{\alpha}}{\partial x^{\sigma}} + \frac{\partial x^{\mu'}}{\partial x^{l}} \frac{\partial x^{m}}{\partial x^{\nu'}} \frac{\partial x^{n}}{\partial x^{\beta}} \Gamma^{l}_{\ mn} W^{\beta} - \frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} \right]$$

notice that in the first term we can write

$$\frac{\partial^2 x^{\mu'}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha} = \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial^2 x^{\mu'}}{\partial x^{\sigma} \partial x^{\alpha}} W^{\alpha}$$

and the second derivative term in the Γ piece is

$$-\frac{\partial x^{\mu'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\alpha}} W^{\alpha}$$

Which, it just so happens we can move the derivatives around such that these two pieces cancel!

What we're left with is

$$D_{\nu'}W^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^{\alpha}}\frac{\partial x^{\sigma}}{\partial x^{\nu'}}\frac{\partial W^{\alpha}}{\partial x^{\sigma}} + \frac{\partial x^{\mu'}}{\partial x^{l}}\frac{\partial x^{m}}{\partial x^{\nu'}}\frac{\partial x^{n}}{\partial x^{\lambda'}}\Gamma^{l}_{mn}\frac{\partial x^{\lambda'}}{\partial x^{\beta}}W^{\beta}$$

$$= \frac{\partial x^{\mu'}}{\partial x^{\alpha}}\frac{\partial x^{\sigma}}{\partial x^{\nu'}}D_{\sigma}W^{\alpha}$$

Which is our transformation law for a $(1,1)$ tensor! This is what folks mean when they say tensor calculus is index gymnastics!

Now we have opened an entire world where we can do some basic calculus within an arbitrary manifold where basis vectors vary from point to point. It all hinges on defining a "connection" Γ. It turns out that this choice is not unique - there can be many equivalent choices for how to connect our tangent and cotangent spaces between points.

In our formalism for differential geometry we have been describing specific components of abstract things: that is to say, we choose a coordinate system and then write things like $v_{i;j}$. This is a coordinate dependent description.

Mathematicians, however, prefer a coordinate free description. They define things without any reference to coordinates or components. In their view the metric is "a symmetric multi-linear form that takes two vectors and produces a real number". We will expand upon this view at the end of our studies. Needless to say the mathematician's abstract notions of maps and diagrams make it hard to see how one can compute with such heady topics: coordinates make things clear (at the expense of having to assume we've made a choice).

At any rate, we have found that, for a general manifold, under a chosen coordinate system, we have derivatives of basis vectors as

$$e_{i,j} = \Gamma^{a}_{ij}e_{a}$$

Suppose we dot this derivative with e_{ν}. Then

$$e_{\nu}\cdot e_{i,j} = e_{\nu}\cdot\Gamma^{a}_{ij}e_{a} = \Gamma^{a}_{ij}e_{\nu}\cdot e_{a} = \Gamma^{a}_{ij}g_{\nu a} = \Gamma_{\nu ij}$$

We can raise the ν index using the metric to find an expression for our Γ by its lonesome in terms of derivatives of our bases

$$\Gamma_{\nu ij}g^{a\nu} = \Gamma^{a}_{ij}$$

Furthermore, given the definition of our basis vectors we see that the lower indices of Γ are symmetric

$$e_a = \frac{\partial}{\partial x^a}$$

then

$$e_{i,j} = \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} = \frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i} = e_{j,i}$$

$$\implies \Gamma^a_{ij} = \Gamma^a_{ji}$$

Be advised that there ARE geometries that violate this condition on the connection we call Γ. In the case where this index symmetry doesn't apply we say there is "torsion"- another topic of study in differential geometry we shall discuss later. Almost everything you see in physics has no geometric torsion so we will make use of this symmetry constantly.

Now, since we found the derivatives of the basis vectors let us study the derivatives of

$$e_i \cdot e_j = g_{ij}$$

$$g_{ij,k} = \frac{\partial}{\partial x^k} e_i \cdot e_j = e_{i,k} \cdot e_j + e_i \cdot e_{j,k}$$

but $e_{i,k} = \Gamma^a_{ik} e_a$ and $e_j \cdot e_{i,k} = \Gamma_{jik}$
so that

$$g_{ij,k} = \Gamma_{jik} + \Gamma_{ijk}$$

but then we also have

$$g_{jk,i} = \Gamma_{kji} + \Gamma_{jki}$$

and

$$g_{ki,j} = \Gamma_{ikj} + \Gamma_{kij}$$

Which means we can find our Γ purely from the metric as

$$\Gamma^a_{ij} = \frac{g^{\mu a}}{2}\left(g_{ai,j} - g_{ij,a} + g_{ja,i}\right)$$

Once again, if you know the metric, you know everything! The connection is given by some derivatives of the metric which depends on your basis vectors!

Usually, we find our metric by writing the line element as

$$ds^2 = g_{ij}dx^i dx^j$$

and this makes evident what the metric components are!

8 The 2-Sphere Again

Recall that for a unit 2-sphere embedded in our Euclidean $3D$ space we had the surface of a 2-sphere parameterized by (θ, ϕ) under the coordinate map

$$\mathbb{X}(\theta,\phi) = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}$$

Next, we define the spherical basis vectors as

$$\hat{e}_i = \frac{\partial \mathbb{X}}{\partial x^i}$$

where $x^1 = \theta$ and $x^2 = \phi$

Therefore we have

$$\hat{e}_\theta = \begin{bmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ -\sin\theta \end{bmatrix}$$

Similarly,

$$\hat{e}_\phi = \begin{bmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{bmatrix}$$

Next, we can find the metric via

$$g_{\theta\theta} = \hat{e}_\theta \cdot \hat{e}_\theta$$

$$g_{\phi\theta} = \hat{e}_\phi \cdot \hat{e}_\theta$$

$$g_{\theta\phi} = \hat{e}_\theta \cdot \hat{e}_\phi$$

$$g_{\phi\phi} = \hat{e}_\phi \cdot \hat{e}_\phi$$

Computing each of these (so many marvelous cancellations!) we find

$$g_{\theta\theta} = 1$$

$$g_{\phi\theta} = 0$$

$$g_{\theta\phi} = 0$$

$$g_{\phi\phi} = \sin^2\theta$$

Therefore,

$$g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

and our line element is

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2$$

This is the same thing we found earlier by starting with Euclidean spherical coordinates and setting the r variable to be constant - which is essentially the same thing we did here, nothing new.

Next we can find our Γ's by finding all the derivatives of the metric

$$g_{ij,\theta} = \partial_\theta \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2\sin\theta\cos\theta \end{bmatrix}$$

$$g_{ij,\phi} = \partial_\phi \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Next we find the inverse metric easily since it is diagonal

$$g^{ij} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2\theta} \end{bmatrix}$$

and now we can put it all together to find the Γ's according to

$$\Gamma^a_{ij} = \frac{g^{\mu a}}{2}\left(g_{ai,j} - g_{ij,a} + g_{ja,i}\right)$$

we find

$$\Gamma^\theta_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & -\sin\theta\cos\theta \end{bmatrix}$$

$$\Gamma^\phi_{ij} = \begin{bmatrix} 0 & \cot\theta \\ \cot\theta & 0 \end{bmatrix}$$

Now, using this we can install a vector field on the sphere. To make things interesting lets consider a constant field and a field that depends on location

$$W^\mu(\theta,\phi) = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

$$V^\mu(\theta,\phi) = \begin{bmatrix} \phi^2 \\ \cos\theta + \cos\phi \end{bmatrix}$$

Then, we can find derivatives on this manifestly curved manifold!

$$D_\nu W^\mu(x) = \partial_\nu W^\mu + \Gamma^\mu_{\rho\nu} W^\rho(x)$$

$$D_\theta W^\theta(x) = \partial_\theta W^\theta + \Gamma^\theta_{\theta\theta} W^\theta(x) + \Gamma^\theta_{\phi\theta} W^\phi(x)$$

$$D_\theta W^\phi(x) = \partial_\theta W^\phi + \Gamma^\phi_{\theta\theta} W^\theta(x) + \Gamma^\phi_{\phi\theta} W^\phi(x)$$

$$D_\phi W^\phi(x) = \partial_\phi W^\phi + \Gamma^\phi_{\theta\phi} W^\theta(x) + \Gamma^\phi_{\phi\phi} W^\phi(x)$$

$$D_\phi W^\theta(x) = \partial_\theta W^\theta + \Gamma^\theta_{\theta\phi} W^\theta(x) + \Gamma^\theta_{\phi\phi} W^\phi(x)$$

Well, not we just plug in our vector and our Γ's

$$D_\theta W^\theta = \partial_\theta 5 + \Gamma^\theta_{\theta\theta} 5 + \Gamma^\theta_{\phi\theta} 10 = 0 + 0 \times 5 + 10 \times 0 = 0$$

$$D_\theta W^\phi = \partial_\theta 10 + \Gamma^\phi_{\theta\theta} 5 + \Gamma^\phi_{\phi\theta} 10 = 0 + 0 + 10 \cot\theta$$

$$D_\phi W^\phi = \partial_\phi 10 + \Gamma^\phi_{\theta\phi} 5 + \Gamma^\phi_{\phi\phi} 10 = 0 + 5\cot\theta + 0$$

$$D_\phi W^\theta(x) = \partial_\theta 5 + \Gamma^\theta_{\theta\phi} 5 + \Gamma^\theta_{\phi\phi} 10 = 0 + 0 + 0 - 10\sin\theta\cos\theta$$

all together we have

$$D_\nu W^\mu(x) = \begin{bmatrix} 0 & 10\cot\theta \\ -10\sin\theta\cos\theta & 5\cot\theta \end{bmatrix}$$

We have a constant vector field in (θ, ϕ) coordinates and its derivative is not just non-zero but it depends on where you are! This means we were too naïve in our understanding what a "constant" vector field on a sphere even is! What's more is, look at what happens near $\theta = 0$: this thing blows up!

This is a manifestation of the hairy ball theorem which says that it is impossible to define a vector field on the sphere unless there is a point at which the vector is zero. The theorem can be visualized by trying to comb the hair on a hairy ball: at some point you're guaranteed to have a cowlick! If our vector field were a "constant" wind at every point then at the two poles we would have a MASSIVE tornados!

What about our other vector field $V^\mu(x)$? I'll let you work that one out as an exercise!

9 Geodesics

One of the most pleasing aspects of our flat everyday Euclidean world is the fact that if you want to travel from point A to point B on a Euclidean plane then you should take the straight line path from A to B.

For us it's one of those "more obvious than obvious" facts but dwelling upon it you find that fact somehow emerges from the Pythagorean theorem. Upon further reflection it has something specific to do with the nature of the metric since

$$ds^2 = dx^2 + dy^2 + dz^2$$

We wish to use THIS to prove THAT. What is it about our Euclidean metric that produces paths of minimal distance? We should be able to find a way to use δ_{ij} to find $y = mx + b$ and if we can find a method that works with this metric then maybe we can generalize it for any metric.

