NOTES OF MATH 451, CSULB

YU DING

Prerequisites: Some familiarity with linear algebra (MATH 247), calculus of several variables (MATH 224). Not really need a course in ODE (MATH 364A or 370A) - you just need to take for granted the existence and uniqueness theorem in ODE - its proof is not given in 364A/370A anyway.

E-MAIL TYPOS, ERRORS TO yu.ding@csulb.edu

An exercise with a star (*) means this exercise needs knowledge from another course (e.g. linear algebra) that is not covered in these notes. They may not be harder than other exercises.

CONTENTS

1. Vectors and curves 3
2. Parametric surfaces 8
3. Some linear algebra 11
4. More linear algebra 15
5. Length, angle, area and volume 19
6. Second fundamental form 26
7. Theorema Egregium 32
8. Triangle Gauss-Bonnet theorem 37
9. Global Gauss-Bonnet theorem 41
10. Higher dimensions: manifolds and tangent vectors 44
11. Geodesics: first variation formula 50
12. Connection 52
13. Theorem of Clairaut 57
14. Levi-Civita as induced connection 58
15. Space of constant negative curvature 61
16. Curvature tensor 65

Date: March 12, 2017.
17. Jacobi field and meaning of curvature 67
18. Bonnet-Myers theorem and meaning of curvature 69
19. General relativity 71
Further readings 72
1. Vectors and curves

The 3-dimensional Euclidean space $\mathbb{R}^3$ contains 3-dimensional vectors, which can be written as a column with three entries:

\begin{equation}
A = \begin{bmatrix} 
  a_1 \\
  a_2 \\
  a_3 
\end{bmatrix}.
\end{equation}

$a_1$ is the $x$-coordinate of $A$, $a_2$ is the $y$-coordinate of $A$, $a_3$ is the $z$-coordinate of $A$. Writing a vector as a column instead of a row is a rather arbitrary choice.\(^1\)

You can think a vector $A$ as above is an arrow whose initial point is the origin in $\mathbb{R}^3$, whose terminal point has coordinate $(a_1, a_2, a_3)$. In this picture, we do not really need a vector to start at the origin, we view two vectors to be identical if they have the same direction and the same length.

For any vector

\begin{equation}
A = \begin{bmatrix} 
  a_1 \\
  a_2 \\
  a_3 
\end{bmatrix}
\end{equation}

its length (or absolute value) is

\begin{equation}
|A| = \sqrt{a_1^2 + a_2^2 + a_3^2}.
\end{equation}

For any vector

\begin{equation}
A = \begin{bmatrix} 
  a_1 \\
  a_2 \\
  a_3 
\end{bmatrix}
\end{equation}

and any number $c$, we can multiply them together and get

\begin{equation}
cA = \begin{bmatrix} 
  ca_1 \\
  ca_2 \\
  ca_3 
\end{bmatrix}.
\end{equation}

If $c \geq 0$, then $cA$ is the vector that is in the same direction of $A$ with length $c|A|$. If $c < 0$, then $cA$ is the vector that is in the opposite direction of $A$ with length $|c| \cdot |A|$.

For two vectors

\begin{equation}
A = \begin{bmatrix} 
  a_1 \\
  a_2 \\
  a_3 
\end{bmatrix}, \\
B = \begin{bmatrix} 
  b_1 \\
  b_2 \\
  b_3 
\end{bmatrix},
\end{equation}

we can add them together:

\begin{equation}
A + B = \begin{bmatrix} 
  a_1 \\
  a_2 \\
  a_3 
\end{bmatrix} + \begin{bmatrix} 
  b_1 \\
  b_2 \\
  b_3 
\end{bmatrix} = \begin{bmatrix} 
  a_1 + b_1 \\
  a_2 + b_2 \\
  a_3 + b_3 
\end{bmatrix}.
\end{equation}

\(^1\)However the fact that the writing system of English (and many other languages) writes from left to right plays a role in this choice.
The geometric meaning of adding two vectors can be seen from the following picture:

![Vector Addition Diagram](image)

For two vectors

\[(1.8)\quad A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},\]

define their dot product to be

\[(1.9)\quad A \cdot B = a_1b_1 + a_2b_2 + a_3b_3.\]

**Exercise 1.** Prove that

\[(1.10)\quad A \cdot A = |A|^2 = a_1^2 + a_2^2 + a_3^2.\]

Moreover, prove that

\[(1.11)\quad A \cdot B = A^T B = B^T A;\]

here in $A^T B$, $B^T A$ we view $A, B$ as $3 \times 1$ matrices; and for any matrix $X$ we use $X^T$ to denote its transpose, i.e. if $X$ is an $m \times n$ matrix whose entry at the $i$-th row, $j$-th column is $x_{ij}$, then $X^T$ is an $n \times m$ matrix whose entry at the $i$-th row, $j$-th column is $x_{ji}$; in particular here $A^T, B^T$ are just to write $A, B$ as rows.

**Lemma 1.12.** In the three dimensional Euclidean space $\mathbb{R}^3$,

\[(1.13)\quad A \cdot B = |A| \cdot |B| \cdot \cos \angle(A, B).\]

Especially, $A \cdot B = 0$ if and only if $A \perp B$.

**Proof.** By law of cosine.

**Exercise 2.** Give detailed proof of the above lemma.

A curve $\gamma(t)$ is just a continuous map from $\mathbb{R}$ to $\mathbb{R}^3$. So a curve is a vector that depends on one single parameter $t$, which we typically interpret as time. Therefore we can write a curve as

\[(1.14)\quad \gamma(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix},\]

where $x, y, z$ are continuous functions of $t$. You remember that in your Calculus class this was called the **parametric equation** of a curve. In the following we will always assume that $x, y, z$ are (infinitely) differentiable in $t$; such curves are called **smooth**.
The derivative of \( \gamma(t) \) at \( t \) is

\[
\gamma'(t) = \lim_{\Delta t \to 0} \frac{\gamma(t + \Delta t) - \gamma(t)}{\Delta t} = \begin{bmatrix} x'(t) \\ y'(t) \\ z'(t) \end{bmatrix},
\]

and \( \gamma'(t_0) \) is a tangent vector of \( \gamma \) at \( \gamma(t_0) \).

In the following we will consider only those curve \( \gamma \) so that

\[
\gamma'(t) \neq 0 \quad \text{for any } t.
\]

**Exercise 3.** What’s wrong with \( \gamma'(t) = 0 \)? Sketch the following two examples:

\[
\alpha(t) = \begin{bmatrix} t^3 \\ t^6 \\ 0 \end{bmatrix}, \quad \beta(t) = \begin{bmatrix} t^3 \\ t^2 \\ 0 \end{bmatrix}.
\]

You can view the curve as a highway, and \( t \) is the time, thus the parametrization \( \gamma(t) \) describes how you drive on the highway: at time \( t \) you are at the location \( \gamma(t) \). In particular, the condition \( \gamma'(t) \neq 0 \) guarantees that you can only drive one-way: you cannot stop, make a U-turn on the highway and travel backwards!

Now with the above highway interpretation, \( \gamma'(t) \) is your velocity. Thus \( |\gamma'(t)| \) is the speed, i.e. magnitude of your velocity. Recall if you take derivative of the distance you drove, you get speed. Thus if you integrate speed, you get the distance. In particular, we have the arclength formula: the length of the curve of the portion \( a \leq t \leq b \) is

\[
\int_a^b |\gamma'(t)| \, dt = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2} \, dt.
\]

On can similarly consider curves in the \( n \)-dimensional space \( \mathbb{R}^n \) as a smooth function from \( \mathbb{R} \) to \( \mathbb{R}^n \) with non-vanishing derivative.

**Lemma 1.18.** Assume \( A = A(t) \), \( B = B(t) \) are vector functions of one single variable, \( t \). Then

\[
\frac{d}{dt}[A(t) \cdot B(t)] = A'(t) \cdot B(t) + A(t) \cdot B'(t).
\]

**Exercise 4.** Give detailed proof of the above lemma.

**Corollary 1.20.** Assume \( A = A(t) \) is a vector functions of one single variable, \( t \). If its length \( |A(t)| \) is a constant, then

\[
A'(t) \perp A(t).
\]
Exercise 5. Give detailed proof of the above lemma.

A curve $\gamma(t)$ has an arclength parameter $s$, where

$$s = \int_0^t |\gamma'(\tau)|d\tau. \tag{1.22}$$

Therefore

$$\frac{ds}{dt} = |\gamma'(t)| \neq 0. \tag{1.23}$$

So we have

$$\frac{dt}{ds} = \frac{1}{|\gamma'(t)|}. \tag{1.24}$$

So we can view $\gamma$ as a function of $s$,

$$\gamma'(s) = \frac{d\gamma}{ds} = \frac{d\gamma}{dt} \frac{dt}{ds} = \gamma'(t) \cdot \frac{1}{|\gamma'(t)|}. \tag{1.25}$$

In particular,

$$|\gamma'(s)| = |\gamma'(t)| \cdot \frac{1}{|\gamma'(t)|} = |\gamma'(t)| \cdot \frac{1}{|\gamma'(t)|} = 1. \tag{1.26}$$

We will use arclength parameter in theoretical arguments. The unit tangent vector of $\gamma$ is

$$T = \gamma'(s). \tag{1.27}$$

So $|T| = 1$ everywhere.  

