Introduction to Differential Geometry

Walker H. Stern

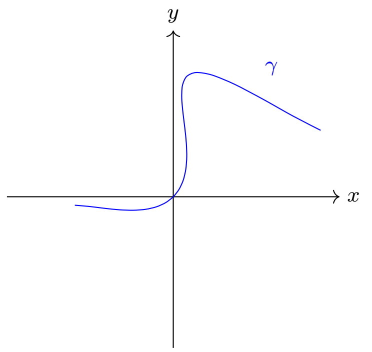
These notes were prepared for a lecture course held during fall semester of 2022 at the University of Virginia. In these notes, I have freely and heavily used the references cited, and the notes should not be considered original. The figures were mostly made in Asymptote with some made in Mathematica.

Contents

Overture: What is curvature?

Throughout the coming semester, we will repeatedly return to the concept of *curvature*. At a purely heuristic level, curvature should be a measure of "how much a geometric object bends in a given direction." Given a curve

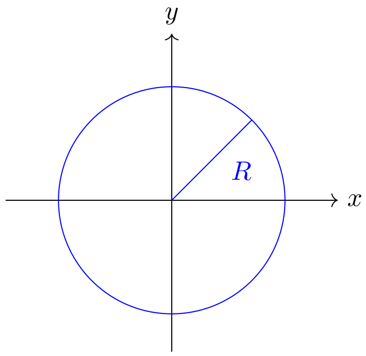
$$\gamma\colon\ [a,b] \longrightarrow \mathbb{R}^2$$



we want to be able to assign a number $\kappa(t)$ to each point $\gamma(t)$ on the curve such that

1. When $|\kappa(t)|$ is larger, the curve γ is "more bendy" at the point $\gamma(t)$.
2. When $\kappa(t) = 0$, the curve is a straight line in a small neighborhood of $\gamma(t)$.

To try and make sense of this, lets consider a *very* special curve: the circle of radius R about the origin in $\mathbb{R}^2$.



We can immediately notice that $\frac{1}{R}$ could be precisely the number we are looking for. When R is very small, the circle bends very quickly, and $\frac{1}{R}$ gets higher and higher. When R goes to ∞, a little arc of the circle starts to look a lot like a straight line, and $\frac{1}{R}$ goes to 0.

Since we now have a notion of curvature for a circle, we can try to generalize this to an arbitrary curve γ. When discussing the *slope* of an arbitrary curve, we define the *tangent line* — the line which best approximates the curve — because we have a well-defined notion of slope for lines. In a similar fashion, we will try to define the curvature of γ at a point $t \in [a,b]$ by finding the circle which best approximates γ at the point $\gamma(t_0)$.[1]

[1] Such a circle is called the *osculating circle* for γ at t_0, from the latin word for "kissing".

Our starting data is thus a curve

$$\gamma\colon\ [a,b] \longrightarrow \mathbb{R}^2$$

and we wish to find a parameterized circle

$$\begin{aligned} \psi\colon [0,\ell] &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (R\cos(f(x)) + c_1, R\sin(f(x)) + c_2) \end{aligned}$$

with a value $x_0 \in [0,\ell]$ such that $\psi(x_0) = \gamma(t_0)$, and such that as many derivatives as possible of ψ and γ agree at x_0 and t_0, respectively.

To make this a little more tractable, we will make some SIMPLIFYING ASSUMPTIONS:

1. We will assume that both $\psi'(x)$ and $\gamma'(t)$ have length 1 for every value $x \in [0,\ell]$ or $t \in [a,b]$, respectively.[2] That is,

$$|\gamma'(t)| = 1 = |\psi'(x)|$$

2. The second derivatives $\psi''(x_0)$ and $\gamma''(t_0)$ are non-zero.[3]

These first two assumptions have some immediate consequences. For example, we then have the inner product[4]

$$\langle\gamma'(t), \gamma'(t)\rangle|\gamma'(t)|^2 = 1$$

taking a derivative yields

$$2\langle\gamma''(t), \gamma'(t)\rangle = 0$$

so that $\gamma'(t)$ and $\gamma''(t)$ are orthogonal. By our second assumption, this means that $\{\gamma'(t_0), \gamma''(t_0)\}$ forms an orthonormal basis of $\mathbb{R}^2$.

While similar facts hold for ψ, we can actually use our assumptions to give an explicit form for ψ. Our assumptions, plus the parameterization of ψ, imply that

$$|\psi'(x)| = R|f'(x)| = 1$$

so that

$$f'(x) = \pm\frac{1}{R}.$$

This in turn means that, assuming $f(0) = 0$

$$f(x) = \epsilon\frac{x}{R}.$$

for $\epsilon \in \{+1, -1\}$.

We will also make one final simplifying assumption, namely

3. At the value t_0, we have[5]

$$\begin{aligned} \gamma(t_0) &= (0,0) \\ \gamma'(t_0) &= (-1,0). \end{aligned}$$

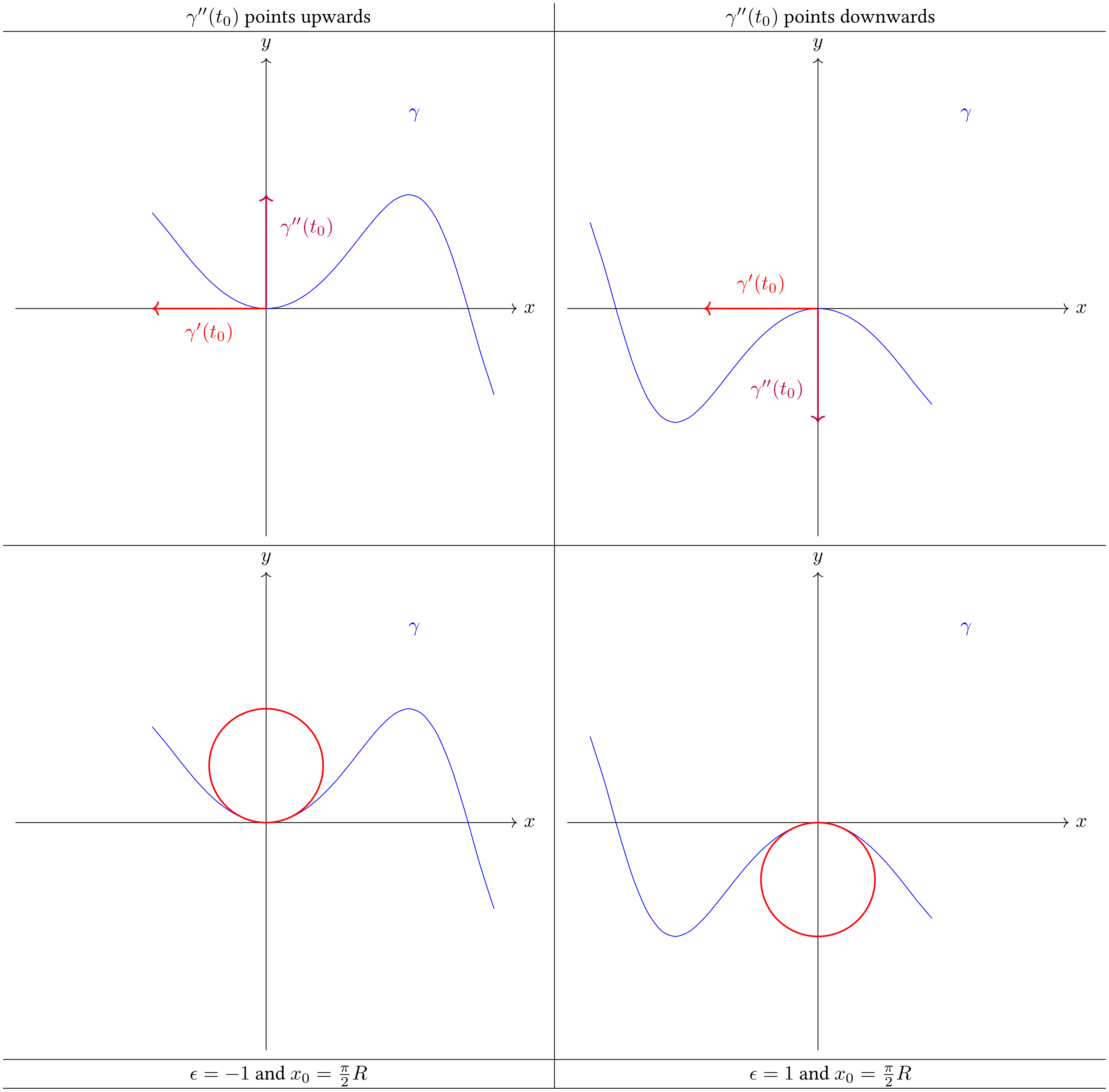
This tells us we have one of two scenarios:

[2] It will turn out that, for well-behaved curves, we can always reparameterize our curves so that this is the case. We will not show this yet, however.

[3] This, as it turns out, excludes the case where the curvature is 0. This is because I don't currently want to discuss "circles of infinite radius".

[4] Also called the dot product.

[5] This is also justified in general, as we can translate and rotate the curve γ without changing the curvature at a given point.


$\gamma''(t_0)$ points upwards
$\gamma''(t_0)$ points downwards
y
x
γ
$\gamma''(t_0)$
$\gamma'(t_0)$
$\epsilon = -1$ and $x_0 = \frac{\pi}{2}R$
$\epsilon = 1$ and $x_0 = \frac{\pi}{2}R$

More formally, we can set

$$\gamma^{(k)}(t_0) = \psi^{(k)}(x_0)$$

for $k = 0, 1, 2$, yielding

$$\left(R\cos\left(\epsilon\frac{x_0}{R}\right) + c_1, R\sin\left(\epsilon\frac{x_0}{R}\right) + c_2\right) = (0,0) \tag{1}$$

$$\left(-\epsilon\sin\left(\epsilon\frac{x_0}{R}\right), \epsilon\cos\left(\epsilon\frac{x_0}{R}\right)\right) = (-1,0) \tag{2}$$

$$\left(-\frac{1}{R}\cos\left(\epsilon\frac{x_0}{R}\right), -\frac{1}{R}\sin\left(\epsilon\frac{x_0}{R}\right)\right) = \gamma''(t_0). \tag{3}$$

If $\epsilon = +1$, then $\{\psi'(x_0), \psi''(x_0)\}$ is a positively oriented basis of $\mathbb{R}^2$, and so equations (2) and (3) imply that $\gamma''(t_0)$ points downwards. On the other hand, if $\epsilon = -1$, then $\{\psi'(x_0), \psi''(x_0)\}$ is a negatively oriented basis, and so $\gamma''(t_0)$ must be pointing upwards. In either case, we see that $x_0 = \frac{\pi R}{2}$.

Taking the norm of equation (3), we then see that

$$\frac{1}{R} = |\gamma''(t_0)|,$$

so that

$$R = \frac{1}{|\gamma''(t_0)|}$$

Moreover, for equation (1) to be satisfied, we must have $(c_1, c_2) = (0, -\epsilon R)$.

We thus see that, setting $\epsilon = \det(\gamma'(t_0), \gamma''(t_0))$ the circle

$$\psi(x) = \left(\frac{1}{|\gamma''(t_0)|}\cos(\epsilon x|\gamma''(t_0)|), \frac{1}{|\gamma''(t_0)|}\left(\sin(\epsilon x|\gamma''(t_0)|) - \epsilon\right)\right)$$

is the unique circle which agrees with γ to second order at t_0. We can thus make the following

Definition 0.1. Let $\gamma : [a,b] \to \mathbb{R}^2$ be a plane curve with $|\gamma'(t)| = 1$ for all $t \in [a,b]$.[6] The *(unsigned) curvature* of a plane curve $\gamma(t)$ at $t_0 \in [a,b]$ is

$$\kappa(t_0) := |\gamma''(t_0)|.$$

[6] If we do not make this assumption, it is substantially more subtle to define the curvature. Fortunately, we can always reparameterize γ so that this condition is satisfied.

1

Curves

We will begin our study of geometry with *curves*, by which we will roughly mean *1-dimensional subsets of* $\mathbb{R}^n$. There are two reasonable perspectives from which we might formally define such subsets – by *parameterizations*, and as *subsets of* $\mathbb{R}^n$. It turns out that — with the appropriate definition of dimension, and under sufficient regularity/smoothness conditions — these two perspectives are equivalent, but for convenience, we will focus on the former.

Let us illustrate, in two dimensions, why we might want to impose the condition of *regularity* on our curves. It is fairly clear why we need smoothness: If we want to use techniques from calculus, we need to have access to derivatives, and for simplicity, we often simply make the assumption that γ has derivatives of all orders.

Regularity, on the other hand, seems less obvious. For instance, we can define a function

$$\begin{aligned} \beta : [0,2] &\longrightarrow [0,2\pi] \\ t &\longmapsto \begin{cases} 2\pi \exp\left(\frac{1}{1-t^2}\right) & t \in [0,1) \\ 0 & t \in [1,2] \end{cases} \end{aligned}$$

which is smooth, surjective, and non-increasing. The curve

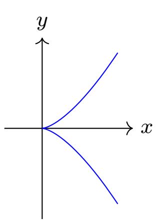
$$\begin{aligned} \gamma : [0,2] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (\cos(\beta(t)), \sin(\beta(t))) \end{aligned}$$

thus describes the unit circle in $\mathbb{R}^2$, but for any $t \in [1,2]$, we have $\gamma'(t) = (0,0)$. So this parameterization is not regular, even though the circle clearly admits a regular parameterization.

The reason to require regularity is that there are curves for which *no* parameterization will be regular. For example, consider the curve

$$\begin{aligned} \gamma : [-1,1] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t^2, t^3) \end{aligned}$$

whose graph looks like



It is an instructive exercise to convince yourself that any smooth parameterization of this curve will not be regular, because of the "fold" at the point $(0,0)$

Definition 1.1. A *parameterized curve* in $\mathbb{R}^n$ is a function

$$\gamma\colon\ [a,b] \longrightarrow \mathbb{R}^n$$

from some closed interval to $\mathbb{R}^n$. We will additionally call γ

- A C^k *(parameterized) curve* if γ has continuous derivatives of all orders $0 \le r \le k$ on $[a,b]$.
- A *smoothly (parameterized) curve* if γ has continuous derivatives of all orders on $[a,b]$. This is also called a C^∞ curve.
- A *regular parameterization* if the first derivative $\gamma'(t)$ is non-zero for all $t \in [a,b]$ This is equivalent to requiring that the norm $|\gamma'(t)|$ is non-zero for all $t \in [a,b]$.

In this chapter, we will *always* take the word *curve* to mean *smoothly parameterized curve* unless specified otherwise. We will consider two parameterized curves to be "the same" when they are related by a change of parameterization.

Definition 1.2. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a smooth curve in $\mathbb{R}^n$. A *change of parameter* is a C^∞ bijection $\phi : [c,d] \to [a,b]$, the inverse of which is also C^∞. The curve $\gamma \circ \phi : [c,d] \to \mathbb{R}^n$ is then called a *reparameterization* of γ.

We call a change of parameter ϕ *orientation-preserving* if $\phi'(t) > 0$ for all $t \in [c,d]$.

Since we are studying geometry, we will be exploring notions of distance, angle, and curvature — these are in some sense the properties which characterize geometry as a subject. We will start by giving a definition of the *length* of a curve or curve segment.

Definition 1.3. Let $\gamma : [a,b] \to \mathbb{R}^2$ be a C^1 parameterized curve. For any $t_0, t_1 \in [a,b]$ with $t_0 < t_1$, we define the *arc length* of γ from t_0 to t_1 to be the integral

$$L(\gamma; t_0, t_1) = \int_{t_0}^{t_1} |\gamma'(t)| dt.$$

We also define the arc length as a function of the parameter t

$$s(t) := L(\gamma; a, t) = \int_a^t |\gamma'(u)| du.$$

The first thing we need to check is that this is well-defined, and does not depend on our choice of parameterization.

Lemma 1.4. *Let $\gamma : [a,b] \to \mathbb{R}^n$ be a C^1 curve, and let $\phi : [c,d] \to [a,b]$ be an orientation-preserving change of parameter. Then*

$$L(\gamma \circ \phi; c, d) = L(\gamma; a, b).$$

Proof. We simply write the definition, and manipulate the resulting integral.

$$\begin{aligned} L(\gamma \circ \phi; c, d) &= \int_c^d \left| \frac{d}{dt}(\gamma \circ \phi)(u) \right| d_u \\ &= \int_c^d |\gamma'(\phi(u))\phi'(u)| \, du \end{aligned}$$

Since ϕ is orientation preserving, $|\phi'(u)| = \phi'(u)$. We can thus change variables, using $v = \phi(u)$, $dv = \phi' du$, to obtain

$$\begin{aligned} L(\gamma \circ \phi; c, d) &= \int_c^d |\gamma'(\phi(u))| \, \phi'(u) du \\ &= \int_a^b |\gamma'(v)| dv \\ &= L(\gamma; a, b). \end{aligned}$$

Which yields the desired equality. □

Example 1.5. Let us consider the example of a helix in $\mathbb{R}^3$. This is the curve

$$\begin{aligned} \gamma \colon [0, 2\pi] &\longrightarrow \mathbb{R}^3 \\ t &\longmapsto (R\cos(t), R\sin(t), bt) \end{aligned}$$

where R and b are positive real parameters.

We can compute the arc length of the helix γ as follows. The first derivative is

$$\gamma'(t) = (-R\sin(t), R\cos(t), b)$$

so

$$|\gamma'(t)| = \sqrt{R^2 + b^2}.$$

Thus

$$s(t) = \int_0^t \sqrt{R^2 + b^2} du = t\sqrt{R^2 + b^2}$$

is the arc length of γ as a function of the parameter t.

Definition. A *polygonal path* in $\mathbb{R}^n$ consists of an ordered tuple $\mathbf{x} := (x^1, \ldots, x^k)$ of points in $\mathbb{R}^n$. We view this as a (continuous) path in $\mathbb{R}^n$ via the map

$$\psi_{\mathbf{x}} : \longrightarrow \mathbb{R}^n$$

which is defined, for $t \in [\ell, \ell+1]$, by

$$\psi_{\mathbf{x}}(t) = (1 - t + \ell)x^\ell + (t - \ell)x^{\ell+1}.$$

The *length* of a polygonal path $\mathbf{x} = (x^1, \ldots, x^k)$ is

$$L(\psi_{\mathbf{x}}) = \sum_{\ell=0}^{k-1} |x^{\ell+1} - x^\ell|$$

Construction. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a continuous curve in $\mathbb{R}^n$. Given $a = t_0 < t_1 < \cdots < t_k = b$, we obtain a polygonal path

$$(\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_k))$$

in $\mathbb{R}^n$, we can view as an approximation of γ. We call such a path a *polygonal approximation* to γ.

Challenge Problem. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a regular C^2 parameterized curve. Show that the arc length is the supremum of the lengths of polygonal approximations to γ, taken over all polygonal approximations to γ. Use this to give a definition of arc length for C^1 curves. Show that your definition yields a well-defined value for any C^1 curve.(Hint: show that the set of lengths of polygonal approximations is bounded above by $(\sup(|\gamma'(t)||b-a|)))$

There is a very special kind of parameterized curve $\gamma : [a,b] \to \mathbb{R}^n$, where the parameter t is precisely the arc length of γ from a to t. We encountered this condition in disguise in the the overture. A key trick we will make use of is to reparameterize curves so that they are parameterized by arc length.

Definition 1.6. We say that a curve $\gamma : [a,b] \to \mathbb{R}^n$ is *parameterized by arc length* if $|\gamma'(t)| = 1$. In this case, the arc length of γ is

$$s(t) = \int_a^t 1 du = t.$$

That is, the parameter t of γ is precisely the arc length.

Lemma 1.7. *Let $\gamma : [a,b] \to \mathbb{R}^n$ be a regular curve.*[1] *Then there is an orientation preserving change of parameter ϕ such that $\gamma \circ \phi$ is parameterized by arc length.*[2]

[1] Remember that we are taking the word "curve" to mean "smooth curve".
[2] Technically, we can weaken the hypotheses here. If we allow our reparameterization ϕ to just be C^1 with C^1 inverse, then the lemma holds whenever γ is regular and C^1.

Proof. The arc length of γ is

$$s(t) = \int_a^t |\gamma'(u)| du$$

and so, by the fundamental theorem of calculus,

$$\frac{ds}{dt} = |\gamma'(t)|.$$

As a consequence, s is a C^∞ monotone increasing function of t, and has non-zero derivative everywhere. Since s is monotone increasing, it has a global inverse $t(s)$. Since it has non-zero derivative, we can apply the inverse function theorem B.1 to see that $t(s)$ is itself C^∞ with

$$\frac{dt}{ds} = \frac{1}{|\gamma'(t(s))|}.$$

We then claim that $\gamma(t(s))$ is parameterized by arc length. To see this, note that

$$\left| \frac{d}{ds}\gamma(t(s)) \right| = \left| \gamma'(t(s)) \frac{dt}{ds} \right| = |\gamma'(t(s))| \frac{1}{|\gamma'(t(s))|} = 1$$

as desired. ☐

1 *Curvature in the plane revisited*

We now return to the study of plane curves, $\gamma : [a,b] \to \mathbb{R}^2$, and their curvature. While we have already discussed this in the overture, we will proceed more systematically here, in a way that will lay the groundwork for our later study of curves in $\mathbb{R}^n$.

ASSUMPTION: Throughout this section, a *curve* will mean a *regular, smooth, plane curve.*

Given a curve $\gamma : [a,b] \to \mathbb{R}^2$, notice that $\gamma'(t)$ is a tangent vector to $\gamma(t)$ for any $t \in [a,b]$. When we were first considering curvature, we used a curve γ which was parameterized by arc length, in which case the vector $\frac{d\gamma}{ds}$ is a *unit* tangent vector. Our contemplation of curvature split into two cases, depending on whether the basis

$$\left\{ \frac{d\gamma}{ds}, \frac{d^2\gamma}{ds^2} \right\}$$

was positively oriented or not.

To try and avoid case-by-case arguments, we will try to define a positively oriented unit basis (for each $t \in [a, b]$) to which we can compare the derivatives of γ.

Definition 1.8. Let $\gamma : [a, b] \to \mathbb{R}^2$ be a curve. A *(smooth) vector field* along γ is a smooth function

$$X\colon\ [a,b] \longrightarrow \mathbb{R}^2$$

where, for each $t \in [a, b]$, we consider $X(t)$ as a vector starting at $\gamma(t)$.[3] We call a vector field X on γ a *tangent vector field* if $X(t)$ is a tangent vector to γ at $\gamma(t)$ for every $t \in [a, b]$.

[3] There is a much better way to formalize this, using *tangent spaces* and *tangent bundles*. We will eventually switch to this formalism, but for the time being, we take the naïve perspective.

Example 1.9. We can view $\gamma'(t)$ and $\gamma''(t)$ as vector fields on γ. In this case, $\gamma'(t)$ is a tangent vector field on γ.

Definition 1.10. Let $\gamma : [a, b] \to \mathbb{R}^2$ be a curve.

- A *moving* 2*-frame* along γ consists of a pair of vector fields along γ which, for each $t \in [a, b]$, form an orthonormal basis of $\mathbb{R}^2$. More precisely, a moving 2-frame consists of two vector fields

$$e_1, e_2\colon\ [a,b] \longrightarrow \mathbb{R}^2$$

  along γ such that, for every $t \in [a, b]$,[4]

$$e_i(t) \cdot e_j(t) = \delta_{i,j}.$$

[4] Here, $\delta_{i,j}$ is simply the Kroenecker delta, i.e.

$$\delta_{i,j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

- A Frenet 2-frame for γ is a moving 2-frame along γ such that

  1. For every $t \in [a, b]$, $e_1(t)$ is a positive scalar multiple of $\gamma'(t)$.
  2. For every $t \in [a, b]$, the basis $\{e_1(t), e_2(t)\}$ is positively oriented.

Conveniently, there is only one Frenet 2-frame. Condition (1) implies that

$$e_1(t) = \frac{1}{|\gamma'(t)|}\gamma'(t).$$

Since γ is C^∞ and $|\gamma'|$ is non-zero, it follows that $e_1(t)$ is C^∞. Condition (2), together with orthonormality, requires that

$$e_2(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} e_1(t).$$

We can thus refer to *the* Frenet 2-frame.

In 2 and 3 dimensions, there is a standard notation for the vectors $e_1(t)$ and $e_2(t)$. One conventionally writes $\mathbf{T}(t) := e_1(t)$ for the *unit tangent vector* and $\mathbf{N}(t) = e_2(t)$ for the *unit normal* vector.

Lemma 1.11. *Let γ be a curve, and let $\{e_1(t), e_2(t)\}$ be the unique Frenet 2-frame. Then the* Frenet equations

$$\gamma'(t) = |\gamma'(t)|e_1(t) \tag{1.1}$$
$$e_1'(t) = \omega(t)e_2(t) \tag{1.2}$$
$$e_2'(t) = -\omega(t)e_1(t) \tag{1.3}$$

hold, where

$$\omega(t) = e_1'(t) \cdot e_2(t).$$

Proof. The Frenet equation (1.1) is the definition of $e_1(t)$. To obtain equations (1.2) and (1.3), we differentiate the rothonormality relation

$$\frac{d}{dt}\left(e_i(t) \cdot e_j(t) = \delta_{i,j}\right)$$

to obtain

$$\begin{aligned} e_1'(t) \cdot e_2 &= -e_2'(t) \cdot e_1 \\ e_1'(t) \cdot e_1 &= 0 \\ e_2'(t) \cdot e_2 &= 0. \end{aligned}$$

Equations (1.2) and (1.3) then simply amount to expanding $e_1'(t)$ and $e_2'(t)$ in the basis $\{e_1(t), e_2(t)\}$. □

We can rewrite equations (1.2) and (1.3) in a more compact form. Letting $e(t)$ be the matrix whose rows are $e_1(t)$ and $e_2(t)$, we can rewrite these equations as

$$e'(t) = \begin{pmatrix} 0 & \omega(t) \\ -\omega(t) & 0 \end{pmatrix} e(t)$$

In the special case where γ is parameterized by the arc length s, we find something interesting. The Frenet equations simplify to

$$\begin{aligned} \gamma'(s) &= e_1(s) \\ e'(s) &= \begin{pmatrix} 0 & \omega(s) \\ -\omega(s) & 0 \end{pmatrix} e(s) \end{aligned}$$

We can then express the curvature we defined in the overture in terms of the quantity ω

$$\kappa(s) := \left|\frac{d^2\gamma}{ds^2}\right| = |e_1'(s)| = |\omega(s)|$$

we might be tempted to define the (signed) curvature of a general curve to be the real number $\omega(s)$. However, as the example in the sidebar shows, this is not independent of the choice of parameterization. The next lemma gives us a notion which *is* independent of reparameterization.

Lemma 1.12. *Let $\gamma : [a,b] \to \mathbb{R}^2$ be a curve, and let $\phi : [c,d] \to [a,b]$ be an orientation-preserving change of parameter. Let $e(t)$ be the Frenet frame for γ, and let $\widetilde{e}(u)$ be the Frenet frame for $\gamma \circ \phi$, and similarly for ω and $\widetilde{\omega}$. Then*

$$\frac{\widetilde{\omega}(u)}{\left|\frac{d}{du}\gamma(\phi(u))\right|} = \frac{\omega(\phi(u))}{\gamma'(\phi(u))}.$$

Example. Let us consider the following two parameterizations of the unit circle

$$\psi(t) = (\cos(t), \sin(t)) \qquad t \in [0, 2\pi]$$

and

$$\phi(u) = (\cos(2u), \sin(2u)) \qquad t \in [0, \pi].$$

We first compute the Frenet 2-frame associated to each parameterization.

For ψ: Note

$$\psi'(t) = (-\sin(t), \cos(t))$$

is a unit vector, so $e_1(t) = \psi'(t)$. Then

$$e_2(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} e_1(t) = (-\cos(t), -\sin(t)).$$

For ϕ: The norm of

$$\phi'(u) = (-2\sin(2u), 2\cos(2u))$$

is 2. So $\widetilde{e}_1(u) = (-\sin(2u), \cos(2u))$. As above, we find

$$\widetilde{e}_2(u) = (-\cos(2u), -\sin(2u))$$

We can then write the Frenet equations for each curve.

For ψ:

$$\begin{aligned} \psi'(t) &= e_1(t) \\ e'(t) &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} e(t) \end{aligned}$$

For ϕ:

$$\begin{aligned} \psi'(u) &= \widetilde{e}_1(u) \\ \widetilde{e}'(u) &= \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \widetilde{e}(u) \end{aligned}$$

Proof. We first note that $\tilde{e}_i(u) = e_i(\phi(u))$ for $i = 1, 2$. Thus $\tilde{e}_i'(u) = e_i'(\phi(u))\phi'(u)$. We then directly compute

$$\begin{aligned}
\frac{\widetilde{\omega}(u)}{\left|\frac{d}{du}\gamma(\phi(u))\right|} &= \frac{\tilde{e}_1'(u)\cdot \tilde{e}_2(u)}{\left|\frac{d}{du}\gamma(\phi(u))\right|} \\
&= \frac{\phi'(u)e_1'(\phi(u))\cdot e_2(\phi(u))}{|\gamma\prime(\phi(u))\phi'(u)|} \\
&= \frac{e_1'(\phi(u))\cdot e_2(\phi(u))}{|\gamma\prime(\phi(u))|} \\
&= \frac{\omega(\phi(u))}{|\gamma\prime(\phi(u))|}
\end{aligned}$$

as desired. □

Definition 1.13. Given a curve $\gamma : [a, b] \to \mathbb{R}^2$, the *signed curvature* of γ at $t \in [a, b]$ is

$$\kappa(t) := \frac{\omega(t)}{|\gamma'(t)|}.$$

The *unsigned curvature* is the absolute value of $\kappa(t)$.

This accords fully with our previous definition: in the case when γ is parameterized by arc length, $|\gamma'(s)| = 1$, so $\kappa(s) = \omega(s)$. This means that the signed curvature is precisely the scalar $\kappa(s)$ such that

$$\frac{d^2\gamma}{ds} = \kappa(s)e_2(s).$$

Even in the case where γ is not parameterized by arc length, there is a relatively easy way to compute the curvature of γ, which does not require us to reparameterize the curve by arc length, nor to explicitly compute the Frenet 2-frame.

Lemma 1.14. *Let $\gamma : [a, b] \to \mathbb{R}^2$ be a curve. Then*

$$\kappa(t) = \frac{\det(\gamma'(t), \gamma''(t))}{|\gamma'(t)|^3}.$$

Proof. As usual, we write $s(t)$ for the arc length of γ as a function of the parameter $t \in [a, b]$, and $t(s)$ for the inverse function. Then

$$e_1(s) = \frac{d}{ds}\gamma(t(s)) = \gamma'(t)\frac{dt}{ds}.$$

Taking an s-derivative, we find

$$e_1'(s) = \gamma'(t)\frac{d^2t}{ds^2} + \gamma''(t)\left(\frac{ds}{dt}\right)^2.$$

Applying the function $\det(\gamma'(t), -)$ to our above equation, we obtain

$$\det(\gamma'(t), e_1'(s)) = 0 + \det(\gamma'(t), \gamma''(t))\left(\frac{ds}{dt}\right)^2.$$

We then can note that $\gamma'(t) = |\gamma'(t)|e_1(t)$ and $e_1'(s) = \kappa(s)e_2(s)$. Thus, the left-hand side becomes

$$\det(\gamma'(t), e_1'(s)) = |\gamma'(t)|\kappa(s(t)) \det(e_1(t), e_2(t)) = |\gamma'(t)|\kappa(s(t)).$$

Moreover, by definition

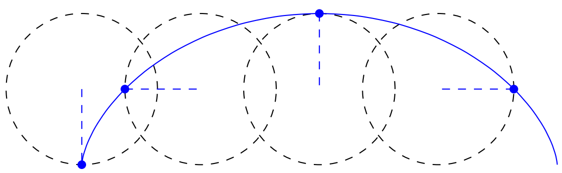
$$\frac{ds}{dt} = \frac{1}{|\gamma'(t)|}$$

so that our equation simplifies to

$$\kappa(t)|\gamma'(t)| = \frac{\det(\gamma'(t), \gamma''(t))}{|\gamma'(t)|^2}.$$

This completes the proof. $\square$

Example 1.15. Consider the *cycloid*, the curve created a point on a rolling unit circle



One can derive a parameterization for this curve:

$$\gamma(t) = (t - \sin(t), 1 - \cos(t)) \qquad t \in [0, 2\pi]$$

which is **not** regular when $t = 0$ or $t = 2\pi$. In the realm where it is regular, however, we may attempt to compute its curvature using Lemma 1.14.

We first compute the derivatives

$$\gamma'(t) = (1 - \cos(t), \sin(t))$$
$$\gamma''(t) = (\sin(t), \cos(t))$$

so that we find

$$\begin{aligned}
\det(\gamma'(t), \gamma''(t)) &= \cos(t) - \cos^2(t) - \sin^2(t) \\
&= \cos(t) - 1 \\
|\gamma'(t)|^2 &= 1 - 2\cos(t) + \cos^2(t) + \sin^2(t) \\
&= 2 - 2\cos(t)
\end{aligned}$$

We thus find that the curvature of the cycloid is given by

$$\kappa(t) = \frac{\cos(t) - 1}{(2 - 2\cos(t))^{\frac{3}{2}}}.$$

With som trigonometric identities, we can simplify this to

$$\kappa(t) = \frac{-2\sin^2\left(\frac{t}{2}\right)}{\left(4\sin^2(\frac{t}{2})\right)^{\frac{3}{2}}} = \frac{-2\sin^2\left(\frac{t}{2}\right)}{8\sin^3\left(\frac{t}{2}\right)} = \frac{-1}{4\sin(\frac{t}{2})}.$$

Notice that this becomes undefined precisely when $t = 0$ or $t = 2\pi$, i.e., the points at which the curve is not regular.

Example. We could also try to compute the curvature of the cycloid using an arc length parameterization. The arc length of the cycloid is

$$\begin{aligned}
s(t) &= \int_0^t \sqrt{2 - 2\cos(u)}du \\
&= \int_0^t 2\sin\left(\frac{u}{2}\right) du \\
&= 4\int_0^{t/2} \sin(v)dv \\
&= -4\cos(v)|_{v=0}^{t/2} \\
&= 4(1 - \cos(t/2)) \\
&= 8\sin^2\left(\frac{t}{4}\right)
\end{aligned}$$

The inverse function $t(s)$ is thus

$$t(s) = 4\arcsin\left(\sqrt{\frac{s}{8}}\right)$$

This then requires us to take two derivatives of the function

$$\begin{aligned}
\gamma(t(s)) = \Bigg(& 8\arcsin\left(\sqrt{\frac{s}{8}}\right) - \sin\left(8\arcsin\left(\sqrt{\frac{s}{8}}\right)\right), \\
& 1 - \cos\left(8\arcsin\left(\sqrt{\frac{s}{8}}\right)\right)\Bigg)
\end{aligned}$$

then take the norm (and possibly reparameterize it until it is again in terms of t). This yields the same result (one can check with, e.g., Mathematica) but the computation is gross, and it is very easy to make a mistake.

We round off our discussion of curves in the plane by showing something remarkable: If we know the interval that a curve γ is defined on and the curvature of that curve at all points, then we know the curve (up to Euclidean isometry).

Proposition 1.16. *Let $\gamma, \psi : [a,b] \to \mathbb{R}^2$ be two curves parameterized by arc length. Let $e(s)$ and $\tilde{e}(s)$ be the Frenet 2-frames of γ and ψ respectively, and let $\kappa(s)$ and $\widetilde{\kappa}(s)$ be the curvatures. If $\kappa(s) = \widetilde{\kappa}(s)$ for all $s \in [a,b]$ then there is an isometry $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f \circ \gamma = \psi$.*

Proof. Note that, since any pair of orthonomal bases are related by a unique orthogonal transformation, there is a (unique) orthogonal matrix A such that

$$Ae_1(a) = \widetilde{e}_1(a) \qquad \text{and} \qquad Ae_2(a) = \tilde{e}_2(a).$$

Now consider the system of linear first-order ODEs

$$\begin{pmatrix} x'_{1,1}(s) \\ x'_{1,2}(s) \\ x'_{2,1}(s) \\ x'_{2,2}(s) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \kappa(s) & 0 \\ 0 & 0 & 0 & \kappa(s) \\ -\kappa(s) & 0 & 0 & 0 \\ 0 & -\kappa(s) & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{1,1}(s) \\ x_{1,2}(s) \\ x_{2,1}(s) \\ x_{2,2}(s) \end{pmatrix}.$$

These are simply the Frenet equations 1.2 and 1.3 for γ (or ψ), so the scalar functions $e_{1,1}(s)$, $e_{1,2}(s)$, $e_{2,1}(s)$, and $e_{2,2}(s)$ satisfy these equations, as do the functions $\widetilde{e}_{1,1}(s)$, $\widetilde{e}_{1,2}(s)$, $\widetilde{e}_{2,1}(s)$, and $\widetilde{e}_{2,2}(s)$. By linearity, we then see that the component functions of $Ae_1(s)$ and $Ae_2(s)$ satisfy this linear system.

However, by construction, $Ae_1(s)$ and $Ae_2(s)$ have the same initial values as $\widetilde{e}_1(s)$ and $\widetilde{e}_2(s)$, respectively. Thus, the uniqueness of solutions to systems of linear ODEs tells us that $Ae_1(s) = \widetilde{e}_1(s)$ and $Ae_2(s) = \widetilde{e}_2(s)$ for all $s \in [a,b]$.

It is an easy exercise to see that the Frenet 2-frame associated to $A\gamma(s)$ is $\{Ae_1(s), Ae_2(s)\}$. Thus, since their Frenet 2-frames agree and they are parameterized by arc length, both γ and ψ satisfy the differential equation

$$\frac{d}{ds}\rho(s) = e_1(s)$$

which is the Frenet equation 1.1, under the assumption that γ is parameterized by arc length. We again see that both $A\gamma(s)$ and $\psi(s)$ satisfy this equation. Let $v \in \mathbb{R}^2$ be the unique vector such that $A\gamma(a) + v = \psi(a)$. Then $A\gamma(s) + v$ and ψ both satisfy this linear system, and have the same initial values. The uniqueness of solutions for systems of linear ODEs then tells us that, defining

$$f(x) = Ax + v,$$

we have $f \circ \gamma = \psi$, as desired. □

Proposition 1.17. *Let $\kappa(s) : [a,b] \to \mathbb{R}$ be a smooth function. Then there exists a (smooth) curve $\gamma : [a,b] \to \mathbb{R}^2$, parameterized by arc length, such that the curvature of γ is $\kappa(s)$.*

Proof. This follows immediately from the existence of solutions to the Frenet equations. The only technicality is the fact that the resulting frame is orthogonal, which we will show more generally in the next section. □

As a final application, lets try to understand what curves have *constant* curvature $\kappa \in \mathbb{R}$.

Corollary 1.18. *Let $\gamma : [a, b] \to \mathbb{R}^2$ be a curve parameterized by arc length such that the curvature $\kappa(s) = \kappa$ is constant.*

1. *If $\kappa = 0$, then γ is a line segment.*
2. *If $\kappa \neq 0$ then γ is an arc of a circle of radius $\frac{1}{|\kappa|}$.*

Proof. It suffices to compute the curvatures of circular arcs and line segments, and then apply Proposition 1.16. A line parameterized by arc length s is given by

$$\psi(s) = (s, 0).$$

and it is immediate that the curvature of ψ is 0. The arc-length parameterization of an arc of a circle of radius R is given by

$$\psi(s) = (R\cos(\pm\frac{s}{R}), R\sin(\pm\frac{s}{R}))$$

And, taking two derivatives, we find that the curvature is

$$\kappa(s) = \pm\frac{1}{R}.$$

This completes the proof. □

2 *Curvature in* $\mathbb{R}^n$

We now want to go back through our 2-dimensional definitions, and generalize them to curves in arbitrarily many dimensions. Our first task is to understand the appropriate generalization of the Frenet frame to this case. Ideally, given a smooth curve

$$\gamma : [a, b] \longrightarrow \mathbb{R}^n,$$

we would be able to define smooth functions

$$e_i : [a, b] \longrightarrow \mathbb{R}^n$$

with properties analogous to the Frenet 2-frame. Roughly speaking, the properties that we used in discussing the Frenet 2-frame should generalize as follows:

- For every $t \in [a, b]$, the ordered tuple of vectors

  $$(e_1(t), e_2(t), \ldots, e_n(t))$$

  form an orthonormal basis of $\mathbb{R}^n$.

- The vector $e_i(t)$ "corresponds to" the i^{th} derivative $\gamma^{(i)}(t)$ of γ at t.

We've been a little vague with what we mean by the second point, so let's develop these ideas further.

Definition 1.19. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a curve. A *(smooth) vector field* along γ is a smooth function

$$X\colon [a,b] \longrightarrow \mathbb{R}^n$$

where, for each $t \in [a,b]$, we consider $X(t)$ as a vector starting at $\gamma(t)$. We call a vector field X on γ a *tangent vector field* if $X(t)$ is a tangent vector to γ at $\gamma(t)$ for every $t \in [a,b]$.

Definition 1.20. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a (smooth) curve.

- A *moving n-frame* along γ consists of a pair of vector fields along γ which, for each $t \in [a,b]$, form an orthonormal basis of $\mathbb{R}^n$. More precisely, a moving n-frame consists of vector fields

  $$e_i\colon [a,b] \longrightarrow \mathbb{R}^n$$

  along γ for every $1 \leq i \leq n$. These vector fields must satisfy the following condition: For every $t \in [a,b]$ and every $1 \leq i,j \leq n$,

  $$e_i(t) \cdot e_j(t) = \delta_{i,j}.$$

  That is, the $e_i(t)$ form an orthonormal basis of $\mathbb{R}^n$.

- A *Frenet n-frame* for γ is a moving n-frame $(e_1, \ldots, e_n)$ along γ such that

  1. For every $k < n$ and every $t \in [a,b]$

     $$\text{Span}(e_1(t), \ldots, e_k(t)) = \text{Span}(\gamma'(t), \ldots, \gamma^{(k)}(t)),$$

     where Span denotes the span of a set of vectors in $\mathbb{R}^n$.
  2. For every $t \in [a,b]$, the basis $(e_1(t), e_2(t), \ldots, e_n(t))$ is positively oriented.
  3. For every $1 \leq k \leq n-1$,

     $$\gamma^{(k)}(t) \cdot e_k(t) > 0$$

It is important to note that not every curve has a Frenet n-frame. Conditions (1) and (3) together imply that $(\gamma'(t), \ldots, \gamma^{(k)}(t))$ is a linearly independent set for $k < n$, and that it has the same orientation as $(e_1(t), \ldots, e_k(t))$.

Definition 1.21. We call a (smooth) curve $\gamma : [a,b] \to \mathbb{R}^n$ a *Frenet curve* when, for every $t \in [a,b]$, the set

$$(\gamma'(t), \ldots, \gamma^{(n-1)}(t))$$

is linearly independent.

Example 1.22. Consider the helix

$$\begin{aligned} \gamma\colon [0,2\pi] &\longrightarrow \mathbb{R}^3 \\ t &\longmapsto (R\cos(t), R\sin(t), bt) \end{aligned}$$

The first two derivatives are

$$\gamma'(t) = (-R\sin(t), R\cos(t), b)$$
$$\gamma''(t) = (-R\cos(t), -R\sin(t), 0)$$

which are clearly linearly independent. Thus γ is a Frenet curve.

More generally, the condition that a regular curve be a Frenet curve is very simple in three dimensions. In this case, the requirement is simply that $\frac{d^2\gamma}{ds}^2$ does not vanish.

Lemma 1.23. *Suppose that $\gamma : [a,b] \to \mathbb{R}^3$ is a regular smooth curve. Then γ is a Frenet curve if and only if $\frac{d^2\gamma}{ds^2}$ is always non-zero.*

Proof. The "only if" direction is immediate. On the other hand, suppose that $\gamma''(t_0)$ is a parallel to $\gamma'(t_0)$ for some $t_0 \in [a,b]$.

Parameterize γ by arc length, and note that

$$\frac{d\gamma}{dt} = \frac{d\gamma}{ds}\frac{ds}{dt}$$
$$\frac{d^2\gamma}{dt^2} = \frac{d\gamma}{ds}\left(\frac{d^2s}{dt^2}\right) + \frac{d^2\gamma}{ds^2}\left(\frac{ds}{dt}\right)^2$$

and that the matrix

$$\begin{pmatrix} \frac{ds}{dt} & 0 \\ \frac{d^2s}{dt^2} & \left(\frac{ds}{dt}\right)^2 \end{pmatrix}$$

has determinant $\left(\frac{ds}{dt}\right)^3 = |\gamma'|^3$, and so is invertible for any $t \in [a,b]$ by regularity. This means that $(\gamma'(t_0), \gamma''(t_0))$ is a basis of $\mathbb{R}^2$ if and only if $\left(\frac{d\gamma}{ds}, \frac{d^2\gamma}{ds^2}\right)$ is a basis of $\mathbb{R}^2$.

Since the speed of an arc-length parameterization is 1, we see that

$$0 = \frac{d}{ds}\gamma' \cdot \gamma' = 2\frac{d\gamma}{ds} \cdot \frac{d^2\gamma}{ds^2}$$

so that $\frac{d\gamma}{ds}$ and $\frac{d^2\gamma}{ds^2}$ are orthogonal. At $s(t_0)$, these two vectors do not form a basis, and so

$$\frac{d^2\gamma}{ds^2}(s(t_0)) = 0.$$

as desired. □

Warning. You cannot apply the criterion of Lemma 1.23 to the second derivative of γ with respect to another parameter. For example, let

$$v := \frac{1}{\sqrt{3}}(1,1,1)$$

and define

$$\gamma : [\tfrac{1}{2}, 1] \longrightarrow \mathbb{R}^3$$
$$t^2 v \longrightarrow$$

So that γ is a quadratic parameterization of a straight line segment. Then note that

$$\gamma''(t) = 2v$$

and

$$\gamma'(t) = 2tv$$

so that γ is not a Frenet curve, by $\gamma''(t)$ is non-zero for all $t \in [\frac{1}{2}, 1]$. The arc-length parameterization of this same curve is

$$\gamma(s) = (s + \frac{1}{4})v$$

where $s \in [0, 3/4]$. Note that $\frac{d^2\gamma}{ds^2} = 0$, as expected.

We know that, for a curve γ to have a Frenet frame, it is necessary that γ be a Frenet curve. However, we have not established necessity. This is, however, not terribly hard, and is a consequence of the Gram-Schmidt orthonormalization process.

Exercise 1. Use the implicit function theorem to show that the final vector field $e_n(t)$ constructed in the proof below is, indeed, a C^∞ vector field. (You can assume without argument that $e_1, e_2, \ldots, e_{n-1}$ are C^∞.)

Proposition 1.24. *Let $\gamma : [a,b] \to \mathbb{R}^n$ be a Frenet curve. Then there is a unique Frenet frame $(e_1, \ldots, e_n)$ for γ.*

Proof. The Gram-Schmidt process uniquely determines e_1 through e_{n-1} to be

$$e_1 = \frac{1}{|\gamma'|}\gamma'$$
$$e_k = \frac{\gamma^{(k)} - \sum_{i=1}^{k-1}(\gamma^{(i)} \cdot e_i)e_i}{\left|\gamma^{(k)} - \sum_{i=1}^{k-1}(\gamma^{(i)} \cdot e_i)e_i\right|}$$

Since $e_1, \ldots, e_{n-1}$ then span an $n-1$-dimensional space, there is a unique unit vector e_n such that $(e_1, \ldots, e_n)$ form an oriented basis. □

We now want to generalize the Frenet-Serre equations to the n-dimensional case. Fortunately, this works in much the same way as the 2-dimensional case.

Proposition 1.25. *Let $\gamma : [a, b] \to \mathbb{R}^n$ be a smooth Frenet curve, and denote by $(e_1, \ldots, e_n)$ the Frenet n-frame for γ. Then γ and the e_i satisfy the differential equations*

$$\gamma'(t) = |\gamma'(t)| e_1(t) \tag{1.4}$$

$$e_i'(t) = \sum_{j=1}^{n} \omega_{i,j}(t) e_j(t). \tag{1.5}$$

where

$$\omega_{i,j}(t) = e_i'(t) \cdot e_j(t) = -\omega_{j,i}(t).$$

Moreover, for $j > i + 1$, we have that $\omega_{i,j}(t) = 0$.

Proof. Equation 1.4 is simply the first step of the Gram-Schmidt orthonormalization procedure, and equation 1.5, together with the definition of $\omega_{i,j}(t)$ follows by expanding $e_i'(t)$ in the basis given by the Frenet frame.