We call the special curves of minimal length on a manifold a "geodesic".

So far we have found out how to find the metric and compute derivatives on a general smooth manifold. Now we shall study curves on the manifold. Our general line element is given by

$$ds^2 = g_{ij} dx^i dx^j$$

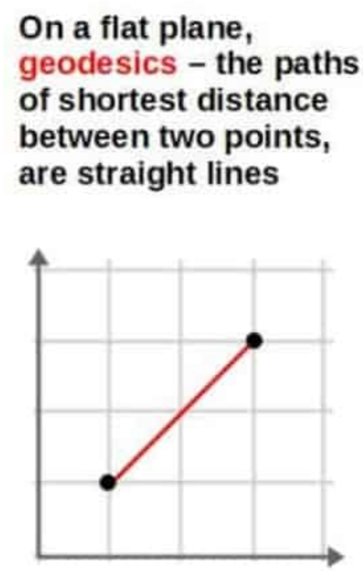
Given some curve $\gamma(t)$ parameterized by t we may integrate along the curve using this expression and all is well (despite the annoying square root).

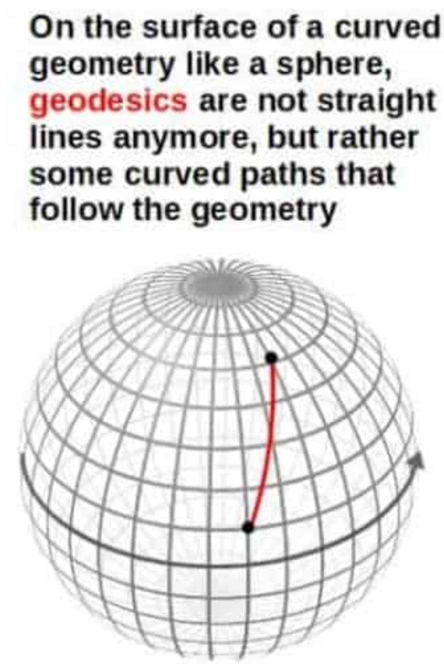
By $\gamma(t)$ I mean a designation on a manifold such as

$$\gamma(t) = \left[\begin{array}{c} x_f(t) = x_i + v_{ix} t + \frac{1}{2} a_x t^2 \\ y_f(t) = y_i + v_{iy} t + \frac{1}{2} a_y t^2 \end{array} \right]$$

in the case of the (x, y) plane. You know all about that particular curve from introductory physics. It's the parabola Galileo gave us and Newton clarified! A parabola is assuredly NOT a geodesic but it is a curve that minimizes something else called "Action" - via the principle of least action!

Suppose we wish to find the curve between two points which minimizes the total distance? This curve seems to be extremely important for the geometry of the manifold. For example, one could say that Euclidean space is the land where the shortest distance between any two points is a straight line. Or for a sphere the analogous curve would be a section of a great circle (that's why airplanes don't fly in straight line!).





How could we formalize this? It sounds exactly like the calculus of variations only now, instead of δ_{ij} we have a general metric and it would be near impossible to figure out how to replace all the implicit deltas with g_{ij}.

The crux of what we've done in flat land is we computed

$$\delta L[f] = 0$$

for some functional L. A functional is a mathematical map that takes a function and produces a number. We might say

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

where

$$f : X \to Y$$

such that

$$L[f] \in \mathbb{R}$$

That's essentially what a functional is. In flatland we studied

$$L[x, \dot{x}] = \int ds$$

If we similarly study this in general then when we minimize the length between two arbitrary points we will have a general equation that functionally depends on the metric as well as the cordinates.

$$L = \int \sqrt{g_{ij} dx^i dx^j}$$

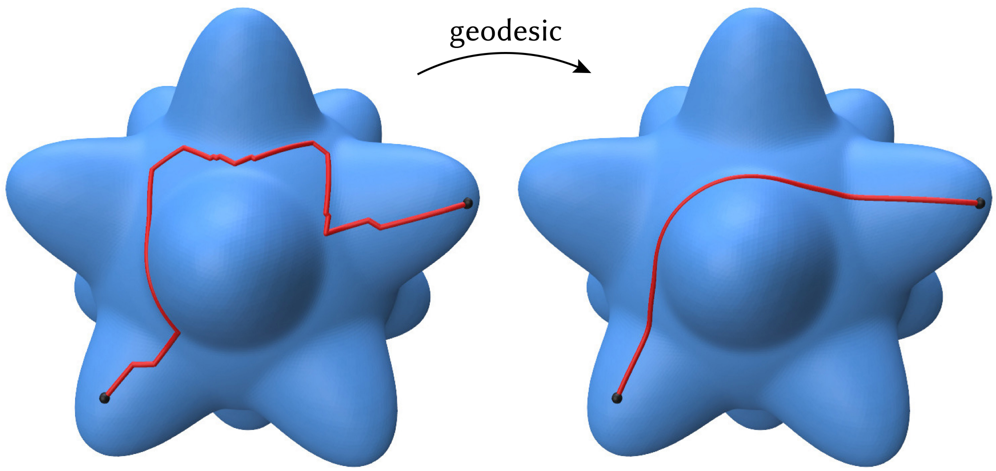
In order to study how to minimize this distance we have to pick a curve, any curve, and parameterize it.

Consider the curve given by $\gamma(s) = X^i(s)$ where now we use s as a measure of "how far along on the curve we are". We will define s such that it varies from 0 to 1 such that at $s = 1$ we have traveled the entire length L of the curve.

That said, then we must have

$$L = \int \sqrt{g_{ij}(X(s)) \frac{dX^i}{ds} \frac{dX^j}{ds}} ds$$

$g_{ij}(X(s))$ means that we ONLY evaluate the metric AT our curve $X^i(s)$. That is to say, the metric is defined EVERYWHERE on our manifold and in order to compute the length of a curve we only need the metric at places the curve passes through. For example, if you had some weird curved geometry but there was one place that was flat and you wanted your curve to be within the flat area then you don't even need to bother with the metric anywhere else - the curve of minimum distance there would be a straight line despite the entire manifold being curved! It simply matter WHERE $\gamma(s)$ is at!



So here's the plan:

1. suppose $\gamma(s)$ is the curve of minimal length

2. then if we change our $\gamma(s)$ for all s by a tiny bit $\delta\gamma(s)$ at each point then if $gamma(s)$ is the minimal curve then L will surely increase

3. therefore, when we vary $\gamma(s)$, while keeping its end points fixed, we demand that δL goes to zero at first order

4. We compute $\delta L = 0$ and this will produce a differential equation for our minimal $\gamma(s)$

All of this is in the exact same spirit as finding the minimum of some function. Given a function $f(x)$ it's minimum is found by setting $f'(x) = 0$ and solving for x. Same concept just a little bit complicated.

Now there's no way around it; we must vary

$$L' = L + \delta L = L \implies \delta L = 0$$

$$\delta L = \int ds\, \delta \sqrt{g_{ij}(X(s)) \frac{dX^i}{ds} \frac{dX^j}{ds}} = 0$$

Therefore we vary our curve by some small functional amount

$$X^i(s) \to X^i(s) + \delta X^i(s)$$

such that

$$\delta X^i(0) = \delta X^i(1) = 0$$

then (we will suppress the dependence on s for clarity)

$$\delta L = \int ds \sqrt{g_{ij}(X + \delta X)\frac{d(X^i + \delta X^i)}{ds}\frac{(X^j + \delta X^j)}{ds}}$$

$$-\sqrt{g_{ij}(X)\frac{dX^i}{ds}\frac{dX^j}{ds}} = L' - L = 0$$

Next, you will expand EVERYTHING out to first order following Taylor as an exercise and then gather up like terms and see many cancellations. You will then integrate by parts making use of the fact that the boundary term is ZERO since $\delta X^i(0) = \delta X^i(1) = 0$.

Next you will come to a piece that looks like

$$\frac{d\delta X^j}{ds}$$

Try to prove that

$$\frac{d\delta X^j}{ds} = \frac{\delta dX^j}{ds} = \delta\frac{dX^j}{ds}$$

that is to say δ and d commute. From here you will have the integral of an integrand equal 0 and argue that the only way that can happen is if the integrand is zero. Then you're done.

All of this proceeds exactly like when we varied the action and obtained the Euler-Lagrange equation

$$S = \int L(x, \dot{x})dt$$

and

$$\delta S = 0 \implies \frac{d}{ds}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$$

In our case the role of the Lagrangian is played by

$$L(x, \dot{x}) = \sqrt{g_{ij}(X(s))\dot{X}^i\dot{X}^j}$$

We're welcome to minimize L^2 rather than simply L for convenience: the results will be the same. Therefore, chug along

then

$$\frac{\partial L^2}{\partial \dot{x}^i} = \frac{1}{2} g_{ij}\dot{X}^j$$

$$\frac{\partial L^2}{\partial x^i} = g_{\alpha\beta,i}\dot{X}^\alpha \dot{X}^\beta$$

Then our Euler-Lagrange equation is

$$\frac{d}{ds}\frac{\partial L^2}{\partial \dot{x}^i} - \frac{\partial L^2}{\partial x^i} = 0$$

$$\frac{d}{ds}\frac{1}{2} g_{ij}\dot{X}^j - g_{\alpha\beta,i}\dot{X}^\alpha \dot{X}^\beta = 0$$

Now we need to apply the s derivative remembering that $g_{ij}(X(s))$. We then have

$$\frac{d}{ds}\frac{1}{2} g_{ij}\dot{X}^i = \frac{1}{2} g_{ij}\ddot{X}^i + \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\frac{dX^k}{ds}\dot{X}^j$$

then, all together, we have

$$\frac{1}{2} g_{ij}\ddot{X}^i + \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{X}^k\dot{X}^j - g_{\alpha\beta,i}\dot{X}^\alpha \dot{X}^\beta = 0$$

This equation is just dandy: its solution solves our problem. However, we can clean it up a bit by doing the following:

We contract with the inverse metric $g^{\mu j}$ and make use of our identity connecting derivatives of g_{ij} to Γ and so we finally arrive at the differential equation which yields the curve $X^\mu(s)$ of minimal length between two points, the geodesic, as

$$\frac{d^2 X^\mu}{ds^2} + \Gamma^\mu_{ij}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

This is called the geodesic equation and curves of minimal length are solutions of this equation. Given a metric for a manifold $\mathcal{M}$ and a parameterization you can solve this second order differential equation. It may require fancy methods for fancy manifolds but at the very least numerical methods can approximate its solutions.

Now it should DEFINIITELY solve the shortest distance between two points in the (x, y) - if it doesn't then we've certainly screwed up!

For a flat Euclidean plane all Γ's are zero so then the equation we must solve is

$$\frac{d^2X^\mu}{ds^2} = 0$$

Come on, that's too easy. The solution is a curve $Y = as + b$ and $X = cs + d$ that you're welcome to parameterize any way you'd like or you could combine them into $Y = mX + b$ as is more standard.