Define

$$\kappa = |T'(s)| \tag{1.28}$$

to be the curvature of $\gamma$. If $k \neq 0$, call

$$n = T'(s)/|T'(s)| \tag{1.29}$$
	he normal vector of $\gamma$. So

$$T'(s) = \kappa n, \quad |T| = |n| = 1, \quad T \perp n. \tag{1.30}$$

In particular, $T'(s) \perp T$; i.e. if you travel at constant speed, then your acceleration $a$ is perpendicular to your velocity $v$. In fact, if $a \perp v$ then the acceleration $a$ is used only to change the direction of $v$, it cannot change the speed.
Exercise 6. *Find a formula of curvature that does not involve \( s \).*

*Answer:*

(1.31) \[ \kappa = \left| \frac{\gamma'(t)\gamma''(t) - \gamma'(t) \cdot \gamma''(t)}{|\gamma'(t)|^3} \right|. \]

Exercise 7. *Compute the curvature of a straight line,*

(1.32) \[ \gamma(t) = \begin{bmatrix} at \\ bt \\ ct \end{bmatrix}. \]

*Answer: 0.*

Exercise 8. *Compute the curvature of the circle of radius \( R \):*

(1.33) \[ \gamma(t) = \begin{bmatrix} R \cos t \\ R \sin t \end{bmatrix}. \]

*Answer: 1/R.*

Exercise 9. *Compute the curvature of*

(1.34) \[ \gamma(t) = \begin{bmatrix} \cos t \\ \sin t \\ t \end{bmatrix}. \]

Exercise 10. *Compute the curvature of the parabola \( y = x^2 \).*

Exercise 11. *Compute the curvature of \( y = x^3 \).*

Curvature measures “how curved” a curve is. We can consider the 2-dimensional case. Notice \( T \) is a unit vector, we can write

(1.35) \[ T(s) = \begin{bmatrix} \cos \theta(s) \\ \sin \theta(s) \end{bmatrix}. \]

Here \( \theta = \theta(s) \) is the angle between the \( x \)-axis and \( T \). So by the chain rule,

(1.36) \[ T'(s) = \begin{bmatrix} -\sin \theta(s) \\ \cos \theta(s) \end{bmatrix} \frac{d\theta}{ds}, \]

thus

(1.37) \[ \kappa = |T'(s)| = \left| \frac{d\theta}{ds} \right|. \]

Since \( \theta \) is the direction of the curve, we see \( \kappa \) is the “rate of change in direction”. So \( \kappa \) measures how curved \( \gamma \) is.

**Lemma 1.38.** Assume \( \gamma \) is a closed convex loop, let \( |\gamma| \) be the total length of \( \gamma \), then

(1.39) \[ \int_{\gamma} \kappa ds = 2\pi. \]
Proof. We go along $\gamma$ counterclockwise, then $\theta$ is an increasing function of $s$, thus $\frac{d\theta}{ds} > 0$. So $\kappa = \frac{d\theta}{ds}$. So

$$\int_\gamma \kappa ds = \int_0^{\left|\gamma\right|} \frac{d\theta}{ds} ds = \theta\left|\gamma\right|_0 = 2\pi.$$

This is the first global result.

Finally, if you remember the Taylor series, we can expand a curve into power series. Assume for the curve $\gamma$ we use the arclength parameter $s$, then

$$\gamma(s) = \gamma(0) + \gamma'(0)s + \frac{\gamma''(0)}{2}s^2 + \int_0^s (s - \theta)^2 \frac{\gamma^{(3)}(\theta)}{2} d\theta$$

We can check, because $\gamma^{(3)}$ is continuous, for some constant $C$ we have the integral is $< C|s|^3$ (Exercise!). Notice

$$\gamma'(s) = T, \quad \gamma''(s) = T'_x = \kappa n,$$

we see

$$\gamma(s) = \gamma(0) + Ts + \frac{\kappa n}{2}s^2 + O(s^3),$$

where $O(s^3)$ (the integral) is a quantity $< C|s|^3$. So if we ignore the error term $O(s^3)$, which is much smaller than other terms when $s$ is small, we see

$$\gamma(s) \approx \gamma(0) + Ts + \frac{\kappa n}{2}s^2,$$

This is a parabola, lies in the plane passing through $\gamma(0)$ that contains the (perpendicular) vectors $T$ and $n$. This plane is called the osculating plane of $\gamma$ at $\gamma(0)$.

2. Parametric surfaces

Assume $\sigma : \Omega \rightarrow \mathbb{R}^3$,

$$\sigma(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}$$

is a (local) parametric equation for a surface $\Sigma$; here $\Omega$ is an open domain in the $u, v$ plane. We will also call it a 2-surface when there might be confusions. We require that $x, y, z$ are, for convenience, infinite differentiable. We also assume that the map $(u, v) \rightarrow \sigma(u, v)$ is one-to-one.

In general, a “global” surface does NOT have a coordinate system $u, v$ that covers all of it. We can only be satisfied with “local” coordinates. You can convince yourself on this by looking at a sphere or a torus.

Exercise 12. Write a surface of the form $z = f(x, y)$ is parametric form.
Example 2.2. We can write the sphere $x^2 + y^2 + z^2 = 1$ in the following parametric form:

\[
\sigma(\phi, \theta) = \begin{bmatrix}
\sin \phi \cos \theta \\
\sin \phi \sin \theta \\
\cos \phi
\end{bmatrix}.
\]

Exercise 13. What does the following surface look like?

Exercise 14. Prove that the surface $x^2 + y^2 - z^2 = 1$ is a ruled surface.

Exercise 15. (optional) use software, e.g. Matlab, Octave, Maple, etc. to sketch any surfaces you can find on internet.

For example, the Enneper surface

\[
\sigma(u, v) = \begin{bmatrix}
(1 - \frac{u^2}{3} + \frac{v^2}{3})u \\
-(1 - \frac{v^2}{3} + \frac{u^2}{3})v \\
\frac{1}{3}(u^2 - v^2)
\end{bmatrix}.
\]

And Boy’s surface, see http://en.wikipedia.org/wiki/Boy’s_surface

More generally, we consider an n dimensional surface $\Sigma$ in the m-dimensional Euclidean space defined by the parametric equation

\[
\sigma(x_1, x_2, ..., x_n) = \begin{bmatrix}
y_1(x_1, ..., x_n) \\
y_2(x_1, ..., x_n) \\
y_3(x_1, ..., x_n) \\
......
\end{bmatrix},
\]

where $(x_1, ..., x_n)$ are in some domain $\Omega \subset \mathbb{R}^n$. For simplicity, we say $\Sigma$ is an n-surface in $\mathbb{R}^m$.

Example 2.6. We can write the graph of the function $f(x_1, ..., x_n)$ in the following parametric form:

\[
\sigma(x_1, ..., x_n) = \begin{bmatrix}
x_1 \\
x_2 \\
......
\end{bmatrix},
\]

this is an n-surface in $\mathbb{R}^{n+1}$. 
Example 2.7. Consider the surface 

\[ \sigma(\phi, \theta) = \begin{bmatrix} \cos \phi \\ \sin \phi \\ \cos \theta \\ \sin \theta \end{bmatrix} \].

Here \( 0 \leq \phi, \theta < 2\pi \). You can check this is a torus. Later we will see this is a flat torus.

In the following, we will discuss the theory of surface in \( \mathbb{R}^3 \) and the theory of \( n \)-surface in \( \mathbb{R}^m \) in parallel. You will see that the former is just a special case of the latter.

For an \( n \)-surface in \( \mathbb{R}^m \), write

\[ \sigma_1 = \begin{bmatrix} \partial y_1/\partial x_1 \\ \partial y_2/\partial x_1 \\ \vdots \\ \partial y_m/\partial x_1 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} \partial y_1/\partial x_2 \\ \partial y_2/\partial x_2 \\ \vdots \\ \partial y_m/\partial x_2 \end{bmatrix}, \quad \ldots, \quad \sigma_n = \begin{bmatrix} \partial y_1/\partial x_n \\ \partial y_2/\partial x_n \\ \vdots \\ \partial y_m/\partial x_n \end{bmatrix}. \]

For a 2-surface, we will use slight different notations:

\[ \sigma_u = \begin{bmatrix} \partial x/\partial u \\ \partial y/\partial u \\ \partial z/\partial u \end{bmatrix}, \quad \sigma_v = \begin{bmatrix} \partial x/\partial v \\ \partial y/\partial v \\ \partial z/\partial v \end{bmatrix}. \]

For an \( n \)-surface \( \Sigma \) in \( \mathbb{R}^m \), we assume \( \sigma_1, \ldots, \sigma_n \) are linearly independent, i.e. at any point, if for some \( c_1, \ldots, c_n \) we have

\[ c_1 \sigma_1 + \ldots + c_n \sigma_n = 0, \]

then \( c_1 = c_2 = \ldots = c_n = 0 \). The reason we require this condition is to guarantee that \( \Sigma \) is an “\( n \)-manifold” - later we will talk more about this condition, and other conditions that are equivalent to it.

