

Math 136 Class Notes: based on the course taught by Dr. Puskar Mondal

Wittmann Goh (wgoh@college.harvard.edu)

Fall 2023

Contents

1	(9/1) Course Overview	4
1.1	Curves in $\mathbb{R}^n$	4
1.2	Surfaces in $\mathbb{R}^n$	6
1.3	Intrinsic versus extrinsic geometry	6
1.4	Geometry of Submanifolds	6
1.5	What's after this class?	7
2	(9/6) Curves in $\mathbb{R}^n$	7
2.1	Parameterized Curves	7
2.2	Integral Curves	8
2.3	Examples	9
3	(9/8) Phase Space, Fixed Points and Regular Curves	10
3.1	Recap	10
3.2	(Con't) Particle in a Potential	10
3.3	Fixed Points, Regular Curves	12
4	(9/13) Reparameterization, Length Framework	12
4.1	Reparameterizations	12
4.2	Invariant Property: Curve Length	14
5	(9/15): Invariants (Con't)	15
5.1	Arclength	15
5.2	Energy	17
5.3	More Invariants	18
6	(9/20): Invariants: Frames	18
6.1	Recap	18

6.2	Frames	19
6.3	Change of Frame Along the Curve	21
7	(9/22): Frames (Con't)	21
7.1	Derivative of a Frame	21
7.2	Curvature	23
8	(9/27): Curvature	23
8.1	Wrapping up curvature	23
8.2	Going from Curvature back to the Curve	23
9	(9/28): Curvature, Pt. 2: The case of $\mathbb{R}^2$	25
9.1	Curvature in $\mathbb{R}^2$	25
9.2	Curves of Constant, Non-zero Curvature in $\mathbb{R}^2$	27
10	(10/4): Some logistics, Curves in $\mathbb{R}^3$	28
10.1	Midterm Format	28
10.2	Space Curves (Curves in $\mathbb{R}^3$)	29
10.3	Representing the curve locally	29
10.4	Projections onto planes in the frame	30
11	(10/6): Wrapping up Curvature; Transformations	31
11.1	More on the Taylor Expansion	31
11.2	Transformations	32
11.3	Topology Fundamentals	32
12	(10/11) : Topology Fundamentals, Con't	33
12.1	Point-Set Topology on $\mathbb{R}^n$	33
12.2	Injectivity, Surjectivity, Bijectivity	34
12.3	Homeomorphisms, Simply-connectedness	34
13	(10/13) : Diffeomorphisms	35
13.1	Announcements	35
13.2	Wrapping up Topology	36
13.3	Differentials, Diffeomorphisms	36
14	(10/18) : More on Diffeo/Homeos, Topological Spaces	38
14.1	Diffeomorphisms	38
14.2	Homeomorphisms and Diffeomorphisms: The Bigger Picture	39
14.3	Topology	39
15	(10/20) : Manifolds	40

15.1	Wrapping up topological spaces	40
15.2	Manifolds: Motivation	40
16	(10/25) : Charts and Atlases	41
16.1	Charts, Atlases	41
16.2	Atlas of S^2 , Stereographic Projection	42
17	(10/27) : Maps between Manifolds, “Bootstrapping”	43
17.1	Maps between Manifolds	43
17.2	Homeo/Diffeomorphisms of Manifolds	45
17.3	Road map for next few lectures	45
18	(11/1) : Tangent Spaces, Equivalence Relation on Integral Curves	46
18.1	Recap	46
18.2	Tangent Spaces	46
18.3	Equivalence of Curves by Tangency	47
19	(11/8) : Defining the Tangent Space	48
19.1	Midterms, Other logistics	48
19.2	Tangent Vectors	48
19.3	Construction of the Tangent Space	48
20	(11/10) : Tangent Bundles, Derivations	49
20.1	More on Tangent Spaces	49
20.2	Tangent Bundles	50
20.3	Constructing a basis, Derivations	51
20.4	Teaser for next week	52
21	(11/15): Basis of the Tangent Space, Dual Space	52
21.1	Differentials as the Basis of the Tangent Space	52
21.2	Dual Space	54
22	Co-tangent spaces, co-vectors, Tensors,.....	54
23	Metric	59
24	Lengths	61
24.1	Distances on a Manifold	62
25	(11/22) : (TODO:)	63
25.1	Some logistics	63
25.2	Towards an Inner product	63

25.3 Computing the distance	65
26 (11/29) : Derivatives of Tangent Vectors	66
26.1 Connections and Connection Coefficients	67
26.2 Metric compatibility	68
27 (12/1): Last Lecture!	68
27.1 Metric Compatibility	68
27.2 Obtaining the Connection from a Metric	69
27.3 Geodesics	70
28 (11/17): (TODO:)	73
28.1 Change of Charts on the Tangent Space	73
28.2 Metrics	75
29 Differentials in the space of matrices	76
29.1 Applications	76
29.2 Course Feedback	77

1 (9/1) Course Overview

Score breakdown: 50% homework, 50% exams. Homework will be fortnightly, tentatively released on Tuesdays or Thursdays. You are free to take a couple of late days here and there.

Puskar's office hours will be 8-9:30pm on Tuesdays and 7:30-9:00 pm Thursdays. Mine are TBD, but I will confirm them hopefully by end of next week. We will begin office hours starting the week of 12th Sept, once the first homework is released.

Here are some topics and questions that we will study over this course:

1.1 Curves in $\mathbb{R}^n$

A curve in $\mathbb{R}^n$ is a map $I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$. Can think of this as taking the unit interval, inserting it into $\mathbb{R}^n$, and deforming it in some fashion.

Physical interpretation. A curve can be treated as the trajectory of a point particle. Let $x(t)$ be this path. The derivative at any point $\dot{x}(t)$ is the velocity of the particle. Taken over the entire space, $\dot{x}(t)$ can be treated as a vector field. One important property is the fixed point, or where the velocity vanishes ($\dot{x}(t) = 0$). It turns out that studying these fixed points can reveal properties about the topology of the global space.

Let's look at the sphere S^2 (note: this is the 2-dimensional boundary of the unit solid sphere). An interesting theorem is the following:

Theorem 1.1 (Hairy ball theorem). On the sphere S^2 , any vector field must vanish at at least 1 point on the sphere. i.e. for any vector field $\vec{F}$, there exists a point $x \in S^2$ such that $\vec{F}(x) = \vec{0}$.

Remark. This isn't the precise statement of the Hairy ball theorem. In particular, we will require that $\vec{F}$ is always tangent to the sphere.

Physical interpretation: A point particle living in $\mathbb{R}^n$ is described by its position and velocity, the pair $(x(t), \dot{x}(t))$.

Example 1.2. If we have a force field $\vec{F}$ and a particle moving inside the force field. What does it mean in this context when $\dot{x}(t) = 0$?

Physically, this means that the particle has run out of energy.

Another scenario: particle on a sinusoidal surface. From classical mechanics, we know that a particle starting on the slope will oscillate.

How does the particle's trajectory look like in position-velocity space? It should be a circle (or more generally an ellipse) - see Figure 1.

What happens if we give the particle just enough energy to overcome the potential energy and pass over the crest? For example we can start the particle moving from the top of the crest with 0 velocity. Then when we release it, when it reaches the next peak it will just have enough energy to cross over and keep moving to the right. Let's see how this particle looks in phase space (also Figure 1).

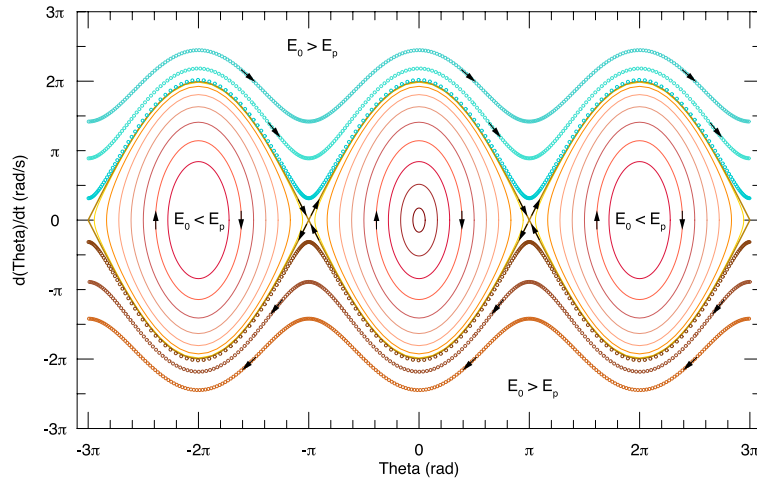


Figure 1: This is a plot of velocity ($\dot{x}$) against position (x). The red curves are the trajectories of particles which have insufficient energy to cross the crest. The brown curves (which periodically meet the $\dot{x} = 0$ line) are the trajectories of particles with just enough energy to cross the crest. [Image taken from this page](#)

Question 1.3. Are the intersections with the position-axis considered equilibrium points?

No, because those points are in phase space. These intersections are points where $\dot{x}(x(t)) = 0$, which is not the same as the condition for equilibrium (which is $\dot{x}(t) = 0$).

The confusion from above is essentially the cause of us looking at different spaces. The phase space is the space of pairs $(x, \dot{x})$. For the example considered earlier, the phase space is $\mathbb{R}^2 \times \mathbb{R}^2$. We can project this onto configuration space, which gives the physical trajectory x of the particle, and this space in our example is $\mathbb{R}^2$.

One interesting thing to note is that curves in these particular phase spaces are non-intersecting. Physical explanation: Given a force field, a particle's trajectory is uniquely determined by $(x, \dot{x})$. Therefore given any point in phase space, there is only one possible direction it can travel in. Contrast this to the trajectory in configuration space, which are clearly allowed to intersect.

1.2 Surfaces in $\mathbb{R}^n$

We saw previously that we can think of curves as maps $I \rightarrow \mathbb{R}^n$. We can treat surfaces similarly. For clarity we define surfaces to be 2-dimensional objects (in some precise sense that we will eventually see) in $\mathbb{R}^n$. Then, we can think of surfaces as maps

$$\begin{aligned} F : U \subset \mathbb{R}^2 &\longrightarrow \mathbb{R}^n \\ (u, v) &\longmapsto (x^1(u, v), x^2(u, v), \dots, x^n(u, v)) \end{aligned}$$

where each x^i is a function $\mathbb{R}^2 \rightarrow \mathbb{R}$.

Sphere

We know that algebraically the sphere $S^2 \in \mathbb{R}^3$ is the set

$$S^2 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$$

This formulation, we will see, is not very useful for differential geometry.

1.3 Intrinsic versus extrinsic geometry

There are a number of questions we can ask. If we are in $\mathbb{R}^3$, we can easily answer the question of what the distance is between two points. The same question is not as obvious if we constrain the two points to be on a sphere. What “distance” are we talking about? Of course, we could still consider the straight-line distance that passes into the interior of a sphere. But a more interesting question is: what if we constrained the path to lie on the sphere? Then the answer is not so clear.

Well, some might know the solution to this question already. The shortest distance is given by the arclength of the great circle that contains both endpoints. But what about an arbitrary surface? What is the distance in this case? Here is how we can define the distance. Let p be a path between the endpoints x and y , and let P be the set of all possible paths. Then we can simply define the distance to be the minimum length of all such paths:

$$d(x, y) = \inf_{p \in P} \|p\|$$

Definition 1.4. The path that minimizes the distance is called a geodesic.

But we have just deferred the problem: how do we calculate the length of any such path p ? Here is where the notion of intrinsic geometry comes into play. We don’t have the luxury of using the calculus of $\mathbb{R}^n$, instead we have to do calculus on the manifold itself.

Definition 1.5 (Informal). An n -dimensional manifold is something that looks like $\mathbb{R}^n$ locally, in the sense that for every point on the manifold, if we zoom in enough, then we should see something that looks like $\mathbb{R}^n$.

Take a curve in $\mathbb{R}^2$. It looks curved. But we will see that a curve is really in fact flat in an intrinsic sense, and that the curvature only exists in an extrinsic sense. Roughly speaking, the extrinsic curvature depends on how we embed an object (curve, surface, etc) into a larger Euclidean space.

1.4 Geometry of Submanifolds

Once we set up the machinery for doing calculus on manifolds, we will see that it is not difficult to extend the tools to deal with submanifolds, which can be thought of simply as subsets of manifolds. We will see that the intrinsic perspective will be very helpful here.

1.5 What's after this class?

This class will serve as a solid foundation for Math 230A/B. 230A will focus on bundles over manifolds, and 230B will focus on general relativity.

2 (9/6) Curves in $\mathbb{R}^n$

As promised, we will start from the basics: parametrized and integral curves in $\mathbb{R}^n$.

We will use the following notation: A point $P \in \mathbb{R}^n$ has the coordinates $(x^1, x^2, \dots, x^n) \equiv x$.

2.1 Parameterized Curves

Let's begin with a preliminary definition:

Definition 2.1 (First attempt). Let $I \subseteq \mathbb{R}^n$. A parameterized curve is a map

$$\begin{aligned} I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto (x^1(t), x^2(t), \dots, x^n(t)) \equiv x(t) \end{aligned}$$

What's wrong with this definition? We haven't said anything about even the continuity of the functions. We could end up with a bunch of discontinuous segments and it would count as a curve under our definition. To restrict ourselves to this nicer family of curves, we should demand that each of the coordinate functions $x^1, \dots, x^n$ are continuous.

Are we done yet? The revised definition still allows curves with "sharp" corners, where the curve abruptly changes directions. While these curves are certainly interesting to study, we will not deal with them in this course.

Definition 2.2. A curve is said to be:

- C^0 , if the curve is continuous (eg: the graph $(x, |x|) \subset \mathbb{R}^2$).
- C^1 , if the curve once-differentiable (note that this implies continuity).
- C^k if the curve is k -times differentiable
- C^∞ if the curve is infinitely differentiable, aka "smooth"

So we adjust our definition of a parameterized curve to the following:

Definition 2.3. Let $I \subseteq \mathbb{R}^n$ be a connected set (see Qn 2.5). A parameterized curve is a smooth map

$$\begin{aligned} I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto (x^1(t), x^2(t), \dots, x^n(t)) \equiv x(t) \end{aligned}$$

Question 2.4. Why do we write subscripts in superscript?

Answer. The superscripts denote a contravariant elements, and subscripts are reserved for covariant elements. This will make a difference when we get to manifolds, but that is for later.

Question 2.5. Can I be the integers $\mathbb{Z}$ in the above definition?

Answer. Well, because we are dealing with notions of continuity and differentiable, we want I to be an open interval (we want to be able to take left/right limits). $\mathbb{Z} \subset \mathbb{R}$ is closed, so that is not an interval. Definition 2.3 amended to reflect this.

Let's see if we need to further refine this definition. The current definition still allows self-intersections! This will be a problem if we want the curve to have a well-defined tangent at every point.

Definition 2.6. The velocity vector $v(t)$ at a point $x(t)$ is the derivative of the parameterized curve:

$$\begin{aligned} v(t) &:= \frac{dx(t)}{dt} \\ &= \left(\frac{dx^1(t)}{dt}, \dots, \frac{dx^n(t)}{dt} \right) \end{aligned}$$

Remark. Note that the derivatives are all well-defined since we imposed smoothness.

2.2 Integral Curves

Let's leave the notion of parameterized curves behind for a while and look at integral curves. To define this notion we need to look at vector fields in $\mathbb{R}^n$.

Let's recall: Given two vectors in $\mathbb{R}^n$, based at different points, how do you find their vector sum? You can do parallel transport and move the base of one of the vectors to the base of the other. Once they are at the same base you can use the parallelogram rule.

Notion of parallel transport: The transport is really given by a vector field, which tells us how the vector changes as we move it from one base point to another. For the spaces we are used to, this vector field is uninteresting (it is just the same vector everywhere)

Definition 2.7. A vector field on $\mathbb{R}^n$ (or a subset $U \subset \mathbb{R}^n$) is an assignment of a vector to each point $x \in \mathbb{R}^n$ (or $x \in U$). In other words, it is a function y that looks like the following:

$$\begin{aligned} y : U \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (x^1, \dots, x^n) &\longmapsto (y^1(x^1, \dots, x^n), \dots, y^n(x^1, \dots, x^n)) \end{aligned}$$

and also satisfies the condition that each y^i is C^∞ in each of the components x^j .

Now we can ask the following:

Given a vector field y , can we construct a curve such that at each point x of the curve, $y(x)$ is the velocity vector to the curve at that point? Is this even possible locally, in a small interval?

To answer this, we return to the definition of parameterized curves momentarily. Since y is smooth it is in particular integrable. We see that we can let

$$\begin{aligned} \frac{dx^1(t)}{dt} &= y^1(x^1(t), \dots, x^n(t)) \\ &\vdots \\ \frac{dx^n(t)}{dt} &= y^n(x^1(t), \dots, x^n(t)) \end{aligned} \tag{*}$$

Further given some initial condition (the coordinates of $x(0)$), the question reduces to the question of whether this system of differential equations is solvable. We will see (and prove) that the answer is yes, but only locally.

Remark. We would like/expect the following conditions:

1. Existence (locally)
2. Uniqueness (each initial condition produces exactly one solution)
3. Cauchy stability: Smooth dependence of the curve on the initial condition (If we change the initial condition slightly, we expect the curves to be similar, at least for a sufficiently small time interval)

Theorem 2.8 (Local well-posedness of integral curves, or LWT). Let $x(0) = \alpha$ be the initial data/condition for the system of differential equations $(*)$ such that $|x(0)| < \infty$. Then there exists a solution to $(*)$ on the interval $t \in [0, t^*)$ such that

$$(x^1(0), \dots, x^n(0)) \mapsto (x^1(t), \dots, x^n(t))$$

is continuous. Furthermore, either $t^* = \infty$ or at least one of the functions blows up: $\limsup_{t \rightarrow t^*} \left(\max_{j=1, \dots, n} \|x^j(t)\| \right) = \infty$. t^* is called the lifetime of the solution.

We conclude by giving a formal definition of integral curves.

Proof. We will see, maybe next class. Sneak peek: It uses the contraction mapping principle and Picard iterations □

Question 2.9. Is the condition of Cauchy stability necessary given that we are already assuming C^∞ of the curve?

Answer. Smoothness on the vector field of course tell you that you can't have arbitrary curves given two initial data close to each other. Think of this as a technical point at the moment: This is provable but really quite painful, since taking n -th derivatives of multivariate functions gets out of hand quickly. We shall see later that this is tied to the issue of *Chaos* i.e., given two initial data that are close to each other, the corresponding curves stay close for short time (locally) but can diverge in long time.

Question 2.10. Could you define $\limsup$?

Answer. The limit supremum $\limsup_{t \rightarrow \infty} f$ is the limit of $\sup f$ over the domain (t, ∞) as t goes to ∞ .

Remark. the construction of integral curves automatically forbids self-intersections.

Definition 2.11 (Integral curves). Add to the definition of a parameterized curve, the requirement that it is tangent to some vector field.

2.3 Examples

Example 2.12 (Finite blow-up). Consider the vector field $y = x^2$ in $\mathbb{R}$. Then we want to solve the DE $\frac{dx}{dt} = x^2$, with the initial condition α . By separation of variables, we have

$$\begin{aligned} \int_{x(0)}^{x(T)} \frac{dx}{x^2} &= \int_0^T dt \\ \implies T &= \frac{1}{\alpha} - \frac{1}{x(T)} \\ \implies x(T) &= \frac{\alpha}{1 - \alpha T} \end{aligned}$$

Assume $\alpha > 0$. Then this solution blows up at $T = 1/\alpha$. Formally, take a sequence of points $\{x(T_k)\}_{k=1}^\infty$ such that $\lim_{k \rightarrow \infty} T_k = 1/\alpha$. Then $\lim_{k \rightarrow \infty} x(T_k) = \infty$. Another way to state this: the sequence leaves every compact set. (Compact = closed and bounded in $\mathbb{R}^n$)

Remark. This shows that a finite time interval can map to an “infinite”, unbounded curve.

Let's end of with an example, while introducing the concepts of phase space and configuration space.

Example 2.13 (Newton's Law on $\mathbb{R}$). Recall that a useful picture is to think of a curve as the trajectory of a point particle. Let's think of the potential $V(x) = -E \cos x$ (this arises from a sinusoidal surface).

Newton's Law: $F = ma = m \frac{d^2x(t)}{dt^2}$. Now, given a potential, the force is $F = -\frac{\partial V(x)}{\partial x}$. So considering our scenario, we have

$$m \frac{d^2x}{dt^2} = -\frac{\partial V(x)}{\partial x} = -E \sin x \implies \frac{d^2x}{dt^2} = -\frac{E}{m} \sin x$$

We want to construct an integral curve from this. But what is the issue here? We have a second-order DE! How do we convert this into a form we can use?

We can define $u := \frac{dx}{dt}$. This allows us to write the second-order DE as a system of two first-order DEs:

$$\begin{aligned} \frac{dx}{dt} &= u && \text{(by definition)} \\ \frac{d^2x}{dt^2} &= \frac{du}{dt} = -\frac{E}{m} \sin x && \text{(by } F = ma) \end{aligned}$$

What have we done? We increased the dimension of the target space (where the curve lives) from 1 to 2. This, we will see in the next class, is a lift from configuration space into phase space (to be defined).

3 (9/8) Phase Space, Fixed Points and Regular Curves

3.1 Recap

So far we have looked at parameterized curves and integral curves. We have also seen that integral curves have nicer properties than a parameterized curve.

Let $t \mapsto (x^1(t), \dots, x^n(t))$ be a curve and $y : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ which sends

$$(x^1, \dots, x^n) \mapsto (y^1(x^1, \dots, x^n), \dots, y^n(x^1, \dots, x^n))$$

be a vector field. Then we have seen that the integral curve must satisfy the system of differential equations

$$\frac{dx^1}{dt} = y^1, \dots, \frac{dx^n}{dt} = y^n$$

We have also seen (but not proven) that given an initial condition $x(0) = \alpha$, a unique solution always exists to this system, for some time interval $[0, t^*)$.

3.2 (Con't) Particle in a Potential

Let us return to the last example from the previous class. Recall the setup: $t \mapsto x(t)$ is a point particle in the potential $V(x) = -E \cos x$ (see Figure 2). Newton's law gives us the 2nd-order DE

$$m \frac{d^2x}{dt^2} = -E \sin x$$

We can transform this into a system of 1st-order DEs by defining $\frac{dx}{dt} = u$. This gives us the system

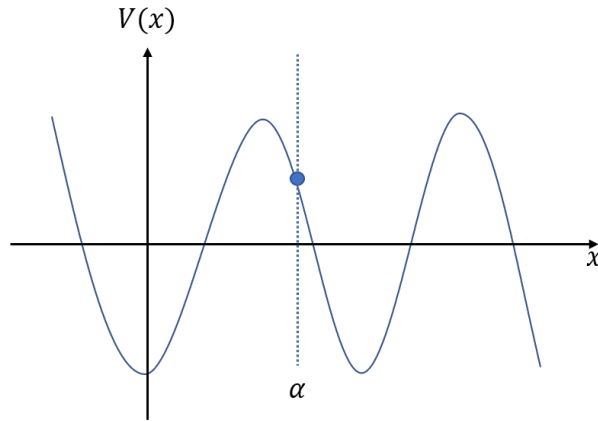


Figure 2: Picture

$$\frac{dx}{dt} = u, \quad \frac{du}{dt} = -\frac{E}{m} \sin x$$

We have increased the dimension of the curve by 1. Instead of a map $t \mapsto x(t)$, now we have a map $t \mapsto (x(t), u(t))$, which is a map $I \rightarrow \mathbb{R}^2$.

Remark. Intuitively, why don't we want to work with higher order differential equations? One way to think about it is that we want to avoid ugly self-intersections as much as possible (consider the fact that the trajectory on the x -axis will intersect itself many times). To “unwrap” the curve, we necessarily must introduce more dimensions.

Let us re-frame the system in terms of our integral curves definition: We have $(x^1, x^2) = (x, u)$ and $y = (y^1, y^2) = (u, -\frac{E}{m} \sin x)$. How do we know that y is a well-defined, smooth vector field? Well, each component is a smooth function of x (and u , trivially, since each component has no u -dependence), so it satisfies the requirements.

Remark. For those of you who are reading ahead, what happened is that we started from a curve on a manifold ($t \mapsto x(t)$) and we constructed a curve on the tangent bundles of the manifold. For those of you who haven't encountered this terminology, it's ok! We will get there in class eventually, and it is not all that scary.

We can plot this qualitatively (Figure 3) (aka not solving the system of differential equations, just relying on intuition) (I missed the exposition related to how the diagram was plotted, but hopefully most of it is clear/makes intuitive sense)

Remark. In this scenario, what makes the first curve $t \mapsto x(t)$ only a parameterized curve and not an integral curve? Each point (other than the left/right bounds) has two velocities associated to it (leftwards and rightwards). Then the tangents cannot come from a vector field, since a vector field is in particular a function.

Remark. It is worth stating that integral curves are a subset of parameterized curves! (One way to see this: note that an integral curve must satisfy the definition of a parameterized curve, but also has additional properties)

Question 3.1. Let's take a step back: why are we doing this? Aren't we trying to do differential geometry?

Answer. Curves are the simplest, well, curved objects. We will need to get comfortable with the simple case before we can start handling higher-dimensional curved objects. In particular, we will be able to handle more interesting manifolds such as the $(n-)$ holed torus.

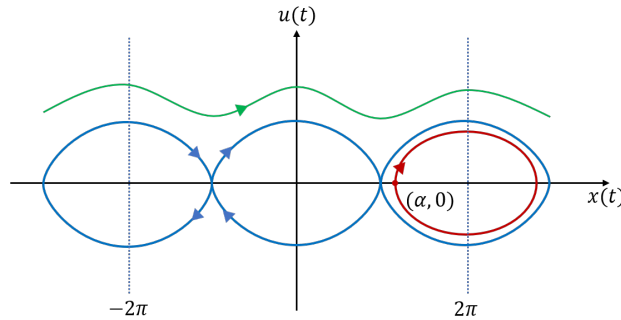


Figure 3: The 3 curves in phase space corresponding the scenarios we discussed

Let's put a particle (almost) at rest at $x(0) = -\pi$, with very small initial velocity towards the right. What happens now? It will roll down, reach the next peak at (essentially) 0 velocity, and then continue to the next valley. This is the blue curve in Figure 3.

Lastly, let's consider if we give it a good shove starting at the peak. This gives rise to the green curve in Figure 3.

Question 3.2. But what is going on where the blue curves intersect? Aren't there 2 different vector fields there? Does that mean that this isn't actually an integral curve?

Answer. Well, let's take a look. Recall that $y = (y^1, y^2) = (u, -\frac{E}{m} \sin x)$. What happens if $y = (0, 0)$? This means $u = 0$ and $\sin x = 0$, which has solution $(x, u) = (n\pi, 0)$, $n \in \mathbb{Z}$. If we now look at the phase plot, we see that for n even, these correspond exactly to the intersection points of the blue curve!

This is a good segue into our next important definition:

3.3 Fixed Points, Regular Curves

Definition 3.3 (Fixed points). A fixed point of an integral curve is a zero of its associate vector field.

Remark. We can unpack this a bit. Let $y = (y^1, \dots, y^n)$. Recall that the integral curve is defined by the system $\frac{dx^i}{dt} = y^i$. If y vanishes, then this implies that each x^i is constant, and so the particle is fixed at that point.