What about a sphere? What are the geodesics? This is a VERY practical matter for airlines and shipping companies! The differential equation will be much more complicated but we should be able to do it. Someone has to, right?

9.1 Back to the 2-Sphere

Let's figure out what we know already - extremal curves on the sphere are great circles. We must solve

$$\frac{d^2X^\mu}{ds^2} + \Gamma^\mu_{ij}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

For $X^\mu(s)$.

Let's do this for our 2-sphere where we've found both the metric and all the Γ's. We have

$$g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}$$

$$\Gamma^\theta_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & -\sin\theta\cos\theta \end{bmatrix}$$

$$\Gamma^\phi_{ij} = \begin{bmatrix} 0 & \cot\theta \\ \cot\theta & 0 \end{bmatrix}$$

So, we now write our differential equation for the θ part of X and the ϕ part as

$$\mathbf{X}(s) = \begin{bmatrix} X^\theta(s) = \theta(s) \\ X^\phi(s) = \phi(s) \end{bmatrix}$$

Therefore we have two differential equations to solve

$$\frac{d^2\theta}{ds^2} + \Gamma^\theta_{ij}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

$$\frac{d^2\phi}{ds^2} + \Gamma^\phi_{ij}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

So, we simply plug in what we know for each equation and hope for the best. We have

$$\frac{d^2\theta}{ds^2} + \begin{bmatrix} 0 & 0 \\ 0 & -\sin\theta\cos\theta \end{bmatrix}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

$$\frac{d^2\phi}{ds^2} + \begin{bmatrix} 0 & \cot\theta \\ \cot\theta & 0 \end{bmatrix}\frac{dX^i}{ds}\frac{dX^j}{ds} = 0$$

In general this could produce quite complicated equations indeed but since the 2-sphere has so many zero terms in Γ then these equations reduce nicely to the pair

$$\frac{d^2\theta}{ds^2} - \sin\theta\cos\theta\frac{d\phi}{ds}\frac{d\phi}{ds} = 0$$

$$\frac{d^2\phi}{ds^2} + 2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds} = 0$$

These two equations look particularly difficult but you might notice that the second one can be re-written as

$$\frac{d}{ds}\left(\sin^2\theta\frac{d\phi}{ds}\right) = \frac{d^2\phi}{ds^2} + 2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds} = 0$$

I wouldn't have noticed that! Some folks can though, I hope you're one of them but if not you could apply the usual differential equation methods to solve the coupled pair.

So then this clever observation implies that

$$\sin^2\theta\frac{d\phi}{ds} = \text{constant} = \alpha$$

so that

$$\frac{d\phi}{ds} = \frac{\alpha}{\sin^2\theta}$$

or

$$\phi(s) = \int \frac{\alpha}{\sin^2\theta(s)} ds$$

Not much else to do with this guy unless we can find $\theta(s)$
Now, we've figured out that

$$\frac{d\phi}{ds} = \frac{\alpha}{\sin^2\theta}$$

So then we can plug this into our $\theta(s)$ equation

$$\frac{d^2\theta}{ds^2} - \sin\theta\cos\theta\frac{\alpha}{\sin^2\theta}\frac{\alpha}{\sin^2\theta} = 0$$

$$\frac{d^2\theta}{ds^2} - \cot\theta\frac{\alpha^2}{\sin^2\theta} = 0$$

This $\theta(s)$ equation is MUCH more difficult. The issue is that two arbitrary points may have different (θ, ϕ) coordinates and the resulting integrals we must solve are generally non-analytic elliptical integrals! Good luck! Despite this complexity we KNOW the geodesics are simple great circles, so what gives?

We can bypass all this elliptical mumbo-jumbo by leveraging the symmetry of the sphere. Don't like that two points have different (θ, ϕ) values? No problem, just rotate your coordinates such that both points lie on the same meridian! Ain't symmetry great?!

What does this mean for our differential equations? Well, upon judicious choice of coordinates now our $\phi(s)$ is constant so that $d\phi = 0$!

This means our $\theta(s)$ equations goes from

$$\frac{d^2\theta}{ds^2} - \sin\theta\cos\theta\frac{d\phi}{ds}\frac{d\phi}{ds} = 0$$

to

$$\frac{d^2\theta}{ds^2} = 0$$

and lo and behold we get our great circles! If you're really fussy about your coordinates then you can always numerically integrate and bleed and sweat but I think you agree that symmetry should always be leveraged where possible lest we get nowhere!

Additionally, notice we have a "hidden" relationship: given that

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2$$

we can ALWAYS choose to parameterize our curve in terms of length. If we choose units such that the total length equals 1 then we manifestly have

$$\left(\frac{ds}{ds}\right)^2 = \left(\frac{d\theta}{ds}\right)^2 + \sin^2\theta\left(\frac{d\phi}{ds}\right)^2 = 1$$

Therefore we have three differential equations we can leverage when solving for these $2D$ spherical geodesics. Any particular curve should be parameterization independent but choosing a parameter given by the line element itself is quite judicious. This acts like a conservation law: if we chose time as a parameter then we would find the energy would be conserved (modulo closed system).

What about a 3-sphere S^3? Can you figure out the metric and then find its geodesics?

9.1.1 Covariant Derivative

We could have also defined geodesic curves to be such that the covariant derivative of a vector defined at some point is equal to zero everywhere along the curve as it is parallel transported. The geodesic is such that under parallel transport the vector is always unchanged with respect to the tangent vector of the geodesic.

This assumption would also yield the geodesic equation as follows:

Given an arbitrary curve γ defined by $X^\mu(\lambda)$ (where λ parameterizes the curve) and given a single vector at some point λ_0 on the curve $Q^\mu(\lambda_0)$ then we may determine the curve by demanding that the covariant derivative with respect to the curve is always zero along the curve.

Given a single vector $Q^\mu(\lambda_0)$ we can determine it all all other points along the curve via parallel transport.

To do this operation we must first figure out how to covariantly differentiate with respect to a curve. This is a MUCH different thing than the derivative with respect to coordinates we've developed so far. Recall that we found

$$D_\nu W^\mu(x) = \frac{DW^\mu}{Dx^\nu}$$

but here we hope for something like

$$\frac{DW^\mu(\lambda)}{D\lambda}$$

This is a MUCH different beast. Let's reason our way through it.

Naively, in flatland, we would usually just say "easy, it's just $\frac{dW^\mu(\lambda)}{d\lambda}$" a garden variety derivative and in flat space we would be totally right but here in general curvy land we expect something like Γ or the metric to show up in some potentially complicated fashion.

The biggest insight that allows us to proceed, which we could've leveraged in the past is to say "we hope $\frac{dW^\mu(\lambda)}{d\lambda}$ is a vector".

It has an upper index after-all! Now, we know something is a vector if it transforms in the linear fashion we demand vectors to transform under coordinate changes. If we get something non-linear then we no longer have a vector.

So we ask "is $\frac{dW^\mu(\lambda)}{d\lambda}$ a vector under a change in coordinates?"

First we know W^μ is a vector because

$$W'^\mu(\lambda) = \frac{\partial x'^\mu}{\partial x^i} W^i(\lambda) = S^\mu_i W^i(\lambda)$$

But recall, most generally S^μ_i is a function of position on the manifold and here we should be evaluating it ON our curve γ given by $X(\lambda)$ so that we have

$$W'^\mu(\lambda) = S^\mu_i(X(\lambda)) W^i(\lambda)$$

Now, if we happen to discover that

$$\frac{dW'^\mu(\lambda)}{d\lambda} = S^\mu_i \frac{dW^i(\lambda)}{d\lambda}$$

then we're done. That would mean that this naive derivative is a vector under coordinate changes which is our new definition of what it means to be a vector!

So, let's take the λ derivative of our expression for $W'^\mu(\lambda)$

$$\frac{d}{d\lambda} W'^\mu(\lambda) = \frac{d}{d\lambda}[S^\mu_i(X(\lambda)) W^i(\lambda)]$$

A-ha! The transformation S depends on λ so we have to invoke the product rule

$$\frac{d}{d\lambda} W'^\mu(\lambda) = \frac{dS^\mu_i(X(\lambda))}{d\lambda} W^i(\lambda) + S^\mu_i(X(\lambda)) \frac{dW^i(\lambda)}{d\lambda}$$

How do we evaluate this? Well the derivative on S is a chain rule, viz

$$\frac{dS^\mu_i(X(\lambda))}{d\lambda} = \frac{\partial S^\mu_i(X(\lambda))}{\partial X^\nu} \frac{dX^\nu}{d\lambda}$$

but

$$\frac{dX^\nu}{d\lambda} = V^\nu(\lambda)$$

is the velocity or tangent vector along the curve. Then all together we have

$$\frac{d}{d\lambda} W'^\mu(\lambda) = \frac{\partial S^\mu_i(X(\lambda))}{\partial X^\nu} \frac{dX^\nu}{d\lambda} W^i(\lambda) + S^\mu_i(X(\lambda)) \frac{dW^i(\lambda)}{d\lambda}$$

This is clearly NOT a vector. $\frac{\partial S_i^\mu(X(\lambda))}{\partial X^\nu}$ spoils the transformation.

Well, we've seen stuff like this before: instead of the ordinary derivative, perform the covariant derivative. In that case everything will transform just fine

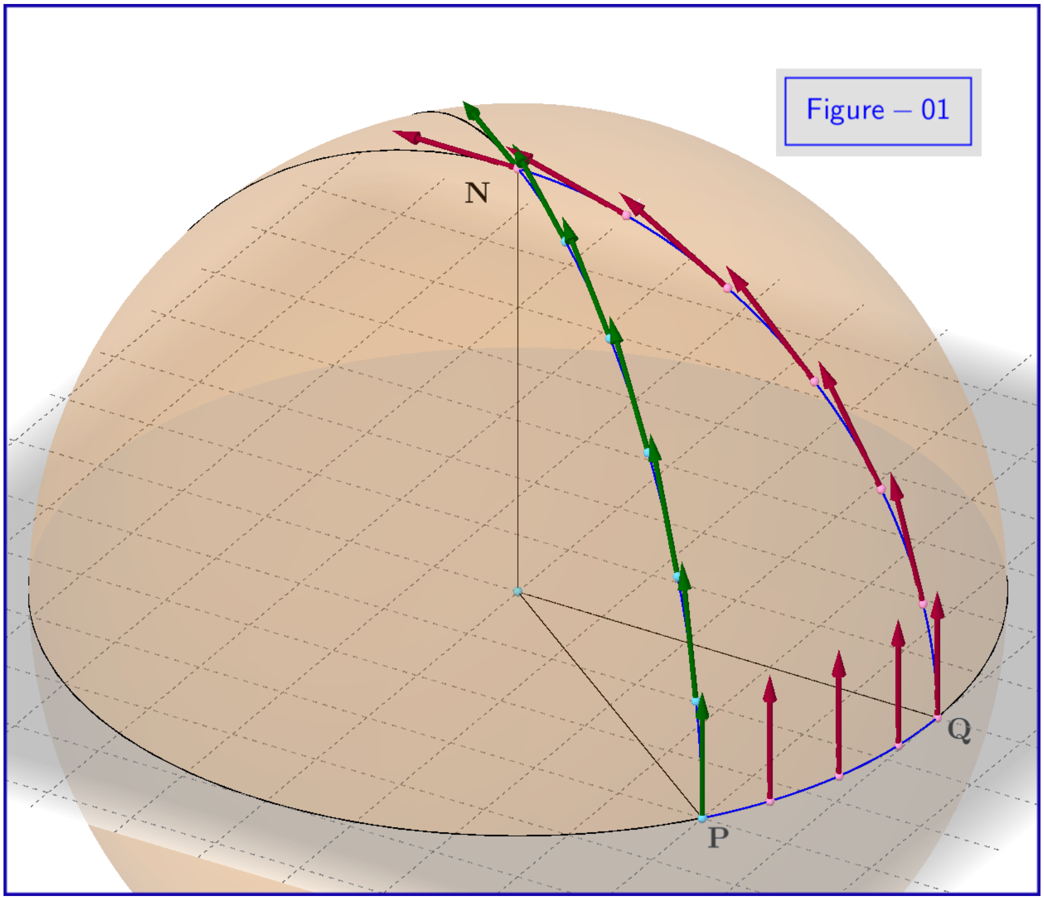
$$\frac{DW^\mu(\lambda)}{D\lambda} = \frac{dW^\mu(\lambda)}{d\lambda} + \Gamma^\mu_{\rho\nu}(X(\lambda))\frac{dX^\rho}{d\lambda}W^\nu(x)$$