Exercise 16. For a 2-surface, the above condition becomes

\[ \sigma_u \times \sigma_v \neq 0, \]

where \( \times \) is the cross product in \( \mathbb{R}^3 \).

Furthermore, prove that the above is equivalent to that \( \sigma_u, \sigma_v \) are not in the same line.

A tangent vector of \( \Sigma \) at \( p \in \Sigma \) is a vector that is tangent to some curve in \( \Sigma \) that passes through \( p \in \Sigma \).

Lemma 2.11. Any tangent vector of an \( n \)-surface \( \Sigma \) at \( \sigma(x_1, \ldots, x_n) \) is a linear combination of \( \sigma_1, \ldots, \sigma_n \).

Proof. In local coordinate \( x_1, \ldots, x_n \), we can represent a curve by \( (x_1(t), \ldots, x_n(t)) \), i.e. the curve has parametric equation

\[ \gamma(t) = \begin{bmatrix} y_1(x_1(t), \ldots, x_n(t)) \\ y_2(x_1(t), \ldots, x_n(t)) \\ \vdots \\ y_m(x_1(t), \ldots, x_n(t)) \end{bmatrix}. \]
Thus

\[
\gamma'(t) = \begin{bmatrix}
\frac{\partial y_1}{\partial x_1} x_1'(t) + \ldots + \frac{\partial y_1}{\partial x_n} x_n'(t) \\
\frac{\partial y_2}{\partial x_1} x_1'(t) + \ldots + \frac{\partial y_2}{\partial x_n} x_n'(t) \\
\vdots \\
\frac{\partial y_m}{\partial x_1} x_1'(t) + \ldots + \frac{\partial y_m}{\partial x_n} x_n'(t)
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial y_1}{\partial x_1} \\
\frac{\partial y_2}{\partial x_1} \\
\vdots \\
\frac{\partial y_m}{\partial x_1}
\end{bmatrix} x_1'(t) + \ldots + 
\begin{bmatrix}
\frac{\partial y_1}{\partial x_n} \\
\frac{\partial y_2}{\partial x_n} \\
\vdots \\
\frac{\partial y_m}{\partial x_n}
\end{bmatrix} x_n'(t)
\]

= x_1'(t) \cdot \sigma_1 + x_2'(t) \cdot \sigma_2 + \ldots + x_n'(t) \cdot \sigma_n.

\]

□

In particular, any tangent vector of a 2-surface \( \Sigma \) at \( \sigma(u, v) \) is a linear combination of \( \sigma_u, \sigma_v \).

Generally, the tangent plane of \( \Sigma \) at \( p \in \Sigma \) is the plane spanned by \( \sigma_1, \ldots, \sigma_n \) at \( p \).

3. Some linear algebra

First, we review some notions in linear algebra. Consider the \( n \)-dimensional Euclidean space \( \mathbb{R}^n \), which contains all \( n \)-dimensional vectors; we can also think it contains all \( n \times 1 \) matrices (i.e. “\( n \)-dimensional column vectors”). Take \( n \) many vectors \( v_1, v_2, \ldots, v_n \), they space a “parallelepiped"

\[
P = \{ t_1 v_1 + \ldots + t_n v_n | 0 \leq t_1 \leq 1, \ 0 \leq t_2 \leq 1, \ \ldots, 0 \leq t_n \leq 1 \};
\]

define \( A(v_1, v_2, \ldots, v_n) \) to be the “\( n \)-dimensional volume” of \( P \). We put a quote sign here since we will see that it is most convenient to allow the “volume” to be negative sometimes. The function \( A \) shall satisfy the following conditions:

i. \( A \) is linear in each variable: for any \( k \), and any \( a, b \in \mathbb{R} \),

\[
A(v_1, \ldots, av_k + bv_k, \ldots, v_n) = aA(v_1, \ldots, u_k, \ldots, v_n) + bA(v_1, \ldots, v_k, \ldots, v_n).
\]

ii. If any two among \( v_1, v_2, \ldots, v_n \) are identical, then the \( A \) value is 0:

\[
A(\ldots, u, \ldots, u, \ldots) = 0.
\]

iii. For the “standard basis” of \( \mathbb{R}^n \),

\[
e_1 = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}, \quad e_2 = \begin{bmatrix}
0 \\
1 \\
\vdots \\
0
\end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix}
0 \\
0 \\
\vdots \\
1
\end{bmatrix},
\]

we have

\[
A(e_1, \ldots, e_n) = 1.
\]

You can draw a picture to convince yourself that i. must be true, at least when \( a, b > 0 \). It is most convenient to allow any \( a, b \), even negative (being linear is too good to abandon!). Then \( A \) must sometimes take negative values. Condition ii. says that a “degenerate” parallelepiped has volume 0. Condition iii. says that the volume of a standard cube is 1 (essentially this fixes the unit of volume - we have to assign volume 1 to some standard shape).
Exercise 17. Prove that the second condition ii. implies that $A$ is anti-symmetric:

$$A(..., v, ..., u, ...) = - A(..., u, ..., v, ...)$$

Exercise 18. Prove that the above conditions decides the $A$ function uniquely.

Prove that when $n = 2$,

$$A \left( \begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix} \right) = ad - bc.$$

In general, define the determinant of an $n \times n$ matrix $F$ to be

$$|F| = \det F = A(f_1, ..., f_n),$$

where $f_1, ..., f_n$ are the $n$ columns of $F$.

Exercise 19. * Get a general formula for $|F|$: prove that

$$|F| = \sum_{\sigma} (-1)^{\text{sign} \sigma} f_{1, \sigma(1)} f_{2, \sigma(2)} ... f_{n, \sigma(n)}.$$

Here $f_{i,j}$ is the entry on the $i$-th row and $j$-th column of $F$, the summation runs through all permutations of $\{1, 2, ..., n\}$ - a permutation is a one-to-one map $\sigma$ from the set $\{1, 2, ..., n\}$ to itself, and $\text{sign} \sigma$ is the number of pairs $(i, j)$ so that $i < j$ but $\sigma(i) > \sigma(j)$.

Exercise 20. * Prove that $|F^T| = |F|$, where $F$ is the transpose of $F$.

Exercise 21. * Take any row of $F$, say the $k$-th row, then we can compute $|F|$ by “expanding” this row:

$$|F| = f_{k,1}(-1)^{k+1}|F_{k,1}| + f_{k,2}(-1)^{k+2}|F_{k,2}| + ... + f_{k,n}(-1)^{k+n}|F_{k,n}|,$$

where $F_{i,j}$ is the $(n-1) \times (n-1)$ matrix obtained by deleting the $i$-th row and $j$-th column from $F$.

Exercise 22. * $F$ is invertible if and only if $|F| \neq 0$, in that case the $(i, j)$ entry of $F^{-1}$ is

$$(-1)^{i+j}|F_{j,i}|/|F|.$$

The proof of the above exercises can be found in linear algebra texts, so we omit their proofs.

Next, we take the opportunity to discuss a special features in dimension three. For two vectors

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

define their cross product to be

$$A \times B = \begin{vmatrix} i & a_1 & b_1 \\ j & a_2 & b_2 \\ k & a_3 & b_3 \end{vmatrix} = a_2 b_3 \begin{vmatrix} i & a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + a_1 b_2 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} + a_1 b_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$
Here the notation is,

\[(3.6)\]
\[
i = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad j = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},
\]
we formally “mix” them with the entries \(a_i, b_j\) in a big matrix. So

\[(3.7)\] \(A \times B = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}.

**Lemma 3.8.** We have:

\(i.\) \(A \times B = -(B \times A),\)

\(ii.\) \((A + kB) \times C = A \times C + kB \times C,\)

\(iii.\) \(|A \times B| = |A| \cdot |B| \cdot \sin \theta,\) where \(\theta\) is the angle between \(A, B.\)

\(iv.\) \((A \times B) \perp A,\) and \((A \times B) \perp B.\)

**Proof.** For iii.,

\[(3.9)\]
\[|A \times B|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2
\]
\[= a_2^2b_3^2 + a_3^2b_2^2 - 2a_2a_3b_2b_3 + a_3^2b_1^2 + a_1^2b_2^2 - 2a_1a_3b_1b_3 + a_1^2b_2^2 + a_2^2b_3^2 - 2a_1a_2b_1b_2
\]
\[= a_2^2b_3^2 + a_3^2b_2^2 + a_3^2b_1^2 + a_1^2b_2^2 + a_1^2b_1^2 + a_2^2b_2^2 + a_2^2b_1^2 + a_3^2b_3^2
\]
\[- 2a_1a_3b_1b_3 - 2a_2a_3b_2b_3 - 2a_1a_2b_1b_2 - a_2^2b_1^2 - a_2^2b_2^2 - a_3^2b_3^2
\]
\[= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2
\]
\[= |A|^2 \cdot |B|^2 - (A \cdot B)^2 = |A|^2 \cdot |B|^2 - (|A| \cdot |B| \cdot \cos \theta)^2
\]
\[= (|A| \cdot |B| \cdot \sin \theta)^2.
\]

For iv.,

\[(3.10)\] \((A \times B) \cdot A = (a_2b_3 - a_3b_2)a_1 + (a_3b_1 - a_1b_3)a_2 + (a_1b_2 - a_2b_1)a_3 = 0.
\]
So \((A \times B) \perp A.\) Similarly \((A \times B) \perp B.\)

**Exercise 23.** Furnish all details.

In particular, \(A \times B = 0\) if and only if \(A\) is parallel to \(B\) (note zero vector is regarded as being parallel to any vectors).