Remark. How is this consistent with Cauchy stability (the condition that nearby initial conditions lead to similar curves)? Well, it is kind of a special case, because a fixed point doesn't even give rise to a curve.

Definition 3.4 (Regular (integral) curve). A regular curve is an integral curve whose vector field never vanishes.

Remark. So the blue curves are regular only if the fixed points are excluded.

4 (9/13) Reparameterization, Length Framework

4.1 Reparameterizations

For the past few lectures we have been working with parameterizations of curves: functions $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ mapping $t \mapsto (x^1(t), \dots, x^n(t))$. Also recall that t is called the parameter. By assumption of smoothness, we can take as many derivatives with respect to t we want.

What is the purpose of t ? It allows us to keep track of where along the curve we are at. But we don't always have to stick to the same parameterization. Define $S := s(t)$, where $s : \mathbb{R} \rightarrow \mathbb{R}$. This is a change of variables for the parameterization of the curve. Under certain conditions, this change of parameters does not change the geometric properties of the curve (eg. length).

The key intuition is that the parameter is just a way of labelling the points. This labelling does not (and should not) affect the geometry of the curve!

Formally, consider two curves given by the parameterizations

$$\begin{aligned} c_1 : I \subseteq \mathbb{R} &\rightarrow \mathbb{R}^n, & t &\mapsto c_1(t) \\ c_2 : J \subseteq \mathbb{R} &\rightarrow \mathbb{R}^n, & s &\mapsto c_2(s) \end{aligned}$$

along with a function that sends $s \mapsto t(s)$.

Definition 4.1 (Reparameterization). Formally, consider two curves given by the parameterizations:

$$\begin{aligned} c_1 : I \subseteq \mathbb{R} &\rightarrow \mathbb{R}^n, & t &\mapsto c_1(t) \\ c_2 : J \subseteq \mathbb{R} &\rightarrow \mathbb{R}^n, & s &\mapsto c_2(s) \end{aligned}$$

A reparameterization of a regular curve $c_2 : s \mapsto c_2(s)$ is a change of variables $s \mapsto t(s)$ such that the following diagram commutes: **(TODO: draw diagram)**

In other words, the identity $c_2(s) = c_1(t(s))$ holds.

Question 4.2. When we say that length is preserved, is that an extrinsic or intrinsic property of the curve?

Answer. We will see. Currently we have not even defined length yet.

Example 4.3. Consider the curve $t \mapsto (t, t)$ for the parameter interval $t \in (0, 1)$. Suppose we now reparameterize by $s(t) = \tan\left(\frac{\pi t}{2}\right)$. Solving for t , this is the same as $t(s) = \frac{2}{\pi} \tan^{-1}(s)$. The curve can now be given as $s \mapsto (t(s), t(s)) = \left(\frac{2}{\pi} \tan^{-1}(s), \frac{2}{\pi} \tan^{-1}(s)\right)$. Applying $s(t)$ to the interval itself, we see that the interval for the new parameterization is $s \in (0, \infty)$.

Question 4.4. What if the function looks like $s(t) = |t|$?

Answer. We will put some restrictions on s such that these cases are not allowed. (Intuitively, we want to ensure that the reparameterization is also smooth)

We want the resulting parameterization to be smooth. Start with $c_1(t(s)) = c_2(s)$. Taking derivatives, we have $\frac{dc_1(t)}{ds} = c'_1 \cdot \frac{dt(s)}{ds}$, where we write $c'_1 := \frac{dc_1}{dt}$. Taking norms, we have

$$\left| \frac{dc_1(t(s))}{ds} \right| = |c'_1| \cdot \left| \frac{dt(s)}{ds} \right| \implies \left| \frac{dt(s)}{ds} \right| = \frac{1}{|c'_1|} \left| \frac{dc_1(t)}{ds} \right|$$

Since we want the derivative to always exist, $|c'_1|$ can never vanish. This is always true since c_1 is required by the definition to be a regular curve (the derivative wrt the parameter is given by the vector field). We also want

1. $\left| \frac{dt(s)}{ds} \right| < \infty$
2. $\frac{dt(s)}{ds} \neq 0$
3. t is a continuously differentiable function $J \rightarrow I$.

If these conditions hold, we see that $\frac{dc_1(t(s))}{ds}$ never vanishes, and so $c_1(t(s)) := c_2(s)$ will also be a regular curve. Here, t is an example of a " C^1 -diffeomorphism", or a C^1 change of variables. If the first two conditions are replaced with $0 < \left| \frac{d^n t(s)}{ds^n} \right| < \infty$, then t would be a C^n -diffeomorphism/change of variables.

Remark. The reparameterized curve can be as smooth as the original smooth curve, depending on the smoothness of the change of variables function t .

Question 4.5. Can we still say that the two curves are the same if the parameterizations belong to different C^k ?

Answer. The condition that two curves are the same are just the fact that the images of the parameterizations coincide (they map their corresponding intervals to the same subset of space). It does not require anything about their derivatives.

4.2 Invariant Property: Curve Length

From now on, we stay in the world where all parameterizations, as well as changes of variables, are C^∞ .

Question: how many smooth, bijective functions $s \mapsto t(s)$ are there in $C^\infty(\mathbb{R})$? Uncountably many.

(Fun functional-analytical fact: Let H^k be the space of functions whose k th derivatives are square-integrable. The space $C_0^\infty(\mathbb{R})$ of compactly-supported smooth functions (functions that vanish except on a closed bounded subset) is dense in H^k .)

This is not very helpful for us. We want an invariant property of a curve: some property that is unchanged after taking any smooth reparameterizations. The first such property is the length of the curve.

Let $I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ acting by $t \mapsto c_1(t) = (x^1(t), \dots, x^n(t))$. How do you measure distances between two points? Pythagoras. But to calculate the length of the curve, you cannot just calculate the distance between the endpoints. How do we deal with this? At every point on the curve, the curve locally looks like a straight line!

If we zoom in to a small line segment with start point (x^1, x^2) and endpoint $(x^1 + dx^1, x^2 + dx^2)$ (why are we allowed to do this? Smoothness assumption), we can write the length of the segment as

$$d\ell = \sqrt{(dx^1)^2 + (dx^2)^2} = \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \left(\frac{dx^2}{dt}\right)^2} dt$$

Remark. We will see once we move to doing more complicated manifolds, we can repeat a similar procedure to define the length, but the expression for the infinitesimal length will differ.

Definition 4.6 (Length). Let **Curve** be the space of (smooth) curves. The length of a curve $c : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ is a function

$$\begin{aligned} \ell : \mathbf{Curve} &\longrightarrow \mathbb{R}_{\geq 0} \\ c &\longmapsto \ell(c) := \int_I \left| \frac{dc}{dt} \right| dt \end{aligned}$$

Well, we claimed that the length is an invariant property. Let's prove it:

Proposition 4.7. The length function is invariant under reparameterization

Proof. Let $c_1 : (0, T) \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ sending $s \mapsto c_1(s)$, and let $s \mapsto t(s)$ be a reparameterization. The “new” curve is given by $c_2 : s \mapsto c_2 \circ t(s)$. The length of c_2 is

$$\begin{aligned} \ell(c_2) &= \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_2}{ds} \right| ds = \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_1}{dt} \frac{dt}{ds} \right| ds \\ &= \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_1}{dt} \right| \left| \frac{dt}{ds} \right| ds \end{aligned}$$

There are two cases: $\frac{dt}{ds} > 0$ and $\frac{dt}{ds} < 0$. In the first case, we can infer $s(0) < s(T)$ (the change of variables is orientation-preserving), and in the second case we have $s(0) > s(T)$ (orientation reversing). Assuming the first case, we can further simplify:

$$\begin{aligned} \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_1}{dt} \right| \left| \frac{dt}{ds} \right| ds &= \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_1}{dt} \right| \frac{dt}{ds} ds \\ &= \int_{t=0}^{t=T} \left| \frac{dc_1}{dt} \right| dt \end{aligned}$$

In the second case, we flip sign twice (once when we drop the $|\cdot|$, once to swap the integration limits; note we integrate from $s(t=T)$ to $s(t=0)$ since $s(t=0) > s(t=T)$ in the second case)

$$\begin{aligned} \int_{s(t=T)}^{s(t=0)} \left| \frac{dc_1}{dt} \right| \left| \frac{dt}{ds} \right| ds &= \int_{s(t=T)}^{s(t=0)} \left| \frac{dc_1}{dt} \right| \left(-\frac{dt}{ds} \right) ds \\ &= - \int_{s(t=0)}^{s(t=T)} \left| \frac{dc_1}{dt} \right| \left(-\frac{dt}{ds} \right) ds \\ &= \int_{t=0}^{t=T} \left| \frac{dc_1}{dt} \right| dt \end{aligned}$$

□

5 (9/15): Invariants (Con't)

5.1 Arclength

Recall the setup:

$$\begin{array}{ccc} I \subseteq \mathbb{R} & \xrightarrow{c_1} & \mathbb{R}^n \\ t \uparrow & \nearrow c_2 & \\ J \subseteq \mathbb{R} & & \end{array}$$

This diagram summarizes what we want from a change of variables: the diagram above commutes, aka $c_1 \circ t = c_2$.

We proved last lecture that the length is an invariant that is preserved by a change of variables. There was some confusion regarding the case, where s flips the orientation of the interval. Since $s(0) > s(T)$, in the length integral we must first flip the limits of integration:

$$\ell(c_2) = \int_{s(T)}^{s(0)} \left| \frac{dc_1}{dt} \right| \left| \frac{dt}{ds} \right| ds$$

After replacing $\left| \frac{dt}{ds} \right| = -\frac{dt}{ds}$, we arrive at the same result as the first case.

Q: Why can't we have $\frac{dt}{ds} > 0$ on parts of the interval, and $\frac{dt}{ds} < 0$ elsewhere? That's because if that's the case, then there exists some point where $\frac{dt}{ds} = 0$ (direct consequence of the intermediate value theorem), which contradicts our definition for a change of variables (violates the regularity assumption).

Definition 5.1. When $\frac{dt}{ds} > 0$, we call such a change of variables orientation-preserving. When $\frac{dt}{ds} < 0$, it is orientation-reversing.

Remark. The objects dt, ds are called differential forms. We will see that the orientation of the change of variables is related to certain positivity property of these differential forms.

We have now seen that length is something that is preserved by change of variables. Now we can ask, is there a special parameterization that “encodes” this invariant? Could we choose a parameterization c on the interval $I = (0, s')$ such that s is the length of the curve from $c(0)$ to $c(s)$?

Yes! Looking at the length formula

$$\ell(c) = \int_{s_1}^{s_2} \left| \frac{dc}{ds} \right| ds$$

if we choose c such that $\left| \frac{dc}{ds} \right| = 1$ identically, this will give us a parameterization by arclength. Let's see this formally:

Lemma 5.2 (Parameterization by Arclength). Every regular curve can be parameterized (in an orientation-preserving manner) by its arclength, and such a parameterization has unit speed ($\left| \frac{dc}{ds} \right| = 1$).

Proof. We can prove this constructively. Suppose we have a regular curve parameterized by the function c_2 on the interval $J = (0, s)$. We want to use this to construct an arclength parameterization.

The arclength of the curve from the endpoints $c_2(0)$ and $c_2(s)$ is given by $\int_0^s \left| \frac{dc_2}{ds} \right| ds$. We can define the new parameterization

$$t(s) = \int_0^s \left| \frac{dc_2}{ds'} \right| ds'$$

Clearly, if this is a valid parameterization, then it is an arclength parameterization. However, this is only the case if $\left| \frac{dc_2}{ds} \right|$ is integrable over the entire interval. As an example, if $\left| \frac{dc_2}{ds} \right| = 1/s$ and the parameter interval is $(0, \infty)$, then the integral will not converge. Formally, we require that

$$\left| \frac{dc_2}{ds} \right| \in L^1(J \subseteq \mathbb{R})$$

the space of L^1 -integrable functions over the interval J .

Clarification: we are not saying that infinite-length curves are not allowed. The issue with the $1/s$ example is that the function $t(s)$ will be ill-defined for any $s > 0$, since the blowup happens at $s = 0$.

Now, we want to show that this parameterization has unit speed. Since $c_2 = c_1 \circ t$

$$\frac{dc_2}{ds} = \frac{dc_1}{dt} \frac{dt}{ds} \implies \left| \frac{dc_2(s)}{ds} \right| = \left| \frac{dc_1}{dt} \right| \frac{dt}{ds}$$

By the fundamental theorem of calculus, $\frac{dt}{ds} = \left| \frac{dc_2}{ds} \right|$. So, the equation becomes

$$\left| \frac{dc_2}{ds} \right| = \left| \frac{dc_1}{dt} \right| \frac{dc_2}{ds}$$

Since $\frac{dc_2}{ds} \neq 0$ by the regularity of c_2 , we can cancel both sides to get $\frac{dc_1}{dt} = 1$. □

Example 5.3. Let $c_2 : J \longrightarrow \mathbb{R}^3$ be given by $s \mapsto (as, bs, s)$. The speed is

$$\left| \frac{dc_2}{ds} \right| = \sqrt{a^2 + b^2 + 1} \geq 1$$

How do we convert this to an arc-length parameterization? Following the steps of the proof, we can define the change of variables

$$s \mapsto t(s), \quad t(s) = \int_0^s \left| \frac{dc_2}{ds'} \right| ds' = s\sqrt{a^2 + b^2 + 1}$$

Then, the map

$$t \mapsto c_2(t(s)) = \left(\frac{at}{\sqrt{a^2 + b^2 + 1}}, \frac{bt}{\sqrt{a^2 + b^2 + 1}}, \frac{t}{\sqrt{a^2 + b^2 + 1}} \right)$$

is an arclength parameterization.

Example 5.4. Let $c_2 : J \rightarrow \mathbb{R}^3$ be given by $s \mapsto (a \sin(\alpha s), a \cos(\alpha s), \alpha s)$. The derivative is

$$\left| \frac{dc_2}{ds} \right| = \sqrt{a^2 \alpha^2 + \alpha^2}$$

So we can define the change of variables

$$t(s) = \int_0^s \left| \frac{dc_2}{ds'} \right| ds' = s|\alpha|\sqrt{a^2 + 1}$$

5.2 Energy

Let us consider an integral curve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$, $t \mapsto c(t)$. The energy (or the total energy) E_c of this integral curve is defined as follows

$$E_c := \frac{1}{2} \int_I \left| \frac{dc}{dt} \right|^2 dt \quad (1)$$

Lemma 5.5. The energy of a curve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$, $t \mapsto c(t)$ is not invariant under reparametrization

Proof. Consider the following smooth reparametrization (diffeomorphism of the interval I) of the curve $c_1(t)$

$$s : I \subset \mathbb{R} \rightarrow J \subset \mathbb{R} \quad (2)$$

$$t \mapsto s(t) \quad (3)$$

The new curve is $c_2(s(t)) = c_1(t)$ and therefore by chain rule

$$\frac{dc_1}{dt} = \frac{dc_2}{ds} \frac{ds}{dt} \quad (4)$$

which yields

$$E_{c_1} := \frac{1}{2} \int_I \left| \frac{dc_1}{dt} \right|^2 dt = \frac{1}{2} \int_I \left| \frac{dc_2}{ds} \right|^2 \left| \frac{ds}{dt} \right|^2 dt = \frac{1}{2} \int_J \left| \frac{dc_2}{ds} \right|^2 \frac{ds}{dt} ds \neq \frac{1}{2} \int_J \left| \frac{dc_2}{ds} \right|^2 ds = E_{c_2} \quad (5)$$

□

Lemma 5.6. consider an integral curve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$, $t \mapsto c(t)$ that has a fixed length l_c . The energy E_c of the integral curve verifies the following

$$\lim_{|I| \rightarrow 0} E_c = \infty \quad (6)$$

Remark. Let f, g be two functions on $\mathbb{R}^n$ i.e., $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$. The Cauchy-Schwartz inequality reads

$$\int_{\mathbb{R}^n} fg d^n x \leq \left(\int_{\mathbb{R}^n} f^2 d^n x \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} g^2 d^n x \right)^{\frac{1}{2}} \quad (7)$$

which is a special case of the Hölder's inequality

$$\int_{\mathbb{R}^n} fg d^n x \leq \left(\int_{\mathbb{R}^n} |f|^p d^n x \right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} |g|^q d^n x \right)^{\frac{1}{q}}, \quad p^{-1} + q^{-1} = 1, \quad 0 < p, q < \infty \quad (8)$$

with $p = q = 2$. Notice that this is nothing but scaling.

Proof. Let us recall the expression for the length of a curve c

$$l_c := \int_I \left| \frac{dc}{dt} \right| dt. \quad (9)$$

Now use the Cauchy-Schwartz inequality on l_c after choosing $f = \left| \frac{dc}{dt} \right|$ and $g = 1$ to yield

$$l_c = \int_I \left| \frac{dc}{dt} \right| dt \leq \left(\int_I \left| \frac{dc}{dt} \right|^2 dt \right)^{\frac{1}{2}} \left(\int_I 1^2 dt \right)^{\frac{1}{2}} = (2E_c)^{\frac{1}{2}} |I|^{\frac{1}{2}} \quad (10)$$

and since both left and right-hand sides are positive, we can square it to obtain

$$l_c^2 \leq 2|I|E_c \implies E_c \geq \frac{l_c^2}{2|I|}. \quad (11)$$

since the length l_c is fixed we may write

$$\lim_{|I| \rightarrow 0} E_c \geq l_c^2 \lim_{|I| \rightarrow 0} \frac{1}{|I|} = \infty \quad (12)$$

□

Remark. The equality holds in Cauchy-Schwartz when f and g are proportional to each other (check it by substituting $f = \alpha g$ for a non-zero constant α). Therefore in the current example, if we have $\left| \frac{dc}{dt} \right| = 1$ (choosing the proportionality constant to be 1) i.e., make the curve parametrized by arclength, we obtain the equality. More precisely

$$l_c = \int_I dt = |I|, E_c = \frac{l_c^2}{2|I|} = \frac{|I|}{2}. \quad (13)$$

5.3 More Invariants

Let's imagine the path from the SC507 to Annenberg. What about the path is invariant? Certainly not the speed. Sometimes we run, sometimes we walk etc. If we think about it, the curvature of the path is an invariant. Another invariant that is a bit harder to state is the orientation of the frame: for instance, if you left e_1 be the tangent vector, e_2 be the vector pointing towards the left (along your outstretched left arm, for example), and e_3 be upwards. Then moving along the curve, the orientation of this frame is invariant. We will see next lecture how to define such a frame at every point along the curve and see that it is an invariant. (Spoiler: this involves Gram-Schmidt normalization)

Note that if we consider the frames along the same line, all the vectors remain parallel. However, along an arbitrary curve, even though the orientation is preserved, it still rotates in space. Why? Well, the straight line is straight, and the curved line is curved. Informally, if $\mathbf{e}$ is the frame, then " $\frac{d\mathbf{e}}{ds}$ " encodes information about the curvature of the curve.

6 (9/20): Invariants: Frames

6.1 Recap

Last class, we constructed a special parameterization called the arclength parameterization (see Lemma 5.2). From this point forward, we will only consider curves parameterized by their arclength. In other words, we demand $\left| \frac{dc}{dt} \right| = 1$. This turns out to be a convenient choice, as we will see.

Remark. The condition $\left| \frac{dc}{dt} \right| = 1$ comes automatically if we choose an arclength parameterization.

6.2 Frames

Let's begin with the usual setup. Let $c : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ be an (integral, arclength-parameterized) curve. Pictorially, the frames along the curve look like Figure 4.

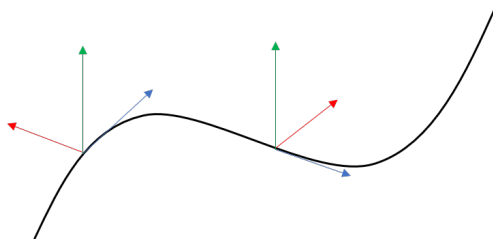


Figure 4: Two frames along some integral curve.

Let's try to arrive at a formal definition. We can start with $e_1(t)$, which is just tangent to the curve at each point. This is something we are familiar with: it is just the tangent vector field to the curve $c(t)$. In this scenario we don't need the whole tangent vector field, just the restriction of it to the integral curve. Next, we can consider $e_2(t), \dots, e_n(t)$. These too can be thought of as vector fields restricted to the curve.

We thus have the following definition:

Definition 6.1 (Frames). A moving frame is a collection of smooth maps

$$\begin{aligned} A = 1, 2, \dots, n : \quad e_A : I \subseteq \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto (e_A^1(t), \dots, e_A^n(t)) \end{aligned}$$

such that the inner product $\langle e_A(t), e_B(t) \rangle = \delta_{AB}$ (δ_{AB} is the Kronecker delta, which is 1 if $A = B$ and 0 otherwise)

Remark. This essentially is a moving orthonormal basis that everywhere on the curve looks the same to an observer on the curve.

Question 6.2. Isn't the frame generated by Gram-Schmidt not unique?

Answer. Yes, there is a lot of redundancy. We will later see that some interesting properties concerning frames are invariant under isometries, or rigid rotations/reflections.

We have one of the vectors for free: as we observed, we can just let $e_1(t)$ be the tangent vector field of $x(t)$ (why does this work? in particular, how do we know $\langle e_1(t), e_1(t) \rangle = 1$?). So we let $e_1(t) = (\dot{x}^1(t), \dots, \dot{x}^n(t))$. What happens if we take more derivatives? We can obtain the vectors

$$\begin{aligned} (new)_1 &= \left(\frac{d^2 x^1}{dt^2}, \dots, \frac{d^2 x^n}{dt^2} \right) \\ &\vdots \\ (new)_{n-1} &= \left(\frac{d^{n-1} x^1}{dt^{n-1}}, \dots, \frac{d^{n-1} x^n}{dt^{n-1}} \right) \end{aligned}$$

The question now is: can we use this as a basis of $\mathbb{R}^n$? Yes, if they are linearly independent.

Definition 6.3 (Frenet curves). A curve which has n linearly independent derivatives $\frac{dc}{dt}, \dots, \frac{d^n c}{dt^n}$ along the whole curve is called a Frenet curve.

Remark. Fun fact: in infinite dimensional spaces, a closed bounded set (such as the unit ball) is not compact.

Assuming that we have a Frenet curve, we can use Gram-Schmidt in order to compute a frame:

Proposition 6.4. On every Frenet curve, one can construct a moving frame.

Proof. Let $t \mapsto c(t) := (x^1(t), \dots, x^n(t))$ be a Frenet curve. By its definition, we can construct n independent vectors by taking derivatives:

$$\begin{aligned}\tilde{e}_1 &= \frac{dc}{dt} = \left(\frac{dx^1}{dt}, \dots, \right) \\ \tilde{e}_2 &= \frac{d^2c}{dt^2} = \left(\frac{d^2x^1}{dt^2}, \dots, \right) \\ &\vdots \\ \tilde{e}_n &= \frac{d^nc}{dt^n} = \left(\frac{d^nx^1}{dt^n}, \dots, \right)\end{aligned}$$

Since c was chosen to be an arclength parameterization, we have $\|\tilde{e}_1\| = 1$, so we can set $e_1 = \tilde{e}_1$.

We can construct the rest by Gram-Schmidt. Let's consider how to get e_2 . Given the vector $\tilde{e}_2$ that is different from e_1 , we can calculate the projection onto e_1 to be $\langle \tilde{e}_2, e_1 \rangle e_1$. Subtracting this from $\tilde{e}_2$ gives a vector $\tilde{\tilde{e}}_2 = \tilde{e}_2 - \langle \tilde{e}_2, e_1 \rangle e_1$ that is orthogonal to e_1 (you can check this by computing their inner product to be 0). However, this doesn't mean that the norm of $\tilde{\tilde{e}}_2$ is 1. To fix that, we can simply normalize this vector and define $e_2 = \frac{\tilde{\tilde{e}}_2}{\|\tilde{\tilde{e}}_2\|}$. To compute the rest of the orthonormal basis, we simply extend the Gram-Schmidt process. For example,

$$\tilde{\tilde{\tilde{e}}}_3 = \tilde{e}_3 - \langle \tilde{e}_3, e_1 \rangle e_1 - \langle \tilde{e}_3, e_2 \rangle e_2, \quad e_3 = \frac{\tilde{\tilde{\tilde{e}}}_3}{\|\tilde{\tilde{\tilde{e}}}_3\|}$$

and in general,

$$\tilde{\tilde{\tilde{\tilde{e}}}}_k = \tilde{e}_k - \sum_{i=1}^{k-1} \langle \tilde{e}_k, e_i \rangle e_i, \quad e_k = \frac{\tilde{\tilde{\tilde{\tilde{e}}}}_k}{\|\tilde{\tilde{\tilde{\tilde{e}}}}_k\|}$$

□

Question 6.5. Is the converse true? Given a moving frame, can we conclude that the underlying curve is Frenet?

Answer. Not necessarily. Consider a moving frame going in the straight line $t \mapsto (t/\sqrt{3}, t/\sqrt{3}, t/\sqrt{3})$. In this case, the underlying curve is a straight line, whose 2 and 3 derivatives are all 0. But we can under "certain" restriction fit a "curve" through a moving frame (we'll shortly see what these are).

Question 6.6. Isn't it possible to always construct a moving frame, if we let $\{e_i\}$ just be the standard basis on $\mathbb{R}^n$?

Answer. Yes, but this frame is useless since it doesn't encode any information about the curve. By carefully constructing moving frames, we can use these frames to understand invariant properties of the curve, like curvature.

6.3 Change of Frame Along the Curve

Consider the frame $\{e_A(t)\}_{A=1,\dots,n}$. How do we understand how the frame changes? Our intuition demands that we take derivatives, and our intuition is correct.

Teaser for next class:

Proposition 6.7. Assuming that $\{e_A(t)\}_{A=1,\dots,n}$ is the moving frame constructed as in the previous section for a Frenet curve.

$$\begin{aligned}\frac{dc}{dt} &= \sum_{i=1}^n \alpha_i e_i \\ \frac{de_A}{dt} &= \sum_{B=1}^n \omega_{AB} e_B\end{aligned}$$

where $\omega_{AB} = -\omega_{BA}$ and $\omega_{AB} = 0$ for all $B > A + 1$.

7 (9/22): Frames (Con't)

Recall that frames are simply a set of vector fields that together assign an orthonormal basis at each point along the curve. One type of frame we are particularly interested in is a frame where the first vector field $e_1(t)$ is set to be the velocity of the curve (arclength parameterized)

For a class of curves c known as Frenet curves, the n time derivatives $\{\frac{d^k c}{dt^k}\}_{k \in [n]}$ are linearly independent. Using Gram-Schmidt, we can turn these derivatives into a moving frame of the curve.

7.1 Derivative of a Frame

Proposition 7.1. Let $c : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ be a Frenet curve, (and the e_I be the moving frame constructed from last time). Then the following holds:

$$\begin{aligned}\frac{dc}{dt} &= \sum_{I=1}^n \alpha_I e_I \\ \frac{de_I}{dt} &= \sum_{J=1}^n \omega_{IJ} e_J,\end{aligned}$$

where $\omega_{IJ} = -\omega_{JI}$ and $\omega_{IJ} = 0$ for all $J > I + 1$ and α_I are to be determined.