This is often written as

$$\frac{DW^\mu(\lambda)}{D\lambda} = \frac{dW^\mu(\lambda)}{d\lambda} + \Gamma^\mu_{\rho\nu}(X(\lambda))V^\rho(\lambda)W^\nu(\lambda)$$

Great - we've found a covariant derivative with respect to a curve.

Going back to our motivation we claimed a geodesic is the unique curve such that if you parallel transport a vector it will not change one bit. The geodesic is the only type of curve which accomplishes this. So we define a single vector at λ_0 and transport it along a candidate curve and if we find that it changes then we throw away that curve and try again.



Eventually we will find the unique curve between two points on our manifold such that parallel transporting our candidate vector is unchanging for all λ. Try it with a vector from the north pole to the equator along a great circle. It doesn't change!

Therefore, we wish to solve the differential equation

$$\frac{DW^\mu(\lambda)}{D\lambda} = 0$$

$$= \frac{dW^\mu(\lambda)}{d\lambda} + \Gamma^\mu_{\rho\nu}(X(\lambda))V^\rho W^\nu(\lambda) = 0$$

or, briefly, (without expressing the fact that we only evaluate the metric on the curve)

$$\frac{DW^\mu}{D\lambda} = \frac{dW^\mu}{d\lambda} + \Gamma^\mu_{\rho\nu} V^\rho W^\nu = 0$$

notice that

$$\frac{dW^\mu}{d\lambda} = \frac{dX^\rho}{d\lambda}\frac{\partial W^\mu}{\partial X^\rho}$$

$$\frac{DW^\mu}{D\lambda} = \frac{dX^\rho}{d\lambda}\partial_\rho W^\mu + \Gamma^\mu_{\rho\nu} V^\rho W^\nu = 0$$

Notice how similar this appears to our covariant derivative
but again

$$\frac{dX^\rho}{d\lambda} = V^\rho$$

so that

$$\frac{DW^\mu}{D\lambda} = V^\rho \partial_\rho W^\mu + \Gamma^\mu_{\rho\nu} V^\rho W^\nu = 0$$

now compare with our fully covariant derivative

$$D_\rho W^\mu(x) = \partial_\rho W^\mu + \Gamma^\mu_{\rho\nu} W^\nu(x)$$

The only thing distinguishing the two is the tangent vector to the curve V^ρ!

therefore we find that

$$\frac{DW^\mu}{D\lambda} = V^\rho D_\rho W^\mu$$

Therefore a geodesic MUST satisfy

$$V^\rho(\lambda) D_\rho W^\mu(\lambda) = 0$$

This equation will produce a differential equation you can solve for the unique curve $\gamma(\lambda)$. We have two different tools now: either are just dandy. Sometimes one of the methods is easier than the other and vice versa.

10 Covariant Operators

Now that we've defined the covariant derivative perhaps we can use it to construct our favorite operators from vector calculus: divergence, gradient, curl, etc.

Our covariant derivative is the operator that fundamentally depends on its underlying manifold such that

$$D_\nu = \partial_\nu + \Gamma^\mu_{\rho\nu}$$

where it acts on anything we please. If we apply this to a vector then

$$D_\nu W^\mu(x) = W^\mu_{,\nu} + \Gamma^\mu_{\rho\nu} W^\rho(x) = W^\mu_{;\nu}(x)$$

This covariant derivative of a vector manifestly produces a tensor M^μ_ν: prove that this transforms exactly as a tensor under coordinate transformation.

Let's go through our list of geometric objects and the operations we can do on them.

10.1 Scalars

A scalar field on our manifold $\mathcal{M}$ is a number defined at all points given by the coordinate transformation law

$$\phi'(x') = \phi(x)$$

Applying D^μ we have

$$D_\nu \phi(x) = \partial_\nu \phi(x)$$

In the case of a scalar function there is no sense in which the values of $\phi(x)$ have "direction" and therefore when we try and contract a Γ with it we simply cant - there is no index to match up!

Therefore $D_\nu \phi(x)$ is our good old fashioned gradient of a function $\nabla\phi(x)$

10.2 Vectors

We've seen this already: we simply have

$$D_\nu W^\mu(x) = W^\mu_{\ ,\nu}(x) + \Gamma^\mu_{\rho\nu} W^\rho(x) = W^\mu_{\ ;\nu}(x)$$

But now we have options. We could match the μ and ν indices and get the covariant divergence

$$D_\mu W^\mu(x) = W^\mu_{\ ,\mu} + \Gamma^\mu_{\rho\mu} W^\rho(x) = W^\mu_{\ ;\mu}(x)$$

In this case we often write

$$\nabla_\mu W^\mu = W^\mu_{\ ;\mu}$$

Indeed, often you will see ∇ written instead of D - if a "flat" divergence in needed then it will be written $\partial_\mu W^\mu$

$$\nabla_\mu W^\mu = \partial_\mu W^\mu + \Gamma^\mu_{\rho\mu} W^\rho$$

Incidentally show that this can conveniently be written (due to the Γ's identities) as

$$\nabla_\mu W^\mu = \partial_\mu W^\mu + \left(\frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} \right) W^\mu$$

10.3 Co-vectors

We can also find covariant derivatives of co-vectors in a similar manner. Pursuing the same ideas we find

$$D_\nu U_\mu = U_{\mu,\nu} - \Gamma^\rho_{\mu\nu} U_\rho = U_{\mu;\nu}$$

Similarly you can find the divergence in the same manner. Note the crucial minus sign and the arrangement of the indices!

10.4 The Product Rule

We can find covariant derivatives of products in the same way as our flat calculus

First, we know that the dot product of two vectors is a scalar and we know how covariant derivatives act on scalar fields. Let's study the covariant derivative of a dot product of vectors.

$$D_\nu(U_\mu W^\mu) = (D_\nu U_\mu)W^\mu + U_\mu(D_\nu W^\mu)$$

$$= (U_{\mu,\nu} - \Gamma^\rho_{\mu\nu}U_\rho)W^\mu + U_\mu(W^\mu_{\ ,\nu} + \Gamma^\mu_{\rho\nu}W^\rho)$$

Notice the two Γ terms, we have

$$-\Gamma^\rho_{\mu\nu}U_\rho W^\mu$$

and

$$U_\mu\Gamma^\mu_{\rho\nu}W^\rho$$

since μ and ρ are being summed over we are allowed to rename them as we please. In the first term turn all the μ's into ρ's and vice versa. Then we have

$$-\Gamma^\mu_{\rho\nu}U_\mu W^\rho$$

This is identical to the second Γ term and hence they annihilate! We have found then

$$D_\nu(U_\mu W^\mu) = (U_\mu W^\mu)_{,\nu}$$

which is exactly what we expect for a scalar function!

Now, for an arbitrary product between vector and co-vector we have the product rule

$$D_\lambda(U_\nu W^\mu) = (D_\lambda U_\nu)W^\mu + U_\nu(D_\lambda W^\mu)$$

$$= (U_{\nu,\lambda} - \Gamma^{\rho}_{\nu\lambda} U_{\rho}) W^{\mu} + U_{\mu} (W^{\mu}_{\ ,\lambda} + \Gamma^{\mu}_{\rho\lambda} W^{\rho})$$

where now there is no particular cancellations. The main thing to appreciated is that the product rule holds in this more general environment.

10.5 Tensors

What we've done so far hints at how we can covariantly differentiate tensors. We follow the rule of having a minus sign against Γ for each lower index and a plus sign for each upper index. A tensor should "behave" like a product of vectors and co-vectors.

We have, generally

$$T^{\alpha\beta}_{\gamma;i} = T^{\alpha\beta}_{\gamma,i} + \Gamma^{\alpha}_{ij} T^{j\beta}_{\gamma} + \Gamma^{\beta}_{ij} T^{\alpha j}_{\gamma} - \Gamma^{k}_{\gamma i} T^{\alpha\beta}_{k}$$

Yes, keeping track of indices is a nightmare we call "index gymnastics"! The pattern follows for arbitrary tensors.

The most important tensor you've seen is the metric itself.

What is $\nabla_k g_{ij} = g_{ij;k}$

$$\nabla_k g_{ij} = g_{ij,k} - \Gamma^{m}_{jk} g_{im} - \Gamma^{m}_{ik} g_{mj} = 0$$

Using the definition and identities of Γ show that the metric vanishes upon covariant differentiation.

This will be massively important later.

10.6 Commutators

Given two vector fields we may study the commutator as an operation in the same spirit as we do in quantum theory.

If

$$U(x) = U^\nu(x)\partial_\nu = U^\nu(x)e_\nu$$

and

$$V(x) = V^\alpha(x)\partial_\alpha = V^\alpha(x)e_\alpha$$

Then

$$W^\mu = [U, V]^\mu = (UV - VU)^\mu$$

$$= U^\nu\partial_\nu V^\alpha\partial_\alpha - V^\alpha\partial_\alpha U^\nu\partial_\nu$$

but since $\partial_\nu\partial_\alpha = \partial_\alpha\partial_\nu$ and most importantly the fact that we may rename summed over indices we can "factor out" one of the partials

$$W(x) = W^\mu\partial_\mu = [U^\mu(\partial_\mu V^\nu) - V^\mu(\partial_\mu U^\nu)]\partial_\mu$$

or

$$W^\mu = U^\mu(\partial_\mu V^\nu) - V^\mu(\partial_\mu U^\nu)$$

but given the symmetry of the Γ's show that we also have

$$W^\mu = U^\mu(D_\mu V^\nu) - V^\mu(D_\mu U^\nu)$$

We learn that on an arbitrary manifold the order of vectors matters and there exists a function

$$[\ \ ,\ \ ] : T_xM \times T_xM \to T_xM$$

This function is called a "Lie Bracket" and behaves sort of like addition or multiplication of the real numbers. For example

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

We will see later that this operation is fundamentally connected to curvature and "flow" of vector fields (defined below) which we have yet to define. For now think of it as just another operation we can use on vectors and co-vectors.