**Corollary 3.11.** The vectors \(A, B\) and \(A \times B\) satisfies the right hand rule.

**Proof.** (Sketch) this is a little technical. You can check that the definition of \(A \times B\) does not depend on the choice of coordinate system (of course you can only use “right handed” coordinate system - it is hard to explain what is “right handed”, but you can use the three axes for a “standard” coordinate system, and you call another coordinate system “right handed if the three axes of that one can be obtained from the “standard one” by a single continuous rotation). Then you can choose some coordinate under which \(A\) is
in the positive $x$-axis direction, $B$ is in the $x,y$ plane, then you verify the conclusion directly.

The unit normal vector field for a 2-surface is

$$N = \frac{\sigma_u \times \sigma_v}{|\sigma_u \times \sigma_v|}.$$  

For an $n$-surface $\Sigma$ in $\mathbb{R}^m$, if $m = n + 1$ (called the “codimension one” case), then we can still define a unit normal vector field; but the “cross product” becomes more complicated, involving minors of a certain matrix. So we will delay the expression of $N$ in higher dimensions to a later time.

If $m > n + 1$, there are infinitely many directions that are perpendicular to $\Sigma$: for example, if $n = 1$ and $m = 3$, i.e. a curve in $\mathbb{R}^3$, then at any point of the curve, you can rotate any normal vector to obtain more normal vectors (here normal vector simply means a vector that is perpendicular to the curve, not the $T_s^\prime$s/|$T_s^\prime$s we defined before).

Recall that if $(x_0, y_0, z_0)$ is a point on the surface, and $\sigma_u \times \sigma_v = [a, b, c]^T$ at $(x_0, y_0, z_0)$, then the tangent plane of the surface at $(x_0, y_0, z_0)$ is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0;$$

i.e. every vector $[x - x_0, y - y_0, z - z_0]^T$ should be perpendicular to $[a, b, c]^T$.

**Exercise 24.** Find tangent plane and normal vector of the surface $z = 3x + 2y$.

**Exercise 25.** Find tangent plane and normal vector of the surface $z = x^2 - y^2$.

**Exercise 26.** Find the tangent plane and normal vector of the sphere

$$\sigma(u,v) = \begin{bmatrix} R \cos u \cos v \\ R \cos u \sin v \\ R \sin u \end{bmatrix}.$$  

**Exercise 27.** Find tangent plane and normal vector of the surface

$$\sigma(u,v) = \begin{bmatrix} R \cosh u \cos v \\ R \cosh u \sin v \\ R \sinh u \end{bmatrix}.$$  

*Recall* $\cosh u = \frac{1}{2}(e^x + e^{-x})$ and $\sinh u = \frac{1}{2}(e^x - e^{-x})$.

In higher dimensions, the most convenient way to describe the tangent plane is to use the parametric equations. At $\sigma(a_1, ..., a_n)$, the tangent plane is

$$\begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_m \end{bmatrix} = \sigma(a_1, ..., a_n) + t_1\sigma_1(a_1, ..., a_n) + ... + t_n\sigma_n(a_1, ..., a_n).$$

We cannot write the tangent plane in a single equation like (3.13), unless in the codimension one case.
4. More linear algebra

And as in (1.11), we define for two vectors in \( \mathbb{R}^n \),

\[
a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},
\]

the dot product

\[
a \cdot b = a^T b = b^T a = a_1 b_1 + \ldots + a_n b_n;
\]

where after the first equal sign \( a, b \) are being viewed as \( n \times 1 \) matrices, and the multiplications are matrix multiplications. The following are clear:

1. \( a \cdot b = b \cdot a \).
2. \((ca) \cdot b = c(a \cdot b)\).
3. \((a + b) \cdot z = a \cdot z + b \cdot z\).
4. \( a \cdot a \geq 0. \ a \cdot a = 0 \) if and only if \( a = 0 \).

**Theorem 4.1** (Cauchy-Schwartz inequality). \( |x \cdot y| \leq |x| \cdot |y| \).

**Proof.** We can assume \( y \neq 0 \) - otherwise the inequality is obvious. For any \( t \),

\[
0 \leq |x - ty|^2 = (x - ty) \cdot (x - ty) = |x|^2 - 2(x \cdot y)t + |y|^2 t^2;
\]

rewrite this as

\[
|y|^2 t^2 - 2\frac{x \cdot y}{|y|} \cdot |y| + \left(\frac{x \cdot y}{|y|}\right)^2 - \left(\frac{x \cdot y}{|y|}\right)^2 + |x|^2 \geq 0,
\]

i.e.

\[
\left[|y|^2 t^2 - \left(\frac{x \cdot y}{|y|}\right)^2\right] - \left(\frac{x \cdot y}{|y|}\right)^2 + |x|^2 \geq 0.
\]

Since \( t \) is arbitrary,

\[
-\left(\frac{x \cdot y}{|y|}\right)^2 + |x|^2 \geq 0.
\]

\( \square \)

**Corollary 4.6** (Triangle inequality). \( |x + y| \leq |x| + |y| \).

**Proof.**

\[
|x + y|^2 = (x + y) \cdot (x + y) = |x|^2 + 2x \cdot y + |y|^2 \\
\leq |x|^2 + 2|x| \cdot |y| + |y|^2 = (|x| + |y|)^2.
\]

\( \square \)

Define the length of the vector \( a = (a_1, \ldots, a_n)^T \) to be

\[
|a| = \sqrt{a_1^2 + \ldots + a_n^2}.
\]
This is clear when \( n \leq 3 \). When \( n = 4 \), we can justify this if we agree \( e_4 = (0, 0, 0, 1)^T \) is perpendicular to \( e_1, e_2, e_3 \). Here “perpendicular” means we can use the Pythagorean theorem, i.e.

\[
\text{length}(a_1, a_2, a_3, a_4) = \sqrt{[\text{length}(a_1, a_2, a_3, 0)]^2 + [\text{length}(0, 0, 0, a_4)]^2};
\]

and we view \( \{(x_1, x_2, x_3, 0)\mid x_1, x_2, x_3 \in \mathbb{R}\} \) as an exact copy of \( \mathbb{R}^3 \), then

\[
\text{length}(a_1, a_2, a_3, 0) = \sqrt{a_1^2 + a_2^2 + a_3^2};
\]

view \( \{(0, 0, 0, x_4)\mid x_4 \in \mathbb{R}\} \) as an exact copy of \( \mathbb{R} \), then

\[
\text{length}(0, 0, 0, a_4) = |a_4|.
\]

We can argue similarly for \( n > 4 \).

Thus the distance from \( \mathbf{a} \) to \( \mathbf{b} \) is \( |\mathbf{a} - \mathbf{b}| \).

Then we have \( |\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a} \). For the moment if \( \mathbf{a} \cdot \mathbf{b} = 0 \), we say \( \mathbf{a}, \mathbf{b} \) are perpendicular; we will justify this later.

Now there is a wonderful thing: the group of orthogonal matrices \( O(n) \), which is the set of all \( n \times n \) matrices \( Q \) so that

\[
Q^T Q = I.
\]

Thus \( Q \) is invertible, \( Q^{-1} = Q^T \).

**Exercise 28.** Prove that \( O(n) \) is a group, i.e.

i. if \( P, Q \in O(n) \) then \( PQ \in O(n) \),

ii. \( P \in O(n) \) then \( P^T = P^{-1} \in O(n) \).

**Exercise 29.** Prove that if \( Q \in O(n) \) then \( |Q| = \pm 1 \).

**Exercise 30.** Prove that those \( Q \in O(n) \) with \( |Q| = 1 \) form a subgroup of \( O(n) \). It called the special orthogonal group, denoted by \( SO(n) \).