Proof. From our moving frame construction, $\alpha_1 = 1$ and $\alpha_{i>1} = 0$, since we chose $e_1 = \frac{dc}{dt}$.

Linear algebra fact: every vector (in a finite-dimensional space) can be written as a linear combination of the basis vectors. Also, derivatives of vectors are just vectors. Since $\{e_I\}$ are a basis, we can always write

$$\frac{de_I}{dt} = \sum_{J=1}^n \beta_{IJ} e_J$$

How do we determine the coefficients β_{IJ} ? We can take the inner product of the LHS with each of the basis vectors!

Linear algebra recap: the inner product is bilinear: aka for scalars a, b and vectors u, v, w , we have $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$, and analogously for the second argument.

So choose an index $k \in [n]$, we can take the dot product of both sides with e_k to get:

$$\begin{aligned} \left\langle \frac{de_I}{dt}, e_k \right\rangle &= \left\langle \sum_{J=1}^n \beta_{IJ} e_J, e_k \right\rangle \\ &= \sum_{J=1}^n \beta_{IJ} \langle e_J, e_k \rangle \\ &= \beta_{Ik} \end{aligned} \quad (\text{linearity})$$

So we can rewrite original vector, replacing the β_J 's with this new expression:

$$\frac{de_I}{dt} = \sum_{J=1}^n \left\langle \frac{de_I}{dt}, e_J \right\rangle e_J$$

Let us define $\omega_{IJ} := \left\langle \frac{de_I}{dt}, e_J \right\rangle$. We shall prove that it satisfies the properties stated in the proposition.

$\omega_{IJ} = -\omega_{JI}$: Recall that $\langle e_I, e_J \rangle = \delta_{IJ}$, which is constant for any choice of indices. Therefore its derivative is 0. We can use the bilinearity of the dot product to obtain

$$\frac{d}{dt} \langle e_I, e_J \rangle = \left\langle \frac{de_I}{dt}, e_J \right\rangle + \left\langle e_I, \frac{de_J}{dt} \right\rangle = \omega_{IJ} + \omega_{JI} = 0$$

Moving the second term to the RHS gives us exactly what we want.

Remark. The above derivation involving the dot product is not guaranteed to work if we are not in Euclidean $\mathbb{R}^n$. Fortunately, that is where we are right now.

$\omega_{IJ} = 0, \quad \forall J > I + 1$: We use the shorthand $c^I = \frac{d^I c}{dt^I}$. Recall the construction of our frame:

$$e_I = \frac{c^I - \sum_{J=1}^{I-1} \langle c^I, e_J \rangle e_J}{\|c^I - \sum_{J=1}^{I-1} \langle c^I, e_J \rangle e_J\|} \quad (*)$$

We can rearrange this to get something of the form $c^I = \sum_{J=1}^I \alpha^J e_J$, which in particular implies $c^I \in \text{span}\{e_1, \dots, e_I\}$.

At the same time, using an inductive argument, we have that $e_I \in \text{span}(c^1, \dots, c^I)$ (Assume that this is true for indices up to $I - 1$. Then we can substitute this into $(*)$ to express e_I as a linear combination of $c^1, \dots, c^I$). Taking one more derivative, we have

$$\frac{de_I}{dt} \in \text{span}\{c^1, \dots, c^I, c^{I+1}\}$$

Since $c^{I+1} \in \text{span}(e_1, \dots, e_{I+1})$, this implies $\frac{de_I}{dt} \in \text{span}(e_1, \dots, e_{I+1})$. Taking the dot products with the basis vectors, we have $\omega_{IJ} = 0$ if $J > I + 1$. $\square$

Here is a visualization of the coefficient matrix (ω_{IJ}) :

$$(\omega_{IJ}) = \begin{pmatrix} 0 & \omega_{1,2} & 0 & \cdots & 0 \\ -\omega_{1,2} & 0 & \omega_{2,3} & \cdots & 0 \\ 0 & -\omega_{2,3} & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & -\omega_{n-1,n} & 0 \end{pmatrix}$$

Remark. Note that all of these coefficients depend on t : really, $\omega_{IJ}(t) = \left\langle \frac{de_I(t)}{dt}, e_J(t) \right\rangle$.

7.2 Curvature

So recall why we went through all this trouble to work with frames: we wanted to define curvature.

Definition 7.2 (Curvature). The curvature of a curve $c : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ is defined as

$$\kappa_I = \frac{\omega_{I,I+1}}{|dc/dt|} = \omega_{I,I+1} \text{ (for an arclength param)}$$

κ_I is the curvature in the direction of e_{I+1} .

Why are there different curvatures in different directions? You can imagine that viewing a curve from different points of view, you would see the curve bending in a different way.

8 (9/27): Curvature

8.1 Wrapping up curvature

To clarify, we are restricting once again to considering only Frenet curves and the associated construction of the moving frame.

Definition 8.1 (Curvature). Let $t \mapsto c(t)$ be a Frenet curve. The curvature functions are defined by

$$\kappa_I(t) = \frac{\omega_{I,I+1}}{|dc/dt|}, \quad 1 \leq I \leq n-1$$

Remark. In the coefficient matrix for (ω_{ij}) for a curve parameterized by arclength, we can substitute k_I for $\omega_{I,I+1}$ to get the following coefficient matrix:

$$K := \begin{pmatrix} 0 & k_1 & 0 & \cdots \\ -k_1 & 0 & k_2 & \cdots \\ 0 & -k_2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Question 8.2. Can we build a curve given the information of its curvature? i.e. given the functions $t \mapsto (k_1(t), k_2(t), \dots, k_{n-1}(t))$ on the interval $t \in [0, 1]$, can we fit a curve with those curvatures?

Intuitively, if we are given a rope curved in some form, and then we are asked to take another rope and make it of the same shape, what would you do? One could imagine taking the start of your rope, putting it next to the rope, and then start “matching” your rope to the reference rope.

This intuition can extend to the mathematical case. We first have to be given an initial condition for the start position of the curve. Once we do that, we can fit the rest of the curve by solving the differential equations of the frame!

But another question arises: can we always solve the system of frame equations? Well, we have seen this type of equations before in integral curves, in particular the definition of the tangent vector field! However there is a catch here: the coefficients ω_{IJ} do not have to be constant with time. Another difference we have: how do we know that the frame vectors remain orthonormal throughout?

8.2 Going from Curvature back to the Curve

Theorem 8.3. Let $t \mapsto (k_1(t), k_2(t), \dots, k_{n-1}(t))$ be a smooth function $\mathbb{R} \rightarrow \mathbb{R}^{n-1}$. Then there exists a curve $t \mapsto c(t) \subset \mathbb{R}^n$ for the parameter interval $t \in [0, t^*)$, $t^* > 0$, such that the functions $k_1(t), \dots, k_{n-1}(t)$

are its curvatures. In other words, the system of ODEs

$$\frac{de_I}{dt} = \sum_{J=1}^n \omega_{I,J} e_J$$

with initial condition $e_I(0)$ has a solution for $t \in [0, t^*)$. Furthermore the basis $\{e_I(t)\}$ remains orthonormal for all $t \in [0, t^*)$.

Proof. Consider a canonical basis $\{\hat{1}, \hat{2}, \dots, \hat{n}\} \subset \mathbb{R}^n$, where

$$\hat{1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \hat{2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \hat{n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Then we can write each frame vector as a linear combination of this canonical basis:

$$\begin{aligned} e_1(t) &= e_1^1(t) \hat{1} + e_1^2(t) \hat{2} + \dots + e_1^n(t) \hat{n} \\ &\vdots \\ e_n(t) &= e_n^1(t) \hat{1} + e_n^2(t) \hat{2} + \dots + e_n^n(t) \hat{n} \end{aligned}$$

Recall we have the system of ODEs:

$$\begin{aligned} \frac{de_1}{dt} &= \omega_{1,1} e_1 + \omega_{1,2} e_2 + \dots + \omega_{1,n} e_n \\ &\vdots \\ \frac{de_n}{dt} &= \omega_{n,1} e_1 + \omega_{n,2} e_2 + \dots + \omega_{n,n} e_n \end{aligned}$$

Just like in the bonus of the first HW, we can rewrite this in matrix form:

$$\frac{d}{dt} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 & \omega_{1,2} & \dots & 0 \\ -\omega_{1,2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

We can expand the left hand side by writing each frame vector in its components:

$$E := \frac{d}{dt} \begin{pmatrix} e_1^1 & e_1^2 & \dots & e_1^n \\ e_2^1 & e_2^2 & \dots & e_2^n \\ \vdots & \vdots & \ddots & \vdots \\ e_n^1 & e_n^2 & \dots & e_n^n \end{pmatrix}$$

Substituting this and noting that each ω is the corresponding curvature multiplied by $|\dot{c}|$, we have the system

$$\frac{dE}{dt} = |\dot{c}| K E$$

along with the initial condition $E(0)$. Then, we have the first part of the result:

1. There exists a solution $t \mapsto e_I(t)$ in the interval $t \in [0, t^*)$, $t^* > 0$ from ODE theory (local well-foundedness theorem)

Intuition: for a short enough time, K is constant, and so the above system of equations have a canonical exponential solution. However, this niceness is not guaranteed to hold for long times.

Question 8.4. Why must we work in the canonical basis rather than the frame basis itself?

Answer. We don't actually know that the frame vectors actually form a basis for time > 0 . That is what we are trying to prove!

With that done, we still need to show that the solutions $e_I(t)$ form an orthonormal set for all $t \in [0, t^*)$. This is equivalent to showing that each pair of rows of the matrix E are orthonormal to each other. Such a matrix is known as an orthogonal matrix.

A well-known result from linear algebra says that E consisting of orthonormal rows is equivalent to the statement that $E^T E = \text{id}$ (write out small matrices and try it out if this fact is unfamiliar to you). So we want to show that

$$E^T(t)E(t) = E^T(t)E(t) = \text{id}, \quad \forall t \in [0, t^*)$$

We know that $E^T(0)E(0) = \text{id}$. How do we show that this remains true for all subsequent times? We can take the derivative! Just like in HW1 where we showed that the energy was constant, if the derivative vanishes then $E^T(t)E(t)$ will remain equal to id for all time. So, dropping the explicit t dependence,

$$\begin{aligned} \frac{d}{dt}(E^T E) &= \frac{dE^T}{dt} E + E^T \frac{dE}{dt} \\ &= (|\dot{c}| E^T K^T) E + E^T (|\dot{c}| K E) \\ &= |\dot{c}| (E^T (K^T + K) E) \\ &= |\dot{c}| (E^T (-K + K) E) && (\text{since } K^T = -K) \\ &= 0 \end{aligned}$$

□

9 (9/28): Curvature, Pt. 2: The case of $\mathbb{R}^2$

Over the past few lectures we set up the machinery for dealing with curvature of a curve in a general space $\mathbb{R}^n$. The major result from Tuesday: given a set of functions $(k_1(t), \dots, k_{n-1}(t))$ describing the curvature and an initial condition $E(0)$, we can always fit a curve for time $t \in [0, t^*)$ such that the moving frame $E(t)$ satisfies the differential equation $\frac{dE}{dt} = |\dot{c}| K E$. Today we will restrict our attention to $n = 3$ to get some concrete results.

9.1 Curvature in $\mathbb{R}^2$

Recall the general setup of the moving frame. We assume an arc-length parameterization of the curve. In 2 dimensions the frame is described by $E(t) = (e_1(t), e_2(t))$. The matrix ω_{IJ} is given simply by

$$\omega_{IJ} = \begin{pmatrix} 0 & \omega_{1,2} \\ -\omega_{1,2} & 0 \end{pmatrix}$$

Recalling our definition of $k_i = \omega_{i,j}$, this means a curve in 2 dimensions only has 1 curvature. This should make intuitive sense, because in 2 dimensions the only information we need is the magnitude of the curvature (for eg how concave/convex it is). And the general frame equation $\frac{dE}{dt} = |\dot{c}|KE$ reduces to:

$$\frac{d}{dt} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 0 & \omega_{1,2} \\ -\omega_{1,2} & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \implies \frac{de_1}{dt} = \omega_{1,2}e_2, \quad \frac{de_2}{dt} = -\omega_{1,2}e_1$$

In order to understand the rotation of the frame, we need to choose some external reference frame (it would make no sense to consider the rotation of the frame with respect to itself - a frame is always stationary from its own point of view).

We can take the canonical basis on $\mathbb{R}^2$ as a reference. But the canonical basis is for a vector space centered at the origin, whereas the origin of the frame moves along the curve. So as vector spaces they are not directly commensurable. However, since we are in $\mathbb{R}^2$ (in particular a flat euclidean space), we can perform parallel transport of our canonical basis so that their origin coincides with that of the moving frame. Explicitly, we will perform the following steps:

1. Choose a fixed basis (v_1, v_2) , for example the standard basis $\{(0, 1), (1, 0)\}$.
2. Translate the basis along the curve in such a way that it remains the same (parallel to its initial state). See Figure 5 for an illustration.
3. Compute the angle $\theta(t)$ between the chosen basis and the frame.

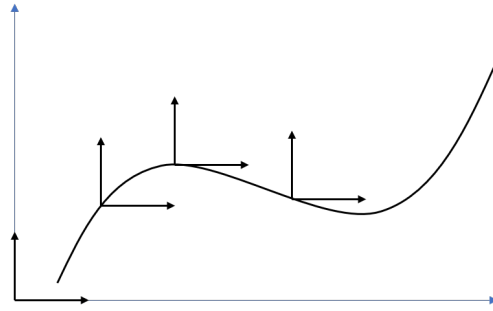


Figure 5: The standard basis and some of its parallel transports

To compute the angle between the reference basis and the frame, we can just compute dot products:

$$\begin{aligned} \cos \theta(t) &= \frac{\langle v_1, e_1(t) \rangle}{|v_1||e_1(t)|} = \frac{\langle v_1, e_1(t) \rangle}{|v_1|} \\ \sin \theta(t) &= \dots = \frac{\langle v_1, e_2(t) \rangle}{|v_1|} \end{aligned}$$

Question 9.1. Why do we not just choose a reference basis with unit length?

Answer. We certainly could!

So what do we do in differential geometry? We take derivatives!

$$\begin{aligned} \frac{d}{dt} \cos \theta(t) &= -\sin \theta(t) \frac{d\theta(t)}{dt} = \frac{\langle v_1, \frac{d}{dt} e_1(t) \rangle}{|v_1|} = \left\langle \frac{v_1}{|v_1|}, \frac{d}{dt} e_1(t) \right\rangle \\ \cos \theta(t) \frac{d\theta(t)}{dt} &= \dots = \left\langle \frac{v_1}{|v_1|}, \frac{d}{dt} e_2(t) \right\rangle \end{aligned}$$

Here we used the fact that the differential operator commutes with the dot product in Euclidean space:

$$\frac{d}{dt}\langle \cdot, \cdot \rangle = \left\langle \frac{d}{dt} \cdot, \cdot \right\rangle + \left\langle \cdot, \frac{d}{dt} \cdot \right\rangle$$

Back to the derivation. We can substitute $\frac{de_1(t)}{dt} = \omega_{1,2}$ and likewise to get

$$\begin{aligned} -\sin \theta(t) \frac{d\theta(t)}{dt} &= \left\langle \frac{v_1}{|v_1|}, \omega_{1,2} e_2(t) \right\rangle = \left\langle \frac{v_1}{|v_1|}, e_2(t) \right\rangle \omega_{1,2} \\ \cos \theta(t) \frac{d\theta(t)}{dt} &= \left\langle \frac{v_1}{|v_1|}, -\omega_{1,2} e_1(t) \right\rangle = -\left\langle \frac{v_1}{|v_1|}, e_1(t) \right\rangle \omega_{1,2} \end{aligned}$$

But recall our definitions for $\cos \theta(t)$ and $\sin \theta(t)$. Replacing the dot products with the trig terms gives us

$$\begin{aligned} -\sin \theta(t) \frac{d\theta(t)}{dt} &= \sin \theta(t) \omega_{1,2} \\ -\cos \theta(t) \frac{d\theta(t)}{dt} &= \cos \theta(t) \omega_{1,2} \end{aligned}$$

Cancelling out the trig factor in either term leads to the conclusion that $\frac{d\theta(t)}{dt} = \omega_{1,2}$. The only potential issue is when $\sin \theta(t) = \cos \theta(t) = 0$ simultaneously, because in that case we would have division by 0 for both equations. Fortunately, $\sin \theta(t)$ and $\cos \theta(t)$ cannot vanish at the same time. Therefore, we are guaranteed that the curvature satisfies

$$\frac{d\theta(t)}{dt} = -\omega_{1,2}$$

Theorem 9.2 (Frame rotation in 2D). The rate of rotation of the moving frame of a curve $c : \mathbb{R} \rightarrow \mathbb{R}^2$ is exactly the curvature.

Remark. Note that $\theta(t), \cdot \theta(t)$ are both C^∞ by assumption that $\omega_{1,2}(t)$ is C^∞ .

Example 9.3. Let $\omega_{1,2} = 1$. Then $\frac{d\theta}{dt} = -1 \implies \theta(t) = -t + \theta(0)$

Question 9.4. What changes when we start considering plane curves (curves that lie entirely on a plane) embedded in $\mathbb{R}^n$?

Answer. For a plane curve in particular, the frame can be constructed such that only the first two vectors are in the plane, and so these are the ones that rotate. The rest of the frame vectors are orthogonal to the plane, and do not change in time. So if we reformulate our theorem with more specificity it also applies in this place. However we will need a more general theorem if we are not considering plane curves.

Question 9.5. So this analysis doesn't say anything about where the curve is located, just its shape?

Answer. Yes, in $\mathbb{R}^n$ that is okay because of parallel transport.

9.2 Curves of Constant, Non-zero Curvature in $\mathbb{R}^2$

We will see that in $\mathbb{R}^2$, a curve has constant non-zero curvature iff it is a circle. (Technically the case where $k = 0$ can be included in the statement if we consider a straight line as a degenerate circle with infinite radius.)

If the curve has constant curvature, then $\omega_{1,2}(t) = \text{const.} = \frac{\varepsilon}{r}$ where $\varepsilon = \pm 1$ and $r > 0$. Let us compute the curve that it defines. Theorem 8.3 suggests we can try to fit a curve to this curvature function by writing

the frame equations.

$$\begin{aligned}\frac{dc}{dt} &= e_1 \\ \frac{de_1}{dt} &= \frac{\varepsilon}{r} e_2 \\ \frac{de_2}{dt} &= -\frac{\varepsilon}{r} e_1\end{aligned}$$

With the above notation, we can write the proposition proper:

Proposition 9.6. Let $c : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^2$ be a Frenet curve with constant curvature with magnitude $\frac{1}{r}$. Then $c(t)$ is a parameterization of a circle. i.e., there exists a constant vector y_0 such that

$$\|c(t) - y_0\| = r = \sqrt{(x^1(t) - y_0^1(t))^2 + (x^2(t) - y_0^2(t))^2}$$

Proof. Consider the vector $c(t) + \varepsilon r e_2$. Taking its derivative, we have

$$\frac{d}{dt}(c(t) + \varepsilon r \frac{de_2}{dt}) = \frac{dc}{dt} + \varepsilon r \frac{-\varepsilon e_1}{r} = e_1 - \varepsilon^2 e_1 = 0$$

This implies $c(t) + \varepsilon r e_2 = y_0$ for some constant vector $y_0 \in \mathbb{R}^2$. So we have $c(t) - y_0 = -\varepsilon r e_2$. Taking magnitudes, we have

$$\|c(t) - y_0\| = |-\varepsilon| r \|e_2\| = r$$

and we are done. □

Question 9.7. Isn't a hyperbola a surface of negative curvature?

Answer. That refers to the Gaussian curvature (intrinsic) of hyperboloids (dimension greater than or equal to 2). In a plane, the curvature (that we have been dealing with) of a hyperbola is not constant - as we travel towards the asymptote, the curvature tends towards 0.

Random fact about Kerr black holes: one can travel backwards in time by traveling in circles around the origin.

Next class we will look at curves in 3D, where we will start to see how to generalize the methods today to compute the rate of change of the moving frame.

10 (10/4): Some logistics, Curves in $\mathbb{R}^3$

10.1 Midterm Format

Around 3 weeks from now. Exact date to be confirmed. Open notes.

Questions will be similar to homework, not going to be calculation heavy Eg. Modified potential, deformation of the disk in the phase plots.

Practice questions to be released around 1.5 weeks before midterm.

Extra points from the homework can be used to compensate for other homeworks/exam.

10.2 Space Curves (Curves in $\mathbb{R}^3$)

Recall the familiar setup: We have a Frenet curve $c : I \rightarrow \mathbb{R}^n$ given by $c(t) = (x^1(t), \dots, x^n(t))$, its moving frame $\{e_I\}_{I \in 1, \dots, n}$, and the frame equations

$$\frac{de_I(t)}{dt} = \sum_{J=1}^n \omega_{I,J}(t) e_J(t)$$

satisfying $\omega_{I,J} = -\omega_{J,I}$, $\omega_{I,I} = 0$ for $J > I + 1$, and $e_1(t) = \frac{dc}{dt}$ has unit length.

Today we will specialize to the case $n = 3$, one dimension up from plane curves, and see what we can learn.

In 3 dimensions, the frame equations are

$$\frac{de}{dt} = \begin{pmatrix} 0 & k_1(t) & 0 \\ -k_1(t) & 0 & k_2(t) \\ 0 & -k_2(t) & 0 \end{pmatrix} e$$

Let's consider a special case: what if the curve lies entirely on the plane $x_1 - x_2$? We would expect that the frame vector orthogonal to the $x_1 - x_2$ plane does not change. How does this affect the matrix ω ? We must have $k_2(t) = 0$, resulting in

$$\omega = \begin{pmatrix} 0 & k_1(t) & 0 \\ -k_1(t) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We see that the matrix reduces to that of a plane curve!

Based on this, what is the role of $k_2(t)$? Intuitively, we would expect $k_2(t)$ to measure the failure of the curve to be a plane curve. One could think of it as related to the “torque” that lifts a curve off of the plane it was on. You can treat it analogous to the role of $k_1(t)$ measuring the failure of the curve to be a straight line.

For $n = 3$, the two k -terms have special names:

- k_1 : curvature
- k_2 : torsion

Remark. What happens if we choose a different sort of orthonormal frame (eg by using Gram-Schmidt in a different order)? We will see that even though the individual curvatures will change, the total curvature (which we will define soon) is invariant.

10.3 Representing the curve locally

We continue with the notation as before. Pick some time t_0 . Our goal is to represent the curve locally around the time $t = t_0$. We can do that by Taylor expanding:

$$c(t) = c(t_0) + \frac{dc}{dt}(t_0)(t - t_0) + \frac{d^2c}{dt^2}(t_0)\frac{(t - t_0)^2}{2} + \frac{d^3c}{dt^3}(t_0)\frac{(t - t_0)^3}{6} + O((t - t_0)^4)$$

But what is a possible danger? Not all smooth functions are analytic.

Definition 10.1. An analytic function f is a smooth function such that its Taylor series has a non-zero radius of convergence.

So let's impose the additional assumption that $c(t)$ is analytic. This will allow us to work with the Taylor series in some small neighborhood of t_0 . Now, how do we relate this back to the curvatures? Recall that the frame vectors were constructed by applying Gram-Schmidt to the derivatives.

We have the frame equations

$$\begin{aligned}\frac{dc}{dt} &= e_1(t) \\ \frac{de_1}{dt} &= k_1(t)e_2(t) \\ \frac{de_2}{dt} &= -k_1(t)e_2(t) + k_2(t)e_3(t) \\ \frac{de_3}{dt} &= -k_2(t)e_3(t)\end{aligned}$$

Let's recover the derivatives from here. We already have the first derivative. Taking derivatives, we have

$$\begin{aligned}\frac{d^2c}{dt^2} &= \frac{de_1(t)}{dt} = k_1(t)e_2(t) \\ \frac{d^3c}{dt^3} &= \frac{dk_1}{dt}e_2(t) + k_1(t)\frac{de_2}{dt} = \frac{dk_1}{dt}e_2(t) + k_1(t)(-k_1(t)e_2(t) + k_2(t)e_3(t))\end{aligned}$$

So we can just substitute this into the Taylor series

$$c(t) = c(t_0) + (t - t_0)e_1(t_0) + \frac{1}{2}k_1(t_0)e_2(t_0)(t - t_0)^2 + \frac{1}{6}\left(-\frac{dk_1}{dt}(t_0)e_2(t_0) + k_1(t_0)(-k_1(t_0)e_2(t_0) + k_2(t_0)e_3(t_0))\right)$$

Collecting terms by the frame vectors, we have the following proposition:

Proposition 10.2 (Normal form of a curve). Every analytic Frenet curve can be represented as follows:

$$\begin{aligned}c(t) &= c(t_0) + \left((t - t_0) - \frac{k_1(t_0)^2}{6}(t - t_0)^3\right)e_1(t_0) + \left(\frac{k_1(t_0)}{2}(t - t_0)^2 - \frac{\dot{k}_1(t_0)}{6}(t - t_0)^3\right)e_2(t_0) \\ &\quad + \left(\frac{k_1(t_0)k_2(t_0)}{6}(t - t_0)^3\right)e_3(t_0) + O(|t - t_0|^4)\end{aligned}$$

in some neighborhood of $c(t_0)$.

10.4 Projections onto planes in the frame

Let us now consider what happens if we project onto the $e_1 - e_2$, $e_1 - e_3$ and $e_2 - e_3$ planes? (Figure 6)

$e_1 - e_2$ plane: Locally looks like a graph of $y = x^2$

$e_1 - e_3$ plane: Locally looks like a graph of $y = x^3$

$e_2 - e_3$ plane: Locally looks like $y^2 = x^3$ (the graph has a kink at $c(t_0)$).

The planes are called the osculating, rectifying, and normal planes respectively.

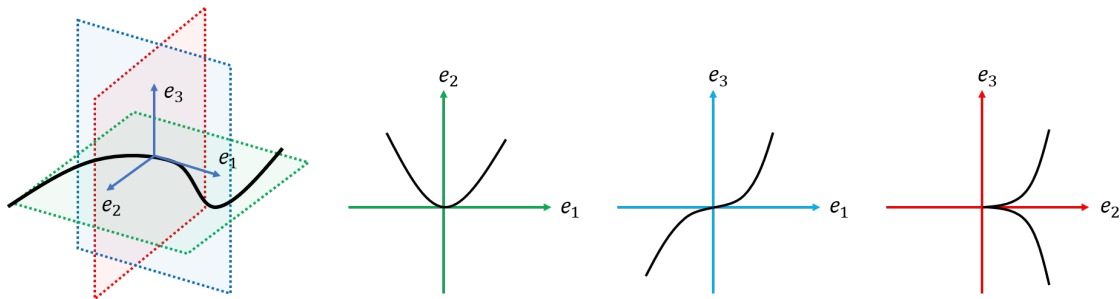


Figure 6: (Left) A space curve, its frame, and the 3 planes (green = $e_1 - e_2$, blue = $e_1 - e_3$, red = $e_2 - e_3$) (Right): Projections onto each of the planes (osculating, rectifying, and normal) in a small neighborhood of $c(t_0)$

11 (10/6): Wrapping up Curvature; Transformations

11.1 More on the Taylor Expansion

We continue exploring Proposition 10.2. Recall the Taylor expansion

$$c(t) = c(t_0) + \left((t - t_0) - \frac{k_1(t_0)^2}{6}(t - t_0)^3 \right) e_1(t_0) + \left(\frac{k_1(t_0)}{2}(t - t_0)^2 - \frac{\dot{k}_1(t_0)}{6}(t - t_0)^3 \right) e_2(t_0) + \left(\frac{k_1(t_0)k_2(t_0)}{6}(t - t_0)^3 \right) e_3(t_0) + O(|t - t_0|^4)$$

as well as the 3 projections in Figure 6. How do we understand the shape of each of these graphs?