If we differentiate the orthonormality equation

$$e_i \cdot e_j = \delta_{i,j}$$

we obtain the relation

$$e_i' \cdot e_j = -e_i \cdot e_j'.$$

which gives the skew-symmetry of the $\omega_{i,j}$. Finally, we note that e_i is, by construction, a linear combination

$$e_i(t) = \sum_{k=1}^{i} \lambda_k(t) \gamma^{(k)}(t).$$

Differentiating shows that

$$e_i'(t) = \sum_{k=1}^{i} \lambda_k(t) \gamma^{(k+1)}(t).$$

so that, in particular, e_i' must be a linear combination of $e_1, \ldots, e_{i+1}$. □

Note that we can rewrite equation 1.5 in matrix form

$$\begin{pmatrix} e_1' \\ e_2' \\ \vdots \\ e_n' \end{pmatrix} = \begin{pmatrix} 0 & \omega_{1,2} & 0 & 0 & \cdots & 0 \\ -\omega_{1,2} & 0 & \omega_{2,3} & 0 & \cdots & 0 \\ 0 & -\omega_{2,3} & 0 & \omega_{3,4} & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & -\omega_{n-1,n} & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

As before, we want to find a version of the $\omega_{i,j}$ which do *not* depend on our choice of parameterization:

Lemma 1.26. *Let $\gamma : [a, b] \to \mathbb{R}^n$ be a Frenet curve, and let $\phi : [c, d] \to [a, b]$ be an orientation-preserving change of parameter. Let $e(t)$ be the Frenet frame for γ, and let $\widetilde{e}(u)$ be the Frenet frame for $\gamma \circ \phi$, and similarly for $\omega_{i,j}$ and $\widetilde{\omega}_{i,j}$. Then*

$$\frac{\widetilde{\omega}_{i,j}(u)}{\left|\frac{d}{du}\gamma(\phi(u))\right|} = \frac{\omega_{i,j}(\phi(u))}{\gamma'(\phi(u))}.$$

Proof. This is identical to the proof of Lemma 1.12. □

Definition 1.27. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a Frenet curve with Frenet n-frame $\{e_i(t)\}$. The smooth function

$$\kappa_i(t) = \frac{\omega_{i,i+1}(t)}{|\gamma'(t)|}$$

is called the i^{th} *curvature* of γ.

Remark 1.28. Note that for $n > 2$, if γ is parameterized by the arc length s, we have

$$\kappa_1(s) = \frac{\omega_{1,2}(s)}{|\gamma'(s)|} = \frac{e_1'(s) \cdot e_2}{1} = \frac{\gamma''(s) \cdot e_2(s)}{1} = |\gamma''(s)| > 0$$

Notice that, unlike in two dimensions, the curvature $\kappa_1(s)$ is always positive, because we have taken the convention that (e_1, e_2) has the same orientation as $(\gamma'(s), \gamma''(s))$.

We can reformulate the Frenet equations in terms of the curvatures κ_i, yielding our final form of the equations. First, though, we need to fix some notation.

Definition 1.29. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a Frenet curve. We denote by $E(t)$ the matrix whose rows are the vectors $e_i(t)$ which make up the Frenet n-frame. That is, we write

$$E(t) := \begin{pmatrix} e_1(t) \\ e_2(t) \\ \vdots \\ e_n(t) \end{pmatrix}.$$

Now, given a Frenet curve $\gamma : [a,b] \to \mathbb{R}^n$ *parameterized by arc length*, we note that $\omega_{i,i+1}(s) = \kappa_i(s)$. We thus can reformulate the Frenet equations as the matrix ODE

$$E'(s) = K(s)E(s) \tag{1.6}$$

where $K(s)$ is the skew-symmetric matrix

$$K(s) := \begin{pmatrix} 0 & \kappa_1 & 0 & 0 & \cdots & 0 \\ -\kappa_1 & 0 & \kappa_2 & 0 & \cdots & 0 \\ 0 & -\kappa_2 & 0 & \kappa_3 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & -\kappa_{n-1} & 0 \end{pmatrix}.$$

As before, we have existence and uniqueness theorems for curves with given curvature functions. The proofs are more or less the same as the corresponding proofs in two dimensions. We will state the uniqueness theorem, and then prove the existence theorem, along the way filling in the gap in the proof of the 2-dimensional existence theorem.

Proposition 1.30. *Let $\gamma, \psi : [a,b] \to \mathbb{R}^n$ be two Frenet curves parameterized by arc length with Frenet n-frames given by $\{e_i(s)\}$ and $\{\widetilde{e}_i(s)\}$. Suppose that γ and ψ have the same curvature functions $\kappa_i(s)$ for $1 \leq i \leq n-1$. Then there is a unique isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ such that $f \circ \gamma = \psi$.*

Proof. The proof is virtually identical to that of Proposition 1.16. □

Finally, we come to existence.

Proposition 1.31. *Let $\kappa_i : [a, b] \to \mathbb{R}$ for $1 \leq i \leq n-1$ be smooth functions such that $\kappa_i(s) > 0$ for $1 \leq i \leq n-2$. Then there exists a Frenet curve $\gamma : [a, b] \to \mathbb{R}^n$ parameterized by arc length such that the κ_i are the curvature functions of γ.*

Proof. The argument hinges on the existence of solutions to the linear ODE

$$E'(s) = K(s)E(s).$$

By [5, Theorem 6.2.3], for any choice of an initial orthonormal frame $E(a)$, a solution to this equation exists on the interval $[a, b]$. It is immediate from the fact that $K(s)$ is C^∞ that the solution $E(s)$ must also be C^∞.

Before continuing, however, we must show that $E(s)$ defines an orthnormal frame. We will do this by showing that the matrix $E(s)^T E(s)$ is constant. Since the initial data $E(a)$ consist of an orthogonal matrix, this will suffice to prove that $E(s)$ is orthogonal. Note, before we begin, that $K(s)^T = -K(s)$. We compute

$$\begin{aligned}
\frac{d}{ds}(E(s)^T E(s)) &= (E'(s))^T E(s) + E(s)^T E'(s) \\
&= (K(s)E(s))^T + E(s)K(s)E(s) \\
&= E(s)^T K(s)^T E(s) + E(s)K(s)E(s) \\
&= -E(s)^T K(s)E(s) + E(s)K(s)E(s) \\
&= 0
\end{aligned}$$

so that $E(s)^T E(s)$ is constant, as desired. For an arbitary choice of inital value $\gamma(a)$, we can then solve the linear ODE

$$\gamma'(s) = e_1(s).$$

to obtain a smooth curve $\gamma : [a, b] \to \mathbb{R}^n$.

To see that this is a Frenet curve, we first note that regularity follows from the equation

$$\gamma'(s) = e_1(s)$$

Taking successive derivatives of γ shows that $\gamma^{(i)}$ lies in the span of $e_1, \ldots, e_i$. In particular, the coefficient of e_i in the expansion of $\gamma^{(i)}$ in terms of the basis $(e_1, \ldots, e_i)$ is

$$\kappa_1(s)\kappa_2(s) \cdots \kappa_i(s)$$

and so

$$\gamma^{(i)} \cdot e_i > 0.$$

As a consequence, the vectors $\gamma', \ldots, \gamma^{(i)}$ are linearly independent, completing the proof. □

Exercise 2. Let $\gamma : [a, b] \to \mathbb{R}^n$ be a Frenet curve such that κ_{n-1} vanishes identically on $[a, b]$. Show that γ is contained in a hyperplane of dimension $(n-1)$.

3 *Three dimensions, osculating circles*

Now that we have a theory of curves and curvatures in arbitrary dimensions, let us specialize to a more visualizable case: dimension 3. In this case, a regular smooth curve γ is Frenet if and only if $\frac{d^2\gamma}{ds^2} \neq 0$, so the condition that γ be Frenet is not terribly restrictive.

Example 1.32. Let's return to the helix of Example 1.5. We will work directly with the arc-length parameterization,

$$\gamma(s) = \left(R\cos\left(\frac{s}{d}\right), R\sin\left(\frac{s}{d}\right), b\frac{s}{d}\right) \qquad s \in [0, 6\pi d]$$

where $R > 0$ and $d = \sqrt{R^2 + b^2}$.

We first compute the Frenet frame of γ. The first two derivatives are

$$\gamma'(s) = \left(-\frac{R}{d}\sin\left(\frac{s}{d}\right), \frac{R}{d}\cos\left(\frac{s}{d}\right), \frac{b}{d}\right)$$
$$\gamma''(s) = \left(-\frac{R}{d^2}\cos\left(\frac{s}{d}\right), -\frac{R}{d^2}\sin\left(\frac{s}{d}\right), 0\right)$$

We already know that $e_1 = \gamma'(s)$, so it remains for us to compute the norm

$$\kappa_1(s) := |\gamma''(s)| = \sqrt{\frac{R^2}{d^4}} = \frac{R}{d^2} = \frac{R}{R^2 + b^2}$$

The second Frenet vector, is thus

$$e_2(s) = \frac{d^2}{R}\gamma''(s) = \left(-\cos\left(\frac{s}{d}\right), -\sin\left(\frac{s}{d}\right), 0\right)$$

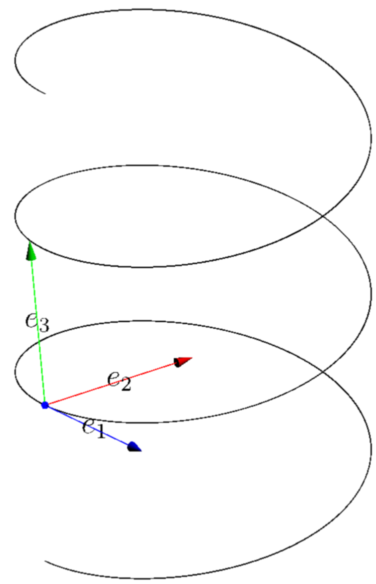
The third vector can be conveniently obtained using the cross product

$$e_3(s) = e_1(s) \times e_2(s) = \left(\frac{b}{d}\sin\left(\frac{s}{d}\right), -\frac{b}{d}\cos\left(\frac{s}{d}\right), \frac{R}{d}\right)$$

We can then compute the second curvature

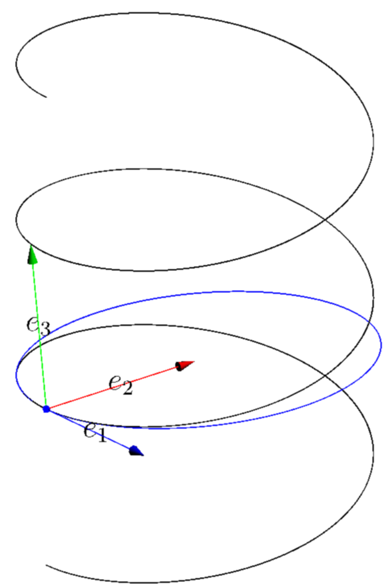
$$\kappa_2(s) = e_2'(s) \cdot e_3(s) = \frac{b}{d^2} = \frac{b}{R^2 + b^2}$$

Let's picture the Frenet vectors at the point $s = 2\pi d$ on the helix. Visually, we obtain



The vector e_1 is, of course, the tangent vector, and precisely as in the 2-dimensional case, the vector e_2 points in the direction the curve is bending. The vector e_3 is then determined by e_1 and e_2.

If we also include the osculating circle, we obtain



Our discussion of osculating circles from the overture can be carried over nearly verbatim in this context. Consider a Frenet curve

$$\gamma\ [a, b] \longrightarrow \mathbb{R}^3$$

parameterized by arc length, and choose $s_0 \in [a, b]$. We abbreviate $\kappa_1(s_0) = \kappa_1$, $e_i(s_0) = e_1$, etc. Then define an arc-length parameterized circle in e_1-e_2 plane by

$$\phi(u) = \tfrac{1}{\kappa_1}\left(-\cos(\kappa_1 u)e_2 + \sin(\kappa_1 u)e_1\right) + \gamma(s_0) + \tfrac{1}{\kappa_1}e_2$$

Notice that $\phi(0) = \gamma$. We can compute the derivatives

$$\phi'(u) = \sin(\kappa_1 u)e_2 + \cos(\kappa_1 u)e_1$$
$$\phi''(u) = -\kappa(-\cos(\kappa_1 u)e_2 + \sin(\kappa_1 u)e_1)$$

We then see that

$$\phi'(0) = e_1 = \gamma'(s_0)$$
$$\phi''(0) = \kappa_1 e_2 = \gamma''(s_0)$$

As such, the circle ϕ has contact of order 2 with the curve γ.

Definition 1.33. Let $\gamma : [a, b] \to \mathbb{R}^3$ be a curve. The *osculating plane* of γ at $s_0 \in [a, b]$ is the plane containing $\gamma(s_0)$ and spanned by $e_1(s_0)$ and $e_2(s_0)$. The *osculating circle* of γ at s_0 is the best circular approximation of γ at s_0, and is the circle in the osculating plane with radius $\frac{1}{\kappa_1(s_0)}$ and center

$$c = \gamma(s_0) + \frac{1}{\kappa_1(s_0)} e_2(s_0).$$

Exercise 3. Show that, for an arbitrary Frenet curve $\gamma : [a, b] \to \mathbb{R}^3$, the identity

$$\kappa_1(t) = \frac{|\gamma'(t) \times \gamma''(t)|}{|\gamma'(t)|^3}$$

holds.

Exercise 4. Characterize all three-dimensional Frenet curves with constant curvatures κ_1 and κ_2, as in Cor 1.18.

Remark 1.34. We should here remark that, in dimension 3, a slightly different notation is more common. Given a Frenet curve γ, one typically writes the unit tangent vector $e_1(t)$ as $\mathbf{T}(t)$. We call the vector $e_2(t)$ the *unit normal* vector, and denote it by $\mathbf{N}(t)$, and we call the vector $e_3(t)$ the *binormal*, and denote it by $\mathbf{B}(t)$.

Moreover, in dimension 3, the first curvature $\kappa_1(t)$ is usually simply called the *curvature*, and denoted $\kappa(t)$. The second curvature $\kappa(2)$ is usually called the *torsion*, and denoted by τ.

4 *Curves in 2d Minkowski space*

So far, we have restricted ourselves to the conventional Euclidean metric on $\mathbb{R}^n$ — the distance between points x and y is given by

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

This distance is defined via the *Euclidean scalar product* — also called the *dot product* — which we write as

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i w_i.$$

We can then define distance by

$$d(x, y)^2 = \langle (x - y), (x - y) \rangle.$$

In physics — more precisely the physics of general relativity — it is common to consider a different inner product: the *Minkowski inner product.* In this section, we will consider a 2-dimensional toy model of this inner product, and attempt to understand the geometry of curves in this context.

Definition 1.35. Let $v, w \in \mathbb{R}^2$ be two vectors. The *Minkowski inner product* of v and w is defined to be

$$\langle v, w \rangle_{1,1} := v_1 w_1 - v_2 w_2.$$

Notice that there are non-zero vectors v such that $\langle v, v \rangle_{1,1} = 0$.[5] We denote the space $\mathbb{R}^2$ equipped with the inner product $\langle -, - \rangle_{1,1}$ by $\mathbb{R}^{1,1}$.

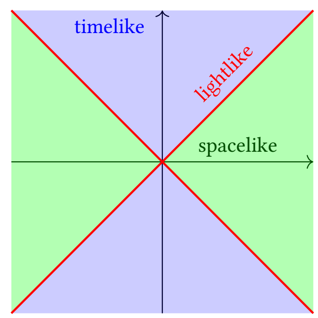
[5] We sometimes say that the *bilinear form* $\langle -, - \rangle_{1,1}$ is *degenerate*.

There are three types of vectors in $\mathbb{R}^{1,1}$:

- A vector $v \in \mathbb{R}^{1,1}$ is called *timelike* if $\langle v, v \rangle_{1,1} < 0$.
- A vector $v \in \mathbb{R}^{1,1}$ is called *spacelike* if $\langle v, v \rangle_{1,1} > 0$.
- A vector $v \in \mathbb{R}^{1,1}$ is called *lightlike* (or *null*) if $\langle v, v \rangle_{1,1} = 0$.

Corresponding to these three cases, we have three corresponding special types of curves.

We draw the spacelike, timelike, and lightlike vectors in $\mathbb{R}^{1,1}$.



Definition 1.36. A smooth curve $\gamma : [a, b] \to \mathbb{R}^{1,1}$ is called *regular* when γ' is always a non-zero vector. We call a regular curve γ timelike, spacelike, or lightlike, respectively, when $\gamma'(t)$ is timelike, spacelike, or lightlike, respectively.

Exercise 5. Show that if γ is a regular smooth null curve, then γ is a straight line segment with slope ± 1.

For spacelike or timelike curves, we can define a reasonable notion of arc length, using more or less the same definitions we used in Euclidean space.

Definition 1.37. Let $\gamma : [a, b] \to \mathbb{R}^{1,1}$ be a smooth regular curve.

- If γ is timelike, we define the *arc length* of γ to be

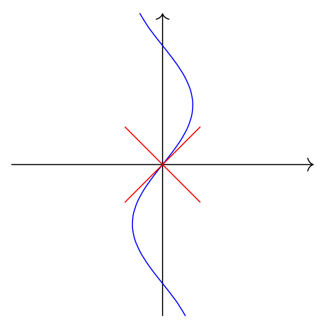
$$L(\gamma; a, b) := \frac{1}{i} \int_a^b \sqrt{\langle \gamma'(u), \gamma'(u) \rangle_{1,1}} du$$

  where i denotes the imaginary unit.

- If γ is spacelike, we define the *arc length* of γ to

$$L(\gamma; a, b) := \int_a^b \sqrt{\langle \gamma'(u), \gamma'(u) \rangle_{1,1}} du$$

Timelike curves must have timelike tangent vectors for every value their parameters, so at each point in their trajectory, they must be bounded by the two lightlike lines through that point. For instance the following is a Timelike curve.



Similar considerations apply for spacelike curves.

4.1 Hyperbolic trigonometric functions

Our first aim in exploring $\mathbb{R}^{1,1}$ is to understand the relation of arc length and angle to the sin and cosine functions.

Let us briefly recall how sine and cosine are defined. We define a curve, the unit circle, by requiring that

$$\langle x, x \rangle = 1$$

Parameterizing this curve by arc length then yields

$$\gamma(s) = (\cos(s), \sin(s)).$$

We will now attempt to do the same in Minkowski space. The equation corresponding to the circle is

$$\langle x, x\rangle_{1,1} = 1$$

or, in more familiar notation,

$$x_1^2 - x_2^2 = 1.$$

This is a hyperbola, whose asymptotes are the lightlike curves through the origin.[6]

We can parameterize one sheet of the hyperbola as

$$\gamma(t) = (\sqrt{t^2+1}, t) \qquad t \in \mathbb{R}.$$

and so our first objective is to parameterize this curve by (Minkowski) arc-length. Taking a derivative, we find

$$\gamma'(t) = \left(\frac{t}{\sqrt{t^2+1}}, 1\right)$$

So that

$$\langle \gamma'(t), \gamma'(t)\rangle_1 = \frac{t^2}{t^2+1} - 1 = \frac{-1}{t^2+1}$$

This tells us that this curve is timelike.

To compute the arclength of the curve, we consider the integral

$$s(t) = \int_0^t \sqrt{-\langle \gamma'(y), \gamma'(y)\rangle_1}du = \int_0^t \frac{1}{\sqrt{y^2+1}}dy$$

We can perform a somewhat odd substitution:

$$\begin{aligned} u = \sqrt{y^2+1} + y du \quad &= \frac{\sqrt{y^2+1}+y}{\sqrt{y^2+1}}dy \\ = \frac{u}{\sqrt{y^2+1}}dy. \end{aligned}$$

We thus can simplify the integral to

$$\int_1^{\sqrt{t^2+1}+t} \frac{du}{u} = \ln\left(\sqrt{t^2+1}+t\right).$$

On final assessment, therefore, we obtain

$$s(t) = \ln\left(\sqrt{t^2+1}+t\right)$$

We can then solve for t

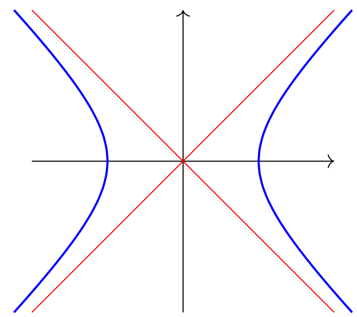
$$t = \frac{e^{2s}-1}{2e^s} = \frac{e^s - e^{-s}}{2}$$

This is the y-coordinate of the corresponding point on the hyperbola. We thus can think of it as our analogue of the sine function in Minkowski space.

Definition 1.38. The *hyperbolic sine function* is the function

$$\sinh(s) := \frac{e^s - e^{-s}}{2}.$$

[6] The plot, complete with asymptotes, looks like this.



We can also compute the x-coordinate — our cosine analogue — by plugging $t(s)$ into the first component of γ:

$$\begin{aligned} x(s) &= \sqrt{\left(\frac{e^s - e^{-s}}{2}\right)^2 + 1} \\ &= \sqrt{\frac{e^{2s} - 2 + e^{-s}}{4} + 1} \\ &= \sqrt{\frac{e^{2s} + 2 + e^{-s}}{4}} \\ &= \frac{e^s + e^{-s}}{2} \end{aligned}$$

Definition 1.39. The *hyperbolic cosine function* is the function

$$\cosh(s) := \frac{e^s + e^{-s}}{2}.$$

One of the odd features of Minkowski space is that our circle analogue — the hyperbola — has two pieces. We will refer to these as *path components*, since no continuous path which lies in one of these pieces can reach into the other. One of the effects of this fact is explored in the following exercise.

Exercise 6. Show that for any point (a, b) on the hyperbola defined by

$$x_1^2 - x_2^2 = 1,$$

there is a unique $\alpha \in \mathbb{R}$ and a unique sign $\pm$ such that

$$(a, b) = (\pm \cosh(\alpha), \sinh(\alpha)).$$

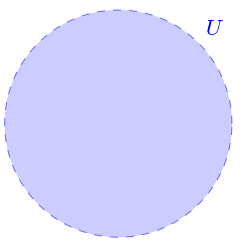
2

Parameterized submanifolds

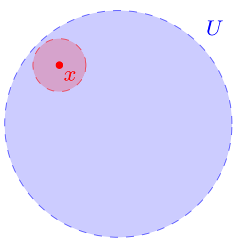
So far, we have considered only the geometry of curves: 1-dimensional subspaces of $\mathbb{R}^n$. Our next goal is to generalize this to k-dimensional subspaces of $\mathbb{R}^n$. Our approach does not need to change much. Jumping straight to dimension k from 1 may seem a little extreme, but we will quickly specialize to easier cases: first to the case of *hypersurfaces* ($k = n - 1$), and then to *surfaces in 3d space* ($k = 2$ and $n = 3$). We work in greater generality initially to emphasize that many of the concepts we will define work equally well in any dimension.

Some results and constructions will, of course, depend on the dimensions n and k (e.g., the Gauß map), and we will comment explicitly when this occurs. We will begin by recalling some properties of subsets of $\mathbb{R}^n$.

Pictorially, an open set is one that does not contain its boundary. E.g.



Given a point $x \in U$, we can always draw a little ball around x which always stays in U.



Definition 2.1. We call a subset $U \subseteq \mathbb{R}^n$ *open* when, for every point $x \in U$, there is some $r > 0$ such that the *open ball*

$$B_r(x) := \{y \in \mathbb{R}^n \mid |y - x| < r\}$$

is contained in U. Intuitively (and in a way which can be made rigorous) an open set is a set which does not contain its boundary.

Example 2.2.

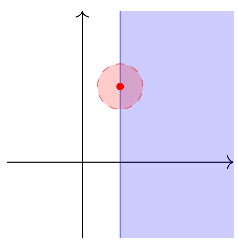
1. In $\mathbb{R}$, every open interval (a, b) is an open set.

2. The set $\mathbb{R}^n \subseteq \mathbb{R}^n$ is open.

3. The set

$$H := \{x \in \mathbb{R}^2 \mid x_1 \geq 1\}$$

is *not* open. The reason for this is that, for any point $(a, 1) \in H$ and any $r > 0$, the ball $B_r(a, 1)$ will contain points whose first coordinate is less than 1, as pictured below.



4. The set
$$\{x \in \mathbb{R}^2 \mid x_1 > 1\}$$
is open.

5. For any point $x \in \mathbb{R}^n$ and any $r > 0$, the open ball $B_r(x)$ of radius r around x is open.

Notice that the majority of our key analytic theorems from **??** assume that we are working on some open set. It is for this reason that we will, throughout this course, tend to work with open sets.

Before we can start in on the definition of submanifolds, we need to better understand what the differentials of multivariable functions mean.

1 Differentiation on $\mathbb{R}^n$

I'm not going to aim for a rigorous exposition of differentiation and integration in these notes. Such matters are better left to an analysis course. I will, however, try to develop some computational tools, as well as some of the underlying intuitions.

The key point of a derivative is to approximate an arbitrary function by something linear. In single-variable calculus, when we differentiate a function $f : \mathbb{R} \to \mathbb{R}$ at a point x, the defining property is that

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

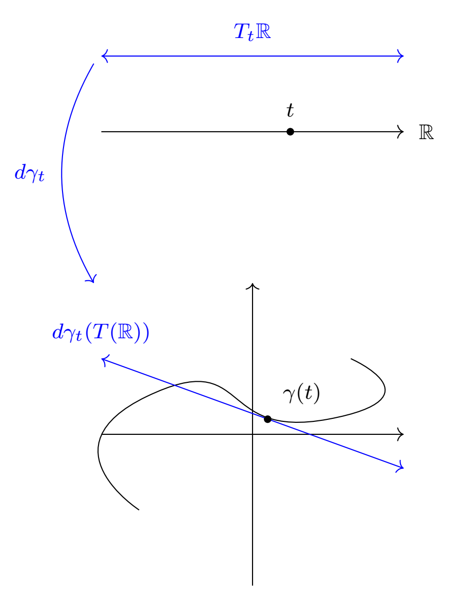
for x close to x_0. We make this formal by defining the derivative $f'(x_0)$ to be the real number (if one exists) such that

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0 + h)}{h} = 0.$$

Equivalently, we can define

$$f'(x_0) = \lim_{h \to 0} .$$

Pictorially



Our first question to answer is *what does the number* $f'(x_0)$ *mean?* We can think of it as the slope of the tangent line to $f(x)$, or as a velocity. More useful, however, is to view $f'(x_0)$ as a *linear transformation*. We can think of $f'(x_0)$ as the linear transformation which turns possible velocities at the point x_0 to possible velocities at the point $f(x_0)$. We typically formalize this viewpoint by defining the *tangent space* to $\mathbb{R}$ at x_0 to be $T_{x_0}\mathbb{R} := \mathbb{R}$, which we view as the possible velocities of linear paths through x_0 in $\mathbb{R}$. We then

Even clearer is when we consider a curve $\gamma : \mathbb{R} \to \mathbb{R}^2$. To each point $z \in \mathbb{R}^2$ we associate a tangent space $T_z\mathbb{R}^2 := \mathbb{R}^2$ — viewed as the possible velocities of linear paths through z — The derivative $\gamma'(t) = (\gamma_1'(t), \gamma_2'(t))$ can then be viewed as a transformation

$$d\gamma_t : T_t\mathbb{R} \longrightarrow T_{\gamma(t)}\mathbb{R}^2,$$

The image of this transformation is a linear subspace, which can be identified with the space of all possible tangent vectors to the image of γ at the point $\gamma(t)$ — The *tangent space of the curve* γ at $\gamma(t)$.

To make this intuition formal, we make the following definitions.

Definition 2.3. Let $U \subset \mathbb{R}^n$ be an open set. For any $x \in U$, the *tangent space of U at x* is the vector space

$$T_xU := \mathbb{R}^n.$$

The *tangent bundle* of U is the set TU consisting of pairs (x, v) where $x \in U$ and $v \in T_xU$. We call (x, v) (or $v \in T_xU$) a *tangent vector to U at x*. Notice that there is a canonical identification

$$TU \cong U \times \mathbb{R}^n,$$

and so we can view TU as a subset of $\mathbb{R}^{2n}$.

Definition 2.4. Let $U \subset \mathbb{R}^n$ be an open subset, and $f : U \to \mathbb{R}^m$ a function. The *(total) derivative Df_x* at a point $x \in U$ is the linear map

$$(Df)_x : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

such that

$$\lim_{y \to x} \frac{|f(y) - f(x) - Df_x(y - x)|}{|y - x|} = 0.$$

If the total derivative of f exists at every point $x \in U$, we call f *differentiable on U*, and we define a map

$$\begin{aligned} df :\ TU &\longrightarrow T\mathbb{R}^m \\ (x, v) &\longmapsto (f(x), Df_x(v)) \end{aligned}$$

called the *differential* of f.

Remark 2.5. Under the canonical identifications $T_xU \cong U \times \mathbb{R}^n$, we can identify df_x with Df_x. Moreover, we can equivalently view df as a map

$$\begin{aligned} U &\longrightarrow \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^m) \\ x &\longmapsto Df_x \end{aligned}$$

from U to linear maps $\mathbb{R}^n \to \mathbb{R}^m$.

Remark 2.6. Notice that if the differential of $f : U \to \mathbb{R}^m$ exists at every point, then f is continuous.

One of the key questions we need to answer is ”how do we compute using total derivatives and differentials?” The solution is the following lemmata.

Lemma 2.7. *Let $U \subset \mathbb{R}^n$ be open, and let $f : U \to \mathbb{R}^m$. Denote by $f_i : U \to \mathbb{R}$ for $1 \leq i \leq m$ the i^{th} component function. If f_i is differentiable for every $1 \leq i \leq m$, then f is differentiable*

Proof. It suffices to show the lemma at a point $x \in U$. Suppose the differential of f_i at x exists, and represent it by a $1 \times n$ matrix A_i, so that, for any $1 \leq i \leq m$,

$$\lim_{y \to x} \frac{|f_i(y) - f_i(x) - A_i(y - x)|}{|y - x|} = 0.$$

For $\epsilon > 0$, choose $\delta > 0$ such that, for any $y \in U$ with $|y - x| < \delta$,

$$|f_i(y) - f_i(x) - A_i(y - x)| < \frac{\epsilon}{\sqrt{m}}|y - x|$$

for every $1 \leq i \leq m$.

Define a $m \times n$ matrix A by

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}.$$

We then see that for $y \in U$ such that $|y - x| < \delta$, we have

$$\begin{aligned} |f(y) - f(x) - Ay - x|^2 &= \sum_{j=1}^{m} (f_i(y) - f_i(x) - A_i(y - x))^2 \\ &< \sum_{i=1}^{m} \frac{\epsilon^2}{m}|y - x|^2 = \epsilon^2|y - x|^2. \end{aligned}$$

We thus see that

$$\lim_{y \to x} \frac{|f(y) - f(x) - A(y - x)|}{|y - x|} = 0,$$

as desired. □

Our second lemma is a generalization of the chain rule from single-variable calculus.

Exercise 7. Let B be an $m \times n$ matrix. Define

$$\|B\| := \sup_{v \in \mathbb{R}^n} \frac{|Bv|}{|v|}.$$

Show that $\|B\|$ is always finite. Note that for any $v \in \mathbb{R}^n$,

$$|Bv| \leq \|B\||v|.$$

Lemma 2.8. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$ be differentiable. Then $g \circ f$ is differentiable, and*

$$D(g \circ f)_x = (Dg)_{f(x)} \circ Df_x$$

for any $x \in \mathbb{R}^n$

Proof. Fix $x \in \mathbb{R}^n$, and denote by A the $m \times n$ matrix representing Df_x, and by B the matrix representing $Dg_{f(x)}$. By the triangle inequality

$$\begin{aligned} |g(f(y)) - g(f(x)) - BA(y - x)| \leq &|g(f(y)) - g(f(x)) - B(f(y) - f(x))| \\ &+ |Bf(y) - B(f(x)) - BA(y - x)| \end{aligned}$$

Since f is differentiable at x, for any $\epsilon > 0$ we may choose $\delta_1 > 0$ such that for $|y - x| < \delta_1$, we have

$$\frac{|f(y) - f(x)|}{|y - x|} < \|A\| + \epsilon$$

and

$$|f(y) - f(x) - A(y - x)| < \epsilon|y - x|.$$

Similarly, since g is differentiable at $f(x)$ and f is continuous, we may choose $\delta_2 > 0$ such that when $|y - x| < \delta_2$,

$$|g(f(y)) - g(f(x)) - B(f(x) - f(y))| < \epsilon|f(y) - f(x)|.$$

Taking $\delta = \min(\delta_1, \delta_2)$ we see that for $|y - x| < \delta$, we have

$$\begin{aligned}|g(f(y)) - g(f(x)) - BA(y - x)| \leq & |g(f(y)) - g(f(x)) - B(f(y) - f(x))| \\ & + |Bf(y) - B(f(x)) - BA(y - x)| \\ < & \epsilon|f(y) - f(x)| + \epsilon\|B\||y - x|\end{aligned}$$

Dividing through by $|y - x|$, we obtain

$$\frac{|g(f(y)) - g(f(x)) - BA(y - x)|}{|y - x|} < \epsilon\left(\frac{|f(y) - f(x)|}{|y - x|} + \|B\|\right)$$

And apply the property of δ_1 to see that

$$\frac{|g(f(y)) - g(f(x)) - BA(y - x)|}{|y - x|} < \epsilon\left(\|A\| + \epsilon + \|B\|\right).$$

Since ϵ can be chosen arbitrarily small, this completes the proof. □

Remark 2.9. When considering differentials, the statement of this lemma can be simplified even further. Let $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ and $W \subset \mathbb{R}^k$ be open subsets, and $g : U \to V$ and $f : V \to W$ be smooth maps. Then

$$d(f \circ g) = df \circ dg$$

as maps $TU \to TW$.

This lemma has an immediate corollary, allowing us to give a matrix representing Df_x.

Corollary 2.10. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function, and denote by $x_i : \mathbb{R} \to \mathbb{R}^n$ the i^{th} coordinate function*

$$\gamma_i(t) = (0, \ldots, 0, \underbrace{t}_{i^{th}}, 0 \ldots, 0).$$

Denote by A the matrix representing Df_0 with respect to the standard bases. Then

$$A_{i,j} = \frac{d}{dt}(f_i \circ \gamma_j)|_{t=0}.$$

The implication of this corollary is that we can compute the matrix representation of $A_{i,j}$ using only the techniques of 1-variable calculus.

Definition 2.11. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function, and let $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Define a curve in $\mathbb{R}^n$ through a by

$$\gamma_i(t) = (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n).$$

The i^{th} *partial derivative of* f[1] at a is

$$\frac{\partial f}{\partial x_i}(a) := \frac{d}{dt} f \circ \gamma_i|_{t=0}.$$

[1] The partial derivative function $\frac{\partial f}{\partial x^i}$ can be computed by applying the usual differentiation rules to an expression for f, treating variables other than x_i as constants.

Given a differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$, and $x \in \mathbb{R}^n$ the matrix representing Df_x is called the *Jacobian* of f at x, and is denoted by Jf_x. By the corollary, we have

$$(Jf_x)_{i,j} = \frac{\partial f_i}{\partial x^j}.$$

We can rewrite Lemma 2.8 in terms of Jacobians.

$$(J(f \circ g))_{i,j} = \frac{\partial (f \circ g)_i}{\partial_x^j} = \sum_{k=1}^{m} \frac{\partial f_i}{\partial y^k} \frac{\partial g_k}{\partial x^j}.$$

Which is the usual chain rule for partial derivatives.

We will also make use of the notion of *smoothness.* This is more or less the same as the corresponding notion for single-variable functions.

Definition 2.12. Let $U \subset \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^m$ a function.

- We call f a C^1 *function* if f is differentiable at every point $x \in U$, and the map

$$\begin{aligned} Df : U &\longrightarrow \mathbb{R}^{n \times m} \\ x &\longmapsto (Jf)_x \end{aligned}$$

  is continuous.

- We call f a C^2 *function* if f is C^1 and the map

$$\begin{aligned} Df : U &\longrightarrow \mathbb{R}^{n \times m} \\ x &\longmapsto (Jf)_x \end{aligned}$$

  is C^1.

- $\vdots$

We say that f is C^∞ (or *smooth*) if it is C^k for any k.

Remark 2.13. If $f : U \to V$ is smooth, then its differential, viewed as a map

$$df : U \times \mathbb{R}^n \longrightarrow V \times \mathbb{R}^m,$$

is smooth. Indeed, f is smooth if and only if f is C^1 and df is smooth.

We will not focus on proving smoothness here, but will rather note some facts which allow us to check when functions are smooth.

Fact 2.14.
- Any composite of smooth functions is smooth.
- Any sum or difference of smooth functions is smooth.

- Any rational function of n variables is smooth on its domain.
- The function $\sqrt{x}$ is smooth on $(0, \infty)$. The functions exp, sin, cos, and ln are all smooth on their respective (open) domains.

As our final definition, we want a smooth notion of "sameness"

Definition 2.15. Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open subsets. We call a function $f : U \to V$ a *diffeomorphism* if f is a C^∞ bijection, and the function f^{-1} is also C^∞.

Exercise 8. Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be nonempty open subsets. Show that if $f : U \to V$ is a diffeomorphism, then $n = m$.

2 Submanifolds, first steps

Now that we've established how to work with derivatives, we can get back to geometry. We're going to define embedded submanifolds in a fairly high degree of generality, but most of our examples will live in $\mathbb{R}^2$ or $\mathbb{R}^3$. The image you should keep in mind throughout this chapter is that of a surface which is curved into three dimensions, like a torus or a sphere.

Definition 2.16. Let $U \subset \mathbb{R}^k$ be an open subset, and let $k < n$. A *(smooth) parameterization* is a smooth map

$$\phi : U \longrightarrow \mathbb{R}^n.$$

We say that ϕ is *regular* if the differential $df_x : T_xU \to T_{f(x)}\mathbb{R}^n$ is injective for every $x \in U$ (or equivalently if the Jacobian $(Jf)_x$ has maximal rank). We call the point $\phi(x)$ *singular* if df_x does not have full rank.

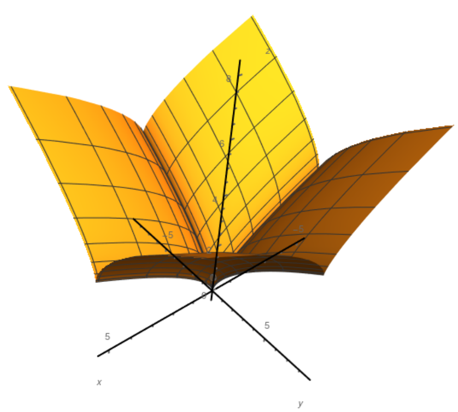
Before continuing, let's examine why we want to require parameterizations to be regular. Consider the open set $U = (-2,2) \times (-2,2) \subset \mathbb{R}^2$, and define a parameterization

$$\begin{aligned} \phi \quad U &\longrightarrow \mathbb{R}^3 \\ (x_1, x_2) &\longmapsto (x_1^3, x_2^3, (x_2^2 + x_2^2)). \end{aligned}$$

The Jacobian of this function is

$$J\phi = \begin{pmatrix} 3x_1^2 & 0 \\ 0 & 3x_2^2 \\ 2x_1 & 2x_2 \end{pmatrix}$$

And we notice that this matrix only has full rank when x_1 and x_2 are both non-zero. We might initially think that we could find some other parameterization of the same surface which is non-singular, but looking at a plot

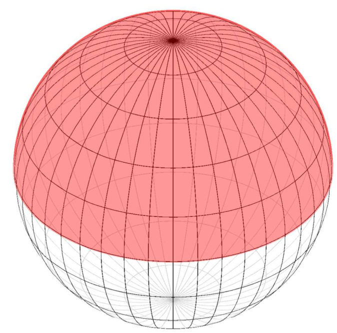


we see that there are "creases" when x_1 or x_2 is zero. These are precisely what we seek to avoid when we require that our parameterizations be non-singular.

Let's think about some examples of submanifold elements. Let $U \subset \mathbb{R}^2$ be the open unit ball around the origin, $U = \{x \in \mathbb{R}^2 \mid |x| < 1\}$. We can define a map

$$\begin{aligned} \phi : U &\longrightarrow \mathbb{R}^3 \\ x &\longmapsto (x_1, y_2, \sqrt{1 - x_1^2 - x_2^2}) \end{aligned}$$

Which parameterizes the upper hemisphere of the 2d sphere.



Computing the Jacobian, we obtain

$$J\phi := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{-x_1}{\sqrt{1-x_1^2-x_2^2}} & \frac{-x_2}{\sqrt{1-x_1^2-x_2^2}} \end{pmatrix}$$

And clearly, this has maximal rank.

We can give another parameterization of (part of) the upper hemisphere, using spherical coordinates. Let $U = (-\pi/2, \pi/2) \times (-\pi, \pi)$, and define

$$\begin{aligned} \psi : \quad U &\longrightarrow \mathbb{R}^3 \\ (\phi, \theta) &\longmapsto (\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi)) \end{aligned}$$

The corresponding Jacobian is

$$J\psi = \begin{pmatrix} \cos(\phi)\cos(\theta) & -\sin(\phi)\sin(\theta) \\ \cos(\phi)\sin(\theta) & \sin(\phi)\cos(\theta) \\ -\sin(\phi) & 0 \end{pmatrix}$$

At most points, this has maximal rank. However, when $\phi = 0$ we get

$$(J\psi)_{0,\theta} = \begin{pmatrix} \cos(\theta) & 0 \\ \sin(\theta) & 0 \\ 0 & 0 \end{pmatrix}$$

Which has, at most, rank 1. As we saw when we studied curves, this is a kind of "fake" singularity. There are parameterizations of the same surface which do not have a singularity at this point.

Another reason we want to restrict ourselves to regular parameterizations is the notion of dimension. Technically, the map $(u, v) \mapsto (0, 0, v)$ is a smooth parameterization, but the image is a one-dimensional object — a curve. The parameterization is obviously not regular, and so we can safely discard it.

We now seek to understand what it means for different parameterizations to define the *same* submanifold. To understand this, let's give two parameterizations of parts of the unit sphere

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

We can then try to compare them.

- We can parameterize the lower hemisphere over the unit disk $B_1(0)$ by

$$\begin{aligned} \chi : U &\longrightarrow \mathbb{R}^3 \\ x &\longmapsto (x_1, y_2, -\sqrt{1 - x_1^2 - x_2^2}) \end{aligned}$$

- We can use the inverse stereographic projection to parameterize all of S^3 except one pole

$$\begin{aligned} \sigma : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x_1, x_2) &\longrightarrow \left(\tfrac{2x_1}{1+x_1^2+x_2^2}, \tfrac{2x_2}{1+x_1^2+x_2^2}, \tfrac{x_1^2+x_2^2-1}{1+x_1^2+x_2^2}\right) \end{aligned}$$

Exercise 9. Verify that the three parameterizations given above are smooth and regular on their given domains.

If we want to compare two of these parameterizations, we need to think about the points where they describe the same points of the sphere. Let's begin by comparing χ and σ. Since σ hits every point except the north pole precisely once, and χ hits every point below the equator exactly once, These parameterizations overlap precisely on the lower hemisphere. Let's denote the lower hemisphere by

$$H_L := \{(x_1, x_2, x_3) \in S^2 \mid x_3 < 0\}.$$

The map σ is a bijection between $\mathbb{R}^2$ and $S^2 \setminus \{(0, 0, 1)\}$, with inverse

$$\sigma^{-1}(x, y, z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right)$$

The image of H_L under σ^{-1} is the unit disk,

$$B_1(0) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 < 1\}$$

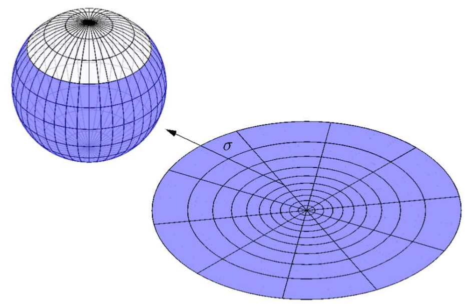
Similarly, χ is a bijection between U and H_L, with inverse

$$\chi^{-1}(x, y, z) = (x, y).$$

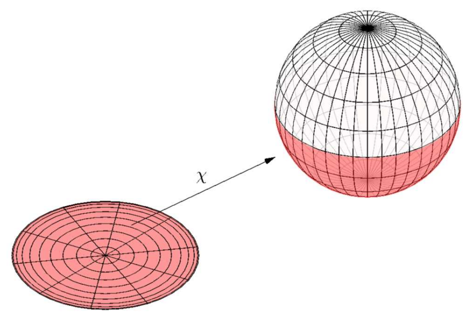
As a result, we get mutually inverse functions – the *transition functions* — from the disk to itself

$$\psi : B_1(0) \xrightarrow{\chi} H_L \xrightarrow{\sigma^{-1}} B_1(0)$$

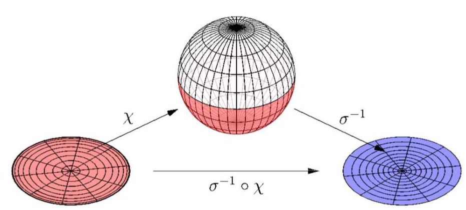
In both of the following figures, we draw the images in the disk of gridlines in the sphere, to aid in visualizing the way that each map transforms the sphere. We draw the stereographic projection on a restricted domain, since drawing the whole plane is tricky. First, the stereographic projection



and then the Cartesian parameterization



Finally, the "transition function" $\psi := \sigma^{-1} \circ \chi$ looks like



and

$$\gamma : B_1(0) \xrightarrow{\sigma} H_L \xrightarrow{\chi^{-1}} B_1(0)$$

With

$$\psi(x_1, x_2) = \left(\frac{x_1}{1 + \sqrt{1 - x_1^2 - x_2^2}}, \frac{x_2}{1 + \sqrt{1 - x_1^2 - x_2^2}} \right)$$

and

$$\gamma(x_1, x_2) = \left(\frac{2x_1}{1 + x_1^2 + x_2^2}, \frac{2x_2}{1 + x_1^2 + x_2^2} \right).$$

Both of these functions are smooth $B_1(0)$, and so ϕ and γ are mutually inverse *diffeomorphisms* from $B_1(0)$ to $B_1(0)$.

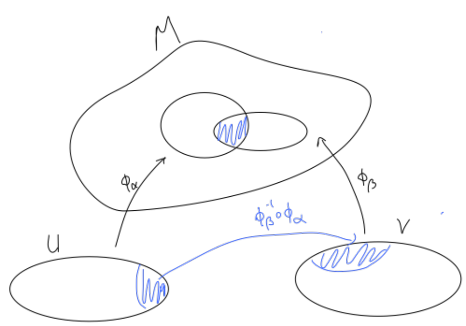
Exercise 10. Compare the parameterizations σ and χ for pieces of the sphere to the parameterization given by spherical coordinates:

$$\begin{aligned} \phi : (0, \pi) \times (0, 2\pi) &\longrightarrow \mathbb{R}^3 \\ (\phi, \theta) &\longmapsto (\sin(\phi)\cos(\theta), \sin(\phi)\sin(\theta), \cos(\phi)). \end{aligned}$$

Find the image of ϕ, and verify that the corresponding transition functions with χ and σ are both diffeomorphisms.

This motivates our definition of a *compatible* parameterizations. The point here is to guarantee that things that look smooth with respect to one chart also look smooth with respect to another. It will turn out that, under the hypotheses of smoothness and regularity, *all* charts of the same dimension are compatible, but this is something we need to prove.

The generic picture of manifold charts is very similar to the picture of the specific charts we drew for S^2:



Definition 2.17. Let $\psi : U \to \mathbb{R}^n$ and $\phi : V \to \mathbb{R}^n$ be two regular parameterizations which are *injective*. Set $M := \psi(U) \cap \phi(V)$. We say that ψ and ϕ are *compatible parameterizations* if

- The sets $\psi^{-1}(M) \subset U$ and $\phi^{-1}(M) \subset V$ are open.
- The composites

  $$\psi^{-1} \circ \phi : \phi^{-1}(M) \longrightarrow \psi^{-1}(M)$$

  and

  $$\phi^{-1} \circ \psi : \psi^{-1}(M) \longrightarrow \phi^{-1}(M)$$

  are (mutually inverse) diffeomorphisms.

We call an injective, regular parameterization a *chart*.

Remark 2.18. Notice that any two charts whose images don't intersect are compatible, using this definition.

We are now ready to define *submanifolds of* $\mathbb{R}^n$.

Definition 2.19. Let $M \subset \mathbb{R}^n$. We call a subset $U \subset M$ *open* if there is an open set $V \subset \mathbb{R}^n$ such that $V \cap M = U$.

We call the subset $M \subset \mathbb{R}^n$ together with a collection of charts a collection $\{\phi_\alpha\}_{\alpha \in I}$ of charts

$$\phi_\alpha : U_\alpha \longrightarrow \mathbb{R}^n$$

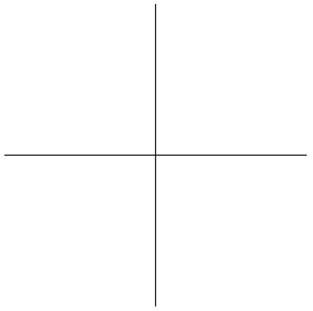
where $U_\alpha \subset \mathbb{R}^k$ an k*-dimensional submanifold* if the following conditions ae satisfied

1. For every $x \in M$, there is an open subset $V \subset M$ with $x \in V$ and an $\alpha \in I$ such that $V \subset \phi_\alpha(U_\alpha)$. (This amounts to requiring that every point x is covered by a chart[2]).