**Exercise 31.** Prove that if \( Q \in SO(2) \), then there is \( \theta \) so that

\[
Q = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}.
\]

Matrices in \( O(n) \) preserve the length of a vector:

\[
|Q\mathbf{a}|^2 = (Q\mathbf{a}) \cdot (Q\mathbf{a}) = (Q\mathbf{a})^T (Q\mathbf{a}) = \mathbf{a}^T Q^T Q \mathbf{a} = \mathbf{a}^T \mathbf{a} = |\mathbf{a}|^2.
\]

In particular, \( Q \) preserve the distance from \( \mathbf{u} \) to \( \mathbf{v} \):

\[
d(Q\mathbf{u}, Q\mathbf{v})^2 = |Q\mathbf{u} - Q\mathbf{v}|^2 = |Q(\mathbf{u} - \mathbf{v})|^2 = |\mathbf{u} - \mathbf{v}|^2 = d(\mathbf{u}, \mathbf{v})^2.
\]

Thus \( Q \) is an isometry (i.e. rigid motion) of \( \mathbb{R}^n \) that moves 0 to 0. In particular, this allows us to define the angle between two vectors \( \mathbf{w}_1, \mathbf{w}_2 \): we see that the plane

\[
P_{12} = \{(x_1, x_2, 0, 0, ..., 0) \in \mathbb{R}^n \mid x_1, x_2 \in \mathbb{R}\}
\]

is an exact copy of of the ordinary \( x, y \) plane \( \mathbb{R}^2 \), thus we can define angle between vectors in \( P_{12} \); so we just need to use some \( S \in O(n) \) to move \( \mathbf{w}_1, \mathbf{w}_2 \) into \( P_{12} \), then measure the
angle between $Sw_1, Sw_2$, and that would be the angle between two vectors $w_1, w_2$. We will see how this can be done.

Write
$$Q = [u_1, \ldots, u_n],$$
where $u_k$ is the $k$-th column of $Q$.

**Lemma 4.8.** $Q \in O(n)$ if and only if all $u_j$ are of length 1 and are mutually perpendicular.

**Proof.** We have
$$I = Q^TQ = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} [u_1, \ldots, u_n] = \begin{bmatrix} u_1^Tu_1, & u_1^Tu_2, & \ldots \\ u_2^Tu_1, & u_2^Tu_2, & \ldots \\ \vdots & \vdots & \ddots \\ u_n^Tu_1, & u_n^Tu_2, & \ldots \end{bmatrix},$$
this implies $u_i \cdot u_j = 1$ if $i = j$ and $u_i \cdot u_j = 1$ if $i \neq j$. □

Since $Q^{-1}$ exists, we know $u_1, \ldots, u_n$ is a basis of $\mathbb{R}^n$, we call such a basis an **orthonormal basis**. For a general basis, it is time consuming to write a vector $x$ as a linear combination of this basis. But for an orthonormal basis this is easy. Assume
(4.9) 
$$x = x_1u_1 + \ldots + u_n.$$  
Then for any $k$,
(4.10) 
$$u_k^Tx = x_1u_k^Tu_1 + \ldots + x_k u_k^Tu_k + \ldots + x_n u_k^Tu_n = x_k.$$  
So we readily get the coefficient
(4.11) 
$$x_k = u_k^Tx.$$

Given $k$ many vectors $w_1, \ldots, w_k$ in $\mathbb{R}^n$. Let
(4.12) 
$$u_1 = \frac{w_1}{|w_1|}.$$  
In particular, $u_1$ is a **unit vector**, i.e. its length is 1. Then, let
(4.13) 
$$u_2' = w_2 - (w_2 \cdot u_1)u_1.$$  
Assume $u_2' \neq 0$. Compute
(4.14) 
$$u_1 \cdot u_2' = u_1 \cdot w_2 - (w_2 \cdot u_1)|u_1|^2 = 0,$$
so $u_2' \perp u_1$. Let
(4.15) 
$$u_2 = \frac{u_2'}{|u_2'|}.$$  
In general, assume we already found mutually perpendicular, unit vectors $u_1, \ldots, u_j$, so that
(4.16) 
$$\text{span } (u_1, \ldots, u_j) = \text{span } (w_1, \ldots, w_j).$$  
Let
(4.17) 
$$u_{j+1}' = w_{j+1} - (w_{j+1} \cdot u_1)u_1 - (w_{j+1} \cdot u_2)u_2 - \ldots - (w_{j+1} \cdot u_j)u_j.$$
Assume \( u_{j+1}' \neq \mathbf{0} \). Compute, for any \( i \leq j \),

\[(4.18) \quad u_i \cdot u_{j+1}' = u_i \cdot w_{j+1} - [w_{j+1} \cdot u_i] u_i \cdot u_i - ... - [w_{j+1} \cdot u_n] u_i \cdot u_n.\]

So

\[(4.19) \quad u_i \cdot u_{j+1}' = u_i \cdot w_{j+1} - (w_{j+1} \cdot u_i) u_i \cdot u_i = u_i \cdot w_{j+1} - (w_{j+1} \cdot u_i) = 0.\]

So \( u_{j+1} \perp u_i \) for \( i = 1, 2, ..., j \). Let

\[(4.20) \quad u_{j+1} = \frac{u_{j+1}'}{|u_{j+1}'|}.\]

Continue like this. In conclusion, we can find \( k \) many unit vectors \( u_1, ..., u_k \) that are mutually perpendicular, and for each \( j \leq k \),

\[(4.21) \quad \text{span}(u_1, ..., u_j) = \text{span}(w_1, ..., w_j).\]

In particular, we found a “better” basis for \( V = \text{span}(w_1, ..., w_k) \). This procedure is called the Gram-Schmidt process.

**Exercise 32.** In the above process, if \( u_{j+1}' = 0 \), show that in this case \( w_{j+1} \) is dependent on \( w_1, ..., w_j \).

Thus we can throw away \( w_{j+1} \) and consider \( w_{j+2} \) instead... and (4.21) should be modified into

\[(4.22) \quad \text{span}(u_1, ..., u_j) \supset \text{span}(w_1, ..., w_j).\]

**Lemma 4.23 (“QR” decomposition).** Any real matrix \( A \) of size \( n \times k \), where \( k \leq n \), can be written as \( A = QR \), where \( Q \in O(n) \) and \( R \) is an \( n \times k \) matrix in which all \( i, j \) entries are 0 whenever \( i > j \). (upper triangular if \( k = n \)).

**Proof.** Let the columns of \( A \) be \( w_1, ..., w_k \). We can find vectors \( w_{k+1}, ..., w_{n'} \) so that

\[(4.24) \quad \text{span } \mathbb{R}^n \text{ (here we can make } n' = n \text{ if } w_1, ..., w_k \text{ are independent, and } n' > n \text{ otherwise). Do the Gram-Schmidt process and get } u_1, ..., u_n, \text{ so } Q = [u_1, ..., u_n] \text{ is an orthogonal matrix. Let } R = Q^{-1} A, \text{ so } A = QR. \text{ Let the } j-\text{th column of } R \text{ be } (r_{1j}, ..., r_{jj}, ..., r_{nj})^T, \text{ then the } j-\text{th column of } A \text{ is}

\[(4.25) \quad w_j = Q(r_{1j}, ..., r_{jj}, ..., r_{nj})^T = [u_1, ..., u_n](r_{1j}, ..., r_{jj}, ..., r_{nj})^T = r_{1j}u_1 + ... + r_{jj}u_j + r_{j+1j}u_{j+1} + r_{nj}u_n.\]

Since \( \text{span}(u_1, ..., u_j) = \text{span}(w_1, ..., w_j) \), \( w_j \) is a linear combination of \( u_1, ..., u_j \), in particular,

\[(4.26) \quad r_{j+1j} = ... = r_{nj} = 0.\]

This implies all the \((i, j)\) entries in \( R \) are 0 whenever \( i > j \). \( \square \)

Given \( w_1, w_2 \), by QR we can find an an orthonormal basis \( u_1, ..., u_n \) so that

\( \text{span } (w_1, w_2) \subset \text{span } (u_1, u_2), \)
and \( Q = [u_1, u_2, \ldots, u_n] \in O(n) \). Observe \( Qe_1 = u_1 \), \( Qe_2 = u_2 \), and we view \( \text{span} \ (e_1, e_2) \) as an exact copy of \( \mathbb{R}^2 \). We can assume \( w_1 = Qv_1 \), \( w_2 = Qv_2 \), where \( v_1, v_2 \in \text{span} \ (e_1, e_2) = \mathbb{R}^2 \). So we see
\[
Qe_1 = u_1, \quad Qe_2 = u_2,
\]
and we view \( \text{span} \ (e_1, e_2) \) as an exact copy of \( \mathbb{R}^2 \). We can assume
\[
w_1 \cdot w_2 = Qv_1 \cdot Qv_2 = v_1 \cdot v_2 = |v_1| \cdot |v_2| \cos \angle(v_1, v_2).
\]
Here \( \angle(v_1, v_2) \) makes sense since \( v_1, v_2 \in \mathbb{R}^2 \). Now define
\[
\angle(w_1, w_2) = \angle(v_1, v_2).
\]
Recall \( |w_i| = |Qv_i| = |v_i| \). This implies the important relation
\[
w_1 \cdot w_2 = |w_1| \cdot |w_2| \cos \angle(w_1, w_2).
\]
Thus the \( n \)-dimensional dot product works just like the 3-dimensional one.