10.7 Differentiation Along a Curve

We previously figured out how to covariantly differentiate along a curve parameterized as $X(t)$ as

$$\frac{DW^\mu}{Dt} = \frac{dW^\mu}{dt} + \Gamma^\mu_{\rho\nu} V^\rho W^\nu$$

where

$$V^\rho = \frac{dX^\rho}{dt}$$

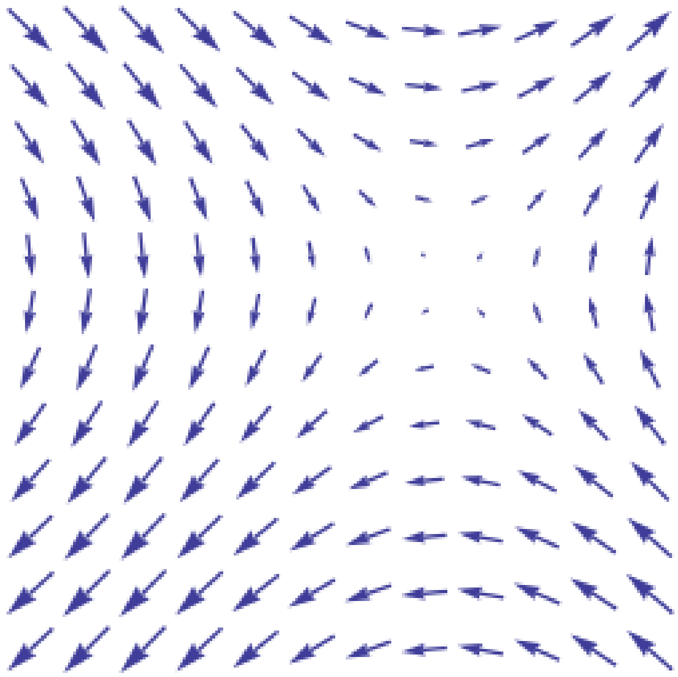
is the velocity of the curve/trajectory.

10.8 Along a Vector Field - Lie Derivative

10.8.1 Hamilton

Let's take a step back away from this generalized geometry and touch base with something we've seen before in order to prepare our minds for the generalization to come.

For visualization consider the (x, y) plane $\mathbb{R}^2$. Suppose we have a static vector field $\vec{V}(x, y)$ defined on our plane. For example,



Our vector field is a vector valued function of the coordinates as

$$\vec{V}(x,y) = \begin{bmatrix} V_x(x,y) \\ V_y(x,y) \end{bmatrix}$$

Now, given a vector field there exists a special curve $\gamma(t)$ called an "integral curve" of $\vec{V}$. For example,

$$\gamma(t) = \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}$$

is a curve parameterized by the variable t. So far there's nothing special about this curve but if we consider that this is a special curve that has $\vec{V}$ as its tangent then it becomes special indeed.

We say $\gamma(t)$ is an integral curve of $\vec{V}$ if it solves the following differential equations:

$$\frac{dX(t)}{dt} = V_x(x,y)$$

$$\frac{dY(t)}{dt} = V_y(x,y)$$

This is to say there is a special curve "that goes with the flow" so to speak.

Given this we see how to generalize to arbitrary dimensions but it's not so clear how to go about arbitrary curved manifolds.

Briefly we will write that an integral curve $\gamma(t)$ satisfies

$$\frac{d\gamma(t)}{dt} = V(\gamma(t))$$

where $V(\gamma(t))$ means you take the field value at that point on the curve.

10.8.2 Example

We're very familiar with the classical harmonic oscillator. Let's study the integral curves of the SHO.

First,

$$H(q,p) = \frac{1}{2}(p^2 + q^2)$$

Next, we have Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p} = p$$

$$\dot{p} = -\frac{\partial H}{\partial q} = -q$$

Comparing to our formalism above we notice that Hamilton's equations define a vector field on the (q,p) plane.

$$\vec{V}(p,q) = \begin{bmatrix} p \\ -q \end{bmatrix}$$

our goal is to solve the differential equations to determine the integral curves of this vector field.

In our notation above we solve the coupled differential equations given by

$$\frac{dq(t)}{dt} = p$$

$$\frac{dp(t)}{dt} = -q$$

Easy peasy- we have

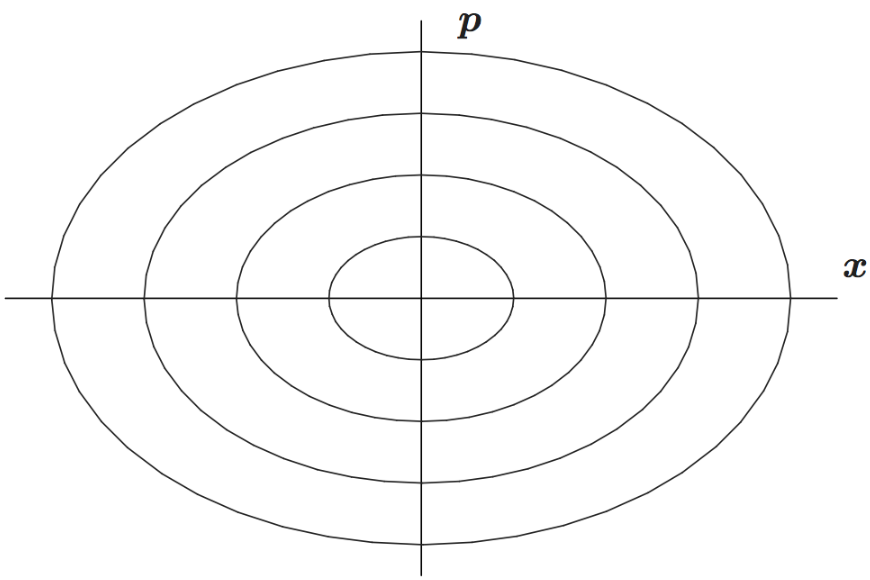
$$q(t) = q_0 \cos t + p_0 \sin t$$

$$p(t) = p_0 \cos t - q_0 \sin t$$

Therefore, we say our integral curve to the given vector field

$$\vec{V}(p,q) = \begin{bmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{bmatrix}$$

$$\gamma(t) = \begin{bmatrix} q_0 \cos t + p_0 \sin t \\ -q_0 \sin t + p_0 \cos t \end{bmatrix}$$



In fact, this was all we were ever doing when studying Hamiltonian mechanics. Given a Hamiltonian we could "generate" the vector field above and then solve for the integral curves.

Notice what $\gamma(t)$ does. It says "if you tell me your initial position in (q,p) space then I will tell you where it goes".

Notice what's staring at us in the face: that $\gamma(t)$ "looks like" we could factor out a rotation matrix. I.e.

$$\gamma(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} q_0 \\ p_0 \end{bmatrix}$$

This is true for the initial point (q_0, p_0). However, isn't it the case that we can choose ANY point in the (q, p) plane to be our initial point?

Then we find that if we're given some point it will GLOBALLY "flow" to another point by

$$\begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} q(t) \\ p(t) \end{bmatrix}$$

This is a subtly different concept than an integral curve. This is actually a "family" of maps that evolves ALL points (q, p). We call this family of maps a 1-parameter group where here t is the parameters. Thus we write

$$\phi_t(q, p) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} q(t) \\ p(t) \end{bmatrix}$$

where

$$\phi_t = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

This describes a rotation but with time as the parameter. We know all about rotations. For example, it's true that rotating by 5 degrees and then 10 degrees is the same as rotating by 15. Practical stuff. so this "flow" has the property that

$$\phi_s \circ \phi_t = \phi_{s+t}$$

and

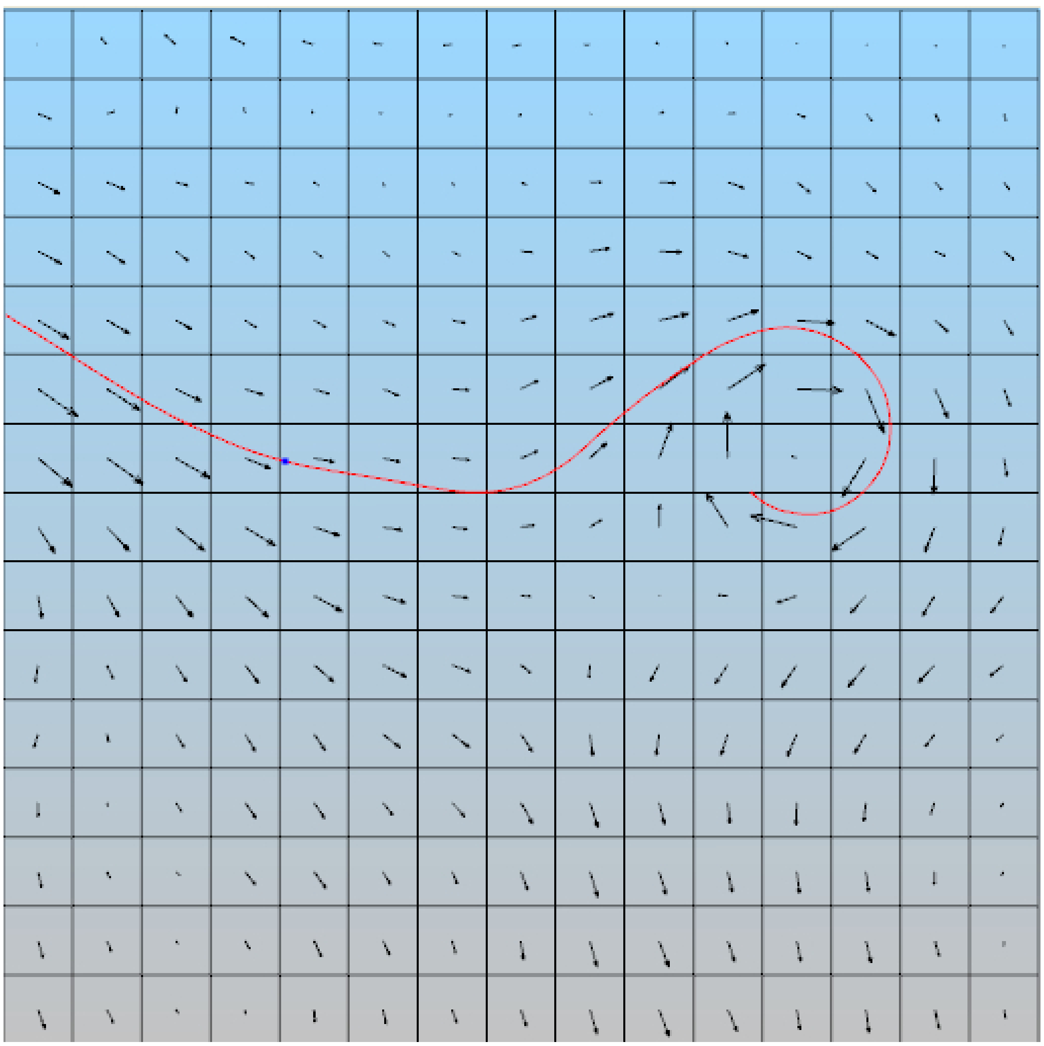
$$\phi_{-t} \circ \phi_t = \phi_0 = \mathrm{id}$$

therefore

$$\phi_{-t} = \phi_t^{-1}$$

Appreciate the nuance of what we've done. We first found AN integral curve. This was a readily solvable coupled differential equation. This was just a single curve that started at (q_0, p_0). In general there is absolutely zero reason to expect a nearby curve to have the same solution: that depends on the entire vector field.