2. For any $\alpha, \beta \in I$, the charts ϕ_α and ϕ_β are compatible.

Where the charts ϕ_α and ϕ_β overlap, we call the functions $\phi_\alpha^{-1} \circ \phi_\beta$ and $\phi_\beta^{-1} \circ \phi_\alpha$ the *transition functions* between the charts.

[2] There is one more technical consideration involved, which we won't go into in detail, related to the the fact that the subset in question must be open. If we let $M \subset \mathbb{R}^2$ be the cross formed by the two axes in $\mathbb{R}^2$



We can see fairly easily that every point in M other than the origin has an open set around it and a chart to that open set. The origin, however, has a problem: any open set around it will contain a little cross, and so we can't define a chart from $\mathbb{R}$ on that open set. In principle, such points aren't too much of an issue in our setting, but we exclude them to better match the more abstract definition of manifold.

Remark 2.20. Changing between charts is effectively the same as the change of parameters from our study of curves. If $\psi : U \to \mathbb{R}^n$ is a chart, and $\phi : V \to U$ is a diffeomorphism, then $\psi \circ \phi$ is a new chart, with the same image, and $\psi \circ \phi$ is compatible with ψ. On the other hand, given two compatible charts, they are related on their intersection by such a diffeomorphism.

Remark 2.21. Since we want to study the geometry of the submanifolds $M \subset \mathbb{R}^n$, one of our major tasks will be verifying that each of our constructions does *not* depend on the choice of chart — any compatible chart should give us the same result.

Having gone through the rigamarole of defining compatibility, we are now going to step back, and show that we don't need to worry overmuch about it.

Proposition 2.22. *Let $\phi : U \to \mathbb{R}^n$ and $\psi : V \to \mathbb{R}^n$ be two smooth,regular, injective functions from open sets of $\mathbb{R}^k$ with the same image. Then the composite $\phi^{-1} \circ \psi$ is smooth, and thus is a diffeomorphism.*

Proof. We can check smoothness around any point $y \in V$ by applying the implicit function theorem to

$$U \times V \longrightarrow \mathbb{R}^n$$

$$(x, y) \longmapsto \phi(x) - \psi(y).$$

and the result follows. □

Corollary 2.23. *Any two k-dimensional charts to $\mathbb{R}^n$ are compatible.*

Remark 2.24. One might reasonably wonder why on earth we worried about compatibility in the first place. The answer is that, *a priori*, we had no reason to assume that we couldn't get smooth

Exercise 11. Let $\gamma : (a, b) \to \mathbb{R}^2$ be a smooth regular curve which is injective.

1. Show that the image of γ is a smooth 1-dimensional submanifold of $\mathbb{R}^2$.

2. Suppose that the first coordinate of γ is always positive. Define the corresponding *surface of revolution* R_γ as follows. Say that a point $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ lies in R_γ precisely when the point

$$\left(\frac{\langle (x_1, x_2, x_3), (x_1, x_2, 0) \rangle}{|(x_1, x_2, 0)}, x_3 \right)$$

lies in the image of γ. Give a geometric interpretation (in words) of this condition.

3. Show that R_γ can be equipped with the structure of a 2-dimensional submanifold of $\mathbb{R}^3$.

Notation 2.25. Throughout the rest of the course, we will often shorten the term "submanifold of $\mathbb{R}^n$" to simply *manifold*. A k-dimensional submanifold will sometimes be called simply a *k-manifold*.

3 *Smooth functions and maps*

Our next question is: can we define differentiable functions *on manifolds*? If we have a k-manifold $M \subset \mathbb{R}^n$, we can define functions $f : M \to \mathbb{R}$ as maps of the underlying sets, but it is difficult to talk about differentiability in this context. The main reason is that f may not depend on all of the coordinates of $\mathbb{R}^n$, and so some parts of the derivative may not be defined. The solution is call a function differentiable when it is differentiable in any of our charts.

Definition 2.26. Let $(M, \{(\phi_\alpha, U_\alpha)\})$ be a manifold, and let $f : M \to \mathbb{R}$ be a function.

- We say that f is C^k around a point $x \in M$ if there is a chart $\phi_\alpha : U_\alpha \to M$ such that $x \in \phi_\alpha(U_\alpha)$ and the composite map

$$f \circ \phi_\alpha$$

  is C^k

- We say that f is a C^k function if it is C^k around every point in M

Notation 2.27 (Important notational convention.)**.** Coordinates and component functions will, from here on out, be denoted with superscripts. For instance, given a function $\phi : \mathbb{R}^n \to \mathbb{R}^m$, we will denote the component functions

$$\phi = (\phi^1, \dots, \phi^m).$$

Similarly, the coordinates of $x \in \mathbb{R}^n$ will be written

$$x = (x^1, \dots, x^n).$$

This makes writing exponents somewhat trickier, but in the long run, is a very useful convention.

Example 2.28. In the following examples, we will define functions from $S^2 \subset \mathbb{R}^3$ to $\mathbb{R}$ using spherical coordinates $(\phi, \theta) \in [0, \pi] \times [0, 2\pi)$. As we will see, because the spherical coordinate chart is singular at the poles, it is not sufficient to check using spherical coordinates that a function is smooth. More generally, we always will need at least two charts to check that a function is smooth on the sphere.

1. Define a function $f : S^2 \mapsto \mathbb{R}$ by

$$f(\phi, \theta) = \phi$$

This is clearly well-defined on the whole sphere, and clearly smooth on the chart defined by spherical coordinates. However, if we want to show that the function is smooth at the north point $(0, 0, 0) \in S^2$, we need to use a different chart. Since the chart

$$\begin{aligned} \chi : \quad U &\longrightarrow \mathbb{R}^3 \\ (x^1, x^2) &\longmapsto (x^1, x^2, \sqrt{1-(x^1)^2-(x^2)^2}) \end{aligned}$$

contains the north pole, we can try to compose f with χ. Notice that the polar angle of a point $\chi(x^1, x^2)$ is given by $\phi = \arccos(\sqrt{1-(x^1)^2-(x^2)^2})$. We thus see that

$$(f \circ \chi)(x^1, x^2) = \arccos(\sqrt{1-(x^1)^2-(x^2)^2}).$$

If we consider the (smooth) path in $\mathbb{R}^2$ given by $\gamma(t) = (t, 0)$, we see that, if $f \circ \chi$ is smooth around $(0, 0)$, the function $f \circ \chi \circ \gamma$ will be smooth at the point $t = 0$. However, the derivative of $\arccos(\sqrt{1-t^2})$ is discontinuous at $t = 0$, so this f cannot even be C^1.

2. We define a function $g : S^2 \mapsto \mathbb{R}$ by

$$g(\phi, \theta) = \cos(\phi).$$

This is also well-defined on the sphere, and clearly differentiable away from the poles. Composing with the chart χ from the first example, we see that the corresponding map $g \circ \chi$ is given on the open disk $B_1(0)$ by

$$(g \circ \chi)(x^1, x^2) = \sqrt{1-(x^1)^2-(x^2)^2}$$

which is smooth on its domain. A similar argument shows that f is smooth about the south pole, and so f is a smooth function defined of S^2.

More generally, we can define smooth maps between two different manifolds.

Definition 2.29. Let M and N be manifolds. A function $f : M \to N$ is said to be C^k *at* $x \in M$ if there is a chart $\phi_\alpha : U_\alpha \to M$ of M containing x, and a chart $\psi_\beta : U_\beta \to N$ of N containing y such that $\psi_\beta^{-1} \circ f \circ \phi_\alpha$ is a smooth function.

We say that f is C^k if it is C^k at every point in M.

Example 2.30. We define a few smooth maps.

1. Let $M \subset \mathbb{R}^n$ be k-dimensional submanifold. The set $\mathbb{R}^n$ can be viewed as a submanifold of $\mathbb{R}^n$ via the single chart $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$. The inclusion map
$$\iota : M \hookrightarrow \mathbb{R}^n$$
is smooth. As a consequence, any smooth map $f : \mathbb{R}^n \to \mathbb{R}^m$ induces a smooth map $f \circ \iota : M \to \mathbb{R}^m$.

2. For any manifold M, the identity map $\mathrm{id} : M \to M$ is smooth.

Notation 2.31. We will rarely specify the composition with the charts when writing a smooth function on a chart. Suppose we have a smooth function $f : M \to \mathbb{R}$, and a chart $\phi : U \to M$. Write the coordinates on U as $x = (x^1, \ldots, x^k)$. We often abuse notation and write
$$f(x^1, \ldots, x^k)$$
to mean $f(\phi(x^1, \ldots, x^k))$.

4 *Tangents*

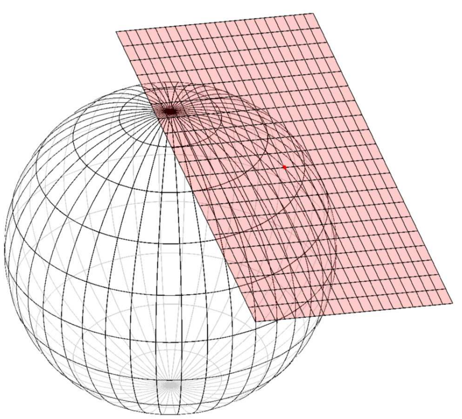
Now that we have in mind our objects of study — submanifolds of $\mathbb{R}^n$, we need to begin to develop some technology to understand them. Our initial understanding of curves came from understanding their tangent vectors. Indeed, in two dimensions, the Frenet frame is completely determined by the tangent vector. In all cases, the first curvature only really depends on the tangent vector and its derivative. We hope to generalize this perspective, and thus we need to understand tangent vectors to a submanifold.

Recall that for $U \subset \mathbb{R}^n$ an open subset, the tangent space T_xU is defined to be a copy of $\mathbb{R}^n$, and the tangent bundle TU is the set of pairs (x, v) with $x \in U$ and $v \in T_xU$. We can identify a point $(x, v) \in TU$ with a vector in $\mathbb{R}^n$ starting from x with the same direction and magnitude of as v.

Definition 2.32. Let M be a k-dimensional submanifold of $\mathbb{R}^n$, and let $\phi : U \to M$ be one of the charts of M. For any $x \in U$, we define the *tangent space* $T_{\phi(x)}M$ of M at $\phi(x)$ to be the image
$$d\phi_x(T_xU) \subset T_{\phi(x)}\mathbb{R}^n.$$
We define the *tangent bundle* of M to be the subset
$$TM := \{(x, v) \in T\mathbb{R}^n \mid v \in T_xM\} \subset T\mathbb{R}^n.$$

We can view the vectors in $T_{\phi(x)}M$ as tangent vectors to curves in M as follows. Consider any vector $v \in T_xU$. This uniquely defines a straight-line path
$$\gamma(t) = x + vt$$
in U. The path $\phi \circ \gamma$ is then a path *in* $M \subset \mathbb{R}^n$, and the tangent vector of this curve is
$$D\phi_x(\gamma'(0)) = D\phi_x(v)$$

Let's consider tangent spaces on the upper hemisphere of S^1. We'll use the now-familiar chart
$$\chi : \begin{array}{ccc} U & \longrightarrow & \mathbb{R}^3 \\ (x^1, x^2) & \longmapsto & (x^1, x^2, \sqrt{1-(x^1)^2-(x^2)^2}) \end{array}$$
The Jacobian of this map is
$$J\chi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{-x^1}{\sqrt{1-(x^1)^2-(x^2)^2}} & \frac{-x^2}{\sqrt{1-(x^1)^2-(x^2)^2}} \end{pmatrix}$$
and so, for any point $p = \chi(x^1, x^2)$ in the upper hemisphere, the tangent plane is spanned by the vectors
$$\left(1, 0, \frac{-x^1}{\sqrt{1-(x^1)^2-(x^2)^2}}\right)$$
and
$$\left(0, 1, \frac{-x^2}{\sqrt{1-(x^1)^2-(x^2)^2}}\right)$$
Viewing this plane as a subspace of $\mathbb{R}^3$, we obtain.



As a result, every vector in the image of $D\phi_x(v)$ can be realized as the tangent vector to a curve in M.

On the other hand, given a smooth curve γ in $\phi(U) \subset M$ which passes through $\phi(x)$ at $t = 0$, we can write

$$\frac{d\gamma}{dt} = \frac{d}{dt}\left(\phi \circ (\phi^{-1} \circ \gamma)\right) = D\phi_x\left(\frac{d(\phi^{-1} \circ \gamma)}{dt}(0)\right)$$

so that the tangent vector to γ at $t = 0$ is in the image of $D\phi_x$.

Having now defined the tangent bundle, and partially understood the definition, we need to see that it does *not* depend on the choice of chart.

Lemma 2.33. *Let $M \subset \mathbb{R}^n$ be a k-submanifold of $\mathbb{R}^n$, let $\phi : U \to M$ be a chart of M, and let $\psi : V \to U$ be a diffeomorphism for $V \subset \mathbb{R}^k$. Then for any $x \in U$,*

$$d\phi_x(T_xU) = d(\phi \circ \psi)_{\psi^{-1}(x)}(T_xV).$$

Proof. This follows immediately from the fact that

$$d(\phi \circ \psi)_{\psi^{-1}(x)} = d\phi_x \circ d\psi_{\psi^{-1}(x)}$$

and, since ψ is a diffeomorphism, $d\psi_{\psi^{-1}(x)}$ is an isomorphism of vector spaces. □

Our next observation is subtler. For a smooth k-submanifold $M \subset \mathbb{R}^n$, the tangent bundle

$$TM \subset T\mathbb{R}^n$$

can be identified with a subset of $\mathbb{R}^{2n}$ under the identification $T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$.

Proposition 2.34. *Let $(M, \{(\phi_\alpha, U_\alpha)\})$ be a smooth k-submanifold of $\mathbb{R}^n$. Then $TM \subset \mathbb{R}^{2n}$ is a smooth $2k$-submanifold.*

Proof. We will show that the maps

$$d\phi_\alpha : TU_\alpha \cong U_\alpha \times \mathbb{R}^k \longrightarrow TM \subset \mathbb{R}^{2n}$$

are charts displaying a smooth structure on TM. Notice that, since ϕ_α is assumed to be smooth, each of these maps is smooth, and it is easy to see that they are injective. We need to show regularity, and show compatibility, and that the $d\phi_\alpha$ appropriately cover TM.

To see regularity, we first note that the components of $d\phi_\alpha$ can be written as

$$(d\phi_\alpha)(x, v) = \left(\phi_\alpha^1(x), \phi_\alpha^n(x), \sum_{k=1}^n \frac{\partial \phi_\alpha^1}{\partial x^k}(v^k), \ldots, \sum_{k=1}^n \frac{\partial \phi_\alpha^n}{\partial x^k}(v^k)\right)$$

So the Jacobians are

$$J(d\phi_\alpha)_{(x,v)}) = \begin{pmatrix} (J\phi_\alpha)_x & 0 \\ H & (J\phi_\alpha)_x \end{pmatrix}$$

For a matrix H comprised of sums of second partial derivatives. Since $(J\phi_\alpha)_x$ has maximal rank k, we see that the rank of this matrix is $2k$. Thus, $d\phi_\alpha$ is regular.

Next, suppose that $(x, v) \in TM$. Then there is some open subset $W \subset \mathbb{R}^n$ and ϕ_α a chart of M such that $x \in W$ and $W \cap M \subset \phi_\alpha(U_\alpha)$. But then, setting $V = W \times \mathbb{R}^n \subset \mathbb{R}^{2n}$, we have that $(x, v) \in V$, and $V \subset d\phi_\alpha(U_\alpha \times \mathbb{R}^k)$. We thus see that the charts $d\phi_\alpha$ cover TM.[3]

We do not need to show compatibility, per Corollary 2.23. However, it is an instructive exercise in manipulating differentials to do so. Let $\phi : U \to M$ be a chart, and $\gamma : V \to U$ a diffeomorphism. We then notice that $d\gamma \circ d(\gamma^{-1}) = \mathrm{id}_{TU}$ and $d(\gamma^{-1}) \circ d\gamma = \mathrm{id}_{TV}$, so that $d\gamma$ is invertible. Since γ and γ^{-1} are smooth, so are $d\gamma$ and $d\gamma^{-1}$. Since

$$d(\phi \circ \gamma) = d\phi \circ d\gamma,$$

the charts $d\phi$ and $d(\phi \circ \gamma)$ are compatible, as desired. □

Corollary 2.35. *For a k-submanifold $M \subset \mathbb{R}^n$, the projection*

$$\begin{aligned} \pi : TM &\longrightarrow M \\ (x, v) &\longmapsto x \end{aligned}$$

is a smooth map.

Proof. For any chart $\phi : U \to M$ and the corresponding chart $d\phi : TU \to TM$,, the map $\phi^{-1} \circ \pi \circ d\phi$ is given by forgetting the first k coordinates, and thus is smooth. □

4.1 Vector fields

Throughout this section, we fix a smooth k-submanifold $M \subset \mathbb{R}^n$, and denote the projection from the tangent bundle by $\pi : TM \to M$.

Definition 2.36. A *smooth vector field* on M is a smooth map $X : M \to T\mathbb{R}^n$ such that, for each $p \in M$, $X(p) \in T_p\mathbb{R}^n$. A smooth *tangent vector field* on M is a smooth map $X : M \to TM$ such that $\pi \circ X = \mathrm{id}_M$. A *normal vector field* on M is a vector field X on M such that, for all $p \in M$ and $v \in T_pM$, $X(p)$ is orthogonal to v.

On any open set $U \subset M$, we can further define a *tangent vector field* on U to be a map $X : U \to TM$ such that $X(p) \in T_pM$ for any $p \in U$.

Notation 2.37. We will often represent vector field in coordinates. Let $\phi : U \to M$ be a chart of M, and suppose that $x = (x^1, \dots, x^k)$ are the coordinates of U. We will often specify smooth vector fields on M locally as a function of the coordinates x^i, writing, e.g. $X(x^1, \dots, x^n)$ instead of the (technically more accurate) $X(\phi(x^1, \dots, x^n))$.

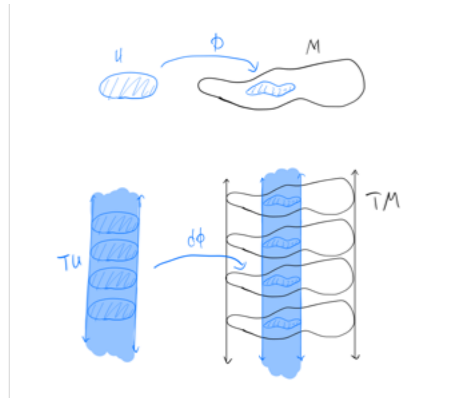
Definition 2.38. A *local basis* of TM over an open subset $U \subset M$ is a collection $X_1, \dots, X_k : U \to TM$ of tangent vector fields on U such that, for every $p \in U$, the set $(X_1(p), \dots, X_2(p))$ is a basis.

Example 2.39. There is a canonical example of a local basis. Consider $U \subset \mathbb{R}^k$, and define vector fields

$$\begin{aligned} e_{x^i} : U &\longrightarrow TU \\ x &\longmapsto (x, e_i) \end{aligned}$$

where the latter e_i is the standard basis vector in $\mathbb{R}^n \cong T_xU$. It is immediate that this is a smooth map, and so the set $(e_{x^1}, \dots, e_{x^k})$ is a local basis for U.

[3] The schematic picture we want to keep in mind here is:



Notice that this really is only schematic. The tangent bundle TM is, in general, not the same as $M \times \mathbb{R}^k$.

Construction 2.40. Let $\phi : U \to M$ be a chart of M, and let $x \in U$ and set $p = \phi(x)$. By definition, the tangent vectors $d\phi_x(e_{x^i}(x))$ form a basis of T_pM. As a result, the vector fields

$$\partial_{x^i}\phi := \partial_i\phi := d\phi \circ e_{x^i}$$

form a local basis on $\phi(U)$. Explicitly, viewing ϕ_{x^i} as a vector in $\mathbb{R}^n$, we have

$$\partial_i\phi = \frac{\partial\phi}{\partial x^i} = \left(\frac{\partial\phi^1}{\partial x^i}, \dots, \frac{\partial\phi^n}{\partial x^i}\right)$$

We call $\partial_i\phi$ a *coordinate vector field*

Lemma 2.41. *Let $\phi : U \to M$ be a chart, let $\rho : V \to U$ be a diffeomorphism, and let $\psi = \phi \circ \rho$ be the composite chart. Write x^i for the coordinates on V, and write y^i for the coordinates on U. Then*

$$\partial_i\psi = \sum_{\ell=1}^{k} \frac{\partial\rho^\ell}{\partial x^i}\partial_\ell\phi$$

Proof. We simply compute using the chain rule

$$\begin{aligned}
\partial_i\psi &= \left(\frac{\partial\psi^1}{\partial y^\ell}, \dots, \frac{\partial\psi^n}{\partial x^i}\right) \\
&= \left(\sum_{\ell=1}^{k} \frac{\partial\phi^1}{\partial y^\ell}\frac{\partial\rho^\ell}{\partial x^i}, \dots, \sum_{\ell=1}^{k} \frac{\partial\phi^n}{\partial y^\ell}\frac{\partial\rho^\ell}{\partial x^i}\right) \\
&= \sum_{\ell=1}^{k} \frac{\partial\rho^\ell}{\partial x^i}\left(\frac{\partial\phi^1}{\partial y^\ell}, \dots, \frac{\partial\phi^n}{\partial x^i}\right) \\
&= \sum_{\ell=1}^{k} \frac{\partial\rho^\ell}{\partial x^i}\partial_\ell\phi
\end{aligned}$$

completing the proof. □

Notation 2.42. Now is a good moment to introduce the *Einstein summation convention*, which allows us to shorten our written computations significantly. In this convention, whenever the symbol for an index appears twice, with one index occurring as a superscript, and the other as a subscript, we sum over that index. So, for instance

$$\frac{\partial\rho^\ell}{\partial x^i}\partial_\ell\phi$$

would be interpreted as meaning

$$\sum_{\ell=1}^{k} \frac{\partial\rho^\ell}{\partial x^i}\partial_\ell\phi.$$

As you can see, this has the great advantage of substantially shortening computations on the page, without losing specificity.

Notation 2.43. When two coordinate $\phi : U \to M$ and $\psi : V \to M$ are related by a diffeomorphism $\rho : V \to U$, we will often view the coordinate y on U as a function of the coordinate x on V via the function ρ, and view x as a function of y via ρ^{-1}. This makes our formulas much more transparent, particularly when transforming vector fields into different coordinate systems. Using this convention, our formula relating the coordinate vector fields of ϕ and ψ becomes (also using the Einstein summation convention)

$$\partial_i\psi = \frac{\partial y^\ell}{\partial x^i}\partial_\ell\phi.$$

This has the advantage of being symmetric, inasmuch as we then have

$$\partial_i\phi = \frac{\partial x^\ell}{\partial y^i}\partial_\ell\psi.$$

Given a chart $\phi : U \to M$, we can express any tangent vector field on $\phi(U)$ in terms of the coordinate vector fields $\partial_i\phi$. Let X be a vector field on $\phi(U)$. Since $d\phi_x$ is an isomorphism for every $x \in U$, we can compose with the inverse to get a smooth vector field $\overline{X}$ on U such that $d\phi \circ \overline{X} = X$. In terms of the standard basis $\{e_{x^i}\}$ corresponding to the coordinates x of U, we can write

$$\overline{X}(x) = X^i(x)e_{x^i}$$

Applying $d\phi$, we get

$$X = d\phi \circ \overline{X} = X^i(x)\partial_i\phi$$

Definition 2.44. Let $\phi : U \to M$ be a chart, and let X be a smooth tangent vector field on $\phi(U)$. We call the smooth functions X^i such that

$$X = X^i\partial_i\phi$$

the *component functions of X with respect to ϕ*.

Corollary 2.45. *Let $\phi : U \to M$ and $\psi : V \to M$ be charts with the same image, and let V be a vector field. Let x denote the coordinate on V and y the coordinate on U. Write V^i for the components of V with respect to ϕ, and $\widetilde{V}^i$ for the components of V with respect to ψ. Then*

$$V^i = \frac{dy^i}{dx^j}\widetilde{V}^j.$$

Proof. This follows immediately from Lemma 2.41. □

Example 2.46. Let's first compute explicitly the coordinate vector fields on the sphere $S^2 \subset \mathbb{R}^3$ defined by the spherical coordinate chart

$$\phi(u^1, u^2) = \left(\sin(u^2)\cos(u^1), \sin(u^2)\sin(u^1), \cos(u^2)\right)$$

A quick computation of partial derivatives tells us that

$$\partial_1\phi = \left(-\sin(u^2)\sin(u^1), \sin(u^2)\cos(u^1), 0\right)$$

and

$$\partial_2\phi = \left(\cos(u^2)\cos(u^1), \cos(u^2)\sin(u^1), -\sin(u^2)\right).$$

A priori, these vector fields are defined everywhere except for the poles. Let's check whether either of these vector fields extend to vector fields on the whole sphere.

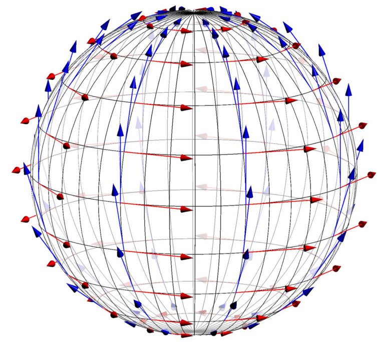
Considering $\partial_2\phi$, we notice that for any (u^1, u^2), $|\partial_2\phi(u^1, u^2)| = 1$. As such, by continuity, we see that, if $\partial_2\phi$ had a continuous extension to the poles, then the vector assigned to, e.g., the south pole $(0, 0, -1)$ would have to be a unit vector. If we consider the coordinate curves approaching the south pole defined by $u^1 = 0$ and $u^1 = \pi$, however, we see that the corresponding limits are $(1, 0, 0)$ and $(-1, 0, 0)$ respectively. Thus, $\partial_2\phi$ cannot extend to the south pole.

To check if we can extend $\partial_1\phi$ to a vector field on all of S^2, we change coordinates to the chart $\chi(x^1, x^2) = (x^1, x^2, \sqrt{1-(x^1)^2-(x^2)^2})$. In this chart, we see

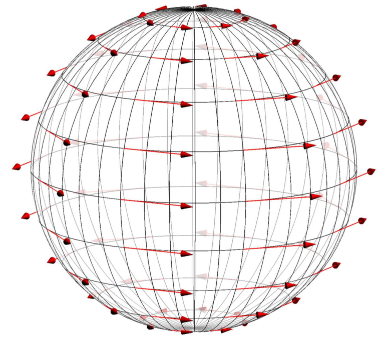
$$\partial_1\phi(x^1, x^2) = (-x^2, x^1, 0).$$

So this does extend to a smooth function, which is zero at the poles.

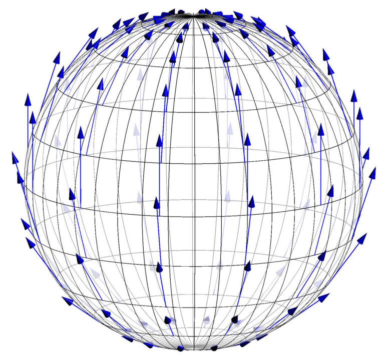
We can visualize the vector fields from Example 2.46. Notice that each vector $\partial_1\phi$ can be viewed as a tangent vector to a curve with constant u_2, and similarly for $\partial_2\phi$. Plotting both $\partial_1\phi$ and $\partial_2\phi$ over the surface of the sphere, we obtain



Notice that the vectors $\partial_1\phi(u^1, u^2)$ and $\partial_1\phi(u^1, u^2)$ form a basis at each point where they are both defined and non-zero. Separating the two fields, we have $\partial_1\phi$



and $\partial_2\phi$



3

The geometry of submanifolds

1 The first fundamental form

When we examined curves, we obtained the arc length of a curve by measuring the tangent vector using the usual inner product on $\mathbb{R}^n$. In practice, though, we did not need to know the inner product on all of $\mathbb{R}^n$. We *only* needed to know it on the tangent spaces. It is this observation which will lead to to the *first fundamental form.*

1.1 The first fundamental form, length and volume

Suppose we have a k-submanifold $M \subset \mathbb{R}^n$. Each tangent space T_pM can be canonically considered as a subset of $\mathbb{R}^n$. We can restrict the Euclidean inner product to obtain a bilinear form on T_pM. This bilinear form is enough to allow us to measure lengths and angles of *tangent vectors* to M, but forgets anything not concerned with the tangent bundle.

Definition 3.1. Let $M \subset \mathbb{R}^n$ be a k-submanifold. The *first fundamental form* is the bilinear form

$$\mathbf{I} : T_pM \times T_pM \longrightarrow \mathbb{R}$$

obtained by restricting the Euclidean inner product.

With respect to a chart $\phi : U \to M$ with coordinate $x \in U$, it is given by a symmetric matrix g with components

$$g_{i,j} = \mathbf{I}(\partial^i \phi, \partial_j \phi) = \langle \partial_i \phi, \partial_j \phi \rangle$$

Lemma 3.2. *Let $\phi : U \to M$ be a chart, and let $\psi : V \to U$ be a change of parameters. Write $\tilde{g}$ for the matrix of $\mathbf{I}$ associated to $\phi \circ \psi$ and g for the matrix associated to ϕ. Then*

$$\tilde{g} = (J\psi)^T g(J\psi).$$

Proof. This is a straightforward computation. Let $x = (x^1, \ldots, x^k)$ be the coordinate on

V. Then, using the Einstein summation convention, Notation 2.42, we have

$$\begin{aligned}
\widetilde{g}_{i,j} &:= \langle \partial_i(\phi \circ \psi), \partial_j(\phi \circ \psi) \rangle \\
&= \left\langle \frac{\partial \psi^\ell}{\partial x^i} \partial_\ell \phi, \frac{\partial \psi^r}{\partial x^j} \partial_r \phi \right\rangle \\
&= \frac{\partial \psi^\ell}{\partial x^i} \frac{\partial \psi^r}{\partial x^j} \langle \partial_\ell \phi, \partial_r \phi \rangle \\
&= \frac{\partial \psi^\ell}{\partial x^i} \frac{\partial \psi^r}{\partial x^j} g_{\ell,r} \\
&= \left[(J\psi)^T g (J\psi) \right]_{i,j}
\end{aligned}$$

as desired. □

Remark 3.3. Writing x and y for the coordinates on V and U respectively, we can rewrite the relation of Lemma 3.2 in the congenial form

$$\widetilde{g}_{i,j} = \frac{\partial y^\ell}{\partial x^i} \frac{\partial y^r}{\partial x^j} g_{\ell,r}$$

The use of the first fundamental form is that, as with curves, it allows us to compute distances and volumes.

Definition 3.4. Let $M \subset \mathbb{R}^n$ be a k-manifold. A *curve* in M is a smooth map

$$\gamma : [a,b] \longrightarrow M.$$

The length of a curve γ can be computed directly using the first fundamental form. The derivative $\gamma'(t_0) \in \mathbb{R}^n$ is identified with the tangent vector $d\gamma_{t_0}(1) \in T_{\gamma(t_0)}M$.[1] We thus have

$$L(\gamma; a, b) = \int_a^b |\gamma'(u)| du = \int_a^b \sqrt{\mathbf{I}(d\gamma_u(1), d\gamma_u(1))} du.$$

[1] Here, I'm using $1 \in T_{t_0}[a,b] \cong \mathbb{R}$ to denote a tangent vector on $[a,b]$.

More generally, though, the first fundamental form gives us a way to integrate over all or part of M.

Definition 3.5. Let $M \subset \mathbb{R}^n$ be a k-manifold, and let $f : M \to \mathbb{R}$ be a continuous function[2]. Let $\phi : U \to M$ be a chart and $A \subset \phi(U)$ a subset. We define the *integral of f over A* to be

$$\int_A f dV := \int_{\phi^{-1}(A)} (f \circ \phi)(x) \sqrt{\det(g)} dx$$

whenever the latter is defined.

[2] Technically, one only needs a weaker integrability condition here, but for our purposes, this will suffice.

Similarly, let $A \subset M$ be a subset such that

$$A = A_1 \cup A_2 \cup \cdots \cup A_\ell$$

where the A_i are disjoint, and there are charts $\phi_i : U_i \to M$ such that $A_i \subset \phi_i(U_i)$. We then define

$$\int_A f dV := \sum_{i=1}^{\ell} \int_{A_i} f dV.$$

As a special case, we define the *volume of $A \subset M$* to be the integral of the constant function 1 over that region, i.e.

$$V(A) := \int_A 1 dV.$$

Before we show that this definition is independent of the choice of charts, we first do an example.

Example 3.6. Let T^2 be the torus parameterized by

$$\begin{aligned} \phi : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (u^1, u^2) &\longmapsto \begin{pmatrix} \cos(u^1)(2+\cos(u^2)) \\ \sin(u^1)(2+\cos(u^2)) \\ \sin(u^2) \end{pmatrix} \end{aligned}$$

We can compute the tangent vectors $\partial_1\phi$ and $\partial_2\phi$ simply by taking partial derivatives:

$$\partial_1\phi = \big(-\sin(u^1)(2+\cos(u^2)), \cos(u^1)(2+\cos(u^2)), 0\big)$$

$$\partial_2\phi = \big(-\cos(u^1)\sin(u^2), -\sin(u^1)\sin(u^2), \cos(u^2)\big)$$

and so the matrix g is

$$g = \begin{pmatrix} (2+\cos(u^2))^2 & 0 \\ 0 & 1 \end{pmatrix}$$

and so

$$\sqrt{\det(g)} = 2 + \cos(u^2)$$

We can thus compute the volume (surface area) of T^2 as[3]

$$V(T^2) = \int_{[0,2\pi]\times[0,2\pi]} (2+\cos(u^2))du.$$

Applying Fubini's Theorem, we obtain

$$V(T^2) = \int_0^{2\pi}\int_0^{2\pi} (2+\cos(u^2))du^2du^1 = 8\pi^2.$$

To properly understand invariance of integrals under changes of parameterization, we need to give additional conditions on the diffeomorphims involved.

Definition 3.7. A diffeomorphism $\phi : U \to V$ between subsets of $\mathbb{R}^k$ is said to be *orientation-preserving* if $\det(J\phi) > 0$ everywhere in U, and *orientation reversing* otherwise. Notice that if U is path-connected (i.e., if there is a path in U connecting any two points in U) then ϕ can *only* be orientation-preserving or *orientation-reversing*.

Proposition 3.8. *Let $M \subset \mathbb{R}^n$ be a k-manifold, let $A \subset M$, and let $f : M \to \mathbb{R}$ be a continuous function. The value of $\int_A f dV$ is invariant under orientation-preserving changes of parameters.*

Proof. Let $\phi : U \to M$ be a chart such that $A \subset \phi(U)$, and let $\psi : V \to U$ be an orientation-preserving diffeomorphism between subsets of $\mathbb{R}^k$. Let g be the matrix of the first fundamental form with respect to ϕ, and $\widetilde{g}$ the matrix of the first fundamental form

[3] You may object that we are double-counting the integral when, e.g., $u^1 = 0$. This is, indeed, true, but since this is a curve — i.e. a 1-dimensional subset of our 2-dimensional manifold — it will turn out to contribute nothing to the integral. We will not delve into the analysis of integration, nor into measurable sets here, but the interested student can find out more in a book on measure theory.

with respect to $\phi \circ \psi$. Computing the integral with $\phi \circ \psi$ we obtain

$$\begin{aligned}\int_A f dV &:= \int_{\psi^{-1}(\phi^{-1}(A))} (f \circ \phi \circ \psi)\sqrt{\det(\widetilde{g})}dx \\ &= \int_{\psi^{-1}(\phi^{-1}(A))} (f \circ \phi \circ \psi)\sqrt{\det((J\psi)^T g(J\psi))}dx \\ &= \int_{\psi^{-1}(\phi^{-1}(A))} (f \circ \phi \circ \psi)|\det(J\psi)|\sqrt{\det(g)}dx \\ &= \int_{\psi^{-1}(\phi^{-1}(A))} (f \circ \phi \circ \psi)\det(J\psi)\sqrt{\det(g)}dx\end{aligned}$$

Taking the change-of-variables $y = \psi(x)$, and $dy = \det(J\psi)dx$,[4] we obtain

$$\int_{\phi^{-1}(A)} f \circ \phi \sqrt{\det(g)}dy$$

so that the value of the integral does not change. □

[4] We are now simply using the change-of variables rule for integration on $\mathbb{R}^n$.

Remark 3.9. Notice that, if ψ were instead orientation-reversing, we would simply have introduced a negative sign into the computation. In one dimension, this phenomenon is simply the relation

$$\int_a^b f(x)dx = -\int_b^a f(y)dy$$

using the orientation-reversing change-of-variable $y = b + a - x$.

2 *The intrinsic and the extrinsic*

A major theme for the remainder of the course will be the relation between *intrinsic* and *extrinsic* quantities associated to a submanifold M. Loosely speaking, *intrinsic* quantities are measurements which an observer living on M could measure: a person stuck on the surface of the sphere S^2, for instance, could measure the angles between paths on the sphere, distances along the surface of the sphere, areas on the sphere, or speeds of paths on the sphere, etc. But they might not be able to measure 3-dimensional quantities and relations in the ambient space $\mathbb{R}^3$.[5]

[5] The perceptual difficulties inherent in living in a space of fixed dimension are explored in Edwin Abbott's novel *Flatland*, published in 1884.

On the other hand, *extrinsic* quantities are those which can be measured using *all* of the information available to us: the submanifold M *and* the ambient space $\mathbb{R}^n$.

To formalize what we mean by intrinsic an extrinsic, we make the following definitions:

Definition 3.10. Let $M \subset \mathbb{R}^n$ be a smooth k-submanifold. We call a quantity, computation, or definition *intrinsic* if it relies only on the following data:

- The set M and the smooth regular charts $\phi : U \to M$.
- The tangent bundle TM of M.
- The first fundamental form $\mathbf{I} : T_pM \times T_pM \to \mathbb{R}$.

Otherwise, we call a quantity *extrinsic*.

To understand why we consider only these three types of data intrinsic, lets break each one down.

- The set $M \subset \mathbb{R}^n$ is nothing more than the points where an observer confined to M is allowed to be. The charts tell us how M fits together smoothly.
- The tangent bundle tells us the possible velocity vectors for paths in M, and so are measurements an observer in M should have access to.
- The first fundmental form **I** tells us how to measure (1) angles between tangent vectors/ paths in M, and (2) how to get a speed out of a velocity for a path in M (i.e., the *norm* of the velocity). Thus, our observer in M should be able to measure speeds and angles.

One of our major theorems in this course will show that certain notions of curvature are *intrinsic*. That is, to know how M curves in $\mathbb{R}^3$, we do *not* need to know the way in which M sits inside $\mathbb{R}^3$, we only need to know how to measure distances, angles, etc. *in* M.

Remark 3.11. It is worthwhile to convince yourself that the notions we defined in the previous section are all intrinsic.

We can write curves in $M \subset \mathbb{R}^n$ either by composing with a chart, or by giving a smooth map $\gamma : (a, b) \to \mathbb{R}^n$ whose image lies in M. For instance, working in the torus $T^2 \subset \mathbb{R}^3$ parameterized by

$$\phi(u^1, u^2) = \big(\cos(u^1)(2 + \cos(u^2)), \sin(u^1)(2 + \cos(u^2)), \sin(u^2)\big)$$

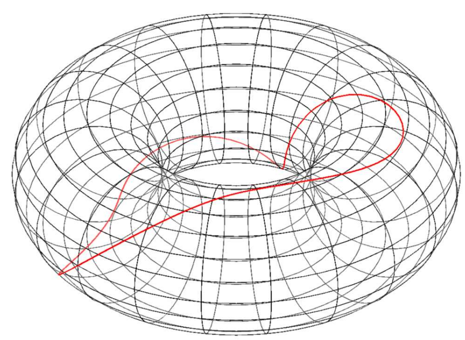
We can define a smooth curve in T^2 either by giving the map

$$\begin{aligned} \rho : [0, 2\pi] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (t, t + 0.3\sin(2t)) \end{aligned}$$

or the composite map $\gamma = \phi \circ \rho$.

$$\gamma(t) = (\cos(t)(2 + \cos(t + 0.3\sin(2t))), \sin(t)(2 + \cos(t + 0.3\sin(2t))$$

The curve described by this function can be visualized as



3 Curves and geodesics

Since we already have a good understanding of the theory of curves and their curvatures, it makes sense that we would try to first study manifolds by studying curves in them. We will not yet seek to define curvatures for manifolds, but instead will try to connect notions of length to our previous understanding of curves.

Recollection 1. A *curve* in a k-submanifold $M \subset \mathbb{R}^n$ is a smooth map

$$\gamma : [a, b] \longrightarrow M$$

We say that γ is *regular* if the tangent vector $\dot{\gamma}(t) := d\gamma_t(1) \in T_pM$ is always nonzero, or, equivalently if $\mathbf{I}(\dot{\gamma}, \dot{\gamma})$ is always nonzero.

Remark 3.12. Since the inclusion $M \subset \mathbb{R}^n$ is smooth, we can view any curve in M as a smooth curve in $\mathbb{R}^n$.

Instead of immediately studying curvature, our aim is to study *straightness*. Our guiding example in doing so will be the following dictum

> *A curve from x to y in the Euclidean space $\mathbb{R}^n$ is a straight line if and only if its length is the shortest length of a curve from x to y.*

Our aim will be use this as a *definition* of straightness in a submanifold. However, to make actual use of this, we need to understand what we mean by the *shortest length*.

Example 3.13. Consider the sphere $S^2 \subset \mathbb{R}^3$, and consider the points $y = (1, 0, 0)$ and $x = (0, 0, 1)$. We can consider two curves in S^2:

$$\begin{aligned} \gamma : [0, 3\pi/2] &\longrightarrow S^2 \subset \mathbb{R}^3 \\ t &\longmapsto (\sin(t), 0, -\cos(t)) \end{aligned}$$

and

$$\begin{aligned} \rho : [0, \pi/2] &\longrightarrow S^2 \subset \mathbb{R}^3 \\ t &\longmapsto (\sin(t), 0, \cos(t)) \end{aligned}$$

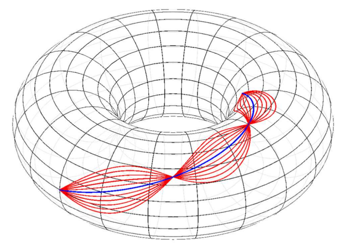
Both of these curves self-evidently go from x to y.

If we, for a moment, assume that ρ is the shortest path from x to y staying in S^2[6], then we notice something odd. By symmetry, γ is composed of the shortest path from x to $(-1, 0, 0)$, then the shortest path from $(-1, 0, 0)$ to $(0, 0, -1)$, then the shortest path from $(0, 0, -1)$ to y. However, γ is clearly longer than ρ. This suggests that there is some property *weaker* than being the shortest path from x to y that both of these curves have.

[6] It is, but proving this will take some time.

Because of examples like this, we will first focus on characterizing curves that are *critical points* of length. To make this formal, we make the following definition:

Below is a picture of a proper variation (in red) of a curve γ (in blue)in the torus.



Definition 3.14. Let $\gamma : [a, b] \to M \subset \mathbb{R}^n$ be a smooth curve in M. A *variation* of γ is a smooth map

$$H : (-c, c) \times [a, b] \longrightarrow M$$

for some $c > 0$ such that the curve $H(0, t) = \gamma(t)$. We call a variation *proper* if $H(u, a) = \gamma(a)$ and $H(u, b) = \gamma(b)$ for all $u \in (-c, c)$. We will denote the first coordinate of a variation by u, and the second by t.

The idea of studying variations of curves is that, since we have a family of curves varying according to the parameter u, we can take a derivative of the length, and check if $\gamma(t) = H(0, t)$ might represent a local minimum. To ease notation, lets write $\gamma_u(t) := H(u, t)$.

To do this, we consider the *functional* of u

$$L(u) := \int_a^b \sqrt{\mathbf{I}(\dot{\gamma}_u(t), \dot{\gamma}_u(t))}dt$$

We can then take the derivative of $L(u)$ with respect to u

$$\begin{aligned} \frac{dL}{du}(0) &= \frac{d}{du} \int_a^b \sqrt{\mathbf{I}(\dot{\gamma}_u(t), \dot{\gamma}_u(t))}dt\Big|_{u=0} \\ &= \int_a^b \frac{\partial}{\partial u} \sqrt{\mathbf{I}(\dot{\gamma}_u(t), \dot{\gamma}_u(t))}\Big|_{u=0}dt \end{aligned}$$

We then rewrite the integrand in terms of the Euclidean inner product.

$$\frac{\partial}{\partial u}\sqrt{\langle \dot{\gamma}_u(t), \dot{\gamma}_u(t)\rangle} = \frac{1}{\sqrt{\langle \dot{\gamma}_u(t), \dot{\gamma}_u(t)\rangle}} \left\langle \frac{\partial}{\partial u}\frac{\partial}{\partial t}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle$$

We make the simplifying assumption that $\gamma_0(t) = \gamma(t)$ is unit speed. Evaluating at $u = 0$ thus yields

$$\frac{\partial}{\partial u}\sqrt{\langle \dot{\gamma}_u(t), \dot{\gamma}_u(t)\rangle}\Big|_{u=0} = \left\langle \frac{\partial}{\partial u}\frac{\partial}{\partial t}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle \Big|_{u=0}$$

We can then apply the equality of mixed partials for smooth functions to get

$$\frac{dL}{du}(0) = \int_a^b \left\langle \frac{\partial}{\partial t}\frac{\partial}{\partial u}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle |_{u=0} dt$$

Finally, we notice that

$$\frac{\partial}{\partial t}\left\langle \frac{\partial}{\partial u}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle = \left\langle \frac{\partial}{\partial t}\frac{\partial}{\partial u}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle + \left\langle \frac{\partial}{\partial u}\gamma_u(t), \frac{\partial^2}{\partial t^2}\gamma_u(t) \right\rangle$$

so that

$$\frac{\partial}{\partial t}\left\langle \frac{\partial}{\partial u}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle - \left\langle \frac{\partial}{\partial u}\gamma_u(t), \frac{\partial^2}{\partial t^2}\gamma_u(t) \right\rangle = \left\langle \frac{\partial}{\partial t}\frac{\partial}{\partial u}\gamma_u(t), \frac{\partial}{\partial t}\gamma_u(t) \right\rangle$$

Applying the fundamental theorem of calculus. We thus can rewrite $\frac{dL}{dt}(0)$ as

$$\frac{dL}{dt}(0) = \left\langle \frac{\partial}{\partial u}\gamma_u(t)\big|_{u=0}, \frac{\partial}{\partial t}\gamma_0(t) \right\rangle \Bigg|_{t=a}^{b} - \int_a^b \left\langle \frac{\partial}{\partial u}\gamma_u(t)\big|_{u=0}, \frac{\partial^2}{\partial t^2}\gamma_0(t) \right\rangle dt$$

For a variation $\gamma_u(t)$ of a curve $\gamma(t) = \gamma_0(t)$, we will often denote $\frac{dL}{dt}(0)$ by δL.

Definition 3.15. We call the derivative δL the *first variation of arc length*

We call a smooth curve $\gamma : [a, b] \to M \subset \mathbb{R}^n$ a *geodesic* if, for every variation H of γ, the corresponding derivative ∂L vanishes.

Notice that, per our definition, a geodesic is not necessarily a local minimum for length, or even a local extremum, but merely a critical point.

Remark 3.16. Notice that, *a priori*, the formula for δL is not intrinsic. The second derivative $\frac{d}{dt}\gamma(t)$ need not live in the tangent space $T_{\gamma(t)}M$, and so is an extrinsic quantity.

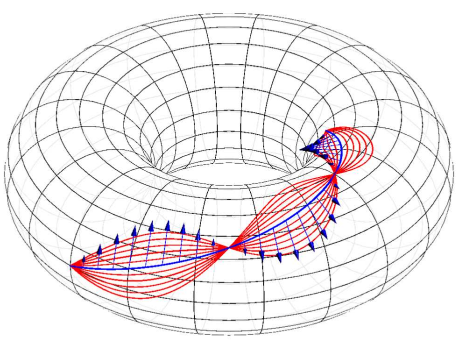
We can now notice that there is a vector field lurking in our formula for δL.

Definition 3.17. Let $H : (-c, c) \times [a, b] \to M \subset \mathbb{R}^n$ be a proper variation of a curve γ in M. The vectors

$$V(t) := \frac{\partial}{\partial u}\gamma_u(t)\Bigg|_{u=0} = dH(e_u)$$

form a tangent vector field along γ, i.e. a smooth map $V : [a, b] \to TM$ such that, for each $t \in [a, b]$, $V(t) \in T_{\gamma(t)}M$. We call this tangent vector field the *variation field* of H.

