Differential Geometry Prospectus

Alex Nelson*

October 17, 2022

Contents

*This is a page from https://pqnelson.github.io/notebk/
Compiled: November 6, 2022 at 7:41pm (PST)

1 Tangent Vectors

1. The motivation underlying differential geometry is: we want to generalize vector calculus. A first step is to generalize the notion of a "vector". And really, if we look at a vector *in* vector calculus, there are two components to it:

1. The base point, and
2. The vector part.

Linear algebra studies "the vector part", assuming all the vectors live at the same base point (i.e., within the same vector space).

1.1. Remark. We will introduce the concepts, like paint on a canvas, in layers. The first pass ("primer") will be in $\mathbb{R}^3$ (but easily could work in $\mathbb{R}^n$ for any $n \in \mathbb{N}$). The second pass will be on a surface $\Sigma \subset \mathbb{R}^3$ (again, it could be generalized to a surface $\Sigma \subset \mathbb{R}^n$ for any $n \in \mathbb{N}$). The third pass "breaks the bottle", cutting the cord with an "ambient $\mathbb{R}^n$", and works with manifolds. Since this is a prospectus, we will only look at the first pass; this truly is a "primer".

2. Definition. A **"Tangent Vector in $\mathbb{R}^3$"** is an ordered pair $\boldsymbol{v}_{\boldsymbol{p}} := (\boldsymbol{p}, \boldsymbol{v})$ where $\boldsymbol{p} \in \mathbb{R}^3$ is the base point, and $\boldsymbol{v} \in \mathbb{R}^3$ is the vector part.

2.1. Remark. In ordinary vector algebra/calculus, we often ignore the base point (and whenever we add two vectors with different base points, we just transport them to the same base point). But in differential geometry, we must be more careful:

1. We can only add two tangent vectors if they live at the same base point. So $\boldsymbol{v}_{\boldsymbol{p}} + \boldsymbol{w}_{\boldsymbol{q}}$ makes sense provided $\boldsymbol{p} = \boldsymbol{q}$.
2. We have $\boldsymbol{v}_{\boldsymbol{p}} = \boldsymbol{v}_{\boldsymbol{q}}$ if and only if $\boldsymbol{p} = \boldsymbol{q}$.

In vector calculus, we typically freely transport tangent vectors along to share the same base point. This is not longer "free" in differential geometry, so care must be taken.

3. Definition. The tangent vectors at a given base point $\boldsymbol{p} \in \mathbb{R}^n$ form a vector space $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ called the **"Tangent Space of $\mathbb{R}^n$ at $\boldsymbol{p}$"**.

3.1. Remark. Note that $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ is isomorphic (as a vector space) to $\mathbb{R}^n$ as a vector space, but we must remember that at different base points $\boldsymbol{p} \neq \boldsymbol{q}$ we have $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n \neq \mathrm{T}_{\boldsymbol{q}}\mathbb{R}^n$ different tangent spaces.

1.1 Vector Fields

4. Definition. A **"Vector Field"** V on $\mathbb{R}^n$ assigns to each point $\boldsymbol{p} \in \mathbb{R}^n$ a tangent vector $V(\boldsymbol{p})$ at $\boldsymbol{p}$.
We call $\mathrm{Vect}(\mathbb{R}^n)$ the set of all vector fields on $\mathbb{R}^n$.

5. Structure of $\mathrm{Vect}(\mathbb{R}^n)$. Note that $\mathrm{Vect}(\mathbb{R}^n)$ is an infinite-dimensional vector space over the ring of [smooth] real-valued functions. If we let V, $W \in \mathrm{Vect}(\mathbb{R}^n)$, $\boldsymbol{p} \in \mathbb{R}^n$, and $f\colon \mathbb{R}^n \to \mathbb{R}$, observe $V(\boldsymbol{p}) \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$, and

$$(V + W)(\boldsymbol{p}) = V(\boldsymbol{p}) + W(\boldsymbol{p}) \tag{1.1a}$$

$$(fV)(\boldsymbol{p}) = f(\boldsymbol{p})V(\boldsymbol{p}). \tag{1.1b}$$

Note that the operations on the right-hand side of these equations are the vector addition and scalar multiplication in $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$, whereas the operations on the left-hand side are the newly-defined binary operators for $\mathrm{Vect}(\mathbb{R}^n)$.

The natural question to ask, what about a "basis vector field"?

6. Definition. There are n obvious vector fields $U_1, U_2, \ldots, U_n \in \mathrm{Vect}(\mathbb{R}^n)$ defined by

$$\begin{aligned} U_1(\boldsymbol{p}) &= (1,0,0,\ldots,0)_{\boldsymbol{p}} \\ U_2(\boldsymbol{p}) &= (0,1,0,\ldots,0)_{\boldsymbol{p}} \\ &\vdots \\ U_n(\boldsymbol{p}) &= (0,0,0,\ldots,1)_{\boldsymbol{p}} \in \mathrm{Vect}(\mathbb{R}^n) \end{aligned} \tag{1.2}$$

called the **"Natural Frame Field"**.

7. Definition. More generally, in $\mathbb{R}^n$, a **"Frame Field"** is a list of n vector fields $W_1, W_2, \ldots, W_n \in \mathrm{Vect}(\mathbb{R}^n)$ such that at each $p \in \mathbb{R}^n$ the vectors $W_1(\boldsymbol{p}), W_2(\boldsymbol{p}), \ldots, W_n(\boldsymbol{p})$ form a basis of $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$.

8. Lemma. *Given any $V \in \mathrm{Vect}(\mathbb{R}^n)$, we can write*

$$V = v_1U_1 + v_2U_2 + \cdots + v_nU_n, \tag{1.3}$$

where U_i are natural frame fields and $v_i\colon \mathbb{R}^n \to \mathbb{R}$ for $i = 1, 2, \ldots, n$.

Proof. At each $\boldsymbol{p} \in \mathbb{R}^n$, the tangent vector $V(\boldsymbol{p}) \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ can be written in coordinates as

$V(\boldsymbol{p})$

= | since $V(\boldsymbol{p})$ is a vector and can be written in components relative to a canonical basis (a)

$(v_1(\boldsymbol{p}), \ldots, v_n(\boldsymbol{p}))_{\boldsymbol{p}}$

= | because $U_1(\boldsymbol{p}), \ldots, U_n(\boldsymbol{p})$ form a [canonical] basis of $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ (b)

$v_1(\boldsymbol{p})U_1(\boldsymbol{p}) + \cdots + v_n(\boldsymbol{p})U_n(\boldsymbol{p})$

= | since $U_1, \ldots, U_n$ are vector fields, using scalar multiplication (§5) (c)

$(v_1U_1 + \cdots + v_nU_n)(\boldsymbol{p})$

Since we were using $\boldsymbol{p} \in \mathbb{R}^n$ arbitrary, we have $V = v_1U_1 + \cdots + v_nU_n$. □

8.1. Remark (Definition). The functions $v_i\colon \mathbb{R}^n \to \mathbb{R}$ are called the **"Coordinate Functions"** of V relative to the frame field U_i.

9. Definition. We say a vector field is **"Differentiable"** if all its coordinate functions (with respect to the natural frame field) are differentiable. Similarly, we call a vector field **"Smooth"** (C^∞) if its coordinate functions are C^∞ (i.e., they have all partial derivatives [including mixed partial derivatives] of all orders).

9.1. Remark. We could specify the subset of smooth vector fields as $C^\infty\,\mathrm{Vect}(\mathbb{R}^n)$, but we will implicitly assume everything is smooth from now on.

9.2. Remark. We could also work with C^k vector fields by demanding only the first $k \in \mathbb{N}$ order partial derivatives of the coordinate functions exist and be continuous. Or we could demand they be analytic functions, and we could work analytic vector fields, traditionally denoted $C^\omega\,\mathrm{Vect}(\mathbb{R}^n)$.

1.2 Directional Derivatives

10. Definition. Given a tangent $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$, we can use it to differentiate a smooth function $f \in C^\infty(\mathbb{R}^n)$. We define the **"Directional Derivative"** of f in the $\boldsymbol{v}_{\boldsymbol{p}}$ direction [or, the directional derivative of f with respect to $\boldsymbol{v}_{\boldsymbol{p}}$] as:

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p} + t\boldsymbol{v})\Big|_{t=0}, \tag{1.4}$$

which represents the rate of change of f in the $\boldsymbol{v}$ direction at the point $\boldsymbol{p}$.

10.1. Remark. This lets us think of a tangent vector as a map

$$\begin{aligned} \boldsymbol{v}_{\boldsymbol{p}} \colon C^{\infty}(\mathbb{R}^n) &\to \mathbb{R} \\ f &\mapsto \boldsymbol{v}_{\boldsymbol{p}}[f]. \end{aligned} \tag{1.5}$$

11. Example. Assume $f \in C^{\infty}(\mathbb{R}^3)$, $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3$, and let us write $\boldsymbol{p} = (p_x, p_y, p_z)$, $\boldsymbol{v}_{\boldsymbol{p}} = (v_x, v_y, v_z)_{\boldsymbol{p}}$. Then we can explicitly determine what the directional derivative of f with respect to $\boldsymbol{v}_{\boldsymbol{p}}$ looks like:

$$\boldsymbol{v}_{\boldsymbol{p}}[f]$$

$$= \quad | \quad \text{by Definition 10} \tag{a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p} + t\boldsymbol{v})\Big|_{t=0}$$

$$= \quad | \quad \text{unfold components} \tag{b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f(p_x + tv_x, p_y + tv_y, p_z + tv_z)\Big|_{t=0}$$

$$= \quad | \quad \text{chain-rule and linearity of derivative} \tag{c}$$

$$\left(\frac{\partial f}{\partial x}(\boldsymbol{p} + t\boldsymbol{v})\right) v_x + \left(\frac{\partial f}{\partial y}(\boldsymbol{p} + t\boldsymbol{v})\right) v_y + \left(\frac{\partial f}{\partial z}(\boldsymbol{p} + t\boldsymbol{v})\right) v_z\Big|_{t=0}$$

$$= \quad | \quad \text{evaluate at } t = 0\text{, then use the dot product} \tag{d}$$

$$(v_x, v_y, v_z) \cdot \left(\frac{\partial f}{\partial x}(\boldsymbol{p}), \frac{\partial f}{\partial y}(\boldsymbol{p}), \frac{\partial f}{\partial z}(\boldsymbol{p})\right)$$

$$= \quad | \quad \text{folding back into vector form} \tag{e}$$

$$\boldsymbol{v} \cdot (\boldsymbol{\nabla} f(\boldsymbol{p})).$$

This should look familiar: it's the directional derivative from vector calculus.

12. Definition. Let $f \in C^{\infty}(\mathbb{R}^n)$ and $V \in \mathrm{Vect}(\mathbb{R}^n)$. We can take **"Directional Derivative"** of f with respect to the vector *field* V denoted $V[f] \colon \mathbb{R}^n \to \mathbb{R}$ given by, for any $\boldsymbol{p} \in \mathbb{R}^n$,

$$(V[f])(\boldsymbol{p}) = V(\boldsymbol{p})[f]. \tag{1.6}$$

13. Proposition. *Let $f, g, h \in C^{\infty}(\mathbb{R}^n)$, $a, b \in \mathbb{R}$, and $V, W \in \mathrm{Vect}(\mathbb{R}^n)$.*

1. $(fV + gW)[h] = fV[h] + gW[h]$
2. *Linearity:* $V[af + bg] = aV[f] + bV[g]$
3. *Product rule:* $V[fg] = V[f]g + fV[g]$

1.3 Differential Forms

14. Definition. Given some (real) vector space W, a **"Covector"** (or *dual vector*) is a linear map $\varphi \colon W \to \mathbb{R}$.

15. Notation change: Indices, superscripts, subscripts. I am going to try to be consistent from now on, with coordinates to tangent vectors being written with superscipted indices, frame fields with subscripted indices (so $V = v^1 U_1 + v^2 U_2 + \cdots + v^n U_n$), covector bases written with superscripted indices, and components/coordinates of covectors with subscripted indices. This seems random and insane at first (and it probably is), but this is the convention which physicists use. It also lends itself to the Einstein summation convention, where we sum over repeated indices, for example:

$$v^i U_i := \sum_{i=1}^{n} v^i U_i. \tag{1.7}$$

We will use explicit summation in our notes, for explicit clarity.

16. Example. Consider $W = \mathbb{R}^3$ and define $\varphi\colon W \to \mathbb{R}$ by

$$\varphi\left(\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix}\right) = w^1. \tag{1.8}$$

It's linear (it's just the projection map onto the first coordinate, which is famously linear). Thus it's a covector.

17. Example. Consider $W = \mathbb{R}^3$ and define $\varphi\colon W \to \mathbb{R}$ by

$$\varphi\left(\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix}\right) = \alpha_1 w^1 + \alpha_2 w^2 + \alpha_3 w^3, \tag{1.9}$$

where α_1, α_2, $\alpha_3 \in \mathbb{R}$ are fixed constants. We see this is linear, and we could write it as:

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \alpha_1 w^1 + \alpha_2 w^2 + \alpha_3 w^3. \tag{1.10}$$

The row vectors are called covectors, the column vectors are vectors.

18. Definition. If $W = \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$, then a covector is called a **"Cotangent Vector"**. The space of all cotangent vectors to $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ is called the **"Cotangent Space with Base Point $\boldsymbol{p}$"** and denoted $\mathrm{T}^*_{\boldsymbol{p}}\mathbb{R}^n$.

19. Example. Let $\boldsymbol{p} \in \mathbb{R}^n$, define $\varphi\colon \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n \to \mathbb{R}$ by

$$\varphi(\boldsymbol{v}_{\boldsymbol{p}}) = v^1, \tag{1.11}$$

where $\boldsymbol{v}_{\boldsymbol{p}} = (v^1, \dots, v^n)_{\boldsymbol{p}}$. This is linear:

$$\begin{aligned} \varphi(a\boldsymbol{v}_{\boldsymbol{p}} + b\boldsymbol{w}_{\boldsymbol{p}}) &= \varphi\left((a\boldsymbol{v} + b\boldsymbol{w})_{\boldsymbol{p}}\right) && \text{(1.12a)} \\ &= (a\boldsymbol{v} + b\boldsymbol{w})^1 && \text{(1.12b)} \\ &= a\varphi(\boldsymbol{v}_{\boldsymbol{p}}) + b\varphi(\boldsymbol{w}_{\boldsymbol{p}}). && \text{(1.12c)} \end{aligned}$$

20. Definition. A **"One-Form"** (or "*Covector Field*") φ on $\mathbb{R}^n$ assigns to each point $p \in \mathbb{R}^n$ a covector $\varphi_{\boldsymbol{p}}\colon \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n \to \mathbb{R}$.

20.1. Remark. A one-form is one way of getting certain information out of the vector field.

If V is a vector field and φ is a one-form, we get (at each point $\boldsymbol{p} \in \mathbb{R}^n$):

$$\varphi_{\boldsymbol{p}}(V(\boldsymbol{p})) \in \mathbb{R}. \tag{1.13}$$

So at each point we get a number; so φ and V give us a function. That is, we may think of φ as a map

$$\begin{aligned} \varphi\colon\ &\mathrm{Vect}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n), \\ &V \mapsto \varphi(V), \end{aligned} \tag{1.14}$$

as defined by

$$\left(\varphi(V)\right)(\boldsymbol{p}) = \varphi_{\boldsymbol{p}}(V(\boldsymbol{p})). \tag{1.15}$$

The main way to obtain a 1-form is to take a differential (of a function).

21. Definition. Given a smooth function $f \in C^\infty(\mathbb{R}^n)$, define the **"Differential of f"** to be the map

$$\mathrm{d}f\colon\ \mathrm{Vect}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n) \tag{1.16}$$

given by, for any $V \in \mathrm{Vect}(\mathbb{R}^n)$,

$$\mathrm{d}f[V] = V[f], \tag{1.17}$$

i.e., given by the directional derivative of f in the direction of V (§12).

22. Differential forms are one-forms. In fact, $\mathrm{d}f$ really is a one-form. We check linearity, letting $V, W \in \mathrm{Vect}(\mathbb{R}^n)$ be arbitrary,

$$\begin{aligned} \mathrm{d}f[V+W] &= (V+W)[f] && (1.18a)\\ &= V[f] + W[f] && (1.18b)\\ &= \mathrm{d}f[V] + \mathrm{d}f[W] && (1.18c) \end{aligned}$$

for arbitrary $c \in \mathbb{R}$,

$$\begin{aligned} \mathrm{d}f[cV] &= (cV)[f] && (1.19a)\\ &= c(V[f]) && (1.19b)\\ &= c\,\mathrm{d}f[V] && (1.19c) \end{aligned}$$

For arbitrary $h\colon \mathbb{R}^n \to \mathbb{R}$ smooth,

$$\begin{aligned} \mathrm{d}f[hV] &= (hV)[f] && (1.20a)\\ &= h \cdot (V[f]) && (1.20b)\\ &= h\,\mathrm{d}f[V]. && (1.20c) \end{aligned}$$

23. Example. Let $x^1, \dots, x^n$ be the standard coordinate functions on $\mathbb{R}^n$. Then we see

$$\begin{aligned} \mathrm{d}x^j[U_i] &= U_i[x^j] && (1.21a)\\ &= \frac{\partial}{\partial x^i} x^j && (1.21b)\\ &= \delta_i{}^j = \delta^j{}_i && (1.21c) \end{aligned}$$

is the Kronecker delta $\delta^j{}_i = 0$ if $i \neq j$ and $\delta^j{}_i = 1$ if $i = j$. Then for any vector field $V \in \mathrm{Vect}(\mathbb{R}^n)$, we have,

$$\begin{aligned} \mathrm{d}x^j[V] &= \mathrm{d}x^j\left[\sum_i v^i U_i\right] && (1.22a)\\ &= \sum_i \mathrm{d}x^j[v^i U_i] && (1.22b)\\ &= \sum_i v^i\,\mathrm{d}x^j[U_i] && (1.22c)\\ &= \sum_i v^i \delta^j{}_i && (1.22d)\\ &= v^j. && (1.22e) \end{aligned}$$

So $\mathrm{d}x^j$ just picks out the j^{th} component of the vector field at the point.

24. Rosetta Stone. At this point, it's useful to write a "Rosetta Stone" relating "Stuff" and "Co-Stuff".

Stuff	Co-Stuff
• Tangent vector $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ (vectors with base points)	• Cotangent vectors $\phi_{\boldsymbol{p}} \in \mathrm{T}^*_{\boldsymbol{p}}\mathbb{R}^n$ is a linear map $\phi_{\boldsymbol{p}}\colon \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n \to \mathbb{R}$
• Vector Field V gives a tangent vector at each point $V(\boldsymbol{p}) \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$	• One-Forms, or Covector fields, ϕ assigns a cotangent vector $\phi_{\boldsymbol{p}}$ at each point $\boldsymbol{p} \in \mathbb{R}^n$

• Vector fields act on functions (by the directional derivative) to give new functions: $(V[f])(\boldsymbol{p}) = \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p}+t\boldsymbol{v})\Big\vert_{t=0}$	• One-forms act on vector fields to give functions $(\phi[V])(\boldsymbol{p}) = \phi_{\boldsymbol{p}}[V(\boldsymbol{p})]$. • Any smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$ gives a one-form $\mathrm{d}f$ by the rule $(\mathrm{d}f)_{\boldsymbol{p}}[\boldsymbol{v}_{\boldsymbol{p}}] = \boldsymbol{v}_{\boldsymbol{p}}[f]$, or $\mathrm{d}f[V] = V[f]$.
• The *natural frame field* on $\mathbb{R}^n$ is given by the standard coordinate vector fields $U_1(\boldsymbol{p}) = (1,0,\dots,0)_{\boldsymbol{p}}$ and so on. These are special because $U_i[f] = \frac{\partial f}{\partial x^i}(\boldsymbol{p})$.	• The *differentials of the coordinate functions* $x^1, \dots, x^n \mapsto \mathrm{d}x^1, \dots, \mathrm{d}x^n$ are special because $\mathrm{d}x^i[U_j] = \delta^i_j$.
• We proved in Lemma 8 that any vector field can be written as $V = \sum_i v^i U_i$ where the smooth functions v^i are the *coordinate functions of* V *relative to the frame field* U_i.	• We will show that any covector field (i.e., one-form) ϕ can be written as $\phi = \sum_i f_i \,\mathrm{d}x^i$ where the f_i are called the **"Coordinate functions of ϕ relative to the coframe field"** given by $\mathrm{d}x^j$.
Stuff	Co-Stuff

25. Proposition. *Every one-form is a $C^\infty(\mathbb{R}^n)$ linear combination of the* $\mathrm{d}x^i$.

Proof. Suppose ϕ is a one-form Then for any vector field $V \in \mathrm{Vect}(\mathbb{R}^n)$,

$$
\begin{aligned}
& \phi[V] \\
= & \quad \big\vert \ \text{because } V = \textstyle\sum_i v^i U_i && \text{(a)} \\
& \phi\left[\textstyle\sum_i v^i U_i\right] \\
= & \quad \big\vert \ \text{by linearity of } \phi && \text{(b)} \\
& \textstyle\sum_i v^i \phi[U_i] \\
= & \quad \big\vert \ \text{recall } \mathrm{d}x^i[V] = v^i && \text{(c)} \\
& \textstyle\sum_i \mathrm{d}x^i[V]\,\phi[U_i] \\
= & \quad \big\vert \ \text{commutativity of multiplication of real numbers} && \text{(d)} \\
& \textstyle\sum_i \phi[U_i]\,\mathrm{d}x^i[V] \\
= & \quad \big\vert \ \text{linearity} && \text{(e)} \\
& \left(\textstyle\sum_i \phi[U_i]\,\mathrm{d}x^i\right)[V].
\end{aligned}
$$

Since this is true for all V, that means $\phi = \sum_i \phi[U_i]\,\mathrm{d}x^i$. That's what we wanted to prove. In fact, we have an explicit formula for the coefficients. □

26. Corollary. *If $\phi = \mathrm{d}f$, then*

$$\mathrm{d}f = \sum_i \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i.$$

Proof. We can compute this directly,

$$\mathrm{d}f = \sum_i \mathrm{d}f[U_i]\,\mathrm{d}x^i \tag{1.23a}$$

$$= \sum_i U_i[f]\,\mathrm{d}x^i \tag{1.23b}$$

$$= \sum_i \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i. \tag{1.23c}$$

Hence the result. □

27. Definition: Differential Operator. We actually have something more. We have the **"Differential"** is a map

$$\mathrm{d}\colon \{\text{functions}\} \to \{\text{one-forms}\}, \tag{1.24a}$$

defined by

$$\mathrm{d}f := \sum_i \frac{\partial f}{\partial x^i}\mathrm{d}x^i. \tag{1.24b}$$

Let us now prove properties about the differential operator.