$e_1 - e_2$: By making this projection we are setting the e_3 component to be 0. Since we are looking in a small neighborhood of $c(t_0)$, for each remaining component e_1, e_2 , we can keep only the leading order term. Our expansion thus becomes

$$c(t) \approx e_1(t_0)(t - t_0) + e_2(t_0)(t - t_0)^2 \frac{k_1(t_0)}{2} + \dots$$

From this, letting y be the e_2 component and x be the e_1 component, then $y \propto x^2$. This explains the parabolic shape of the projection onto the osculating plane.

Question 11.1. What is the difference between smooth and analytic functions?

Answer. A smooth function f is a function which is infinitely differentiable. I.e. $\frac{d^k f}{dx^k}$ is continuous for all $k \in \mathbb{N}$. An analytic function is a smooth function f such that its Taylor series $T(f)$ converges to f over some neighborhood.

An example of a smooth but non-analytic function $f(x) = e^{-1/x}$. We can compute its first couple of derivatives

$$f'(x) = \frac{1}{x^2} e^{-1/x}, \quad f''(x) = -\frac{x}{x^3} e^{-1/x} + \frac{1}{x^4} e^{-1/x}$$

and so on. Now let's look at the Taylor expansion around $x = 0$. Using the fact that $\lim_{x \rightarrow 0} \frac{1}{x^n} e^{-1/x} = 0$, we see that the in the Taylor expansion

$$f(x) = f(0) + f'(0)x + f''(0)x^2 + \dots$$

each term in the right hand side just evaluates to 0, and so the Taylor expansion converges to zero polynomial. But $e^{-1/x}$ is clearly not the zero polynomial.

Back to the osculating plane: it is the plane that “best approximates” the space curve, in the sense that the projection onto the osculating plane looks the most similar to the space curve.

$e_1 - e_3$: Dropping the e_2 term and keeping only terms up to quadratic, we have

$$c(t) \approx c(t_0) + (t - t_0)e_1(t_0) + O((t - t_0)^3)$$

Question 11.2. Why is the normal plane not very useful?

Answer. The projection onto the normal plane has a cusp. So we lose the smoothness property of our curve.

Question 11.3. What is the motivation for studying these kinds of projections?

Answer. We can use it in a number of ways. Projecting onto the osculating plane, we can determine if the space curve is really a plane curve by checking if the projection is the same as the original. Another thing we can do is to calculate the quadratic term of the projection onto the osculating plane to determine the k_1 curvature term.

11.2 Transformations

Given a curve $c : \mathbb{R} \rightarrow \mathbb{R}^n$, we can calculate its length, curvature, and frame. What happens if the curve is rigidly translated in space? Nothing will change.

What happens if the curve is rigidly rotated around some point? The length and curvature do not change. The frame changes by a rotation.

However, what happens if we take some section of the curve and deform it? Then none of the properties will remain the same.

What is an equivalent transformation that has the same effect as translating the curve? We can equally shift the axes of the graph to obtain the same effect! Likewise, to rotate a curve, we can equally rotate the axes instead. In other words, you can either

1. Rotate or translate or otherwise change the curve, or
2. You can transform the underlying space

We will focus on the second type. But how does one transform the underlying space?

11.3 Topology Fundamentals

Let’s consider translating all of the points in $\mathbb{R}^2$ which are ℓ units in the positive y direction. Instead of moving “all” the points, what is a simpler way? One observation is that we do not need to translate all the points “at the same time”. So rather than translating the whole space, we can choose some covering of the space by subsets, translating each piece individually, and then taking the union.

In other words, we want to perform the mathematical operations (corresponding to translation, rotation, etc) on subsets of the whole space. Afterwards, we can “glue” the results together, across their intersections, to get the transformed space.

But we need to make precise the notion of “gluing”.

What is the simplest subset in the plane? Circular disks! (Squares are almost good enough, but at the corners the square is not smooth)

Definition 11.4. A ball $B_r(x_0)$ with radius r centered at x_0 in $\mathbb{R}^n$ is defined as the subset

$$B_r(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| \leq r\}$$

The ball can be further split into two subsets: the interior $B_r^\circ(x_0)$ and the boundary $\partial B_r(x_0)$, defined by

$$\begin{aligned} B_r^\circ(x_0) &:= \{x \in \mathbb{R}^n : \|x - x_0\| < r\} \\ \partial B_r(x_0) &:= \{x \in \mathbb{R}^n : \|x - x_0\| = r\} \cong S^{n-1} \end{aligned}$$

where S^{n-1} is the $n - 1$ -sphere in $\mathbb{R}^n$. The interior is also called the open ball

Let's look closer at the interior $B_r^\circ(0)$:

Question 11.5. Take any point $x \in B_r^\circ(0)$. Is it always possible to construct a smaller ball $B_{r'}(x)$ that lies entirely in $B_r^\circ(0)$?

Answer. Yes! Intuitively, regardless of how close to the “outside” of the open ball, we can always fit a very small ball,

Now let's consider an arbitrary open set. We can define it similarly, using the property we noted about the open ball.

Definition 11.6 (Open subsets of $\mathbb{R}^n$). Let U be a subset of $\mathbb{R}^n$. Then it is open if for all $x \in U$, there exists a ball centered at x which lies entirely in U . Formally,

$$\forall x \in U, \exists r > 0 \text{ st. } B_r(x) \subset U$$

Definition 11.7. (Interior of an open subset) The Interior of an open subset U is defined as the union of all open subsets contained in U

Question 11.8. Can you make an open subset closed?

Answer. Yes, by “filling in the outside”, or equivalently adding the boundary.

Definition 11.9 (Closure). The closure of a set $U \subset \mathbb{R}^n$ is the union $\overline{U} := U \cup \partial U$.

Definition 11.10. (Closure) Another way to define the closure $\overline{U}$ of a set U is the intersection of all the closed sets containing U .

Definition 11.11. (Boundary of a set) Once we have defined the closure and interior of a set U , we can define the boundary as the formal subtraction (at the level of sets) $\partial U := \overline{U} - \text{Interior}(U)$

Remark. To avoid conflict of notation, use a superscript c to denote set complements rather than an overline.

12 (10/11) : Topology Fundamentals, Con't

12.1 Point-Set Topology on $\mathbb{R}^n$

Recall that we defined:

1. The open balls $B_1^\circ(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| < 1\}$
2. The boundaries of the open balls $\partial B_1^\circ(x_0) := \{x \in \mathbb{R}^n : \|x - x_0\| = 1\}$
3. The closed ball, defined in terms of the closure of the open ball: $\overline{B_1}(x_0) := B_1^\circ(x_0) \cup \partial B_1(x_0)$

Question 12.1. How can we compare two sets at the most basic level?

Answer. We could for example look at which elements are in the intersection of the two sets, or the number of points in the set difference.

But how about uncountable sets, like intervals on the real line? We wouldn't be able to just count up the elements. Instead, we will have to determine the relative size of sets in terms of functions from one set to the other.

12.2 Injectivity, Surjectivity, Bijectivity

Definition 12.2. A map $f : A \rightarrow B$ is injective if $\forall x_1, x_2 \in A, f(x_1) = f(x_2) \implies x_1 = x_2$.

Example 12.3. An integral curve is an injective map from $\mathbb{R}$ to $\mathbb{R}^n$. This comes from the observation that an integral curve cannot have self-intersections (If it did, then there would be two times $t_1 \neq t_2$ such that $c(t_1) = c(t_2)$)

Definition 12.4. A map $f : A \rightarrow B$ is surjective if $\forall y \in B, \exists x \in A$ such that $f(x) = y$.

Definition 12.5. A map $f : A \rightarrow B$ is bijective if it is both injective and surjective.

Bijectivity is a useful tool in comparing the sizes of sets. If there is a bijection between two sets A and B , then we say that they are sets of the same size.

Definition 12.6 (Isomorphism of sets). If A and B are thought of as sets of points, a bijection between A and B is called an isomorphism. We say that A and B are isomorphic to each other as sets.

However, we want to talk about maps that have more structure, such as continuous or differentiable maps. The notion of isomorphism of sets is too weak, as we see in the following example:

Example 12.7. The intervals $(0, 1)$ and $(0, \infty)$ are isomorphic at the level of sets. To show that they are isomorphic sets, we just need to exhibit a bijection. Some examples are

$$\begin{aligned}x &\mapsto \tan\left(\frac{x\pi}{2}\right) \\x &\mapsto -\log x \\x &\mapsto \frac{1}{x} - 1\end{aligned}$$

Question 12.8. In the example of stretching the jacket, how do we do so without introducing new points? Doesn't the area/volume change when we do so?

Answer. We can do so, because area/volume is a geometrical concept, not a topological one. In topology, we only care about shapes.

Example 12.9. Is the interval $(0, 1)$ isomorphic (at the level of sets) to two disjoint copies of $(0, \infty)$? It turns out that the answer is yes. However, this requires constructing a bijection that has very nasty properties. In particular it is discontinuous.

12.3 Homeomorphisms, Simply-connectedness

We want to restrict our attention to maps between sets that do not “delete” points, or “break apart” the set into disjoint pieces. Formally, we want bijections that are continuous. This leads us to the definition of a homeomorphism.

Definition 12.10 (Homeomorphism). Let $f : A \rightarrow B$ where $A, B \subset \mathbb{R}^n$ are open subsets. Then $f : A \rightarrow B$ is a homeomorphism if:

- f is a bijection of sets
- f, f^{-1} are both continuous maps.

Question 12.11. Is the condition that f^{-1} is continuous implied by the conditions that f is a continuous bijection?

Answer. $f : [0, 1] \cup (2, 3] \rightarrow [0, 2]$ s.t. $f(x) = x, x \in [0, 1]$ and $x - 1$, otherwise would be such an example where f continuous bijection while f^{-1} is not continuous.)

Question 12.12. Can we use this definition to identify open subsets of $\mathbb{R}^n$? Can we claim that every open connected subset is homeomorphic to an open ball?

Answer. No. Some counterexamples: 1) $\emptyset$ 2) Open balls of different dimension. These are good counterexamples, but let us consider the following:

Example 12.13. Consider the sets $\{x : \|x - x_0\| < 1\}$ and $\{x : 0.5 < \|x - x_0\| < 1 \text{ in } \mathbb{R}^2\}$. One of them is the open disk, and the other is an open annulus. We claim that these two sets are not homeomorphic.

One way to see this: Consider the loop defined by the equation $\|x - x_0\| = 0.75$. On the open disk, there is a continuous deformation (that stays on the disk) of this loop onto a single point. No such continuous deformation exists on the annulus. See Figure 7

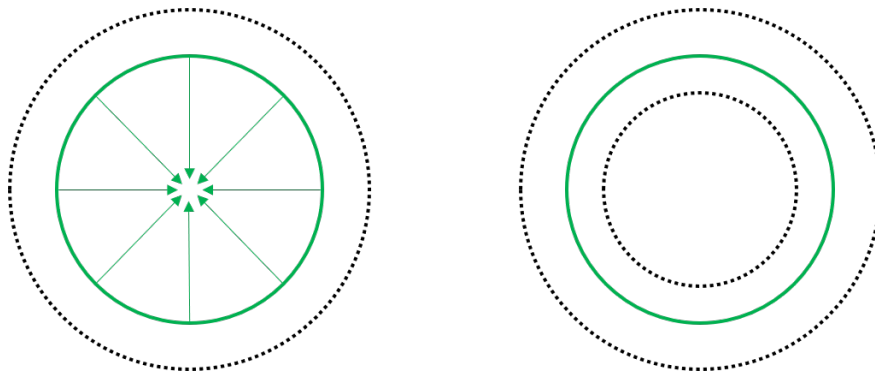


Figure 7: A disk is simply connected while an annulus is not

The notion of shrinking a loop to a point is a central one in topology. We have a definition for this:

Definition 12.14 (Simply connected aka. no holes). A subset $U \subset \mathbb{R}^n$ is called simply connected if every closed loop (a curve with the same start and end point) in U can be continuously shrunk to a point.

Example 12.15. The subset $\mathbb{R}^2 - \{0\} \subset \mathbb{R}^2$ is not simply-connected, because a loop that encloses the origin cannot be contracted onto a point. Is it homeomorphic to an annulus? turns out yes!

Example 12.16. The subset $\mathbb{R}^3 - \{0\} \subset \mathbb{R}^3$ is simply-connected.

13 (10/13) : Diffeomorphisms

13.1 Announcements

1. PSET 3 will be dropped sometime before tomorrow night

2. Midterm practice PSET: next week
3. Nov 3: In-class midterm.

13.2 Wrapping up Topology

Clearing up some confusion between isomorphisms and homeomorphisms:

- An isomorphism is a bijection at the level of sets.
- A homeomorphism is an isomorphism that has the additional property that: both the map and its inverse is continuous.

Example 13.1. The interval $(0, 1)$ is isomorphic to $[0, \infty)$. However, they are not homeomorphic, because there is no bijection from $(0, 1) \rightarrow [0, \infty)$ that are both continuous and have continuous inverse.

13.3 Differentials, Diffeomorphisms

Let us start with another example:

Example 13.2. Consider the closed unit square $U \subset \mathbb{R}^2$ and the closed unit disk $V \subset \mathbb{R}^2$.

- Are U and V homeomorphic to each other? Yes. Without going into the technicalities, one can construct a bijection $f : U \rightarrow V$ that is continuous and has continuous inverse.
- Are $f : U \rightarrow V$ and $f^{-1} : V \rightarrow U$ differentiable? No. It turns out that we will run into trouble at the corners of the square. (Think: absolute value function)

Things that are not differentiable are useless for doing calculus. So we want to consider homeomorphisms that are also differentiable both ways.

Let's take a step back and define differentials, starting from a low-dimensional case: Let $U, V \subset \mathbb{R}^2$ be open sets of the same dimension, and $f : U \rightarrow V$ be a map given by $(x^1, x^2) \mapsto (f^1(x^1, x^2), f^2(x^1, x^2))$. Recall the case of single-variable calculus: the derivative gives the best linear approximation to a curve. $f(x) = f(x_0) + f'(x_0)(x - x_0) + O((x - x_0)^2)$. To define differentiability, we need to extend the idea of a linear approximation to higher dimensions.

Definition 13.3 (Differentiability at a point). A map $f : U \rightarrow V$ is said to be differentiable at $y = (y^1, y^2)$ if there exists a linear map $Df|_y$, such that:

$$\forall \|x - y\| < \varepsilon(\delta), \frac{\|f(x) - f(y) - Df|_y(x - y)\|}{\|x - y\|} < \delta$$

This is just Taylor's theorem but reformulated. Now let us understand what the differential is doing.

$$\begin{aligned} f^1(x) &\approx f^1(y) + (Df|_y(x - y))^1 \\ f^2(x) &\approx f^2(y) + (Df|_y(x - y))^2 \end{aligned}$$

$Df|_y(x - y)$ has two components, so it lies in $\mathbb{R}^2$.

Example 13.4. Consider the map $(x^1, x^2) \mapsto (f^1(x^1, x^2), f^2(x^1, x^2))$. Then,

$$\begin{aligned} df^1 &= \frac{\partial f^1}{\partial x^1} dx^1 + \frac{\partial f^1}{\partial x^2} dx^2 \\ df^2 &= \frac{\partial f^2}{\partial x^1} dx^1 + \frac{\partial f^2}{\partial x^2} dx^2 \end{aligned}$$

We can write this in matrix form as

$$\begin{pmatrix} df^1 \\ df^2 \end{pmatrix} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} \\ \frac{\partial f^2}{\partial x^1} & \frac{\partial f^2}{\partial x^2} \end{pmatrix} \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix}$$

The matrix here is our linear map Df . Then, the value of the differential at the point y looks like:

$$Df|_y = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}|_y & \frac{\partial f^1}{\partial x^2}|_y \\ \frac{\partial f^2}{\partial x^1}|_y & \frac{\partial f^2}{\partial x^2}|_y \end{pmatrix}$$

Taking stock, given open sets of the same dimension $U, V \subset \mathbb{R}^2$ and a map $f : U \rightarrow V$ which acts by $(x^1, x^2) \mapsto (f^1(x^1, x^2), f^2(x^1, x^2))$, we obtain a family of maps indexed by $x_0 \in U$:

$$Df|_{x_0} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : \begin{pmatrix} h^1 \\ h^2 \end{pmatrix} \mapsto Df|_{x_0} \begin{pmatrix} h^1 \\ h^2 \end{pmatrix}$$

Note that we do not need to specify the domain of $Df|_{x_0}$ since it is a linear map. Generalizing, we are now in a position to state the definition of a diffeomorphism

Definition 13.5 (Diffeomorphism). Let $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ be a homeomorphism. f is a diffeomorphism if in addition f, f^{-1} are continuously differentiable.

Definition 13.6. A map f satisfying the above conditions is also called a C^1 -diffeomorphism.

Definition 13.7. A diffeomorphism f is a C^∞ -diffeomorphism if both f, f^{-1} are smooth.

Example 13.8. Consider the map $f : x \mapsto x^3$. Then we can compute

$$\begin{aligned} Df|_0 \cdot y &= 3x^2 \cdot y \\ \implies Df|_0 &= 0 \\ f^{-1}(z) &= z^{1/3} \\ Df^{-1} \cdot k &= \frac{1}{3} z^{-2/3} \cdot k \\ \implies Df^{-1}|_0 &= \infty \end{aligned}$$

So what prevented this map from being a diffeomorphism? Here we see 2 extremes: $Df|_0 = 0$ and $Df^{-1}|_0 = \infty$. It turns out that their inverse relation is not a coincidence. Before we proceed to the statement and proof, it would be useful to state the chain rule for the differential:

Definition 13.9. Let $f : U \rightarrow V$ and $g : W \rightarrow U$. Then the chain rule is

$$D(f \circ g)(x) = Df(g(x)) \cdot Dg(x)$$

Now, we are ready:

Proposition 13.10. Let $f : U \rightarrow V$ be a diffeomorphism. Then the following equality of Jacobians hold

$$Df^{-1}|_{f(x)} = (Df|_x)^{-1}$$

Proof. Let $x \in U$ be an arbitrary point. Then we have $f^{-1} \circ f(x) = x$ by definition. Taking the differential on both sides, we have

$$D(f^{-1} \circ f)(x) = \text{id}$$

where the right hand side is the differential of the identity map $x \mapsto x$. We simplify the left hand side using the chain rule to get

$$D(f^{-1} \circ f)(x) = Df^{-1}(f(x)) \cdot Df(x) = \text{id} \implies Df^{-1}(f(x)) = (Df(x))^{-1}$$

(Notation: Here we use $Df(x)$ to denote $Df|_x$ to emphasize the fact that Df is a function that returns the Jacobian at a point, and also to make the chain rule clearer.

□

To summarize, for f to be a diffeomorphism,

- f is a homeomorphism
- Df is continuous and invertible (i.e. $\det(Df) \neq 0$).

Finally, let us set up some concepts for next time:

Definition 13.11 (General linear group). The space $\text{GL}_n(\mathbb{R})$ is the set of $n \times n$ invertible matrices with real coefficients.

So another way to state our definition is:

Definition 13.12 (Diffeomorphism). A map $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ is a diffeomorphism if $Df|_x \in \text{GL}_n(\mathbb{R})$ for all $x \in U$.

14 (10/18) : More on Diffeo/Homeos, Topological Spaces

14.1 Diffeomorphisms

Recall the definition from last week: of a diffeomorphism. From now on, when we say that something is a diffeomorphism, we also impose the requirement that it is smooth.

Example 14.1. Here are some diffeomorphisms from $\mathbb{R}$ to $\mathbb{R}$:

1. The identity map

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x \end{aligned}$$

is a diffeomorphism because $Df = 1$.

2. Reparameterizations that have non-vanishing derivatives.

$$\begin{aligned} f : I \subset \mathbb{R} &\longrightarrow J \subset \mathbb{R} \\ x &\longmapsto f(x) \\ Df &\neq 0 \end{aligned}$$

So another way we can think of reparameterizations of curves is that they are diffeomorphisms of open intervals.

Examples of diffeomorphisms from $\mathbb{R}^2 \rightarrow \mathbb{R}^2$

1. Flows along smooth vector fields. (eg. the HW1 problem)
2. Shrinking or expanding an open set.

Example 14.2. Transforming an open square with corners $(\pm 1, \pm 1)$ to all of $\mathbb{R}^2$ using the map

$$(x, y) \mapsto \left(\frac{x}{1-x^2}, \frac{y}{1-y^2} \right)$$

is a diffeomorphism

Proof sketch. The function is continuous because it is a composition of continuous functions. We can ignore the blowup at $(\pm 1, \pm 1)$ because those points are not in the open square.

Now we want to check that $Df|_{x,y} \neq 0$ for all points in the open square. We can compute:

$$Df = \left| \begin{pmatrix} \frac{1+x^2}{(1-x^2)^2} & 0 \\ 0 & \frac{1+y^2}{(1-y^2)^2} \end{pmatrix} \right| = \frac{1+x^2}{(1-x^2)^2} \frac{1+y^2}{(1-y^2)^2}$$

This is well-defined on the open square and does not vanish for any real (x, y) . □

14.2 Homeomorphisms and Diffeomorphisms: The Bigger Picture

Now is a good time to ask: why did we bother defining these concepts of homeomorphisms and diffeomorphisms?

Here is a practical example: transferring solution vector fields

Example 14.3. Consider an open square with a vector field, and suppose that the vector field is the solution of some differential equation (for example). Now, if we asked to solve the same system of equations for an arbitrary open subset on the plane, it is much harder. But if we have a diffeomorphism f from the open square to the open subset, we have a much easier route. Since Df is a linear map, we can apply Df to the vector field to get the solution on the open set!

On a more abstract level, homeomorphisms and diffeomorphisms “preserves nice properties”. Because they have the same nice properties, we can just choose to work with one of them and use the homeo/diffeo to transport to the other set. This identifies open sets “in some sense”. Homeomorphisms identify sets that have the same topology.

Remark. If we have two diffeomorphisms f_1, f_2 , then $f_1 \circ f_2$ is a diffeomorphism! This follows from the chain rule: $D(f_1 \circ f_2) = Df_1 \cdot Df_2$, so if neither of those are 0, then the product is also non-zero.

14.3 Topology

Consider two open sets: U a simply connected set and V an annulus-shaped set. Suppose also we are given a number of open sets $\{U_i\}_{i=1}^n$ and $\{V_i\}_{i=1}^n$ and you are asked to “patch” the corresponding set with these smaller open sets. Can you always perform this patching? Yes. Will you perform the same way of patching for the set U and the set V ? No, because if you did you would achieve exactly the same shape.

Definition 14.4 (Topological space). A topological space is a set S and a collection of subsets of S called $\mathcal{O}$, or the open subsets, such that the following holds

1. $\emptyset \in \mathcal{O}, S \in \mathcal{O}$

2. For $U_1, U_2 \in \mathcal{O}$, then $U_1 \cap U_2 \in \mathcal{O}$ (Finite intersections of open sets are open)
3. For any collection $T \subset \mathcal{O}$, we have $\bigcup_{U_i \in T} U_i \in \mathcal{O}$ (Arbitrary unions of open sets are open)

Remark. The choice of $\mathcal{O} \subset 2^S$ determines the topology imposed on S .

Example 14.5. $\mathbb{R}$ with the “usual topology”, where $\mathcal{O} = \{\text{all open intervals}\}$. Then one can check that the 3 properties of a topological space are obeyed.

Example 14.6. $\mathbb{R}^2$ with the “usual topology”

Example 14.7. The unit sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 : \|(x, y, z)\| = 1\}$. We can give S^2 a topology by taking the open subsets of S^2 to be the intersection of the open subsets of $\mathbb{R}^3$ with S^2 .

15 (10/20) : Manifolds

15.1 Wrapping up topological spaces

Recall the definition of a topology from the previous lecture. Let’s revisit some simple examples to illustrate some properties

Example 15.1 (Infinite intersections need not be open). Let $\mathbb{R}$ be the real line with open sets $\mathcal{O}$ given by all open intervals. Then as one can check, arbitrary unions of open sets are open. So are finite intersections. However, are infinite intersections always open? No. Consider the family of open intervals $(-\frac{1}{n}, \frac{1}{n})$ for $n \in \mathbb{Z}_+$. Then the infinite intersection

$$\bigcap_{n \in \mathbb{Z}_+} (-\frac{1}{n}, \frac{1}{n}) = (-1, 1) \cap (-\frac{1}{2}, \frac{1}{2}) \cap \dots = \{0\}$$

which is *not* open.

Example 15.2 (Subspace topology). Consider the unit sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\} \subset \mathbb{R}^3$. How can one give S^2 a topology? One way is to define the open sets $\mathcal{O}_{S^2}$ in reference to the open sets $\mathcal{O}_{\mathbb{R}^3}$ of $\mathbb{R}^3$. For every $x_0 \in S^2$, we construct the open balls $B_\varepsilon(x_0) \subset \mathbb{R}^3$. Now we can let the open sets on S^2 be $S^2 \cap B_\varepsilon(x_0)$.

Remark. On a tangential note, the sphere has no boundary. The reason for this is that it itself is the boundary of another shape (the unit ball). If we let ∂ be the boundary operator, then it is a very important result that $\partial \circ \partial = 0$ (i.e. the boundary of a boundary is empty). Take Algebraic Topology to learn more about this! There is also an analog with differential forms (the differential of a differential is 0).

Remark. The intersection of a closed-without-boundary set U with an open set V is always open when restricted to U .

15.2 Manifolds: Motivation

Let’s begin with a concrete scenario. Suppose we had a unit sphere in $S^2 \subset \mathbb{R}^3$ and we had a particle on the sphere. Suppose also that the particle is restricted to move on the sphere. If you were asked to solve for the motion on the sphere, what would you do? One natural option would be to use spherical coordinates, $(\theta, \varphi) \in (0, \pi) \times (0, 2\pi)$.

How many parameters are needed to describe a curve? We need just 1, the “time” parameter. What about the unit sphere? We need two parameters, just as in spherical coordinates.

How does this work? Really, we identified the unit sphere with the open rectangle $(0, \pi) \times (0, 2\pi) \subset \mathbb{R}^2$. When we work with spherical coordinates, we are really transporting the problem back into $\mathbb{R}^2$ and then applying our mathematical machinery there (e.g. doing calculus).

Another intuition for why we say S^2 is a 2-dimensional manifold: Locally, the land looks flat, or homeomorphic to an open disk in $\mathbb{R}^2$! Without advanced technology and just looking around us, we would have thought that the Earth was flat! So another approach for doing calculus on the sphere is to chop it up into little open sets, which each look like $\mathbb{R}^2$. Then, we do the calculations of each of these little open sets and glue them back together to reassemble the sphere.

We are now ready to introduce the formal definition of a topological and differential manifolds.