Pictorially, the variation field is, unsurprisingly, the tangent vector field which points in the direction that $H(u, t)$ changes the curve. For instance, for the proper variation in the torus we drew before, we get



We can then rewrite our formula of above in a simpler form:

$$\delta L = \langle V, \gamma'(t)\rangle \Bigg|_{t=a}^{b} - \int_a^b \left\langle V, \frac{d^2\gamma}{dt^2} \right\rangle dt$$

We thus can compute the derivative δL in terms of *only* the following information:

- The tangent vector $\gamma'(t) = d\gamma_t(1)$ to the curve γ.
- The second derivative $\frac{d^2\gamma}{dt^2}(t) \in T_{\gamma(t)}\mathbb{R}^n$.

- The tangent vector field $V(t)$ along γ associated to the variation.

Ideally, we'd like to determine if the curve γ satisfies $\delta L = 0$ for every proper variation (i.e., is a geodesic) without having to test every possible variation of γ. To have any hope of doing this, we need to understand what kind of vector fields can arise as $dH(e_u)$ for some variation H.

Proposition 3.18. *Let $\gamma : [a,b] \to M \subset \mathbb{R}^n$ be a smooth curve, and let $X : [a,b] \to TM$ be a tangent vector field along γ with $X(a) = 0 = X(b)$. Then there is a proper variation H of γ whose variation field is X.*

Proof. We will prove this proposition in the special case where the image of γ is contained in a single chart. The more general case follows similarly, but with rather more effort.

Let $\phi : U \subset \mathbb{R}^k \to M$ be a chart such that $\gamma([a,b]) \subset \phi(U)$. In particular, we can write $\gamma = \phi \circ \rho$ where $\rho : [a,b] \to U$ is a smooth curve in U. Since $d\phi : TU \to TM$ is an isomorphism at every point, there is a smooth vector field $W : [a,b] \to TU$ along the curve ρ such that $d\phi \circ W = V$.

We then define a smooth map $\psi : (-1,1) \times [a,b] \to \mathbb{R}^k$ by

$$\psi(u,t) = \rho(t) + uW(t)$$

Notice that when $u = 0$, we obtain simply the curve ρ. Notice, too, that $\psi(u,a) = \rho(a)$ and $\psi(u,b) = \rho(b)$ are constant, so that ψ is a proper variation of the curve ρ.

Choose $\epsilon > 0$ such the image of $(-\epsilon,\epsilon) \times [a,b]$ under ψ is contained in U. Then $H = \phi \circ \psi$ is a proper variation of γ. Moreover, the variation field of this variation has components

$$\left(\frac{\partial}{\partial u} (\phi(\rho(t) + uW(t)) \big|_{u=0} \right)^j = \frac{\partial \phi^j}{\partial x^i} W^i(t)$$

so that the variation field of H is $d\phi \circ W = V$, as desired. □

Definition 3.19. Let $M \subset \mathbb{R}^n$ be a k-submanifold, and let $p \in M$. Write $T_pM^\perp$ for the collection of all vectors in $T_p\mathbb{R}^n$ orthogonal to every tangent vector to M at p. We call $T_pM^\perp$ the *normal space* to M at p. We can uniquely write every vector $w \in T_p\mathbb{R}^n$ as

$$w = w^T + w^N$$

where $w^T \in T_pM$, and $w^N \in T_pM^\perp$. We call w^T the *tangential component of w* and w^N the *normal component of w*.

Given a vector field X on $U \subset M$, we obtain two new vector fields X^T and X^N by taking the tangential and normal components of X at each point. These satisfy

$$X^T(p) \in T_pM \quad \text{and} \quad X^n(p) \in T_pM^\perp$$

and $X = X^T + X^N$.

Corollary 3.20. *Let $\gamma : [a,b] \to M \subset \mathbb{R}^n$ be a smooth curve in M. Then γ is a geodesic if and only if*

$$\left(\frac{d^2\gamma}{dt^2} \right)^T (t) = 0$$

for every $t \in [a,b]$.

Proof. Suppose first that the tangential component of $\frac{d^2\gamma}{dt^2}$ is zero. We then notice that, for any proper variation H of γ, $V(a) = V(b) = 0$. Thus,

$$\delta L = -\int_a^b \left\langle V, \frac{d^2\gamma}{dt^2} \right\rangle dt$$

However, since V is a tangent vector field, and the tangential component of $\frac{d^2\gamma}{dt^2}$ vanishes, this becomes $\delta L = 0$ as desired.

On the other hand, suppose that γ is a geodesic. The vector field

$$\left(\frac{d^2\gamma}{dt^2}\right)^T$$

is a tangent vector field along γ. We can then define a tangent vector field

$$X(t) = (t-a)(b-t)\left(\frac{d^2\gamma}{dt^2}\right)^T$$

along γ, which satisfies $X(a) = 0 = X(b)$. By Proposition 3.18, there exists a variation H with X as its variation field. Applying our formula for δL to H, we thus obtain

$$\begin{aligned} \delta L &= -\int_a^b \left\langle X, \frac{d^2\gamma}{dt^2} \right\rangle dt \\ &= -\int_a^b (t-a)(b-t)\left|\left(\frac{d^2\gamma}{dt^2}\right)^T\right|^2 dt \end{aligned}$$

The integrand is continuous and always non-negative, and so since $\delta L = 0$ the integrand must vanish identically. This implies that

$$\left(\frac{d^2\gamma}{dt^2}\right)^T = 0$$

except possibly when $t = a$ or $t = b$. However, since

$$\left(\frac{d^2\gamma}{dt^2}\right)^T$$

is continuous, it must vanish identically on all of $[a, b]$, completing the proof. $\square$

Example 3.21. Let us first consider the case of Euclidean space $\mathbb{R}^n$ as a submanifold of $\mathbb{R}^n$. We wish to understand what the geodesics are. Let $\gamma : [a, b] \to \mathbb{R}^n$. Since $\mathbb{R}^n$ is its own ambient space, it is easy to see that *all* of $\frac{d^2\gamma}{dt^2}$ is tangential. Thus, γ is a geodesic if and only if $\gamma''(t) = 0$. Integrating, we thus find that there are vectors $a, v \in \mathbb{R}^n$ such that

$$\gamma(t) = a + tv$$

i.e., γ is a straight line.

Example 3.22. We now consider curves on the sphere S^2. Let $\gamma : [a, b] \to S^2$ be a geodesic, and assume without loss of generality that γ has unit speed. Since the tangent plane to S^2 at any point $p \in S^2$ is orthogonal to p (considered as a vector in $T_p\mathbb{R}^3$), we can thus notice that $\gamma''(t) = \lambda(t)\gamma(t)$, i.e. the second derivative of γ is always a scalar multiple of γ.

Since $\gamma(t)$ is always a unit vector, we see that $\gamma'(t)$ is orthogonal to $\gamma(t)$, so γ is a Frenet curve[7] , and the Frenet vectors $e_1(t)$ and $e_2(t)$ are $\gamma'(t)$ and $\gamma(t)$, respectively. The third Frenet vector is

$$e_3(t) = e_1(t) \times e_2(t)$$

and its derivative is

$$\begin{aligned} e_3'(t) &= e_1'(t) \times e_2(t) + e_1(t) \times e_2'(t) \\ &= \gamma''(t) \times \gamma(t) + \gamma'(t) \times \gamma'(t) \\ &= 0 \end{aligned}$$

We thus see that e_3 is constant, and so the second curvature $\kappa_2(t)$ vanishes. Thus, γ is contained entirely in a plane in $\mathbb{R}^3$. By symmetry, we may assume that this plane is a plane with constant z-coordinate, and we may exclude the cases of the poles, since we are assuming that the image of γ is more than simply a point. Hence, we may write γ in spherical coordinates as

$$\gamma(t) = (\sin(\alpha)\cos(u(t)), \sin(\alpha)\sin(u(t)), \cos(\alpha))$$

The derivatives of this expression are

$$\begin{aligned} \gamma'(t) &= (-\sin(\alpha)\sin(u(t))u'(t), \sin(\alpha)\cos(u(t))u'(t), 0) \\ \gamma''(t) &= \left(-\sin(\alpha)\cos(u(t))(u'(t))^2, -\sin(\alpha)\sin(u(t))(u'(t))^2, 0\right) \end{aligned}$$

Since we assume that γ has unit speed, this means that

$$(u'(t))^2 \sin^2(\alpha) = 1$$

that is, $u'(t)$ is constant.

We thus see that

$$|\gamma''(t)|^2 = \sin^2(\alpha)(u'(t))^4$$

$$\langle \gamma(t), \gamma''(t) \rangle = -\sin^2(\alpha)(u'(t))^2$$

For $\gamma''(t)$ to have no tangential component, we must thus have

$$\sin^2(\alpha)(u'(t))^2 = \sin(\alpha)(u'(t))^2$$

i.e.,

$$\sin(\alpha) = \sin^2(\alpha)$$

Since, by assumption $0 < \alpha < \pi$, this means that $\alpha = \pi/2$ and $u(t) = t$. We thus see that the only geodesics are arcs of great circles on the sphere.

[7] Technically, we must show that $\gamma''(t)$ is non-zero. To do this, consider $t_0 \in [a, b]$. By symmetry, we may assume that $\gamma(t_0) = (0, 0, 1)$, and so assume that the restriction of γ to $[t_0 - \epsilon, t_0 + \epsilon]$ lies in the image of the upper hemisphere coordinate chart. Without loss of generality, we can thus write

$$\gamma(t) = \left(x(t), y(t), \sqrt{1 - x(t)^2 - y(t)^2}\right)$$

The condition that $|\gamma'(t)| = 1$ gives us

$$\frac{-2x(t)y(t)x'(t)y'(t) + \left(y(t)^2 - 1\right)x'(t)^2 + \left(x(t)^2 - 1\right)y'(t)^2}{x(t)^2 + y(t)^2 - 1} = 1$$

On the other hand, we can compute $\gamma''(t)$, yielding

$$\left(x''(t), y''(t), \frac{-2x(t)x''(t) - 2x'(t)^2 - 2y(t)y''(t) - 2y'(t)^2}{2\sqrt{-x(t)^2 - y(t)^2 + 1}} - \frac{(-2x(t)x'(t) - 2y(t)y'(t))^2}{4\left(-x(t)^2 - y(t)^2 + 1\right)^{3/2}}\right)$$

Suppose that $x''(t_0)$ and $y''(t_0)$ are zero, then the final component is

$$\frac{-2x(t_0)y(t_0)x'(t_0)y'(t_0) + \left(y(t_0)^2 - 1\right)x'(t_0)^2 + \left(x(t_0)^2 - 1\right)y'(t}{(-x(t_0)^2 - y(t_0)^2 + 1)^{3/2}}$$

Which, by the condition that $|\gamma'(t_0)| = 1$ is simply

$$-\frac{1}{\sqrt{-x(t_0)^2 - y(t_0)^2 + 1}}$$

which is non-zero.

4 *Directional derivatives*

We will now briefly digress, and discuss how we can think about tangent fields as "directions to take derivatives on a submanifold". We will throughout work in the image of a chart $\phi : U \to M$ The basic idea here is that, given a smooth function $f : M \to \mathbb{R}$, the vector tangent vector $df_p(\partial_i \phi) \in T_{f(p)}\mathbb{R} \cong \mathbb{R}$ can be computed as

$$df_p(\partial_i \phi) = \frac{\partial f}{\partial x^i}$$

or, more formally, as $\frac{\partial}{\partial x^i}(f \circ \phi)$. Moreover, since the partial derivative is defined as the derivative of a function composed with a coordinate curve, we can view $\frac{\partial f}{\partial x^i}$ as a *directional derivative* in the x^i-direction on M.

To generalize this, let

$$\gamma : (-a, a) \longrightarrow U$$

be a curve in U. Write $q = \gamma(0)$ and $p = \phi(q)$. Then we can take a derivative of f along the curve γ exactly as we would along along one of the coordinate curves:

$$\frac{d}{dt}(f \circ \phi \circ \gamma)|_{t=0} = \frac{\partial (f \circ \phi)}{\partial x^i} \frac{d\gamma^i}{dt}|_{t=0} = \frac{d\gamma^i}{dt}|_{t=0} \frac{\partial f}{\partial x^i}(q)$$

Since this only depends on the tangent vector $\gamma'(0)$ and the point p, for any tangent vector

$$V = V^i (\partial_i \phi)(p)$$

we can define the *directional derivative* in the V-direction at p to be

$$V(f) = V^i \frac{\partial f}{\partial x^i}(p).$$

More generally still, we can extend this construction to vector fields:

Definition 3.23. Let $f : M \to \mathbb{R}$ be a smooth function, and let $X = X^i \partial_i \phi$ be a smooth tangent vector field on $\phi(U)$. We define a new smooth function on $\phi(U)$

$$X(f) := X^i \frac{\partial f}{\partial x^i}.$$

This is sometimes called the *action of X on f*.

Notice that, if we change coordinates to a new coordinate y, the function $X(f)$ is unchanged. The components of X with respect to the coordinates y are

$$\overline{X}^i = \frac{\partial y^i}{\partial x^j} X^j$$

and so

$$\overline{X^i} \frac{\partial f}{\partial y^i} = X^j \frac{\partial y^i}{\partial x^j} \frac{\partial f}{\partial y^i} = X^j \frac{\partial f}{\partial x^j}$$

as expected.

Exercise 12. Show that, for tangent vector fields X, Y on $\phi(U)$ and smooth functions $f, g : M \to \mathbb{R}$,

- $X(f+g) = X(f) + X(g)$
- $X(fg) = X(f)g + fX(g)$
- $(X+Y)(f) = X(f) + Y(f)$

We now might wonder what might happen if we try to take $X(Y(f))$ — A kind of directional second derivative. We can compute

$$\begin{aligned} X(Y(f)) &= X\left(Y^i \frac{\partial f}{\partial x^i}\right) \\ &= X^j \frac{\partial Y^i}{\partial x^j}\frac{\partial f}{\partial x^i} + X^j Y^i \frac{\partial^2 f}{\partial x^j \partial x^i} \end{aligned}$$

This has two different components — one which looks like a second derivative, and one which looks like a directional derivative. We will consider second derivatives more generally in the next section, but for the time being we are interested in trying to extract the part that looks like a directional derivative. We could simply forget about the tangential component, but this would not give us a vector field independent of the choice of coordinates, for instance, we can compute

$$\begin{aligned} \overline{X}^j \frac{\partial \overline{Y}^i}{\partial y^j}\frac{\partial f}{\partial y^i} &= X^k \frac{\partial y^j}{\partial x^k}\frac{\partial}{\partial y^j}\left(Y^r \frac{\partial y^i}{\partial x^r}\right)\frac{\partial f}{\partial y^i} \\ &= X^k \frac{\partial y^j}{\partial x^k}\left(\frac{\partial Y^r}{\partial x^m}\frac{\partial x^m}{\partial y^j}\frac{\partial y^i}{\partial x^r} + Y^r \frac{\partial y^i}{\partial x^m \partial x^r}\frac{\partial x^m}{\partial y^j}\right)\frac{\partial f}{\partial y^i} \\ &= X^k \frac{\partial y^j}{\partial x^k}\frac{\partial x^m}{\partial y^j}\frac{\partial Y^r}{\partial x^m}\frac{\partial y^i}{\partial x^r}\frac{\partial f}{\partial y^i} + X^k \frac{\partial y^j}{\partial x^k}\frac{\partial x^m}{\partial y^j} Y^r \frac{\partial^2 y^i}{\partial x^m \partial x^r}\frac{\partial f}{\partial y^i} \\ &= X^k \frac{\partial Y^r}{\partial x^k}\frac{\partial f}{\partial x^r} + X^k Y^r \frac{\partial^2 y^i}{\partial x^k \partial x^r}\frac{\partial f}{\partial y^i} \end{aligned}$$

Since this is not the same result as directly computing with respect to the coordinate x, this is not a well-defined vector field.

However, we can notice that the "error term" is symmetric in X and Y. Thus, we can make the following definition.

Definition 3.24. Let $X = X^i \partial_i \phi$ and $Y = Y^i \partial_i \phi$ be tangent vector fields on $\phi(U)$. The *Lie Bracket* of X and Y is the tangent vector field

$$[X, Y] := \left(X^k \frac{\partial Y^i}{\partial x^k} - Y^k \frac{\partial X^i}{\partial x^k}\right)\partial_i \phi$$

on $\phi(U)$.

Exercise 13. Show that, for a smooth function $f : M \to \mathbb{R}$,

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

5 *The covariant derivative*

We now understand that the (not obviously intrinsic) quantity

$$(\gamma'')^T$$

is a significant part of understanding the geometry of curves in a manifold. Our next goal is to try to understand this as a kind of derivative in its own right — the *covariant derivative*. On a heuristic level, this will measure how much a curve γ is bending in the tangent plane. However, there is a more general notion of covariant derivative, which takes as input a general vector field.

Lets consider a coordinate chart $\phi : U \to M$, with coordinate x on U and a tangent vector field

$$V = V^i \frac{\partial \phi}{\partial x^i}$$

on $\phi(U)$. Differentiating V with respect to x^j, we find

$$\frac{\partial V}{\partial x^j} = \frac{\partial V^i}{\partial x^j} \frac{\partial \phi}{\partial x^i} + v^i \frac{\partial^2 \phi}{\partial x^j \partial x^i}$$

This derivative is not strictly tangential — the first component is, but linear combination of mixed partials

$$v^i \frac{\partial^2 \phi}{\partial x^j \partial x^i}$$

may not be. We want to consider only the tangential component of this derivative.

At any point $p \in \phi(U)$, we can extend the basis $\partial_1 \phi(p), \ldots, \partial_k \phi(p)$ to a basis

$$(\partial_1 \phi(p), \ldots, \partial_k \phi(p), \beta_{k+1}, \ldots, \beta_n)$$

such that each of the β_j are orthogonal to each of the $\partial_i \phi$. We can thus write

$$\frac{\partial^2 \phi}{\partial x^j \partial x^i} = \Gamma^k_{i,j}(p) \partial_k \phi(p) + m^i \beta.$$

The first term of this expression is the (unique) tangential component of $\frac{\partial^2 \phi}{\partial x^i \partial x^j}$.

Definition 3.25. Let $\phi : U \to M$ be a coordinate chart. The *Christoffel symbols* with respect to ϕ are the unique functions $\Gamma^k_{i,j} : \phi(U) \to \mathbb{R}$ such that

$$\frac{\partial^2 \phi}{\partial x^i \partial x^j} = \Gamma^k_{i,j}(p) \partial_k \phi(p) + u.$$

where u is a vector orthogonal to $T_p M$.

The *covariant derivative* of a vector field $V = V^i \partial_i \phi$ in the $\partial_j \phi$-direction is the tangential component of $\frac{\partial V}{\partial x^j}$, that is

$$\nabla_j V = \frac{\partial V^i}{\partial x^j} \partial_i \phi + V^i \Gamma^k_{j,i} \partial_k \phi$$

More generally, given a tangent vector field $X = X^i \partial_i \phi$ on U, the *covariant derivative of V along X* is the tangent vector field

$$\nabla_X V = X^j \nabla_j V = X^j \frac{\partial V^i}{\partial x^j} \partial_i \phi + X^j V^i \Gamma^k_{j,i} \partial_k \phi$$

Before proving results about the covariant derivative, let's connect it back to geodesics. Let $\gamma = \phi \circ \rho$, where ρ is a curve in U, and consider the vector field

$$\gamma'(t) = \frac{d\rho^i}{dt}\partial_i\phi = V^i\partial_i\phi.$$

The second derivative can then be written as

$$\frac{d^2\gamma}{dt^2} = \frac{dV^i}{dt}\partial_i\phi + V^i\frac{\partial^2\phi}{\partial x^i\partial x^j}\frac{d\rho^j}{dt}$$

the tangential component is thus

$$\left(\frac{d^2\gamma}{dt^2}\right)^T = \frac{dV^i}{dt}\partial_i\phi + V^iV^j\Gamma^k_{i,j}\partial_k\phi$$

Supposing we can extend the V^j to functions on an open neighborhood of our curve, we can rewrite the first term in terms of $V^i(\rho(t))$, and thus obtain

$$\begin{aligned}\left(\frac{d^2\gamma}{dt^2}\right)^T &= \frac{\partial V^i}{\partial x^\ell}\frac{d\rho^\ell}{dt}\partial_i\phi + V^iV^j\Gamma^k_{i,j}\partial_k\phi \\ &= \frac{\partial V^i}{\partial x^\ell}V^\ell\partial_i\phi + V^iV^j\Gamma^k_{i,j}\partial_k\phi \\ &= \nabla_{\gamma'}\gamma'\end{aligned}$$

Now, this last step is not fully justified — there is no reason to assume that the V^i can be extended to smooth functions beyond the image of γ.[8] We instead use this as a *definition* for the covariant derivative of a tangent vector field V along a curve γ:

[8] This is, however, true, though we will not prove it in this course.

Definition 3.26. Let $\gamma : [a, b] \to M$ be a smooth curve, with corresponding tangent vector field γ' along γ. Let V be a vector field along γ. The *covariant derivative of V along γ* is the tangent vector field

$$\nabla_{\gamma'}V = \frac{dV^i}{dt}\partial_i\phi + V^i\frac{d\rho^j}{dt}\Gamma^k_{i,j}\partial_k\phi.$$

With these definitions in hand, we now prove some properties of the covariant derivative and the Christoffel symbols.

Lemma 3.27. *The Christofel symbols $\Gamma^k_{i,j}$ are intrinsic.*

Proof. We work in a chart $\phi : U \to M$. Denote the matrix of the first fundamental form with respect to the coordinates x on U by g. Then notice that

$$\begin{aligned}\frac{\partial}{\partial x^k}g_{i,j} &= \frac{\partial}{\partial x^k}\langle\partial_i\phi, \partial_j\phi\rangle \\ &= \left\langle\frac{\partial^2\phi}{\partial x^k\partial x^i}, \partial_j\phi\right\rangle + \left\langle\partial_i\phi, \frac{\partial^2\phi}{\partial x^k\partial x^j}\right\rangle\end{aligned}$$

Since taking inner products with $\partial_j\phi$ only detects the tangential components of vector fields, we then find

$$\begin{aligned}\frac{\partial}{\partial x^k}\langle\partial_i\phi, \partial_j\phi\rangle &= \left\langle\frac{\partial^2\phi}{\partial x^k\partial x^i}, \partial_j\phi\right\rangle + \left\langle\partial_i\phi, \frac{\partial^2\phi}{\partial x^k\partial x^j}\right\rangle \\ &= \Gamma^\ell_{k,i}\langle\partial_\ell\phi, \partial_j\phi\rangle + \Gamma^\ell_{k,j}\langle\partial_\ell\phi, \partial_i\phi\rangle \\ &= \Gamma^\ell_{k,i}g_{\ell,j} + \Gamma^\ell_{k,j}g_{\ell,i}\end{aligned}$$

Noting that, by the equality of mixed partials, $\Gamma^\ell_{i,j} = \Gamma^\ell_{j,i}$, we then see that

$$2g_{\ell,k}\Gamma^\ell_{i,j} = \frac{\partial g_{j,k}}{\partial x^i} + \frac{\partial g_{k,i}}{\partial x^j} - \frac{\partial g_{i,j}}{\partial x^k}$$

Since g is an invertible matrix, we can write the components of its inverse as $g^{i,j}$. We thus find that

$$\boxed{\Gamma^r_{i,j} = \frac{1}{2}g^{r,k}\left(\frac{\partial g_{j,k}}{\partial x^i} + \frac{\partial g_{k,i}}{\partial x^j} - \frac{\partial g_{i,j}}{\partial x^k}\right)}$$

This means that the Christoffel symbols can be expressed solely in terms of derivatives of the first fundamental form's matrix and local coordinates. In other words, the Christoffel symbols are intrinsic. □

Corollary 3.28. *The covariant derivative $\nabla_X Y$ is intrinsic.*

Remark 3.29. Our key motivation for defining the covariant derivative is to understand geodesics. In our new notation, a curve $\gamma : [a, b] \to M$ is a geodesic if and only if

$$\boxed{\nabla_{\gamma'}\gamma' = 0.}$$

Thus, we can understand our analogue of "straight lines" by studying the covariant derivative.

We now prove that our definition of the covariant derivative is independent of our choice of coordinates.

Lemma 3.30. *Let $\phi : U \to M$ and $\psi : V \to M$ be two charts with the same image, and coordinates y and x respectively. Write $\widetilde{\Gamma}^k_{i,j}$ and $\Gamma^k_{i,j}$ for the Christoffel symbols of ψ and ϕ respectively. Then*

$$\boxed{\widetilde{\Gamma}^k_{i,j} = \frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i}\frac{\partial x^k}{\partial y^m}\Gamma^m_{\ell,r} + \frac{\partial x^k}{\partial y^m}\frac{\partial^2 y^m}{\partial x^i \partial x^j}}$$

Proof. Once again, this is simply a computation. We compute

$$\begin{aligned}
\frac{\partial^2 \psi}{\partial x^i \partial x^j} &= \frac{\partial^2 \phi}{\partial y^\ell \partial y^r}\frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i} + \frac{\partial \phi}{\partial y^\ell}\frac{\partial^2 y^\ell}{\partial x^i \partial x^j} \\
&= \frac{\partial^2 \phi}{\partial y^\ell \partial y^r}\frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i} + \frac{\partial^2 y^\ell}{\partial x^i \partial x^j}\partial_\ell \phi
\end{aligned}$$

taking the tangential component, we obtain

$$\left(\frac{\partial^2 \psi}{\partial x^i \partial x^j}\right)^T = \frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i}\Gamma^k_{\ell,r}\partial_k \phi + \frac{\partial^2 y^k}{\partial x^i \partial x^j}\partial_k \phi$$

and applying Lemma 2.41, we see that

$$\left(\frac{\partial^2 \psi}{\partial x^i \partial x^j}\right)^T = \frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i}\Gamma^k_{\ell,r}\frac{\partial x^k}{\partial y^m}\partial_k \psi + \frac{\partial^2 y^m}{\partial x^i \partial x^j}\frac{\partial x^k}{y^m}\partial_k \psi$$

so that the coefficient of $\partial_k \psi$ is

$$\widetilde{\Gamma}^k_{i,j} = \frac{\partial y^\ell}{\partial x^j}\frac{\partial y^r}{\partial x^i}\frac{\partial x^k}{\partial y^m}\Gamma^m_{\ell,r} + \frac{\partial x^k}{\partial y^m}\frac{\partial^2 y^m}{\partial x^i \partial x^j}$$

as desired. □

Lemma 3.31. *Let V and X be vector field on an open subset of M. Then the vector field $\nabla_X V$ is independent of the choice of chart.*

Proof. Exercise. □

6 *Parallel vector fields*

Our next goal is to understand how we make use of the covariant derivative. The basic idea is that ∇_i is the analogue of the usual partial derivative operators that is adapted to work on tangent vector fields. Given $X = X^i \partial_i \phi$, we can view X as a differential operator on smooth functions, and the analogous differential operator on vector fields is ∇_X.

For one example of the way in which this is true, we will connect partial derivatives to covariant derivatives via a form of the product rule. Let $X = X^i \partial_i \phi$, $V = V^i \partial_i \phi$, and $W = W^i \partial_i \phi$ be tangent vector fields on a k-submanifold $M \subset \mathbb{R}^n$. We can define a smooth function

$$\begin{aligned} \langle V, W\rangle : M &\longrightarrow \mathbb{R} \\ p &\longmapsto \mathbf{I}(V_p, W_p) = \langle V_p, W_p\rangle \end{aligned}$$

via the first fundamental form. We can take the X-derivative of this function using the usual product rule for functions into $\mathbb{R}^n$:

$$\begin{aligned} X(\langle V, W\rangle) &= X^i \frac{\partial}{\partial x^i}\langle V, W\rangle \\ &= X^i \left(\left\langle \frac{\partial}{\partial x^i} V, W \right\rangle + \left\langle V, \frac{\partial}{\partial x^i} W \right\rangle \right) \\ &= \left\langle X^i \frac{\partial}{\partial x^i} V, W \right\rangle + \left\langle V, X^i \frac{\partial}{\partial x^i} W \right\rangle \\ &= \left\langle \left(X^i \frac{\partial}{\partial x^i} V \right)^T, W \right\rangle + \left\langle V, \left(X^i \frac{\partial}{\partial x^i} W \right)^T \right\rangle \\ &= \langle \nabla_X V, W\rangle + \langle V, \nabla_X W\rangle \end{aligned}$$

So the product rule still holds when we replace the partial derivative with the covariant derivative, we thus get an expression in terms of only intrinsic quantities

$$X(\langle V, W\rangle) = \langle \nabla_X V, W\rangle + \langle V, \nabla_X W\rangle$$

Analogously, if $\gamma : [a, b] \to M$ is a smooth curve, we have

$$\frac{d}{dt}(\langle V_{\gamma(t)}, W_{\gamma(t)}\rangle) = \langle \nabla_{\gamma'} V, W\rangle + \langle V, \nabla_{\gamma'} W\rangle$$

It is this latter equation which motivates our next definition.

Definition 3.32. We say that a tangent vector field V is *parallel* along a curve γ when

$$\nabla_{\gamma'} V = 0.$$

The basic idea here is that, for any two vector fields along γ, then

$$\frac{d}{dt}(\langle V_{\gamma(t)}, W_{\gamma(t)}\rangle) = \langle \nabla_{\gamma'} V, W\rangle + \langle V, \nabla_{\gamma'} W\rangle = 0$$

so that the length of the vectors is constant along γ, and the angle between them is constant along γ.

Intuitively, we would expect that given any tangent vector $X_{\gamma(a)}$ in $T_{\gamma(a)}M$, we can choose a parallel vector field along γ by "dragging $X_{\gamma(a)}$ along γ". More rigorously, we have the following.

Lemma 3.33. *Let $\gamma : [a, b] \to M$ be a smooth curve, and let $v \in T_{\gamma(a)}M$ be a tangent vector. There is a unique parallel vector field Y along γ such that $Y_{\gamma(a)} = v$.*

Proof. We expand the equation $\nabla_{\gamma'} Y = 0$, yielding

$$\frac{dY^i}{dt}\partial_i\phi + Y^i\frac{dx^j}{dt}\Gamma^{\ell}_{i,j}\partial_\ell\phi = 0.$$

We can collect the coefficients of $\partial_\ell\phi$ to get an equivalent system of equations:

$$\frac{dY^\ell}{dt} + Y^i\frac{dx^j}{dt}\Gamma^{\ell}_{i,j} = 0 \qquad 1 \leq \ell \leq k$$

This is a linear first-order system of ODEs with smooth coefficients, so that subject to the initial conditions $Y^i(\gamma(a)) = v$ there is a unique solution on $[a, b]$. ☐

Definition 3.34. Given γ and v as in the lemma, we call the resulting parallel vector field Y the *parallel transport* of v along γ.

Example 3.35. We compute a parallel transport on the sphere $S^2 \subset \mathbb{R}^3$, using polar coordinates

$$\phi(u^1, u^2) = (\cos(u^1)\cos(u^2), \cos(u^1)\sin(u^2), \sin(u^1)).$$

We use the curve

$$\gamma(t) = \left(\frac{1}{\sqrt{3}}\cos(t), \frac{1}{\sqrt{3}}\sin(t), \frac{2}{\sqrt{3}}\right)$$

and note that $\gamma = \phi \circ \rho$, where $\rho(t) = \big(\arccos(1/\sqrt{3}), t\big)$.

To write down the desired ODEs, we first compute the first fundamental form

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \cos^2(u^1) \end{pmatrix}$$

and the Christoffel symbols

$$\Gamma^1_{i,j} = \begin{pmatrix} 0 & 0 \\ 0 & \cos(u^1)\sin(u^1) \end{pmatrix} \qquad \Gamma^2_{i,j} = \begin{pmatrix} 0 & -\tan(u^1) \\ -\tan(u^1) & 0 \end{pmatrix}$$

We can also compute the derivatives $\frac{du^1}{dt} = 0$ and $\frac{du^2}{dt} = 1$. Let us parallel transport the vector

$$v = 2\partial_1\phi + \partial_2\phi.$$

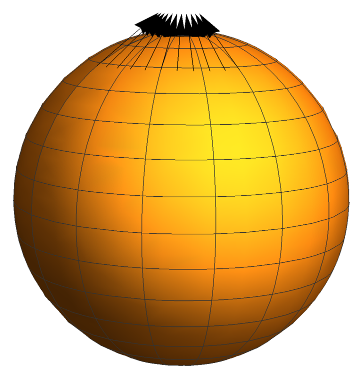
Our initial value problem for Y is thus

$$\begin{aligned} \frac{dY^1}{dt} + \frac{\sqrt{2}}{3}Y^2 &= 0 \\ \frac{dY^2}{dt} + \sqrt{2}Y^1 &= 0 \\ Y^1(0) &= 2 \\ Y^2(0) &= 1 \end{aligned}$$

and so the coefficients of the parallel transport are[9]

$$\begin{aligned} Y^1(t) &= 2\cos\left(\frac{\sqrt{2}}{3}t\right) - \frac{\sqrt{3}}{3}\sin\left(\sqrt{\frac{2}{3}}t\right) \\ Y^2(t) &= 1\cos\left(\frac{\sqrt{2}}{3}t\right) + 2\sqrt{3}\sin\left(\sqrt{\frac{2}{3}}t\right). \end{aligned}$$

[9] Pictorially, the parallel transport of v along γ looks something like



where I have rescaled the vectors $Y_{\gamma(t)}$ to have length $\frac{1}{2}$, for ease of viewing.

4

Curvature

We now come to the heart of the course: the definition of curvature for a submanifold of $\mathbb{R}^n$. The remainder of the course will be a process of specialization: first to $n-1$-dimensional submanifolds of $\mathbb{R}^n$, and eventually to *surfaces* — 2-dimensional submanifolds of $\mathbb{R}^3$. We will work in the following special case:

Definition 4.1. A *hypersurface* in $\mathbb{R}^{n+1}$ is an n-dimensional submanifold of $\mathbb{R}^{n+1}$.

1 The Gauß map

Now that we are working with hypersurfaces, we have a much better way of controlling normal vectors to our chosen hypersurface $M \subset \mathbb{R}^{n+1}$. Since each tangent space $T_pM \subset T_p\mathbb{R}^{n+1}$ is an n-dimensional subspace[1] The space $T_pM^\perp$ is one-dimension. In particular, there are precisely two unit vectors: $\pm n$ in $\pm n \in T_pM^\perp$.

[1] We say that such a subspace (or the manifold M itself) has *codimension* 1. This simply means that the difference in dimensions between M and the embedded space is 1.

Now suppose that $\phi : U \to M$ is a chart on our hypersurface. On $\phi(U)$, we can define a unique normal vector field

$$n : \phi(U) \longrightarrow T\mathbb{R}^{n+1}$$

satisfying the following conditions:

- At every $p \in \phi(U)$, $n(p)$ is orthogonal to T_pM.
- At every $p \in \phi(U)$, $n(p)$ is a unit vector.
- The basis
$$(\partial_1\phi, \ldots, \partial_n\phi, n)$$
of $T_p\mathbb{R}^{n+1}$ is always positively oriented.

Definition 4.2. We call the unique smooth vector field $n : \phi(U) \to T\mathbb{R}^{n+1}$ constructed above the *normal field* on $\phi(U)$. Since we can view $n(p) \in T_p\mathbb{R}^{n+1}$ as a unit vector in $\mathbb{R}^{n+1}$, i.e., as a point in S^n. We thus obtain a smooth map

$$n : \phi(U) \longrightarrow S^n \subset \mathbb{R}^{n+1}$$

In this form, we call n the *Gauß map* on U.

Example 4.3. If $M \subset \mathbb{R}^3$ is a 2-dimensional submanifold of $\mathbb{R}^3$ — what we call a *surface* — then the Gauß map can be computed on the image of a chart ϕ as

$$n = \frac{\partial_1 \phi \times \partial_2 \phi}{|\partial_1 \phi \times \partial_2 \phi|}$$

So, for instance, we can consider the torus $T^2 \subset \mathbb{R}^3$, with parameterization

$$\phi(u^1, u^2) = \big(\sin(u^1)(\cos(u^2) + 2), \cos(u^1)(\cos(u^2) + 2), \sin(u^2)\big) \,.$$

The coordinate vector fields are

$$\partial_1 \phi = \big(\cos(u^1)\left(\cos(u^2) + 2\right), -\sin(u^1)(\cos(u^2) + 2), 0\big)$$

and

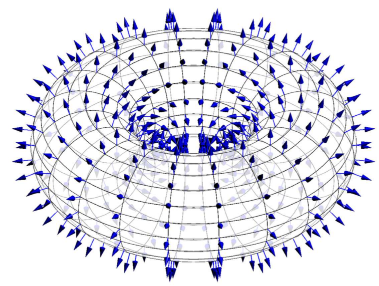
$$\partial_2 \phi = \big(-\sin(u^1)\sin(u^2), -\cos(u^1)\cos(u^2), \cos(u^2)\big)$$

The unit normal is thus

$$n = (-\sin(u^1)\cos(u^2), -\cos(u^1)\cos(u^2), -\sin(u^2))$$

Notice that this is actually a well-defined smooth map on the entire torus, since it is 2π-periodic in both parameters.

Plotted out, one possible normal field for the torus (the negative of the one derived in the example) looks like



Example 4.4. It is also possible to define a global normal field on the n-sphere $S^n \subset \mathbb{R}^{n+1}$. In this case, the Gauß map

$$n : S^n \longrightarrow S^n$$

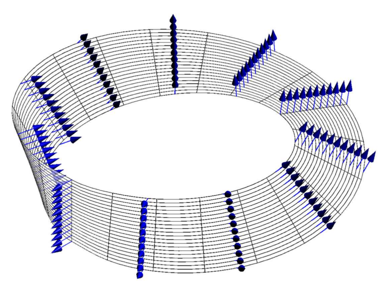
is either the identity map or its negative, depending on which convention is chosen.

A natural next question to ask is: can we always define a *global* normal field/Gauß map? That is, can we define n smoothly on the entire manifold M? The answer to this question is no, as the next example demonstrates.

Example 4.5. We consider the Möbius band $M \subset \mathbb{R}^3$. This is the surface given by the parameterization

$$\psi(u^1, u^2) = \begin{pmatrix} 2\cos(u^2) - u^1\cos(u^2)\sin(u^2/2) \\ 2\sin(u^2) - u^1\sin(u^2)\sin(u^2/2) \\ u^1\cos(u^2/2) \end{pmatrix}$$

If we plot the normal field once around the Möbius band, we can more immediate *see* the problem with extending n to all of M:



From this parameterization, one can compute that the normal field should be

$$n(u^1, u^2) = (-\cos(u^2/2) * \cos(u^2), -\cos(u^2/2) * \sin(u^2), -\sin(u^2/2)).$$

This causes a problem: the parameters $(0, 0)$ and $(0, 2\pi)$ specify the same point in M, but the corresponding normal vectors are opposite. We thus see that we can't extend the normal field smoothly to all of M.

Notice that after looping once around the Möbius band, the normal field must change direction. We thus get a discontinuity when we return to our original point.

Definition 4.6. We call a hypersurface $M \subset \mathbb{R}^{n+1}$ *orientable* if there is a smooth global normal field on M. Otherwise, we call M *non-orientable*.

2 *The shape operator*

We now come to the heart of our study of hypersurfaces: curvature. Throughout this section we will work locally on a single coordinate chart, $\phi : U \to M \subset \mathbb{R}^{n+1}$, on our chosen hypersurface.

Our aim is to define a notion of curvature for hypersurfaces. One reasonable notion is by taking derivatives of the normal field, as the following example demonstrates.

Example 4.7. Let A be an $(n+1) \times n$ matrix of rank n. Then the image of

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$$

is a hyperplane $H \subset \mathbb{R}^{n+1}$ through the origin in $\mathbb{R}^{n+1}$. Thus, there is a unit vector $v \in \mathbb{R}^{n+1}$ such that

$$H = \{x \in \mathbb{R}^n \mid \langle x, v \rangle = 0\}.$$

Since the map A is injective, smooth, and regular (the Jacobian of A at any point is simply A), we can view H as a hypersurface in $\mathbb{R}^{n+1}$. The associated Gauß map

$$n : H \longrightarrow S^n$$

is the constant map on v, i.e. $n(p) = v$ for all $p \in H$. We thus see that, taking the derivative of n in any direction yields the zero vector 0.

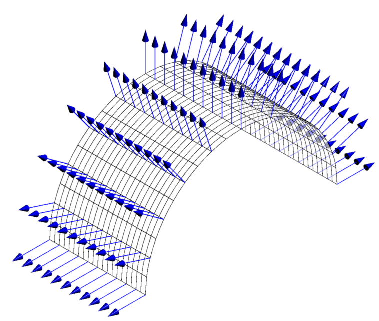
Example 4.8. The Gauß map n for $S^2 \subset \mathbb{R}^3$ is the identity map. If we take a smooth curve $\gamma : [a, b] \to S^2$ and compute the derivative, we get

$$\frac{d}{dt} n(\gamma(t)) = \gamma'(t)$$

And thus, in general, this vector is non-zero.

More generally, if we draw a surface, we see that the derivative of the normal field will change in a given direction precisely when the surface is curved in that direction.[2] Thus we might expect that a good notion of curvature will involve derivatives of the normal field.

[2] For example, examining the normal field on the following half-cylinder



We see that taking derivatives of n along a curve will yield zero along the "straight" coordinate curves, and something non-zero along any other direction.

Our problem now is: which direction do we take a derivative in? Fortunately, there is a devilishly simple solution: all of them! We will consider the *differential* of the Gauß map: $dn_p : T_pM \to T_pS^n$.

However, we will reinterpret the Gauß map as follows. Notice that since n is a unit vector, $\langle n, n \rangle = 1$. Taking a partial derivative of this relation (with respect to a coordinate x^i) yields

$$2\langle \frac{\partial n}{\partial x^i}, n \rangle = 0$$

That is, $\frac{\partial n}{\partial x^i}$ is orthogonal to the unit normal, and thus can be viewed as a tangent vector on n. We can then note that

$$dn \circ (\partial_i \phi) = \frac{\partial n}{\partial x^i}$$

so we can view dn as a linear map

$$dn : T_pM \longrightarrow T_pM.$$

Definition 4.9. Let $M \subset \mathbb{R}^{n+1}$ be a hypersurface, and $\phi : U \to M$ a chart. Let n be the Gauß map on $\phi(U)$. For $p \in \phi(U)$, the *shape operator*[3] is the map

[3] Also sometimes called the Weingarten map.

$$\begin{aligned} L : T_pM &\longrightarrow T_pM \\ v &\longmapsto dn(v) \end{aligned}$$

Proposition 4.10. *The shape operator is self-adjoint with respect to the first fundamental form.*

Proof. It suffices to show this on a basis, i.e., to check that

$$\langle dn(\partial_i\phi), \partial_j\phi\rangle = \langle \partial_i\phi, dn(\partial_j\phi)\rangle$$

Note first that

$$dn(\partial_j\phi) = \frac{\partial n}{\partial x^j}$$

Moreover, since $\partial_i\phi$ is a tangent vector, and n a normal vector, we have

$$\langle \partial_i\phi, n\rangle = 0$$

Taking a derivative with respect to x_j we obtain

$$\left\langle \partial_i\phi, \frac{\partial n}{\partial x^j}\right\rangle + \left\langle \frac{\partial^2\phi}{\partial x^j \partial x^i}, n\right\rangle = 0$$

and, similarly,

$$\left\langle \partial_j\phi, \frac{\partial n}{\partial x^i}\right\rangle + \left\langle \frac{\partial^2\phi}{\partial x^i \partial x^j}, n\right\rangle = 0$$

applying the equality of mixed partials for smooth functions, we thus obtain

$$\left\langle \partial_i\phi, \frac{\partial n}{\partial x^j}\right\rangle = \left\langle \partial_j\phi, \frac{\partial n}{\partial x^i}\right\rangle$$

as desired. □

Definition 4.11. Since L is a self-adjoint linear map, there is an orthonormal[4] basis of T_pM which diagonalizes L. We call the vectors in this orthonormal eigenbasis the *principal directions* of M at p. The corresponding eigenvalues are called the *principal curvatures* of M at p.

[4] Orthonormal with respect to the first fundamental form.

Example 4.12. For $S^n \subset \mathbb{R}^{n+1}$, the shape operator is the identity. Thus, any orthonormal basis of T_pS^n can be considered the principle directions, and the principal curvatures are all 1.

More generally, for $r > 0$, we can consider the *r-scaled n-sphere*:

$$rS^n = \{x \in \mathbb{R}^{n+1} \mid |x| = r\} \subset \mathbb{R}^{n+1}.$$

It is quite easy to see that the Gauß map is the map which sends $x \in rS^n$ to $\frac{x}{r}$. Thus, the principal curvatures of rS^n at any point are all $\frac{1}{r}$.

Example 4.13. We again consider the torus, $T^2 \subset \mathbb{R}^3$, using the coordinate chart

$$\phi(u^1, u^2) = \left(\sin(u^1)(\cos(u^2) + 2), \cos(u^1)(\cos(u^2) + 2), \sin(u^2)\right).$$

The coordinate vector fields are

$$\partial_1 \phi = \left(\cos(u^1)\left(\cos(u^2) + 2\right), -\sin(u^1)(\cos(u^2) + 2), 0\right)$$

and

$$\partial_2 \phi = \left(-\sin(u^1)\sin(u^2), -\cos(u^1)\sin(u^2), \cos(u^2)\right)$$

The unit normal is

$$n(u^1, u^2) = (\sin(u^1)\cos(u^2), \cos(u^1)\cos(u^2), \sin(u^2))$$

We can compute the differential of n using the Jacobian

$$Jn = \begin{pmatrix} \cos(u^1)\cos(u^2) & -\sin(u^1)\sin(u^2) \\ -\sin(u^1)\cos(u^2) & -\cos(u^1)\sin(u^2) \\ 0 & \cos(u^2) \end{pmatrix}$$

The corresponding images of the coordinate vector fields are

$$L(\partial_1 \phi) = (\cos(u^1)\cos(u^2), -\sin(u^1)\cos(u^2), 0)$$

and

$$L(\partial_2 \phi) = (-\sin(u^1)\sin(u^2), -\cos(u^1)\sin(u^2), \cos(u^2))$$

Expressing these vectors in terms of the basis $\partial_1 \phi, \partial_2 \phi$, we find

$$L(\partial_1 \phi) = (\cos(u^1)\cos(u^2), -\sin(u^1)\cos(u^2), 0) = \frac{\cos(u^2)}{2 + \cos(u^2)} \partial_1 \phi$$

and

$$L(\partial_2 \phi) = (-\sin(u^1)\sin(u^2), -\cos(u^1)\sin(u^2), \cos(u^2)) = \partial_2 \phi$$

This is quite convenient, especially since the vector fields $\partial_1 \phi$ and $\partial_2 \phi$ are already orthogonal. Moreover, $\partial_2 \phi$ is already a unit vector field. Normalizing $\partial_1 \phi$ does not change the matrix representation, so our principal curvatures are

$$\kappa_1 = \frac{\cos(u^2)}{2 + \cos(u^2)}$$

and

$$\kappa_2 = 1.$$

We now want to get a single function from the shape operator which measures "how curvy" a hypersurface is at a given point. There are two ways to do this.

Definition 4.14. The *Gaußian curvature* of a hypersurface M is the determinant of the shape operator.

$$K = \det(L).$$

The *mean curvature* is a normalized version of the trace:

$$H = \frac{1}{n} \operatorname{tr}(L)$$

where n is the dimension of M.

Example 4.15. The Gaußian curvature of the r-scaled n-sphere $rS^n \subset \mathbb{R}^{n+1}$ is

$$K = \frac{1}{r^n}.$$

The mean curvature of rS^n is

$$H = \frac{1}{r}.$$

The Gaußian curvature of the torus $T^2 \subset \mathbb{R}^3$ is

$$K = \frac{\cos(u^2)}{2 + \cos(u^2)}$$

The mean curvature of T^2 is

$$H = \frac{3 + 2\cos(u^2)}{4 + 2\cos(u^2)}.$$

3 *The second fundamental form*

We now are in a position to deepen our understanding of curvature, and to tie the covariant derivative together with the shape operator.

Definition 4.16. Let $M \subset \mathbb{R}^{n+1}$ be a hypersurface, and $\phi : U \to M$ a chart. The *second fundamental form* at $p \in \phi(U)$ is the symmetric bilinear form

$$\mathbf{II} : T_pM \times T_pM \longrightarrow \mathbb{R}$$

defined by

$$\mathbf{II}(v, w) = \mathbf{I}(L(v), w) = \langle L(v), w \rangle.$$

We use $h_{i,j}$ to denote the matrix of $\mathbf{II}$ with respect to the basis $\partial_i \phi$ of T_pM, i.e.

$$\mathbf{II}(\partial_i \phi, \partial_j \phi) = h^{i,j}.$$

For $X = X^i \partial_i \phi$, $Y = Y^i \partial_i \phi$, the second fundamental form can then be written as

$$\mathbf{II}(X, Y) = h_{i,j} X^i Y^j.$$

A priori, the second fundamental form may not seem to have a particularly geometric interpretation. To correct this, we first connect the second fundamental form to the covariant derivative.