It so happened in our example that the vector field had a rotational symmetry and if we knew that then solving for one trajectory is the same as solving them all. But without that symmetry we'd be stuck computing integral curves for every point in the plane.



The ultimate goal in solving a problem is to find the symmetry: then everything goes swimmingly.

So, once we solved for an integral curve a symmetry was screaming in our face. The rotation matrix was right there and evident. It was the underlying symmetry - rotations in "time". Not exactly a familiar concept but certainly plausible given that we know the harmonic oscillator is connected to circular motion in time!

The integral curves are LOCAL properties of the vector field
The "flow" is a GLOBAL property of the entire system.

Now, you should be recalling what we did with rotations in quantum mechanics. We followed Lie and said there is a "generator" of $SO(2)$ such that

$$\phi_t = e^{tX_H}$$

where X_H is the infinitesimal generator of rotation

$$X_H = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Notice that

$$X_H = \left.\frac{d\phi_t(x)}{dt}\right|_{t=0} = \frac{d}{dt}\left.\left[\begin{array}{cc} \cos t & \sin t \\ -\sin t & \cos t \end{array}\right]\right|_{t=0}$$

Which says, if you know ϕ_t then you also know your vector field and vice versa.

What's more is that this flow defines how ANY function evolves over time

$$f(q(t), p(t)) = f(\phi_t(q_0, p_0))$$

Then

$$\frac{df}{dt} = \{f, H\}$$

where the braces are what's called the Poisson bracket defined as

$$\{f, g\} \ := \ \sum_{i=1}^{N}\left[\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right].$$

It is tantalizingly closely related with the quantum commutator you've seen before!

If the system is evolving in time then

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \tag{1}$$

Furthermore we have that X_H itself is what generates the flow - it is not only a matrix but it is also the vector field given by

$$X_H = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}$$

$$X_H = \vec{V}(p,q) = \begin{bmatrix} \frac{\partial H}{\partial p} \\ -\frac{\partial H}{\partial q} \end{bmatrix} = \begin{bmatrix} p \\ -q \end{bmatrix}$$

We're beginning to see intimations of some underlying geometric manifold structure of Hamiltonian mechanics itself! It is known as a "symplectic manifold". Everything we've done here, just by following our noses, can be generalized for arbitrary manifolds.

Notice that for an oscillator (x, y, z) space is flat and Euclidean. But the physics of an oscillator itself has some more abstract notion of geometry in (q, p) coordinates. It's similar to how we discovered all planets move in perfect circles only (but in velocity space!).

Hamiltonian mechanics has an underlying geometry given by a specific type of manifold- then the rest is studying particular vector, scalar, and tensor fields on this "symplectic" manifold.

Here, the generator of our fundamental, global symmetry group ϕ_t can act on either individual points (q, p) or functions. When acting on points it is a matrix and when acting on functions it is a vector field

$$X_H = p\frac{\partial}{\partial q} - q\frac{\partial}{\partial p}$$

Which has components $X_H^q = p$ $X_H^p = q$ and basis vectors given by

$$\frac{\partial}{\partial q} = e_q$$

$$\frac{\partial}{\partial p} = e_p$$

I intentionally left the minus sign there. that minus sign comprises the "symplectic structure" of the manifold. Given some technical details we haven't covered yet we will refrain from defining this type of manifold. However, it's looking more and more like our general considerations of arbitrary manifold machinery!

Finally, in general we LOVE to study "flow" because

1. it describes ALL trajectories
2. it encodes global transformations
3. it's a group
4. it reveals a fundamental symmetry of the vector field
5. it's deeply connect to Lie groups and generators
6. it governs global observables

Integral curves, however important, do not give us any leverage into global aspects of our manifold: they are individual trajectories. Solving for one of them might suggest the flow and generators (like with what happened for the SHO) - but in general we are not so fortunate.

Now, let's define "flow" more generally

10.9 Flow

First, let us define a notion of the "flow" of a vector field on a manifold $\mathcal{M}$. The easiest way to study this is to imagine a stream of water flowing along a riverbed (our manifold).

We can define a velocity vector field for all points on the riverbed. Connecting all the vectors results in the introduction of "streamlines" or integral curves. If you place an object on a particular streamline then it will "flow" along that particular streamline. What's more is that streamlines cannot cross since that would result in two different velocities at the same point (which would vectorially add into a single velocity vector anyways!).

Given this germ of an idea we can then define a family of maps that depend on a single parameter we call t where this map takes a "molecule" located at $x = p \in \mathcal{M}$ at $t = 0$ to a new position $p' \in \mathcal{M}$ some time t later.

This one parameter map is

$$\phi_t : \mathcal{M} \times \mathbb{R} \to \mathcal{M}$$

This function takes points in $\mathcal{M}$ and a "time" t and then spits out other points (maybe even the same point) in $\mathcal{M}$. That is

$$\phi_t(p) = p'$$

We expect further that if we move from p to p'' then that should be the same as moving from p to p' and then from p' to p''. So we demand the rule

$$\phi_s(\phi_t(p)) = \phi_s(p') = \phi_{t+s}(p) = \phi_t(\phi_s(p)) = p''$$

Furthermore we can also "undo" the flow by

$$\phi_{-t}(\phi_t(p)) = \phi_{-t}(p') = p$$

therefore we have

$$\phi_{-t} = \phi_t^{-1}$$

Given this we say this forms a 1-parameter group of maps. These are maps that transport us along with the flow of the vector field - they are the streamlines all together. Such gory detail for something that seems pretty natural.

Recall we had "streamlines" when we studied phase space in Hamiltonian mechanics. The Hamiltonian itself generated "flow" along a trajectory of constant energy. The same thing is happening here but in more generality.

Now, if we a priori know these maps/group then we may deduce the vector field creating the flow by

$$\vec{V}(x) = \left.\frac{d\phi_t(x)}{dt}\right|_{t=0}$$

The vector field is the velocity of the flow at each point in $\mathcal{M}$

Then

$$\phi_t = e^{tV}$$

If we know the map, then we know the vector field. If we know the vector field, then we know the map. There are special situations where this idea breaks down globally, but for most well behaved fields this is generally true. If you do have trouble then you're always welcome to restrict to a nicer subset of your manifold.

These streamlines are the special curves $\gamma(t)$ (that we also write in terms of coordinates $X^\mu(t)$) such that the vector field is always tangent to the curve.

We have the result that given some vector field there exists a unique 1-parameter family of maps ϕ_t such that the vector field components are given by

$$V^\mu(p) = \frac{dX^\mu(\phi_t(p))}{dt} = \frac{dX^\mu}{dt}$$

Where in the last expression we simplified the notation as $V^\mu(x) = \frac{dX^\mu}{dt}$. This means that $\phi_t(p)$ can be found by solving the set of differential equations

$$V^\mu(x) = \frac{dX^\mu}{dt}$$

such that $x(0) = p$

Now our velocity vector field is given by

$$\vec{V} = V^\mu(x)\frac{\partial}{\partial x^\mu}$$

and upon defining ϕ_t we then have

$$\vec{V} = \frac{dX^\mu}{dt}\frac{\partial}{\partial x^\mu}$$

If we are given some function $f(x)$ then

$$\vec{V}(f) = \frac{dX^\mu}{dt}\frac{\partial f}{\partial x^\mu} = \frac{df(\phi_t(P))}{dt}$$

which we say is the "derivative of f along the stream line through P"

Every vector field on $\mathcal{M}$ defines a flow ϕ_t having $\vec{V}$ as its velocity field (generator).

Now, suppose we have a manifold with coordinates x^μ and a pair of vector fields defined on the manifold $\mathcal{M}$.
That is we have

$$Q^\mu(x) \quad \text{and} \quad W^\mu(x)$$

and each of these fields has its own flow maps ϕ_t^Q and ϕ_s^W

Our goal is to compare W^μ at a point ϕ_t^Q with the value of W^μ at the streamline of Q^μ when $t = 0$. Got that? We follow the streamlines of Q and see how W changes along THOSE streamlines of Q

We are looking for something like

$$\lim_{t \to 0} \frac{W|_{\phi_t^Q} - W|_x}{t}$$

The problem, of course, is that these two vectors live in completely different tangent space. So we need some function that will "push forward" $W(x)$ to the point ϕ_t^Q.

This is entirely analogous to parallel transport only now we are setting this up in a more abstract way which doesn't depend on the metric explicitly.

Define the push-forward as

$$\phi_{t*} : T_p M \to T_{\phi_t(p)} M$$

This particular function takes a vector as input and spits out a vector and should be related to ϕ_t. We shall call it ϕ_{t*}^Q for our vector field $Q(x)$. In general it will take a vector from one tangent space to another along the flow of a vector field. - once it's there we can compare vectors. We can then define a derivative of a vector field against another vector field. We just go with Q's flow

$$\mathcal{L}_Q W = \lim_{t \to 0} \frac{W_{\phi_t^Q} - \phi_t^{*Q}(W_x)}{t}$$

Now we have a bonafide derivative looking thing. This is called the Lie Derivative of W with respect to Q. This form is not particularly handy to use so let's keep going.

Now, recall parallel transport: Suppose we have a curve γ defined by $X^\mu(\lambda)$ where λ parameterizes the curve. This can be any curve through the manifold. Next suppose we have a single vector W^μ defined at a point on the curve γ at $\lambda = \lambda_0$. Then we previously found that we can determine the parallel transport of W^μ along our curve by integrating the differential equation given by

$$\frac{dW^\mu(\lambda)}{d\lambda} + \Gamma^\mu_{\rho\nu}(X^\mu(\lambda)) V^\rho(\lambda) W^\nu(\lambda) = 0$$

Here we stress that all of this should be evaluated ONLY on our curve $X^\mu(\lambda)$. If we keep this in mind we can write the parallel transport equation more tidily as

$$\frac{dW^\mu}{d\lambda} + \Gamma^\mu_{\rho\nu} V^\rho W^\nu = 0$$

Don't forget - in this case W^μ was NOT a field. We defined it to ONLY be at λ_0. If we have a field then we can ask how a particular value at some specific point transports as we follow some path.