**Exercise 33.** Show that the law of cosine works in the \( n \)-dimensional space.

5. LENGTH, ANGLE, AREA AND VOLUME

Assume we have an \( n \)-surface \( \Sigma \), with parametric equation \( \sigma(x_1, \ldots, x_n) \). Recall the derivative \(^3\) of \( \sigma \) is
\[
D\sigma = \begin{bmatrix}
\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\
\frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n}
\end{bmatrix} = [\sigma_1, \sigma_2, \ldots, \sigma_n];
\]
it is an \( m \times n \) matrix.

In local coordinate \( x_1, \ldots, x_n \), we can represent a curve \( \gamma \) on \( \Sigma \) by a curve on the parameter domain,
\[
x(t) = \begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_n(t)
\end{bmatrix},
\]
that is, \( \gamma \) has the parametric equation
\[
\gamma(t) = \sigma(x(t)) = \begin{bmatrix}
y_1(x_1(t), \ldots, x_n(t)) \\
y_2(x_1(t), \ldots, x_n(t)) \\
\vdots \\
y_m(x_1(t), \ldots, x_n(t))
\end{bmatrix}.
\]
Thus by the chain rule, we have
\[
\gamma'(t) = (D\sigma)x;
\]
\(^3\)According to Dieudonné, the derivative of a vector function of several variables has to be understood as a **linear transformation**, represented by a matrix. If you take this point of view, the chain rule would be much easier to understand. One should not view derivative merely as a collection of numbers (i.e. partial derivatives).
here $\dot{x}$ is derivative:

$$
\dot{x}(t) = \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ \vdots \\ x'_n(t) \end{bmatrix},
$$

and $(D\sigma)\dot{x}$ is matrix multiplication: $D\sigma$ is an $m \times n$ matrix, while $\dot{x}$ is an $n \times 1$ matrix.

Recall that we can view $\gamma'$ as velocity and $|\gamma'|$ as speed, integration of which gives arclength. Assume $a \leq t \leq b$, so we see the length of $\gamma$ is

$$
|\gamma| = \int_a^b |\gamma'(t)| dt = \int_a^b \sqrt{\gamma(t) \cdot \gamma'(t)} dt
$$

(5.3)

Notice

$[(D\sigma)\dot{x}] \cdot [(D\sigma)\dot{x}] = [(D\sigma)\dot{x}]^T [(D\sigma)\dot{x}] = \dot{x}^T (D\sigma)^T (D\sigma) \dot{x}$;

write

$$
g = (D\sigma)^T (D\sigma),
$$

we get $[(D\sigma)\dot{x}] \cdot [(D\sigma)\dot{x}] = \dot{x}^T g \dot{x}$; thus the arclength is

$$
|\gamma| = \int_a^b \sqrt{\dot{x}^T g \dot{x}} dt.
$$

(5.4)

Compute

$$
g = (D\sigma)^T D\sigma = \begin{bmatrix} \sigma_1^T \\ \sigma_2^T \\ \vdots \\ \sigma_n^T \end{bmatrix} \begin{bmatrix} \sigma_1, \sigma_2, \ldots, \sigma_n \end{bmatrix} = \begin{bmatrix} \sigma_1^T \sigma_1 & \sigma_1^T \sigma_2 & \cdots \\ \sigma_2^T \sigma_1 & \sigma_2^T \sigma_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix},
$$

so we have

$$
g = (D\sigma)^T D\sigma = \begin{bmatrix} \sigma_1 \cdot \sigma_1 & \sigma_1 \cdot \sigma_2 & \cdots \\ \sigma_2 \cdot \sigma_1 & \sigma_2 \cdot \sigma_2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.
$$

(5.5)

Here $\sigma_i \cdot \sigma_j$ means the dot product in dimension $n$. If we write

$$
g = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots \end{bmatrix},
$$

(5.6)

then

$$
g_{ij} = \sigma_i \cdot \sigma_j.
$$

In particular, $g$ is symmetric, i.e.

$$
g_{ij} = g_{ji}.$$
Lemma 5.7. \( g \) is positive definite, i.e. if \( \mathbf{a} \neq 0 \), then
\[
\mathbf{a}^T g \mathbf{a} > 0.
\]

Proof. Observe
\[
\mathbf{a}^T g \mathbf{a} = \mathbf{a}^T (D\sigma)^T (D\sigma) \mathbf{a} = [(D\sigma) \mathbf{a}]^T (D\sigma) \mathbf{a} = [(D\sigma) \mathbf{a}] \cdot [(D\sigma) \mathbf{a}] \geq 0.
\]
Notice the dot product in the last step. The above equals to 0 iff \((D\sigma) \mathbf{a} = 0\). Write out
\[
(D\sigma) \mathbf{a} = a_1 \sigma_1 + a_2 \sigma_2 + \ldots + a_n \sigma_n.
\]
Recall that we always assume \( \sigma_1, \sigma_2, \ldots, \sigma_n \) are linearly independent, thus \((D\sigma) \mathbf{a} = 0\) implies \(\mathbf{a} = 0\).

Exercise 34. Work out the arclength formula using coordinates:
\[
|\gamma| = \int_a^b \left( \sum_{i,j=1}^n g_{ij} x'_i(t) x'_j(t) \right)^{\frac{1}{2}} dt.
\]

Assume at the point \( \sigma(x_1, \ldots, x_n) \),
\[
A = a_1 \sigma_1 + \ldots + a_n \sigma_n, \quad B = a_1 \sigma_1 + \ldots + a_n \sigma_n;
\]
therefore, if we write
\[
\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},
\]
we have
\[
(5.8) \quad A = (D\sigma) \mathbf{a}, \quad B = (D\sigma) \mathbf{b}.
\]
First of all, we can compute length:
\[
|A|^2 = [(D\sigma) \mathbf{a}] \cdot [(D\sigma) \mathbf{a}] = [(D\sigma) \mathbf{a}]^T [(D\sigma) \mathbf{a}] = \mathbf{a}^T (D\sigma)^T (D\sigma) \mathbf{a} = \mathbf{a}^T g \mathbf{a};
\]
similarly \(|B|^2 = \mathbf{b}^T g \mathbf{b}\).

Next, compute
\[
A \cdot B = [(D\sigma) \mathbf{a}] \cdot [(D\sigma) \mathbf{b}] = [(D\sigma) \mathbf{a}]^T [(D\sigma) \mathbf{b}] = \mathbf{a}^T (D\sigma)^T (D\sigma) \mathbf{b} = \mathbf{a}^T g \mathbf{b}.
\]
Thus
\[
\angle(A, B) = \arccos \frac{A \cdot B}{|A||B|} = \arccos \frac{\mathbf{a}^T g \mathbf{b}}{\sqrt{\mathbf{a}^T g \mathbf{a}} \sqrt{\mathbf{b}^T g \mathbf{b}}}.
\]

Next, we compute volume (area is just “2-dimensional volume”) of the \( n \)-surface \( \Sigma \) in \( \mathbb{R}^m \), defined by the equation
\[
y_i = y_i(x_1, \ldots, x_n), \quad i = 1, 2, \ldots, m,
\]
with \((x_1, \ldots, x_n) \in \Omega \subset \mathbb{R}^n\). Now we cut \( \Omega \) into many small cubes (ignore irregularities near the boundary \( \partial \Omega \), or just assume \( \Omega \) is a big box). Each small cube \( S_k \) is indeed a parallelepiped generated by
\[
\mathbf{e}_1 \Delta x_1, \mathbf{e}_2 \Delta x_2, \ldots, \mathbf{e}_n \Delta x_n;
\]
where \( e_i \) be the vector with the \( i \)-th entry 1 and all other entries 0. \( \sigma \) maps this cube \( S_k \) to a small piece \( \sigma S_k \) in the surface \( \Sigma \); since \( \sigma \) is smooth, \( \sigma(S_k) \) is very close to the parallelepiped generated by
\[
D\sigma(e_1 \Delta x_1), \ D\sigma(e_2 \Delta x_2), \ldots, \ D\sigma(e_n \Delta x_n).
\]
Recall \( D\sigma(e_j) = \sigma_j \). Thus \( \sigma(S_k) \) is very close to the parallelepiped generated by
\[
\Delta x_1 \sigma_1, \ \Delta x_2 \sigma_2, \ldots, \ \Delta x_n \sigma_n.
\]
Its volume is
\[
\text{Vol}(\sigma(S_k)) \approx \text{Vol}(\Delta x_1 \sigma_1, \ \Delta x_2 \sigma_2, \ldots, \ \Delta x_n \sigma_n) = \text{Vol}(\sigma_1, \sigma_2, \ldots, \sigma_n) \Delta x_1 \Delta x_2 \ldots \Delta x_n.
\]
Apply the QR decomposition to \( D\sigma \), we get
\[
D\sigma = \begin{bmatrix} \sigma_1, \sigma_2, \ldots, \sigma_n \end{bmatrix} = QR.
\]
Here \( Q \in O(m) \). Let \( R = (r_1, \ldots, r_n) \), where \( r_j \) is the \( j \)-th column of \( R \). So \( Q r_j = \sigma_j \).