Proposition 4.17. *For tangent vector fields X, Y on $\phi(U)$,*

$$X^i \frac{\partial}{\partial x^i} Y = \nabla_X Y - \mathbf{II}(X, Y) n.$$

In particular,

$$h_{i,j} = \mathbf{II}(\partial_i \phi, \partial_j \phi)$$

is the length of the normal component of $\frac{\partial^2 \phi}{\partial x^i \partial x^j}$.

Proof. We first compute

$$\begin{aligned} X^i \frac{\partial}{\partial x^i} Y &= X^i \frac{\partial}{\partial x^i}(Y^j \partial_j \phi) \\ &= X^i \frac{\partial Y^j}{\partial x^i} \partial_j \phi + X^i Y^j \frac{\partial^2 \phi}{\partial x^i \partial x^j} \\ &= X^i \frac{\partial Y^j}{\partial x^i} \partial_j \phi + X^i Y^j \Gamma^k_{i,j} \partial_k \phi + X^i Y^j u_{i,j} n \\ &= \nabla_X Y + X^i Y^j u_{i,j} n \end{aligned}$$

where $u_{i,j}n$ is the normal component of $\frac{\partial^2 \phi}{\partial x^i \partial x^j}$. It thus suffices to show that $-u_{i,j}$ is the second fundamental form applied to $\partial_i \phi$ and $\partial_j \phi$.

From the proof of Proposition 4.10, we have that

$$\begin{aligned} h_{i,j} &= \langle L(\partial_i \phi), \partial_j \phi \rangle \\ &= \langle dn \circ \partial_i \phi, \partial_j \phi \rangle \\ &= -\left\langle n, \frac{\partial^2 \phi}{\partial x^i \partial x^j} \right\rangle \end{aligned}$$

completing the proof. □

This means that we can reinterpret the second fundamental form $\mathbf{II}(X, Y)$ as measuring the normal component of the rate of change of Y in the X direction.

Notation 4.18. Given a non-degenerate bilinear form $B : T_pM \times T_pM \to \mathbb{R}$ with corresponding matrix $b = \{b_{i,j}\}$ with respect to a coordinate basis, we denote by $b^{i,j}$ the entries of the matrix b^{-1}.

We can relate the Gaußian curvature to the first and second fundamental forms as follows.

Proposition 4.19. *Let ℓ_i^j be the coefficients of the shape operator with respect to the coordinate basis, i.e.*

$$L(\partial_i \phi) = \ell_i^j \partial_j \phi.$$

Then

$$\ell_i^j = g^{j,k} h_{k,i}.$$

Proof. The form $\mathbf{I}$ is a symmetric bilinear form, L is a self-adjoint operator with respect to $\mathbf{I}$, and $\mathbf{II}$ is the symmetric bilinear form associated to L. Thus, the proposition follows immediately from Lemma C.17. □

Corollary 4.20. *The determinant of the shape operator L is given by*

$$\det(L) = \det(h) \det(g^{-1}).$$

where g and h are the matrices of the first and second fundamental forms, respectively

Corollary 4.21. *For any i,*

$$\frac{\partial n}{\partial x^i} = g^{j,k} h_{k,i} \partial_j \phi$$

4 *Ways of computing curvature*

In our previous discussions, we actually came up with several *different* ways of computing the curvatures of a hypersurface:

1. We can directly compute the shape operator with respect to a basis as follows. We first compute the coordinate vector fields $\partial_i \phi$, and then we compute the unit normal n. We then take derivatives $\frac{\partial n}{\partial x^i}$, and express these as linear combinations of the coordinate vector fields. We thus obtain a matrix representation of L, and can diagonalize it to obtain principal curvatures and principal directions.

2. As before, we can compute the coordinate vector fields and the unit normal. However, we can instead directly compute the second fundamental form. To do this, we recall that
$$\mathbf{II}(\partial_i \phi, \partial_j \phi) = -\left\langle n, \frac{\partial^2 \phi}{\partial x^i \partial x^j} \right\rangle .$$
We can then compute the Gaussian and mean curvatures in terms of the cooeficients $g^{i,j}$ and $h_{i,j}$.

The second of these methods is often more computationally intensive, but can still be useful. We provide an example of such a computation.

Example 4.22. We return to the torus $T^2 \subset \mathbb{R}^3$, using the coordinate chart

$$\phi(u^1, u^2) = \left(\sin(u^1)(\cos(u^2)+2), \cos(u^1)(\cos(u^2)+2), \sin(u^2)\right) .$$

The coordinate vector fields are

$$\partial_1 \phi = \left(\cos(u^1)\left(\cos(u^2)+2\right), -\sin(u^1)(\cos(u^2)+2), 0\right)$$

and

$$\partial_2 \phi = \left(-\sin(u^1)\sin(u^2), -\cos(u^1)\sin(u^2), \cos(u^2)\right) .$$

The unit normal is

$$n(u^1, u^2) = (\sin(u^1)\cos(u^2), \cos(u^1)\cos(u^2), \sin(u^2)).$$

The second derivatives of the chart are

$$\begin{aligned}
\frac{\partial^2 \phi}{\partial u^1 \partial u^2} &= (-\cos(u^1)\sin(u^2), \sin(u^1)\sin(u^2), 0) \\
\frac{\partial^2 \phi}{\partial u^1 \partial u^1} &= \left(-\sin(u^1)\left(\cos(u^2)+2\right), -\cos(u^1)(\cos(u^2)+2), 0\right) \\
\frac{\partial^2 \phi}{\partial u^2 \partial u^2} &= \left(-\sin(u^1)\cos(u^2), -\cos(u^1)\cos(u^2), -\sin(u^2)\right)
\end{aligned}$$

The first fundamental form with n then yields

$$h = \begin{pmatrix} \cos^2(u^2)(\cos(u^2)+2) & 0 \\ 0 & -1 \end{pmatrix}$$

The metric is given by

$$g = \begin{pmatrix} (\cos(u^2)+2)^2 & 0 \\ 0 & 1 \end{pmatrix}$$

so we have

$$g^{-1} = \begin{pmatrix} \frac{1}{(\cos(u^2)+2)^2} & 0 \\ 0 & 1 \end{pmatrix}$$

We thus see that

$$L = g^{-1}h = \begin{pmatrix} \frac{\cos^2(u^2)}{(\cos(u^2)+2)} & 0 \\ 0 & 1 \end{pmatrix}$$

precisely as we previously calculated.

5 *Interpreting curvature*

Now that we have defined curvature, let us try to understand more precisely what it means. The first thing to note is that, for an arbitrary tangent vector field $X = X^i \partial_i \phi$,

$$dn \circ X = X^i dn \circ \partial_i \phi = X^i \frac{\partial n}{\partial x^i}.$$

that is, $L(X(p))$ is the derivative of n in the $X(p)$ direction. Because of this, the shape operator L captures *all* of the ways in which the normal direction can change at a point.

The principal directions and curvatures, however, are special. For $v \in T_pM$ to be an eigenvector of L at $p \in M$ means that, in the v direction, the normal is also changing in precisely the v-direction. As such, the principal directions at p are the "directions in which M bends at p," and the corresponding principal curvatures are "how much M is bending in the v direction at p". More can be said, however.

Proposition 4.23. *The unit vectors $v \in T_pM$ on which $\mathbf{II}(v,v)$ its maximal and minimal values are eigenvectors of L.*

Proof. This is actually a much more general statement. Let $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a symmetric, positive-definite bilinear form, and let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a self-adjoint map. We will show that the maxima and minima of $B(L(v), v)$ on

$$\{v \in \mathbb{R}^n \mid B(v,v) = 1\}$$

are eigenvectors of L.

Without loss of generality, we may assume that B is the Euclidean inner product, by choosing a B-orthonormal basis. We may also, without loss of generality, assume that L is represented by a diagonal matrix A, and that the eigenvectors are the standard basis vectors. We are left with the task of finding the maxima and minima of

$$f(v) = \langle Av, v \rangle = \sum_i a_{i,i} v^i v^i$$

subject to the constraint that

$$g(v) = \langle v, v \rangle = \sum_i v^i v^i = 1.$$

We apply the method of Lagrange multipliers, i.e. we seek to find values $\lambda \in \mathbb{R}$ such that

$$(Jf) - \lambda(Jg) = 0$$

for any $v \in S^{n-1}$.

Computing the Jacobians, we obtain

$$\begin{pmatrix} 2a_{1,1}v^1 - 2\lambda_1 v^1 \\ \cdots \\ 2a_{n,n}v^n - 2\lambda_n v^n \end{pmatrix} = 0$$

The only possible solutions of this equation are $\lambda = a_{i,i}$ for some i, and v is a linear combination of the standard basis vectors e_j for which $a_{j,j} = a_{i,i}$. Thus, the only possible extrema on S^{n-1} are eigenvectors of L, proving the proposition. □

We typically write $\kappa_1, \ldots, \kappa_n$ for the principal curvatures, with

$$\kappa_1 \leq \kappa_2 \leq \cdots \leq \kappa_n.$$

Thus, κ_1 is always the smallest possible value of the curvature, and κ_n the largest.

When we consider a surface — a hypersurface $M \subset \mathbb{R}^3$ — we only have two principal curvatures: κ_1 and κ_2. We thus only have two possibilities at $p \in M$:

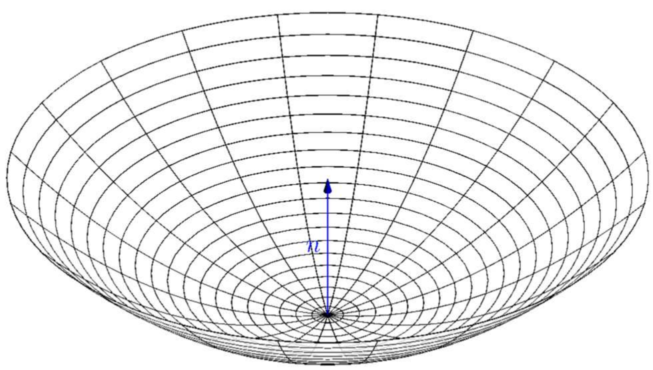
1. $\kappa_1 = \kappa_2$, and the shape operator is multiplication by a constant value. The surface curves the same amount in every direction at p.

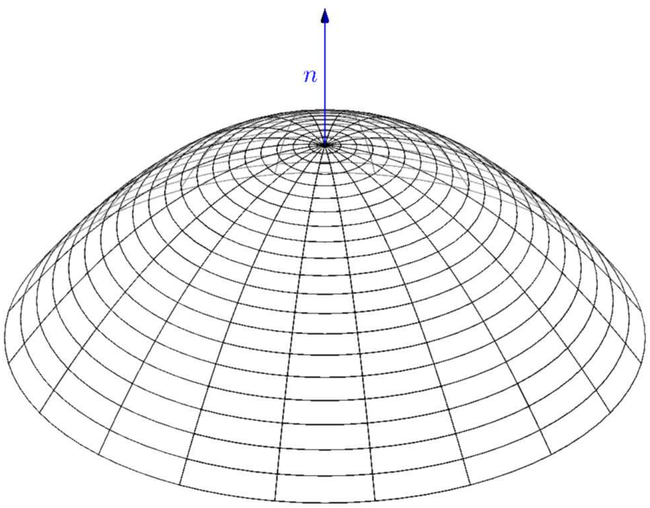
2. $\kappa_1 < \kappa_2$. The curvature is smallest in the first principal direction, and largest in the second principal direction.

In some sense, only the absolute value of curvature corresponds to how much the surface curves. The sign of the curvature corresponds to which direction the surface is bending in. The sign of the Gaußian curvature thus tells us whether the surface is bending the same way in all directions, or whether it bends both "upwards"[5] and "downwards" depending on the direction. In all, there are four cases to consider for a surface $M \subset \mathbb{R}^3$.

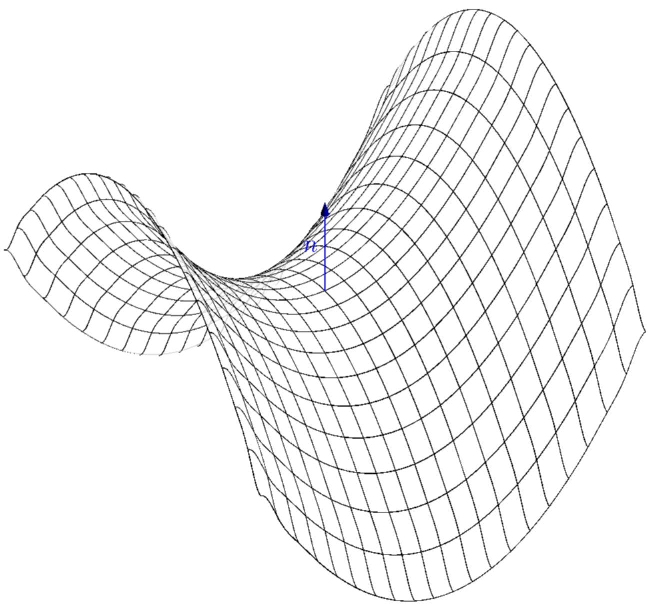
[5] towards the unit normal.

Case I, $K > 0$: In this case, κ_1 and κ_2 have the same sign, and are both non-zero. Pictorially, we have one of the following cases:

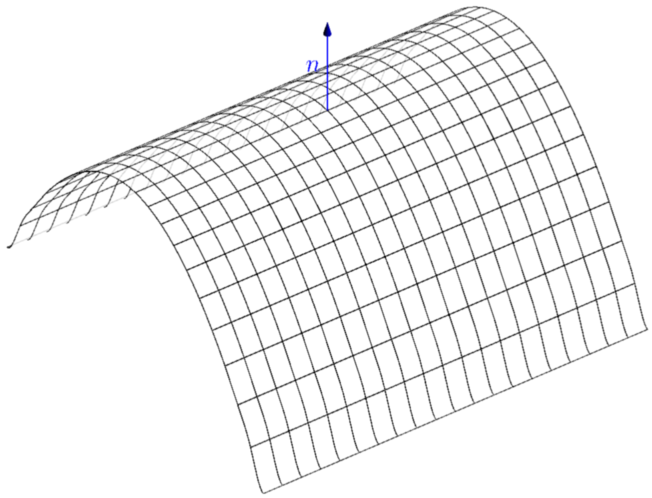




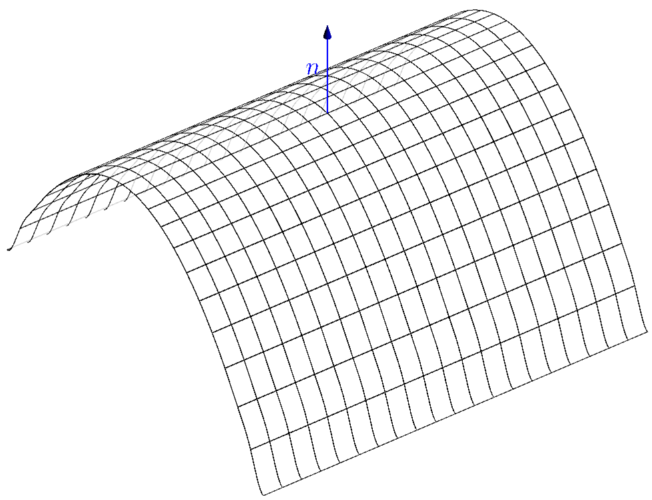
Case II, $K < 0$: In this case, κ_1 and κ_2 have opposite signs. Pictorially:



Case III, $K = 0$ $\kappa_1 \neq 0$ or $\kappa_2 \neq 0$: In this case, there is a direction in which the surface does not bend, and a direction in which it does. Pictorially:



Case IV, $K = \kappa_1 = \kappa_2 = 0$: In this case, the surface is flat at p (though not necessarily more generally)



6 *Surfaces and the Theorema Egregium*

Before moving on to our first main theorem, let us briefly summarize what we know about the intrinsic vs. extrinsic natures of our various constructions.

- By definition, the hypersurface M and the first fundamental form $\mathbb{I}$ (or, equivalently, the matrix-valued function g) are intrinsic.

- We defined the covariant derivative ∇_X in an extrinsic way, as the tangential component of a derivative in the ambient space. However, we showed that the Christoffel symbols are intrinsic, and thus, in Corollary 3.28, that the covariant derivative is intrinsic.
- The shape operator is the differential of the unit normal field, and thus is manifestly extrinsic.
- The second fundamental form is the normal component of a derivative in the ambient space, and thus is extrinsic.
- The principal directions, principal curvatures, Gaußian curvature, and mean curvature are defined via the shape operator (or equivalently the second fundamental form). As a result these are extrinsic quantities, telling us how much our hypersurface bends in the ambient space.

We now come to the first punchline of this course: the Gaußian curvature, which we defined purely extrinsically, is actually intrinsic — it doesn't depend on the ambient space, but only on the first fundamental form. While this is true (up to a sign) in higher dimensions, to prove this would go beyond the scope of this course. We will therefore make one final specialization: to *surfaces*.

Definition 4.24. A *surface* is a hypersurface in $\mathbb{R}^3$.

Most of the examples we have considered so far in the course are surfaces — partly for ease of visualization. It is also in this context that, Gauß's Theorema Egregium[6] was first proven.[7]

[6] This means, in Latin, something like "Remarkable Theorem".
[7] It is worth pointing out that variants of the Theorema Egregium hold for hypersurfaces in higher dimensions. The proofs, however, are far more complicated, and the precise statement varies with the parity of the dimension of the ambient space. For these reasons, we will not delve into the higher dimensional cases any further.

Theorem 4.25 (Theorema Egregium)**.** *Let $M \subset \mathbb{R}^3$ be a surface. The Gaußian curvature of M is intrinsic.*

To prove this theorem, we will make use of everything we have defined thus far: the covariant derivative, the first and second fundamental forms, and the shape operator. We will attack the Theorema Egregium by expressing combinations of the coefficients of the second fundamental form in terms of Christoffel symbols. These are the *Gauß equations*.

Proposition 4.26 (Gauß equations)**.** *Let $M \subset \mathbb{R}^3$ be a surface[8], and let $\phi : U \to M$ be a chart. Then*

[8] This does not, in fact, require that M be a surface. The argument works for any hypersurface in $\mathbb{R}^{n+1}$.

$$\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi = \mathbf{II}(\partial_j \phi, \partial_k \phi) L(\partial_i \phi) - \mathbf{II}(\partial_i \phi, \partial_k \phi) L(\partial_j \phi)$$

or, equivalently

$$\frac{\partial}{\partial x^k} \Gamma^n_{i,j} - \frac{\partial}{\partial x^j} \Gamma^n_{i,k} + \left(\Gamma^\ell_{i,j} \Gamma^n_{\ell k} - \Gamma^\ell_{i,k} \Gamma^n_{\ell,j} \right) = g^{\ell,n} \left(h_{i,j} h_{k,\ell} - h_{i,k} h_{j,\ell} \right).$$

Proof. While this may look like a nightmare, it is simply two different computations of a tangential component of a third-order partial derivative

$$\frac{\partial^3 \phi}{\partial x^i \partial x^j \partial x^k}.$$

We will compute the version of the Gauß equations using covariant derivatives, and leave the coordinate form to the reader.

The computation amounts to the repeated application of Proposition 4.17. We compute

$$\begin{aligned}
\frac{\partial^3 \phi}{\partial x^i \partial x^j \partial x^k} &= \frac{\partial}{\partial x^i}\left(\frac{\partial^2 \phi}{\partial x^j x^k}\right) \\
&= \frac{\partial}{\partial x^i}\left(\nabla_j(\partial_k \phi) - \mathbf{II}(\partial_j, \partial_k) n\right) \\
&= \nabla_i \nabla_j(\partial_k \phi) - \mathbf{II}(\partial_i, \nabla_j(\partial_k \phi)) n - \left(\frac{\partial}{\partial x^i}\mathbf{II}(\partial_j, \partial_k)\right) n - \mathbf{II}(\partial_j, \partial_k)\frac{\partial n}{\partial x^i}
\end{aligned}$$

We then take the tangential component, eliminating multiples of n, this leaves

$$\left(\frac{\partial^3 \phi}{\partial x^i \partial x^j \partial x^k}\right)^T = \nabla_i \nabla_j(\partial_k \phi) - \mathbf{II}(\partial_j, \partial_k)\frac{\partial n}{\partial x^i}$$

An identical computation shows

$$\left(\frac{\partial^3 \phi}{\partial x^j \partial x^i \partial x^k}\right)^T = \nabla_j \nabla_i(\partial_k \phi) - \mathbf{II}(\partial_i, \partial_k)\frac{\partial n}{\partial x^j}.$$

The equality of mixed partials for smooth functions means that these two quantities are equal, and thus we have

$$\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi = \mathbf{II}(\partial_j \phi, \partial_k \phi)\frac{\partial n}{\partial x^i} - \mathbf{II}(\partial_i \phi, \partial_k \phi)\frac{\partial n}{\partial x^j}$$

Finally, noting that $L(\partial_i \phi) = \frac{\partial n}{\partial x^i}$, we have that

$$\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi = \mathbf{II}(\partial_j \phi, \partial_k \phi) L(\partial_i \phi) - \mathbf{II}(\partial_i \phi, \partial_k \phi) L(\partial_j \phi)$$

The second version of the Gauß equations simply amounts to considering the cooefficient of $\partial_m \phi$ in the coordinate expansion of each side. We leave the task of expanding this expression to the reader. □

Corollary 4.27. *For any indices i, j, k, ℓ, the quantity*

$$h_{i,j} h_{k,\ell} - h_{i,k} h_{j,\ell}$$

is intrinsic. In particular,

$$\det(h) = h_{1,1} h_{2,2} - h_{1,2} h_{2,1}$$

is an intrinsic quantity.

Proof. Multiplying the Gauß equations by g shows that the desired quantity can be expressed purely in terms of Christoffel symbols and the first fundamental form g, both of which are intrinsic. □

Proof of the Theorema Egregium. Applying Proposition 4.20, we see that the Gaußian curvature K is

$$K = \det(L) = \det(h) \det(g)^{-1}.$$

However, g is intrinsic and, by Corollary 4.27 $\det(h)$ is intrinsic. Thus, K is intrinsic, as desired. □

7 *The Riemannian curvature*

In our proof of the *Theorema Egregium*, we related the Gaußian curvature to intrinsic quantities: $\det(h)$ and $\det(g)$. On the one hand, this is a satisfactory result, it showed us that the Gaußian curvature was intrinsic. However, it is somewhat unsatisfactory that we do not have an intuitive interpretation of the quantity $\det(h)$.

To rectify this, we will examine the expression on the left-hand side of the Gauß equations:

$$\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi$$

The basic idea here is that this expression measures the failure of parallel transport around a very small loop to be the identity map.

Before we do this, however, we need to note something a bit odd. We can take three arbitrary vector fields, and feed them into the coefficients

$$\frac{\partial}{\partial x^k}\Gamma^n_{i,j} - \frac{\partial}{\partial x^j}\Gamma^n_{i,k} + \left(\Gamma^\ell_{i,j}\Gamma^n_{\ell k} - \Gamma^\ell_{i,k}\Gamma^n_{\ell,j}\right)$$

defined by the expression $\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi$. When we do this for $X = X^i \partial_i \phi$, $Y = Y^i \partial_i \phi$, and $Z = Z^i \partial_i \phi$, we get

$$\left(\frac{\partial}{\partial x^k}\Gamma^n_{i,j} - \frac{\partial}{\partial x^j}\Gamma^n_{i,k} + \left(\Gamma^\ell_{i,j}\Gamma^n_{\ell k} - \Gamma^\ell_{i,k}\Gamma^n_{\ell,j}\right)\right) X^k Y^j Z^i \partial_n \phi$$

A tedious computation shows that this is *not* the coordinate representation of

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z.$$

Rather, a somewhat tedious computation shows that we get

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

Definition 4.28. The *Riemann curvature tensor*[9] is the map which sends three vector fields to a vector field, given by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

[9] We are not going to define tensors in this course. Fortunately, we will not need to in our exploration of the Riemannian curvature.

From our coordinate formula for the Lie bracket, we see that

$$[\partial_i \phi, \partial_j \phi] = \left(\delta^\ell_i \frac{\partial \delta^k}{\partial x^\ell} - \delta^\ell_j \frac{\partial \delta^k_i}{\partial x^\ell}\right) \partial_k \phi = 0$$

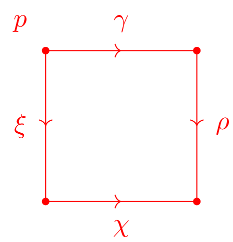
since derivatives of the Kronecker delta are identically zero. Thus, wee see that the Riemann tensor simplifies to the expression

$$\nabla_i \nabla_j \partial_k \phi - \nabla_j \nabla_i \partial_k \phi$$

when applied to coordinate vector fields.

We will give a heuristic interpretation of this expression in terms of parallel transport, to try and explain why it might appear in a description of curvature.[10] We consider the following setup. Consider a small loop in M comprised of four curves, γ, ρ, ξ and χ as in the following figure.

[10] This explanation is cribbed from physics courses, and is not fully rigorous. Nonetheless, it does give a sense of what $R(X,Y)Z$ should represent.



Suppose that γ and χ are defined for $t \in [0, \lambda]$, and ξ, ρ are defined for $u \in [0, \mu]$.

Write $\gamma'(0) = V$, $\xi'(0) = W$, and suppose that $\rho'(0)$ is the parallel transport of W along γ, and similarly that $\chi'(0)$ is the parallel transport of V along ξ. We will start with a tangent vector X at p, and parallel transport it along γ and ρ, and the compare the result to the parallel transport of X along ξ and χ.

Throughout, we will also tacitly act as though (t, u) are coordinates on the little rectangle bounded by our curves.

If we extend the vector X to a parallel vector field along γ, we see that

$$\nabla_{\gamma'} X = 0$$

or, in coordinates

$$\frac{dX^k}{dt} = -X^i \frac{d\gamma^j}{dt} \Gamma^k_{i,j}$$

And similarly for the parallel transport of W along γ. We then Taylor expand X, W, and the Christoffel symbols in t, to approximate their values at $(\lambda, 0)$.

$$\begin{aligned}
X^k|_{(\lambda,0)} &= X^k|_{(0,0)} - \left(X^i \frac{d\gamma^j}{dt} \Gamma^k_{i,j} \right)|_{(0,0)} \lambda + \text{H.O.T.} \\
&= X^k|_{(0,0)} - X^i|_{(0,0)} V^j|_{(0,0)} \Gamma^k_{i,j}|_{(0,0)} \lambda + \text{H.O.T.} \\
W^k|_{(\lambda,0)} &= W^k|_{(0,0)} - W^i|_{(0,0)} V^j|_{(0,0)} \Gamma^k_{i,j}|_{(0,0)} \lambda + \text{H.O.T.} \\
\Gamma^k_{i,j}|_{(\lambda,0)} &= \Gamma^k_{i,j}|_{(0,0)} + \frac{\partial \Gamma^k_{i,j}}{\partial x^\ell}|_{(0,0)} V^\ell|_{(0,0)} \lambda + \text{H.O.T}
\end{aligned}$$

Applying the same procedure to parallel transport $X|_{(\lambda,0)}$ along ρ, we obtain (up to second order)

$$X^k|_{(\lambda,\mu)} = X^k - X^i V^j \Gamma^k_{i,j} \lambda - \left[\left(X^i - X^\ell V^r \Gamma^i_{\ell,r} \lambda \right) \left(W^j - W^\ell V^r \Gamma^j_{\ell,r} \lambda \right) \left(\Gamma^k_{i,j} + \frac{\partial \Gamma^k_{i,j}}{\partial x^p} V^p \lambda \right) \right] \mu + \text{H.O.T}$$

where every coefficient on the right-hand side is evaluated at $(0, 0)$. Simplifying this expression, we then obtain a vector which we denote

$$\begin{aligned}
X^k_{\|\gamma,\rho} = X^k - X^i V^j \Gamma^k_{i,j} \lambda - X^i W^j \Gamma^k_{i,j} \mu \\
+ X^i W^\ell V^r \left(\Gamma^j_{\ell,r} \Gamma^k_{i,j} - \frac{\partial \Gamma^k_{i,j}}{\partial x^r} + \Gamma^n_{i,r} \Gamma^k_{n,\ell} \right) + \text{H.O.T.}
\end{aligned}$$

Performing the same procedure along ξ and χ, we obtain a vector $X^k_{\|\xi\chi}$. We then take the difference of these two vectors. A brief computation shows

$$\begin{aligned} X^k_{\|\gamma,\rho} - X^k_{\|\xi,\chi} &= X^i W^\ell V^r \left(\Gamma^n_{i,r}\Gamma^k_{n,\ell} - \Gamma^n_{i,\ell}\Gamma^k_{n,r} + \frac{\partial \Gamma^k_{i,r}}{\partial x^\ell} - \frac{\partial \Gamma^k_{i,\ell}}{\partial x^r} \right) \lambda\mu + \text{H.O.T.} \\ &= X^i W^\ell V^r R^k_{\ell,r,i} \lambda\mu + \text{H.O.T} \end{aligned}$$

Thus, to second order, the difference between these two parallel transports of X is

$$X_{\|\gamma,\rho} - X_{\|\xi,\chi} = R(W,V)X\lambda\mu$$

As such, we can interpret the Riemann curvature as measuring "how much parallel transport around a small loop fails to be the identity".

8 *Some computations of the curvature of surfaces*

In this short section, we work through a number of examples of curvatures for surfaces, to provide some additional context for the rest of the chapter.

Example 4.29. We consider the *helicoid*[11], a surface defined by a single chart

$$\phi(u^1, u^2) = (u^2 \cos(u^1), u^2 \sin(u^1), u^1)$$

which we will treat as defined on $(-\infty,\infty) \times (0,\infty)$. The coordinate vector fields are then

$$\begin{aligned} \partial_1\phi &= (-u^2 \sin(u^1), u^2\cos(u^1), 1) \\ \partial_2\phi &= (\cos(u^1), \sin(u^1), 0) \end{aligned}$$

The first fundamental form is then given by

$$g = \begin{pmatrix} (u^2)^2 + 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We can compute the unit normal using the cross product on $\mathbb{R}^3$:

$$n = \frac{1}{\sqrt{1+(u^2)^2}} \left(-\sin(u^1), \cos(u^1), -v \right)$$

We thus see that

$$\frac{\partial n}{\partial u^1} = \frac{1}{\sqrt{1+(u^2)^2}} \left(-\cos(u^1), -\sin(u^1), 0 \right) = -\frac{1}{\sqrt{1+(u^2)^2}} \partial_2\phi$$

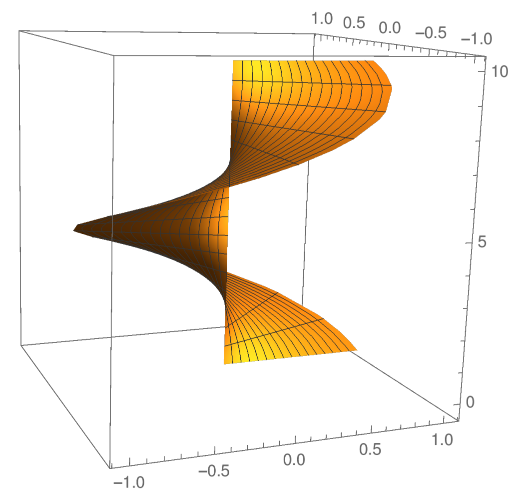
and

$$\frac{\partial n}{\partial u^2} = -\frac{1}{\sqrt{1+(u^2)^2}^3} \partial_1\phi.$$

Thus, the matrix representation of the shape operator[12] is

$$L = \begin{pmatrix} 0 & -\frac{1}{(1+(u^2)^2)^{3/2}} \\ -\frac{1}{\sqrt{1+(u^2)^2}} & 0 \end{pmatrix}$$

[11] An image is



[12] We may notice that the matrix we obtain is not symmetric. The reason for this is that the shape operator is self adjoint *with respect to the first fundamental form*, i.e.

$$l^k_i g_{k,j} = \ell^k_j g_{i,k}.$$

In this case, this translates to

$$-(1+(u^2)^2)\frac{1}{(1+(u^2)^2)^{3/2}} = -\frac{1}{\sqrt{1+(u^2)^2}}$$

which clearly holds.

So we see that the Gaussian curvature is

$$K = -\frac{1}{(1+(u^2)^2)^2}$$

and the mean curvature is $H = 0$. Notice that this is enough for us to deduce the principal curvatures without having to diagonalize L.

Example 4.30. Let $\gamma(t)$ be a unit speed curve with first coordinate strictly positive. We can parameterize the resulting surface of revolution as

$$\phi(u^1, u^2) = \left(\gamma^1(u^2)\cos(u^1), \gamma^1(u^2)\sin(u^1), \gamma^2(u^2)\right)$$

We will use dots above functions to denote derivatives, to ease notation. We thus have

$$\partial_1\phi = (-\gamma^1(u^2)\sin(u^1), \gamma^1(u^2)\cos(u^1), 0)$$

and

$$\partial_2\phi = (\dot\gamma^1(u^2)\cos(u^1), \dot\gamma^1(u^2)\sin(u^1), \dot\gamma^2(u^2))$$

We then compute the cross product of our coordinate vector fields

$$\partial_1\phi \times \partial_2\phi = \begin{pmatrix} \gamma^1(u^2)\dot\gamma^2(u^2)\cos(u^1) \\ \gamma^1(u^2)\dot\gamma^2(u^2)\sin(u^1) \\ -\gamma^1(u^2)\dot\gamma^1(u^2) \end{pmatrix}$$

The norm of this vector is $\gamma^1(u^2)$, so our unit normal is

$$n = \left(\dot\gamma^2(u^2)\cos(u^1), \dot\gamma^2(u^2)\sin(u^1), \dot\gamma^1(u^2)\right).$$

Taking derivatives yields

$$\frac{\partial n}{\partial u^1} = \left(-\dot\gamma^2(u^2)\sin(u^1), \dot\gamma^2(u^2)\cos(u^1), 0\right) = \frac{\dot\gamma^2(u^2)}{\gamma^1(u^2)}\partial_1\phi$$

and

$$\frac{\partial n}{\partial u^2} = \left(\ddot\gamma^2(u^2)\cos(u^1), \ddot\gamma^2(u^2)\sin(u^1), \ddot\gamma^1(u^2)\right)$$

However, since γ is unit speed, we have that $\dot\gamma$ and $\ddot\gamma$ are orthogonal, hence

$$\ddot\gamma^1 = -\frac{\ddot\gamma^2\dot\gamma^2}{\dot\gamma^1}$$

Thus

$$\frac{\partial n}{\partial u^2} = \left(\ddot\gamma^2(u^2)\cos(u^1), \ddot\gamma^2(u^2)\sin(u^1), \frac{\ddot\gamma^2(u^2)\dot\gamma^2(u^2)}{\dot\gamma^1(u^2)}\right) = \frac{\ddot\gamma^2(u^2)}{\dot\gamma^1(u^2)}\partial_2\phi.$$

The matrix of the shape operator with respect to the coordinate basis is thus

$$L = \begin{pmatrix} \frac{\dot\gamma^2(u^2)}{\gamma^1(u^2)} & 0 \\ 0 & \frac{\ddot\gamma^2(u^2)}{\dot\gamma^1(u^2)} \end{pmatrix}$$

The Gaußian curvature is then

$$K = \frac{\ddot\gamma^2\dot\gamma^2}{\dot\gamma^1\gamma^1} = -\frac{\ddot\gamma^1}{\gamma^1}$$

5

The Gauß-Bonnet Theorem

The second major theorem we will discuss is the *Gauß-Bonnet Theorem*, which defines a *topological* invariant of surfaces based on curvatures. Along the way, we will define further curvatures of submanifolds, and develop some more techniques to compute derivatives and integrals on submanifolds. Much of what we do in this chapter holds in higher dimensions, and so we will temporarily move away from surfaces before returning to the 2-dimensional setting for the Gauß-Bonnet Theorem itself.

1 *Geodesic curvature*

We first return to the theme of covariant derivatives and parallel transport. Recall that, for a hypersurface $M \subset \mathbb{R}^{n+1}$, a curve $\gamma : [a, b] \to M$ from p to q, and a vector $v \in T_pM$, we defined the parallel transport of v along γ to be the *unique* vector field X along γ such that

$$\nabla_{\gamma'} X = 0 \qquad \text{and} \qquad X_{\gamma(a)} = v.$$

A geodesic is then a curve whose tangent field is parallel along γ, i.e., such that $\nabla_{\gamma'}\gamma' = 0$.

Taking an arbitrary (not necessarily geodesic) curve $\gamma : [a, b] \to M$, we can still consider the vector field

$$\nabla_{\gamma'}\gamma'$$

along γ. At each point along γ this vector will tell us the tangential rate of change of the tangent field of γ. In some sense, this measures how far from being a geodesic.

Definition 5.1. The *geodesic curvature* of a unit speed curve $\gamma : [a, b] \to M$ is the norm

$$\kappa_g(\gamma) = |\nabla_{\gamma'}\gamma'|.$$

This is a smooth function of the parameter of γ so long as γ is regular.

On the other hand, we define the *normal curvature* of the unit speed curve γ in $\mathbb{R}^n$ to be the length of the normal component of the second derivative. More precisely

$$\kappa_n(\gamma) = \langle \gamma'\prime, n \rangle$$

Remark 5.2. There are a number of interesting relations between four different notions of curvature we have now defined.

- Our first setup is as follows: Suppose that $\gamma : (a, b) \to M$ is a unit speed curve such that, when viewed as a curve $\gamma : (a, b) \to \mathbb{R}^{n+1}$, γ is a Frenet curve. Denote by κ_1 the first curvature of γ as a Frenet curve in $\mathbb{R}^{n+1}$. $\nabla_{\gamma'}\gamma'$ is the tangential component of the second derivative, and $\langle \gamma'', n\rangle$ is the normal component. We thus have the relation

$$\gamma'' = \nabla_{\gamma'}\gamma' + \langle \gamma'', n\rangle n$$

Since $\nabla_{\gamma'}\gamma'$ and $\langle \gamma'', n\rangle n$ are orthogonal, we get the Pythagorean identity

$$|\gamma''|^2 = |\nabla_{\gamma'}\gamma'|^2 + \langle \gamma'', n\rangle^2$$

or, more simply

$$\kappa_1^2 = \kappa_g^2 + \kappa_n^2$$

- On the other hand, let γ be a unit speed curve such that $\gamma'(0) = v$ is a principal direction at $\gamma(0)$. Recall from Section 5 that, writing $\gamma = \phi \circ \rho$ for some chart ϕ, we have

$$\gamma''(0) = \frac{d^2\rho^i}{dt^2}\partial_i\phi + v^i v^j \frac{\partial^2\phi}{\partial x^i \partial x^j}$$

Thus, the normal component of $\gamma''(0)$ is

$$\kappa_n(0) = \left\langle v^i v^j \frac{\partial^2\phi}{\partial x^i \partial x^j}, n_0 \right\rangle$$

equivalently, this is

$$\kappa_n(0) = v^i v^j \left\langle \frac{\partial^2\phi}{\partial x^i \partial x^j}, n_0 \right\rangle = -v^i v^j h_{i,j} = -\mathbf{II}(v, v)$$

i.e., the negative of the principal curvature corresponding to the principal direction v.

Example 5.3.

1. We first consider $\mathbb{R}^n$ as a submanifold of itself. In this case, there is no normal direction in the ambient space, and so the covariant derivative agrees with the usual partial derivative in the ambient space. If we consider a unit-speed Frenet curve $\gamma : [a, b] \to \mathbb{R}^n$, then we find that $\kappa_g(\gamma)$ is simply

$$\left|\frac{d^2\gamma}{dt^2}\right|$$

which is the first curvature of γ, considered as a Frenet curve in $\mathbb{R}^n$.

2. Consider a circle in $S^2 \subset \mathbb{R}^3$ defined by $z = \sin(\theta)$ for some $\theta \in (0, \pi/2)$. An arc-length parameterization of such a circle would be

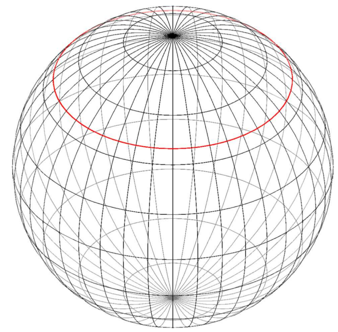
$$\gamma(t) = \left(\cos(\theta)\cos\left(\frac{t}{\cos(\theta)}\right), \cos(\theta)\sin\left(\frac{t}{\cos(\theta)}\right), \sin(\theta)\right)$$

This factors through the chart

$$\phi(u^1, u^2) = (\cos(u^1)\cos(u^2), \cos(u^1)\sin(u^2), \sin(u^1))$$

as $\gamma = \phi \circ \rho$ where $\rho(t) = \left(\theta, \frac{t}{\cos(\theta)}\right)$. Notice that $\gamma'(t) = \cos(\theta)\partial_2\phi(\gamma(t))$.

The curves we are considering look like



From Example 3.35, we have the Christoffel symbols of this chart:

$$\Gamma^1_{i,j} = \begin{pmatrix} 0 & 0 \\ 0 & \cos(u^1)\sin(u^1) \end{pmatrix} \qquad \Gamma^2_{i,j} = \begin{pmatrix} 0 & -\tan(u^1) \\ -\tan(u^1) & 0 \end{pmatrix}$$

We thus can compute

$$\nabla_{\gamma'}\gamma' = -\frac{\sin(\theta)}{\cos^2(\theta)}\partial_2\phi$$

To compute the geodesic curvature, we thus need only compute the norm of

$$\partial_2\phi = \left(-\sin(u^2)\cos(u^1), \cos(u^2)\cos(u^1), 0\right)$$

$$\kappa_g(\gamma) = |\nabla_{\gamma'}\gamma'| = \frac{\sin(\theta)}{\cos(\theta)}.$$

Notice that $\cos(\theta) =: r$ is the radius of the circle γ in the plane $z = \sin(\theta)$, so that we have

$$\kappa_g(\gamma) = \frac{\sqrt{1-r^2}}{r}.$$

2 Orientation and integration

Previously, when we discussed volume integrals in a manifold, we glossed over the fact that the sign of our integrals could depend on a choice of chart ϕ. We now return and to rectify this problem, by applying to the notion of *orientation*.

Recall that, for $M \subset \mathbb{R}^n$ a k-submanifold, and $\phi : U \to M$ a chart, we defined the integral of $f : M \to \mathbb{R}$ over $A \subset \phi(U)$ to be

$$\int_{\phi^{-1}(A)} f dV := \int_{\phi^{-1}(A)} (f \circ \phi)\sqrt{\det(g)}dx.$$

We showed that for a diffeomorphism $\psi : V \to U$ with $\det(J\psi) > 0$, then the value of this integral is the same, regardless of whether we compute using ϕ or $\phi \circ \psi$. This gives us a hint as to the structure we need to define integrals on all of M.

Definition 5.4. Two charts $\phi : U \to M$ and $\psi : V \to M$ are said to be *consistently oriented* if the diffeomorphism

$$\phi^{-1} \circ \psi : \psi^{-1}(\phi(U) \cap \psi(V)) \longrightarrow \phi^{-1}(\phi(U) \cap \psi(V))$$

is orientation preserving, i.e., if $\det(J(\phi^{-1} \circ \psi)) > 0$. An *orientation* of M is a collection of consistently oriented charts $\mathcal{U} := \{(\phi_\alpha, U_\alpha)\}_{\alpha \in I}$ such that every point $p \in M$ is contained in at least one chart $\phi_\alpha(U_\alpha)$. We call M together with a choice of orientation an *oriented submanifold*. We will call a chart $\phi : U \to M$ which is consistently oriented with every chart in $\mathcal{U}$ a *oriented chart*, and we will implicitly assume that $\mathcal{U}$ contains *every* oriented chart.

Proposition 5.5. *Let $M \subset \mathbb{R}^n$ be an oriented manifold and $f : M \to \mathbb{R}$ a continuous function. Then the integral*

$$\int_M f dV$$

is independent of the choice of oriented charts used to compute it.

Proof. This is a corollary of Proposition 3.8. □

We have previously defined orientability in terms of the Gauß map. The following proposition shows that these two notions are the same.

Proposition 5.6. *Let $M \subset \mathbb{R}^{k+1}$ be a hypersurface. The following types of data are equivalent:*

1. *A choice of orientation on M.*

2. *A global Gauß map $n : M \to S^k$.*

Proof. First, suppose that we have an orientation on M. Given an oriented chart $\phi : U \to M$, the vectors $\partial_1\phi, \dots, \partial_k\phi$ provide a basis of T_pM at every point $p \in \phi(U)$. We can thus determine a *unique* unit normal n on $\phi(U)$ by requiring that $\{\partial_1\phi, \dots, \partial_k\phi, n\}$ is a positively oriented basis of $\mathbb{R}^{k+1} \cong T_p\mathbb{R}^{k+1}$.

To show that this is independent of the choice oriented chart, we note that, given a diffeomorphism $\psi : V \to U$, the bases $\partial_i\phi$ and $\partial_i(\phi \circ \psi)$ are related by $J\psi$. Thus, $\{\partial_1\phi, \dots, \partial_k\phi, n\}$ is positively oriented if and only if $\{\partial_1(\phi \circ \psi), \dots, \partial_k(\phi \circ \psi), n\}$ is positively oriented. As such, the unit normals defined by two different oriented charts agree, and so we can define a unit normal globally on M.

On the other hand, suppose that we have a global Gauß map n on M. Then we define an orientation $\mathcal{U}$ on M by including only those charts ϕ such that $\{\partial_1\phi, \dots, \partial_k\phi, n\}$ is a positively oriented basis of $\mathbb{R}^{k+1}$. As above, this immediately implies that the Jacobians relating two charts in $\mathcal{U}$ have positive determinant.

These two constructions are clearly inverse, completing the proof. □

For hypersurfaces, we can thus think about an orientation as a global Gauß map.

3 Compactness

Up until now, we have required very general conditions on our submanifolds — a hyperplane in $\mathbb{R}^n$ was as good as an $(n-1)$-sphere. However, our main result in this chapter pertains to *integration*, and so we need to ensure a certain finiteness going forward. The condition we will use to this end is called *compactness* and more commonly belongs to the realm of topology. Intuitively, the compact manifolds are those which "close up". For instance, among surfaces in $\mathbb{R}^3$, the sphere an torus are compact, while every plane is not compact. More formally:

Definition 5.7. We call a subset $S \subset \mathbb{R}^n$ *compact* if, for every infinite sequence $\{p_i\}_{i\in\mathbb{N}}$ of points in M, there is an infinite subsequence which converges to a point in M.

Notice that this is strictly stronger than requiring that every sequence converges, as the following examples serve to illustrate.

Example 5.8.

1. The submanifold $\mathbb{R}^n \subset \mathbb{R}^n$ is not compact. The sequence given by

   $$p_n = n(1, 0, 0, \ldots, 0)$$

   has no convergent subsequence.

2. The open ball $B_1(0) \subset \mathbb{R}^n$ is not compact. Consider the sequence

   $$p_n = (1 - \frac{1}{n}, 0, 0, \ldots, 0)$$

   in $B_1(0)$. This sequence converges to $(1, 0, \ldots, 0)$, and thus so does every subsequence. However, $(1, 0, \ldots, 0)$ does not lie in $B_1(0)$.

A related, but weaker, notion is that of a closed subset.

Definition 5.9. We say that a subset $S \subset \mathbb{R}^n$ is *closed* for every sequence $\{p_k\}_{k \in \mathbb{N}}$ which lies in S, if p_k converges to a point $p \in \mathbb{R}^n$, then $p \in S$.

This is an easier definition to work with, because continuous functions send limits to limits.

Example 5.10. The n-sphere $S^n \subset \mathbb{R}^{n+1}$ is closed. To see this, notice that we can define a continuous function

$$\begin{aligned} f : \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sum (x^i)^2 \end{aligned}$$

such that $S^n = f^{-1}(1)$. Suppose that p_n is a sequence of points in S^n with limit p. Then we have

$$f(p) = f(\lim_{n \to \infty} p_n) = \lim_{n \to \infty} f(p_n) = \lim_{n \to \infty} 1 = 1$$

so that $p \in S^n$.

More generally, any subset of $\mathbb{R}^n$ defined as a level set of a continuous function is closed. As a result, for instance, the torus $T^2 \subset \mathbb{R}^3$ is also closed.

Example 5.11. The plane P given by $z = 0$ in $\mathbb{R}^3$ is closed. Given any convergent sequence $(a_n, b_n, 0)$ in P, we see that the limit must be of the form $(a, b, 0)$, and thus also contained in p.