Notice also we have the vector

$$V^\rho(\lambda) = \frac{dX^\rho}{d\lambda}$$

This is the velocity or tangent vector of our curve γ.

The solution of this equation allows us to find the parallel transported version of $W^\mu(\lambda_0)$ ANYWHERE on the curve γ. Its solution will look like $W^\mu(\lambda)$ for all λ.

Well, in our abstract case we know the path we wish to follow: it is ϕ_t^X and will have a generating vector field with components $X^\mu(t)$ just as above. Next we know what we wish to transport; that is $W^\mu(p) = W^\mu(x)$. We wish to transport this vector from $x = p$ all the way to ϕ_t^X. Finally we compare this transported vector with $W^\mu(\phi_t^X) = W^\mu(p')$

The velocity vector would be

$$V^\mu = \frac{dX^\mu(\phi_t(P))}{dt}$$

So we solve our parallel transport equation for $W^\mu(t)$ and plug that value into the definition for our Lie Derivative and when all the complicated dust settles we have

$$\mathcal{L}_Q W^\mu(x) = Q^\nu(D_\nu W^\mu) - W^\nu(D_\nu Q^\mu)$$

$$= Q^\nu(\partial_\nu W^\mu) - W^\nu(\partial_\nu Q^\mu)$$

$$= [Q, W]^\mu$$

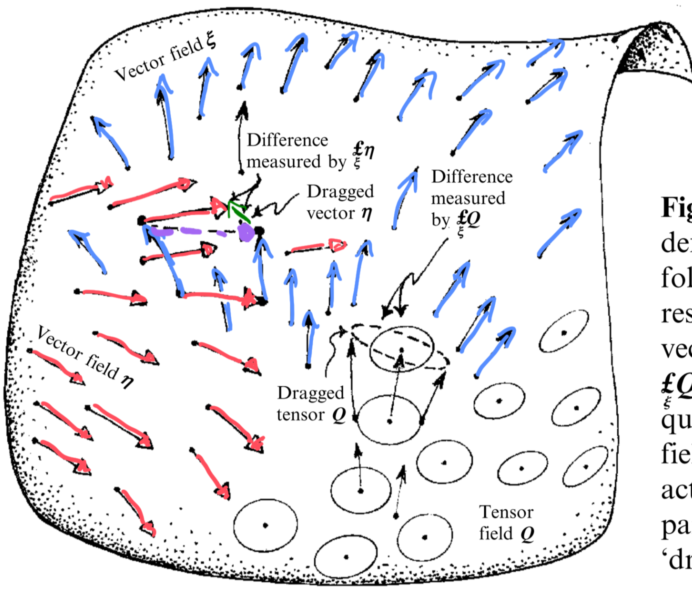


Fig. 14.13 Lie derivative $\underset{\xi}{£}$, defined on a general manifold $\mathcal{M}$, is taken with respect to a given smooth vector field $\boldsymbol{\xi}$ on $\mathcal{M}$. Then $\underset{\xi}{£}\boldsymbol{Q}$ measures how a quantity $\boldsymbol{Q}$ (e.g. a vector field $\boldsymbol{\eta}$ or tensor field $\boldsymbol{Q}$) actually changes, as compared with the quantity 'dragged' by $\boldsymbol{\xi}$.

Interestingly the Lie Derivative is independent of the metric and is expressed as the commutator we know and love!

It seems shocking that here we have two vector fields manifestly living in the tangent spaces of a arbitrary manifold but yet when we differentiate one with respect to the flow of another then all that manifold business disappear. The Lie derivative seems to be more fundamental than the manifold itself. It depends on the smoothness and continuity of the manifold but ultimately it only cares about the relationships established by the flows of the vector fields.

11 Curvature

Recall our ancient experiment where we walk around some path and ask "have we made it home?". For example,

1. walk 100 miles east
2. walk 100 miles north
3. walk 100 miles west

4. walk 100 miles south

On the surface of a sphere, a fundamentally curved manifold, we will not make it back to where we started and we argued that this discrepancy has to do with the curvature of the manifold.

A strategy might be to how the separation between the starting point and the ending point of a path composed of equal parameterized lengths differ and to relate that to the metric some how. This feels cumbersome and ultimately seems to be a path dependent property which for generalization is no good.

Another consideration we made was to start at the north pole with some vector and parallel transport it around a well defined loop $\gamma(s)$ and see how it differs from the vector you started with.

This is a well-defined notion that we can compute and, we hope and expect, it doesn't even depend on the path. We start and end on the same point and so long as our manifold is well defined everywhere it doesn't matter: either the vector is the same or the vector is different. When we did this experiment on our Earth from the north pole to the equator, around by 90° and back to the north pole we found our vector was substantially different - perhaps maximally. This in a sense is path dependent - a shorter path would produce much smaller difference between our vector and its parallel transported brother. However, we have that if there is some curvature then the parallel transported vector will be different and if there is no curvature it will be the exact same.

We can also hope, a la Newton, that if we compute an infinitesimally tiny path and compare vectors we will have a notion of "curvature AT a point" on our manifold and this local curvature could potentially vary from point to point. On our 2-sphere the curvature is constant across the whole sphere. Parallel transporting can potentially solve all our problems.

We shall transport a co-vector since it happens to come with a $-\Gamma$ and will make the algebra a tiny bit nicer

Strategy:

1. Start with a covector W_μ at $x \in \mathcal{M}$

2. Define a parameterized path $\gamma(t)$ from t_0 to t_f such that $X^\mu(t_0) = X^\mu(t_f)$

3. Parallel transport our covector around $\gamma(t)$ and study $\Delta W_\mu = W_\mu(t_f) - W_\mu(t_0)$. Whatever this is we will expect it to be related to the curvature.

4. Take the limit as the as the "size" of $\gamma(t)$ goes to 0

That's our plan. Now what can we do? Suppose we have some curve $\gamma(t)$ and a covector W_μ at x

Then we can find $W_\mu(t)$ at any point on the curve by solving our parallel transport equation:

$$\frac{dW_\mu}{dt} - \Gamma^\nu_{\mu\rho} V^\rho W_\nu = 0$$

where everything is evaluated on the curve $\gamma(t)$ given by the coordinated functions $X^\mu(t)$. and

$$V^\sigma = \frac{dX^\sigma}{dt}$$

i.e.

$$\frac{dW_\mu(t)}{ds} - \Gamma^\nu_{\mu\rho}(X(t))\frac{dX^\rho(t)}{dt} W_\nu(t) = 0$$

Please be VERY careful. We are transporting a SINGLE covector W_μ - it is NOT a covector field $W(x)$. Everywhere in our equations is $W_\mu(t)$ rather than $W_\mu(X(t))$. As we move along $\gamma(t)$ our value of W_μ changes in relation to the manifold's metric!

Now we try to compute

$$\Delta W_\mu = W_\mu(t_f) - W_\mu(t_0)$$

$$= \int_{t_0}^{t_f} \frac{dW_\mu(t)}{dt} dt = \int_{t_0}^{t_f} \Gamma^\rho_{\mu\sigma}(X(t))\frac{dX^\sigma(t)}{dt} W_\rho(t) dt$$

$$= \int_{X_0}^{X_f} \Gamma^\rho_{\mu\sigma}(X(t)) W_\rho(t) dX^\sigma$$

Now we have a couple routes we could take. First, let's study this change infinitesimally to first order. Then, we have

$$X^\mu(t + dt) = X^\mu(t) + \delta X^\mu(t)$$

Next, we Taylor expand everything keeping only the first order terms. Consider $t = t_0 + dt$ and $X(t) = X(t_0) + \delta X$.

Call

$$X(t_0) = X_0$$

then we have

$$\Gamma^\rho_{\mu\sigma}(X(t)) = \Gamma^\rho_{\mu\sigma}(X(t_0)) + \partial_\lambda \Gamma^\rho_{\mu\sigma}(X_0)(X - X_0)^\lambda$$

$$W_\rho(t + dt) = W_\rho(t_0) + \frac{dW_\rho(t_0)}{dt}(t - t_0)$$

But we know

$$\frac{dW_\rho(t_0)}{dt} = \Gamma^\kappa_{\rho\lambda}(X_0)V^\lambda W_\kappa(t_0)$$

Then all together
Making these substitutions yields

$$\Delta W_\mu = [\partial_\lambda \Gamma^\rho_{\mu\sigma}(X_0) + \Gamma^\kappa_{\mu\sigma}\Gamma^\rho_{\kappa\lambda}]W_\rho(X_0)\int_{X_0}^{X_f} X^\sigma dX^\lambda$$

or

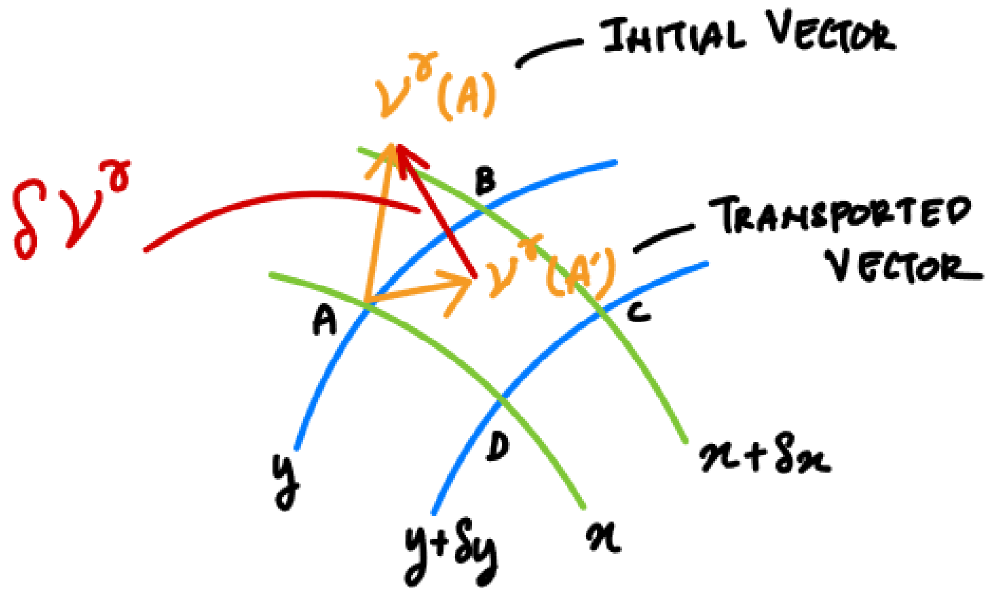
$$\Delta W_\mu = \frac{1}{2}R^\rho_{\mu\sigma\lambda}W_\rho a^{\sigma\lambda}$$

where

$$a^{\sigma\lambda} = \int_{X_0}^{X_f} X^\sigma dX^\lambda$$

is the signed area enclosing the curve. $a^{\sigma\lambda} = -a^{\lambda\sigma}$
and

$$R^\rho_{\mu\sigma\lambda} = [\partial_\sigma \Gamma^\rho_{\mu\lambda} + \Gamma^\rho_{\kappa\sigma}\Gamma^\kappa_{\mu\lambda}] - [\partial_\lambda \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\kappa\lambda}\Gamma^\kappa_{\mu\sigma}]$$



This was a very tricky derivation but we have arrived at the Riemann Curvature Tensor. Notice that curvature depends on the derivatives of the connections and that they themselves are derivatives of the metric. Therefore curvature is inherently a bunch of second derivatives of our metric. This is akin to freshman calculus where you learned the curvature of a function was controlled by the second derivative of a function.