28. Proposition (Leibniz rule)**.** *For any $f, g \in C^\infty(\mathbb{R}^n)$, we have $\mathrm{d}(fg) = g\,\mathrm{d}f + f\,\mathrm{d}g$.*

Proof. Let $V \in \mathrm{Vect}(\mathbb{R}^n)$ be arbitrary, then we have

$$\begin{aligned}
\mathrm{d}(fg)[V] &= V[fg] && (1.25\text{a})\\
&= V[f]\,g + f\,V[g] && (1.25\text{b})\\
&= g\,\mathrm{d}f[V] + f\,\mathrm{d}g[V] && (1.25\text{c})\\
&= (g\,\mathrm{d}f + f\,\mathrm{d}g)[V]. && (1.25\text{d})
\end{aligned}$$

Since V was arbitrary, the result follows. □

29. Proposition (Chain rule)**.** *Let $f\colon \mathbb{R}^n \to R$ and $h\colon \mathbb{R} \to \mathbb{R}$. Then for any $\boldsymbol{p} \in \mathbb{R}^n$, $\mathrm{d}(h\circ f)_{\boldsymbol{p}} = h'(f(\boldsymbol{p}))\,\mathrm{d}f_{\boldsymbol{p}}$.*

Proof. Let $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$ be arbitrary, then

$$\begin{aligned}
\mathrm{d}(h\circ f)_{\boldsymbol{p}}[\boldsymbol{v}_{\boldsymbol{p}}] &= \boldsymbol{v}_{\boldsymbol{p}}[h\circ f] && (1.26\text{a})\\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}h(f(\boldsymbol{p}+t\boldsymbol{v}))\right|_{t=0} && (1.26\text{b})\\
&= \left.h'(f(\boldsymbol{p}+t\boldsymbol{v}))\frac{\mathrm{d}}{\mathrm{d}t}f(\boldsymbol{p}+t\boldsymbol{v})\right|_{t=0} && (1.26\text{c})\\
&= h'(f(\boldsymbol{p}))\boldsymbol{v}_{\boldsymbol{p}}[f] && (1.26\text{d})\\
&= h'(f(\boldsymbol{p}))\,\mathrm{d}f[\boldsymbol{v}_{\boldsymbol{p}}]. && (1.26\text{e})
\end{aligned}$$

Since this is for arbitrary tangent vectors, the result follows. □

30. Example. Let's work in $\mathbb{R}^2$, let $x^1 = x$ and $x^2 = y$. Consider

$$f(x,y) = x^2\sin(y) + y^3 x. \tag{1.27}$$

We compute the one-form $\mathrm{d}f$:

$$\begin{aligned}
&\mathrm{d}f\\
=\ &\quad|\ \text{unfold the definition of } f && (\text{a})\\
&\mathrm{d}(x^2\sin(y) + y^3x)\\
=\ &\quad|\ \text{linearity of d} && (\text{b})\\
&\mathrm{d}(x^2\sin(y)) + \mathrm{d}(y^3x)\\
=\ &\quad|\ \text{Leibniz rule} && (\text{c})\\
&\big(x^2\,\mathrm{d}(\sin(y)) + \sin(y)\,\mathrm{d}(x^2)\big) + \big(y^3\,\mathrm{d}(x) + x\,\mathrm{d}(y^3)\big)\\
=\ &\quad|\ \text{calculus} && (\text{d})\\
&(x^2\cos(y)\,\mathrm{d}y + 2x\sin(y)\,\mathrm{d}x) + (y^3\,\mathrm{d}x + 3y^2x\,\mathrm{d}y)\\
=\ &\quad|\ \text{gather terms} && (\text{e})
\end{aligned}$$

$$(2x\sin(y) + y^3)\,\mathrm{d}x + (x^2\cos(y) + 3y^2x)\,\mathrm{d}y.$$

This coincides with $\mathrm{d}f = (\partial_x f)\,\mathrm{d}x + (\partial_y f)\,\mathrm{d}y$.

31. Example. Working over $\mathbb{R}^2$ with $x^1 = x$, $x^2 = y$, consider $g(x,y) = \cos\sqrt{xy}$. Compute the one-form $\mathrm{d}g$. The trick is to consider $h(z) = \cos\sqrt{z}$ and $f(x,y) = xy$, so $g = h\circ f$. This is because we will use the chain-rule for differential forms, which requires computing $h'(f(x,y))$ and $\mathrm{d}f$. We start with

$$\begin{aligned}
&\mathrm{d}g\\
=\ &\mid\ \text{unfold the definition of } g && \text{(a)}\\
&\mathrm{d}(h\circ f)\\
=\ &\mid\ \text{chain rule (§29)} && \text{(b)}\\
&\left.\frac{\mathrm{d}h(z)}{\mathrm{d}z}\right|_{z=f(x,y)}\mathrm{d}f
\end{aligned}$$

Now we need to compute $h'(f(x,y))$,

$$\begin{aligned}
&\left.\frac{\mathrm{d}h(z)}{\mathrm{d}z}\right|_{z=f(x,y)}\\
=\ &\mid\ \text{unfold definition of } h && \text{(c)}\\
&\left.\frac{\mathrm{d}\cos\sqrt{z}}{\mathrm{d}z}\right|_{z=f(x,y)}\\
=\ &\mid\ \text{chain rule} && \text{(d)}\\
&\left.-\sin\sqrt{z}\frac{\mathrm{d}\sqrt{z}}{\mathrm{d}z}\right|_{z=f(x,y)}\\
=\ &\mid\ \text{power rule} && \text{(e)}\\
&\left.-\sin\sqrt{z}\frac{1}{2\sqrt{z}}\right|_{z=f(x,y)}\\
=\ &\mid\ \text{substitution, simplify numerator} && \text{(f)}\\
&\frac{-\sin\sqrt{f(x,y)}}{2\sqrt{f(x,y)}}\\
=\ &\mid\ \text{unfold definition of } f && \text{(g)}\\
&\frac{-\sin\sqrt{xy}}{2\sqrt{xy}}
\end{aligned}$$

We similarly can compute $\mathrm{d}f$,

$$\begin{aligned}
&\mathrm{d}f\\
=\ &\mid\ \text{unfold definition of } f && \text{(h)}\\
&\mathrm{d}(xy)\\
=\ &\mid\ \text{Leibniz rule} && \text{(i)}\\
&x\,\mathrm{d}y + y\,\mathrm{d}x
\end{aligned}$$

We can combine everything together:

$$\begin{aligned}
&\mathrm{d}g\\
=\ &\mid\ \text{chain rule (§29)}
\end{aligned}$$

$$\left.\frac{\mathrm{d}h(z)}{\mathrm{d}z}\right|_{z=f(x,y)}\mathrm{d}f$$

$$= \quad \Big| \text{ from our previous calculations} \tag{j}$$

$$\frac{-\sin\sqrt{xy}}{2\sqrt{xy}}(x\,\mathrm{d}y + y\,\mathrm{d}x).$$

Thus we conclude

$$\mathrm{d}f = \frac{\sin\sqrt{xy}}{2\sqrt{xy}}(x\,\mathrm{d}y + y\,\mathrm{d}x), \tag{1.28a}$$

or rearranging factors,

$$\mathrm{d}f = \left(\frac{-1}{2}\frac{y}{\sqrt{xy}}\sin\sqrt{xy}\right)\mathrm{d}x + \left(\frac{-1}{2}\frac{x}{\sqrt{xy}}\sin\sqrt{xy}\right)\mathrm{d}y. \tag{1.28b}$$

1.4 Algebra of Differential Forms

32. We know how to add one-forms together, and we know how to multiply one-forms by "scalars" (i.e., smooth functions). But what about the multiplication of one-forms *by other one-forms?*

Now let's formally define a definition of multiplication of one-forms called the **"Wedge Product"** denoted "$\wedge$". By "formal", we mean we'll just produce a series of rules that the wedge product satisfies.

Given two one-forms ϕ and ψ, their **"Wedge Product"**

$$\phi\wedge\psi \tag{1.29}$$

is called a **"Two-Form"**. More generally, we can multiply two-forms by scalars and add them together, so a generic two-form looks like

$$f\phi\wedge\psi + \cdots + g\mu\wedge\lambda. \tag{1.30}$$

More generally, we could take the wedge product of 3 one-forms to produce a three-form, or the wedge produce of n one-forms to produce an n-form.

33. Example. On $\mathbb{R}^{15}$, we have $\mathrm{d}x^1\wedge\mathrm{d}x^2 + \sin(x^1)\,\mathrm{d}x^2\wedge\mathrm{d}x^{10}$ be a perfectly good 2-form.

34. Axioms of Wedge Product. We stipulate the wedge product satisfies the following axioms:

1. Associativity: $\phi\wedge(\psi\wedge\mu) = (\phi\wedge\psi)\wedge\mu$
2. Left distributivity over addition: $\phi\wedge(\psi+\mu) = \phi\wedge\psi + \phi\wedge\mu$
3. Anticommutativity on 1-forms: $\phi\wedge\psi = -\psi\wedge\phi$.
4. Scalar multiplication: $f\cdot(\phi\wedge\psi\wedge\dots) = (f\phi)\wedge\psi\wedge(\dots)$.

35. Consequences. From these axioms, we have two consequences:

1. Right distributivity over addition: $(\psi+\mu)\wedge\phi = \psi\wedge\phi + \mu\wedge\phi$
2. Scalar distributivity: $(f\phi)\wedge\psi = f\cdot(\phi\wedge\psi) = \phi\wedge(f\psi)$
3. Nilpotence: $\phi\wedge\phi = 0$.

Proof. These are straightforward calculations.

$$(\psi+\mu)\wedge\phi$$

$$= \quad \Big| \text{ anti-commutativity} \tag{a}$$

$$-\phi\wedge(\psi+\mu)$$

$$= \quad \Big| \text{ left distributivity} \tag{b}$$

$$-\phi\wedge\psi - \phi\wedge\mu$$

$$= \quad \Big| \text{ anti-commutativity} \tag{c}$$

$$\psi\wedge\phi + \mu\wedge\phi.$$

Similarly, for scalar distributivity,

$$
\begin{aligned}
& (f\phi)\wedge\psi \\
= \quad & \big| \quad \text{scalar multiplication} && \text{(d)} \\
& f\cdot(\phi\wedge\psi) \\
= \quad & \big| \quad \text{anticommutativity} && \text{(e)} \\
& f\cdot(-\psi\wedge\phi) \\
= \quad & \big| \quad \text{scalar multiplication} && \text{(f)} \\
& -(f\psi)\wedge\phi \\
= \quad & \big| \quad \text{anticommutativity} && \text{(g)} \\
& \phi\wedge(f\psi).
\end{aligned}
$$

Nilpotence follows from anticommutativity, since the only number for which $\phi\wedge\phi = -\phi\wedge\phi$ is when $\phi\wedge\phi = 0$. (If you don't believe me, add $\phi\wedge\phi$ to both sides and divide by 2.) □

36. Proposition. *Any two-form on $\mathbb{R}^3$ may be written in the form*

$$f\,\mathrm{d}x\wedge\mathrm{d}y + g\,\mathrm{d}y\wedge\mathrm{d}z + h\,\mathrm{d}z\wedge\mathrm{d}x, \tag{1.31}$$

where $f, g, h \in C^\infty(\mathbb{R}^3)$.

Proof. Let's show if ϕ, ψ are one-forms, then $\phi\wedge\psi$ can be written like Eq (1.31). Let

$$\phi = \sum_{i=1}^{3} f_i\,\mathrm{d}x^i \tag{1.32a}$$

$$\psi = \sum_{j=1}^{3} g_j\,\mathrm{d}x^j \tag{1.32b}$$

We can compute:

$$
\begin{aligned}
& \phi\wedge\psi \\
= \quad & \big| \quad \text{unfolding } \phi,\ \psi && \text{(a)} \\
& \left(\textstyle\sum_{i=1}^{3} f_i\,\mathrm{d}x^i\right)\wedge\left(\textstyle\sum_{j=1}^{3} g_j\,\mathrm{d}x^j\right) \\
= \quad & \big| \quad \text{distributivity, linearity, anticommutativity} && \text{(b)} \\
& (f_1g_2 - f_2g_1)\,\mathrm{d}x^1\wedge\mathrm{d}x^2 + (f_2g_3 - f_3g_2)\,\mathrm{d}x^2\wedge\mathrm{d}x^3 + (f_3g_1 - f_1g_3)\,\mathrm{d}x^3\wedge\mathrm{d}x^1
\end{aligned}
$$

Since an arbitrary two-form is some linear combination of wedge products of one-forms, we just have to use this result and collect terms. □

36.1. Remark. The preceding formula looks a lot like the cross-product of vectors, and it would be if:

1. The $\mathrm{d}x^i$ were "orthonormal basis vectors"
2. We replaced the wedge products with the following basis vectors $\mathrm{d}x^1\wedge\mathrm{d}x^2 \to \boldsymbol{e}_3$, $\mathrm{d}x^2\wedge\mathrm{d}x^3 \to \boldsymbol{e}_1$, $\mathrm{d}x^3\wedge\mathrm{d}x^1 \to \boldsymbol{e}_2$.

36.2. Remark. Similar results hold for k-forms in $\mathbb{R}^n$.

37. Example. Every 2-form on $\mathbb{R}^4$ can be written as a linear combination of $\mathrm{d}x^1\wedge\mathrm{d}x^2$, $\mathrm{d}x^1\wedge\mathrm{d}x^3$, $\mathrm{d}x^1\wedge\mathrm{d}x^4$, $\mathrm{d}x^2\wedge\mathrm{d}x^3$, $\mathrm{d}x^2\wedge\mathrm{d}x^4$, $\mathrm{d}x^3\wedge\mathrm{d}x^4$.

38. Example. In general, every k-form on $\mathbb{R}^n$ is a linear combination of $\binom{n}{k}$ basis k-forms. In particular, for $k > n$, every k-form is zero.

39. Wedge Product of Forms. The wedge product is a map

$$\begin{aligned} \wedge\colon \{k\text{-forms}\} \times \{\ell\text{-forms}\} &\to \{(k+\ell)\text{-forms}\} \\ (\omega, \lambda) &\mapsto \omega \wedge \lambda. \end{aligned} \tag{1.33}$$

In fact, it's useful to think of smooth functions as "0-forms" to complete the picture, where we define the wedge product as just the scalar product $f \wedge \phi = f\phi$. So we have, in $\mathbb{R}^n$,

- 0-forms: smooth functions
- 1-forms: covector fields
- ...
- n-forms

The "k" in "k-form" is called the **"Degree"** of the form, written $\deg(\phi)$.

40. Theorem. *For any differential forms ϕ, ψ, we have*

$$\phi \wedge \psi = (-1)^{(\deg\ \phi)(\deg\ \psi)}(\psi \wedge \phi). \tag{1.34}$$

Proof. It suffices to prove this for monomials $\phi = f\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and $\psi = g\,\mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_\ell}$. The trick is to do this by induction on ℓ (the degree of ψ).

Base Case: $\ell = 1$, we see

$$\phi \wedge \psi = (f\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}) \wedge (g\,\mathrm{d}x^{j_1}). \tag{1.35}$$

We can move the g out in front without a problem, then we must move the $\mathrm{d}x^{j_1}$ in front of k one-forms, which will cost us a factor of $(-1)^k$, giving us:

$$\phi \wedge \psi = (-1)^k fg\,\mathrm{d}x^{j_1} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}. \tag{1.36}$$

And since $\deg(\phi)\deg(\psi) = k$, we're good.

Inductive Hypothesis: we now assume this works for arbitrary $\ell \in \mathbb{N}$.

Inductive Case: we now will prove this is the case for $(\ell+1)$-forms. We write $\psi = \psi^{(\ell)} \wedge \mathrm{d}x^{j_{\ell+1}}$. Then we have

$$\phi \wedge \psi = \phi \wedge (\psi^{(\ell)} \wedge \mathrm{d}x^{j_{\ell+1}}). \tag{1.37}$$

We invoke associativity to write the right-hand side as

$$\phi \wedge \psi = (\phi \wedge \psi^{(\ell)}) \wedge \mathrm{d}x^{j_{\ell+1}}. \tag{1.38}$$

Now look, we have precisely our base case. What's more: $g = 1$ in our current situation. So we use the inductive hypothesis to rewrite

$$(\phi \wedge \psi^{(\ell)}) \wedge \mathrm{d}x^{j_{\ell+1}} = ((-1)^{(\deg\ \phi)\ell}\psi^{(\ell)} \wedge \phi) \wedge \mathrm{d}x^{j_{\ell+1}} \tag{1.39}$$

and invoking the base case to rewrite the right-hand side as

$$\begin{aligned} ((-1)^{(\deg\ \phi)\ell}\psi^{(\ell)} \wedge \phi) \wedge \mathrm{d}x^{j_{\ell+1}} &= (-1)^{(\ell+\deg\ \phi)1}\mathrm{d}x^{j_{\ell+1}} \wedge ((-1)^{(\deg\ \phi)\ell}\psi^{(\ell)} \wedge \phi) \\ &= (-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}\mathrm{d}x^{j_{\ell+1}} \wedge (\psi^{(\ell)} \wedge \phi). \end{aligned} \tag{1.40}$$

We have to move $\mathrm{d}x^{j_{\ell+1}}$ behind the $\psi^{(\ell)}$, so we use associativity

$$(-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}\mathrm{d}x^{j_{\ell+1}} \wedge (\psi^{(\ell)} \wedge \phi) = (-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}(\mathrm{d}x^{j_{\ell+1}} \wedge \psi^{(\ell)}) \wedge \phi. \tag{1.41}$$

Then we can use the inductive hypothesis setting $\phi = \mathrm{d}x^{j_{\ell+1}}$

$$\begin{aligned} (-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}(\mathrm{d}x^{j_{\ell+1}} \wedge \psi^{(\ell)}) \wedge \phi &= (-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}((-1)^{\ell}\psi^{(\ell)} \wedge \mathrm{d}x^{j_{\ell+1}}) \wedge \phi \\ &= (-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell}((-1)^{\ell}\psi^{(\ell+1)}) \wedge \phi. \end{aligned} \tag{1.42}$$

Now we just need to prove that

$$(-1)^{(\ell+\deg\ \phi)+(\deg\ \phi)\ell+\ell} = (-1)^{(\deg\ \phi)(\ell+1)}. \tag{1.43}$$

But this is trivial, since $(\ell + \deg\ \phi) + (\deg\ \phi)\ell + \ell = 2\ell + \deg(\phi) + \ell \cdot \deg(\phi)$ and $(-1)^{2\ell} = 1$. ∎

1.5 Exterior Derivative

41. The goal is to "extend" the differential d so it can work on any differential form. We know if $f \in C^\infty(\mathbb{R}^n)$, then f is a zero-form and $\mathrm{d}f$ is a one-form. In general we want to differential to produce $(k+1)$-forms from k-forms,

$$\mathrm{d}\colon \{k\text{-forms}\} \to \{(k+1)\text{-forms}\} \tag{1.44}$$

which satisfy

1. d acting on a zero-form remains the same as before.
2. d is $\mathbb{R}$-linear — we can pull out constants but not functions,
3. Graded Leibniz property: for any differential forms ϕ and ψ, we want $\mathrm{d}(\phi\wedge\psi) = (D\phi)\wedge\psi + (-1)^{\deg(\phi)}\phi\wedge \mathrm{d}\psi$.
4. For *any* form ϕ, we want $\mathrm{d}(\mathrm{d}\phi) = 0$.

Graded-Leibniz property may seem odd, but consider the following situation: let ϕ and ψ be one-forms, we better get the same answer if we compute $\mathrm{d}(\phi\wedge\psi)$ or $\mathrm{d}(-\psi\wedge\phi)$.

▸ **Exercise 1.** Prove, without the graded Leibniz property, $\mathrm{d}(\phi\wedge\psi) \neq \mathrm{d}(-\psi\wedge\phi)$.

42. Regarding the $\mathrm{d}(\mathrm{d}\phi) = 0$ property, the analogy which should spring to mind is that

$$\mathrm{d}f \sim \text{gradient of } f,$$

and

$$\mathrm{d}(\text{1-form}) \sim \text{curl},$$

so

$$\mathrm{d}(\mathrm{d}f) \sim \nabla\times(\boldsymbol{\nabla} f) = 0. \tag{1.45}$$

43. Computing Exterior Derivative. How do we calculate the exterior derivative of a k-form ω? (Well, if it's a k-form on $\mathbb{R}^k$, it's zero, so let's assume we're on $\mathbb{R}^n$ for $n > k$.)

1. Write ϕ as a linear combination of monomials like $f\,\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}$
2. Use linearity to do the calculation term-by-term.
3. For each term, it looks like: $\mathrm{d}(f\,\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}) = \mathrm{d}f\wedge(\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k})$. (Since the other term from graded Leibniz is $(-1)^0 f\wedge\mathrm{d}(\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}) = f\wedge(0) = 0$.)

Exercises

1. Let $\boldsymbol{v} = (-2,0,4)$, $\boldsymbol{p} = (2,1,-3)$. Working directly from the definition, compute the directional derivative $\boldsymbol{v_p}[f]$ where
   (a) $f = z/y$
   (b) $f = \ln(1+x^2)\tan(z)$.
2. Let $V = -z^2U_1 + \cos(xy)U_3$, $f = x^2y^5z^3$, $g = (x/y)\sin(z)$. Compute:
   (a) $V[f]$
   (b) $V[fg]$.
3. Prove or find a counter-example: if V, W are vector fields on $\mathbb{R}^n$ such that every $f\colon\mathbb{R}^n\to\mathbb{R}$ we have $V[f] = W[f]$, then $V = W$.

Homework on Frame Fields

Mathematics 116 — Differential Geometry
Spring 2008
Derek Wise

Define two vector fields on $\mathbb{R}^2$ by

$$E_1 = \frac{1}{4}U_1 + \frac{(x^1)^2}{4}U_2$$
$$E_2 = U_2$$

(Note: the superscript "1" is just a label, indicating the first of the coordinate functions x^1, x^2 of $\mathbb{R}^2$. The superscript outside parentheses indicates the *square* of the coordinate.)