To show that a subset S is compact seems rather harder than showing that S is closed. However, it turns out that closure and compactness are closely related.

Theorem 5.12 (Heine-Borel). *A subset $S \subset \mathbb{R}^n$ is compact if and only if S is closed, and there is a number $C > 0$ such that, for all $x \in S$, $|x| < C$. We say that such a set is closed and bounded.*

We will not give a proof of the Heine-Borel Theorem here, but the interested student can find proofs in any of a dozen references.[1] As a consequence, we immediately see that the torus, the n-sphere, and numerous other examples are compact.

[1] Or on Wikipedia.

But why to we care about compactness? There are many reasons compactness appears in mathematics, but in our case it is a way to guarantee that integrals over our submanifolds will be finite.

Lemma 5.13. *Suppose that $S \subset \mathbb{R}^n$ is a compact subset, and that $f : S \to \mathbb{R}$ is a continuous function. Then f is bounded on S.*

Proof. Suppose, to the contrary, that f is unbounded above.[2] Then, for every $m \in \mathbb{N}$, the set

[2] An identical proof works in the case where f is unbounded below.

$$\{p \in S \mid f(p) > m\}$$

is non-empty. Choose a sequence $\{p_m\}_{m \in \mathbb{N}}$ of points in S such that $f(p_m) > m$. Since S is compact, there is some convergent subsequence $\{p_{m_i}\}_{i \in \mathbb{N}}$, with

$$\lim_{i \to \infty} p_{m_i} = p \in S.$$

However, this would imply

$$\lim_{i \to \infty} m_i \leq \lim_{i \to \infty} f(p_{m_i}) = f(p).$$

Since the former limit goes to ∞, this is a contradiction. □

Corollary 5.14. *Let $S \subset \mathbb{R}^n$ be a compact subset, and let $f : S \to \mathbb{R}$ be a continuous function. Then*

$$\int_S f dx < \infty$$

whenever the former is well-defined.

Proof. Denote by C an upper bound for $|f(x)|$ on C. Then the value of the integral is bounded above by $\mathrm{Vol}(S) \cdot C$. Since S is a bounded set, its volume is bounded, completing the proof. □

We will make use of the following fact without proof. The interested student can find the proof in most books on point-set topology.

Lemma 5.15. *Let $S \subset \mathbb{R}^n$ be a compact subset. Let $\{U_\alpha\}_{\alpha \in I}$ be a (possibly infinite) collection of open subsets of S such that*

$$\bigcup_{\alpha \in I} U_\alpha = S.$$

Then there is a finite set of indices $\alpha_1, \ldots, \alpha_m$ such that

$$\bigcup_{i=1}^{m} U_{\alpha_i} = S.$$

Theorem 5.16. *Let $M \subset \mathbb{R}^n$ be a compact k-submanifold, and let $f : M \to \mathbb{R}$ be a continuous function on M. Then*

$$\int_M f dV < \infty.$$

Sketch. Firstly, since M compact and f is continuous, f is bounded, i.e. $|f| < C$.

Now we want to subdivide M into pieces with which we can compute the integral. For each $p \in M$, choose a coordinate chart $\phi_p : B_2(0) \to M$ such that $\phi_p(0) = p$. Then the restrictions $\phi_p|_{B_1(0)}$ cover M. By compactness, there is a finite subset of these charts $\phi_1, \ldots, \phi_r$ such that

$$M = \bigcup_{i=1}^{r} \phi_i(B_1(0)).$$

Subdivide M as a disjoint union

$$M = \bigcup_{j=1}^{\ell} A_j$$

where each A_j lies entirely in $\phi_{c(j)}(B_1(0))$ for $1 \leq c(j) \leq r$. We then have

$$\int_M f dV = \sum_{j=1}^{\ell} \int_{A_j} (f \circ \phi_{c(j)}) \sqrt{\det(g)} dx$$

Since $\sqrt{\det(g)}$ defines, via ϕ_i, a continuous function on $B_2(0)$, it is in particular a continuous function on $\overline{B_1(0)}$. We thus see $\sqrt{\det(g)}$ is bounded on $B_1(0)$, say by β_i. Thus, we have

$$\begin{aligned}\int_M f dV &= \sum_{j=1}^{\ell} \int_{A_j} (f \circ \phi_{c(j)}) \sqrt{\det(g)} dx \\ &\leq \sum_{j=1}^{\ell} C \beta_{c(j)} < \infty\end{aligned}$$

as desired. □

4 *Line integration*

We now return to integrals of and along curves. Recall that we defined the length of a smooth curve $\gamma : [a, b] \to M \subset \mathbb{R}^n$ to be

$$L(\gamma) = \int_a^b |\gamma'(t)| dt = \int_a^b \sqrt{g_{i,j} \frac{d\rho^i}{dt} \frac{d\rho^j}{dt}} dt$$

where $\gamma = \phi \circ \rho$ for some chart ϕ. We showed that the value of this integral is independent of orientation-preserving reparameterization of γ, and since the first expression is entirely independent of the chart, the second does not depend on the chart. We now want to generalize this to define integrals of functions or vector fields along a curve.

The factor $|\gamma'|$ in the integral is precisely what causes this expression to be invariant under reparameterization. More conceptually, including the speed in the integral ensures that small subdivisions in the corresponding Riemann sum are given a value approximately equal to the length of a small segment of the curve in $\mathbb{R}^n$.

Definition 5.17. Let $f : M \to \mathbb{R}$ be a continuous function, and X a smooth tangent field on M. The *line integral of f along* $\gamma : [a, b] \to M$ is the quantity

$$\int_\gamma f ds = \int_a^b f(\gamma(t)) |\gamma'(t)| dt = \int_a^b f(\gamma(t)) \sqrt{g_{i,j} \frac{d\rho^i}{dt} \frac{d\rho^j}{dt}} dt.$$

The *line integral of X along γ* is the quantity

$$\int_\gamma X \cdot ds = \int_a^b \langle X_{\gamma(t)}, \gamma'(t)\rangle dt = \int_a^b g_{i,j} X^i \frac{d\rho^j}{dt} dt$$

Our first order of business is to check that neither of these notions depend on orientation-preserving reparameterizations of γ. Notice that, once again, the first (extrinsic) formula in each case shows that the integral in the second formula does not depend on the choice of chart.

Exercise 14. Show that the line integrals in the definition above are independent of orientation preserving reparameterizations of γ. Show further that orientation-reversing reparameterizations only reverse the sign.

Remark 5.18. Very often, given coordinates $x^1, \dots, x^k$ on M, a line integral of a vector field X along γ is written as

$$\int_\gamma (g_{i,j} X^j dx^i)$$

under the implicit convention that $dx^i = \frac{d\rho^i}{dt} dt$. In two dimensions, for instance, one might write

$$\int_\gamma g_{1,j} X^j dx^1 + g_{2,j} X^j dx^2.$$

The meaning of such a notation can be made more precise using differential forms.

There is also an immediate relation between our two definitions.

Lemma 5.19. *Define a function f on the image of γ by*

$$f(\gamma(t)) = \frac{\langle X_{\gamma(t)}, \gamma'(t)\rangle}{|\gamma'(t)|}$$

i.e., the projection of $X_{\gamma(t)}$ onto $\gamma'(t)$.[3] *Then*

$$\int_\gamma f ds = \int_\gamma X \cdot ds.$$

Conversely, given a function f on the image of γ, define a vector field $X = f(\gamma(t))\frac{\gamma'(t)}{|\gamma'(t)|}$ along γ. Then

$$\int_\gamma X \cdot ds = \int_\gamma f ds.$$

[3] Or the component of X in the γ-direction.

In the special case of surfaces, we have another kind of line integral, which will be key to our version of the divergence theorem. We will need to return to the notion of *orientation* to define this second kind of line integral.

Definition 5.20. Let $M \subset \mathbb{R}^3$ be an oriented surface with Gauß map n, and let $\gamma : [a, b] \to M$ be a smooth regular curve. Suppose that ϕ is an oriented chart containing the image of γ, and X is a vector field along γ. We define

$$\int_\gamma X \cdot ds^\perp := \int_a^b \langle X, \gamma' \times n\rangle \, dt$$

We will sometimes refer to this integral as the *(rightward) flow of X across γ*.

As before, the definition does not involve charts, and so the quantity in question is independent of chart. Before deriving a coordinate formula for $\int_\gamma X \cdot ds^\perp$, we show that it does not depend on the choice of parameterization of γ.

Lemma 5.21. *The integral*

$$\int_\gamma X \cdot ds^\perp$$

is invariant under orientation-preserving reparameterizations of γ.

Proof. We first notice that, by the symmetries of the triple product in $\mathbb{R}^3$, we have

$$\langle X, \gamma' \times n \rangle = \langle \gamma', n \times X \rangle$$

If $\gamma = \rho(u(t))$, where u is an orientation-preserving reparameterization of $\mathbb{R}$, then we see that

$$\begin{aligned}
\int_a^b \langle \gamma', n \times X \rangle dt &= \int_a^b \left\langle \rho'(u(t))\frac{du}{dt}, n \times X \right\rangle dt \\
&= \int_a^b \langle \rho'(u(t)), n \times X \rangle \frac{du}{dt} dt \\
&= \int_{u(a)}^{u(b)} \langle \rho', n \times X \rangle du
\end{aligned}$$

and so the integral is invariant under orientation-preserving reparameterizations of γ. $\square$

We now aim to derive a coordinate form of the integral

$$\int_\gamma X \cdot ds^\perp.$$

To do this, we need only find a coordinate form for the triple product $\langle X, \gamma' \times n \rangle$.

The problem we encounter, in a nutshell, is that the triple product is easy to compute when our vectors are expressed in terms of an orthonormal basis, but the basis $\Phi := \{\partial_1 \phi, \partial_2 \phi, n\}$ is not necessarily orthonormal. To rectify this, we simply transform Φ into an orthonormal basis $E = \{e_1, e_2, e_3\}$ using the Gram-Schmidt process.

We first set

$$e_1 = \frac{\partial_1 \phi}{|\partial_1 \phi|} = \frac{\partial_1 \phi}{\sqrt{g_{1,1}}}$$

We then compute

$$\begin{aligned}
u_2 &= \partial_2 \phi - \langle \partial_2 \phi, e_1 \rangle e_1 \\
&= \partial_2 \phi - \frac{1}{g_{1,1}} \langle \partial_2 \phi, \partial_1 \phi \rangle \partial_1 \phi \\
&= \partial_2 \phi - \frac{g_{1,2}}{g_{1,1}} \partial_1 \phi
\end{aligned}$$

We then compute

$$|u_2|^2 = g_{2,2} - \frac{g_{1,2}^2}{g_{1,1}} = \frac{1}{g_{1,1}} \det(g)$$

Thus,

$$e_2 = \frac{\sqrt{g_{1,1}}}{\sqrt{\det g}} \left(\partial_2 \phi - \frac{g_{1,2}}{g_{1,1}} \partial_1 \phi \right)$$

Finally, since n was already orthogonal to the other two, and is a unit vector, we have

$$e_3 = n.$$

The corresponding change-of-basis matrix is

$$A = \begin{pmatrix} \frac{1}{\sqrt{g_{1,1}}} & -\frac{g_{1,2}}{\sqrt{g_{1,1}\det(g)}} & 0 \\ 0 & \frac{\sqrt{g_{1,1}}}{\sqrt{\det(g)}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

That is, given a vector $v \in T_pM$, and writing v_E and v_Φ for the expressions for v with respect to the bases E and Φ respectively, we have

$$Av_E = v_\Phi.$$

We can use A to compute triple products in T_pM. Given three vectors u, v, w, the triple product can be computed as

$$\langle u, v \times w \rangle = \det\left(u_E \mid v_E \mid w_E\right)$$

since E is orthonormal. Thus, we can compute

$$\begin{aligned} \det\left(u_E \mid v_E \mid w_E\right) &= \det\left(A^{-1}u_\Phi \mid A^{-1}v_\Phi \mid A^{-1}w_\Phi\right) \\ &= \det(A^{-1})\det\left(u_\Phi \mid v_\Phi \mid w_\Phi\right) \end{aligned}$$

It is easy to see that the determinant of A is

$$\det(A) = \frac{1}{\sqrt{\det(g)}},$$

and so we obtain:

Proposition 5.22. *With respect to a coordinate chart $\phi : U \to M$,*

$$\int_\gamma X \cdot ds^\perp = \int_a^b \sqrt{\det(g)}\left(-X^2\frac{d\rho^1}{dt} + X^1\frac{d\rho^2}{dt}\right)dt.$$

Proof. By our previous work we see that

$$\begin{aligned} \langle X, \gamma' \times n \rangle &= \sqrt{\det(g)}\det\begin{pmatrix} X^1 & \frac{d\rho^1}{dt} & 0 \\ X^2 & \frac{d\rho^2}{dt} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \sqrt{\det(g)}\left(-X^2\frac{d\rho^1}{dt} + X^1\frac{d\rho^2}{dt}\right) \end{aligned}$$

precisely as desired. □

5 *Geodesic coordinates and Geodesic curvature revisited*

We now specialize back to the case of surfaces. In this context, we have a number of useful computational tools at our disposal, which we will develop in this and the following sections.

The first of these tools is a type of coordinate system which dramatically simplifies the form of the metric g. These *geodesic coordinates* have the key property the metric g is diagonal. It will turn out that we can always choose geodesic coordinates.

Definition 5.23. Let $M \subset \mathbb{R}^3$ be a surface. We call a chart $\phi : U \to M$ with coordinates (u^1, u^2) a *geodesic (orthogonal) coordinate chart* if

1. With respect to the coordinate basis $\partial_1\phi\ \partial_2\phi$ the metric has the form
$$g = \begin{pmatrix} 1 & 0 \\ 0 & g_{2,2} \end{pmatrix}$$
2. The coordinate curves with constant u^2 are geodesics.

We have already met geodesic coordinates: the polar coordinates we defined via the exponential map. We will use without proof the following lemma, which can be proven directly using the theory of partial differential equations.

Lemma 5.24. *Let $M \subset \mathbb{R}^3$ be a surface, and $p \in M$ a point. There is a geodesic orthogonal coordinate chart $\phi : U \to M$ containing p.*

Exercise 15. Show that, with respect to a geodesic orthogonal chart $\phi : U \to M$,
$$\Gamma^1_{1,1} = \Gamma^1_{1,2} = \Gamma^2_{1,1} = 0$$
and
$$\Gamma^2_{1,2} = \frac{\partial}{\partial u^1} \ln(\sqrt{g_{2,2}}).$$
Further show that, with respect to geodesic orthogonal coordinates, the Gaußian curvature is given by
$$K = -\frac{1}{\sqrt{g_{2,2}}} \frac{\partial^2}{\partial (u^1)^2} \left(\sqrt{g_{2,2}}\right).$$

One of the further useful characteristics of geodesic orthogonal coordinates is that they define a canonical *orthonormal frame*: a pair of tangent vector field on $\phi(U)$ which, at every point p in $\phi(U)$, determine an orthonormal basis of T_pM. In general, when working with geodesic coordinates, we will set
$$E_1 = \partial_1\phi \qquad E_2 = \frac{1}{|\partial_2\phi|}\partial_2\phi$$
and use this as a canonical orthonormal frame for our computations.

We now want to explore how we can compute the geodesic curvature of a curve in geodesic coordinates. Before we can do this, we need to refine our definition of geodesic curvature in two dimensions.

Definition 5.25. Let $M \subset \mathbb{R}^3$ be an oriented surface with Gauß map n, and $\gamma : [a, b] \to M$ a unit speed curve in M. The *geodesic curvature* of γ is the unique function κ_g defined along γ such that
$$\nabla_{\gamma'}\gamma' = \kappa_g(n \times \gamma')$$

This merely amounts to adding a sign to κ_g, depending on whether the curve γ is turning "right" or "left" with respect to the orientation of M.

We now aim to understand how we can express κ_g using geodesic coordinates. First, we warm up with an exercise:

Exercise 16. Define

$$\begin{aligned} f : \mathbb{R} &\longrightarrow S^1 \subset \mathbb{R}^2 \\ t &\longmapsto (\cos(t), \sin(t)) \end{aligned}$$

and let $\mu : [a, b] \to S^1$ be a smooth function.

1. Show that, given $\tilde{p} \in \mathbb{R}$ and a semi-circle $U \subset S$ containing $p := f(\tilde{p})$, there is a *unique* open interval $(c, c+\pi) \subset \mathbb{R}$ containing p such that

$$f|_{(c,c+\pi)} : (c, c+\pi) \longrightarrow U$$

   is a diffeomorphism.

2. Show that there is a smooth function $\theta : [a, b] \to \mathbb{R}$ such that

$$\mu(t) = (\cos(\theta(t)), \sin(\theta(t))).$$

   Further show that any two such functions differ by a constant multiple of 2π.

Proposition 5.26. *Let $M \subset \mathbb{R}^3$ be an oriented surface with Gauß map $n : M \to S^2$. Let $\phi : U \to M$ be an oriented geodesic coordinate chart with coordinates (u^1, u^2) on U. Let $\gamma : [a, b] \to M$ be a unit speed curve which factors through ϕ as $\gamma = \phi \circ \rho$. Then*

1. *There is a smooth function*

$$\theta : [a, b] \longrightarrow \mathbb{R}$$

   such that

$$\gamma'(t) = \cos(\theta(t))E_1 + \sin(\theta(t))E_2.$$

2. *If θ and $\overline{\theta}$ are two functions satisfying part (1), then the difference $\theta(t) - \overline{\theta}(t)$ is constant on a multiple of 2π.*

3. *The geodesic curvature is given by*

$$\kappa_g = \frac{d\theta}{dt} + \left(\frac{\partial}{\partial u^1}\sqrt{g_{2,2}}\right)\frac{d\rho^2}{dt}$$

Proof. We first notice that the frame $\{E_1, E_2\}$ allows us to uniquely identify $T_pM \cong \mathbb{R}^2$ by sending

$$v \longmapsto (\langle v, E_1\rangle, \langle v, E_2\rangle).$$

More generally, this gives a diffeomorphism

$$\begin{aligned} \psi : TM|_{\phi(U)} &\longrightarrow \phi(U) \times \mathbb{R}^2. \\ (p, v) &\longmapsto (p, (\langle v, E_1\rangle, \langle v, E_2\rangle)) \end{aligned}$$

We can thus consider $\psi(\gamma'(t))$ as giving a smooth map $[a,b] \to \mathbb{R}^2$ which takes values in $S^1 \subset \mathbb{R}^2$. The first part of our proposition then follows immediately from Exercise 16.

For part (3), we compute the curvature explicitly. First note that

$$\nabla_{\gamma'}\gamma' = \theta'(t)\,(n \times \gamma') - \sin(\theta(t)E_1 + \cos(\theta(t))E_2) + \cos(\theta(t))\nabla_{\gamma'}E_1 + \sin(\theta(t))\nabla_{\gamma'}E_2$$

Moreover, taking the cross product using the orthonormal basis $\{E_1, E_2, n\}$ yields

$$n \times \gamma' = -\sin(\theta(t))E_1 + \cos(\theta(t))E_2$$

so that

$$\langle \nabla_{\gamma'}\gamma', n \times \gamma' \rangle = \theta'(t) + \langle \cos(\theta(t))E_2) + \cos(\theta(t))\nabla_{\gamma'}E_1 + \sin(\theta(t))\nabla_{\gamma'}E_2, n \times \gamma' \rangle$$

We are thus left to compute the inner products $\langle \nabla_{\gamma'}E_i, E_j \rangle$. However, if we differentiate the relation

$$\langle E_i, E_j \rangle = \delta_{i,j}$$

we obtain

$$\langle \nabla_{\gamma'}E_i, E_j \rangle = \langle E_i, \nabla_{\gamma'}E_j \rangle$$

so that our expression for the geodesic curvature simplifies to

$$\kappa_g = \theta'(t) + \langle \nabla_{\gamma'}E_1, E_2 \rangle.$$

Moreover, we see that

$$\nabla_{\gamma'}E_1 = \frac{d\rho^i}{dt}\Gamma^k_{1,i}\partial_k\phi$$

and, using the computation of the Christoffel symbols in geodesic coordinates from Exercise 15, we see that this implies

$$\langle \nabla_{\gamma'}E_1, E_2 \rangle = \frac{d\rho^2}{dt}\frac{1}{2\sqrt{g_{2,2}}}\frac{\partial g_{2,2}}{\partial u^1}$$

completing the computation. □

We can interpret this as follows:

1. The term $\theta'(t)$ is simply the angular rate of change of the unit tangent vector of γ. If we were working in the Euclidean plane $\mathbb{R}^2$ this really would be the curvature.

2. The term
$$\frac{d\rho^2}{dt}\left(\frac{\partial}{\partial u^1}\sqrt{g_{2,2}}\right)$$
compensates for the failure of the surface M to be flat. In this sense it can be taken to measure how much of $\theta'(t)$ consists of changes in a non-tangential direction.

6 Polygons & triangulations

We now would like to divide our manifold into pieces which are easier to integrate over. The pieces we will choose are polygons, and their boundaries will be composed of collections of smooth curves.

Definition 5.27. A *convex combination* of points[4] $x_1, \dots x_n \in \mathbb{R}^2$ is a sum

$$\sum_{i=1}^{n} \lambda_i x_i$$

where $\sum_i \lambda_i = 1$ and $0 \leq \lambda_i \leq 1$ for every $1 \leq i \leq n$. The *convex hull* of a subset $S = \{x_1, \dots, x_n\} \subset \mathbb{R}^2$ is the set of *all* convex combinations of the points in S, i.e.

$$\mathrm{Conv}(S) := \left\{ \sum_{i=1}^{n} \lambda_i x_i \middle| \sum_{i=1}^{n} \lambda_i = 1 \quad \text{and} \quad 0 \leq \lambda_i \leq 1 \right\}$$

A n-*gon* (or *polygon* with n sides) in $\mathbb{R}^2$ is the convex hull of a set S such that

1. The cardinality of S is n.

2. For every $x \in S$, $x \notin \mathrm{Conv}(S \setminus \{x\})$.

We call the elements of S the *vertices* of the n-gon $P = \mathrm{Conv}(S)$.

Exercise 17. Let $x_1, \dots, x_k \in \mathbb{R}^n$ be a set of points. A *convex combination* of $x_1, \dots, x_k$ is a sum

$$\sum_{i=1}^{k} \lambda_i x_i$$

where $0 \leq \lambda_i \leq 1$ and

$$\sum_{i=1}^{k} \lambda_i = 1.$$

The *convex hull* of the set $S = \{x_1, \dots, x_k\}$ is the set $\mathrm{Conv}(S) \subset \mathbb{R}^n$ of all convex combinations of points in S.

1. Show that, for any $y, z \in \mathrm{Conv}(S)$, the line segment $\overline{yz}$ from y to z lies in $\mathrm{Conv}(S)$.

2. Show that $\mathrm{Conv}(S)$ is compact.

Exercise 18. Let $S \subset \mathbb{R}^2$, and let $x \in \mathrm{Conv}(S)$. We say that a line

$$L = \{v \in \mathbb{R}^2 \mid \langle v - x, n \rangle = 0\}$$

through x in $\mathbb{R}^2$ *supports* S when $y \in \mathrm{Conv}(S)$ implies that $\langle y - x, n \rangle \geq 0$.

1. Let $S \subset \mathbb{R}^2$. Show that for any $y \notin \mathrm{Conv}(S)$, there is a line $L = \{\langle v - x, n \rangle = 0\}$ which supports S and such that $\langle y - x, n \rangle < 0$. (Hint: let x be the point of $\mathrm{Conv}(S)$ whose distance to y is minimal.)

[4] The basic idea of a convex combination is iterative. For the first iteration, take $L_1(S)$ to be the set of all points which lie on line segments between points in S. This will have the form

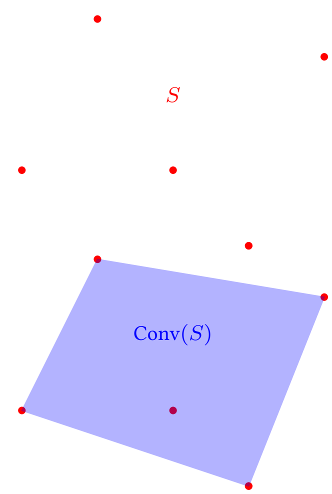
$$L_1(S) = \{tx_i + (1-t)x_j \mid t \in [0,1]\} \subset \mathrm{Conv}(S).$$

One then takes $L_2(S) = L_1(L_1(S))$, so on and so forth. Then

$$\mathrm{Conv} = \bigcup_{i=1}^{\infty} L_i(S).$$

The basic idea is that a point in a convex combination is something like a point on a segment between points on segments ... between points on segments between points in S.

Pictorially, the convex hull of a set $S \subset \mathbb{R}^2$ looks like



Note that the set pictured is *not* convex independent.

2. We say that S is *convex independent* if, for every $x_i \in S$, $x_i \notin \operatorname{Conv}(S \setminus \{x_i\})$. If $S \subset \mathbb{R}^2$ is convex independent, show that, for every three distinct indices i, j, k, the vectors $x_j - x_i$ and $x_k - x_i$ are linearly independent.

3. Let $S \subset \mathbb{R}^2$ be a convex independent subset. Show that if
$$x = \sum_{i=1}^{\ell} \lambda_i x_i$$
is a convex combination in which at least three of the λ_i are non-zero, then there is some $\epsilon > 0$ such that $B_\epsilon(x) \subset \operatorname{Conv}(S)$.

Definition 5.28. Let $U \subset \mathbb{R}^n$ be a closed subset. We say that a point $x \in U$ is

1. An *interior point* if there is an $\epsilon > 0$ such that $B_\epsilon(x) \subset U$.

2. A *boundary point* if for every $\epsilon > 0$, $B_\epsilon(x) \cap (\mathbb{R}^n \setminus U) \neq \varnothing$.

We write the set of interior points of U as $\mathring{U}$, and call this the *interior* of U. We denote the set of boundary points of U by ∂U, and call this the *boundary* of U. Note that U is the disjoint union of its interior and its boundary.

We will make use of the following proposition without proof:

Proposition 5.29. *Let $S \subset \mathbb{R}^2$ be a convex independent set with $|S| \geq 3$.*

1. *A point $x \in \operatorname{Conv}(S)$ lies in $\partial \operatorname{Conv}(S)$ if and only if there is a line through x supporting S.*

2. *There is an ordering $x_1, \ldots, x_k$ of the elements of S such that the boundary $\partial \operatorname{Conv}(S)$ consists of precisely the line segments $\overline{x_1 x_2}, \overline{x_2 x_3}, \ldots, \overline{x_k x_1}$.*

Definition 5.30. Let $\gamma : [a, b] \to M \subset \mathbb{R}^3$ be a continuous curve. We call γ

1. *piecewise smooth* If there are values
$$a = t_0 < t_1 < \ldots < t_k = b$$
such that $\gamma|_{[t_i, t_{+1}]}$ is smooth for each i;

2. *piecewise linear* If there are values
$$a = t_0 < t_1 < \ldots < t_k = b$$
such that $\gamma|_{[t_i, t_{+1}]}$ is linear for each i;

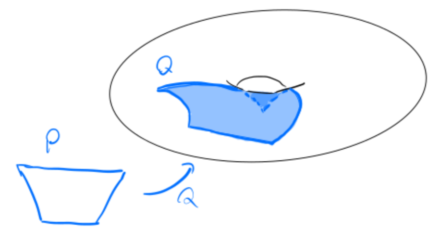
3. *closed* if $\gamma(a) = \gamma(b)$; or

4. *simple closed* if γ is closed, and $\gamma(t) = \gamma(s)$ if and only if either $t = s$ or $t = a$ and $s = b$.

Notation 5.31. Notice that the boundary of an n-gon P can be thought of as a simple, closed, piecewise linear curve. We will always parameterize ∂P *counter-clockwise.* We will call the lines $\overline{x_i x_{i+1}}$ which comprise the boundary the *edges* of P, and the points x_i the *vertices* of P.

Our reason for defining convex n-gons in the plane is to talk about polygons in surfaces.

Pictorially, a polygon in a surface is precisely what one would expect:



Definition 5.32. Let $M \subset \mathbb{R}^3$ be an oriented surface. A *polygon in* M consists of an n-gon $P = \mathrm{Conv}(S) \subset \mathbb{R}^2$ and a smooth, regular, injective map $Q : P \to M$. We will abuse notation by writing Q for the image of P under the map Q, and by writing ∂Q for the image of ∂P under Q. We call a polygon Q in M *oriented* if Q is compatible with the orientation of M.

Polygons in surfaces are particularly nice subsets to integrate over. Since the boundary of a polygon has codimension 1, we have

$$\int_{\partial Q} f dV = 0.$$

Thus, if we can cover a surface M by polygons which only overlap on their boundaries, we can compute integrals over M by summing the integrals over the polygons. To make this precise, we make the following definition:

Definition 5.33. A *polygonal subdivision* $\mathcal{T}$ of an oriented surface $M \subset \mathbb{R}^3$ is a finite set $\{Q_i\}_{i=1}^r$ of oriented polygons in M with the following properties

- The polygons cover M, i.e.,
$$\bigcup_{i=1}^r Q_i = M$$
- For any two polygons Q_i and Q_j with $i \neq j$, the intersection $Q_i \cap Q_i$ is a single edge of Q_i and Q_j.

We call the polygons Q_i the *faces* of $\mathcal{T}$, the collection of all edges of polygons in $\mathcal{T}$ the *edges* of the subdivision, and the collection of all vertices of polygons in $\mathcal{T}$ the *vertices* of the subdivision.

We will make use of the following theorem without proof.

Theorem 5.34. *Every oriented surface admits a polygonal subdivision.*

One key ingredient in the Gauß-Bonnet Theorem is the *Euler Characteristic* of a subdivision $\mathcal{T}$. We will eventually prove that the Euler Characteristic doesn't depend on the choice of subdivision, but rather only on the oriented surface M.

Definition 5.35. Let $M \subset \mathbb{R}^3$ be an oriented surface, and let $\mathcal{T}$ be a subdivision of M. The *Euler characteristic of* $\mathcal{T}$ is equal

$$\chi(\mathcal{T}) = V - E + F$$

where V is the number of vertices of $\mathcal{T}$, E the number of edges, and F the number of faces.

7 *The Divergence Theorem*

We now come to a technical result we will make use of in our proof of Gauß-Bonnet: *the Divergence Theorem.* Once definitions are unwound, we will see that this is actually a corollary of Green's Theorem from multivariable calculus.

Before we can get to the Divergence Theorem, however, we must define the divergence. We will do this in arbitrary dimensions, before again specializing to surfaces for the Divergence Theorem.

Let $M \subset \mathbb{R}^n$ be a k-submanifold, and let X be a vector field on M. At any point $p \in M$, we can use the covariant derivative and X to define a linear map

$$\begin{aligned} C_X : T_pM &\longrightarrow T_pM \\ V &\longmapsto (\nabla_V X)_p \end{aligned}$$

What does this linear map represent? It sends a direction to the (tangential) rate of change of X in that direction. The linear map thus encodes the rates of change of X in all directions.

Before moving on, let us note that this really is well-defined: Given vector fields W and X, the covariant derivative

$$(\nabla_W X)_p = W^i_p(\frac{dX^j}{dx^i})_p \partial_j \phi + W^i_p X^j_p \Gamma^k_{i,j} \partial_k \phi$$

only depends on the value W_p of W at p. Thus, we get a well-defined map of tangent spaces. It is easy to see that the covariant derivative $\nabla_V X$ is linear in V.

Definition 5.36. The *divergence of* X at $p \in M$ is the trace of C_X, i.e.,

$$\text{div}(X)(p) := \text{tr}(C_X).$$

Heuristically, this should measure "how much X is flowing outwards at p."

Our first order of business is to give a coordinate expression for the divergence.

Lemma 5.37. *With respect to a chart $\phi : U \to M$, the divergence of X is given by*

$$\text{div}(X) = \frac{1}{\sqrt{\det(g)}} \sum_{i=1}^{k} \frac{\partial}{\partial x^i} \left(X^i \sqrt{\det(g)} \right)$$

Proof. We note that, with respect to the coordinate basis, the diagonal entries of the matrix representing C_X are

$$\frac{dX^i}{dx^i} + X^j \Gamma^i_{i,j}.$$

The definition of the divergence thus yields

$$\text{div}(X) := \sum_{i=1}^{k} \left(\frac{dX^i}{dx^i} + X^j \Gamma^i_{i,j} \right)$$

We then note that by Lemma 3.27,

$$\Gamma^i_{i,j} = \frac{1}{2} g^{i,\ell} \left(\frac{\partial g_{j,\ell}}{\partial x^i} + \frac{\partial g_{\ell,i}}{\partial x^j} - \frac{\partial g_{i,j}}{\partial x^\ell} \right)$$

Summing over i, the first and last terms cancel each other out, and we are left with

$$\sum_{i=1}^{k} X^j \Gamma^i_{i,j} = \frac{1}{2} X^j g^{i\ell} \frac{\partial g_{\ell,i}}{\partial x^j} = \frac{1}{2} X^j \operatorname{tr}\left(g^{-1} \frac{\partial g}{\partial x^j}\right)$$

By Jacobi's formula[5] we find that

$$\frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial x^j} \sqrt{\det(g)} = \frac{1}{2} \operatorname{tr}\left(g^{-1} \frac{\partial g}{\partial x^j}\right)$$

[5] This says that, for an $n \times n$ invertible matrix $A(t)$ which is a function of a real variable t,

$$\frac{d}{dt} \det(A(t)) = \det(A(t)) \operatorname{tr}\left(A(t)^{-1} \frac{dA}{dt}\right)$$

So we see that

$$\operatorname{div}(X) = \frac{\partial X^i}{\partial x^i} + \frac{X^i}{\sqrt{\det(g)}} \frac{\partial}{\partial x^i} \sqrt{\det(g)} = \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial x^i} \left(X^i \sqrt{\det(g)}\right)$$

as desired. □

Remark 5.38. Notice that when our manifold is just $\mathbb{R}^k$, the metric is the identity matrix, so the divergence simplifies to

$$\operatorname{div}(X) = \sum_{i=1}^{k} \frac{\partial}{\partial x^i} X^i.$$

This is precisely the usual notion of the divergence of a vector field on $\mathbb{R}^k$.

Example 5.39 (Key example)**.** Let $M \subset \mathbb{R}^3$ be a surface, and let $\phi : U \to M$ be a geodesic coordinate chart, i.e., a chart such that

$$g = \begin{pmatrix} 1 & 0 \\ 0 & g_{2,2} \end{pmatrix}$$

Consider the vector field

$$X = -\left(\frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^1} \sqrt{\det(g)}\right) \partial_1 \phi$$

on $\phi(U)$. We then take the divergence of X, yielding

$$\operatorname{div}(X) = \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^1} \left(\frac{\partial}{\partial u^1} \sqrt{\det(g)}\right)$$

$$\operatorname{div}(X) = \frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^1} \left(\frac{1}{2\sqrt{\det(g)}} \operatorname{tr}\left(g^{-1} \frac{dg}{dt}\right)\right)$$

Exercise sheet 9 then shows that the divergence of this vector field is the Gaußian curvature K. Notice that the vector field X may have a very different expression in other coordinate systems, even other geodesic coordinate systems.

We now can state the Divergence Theorem.

Theorem 5.40 (The Divergence Theorem)**.** *Let $Q : P \to M$ be an oriented polygon in M and let X be a smooth tangent field on M. Then*

$$\int_Q \operatorname{div}(X) dV = \int_{\partial Q} X \cdot ds^{\perp}.$$

Proof. We use Q as the chart under which we compute the integral. Let (u^1, u^2) be the coordinates on P. The left-hand side is then

$$\begin{aligned}\int_Q \operatorname{div}(X) dV &= \int_P \left(\frac{1}{\sqrt{\det(g)}} \frac{\partial}{\partial u^i} \left(X^i \sqrt{\det(g)} \right) \right) \sqrt{\det(g)} du \\ &= \int_P \frac{\partial}{\partial u^i} \left(X^i \sqrt{\det(g)} \right) du.\end{aligned}$$

If we define smooth functions Y_1 and Y_2 on P by

$$Y_1 = -X^2 \sqrt{\det(g)} \qquad Y_2 = X^1 \sqrt{\det(g)}$$

This expression then becomes

$$\int_P \frac{\partial Y_2}{\partial u^1} - \frac{\partial Y_1}{\partial u^2} du$$

Since we are integrating over a polygon in the plane, we can apply Green's Theorem to obtain

$$\int_{\partial P} Y_1 du^1 + Y_2 du^2 = \sum_{\ell=1}^{n} \int_{a_\ell}^{b_\ell} (-X^2 \frac{d\rho_\ell^1}{dt} + X^1 \frac{d\rho_\ell^2}{dt}) \sqrt{\det(g)} dt$$

However, this latter expression is simply the coordinate expression for

$$\int_{\partial Q} X \cdot ds^\perp$$

from Proposition 5.22. □

8 *Tangents and loops*

Our last preliminary before we prove the Gauß-Bonnet Theorem is the *Theorem of Turning Tangents*. The setup for this theorem is the following:

- $M \subset \mathbb{R}^3$ is an oriented surface.
- $Q : P \to M$ is an oriented polygon in M, contained in a geodesic coordinate chart ϕ.
- $\rho_i : [t_i, t_{i+1}] \to P$ are (counterclockwise) parameterizations of the edges of P, and $\gamma_i = \phi \circ \rho_i$ are the corresponding parameterizations of the edges of Q. We will assume that γ_i is unit speed.
- $\theta_i : [t_i, t_{i+1}] \to \mathbb{R}$ are the smooth functions guaranteed by Proposition 5.26, such that
$$\gamma_i'(t) = \cos(\theta_i(t)) E_1 + \sin(\theta_i(t)) E_2.$$
- For each vertex of P, $\alpha_i \in [-\pi, \pi]$ is the corresponding *external angle*, i.e., the angle such that
$$\cos(\alpha_i) = \langle \gamma_i'(t_{i+1}), \gamma_{i+1}'(t_{i+1}) \rangle$$

Theorem 5.41 (Turning tangents)**.** *The equality*

$$\sum_{i=1}^{\ell} \int_{t_i}^{t_{i+1}} \theta_i'(t) dt = 2\pi - \sum_{i=1}^{\ell} \alpha_i$$

holds.

Proof. We will first prove the theorem when the metric is the identity. In this case, ϕ strictly preserves angles, and so the theorem becomes a statement about the polygon P. In this case, each of the functions $\theta_i(t)$ is necessarily constant, since the ρ_i are straight lines. Thus, the statement amounts to the claim that

$$\sum_{i=1}^{\ell} \alpha_i = 2\pi.$$

However, the internal angles $\pi - \alpha_i$ of the ℓ-gon P add up to $\pi(\ell - 2)$. As such, we see that

$$\ell\pi - \sum_{i=1}^{\ell} \alpha_i = \pi\ell - 2\pi$$

and so

$$\sum_{i=1}^{\ell} \alpha_i = 2\pi.$$

We now define matrix-valued functions on P using the metric g with respect to ϕ:

$$G(r) = \begin{pmatrix} 1 & 0 \\ 0 & r + (1-r)g_{2,2} \end{pmatrix}$$

for $r \in [0, 1]$. We define $\theta_i(r, t)$ and $\alpha_i(r)$ by using G in place of the metric, i.e. we set

$$E_1(r) = \partial_1\phi \qquad E_2(r) = \frac{1}{\sqrt{r + (1-r)g_{2,2}}}\partial_2\phi$$

and define $\theta_i(r, t)$ to be the function such that

$$\cos(\theta_i(r,t))) = \frac{1}{|\gamma_i'|_G}\left\langle \frac{d\gamma_i}{dt}, E_1(r)\right\rangle_G = \frac{1}{|\gamma_i'|_G} G(r)_{1,j}\frac{d\rho_i^j}{dt} = \frac{1}{|\gamma_i'|_G}\frac{d\rho_i^1}{dt}$$

and

$$\sin(\theta_i(r,t))) = \frac{1}{|\gamma_i'|_G}\left\langle \frac{d\gamma_i}{dt}, E_2(r)\right\rangle_G = \frac{1}{|\gamma_i'|_G} G(r)_{2,j}\frac{d\rho^j}{dt} = \frac{1}{|\gamma_i'|_G}\frac{1}{\sqrt{r+(1-r)g_{2,2}}}\frac{d\rho_i^2}{dt}.$$

We define $\alpha_i(r)$ similarly.

Notice that, by construction,[6]

$$\theta_i(r, t_{i+1}) - \theta_{i+1}(r, t_{i+1}) + \alpha_i(r)$$

is some multiple of π. We thus can choose our functions θ_i, θ_{i+1}, and α_i so that the sums

$$\theta_i(r, t_{i+1}) - \theta_{i+1}(r, t_{i+1}) + \alpha_i(r)$$

are constant in r.

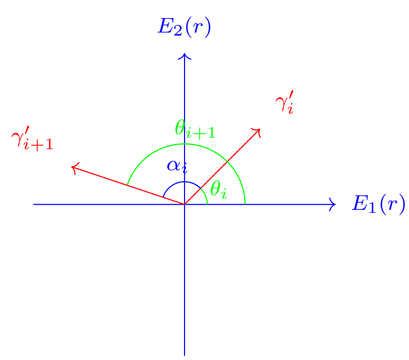
We then reformulate the integral we are interested in:

$$\begin{aligned}\sum_{i=1}^{\ell}\int_{t_i}^{t_{i+1}} \theta_i'(r,t)dt + \sum_{i=1}^{\ell}\alpha_i(r) &= \sum_{i=1}^{\ell}(\theta_i(r,t_{i+1}) - \theta_i(r,t_i) + \alpha_i(r)) \\ &= \sum_{i=1}^{\ell}(\theta_i(r,t_{i+1}) - \theta_{i+1}(r,t_{i+1}) + \alpha_i(r))\end{aligned}$$

[6] This is because $E_1(r)$, $E_2(r)$, $\gamma_i'(t_{i+1})$ and $\gamma_{i+1}'(t_{i+1})$ are vectors in T_pM, equipped with the inner product $G(r)$. With respect to this inner product,

- $\alpha_i(r)$ is the angle between $\gamma_i'(t_{i+1})$ and $\gamma_{i+1}'(t_{i+1})$
- $\theta_i(r, t_{i+1})$ is the angle between $E_1(r)$ and $\gamma_i'(t_{i+1})$
- $\theta_{i+1}(r, t_{i+1})$ is the angle between $E_1(r)$ and $\gamma_{i+1}'(t_{i+1})$.

Schematically:



where the last step follows by cyclically shifting terms downwards in the sum. Hwoever, each of these terms is constant in r, and so the value

$$\sum_{i=1}^{\ell}\int_{t_i}^{t_{i+1}}\theta_i'(r,t)dt+\sum_{i=1}^{\ell}\alpha_i(r)$$

is constant in r. When $r=1$, we saw that this value is 2π. When $r=0$, this is precisely the integral of the theorem. Thus, taking $r=0$, we find

$$\sum_{i=1}^{\ell}\int_{t_i}^{t_{i+1}}\theta_i'(t)dt=2\pi-\sum_{i=1}^{\ell}\alpha_i$$

precisely as desired. □

9 *The Gauß-Bonnet Theorem*

We have now reached our goal for this chapter: The Gauß-Bonnet Theorem. This theorem gives an *invariant* of oriented surfaces by integrating curvatures over a subdivision of the surface.

Theorem 5.42 (Gauß-Bonnet)**.** *Let $\mathcal{T}=\{Q_1,\dots,Q_k\}$ be a polygonal subdivision of a compact oriented surface M. Then*

$$\int_M KdV=2\pi\chi(\mathcal{T})$$

where K denotes the Gaußian curvature of M, and $\chi(\mathcal{T})$ is the Euler characteristic.

We will prove this theorem in steps, first proving more local results about individual polygons in M, and then piecing them together to prove the Gauß-Bonnet Theorem.

We first consider an oriented polygon $Q:P\to M$ in M. As usual, we will parameterize the boundary ∂P counterclockwise, as the piecewise linear curve determined by the linear curves $\rho_1,\dots,\rho_n$ in $\mathbb{R}^2$. We will denote the resulting curves in M by $\gamma_i=Q\circ\rho_i$.

STEP 1: We first define *external angles* at each corner of Q. At the vertex $p_i\in Q$ where γ_i and γ_{i+1} meet, we have two tangent vectors, γ_i' and γ_{i+1}'. We define the *external angle between γ_i and γ_{i+1}* to be th $\alpha_i\in[-\pi,\pi]$ such that

$$\cos(\alpha_i)=\frac{1}{|\gamma_i'||\gamma_{i+1}'|}\langle\gamma_i',\gamma_{i+1}'\rangle.$$

Using this first step, we can state a local form of the Gauß-Bonnet Theorem

Proposition 5.43 (Local form of the Gauß-Bonnet Theorem)**.** *With the setup above, the relation*

$$\int_Q KdV+\int_{\partial Q}\kappa_g ds+\sum\alpha_i=2\pi$$

holds.

We will prove this proposition in two steps: first making a simplifying assumption on coordinate charts, and then making no assumption.

TEMPORARY ASSUMPTION: We first assume that Q factors through a single geodesic coordinate chart ϕ, i.e. that $Q = \phi \circ \mu$ for some smooth, regular, orientation-preserving, injective map $\mu : P \to U \subset \mathbb{R}^2$.

Proof (under our assumption). We consider the vector field X on $\phi(U)$ given by

$$X = -\left(\frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial u^1}\sqrt{\det(g)}\right)\partial_1\phi.$$

By Example **??**,

$$\operatorname{div}(X) = K$$

is the Gaußian curvature of M. The divergence theorem thus tells us that

$$\int_Q K dV = \int_Q \operatorname{div}(X) dV = \int_{\partial Q} X \cdot ds^{\perp}.$$

our special case will thus follow by computing

$$\int_{\partial Q} X \cdot ds^{\perp}.$$

Writing this integral in coordinates, and applying Proposition 5.26, we obtain

$$\begin{aligned}\int_{\partial Q} X \cdot ds^{\perp} &= -\int_{\partial P}\sqrt{\det(g)}\left(\frac{\partial}{\partial u^1}\sqrt{g}\right)\frac{d\rho^2}{dt}dt \\ &= -\int_{\partial Q}\kappa_g ds + \int_{\partial P}\theta'(t)dt.\end{aligned}$$

We thus need only to show

$$\int_{\partial P}\theta'(t)dt = 2\pi - \sum \alpha_i.$$

However, this is precisely the statement of Theorem 5.41. ☐

We now relax our assumption, no longer requiring that Q be contained in a single geodesic coordinate chart. In this case, we will reduce to the previous case by subdividing the polygon P. Visually, given a polygon P, we can break it into two smaller polygons P_1 and P_2 which share a single edge and two vertices:

If we keep subdividing P, we can eventually get the individual polygons to be small enough that they lie in a geodesic coordinate system. Thus, the local Gauß-Bonnet theorem will follow, by induction, by showing that if it holds for P_1 and P_2 above, it holds for P.

Proof (of local Gauß-Bonnet). We show that if the theorem holds for two small polygons Q_1 and Q_1 which share a single edge, and whose union is a larger polygon Q, then it holds for Q. We first compute

$$\int_Q KdV = \int_{Q_1} KdV + \int_{Q_2} KdV = 4\pi - \int_{\partial Q_1} \kappa_g ds - \int_{\partial Q_2} \kappa_g ds - \sum \alpha_i - \sum \beta_i$$

by applying Gauß-Bonnet on the two smaller polygons. The edge shared by Q_1 and Q_2 appears in each of the two boundary integrals with opposite sign, so

$$\int_{\partial Q_1} \kappa_g ds + \int_{\partial Q_2} \kappa_g ds = \int_{\partial Q} \kappa_g ds.$$

Furthermore, denoting the exterior angles of Q by ζ_i, and those of Q_1 and Q_2 by α_i and β_i, respectively, at each shared vertex v we have

$$\alpha_v + \beta_v - \zeta_v = \pi.$$

At each unshared vertex, we have that ζ_v is either α_v or β_v, depending on whether v is a vertex of Q_1 or Q_2. Thus

$$\sum \alpha_i + \sum \beta_i = 2\pi + \sum \zeta_i.$$

We therefore see that

$$\int_Q KdV = 2\pi - \int_{\partial Q} \kappa_g ds - \sum \zeta_i$$

as desired. □

Finally, we prove the *global* Gauß-Bonnet Theorem.

Proof (Gauß-Bonnet). We now let $\mathcal{T}$ be a subdivision of a compact oriented surface $M \subset \mathbb{R}^3$. By the local form of Gauß-Bonnet

$$\int_M KdV = \sum_{Q \in \mathcal{T}} \int_Q KdV = 2\pi|\mathcal{T}| - \sum_{Q \in \mathcal{T}} \int_{\partial Q} \kappa_g ds - \sum_{Q \in \mathcal{T}} \sum_{v \in V_Q} \alpha_v.$$

We will consider each of the terms on the left-hand side separately.