Definition 15.3 (Topological Manifold). A topological n -manifold is a topological space that is locally homeomorphic to open subsets of $\mathbb{R}^n$.

Remark. The dimension of a manifold can be thought of as the number of degrees of freedom a particle moving on the manifold has.

In the same vein:

Definition 15.4 (Differential Manifold). A differential n -manifold is a topological space that is locally diffeomorphic to open subsets of $\mathbb{R}^n$.

Let M be an n -dimensional manifold. Let $x \in M$ and $U \ni x$ be an open subset. Then we can construct a homeomorphism $\varphi : U \rightarrow V \subset \mathbb{R}^n$ where V is an open set (Figure 8). We repeat this for more open subsets on M until the whole of M is covered by open subsets. We also keep track of all the φ 's that map each of these open subsets homeomorphically to open subsets of $\mathbb{R}^n$.

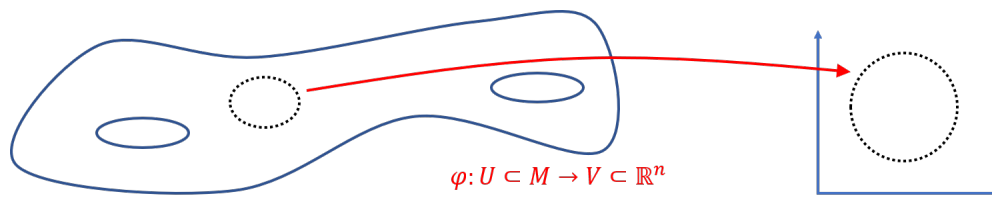


Figure 8: A chart on M

Question 15.5. For two $U_1, U_2 \subset M$, what happens in the intersection (Figure 9)? Recall that we are using these maps (charts) as convenience tools: so that we can do calculus on it. However, we must choose these charts φ_1, φ_2 such that their images are consistent in the following informal sense: Let $x, y \in U_1 \cap U_2$. Then it cannot be the case that $\varphi_1(x)$ and $\varphi_1(y)$ are nearby whereas $\varphi_2(x)$ and $\varphi_2(y)$ are far apart.

Formally, we require that whenever there are two charts $\varphi_1 : U_1 \rightarrow V_1$ and $\varphi_2 : U_2 \rightarrow V_2$ such that the intersection of their domains $U_1 \cap U_2$ is non-empty, then the image of this intersection under each of φ_1 and φ_2 are homeo/diffeomorphic to each other. I.e. there exists a homeo/diffeomorphism:

$$f : \varphi_1(U_1 \cap U_2) \xrightarrow{\sim} \varphi_2(U_1 \cap U_2)$$

Such a map is given by $\varphi_2 \circ \varphi_1^{-1}$.

16 (10/25) : Charts and Atlases

16.1 Charts, Atlases

Question 16.1. Suppose you are given a globe and are asked to identify different countries. What would you do?

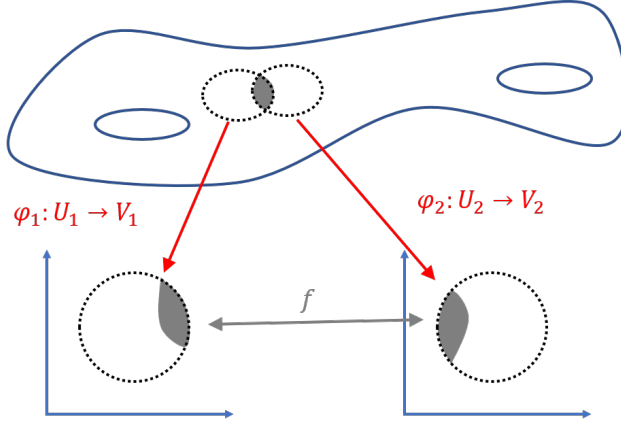


Figure 9: The intersection should map to homeo/diffeomorphic sets

Answer. We would identify the different countries with different flat pieces and try to stick them onto the globe. Intuitively, we are trying to cover the globe with these little charts.

Definition 16.2 (Charts). A chart of a topological (differential) manifold is an ordered pair (U, φ) such that $\varphi : U \rightarrow \text{im } \varphi \subset \mathbb{R}^n$ is a homeomorphism (diffeomorphism).

Definition 16.3 (Atlas). The family of charts $\{(U_i, \varphi_i)\}_{i \in \mathbb{Z}_+}$ is an atlas of a T/D manifold M if the sets cover M . In other words, $\bigcup_{i \in \mathbb{Z}_+} U_i \supset M$.

Question 16.4. Can the atlas be uncountable?

Answer. Technically yes depending on topology you choose, but defining an atlas with uncountably many sets poses some problems down the road when defining things like integration. Ideally we want as simple an atlas as possible.

16.2 Atlas of S^2 , Stereographic Projection

Let's see some examples.

Example 16.5 (Atlas of S^2). Recall that we define the unit sphere as $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. Is S^2 homeomorphic to $\mathbb{R}^2$? No. So 1 chart is insufficient to form an atlas of S^2 . But it turns out that we can construct an atlas with 2 charts. A well-known construction involves using stereographic projections.

Figure 10 shows that the sphere without the north pole is homeomorphic to $\mathbb{R}^2$. If we let $U_1 = S^2 - \{(0, 0, 1)\}$ and φ_1 be this stereographic projection, we can let (U_1, φ_1) be a chart of S^2 . Similarly, we can let $U_2 = S^2 - \{(0, 0, -1)\}$ and φ_2 be its associated stereographic projection, and construct another chart (U_2, φ_2) . For the two charts to be consistent, we need to ensure that φ_1 and φ_2 map the intersection $U_1 \cap U_2$ to homeomorphic sets!

We can do so by constructing the projections explicitly. We use the notation $N = (0, 0, 1)$ and $S = (0, 0, -1)$. Consider a point $(x, y, z) \in S^2$ and let $(p, q, 0)$ be its stereographic projection from N onto the plane $z = 0$. By construction, the points N , (x, y, z) and $(p, q, 0)$ are all collinear. Therefore,

$$\frac{p-0}{x-0} = \frac{q-0}{y-0} = \frac{0-1}{z-1} \implies p = \frac{x}{1-z}, q = \frac{y}{1-z}$$

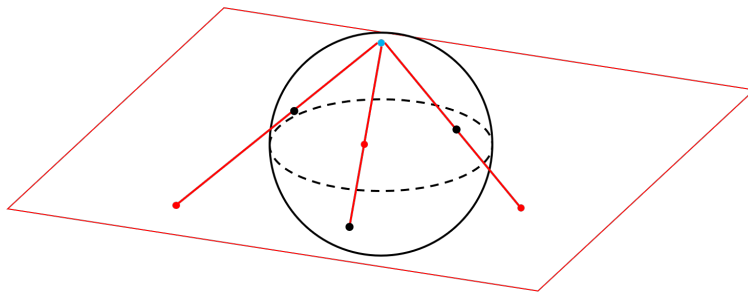


Figure 10: Stereographic projection of the punctured sphere to the whole plane (Red: on the plane, Black: on the sphere).

This means we can construct a map (and inverse too)

$$i_N : S^2 - N \longrightarrow \mathbb{R}^2$$

$$(x, y, z) \longmapsto \left(\frac{x}{1-z}, \frac{y}{1-z} \right)$$

We claim that i_N is a homeomorphism from $S^2 - N \cong \mathbb{R}^2$. There are several ways to argue this. One way is to show that i_N is a composition of continuous maps $z \mapsto 1-z \mapsto \frac{1}{1-z} \mapsto \frac{x}{1-z}$ away from $z=1$, and so the composite is a continuous map (technically you also should compute the inverse and show that inverse is continuous but that's easy since these are rational functions with no infinities). By a similar argument, we can actually show that i_N is a diffeomorphism by computing the differential and showing it does not vanish.

In a similar way, we can define the second chart as the stereographic projection from $S^2 - S \rightarrow \mathbb{R}^2$.

Question 16.6. Going back to the motivation of using charts (solving differential equations on manifolds), how does this make life easier?

Answer. Because we defined the charts to be homeomorphisms to patches of $\mathbb{R}^n$ such that they agree on the intersection, the corresponding differential equations (i.e. vector fields) are also transformed accordingly. We will see that by the way we constructed the atlas, the solutions to the differential equations match up automatically.

Remark. This result that $S^2 - N \cong \mathbb{R}^2$ is somewhat related to the hairy ball theorem.

Lemma 16.7. S^2 is a topological (differential) manifold of dimension 2 equipped with two charts: $(S^2 - N, i_N)$ and $(S^2 - S, i_S)$.

Next lecture: if we have a map f between manifolds, what does the map Df do? It maps vectors to vectors, but on an arbitrary manifold how do we understand how these vectors look like? We will need to introduce the concept of tangent spaces

17 (10/27) : Maps between Manifolds, “Bootstrapping”

Last time we defined what charts and atlases are, and gave an explicit construction for S^2 .

17.1 Maps between Manifolds

We have already been talking about them without referring to them explicitly:

1. $\mathbb{R}^2 \longrightarrow \{\text{manifolds}\}$, i.e. charts, are maps between manifolds
2. The stereographic projection $S^2 - N \longrightarrow \mathbb{R}^2$ is a map between manifolds

Let M be an m -manifold and N be an n -manifold, with $m, n \geq 1$. Does it make sense to talk about a map $f : M \longrightarrow N$? If we drop any requirements about the map (i.e. homeo/diffeo, or even injectivity/surjectivity), of course we can! We just need to send every point in M to points in N .

However, we understand our manifolds using the atlases on them. How do we understand what the map f is doing?

Idea 1: Since M is an m -manifold, we can describe it in m coordinates. if $x = (x^1, \dots, x^m)$, then maybe we can describe f in the following form:

$$f : (x^1, \dots, x^m) \mapsto (f^1(x^1, \dots, x^m), \dots, f^n(x^1, \dots, x^m))$$

That is a good idea. However, how do we make sense of the components $f^i(x^1, \dots, x^m)$? These functions don't make sense on their own. We need to define these functions in relation to the charts on M and N .

We want the domain and range of f to be something well understood. In particular it would be nice if these spaces were Euclidean space. Therefore, in order to make sense of how f acts on a small patch $U \subset M$, we send it to $\mathbb{R}^m$ via a chart on M , and we also send its image $f(U) = V \subset N$ to $\mathbb{R}^n$. We can definitely understand maps $\mathbb{R}^m \rightarrow \mathbb{R}^n$! In Figure 11, we give a definition to f by defining the maps $\psi \circ f \circ \varphi^{-1}$ for every pair of charts.

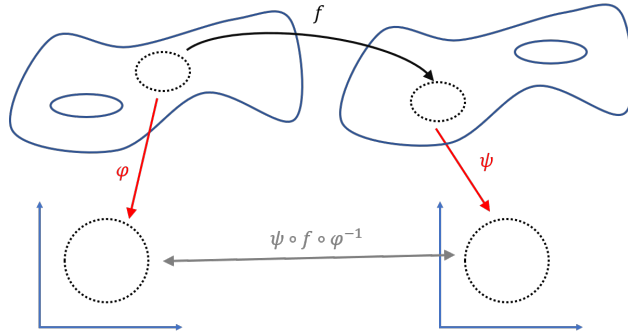


Figure 11: Defining maps between manifolds in terms of maps between their charts

In other words, we have the following commutative diagram on local objects (at the level of charts)

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ \downarrow \varphi & & \downarrow \psi \\ \mathbb{R}^m & \xrightarrow{\psi \circ f \circ \varphi^{-1}} & \mathbb{R}^n \end{array}$$

Remark. At this point, we still do not have any requirements on f .

Question 17.1. How is this useful since the definition of f is self-referential?

Answer. Say we want to define the function $f : S^2 \rightarrow \mathbb{R}$ which tells us the population density at any point on the globe. How would we do this? We would define f in terms of polar coordinates. But this is exactly defining how f acts on the charts of S^2 , since the choice of using polar coordinates (θ, ϕ) is exactly using a chart φ to bring a hemisphere of S^2 to $\mathbb{R}^2$, and then defining the function $f \circ \varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ which tells us the density as a function of the polar coordinates.

17.2 Homeo/Diffeomorphisms of Manifolds

Now, suppose we want to impose the nice properties of f .

Question 17.2. What are the conditions for f to be homeomorphic?

Answer. We need $n = m$, since the map $\psi \circ f \circ \varphi^{-1}$ is a map between open subsets of $\mathbb{R}^m$ and open subsets of $\mathbb{R}^n$. We must have $n = m$ for this to be a homeomorphism.

Of course, there are other conditions.

Question 17.3. When is f diffeomorphic?

Answer. Of course we need all the conditions for f to be homeomorphic. But we also need $D(\psi \circ f \circ \varphi^{-1})|_{\varphi(x)}$ to be invertible for all $x \in M$.

Recall that if $g : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$, then the linear map $Dg|_x$ is a map $\mathbb{R}^n \rightarrow \mathbb{R}^n$. In particular, the domain of Dg is the whole of $\mathbb{R}^n$: It maps vectors at x to vectors at $g(x)$.

In the same way, the notion that $D(\psi \circ f \circ \varphi^{-1})\varphi(x)$ has to be invertible is well defined. Referring to Figure 11 again, we see that $\psi \circ f \circ \varphi^{-1}$ is a map $\mathbb{R}^n \rightarrow \mathbb{R}^n$ (recalling that we require $n = m$). Therefore, $D(\psi \circ f \circ \varphi^{-1})|_{\varphi(x)}$ is a linear map between vectors at $\varphi(x)$ and vectors at $\psi \circ f(x)$.

Question 17.4. Now we have a way of mapping vectors from charts to charts. Is there a way to “lift” these vectors and speak of vectors on the manifold themselves?

Answer. To do so, we first need a well-defined procedure of constructing a vector space at each point on M . We do a similar bootstrapping as before. We construct “fictitious” vectors in M , and send them down to $\mathbb{R}^n$ via the differential of the chart maps. Here, we have the advantage that $\mathbb{R}^n$ has all the nice vector space properties. If we show that the “fictitious” vectors behave nicely when mapped down to $\mathbb{R}^n$, then we can infer indirectly that these “fictitious” vectors obey a vector space structure.

17.3 Road map for next few lectures

So, the key question we have to answer is the following:

Question 17.5. How do we define a vector at a base point $x \in M$.

Answer. We can try doing so in terms of Frenet curves. If we have a Frenet curve passing through x , then the moving frame of this curve at x defines a vector space.

Consider a curve $c : I \subset \mathbb{R} \rightarrow M$. How do we make sense of this curve? The same way we understand maps between manifolds, by looking at the charts φ of M . The maps $t \mapsto \varphi \circ c(t)$ are then curves on $\mathbb{R}^n$ which we can make sense of.

In $\mathbb{R}^n$, the tangent of the curve c is given by $\frac{d}{dt}(\varphi \circ c)$. In order to build up a vector space, we can think about considering the tangents of all the different curves passing through x . This gives us a number of problems. The space of possible curves is very large, and we could end up with a lot of “redundant” curves that define the same vector. We will see that we can define an equivalence relation to get rid of these extra curves, which will give us a well-defined procedure of building the tangent space. The basis vectors will end up being of the form $\frac{\partial}{\partial x^j}$.

18 (11/1) : Tangent Spaces, Equivalence Relation on Integral Curves

18.1 Recap

Recall that last week, we discussed what it means to define a map g between manifolds, and we did so in terms of the atlases of these manifolds. We also began discussing how to define the differential dg of a map between manifolds. Such a map should send vectors on a manifold to vectors on the other manifold (just as it does in the case of maps from $\mathbb{R}^n$ to $\mathbb{R}^n$). In order to do so, we need a well-defined way of defining a vector space at each point on the manifolds.

18.2 Tangent Spaces

Before going into definitions, let's motivate things with an example.

Example 18.1 (Tangents on a Sphere). Recall our favorite space $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. Suppose we draw the meridian and equator of the sphere. We can easily construct a vector field along the meridian that lies on the sphere.

Let us consider the vector field at the North pole N . Since the vector lies on the sphere, it in particular lies on the tangent plane to the sphere at N . And we can let the vector field at N be one of the basis vectors of this plane!

How do we obtain the rest of the basis? We can consider more curves that pass through N , and repeat this construction to obtain a full basis of the tangent space. (Figure 12)

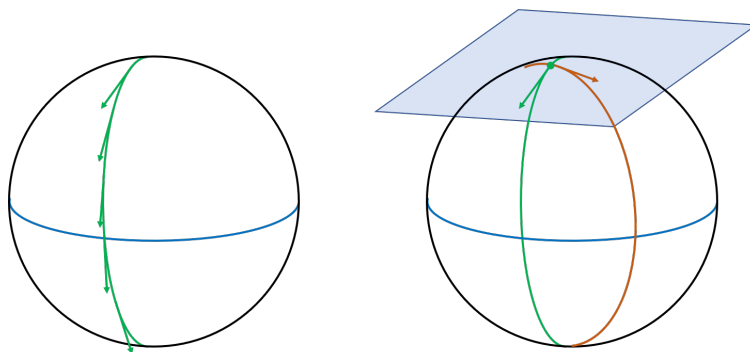


Figure 12: Left: Vector field along the meridian. Right: Constructing the tangent plane at the north pole by taking tangents of curves.

Question 18.2. What is the problem with this approach?

Answer.

1. The vectors we get by considering tangents can be linearly dependent. This happens because of there are many curves on a sphere that share the same tangent vector at a point.
2. The directional derivatives can be undefined. (We will see more of this at a later point)

18.3 Equivalence of Curves by Tangency

Let's address the first point. We want to come up with a way to get rid of the redundancy that comes with having multiple curves with the same tangent at N . What we want to do is to define an equivalence relation such that two curves are considered equivalent if their tangents at N are identical. Let us set up this equivalence relation formally.

Definition 18.3 (Equivalence of Curves). Let M be an m -manifold, and (U, φ) be a chart of M . Let $P \in U$ be a point on the manifold, and consider 2 curves (integral, parameterized by arc-length) passing through P :

$$\begin{aligned} c_1 : I_1 \subseteq \mathbb{R} &\longrightarrow M \\ c_2 : I_2 \subseteq \mathbb{R} &\longrightarrow M \end{aligned}$$

WLOG let $c_1(0) = c_2(0) = P$. We define two curves to be tangent to each other at P if they have the same derivative at $t = 0$. This works because φ is a diffeomorphism, so tangency is preserved by the chart. In other words, two curves are equivalent if

$$\left. \frac{d}{dt}(\varphi \circ c_1)(t) \right|_{t=0} = \left. \frac{d}{dt}(\varphi \circ c_2)(t) \right|_{t=0} \quad (*)$$

For the above equation to make sense, we also require that the curves are related to each other by orientation-preserving diffeomorphisms, and that φ itself is orientation-preserving. Otherwise, the derivatives of two curves tangent to each other are only equal up to sign.

Proposition 18.4. $(*)$ defines an equivalence relation among the curves passing through P . In particular, this equivalence relation is independent of the chart containing P .

Proof. We first verify the properties of an equivalence relation:

1. Reflexivity ($c \sim c$) : Since c is an integral curve, it has a unique tangent at P . Clearly it is tangent to itself.
2. Symmetry ($c_1 \sim c_2 \iff c_2 \sim c_1$) : Equality of the derivatives is symmetric.
3. Transitivity ($c_1 \sim c_2, c_2 \sim c_3 \implies c_1 \sim c_3$) : Equality of the derivatives is transitive.

Now, let us show that this equivalence relation is independent of the chart. Formally, if we let (U, φ) and (V, ψ) be two charts such that $P \in U, V$, then we want to show the following:

$$\left. \frac{d}{dt}(\psi \circ c_1)(t) \right|_{t=0} = \left. \frac{d}{dt}(\psi \circ c_2)(t) \right|_{t=0} \iff \left. \frac{d}{dt}(\varphi \circ c_1)(t) \right|_{t=0} = \left. \frac{d}{dt}(\varphi \circ c_2)(t) \right|_{t=0}$$

Suppose that we know c_1 and c_2 are tangent on the chart (V, ψ) . Since φ is a diffeomorphism, we have $\varphi^{-1} \circ \varphi = \text{id}$. Therefore, composing by identity between ψ and c , and afterwards applying the chain rule, we have

$$\begin{aligned} &\left. \frac{d}{dt}(\psi \circ c_1) \right|_{t=0} = \left. \frac{d}{dt}(\psi \circ c_2) \right|_{t=0} \\ \implies &\left. \frac{d}{dt}(\psi \circ \varphi^{-1})(\varphi \circ c_1) \right|_{t=0} = \left. \frac{d}{dt}(\psi \circ \varphi^{-1})(\varphi \circ c_2) \right|_{t=0} \\ \implies &\left. d(\psi \circ \varphi^{-1}) \right|_{\varphi \circ c_1(0)} \left. \frac{d}{dt}(\psi \circ c_1) \right|_{t=0} = \left. d(\psi \circ \varphi^{-1}) \right|_{\varphi \circ c_2(0)} \left. \frac{d}{dt}(\psi \circ c_2) \right|_{t=0} \end{aligned} \quad (1)$$

Since $\varphi \circ c_1(0) = \varphi \circ c_2(0)$, we have that

$$d(\psi \circ \varphi^{-1}) \Big|_{\varphi \circ c_1(0)} = d(\psi \circ \varphi^{-1}) \Big|_{\varphi \circ c_2(0)} := J$$

In particular, since J is the Jacobian of a composite of diffeomorphisms, J satisfies $\det J \neq 0$, and J^{-1} exists. Rewriting (1), we now have

$$J \cdot \frac{d}{dt}(\psi \circ c_1) \Big|_{t=0} = J \cdot \frac{d}{dt}(\psi \circ c_2) \Big|_{t=0}$$

Applying J^{-1} on the left on both sides, we get the desired conclusion. $\square$

19 (11/8) : Defining the Tangent Space

19.1 Midterms, Other logistics

Range of scores is [38, 52] out of 55. Will have 2 more psets (4 and 5) before the end of the semester.

19.2 Tangent Vectors

Recall the setup from last week: Suppose we have an n -manifold M . For a chart (U, φ) and a point $P \in U$, we defined an equivalence relation of curves passing through P in M . WLOG let the curves $c_1 : I \rightarrow M \sim c_2 : J \rightarrow M$ pass through P at $t = 0$. Then the equivalence relation is

$$c_1 \sim c_2 \iff \frac{d}{dt}(\varphi \circ c_1)(0) = \frac{d}{dt}(\varphi \circ c_2)(0) \quad (*)$$

Definition 19.1 (Tangent vector). A tangent vector at a point $P \in M$ is an equivalence class under the equivalence given by $(*)$.

If M is n -dimensional, we should be able to find n linearly independent tangent vectors at any point $P \in M$. This is not necessarily true, of course we can choose tangent vectors that are linearly dependent. We are now going to look at how we can find this independent set. We also saw that the definition of the tangent vectors is independent of the choice of chart! This is a very important property as it allows us to transform our local basis.

19.3 Construction of the Tangent Space

Our goal now is to construct a vector space at each $P \in M$. Consider curves c_1, c_2 which have the same orientation and are parameterized by arclength. Also assume that they are in different equivalence classes. Note that c_1, c_2 are currently abstract objects with no definition. For a curve c , let ξ_c denote its tangent vector on M . Given a chart (U, φ) , we can look at the images of c_1 and c_2 under φ . We can construct a map θ_φ that takes a tangent vector and maps it to the tangent to the image of the curve in $\mathbb{R}^n$. Concretely,

$$\theta_\varphi : \xi_c \mapsto \frac{d}{dt}(\varphi \circ c) \Big|_{t=0}$$

We would like to use θ_φ to define the tangent vectors ξ_c upstairs. And so we will need to show that this map is a bijection.

Proposition 19.2. θ_φ is a bijection if the equivalence relation $(*)$ is imposed.

Proof. Note that ξ_c is really the same as $\frac{dc}{dt}|_{t=0}$.

Injectivity: Consider curves c, c_1 such that $\theta_\varphi(\xi_c) = \theta_\varphi(\xi_{c_1})$. Then we need to show that $\xi_c = \xi_{c_1}$. Expanding definitions, we have

$$\left. \frac{d}{dt}(\varphi \circ c) \right|_{t=0} = \theta_\varphi(\xi_c) = \theta_\varphi(\xi_{c_1}) = \left. \frac{d}{dt}(\varphi \circ c_1) \right|_{t=0}$$

But this is precisely the same expression as the equivalence relation. Therefore the tangents ξ_c and ξ_{c_1} that the curves define are the same.

Surjectivity: Given a vector downstairs, can we construct a curve upstairs that maps to it under θ_φ ? We can construct a curve downstairs and lift it up to M via φ^{-1} .

Given a vector $v \in \mathbb{R}^n$, let us consider the curve $t \mapsto tv + \varphi(0)$. Lifting it, we have the curve $\alpha : t \mapsto \varphi^{-1}(tv + \varphi(0))$. Let's check that the tangent of this curve is mapped to v under θ_φ :

$$\begin{aligned} \theta_\varphi(\xi_\alpha) &= \left. \frac{d}{dt}(\varphi \circ \alpha) \right|_{t=0} = \left. \frac{d}{dt}(\varphi \circ (\varphi^{-1}(tv + \varphi(0)))) \right|_{t=0} \\ &= d\varphi|_{\varphi^{-1} \circ \varphi(0)} \cdot d\varphi^{-1}|_{\varphi(0)} \cdot v \\ &= Iv = v. \end{aligned}$$

Therefore θ_φ is a bijection between the tangent vectors at P and the unit vectors in $\mathbb{R}^n$. □

We will eventually show that θ_φ is a diffeomorphism. But for now, bijection will suffice. Now we can impose a vector space structure on the tangent space by simply imposing the vector space relations:

Definition 19.3 (Vector space structure of the tangent space). Let ξ, η be two distinct tangents at $P \in M$ and $\lambda \in \mathbb{R}$. Then define

1. $\xi + \eta := \theta_\varphi^{-1}(\theta_\varphi(\xi) + \theta_\varphi(\eta))$
2. $\lambda\xi = \theta_\varphi^{-1}(\lambda\theta_\varphi(\xi))$

Using this, we can construct a basis at each point:

Definition 19.4 (Tangent space at a point). The tangent space at $P \in M$, denoted by $T_P M$, is the set of tangent vectors defined through $(*)$.

So far we have been working with the tangents to a specific point. But how do we do so in a consistent manner over the entire manifold? We will see eventually that this will involve the concept of directional derivatives!

20 (11/10) : Tangent Bundles, Derivations

20.1 More on Tangent Spaces

Recall the new definition of $T_P M$: it is the set of tangent vectors at a point $P \in M$ under the equivalence relation $(*)$ in last class. Let's look at some examples:

Example 20.1. Consider the unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ and a point $P = (z_1, z_2)$ on the circle. What is the tangent space $T_P S^1$?

In this example, we can explicitly write down the tangent space. Since S^1 is a 1-manifold, the tangent space is 1-dimensional, and so has only 1 basis vector. If we let $\vec{r} = (z_1, z_2)$ and $\vec{v}$ be the basis of $T_P S^1$, then the tangent space is defined by $\langle \vec{v}, \vec{r} \rangle = 0$.