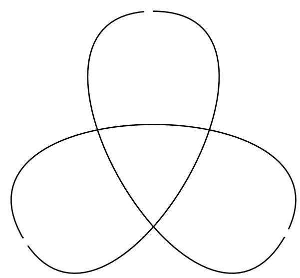
1. Draw the vector fields E_1, E_2
2. Recall that a **frame field** on $\mathbb{R}^n$ is a list of n smooth vector fields such that at each point they give a basis of the tangent space. Show E_1, E_2 is a frame field on $\mathbb{R}^2$.
3. Frame fields in differential geometry play essentially the same role that bases do in linear algebra. For example, given a vector field $V = v^1U_1 + v^2U_2$ (where v^1, v^2 are smooth functions), defined in terms of the natural frame field, we can rewrite V in terms of the frame field E_1, E_2. Do it!
4. The 1-forms $\mathrm{d}x^i$ on $\mathbb{R}^n$ are called **dual** to the vector fields U_j in the natural frame field, meaning that $\mathrm{d}x^i[U_j] = \delta^i_j$. That is, $\mathrm{d}x^i[U_j]$ equals the constant function 1 if $i = j$, or the constant function 0 if $i \neq j$. Find a pair of 1-forms θ^1, θ^2 that are dual to the vector fields E_1, E_2. That is, θ^1, θ^2 should satisfy $\theta^i[E_j] = \delta^i_j$. [Hint: write the θ^i as a linear combination of $\mathrm{d}x^i$ and solve for the coefficients.]

2 Differential Geometry of Curves in $\mathbb{R}^3$ (or $\mathbb{R}^n$)

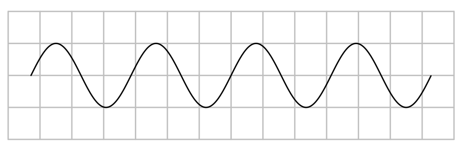
44. The idea is that we have introduced the basic gadgetry of differential geometry, but in the setting of $\mathbb{R}^n$. Now we will consider curves in $\mathbb{R}^n$, and use the gadgetry we've introduced to study properties of curves (for example, how to vector fields on a curve, and what do they tell us). This appears in classical mechanics (especially Lagrangian and Hamiltonian mechanics).

45. Definition. An **"Unparametrized Curve"** in $\mathbb{R}^3$ is a "one-dimensional subset" of points.

46. Example. The doodle below is a closed unparametrized curve — "closed" meaning it forms a "loop" (eventually):



47. Example. Here is a happy open unparametrized curve — "open" meaning it is "not closed":



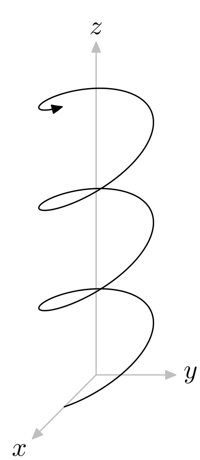
48. Definition. Let $I = (a, b)$ be an open interval (it is possible $a = -\infty$, or $b = +\infty$, or both). We define a **"(Parametrized) Curve"** in $\mathbb{R}^3$ to be a smooth function $\alpha\colon I \to \mathbb{R}^3$

48.1. Remark. Henceforth, we will reserve the term "curve" for an unparametrized curve, and "path" for a parametrized curve.

- "Path" = "Parametrized Curve"
- "Curve" = "Unparametrized Curve".

In fact, almost always we care about curves, so unless otherwise stated, all curves are parametrized.

49. Example. Consider the **"Elliptic Helix"**, a parametrized curve $\alpha(t) = (a\cos(t), b\sin(t), ct)$ where $a, b, c \in \mathbb{R}$ are positive constants, and $t \in [0, +\infty)$ looks like:



We could consider a different parametrization of the same curve, like $\beta(t) = (a\cos(3t), b\sin(3t), 3ct) = \alpha(3t)$. How do we know this is the same curve? Well, one way is to establish a bijection of points $\beta(t/3) = \alpha(t)$ for all $t \in [0, \infty)$.

50. Definition. Let $\alpha\colon I \to \mathbb{R}^3$ be a path with components $(\alpha_1, \alpha_2, \alpha_3)$. The **"Velocity"** of α at time $t \in I$ is the tangent vector

$$\alpha'(t) = \left(\frac{\mathrm{d}\alpha(t)}{\mathrm{d}t}\right)_{\alpha(t)} \in \mathrm{T}_{\alpha(t)}\mathbb{R}^3 \tag{2.1a}$$

at base point $\alpha(t)$, whose components are

$$\alpha'(t) = \left(\frac{\mathrm{d}\alpha_1(t)}{\mathrm{d}t}, \frac{\mathrm{d}\alpha_2(t)}{\mathrm{d}t}, \frac{\mathrm{d}\alpha_3(t)}{\mathrm{d}t}\right). \tag{2.1b}$$

50.1. Remark. The velocity defines a sort of vector field, but defined only on the curve and not all of $\mathbb{R}^3$.

51. Example. For the elliptical helix, $\alpha(t) = (a\cos(t), b\sin(t), ct)$ for $t \in \mathbb{R}$, we have its velocity be

$$\alpha'(t) = (-a\sin(t), b\cos(t), c)_{\alpha(t)}. \tag{2.2}$$

52. Example. Let $\boldsymbol{p}, \boldsymbol{v} \in \mathbb{R}^3$ be constants. Consider the curve $\beta(t) = \boldsymbol{p} + t\boldsymbol{v}$. This is the straight line with initial position $\beta(0) = \boldsymbol{p}$ and initial velocity $\beta'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. We've used this before when we've worked with the directional derivative of $f\colon \mathbb{R}^3 \to \mathbb{R}$,

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p} + t\boldsymbol{v})\bigg|_{t=0}. \tag{2.3}$$

In fact, we could use *any* curve α with $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$ to define $\boldsymbol{v}_{\boldsymbol{p}}[f]$.

53. Theorem. *Let $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3$, I be an interval containing zero, and let $\alpha\colon I \to \mathbb{R}^3$ be such that $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. Then*

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\alpha(t))\bigg|_{t=0}. \tag{2.4}$$

Proof. We know $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. So, let us calculation

$$\begin{aligned}
& \boldsymbol{v}_{\boldsymbol{p}}[f] \\
= \quad & | \quad \text{definition of directional derivative} && \text{(a)} \\
& \tfrac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p} + t\boldsymbol{v})\big|_{t=0} \\
= \quad & | \quad \text{since } \alpha(t) = \boldsymbol{p} + t\boldsymbol{v} && \text{(b)} \\
& \tfrac{\mathrm{d}}{\mathrm{d}t} f(\alpha(t))\big|_{t=0} \\
= \quad & | \quad \text{chain rule} && \text{(c)} \\
& \textstyle\sum_j \frac{\partial f}{\partial x_j}(\alpha(t)) \frac{\mathrm{d}\alpha^j(t)}{\mathrm{d}t}\Big|_{t=0} \\
= \quad & | \quad \text{since } \mathrm{d}\alpha^j(0)/\mathrm{d}t = v^j,\ \alpha(0) = \boldsymbol{p} && \text{(d)} \\
& \textstyle\sum_j \frac{\partial f}{\partial x_j}(\boldsymbol{p}) v^j
\end{aligned}$$

On the other hand, repeating the last three steps with

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\alpha(t))\bigg|_{t=0}$$

gives the same result since the only facts used were $\alpha(0) = \boldsymbol{p}$, $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. ∎

54. Corollary. *For any curve α and smooth function $f\colon \mathbb{R}^3 \to \mathbb{R}$, we have*

$$\alpha'(t)[f] = \frac{\mathrm{d}}{\mathrm{d}s} f(\alpha(s))\Big|_{s=t}. \tag{2.5}$$

That is to say, the directional derivative of f with respect to the velocity vector field is the rate of change of f as we move along the curve α.

2.1 Metric, Distances, Angles

55. In linear algebra, the key geometric tool is the concept of the inner product ("dot product"). Any vector space with an inner product automatically gets notations of:

- **magnitude** of vectors $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$, and
- **angle** between vectors $\cos(\theta) = \langle \boldsymbol{v}, \boldsymbol{w}\rangle/(\|\boldsymbol{v}\|\cdot\|\boldsymbol{w}\|)$.

In differential geometry, we have not just *one* vector space, but a vector space *at each point* (e.g., for each $\boldsymbol{p}\in\mathbb{R}^3$, we have $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3$). We can, in principle, put a different inner product at each of these tangent spaces. In other words, at each point we may have a different notion of magnitude and angle.

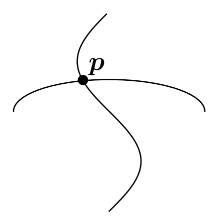
56. Definition. An assignment of an inner product to each $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3$ (varying smoothly with $\boldsymbol{p}\in\mathbb{R}^3$) is called a **"Riemannian metric"**.

57. Example. For now, we will be sticking with the usual Riemannian metric on $\mathbb{R}^3$ given by

$$\underbrace{\langle \boldsymbol{v}_{\boldsymbol{p}}, \boldsymbol{w}_{\boldsymbol{p}}\rangle_{\boldsymbol{p}}}_{\text{inner product on } \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3} = \overbrace{\boldsymbol{v}\cdot\boldsymbol{w}}^{\text{usual dot product}} \tag{2.6}$$

58. Example. Suppose we have two curves $\alpha\colon I\to\mathbb{R}^n$ and $\beta\colon J\to\mathbb{R}^n$ which intersects at a point

$$\boldsymbol{p} = \alpha(s_0) = \beta(t_0). \tag{2.7}$$



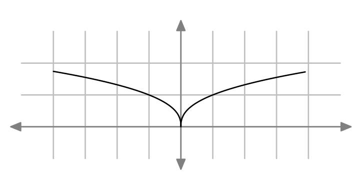
We can define the angle between α and β at $\boldsymbol{p}$ by

$$\cos(\theta) = \frac{\langle \alpha'(s_0), \beta'(t_0)\rangle}{\|\alpha'(s_0)\|\cdot\|\beta'(t_0)\|}. \tag{2.8}$$

This breaks down if $\|\beta'(t_0)\| = 0$ or $\|\alpha'(s_0)\| = 0$. For this reason, we usually work with **"Regular Curves"** which always have nonzero velocity.

59. Definition. We call a curve $\alpha\colon I\to\mathbb{R}^n$ **"Regular"** if its velocity is never zero $\alpha'(t)\neq 0$ for all $t\in I$. If α is not regular, then we call it **"Singular"**.

60. Example. An example of a singular curve would be one with a cusp, for example, $x^2 - y^5 = 0$



61. Measuring Distance. The metric also allows us to measure distances along regular curves. The distance is found by integrating the **"Speed"** or magnitude of velocity. Let $\alpha\colon (a,b) \to \mathbb{R}^n$ be a regular curve, then

$$\int_a^b \|\alpha'(t)\|\,\mathrm{d}t. \tag{2.9}$$

A key fact in this for this to be well-defined, we need to check it's independent of (regular) parametrization. That is to say, suppose we have

$$t = t(s) \tag{2.10}$$

where s is another parameter, and define a **"Reparametrization"** of the curve by

$$\beta(s) = \alpha(t(s)) \tag{2.11}$$

where

$$\frac{\mathrm{d}t(s)}{\mathrm{d}s} > 0 \tag{2.12}$$

for all $s \in J$, so β is regular. Suppose $J = (a', b')$. Then we have to integrate from a' to b' of the magnitude of the velocity of β and demand they be equal

$$\int_{a'}^{b'} \|\beta'(s)\|\,\mathrm{d}s = \int_a^b \|\alpha'(t)\|\,\mathrm{d}t. \tag{2.13}$$

We see, actually, we could just unfold the definition of β and use the chain rule,

$$\int_{a'}^{b'} \|\beta'(s)\|\,\mathrm{d}s = \int_{a'}^{b'} \|\alpha'(t(s))\frac{\mathrm{d}t(s)}{\mathrm{d}s}\|\,\mathrm{d}s. \tag{2.14}$$

Then using substitution rule for calculus on the right-hand side

$$\int_{a'}^{b'} \|\alpha'(t(s))\frac{\mathrm{d}t(s)}{\mathrm{d}s}\|\,\mathrm{d}s = \int_a^b \|\alpha'(t)\|\,\mathrm{d}t. \tag{2.15}$$

We have a particularly useful parametrization:

62. Proposition. *Any regular curve α has a reparametrization β such that*

$$\|\beta'(s)\| = 1 \tag{2.16}$$

*for all s, called the **"Unit Speed Reparametrization"**.*

Proof. Let $\alpha\colon I \to \mathbb{R}^3$ be our given curve. Suppose $I = (a,b)$. Define the distance function

$$s(t) = \int_a^t \|\alpha'(t)\|\,\mathrm{d}t. \tag{2.17}$$

Since $\|\alpha'(t)\| \neq 0$ (and speed is never negative, it follows the speed is always positive), hence s is strictly increasing. In particular, this means

$$\frac{\mathrm{d}s}{\mathrm{d}t} > 0. \tag{2.18}$$

So far, so good.

This also means $s(t)$ has an inverse. So we can write t as a function of s,

$$t = t(s). \tag{2.19}$$

Then we can reparametrize by letting

$$\beta(s) = \alpha(t(s)), \tag{2.20}$$

and

$$
\begin{aligned}
&\|\beta'(s)\| \\
= \quad & \Big| \quad \text{chain rule} && \text{(a)} \\
&\left\|\alpha'(t(s))\frac{\mathrm{d}t(s)}{\mathrm{d}s}\right\| \\
= \quad & \Big| \quad \text{since } \mathrm{d}t(s)/\mathrm{d}s > 0 \text{ always} && \text{(b)} \\
&\|\alpha'(t(s))\| \frac{\mathrm{d}t(s)}{\mathrm{d}s} \\
= \quad & \Big| \quad \text{differentiating under the integral sign} && \text{(c)} \\
&\frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s} \\
= \quad & \Big| \quad \text{basic calculus} && \text{(d)} \\
&1.
\end{aligned}
$$

Hence β has unit-speed. □

62.1. Remark. For any regular curve, we get a vector field on the curve called the **"Unit Tangent Field"** defined by taking the velocity of a unit speed reparametrization.

2.2 Frenet Frame

63. The basic idea is to study a curve by using a different frame at each point, suitably chosen at each point. Towards that end, we should probably make rigorous what we mean by a "vector field along a curve" and whatnot.

64. Definition. Let $\alpha\colon I \to \mathbb{R}^3$ be a (regular) curve. A **"Vector Field on α"** Y is an assignment of a tangent vector $Y(t) \in \mathrm{T}_{\alpha(t)}\mathbb{R}^3$ for each $t \in I$.

An **"(Orthonormal) Frame Field on a curve"** α is a triple of vector fields such that for each $t \in I$ they restrict to an (orthonormal) basis of $\mathrm{T}_{\alpha(t)}\mathbb{R}^3$.

We assume, without loss of generality, that frame fields are orthonormal unless explicitly stated otherwise.

65. For now, we restrict attention to unit speed curves.

We have constructed on natural unit vector field on β:

$$T(s) = \beta'(s) \in \mathrm{T}_{\beta(s)}\mathbb{R}^3. \tag{2.21}$$

We call $T(S)$ the **"Unit Tangent"**. To get a frame field, we need two more vector fields on β.

First note,

$$T(s) \cdot T(s) = 1, \tag{2.22a}$$

because $T(s)$ is a unit tangent vector. It follows then that

$$\frac{\mathrm{d}}{\mathrm{d}s}(T(s) \cdot T(s)) = \frac{\mathrm{d}T(s)}{\mathrm{d}s} \cdot T(s) + T(s) \cdot \frac{\mathrm{d}T(s)}{\mathrm{d}s} \tag{2.22b}$$

$$= 0, \tag{2.22c}$$

and in particular

$$\boxed{T(s) \cdot T'(s) = 0.} \tag{2.22d}$$

If $T'(s) \neq 0$, then define the **"Principal Normal"** field on β as

$$N(s) := \frac{T'(s)}{\|T'(s)\|}. \tag{2.23}$$

We see (since the right-hand side is a vector divided by its norm) $N(s)$ is a unit vector field, i.e., for any s we have

$$\|N(s)\| = 1, \tag{2.24}$$

and it is orthogonal to $T(s)$ by construction.

Now, to obtain one last unit vector field on β, we can just take the cross-product of T and N to obtain the **"Binormal Field"** on β as

$$B(s) := T(s) \times N(s). \tag{2.25}$$

66. Definition. Let β be a unit speed curve, then we define the **"Curvature"** of β as

$$\kappa(s) := \|T'(s)\|. \tag{2.26}$$

66.1. Remark. If the curvature of β is zero, we get nonunique frames from the construction we have just sketched.

67. Definition. Along a regular curve β, we define the **"Frenet Frame"** to consist of $T(s) := \beta'(s)$, $N(s) := T'(s)/\|T'(s)\|$, and $B(s) := T(s) \times N(s)$.

68. For any regular curve with unit-speed parametrization β, we know that $T'(s)$ is proportional to $N(s)$. We claim that $B'(s)$ is also proportional to $N(s)$. How can we see this? We will prove $B'(s)$ is orthogonal to both $B(s)$ and $T(s)$, which means it's either zero or directly proportional to the remaining orthonormal unit vector $N(s)$.

First, observe that $B'(s)$ is orthogonal to $B(s)$. How? We see from it being a unit vector,

$$B(s) \cdot B(s) = 1, \tag{2.27a}$$

taking the derivative with respect to s of both sides,

$$\frac{\mathrm{d}}{\mathrm{d}s}(B(s) \cdot B(s)) = 0, \tag{2.27b}$$

the left-hand side expands according to the product rule as

$$\frac{\mathrm{d}}{\mathrm{d}s}(B(s) \cdot B(s)) = \frac{\mathrm{d}B(s)}{\mathrm{d}s} \cdot B(s) + B(s) \cdot \frac{\mathrm{d}B(s)}{\mathrm{d}s}. \tag{2.27c}$$

Thanks to commutativity of the dot product, the right hand side simplifies to:

$$\frac{\mathrm{d}}{\mathrm{d}s}(B(s) \cdot B(s)) = 2\frac{\mathrm{d}B(s)}{\mathrm{d}s} \cdot B(s). \tag{2.27d}$$

Thus we find

$$2\frac{\mathrm{d}B(s)}{\mathrm{d}s} \cdot B(s) = 0 \implies \frac{\mathrm{d}B(s)}{\mathrm{d}s} \cdot B(s) = 0. \tag{2.27e}$$

Thus we conclude $B'(s)$ is orthogonal to $B(s)$.

Our second step is to show $B'(s)$ is orthogonal to $T(s)$. We have

$$\frac{\mathrm{d}}{\mathrm{d}s}(B \cdot T) = B'(s) \cdot T(s) + B(s) \cdot T'(s). \tag{2.28a}$$

Now since

$$T'(s) = \kappa(s)N(s), \tag{2.28b}$$

we find

$$B'(s) \cdot T(s) + B(s) \cdot T'(s) = B'(s) \cdot T(s) + B(s) \cdot (\kappa(s)N(s)). \tag{2.28c}$$

But since

$$B(s) \cdot (\kappa(s)N(s)) = \kappa(s)(B(s) \cdot N(s)) = \kappa(s)(0) = 0, \tag{2.28d}$$

we find

$$B'(s)\cdot T(s) + B(s)\cdot T'(s) = B'(s)\cdot T(s). \tag{2.28e}$$

But remember, $B(s)\cdot T(s) = (T(s)\times N(s))\cdot T(s)$ using the definition of the binormal field, and recalling the basic property of the cross product (it produces a vector orthogonal to its factors), we find

$$B(s)\cdot T(s) = 0. \tag{2.28f}$$

Therefore its derivative with respect to s vanishes, and we find

$$B'(s)\cdot T(s) = 0. \tag{2.28g}$$

Hence $B'(s)$ is orthogonal to $T(s)$.

Since $B'(s)$ is orthogonal to both $T(s)$ and $B(s)$, we conclude it must be directly proportional to $N(s)$:

$$B'(s) = -\tau(s)N(s) \tag{2.29}$$

where "$-\tau(s)$" is the constant of proportionality. We call $\tau(s)$ the **"Torsion"** of the curve β.

69. Theorem (The Frenet Formulas)**.** *Given a unit speed curve β, whose curvature κ is nonvanishing and whose torsion is τ, we have the* ***"Frenet Formulas"****:*

$$T'(s) = \kappa(s)N(s) \tag{2.30a}$$

$$N'(s) = -\kappa(s)T(s) + \tau(s)B(s) \tag{2.30b}$$

$$B'(s) = -\tau(s)N(s), \tag{2.30c}$$

or, using matrix multiplication,

$$\begin{bmatrix} T'(s)\\ N'(s)\\ B'(s)\end{bmatrix} = \begin{bmatrix} 0 & \kappa(s) & 0\\ -\kappa(s) & 0 & \tau(s)\\ 0 & -\tau(s) & 0\end{bmatrix}\begin{bmatrix} T(s)\\ N(s)\\ B(s)\end{bmatrix}. \tag{2.31}$$

Proof. We know the first and third formulas already, so we just need to prove the second formula. Since the Frenet field forms an orthonormal basis, we know

$$N'(s) = a(s)T(s) + b(s)N(s) + c(s)B(s). \tag{2.32}$$

We just need to determine the coefficients. We take the inner product of both sides with T and B.

We know that

$$N'(s)\cdot T(s) = a(s). \tag{2.33a}$$

We also know $N(s)$ and $T(s)$ are orthonormal vector, in particular,

$$N(s)\cdot T(s) = 0. \tag{2.33b}$$

But taking the derivative of both sides, we find

$$\frac{\mathrm{d}}{\mathrm{d}s}(N(s)\cdot T(s)) = N'(s)\cdot T(s) + N(s)\cdot T'(s). \tag{2.33c}$$

We can use the Frenet formula $T'(s) = \kappa(s)N(s)$ to rewrite the right-hand side

$$N'(s)\cdot T(s) + N(s)\cdot T'(s) = N'(s)\cdot T(s) + N(s)\cdot(\kappa(s)N(s) \tag{2.33d}$$

$$= N'(s)\cdot T(s) + \kappa(s), \tag{2.33e}$$

and further, since $N'(s)\cdot T(s)=a(s)$, we conclude

$$\begin{aligned} N'(s)\cdot T(s)+N(s)\cdot T'(s) &= N'(s)\cdot T(s)+\kappa(s) \\ &= a(s)+\kappa(s). \end{aligned} \tag{2.33f}$$

Returning to our original statement, we find

$$\frac{\mathrm{d}}{\mathrm{d}s}(N(s)\cdot T(s)) = a(s)+\kappa(s) = 0. \tag{2.33g}$$

Hence, in particular

$$\boxed{a(s) = -\kappa(s).} \tag{2.33h}$$

We know $b(s)=0$ since

$$N(s)\cdot N(s)=1 \implies \frac{\mathrm{d}}{\mathrm{d}s}(N(s)\cdot N(s))=0. \tag{2.34a}$$

This gives us $2b(s)=0$, which implies

$$\boxed{b(s)=0.} \tag{2.34b}$$

Now for the last coefficient. Very similar to the first, since $N(s)\cdot B(s)=0$ (thanks to their being orthonormal vectors) we find their derivative with respect to s is zero. (We will use the fact that $N(s)\cdot B'(s)=0$.) But we find

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}(N(s)\cdot B(s)) &= N'(s)\cdot B(s)+N(s)\cdot B'(s) && (2.35a)\\ &= N'(s)\cdot B(s) && (2.35b)\\ &= \tau(s) && (2.35c) \end{aligned}$$

hence

$$\boxed{c(s)=\tau(s).} \tag{2.35d}$$

This proves the remaining Frenet formula. □

70. Proposition. *A unit speed curve is a straight line if and only if its curvature vanishes* $\kappa(s)=0$.

Our proof will consist of two direct proofs, one in the forward direction (straight line implies zero curvature), the other in the backwards direction (zero curvature implies straight line).