Each edge of the subdivision occurs as a boundary edge of one of the polygons, each time with opposite orientation. Thus, the term

$$\sum_{Q \in \mathcal{T}} \int_{\partial Q} \kappa_g ds$$

vanishes.

To analyze the term

$$\sum_{Q \in \mathcal{T}} \sum_{v \in V_Q} \alpha_v$$

Suppose that k polygons of $\mathcal{T}$ meet at the vertex v. For each polygon Q, write the *internal* angle at v as

$$\beta_v^Q = \pi - \alpha_v^Q$$

Summing all of the internal angles around v yields

$$2\pi = \sum_{Q \text{s.t.} v \in Q} \beta_v^Q = \pi\#(\text{edges attached to v}) - \sum_{Q \text{s.t.} v \in Q} \alpha_v^Q$$

Thus we see that

$$\sum_{Q \in \mathcal{T}} \sum_{v \in V_Q} \alpha_v = \pi(2E_{\mathcal{T}}) - 2\pi(V_{\mathcal{T}}).$$

Putting this all together we get

$$\int_M K dV = 2\pi F_{\mathcal{T}} - 2\pi E_{\mathcal{T}} + 2\pi V_{\mathcal{T}} = 2\pi\chi(\mathcal{T})$$

completing the proof. ☐

Corollary 5.44. *The Euler characteristic is independent of the choice of polygonal subdivision.*

This justifies the following definition:

Definition 5.45. The *Euler characteristic* $\chi(M)$ of a surface $M \subset \mathbb{R}^3$ is the Euler characteristic of any polygonal subdivision of M.

Definition 5.46. Let $M \subset \mathbb{R}^3$ and $N \subset \mathbb{R}^3$ be surfaces. A *diffeomorphism* between M and N is a map $f : M \to N$ such that f and f^{-1} are both smooth. If there is a diffeomorphism from M to N, then we call M and N *diffeomorphic* and write $M \cong N$.

Remark 5.47. Diffeomorphisms in a sense "preserve the smooth structure" of surfaces, but do *not* need to preserve any of the geometric structure. In particular, diffeomorphisms do not preserve the first fundamental form, or any of the curvatures.

Proposition 5.48. *The Euler characteristic is a diffeomorphism invariant. That is, if $M \cong N$, then $\chi(M) = \chi(N)$.*

Proof. Notice that, if $f : M \to N$ is a diffeomorphism and $\mathcal{T} = \{Q_i\}_{i=1}^r$ is a subdivision of M, then $f(\mathcal{T}) = \{f \circ Q_i\}_{i=1}^r$ is a subdivision of N. It is not hard to check that $\chi(\mathcal{T}) = \chi(f(\mathcal{T}))$. ☐

Corollary 5.49. *The integral*

$$\int K dV$$

is a diffeomorphism invariant.

Example 5.50. Let us consider the sphere $S^2 \subset \mathbb{R}^3$ and the ellipsoid E^2 given by the equation

$$a^2x^2 + a^2y^2 + z^2 = 1$$

for some chosen $a > 0$.

I claim that these two surfaces are diffeomorphic. To see this, notice that the map

$$\begin{aligned} f : \quad \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x^1, x^2, x^3) &\longmapsto (\tfrac{1}{a}x^1, \tfrac{1}{a}x^2, x^3) \end{aligned}$$

defines a diffeomorphism (indeed, a linear isomorphism) from $\mathbb{R}^3$ to $\mathbb{R}^3$, and sends S^2 bijectively to E^2. Thus, the surfaces are diffeomorphic. However, the Gaußian curvatures of these two surfaces are quite different. We know that, for the sphere, the Gaußian curvature is constant on 1, so let's compute the Gaußian curvature of E^2.

Away from the poles, E^2 is the surface of rotation with profile curve

$$f(u) = \sqrt{\frac{1-u^2}{a^2}}$$

We can thus compute the Gaußian curvature using the formula

$$K = -\frac{f''(u)}{f(u)(f'(u)+1)^2}$$

from Exercise sheet 7. We compute

$$f'(u) = -\frac{u^2}{a^2 f(u)}$$

and

$$f''(u) = \frac{f(u)}{(u^2-1)^2}$$

So that the Gaußian curvature of E^2 is

$$K = -\frac{\frac{f(u)}{(u^2-1)^2}}{f(u)\left(1-\frac{u^2}{a^2 f(u)}\right)^2} = \frac{1}{(a^2 f(u) - u^2)^2}$$

which is clearly not constant.

We have one final corollary about the Euler characteristic.

Corollary 5.51. *Suppose that M is a surface such that $\chi(M) < 0$. Then there are points on M where the Gaußian curvature is negative.*

6

Differential Forms

So far, we have managed to avoid ever explicitly talking about differential forms. However, differential forms are a key modern[1] innovation in the study of calculus on manifolds, and so it is now time for us to begin exploring them. The basic idea is that differential k-forms are "k-dimensional measuring tools". A 1-form is a way of measuring lengths (a yardstick, so to speak), a 2-form is a way of measuring areas, and so on. From a related perspective, a k-form can be seen as a "thing that can be integrated over a k-dimensional submanifold."

[1] Relatively speaking. The formal theory of differential forms is sometimes said to have begun with Élie Cartan's Sur certaines expressions différentielles et le problème de Pfaff in 1899.

Both of these heuristic ideas are realized in a common algebraic framework. Our first task in this section will be to build up this framework on $\mathbb{R}^n$. Once we have done this, we will proceed to defining differential forms on manifolds, and then to applications.

1 *Rewriting integration*

Before getting to the definition of forms, let's take a step back, and try to remember what we are doing when we are integrating something. The basic idea we will work with, familiar from high-school calculus, is that of a Riemann sum. Suppose first we have a function $f : \mathbb{R} \to \mathbb{R}$. When we integrate f on the interval $[a, b]$, we take some subdivisions of $[a, b]$ into N little pieces $[t_i, t_{i+1}]$, and then take the limit of Riemann sums

$$\int_a^b f dt = \lim_{N\to\infty} \sum_{i=1}^{N} f(t_i)\Delta_i t,$$

where $\Delta_i t = t_{i+1} - t_i$.

If we take this into n-dimensions, we can suppose that we have a curve $\gamma : [a, b] \to \mathbb{R}^n$, and want to integrate a tangent vector field X along γ. In this case, we can write our integral as a limit

$$\int \gamma f = \lim_{N\to\infty} \sum_{i=1}^{N} \left\langle X_{\gamma(t_i)}, \Delta_i\gamma \right\rangle$$

where $\Delta_i\gamma = \gamma(t_{i+1}) - \gamma(t_i)$.

How do we make sense of the expression $\left\langle X_{\gamma(t_i)}, \Delta_i\gamma \right\rangle$? Well, the first step is to reinterpret how we think of $\Delta_i\gamma$. *A priori*, this is simply a vector $\gamma(t_{i+1}) - \gamma(t_i)$. However, under the identification $T_{\gamma(t_i)}\mathbb{R}^n \cong \mathbb{R}^n$, we can think of this as a *tangent vector at* $\gamma(t_i)$, which, as the mesh of the limit we are taking gets finer, becomes close to being a tangent vector to γ. We can thus think of the expression

$$\langle X_p, -\rangle$$

as being a linear map $T_p\mathbb{R}^n \to \mathbb{R}$. That is, a linear map which takes a tangent vector at p and gives us a scalar.

On the other hand, suppose that, for each $p \in \mathbb{R}^n$, we have a linear map $\omega_p : T_p\mathbb{R}^n \to \mathbb{R}$, and that the map ω which sends $p \mapsto \omega_p$ is in some sense smooth.[2] Then along any curve γ, we can define an integral

[2] We'll get to what this means later.

$$\int_\gamma \omega = \lim_{N\to\infty} \sum_{i=1}^{N} \omega_{\gamma(t_i)}(\Delta_i\gamma) = \int_a^b \omega_{\gamma(t)}(\gamma'(t))dt.$$

The second of these equalities requires some justification, but it should not be hard to convince yourself that it holds.

The fact that this map is *linear* in the tangent vector is very useful. In particular, it should be easy to convince yourself using linearity that the following expected integral equalities hold:[3]

[3] Here we write $\overline{\gamma}$ for the curve which traces γ backwards.

$$\int_\gamma (\omega + \nu) = \int_\gamma \omega + \int_\gamma \nu$$
$$\int_{\overline{\gamma}} \omega = -\int_\gamma \omega = \int_\gamma (-\omega)$$
$$\int_\gamma (c\omega) = c\int_\gamma \omega$$

where γ is a curve, $-\gamma$ is the curve that traces γ backwards, ω and ν are our chosen "smooth assignments of linear maps $T_p\mathbb{R}^n \to \mathbb{R}$", and $c \in \mathbb{R}$ is a constant.

We can turn this reformulation into a definition.

Definition 6.1. A *covector* at $p \in \mathbb{R}^n$ is a linear map

$$\omega_p : T_p\mathbb{R}^n \longrightarrow \mathbb{R}.$$

The *cotangent space* to $\mathbb{R}^n$ at p is the $\mathbb{R}$-vector space

$$T_p^*\mathbb{R}^n := \mathrm{Lin}(T_p\mathbb{R}^n, \mathbb{R}).$$

Our next goal is to formalize what it means to assign a covector in a *smooth* way to every point in $\mathbb{R}^n$. We will call such an assignment ω a *(smooth) 1-form.*

First, let's try to think about what the space $\mathrm{Lin}(T_p\mathbb{R}^n, \mathbb{R})$ is in terms of coordinates. If $x^1, \ldots, x^n$ are our coordinates on $\mathbb{R}^n$, we have a corresponding basis of tangent fields $\partial_{x^1}, \ldots, \partial_{x^n}$ on $\mathbb{R}^n$. At each $p \in \mathbb{R}^n$, we can take the *dual basis* of $\{\partial_{x^1}, \ldots, \partial_{x^n}\}$. The dual basis consists of the linear maps

$$\begin{aligned} \mathrm{d}x^i_p : T_p\mathbb{R}^n &\longrightarrow \mathbb{R} \\ \partial_{x^j} &\longmapsto \delta^i_j \end{aligned}$$

where δ^i_j is the Krönecker delta. From linear algebra, this is a basis of T_p^*M. As such, we have an identification $T_p^*M \cong \mathbb{R}^n$.

Definition 6.2. We define the *cotangent bundle* of $\mathbb{R}^n$ to be the space

$$T^*\mathbb{R}^n \cong \{(p, \nu_p) \mid \nu_p \in T_p^*\mathbb{R}^n\} \cong \mathbb{R}^n \times \mathbb{R}^n$$

and we view it as a smooth manifold via the latter bijection. Notice that $T^*\mathbb{R}^n$ comes equipped with a smooth projection $\pi : T^*\mathbb{R}^n \to \mathbb{R}^n$.

A *(smooth) 1-form on* $\mathbb{R}^n$ is a smooth map

$$\omega : \mathbb{R}^n \longrightarrow T^*\mathbb{R}^n$$

such that $\pi \circ \omega = \mathrm{id}_{\mathbb{R}^n}$. That is, such that $\omega_p \in T_p^*\mathbb{R}^n$.

Example 6.3. In particular, we have the coordinate 1-forms $\mathrm{d}x^i$ which assign to every point $p \in \mathbb{R}^n$ the covector $\mathrm{d}x_p^i$. By construction, for any smooth 1-form ω on $\mathbb{R}^n$, there are unique smooth functions $\omega_i : \mathbb{R}^n \to \mathbb{R}$ such that

$$\omega = \omega_i \mathrm{d}x^i = \omega_1 \mathrm{d}x^1 + \cdots + \omega_n \mathrm{d}x^n,$$

where we again use the Einstein summation convention.

Since the entire point of defining 1-forms was that they should be "things we can integrate along curves", let's think about what the integral of such a 1-form along a curve $\gamma : [a, b] \to \mathbb{R}^n$ should be. From our definition, we should have

$$\begin{aligned} \int_\gamma \omega &= \int_\gamma \omega_i \mathrm{d}x^i \\ &= \int_a^b \omega_i(\gamma(t)) \mathrm{d}x^i\,(\gamma')\, dt. \end{aligned}$$

But we can write $\gamma'(t) = \frac{d\gamma^i}{dt} \partial_{x^i}$, so by the definition of $\mathrm{d}x^i$ we have

$$\int_a^b \omega_i(\gamma(t)) \mathrm{d}x^i\,(\gamma')\, dt = \int_a^b \omega_i(\gamma(t)) \frac{d\gamma^i}{dt} dt$$

In particular, if we are simply integrating the form $\omega = f(x)\mathrm{d}x^1$, then we are, effectively, integrating f along γ in only the x^i-direction.

Remark 6.4. It is natural at this point to ask "What about arc-length integrals?" If 1-forms are "things we can integrate along curves", shouldn't the arc-length integral also be a 1-form? The answer sadly, is no, and comes down to a subtlety in our definition. The notion of a 1-form is fundamentally bound up with the *linearity* of path integrals in the tangent vector $\gamma'(t)$. In particular, given a 1-form ω, we want

$$\int_{\overline{\gamma}} \omega = -\int_\gamma \omega,$$

a condition which is embodied by the fact that $\omega(p) : T_p\mathbb{R}^n \to \mathbb{R}$ is *linear.* However, for an arc-length integral, we have

$$\begin{aligned}\int_{\overline{\gamma}} f ds &= \int_a^b f(\gamma(t))|\overline{\gamma}'(t)|dt \\ &= \int_a^b f(\gamma(t))|-\gamma'(t)|dt \\ &= \int_a^b f(\gamma(t))|\gamma'(t)|dt \\ &= \int_\gamma f ds,\end{aligned}$$

so the desired linearity does not hold. What this means is that 1-forms formalize notions of integration *which depend on orientation*. This will be of crucial importance when we discuss higher forms.

Notation 6.5 (Important special case)**.** Consider a curve $\gamma : [a, b] \hookrightarrow \mathbb{R}$ which is simply the inclusion of the interval $[a, b]$ into $\mathbb{R}$, and a 1-form

$$\omega = f(t)\mathrm{d}t$$

on $\mathbb{R}$. We write

$$\int_{[a,b]} \omega := \int_\gamma \omega := \int_a^b f(t)dt$$

for the corresponding path integral. Since 1-forms defined on $[a, b]$ are in one-to-one correspondence with smooth real-valued functions on $[a, b]$, this is simply the usual notion of an integral over an interval.

Definition 6.6. Consider a smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$. Let $\omega \in T_p^*\mathbb{R}^m$ be a covector on $\mathbb{R}^m$. The *pullback of* ω *along* ϕ is the 1-form $(\phi^*)_p\omega$ defined by

$$(\phi_p^*\omega)(v) = \omega(d\phi_p(v))$$

for any tangent vector $v \in T_p\mathbb{R}^n$. Note that this defines a linear map

$$\phi_p^* : T^*_{\phi(p)}\mathbb{R}^m \longrightarrow T_p^*\mathbb{R}^n$$

for each $p \in \mathbb{R}^n$. Given a smooth 1-form ν on $\mathbb{R}^m$, we will denote by $\phi^*\nu$ the 1-form with

$$(\phi^*\nu)(p) = \phi_p^*(\nu(p)).$$

We will shortly show that this is a smooth 1-form on $\mathbb{R}^n$.

Lemma 6.7. *Let* $\phi : \mathbb{R}^n \to \mathbb{R}^m$ *and* $\psi : \mathbb{R}^m \to \mathbb{R}^k$ *be smooth maps, and write* x^i, y^i, *and* z^i *for the coordinates on* $\mathbb{R}^n$, $\mathbb{R}^m$, *and* $\mathbb{R}^k$ *respectively. For any* $p \in \mathbb{R}^n$, *then*[4]

$$(\psi \circ \phi)_p^* = \phi_p^* \circ \psi^*_{\phi(p)}.$$

[4] Mathematicians in my discipline would say the the pullback is *functorial*.

Proof. It suffices to check that the two sides agree on an arbitrary covector $\omega \in T_{\psi(\phi(p))}\mathbb{R}^k$, i.e., that for any such ω,

$$(\psi \circ \phi)_p^*(\omega) = \phi_p^*\left(\psi^*_{\phi(p)}(\omega)\right).$$

However, this is a statement about the equality of linear maps, so it further suffices to check that this holds for any vector $v \in T_p\mathbb{R}^n$. We then compute

$$\begin{aligned}(\psi \circ \phi)_p^*(\omega)(v) &= \omega\left(d(\psi \circ \phi)_p(v)\right) \\ &= \omega\left(d\psi_{\phi(p)}\left(d\phi_p(v)\right)\right) \\ &= \psi_{\phi(p)}^*(\omega)\left(d\phi_p(v)\right) \\ &= \phi_p^*\left(\psi_{\phi(p)}^*(\omega)\right)(v).\end{aligned}$$

as such, the two sides of the equation agree on every tangent vector, and must describe the same map, as desired. ☐

Proposition 6.8. *Consider a smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$, and write x^i and y^j for the coordinates on the source and target of ϕ, respectively. Let*

$$\omega = \omega_i \mathrm{d}y^i$$

be a smooth 1-form on the target. Let $\gamma : [a, b] \to \mathbb{R}^n$ be a smooth curve.

1. *With respect to the coordinates x^i, we have*

$$\phi^* \mathrm{d}y^i = \frac{\partial \phi^i}{\partial x^\ell} \mathrm{d}x^\ell$$

or, more generally

$$\phi^*(\omega) = (\omega_i \circ \phi)\frac{\partial \phi^i}{\partial x^\ell} \mathrm{d}x^\ell.$$

2. *For a smooth 1-form*

$$\nu = \nu_i \mathrm{d}x^i$$

on $\mathbb{R}^n$,

$$\int_\gamma \nu = \int_{[a,b]} \phi^* \nu.$$

In particular, we have

$$\int_{\phi\circ\gamma} \omega = \int_\gamma \phi^* \omega$$

Proof. We first prove (1). Again, we test against an arbitrary tangent vector

$$V = V^i \partial_{x^i}.$$

Note that

$$d\phi(V) = V^i \frac{\partial \phi^j}{\partial x^i} \partial_{y^j}.$$

We then compute (suppressing the point p in our computations)

$$\begin{aligned}\phi^*(\omega)(V) &= \omega(d\phi(V)) \\ &\omega\left(V^i \frac{\partial \phi^j}{\partial x^i} \partial_{y^j}\right) \\ &= \omega_k \mathrm{d}y^k \left(V^i \frac{\partial \phi^j}{\partial x^i} \partial_{y^j}\right) \\ &= \omega_k V^i \frac{\partial \phi^j}{\partial x^i} \delta_j^k \\ &= \omega_k V^i \frac{\partial \phi^k}{\partial x^i}.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}(\omega_i \circ \phi)\frac{\partial \phi^i}{\partial x^\ell}\mathrm{d}x^\ell\left(V^k \partial_{x^k}\right) &= \omega_i \frac{\partial \phi^i}{\partial x^\ell} V^k \delta_k^\ell \\ &= \omega_i V^\ell \frac{\partial \phi^i}{\partial x^\ell}.\end{aligned}$$

Reindexing, we see that these expressions are equal.

We now turn out attention to (2). Let u denote the coordinate on $[a, b]$. By definition

$$\begin{aligned}\int_\gamma \nu &= \int_a^b \nu(\gamma'(t))dt \\ &= \int_a^b \nu(d\gamma_t(\partial_u))dt \\ &= \int_a^b (\gamma^*\nu)(\partial_u)dt \\ &= \int_{[a,b]} \gamma^*\nu\end{aligned}$$

as desired. To see that this implies our more general statement, note that

$$\begin{aligned}\int_{\phi\circ\gamma} \omega &= \int_{[a,b]} (\phi \circ \gamma)^* \omega \\ &= \int_{[a,b]} \gamma^* (\phi^*\omega) \\ &= \int_\gamma \phi^*\omega\end{aligned}$$

completing the proof. □

Example 6.9. Let us consider the curve

$$\begin{aligned}\gamma : \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ \theta &\longmapsto (\sin(\theta), \cos(\theta))\end{aligned}$$

tracing out the unit circle in $\mathbb{R}^2$. First consider a 1-form

$$\omega = \omega_1 \mathrm{d}x^1 + \omega_2 \mathrm{d}x^2$$

on $\mathbb{R}^2$. We can pull this back to a 1-form $\gamma^*\omega$ on $\mathbb{R}$, explicitly given by

$$\gamma^*\omega = \omega_i \frac{\partial \gamma^i}{\partial \theta}\mathrm{d}\theta = (-\omega_1 \sin(\theta) + \omega_2 \cos(\theta))\,\mathrm{d}\theta.$$

We can then notice that the coefficient of $\mathrm{d}\theta$ is a 2π-periodic function, regardless of the form we started with. Our takeaway from this is that, if we want something to integrate over the circle which is (1) smooth and (2) globally defined, then it must be defined by a smooth function on the circle.[5]

[5] There are many ways of phrasing this observation, with varying levels of technology involved in the statement. Using terminology we haven't yet defined, we could say (in increasing order of technicality)

- Smooth 1-forms on S^1 are in bijection with smooth functions on S^1.
- The cotangent bundle of S^1 is a trivial vector bundle.

Time permitting, we will return to this example, and unpack both of these rephrasings.

2 2-forms and k-forms

Our work so far as been interesting, but from a practical standpoint, it is unsatisfying to have a theory of integration which can only handle 1-dimensional integrals. To rectify this, let us try to consider what a *2-form* might be.

From the name, we expect that we will be trying to define "things we can integrate over 2-dimensional submanifolds." We could try to repeat our analysis of Riemann sums from the first section, but instead, let's try to apply broad principles to figure out what a 2-form should be.

1. Firstly, we thought of a covector ω as a sort of "yardstick" by which we could measure tangent vectors. Loosely speaking, we think of tangent vectors at p as "infinitesmal displacements" at p, and the covector ω tells us how big of a contribution such an infinitesmal displacement should make to the integral. In our new, two dimensional setting, we don't just want to measure "infinitesmal displacements" at p, but rather, we want to measure "infinitesmal areas" at p. To make this formal, we think of an "infinitesmal area element" at p to be the parallelogram in $T_p\mathbb{R}^n$ defined by a *pair* of tangent vectors. We thus can presume that a 2-form ν should take *two* tangent vectors as arguments, and should give us a number representing "how much the infinitesmal area element contributes to the integral". That is, we would expect that a 2-form ν must define a map
$$T_p\mathbb{R}^n \times T_p\mathbb{R}^n \longrightarrow \mathbb{R}.$$
2. The map associated to the 2-form ν at p should represent something like "measure of area times a function value". As a result, it should be additive and scale like a measurement of the area of parallelograms in $T_p\mathbb{R}^n$. In particular, we expect that it should be linear in each argument, i.e., a bilinear map
$$T_p\mathbb{R}^n \times T_p\mathbb{R}^n \longrightarrow \mathbb{R}$$
or, equivalently, a linear map
$$T_p\mathbb{R}^n \otimes T_p\mathbb{R}^n \longrightarrow \mathbb{R}.$$
3. Per Remark 6.4, the notions of integration that we are formalizing depend on the orientation. This tells us that, given two tangent vectors $v, w \in T_p\mathbb{R}^n$, we should have $\nu(v, w) = -\nu(w, v)$. Another way to see this is that our map is bilinear, and $\nu(v, v)$ should always equal zero, since the corresponding parallelogram always is zero.

We thus have arrived at something resembling a working definition.

Definition 6.10. A *2-form* at $p \in \mathbb{R}^n$ is map
$$\nu(p) : T_p\mathbb{R}^n \times T_p\mathbb{R}^n \longrightarrow \mathbb{R}$$
which is

1. *bilinear*, i.e. linear in each argument; and

2. *alternating*, i.e., $\nu(v, w) = -\nu(w, v)$.

More generally, the same line of reasoning yields

Definition 6.11. A k-*form* at $p \in \mathbb{R}^n$ is map

$$\nu(p) : \underbrace{T_p\mathbb{R}^n \times \cdots \times T_p\mathbb{R}^n}_{\times k} \longrightarrow \mathbb{R}$$

which is

1. k-*multilinear*, i.e. linear in each argument; and

2. *alternating*, i.e.,

$$\nu(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\nu(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$$

To turn this definition into something we can work with, we need to develop some algebraic techniques.

Definition 6.12. Let V be a vector space. Let $F_k(V)$ be the vector space whose (uncountable) basis is

$$V^{\times k} := \underbrace{V \times \cdots \times V}_{\times k}.$$

The k^{th} *exterior power* of V is defined to be the quotient of $F_k(V)$ by the relations:

$$\begin{aligned} (v_1, \ldots, v_i + w_i, \ldots, v_k) &\sim (v_1, \ldots, v_i, \ldots, v_k) + (v_1, \ldots, v_i, \ldots, v_k) \\ (v_1, \ldots, \lambda v_i, \ldots, v_k) &\sim \lambda(v_1, \ldots, v_i, \ldots, v_k) \\ (v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) &\sim -(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) \end{aligned}$$

We denote the k^{th} exterior power of V by

$$\bigwedge^k V.$$

The idea of the k^{th} exterior power of V is that, just linear maps out of the tensor product correspond to multilinear maps, linear maps out of the exterior power correspond to *alternating* k-*multilinear maps*. We first make this precise, and then provide some computational tools for dealing with exterior powers.

Lemma 6.13. *Let V and W be a vector space, and denote by $\mathrm{Alt}^k(V, W)$ the vector space of alternating k-multilinear maps from $V^{\times k}$ to W. Then there is an isomorphism of vector spaces*

$$\mathrm{Alt}^k(V, W) \cong \mathrm{Lin}(\bigwedge^k V, W).$$

Proof. The quotient map

$$q : F_k(V) \longrightarrow \bigwedge^k V$$

is linear and surjective by construction, and so composing with q defines an injective linear map

$$q^* : \mathrm{Lin}(\bigwedge^k V, W) \longrightarrow \mathrm{Lin}(F_k(V), W).$$

The image of this map is precisely the subspace of $\mathrm{Lin}(F_k(V), W)$ consisting of those linear maps $f : F_k(V) \to W$ such that f vanishes on the kernel K of q. This kernel is generated by the three relations given above.

Since $V^{\times k}$ is a basis for $F_k(V)$, linear maps out of $F_k(V)$ are equivalently maps of sets $V^{\times k} \to W$. Given a map $f : V^{\times k} \to W$, f corresponds to the unique $\mathbb{R}$-linear extension of f to $F_k(V)$. This extension of f will vanish on K precisely when it vanishes on the generators of K, i.e., when

$$\begin{aligned} f(v_1, \dots, v_i + w_i, \dots, v_k) &= f(v_1, \dots, v_i, \dots, v_k) + f(v_1, \dots, v_i, \dots, v_k) \\ f(v_1, \dots, \lambda v_i, \dots, v_k) &= \lambda f(v_1, \dots, v_i, \dots, v_k) \\ f(v_1, \dots, v_i, \dots, v_j, \dots, v_k) &= -f(v_1, \dots, v_j, \dots, v_i, \dots, v_k) \end{aligned}$$

that is, precisely when f is an alternating k-multilinear map. □

Proposition 6.14. *Let V have basis $\{v_1, \dots, v_n\}$. Denote the image of $(w_1, \dots, w_k) \in F_k(V)$ under the quotient map by*

$$w_1 \wedge w_2 \wedge \cdots \wedge w_k \in \bigwedge^k V.$$

1. *For a permutation $\sigma \in S_k$,*

$$w_1 \wedge \cdots \wedge w_k = \mathrm{sgn}(\sigma) w_{\sigma(1)} \wedge \cdots \wedge w_{\sigma(k)}.$$

2. *An element*

$$w_1 \wedge w_2 \wedge \cdots \wedge w_k \in \bigwedge^k V$$

is zero if and only if the vectors $w_1, \dots, w_k$ are linearly dependent in V.

3. *The elements*

$$v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k}$$

defined by k-tuples of distinct integers $1 \leq i_\ell \leq n$ such that $i_r \leq i_\ell$ whenever $r \leq \ell$ form a basis of $\bigwedge^k V$. As a result

$$\dim\left(\bigwedge^k V\right) = \binom{n}{k}.$$

Proof. To see (1), we simply need note that, under the equivalence relation defined, we have

$$w_1 \wedge \cdots \wedge w_i \wedge \cdots \wedge w_j \wedge \cdots \wedge w_k = -w_1 \wedge \cdots \wedge w_j \wedge \cdots \wedge w_i \wedge \cdots \wedge w_k$$

so that decomposing a permutation into flips, we obtain the desired result.

To see (2), suppose first that the w_i are linearly dependent. Then write

$$w_1 = \sum_{j=2}^{k} \lambda_i w_i.$$

We then have

$$\begin{aligned} w_1 \wedge w_2 \wedge \cdots \wedge w_k &= \left(\sum_{j=2}^{k} \lambda_i w_i \right) \wedge w_2 \wedge \cdots \wedge w_k \\ &= \sum_{j=2} \lambda_i \left(w_j \wedge w_2 \wedge \cdots \wedge w_k \right) \end{aligned}$$

however, each of the expressions in the sum is the equivalence class represented by a vector $(w_j, w_2, \ldots, w_k)$, with $2 \leq j \leq k$, and so contains a repeated entry. As such, this element is zero in the quotient. The other direction — that if the w_i are linearly independent, the $w_1 \wedge \cdots \wedge w_k \neq 0$ — follows directly from part (3).

Finally, to see (3), first note that, by (1) and the fact that $\{v_1, \ldots, v_k\}$ is a basis, the given set spans $\bigwedge^k V$. To prove that it is a basis, we need two steps. When $k = n$, the space of alternating multilinear maps $V^{\times n} \to \mathbb{R}$ is 1-dimensional (by the existence and uniqueness of the determinant). Thus, $\bigwedge^n V$ is 1-dimensional by the preceding lemma, and so $v_1 \wedge \cdots \wedge v_n$ is a basis.

Suppose that there is a relation

$$\sum_{I \subset \{1, \ldots, n\}} \lambda_I v_{i_1} \wedge \cdots \wedge v_{i_k} = 0.$$

Then the element

$$W = \sum_{I \subset \{1, \ldots, n\}} \lambda_I (v_{i_1}, \ldots, v_{i_k})$$

in $F_k(V)$ can be built out of our three relations. For a fixed subset $J \subset \{1, \ldots, n\}$, let $J^c = \{1, \ldots, n\} \setminus J$, and denote the elements of J^c by $j_1, \ldots, j_{n-k}$. Form an element of $F_n(V)$

$$U = \sum_{I \subset \{1, \ldots, n\}} \lambda_I (v_{i_1}, \ldots, v_{i_k}, v_{j_1}, \ldots, v_{j_{n-k}})$$

By applying the same operations which showed W to be equivalent to zero (fixing the last $n - k$-indices) we see that $U = 0$. However, every term of the sum other than that corresponding to J contains a repeated index, and so vanishes identically in the quotient. Thus, we find that

$$\pm \lambda_J (v_1, \ldots, v_n) \sim 0$$

But $v_1 \wedge \cdots \wedge v_n$ is a basis of $\bigwedge V$, and thus $\lambda_J = 0$. Since J was arbitrary, this completes the proof. □

Proposition 6.15. *Let V be a finite-dimensional vector space. There is a canonical isomorphism*

$$\bigwedge^k \operatorname{Lin}(V, \mathbb{R}) \cong \operatorname{Alt}^k(V, \mathbb{R})$$

Proof. We first define a map

$$\Psi : \bigwedge^k \mathrm{Lin}(V, \mathbb{R}) \longrightarrow \mathrm{Alt}^k(V, \mathbb{R})$$

by defining it on expressions of the form $f_1 \wedge \cdots \wedge f_k$. We define

$$\Psi(f_1 \wedge \cdots \wedge f_k)(v_1, \ldots, v_k) := \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) f_1(v_{\sigma(1)}) f_2(v_{\sigma(2)}) \cdots f_k(v_{\sigma(k)}).$$

It is a straightforward check to see that the linearity of the f_i and the inclusion of $\mathrm{sgn}(\sigma)$ in the sum together imply that $\Psi(f_1 \wedge \cdots \wedge f_k)$ is an alternating linear map. Similar arguments imply that $\Psi(f_1 \wedge \cdots \wedge f_k)$ is well-defined.

We now claim that this map is an isomorphism. Fix a basis $\{v_1, \ldots, v_n\}$ of V, and let $\{\omega_1, \ldots, \omega_n\}$ be the dual basis for $\mathrm{Lin}(V, \mathbb{R})$. We compute the image of

$$u = \omega_{i_1} \wedge \cdots \wedge \omega_{i_k}$$

under Ψ. Applying this to a basis element of $\bigwedge^k V$ of the form $v_{j_1} \wedge \cdots \wedge v_{j_k}$, we see that

$$\Psi(u)(v_{j_1} \wedge \cdots \wedge v_{j_k}) = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \omega_{i_1}(v_{j_{\sigma(1)}}) \cdots \omega_{i_k}(v_{j_{\sigma(k)}})$$

However, since the i_ℓ's must be distict, as must the j_ℓ's, we see that this is non-zero if and only if there is some permutation σ such that $i_\ell = j_{\sigma(\ell)}$ for all $1 \leq \ell \leq k$. Since we assume that the i's and j's are each sequences of indices in order, this implies

$$\Psi(u)(v_{j_1}, \ldots, v_{j_k}) = \begin{cases} 1 & i_\ell = j_\ell \forall \ell \\ 0 & \text{else.} \end{cases}$$

However, this means that under the identification

$$\mathrm{Alt}^k(V, \mathbb{R}) \cong \mathrm{Lin}(\bigwedge^k V, \mathbb{R}),$$

the alternating map $\Psi(u)$ is identified with an element of the dual basis of $\mathrm{Lin}(\bigwedge^k V, \mathbb{R})$. Thus, Ψ sends a basis to a basis, and so is an isomorphism. ☐

Note 6.16. Let us briefly think about what the linear map associated to an element of $\bigwedge^k \mathrm{Lin}(V, \mathbb{R})$ represents. We will first consider an element of the form

$$\omega = \omega_1 \wedge \cdots \wedge \omega_k$$

where each ω_i is a linear map $V \to \mathbb{R}$. Assuming that $\omega \neq 0$, we see that the ω_i are linearly independent in $\mathrm{Lin}(V, \mathbb{R})$. Write $K = \bigcup_{i=1}^k \ker(\omega_i)$ for the intersection of the kernels of the ω_i. Since the ω_i's are linearly independent, the dimension of K is precisely $n - k$.

The quotient space V/K thus has dimension k, and each ω_i is uniquely determined the map

$$\widetilde{\omega}_i : V/K \longrightarrow \mathbb{R}$$

to which it descends. We can further see that ω is uniquely determined by the induced element

$$\widetilde{\omega} := \widetilde{\omega}_1 \wedge \cdots \wedge \widetilde{\omega}_k \in \bigwedge^k \mathrm{Lin}(V/K, \mathbb{R}).$$

Since ω is non-zero, so is $\widetilde{\omega}$. Thus, ω is uniquely determined by the subspace K and a nonzero, alternating k-multilinear map $(V/K)^{\times k} \to \mathbb{R}$. By a standard result, all such maps are scalar multiples of a determinant function, and thus are ways of measuring signed k-dimensional volumes in V/K.

We can actually say more: in the presence of a chosen inner product on V, we can uniquely define the orthogonal complement $K^\perp$ of K, which is then canonically identified with V/K. Thus, in the presence of an inner product, the form ω is uniquely determined by the k-dimensional subspace $K^\perp$ together with a way of measuring signed k-dimensional volumes in $K^\perp$.

More general elements of $\bigwedge^k \mathrm{Lin}(V, \mathbb{R})$ will be formal linear combinations of such gadgets. Fixing an inner product and an orthonormal basis $\{v_1, \ldots, v_n\}$ of V, and letting ω_i be the elements of the dual basis, every element of $\bigwedge^k \mathrm{Lin}(V, \mathbb{R})$ will be of the form

$$\omega = \sum_{i_1 < \cdots < i_k} \lambda_I \omega_{i_1} \wedge \cdots \wedge \omega_{i_k}.$$

We can interpret this as follows: the term $\lambda_I \omega_{i_1} \wedge \cdots \wedge \omega_{i_k}$ chooses the subspace spanned by the basis elements $v_{i_1}, \ldots, v_{i_k}$, and assigns the cube defined by these basis elements the volume λ_I. Applying ω to a k-tuple of vectors $u_1, \ldots, u_k$ can be computed as follows:

1. Project the u_i onto each k-dimensional subspace spanned by basis vector.

2. Compute the signed volume according to Λ of the resulting parallelpiped in the chosen subspace.

3. Sum together all of the results.

If we put together everything we have shown about exterior powers, we can give a nice description of k-forms at a point. A k-form ω at $p \in \mathbb{R}^n$ is an alternating multilinear map

$$\omega : (T_p\mathbb{R}^n)^{\times k} \longrightarrow \mathbb{R}$$

so, equivalently, it is a linear map

$$\omega : \bigwedge^k T_p\mathbb{R}^n \longrightarrow \mathbb{R}$$

and, finally, this means that it is an element

$$\omega \in \bigwedge^k \mathrm{Lin}(T_p\mathbb{R}^n, \mathbb{R}) = \bigwedge^k T_p^*\mathbb{R}^n.$$

We have the coordinate basis $\{\mathrm{d}x^1, \ldots, \mathrm{d}x^n\}$ for $T_p^*\mathbb{R}^n$, and thus a basis for $\bigwedge^k T_p^*\mathbb{R}^n$ is given by the wedge products

$$\mathrm{d}x^{i_1} \wedge \cdots \mathrm{d}x^{i_k}$$

defined by k-tuples of distinct, ordered, integers between 1 and n. We can thus write any k-form at $p \in \mathbb{R}^n$ as

$$\omega = \sum_{i_1 < i_2 < \cdots < i_k} \omega_{i_1, i_2, \ldots, i_k} \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k},$$

where

$$\omega_{i_1, i_2, \ldots, i_k} = \omega\left(\partial_{x^{i_1}}, \ldots, \partial_{x^{i_k}}\right).$$

Definition 6.17. The k^{th} *exterior power of the cotangent bundle* is the set

$$\bigwedge^k T^*\mathbb{R}^n := \{(p, \omega) \mid \omega \in \bigwedge^k T_p^*\mathbb{R}^n\} \cong \mathbb{R}^n \times \mathbb{R}^{\binom{n}{k}}$$

which we view as a smooth manifold via the latter identification. A *smooth k-form* on $\mathbb{R}^n$ is a smooth map

$$\omega : \mathbb{R}^n \longrightarrow \bigwedge^k T^*\mathbb{R}^n$$

such that $\omega(p) \in \bigwedge^k T_p^*\mathbb{R}^n$ for each $p \in \mathbb{R}^n$.

Now that we have a definition of a k-form, let's see how we integrate a k-form over a k-dimensional submanifold.

Definition 6.18. Suppose first that $C \subset \mathbb{R}^k$ is a compact subset, and ν is a smooth k-form on $\mathbb{R}^k$. Then there is a unique smooth coefficient function

$$\nu = f(x^1, \ldots, x^k)\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^k$$

We define the integral of ν over C to be

$$\int_C \nu := \int_C f(x^1, \ldots, x^k)\, dx^1 dx^2 \cdots dx^k.$$

To define the integral of a k-form over a submanifold of $\mathbb{R}^n$, we need to extend our definition of the pullback.

Definition 6.19. Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function, and let ω be a k-form at $\phi(p) \in \mathbb{R}^m$. The *pullback of ω along ϕ* is the k-form $\phi_p^*\omega$ at $p \in \mathbb{R}^n$ defined by

$$(\phi_p^*\omega)(v_1, \ldots, v_k) = \omega(d\phi_p(v_1), \ldots, d\phi_p(v_k))$$

for any tangent vectors $v_1, \ldots, v_k \in T_p\mathbb{R}^n$. This defines a linear map

$$\phi_p^* : \bigwedge^k T_p^*\mathbb{R}^m \longrightarrow \bigwedge^k T_p^*\mathbb{R}^n$$

for each $p \in \mathbb{R}^n$. We define the pullback of a smooth k-form ω on $\mathbb{R}^m$ by

$$(\phi^*\omega)(p) = \phi_p^*(\omega(\phi(p))).$$

The formal properties of the pullback are very similar to those for the pullback of 1-forms.

Exercise 19. Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ and $\psi : \mathbb{R}^m \to \mathbb{R}^\ell$ be smooth maps. Let ω be a smooth k-form on $\mathbb{R}^\ell$. Show that
$$(\psi \circ \phi)^* \omega = \phi^*(\psi^* \omega).$$

We then make the definition

Definition 6.20. Let $C \subset \mathbb{R}^k$ be a compact set, and $\phi : C \to \mathbb{R}^n$ a smooth, regular, injective map. Let ω be a k-form on $\mathbb{R}^n$. The *integral of* ω *over* $\phi(C)$ is defined to be
$$\int_{\phi(C)} \omega = \int_C \phi^* \omega.$$

3 *The algebra of forms*

To properly understand the properties of the integral, we will need more understanding of the algebraic structure of forms. The first step is to relate the differentials of smooth maps to differential forms. The second step is to understand the "product" operation on forms. To this end, we first introduce some notation.

Notation 6.21. We denote the vector space of smooth k-forms on $\mathbb{R}^n$ by
$$\Omega^k(\mathbb{R}^n).$$

We will sometimes write $\Omega^0(\mathbb{R}^n)$ to mean the vector space of smooth functions $\mathbb{R}^n \to \mathbb{R}$, for reasons which will become clear in time.

With this definition in hand, our first order of business is to resolve a notational contradiction. Given a smooth map $f : \mathbb{R}^n \to \mathbb{R}^m$, we have been using the notation $df : T\mathbb{R}^n \to T\mathbb{R}^m$ to denote the differential of f. However, this looks very similar to the notation $\mathrm{d}x^i$ we use for the coordinate 1-forms. As it turns out, there is a good reason for this collision.

Construction 6.22. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth map, and consider the differential
$$df : T\mathbb{R}^n \longrightarrow T\mathbb{R}.$$

We can view this as assigning to each point $p \in \mathbb{R}^n$ a map
$$df_p : T_p\mathbb{R}^n \longrightarrow T_{f(p)}\mathbb{R} \cong \mathbb{R}$$

that is, a covector at p. In particular, this covector acts on the standard basis vectors ∂_{x^i} by
$$df_p(\partial_{x^i}) = \frac{\partial f}{\partial x_i}|_p.$$

We can thus view df as a smooth 1-form $\mathrm{d}f$, which has coordinate representation
$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\mathrm{d}x^i.$$

Before continuing, let us try to understand the meaning of this 1-form. If one thinks of $\mathrm{d}x^i$ as a way of measuring an infinitesmal change in the x^i-direction, then $\mathrm{d}f$ is a way of measuring the *infinitesmal change in the value of f* cause by an infinitesmal movement in a given tangent direction. Heuristically,

$$\Delta f \approx \frac{\partial f}{\partial x^i} \Delta x^i.$$

Notice that f is constant on $\mathbb{R}^n$ if and only if $\mathrm{d}f$ is a identically zero.

Exercise 20. Let $\gamma : [a, b] \to \mathbb{R}^n$ be a smooth curve, and $f : \mathbb{R}^n \to \mathbb{R}$ a smooth function. Show that

$$\int_\gamma \mathrm{d}f = f(\gamma(b)) - f(\gamma(a)).$$

Exercise 21. Show that the differential defines an $\mathbb{R}$-linear map

$$\mathrm{d} : \Omega^0(\mathbb{R}^n) \longrightarrow \Omega^1(\mathbb{R}^n)$$

Exercise 22. Let x^i denote the coordinates on $\mathbb{R}^n$, and denote by $\mathbf{x}^i : \mathbb{R}^n \to \mathbb{R}$ the coordinate function, i.e., the smooth function which sends a point to its i^{th} coordinate. Show that the differential of $\mathbf{x}^i$ is $\mathrm{d}x^i$.

The second thing we need to note about forms is that there is a canonical product, the *wedge product* of forms, which combines lower-dimensional forms into higher dimensional forms.

Definition 6.23. Given basis k- and ℓ-forms

$$\omega = \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \quad \text{and} \quad \nu = \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_\ell}$$

on $\mathbb{R}^n$, we define their *wedge product* to be

$$\omega \wedge \eta := \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_\ell}.$$

We can extend this uniquely by bilinearity to give a bilinear map

$$\wedge : \Omega^k(\mathbb{R}^n) \times \Omega^\ell(\mathbb{R}^n) \longrightarrow \Omega^{k+\ell}(\mathbb{R}^n)$$

Notation 6.24. For $I = \{i_1, \ldots, i_k\}$ we write

$$\mathrm{d}x^I := \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

We will employ the Einstein summation convention for such *multi-indices* as with single indices. That is

$$\omega_I \mathrm{d}x^I := \sum_I \omega_I \mathrm{d}x^I$$

where I ranges through all ordered multi-indices.

Exercise 23. Let

$$\omega = \omega_I \mathrm{d}x^I \in \Omega^k(\mathbb{R}^n) \quad \text{and} \quad \nu = \nu_J \mathrm{d}x^J \in \Omega^r(\mathbb{R}^n)$$

Show that

$$\omega \wedge \nu = \frac{1}{k!r!} \sum_{L} \left(\sum_{\sigma \in S_L} \operatorname{sgn}(\sigma) \omega_{\ell_{\sigma(1)},\dots,\ell_{\sigma(k)}} \nu_{\ell_{\sigma(k+1)},\dots,\ell_{\sigma(k+r)}} \right) \mathrm{d}x^L$$

Conclude that

$$\omega \wedge \nu = (-1)^{k+r} \nu \wedge \omega.$$

Further convince yourself that the wedge product is associative, i.e.,

$$\omega \wedge (\nu \wedge \eta) = (\omega \wedge \nu) \wedge \eta.$$

The wedge product connects *all* of the spaces $\Omega^k(\mathbb{R}^n)$ together in a single algebraic object. We can, in fact, connect these spaces in one other way: via derivatives.

Definition 6.25. The *exterior derivative* is the $\mathbb{R}$-linear map

$$\mathrm{d} : \Omega^k(\mathbb{R}^n) \longrightarrow \Omega^{k+1}(\mathbb{R}^n) \qquad \omega_I \mathrm{d}x^I \longmapsto \tfrac{\partial \omega_I}{\partial x^j} \mathrm{d}x^j \wedge \mathrm{d}x^I.$$

notice that when $k = 0$, this is just the differential.

Exercise 24. Show that the exterior derivative has the following properties:

1. Linearity

$$\mathrm{d}(\omega + \nu) = \mathrm{d}\omega + \mathrm{d}\nu$$

   and

$$\mathrm{d}(\lambda\omega) = \lambda \mathrm{d}\omega$$

   for $\lambda \in \mathbb{R}$.

2. The *graded Leibnitz rule*

$$\mathrm{d}(\omega \wedge \nu) = (\mathrm{d}\omega) \wedge \nu + (-1)^{\deg(\omega)} \omega \wedge \mathrm{d}\nu.$$

   Where $\deg(\omega) = k$, for ω a k-form.

3. Chain complex:

$$\mathrm{d} \circ \mathrm{d} = 0$$

Example 6.26. We consider a collection of examples in $\mathbb{R}^3$, to see that the differential encodes a number of familiar examples. Firstly, if we consider a 0-form, i.e. a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$, we see that

$$\mathrm{d}f = \frac{\partial f}{\partial x^1} \mathrm{d}x^1 + \frac{\partial f}{\partial x^2} \mathrm{d}x^2 + \frac{\partial f}{\partial x^3} \mathrm{d}x^3$$

If we (somewhat improperly) view the resulting 1-form as a vector field on $\mathbb{R}^3$, identifying 1-forms and vector fields via the metric, we see that this is precisely the gradient of f.

Moving on, lets consider a 1-form

$$\omega = \omega_1 \mathrm{d}x^1 + \omega_2 \mathrm{d}x^2 + \omega_3 \mathrm{d}x^3$$

which, again, we interpret as a vector field via the metric. Then the exterior derivative is (neglecting terms which become zero)

$$\begin{aligned}
\mathrm{d}\omega &= \frac{\partial \omega_1}{\partial x^2}\mathrm{d}x^2 \wedge \mathrm{d}x^1 + \frac{\partial \omega_1}{\partial x^3}\mathrm{d}x^3 \wedge \mathrm{d}x^1 + \frac{\partial \omega_2}{\partial x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \frac{\partial \omega_2}{\partial x^3}\mathrm{d}x^3 \wedge \mathrm{d}x^2 \\
&\quad + \frac{\partial \omega_3}{\partial x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^3 + \frac{\partial \omega_3}{\partial x^2}\mathrm{d}x^2 \wedge \mathrm{d}x^3 \\
&= \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2}\right)\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \left(\frac{\partial \omega_3}{\partial x^1} - \frac{\partial \omega_1}{\partial x^3}\right)\mathrm{d}x^1 \wedge \mathrm{d}x^3 + \left(\frac{\partial \omega_2}{\partial x^3} - \frac{\partial \omega_3}{\partial x^2}\right)\mathrm{d}x^2 \wedge \mathrm{d}x^3
\end{aligned}$$

As it turns out, we can also identify 2-forms with vector fields on $\mathbb{R}^3$. The idea is that we identify a 2-form ν with the (unique) vector $w \in \mathbb{R}^3$ such that the triple product

$$\langle (x \times y), w \rangle = \nu(x, y).$$

Under this identification, $\mathrm{d}x^1 \wedge \mathrm{d}x^2$ corresponds to ∂_{x^3}, $\mathrm{d}x^2 \wedge \mathrm{d}x^3$ corresponds to ∂_{x^1}, and $\mathrm{d}x^1 \wedge \mathrm{d}x^3$ corresponds to $-\partial_{x^2}$. Making this identification, we see that the vector field which results is

$$\begin{pmatrix} \frac{\partial \omega_3}{\partial x^1} - \frac{\partial \omega_1}{\partial x^3} \\ \frac{\partial \omega_1}{\partial x^3} - \frac{\partial \omega_3}{\partial x^1} \\ \frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2} \end{pmatrix}$$

i.e., the *curl* of the vector field $(\omega_1, \omega_2, \omega_3)$.