Remark. We have the general fact that $T_P M \cong \mathbb{R}^m$, where $\dim M = m$. Is this circular? No. Recall that the dimension of the manifold is defined in terms of the charts: If the charts on M give diffeomorphisms to $\mathbb{R}^m$, then we define $\dim M = m$. Then it makes sense for $\dim T_P M = m$ since the tangents are just linear approximations on the manifold.

20.2 Tangent Bundles

Let's begin with a question:

Question 20.2. Where do the vectors lie on a manifold?

Answer. By construction, they lie in the tangent spaces of the manifold.

Notice that the tooling we have set up so far only allows us to compare vectors that lie in the same tangent space. For example, if we have two points P, Q in the same chart (U, φ) , the vector difference $\varphi(P) - \varphi(Q)$ in $\mathbb{R}^m$ does not translate directly back to a vector difference in the manifold, because we currently don't have the machinery to compare the tangent spaces $T_P M$ and $T_Q M$! It turns out that we will have to generalize the notion of parallel transport of vectors, and this will lead to the concept of connections.

So we want to ultimately be able to treat the vectors in all the different tangent spaces as belonging to the same overall "space". Well, what happens if we take the union of all the tangent spaces?

Definition 20.3 (Tangent Bundle). The tangent bundle TM is defined by the following:

$$TM := \bigsqcup_{P \in M} (P, T_P M)$$

The union is disjoint: each pair is disjoint different because the base point (the first coordinate) is different.

Example 20.4 (Tangent Bundle of S^1). What does TS^1 look like? We have the disjoint union of $(P, \mathbb{R})$ for each point P on the circle. If we keep the picture flat on the plane, we get a mess. But if we lift the picture into 3-dimensional space, we can rotate each tangent space by a quarter turn so that they are normal to the plane of the circle. In this way, it is clear that TS^1 is isomorphic to the cylinder $S^1 \times \mathbb{R}$. (Figure 13)

Question 20.5. Seeing the tangent bundle as a cylinder seems weird. If you travel along the cylinder, it can lead to arbitrary and discontinuous changes in the actual tangent spaces.

Answer. That is indeed a problem. We will later define the concept of a section which will turn out to be a way of selecting a vector field from the tangent bundle.

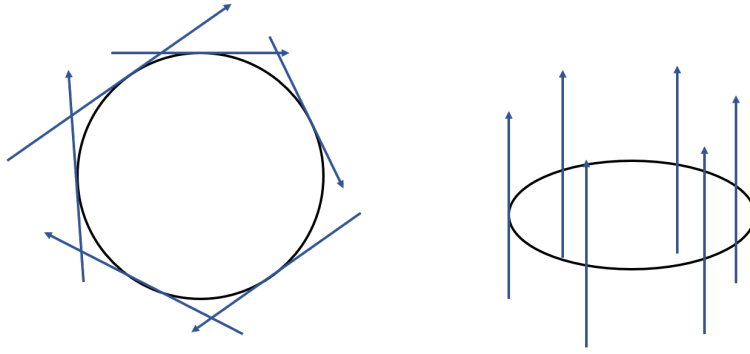


Figure 13: Thinking about TS^1

20.3 Constructing a basis, Derivations

Let's consider a euclidean space $U \subset \mathbb{R}^2$ and a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. How would you compute the directional derivative in the direction of a vector v at a point $(x, y) \in U$? From multivariate calculus this is simply

$$\begin{aligned} \partial_{\vec{v}} f &= \nabla f \cdot v = v^1 \partial_x f + v^2 \partial_y f \\ &= \sum_{i=1}^2 v^i \frac{\partial f}{\partial x^i} && \text{(writing } x^1 = x, x^2 = y) \\ &= \left(\sum_{i=1}^2 v^i \frac{\partial}{\partial x^i} \right) f && \text{(since differentials act linearly)} \end{aligned}$$

Here is where it gets interesting. There is a bijective identification between the sets

$$\{\text{vectors in } \mathbb{R}^2\} \longleftrightarrow \{\text{directional derivatives on } \mathbb{R}^2\}$$

ie, each $v \in \mathbb{R}^2$ can be uniquely associated with a directional derivative $\sum_{i=1}^2 v^i \frac{\partial}{\partial x^i}$.

Definition 20.6. $C^\infty(M)$ is the space of smooth functions on M , ie the set of $f : M \rightarrow \mathbb{R}$.

Definition 20.7 (Vectors as derivations). A tangent vector $v \in T_P M$ is a **linear** map

$$v : C^\infty(M) \longrightarrow \mathbb{R}$$

$$f \longmapsto v(f)$$

and satisfies the product rule

$$\begin{aligned} v_p(fg) &= v_p(f)g + f v_p(g) \\ v_p(cf) &= c v_p(f), \quad c \in \mathbb{R} \end{aligned}$$

v_p is called a derivation.

We can easily check that the directional derivatives satisfy the above definition. We say that the usual directional derivative is one such derivation of a vector.

Remark. Notice that we use v both to describe a vector in the usual sense, as well as to describe the directional derivative associated with it. They are strictly speaking different objects, but they can be identified isomorphically.

Notice that $v_P(1) = 0$ (use the product rule with $f = g = 1$), and by extension $v_P(c) = 0$ (use scalar multiplication).

20.4 Teaser for next week

Notice how we ordinarily express a vector w in terms of its basis as $\sum w_i e^i$. Looking at the identification $v \cong \sum v^i \frac{\partial}{\partial x^i}$, we might want to consider the partial operators as the basis of the tangent space!

Here is a theorem we will prove:

Theorem 20.8. Let $P \in M$. Then, the set $\left\{ \frac{\partial}{\partial x^i} \Big|_P \right\}_{i=1}^m$ forms a basis of $T_P M$. Furthermore,

$$v|_P = \sum_{i=1}^m v^i(P) \frac{\partial}{\partial x^i} \Big|_P$$

21 (11/15): Basis of the Tangent Space, Dual Space

Confusion from last time: We defined the tangent bundle as

$$TM = \bigcup_{P \in M} (P, T_P M)$$

where the union is disjoint. In the case of S^1 , this gives us that the tangent bundle TS^1 is homeomorphic to $S^1 \times \mathbb{R}$, a cylinder. Some people were confused because this the tangent lines of the circle intersect each other, and these intersections are not reflected in the cylindrical structure. However, note that in our treatment, the intersections between these tangent lines is meaningless: it doesn't make sense to compare two vectors from different tangent spaces directly.

21.1 Differentials as the Basis of the Tangent Space

We begin by restating the theorem at the end of last lecture:

Theorem 21.1. Let M be a m -manifold and $P \in M$. Let $\{x^i\} = (x^1, \dots, x^m)$ be a local chart (aka coordinates). Then, a basis of $T_P M$ (which is isomorphic to $\mathbb{R}^m$) can be given by the set $\left\{ \frac{\partial}{\partial x^i} \Big|_P \right\}$ and any vector $v \in T_P M$ can be written as

$$v|_P = \sum v^i(P) \frac{\partial}{\partial x^i} \Big|_P$$

Proof. Let $f : M \rightarrow \mathbb{R}$ be a smooth function, ie. $f \in C^\infty(M)$. Take $P = (0, 0, \dots, 0)$ (under the chart), and let $U \ni P$ be an open set.

Given a smooth function, we can always take its Taylor series. However, there are functions that have 0 radius of convergence. But we can always write, given $x \in U$:

$$f(x) = f(0) + \int_0^1 \frac{d}{dt} f(tx) dt$$

Check that the right hand side agrees by applying the fundamental theorem of calculus. Now, using the chain rule, we have

$$\begin{aligned}
\frac{d}{dt}f(tx) &= \frac{d}{dt}(tx^1, tx^2, \dots, tx^m) \\
&= \frac{d}{dt}(y^1, y^2, \dots, y^m) && \text{(letting } y^j = tx^j) \\
&= \frac{\partial f}{\partial y^1} \frac{dy^1}{dt} + \dots + \frac{\partial f}{\partial y^m} \frac{dy^m}{dt} \\
&= \sum_{i=1}^m \partial_i f(tx) \cdot x^i && \text{(Using } \partial_i = \frac{\partial}{\partial y^i})
\end{aligned}$$

Plugging it back, we have

$$\begin{aligned}
f(x) &= f(0) + \int_0^1 \left(\sum_{i=1}^m x^i \partial_i f(tx) \right) dt \\
&= f(0) + \int_0^1 \left(\sum_{i=1}^m x^i f_i \right) dt && \text{(Letting } f_j = \partial_j f(tx))
\end{aligned}$$

Now, recall we saw last lecture that $v_P(1) = 0$ and $v_P(c) = 0$. We can thus act on the last equation by v_P , making use of the product rule to obtain:

$$v_P f = v_P f(0) + \int_0^1 \sum_{i=1}^m \left(v_P(x^i) f_i(tx)|_P + x^i|_P v_P f_i(tx) \right) dt$$

(Note: this is differentiating under the integral sign. But all the integrals converge, so this is fine).

The first term is 0 because we are applying v_P to a constant. Also, the $x^i|_P = 0$ since we set P to be the origin under the chart (x^i gives the i -th chart coordinate of P). This simplifies the expression to

$$\begin{aligned}
v_P f &= \int_0^1 \left(\sum_{i=1}^m v_P(x^i) \partial_i f|_P \right) dt = \sum_{i=1}^m v_P(x^i) \partial_i f|_P && \text{(Integrand has no } t \text{ dependence)} \\
&= \left(\sum_{i=1}^m v_P(x^i) \frac{\partial}{\partial x^i} \right) \Big|_P f
\end{aligned}$$

The equality holds for arbitrary $f \in C^\infty(M)$. So, we have the vector equality

$$v_P = \sum_{i=1}^m v_P(x^i) \frac{\partial}{\partial x^i} \Big|_P$$

Now, recall that v_P is a vector in the tangent space. This therefore shows that $\{\frac{\partial}{\partial x^i}|_P\}$ is a spanning set of the tangent space. To show that it is a basis, it remains to show that it is linearly independent. In other words, we want the following:

$$\sum_{i=1}^m a^i \frac{\partial}{\partial x^i} = 0 \implies a^i = 0 \quad \forall i$$

We can see what happens if we apply both sides to the coordinate functions:

$$\left(\sum_{i=1}^m a^i \frac{\partial}{\partial x^i} \right) x^j = \sum_{i=1}^m a^i \frac{\partial x^j}{\partial x^i} = a^j = 0$$

Letting j vary from 1 to m , this forces all the coefficients to be 0, as desired. $\square$

Remark. $v^i(P) = v_P(x^i)$ is the i -th component of the vector v at P .

Question 21.2. But this only works at a point P right?

Answer. Yes. To extend it to the whole manifold, we will have to create a vector field that gives the basis at any point.

Question 21.3. So does this work on $\mathbb{R}^n$?

Answer. Yes. For example, $\frac{\partial}{\partial x^1} = (1 \ 0 \ \cdots \ 0)$.

21.2 Dual Space

Now that we have set up a vector space and its basis, we turn our attention to its dual space. This for example, will help us set up the inner product.

Here is a refresher:

Definition 21.4 (Dual Space). Let V be an $\mathbb{R}$ -vector space. Then the dual space V^* is defined as the set of linear operators $V \rightarrow \mathbb{R}$.

$$V^* = \{w : w(v) \in \mathbb{R}, v \in V\}$$

In other words, the dual space is the space of linear functionals on V .

Example 21.5. On $\mathbb{R}^3$, the dual is $(\mathbb{R}^3)^*$, the set of vectors that send vectors $\mathbb{R}^3 \rightarrow \mathbb{R}$. Canonically, if $w \in (\mathbb{R}^3)^*$, then $w(1, 1, 1) = 1 \cdot w^1 + 1 \cdot w^2 + 1 \cdot w^3$, for example.

Next class, we will define the dual space T_P^*M of the tangent space $T_P M$, otherwise known as the cotangent space. We will also see that a basis of T_P^*M is given by $\{dx^i|_P\}$.

22 Co-tangent spaces, co-vectors, Tensors,.....

Let us denote the tangent space of M at p by $T_p M$. We proved that $T_p M$ is a vector space of dimension m and moreover using the notion of ‘derivation’ we constructed a basis of $T_p M$ that is given by $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$ in a local chart $\{x^i\}_{i=1}^m$. Once we have constructed the tangent space, we can talk about its dual vector space. First, we define the notion of a linear functional.

Definition 22.1 (Linear Functional). A linear functional on $T_p M$ is a linear map $\varphi : T_p M \rightarrow \mathbb{R}$, $v \mapsto \varphi(v)$.

Let us denote the space of linear functional on $T_p(M)$ as $\mathcal{L}(T_p M, \mathbb{R})$. Cotangent space $T_p^* M$, the dual space of $T_p M$ is defined as follows.

Definition 22.2 (Cotangent Space). The co-tangent space $T_p^* M$ is $\mathcal{L}(T_p M, \mathbb{R})$.

In finite dimensions, we know that the dimension of a vector space (tangent space) is same as the that of its dual space (co-tangent space). More precisely, $T_p^* M$ and $T_p M$ are isomorphic as vector spaces. But there is no canonical way to construct this isomorphism. We will see soon enough that we can construct *an* isomorphism once we define a metric on M . The next thing we want to do is to define a basis on $T_p M$.

Theorem 22.3. Let $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$ be a basis of $T_p M$. For each $i = 1, 2, 3, \dots, m$ define the linear functionals on the basis $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$ as follows

$$dx^j(\frac{\partial}{\partial x^i}) = \delta_i^j, \delta_i^j = 1 \text{ if } i = j, \delta_i^j = 0, \text{ if } i \neq j, \quad (14)$$

then $\{dx^j\}_{j=1}^m$ is a basis of $T_p^* M$.

Proof. The first thing we want to show that $\{dx^j\}_{j=1}^m$ are linearly independent. To prove this, we have to show that if the following linear combination vanishes i.e.,

$$\sum_{j=1}^m \alpha_j dx^j = 0, \quad (15)$$

then $\alpha_j = 0 \forall j \in 1, \dots, m$. How do we do this? We need to somehow use the condition 14 i.e., act $\sum_{j=1}^m \alpha_j dx^j$ on $\frac{\partial}{\partial x^i}$ i.e.,

$$\begin{aligned} \sum_{j=1}^m \alpha_j dx^j \left(\frac{\partial}{\partial x^i} \right) &= 0 \left(\frac{\partial}{\partial x^i} \right) \\ \implies \sum_{j=1}^m \alpha_j \delta_i^j &= 0 \implies \alpha_i = 0, \end{aligned} \quad (16)$$

where we have used the linearity of the maps $dx^j : T_p^*M \rightarrow \mathbb{R}$. Therefore we have a proof that $\{dx^j\}_{j=1}^m$ are linearly independent. Now we prove that $\{dx^j\}_{j=1}^m$ spans T_p^*M . Let $\beta \in T_p^*M$ and for each j , let β_j denotes $\beta(\frac{\partial}{\partial x^j})$. We claim that the following holds

$$\beta = \sum_{j=1}^m \beta_j dx^j. \quad (17)$$

What this means is that both sides should produce the same result when acting on any arbitrary vector $W \in T_pM$. We can check this on a basis since by appropriate linear combinations, we can get any vector. Therefore if we act both sides by $\frac{\partial}{\partial x^i}$, we get

$$\beta \left(\frac{\partial}{\partial x^i} \right) = \beta_i \quad (18)$$

on the left-hand side. On the right-hand side, we get by linearity

$$\sum_{j=1}^m \beta_j dx^j \left(\frac{\partial}{\partial x^i} \right) = \sum_j \beta_j \delta_i^j = \beta_i. \quad (19)$$

Therefore, 17 agrees on the basis. Therefore it agrees on any vector belonging to T_pM . Therefore we are done. $\square$

Remark. An element of T_p^*M is called a **co-vector**.

Remark. Any vector $V \in T_pM$ can be written as $V = \sum_{i=1}^m V^i(x) \frac{\partial}{\partial x^i}$ and any co-vector α in T_p^*M can be written as $\alpha = \sum_{j=1}^m \alpha_j(x) dx^j$. This is simple linear algebra. V^i s are the components of the vector V along the vectors $\frac{\partial}{\partial x^i}$. Similarly, α_j s are the components of α along the co-vectors dx^j .

Remark. $(T_p^*M)^* \simeq T_pM$ and canonically the basis of T_pM are $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$. Therefore $\frac{\partial}{\partial x^i}(dx^j) = \delta_i^j$.

Example 22.4 (vectors and co-vectors on $\mathbb{R}^2$). Let us consider the manifold $\mathbb{R}^2$ equipped with a standard coordinate chart (x, y) . In this chart, we have the two usual basis vectors $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ of the tangent space $T_{x,y}\mathbb{R}^2$. The basis of co-tangent space $T_{(x,y)}^*\mathbb{R}^2$ is given by dx and dy . Consider the vector $V = 2x \frac{\partial}{\partial x} + \frac{\partial}{\partial y}$ and a co-vector $\alpha = 2dx + ydy$. We can explicitly compute $V(\alpha)$ and $\alpha(V)$

$$V(\alpha) = (2x \frac{\partial}{\partial x} + \frac{\partial}{\partial y})(2dx + dy) = 4x + 1 \quad (20)$$

$$\alpha(V) = (2dx + dy)(2x \frac{\partial}{\partial x} + \frac{\partial}{\partial y}) = 4x + 1 \quad (21)$$

In particular, $\alpha(V) = V(\alpha)$.

Once we have defined the tangent space $T_p M$ and its dual space the co-tangent space $T_p^* M$, we can start constructing new spaces by taking the Cartesian product of these spaces. Define the following Cartesian product of r copies of the tangent space $T_p M$ i.e.,

$$\underbrace{T_p M \times T_p M \times T_p M \times \cdots \times T_p M}_{r \text{ copies}} \quad (22)$$

$$:= \{(V_1, V_2, V_3, \dots, V_r) | V_1 \in T_p M, V_2 \in T_p M, V_3 \in T_p M, \dots, V_r \in T_p M\}.$$

Similarly define the Cartesian product of s copies of $T_p^* M$

$$\underbrace{T_p^* M \times T_p^* M \times T_p^* M \times \cdots \times T_p^* M}_{s \text{ copies}} \quad (23)$$

$$:= \{(\alpha^1, \alpha^2, \alpha^3, \dots, \alpha^s) | \alpha^1 \in T_p^* M, \alpha^2 \in T_p^* M, \alpha^3 \in T_p^* M, \dots, \alpha^s \in T_p^* M\}.$$

Now construct the following map

$$\mathbf{Q} : \underbrace{T_p M \times T_p M \times T_p M \times \cdots \times T_p M}_{r \text{ copies}} \times \underbrace{T_p^* M \times T_p^* M \times T_p^* M \times \cdots \times T_p^* M}_{s \text{ copies}} \rightarrow \mathbb{R} \quad (24)$$

$$(V_1, \dots, V_r, \alpha^1, \dots, \alpha^s) \mapsto \mathbf{Q}(V_1, \dots, V_r, \alpha^1, \dots, \alpha^s) \quad (25)$$

that verifies the multi-linearity property i.e., linearity property in each slot. More explicitly

$$\begin{aligned} & \mathbf{Q}(V_1, V_2, \dots, aV_i + b\widehat{V}_i, \dots, V_r, \alpha^1, \dots, \alpha^s) \\ &= a\mathbf{Q}(V_1, \dots, V_i, \dots, V_r, \alpha^1, \dots, \alpha^s) + b\mathbf{Q}(V_1, \dots, \widehat{V}_i, \dots, V_r, \alpha^1, \dots, \alpha^s) \end{aligned} \quad (26)$$

true for any slot of $\mathbf{Q}$ (both V and α slots). $\mathbf{Q}$ is called a multilinear functional on

$$\underbrace{T_p M \times T_p M \times T_p M \times \cdots \times T_p M}_{r \text{ copies}} \times \underbrace{T_p^* M \times T_p^* M \times T_p^* M \times \cdots \times T_p^* M}_{s \text{ copies}}.$$

If you restrict to $r = 1$ and $s = 0$, you simply recover the co-vectors and if you restrict to $r = 0$ and $s = 1$, then you recover vectors. This leads to the definition of tensors on a manifold.

Definition 22.5 (Tensors). The multilinear functional $\mathbf{Q}$ is called a tensor of type $\binom{s}{r}$. We denote the space of tensors of type $\binom{s}{r}$ by $\mathcal{T}_r^s$.

Remark. In particular, $\binom{0}{1}$ tensors are co-vectors and $\binom{1}{0}$ tensors are vectors.

Intuition: Think of a blender that takes in two types of fruits and gives you juice. Think of one type of fruits (let's say type A) being elements of $T_p M$ (i.e., vectors) and the other type of fruits (let's say type B) being elements of $T_p^* M$ (i.e., co-vectors). Now you can think of a tensor of type $\binom{s}{r}$ as a blender that takes r many fruits of type A (vectors) and s many fruits of type B (co-vectors) and produces Juice (real number; juice is real enough though, ain't it?). Now you can think of multi-linearity as follows: if I double the size of any fruit I would get twice the juice (scalar multiplication) and if I add more fruits, I would get extra juice (addition) i.e., $\mathbf{Q}(\dots, aV_1 + V_2, \dots) = a\mathbf{Q}(\dots, V_1, \dots) + \mathbf{Q}(\dots, V_2, \dots)$. Think of this analogy and you'll never forget tensors.

Now we ask the following question.

Question 22.6. How do we construct a basis of this space $\mathcal{T}_r^s$?

In order to construct a basis of $\mathcal{T}_r^s$, we need to define a notion of tensor product.

Definition 22.7. The tensor product is the linear map

$$\otimes : \mathcal{T}_{r_1}^{s_1} \times \mathcal{T}_{r_2}^{s_2} \rightarrow \mathcal{T}_{r_1+r_2}^{s_1+s_2} \quad (27)$$

defined as follows

$$\begin{aligned}
& (\mathcal{A}, \mathcal{B}) \mapsto \mathcal{A} \otimes \mathcal{B} \\
& (\mathcal{A} \otimes \mathcal{B})(V_1, V_2, \dots, V_{r_1+r_2}, \alpha^1, \alpha^2, \dots, \alpha^{s_1+s_2}) \\
& := \mathcal{A}(V_1, V_2, \dots, V_{r_1}, \alpha^1, \alpha^2, \dots, \alpha^{s_1}) \bullet \mathcal{B}((V_{r_1+1}, V_{r_1+2}, \dots, V_{r_1+r_2}, \alpha^{s_1+1}, \alpha^{s_1+2}, \dots, \alpha^{s_1+s_2}),
\end{aligned}$$

where $\bullet$ is simply the $\mathbb{R}$ -multiplication.

Now we can build our way up using the definition of the tensor product. Suppose I want to study the space $\mathcal{T}_2^0$ i.e., tensors of the type $\binom{0}{2}$. How exactly do we do that? We can use the definition of the tensor product. Notice that the tensor product of two tensors of type $\binom{0}{1}$ gives a tensor of type $\binom{0}{2}$ according to the definition 22.7 of the tensor product. Therefore if I have a basis $\{dx^i\}_{i=1}^m$ for $\mathcal{T}_1^0 = T_p^*M$, then, we can take two copies of $\{dx^i\}_{i=1}^m$ one for each $\mathcal{T}_1^0$ and tensor product them to define a basis for $\mathcal{T}_2^0$. Let us formalize this in the following theorem

Theorem 22.8. Let us denote the space of tensors of the type $\binom{s}{r}$ by $\mathcal{T}_r^s$. Then $\{dx^{i_1} \otimes \dots \otimes dx^{i_r} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_s}}\}_{i_1, \dots, i_r, j_1, \dots, j_s=1, \dots, m}$ is a basis of the space $\mathcal{T}_r^s$.

Proof. We sketch out the proof. This is similar to the proof of the basis for T_p^*M and T_pM . Firstly note that these constitute a linearly independent set. To show this we use the usual definition i.e.,

$$\sum_{i_1, \dots, i_r, j_1, \dots, j_s=1, \dots, m} \alpha_{i_1 \dots i_r}^{j_1 \dots j_s} dx^{i_1} \otimes \dots \otimes dx^{i_r} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_s}} = 0 \quad (28)$$

implies $\alpha_{i_1 \dots i_r}^{j_1 \dots j_s} = 0 \forall i_1, \dots, i_r, j_1, \dots, j_s = 1, \dots, m$. Here the right-hand side 0 is the 0 tensor. The trick is the same. We act both sides on $\frac{\partial}{\partial x^{k_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{k_r}} \otimes dx^{l_1} \otimes \dots \otimes dx^{l_s}$ to yield

$$\alpha_{k_1 \dots k_r}^{l_1 \dots l_s} = 0. \quad (29)$$

The span part is exactly the same as done in the proof of the theorem 22.3. Please do this and check. $\square$

Remark. Now I can even compute the dimension of the space $\mathcal{T}_r^s$. for this, we need to find out how many basis are there. Or even in simpler terms, how many elements are there in the set $\{dx^{i_1} \otimes \dots \otimes dx^{i_r} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{j_s}}\}_{i_1, \dots, i_r, j_1, \dots, j_s=1, \dots, m}$? Notice that i_1 can any number between 1 and m , i_2 can be any number between 1 and m and so on. Therefore, total number of elements of this set is essentially $\underbrace{m \times m \times m \dots \times m}_{r \text{ time}} \times \underbrace{m \times m \times m \dots \times m}_{s \text{ time}} = m^{r+s}$.

Now what is the intuition and motivation behind constructing the tensors? Why go through all the troubles of defining a tensor product and all that? In other words, can we have some *physical* objects described by a tensor? Of course, most trivially, vectors are of type $\binom{0}{1}$ and we know what they mean. What about tensors of type $\binom{0}{2}$?

Look at the diagram. I have a cube $\mathbf{C}$ which is a zoomed-out version of a sufficiently small cube. Now I want to take the cube put it under water. On the cube under water the water pressure will be acting from all directions. Since the cube is chosen to be sufficiently small, I can think of the pressure/stress/force per unit area to be roughly constant over the entire cube. Now if I ask you what is the force per unit area on the face A in the direction of y axis? Firstly, at each point on the cube, we have the tangent space $\mathbb{R}^3$ which is spanned by three vectors $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\}$. Now the face A is identified by its normal which is $\frac{\partial}{\partial x}$. Therefore, to answer the question “what is the force per unit area on the face A in the direction of y axis?”, I need input as two vectors $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$. Our answer should be the amount of force per unit area which is areal number.

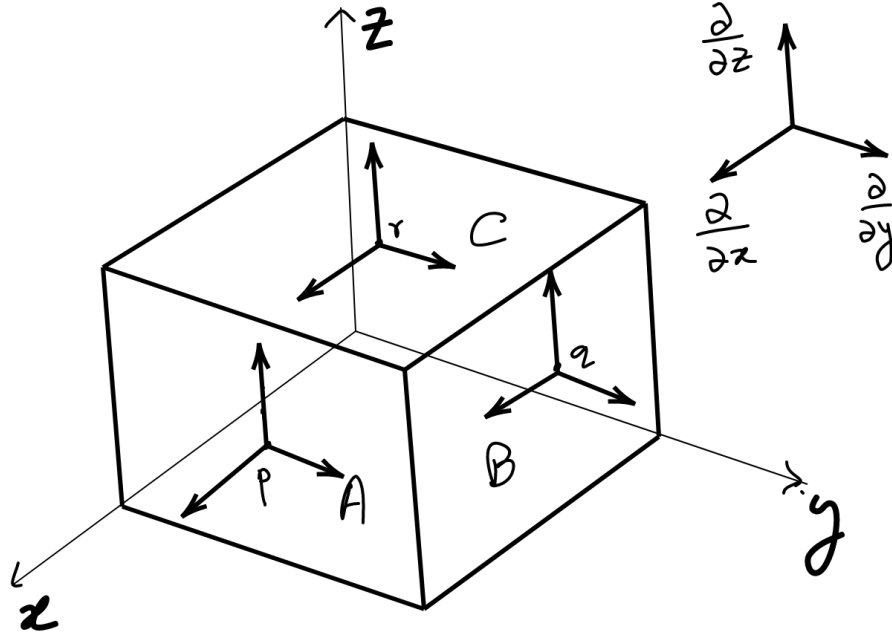


Figure 14: The cube C in $\mathbb{R}^n$. at each point the tangent space is again $\mathbb{R}^3$ and the basis is the standard one $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\}$.