Proof. Let $\beta(s)$ be a unit speed curve.

($\implies$) Assume β is a straight line, then

$$\beta(s) = \boldsymbol{p} + s\boldsymbol{v} \tag{2.36}$$

for some $\boldsymbol{p}$ and (unit) vector $\boldsymbol{v}$. So

$$\begin{aligned} \kappa(s) &= \|T'(s)\| && (2.37a)\\ &= \|\beta''(s)\| && (2.37b)\\ &= \|0\| = 0. && (2.37c) \end{aligned}$$

Hence straight lines have zero curvature.

($\impliedby$) Conversely, assume for all s we have $\kappa(s)=0$. Write out the components of the curve as

$$\beta(s) = (\beta_1(s), \beta_2(s), \beta_3(s)). \tag{2.38a}$$

We find

$$\beta''(s) = (\beta_1''(s), \beta_2''(s), \beta_3''(s)). \tag{2.38b}$$

For the curvature to be zero everywhere, we need the second derivative of each component to vanish, i.e., for each s we have

$$\beta_j''(s) = 0 \tag{2.38c}$$

for $j = 1, 2, 3$. We integrate this equation twice to find

$$\beta_j(s) = p_j + sv_j \tag{2.38d}$$

for some constants of integration v_j, p_j. This means

$$\beta(s) = \boldsymbol{p} + s\boldsymbol{v}, \tag{2.38e}$$

i.e., that β is a straight line. □

71. Proposition. *A unit speed curve $\beta\colon I \to \mathbb{R}^3$ with positive curvature $\kappa(s) > 0$ has vanishing torsion if and only if β is a* ***"Plane Curve"*** *(i.e., its image lies in some plane in $\mathbb{R}^3$).*

Proof. ($\Longrightarrow$) Assume β is a plane curve, i.e., it lies in the plane through $\boldsymbol{p} \in \mathbb{R}^3$ with unit normal vector $\boldsymbol{n}$. For all s, we have

$$\boldsymbol{n} \cdot (\beta(s) - \boldsymbol{p}) = 0. \tag{2.39a}$$

By differentiating with respect to s (and remembering $\boldsymbol{p}$ and $\boldsymbol{n}$ are constants), we find

$$\boldsymbol{n} \cdot \beta'(s) = 0 \implies \boldsymbol{n} \cdot T(s) = 0. \tag{2.39b}$$

Differentiating once more, we find

$$\boldsymbol{n} \cdot \beta''(s) = 0 \implies \boldsymbol{n} \cdot N(s) = 0. \tag{2.39c}$$

But we assumed $\boldsymbol{n}$ is a unit vector (in particular, it's nonzero) and the Frenet frame is a collection of orthonormal vectors. This forces us to conclude,

$$\boldsymbol{n} = \pm B(s). \tag{2.39d}$$

In particular, $B(s)$ is constant. So

$$B'(s) = 0, \tag{2.39e}$$

and by the Frenet formulas

$$B'(s) = -\tau(s)N(s) = 0. \tag{2.39f}$$

Since $N(s) \neq 0$ we conclude,

$$\boxed{\tau(s) = 0.} \tag{2.39g}$$

This is the first half of the proof.

($\Longleftarrow$) Suppose $\tau(s) = 0$. Then by the Frenet formulas, $B'(s) = 0$. In particular, $B(s)$ is a constant unit vector. Pick some $s_0 \in I$, and define

$$f(s) = (\beta(s) - \beta(s_0)) \cdot B. \tag{2.40}$$

We want to prove $f(s) = 0$ for all $s \in I$.

The first step is to consider its derivative, and since B is constant (with respect to s), we find:

$$f'(s) = (\beta'(s) - 0) \cdot B = T(s) \cdot B. \tag{2.41}$$

But we know $T(s) \cdot B = 0$ since they are orthonormal vectors. Hence we conclude

$$f'(s) = 0. \tag{2.42}$$

This just tells us that $f(s)$ is a constant.

But we also know that

$$f(s_0) = 0. \tag{2.43}$$

Hence we conclude, for all $s \in I$, that:

$$f(s) = 0. \tag{2.44}$$

This is precisely the description of a plane, $\{\boldsymbol{x} \in \mathbb{R}^3 \mid (\boldsymbol{x} - \boldsymbol{p}) \cdot B = 0\}$. $\square$

72. Example (Frenet field for circular helix)**.** Let

$$\alpha(t) = (a\cos(t), a\sin(t), bt) \tag{2.45}$$

where $a > 0$, $b > 0$ are constants. Find the Frenet field for this curve.

First, we need to verify the curve is a unit-speed curve (and, if not, find its unit-speed reparamtrization). We find

$$s(t) = \int_0^t \|\alpha'(u)\| \,\mathrm{d}u \tag{2.46a}$$

$$= \int_0^t \sqrt{a^2 + b^2} \,\mathrm{d}u \tag{2.46b}$$

$$= t\sqrt{a^2 + b^2}. \tag{2.46c}$$

Then we can invert this to find t as a function of s:

$$t(s) = \frac{s}{\sqrt{a^2 + b^2}}. \tag{2.46d}$$

For simplicity, we will define

$$c = \sqrt{a^2 + b^2}, \tag{2.46e}$$

so

$$t(s) = s/c. \tag{2.46f}$$

Then we find the unit-speed parametrization,

$$\boxed{\beta(s) = \alpha(t(s)) = (a\cos(s/c), a\sin(s/c), bs/c).} \tag{2.47}$$

OK, the zeroeth step is complete.

Now we may use the Frenet formulas to get the curvature and torsion. The unit tangent vector is

$$T(s) := \beta'(s) = \frac{1}{c}(-a\sin(s/c), a\cos(s/c), b). \tag{2.48}$$

Now we find its derivative

$$T'(s) = \frac{-a}{c^2}(\cos(s/c), \sin(s/c), 0), \tag{2.49}$$

and the curvature is,

$$\kappa(s) := \|T(s)\| = \frac{a}{c^2}. \tag{2.50}$$

That is to say, the curvature is constant.

The next Frenet formulas give us

$$N(s) := \frac{T'(s)}{\kappa(s)} = (-\cos(s/c), -\sin(s, c), 0). \tag{2.51}$$

The binormal is obtained from the cross-product,

$$B(s) := N(s) \times T(s) = \frac{1}{c}(b\sin(s/c), -b\cos(s/c), a). \tag{2.52}$$

The derivative of $B(s)$ with respect to s gives us

$$B'(s) = \frac{b}{c^2}(\cos(s/c), \sin(s/c), 0), \tag{2.53}$$

but we know using the Frenet formula $B'(s) = -\tau(s)N(s)$, and inspection of terms forces us to conclude,

$$\tau(s) = \frac{-b}{c^2}. \tag{2.54}$$

The torsion is also constant!

In fact, any curve with constant [nonzero] curvature and nonzero torsion is either a helix ($\tau \neq 0$) or a degenerate helix (a.k.a., a circle).

2.3 Frenet Approximation at a Point

73. Consider the unit-speed curve β. Expand this in a Taylor series, e.g., about zero. We need to know several values:

$$\beta'(0) := T(0) = T_0 \tag{2.55a}$$

$$\beta''(0) := \kappa(0)N(0) = \kappa_0 N_0 \tag{2.55b}$$

$$\beta'''(0) = (\kappa N)'(0) = \kappa'(0)N_0 + \kappa_0(-\kappa_0 T_0 + \tau_0 B_0). \tag{2.55c}$$

Then the Taylor series expansion yields

$$\beta(s) \approx \beta_0 + s\beta'(0) + \frac{s^2}{2!}\beta''(0) + \frac{s^3}{3!}\beta'''(0). \tag{2.56}$$

We find

$$\begin{aligned}
\beta(s) &\approx \beta_0 + s\beta'(0) + \frac{s^2}{2!}\beta''(0) + \frac{s^3}{3!}\beta'''(0) && (2.57a)\\
&= \beta_0 + sT_0 + \frac{s^2}{2}(\kappa(0)N(0) = \kappa_0 N_0) + \frac{s^3}{3!}(\kappa'(0)N_0 + \kappa_0(-\kappa_0 T_0 + \tau_0 B_0)) && (2.57b)\\
&= \beta_0 + \left(s - \frac{\kappa_0^2 s^3}{3!}\right)T_0 + \left(\frac{\kappa_0}{2}s^2 + \kappa'(0)\frac{s^3}{3!}\right)N_0 + \left(\tau_0\kappa_0\frac{s^3}{3!}\right)B_0. && (2.57c)
\end{aligned}$$

74. We can truncate this series to be scalar multiples of the Frenet vectors at $\beta(0)$, with the following interpretation:

$$\begin{aligned}
\beta(s) &\approx \overbrace{\beta_0 + T_0 s + N_0 s^2\frac{\kappa_0}{2}}^{\begin{pmatrix}\text{Tangent Parabola}\\ \text{in Osculating Plane}\end{pmatrix}} + B_0\tau_0\kappa_0\frac{s^3}{3!}\\
&\approx \begin{pmatrix}\text{Linear}\\ \text{Approximation}\end{pmatrix} + \begin{pmatrix}\text{How Fast the Curve}\\ \text{Deviates From}\\ \text{Tangent Line}\end{pmatrix} + \begin{pmatrix}\text{How fast the curve}\\ \text{moves out of}\\ \text{the osculating plane}\end{pmatrix}
\end{aligned} \tag{2.58}$$

Exercises

1. Consider the hyperbolic helix, $\alpha(t) = (\cosh(t), \sinh(t), t)$. Compute its arclength $s(t)$ and find its unit-speed reparametrization.
2. Consider the curve $\alpha(t) = (1/(1+t^3), t^3, \ln(1+t^2))$ on $I = (0, \infty)$. It passes through $\boldsymbol{p} = (1/4,\ 3,\ \ln(1+3^{2/3}))$ and $\boldsymbol{q} = (1/9,\ 8,\ \ln(5))$. Compute the arclength of the curve between these points.
3. Let $\alpha\colon I \to \mathbb{R}^n$ be a curve. Suppose β_1, β_2 are two unit-speed reparametrizations of α. Prove or find a counter-example: there exists an $s_0 \in \mathbb{R}$ such that for any $s \in (0, \ell(\alpha))$ we have $\beta_1(s) = \beta_2(s + s_0)$ (where $\ell(\alpha)$ is the arclength of α).
4. Prove or find a counter-example: given a unit-speed curve β, consider the vector field $A = \tau T + \kappa B$ on β. The Frenet formulas become:

$$\begin{aligned} T' &= A \times T \\ B' &= A \times B \\ N' &= A \times N. \end{aligned}$$

5. Prove or find a counter-example: if $\alpha\colon I \to \mathbb{R}^n$ is any curve, and $c \in I$ is some [arbitrary but fixed] value, then (a) $\sigma(t) = \int_c^t \|\alpha'(u)\|\, du$ is a perfectly good distance function, (b) which could be inverted to $t = t(\sigma)$, and (c) $\beta(\sigma) = \alpha(t(\sigma))$ is a unit-speed parametrization. [A counter-example would be a curve α for which at least one of these three claims fails to hold.]
6. The unit-speed parametrization of the circle may be written as

$$\gamma(s) = \boldsymbol{c} + r\cos(s/r)\, \boldsymbol{e}_1 + r\sin(s/r)\, \boldsymbol{e}_2 \tag{2.59}$$

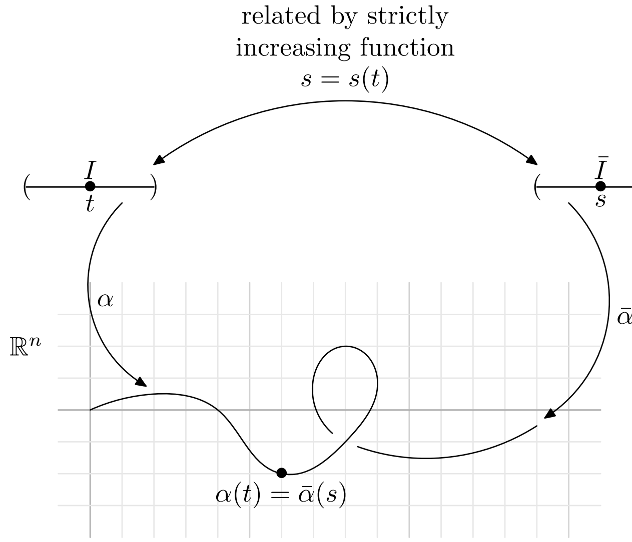
   where $\boldsymbol{e}_i$ are orthonormal unit vectors $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$.
   If $\beta\colon \bar{I} \to \mathbb{R}^3$ is a unit speed curve with (assuming $0 \in \bar{I}$) $\kappa(0) > 0$, then there exists exactly one circle γ which approximates β near $\beta(0)$ in the sense that

$$\gamma(0) = \beta(0), \quad \gamma'(0) = \beta'(0), \quad \gamma''(0) = \beta''(0). \tag{2.60}$$

   Show that γ lies in the osculating plane of β at $\beta(0)$ and find its center $\boldsymbol{c}$ and radius r. The circle is called the **"Osculating Circle"**, and $\boldsymbol{c}$ is called the **"Center of Curvature"** of β at $\beta(0)$. (The same result holds if we replace 0 by any other $s_0 \in \bar{I}$.)

2.4 Frenet Data for Arbitrary Curves

75. We defined the Frenet field for unit-speed curves. But we want the Frenet data T, N, B, τ, κ to be independent of our choice of parametrization. The trick: given some regular curve $\alpha\colon I \to \mathbb{R}^3$, let $\bar{\alpha}\colon \bar{I} \to \mathbb{R}^3$ be its unit-speed parametrization.



We *define* the Frenet data of α to be the Frenet data of $\bar{\alpha}$. The Frenet data of α are unbarred quantities, the Frenet data of $\bar{\alpha}$ are barred quantities. We define them by:

$$T(t) = \bar{T}(s) = \bar{T}(s(t)) \tag{2.61a}$$
$$N(t) = \bar{N}(s) \tag{2.61b}$$
$$B(t) = \bar{B}(s) \tag{2.61c}$$
$$\kappa(t) = \bar{\kappa}(s) \tag{2.61d}$$
$$\tau(t) = \bar{\tau}(s) \tag{2.61e}$$

In principle we can find $\bar{\alpha}$ and compute its Frenet data, but in practice this is usually impossible (analytically) because we must do two things:

1. calculate $s(t) = \displaystyle\int_{t_0}^{t} \|\alpha'(u)\|\,\mathrm{d}u$,
2. invert this to get $t = t(s)$.

Both are hard, sometimes impossible. We need a better way to calculate these things directly.

76. Unit Tangent. We find by direct calculation

$$T(t) = \bar{T}(s) = \bar{\alpha}'(s) \tag{2.62a}$$
$$= \frac{\mathrm{d}}{\mathrm{d}s}\alpha(t(s)) \tag{2.62b}$$
$$= \alpha'(t)\frac{\mathrm{d}t}{\mathrm{d}s}. \tag{2.62c}$$

But

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \|\alpha'(t)\|, \tag{2.63}$$

so

$$T(t) = \frac{\alpha'(t)}{\|\alpha'(t)\|}. \tag{2.64}$$

77. Curvature. We compute directly,

$$\begin{aligned}\kappa(t) = \bar{\kappa}(s) &:= \left\| \frac{\mathrm{d}}{\mathrm{d}s}\bar{T}(s) \right\| &(2.65a)\\ &= \left\| \frac{\mathrm{d}}{\mathrm{d}s}T(t) \right\| &(2.65b)\\ &= \left\| T'(t)\frac{\mathrm{d}t}{\mathrm{d}s} \right\| &(2.65c)\\ &= \frac{\|T'(t)\|}{\|\alpha'(t)\|} &(2.65d)\end{aligned}$$

78. Principal Normal. We find, letting $v(t) = \|\alpha'(t)\|$,

$$N(t) = \bar{N}(s) = \frac{\bar{T}'(s)}{\bar{\kappa}(s)} = \frac{T'(t)/v(t)}{\kappa(t)} = \frac{T'(t)}{\|T'(t)\|}. \tag{2.66}$$

This matches intuition of $N(t) \propto T'(t)$.

79. Binormal. Again, direct computation,

$$B(t) = \bar{B}(s) = \bar{T}(s) \times \bar{N}(s) = T(t) \times N(t). \tag{2.67}$$

This matches intuition of the binormal as cross-product of T and N.

80. Torsion. We know

$$\bar{B}'(s) = -\bar{\tau}(s)\bar{N}(s) \tag{2.68}$$

and so, letting $v(t) = \|\alpha'(t)\|$,

$$\frac{\mathrm{d}}{\mathrm{d}s}B(t) = -\tau(t)v(t)N(t). \tag{2.69}$$

To summarize our results, we have this handy theorem:

81. Theorem. *If $\alpha\colon I \to \mathbb{R}^{33}$ is a regular curve with positive curvature, then up to some factor $v(t) = \|\alpha'(t)\|$ we have,*

$$\begin{aligned}T'(t) &= \kappa(t)v(t)N(t) &(2.70a)\\ N'(t) &= -\kappa(t)v(t)T(t) + \tau(t)v(t)B(t) &(2.70b)\\ B'(t) &= -\tau(t)v(t)N(t) &(2.70c)\end{aligned}$$

Or, using matrices,

$$\begin{bmatrix}T'(t)\\ N'(t)\\ B'(t)\end{bmatrix} = \begin{bmatrix}0 & \kappa(t)v(t) & 0\\ -\kappa(t)v(t) & 0 & \tau(t)v(t)\\ 0 & -\tau(t)v(t) & 0\end{bmatrix}\begin{bmatrix}T(t)\\ N(t)\\ B(t)\end{bmatrix}. \tag{2.71}$$

82. Any vector field Y on a regular curve α could be written as a linear combination of Frenet vector fields:

$$Y(t) = f(t)T(t) + g(t)N(t) + h(t)B(t). \tag{2.72}$$

We find its derivative:

$$\begin{aligned}Y'(t) &= f'T + fT' + g'N + gN' + h'B + hB' &(2.73a)\\ &= f'T + f\kappa vN + g'N + g(t)(-\kappa vT + \tau vB) + h'B + h(t)(-\tau vN) &(2.73b)\\ &= (f' - \kappa vg)T + (f\kappa v + g' - \tau vh)N + (g\tau v + h')B. &(2.73c)\end{aligned}$$

If Y has Frenet components (f, g, h), then the Frenet components of $Y'(t)$ are

$$\begin{bmatrix} f'(t) \\ g'(t) \\ h'(t) \end{bmatrix} + \begin{bmatrix} 0 & -\kappa(t)v(t) & 0 \\ \kappa(t)v(t) & 0 & -\tau(t)v(t) \\ 0 & \tau(t)v(t) & 0 \end{bmatrix} \begin{bmatrix} f(t) \\ g(t) \\ h(t) \end{bmatrix} = Y'(t). \tag{2.74}$$

Note: we use the *transpose* of the matrix from the Frenet formulas.

82.1. Remark. We should observe, when the frame we're working with is not the natural frame field, then we cannot write the derivative of a vector as just the derivatives of the components. We just saw this won't work with the Frenet frame. There was a correction term. This idea underlies the idea of covariant derivatives.

2.5 Covariant Differentiation

83. We want to differentiate vector fields. Suppose we are given some vector field $W \in \mathrm{Vect}(\mathbb{R}^n)$ and a tangent vector $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^n$. There is one obvious way to differentiate W in the direction of $\boldsymbol{v_p}$: consider expressing W using coordinates relative to the natural frame field,

$$W = \sum_j w^j U_j, \tag{2.75}$$

then we just consider

$$\boldsymbol{v_p}[W] = \sum_j \boldsymbol{v_p}[w^j] U_j(\boldsymbol{p}). \tag{2.76}$$

Why not?

If we did this using a vector field $V \in \mathrm{Vect}(\mathbb{R}^n)$ at each point $\boldsymbol{p} \in \mathbb{R}^n$, with $\boldsymbol{v_p} = V(\boldsymbol{p})$, then we get

$$\begin{aligned} \sum_j V(\boldsymbol{p})[w^j] U_j(\boldsymbol{p}) &= \left(\sum_j V[w^j] U_j \right)(\boldsymbol{p}) \\ &= \nabla_V W \in \mathrm{Vect}(\mathbb{R}^n). \end{aligned} \tag{2.77}$$

So we get a vector field which is the natural covariant derivative of W with respect to the vector field V.

84. Definition. Let $V, W \in \mathrm{Vect}(\mathbb{R}^n)$. We define the **"Natural Covariant Derivative"** of W with respect to V is the vector field $\nabla_V W$ defined by coordinates relative to the natural frame field,

$$\nabla_V W = \sum_j V[w^j] U_j, \tag{2.78}$$

where $W = \sum_j w^j U_j$ are the coordinates of W relative to the natural frame field U_j.

84.1. Remark. This is really dependent on the natural frame field, but we would like a notion of covariant differentiation *independent* of the choice of frame field.