Finally, consider a 2-form (again considered as a vector field)

$$\omega = \omega_1 \mathrm{d}x^2 \wedge \mathrm{d}x^3 + \omega_2 \mathrm{d}x^3 \wedge \mathrm{d}x^1 + \omega_3 \mathrm{d}x^1 \wedge \mathrm{d}x^2$$

We compute

$$\begin{aligned}
\mathrm{d}\omega &= \frac{\partial \omega_1}{\partial x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 + \frac{\partial \omega_2}{\partial x^2}\mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}x^1 + \frac{\partial \omega_3}{\partial x^3}\mathrm{d}x^3 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \\
&= \left(\frac{\partial \omega_1}{\partial x^1} + \frac{\partial \omega_2}{\partial x^2} + \frac{\partial \omega_3}{\partial x^3}\right)\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3
\end{aligned}$$

i.e., the *divergence* of f.

Thus, in $\mathbb{R}^3$, the three key differential operators we study can be viewed as special cases of the exterior derivative.

Before returning to integration, we prove a lemma which is one key way to interpret $\mathrm{d}\omega$.[6]

[6] This explanation is adapted, with suitable modification, from Vladimir Arnold's wonderful "Mathematical Methods of Classical mechanics.

Lemma 6.27. *Let ω be a k-form on $\mathbb{R}^n$, and let $v_1, \ldots, v_{k+1}$ be constant[7] vector fields on $\mathbb{R}^n$. Define*

[7] This is only a meaningful thing to say when we identify $\mathbb{R}^n$ with its tangent space. It doesn't have a meaning on a general manifold.

$$F_\omega(v_1, \ldots, v_{k+1}) = \sum_{i=1}^{k+1} \sum_{j=0}^{1} (-1)^{j+i} \omega_{p+jv_i}(v_1, \ldots, \widehat{v_i}, \ldots, v_{k+1})$$

Then

$$\mathrm{d}\omega_p(v_1, \ldots, v_{k+1}) = \lim_{t \to 0} \frac{F_\omega(tv_1, \ldots, tv_{k+1})}{t^{k+1}}.$$

Before we get to the proof, a bit of explanation is in order. Consider the parallelpiped P based at $p \in \mathbb{R}^n$ spanned by $v_1, \ldots, v_{k+1}$, whose vertices are

$$V = \left\{ p + \sum_{i=1}^{k+2} \epsilon_i v_i \middle| \epsilon_i = 0, 1 \right\}.$$

The somewhat complicated expression for F_ω is a way of approximating the (oriented) integral of ω over the boundary of P. In the sum, the term corresponding to an index i with $j = 0$ is an approximation of the integral of ω over the face generated by $v_1, \ldots, \widehat{v}_i, \ldots, v_{k+1}$ based at p (assuming that ω is almost constant along this face). The term with index i and $j = 0$ is an approximation of the integral of ω over the face generated by $v_1, \ldots, \widehat{v}_i, \ldots, v_{k+1}$ based at $p + jv$.

To understand the final statement of the theorem, we make the following observation. Suppose that $F : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ is a map. We want a *multilinear* map $T : \mathbb{R}^\times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is the *best* approximation to F at 0 by a multilinear map. Mimicking the definition of the derivative, we then require that

$$\lim_{t \to 0} \frac{F(tv_1, \ldots, tv_{k+1}) - T(tv_1, \ldots, tv_{k+1})}{t^{k+1}} = 0.$$

The denominator is chosen to be t^{k+1} because any $k+1$-linear part must scale in its arguments in this way. We then see that

$$T(v_1, \ldots, v_{k+1}) = \lim_{t \to 0} \frac{F(tv_1, \ldots, tv_{k+1})}{t^{k+1}}$$

so long as the latter limit exists. We thus see that the lemma claims that the best multilinear approximation to the “surface area” function defined from ω is $d\omega$.

Proof. We first show that the limit so described is alternating and multilinear. For multilinearity, the symmetries of F_ω mean that it suffices to show that the limit is multilinear in v_1. We see that each summand of F_ω is linear in v_1 *except* for the two coresponding to $i = 1$. It thus will suffice to show that the limit of the sum of these two terms is linear in v_1.

To do this, we compute

$$\lim_{t \to 0} \frac{\omega_{p+tv_1}(tv_2, \ldots, tv_{k+1}) - \omega_p(tv_2, \ldots, tv_{k+1})}{t^{k+1}}$$

Pulling the factors of t out of ω this simplifies to

$$\lim_{t \to 0} \frac{\omega_{p+tv_1}(v_2, \ldots, v_{k+1}) - \omega_p(v_2, \ldots, v_{k+1})}{t}$$

However, this is simply the partial derivative in the v_1-direction of the function $p \mapsto \omega_p(v_2, \ldots, v_{k+1})$. Since the construction $v_1 \mapsto \partial_{v_1} f$ is linear in v_1, this completes the proof of multilinearity.

To see that the function is alternating, note first that every term of F_ω except for those corresponding to $i = 1, 2$ changes sign when v_1 and v_2 are swapped. It thus suffices to consider the remaining four terms:

$$\lim_{t \to 0} \frac{\omega_{p+tv_1}(tv_2, \ldots, tv_{k+1}) - \omega_p(tv_2, \ldots, tv_{k+1}) - \omega_{p+tv_2}(tv_1, \ldots, tv_{k+1}) + \omega_p(tv_1, \ldots, tv_{k+1})}{t^{k+1}}$$

which amounts to

$$\partial_{v_1}\omega_p(v_2,\ldots,v_{k+1}) - \partial_{v_2}\omega_p(v_1,v_3,\ldots,v_{k+1})$$

an expression which is clearly alternating in v_1 and v_2.

Now that we have established that the limit is multilinear and alternating, we need only test on basis vectors. Evaluating F_ω at $t\partial_{\ell_1},\ldots,t\partial_{\ell_{k+1}}$, (where the indices appear in order) we find

$$\sum_{i=1}^{k+1}\sum_{j=0}^{1}(-1)^i\omega_{\ell_1,\ldots,\hat{\ell}_i,\ldots,\ell_{k+1}}(p+t\partial_\ell i) - \omega_{\ell_1,\ldots,\hat{\ell}_i,\ldots,\ell_{k+1}}(p)$$

so that the limit becomes

$$\sum_{i=1}^{k+1}(-1)^i\frac{\partial\omega_{\ell_1,\ldots,\hat{\ell}_i,\ldots,\ell_{k+1}}}{\partial x^{\ell_i}}(p)$$

On the other hand, we can compute

$$(\mathrm{d}\omega)_p(\partial_{\ell_1},\ldots,\partial_{\ell_{k+1}}) = \frac{\partial\omega_I}{\partial x^i}(p)\mathrm{d}x^i\wedge\mathrm{d}x^I(\partial_{\ell_1},\ldots,\partial_{\ell_{k+1}})$$

However, the term

$$\mathrm{d}x^i\wedge\mathrm{d}x^I(\partial_{\ell_1},\ldots,\partial_{\ell_{k+1}})$$

is zero if $I\cup\{i\}\neq\{\ell_1,\ldots,\ell_{k+1}\}$, and is $(-1)^j$ when $I=\{\ell_1,\ldots,\ell_{k+1}\}\setminus\{\ell_j\}$ when $i=\ell_j$. Thus, this expression becomes

$$(\mathrm{d}\omega)_p(\partial_{\ell_1},\ldots,\partial_{\ell_{k+1}}) = \sum_{i=1}^{k+1}(-1)^i\frac{\partial\omega_{\ell_1,\ldots,\hat{\ell}_i,\ldots,\ell_{k+1}}}{\partial x^{\ell_i}}(p)$$

completing the proof. □

3.1 Relating the structures

We now have three key operations we can perform on k-forms. We can pull them back along smooth maps, we can take wedge products, and we can take exterior derivatives. These three structures interact in a number of interesting and complex ways, some of which we now explore.

Lemma 6.28. *Let $\phi:\mathbb{R}^n\to\mathbb{R}^m$ be a smooth function, and write x^i and y^i for the coordinates on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Then*

$$\phi^*(\mathrm{d}y^{i_1}\wedge\cdots\wedge\mathrm{d}y^{i_k}) = \left(\phi^*(\mathrm{d}y^{i_1})\wedge\cdots\wedge\phi^*(\mathrm{d}y^{i_k})\right).$$

Proof. We test both sides of the equation on a tuple $\partial_{x^{j_1}},\ldots,\partial_{x^{j_k}}$ of basis vectors whose indices are in order. For the left-hand side,

$$\begin{aligned}\phi^*(\mathrm{d}y^{i_1}\wedge\cdots\wedge\mathrm{d}y^{i_k})(\partial_{x^{j_1}},\ldots,\partial_{x^{j_k}}) &= \mathrm{d}y^{i_1}\wedge\cdots\wedge\mathrm{d}y^{i_k}(d\phi(\partial_{x^{j_1}}),\ldots,d\phi(\partial_{x^{j_k}}))\\ &= \sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)\prod_{z=1}^{k}\mathrm{d}y^{i_z}(d\phi(\partial_{y^{\sigma(j_z)}}))\\ &= \sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)\prod_{z=1}^{k}\frac{\partial\phi^{i_z}}{\partial x^{\sigma(j_z)}}\end{aligned}$$

On the other hand, we have the right-hand side:

$$\begin{aligned}
\left(\phi^*(\mathrm{d}y^{i_1}) \wedge \cdots \phi^*(\mathrm{d}y^{i_k})\right) &= \left(\frac{\partial \phi^{i_1}}{\partial x^{\ell_1}} \mathrm{d}x^{\ell_1}\right) \wedge \cdots \wedge \left(\frac{\partial \phi^{i_k}}{\partial x^{\ell_k}} \mathrm{d}x^{\ell_k}\right) \\
&= \sum_{\substack{\text{ord. subsets} \\ \vec{L} \subset \{1,\ldots,n\}}} \frac{\partial \phi^{i_1}}{\partial x^{\ell_1}} \cdots \frac{\partial \phi^{i_k}}{\partial x^{\ell_k}} \mathrm{d}x^{\ell_1} \wedge \cdots \mathrm{x}x^{\ell_k} \\
&= \sum_{\substack{\text{subsets} \\ L \subset \{1,\ldots,n\}}} \left(\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \frac{\partial \phi^{i_1}}{\partial x^{\ell_{\sigma(1)}}} \cdots \frac{\partial \phi^{i_k}}{\partial x^{\ell_{\sigma(k)}}}\right) \mathrm{d}x^{\ell_1} \wedge \cdots \mathrm{d}x^{\ell_k}.
\end{aligned}$$

Applying this to $\partial_{x^{j_1}}, \ldots, \partial_{x^{j_k}}$, we retain only the term corresponding to $L = J$, that is

$$\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \prod_{z=1}^{k} \frac{\partial \phi^{i_z}}{\partial x^{\sigma(j_z)}}$$

so that the two are equal as desired. □

Lemma 6.29. *Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function.*

1. *For smooth k-forms ν and ω on $\mathbb{R}^m$,*

$$\phi^*(\omega + \nu) = \phi^*(\omega) + \phi^*(\nu).$$

2. *For $\lambda \in \mathbb{R}$ and $\omega \in \Omega^k(\mathbb{R}^m)$,*

$$\phi^*(\lambda\omega) = \lambda\phi^*(\omega).$$

3. *For $\omega \in \Omega^k(\mathbb{R}^m)$ and $\nu \in \Omega^\ell(\mathbb{R}^m)$,*

$$\phi^*(\omega \wedge \nu) = \phi^*(\omega) \wedge \phi^*(\nu).$$

Proof. The first two statements are immediate. To see (3), notice that by the first two statements, it suffices to prove the statement on wedges of 1-forms. Indeed, it is sufficient to show it on the basis $\mathrm{d}x^I$. Letting $I = \{i_1 < i_2 < \cdots < i_k\}$ and $J = \{j_1 < \cdots < j_\ell\}$, we have that

$$\begin{aligned}
\phi^*(\mathrm{d}x^I \wedge \mathrm{d}x^J) &= \phi^*(\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_\ell}) \\
&= \left(\phi^* \mathrm{d}x^{i_1}\right) \wedge \cdots \wedge \left(\phi^* \mathrm{d}x^{i_k}\right) \wedge \left(\phi^* \mathrm{d}x^{j_1}\right) \wedge \cdots \wedge \left(\phi^* \mathrm{d}x^{j_\ell}\right) \\
&= \left(\left(\phi^* \mathrm{d}x^{i_1}\right) \wedge \cdots \wedge \left(\phi^* \mathrm{d}x^{i_k}\right)\right) \left(\wedge \left(\phi^* \mathrm{d}x^{j_1}\right) \wedge \cdots \wedge \left(\phi^* \mathrm{d}x^{j_\ell}\right)\right) \\
&= \phi^*(\mathrm{d}x^I) \wedge \phi^*(\mathrm{d}x^J)
\end{aligned}$$

by applying the associativity of $\wedge$ and the previous lemma. □

Lemma 6.30. *Let $\phi : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function, and ω a smooth k-form on $\mathbb{R}^m$. Write x^i and y^i for the coordinates on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, so that*

$$\omega = \omega_I \mathrm{d}y^i.$$

Then

$$\phi^*\omega = \omega_I \mathrm{d}(\phi^{i_1}) \wedge \cdots \wedge \mathrm{d}(\phi^{i_k}) = \omega_I \mathrm{d}(\mathbf{y}^{i_1} \circ \phi) \wedge \cdots \wedge \mathrm{d}(\mathbf{y}^{i_1} \circ \phi)$$

Proof. It suffices to check this on the elements $\mathrm{d}x^i$, and then apply the previous lemma. This is a straightforward computation. □

Exercise 25. Prove that, for $\phi : \mathbb{R}^n \to \mathbb{R}^m$ a smooth map and $\omega \in \Omega^k(\mathbb{R}^m)$,

$$\phi^*(\mathrm{d}\omega) = \mathrm{d}(\phi^*(\omega)).$$

Exercise 26. Show that for smooth maps $\psi : \mathbb{R}^n \to \mathbb{R}^m$ and $\phi : \mathbb{R}^m \to \mathbb{R}^\ell$, and a smooth k-form $\omega \in \Omega^k(\mathbb{R}^\ell)$,

$$(\phi \circ \psi)^*(\omega) = \psi^*(\phi^*(\omega)).$$

4 *Integration and boundary*

We now return to studying integration. Our first task is to show that the integral of a k-form over an oriented k-dimensional submanifold $M \subset \mathbb{R}^n$ does not depend on the choice of chart. This follows immediately from the following proposition.

Proposition 6.31. *Let $C \subset \mathbb{R}^k$ be a subset, and let $\phi : \mathbb{R}^k \to \mathbb{R}^k$ be an orientation-preserving diffeomorphism. Let $\omega \in \Omega^k(\mathbb{R}^k)$ a smooth k-form, viewed as a form on the target of ϕ. Then*

$$\int_{\phi(C)} \omega = \int_C \phi^*\omega.$$

Proof. We will compute an explicit coordinate representation for the latter integral. Let x^i be the coordinates on the source of ϕ, and y^i the coordinates on the target. Set

$$\omega = f(x)\mathrm{d}y^1 \wedge \cdots \mathrm{d}y^k$$

Then

$$\begin{aligned} \phi^*\omega &= f(\phi(y)) \left(\sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \prod_{z=1}^{k} \frac{\partial \phi^1}{\partial x^{\sigma(1)}} \right) \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k \\ &= f(\phi(x)) \det(J\phi) \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^k \end{aligned}$$

Thus we see that

$$\begin{aligned} \int_C \phi^*\omega &= \int_C f(\phi(x)) \det(J\phi) dx^1 \cdots dx^k \\ &= \int_{\phi(C)} f(y) dy^1 \cdots dy^k \\ &= \int_{\phi(C)} \omega. \end{aligned}$$

□

Exercise 27. For $M \subset \mathbb{R}^n$ an oriented k-submanifold, and $\omega \in \Omega^k(\mathbb{R}^n)$, give a definition

$$\int_M \omega$$

and show that it is independent of any choice of oriented coordinate charts.

Having now established the well-definedness of the integral of a k-form over a k-submanifold, we turn our attention to developing the notion of *manifolds with boundary.*

Definition 6.32. The *upper half-space* $\mathbb{H}^k \subset \mathbb{R}^k$ is the space

$$\mathbb{H}^k := \{(x^1, \dots, x^k) \in \mathbb{R}^k \mid x^1 \geq 0\}.$$

A subset $U \subset \mathbb{H}^k$ is called *open* if it is the intersection of an open subset of $\mathbb{R}^k$ with $\mathbb{H}^k$. The *boundary* of $\mathbb{H}^k$ is the copy of $\mathbb{R}^{k-1} \subset \mathbb{H}^k$ defined by x^0.

A k-dimensional *chart* is a smooth, regular, injective map $\phi : U \to \mathbb{R}^k$ where $U \subset \mathbb{H}^k$ is open. A subset $M \subset \mathbb{R}^n$ is called a *k-dimensional manifold with boundary* if, for every $p \in M$, there is a chart $\phi : U \to M \subset \mathbb{R}^n$ with $p \in \phi(U)$. A point in a manifold with boundary is called a *boundary point* if it is in the image of the boundary of $\mathbb{H}^k$ under some (and thus every) chart of M. We call the set of all boundary points of M the *boundary of M*, and denote it by ∂M.

Example 6.33. Consider the closed unit ball $B_1(0) \subset \mathbb{R}^3$. There is clearly an open chart (the identity) which contains all points in the inside. We will construct a chart (with boundary) which contains points in the bounding sphere S^2, by symmetry, this will show that every point of S^2 can be contained in such a chart, showing that $B_1(0)$ is a manifold with boundary. Consider the polar coordinate chart on the sphere

$$\phi(u^1, u^2) = \big(\cos(u^2)\cos(u^1), \cos(u^2)\sin(u^1), \sin(u^2)\big)$$

defined on $(0, 2\pi) \times (0, \pi) \subset \mathbb{H}^2$. Define a chart

$$\psi(u^0, u^1, u^2) := (1 - u^0)\phi(u^1, u^2)$$

defined on $(0, 2\pi) \times (0, \pi) \times [0, \frac{1}{2}] \subset \mathbb{H}^3$. It is immediate that this is smooth, regular, and injective. Moreover, a point $\psi(u^0, u^1, u^2)$ lies in S^2 precisely if $u^0 = 0$, i.e., if and only if it is the image of a boundary point of $\mathbb{H}^2$. We thus conclude $B_1(0)$ is a 3-submanifold of $\mathbb{R}^3$ *with boundary* S^2.

This example illustrates a more general feature: the boundary of a k-submanifold with boundary is, itself a $(k-1)$-submanifold (whose boundary is empty).

Exercise 28. Let $M \subset \mathbb{R}^n$ be a k-submanifold with boundary ∂M. Show ∂M is a $(k-1)$-submanifold (without boundary), with charts defined by restricting the charts of M to the boundary of $\mathbb{H}^k$.

We define orientations for submanifolds with boundary completely analogously to orientations for manifolds without boundary.

Construction 6.34. Suppose that $M \subset \mathbb{R}^n$ is an oriented k-submanifold with boundary ∂M. Then the orientation of M induces a canonical orientation on the boundary of ∂M as follows.

At a boundary point $q \in \mathbb{H}^k$, we can take the tangent vector $-\partial_{x^1}$ to define a (constant) vector field on the boundary of $\mathbb{H}^k$. Fix a point $p \in \partial M$, and consider a chart $\phi : U \to M$ containing p, and so in particular defining a restricted chart $\psi := \phi|_{x^1=0}$

on ∂M containing p, the domain of this vector field is $V := U \cap \partial\mathbb{H}^k$. The pushforward $d\phi(-\partial x^1)$ defines a vector field V_{out} on $\psi(V)$. Vy construction, the vector field V_{out} takes values in TM, but not in $T\partial M$. We define a chart ζ on ∂M to be oriented if the vector fields $\{V_{\text{out}}, \partial_1\zeta, \dots, \partial_{k-1}\zeta\}$ induce an oriented basis of T_pM for every p in the image of ζ.

The reason to develop manifolds with boundary is to state a broad generalization of the fundamental theorem of calculus.

Theorem 6.35 (Stokes' theorem)**.** *Let $M \subset \mathbb{R}^n$ be a $(k+1)$-manifold with boundary ∂M. Let $\omega \in \Omega^k(\mathbb{R}^n)$. Then*

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

Corollary 6.36. *Let $M \subset \mathbb{R}^n$ be a compact k-sumbanifold with empty boundary, and let $\omega \in \Omega^{k-1}(\mathbb{R}^n)$. Then*

$$\int_M \mathrm{d}\omega = 0.$$

We will not prove this theorem here, but we will explain how it generalizes a wide variety of results you have likely seen in previous courses.

Example 6.37 (The fundamental theorem of calculus)**.** To retrieve the fundamental theorem of calculus, we note that a 0-manifold is a disjoint union of discrete points. An orientation on a zero manifold is an assignment of a $+$ or $-$ sign to each point. A 0-form on a 0-manifold X is simply a function $f : X \to \mathbb{R}$, and the oriented integral is

$$\int_X f = \sum_{x \in X} \operatorname{sgn}(x) f(x).$$

With this special case in mind, let $[a, b] \subset \mathbb{R}$ be a 1-manifold with boundary $\partial[a, b] = \{a, b\}$. We take the usual orientation on $\mathbb{R}$, and notice that the induced orientation assigns $\operatorname{sgn}(a) = -1$ and $\operatorname{sgn}(b) = +1$. Letting $f \in \Omega^0(\mathbb{R})$ be a smooth function, the formula of Stokes' Theorem then yields

$$\int_a^b \frac{df}{dx} dx = \int_{[a,b]} \mathrm{d}f = \int_{\{a,b\}} f = f(b) - f(a).$$

That is, the fundamental theorem of calculus[8]

[8] There are variants of Stokes' theorem for C^k-forms, which recover more general forms of the fundamental theorem of calculus.

Example 6.38 (The fundamental theorem of line integrals)**.** Generalizing the previous example, let $\gamma : [a, b] \to \mathbb{R}^n$ be a smooth curve, and $f \in \Omega^0(\mathbb{R}^n)$ a smooth function. We can use the metric to identify

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

with the vector field

$$X_{\mathrm{d}f} = \left(\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n} \right) = \operatorname{grad}(f).$$

By definition, this means

$$\mathrm{d}f(v) = \langle X_{\mathrm{d}f}, v \rangle.$$

We then notice that

$$\int_\gamma \mathrm{d}f = \int_a^b \mathrm{d}f(\gamma'(t))dt = \int_a^b \langle X_{\mathrm{d}f}, \gamma'(t)\rangle dt = \int_\gamma X_{\mathrm{d}f} \cdot ds.$$

Thus, the statement of Stokes' Theorem

$$\int_\gamma \mathrm{grad}(f) \cdot ds = \int_{\{\gamma(a),\gamma(b)\}} f = f(\gamma(b)) - f(\gamma(a)).$$

This is often known as the fundamental theorem of line integrals.

Example 6.39 (Green's Theorem)**.** Let $M \subset \mathbb{R}^2$ be a 2-submanifold with boundary ∂M. Equip M with the orientation induced by the standard orientation on $\mathbb{R}^2$. The induced orientation on the boundary traces counterclockwise around the bounding curve.[9] Let

$$\omega = \omega_1 \mathrm{d}x^1 + \omega_2 \mathrm{d}x^2$$

be a 1-form on $\mathbb{R}^2$. The exterior derivative of ω is

$$\mathrm{d}\omega = \frac{\partial \omega_1}{\partial x^2}\mathrm{d}x^2 \wedge \mathrm{d}x^1 + \frac{\partial \omega_2}{\partial x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2}\right)\mathrm{d}x^1 \wedge \mathrm{d}x^2.$$

We then compute, for a parameterization $\gamma : [a, b] \to \mathbb{R}^2$ of the boundary of M,

$$\begin{aligned}\int_{\partial M} \omega &= \int_a^b \omega(\gamma'(t))dt \\ &= \int_a^b \omega_1(\gamma(t))\frac{d\gamma^1}{dt} + \omega_2(\gamma(t))\frac{d\gamma^1}{dt}dt.\end{aligned}$$

On the other hand, we see that

$$\int_M \mathrm{d}\omega = \int_M \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2}\right) dx^1 dx^2.$$

Setting these two equal, we obtain Green's Theorem:

$$\int_\gamma \omega_1 dx^1 + \omega_2 dx^2 = \int_M \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2}\right) dx^1 dx^2.$$

Example 6.40. Let $M \subset \mathbb{R}^3$ be an oriented surface with normal n and boundary ∂M. Let

$$\omega = \omega_1 \mathrm{d}x^1 + \omega_2 \mathrm{d}x^2 + \omega_3 \mathrm{d}x^3 \in \Omega^1(\mathbb{R}^3)$$

with exterior derivative

$$\mathrm{d}\omega = \left(\frac{\partial \omega_2}{\partial x^3} - \frac{\partial \omega_3}{\partial x^2}\right)\mathrm{d}x^2 \wedge \mathrm{d}x^3 + \left(\frac{\partial \omega_1}{\partial x^3} - \frac{\partial \omega_3}{\partial x^1}\right)\mathrm{d}x^3 \wedge \mathrm{d}x^1 + \left(\frac{\partial \omega_2}{\partial x^1} - \frac{\partial \omega_1}{\partial x^2}\right)\mathrm{d}x^1 \wedge \mathrm{d}x^2.$$

Recall from Example 6.26 that we can use the metric to identify ω with a vector field X_ω with components ω_1, ω_2, and ω_3, s.t.

$$\omega(v) = \langle X_\omega, v\rangle.$$

[9] There is a subtlety here, since M could have multiple path components, but fortunately, integration treats every path component separately, and so we can without loss of generality assume that M has a single component.

Similarly, we can identify $\mathrm{d}\omega$ with a vector field $Y_{\mathrm{d}\omega}$ such that

$$\mathrm{d}f(v_1, v_2) = \langle Y_{\mathrm{d}\omega}, v_1 \times v_2 \rangle \, .$$

And, in particular, we have $Y_{\mathrm{d}\omega} = \mathrm{curl}(X_\omega)$.

We can thus compute

$$\begin{aligned} \int_{\partial M} X_\omega \cdot ds &= \int_{\partial M} \omega \\ &= \int_M \mathrm{d}\omega \end{aligned}$$

For a positively oriented chart $\phi : U \to M$, we have

$$\begin{aligned} \int_{\phi(U)} \mathrm{d}\omega &= \int_U \mathrm{d}\omega(\partial_1 \phi, \partial_2 \phi) \\ &= \int_U \langle \mathrm{curl}(X_\omega), \partial_1 \phi \times \partial_2 \phi \rangle \, du^1 du^2. \end{aligned}$$

Repurposing notation from calculus III, we see that

$$\int_{\phi(U)} \mathrm{d}\omega = \int_{\phi(U)} \mathrm{curl}(X_\omega) \cdot n \; dA.$$

Putting this all together, we see that we can rewrite Stokes' Theorem as

$$\int_{\partial M} X_\omega \cdot ds = \int_M \mathrm{curl}(X_\omega) \cdot n \; dA.$$

This is what, in calculus III, is called Stokes' Theorem.

Example 6.41 (The Divergence Theorem)**.** Let $M \subset \mathbb{R}^3$ be a 3-manifold with boundary, equipped with the orientation induced by the standard orientation on $\mathbb{R}^3$. Let $\omega \in \Omega^2(\mathbb{R}^3)$ be a 2-form. As in Example 6.26, we can identify this using the metric with a vector field Y_ω such that

$$\omega(v_1, v_2) = \langle Y_\omega, v_1 \times v_2 \rangle \, .$$

We can then note that

$$\mathrm{d}\omega = \mathrm{div}(Y_\omega) \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3.$$

Applying Stokes' Theorem, we then have

$$\begin{aligned} \int_{\partial M} Y_\omega \cdot n \; dA &= \int_{\partial M} \omega \\ &= \int_M \mathrm{d}\omega \\ &= \int_M \mathrm{div}(Y_\omega) \; dV. \end{aligned}$$

which is known as the *divergence theorem*.

Appendices

A

Isometries

Our goal here is to understand *isometries* of the n-dimensional Euclidean space $\mathbb{R}^n$. These are precisely those functions $f : \mathbb{R}^n \to \mathbb{R}^n$ which strictly preserve distances.

Definition A.1. A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is called an *isometry* if, for every $x, y \in \mathbb{R}^n$,

$$d(x, y) = d(f(x), f(y)).$$

It is immediate that f is injective, since, if $f(x) = f(y)$, we have

$$d(x, y) = d(f(x), f(y)) = 0$$

so that $x = y$.

Example A.2. There are some easy examples of isometries.

1. Let $v \in \mathbb{R}^n$. Then there is an isometry

$$\begin{aligned} T_v : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto x + v \end{aligned}$$

   called *translation by* v.

2. Let A be an $n \times n$ *orthogonal* matrix, i.e., a real matrix such that

$$A^T A = I.$$

   Then the linear map

$$\begin{aligned} A : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto Ax \end{aligned}$$

   is an isometry. (Exercise: show that this is true.)

We will show that, in a sense, these are the *only* isometries of $\mathbb{R}^n$. We will first focus on the case of *origin-preserving isometries*, i.e., isometries such that $f(0) = 0$.

Proposition A.3. *Suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is an origin-preserving isometry. Then $f(x) = Ax$ for some orthogonal matrix A.*

We will break this proposition into the following steps:

1. We first show that any isometry f sends lines to lines.
2. We then show that any origin-preserving isometry f must be linear.
3. We show that a linear map is an isometry if and only if it is orthogonal.

To begin with, we will need an exercise.

Exercise 29. Show that, for $x, y, z \in \mathbb{R}^n$, the following are equivalent

1. The points x, y, and z are colinear, and y lies on the line segment from x to z.
2. There is an equality
$$d(x, y) + d(y, z) = d(x, z).$$

Further show that, given a line $L \subset \mathbb{R}^n$ and any two distinct points $x, z \in L$, any point $y \in L$ is uniquely determined by $d(x, y)$ and $d(y, z)$.

With this exercise in hand, we can prove our first lemma

Lemma A.4. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an isometry and let $L \subset \mathbb{R}^n$ be a line. Then $f(L) \subset \mathbb{R}^n$ is a line.*

Proof. We first show that $f(L)$ is contained in a line. Let $x, y, z \in L$ be three distinct points, with y lying on the segment between x and z. Then

$$d(f(x), d(z)) = d(x, z) = d(x, y) + d(y, z) = d(f(x), f(y)) + d(f(y), f(z))$$

so that x, y, z are colinear. Thus, $f(L)$ is contained in the line P determined by x and z.

Now suppose w is a point in this line P. Then w is uniquely determined by $d(f(x), w)$ and $d(f(z), w)$. There is a unique point $y_w \in L$ such that

$$d(x, y_w) = d(f(x), w) \qquad \text{and} \qquad d(z, y_w) = d(f(z), w).$$

But then $d(f(x), f(y_w)) = d(f(x), w)$ and $d(f(z), f(y_w)) = d(f(z), w)$, so that $f(y_w) = w$. Hence, $f(L) = P$ as desired. □

Leveraging this fact, we now prove that any origin-preserving isometry is linear.

Lemma A.5. *Let f be an origin-preserving isometry. Then f is a linear map.*

Proof. First, let $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$. Then we have

$$|x| = |f(x)|$$
$$|\lambda x| = |f(\lambda x)|$$

so that $|f(\lambda x)| = |\lambda||f(x)|$. Moreover,

$$d(f(\lambda x), f(x)) = d(\lambda x, x) = (1 - \lambda)|x|$$

These two conditions uniquely determine the point $f(\lambda x)$ on the line between 0 and $f(x)$. Since $\lambda f(x)$ also satisfies these conditions, we thus have

$$f(\lambda x) = \lambda f(x),$$

so f preserves scalar multiplication.

Now let $x, y \in \mathbb{R}^n$. Since f is an isometry, it must map the midpoint between x and y to the midpoint between $f(x)$ and $f(y)$. That is

$$f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$$

However, since f preserves scalar multiplication, we also have

$$f\left(\frac{x+y}{2}\right) = \frac{1}{2}f(x+y)$$

and thus, $f(x+y) = f(x) + f(y)$, as desired. $\square$

We conclude the proof with an exercise

Exercise 30. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. Show that the following are equivalent:

1. f is an isometry.

2. f is orthogonal.

We have thus proven Proposition A.3. We can finally state and prove the main result we will need about isometries:

Theorem A.6. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an isometry. Then there is a unique vector $b \in \mathbb{R}^n$ and a unique orthogonal matrix A such that*

$$f(x) = Ax + b$$

for all $x \in \mathbb{R}^n$.

Proof. To show existence, let $f(0) = v$. Setting $b = -v$ shows us that

$$T_b \circ f$$

is an origin-preserving isometry, and thus is represented by an orthogonal matrix A. We then see that for any $x \in \mathbb{R}^n$,

$$f(x) + b = Ax$$

proving that f has the desired form. It is an easy exercise to see that this expression is unique. $\square$

B

Implicit and inverse function theorems

There are two key theorems from analysis we will use as black boxes throughout the course.

Theorem B.1 (Inverse function theorem)**.** *Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^n$ be open subsets, and let $f : U \to V$ be a C^k function. Suppose $x_0 \in U$ is a point such that the derivative Df_{x_0} is invertible. Then there is an open neigborhood $U_0 \subset U$ such that $x_0 \in U_0$, f is invertible on U_0, and f^{-1} is C^k on $f(U_0)$.*

Theorem B.2 (Implicit function theorem)**.** *Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open, and let $f : U \times V \to \mathbb{R}^m$ be a C^k function. Let $(a, b) \in U \times V$ be a point such that $f(a, b) = 0$, and write $f_a := f(a, -) : V \to \mathbb{R}^m$. If $D(f_a)$ is invertible at b, then there is are open sets $U_0 \subset U$ and $V_0 \subset V$ such that $a \in U_0$, $b \in V_0$, and a C^k function*

$$g : U_0 \longrightarrow V_0$$

such that $g(a) = b$, and, for $(x, y) \in U_0 \times V_0$, $f(x, y) = 0$ if and only if $y = g(x)$.

Example B.3. Let's do a familiar example, to get a feel for what the theorem is really saying. Let

$$U = \mathbb{R}^2 \qquad V = \mathbb{R}$$

and define

$$\begin{aligned} f : \quad U \times V &\longrightarrow \mathbb{R} \\ ((x, y), z) &\longmapsto x^2 + y^2 + z^2 - 1. \end{aligned}$$

This is a polynomial function, and thus is C^∞. The zero set of f is simply the unit 2-sphere in $\mathbb{R}^3$.

Consider the point $((0, 0), 1) \in U \times V$. We can compute the Jacobian matrix of $f_{(0,0)}$

$$D(f_{(0,0)}) = \left(\frac{\partial f_{(0,0)}}{\partial z} \right) = \left(2z \right)$$

at the point $z = 1$, this is simply the matrix containing 2, and thus is invertible. In consequence, the theorem guarantees the existence of a neighborhood U_0 of $(0, 0)$ in $\mathbb{R}^2$ and a neighborhood V_0 of 1 in $\mathbb{R}$, together with a C^∞ function

$$g : U_0 \longrightarrow V_0$$

such that $g(0,0) = 1$ and, on $U_0 \times V_0$, we have that $f((x,y),z) = 0$ if and only if $z = g(x,y)$. That is, the sphere is the graph of g on U_0.

The statement of theorem does not give us an explicit construction. However, in this case, it is not hard to identify U_0, V_0, and g. We can set

$$U_0 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\} \qquad V_0 = (0,2)$$

and then

$$\begin{aligned} g : \quad U_0 &\longrightarrow V_0 \\ (x,y) &\longmapsto \sqrt{x^2 + y^2 - 1} \end{aligned}$$

is the desired function.

C

Bilinear forms and self-adjoint maps

In our study of the first and second fundamental forms, we will need to recall some linear algebra: the lore of bilinear forms.

1 *Bilinear forms*

We first recall some linear algebra which gives us a general method for working with inner products.

Definition C.1. Let V be a finite-dimensional (real) vector space. A *bilinear form* on V is a linear map

$$B : V \otimes_{\mathbb{R}} V \longrightarrow \mathbb{R}$$

or, equivalently, a bilinear map

$$B : V \times V \longrightarrow \mathbb{R}.$$

We say that a bilinear form is

- *symmetric* if $B(v,w) = B(w,v)$ for all $v,w \in V$;
- *positive definite* if, for all $v \in V$, $B(v,v) > 0$; or
- *non-degenerate* if, whenever $B(v,w) = 0$ for all $w \in V$, we have $v = 0$.

Remark C.2. Note that every positive-definite form in non-degenerate.

Example C.3. 1. The Euclidean inner product

$$\langle -,- \rangle : \mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{R}^n \longrightarrow \mathbb{R}$$

is a non-degenerate, positive definite, symmetric bilinear form on $\mathbb{R}^n$.

2. The Minkowski inner product

$$\langle v,w \rangle_{1,1} = v_1 w_1 - v_2 w_2$$

is a symmetric non-degenerate bilinear form on $\mathbb{R}^2$.

In general, we can represent bilinear forms in terms of matrices. Given a basis $\mathcal{V} = (v_1, \ldots, v_k)$ for V and a bilinear form $B : V \otimes_{\mathbb{R}} V \to \mathbb{R}$, we can form the matrix

$$A^{\mathcal{V}}_{i,j} = B(v_i, v_j)$$

If we represent vectors in V as column vectors with respect to the basis $\mathcal{V}$, we can then write, by bilinearity,

$$B(w, u) = w^T_{\mathcal{V}} A^{\mathcal{V}} v_{\mathcal{V}}.$$

Given another basis $\mathcal{U} = (u_1, \ldots, u_k)$ for V with M the change of basis matrix, it is not hard to check that

$$A^{\mathcal{U}} = M A^{\mathcal{V}} M^T.$$

Notice that if B is symmetric, then for any basis $\mathcal{V}$ of V, the matrix $A^{\mathcal{V}}$ is symmetric.

Definition C.4. Let B be a symmetric bilinear form on V. We call a basis $v_1, \ldots, v_n$ B-*orthonormal* (or just *orthonormal*) if $B(v_i, v_j) = \delta_{i,j}$. Notice that this is equivalent to requiring that the matrix of B with respect to this basis is the identity.

Definition C.5. Let B be a symmetric bilinear form on V, and let $W \subset V$ be a subspace. We denote by $W^\perp \subset V$ the subspace

$$W^\perp = \{v \in V \mid B(v, w) = 0 \; \forall w \in W\}$$

and call W the *orthogonal complement* of W in V.

Lemma C.6. *Let B be a non-trivial symmetric bilinear form on V. Then there is a vector v such that $B(v, v) \neq 0$.*

Proof. Choose $u, w \in V$ such that $B(u, w) \neq 0$, and suppose that $B(u, u) = B(w, w) = 0$. Then

$$B(u + w, u + w) = 2B(u, w) + B(u, u) + B(w, w) = 2B(u, w) \neq 0.$$

□

Lemma C.7. *Let B be a non-trivial symmetric bilinear form on V, and let $v \in V$ such that $B(v, v) \neq 0$. Then every $w \in V$ can be uniquely written as*

$$w = \lambda v + w_2$$

where $B(v, w_2) = 0$ (i.e., $w_2 \in \mathrm{Span}(v)^\perp$). In particular, $\mathrm{Span}(v)^\perp$ has dimension $\dim(V) - 1$.

Proof. Define $w_2 := w - \frac{B(w,v)}{B(v,v)} v$, and note that

$$\begin{aligned} B(w_2, v) &= B(w, v) - \frac{B(w, v)}{B(v, v)} B(v, v) \\ &= 0 \end{aligned}$$

Then by definition

$$w := w_2 + \frac{B(w, v)}{B(v, v)} v$$

as desired.

On the other hand, if

$$w = \lambda v + w_2$$

is such a decomposition, then

$$B(w, v) = \lambda B(v, v) + B(w_2, v) = \lambda B(v, v).$$

proving the uniqueness of the given decomposition. □

Lemma C.8. *Let V be a finite dimensional real vector space, $B : V \otimes_{\mathbb{R}} V \to \mathbb{R}$ a bilinear form, and $W \subseteq V$ a subspace. Then B restricts to a bilinear form $B|_W$ on W which is*

1. *Symmetric when B is.*
2. *Positive definite when B is. In particular, $B|W$ is then non-degenerate.*

Lemma C.9. *Let B be a symmetric bilinear form. Then*

1. *There is a basis in which the matrix representation of B is diagonal.*
2. *If B is additionally positive definite, then there is a basis in which B is represented by the identity matrix, i.e., a B-orthonormal basis.*

Proof. We begin with (1). If B is the trivial bilinear form, we are done, since the corresponding matrix is simply the zero matrix. So we assume, without loss of generality, that B is non-trivial.

For an non-trivial B, we proceed by induction on the dimension. In dimensions 0 and 1, the statement is immediate. Suppose that the statement is true in dimension k, and let $\dim(V) = k + 1$. Choose a vector $v \in K$ such that $B(v, v) \neq 0$. Then B restricts to a symmetric bilinear form on $W = \mathrm{Span}(v)^{\perp}$, and by the inductive hypothesis, W has a basis $(w_1, \ldots, w_k)$ in which $B|_W$ is diagonal. Then $(w_1, \ldots, w_k, v)$ is a basis of V, and B is diagonal with respect to this basis.

To see (2), note that if B is positive definite, then in any basis $\mathcal{V}$ of V, the matrix $A^{\mathcal{V}}$ of B with respect to V has diagonal entries $B(v_i, v_i) > 0$. Rescaling the basis elements by $\sqrt{B(v_i, v_i)}$ then yields that the diagonal entries are all 1. Thus (2) follows from (1). □

Proof. Part (1) is immediate, as $B(v, w) = B(w, v)$ regardless of what space we restrict to. Part (2) is similarly simple. □

Example C.10. Notice that, for $B|W$ to be non-degenerate, it is not sufficient to require that B is simply non-degenerate. Consider, for example, the case when $B = \langle -, - \rangle_{1,1}$ is the Minkowski inner product on $\mathbb{R}^2$ and our chosen subspace is $V \subset \mathbb{R}^2$ given by $V = \{(v^1, v^2) \in \mathbb{R}^2 \mid v^1 = v^2\}$. The restriction of $\langle -, - \rangle_{1,1}$ to V is identically zero, even though the form $\langle -, - \rangle_{1,1}$ is nondegenerate on $\mathbb{R}^2$.

2 Self-adjoint operators

We now interrupt our regularly scheduled programming to digress into linear algebra. Our setup is as follows:

- V is a k-dimensional $\mathbb{R}$-vector space.
- $B(-,-)$ is a symmetric, positive-definite, non-degenerate bilinear form on V.

We refer to the pair (V, B) as an *inner product space.*

Definition C.11. We call a linear map $L : V \to V$ a *self-adjoint operator* if, for every $v, w \in V$,

$$B(L(v), w) = B(v, L(w)).$$

Lemma C.12. *Let $L : V \to V$ be a self-adjoint map. Then $B(L(-), -) : V \times V \to \mathbb{R}$ is a symmetric bilinear form.*

Proof. Bilinearity is immediate from the bilinearity of B and the linearity of L. To see symmetry, note that

$$B(L(v), w) = B(v, L(w)) = B(L(w), v).$$

This completes the proof. □

To explore the properties of self-adjoint operators, we will first establish some further results about V and B.

Definition C.13. The *dual space* of V is the vector space $V^\vee := \mathrm{Lin}(V, \mathbb{R})$ of linear maps from V to $\mathbb{R}$. Given $v \in V$, we denote the linear map

$$\begin{aligned} B(v,-) : V &\longrightarrow \mathbb{R} \\ w &\longmapsto B(v,w) \end{aligned}$$

by $v^* \in V^\vee$. The map

$$\begin{aligned} \beta : V &\longrightarrow V^\vee \\ v &\longmapsto v^* \end{aligned}$$

is a linear map.

Lemma C.14. *The map $\beta : V \to V^\vee$ is an isomorphism.*[1]

Proof. To show surjectivity, let $v_1, \dots, v_k$ be an orthonormal basis of V, and let $\ell \in V^\vee$. Then

$$\ell\left(\sum_{i=1}^k \lambda^i v_i\right) = \sum_{i=1}^k \lambda^i \ell(v_i).$$

We can thus see that

$$\ell = \sum_{i=1}^k \ell(v_i) v_i^* = \beta(\sum_{i=1}^k \ell(v_i) v_i).$$

To see injectivity, suppose that $v^* = 0$. Then, for all $w \in V$,

$$0 = \beta(v)(w) = B(v, w).$$

by the non-degeneracy of B, this implies that $v = 0$. □

[1] This is really a consequence of the non-degeneracy of B.

Proposition C.15. *Let $L : V \to V$ be a self-adjoint operator. Then there is a B-orthonormal basis of V in which L is diagonal.*

Proof. Choose a B-orthonormal basis $v_1, \ldots, v_k$ of V, and let M be the matrix representing L in this basis. Then we have

$$M_{i,j} = B(L(v_i), v_j) = B(v_i, L(v_j)) = M_{j,i}$$

So that M is a symmetric matrix. There thus exists an orthogonal matrix O and a diagonal matrix A such that

$$M = OAO^T.$$

Define a new basis of V by

$$w_j = \sum_{i=1}^{k} O_{i,j} v_i.$$

So that, with respect to this basis, we have

$$L(w_i) = \sum_{j=1}^{k} O_{j,i} L(v_j) = \sum_{j=1}^{k} O_{j,i} \sum_{\ell=1}^{k} M_{j,\ell} v_\ell = \sum_{j=1}^{k} O_{j,i} \sum_{\ell=1}^{k} M_{j,\ell} \sum_{m=1}^{k} O_{\ell,m} w_m$$

so that the matrix of L is $O^T M O = A$.

Moreover, since O is an orthogonal matrix, we have

$$B(w_j, w_\ell) = B\left(\sum_{i=1}^{k} O_{i,j} v_i, \sum_{m=1}^{k} O_{m,\ell} v_m\right) = \sum_{i=1}^{k} \sum_{m=1}^{k} O_{i,j} O_{m,\ell} \delta_{i,m} = (O^T O)_{i,j} = \delta_{i,j}$$

so that $w_1, \ldots, w_k$ is a B-orthonormal. □

Lemma C.16. *Let $L : V \to V$ be a self-adjoint map. Then $D(-,-) = B(L(-),-)$ is a symmetric bilinear form.*

It turns out that this process — obtaining a symmetric bilinear from from a self-adjoint map — is reversable.

Lemma C.17. *Let $L : V \to V$ be a self-adjoint map, and let $D(-,-) = B(L(-),-)$ be the associated symmetric bilinear form. Choose a basis of V, and write ℓ, d, and b for the matrices representing L, D, and B, respectively. Then*

$$\ell = b^{-1} d$$

Proof. We know that

$$D(v, w) = B(L(v), w)$$

for all $v, w \in V$. Then, with respect to our chosen basis v_i, we have

$$d_{i,j} = B\left(\sum_{n=1}^{k} \ell_{n,i} v_n, v_j\right) = \sum_{n=1}^{k} \ell_{n,i} b_{n,j}$$

So

$$d = b\ell$$

Since b is invertible, this means that $\ell = b^{-1} d$, as desired. □

Index

Bibliography

[1] Wilhelm Klingenberg. *A Course in Differential Geometry.* Springer Graduate Texts in Mathematics, Volume 51. ISBN: 978-1461299257

[2] Wolfgang Kühnel. *Differential Geometry: Curves — Surfaces — Manifolds.* AMS. ISBN: 978-1470423209

[3] Manfredo P. Do Carmo. *Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition.* Dover Books on Mathematics. ISBN: 978-0486806990

[4] Ballmann, W. *Introduction to Geometry and Topology.* Birkhäuser. 2018. ISBN: 978-3034809825. Trans: Walker H. Stern.

[5] Lebowitz, Norman R. *Ordinary Differential Equations.* Available on the author's webpage.