Now we shall derive it in a MUCH simpler fashion such that it will be VERY easy to remember its form and how to compute it.

Now, if there were any justice in this world the condition of curvature, like that of flat calculus would be related to the second derivative being non-zero

$$f''(x) = \frac{d^2 f}{dx^2} = \nabla^2 f \neq 0$$

Therefore, we might expect curvature to be related to D^2, say

$$D_\mu D_\nu W_\alpha$$

But since our general covariant derivative doesn't commute then we reckon

$$[D_\mu D_\nu - D_\nu D_\mu]W_\alpha = [D_\mu, D_\nu]W_\alpha$$

Now, for a flat space we would suppose this commutator to identically vanish since our Euclidean derivative commute. Additionally, whatever result we get from this operation on W^α we expect to be able to write the result as a linear combination of W^α. So then

$$[D_\mu, D_\nu]W_\alpha = \tilde{R}^\sigma_{\alpha\mu\nu} W_\sigma$$

Aside from constants this is our only option. We write $\tilde{R}$ since it's not yet cleaner this is our curvature tensor, but it certainly walks and talks like a curvature tensor. Let's check

Consider a rectangular path from $(0,0)$ to $(\delta x^\mu, 0)$ to $(\delta x^\mu, \delta x^\nu)$ then to $(0, \delta x^\nu)$ and back to $(0,0)$. Now, we translate our vector to the first point

$$W_\rho(x + \delta x^\nu) = W_\rho(x) + \delta x^\nu D_\nu W_\rho(x) = (1 + \delta x^\nu D_\nu) W_\rho(x)$$

Next we go along the next leg of our trip

$$W'_\rho(x' + \delta x^\mu) = W'_\rho(x') + \delta x^\mu D_\mu W'_\rho(x')$$

$$= (1 + \delta x^\mu D_\mu)(1 + \delta x^\nu D_\nu) W_\rho(x)$$

Then we traverse the other order (x^μ first then x^ν)

$$= (1 + \delta x^\nu D_\nu)(1 + \delta x^\mu D_\mu) W_\rho(x)$$

Now we take the difference of these two first order expressions and find

$$= \delta x^\mu \delta x^\nu [D_\mu, D_\nu] W_\rho(x)$$

This difference depends on the area of $\gamma(t)$. Now in the limit of γ goes to zero then the difference in paths becomes

$$[D_\mu, D_\nu] W_\rho(x)$$

But we have expressions for D_μ

$$[D_\mu D_\nu - D_\nu D_\mu] W_\rho(x)$$

$$A_{\nu\rho} = D_\nu W_\rho(x) = \partial_\nu W_\rho - \Gamma^\alpha_{\rho\nu} W_\alpha$$

$$B_{\mu\rho} = D_\mu W_\rho(x) = \partial_\mu W_\rho - \Gamma^\alpha_{\rho\mu} W_\alpha$$

$$D_\mu A_{\nu\rho}(x) = \partial_\mu A_{\nu\rho} - \Gamma^\alpha_{\mu\nu} A_{\alpha\rho} - \Gamma^\alpha_{\mu\rho} A_{\nu\alpha} = H_{\mu\nu\rho}$$

$$D_\nu B_{\mu\rho}(x) = \partial_\nu B_{\mu\rho} - \Gamma^\alpha_{\nu\mu} B_{\alpha\rho} - \Gamma^\alpha_{\nu\rho} B_{\mu\alpha} = G_{\nu\mu\rho}$$

then

$$[D_\mu D_\nu - D_\nu D_\mu] W_\rho(x) = H_{\mu\nu\rho} - G_{\mu\nu\rho}$$

Now we simply have to fill in and keep track of ALL our indices

$$A_{\nu\rho} = \partial_\nu W_\rho - \Gamma^\alpha_{\rho\nu} W_\alpha$$

$$A_{\alpha\rho} = \partial_\alpha W_\rho - \Gamma^\beta_{\rho\alpha} W_\beta$$

$$A_{\nu\alpha} = \partial_\nu W_\alpha - \Gamma^\beta_{\alpha\nu} W_\beta$$

You get the idea: it's a MESS! Next,

$$B_{\mu\rho} = \partial_\mu W_\rho - \Gamma^\alpha_{\rho\mu} W_\alpha$$

$$B_{\alpha\rho} = \partial_\alpha W_\rho - \Gamma^\beta_{\rho\alpha} W_\beta$$

$$B_{\mu\alpha} = \partial_\mu W_\alpha - \Gamma^\beta_{\alpha\mu} W_\beta$$

All together we have that following two expressions that must be simplified and subtracted

$$H_{\mu\nu\rho} = \partial_\mu[\partial_\nu W_\rho - \Gamma^\alpha_{\rho\nu} W_\alpha] - \Gamma^\alpha_{\mu\nu}[\partial_\alpha W_\rho - \Gamma^\beta_{\rho\alpha} W_\beta] - \Gamma^\alpha_{\mu\rho}[\partial_\nu W_\alpha - \Gamma^\beta_{\alpha\nu} W_\beta]$$

$$G_{\mu\nu\rho} = \partial_\nu[\partial_\mu W_\rho - \Gamma^\alpha_{\rho\mu} W_\alpha] - \Gamma^\alpha_{\nu\mu}[\partial_\alpha W_\rho - \Gamma^\beta_{\rho\alpha} W_\beta] - \Gamma^\alpha_{\nu\rho}[\partial_\mu W_\alpha - \Gamma^\beta_{\alpha\mu} W_\beta]$$

Notice that because the partials commute then they drop out of our computation. When the dust settles and the indices are worked out we then have

$$[D_\mu, D_\nu]W_\rho = ([\partial_\mu \Gamma^\alpha_{\nu\rho} + \Gamma^\alpha_{\mu\kappa}\Gamma^\kappa_{\nu\rho}] - [\partial_\nu \Gamma^\alpha_{\mu\rho} + \Gamma^\alpha_{\nu\kappa}\Gamma^\kappa_{\mu\rho}])W_\alpha$$

or

$$[D_\mu, D_\nu]W_\rho = R^\alpha_{\rho\mu\nu}W_\alpha$$

$$R^\rho_{\mu\sigma\lambda} = [\partial_\sigma \Gamma^\rho_{\mu\lambda} + \Gamma^\rho_{\kappa\sigma}\Gamma^\kappa_{\mu\lambda}] - [\partial_\lambda \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\kappa\lambda}\Gamma^\kappa_{\mu\sigma}]$$

$R^\rho_{\mu\sigma\lambda}$ is the famous Riemann Curvature Tensor. Everything in its expression can be found from taking derivatives of the metric or the basis vectors which vary from point to point. We hope that if we apply this to our 2-sphere then we will get our usual result that the curvature of the 2-sphere is $1/R^2$ and constant.

12 The 2-sphere yet again!

recall that we found for S^2 the fo

$$g_{ij} = \begin{bmatrix} R & 0 \\ 0 & R^2\sin^2\theta \end{bmatrix}$$

$$g_{ij,\theta} = \begin{bmatrix} 0 & 0 \\ 0 & 2R^2\sin\theta\cos\theta \end{bmatrix}$$

$$g_{ij,\phi} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Gamma^\theta_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & -\sin\theta\cos\theta \end{bmatrix}$$

$$\Gamma^\phi_{ij} = \begin{bmatrix} 0 & \cot\theta \\ \cot\theta & 0 \end{bmatrix}$$

Therefore we have just three non-zero components of Γ

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta$$

$$\Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\theta$$

To find $R^{\rho}_{\mu\sigma\lambda}$ we will need derivatives of these. The only derivatives that are non-zero are θ derivatives

$$\partial_\theta \Gamma^{\theta}_{\phi\phi} = -\partial_\theta(\sin\theta\cos\theta) = -\cos^2(\theta) + \sin^2(\theta) = -\cos(2\theta)$$

$$\partial_\theta \Gamma^{\phi}_{\theta\phi} = \partial_\theta \Gamma^{\phi}_{\phi\theta} = -\partial_\theta \cot\theta = -\csc^2\theta$$

$$R^{\rho}_{\mu\sigma\lambda} = [\partial_\sigma \Gamma^{\rho}_{\mu\lambda} + \Gamma^{\rho}_{\kappa\sigma}\Gamma^{\kappa}_{\mu\lambda}] - [\partial_\lambda \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\kappa\lambda}\Gamma^{\kappa}_{\mu\sigma}]$$

We could write everything out in gory detail or acknowledge that if $\mu = \phi = \lambda$ and $\sigma = \theta$ then

$$R^{\theta}_{\phi\theta\phi} = -\cos(2\theta) + \sin\theta\cos\theta\cot\theta = \sin^2(\theta)$$

$$R^{\phi}_{\theta\phi\theta} = \sin^2(\theta)$$

ALL other components are zero! It was an ENORMOUS task to get just these two components by hand. Typically, for more complicated manifolds one reports to computers but it's always possible by hand if you're very careful!

Well this looks nothing like our $1/R^2$ curvature we've come to know and love since grade school. What gives?

First, this is a 4-index tensorial quantity and generally varies from point to point. What we hope for is a scalar quantity that we can acquire from this Riemann Curvature Tensor. Riemann give you the single most accurate and precise measure of curvature everywhere on the manifold. It is akin to a topographical map - it gives us everything at every point.

Sometimes this is way too much information and we instead wish to know more of an average type of curvature that can tell us how volumes from place to place change.

Given $R^{\rho}_{\mu\sigma\lambda}$ we are free to contract it with our metric as many times as needed to acquire a scalar quantity. Therefore we compute

$$R^{\rho}_{\mu\rho\lambda} = R^{\theta}_{\mu\theta\lambda} + R^{\phi}_{\mu\phi\lambda} = R_{\mu\lambda}$$

This is called the Ricci Tensor. We get

$$R_{\theta\theta} = 1$$

$$R_{\phi\phi} = \sin^2\theta$$

Finally, we may also wish to know a very coarse grained notion of curvature. This is called the scalar curvature and if the scalar produced is positive then we have a manifold that is positively curved like a sphere, if it is negative then it is more akin to a hyperbolic space, and if the scalar curvature is zero then we have a flat Euclidean space.

We get our scalar curvature from tracing over the Ricci tensor.

$$R = g^{\mu\lambda}R_{\mu\lambda} = \mathrm{Tr}\, R_{\mu\lambda} = \frac{2}{R^2}$$

Which, aside from the 2 is the Gaussian curvature of the sphere we've always have known! As R goes to infinity this scalar curvature does indeed go to zero just as we would expect.