85. Example. Let $V = \sum_i v^i U_i$ and $W = \sum_j w^j U_j$ be the vector fields expressed in coordinates relative to the natural frame field U_j. Then

$$\nabla_V W = \sum_j V[w^j]U_j \tag{2.79a}$$

$$= \sum_j \left(\sum_i v^i U_i[w^j]\right) U_j \tag{2.79b}$$

$$= \sum_j \left(\sum_i v^i \frac{\partial w^j}{\partial x_i}\right) U_j \tag{2.79c}$$

$$= \sum_j \underbrace{\sum_i v^i \frac{\partial w^j}{\partial x_i}}_{\text{coordinates of } \nabla_V W} U_j \tag{2.79d}$$

86. Theorem (Essential properties of the covariant derivative)**.** *1.* $\nabla_V(aY + bZ) = a\nabla_V Y + b\nabla_V Z$

2. $\nabla_{fV+gW} Z = f\nabla_V Z + g\nabla_W Z$

3. $\nabla_V(fZ) = V[f]Z + f\nabla_V Z$

4. Metric compatibility: $\nabla_V(Y \cdot Z) = V[Y \cdot Z] = (\nabla_V Y) \cdot Z + Y \cdot (\nabla_V Z)$.

86.1. Remark. More generally, any operation $\bar{\nabla}$ taking two vector fields $V \times W \to Z$ and produces a third, which satisfies the first three properties is called a derivative operation. The fourth property is most geometric, as it deals with angles.

87. Connection Forms. Suppose E_i are some orthonormal frame field on $\mathbb{R}^n$, and we express $W \in \text{Vect}(\mathbb{R}^n)$ in coordinates relative to E_i:

$$W = \sum_i w^i E_i. \tag{2.80}$$

Let $V \in \text{Vect}(\mathbb{R}^n)$. Then we want to find the coordinates of $\nabla_V W$ relative to E_i. We know, using linearity and the Leibniz property,

$$\nabla_V W = \sum_i \nabla_V(w^i E_i) = \sum_i V[w^i]E_i + w^i \nabla_V E_i. \tag{2.81}$$

Now we just need to express $\nabla_V E_i$ in coordinates relative to the frame field E_i. We expect

$$\nabla_V E_i = \sum_j c_{ij} E_j, \tag{2.82}$$

and hope the coefficients c_{ij} somehow depend on V. The usual notation is for these coefficients to be written ω_{ij}, and we would have

$$\nabla_V E_i = \sum_j \omega_{ij}[V]E_j. \tag{2.83}$$

We can get the components by applying $\langle -, E_k\rangle$ to both sides

$$\langle \nabla_V E_i, E_k\rangle = \langle \sum_j \omega_{ij}[V]E_j, E_k\rangle \tag{2.84a}$$

$$= \sum_j \omega_{ij}[V]\langle E_j, E_k\rangle \tag{2.84b}$$

$$= \sum_j \omega_{ij}[V]\delta_{j,k} = \omega_{ik}[V]. \tag{2.84c}$$

88. We can use metric compatibility to find

$$V[\langle E_i, E_j\rangle] = \langle \nabla_V E_i, E_j\rangle + \langle E_i, \nabla_V E_j\rangle \tag{2.85a}$$
$$= \omega_{ij}[V] + \omega_{ji}[V] \tag{2.85b}$$
$$= V[\delta_{ij}] = 0. \tag{2.85c}$$

Hence in particular, we find the coefficients are antisymmetric,

$$\boxed{\omega_{ij}[V] = -\omega_{ji}[V].} \tag{2.86}$$

The diagonal components would satisfy $\omega_{ii}[V] = -\omega_{ii}[V]$, which could only happen if $\omega_{ii}[V] = 0$.

When we look back on our system of equations, we find there are only $n(n-1)/2$ independent components (thanks to antisymmetry).

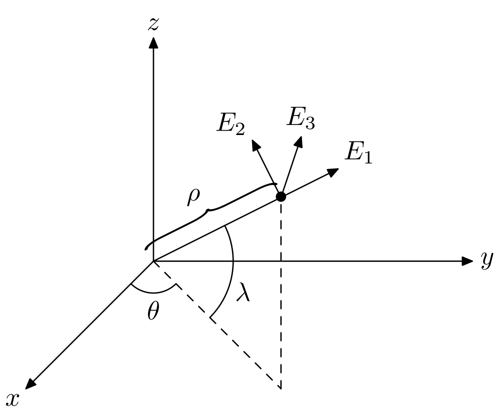
89. For the case when $n = 3$, and we work in $\mathbb{R}^3$, we only needto know $\omega_{12}[V]$, $\omega_{13}[V]$, and $\omega_{23}[V]$ for all $V \in \mathrm{Vect}(\mathbb{R}^3)$. In fact $\omega_{ij}[V]$ depends linearly on V. Consider for arbitrary $f, g \in C^\infty(\mathbb{R}^3)$ and $V, W \in \mathrm{Vect}(\mathbb{R}^3)$,

$$\omega_{ij}[fV + gW] = \langle \nabla_{fV+gW} E_i, E_j\rangle \tag{2.87a}$$
$$= \langle f\nabla_V E_i + g\nabla_W E_i, E_j\rangle \tag{2.87b}$$
$$= f\omega_{ij}[V] + g\omega_{ij}[W]. \tag{2.87c}$$

These coefficients ω_{ij} take a vector field and produce functions. They are one-forms called **"Connection Forms"**.

2.6 Worked Example

90. This is a long example. Consider working in spherical coordinates in $\mathbb{R}^3$. We want to find the spherical frame field on $\mathbb{R}^3$ (or $\mathbb{R}^3$ minus the z axis). We say the order of the basis will be (ρ, θ, λ), as doodled below.



We use the standard Riemannian metric, and the frame field vectors have the following interpretations:

- E_1 is "up"
- E_2 points "East"
- E_3 points "North".

Recall, the coordinates are given by

$$x = \rho\cos(\lambda)\cos(\theta) \tag{2.88a}$$
$$y = \rho\cos(\lambda)\sin(\theta) \tag{2.88b}$$
$$z = \rho\sin(\lambda) \tag{2.88c}$$

91. How do we find E_1? Well, we find a curve passing through (x, y, z) that specifically changes $\rho \to \rho + t$, then its initial velocity unit vector gives us the coefficients for E_1 relative to the natural frame field. So

$$\begin{aligned}\alpha(t) &= (x(\rho+t,\theta,\lambda),\ y(\rho+t,\theta,\lambda),\ z(\rho+t,\theta,\lambda)) \\ &= ((\rho+t)\cos(\lambda)\cos(\theta), (\rho+t)\cos(\lambda)\sin(\theta), (\rho+t)\sin(\lambda).\end{aligned} \tag{2.89}$$

Then we find

$$\alpha'(t) = (\cos(\lambda)\cos(\theta),\ \cos(\lambda)\sin(\theta),\ \sin(\lambda)). \tag{2.90}$$

We observe this *is* a unit vector. Hence

$$\begin{aligned}E_1 &= \alpha'(0)\cdot(U_1, U_2, U_3) \\ &= \cos(\lambda)\cos(\theta)U_1 + \cos(\lambda)\sin(\theta)U_2 + \sin(\lambda)U_3.\end{aligned} \tag{2.91}$$

92. For E_2, a similar calculation with the curve

$$\begin{aligned}\alpha(t) &= (x(\rho,\theta+t,\lambda),\ y(\rho,\theta+t,\lambda),\ z(\rho,\theta+t,\lambda)) \\ &= (\rho\cos(\lambda)\cos(\theta+t),\ \rho\cos(\lambda)\sin(\theta+t),\ \rho\sin(\lambda)).\end{aligned} \tag{2.92}$$

We find its velocity,

$$\alpha'(t) = (-\rho\cos(\lambda)\sin(\theta+t),\ \rho\cos(\lambda)\cos(\theta+t),\ 0). \tag{2.93}$$

We find its unit vector, since its length is

$$\|\alpha'(t)\| = \rho\cos(\lambda) \implies \frac{\alpha'(t)}{\|\alpha'(t)\|} = (-\sin(\theta),\ \cos(\theta),\ 0). \tag{2.94}$$

Its unit vector gives us the coordinates for E_2 relative to the natural frame field,

$$E_2 = -\sin(\theta)U_1 + \cos(\theta)U_2 \tag{2.95}$$

93. The last frame field we need the curve

$$\begin{aligned}\alpha(t) &= (x(\rho,\theta,\lambda+t),\ y(\rho,\theta,\lambda+t),\ z(\rho,\theta,\lambda+t) \\ &= (\rho\cos(\lambda+t)\cos(\theta),\ \rho\cos(\lambda+t)\sin(\theta),\ \rho\sin(\lambda+t)).\end{aligned} \tag{2.96}$$

The velocity vector for this curve,

$$\alpha'(t) = (-\rho\cos(\theta)\sin(\lambda+t),\ -\rho\sin(\theta)\sin(\lambda+t),\ \rho\cos(\lambda+t)). \tag{2.97}$$

Its unit vector

$$\frac{\alpha'(t)}{\|\alpha'(t)\|} = (-\cos(\theta)\sin(\lambda+t),\ -\sin(\theta)\sin(\lambda+t),\ \cos(\lambda+t)). \tag{2.98}$$

Hence we find

$$E_3 = -\sin(\lambda)\cos(\theta)U_1 - \sin(\lambda)\sin(\theta)U_2 + \cos(\lambda)U_3. \tag{2.99}$$

94. Frame field as partial derivatives. We emphasize, for clarity, that if we had a smooth function $f\colon \mathbb{R}^3 \to \mathbb{R}$ that

$$E_1[f] = \frac{\partial f}{\partial \rho} \tag{2.100a}$$

$$E_2[f] = \frac{\partial f}{\partial \theta} \tag{2.100b}$$

$$E_3[f] = \frac{\partial f}{\partial \lambda}. \tag{2.100c}$$

95. Connection Coefficients. Now that we have obtained our spherical frame field, we can compute the connection coefficients $\omega_{ij}[V] = \nabla_V E_i \cdot E_j$. We suppose V has local coordinates relative the E_i,

$$V = \sum_{j=1}^{3} v^j E_j. \tag{2.101}$$

We have three connection components to find.

96. We begin with $\omega_{12}[V] = \langle \nabla_V E_1, E_2 \rangle$. We compute the directional derivative,

$$\begin{aligned}\nabla_V E_1 &= \nabla_V(\cos(\lambda)\cos(\theta)U_1 + \cos(\lambda)\sin(\theta)_2 + \sin(\lambda)U_3) &\text{(2.102a)}\\ &= V[\cos(\lambda)\cos(\theta)]U_1 + V[\cos(\lambda)\sin(\theta)]U_2 + V[\sin(\lambda)]U_3. &\text{(2.102b)}\end{aligned}$$

Now, we use the expansion of V relative to the frame field E_j to find the first summand,

$$\begin{aligned}V[\cos(\lambda)\cos(\theta)] &= (v^1\partial_\rho + v^2\partial_\theta + v^3\partial_\lambda)(\cos(\lambda)\cos(\theta)) &\text{(2.103a)}\\ &= v^1\partial_\rho(\cos(\lambda)\cos(\theta)) + v^2\partial_\theta(\cos(\lambda)\cos(\theta)) + v^3\partial_\lambda(\cos(\lambda)\cos(\theta)) &\text{(2.103b)}\\ &= v^1\cdot 0 - v^2\cos(\lambda)\sin(\theta) - v^3\sin(\lambda)\cos(\theta). &\text{(2.103c)}\end{aligned}$$

Similarly, the second summand

$$\begin{aligned}V[\cos(\lambda)\sin(\theta)] &= (v^1\partial_\rho + v^2\partial_\theta + v^3\partial_\lambda)(\cos(\lambda)\sin(\theta)) &\text{(2.104a)}\\ &= v^2\cos(\lambda)\cos(\theta) - v^3\sin(\lambda)\sin(\theta). &\text{(2.104b)}\end{aligned}$$

The last summand,

$$\begin{aligned}V[\sin(\lambda)] &= (v^1\partial_\rho + v^2\partial_\theta + v^3\partial_\lambda)(\sin(\lambda)) &\text{(2.105a)}\\ &= v^3\cos(\lambda). &\text{(2.105b)}\end{aligned}$$

Now, we combine everything together again,

$$\begin{aligned}\nabla_V E_1 = &-\big(v^2\cos(\lambda)\sin(\theta) + v^3\sin(\lambda)\cos(\theta)\big)\,U_1\\ &+\big(v^2\cos(\lambda)\cos(\theta) - v^3\sin(\lambda)\sin(\theta)\big)\,U_2 + v^3\cos(\lambda)U_3.\end{aligned}\tag{2.106}$$

Now we take the dot product with E_2 (recall Eq (2.95)). We find

$$\begin{aligned}E_2\cdot\nabla_V E_1 = &(-\sin(\theta))(-1)\big(v^2\cos(\lambda)\sin(\theta) + v^3\sin(\lambda)\cos(\theta)\big)\\ &+\cos(\theta)\big(v^2\cos(\lambda)\cos(\theta) - v^3\sin(\lambda)\sin(\theta)\big)\end{aligned}\tag{2.107}$$

After much algebra, this simplifies to

$$E_2\cdot\nabla_V E_1 = v^2\cos(\lambda).\tag{2.108}$$

Thus we conclude $\omega_{12}[V]$ simply picks out the second component of V relative to the E_i frame field and multiplies it by $\cos(\lambda)$, which is precisely what the one-form $\cos(\lambda)\,\mathrm{d}\theta$ does. We summarize this with the result:

$$\boxed{\omega_{12} = \cos(\lambda)\,\mathrm{d}\theta.}\tag{2.109}$$

At this point, the reader is invited to work through computing ω_{13}, ω_{23} for themselves. We will include the calculations later, so as to not tempt the reader.

97. Using the Covariant Derivative. If we want to know how fast the vector field W is changing with respect to the vector field V, we take the covariant derivative $\nabla_V W$. Given vector fields V, W, we can compute $\nabla_V W$ relative to the frame field E_i by this method:

$$\begin{aligned}\nabla_V W &= \nabla_V\left(\sum_i w^i E_i\right)\\ &= \vdots\quad\text{(exercise)}\\ &= \sum_i\left(V[w^i] - \sum_j \omega_{ij}[V]w^j\right)E_i.\end{aligned}\tag{2.110}$$

More explicitly,

$$\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} \mapsto \begin{bmatrix} V[w^1] \\ V[w^2] \\ V[w^3] \end{bmatrix} + \begin{bmatrix} 0 & -\omega_{12}[V] & -\omega_{13}[V] \\ \omega_{12}[V] & 0 & \omega_{23}[V] \\ \omega_{13}[V] & -\omega_{23}[V] & 0 \end{bmatrix} \begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix}. \tag{2.111}$$

Compare this to the Frenet situation,

$$\begin{bmatrix} f \\ g \\ h \end{bmatrix} \mapsto \begin{bmatrix} f' \\ g' \\ h' \end{bmatrix} + \begin{bmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix} \begin{bmatrix} f \\ g \\ h \end{bmatrix}, \tag{2.112}$$

where the (f', g', h') vector is really the directional derivative in the $\alpha(t)$ direction.

2.7 Cartan Structure Equations

(This is really important, but it's hard to grasp until we see some examples.)

98. Given some orthonormal frame fields E_1, E_2, E_3, we have two sets of associated one-forms:
- the coframe field θ^1, θ^2, θ^3 defined by $\theta^i[E_j] = \delta^i_j$.
- The connection forms ω_{ij}.

We may wonder how they are related to each other, and they're related in a very important way called the "Cartan Structure Equations":

$$\mathrm{d}\theta^i = \sum_j \omega_{ij} \wedge \theta^j \tag{2.113a}$$

which describes a "torsion free" condition, and

$$\mathrm{d}\omega_{ij} = \sum_k \omega_{ik} \wedge \omega_{kj}. \tag{2.113b}$$

This latter equation is the most important one, it's what will be generalized in differential geometry.

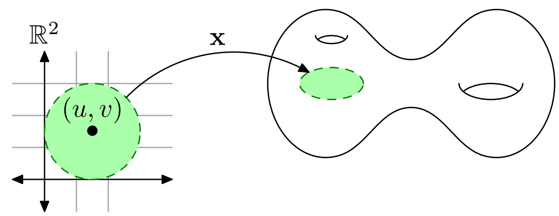
Exercises

1. For the curve $\alpha(t) = (t^4, -2t^2, 3t^{-1})$ for $t > 0$,
   (a) Compute the Frenet data;
   (b) Sketch the curve for $2 \leq t \leq 4$ show T, N, B at $t = 3$;
   (c) Find the limiting values of T, N, and B as $t \to +\infty$ and $t \to 0$.
2. (O'Neill, 2.4#3) The curve $\alpha(t) = (t\cos(t), t\sin(t), t)$ lies on a double cone and passes through the vertex at $t = 0$
   (a) Find the Frenet data of α at $t = 0$;
   (b) Sketch the curve for $-2\pi \leq t \leq 2\pi$, showing T, N, B at $t = 0$.
3. Let $V = -z^2U_1 - 4xU_2 + 5yU_3$, $W = \cos(xy)U_1 + z^2U_2$, compute the following:
   (a) $\nabla_V W$
   (b) $\nabla_V V$
   (c) $\nabla_V (zV - yW)$.
4. Prove or find a counter-example: if W is a vector field with constant length $\|W\| =$ constant, and if V is any vector field, then the covariant derivative $\nabla_V W$ is everywhere orthogonal to W.
5. Prove or find a counter-example: if $\Sigma \subset \mathbb{R}^n$ is a region containing a regular curve $\alpha\colon I \to \mathbb{R}^n$ (i.e., $\alpha(I) \subset \Sigma$), and if W is a vector field defined on Σ, then the mapping $t \mapsto W(\alpha(t))$ is a vector field on α called the **"Restriction"** of W to α and denoted $W|_\alpha$. The claim: $\nabla_{\alpha'(t)} W = (W|_\alpha)'(t)$.

3 Surfaces

99. What is a surface? Intuitively, it's an "$\mathbb{R}^2$" subset of $\mathbb{R}^3$. What does that even mean?

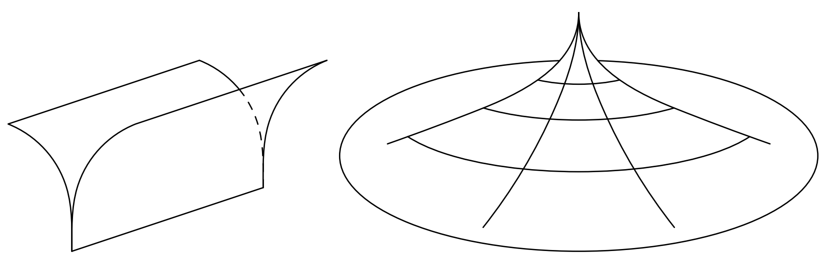
A set should be two-dimensional if it can be built out of pieces that look like open sets of $\mathbb{R}^2$, i.e., with two-dimensional patches of "fabric" we may "sew" together. Conceptually captured by this picture:



The map like the one above gives "coordinates" to each point on a surface. So a map like this is called a *coordinate patch*.[1]

100. Possible Problems With Our Definition. We should pause a moment and ponder if our notion of a surface is really well-defined, or if there are some problems with it. The main issues we should think about:

1. The coordinates could be "degenerate", meaning that different values of (u, v) correspond to the same point on the surface.
   SOLUTION: We demand the coordinate patch be injective[2] to avoid this problem.
2. Even if $\mathbf{x}$ is injective, it could behave badly in other ways and not define a smooth surface. For example, the following "surfaces" are too "pointy" to be smooth:



   SOLUTION: Require that $\mathbf{x}$ be *regular*.

101. Definition. Given a map $F\colon \mathbb{R}^m \to \mathbb{R}^n$ (suppose $m \leq n$), its **"Tangent Map"** at $\boldsymbol{p} \in \mathbb{R}^m$

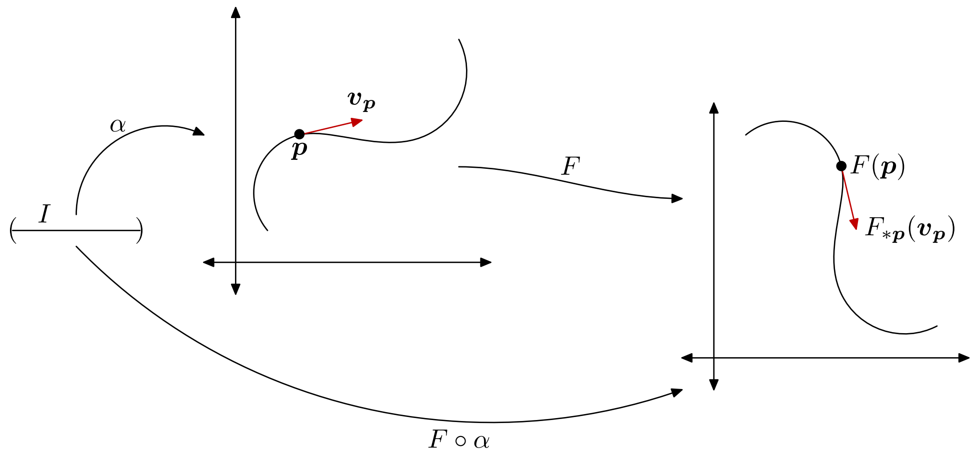
$$F_{*\boldsymbol{p}}\colon \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^m \to \mathrm{T}_{F(\boldsymbol{p})}\mathbb{R}^n$$

is defined as follows: given any $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^m$, pick some curve $\alpha\colon I \to \mathbb{R}^m$ such that $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. Then define

$$F_{*\boldsymbol{p}}(\boldsymbol{v}_{\boldsymbol{p}}) = \frac{\mathrm{d}}{\mathrm{d}t} F\big(\alpha(t)\big)\bigg|_{t=0}. \tag{3.1}$$

[1] This will be made precise shortly, but caution should be given: the literature is inconsistent on which way the arrow goes. Some authors prefer taking the green patch of the surface, and mapping it to some subset of $\mathbb{R}^2$. It is a matter of convention, and either choice is perfectly acceptable.

[2] Recall, a function $f\colon X \to Y$ is injective means for every $x_1, x_2 \in X$ we have $f(x_1) = f(x_2)$ implies $x_1 = x_2$.



101.1. Remark. We should intuitively think of $F_{*\boldsymbol{p}}$ as "the best linear approximation to F at $\boldsymbol{p}$".

101.2. Remark. This definition does not depend on choice of the curve α.

102. Definition. A map $F\colon \mathbb{R}^m \to \mathbb{R}^n$ is **"Regular"** if for every $\boldsymbol{p} \in \mathbb{R}^m$ we have $F_{*\boldsymbol{p}}$ be injective.

102.1. Remark. This is a good definition, because if α is a regular curve, and F is a regular map, then the composition $F \circ \alpha$ is a regular curve. Composing regular stuff together gives us something regular.

103. Definition. A **"(Coordinate) Chart"** in $\mathbb{R}^3$ is an injective regular map $\mathbf{x}\colon D \to \mathbb{R}^3$ where $D \subset \mathbb{R}^2$ is some open subset called the **"Patch"**.