Isn't the answer to this question is equivalent to finding a map that takes in two vector as input and spits out a real number or more precisely

$$\sigma : T_p \mathbb{R}^3 \times T_p \mathbb{R}^3 \rightarrow \mathbb{R} \quad (30)$$

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \mapsto \sigma\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \text{force per unit area on the face A at p.} \quad (31)$$

In addition, if I impose the linearity etc. then isn't σ a tensor of type $\binom{0}{2}$ i.e., takes two vectors and map them to real numbers? Of course to find this object one needs to understand the dynamics or the equations of motion etc. But, σ gives us a way to store the information about the force that is distributed over the whole cube. By the very definition of tensor, then we have $\sigma \in \mathcal{T}_2^0$ and we can write it as follows

$$\sigma := \sum_{i,j=1}^3 \sigma_{ij} dx^i \otimes dx^j, \quad x^1 = x, x^2 = y, x^3 = z \quad (32)$$

and $\sigma_{ij} = \sigma\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$ i.e., σ_{ij} describes the force per unit area in the direction of $\frac{\partial}{\partial x^j}$ on the face whose normal vector is $\frac{\partial}{\partial x^i}$. σ is called the **Stress Tensor** and a vital object in structural engineering. Can you think of any other examples of tensor you have seen before?

23 Metric

Now we want to define a special type of tensor that would be used to define the concept of length on a manifold. Let us look at the heuristics. First we go back to the concept of length of a vector on $\mathbb{R}^2$. Suppose I give you a vector $X = X^1 e_1 + X^2 e_2$ and ask to compute its length squared $|X|^2$. What do you do? You take the inner product of X with itself i.e.,

$$\begin{aligned} |X|^2 &= \langle X, X \rangle = \langle X^1 e_1 + X^2 e_2, X^1 e_1 + X^2 e_2 \rangle \\ &= (X^1)^2 \langle e_1, e_1 \rangle + X^1 X^2 \langle e_1, e_2 \rangle + X^2 X^1 \langle e_2, e_1 \rangle + (X^2)^2 \langle e_2, e_2 \rangle \\ &= \begin{bmatrix} X^1 & X^2 \end{bmatrix} \underbrace{\begin{bmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle \\ \langle e_2, e_1 \rangle & \langle e_2, e_2 \rangle \end{bmatrix}}_G \begin{bmatrix} X^1 \\ X^2 \end{bmatrix}. \end{aligned} \quad (33)$$

In all fairness, we can actually write $|X|^2$ the product of the vector X twice with the matrix G or

$$|X|^2 = G_{ij} X^i X^j. \quad (34)$$

As it happens on $\mathbb{R}^n$ in the usual rectangular chart, G_{ij} is simply the identity matrix and it contains the inner product of the basis vectors which are chosen to be orthonormal. Now how did we construct the basis of the tangent vector? We used curve. More precisely, we pushed forward the unit vector on $\mathbb{R}^n$ by the diffeomorphism φ^{-1} . However, as we saw on the midterm, diffeomorphisms do not necessarily preserve lengths/angles. So the transition from the standard basis to the tangent space

$$\{e_1, \dots, e_m\} \xrightarrow{(\varphi^{-1})^*} \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m} \right) \quad (35)$$

loses the nice orthogonality. What exactly is G doing here? It takes two copies of X and spits out a real number $G_{ij} X^i X^j$. Isn't it simply a tensor of type $\binom{0}{2}$? To generalize to the manifold, we can define this precisely as a $\binom{0}{2}$ tensor that has the similar property as the matrix G_{ij} on $\mathbb{R}^n$.

Definition 23.1 (Metric). The metric on a manifold M is a **symmetric** non-degenerate $\binom{0}{2}$ tensor field. More precisely at each $p \in M$, it is the linear map

$$g : T_p M \times T_p M \rightarrow \mathbb{R} \quad (36)$$

$$(V, W) \mapsto g(V, W) \quad (37)$$

that verifies (a) $g(V, W) = g(W, V)$, (b) $\forall W \in T_p M, g(V, W) = 0 \implies V = 0$. The second condition is called non-degenerate property of g .

Lemma 23.2. The non-degenerate property of g implies $\det(g_{ij}) \neq 0$ i.e., g_{ij} constitute an invertible matrix.

Proof. Recall the definition of non-degenerate property: $\forall W \in T_p M, g(V, W) = 0 \implies V = 0$. In particular we can choose W to be the basis set $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$. Therefore, we have

$$g(V, \frac{\partial}{\partial x^j}) = \sum_{i,k} (g_{ik} dx^i \otimes dx^k) \left(\sum_l V^l \frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^j} \right) = \sum_{i,k,l} g_{ik} V^l \delta_l^i \delta_j^k = \sum_{i,j} g_{ij} V^i \quad (38)$$

Now if $g(V, \frac{\partial}{\partial x^j})$ were to be zero then we have using the preceding calculations

$$g(V, \frac{\partial}{\partial x^j}) = \sum_{i,j} g_{ij} V^i = 0 \quad (39)$$

But this is nothing but a system of linear equations for V^i which have a solution $V^i = 0$ if and only if $\det(g_{ij}) \neq 0$. Therefore we are done. $\square$

In particular, I can use the metric g to define an inner product on T_pM .

Definition 23.3 (Inner product on T_pM). An inner product on T_pM is defined by the metric g as follows

$$\langle \cdot, \cdot \rangle_{T_pM} : T_pM \times T_pM \rightarrow \mathbb{R} \quad (40)$$

$$(V, W) \mapsto \langle V, W \rangle := g(V, W). \quad (41)$$

Remark. Since g is a $\binom{0}{2}$ tensor with the symmetry and non-degeneracy property, $\langle \cdot, \cdot \rangle$ defined through the metric g automatically verifies the axioms of an inner product. Notice one difference here though. We have not defined the norm or inner product to be positive definite i.e., if I interpret the metric components g_{ij} to be $m \times m$ matrix, then it is not necessarily positive definite.

Remark. Notice that the norm squared of a vector $X \in T_pM$ is simply $|X|^2 := g(X, X)$.

Now let us look at some explicit calculations regarding the metric g . First of all g is a $\binom{0}{2}$ tensor and therefore we can write it using theorem 22.8 as follows

$$g := \sum_{i,j} g_{ij} dx^i \otimes dx^j, \quad g_{ij} := g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \quad (42)$$

Now due to symmetry of g , we have $g\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right) = g\left(\frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^k}\right)$. Now $g\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right) = g_{kl}$ and $g\left(\frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^k}\right) = g_{lk}$ and therefore $g_{lk} = g_{kl}$. Now what does the inner product $\langle V, W \rangle$ look like explicitly? Let us use the definition of tensor and evaluate

$$\langle V, W \rangle = g(V, W) = \left(\sum_{i,j} g_{ij} dx^i \otimes dx^j\right) \left(\sum_k V^k \frac{\partial}{\partial x^k}, \sum_l W^l \frac{\partial}{\partial x^l}\right). \quad (43)$$

Now use the definition of the tensor product (definition 22.7) to yield

$$\begin{aligned} \left(\sum_{i,j} g_{ij} dx^i \otimes dx^j\right) \left(\sum_k V^k \frac{\partial}{\partial x^k}, \sum_l W^l \frac{\partial}{\partial x^l}\right) &= \sum_{i,j,k,l} g_{ij} V^k W^l dx^i \left(\frac{\partial}{\partial x^k}\right) dx^j \left(\frac{\partial}{\partial x^l}\right) \\ &= \sum_{i,j,k,l} g_{ij} V^k W^l \delta_k^i \delta_l^j = \sum_{ij} g_{ij} V^i W^j \end{aligned} \quad (44)$$

where we have used the usual relation $dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta_j^i = 1$ if $i = j$ and 0 otherwise. Therefore in terms of the components, we have

$$\langle V, W \rangle = \sum_{i,j} g_{ij} V^i W^j. \quad (45)$$

Remark. What is $\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^i} \rangle$? This is by definition of inner product $g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^i}\right) = g_{ii}$. Now in a chart we constructed the vectors spanning tangent vector T_pM by push forwarding the basis $\{e_i\}_{i=1}^m$ on the chart in $\mathbb{R}^n$ by the inverse diffeomorphism φ^{-1} . In other words, $\frac{\partial}{\partial x^i} = \varphi_*^{-1}(e_i)$ which of course do not remain orthogonal to themselves since a general diffeomorphism does not preserve the inner product as you have already shown in the mid-term (except when the transformation is orthogonal, but for a manifold this definitely won't be the case). Therefore, the functions g_{ij} could be very far from identity in general (In 230A/B, you will see that there exists a unique chart where you can write g as identity matrix at a point (and at point only at once) which is called the geodesic normal chart, but you don't have to worry about that now).

Remark (Einstein summation notation). From now on we do not write the summation notation $\sum$ all the time. Whenever there are repeated indices, they are meant to be summed over. For example $\sum_{ij} g_{ij} V^i W^j$ will simply be denoted by $g_{ij} V^i W^j$.

Now we use the metric to construct an isomorphism between T_pM and T_p^*M . This is called **Musical Isomorphism**.

Theorem 23.4. The map

$$\mu : T_p M \rightarrow T_p^* M \quad (46)$$

$$V \mapsto \mu(V) := g(V, \cdot) \quad (47)$$

is an isomorphism between $T_p M$ and $T_p^* M$.

Proof. First we show that $g(V, \cdot)$ is an element of $T_p^* M$. Let us work explicitly in a chart $\{x^i\}_{i=1}^m$. In this chart, we have

$$g(V, \cdot) = g_{ij} dx^i \otimes dx^j (V^k \frac{\partial}{\partial x^k}, \cdot) = g_{ij} V^i dx^j \quad (48)$$

which is clearly an element of $\mathcal{T}_1^0$ or $T_p^* M$ by theorem 22.3. In other words $g(V, \cdot) = \alpha$ where $\alpha = \alpha_i dx^i$, $\alpha_i = g_{ij} V^j$. Now we show that the map μ is an isomorphism. We only need to show injectivity since we are in finite dimensions and both $T_p M$ and $T_p^* M$ have same vector space dimension. Suppose for $X, Y \in T_p M$ we have $\beta = g(X, \cdot)$ and $\gamma = g(Y, \cdot)$ such that $\beta = \gamma$. We want to show that this implies $X = Y$ i.e., we want to show the following

$$g(X, \cdot) = g(Y, \cdot) \implies X = Y. \quad (49)$$

Now we use the definition of g i.e., using linearity

$$g(X, \cdot) = g(Y, \cdot) \implies g(X - Y, \cdot) = 0 \implies X - Y = 0 \implies X = Y \quad (50)$$

due to the non-degenerate property of g . $\square$

Definition 23.5. The inverse metric g^{-1} of g is defined by the following relation

$$(g^{-1})^{ik} g_{kj} = \delta_j^i \quad (51)$$

Lemma 23.6. $g^{-1} := (g^{-1})^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$ defines an isomorphism between $T_p^* M$ and $T_p M$ as follows

$$\lambda : T_p^* M \rightarrow T_p M \quad (52)$$

$$\alpha \mapsto \lambda_\alpha = g^{-1}(\alpha, \cdot). \quad (53)$$

Proof. Exercise. Use the definition 23.5 of g^{-1} and prove it in similar manner as theorem 23.4. $\square$

24 Lengths

Now we are in a place to define the lengths of curves on a manifold once we have a metric. Note that if I give you a metric you can construct an inner product. Now to define the lengths of curves, let us recall the definition of a curve on a manifold M . In a chart $\{x^i\}_{i=1}^m$, a curve c is the following map you have seen thousand times

$$c : (0, 1) \rightarrow M \quad (54)$$

$$t \mapsto c(t) = (x^1(t), x^2(t), x^3(t), \dots, x^m(t)). \quad (55)$$

Of course, to define a curve, you might need multiple charts but due to the compatibility of chart property, this does not make a difference. First, we compute the tangent vector to this curve. The tangent vector $V = \frac{dc}{dt}$ reads in components

$$V^i = \frac{dx^i}{dt} \quad (56)$$

i.e.,

$$V = V^i \frac{\partial}{\partial x^i} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}. \quad (57)$$

Now we prove the following lemma

Lemma 24.1. Let $t \mapsto c(t) \subset M$ be a curve on M . The tangent vector V reads

$$V = \frac{d}{dt}. \quad (58)$$

Proof. Let us consider a C^∞ function $f : M \rightarrow \mathbb{R}$. Now evaluate the following

$$\frac{df(c(t))}{dt} = df|_{c(t)} \cdot \frac{dc}{dt} = \frac{dx^i}{dt} \frac{\partial f(c(t))}{\partial x^i} = V f(c(t)) \quad (59)$$

and therefore $V = \frac{d}{dt}$ since the above holds for any C^∞ function f . $\square$

Now once we have the tangent vector defined at each point on the curve, we can define the length of the curve as usual

$$l_c := \int_0^1 \left| \frac{dc}{dt} \right| dt = \int_0^1 |V| dt. \quad (60)$$

Now using the metric g , we can compute the squared norm of $|V|$

$$|V|^2 = g(V, V). \quad (61)$$

At this point, we have not said anything about the definiteness of the metric g i.e., sign of $g(V, V)$. Now we make this choice and define a Riemannian metric.

Definition 24.2 (Riemannian Metric). g is called a Riemannian metric if $g(V, V) \geq 0$ and g verifies all the properties of being a metric.

With this definition of a Riemannian metric, the concept of length becomes the usual length concept we are used to. More specifically, we have the following expression for the length of the curve c

$$l_c = \int_0^1 \sqrt{g(V, V)} dt, \quad (62)$$

where $g(V, V)$ is evaluated at each point along the curve. Once we have defined the notion of length of a curve, we can define the distance between two points on a smooth manifold.

24.1 Distances on a Manifold

What is the next step? We want to be able to compute lengths/distances.

Given points $p, q \in M$, how do we compute the distance between p and q ? You want to compute the distances between all possible paths from p to q and take the minimum.

Recall that on the bonus problem in PSET 2, we showed that the straight line gives the shortest distance between 2 points in $\mathbb{R}^n$. We did so by writing the length function and minimizing it by making the derivative vanish. We will be able to use similar steps on a general Riemannian manifold, but it will involve some extra steps.

Let $c : \mathbb{R} \rightarrow M$ be a curve on the m -manifold M given by $c(t) = (x^1(t), \dots, x^m(t))$. Then the length is

$$\ell_c = \int_0^1 \left| \frac{dc}{dt} \right| dt$$

The tangent vector is $v := \frac{dc}{dt} = \left(\frac{dx^1}{dt}, \dots, \frac{dx^m}{dt} \right)$. Suppose we choose a basis $\{\frac{\partial}{\partial x^i}\}_{i=1}^m$ on the tangent space. Then we can write

$$v = v^i \frac{\partial}{\partial x^i} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i}$$

Then we can write the length of v as

$$\begin{aligned} \left| \frac{dc}{dt} \right| &= |v| = \sqrt{\langle v, v \rangle} \\ &= \sqrt{\sum \left\langle \frac{dx^i}{dt} \frac{\partial}{\partial x^i}, \frac{dx^j}{dt} \frac{\partial}{\partial x^j} \right\rangle} \\ &= \sqrt{\sum \frac{dx^i}{dt} \frac{dx^j}{dt} \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle} \\ &= \sqrt{\sum \frac{dx^i}{dt} \frac{dx^j}{dt} g_{ij}} \end{aligned}$$

And so we can rewrite the length as

$$\ell_c = \int_0^1 \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt$$

25 (11/22) : (TODO:)

25.1 Some logistics

PSET 5 will drop tonight, due in 2 weeks. Puskar will hold a review session next Friday instead of regular office hours. It will be recorded. Exam will be the full 3 hours, will have more questions than the midterm (but proportional to the exam duration)

25.2 Towards an Inner product

We have the following diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ \downarrow & & \downarrow \\ T_P M & \xrightarrow{f_* := df} & T_{f(P)} N \end{array}$$

where f is a diffeomorphism. The bottom map acts by

$$[f_* X]^i = \sum_{j=1}^m \frac{\partial f^i}{\partial x^j} X^j$$

Recall we also defined the Riemannian metric to satisfy the following: it is a bilinear map $g|_P : T_P M \times T_P M \rightarrow \mathbb{R}$ that sends $(x, y) \mapsto \langle x, y \rangle$, and satisfies

1. $g|_P(x, x) > 0$ if $x \neq 0$
2. $g|_P(0, 0) = 0$
3. $g|_P(x, y) = g|_P(y, x)$

Remark. This turns out to not be sufficient to define a proper metric: what is missing is the Triangle inequality. We might return to this at some point.

On $\mathbb{R}^n$, for vectors A, B we have $\langle B, A \rangle = B^T A = B^T(I)A$. Therefore, the metric on Euclidean space is just the identity matrix.

Definition 25.1. A Riemannian manifold is a pair (M, g) of a manifold and a Riemannian metric on that manifold.

The inner product of two arbitrary vectors $x^i \partial_i$ and $y^j \partial_j$ (recall Einstein notation):

$$\langle x^i \partial_i, y^j \partial_j \rangle = x^i y^j \langle \partial_i, \partial_j \rangle = x^i y^j g_{ij}$$

Definition 25.2 (Distance). Given points $p, q \in M$, consider a family F of curves $c : \mathbb{R} \rightarrow M$ such that $c(0) = p$ and $c(1) = q$. Then the distance between p and q is given by

$$d(p, q) = \inf_{c \in F} \ell_c$$

Why is the distance defined as an infimum rather than a minimum? This is because the “shortest” path may not be attainable!

Example 25.3. Consider the disk with a closed disk in the center removed, so that we have a hole in the center without the boundary (Figure 15). Then the “shortest” path would pass through the boundary of the hole, but that is not a valid path that stays on the manifold.

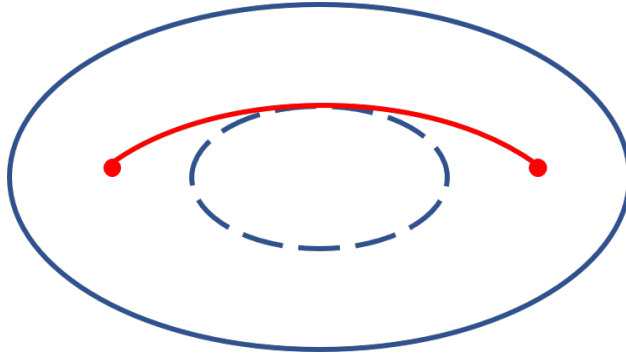


Figure 15: The shortest path may not be attainable

Assume for now that the distance function is well defined. Then the triangle inequality $d(P, Q) \leq d(P, R) + d(R, Q)$ should hold. But what can go wrong? Things can go very wrong in a space that does not have unique limits.

Example 25.4 (Breakdown of triangle inequality). Suppose we have a sequence of points $\{u_n\}$ which has two limits $p \neq q$. Then for any small distance $\varepsilon/2$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then $d(p, u_n) < \varepsilon/2$ and $d(q, u_n) < \varepsilon/2$. Using the triangle inequality, we would have

$$d(p, q) < d(p, u_n) + d(u_n, q) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Since ε can be any arbitrarily small positive number, we must have $d(p, q) = 0$. But for a distance function d , $d(p, q) = 0 \iff p = q$, a contradiction.

The moral of the story: triangle inequality does not hold on non-Hausdorff manifolds. You have already seen this in the PSET4.

From now on, we assume the Hausdorff property of our manifolds.

25.3 Computing the distance

Now we have a working definition of the distance $d(p, q)$. But how do we compute it? Suppose $c(0) = p$ and $c(1) = q$

$$\begin{aligned} \inf_c \int_0^1 \left| \frac{dc}{dt} \right| dt &= \inf_c \int_0^1 \sqrt{g_{ij}(x) \frac{dx^i}{dt} \frac{dx^j}{dt}} dt \\ &\geq \int_0^1 \inf_c \sqrt{g_{ij}(x) \frac{dx^i}{dt} \frac{dx^j}{dt}} dt \end{aligned} \quad (\text{By Fatou's lemma})$$

And notice that $\inf \sqrt{f(x)} = \sqrt{\inf f(x)}$ for positive f . Therefore we just need to minimize the quantity (you were asked in the PSET2 why minimizing length and energy are equivalent and you worked out on $\mathbb{R}^n$ that indeed we can get the same result by minimizing the energy)

$$\int_0^1 g_{ij}(x) \frac{dx^i}{dt} \frac{dx^j}{dt} dt = \int_0^1 g(v, v) dt$$

But this is precisely the energy E_c of the curve! To minimize it, we can take the derivative with respect to the transverse direction x to the family of curves $c(t)$:

$$\begin{aligned} \nabla_x E_c &= \int_0^1 \nabla_x g(v, v) dt \\ &= \int_0^1 \nabla_x \langle v, v \rangle dt \\ &= 2 \int_0^1 \langle \nabla_x v, v \rangle dt \end{aligned}$$

However, at this point the expression $\nabla_x v$ is nonsensical! Here we are trying to take a derivative of a vector field. Let's unpack things:

$$\nabla_x v = \lim_{dx \rightarrow 0} \frac{v(x + dx) - v(x)}{dx}$$

But notice that $v(x + dx) \in T_{x+dx}M$ whereas $v(x) \in T_x M$, so they lie in different tangent spaces! In order for this derivative to make sense, we need a way to compare these spaces. We will do so next week using the notion of connections.

Before going to define the connections in rigorous detail, let us first define a concept we already know. Let us collect all the tangent spaces at every point on the manifold. The tangent bundle is defined in section 20.2. Its essentially the disjoint union of tangent spaces. Now before going into defining the connection, let us define another entity that will come out to be handy. Recall that the vectors on a manifold are constructed as differential operators. Therefore we can think of them as acting on functions as well. In fact that was the main property of derivations i.e., acting on C^∞ functions on M . In other words by construction, we have for a vector field $V = V^i \frac{\partial}{\partial x^i}$

$$Vf = V^i \frac{\partial f}{\partial x^i} \quad (63)$$

Using this idea we can define what is called a commutator $[\cdot, \cdot]$ of two vector fields V and W .

Definition 25.5 (Commutator of vector fields). The commutator $[\cdot, \cdot]$ is the map

$$[\cdot, \cdot] : T_p M \times T_p M \rightarrow T_p M \quad (64)$$

$$(V, W) \mapsto [V, W] \quad (65)$$

that is $\mathbb{R}$ -linear on each slot (called bi-linear) i.e., for $\lambda, \mu \in \mathbb{R}$, $[\lambda V, \mu W] = \lambda \mu [V, W]$ that verifies $[X, X] = 0$ and the Jacobi identity i.e., $X, Y, Z \in T_p M$

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \quad (66)$$

Forgetting about all the technicalities, let us explicitly compute what a bracket looks like for two vector fields. First act $[V, W]$ on a C^∞ function f

$$\begin{aligned} [V, W]f &= V(W(f)) - W(V(f)) = V^i \frac{\partial}{\partial x^i} (W^j \frac{\partial f}{\partial x^j}) - W^j \frac{\partial}{\partial x^j} (V^i \frac{\partial f}{\partial x^i}) \\ &= V^i \frac{\partial W^j}{\partial x^i} \frac{\partial f}{\partial x^j} + V^i W^j \frac{\partial^2 f}{\partial x^i \partial x^j} - W^j \frac{\partial V^i}{\partial x^j} \frac{\partial f}{\partial x^i} - W^j V^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\ &= V^i \frac{\partial W^j}{\partial x^i} \frac{\partial f}{\partial x^j} - W^j \frac{\partial V^i}{\partial x^j} \frac{\partial f}{\partial x^i} \end{aligned} \quad (67)$$

where we have used the fact that the partial derivatives commute i.e., $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$. Now remember that repeated indices are summed over i.e.,

$$V^i \frac{\partial W^j}{\partial x^i} \frac{\partial f}{\partial x^j} := \sum_{i,j} V^i \frac{\partial W^j}{\partial x^i} \frac{\partial f}{\partial x^j} \quad (68)$$

and

$$W^j \frac{\partial V^i}{\partial x^j} \frac{\partial f}{\partial x^i} := \sum_{i,j} W^j \frac{\partial V^i}{\partial x^j} \frac{\partial f}{\partial x^i} = \sum_{i,j} W^i \frac{\partial V^j}{\partial x^i} \frac{\partial f}{\partial x^j} \quad (69)$$

since i, j are summed over if we interchange them both then nothing changes. Therefore we have

$$[V, W]f = (V^i \frac{\partial W^j}{\partial x^i} - W^i \frac{\partial V^j}{\partial x^i}) \frac{\partial f}{\partial x^j} \quad (70)$$

i.e.,

$$[V, W] = (V^i \frac{\partial W^j}{\partial x^i} - W^i \frac{\partial V^j}{\partial x^i}) \frac{\partial}{\partial x^j} \quad (71)$$

i.e., $[V, W] \in T_p M$ and $[V, W]^j = (V^i \frac{\partial W^j}{\partial x^i} - W^i \frac{\partial V^j}{\partial x^i})$. There is a very deep meaning of this brackets called **Lie-Derivatives** of vector fields. The mathematical interpretation of this is asked the Miscellaneous problem set.

These are all the concepts that we need to define a connection

26 (11/29) : Derivatives of Tangent Vectors

(I missed the first 15 minutes bit it seems mostly be recap - If I missed anything important please let me know!)

26.1 Connections and Connection Coefficients

Let $\chi(M)$ be the space of C^∞ vector fields on M , and vectors $X, Y \in \chi(M)$. On the chart $\{x\}$, we have:

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

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

We want to define the notion of “ $\nabla_X Y$ ”. Recall that the naive definition which we tried at the end of the previous lecture does not work because we are trying to compare vectors from different tangent spaces!

Definition 26.1 (Covariant connection/derivative). The covariant connection is given by the following map

$$\begin{aligned} \nabla : \chi(M) \times \chi(M) &\longrightarrow \chi(M) \\ (X, Y) &\longmapsto \nabla_X Y \end{aligned}$$

which satisfies the following conditions

1. C^∞ -(bi)linearity: for $f_1, f_2 \in C^\infty$, we have

$$\nabla_{f_1 X_1 + f_2 X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y$$

and

$$\nabla_X (c_1 Y_1 + c_2 Y_2) = c_1 \nabla_X Y_1 + c_2 \nabla_X Y_2$$

(c_1, c_2 are constants and Note that fX is the pointwise product of functions: $(fX)(x) = f(x)X(x)$, and likewise for similar expressions)

2. Leibniz rule/product rule.

$$\nabla_X (fY) = f \nabla_X Y + X[f]Y$$

where $X[f] = X^i \frac{\partial}{\partial x^i} f$ is the directional derivative.