The advantage of using \( r_1, \ldots, r_n \) is that \( R \) is upper triangular; in particular, by the conditions of volume (3.1), (3.2), we see
\[
A(r_1, \ldots, r_n) = A \begin{bmatrix} r_{11} & 0 & \cdots & 0 \\ 0 & r_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{bmatrix} = r_{11} r_{22} \ldots r_{nn},
\]
where \( r_{ii} \) is the diagonal entries of \( R \), and \( A \) is the signed volume of parallelepiped, \( \text{Vol} = |A| \). Now observe that we can write
\[
R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix},
\]
where \( R_1 \) is an \( n \times n \) upper triangle matrix, with diagonal entries \( r_{11}, r_{22}, \ldots \). Thus we indeed have
\[
A(r_1, \ldots, r_n) = \det R_1.
\]
Note this can be negative. We then observe
\[
\det[(QR)^T (QR)] = \det R^T Q^T Q R = \det R^T I R = \det R^T R = \det [R_1^T, 0^T] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = \det(R_1^T R_1) = \det R_1^T \det R_1 = (\det R_1)^2.
\]
Thus we can compute
\[
\text{Vol}(r_1, \ldots, r_n) = |A(r_1, \ldots, r_n)| = |\det R_1| = \sqrt{(\det R_1)^2} = \sqrt{\det[(QR)^T (QR)]} = \sqrt{\det[(D\sigma)^T D\sigma]}.
\]
By (5.5), we have
\[ \text{Vol}(r_1, \ldots, r_n) = \sqrt{\det g}. \]
So we conclude
\[ \text{Vol}(\sigma(S_k)) \approx \sqrt{\det g} \Delta x_1 \ldots \Delta x_n. \]
So
\[ \text{Vol}(\Sigma) = \sum_k \text{Vol}(\sigma(S_k)) \approx \sum_k \sqrt{\det g} \Delta x_1 \ldots \Delta x_n; \]
by the usual Riemann sum argument, we have the volume formula
\[ \text{Vol}(\Sigma) = \int_\Omega \sqrt{\det g} \, dx_1 \ldots dx_n. \]

**Example 5.9.** We compute the volume (i.e. area) of the torus
\[ \sigma(\phi, \theta) = \begin{bmatrix} \cos \phi \\ \sin \phi \\ \cos \theta \\ \sin \theta \end{bmatrix}. \]
Here \(0 \leq \phi, \theta < 2\pi\). Start with
\[ D\sigma = [\sigma_\phi, \sigma_\theta] = \begin{bmatrix} -\sin \phi & 0 \\ \cos \phi & 0 \\ 0 & -\sin \theta \\ 0 & \cos \theta \end{bmatrix}. \]
So by (5.5),
\[ g = (D\sigma)^T D\sigma = \begin{bmatrix} -\sin \phi & 0 \\ \cos \phi & 0 \\ 0 & -\sin \theta \\ 0 & \cos \theta \end{bmatrix} \begin{bmatrix} -\sin \phi & 0 \\ \cos \phi & 0 \\ 0 & -\sin \theta \\ 0 & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \]
So
\[ \text{Vol}(\Sigma) = \int_0^{2\pi} \int_0^{2\pi} \sqrt{\det g} \, d\phi d\theta = \int_0^{2\pi} \int_0^{2\pi} 1 \, d\phi d\theta = 4\pi^2. \]

For 2-surfaces in \( \mathbb{R}^3 \), the tradition is to use some special notations, for example we use \( u, v \) instead of \( x_1, x_2 \) as coordinates. Also some special notations for the entries in the matrix \( g \): write
\[ (5.10) \quad E = \sigma_u \cdot \sigma_u, \quad F = \sigma_u \cdot \sigma_v, \quad G = \sigma_v \cdot \sigma_v. \]
So by (5.5),
\[ (5.11) \quad g = (D\sigma)^T D\sigma = \begin{bmatrix} \sigma_u^T \\ \sigma_v^T \end{bmatrix} [\sigma_u, \sigma_v] = \begin{bmatrix} E & F \\ F & G \end{bmatrix}. \]
In particular, the surface area formula is
\[ (5.12) \quad \text{Area}(\Sigma) = \int_\Omega \sqrt{\det g} \, dudv = \int_\Omega \sqrt{EG - F^2} \, dudv. \]
Example 5.13. We compute the volume (i.e. area) of the sphere
\[ \sigma(\phi, \theta) = \begin{bmatrix} \sin \phi \cos \theta \\ \sin \phi \sin \theta \\ \cos \phi \end{bmatrix}. \]

Here \(0 \leq \phi \leq \pi, 0 \leq \theta < 2\pi\). Start with
\[ D\sigma = [\sigma_\phi, \sigma_\theta] = \begin{bmatrix} \cos \phi \cos \theta & -\sin \phi \sin \theta \\ \cos \phi \sin \theta & \sin \phi \cos \theta \\ -\sin \phi & 0 \end{bmatrix}. \]

So
\[ g = (D\sigma)^T D\sigma = \begin{bmatrix} \cos \phi \cos \theta & \cos \phi \sin \theta & -\sin \phi \\ -\sin \phi \sin \theta & \sin \phi \cos \theta & 0 \\ \sin \phi & 0 & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \phi \cos \theta & -\sin \phi \sin \theta \\ \cos \phi \sin \theta & \sin \phi \cos \theta \\ -\sin \phi & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \sin^2 \phi \\ 0 & \sin^2 \phi & 0 \\ \sin^2 \phi & 0 & \cos^2 \phi \end{bmatrix}. \]

So
\[ \text{Vol}(\Sigma) = \int_0^{2\pi} \int_0^\pi \sqrt{\det g} \, d\phi d\theta = \int_0^{2\pi} \int_0^\pi \sin \phi \, d\phi d\theta = 4\pi. \]

Exercise 35. In your calculus textbook the area for 2-surfaces in \(\mathbb{R}^3\) is an integral involving cross product. Show that formula is equivalent to the one given here.

From the above basic computations, we see the \(n \times n\) matrix \(g\) is of FUNDAMENTAL IMPORTANCE in the study of geometry. This \(g\) is called the Riemannian metric in general. While \(g\) is a matrix, one should view \(g\) as a bilinear form.

Recall that a bilinear form on \(\mathbb{R}^n\) is a map \(f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\), so that for any \(A, B, C \in \mathbb{R}^n\) and any \(c \in \mathbb{R}\), we have
\begin{align*}
(5.14) \quad f(A + B, C) &= f(A, C) + f(B, C), \quad f(C, A + B) = f(C, A) + f(C, B), \\
(5.15) \quad f(cA, B) &= f(A, cB) = cf(A, B). 
\end{align*}

We say \(f\) is symmetric if \(f(A, B) = f(B, A)\) for any \(A, B \in \mathbb{R}^n\). Assume \(f\) is symmetric, we say \(f\) is positively definite if for any \(A \neq 0\), \(f(A, A) > 0\).

Exercise 36. Prove that the dot product is a positive definite bilinear form on \(\mathbb{R}^n\).

Lemma 5.16. Assume \(f\) is a bilinear form on \(\mathbb{R}^n\). Then there is an \(n \times n\) matrix \(P\) so that for any \(A, B \in \mathbb{R}^n\),
\[ f(A, B) = A^T PB. \]

Here all vectors are written as column vectors (i.e. \(n \times 1\) matrix).

\(f\) is symmetric if and only if \(P\) is a symmetric matrix.
Proof. For any \( i \) with \( 1 \leq i \leq n \), let \( e_i = [0, \ldots, 1, \ldots, 0]^T \), where the only nonzero entry is the \( i \)-th entry whose value is 1.

Then let \( P \) be the matrix whose \((i, j)\) entry is \( f(e_i, e_j) \).

Thus we can view \( g \) as a bilinear form: it maps \((a, b)\) to \( a^T g b \). For the moment \(^4\) we will use the notation

\[
\langle a, b \rangle = a^T g b.
\]

We can also use the notation \( g(a, b) \). Essentially \( g \) is just the dot product on the tangent plane of \( \Sigma \): recall that

\[
a^T g b = a^T (D\sigma)^T D\sigma b = [(D\sigma)a]^T D\sigma b = [(D\sigma)a] \cdot [D\sigma b],
\]

thus

\[
\langle a, b \rangle = [(D\sigma)a] \cdot [D\sigma b].
\]

However, \( g \) is not necessarily diagonal.