Further, we call a chart **"Proper"** if $\mathbf{x}^{-1}\colon \mathbf{x}(D) \to \mathbb{R}^2$ is continuous.

We may abuse language, and refer to the D as the patch, and $\mathbf{x}$ as the chart or parametrization. Technically, the local coordinates refer to the components of the vector-valued function $\mathbf{x}^{-1}\colon \mathbf{x}(D) \to D$ mapping a patch of our surface to Euclidean space (the "space of parameters").

103.1. Remark (Abuse of language)*.* Again, just to reiterate, people mix up what they're referring to when using the terms "chart" and "patch". Undoubtedly *we* will too.

103.2. Remark. We must stress the importance of a patch $\mathbf{x}\colon D \to \mathbb{R}^3$ being regular, which means for any $(u, v) \in D$, the map

$$\mathbf{x}_*\colon \mathrm{T}_{(u,v)}\mathbb{R}^2 \to \mathrm{T}_{\mathbf{x}(u,v)}\mathbb{R}^3 \tag{3.2}$$

is injective.

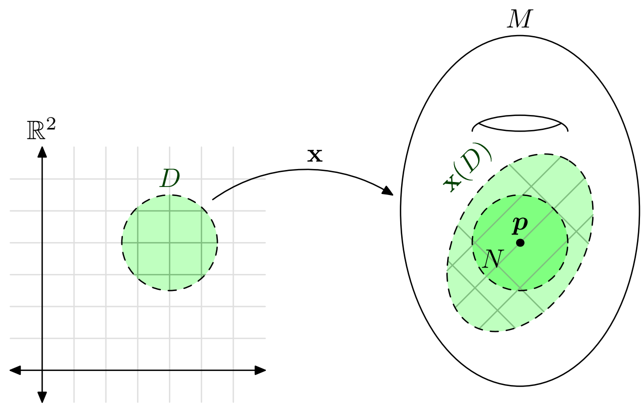
103.3. Remark. "Proper" patches convey topological information.

103.4. Remark. The image of any coordinate patch gives an example of a surface.

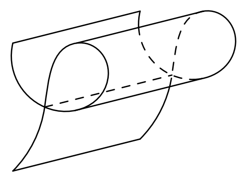
104. Most surfaces cannot be covered by one coordinate patch. The famous example: any coordinates on a sphere is degenerate around the poles. Consequently, we need to use a set of patches to define a surface.

105. Definition. Given a subset $M \subset \mathbb{R}^3$ and a point $\boldsymbol{p} \in M$, a **"Neighborhood"** of $\boldsymbol{p}$ is a set consisting of all points in M whose Euclidean distance to $\boldsymbol{p}$ is less than ε, for some $\varepsilon > 0$ [fixed for the neighborhood].

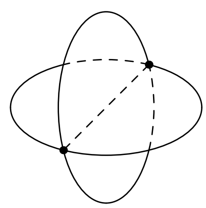
106. Definition. A **"Surface"** in $\mathbb{R}^3$ is a subset $M \subset \mathbb{R}^3$ such that for each point $\boldsymbol{p} \in M$ there exists a neighborhood N of $\boldsymbol{p}$ in M and a proper patch $\mathbf{x}\colon D \to \mathbb{R}^3$ such that $N \subset \mathbf{x}(D) \subset M$.



106.1. Remark. We don't want self-intersecting surfaces, we want to avoid the following doodle:



This is because there's no way to have a neighborhood "near the intersection". It would locally look like:



Why is this a problem?[3] This is a neighborhood of some point on the intersection, say $N(\boldsymbol{p})$. We would like to find a chart $\mathbf{x}\colon D\to\mathbb{R}^3$ such that $N(\boldsymbol{p})\subset\mathbf{x}(D)$. But this is impossible, because $\mathbf{x}(D)$ couldn't contain an intersection (thanks to topology).

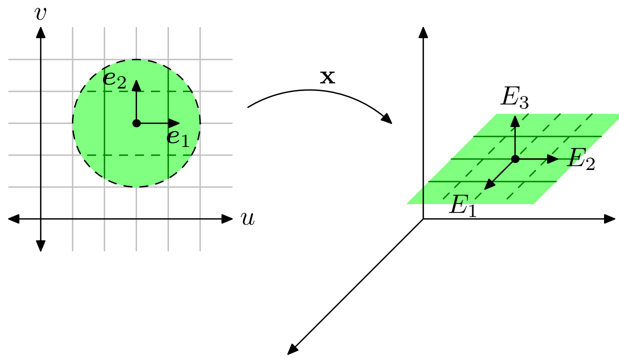
107. Determining if a Patch is Regular. How do we even determine if a patch is regular, anyways? Well, if $F\colon\mathbb{R}^m\to\mathbb{R}^n$ were regular at $\boldsymbol{p}\in\mathbb{R}^m$, then $F_{*\boldsymbol{p}}$ is injective. We know from linear algebra this means the dimension of the image equals the dimension of the domain, i.e.,

$$\dim(\mathrm{T}_{F(\boldsymbol{p})}\mathbb{R}^n)=\dim(\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^m). \tag{3.3}$$

This is equivalent to saying that the rank of $F_{*\boldsymbol{p}}$ is of maximal rank for every $\boldsymbol{p}\in\mathbb{R}^m$. In other words, we know a patch $\mathbf{x}\colon D\to\mathbb{R}^3$ is regular if for each $\boldsymbol{p}\in D$ we have $\mathbf{x}_{*\boldsymbol{p}}$ be of maximal rank. Our strategy for checking this will be to find some frame field $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ defined on D and some frame field E_1, E_2, E_3 on $\mathbf{x}(D)\subset\mathbb{R}^3$. Then we will express $\mathbf{x}_*$ as a 2×3 matrix, and we could use row reduction to find the rank.

What we do is we consider the following diagram:

[3] It's not Hausdorff, that's the problem.



Consider a curve $\alpha\colon I \to \mathbb{R}^2$ such that $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{e}_1$ — i.e., the curve points in the u-direction. We compute $\mathbf{x}_*(\boldsymbol{e}_1)$ to get the components in the first column of the matrix representing $\mathbf{x}_*$, and we find another curve pointing in the $\boldsymbol{e}_2$ (i.e., in the v) direction to find the second column of the matrix of $\mathbf{x}_*$.

We find,

$$\mathbf{x}_*(\boldsymbol{e}_1) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}\big(\alpha(t)\big)\Big|_{t=0} \tag{3.4a}$$

$$= \left(\frac{\partial x^1}{\partial\alpha^1}\frac{\mathrm{d}\alpha^1}{\mathrm{d}t} + \frac{\partial x^1}{\partial\alpha^2}\frac{\mathrm{d}\alpha^2}{\mathrm{d}t}, \frac{\partial x^2}{\partial\alpha^1}\frac{\mathrm{d}\alpha^1}{\mathrm{d}t} + \frac{\partial x^2}{\partial\alpha^2}\frac{\mathrm{d}\alpha^2}{\mathrm{d}t}, \frac{\partial x^3}{\partial\alpha^1}\frac{\mathrm{d}\alpha^1}{\mathrm{d}t} + \frac{\partial x^3}{\partial\alpha^2}\frac{\mathrm{d}\alpha^2}{\mathrm{d}t}\right)\Big|_{t=0} \tag{3.4b}$$

$$= \left(\frac{\partial x^1}{\partial u}, \frac{\partial x^2}{\partial u}, \frac{\partial x^3}{\partial u}\right) = \sum_{j=1}^{3}\frac{\partial x^j}{\partial u}E_j =: \mathbf{x}_u. \tag{3.4c}$$

Similarly, we find

$$\mathbf{x}_*(\boldsymbol{e}_2) = \left(\frac{\partial x^1}{\partial v}, \frac{\partial x^2}{\partial v}, \frac{\partial x^3}{\partial v}\right) = \sum_{j=1}^{3}\frac{\partial x^j}{\partial v}E_j =: \mathbf{x}_v. \tag{3.5}$$

We call the quantities x_u and x_v **"Partial Velocities"**. Hence, if $\boldsymbol{w}_{\boldsymbol{p}} = (w^1, w^2)_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^2$, then

$$\mathbf{x}_*\begin{pmatrix}w^1\\ w^2\end{pmatrix} = \begin{pmatrix}\dfrac{\partial x^1}{\partial u} & \dfrac{\partial x^1}{\partial v}\\ \dfrac{\partial x^2}{\partial u} & \dfrac{\partial x^2}{\partial v}\\ \dfrac{\partial x^3}{\partial u} & \dfrac{\partial x^3}{\partial v}\end{pmatrix}\begin{pmatrix}w^1\\ w^2\end{pmatrix}. \tag{3.6}$$

Now that we have expressed $\boldsymbol{x}_*$ as a matrix, we just need to check there are at least 2 linearly independent rows, which can be done by row reduction. Enough humourless logic, let us look at some examples.

108. Example. Consider the unit sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ in three-dimensions. We have a path formed by "bending" the open unit disc $D^2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$. More explicitly,

$$\mathbf{x}\colon D^2 \to \mathbb{R}^3 \tag{3.7a}$$

defined by

$$\mathbf{x}(u, v) = (u, v, \sqrt{1 - u^2 - v^2}). \tag{3.7b}$$

This is just one possible patch, we could consider another by taking the third component to be $-\sqrt{1 - u^2 - v^2}$, and we can consider others by swapping the third component with either the first or second components.

Now, our patch is clearly injective. If you do not believe it, then just examine $\mathbf{x}(u_1, v_1) = \mathbf{x}(u_2, v_2)$; the first two components reads $u_1 = u_2$ and $v_1 = v_2$. It follows that $(u_1, v_1) = (u_2, v_2)$ and moreover $\mathbf{x}$ is injective.

But is our patch *regular*? We can find the matrix of $\mathbf{x}_*$ relative to the canonical frame fields, which reads

$$\mathbf{x}_* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \text{stuff}_1 & \text{stuff}_2 \end{bmatrix}. \tag{3.8}$$

Since the top 2×2 submatrix is the identity matrix, it follows that $\mathbf{x}_*$ has rank 2. Hence our patch is regular.

108.1. Remark. This example is a special case of a more general fact: if we have a smooth function $f\colon D\to\mathbb{R}$, and we consider its graph $\Gamma(f) = \{(x,y,f(x,y))\in\mathbb{R}^3 \mid (x,y)\in D\}$ (or more generally, for any $D\subset\mathbb{R}^n$, we have $\Gamma(f) = \{(\boldsymbol{x}, f(\boldsymbol{x}))\in\mathbb{R}^{n+1} \mid \boldsymbol{x}\in D\}$), then this graph is a patch of a surface.

Algebraic geometry generalizes this further, by studying the zero sets of functions $\{\boldsymbol{x}\in\mathbb{R}^n \mid f(\boldsymbol{x}) = 0\}$. These generalize the notion of surfaces. A lot of differential geometry is generalized in this manner, it's very deep and profound.

109. Example (Surface of revolution)**.** Undergraduates are taught in integral calculus of a single variable about a surface of revolution by taking a curve $y = f(x)$, then sweeping it out around the x-axis, in the sense that

$$y^2 + z^2 = f(x)^2. \tag{3.9}$$

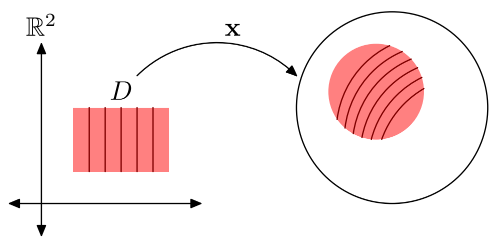
This yields a parametrization in terms of x and θ. Our patch would be

$$\mathbf{x}(x,\theta) = (x, f(x)\cos(\theta), f(x)\sin(\theta)). \tag{3.10}$$

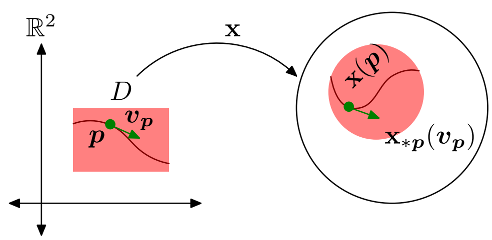
If f is a regular curve, then we have a regular surface.

110. What Patches Give Us. The basic idea for patches is that they let us transfer data on the surface M to data (of various kinds) on the domain $D\subset\mathbb{R}^2$ where we know how to do calculus. What kinds of things do patches give us?

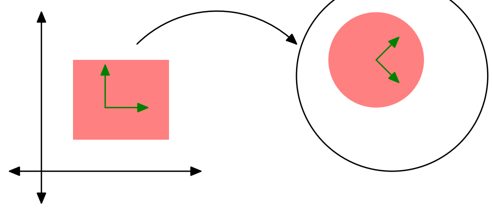
1. "Local coordinates" on M. Grid lines in D are paths like $\alpha(t) = (u_0, v_0+t)$ which pass through the point (u_0, v_0). These induce grid lines on M by $\mathbf{x}\circ\alpha(t) = (u_0, v_0+t)$. (Although these describe grid lines of constant u_0, we can form grid lines of constant v_0 by examining (u_0+t, v_0) for example.)



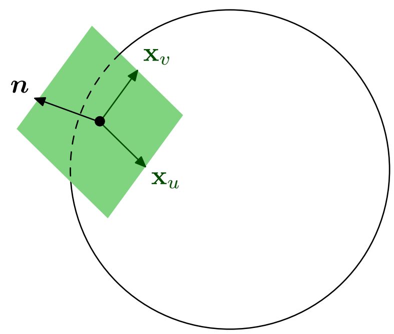
2. Convenient ways to get tangent vectors on M. We have regularity map basis vectors (frame fields) to basis vectors (frame fields) by $\mathbf{x}_{*\boldsymbol{p}}\colon \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^2\to\mathrm{T}_{\mathbf{x}(\boldsymbol{p})}M$. Regularity guarantees we get a whole tangent plane, not just a line.



3. "Local" frame fields on M. This is given to us by the partial velocities $\mathbf{x}_u(u,v)$ and $\mathbf{x}_v(u,v)$.



4. Convenient ways to compute "the normal vector" to a surface. Since we have found $\mathbf{x}_u(u,v)$ and $\mathbf{x}_v(u,v)$ are frame fields for the tangent vectors on the surface, we can consider their cross product $\boldsymbol{n} = \mathbf{x}_u(u,v)\times\mathbf{x}_v(u,v)$ which is normal to the surface. We can do this globally only for "orientable" surfaces (e.g., not for the Möbius strip).



Exercises

Here are some review questions, to make sure you don't forget too quickly what we learned from section 2. (This is an experiment, let me know if you hate this technique. It probably won't happen again in these notes, though.)

1. Recall Parabolic coordinates — we have $0\leq u<\infty$, $0\leq v<\infty$, and $0\leq\varphi<2\pi$, and the Cartesian coordinates are parametrized as

$$x = uv\cos(\varphi) \tag{3.11a}$$

$$y = uv\sin(\varphi) \tag{3.11b}$$

$$z = \frac{1}{2}(u^2 - v^2). \tag{3.11c}$$

   (a) Compute the Parabolic frame field E_1, E_2, E_3
   (b) Compute the connection forms for the parabolic frame field.

2. Let E_1, E_2, E_3 constitute a frame field and $W = \sum_j f_j E_j$. Let V be an arbitrary vector field. Prove or find a counter-example: the covariant derivative satisfies,

$$\nabla_V W = \sum_j \left(V[f_j] + \sum_i f_i \omega_{ij}[V] \right) E_j. \tag{3.12}$$

3. Check the structure equations for the parabolic frame field.

3.1 Calculus on a Surface

OK, we're on the home stretch now. We've generalized calculus in $\mathbb{R}^n$ using the machinery of tangent vectors and differential forms, talked about curves and surfaces. Now our goal is to figure out how to do calculus on surfaces. Ready? Let's go!

111. Our goal is to completely generalize what we know about calculus on $\mathbb{R}^2$ to any surface.

This means we need to define: tangent vectors, vector fields, frame fields, one-forms, differential forms, smooth functions, covariant derivatives, etc., *on a surface.*[4] We'll use this to study various surfaces and properties they have.

What's most fundamental in mathematics is making the correct definitions.

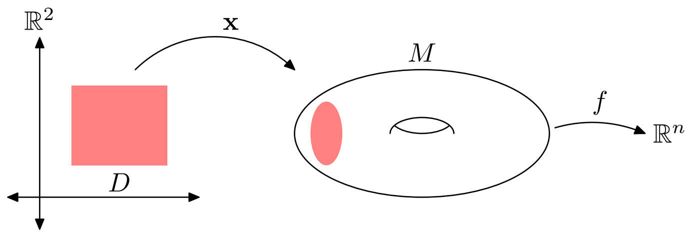


Figure 1: Intuition of a function $f\colon M \to \mathbb{R}^n$ being smooth

112. Definition. Let $M \subset \mathbb{R}^3$ be a surface, $f\colon M \to \mathbb{R}^n$ be some function on the surface. We say that f is **"Smooth"** if, for every patch $\mathbf{x}\colon D \to M$ (where $D \subset \mathbb{R}^2$ is open),

$$f \circ \mathbf{x}\colon D \to \mathbb{R}^n \tag{3.13}$$

is smooth in the usual sense, i.e., $f \circ \mathbf{x} \in C^\infty(D)$. This is schematically doodled in Figure 1.

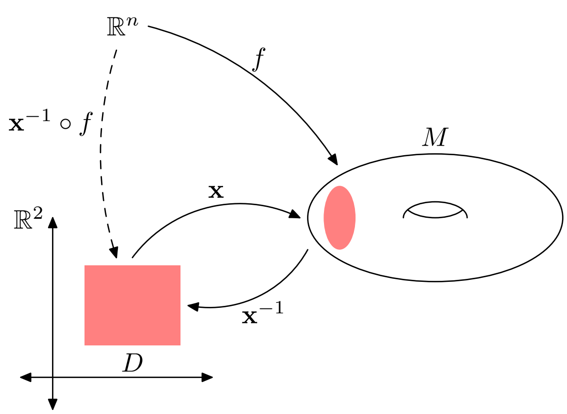


Figure 2: Intuition of a function $f\colon \mathbb{R}^n \to M$ being smooth

113. Definition. Let $M \subset \mathbb{R}^3$ be a surface, $f\colon \mathbb{R}^n \to M$ be a function to the surface. We call f **"Differentiable"** if, for every patch $\mathbf{x}\colon D \to M$, the map

$$\mathbf{x}^{-1} \circ f\colon \mathcal{E} \to D, \tag{3.14}$$

where $\mathcal{E} = f^{-1}(\mathbf{x}(D)) \subset \mathbb{R}^n$ is the preimage of the patch under f, is a smooth (C^∞) function. Note: we *do not* require $f(\mathbb{R}^n) = M$.

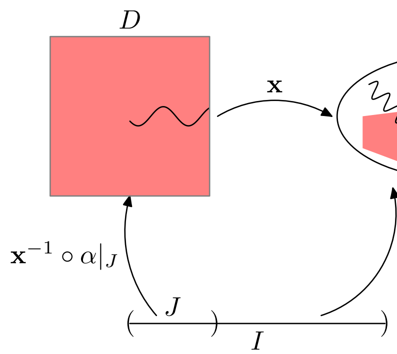
The intuition of this definition is doodled in Figure 2.

[4]The generalization to arbitrary manifolds will be simple.

113.1. Remark. This requires a bit of explanation. Consider the preimage of $\mathbf{x}(D)$ under f, i.e., the set of points $\boldsymbol{x} \in \mathbb{R}^n$ which f maps into the image of the chart $\mathbf{x}(D)$; call this set $\mathcal{E} = \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) \in (D)\}$. We want this to be an open set for topological reasons (this makes f continuous, a necessary prerequisite for derivatives). Now we could consider $f|_{\mathcal{E}} \colon \mathcal{E} \to M$ by restricting f. We know its image will be within the image of the chart, so we then take the preimage of $f(\mathcal{E}) \subset M$ under the chart $\mathbf{x}$ to produce the mapping $\mathbf{x}^{-1} \colon f|_{\mathcal{E}}(\mathcal{E}) \to D$. But this is the same as considering the composition $(x)^{-1} \circ f|_{\mathcal{E}} \colon \mathcal{E} \to D$. The restriction of f to $\mathcal{E}$ has been purely a crutch, the preimage of $\mathbf{x}$ will restrict the composite function for us. So we arrive at our definition.

113.2. Remark. As a quick check, we could consider $M = \mathbb{R}^3$ with $\mathbf{x} = \mathrm{id}$ being the identity function. Then $f \colon \mathbb{R}^n \to \mathbb{R}^3$ being differentiable is the same as $f \in C^\infty(\mathbb{R}^n, \mathbb{R}^3)$. This is good! Our definition of differentiable functions to surfaces coincides with our pre-existing definition of differentiable multivariate vector-valued functions.

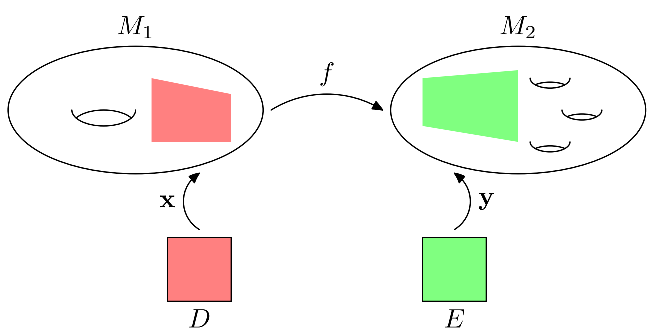
114. Example. Let $M \subset \mathbb{R}^3$ be a surface, and consider a [smooth] path $\alpha \colon I \to \mathbb{R}^3$ such that the curve lies on the surface $\alpha(I) \subset M$. For any patch $\mathbf{x} \colon D \to M$, we could consider the interval $J = \alpha^{-1}(\mathbf{x}(D))$ given by the preimage of the curve which lies in $\mathbf{x}(D)$. The situation is as doodled below:



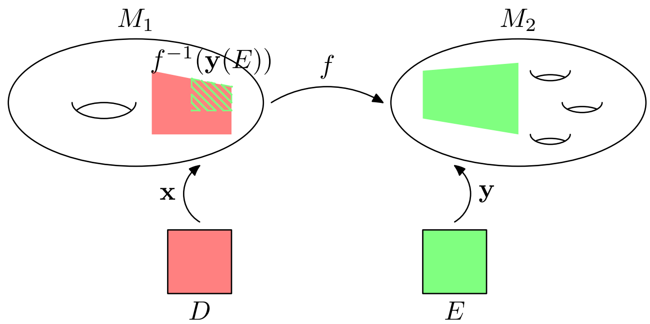
Proving α is smooth on M amounts to proving, for every patch $\mathbf{x} \colon D \to M$ such that $\mathbf{x}(D)$ contains some part of the path, the restriction $\alpha|_J$ is smooth in the preimage of the chart in the familiar way.