What does $\nabla_X Y$ actually mean though? Let us unpack things, using the shorthand $\partial_i = \frac{\partial}{\partial x^i}$:

$$\begin{aligned} \nabla_X Y &= \nabla_{X^i \partial_i} Y^j \partial_j \\ &= X^i \nabla_{\partial_i} Y^j \partial_j && \text{(By linearity in first argument)} \\ &= X^i (Y_j \nabla_{\partial_i} \partial_j + \partial_i[Y^j] \partial_j) && \text{(By product rule + linearity in second argument)} \\ &= X^i \left(Y^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} + \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} \right) \end{aligned}$$

The second summand is easy to compute, since they are usual directional derivatives. The problem is the first term: How do we compute $\nabla_{\partial/\partial x^i} \partial/\partial x^j$? Let’s think back to Frenet curves for a moment. Recall the frame equations

$$\frac{de_i}{dt} = \omega_{ij} e_j, \quad \omega_{ij} := \left\langle \frac{de_i}{dt}, e_j \right\rangle$$

There, we were trying to understand the rate of change of the basis of the Frenet frame.

Quote of the day: “Knitting is a diffeomorphism”

Recall that we defined ω_{ij} to be this inner product between the frame vectors and their derivatives. In a very similar fashion, we can define the ∇ terms in terms of “connection coefficients”.

Definition 26.2 (Connection coefficients). We can write

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$$

Γ_{ij}^k are the connection coefficients.

Question 26.3. Why do we call it connections?

Answer. We are connecting the tangent spaces together, and we want to understand how the basis vectors change as we move across the manifold. The Γ_{ij}^k terms record precisely this information.

For instance, in Euclidean space, when we move the base point x , the basis of the tangent space does not change.

Rewriting, we have

$$\nabla_X Y = X^i \left(Y^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} + \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} \right)$$

Recalling that indices are just dummy variables, we can replace the j 's in the second term with k 's. So,

$$\begin{aligned} \nabla_X Y &= X^i \left(Y^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} + \frac{\partial Y^k}{\partial x^i} \frac{\partial}{\partial x^k} \right) \\ &= \left(\Gamma_{ij}^k X^i Y^j + X^i \frac{\partial Y^k}{\partial x^i} \right) \frac{\partial}{\partial x^k} \\ &= (\nabla_X Y)^k \frac{\partial}{\partial x^k} \end{aligned}$$

where to get to the last line we observe that the left factor is just a scalar component.

26.2 Metric compatibility

So now, how do we actually compute Γ_{ij}^k ? Naively speaking, Γ should depend both on the manifold M and the metric g on the manifold.

We will see in the next class that this can be computed from metric compatibility

$$\begin{aligned} \nabla_X \langle A, B \rangle &= \langle \nabla_X A, B \rangle + \langle A, \nabla_X B \rangle \\ \nabla_X g &= 0 \quad \forall X \in TM \end{aligned}$$

27 (12/1): Last Lecture!

Here is the plan for today: We will first wrap up with geodesics and related concepts. In the second half, we will look at some applications.

27.1 Metric Compatibility

A recap of last lecture: Given M , an m -manifold (which we assume has all the nice properties we want), if $X, Y \in \chi(M)$ are C^∞ vector fields, then we define the connection $\nabla_X Y \in \chi(M)$ to be

$$(X^k \partial_k Y^i + \Gamma_{jk}^i X^j Y^k) \frac{\partial}{\partial x^i}$$

We will look at how we can compute Γ , but we will not go into the full details.

Γ depends on the manifold M , intuitively. We need to make a choice (M, g) of the metric in order to define a Riemannian manifold. Recall that this metric gives us an inner product on the space, and therefore a way to define lengths. We can define the connection coefficients by assuming that ∇ is compatible with g . What this means, is that if we take $X, Y \in \chi(M)$ (Recall that $X = X^i \frac{\partial}{\partial x^i}$, and so on), such that the inner product is $\langle X, Y \rangle = g(X, Y)$, then the following holds:

Definition 27.1 (Metric Compatibility). A connection ∇ and a metric g are metric compatible if for any $Z \in \chi(M)$, we have

$$\nabla_Z(g(X, Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

for any choice of $X, Y \in \chi(M)$.

Remark. When ∇_Z acts on a scalar function, it just results in the directional derivative

Let us actually explain the definition 28.1 in a bit more detail. First of all, recall the definition of the metric g . It's the following map

$$g : T_p M \times T_p M \rightarrow \mathbb{R}, \quad (72)$$

$$(X, Y) \mapsto g(X, Y) \quad (73)$$

i.e., $g(X, Y)$ is a real number that depends on p i.e., it's a function on the manifold M . Therefore $\nabla_Z(g(X, Y))$ is simply the directional derivative of the function $g(X, Y)$ in the direction of Z i.e., $Z(g(X, Y)) = Z^i \frac{\partial}{\partial x^i}(g(X, Y))$. Now $(\nabla_Z g)$ is acting by ∇_Z on the left-hand side of the previous map i.e., acting on $T_p \times T_p M$ or on a tensor of type $\binom{0}{2}$ and returning a new tensor $\nabla_Z g$ of the same type $\binom{0}{2}$ or $\mathcal{T}_2^0$ i.e., more generally, we have the connection Δ as the following map on the space $\mathcal{T}_r^s$ of $\binom{r}{s}$ tensors

$$\nabla : T_p M \times \mathcal{T}_r^s \rightarrow \mathcal{T}_r^s \quad (74)$$

$$(X, T) \mapsto \nabla_X T. \quad (75)$$

Therefore $(\nabla_Z g)(X, Y) \neq \nabla_Z(g(X, Y))$ obviously. Hope this makes things more clear.

We can relate this back to an orthonormal frame. Recall that if we had $\frac{d}{dt}\langle e_1, e_2 \rangle = 0$, then we can move the derivative in and write $\langle \frac{de_1}{dt}, e_2 \rangle + \langle e_1, \frac{de_2}{dt} \rangle = 0$. We can see this by writing out the derivatives explicitly in the coordinates: **(TODO:)** sum of $g_{ij}e_1^i e_2^j$.

In the general case, g is not constant, and so we cannot pass the derivative through. So in general, we would expect

$$\nabla_Z(g(X, Y)) = (\nabla_Z g)(X, Y) + g(\nabla_Z X, Y) + g(X, \nabla_Z Y)$$

Then the condition of metric compatibility is the statement that $\nabla_Z g = 0$.

27.2 Obtaining the Connection from a Metric

Theorem 27.2. Assuming metric compatibility holds, then for a chart x ,

$$\Gamma_{jk}^i = \frac{1}{2}(g^{-1})^{i\ell} \left(\frac{\partial g_{j\ell}}{\partial x^k} + \frac{\partial g_{\ell k}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^\ell} \right)$$

where $(g^{-1})^{i\ell}$ is the inverse of the matrix g_{ij} :

$$(g^{-1})^{i\ell} g_{\ell j} = \delta_j^i$$

(In this context, we use superscripts to index into the entries of the inverse of a matrix)

Proof Sketch. Again using the notation $\partial_i = \frac{\partial}{\partial x^i}$. Recall that the coefficients are defined by the relation $\nabla_{\partial_j} \partial_k = \Gamma_{jk}^i \partial_i$. Then by metric compatibility, for vectors $\partial_j, \partial_k, \partial_\ell$

$$\nabla_{\partial_j} \langle \partial_k, \partial_\ell \rangle = \langle \nabla_{\partial_j} \partial_k, \partial_\ell \rangle + \langle \partial_k, \nabla_{\partial_j} \partial_\ell \rangle \quad (1)$$

We can also swap the roles of the three vectors and obtain

$$\nabla_{\partial_k} \langle \partial_\ell, \partial_j \rangle = \langle \nabla_{\partial_k} \partial_\ell, \partial_j \rangle + \langle \partial_\ell, \nabla_{\partial_k} \partial_j \rangle \quad (2)$$

$$\nabla_{\partial_\ell} \langle \partial_j, \partial_k \rangle = \langle \nabla_{\partial_\ell} \partial_j, \partial_k \rangle + \langle \partial_j, \nabla_{\partial_\ell} \partial_k \rangle \quad (3)$$

Let's look at (1) in detail: We can rewrite the connection terms:

$$\begin{aligned} \nabla_{\partial_j} \partial_k &= \Gamma_{jk}^{\text{Alex}} \partial_{\text{Alex}} \\ \nabla_{\partial_j} \partial_\ell &= \Gamma_{jk}^{\text{Connor}} \partial_{\text{Connor}} \end{aligned}$$

Substituting in, we have

$$\begin{aligned} \nabla_{\partial_j} \langle \partial_k, \partial_\ell \rangle &= \frac{\partial g_{k\ell}}{\partial x^j} = \langle \Gamma_{jk}^{\text{Alex}} \partial_{\text{Alex}}, \partial_\ell \rangle + \langle \partial_k, \Gamma_{jk}^{\text{Connor}} \partial_{\text{Connor}} \rangle \\ &= \Gamma_{jk}^{\text{Alex}} \langle \partial_{\text{Alex}}, \partial_\ell \rangle + \Gamma_{jk}^{\text{Connor}} \langle \partial_k, \partial_{\text{Connor}} \rangle \\ &= \Gamma_{jk}^{\text{Alex}} g_{\text{Alex}, \ell} + \Gamma_{jk}^{\text{Connor}} g_{k, \text{Connor}} \end{aligned}$$

This basically gives a system of linear equations which can be solved to recover the coefficients. □

Summary: $(M, g) + \text{metric compatibility}$ gives $\Gamma = \Gamma[g]$.

27.3 Geodesics

Definition 27.3 (Geodesics (informal definition)). Geodesics are local minimizers of length/energy.

Now using the idea of the connection, we want to obtain the equation that governs geodesic. Look at the figure 27.3. We want to find out the curve that has possibly the minimum length. Similar to PSET2, I can construct a family of curves with the same initial and final points by varying in the transverse direction. Also, assume that all the curves have finite lengths and energy. This amounts to considering a two-parameter family of curves

$$c : (0, 1) \times (0, 1) \rightarrow M \quad (76)$$

$$(t, s) \mapsto c(s, t), \quad (77)$$

$$\gamma : (0, 1) \times (0, 1) \rightarrow M \quad (78)$$

$$(t, s) \mapsto \gamma(s, t). \quad (79)$$

Let us denote the tangent vector fields as follows $V : \text{tangent to the family } c_{s,t}$, $W : \text{tangent to the family } \gamma_{s,t}$. Following the lemma 3.1, then we can write $V = \frac{\partial}{\partial t}$ and $W = \frac{\partial}{\partial s}$. In particular, we have

$$\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial s} \right] = 0 \quad (80)$$

following the definition of the commutator (you of course know that the partial derivatives commute).

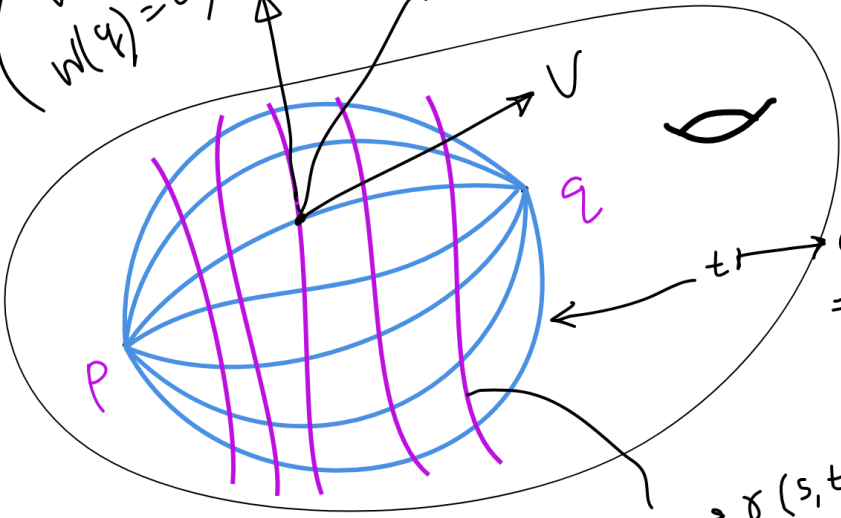


Figure 16: Notice the two families of curves.

Since the curves $t \mapsto c(s, t)$ are fixed at two ends, $W(0) = W(1) = 0$. Now we want to consider the minimization of length. But you know this is the same as minimizing energy (you have seen this in PSET2 so not much to say here). Let us denote the curves $t \mapsto c(s, t)$ by $c_s(t)$. The energy of the family of the curves $t \mapsto c_s(t)$ is defined as follows

$$E_{c_s} := \int_0^1 |V|^2 dt = \int_0^1 g(V, V) dt. \quad (81)$$

Now we want to compute the entity $\frac{\partial E_{c_s}}{\partial s}$ and set it to 0 by the first requirement of minimization

$$\begin{aligned} \frac{\partial E_{c_s}}{\partial s} &= \frac{\partial}{\partial s} \int_0^1 g(V, V) dt = \int_0^1 \frac{\partial}{\partial s} (g(V, V)) dt = \int_0^1 \nabla_W (g(V, V)) dt \\ &= 2 \int_0^1 g(\nabla_W V, V) dt = 2 \int_0^1 g(\nabla_V W + [W, V], V) dt = 2 \int_0^1 g(\nabla_V W, V) dt \\ &= 2 \int_0^1 (\nabla_V (g(W, V)) - g(W, \nabla_V V)) dt = 2 \int_0^1 \frac{\partial}{\partial t} (g(W, V)) dt - 2 \int_0^1 g(W, \nabla_V V) dt \\ &= 2[g(W, V)]_{t=0}^{t=1} - 2 \int_0^1 g(W, \nabla_V V) dt = -2 \int_0^1 g(W, \nabla_V V) dt, \end{aligned} \quad (82)$$

where we have used metric compatibility i.e., $\nabla_X g = 0 \forall X \in T_p M$ and the fact that ∇_V acts on scalars/functions simply by means of the directional derivative. In addition, $W(0) = W(1) = 0$. Therefore we have by setting $\frac{\partial E_c}{\partial s} = 0$

$$\frac{\partial E_c}{\partial s} = 0 = -2 \int_0^1 g(W, \nabla_V V) dt \quad (83)$$

by the preceding formula obtained for $\frac{\partial E_{c_s}}{\partial s}$. Now W is an arbitrary variation vector field transverse to the family of curves $t \mapsto c(s, t)$. Therefore by the non-degenerate property of the metric g ,

$$\int_0^1 g(W, \nabla_V V) dt = 0 \quad \forall W \text{ transverse to } c_s(t) \implies \nabla_V V = 0. \quad (84)$$

This leads to the following definition of a geodesic on a manifold (M, g) .

Definition 27.4 (Geodesics). Let $t \mapsto c(t) \subset M$ is a smooth curve connecting two points $p = c(0)$ and $q = c(1)$ such that its tangent vector field is V . This curve is called a geodesic if the tangent vector field V verifies $\nabla_V V = 0$.

In other words, if I can give you the condition $\nabla_V V = 0$, you should be able to prove that the geodesics are critical points of the energy functional and therefore of the length functional.

Interpretations: This condition corresponds to straight lines in Euclidean space. It also corresponds to the idea that the “acceleration” is 0.

If a curve $c(t)$ with $V := \frac{dc}{dt}$ is a geodesic (if it exists) connecting points $p, q \in M$, then it satisfies $\nabla_V V = 0$.

Let's understand the condition $\nabla_V V = 0$. Recall that $V = V^i \partial_i$. Therefore we can expand $\nabla_V V$ step by step following the definition of the connections.

$$\nabla_V V = \nabla_{V^i \frac{\partial}{\partial x^i}} (V^j \frac{\partial}{\partial x^j}) = V^i \nabla_{\frac{\partial}{\partial x^i}} (V^j \frac{\partial}{\partial x^j}) = V^i \left(\frac{\partial V^j}{\partial x^i} \frac{\partial}{\partial x^j} + V^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \right) \quad (85)$$

where the first term is just the application of the vector field $\frac{\partial}{\partial x^i}$ on the scalar function V^j (look back at the definition 27.1). The second term can be reduced by means of the definition of the connection coefficients i.e.,

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}. \quad (86)$$

Therefore we have

$$\nabla_V V = V^i \frac{\partial V^j}{\partial x^i} \frac{\partial}{\partial x^j} + V^i V^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} = (V^i \frac{\partial V^j}{\partial x^i} + \Gamma_{ik}^j V^i V^k) \frac{\partial}{\partial x^j} \quad (87)$$

where we have used the fact the repeated indices are summed over and they are simply dummies or placeholders i.e., $A^{ij}B_{ij}$ is the same as $A^{ji}B_{ji}$ (be cautious here that the order should be same i.e. $A^{ij}B_{ij} \neq A^{ij}B_{ji}$ i.e., you have to change both simultaneously; this should be obvious by writing down the sum; all the finite sum here so you can interchange the order of summation). Therefore $\nabla_V V = 0$ implies each component should be zero i.e.,

$$\nabla_V V = 0 \implies (\nabla_V V)^j = V^i \frac{\partial V^j}{\partial x^i} + \Gamma_{ik}^j V^i V^k = 0 \quad (88)$$

Now recall the tangent vector V and the lemma 24.1

$$V = V^i \frac{\partial}{\partial x^i} = \frac{dx^i}{dt} \frac{\partial}{\partial x^i} = \frac{d}{dt} \quad (89)$$

and therefore

$$\begin{aligned} (\nabla_V V)^j = V^i \frac{\partial V^j}{\partial x^i} + \Gamma_{ik}^j V^i V^k = 0 &\implies \frac{dV^j}{dt} + \Gamma_{ik}^j V^i V^k = 0 \\ &\implies \frac{d^2 x^j}{dt^2} + \Gamma_{ik}^j \frac{dx^i}{dt} \frac{dx^k}{dt} = 0. \end{aligned} \quad (90)$$

The last expression is known as proper acceleration. Therefore given initial point $x^i(0)$ and velocity $\frac{dx^i}{dt}(t=0)$, you can integrate this second order *ODE* for small enough parameter time t to find the desired geodesic $t \mapsto x^i(t) = c(t)$.

In $\mathbb{R}^n$, $g = \delta$ in usual chart, and so $\Gamma \sim \partial g = 0$. Therefore the equation is simplified to

$$\frac{d^2 x^i}{dt^2} = 0$$

This tells us that the geodesics on $\mathbb{R}^n$ are straight lines. You have already proven this in PSET2. Now in the case of a manifold, you have the extra Γ factor appearing.

Question 27.5. What happens if the manifolds have holes?

Answer. Then the minimization procedure that we discussed breaks down.

28 (11/17): (TODO:)

28.1 Change of Charts on the Tangent Space

Last lecture, we gave the definition of a basis of the tangent space $T_P M$ in terms of the vectors $\frac{\partial}{\partial x^i}$. We did all of this with respect to a choice of chart. What happens if we choose a different chart? Consider a map between charts $x \mapsto y(x)$. This map is a diffeomorphism.

We should keep in mind: a switch of coordinate charts does not change the vectors in the tangent space itself! We can look to $\mathbb{R}^3$ as an example: Given a vector in space, we can describe it in Euclidean coordinates, or using spherical coordinates, and so on. The vector v itself does not change when we change the chart. What changes is the basis, and by extension the components v^1, v^2, v^3 along the basis vectors.

Therefore, given a vector

$$v = \sum_{i=1}^m v_x^i(P) \frac{\partial}{\partial x^i}$$

and a change of chart $x \mapsto y(x)$, we expect to be able to write

$$v = \sum_{i=1}^m v_y^i(P) \frac{\partial}{\partial y^i}$$

How do v_x^i and v_y^i relate? We can use the chain rule! Notice that

$$\frac{\partial f}{\partial x^i} = \sum_{j=1}^m \frac{\partial f}{\partial y^j} \frac{\partial y^j}{\partial x^i} \quad \forall f \implies \frac{\partial}{\partial x^i} = \sum_{j=1}^m \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$$

And so,

$$\begin{aligned} v &= \sum_{i=1}^m v_x^i(P) \frac{\partial}{\partial x^i} \\ &= \sum_{i=1}^m v_x^i(P) \sum_{j=1}^m \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \\ &= \sum_{i=1}^m \sum_{j=1}^m \left(v_x^i \frac{\partial y^j}{\partial x^i} \right) \frac{\partial}{\partial y^j} \\ &= \sum_{j=1}^m \left(\sum_{i=1}^m v_x^i \frac{\partial y^j}{\partial x^i} \right) \frac{\partial}{\partial y^j} && \text{(exchanging finite sums is ok)} \\ &= \sum_{j=1}^m v_y^j \frac{\partial}{\partial y^j} \\ &\implies v_y^j = \sum_{i=1}^m v_x^i \frac{\partial y^j}{\partial x^i} \end{aligned}$$

Question 28.1. Why could we just compare the coefficients in the penultimate line?

Answer. That is because the $\frac{\partial}{\partial y^j}$'s are linearly independent. We saw this by acting with both sides on the coordinate functions y^k .

Now, notice that the form of v_y^j looks suspiciously like a matrix multiplication! If we let A be a matrix where $A_i^j = \frac{\partial y^j}{\partial x^i}$, then we can write

$$v_y^j = \sum_{i=1}^m A_i^j v_x^i$$

Question 28.2. Can we think of this as a Df map?

Answer. Yes! Recall that f is a map between manifolds. For a change of charts, this is just the special case of a map from a manifold to itself.

Remark. The fact that v is the “real” object and that the components v^i are “fake” mirrors our earlier study of curves. Recall that we discussed that the tangent vectors along a curve are not invariant under reparameterization. What is invariant is a geometric property, such as arclength and curvature.

Definition 28.3 (Einstein's Summation notation). To reduce clutter, we write

$$A_j^i v^j := \sum_{j=1}^m A_j^i v^j$$

A more complicated example:

$$A_{j\ell m}^i x^j y^k z^\ell w^m := \sum_j \sum_k \sum_\ell \sum_m A_{j\ell m}^i x^j y^k z^\ell w^m$$

Summation is implied by paired indices.

28.2 Metrics

Now recall: how did we define the basis of the tangent space? We looked at (equivalence classes of) curves and took a collection of tangent vector as the basis. Can we talk about the length of these curves on the manifold?

Let's look at $\mathbb{R}^2$. We have $d\ell = \sqrt{(dx)^2 + (dy)^2}$. Notice that we don't have any cross terms $dx dy$. Another way of writing Pythagoras theorem is the following:

$$d\ell^2 = \begin{pmatrix} dx & dy \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}$$

Why is the middle matrix the identity matrix? Because dx and dy are orthogonal. The precise definition of the matrix is given by

$$\begin{pmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle \\ \langle e_2, e_1 \rangle & \langle e_2, e_2 \rangle \end{pmatrix}$$

However, as we saw on the midterm, diffeomorphisms do not necessarily preserve lengths/angles. So the transition from the standard basis to the tangent space

$$\{e_1, \dots, e_m\} \xrightarrow{\varphi^{-1}} \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m} \right)$$

loses the nice orthogonality.

We end off by defining the notion of a metric. Let $X, Y \in T_P M$. Choose a chart x for the point P . Then we can write

$$\begin{aligned} X &= X^i(P) \frac{\partial}{\partial x^i} \\ Y &= Y^j(P) \frac{\partial}{\partial x^j} \end{aligned}$$

We want to define an inner product on the manifold. Heuristically, we can try defining

$$\langle X, Y \rangle = \langle X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^j} \rangle = X^i Y^j \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$$

There is no reason for $\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$ to be the identity matrix. So let us give it a name: $g_{ij} := \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle$.

Definition 28.4. A Riemannian metric at point $P \in M$ is a non-degenerate symmetric bilinear map

$$\begin{aligned} g_P : T_P M \times T_P M &\longrightarrow \mathbb{R} \\ (g_P)_{ij} &:= \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \Big|_P \end{aligned}$$

which satisfies $g_P(X, X) \geq 0 \quad \forall X \in T_P M$ and $g_P(X, X) = 0 \iff X = 0$.

29 Differentials in the space of matrices

Recall the differential on $\mathbb{R}^n$. You can of course do it on manifolds. You know the definition of the differential of a map $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$. The differential at $x \in U$ is a linear map denoted by the Jacobian $df|_x$ such that the following holds

$$df|_x : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (91)$$

$$h \mapsto df|_x \cdot h \quad (92)$$

and moreover

$$f(x + th) - f(x) = tdf|_x \cdot h + O(t^2) \quad (93)$$

for $t > 0$ sufficiently small. The entity $df|_x \cdot h$ called the differential of f in the direction of h (or the directional derivative) can also be computed as follows

$$df|_x \cdot h = \frac{d}{dt}f(x + th)|_{t=0}. \quad (94)$$

Here $\cdot$ represents the linear action of the differential $df|_x$ on h . The equation 94 just follows from differentiating 93 both sides with respect to t and setting $t = 0$. Using this definition, I want you to compute the derivative in the space of matrices. First, denote by $GL(k, \mathbb{R})$ the space of all invertible matrices (non-zero determinant) with real entries. Now since any matrix belonging to the set $GL(k, \mathbb{R})$ has k^2 real elements. The space $GL(k, \mathbb{R})$ is a subset of $\mathbb{R}^{k^2}$. Let $U \subset GL(k, \mathbb{R})$ open and f is the following map

$$f : U \subset GL(k, \mathbb{R}) \rightarrow V \subset GL(k, \mathbb{R}) \quad (95)$$

$$A \mapsto f(A) := A^2 \quad (96)$$

where A is a matrix and A^2 is the usual matrix multiplication AA . Using 94, compute the differential of f at identity $\mathbb{I}$ in the direction of any real $k \times k$ matrix H i.e., compute $df|_{\mathbb{I}} \cdot H$.

Answer. The differential of the map f at any $A \in U$ can be computed using the equation 94. More explicitly

$$df|_A : M(n, \mathbb{R}) \rightarrow M(n, \mathbb{R}) \quad (97)$$

$$H \mapsto df_A \cdot H = \frac{d}{dt}f(A + tH)|_{t=0} = \frac{d}{dt}(A + tH)(A + tH)|_{t=0} = \frac{d}{dt}(A^2 + tAH + tHA + t^2H^2)|_{t=0} \quad (98)$$

$$= AH + HA. \quad (99)$$

Therefore $df|_{\mathbb{I}} \dot{H} = H + H = 2H$. Here $M(n, \mathbb{R})$ denotes the space of real $n \times n$ matrices.

29.1 Applications

Why do we learn math? We want to acquire a toolkit to solve real-world problems.

- Biology

Humans are a perfect example of manifolds. Proteins are manifolds. Suppose we want to Xray a tumor - everything is curved, so we need differential geometry. In the brain, if we want to find the shortest distance between two points, we need to use geodesics! The straight-line path will pierce through the brain.

- Manifold learning

In machine learning we want to minimize the cost. This is the same as minimizing some metric. Suppose we have a 10-dimensional “statistical manifold” in a 100-dimensional space. We can use geodesics to reduce dimensions and get faster computations. Fisher metrics.

- Physics

Not much needs to be said here.

29.2 Course Feedback

Please provide your honest feedback on the course!

Thanks for attending the lectures and submitting your homework!