115. Smooth Functions Between Surfaces. Suppose now we have two surfaces M_1 and M_2. We can construct a notion of a smooth function $f \colon M_1 \to M_2$ between these surfaces. The solution is to cheat.

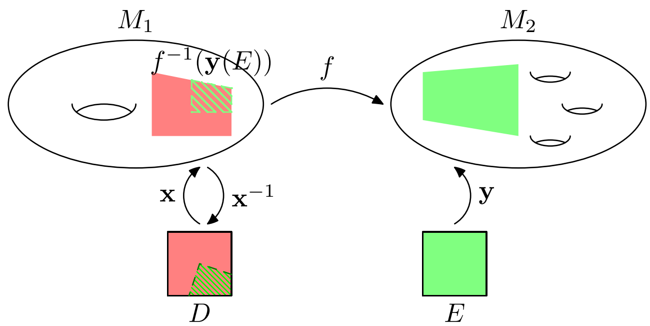
Given arbitrary patches $\mathbf{x} \colon D \to M_1$ on M_1 and $\mathbf{y} \colon E \to M_2$ on M_2, we have the situation as doodled below:



Now we have to make sense of f. We first take the preimage of $\mathbf{y}(E)$ under f, which may or may not intersect $\mathbf{x}(D)$ on M_1. If it doesn't, then we're in the trivial situation, and everything works out fine. So let's examine the exciting case where $f^{-1}(\mathbf{y}(E)) \cap \mathbf{x}(D) \neq \emptyset$. This gives us the etched region doodled below:



We can pull back $f^{-1}(\mathbf{y}(E))$ to the patch D using the preimage of $\mathbf{x}$, which produces the following situation (with the hatched region indicating the $\mathbf{x}^{-1} \circ f^{-1}$ preimage):



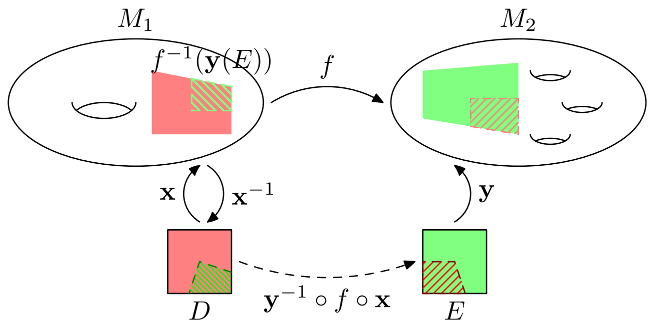
It looks like we're making this more complicated, doesn't it? There is one thing we have not yet exploited: we can move *forward* as well as backward. If we start with $\mathbf{x}|^{-1}_{f^{-1}(E)}(D)$ the portion of the patch which, when charted onto M_1 will be mapped by f to part of $\mathbf{y}(E)$, then move forward along these lines, we end up with a subset

$$(f \circ \mathbf{x})\left(\mathbf{x}|^{-1}_{f^{-1}(E)}(D)\right) \subset \mathbf{y}(E). \tag{3.15}$$

We can take its preimage under $\mathbf{y}$ to get a subset in $E \subset \mathbb{R}^2$. This gives us a mapping, however, from D to E:

$$\mathbf{y}^{-1} \circ f \circ \mathbf{x} \colon D \to E, \tag{3.16}$$

which is a function where we can sensibly discuss smoothness and derivatives. In pictures, we get the situation as follows:



The induced function is drawn with a dashed arrow, and it is the one *we know* how to determine if it's smooth or not (because it's a function of an open subset in $\mathbb{R}^2$ to an open subset in $\mathbb{R}^2$). And if we do this for every possible pair of patches on M_1 and M_2, we end up verifying f is smooth and differentiable.

More precisely, we have $f \circ \mathbf{x}$ be smooth function to M_2 in the sense of Definition 113 We also have $\mathbf{y}^{-1} \circ f$ be a smooth function on M_1, in the sense of Definition 112. Since this is done for every possible charts on M_1 and M_2, we conclude that $f \colon M_1 \to M_2$ is smooth.

Exercises

1. Partial velocities $\mathbf{x}_u$, $\mathbf{x}_v$ are defined for an arbitrary mapping $\mathbf{x}\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$, so we can consider the [real-valued] functions

$$E = \mathbf{x}_u \cdot \mathbf{x}_u, \quad F = \mathbf{x}_u \cdot \mathbf{x}_v, \quad G = \mathbf{x}_v \cdot \mathbf{x}_v \tag{3.17}$$

on D.
(a) Prove $\|\mathbf{x}_u \times \mathbf{x}_v\|^2 = EG - F^2$.
(b) Prove $\mathbf{x}$ is regular if and only if $EG - F^2$ is never zero.

Homework: Stereographic Projection

Mathematics 116 — Differential Geometry

Spring 2008

Derek Wise

Stereographic projection gives a nice coordinate patch on the unit sphere $x^2+y^2+z^2 = 1$. It is defined by

$$\boldsymbol{x}\colon \mathbb{R}^2 \to \mathbb{R}^3$$

where $\boldsymbol{x}(u, v)$ is defined to be the unique point in $\mathbb{R}^3$ that lies both on the unit sphere and the ray from $(0, 0, 1)$ through $(u, v, 0)$.

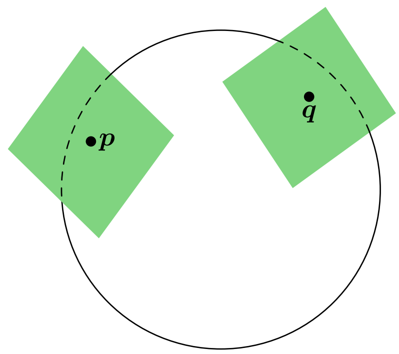
1. Derive an explicit formula for $\boldsymbol{x}(u, v)$. [Hint: use a parameterization of the line, and solve for the time t when it passes through the unit sphere.]
2. Find the matrix of the tangent map $\boldsymbol{x}_*$, relative to the natural frame fields on $\mathbb{R}^2$ and $\mathbb{R}^3$.
3. Prove that $\boldsymbol{x}$ is a patch.
4. Show that $\boldsymbol{x}$ is *conformal*, meaning that it preserves angles. That is, given a pair of tangent vectors w_p, z_p at the same point in $\mathbb{R}^2$, show that the angle between them (defined by the dot product in $\mathbb{R}^2$) is the same as the angle between $\boldsymbol{x}_*(w_p)$ and $\boldsymbol{x}_*(z_p)$ (defined by the dot product in $\mathbb{R}^3$).

3.2 Vectors on Surfaces

116. We require $\mathbf{x}\colon D \to \mathbb{R}^3$ be smooth (C^∞) and $\mathbf{x}^{-1}\colon D \leftarrow \mathbf{x}(D)$ be continuous. We require an additional property for the inverse to be differentiable.

117. Definition. Let M be a surface, let $\boldsymbol{p} \in M$ be some point. We define the **"Tangent Space"** to $\boldsymbol{p}$ in M, denoted $\mathrm{T}_{\boldsymbol{p}}M$, is the set of all vectors $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}\mathbb{R}^3$ such that $\boldsymbol{v}_{\boldsymbol{p}} = \alpha'(0)$ for some curve on the surface $\alpha\colon I \to M$.

118. Base points are important. In $\mathbb{R}^2$, we often just ignored the point of tangency and pretended that $\mathrm{T}_{\boldsymbol{p}}\mathbb{R}^2$ and $\mathrm{T}_{\boldsymbol{q}}\mathbb{R}^2$ are the same just by sliding $\boldsymbol{p}$ to $\boldsymbol{q}$. But for general surfaces (of which $\mathbb{R}^2$ is just the most boring example), we cannot do this. There is no way to slide $\mathrm{T}_{\boldsymbol{p}}M$ to $\mathrm{T}_{\boldsymbol{q}}M$ without embedding M into $\mathbb{R}^2$.



119. Definition. A **"Vector Field"** V on a surface M is an assignment to each point $\boldsymbol{p} \in M$ a vector $V(\boldsymbol{p}) \in \mathrm{T}_{\boldsymbol{p}}M$.

120. Originally, we defined the derivative of f in the direction of some tangent vector $\boldsymbol{v}_{\boldsymbol{p}}$ as

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{p} + t\boldsymbol{v})\Big|_{t=0}. \tag{3.18}$$

This captures the information of how much f changes in the direction of $\boldsymbol{v}$ (at base-point $\boldsymbol{p}$). There's a problem generalizing this to a surface: it doesn't work if f is defined only on M. Why not? Well, the line $\boldsymbol{p} + t\boldsymbol{v}$ will leave the surface, and f is undefined off the surface, so we're out of luck.

But later we proved, if $\alpha\colon I \to \mathbb{R}^n$ passes through $\boldsymbol{p} = \alpha(0)$ and it has velocity $\boldsymbol{v}_{\boldsymbol{p}} = \alpha'(0)$ there, then we could define the directional derivative as:

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\alpha(t))\Big|_{t=0}. \tag{3.19}$$

So if $\alpha\colon I \to M$ has initial position $\alpha(0) = \boldsymbol{p}$ and initial velocity $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}M$, and if $f\colon M \to \mathbb{R}$ is smooth, then we can define the directional derivative of a function on our surface by

$$\boldsymbol{v}_{\boldsymbol{p}}[f] = \frac{\mathrm{d}}{\mathrm{d}t} f(\alpha(t))\Big|_{t=0}. \tag{3.20}$$

This is independent of the choice of such α.

If we do this at every point, we can differentiate functions with respect to vector fields using

$$V[f](\boldsymbol{p}) = V(\boldsymbol{p})[f]. \tag{3.21}$$

This works out perfectly.

120.1. Remark (Boring). This should be boring, because we defined things in a clever way. Generalizations follow easily once we have the right definitions. So if you find this boring, good: it means you have a grasp of the concepts underlying the definitions.

3.3 Differential Forms on Surfaces

121. A one-form ϕ on M assigns to each point $\boldsymbol{p} \in M$ a covector on $\mathrm{T}_{\boldsymbol{p}}M$, i.e., a linear function

$$\phi_{\boldsymbol{p}} \colon \mathrm{T}_{\boldsymbol{p}}M \to \mathbb{R}. \tag{3.22}$$

The most important examples: let $f \colon M \to \mathbb{R}$ be a smooth function, then $\mathrm{d}f$ is a one-form given by

$$\mathrm{d}f[\boldsymbol{v}_{\boldsymbol{p}}] = \boldsymbol{v}_{\boldsymbol{p}}[f], \tag{3.23}$$

for every $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}M$. But let us see what the zoo of differential forms becomes on a surface.

122. Zero-Forms. The 0-forms are just smooth functions on M, i.e., functions like $\phi \colon M \to \mathbb{R}$ such thatfor every patch $\mathbf{x} \colon D \to M$, the function $\phi \colon \mathbf{x} \colon D \to \mathbb{R}$ is smooth.

123. One-Forms. The 1-forms are defined just like in Euclidean space. A **"One-Form"** ϕ is a linear map at each point $\boldsymbol{p} \in M$ taking tangent vectors to real numbers

$$\phi_{\boldsymbol{p}} \colon \mathrm{T}_{\boldsymbol{p}}M \to \mathbb{R} \tag{3.24}$$

in a linear way, and taking vector fields to functions

$$\phi \colon \ \mathrm{Vect}(M) \to C^{\infty}(M). \tag{3.25}$$

Given a zero-form f, the differential $\mathrm{d}f$ is the one-form given by

$$\mathrm{d}f[V] = V[f], \tag{3.26}$$

for any vector field $V \in \mathrm{Vect}(M)$.

124. Two-Forms. Now we have something slightly different. But it tells us what 2-forms *do*. A 2-form η on M is a map at each point $\boldsymbol{p} \in M$ that takes an *ordered pair* of tangent vectors and gives a number, that is to say,

$$\eta_{\boldsymbol{p}} \colon \mathrm{T}_{\boldsymbol{p}}M \times \mathrm{T}_{\boldsymbol{p}}M \to \mathbb{R} \tag{3.27}$$

such that

1. Antisymmetry: $\eta(\boldsymbol{v}, \boldsymbol{w}) = -\eta(\boldsymbol{w}, \boldsymbol{v})$
2. Linearity in first slow: for any $a, b \in \mathbb{R}$, $\eta(a\boldsymbol{u} + b\boldsymbol{v}, \boldsymbol{w}) = a\eta(\boldsymbol{u}, \boldsymbol{w}) + b\eta(\boldsymbol{v}, \boldsymbol{w})$.

It's easy to prove from these two properties that a 2-form is also linear in the second slot; that is to say, it's *bilinear*.

If we use η at every point, we get a mapping

$$\eta \colon \ \mathrm{Vect}(M) \times \mathrm{Vect}(M) \to C^{\infty}(M). \tag{3.28}$$

125. But...the wedge product? Earlier we defined 2-forms in terms of the formal wedge product. Let us now endeavour to produce a definition of the wedge product for differential forms on a surface which is consistent with how we defined 2-forms.

Let ϕ, ψ be two 1-forms on M. We want to make a 2-form out of them, and call it $\phi \wedge \psi$ (and make it a mapping $\mathrm{Vect}(M) \times \mathrm{Vect}(M) \to C^{\infty}(M)$). The most obvious thing we could try is,

$$(\phi \wedge \psi)(V, W) = \phi(V)\psi(W). \tag{3.29}$$

Does it work? No, not by a long shot, since

$$(\phi \wedge \psi)(V, W) = \phi(V)\psi(W) \neq -\phi(W)\psi(V) \text{ in general.} \tag{3.30}$$

Let us try

$$(\phi \wedge \psi)(V, W) \stackrel{???}{=} \phi(V)\psi(W) - \phi(W)\psi(V). \tag{3.31}$$

Does it work?

We can see it is antisymmetric, since

$$\begin{aligned}(\phi\wedge\psi)(V,W) &= \phi(V)\psi(W)-\phi(W)\psi(V) && (3.32a)\\ &= -\phi(W)\psi(V)-(-\phi(V)\psi(W)) && (3.32b)\\ &= -(\phi\wedge\psi)(W,V), && (3.32c)\end{aligned}$$

which is a relief. So this is possibly a good definition.

Now the real moment of truth: is it linear in the first slot? We have something stronger than *mere* linearity, it's linear with respect to arbitrary smooth function $f, g \in C^{\infty}(M)$, we have

$$(\phi\wedge\psi)(fU+gV,W) = f(\phi\wedge\psi)(U,W)+g(\phi\wedge\psi)(V,W). \tag{3.33}$$

This is awesome!

And what's really cute: we had an axiom (§34) that the formal wedge product is anticommutative on 1-forms. We see that

$$\begin{aligned}(\phi\wedge\psi)(V,W) &= \phi(V)\psi(W)-\phi(W)\psi(V) && (3.34a)\\ &= -(-\phi(V)\psi(W)+\phi(W)\psi(V)) && (3.34b)\\ &= -(-\psi(W)\phi(V)+\psi(V)\phi(W)) && (3.34c)\\ &= -(\psi\wedge\phi)(V,W). && (3.34d)\end{aligned}$$

In fact we have

$$\begin{matrix}(\phi\wedge\psi)(V,W) & = & -(\phi\wedge\psi)(W,V)\\ \| & & \| \\ -(\psi\wedge\phi)(V,W) & = & (\psi\wedge\phi)(W,V).\end{matrix} \tag{3.35}$$

It's consistent!

125.1. Remark. We see the 3-form a 2-dimensional surface $M\subset\mathbb{R}^3$ is zero.

126. Corollary: Nilpotence. The reader can verify that, for any one-form ϕ, we have $\phi\wedge\phi=0$. We proved this formally, as a consequence of antisymmetry, but the reader may verify this is true for our concrete realization of the wedge product.

127. Computing 2-form using basis vectors. Let us consider a 2-form η and suppose $\boldsymbol{e}_1$, $\boldsymbol{e}_2\in \mathrm{T}_{\boldsymbol{p}}M$ is a basis. We will do some multilinear algebra: once we know how η acts on all possible linear combinations of our basis vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, then we will know how it acts on any arbitrary vector in $\mathrm{T}_{\boldsymbol{p}}M$.

Consider

$$\begin{aligned}\eta(\alpha\boldsymbol{e}_1+\beta\boldsymbol{e}_2,\gamma\boldsymbol{e}_1+\delta\boldsymbol{e}_2) &= \alpha\eta(\boldsymbol{e}_1,\gamma\boldsymbol{e}_1+\delta\boldsymbol{e}_2)+\beta\eta(\boldsymbol{e}_2,\gamma\boldsymbol{e}_1+\delta\boldsymbol{e}_2) && (3.36a)\\ &= \alpha(\gamma\eta(\boldsymbol{e}_1,\boldsymbol{e}_1)+\delta\eta(\boldsymbol{e}_1,\boldsymbol{e}_2))+\beta(\gamma\eta(\boldsymbol{e}_2,\boldsymbol{e}_1)+\delta\eta(\boldsymbol{e}_2,\boldsymbol{e}_2)) && (3.36b)\\ &= \alpha\gamma\eta(\boldsymbol{e}_1,\boldsymbol{e}_1)+\alpha\delta\eta(\boldsymbol{e}_1,\boldsymbol{e}_2)+\beta\gamma\eta(\boldsymbol{e}_2,\boldsymbol{e}_1)+\beta\delta\eta(\boldsymbol{e}_2,\boldsymbol{e}_2) && (3.36c)\\ &= 0+\alpha\delta\eta(\boldsymbol{e}_1,\boldsymbol{e}_2)-\beta\gamma\eta(\boldsymbol{e}_1,\boldsymbol{e}_2)+0 && (3.36d)\\ &= (\alpha\delta-\beta\gamma)\eta(\boldsymbol{e}_1,\boldsymbol{e}_2) && (3.36e)\\ &= \det\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\eta(\boldsymbol{e}_1,\boldsymbol{e}_2). && (3.36f)\end{aligned}$$

We only need to compute $\eta(\boldsymbol{e}_1,\boldsymbol{e}_2)$ once, and then computing $\eta(\boldsymbol{v},\boldsymbol{w})$ amounts to computing a determinant.

128. Exterior Derivative. Let us talk about one forms on $\mathbb{R}^2$, call such a 1-form ϕ and let $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ be the natural frame field in the u^1, u^2 coordinate directions. By the above formula, to figure out $\mathrm{d}\phi$, we just need to know what $\mathrm{d}\phi(\boldsymbol{e}_1,\boldsymbol{e}_2)$ is. In $\mathbb{R}^2$, we know any one-form can be written as

$$\phi = f_1\,\mathrm{d}u^1+f_2\,\mathrm{d}u^2. \tag{3.37}$$

We have

$$\mathrm{d}\phi(\boldsymbol{e}_1,\boldsymbol{e}_2) = \mathrm{d}(f_1\,\mathrm{d}u^1 + f_2\,\mathrm{d}u^2)(\boldsymbol{e}_1,\boldsymbol{e}_2) = (\mathrm{d}f_1\wedge\mathrm{d}u^1 + \mathrm{d}f_2\wedge\mathrm{d}u^2)(\boldsymbol{e}_1,\boldsymbol{e}_2) \tag{3.38a}$$

$$= \left(\left(\frac{\partial f_1}{\partial u^1}\mathrm{d}u^1 + \frac{\partial f_1}{\partial u^2}\mathrm{d}u^2\right)\wedge\mathrm{d}u^1 + \left(\frac{\partial f_2}{\partial u^1}\mathrm{d}u^1 + \frac{\partial f_2}{\partial u^2}\mathrm{d}u^2\right)\wedge\mathrm{d}u^2\right)(\boldsymbol{e}_1,\boldsymbol{e}_2) \tag{3.38b}$$

$$= \left(\frac{\partial f_1}{\partial u^2}\mathrm{d}u^2\wedge\mathrm{d}u^1 + \frac{\partial f_2}{\partial u^1}\mathrm{d}u^1\wedge\mathrm{d}u^2\right)(\boldsymbol{e}_1,\boldsymbol{e}_2) \tag{3.38c}$$

$$= \left[\left(\frac{\partial f_2}{\partial u^1} - \frac{\partial f_1}{\partial u^2}\right)\mathrm{d}u^1\wedge\mathrm{d}u^2\right](\boldsymbol{e}_1,\boldsymbol{e}_2) \tag{3.38d}$$

$$= \left(\frac{\partial f_2}{\partial u^1} - \frac{\partial f_1}{\partial u^2}\right)\left(\mathrm{d}u^1(\boldsymbol{e}_1)\mathrm{d}u^2(\boldsymbol{e}_2) - \mathrm{d}u^1(\boldsymbol{e}_2)\mathrm{d}u^2(\boldsymbol{e}_1)\right) \tag{3.38e}$$

$$= \left(\frac{\partial f_2}{\partial u^1} - \frac{\partial f_1}{\partial u^2}\right)(1-0) = \left(\frac{\partial f_2}{\partial u^1} - \frac{\partial f_1}{\partial u^2}\right). \tag{3.38f}$$

We stipulated in Eq (3.37) that we could write a one-form ϕ using f_1, f_2. This let us find $\mathrm{d}\phi(\boldsymbol{e}_1,\boldsymbol{e}_2)$. Could we perform our calculation without this stipulation? That is to say, can we write $\mathrm{d}\phi(\boldsymbol{e}_1,\boldsymbol{e}_2)$ in terms of derivatives of ϕ? Yes! We can, thus:

$$\mathrm{d}\phi(\boldsymbol{e}_1,\boldsymbol{e}_2) = \frac{\partial}{\partial u^2}\phi(u^1) - \frac{\partial}{\partial u^1}\phi(u^2). \tag{3.39}$$

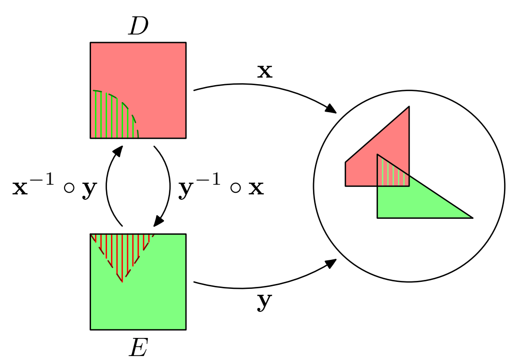
This motivates the following definition of exterior derivatives of one-forms on surfaces, recalling the partial velocities form a frame field on the surface.

129. Definition. On a surface $M\subset\mathbb{R}^3$, in a given patch, we have coordinate vector fields $\mathbf{x}_u$, $\mathbf{x}_v$. We define for any 1-form ϕ on M the 2-form $\mathrm{d}\phi$ on M given by

$$\mathrm{d}\phi(\mathbf{x}_u,\mathbf{x}_v) = \frac{\partial}{\partial u}\phi(\mathbf{x}_v) - \frac{\partial}{\partial v}\phi(\mathbf{x}_u), \tag{3.40}$$

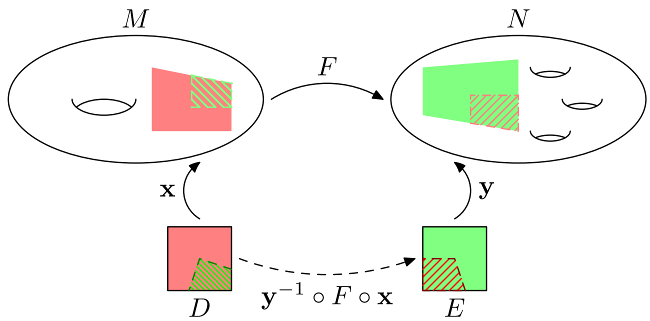
for any patch.

130. Consistency of Definitions. One potential problem is that our definition might not make sense. We demand consistency when the patches overlap, like the situation doodled below:



Whenever we define *anything* in differential geometry, we *must* worry about consistency on the overlap of patches.

131. Maps of Surfaces. Recall our discussion of smooth maps between surfaces (§115). If we have smooth surfaces M and N, is there a smooth function $F\colon M \to N$? In order for F to be smooth, we need for any patch $D \subset \mathbb{R}^2$ with chart $\mathbf{x}\colon D \to M$ and for any patch $E \subset \mathbb{R}^2$ with chart $\mathbf{y}\colon E \to N$, there exists a map $f\colon D \to E$ (defined by $f = \mathbf{y}^{-1} \circ F \circ \mathbf{x}$) which is smooth in the usual sense. We have the situation similar to what we have doodled below:



Now, we have the tangent map in such an approach, defined for $f\colon D \to E$ using Definition 101. We patch them together to induce a tangent map associated for F.

132. Definition. Let M, N be surfaces, $F\colon M \to N$ be a smooth map. The **"Tangent Map"** is defined, for each $\boldsymbol{p} \in M$, as $F_*\colon \mathrm{T}_{\boldsymbol{p}}M \to \mathrm{T}_{F(\boldsymbol{p})}N$.

133. The really slick way to approach this is if we have some path $\alpha(t)$ that goes through the given point $\boldsymbol{p} = \alpha(0)$. Then we have

$$F_*(\alpha'(0)) = \frac{\mathrm{d}}{\mathrm{d}t}F(\alpha(t))\bigg|_{t=0}. \tag{3.41}$$

We call $F_*(\boldsymbol{v_p})$ the **"Pushforward"** of $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M$ along F.

134. Recall that one-forms are maps from tangent vector spaces to $\mathbb{R}$. Given a one-form ϕ on N, and a smooth map $F\colon M \to N$, we can obtain a one-form on M called the **"Pullback"** of ϕ along F, denoted $F^*(\phi)$. This is defined by, for any $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M$,

$$(F^*\phi)[\boldsymbol{v_p}] = \phi[F_*(\boldsymbol{v_p})]. \tag{3.42}$$

In short, any smooth map $F\colon M \to N$ on surfaces gives me two induced maps, the pushforward

$$F_*\colon \{\text{tangent vectors on } M\} \to \{\text{tangent vectors on } N\}$$

and, going in the opposite direction, the pullback which maps one-forms to one-form,

$$F^*\colon \{\text{1-forms on } M\} \leftarrow \{\text{1-forms on } N\}.$$

3.4 Shape Operators

135. Gauss Map. The most important map for differential geometry is the Gauss map. Let us try to describe it.

Given any oriented surface $M \subset \mathbb{R}^3$ (meaning we can choose a specific unit normal vector field, i.e., M is equipped with a chosen unit normal vector field U). There is a canonical map of surfaces ("canonical" meaning we have no need to make arbitrary choices):

$$G\colon M \to S^2, \tag{3.43}$$

where S^2 is the unit sphere in $\mathbb{R}^3$. This is defined by taking U at each point $\boldsymbol{p} \in M$ and translating it to the origin of $\mathbb{R}^3$. More explicitly, if we have $\boldsymbol{n_p} = (\boldsymbol{n})_{\boldsymbol{p}} = U(\boldsymbol{p})$ be the unit normal vector at $\boldsymbol{p}$ with vector part $\boldsymbol{n} \in \mathbb{R}^3$, then $G(\boldsymbol{p}) = \boldsymbol{n}$ is the vector part of $\boldsymbol{n_p}$ *as a point on the unit sphere,* ***not*** *a tangent vector on the unit sphere.* This tells us how the unit normal U rotates as we move $\boldsymbol{p}$ infinitesimally. The intuition is sketched in Figure 3.

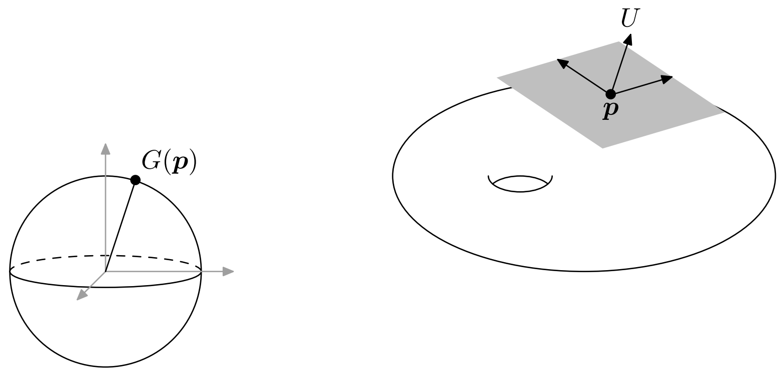


Figure 3: Gauss map for a generic surface. The value of $G(\boldsymbol{p})$ is identical to the vector part of the unit normal vector $U(\boldsymbol{p})$.

136. Example. If M is a plane, then $G(\boldsymbol{p})$ is a constant, since the plane has a constant normal vector (so the Gauss map sends everything in the surface to a single point on the unit sphere).

137. Example. If M is a sphere of radius $r > 0$, then the Gauss map is just a "rescaling" of the sphere together with a translation to the origin.

138. Differential of Gauss Map. For *any* surface $M \subset \mathbb{R}^3$, at any point $\boldsymbol{p} \in M$, the tanget space $\mathrm{T}_{\boldsymbol{p}}M$ and its pushforward along the Gauss map to $\mathrm{T}_{G(\boldsymbol{p})}S^2$, they are parallel. What's more: they are *canonically isomorphic*, which is nicer than "just" isomorphic, because we just have to change basis vectors.

This is a bold claim, so let us prove it more explicitly. Let $\alpha\colon I \to M$ be a map such that $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{v}_{\boldsymbol{p}}$. We find, by definition of the pushforward,

$$G_*(\boldsymbol{v}_{\boldsymbol{p}}) = \frac{\mathrm{d}}{\mathrm{d}t} G\left(\alpha(t)\right)\bigg|_{t=0}. \tag{3.44}$$

But $G\left(\alpha(t)\right)$ "is" $U\left(\alpha(t)\right)$, in the sense that the vector part of $U\left(\alpha(t)\right)$ equals the coordinates of the point $G\left(\alpha(t)\right) \in S^2$. So we have

$$\begin{aligned}
G_*(\boldsymbol{v}_{\boldsymbol{p}}) &= \frac{\mathrm{d}}{\mathrm{d}t} U\left(\alpha(t)\right)\bigg|_{t=0} && \text{(3.45a)}\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i u^i\left(\alpha(t)\right) U_i\right)\bigg|_{t=0} && \text{(3.45b)}\\
&= \sum_i \left(\frac{\mathrm{d}}{\mathrm{d}t} u^i\left(\alpha(t)\right)\bigg|_{t=0} U_i\right) && \text{(3.45c)}\\
&= \sum_i \boldsymbol{v}_{\boldsymbol{p}}[u^i] U_i && \text{(3.45d)}\\
&= \nabla_{\boldsymbol{v}_{\boldsymbol{p}}} U. && \text{(3.45e)}
\end{aligned}$$

139. Definition. Let $M \subset \mathbb{R}^3$ be a surface, $\boldsymbol{p} \in M$ be an arbitrary point. We define a **"Shape Operator"** $S_{\boldsymbol{p}}$ to be a function that maps tangent vectors from $\mathrm{T}_{\boldsymbol{p}}M$ and gives new tangent vectors (i.e., $S_{\boldsymbol{p}}\colon \mathrm{T}_{\boldsymbol{p}}M \to \mathrm{T}_{\boldsymbol{p}}M$) by the following formula:

$$S_{\boldsymbol{p}}(\boldsymbol{v}_{\boldsymbol{p}}) = -\nabla_{\boldsymbol{v}_{\boldsymbol{p}}} U = -G_*(\boldsymbol{v}_{\boldsymbol{p}}) + \text{translation}, \tag{3.46}$$

where U is the unit normal vector field on M and $\boldsymbol{v}_{\boldsymbol{p}} \in \mathrm{T}_{\boldsymbol{p}}M$.

139.1. Remark. We should worry that U is defined only on the surface, so the covariant derivative operation may not be defined everywhere. However, if we demand the vector $\boldsymbol{v_p}$ be on the surface, then we're golden.

139.2. Remark. We see the shape operator is *linear* since the covariant derivative is linear. Also (and this is not obvious) it's symmetric with respect to the dot product [i.e., it's self-adjoint]: $S_{\boldsymbol{p}}(\boldsymbol{w})\cdot\boldsymbol{v} = S_{\boldsymbol{p}}(\boldsymbol{v})\cdot\boldsymbol{w}$.

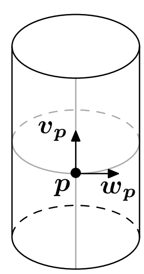
140. Example. Let M be a plane, then $S_{\boldsymbol{p}}(\boldsymbol{v}) = -\nabla_{\boldsymbol{v}}U = 0$ since the normal vector is constant.

141. Example. Let M be a sphere of radius r centered at (x_0, y_0, z_0). We find

$$\begin{aligned}
S_{\boldsymbol{p}}(\boldsymbol{v}) &= -\nabla_{\boldsymbol{v}}U && (3.47a)\\
&= -\nabla_{\boldsymbol{v}}\left(\sum_i \frac{x^i - x^i_0}{r}U_i\right) && (3.47b)\\
&= -\sum_i \boldsymbol{v}\left[\frac{x^i - x^i_0}{r}\right]U_i && (3.47c)\\
&= -\sum_i\sum_j v^j U_j\left[\frac{x^i - x^i_0}{r}\right]U_i && (3.47d)\\
&= -\sum_i\sum_j \frac{v^j}{r}\delta^i{}_j U_i && (3.47e)\\
&= \frac{-1}{r}\sum_i v^i U_i && (3.47f)\\
&= \frac{-\boldsymbol{v}}{r}. && (3.47g)
\end{aligned}$$

Hence $S_{\boldsymbol{p}}(\boldsymbol{v}) = -\boldsymbol{v}/r$ for any tangent vector $\boldsymbol{v}\in \mathrm{T}_{\boldsymbol{p}}M$.

142. Example. We see on the cylinder, pick some frame field as sketched thus:



In the v-direction, it's geometrically "a line"; whereas in the w-direction, it's geometrically "a circle". The reader can verify therefore:

$$\begin{aligned}
S_{\boldsymbol{p}}(\boldsymbol{v_p}) &= 0 && (3.48a)\\
S_{\boldsymbol{p}}(\boldsymbol{w_p}) &= \frac{-1}{r}\boldsymbol{w_p}. && (3.48b)
\end{aligned}$$

143. Shape Operator as Matrix. Let M be a surface in $\mathbb{R}^3$, and suppose we have a frame field E_1, E_2 for M. We can find a matrix representation of the shape operator relative to the frame field in the usual way,

$$\begin{aligned}
S_{\boldsymbol{p}}(E_1) &= s^{1,1}E_1 + s^{1,2}E_2,\\
S_{\boldsymbol{p}}(E_2) &= s^{2,1}E_1 + s^{2,2}E_2.
\end{aligned} \qquad (3.49)$$

Recall the eigenvalues λ_j for a square matrix M contain nearly all possible information about the matrix and further

$$\det(M) = \prod_j \lambda_j \tag{3.50a}$$

$$\operatorname{tr}(M) = \sum_j \lambda_j. \tag{3.50b}$$

For the shape operator, we find the trace and determinant contain crucial geometric information. Specifically, these quantities encode different aspects of the curvature of the surface.

3.5 Curvature of a Surface

144. Proposition. *If $\alpha\colon I \to M$ is any curve on the surface M with $\alpha(0) = \boldsymbol{p}$ and $\alpha'(0) = \boldsymbol{v_p}$, then the normal components of its acceleration is given by $\boldsymbol{v_p} \cdot S_{\boldsymbol{p}}(\boldsymbol{v_p})$.*

Proof. Let M be a surface with unit normal vector field U, let α be an arbitrary curve given by hypothesis. We can restrict U to the curve α, denote it $U_\alpha(t) := U(\alpha(t))$. Now, the velocity vector of α at time $t \in I$ is tangent to M, *not normal to N*. This means $\alpha'(t)$ is orthogonal to the normal vector at $\alpha(t)$. Consequently, we find

$$\alpha'(t) \cdot U_\alpha(t) = 0. \tag{3.51}$$

We can take the time derivative to find

$$\alpha''(t) \cdot U_\alpha(t) + \alpha'(t) \cdot U'_\alpha(t) = 0, \tag{3.52}$$

where

$$U'_\alpha(t) = \nabla_{\alpha'} U = -S_{\alpha'}(\alpha'). \tag{3.53}$$

In particular, we find

$$\alpha'(t) \cdot S_{\alpha(t)}(\alpha'(t)) = \alpha''(t) \cdot U_\alpha(t). \tag{3.54}$$

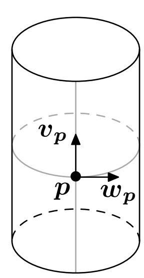
What this tells us is $\alpha''(t) \cdot U_\alpha(t)$ is the component of the acceleration in the direction to the normal of the surface. □

In fact, this motivates the following definition:

145. Definition. Let $M \subset \mathbb{R}^3$ be a surface and $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M$ be a unit vector. We define the **"Normal Curvature"** of M in the direction spanned by $\boldsymbol{v_p}$ is

$$k(\boldsymbol{v_p}) = \boldsymbol{v_p} \cdot S_{\boldsymbol{p}}(\boldsymbol{v_p}). \tag{3.55}$$

146. Example. Recall Example 142 where we worked out the shape operator for a cylinder of radius r. We have two unit vectors $\boldsymbol{v_p}$ and $\boldsymbol{w_p}$ sketched:



We find

$$k(\boldsymbol{v_p}) = \boldsymbol{v_p} \cdot S_{\boldsymbol{p}}(\boldsymbol{v_p}) = 0 \tag{3.56a}$$

$$k(\boldsymbol{w_p}) = \boldsymbol{w_p} \cdot S_{\boldsymbol{p}}(\boldsymbol{w_p}) = \frac{-1}{r}. \tag{3.56b}$$

147. On the Normal "Curvature". If we have our surface $M \subset \mathbb{R}^3$ with unit normal vector field U and, at some point $\boldsymbol{p} \in M$, a [unit] vector $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M$, then we can construct a plane $P \subset \mathbb{R}^3$ spanned by $U(\boldsymbol{p})$ and $\boldsymbol{v_p}$ (so $P = \{c_1 U(\boldsymbol{p}) + c_2\boldsymbol{v_p} \in \mathbb{R}^3 \mid c_1, c_2 \in \mathbb{R}\}$). We obtain a curve given by the intersection of the surface with this plane, $\gamma = P \cap M$. Observe γ passes through $\boldsymbol{p}$. We can further observe the curvature of γ at $\boldsymbol{p}$ is exactly equal to the normal curvature $k(\boldsymbol{v_p})$.

148. Definition. Given a point $\boldsymbol{p} \in M$, we define the **"Principal Curvature"** at $\boldsymbol{p}$ to be:

$$k_1 = \max\{k(\boldsymbol{v_p}) \mid \boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M, \|\boldsymbol{v_p}\| = 1\}, \tag{3.57a}$$

$$k_2 = \min\{k(\boldsymbol{v_p}) \mid \boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}M, \|\boldsymbol{v_p}\| = 1\}. \tag{3.57b}$$

148.1. Remark. If $k_1 \neq k_2$ at a point $\boldsymbol{p} \in M$, then there is a unique vector pointing in the k_1-direction and a unique vector pointing in the k_2-direction. Furthermore, these vectors are orthogonal.

149. Proposition. *Let M be a surface. If $k_1 \neq k_2$ at $\boldsymbol{p} \in M$, and $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathrm{T}_{\boldsymbol{p}}M$ are unit vectors such that $k(\boldsymbol{e}_1) = k_1$ and $k(\boldsymbol{e}_2) = k_2$, then $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are unique up to sign, and $\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = 0$ and $S_{\boldsymbol{p}}(\boldsymbol{e}_1) = k_1\boldsymbol{e}_1$ and $S_{\boldsymbol{p}}(\boldsymbol{e}_2) = k_2\boldsymbol{e}_2$. We call the directions given by $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ the **"Principal Directions"** tangent to $\boldsymbol{p}$.*

150. Why is this so great? We can rewrite the shape operator. The recurring theme of the course is that we may choose a frame that is particularly nice for the situation. We have a nice basis of eigenvectors. This is great because we have the operator (which these are eigenvectors of) be diagonal,

$$S_{\boldsymbol{p}}(\alpha\boldsymbol{e}_1 + \beta\boldsymbol{e}_2) = \alpha k_1\boldsymbol{e}_1 + \beta k_2\boldsymbol{e}_2. \tag{3.58}$$

So with respect to this basis $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ (which are the principal directions) the shape operator acts by

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} \mapsto \begin{bmatrix}k_1\alpha\\ k_2\beta\end{bmatrix} = \begin{bmatrix}k_1 & 0\\ 0 & k_2\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}. \tag{3.59}$$

This leads to a much more beautiful notion of curvature. We'd like to get "just a number" at each point (or something like that) for describing curvature of a surface. We have:

151. Definition. Let M be a surface. We define the **"Gaussian Curvature"** $K = \det(S)$ as the determinant of the shape operator, and the **"Mean Curvature"** $H = \operatorname{tr}(S)/2$ as the trace divided by the number of dimensions of M.

151.1. Remark. The Gaussian curvature is an intrinsic geometric quantity. Often finding the principal curvatures is hard, but computing the Gaussian curvature can be easier.

151.2. Remark. The mean curvature is an extrinsic geometric quantity.

152. Examples. Here are a few examples of the Gaussian and mean curvatures:

1. Cylinder: $K = (0)\cdot(-1/r) = 0$, $H = -1/(2r)$
2. Sphere: $K = 1/r^2$, $H = -1/r$
3. Plane: $K = 0$, $H = 0$.

153. Definition. Let M be a surface.

1. We call M **"Flat"** if its Gaussian curvature is zero, $K = 0$.
2. We call M **"Minimal"** (or *minimum*) if its mean curvature is zero, $H = 0$.

Exercises

1. Compute the Gaussian and mean curvature for:
   (a) the saddle $z = x^2 - y^2$,
   (b) the monkey saddle: $z = x^3 - 3xy$.
2. Let $f\colon \Sigma \to \mathbb{R}$ be a real-valued function on the surface Σ, let $g\colon \mathbb{R} \to \mathbb{R}$ be some smooth function.
   (a) Prove or find a counter-example: for any $\boldsymbol{v_p} \in \mathrm{T}_{\boldsymbol{p}}\Sigma$, we have $\boldsymbol{v_p}[g \circ f] = g'(f(\boldsymbol{p}))\boldsymbol{v_p}[f]$
   (b) Deduce $\mathrm{d}(g \circ f) = (g' \circ f)\,\mathrm{d}f$.
3. Let T^2 be the familiar torus.
   (a) What are the image curves under the Gauss map for meridians and parallels of T^2?
   (b) Are there any points $\boldsymbol{q} \in G(T^2)$ for which exactly two distinct points on the torus $\boldsymbol{p}_1 \in T^2$ and $\boldsymbol{p}_2 \in T^2$ are mapped to $G(\boldsymbol{p}_1) = G(\boldsymbol{p}_2) = \boldsymbol{q}$?
4. Consider the surface Σ defined by $z = xy$.
   (a) What are the image curves under the Gauss map for $x = \text{constant}$ on Σ?
   (b) Are there any points $\boldsymbol{q} \in G(\Sigma)$ for which exactly two distinct points on the surface $\boldsymbol{p}_1 \in \Sigma$ and $\boldsymbol{p}_2 \in \Sigma$ (so $\boldsymbol{p}_1 \neq \boldsymbol{p}_2$) which are mapped to $G(\boldsymbol{p}_1) = G(\boldsymbol{p}_2) = \boldsymbol{q}$?
5. For each of the following surfaces, find the quadratic approximation near the origin:
   (a) $z = x^2 + y^2$
   (b) $z = x^2 - y^2$
   (c) $x^2 + y^2 - z^2 = 0$.
   (d) $z = (2x + 3y)^5$.
6. Recall from linear algebra, if A is any $n \times n$ matrix, the **"Characteristic Polynomial"** of A is the polynomial in λ defined by
   $$p(\lambda) = \det(A - \lambda \cdot I_n),$$
   where I_n is the $n \times n$ identity matrix. **Compute** the characteristic polynomial for the shape operator.

4 References

Great books to consult while reading these notes include:

[1] Barrett O'Neill, *Elementary Differential Geometry*. Revised Second edition, Academic Press, 2006. (**Caution**: O'Neill's notation can be odd at times, like writing the transpose of a matrix as ${}^tM$ instead of M^T or $M^\top$.)

[2] Manfredo do Carmo, *Differential Geometry of Curves and Surfaces*. Revised and Updated Second Edition, Dover, 2016.

The "next steps" would be to read the wonderful books:

[3] Michael Spivak, *Calculus on Manifolds: A Modern Approach To Classical Theorems Of Advanced Calculus*. CRC Press, 1971.

[4] John Milnor, *Topology from the Differentiable Viewpoint*. Princeton University Press, 1997.