In dimension 2, in tradition \( g \) is also called the first fundamental form:

\[(5.18)\]

\[
g = \begin{bmatrix}
E(u,v) & F(u,v) \\
F(u,v) & G(u,v)
\end{bmatrix}.
\]

So \( E, F, G \) are all functions of \( u, v \). For any two vectors

\[(5.19)\]

\[
A = \begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}, \quad B = \begin{bmatrix}
b_1 \\
b_2
\end{bmatrix},
\]

we have

\[(5.20)\]

\[
\langle A, B \rangle = A^T \begin{bmatrix}
E & F \\
F & G
\end{bmatrix} B = [a_1, a_2] \begin{bmatrix}
E & F \\
F & G
\end{bmatrix} \begin{bmatrix}
b_1 \\
b_2
\end{bmatrix}
= E \cdot a_1 b_1 + F \cdot (a_1 b_2 + a_2 b_1) + G \cdot a_2 b_2.
\]

It is an old tradition also to write the first fundamental form as

\[(5.21)\]

\[
Edu^2 + 2Fdudv + Gdv^2.
\]

With \( g \) we can compute length, angle and area, so that we can ignore the fact that the surface \( \Sigma \) is in \( \mathbb{R}^3 \). In fact, in many occasions, it is a hindrance, not a help, to demand that our curved space (e.g. a surface) to be in some Euclidean space. For example on the surface of cylinder we still see the total angle of a (small) triangle is \( \pi \), ... and all theorems of plane geometry hold on the cylinder (assuming the configuration you study is not too big - e.g. nothing “wraps” the cylinder), so from an intrinsic point of view, we cannot distinguish a cylinder from a plane; \(^5\) but the equation for a cylinder certainly looks very different from the equation of a plane.

\(^4\)Later, we will see this way of using the \( \langle \cdot, \cdot \rangle \) is somehow misleading - if we use a different coordinate system, the expression of \( a, b \) will be different. In general one want to put in \( \langle \cdot, \cdot \rangle \) the abstract tangent vectors - not their local coordinates.

\(^5\)Of course, from an extrinsic point of view, i.e. if we are allowed to leave to cylinder and travel to “outside”, then the cylinder is very different from the plane.

A more subtle issues is, take the intrinsic point of view, from a global point of view a cylinder is still different from a plane: if you keep on traveling “straight” in certain direction in a cylinder, then after some time you will return to your starting point - this never happens on a plane.
Thus a general guideline in study intrinsic geometry, i.e. those geometric properties of \( \Sigma \) that can be studied by doing measurement without moving out of \( \Sigma \), is (for convenience stated in dimension 2):

**Everything should be computed in terms of the coordinate \( u, v \) and the metric \( g \), without explicitly mentioning the \( x, y, z \) as functions of \( u, v \).**

This requirement shows the fact intrinsic properties do not depend on particular embedding of the surface into \( \mathbb{R}^3 \), thus filter away all “superficial” information concerning the embedding. In fact, \( g \) gives us complete information on the intrinsic geometry.

As an example, we show that it is possible to recognize that a surface is curved, just from intrinsic measurement. For example, consider the 2-dimensional sphere \( S^2 \), say the surface of earth. Now you measure the area of the land within 100 meters from you, then you get basically \( \pi 100^2 \) square meters, ignoring error in measurement (the land it not exactly flat, etc.). This is what you get on a perfect plane. However, if you measure the area of the land within 1000 kilometers from you, you see the outcome deviates significantly from \( \pi 1000^2 \) square kilometers - your measurement is quite smaller that that - the unevenness of land, e.g. mountains, ... cannot explain the discrepancy. Then we have to conclude that the surface of earth is essentially different from a flat plane. All these can be done without leaving the earth, in principle possible for 2-dimensional beings. In contrast, in the extrinsic way, one proves that the earth is round by taking a trip in spaceship.

However, “extrinsic” geometry is also interesting on its own right. Historically, the most important intrinsic quantity - the Gauss curvature - was first defined in an extrinsic way. We will see this in the next section.

**Exercise 37.** Compute the area of a sphere of radius \( R \).

**Exercise 38.** Compute the first fundamental form of a sphere of radius \( R \). You can choose the \( x, y \) coordinate or the spherical polar coordinate.

**Exercise 39.** Consider the graph of a function \( y = f(x) \), where \( x \in [a, b] \). Rotate it around the \( x \)-axis we get a surface \( \Sigma \). Compute its first fundamental form, and the area of \( \Sigma \).

### 6. Second fundamental form

Define

\[
(6.1) \quad a_{11} = \frac{\partial^2 \sigma}{\partial u^2} \cdot N, \quad a_{12} = a_{21} = \frac{\partial^2 \sigma}{\partial u \partial v} \cdot N, \quad a_{22} = \frac{\partial^2 \sigma}{\partial v^2} \cdot N.
\]

Here we view \( \sigma \) as a vector valued function of two variables:

\[
(6.2) \quad \sigma(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}, \quad \frac{\partial^2 \sigma}{\partial u^2} = \begin{bmatrix} \frac{\partial^2 x(u, v)}{\partial u^2} \\ \frac{\partial^2 y(u, v)}{\partial u^2} \\ \frac{\partial^2 z(u, v)}{\partial u^2} \end{bmatrix}, \quad \frac{\partial^2 \sigma}{\partial u \partial v} = \ldots.
\]
The **second fundamental form**, written as $\mathcal{II}$, is a bilinear form on the two dimensional vector space $\mathbb{R}^2$, i.e. for any two vectors

\begin{equation}
C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix},
\end{equation}

define

\begin{equation}
\mathcal{II}(C, D) = C^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} D = [c_1, c_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = a_{11} \cdot c_1 d_1 + a_{12} \cdot (c_1 d_2 + c_2 d_1) + a_{22} \cdot c_2 d_2.
\end{equation}

Of course, if you choose another coordinate system, it is possible that the normal vector $N$ is the opposite of the one used above. Thus the “second fundamental form” we get from the new coordinate system could be $-\mathcal{II}$, where $\mathcal{II}$ is the second fundamental form we get from the new coordinate system. As we will see later, this ambiguity is not an issue for most purposes.

**Exercise 40.** Prove the following:

\begin{equation}
\mathcal{II}(W_1, W_2) = \mathcal{II}(W_2, W_1), \quad \mathcal{II}(cW_1, W_2) = c\mathcal{II}(W_1, W_2),
\end{equation}

\begin{equation}
\mathcal{II}(W_1 + W_2, W_3) = \mathcal{II}(W_1, W_3) + \mathcal{II}(W_2, W_3).
\end{equation}

**Exercise 41.** Compute the second fundamental form of a plane.

**Exercise 42.** Compute the second fundamental form of the standard sphere of radius $R$.

**Exercise 43.** Compute the second fundamental form of the surface $z = x^2 - y^2$.

**Exercise 44.** Compute the second fundamental form of the Enneper surface.

**Lemma 6.7.** Assume $W$ is a two dimensional vector at $(u_0, v_0)$ so that

\begin{equation}
\langle W, W \rangle = 1.
\end{equation}

Let $\alpha$ be a curve so that $\alpha(0) = (u_0, v_0)$, $\alpha'(0) = W$. Let $\gamma = \sigma \circ \alpha$. Then

\begin{equation}
\mathcal{II}(W, W) = \gamma''(0) \cdot N = \kappa n \cdot N,
\end{equation}

here $k$ is the curvature of $\gamma$ at $t = 0$, $n$ is the normal vector (if exists) of $\gamma$ at $t = 0$.

**Proof.** Let $s$ be the arclength parameter of $\gamma$,

\begin{equation}
\gamma' \cdot \gamma' = \langle \alpha', \alpha' \rangle = 1.
\end{equation}

So

\begin{equation}
\kappa n \cdot N = \gamma''(0) \cdot N = (D \sigma \cdot \alpha')' \cdot N.
\end{equation}

Recall

\begin{equation}
\sigma_u = \frac{\partial \sigma}{\partial u}, \quad \sigma_v = \frac{\partial \sigma}{\partial v}.
\end{equation}
Compute

\[ (D\sigma \cdot \alpha')' = \frac{d(D\sigma)}{ds} \alpha'(s) + D\sigma \cdot \alpha''(s) = \frac{d(\sigma_u, \sigma_v)}{ds} \alpha'(s) + D\sigma \cdot \alpha''(s) \]

\[ = \left( \frac{d\sigma_u}{ds}, \frac{d\sigma_v}{ds} \right) \alpha'(s) + D\sigma \cdot \alpha''(s) \]

\[ = \left( \frac{\partial^2 \sigma}{\partial u \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial u \partial v} \frac{\partial v}{\partial s} \right) \alpha'(s) + \left( \frac{\partial^2 \sigma}{\partial v \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial v \partial v} \frac{\partial v}{\partial s} \right) \alpha'(s) + \left( \frac{\partial \sigma}{\partial u}, \frac{\partial \sigma}{\partial v} \right) \alpha''(s), \]

where

\[ \alpha'(s) = \begin{bmatrix} u'(s) \\ v'(s) \end{bmatrix}, \quad \alpha''(s) = \begin{bmatrix} u''(s) \\ v''(s) \end{bmatrix}. \]

Notice

\[ \frac{\partial \sigma}{\partial u} \perp N, \quad \frac{\partial \sigma}{\partial v} \perp N, \]

therefore

\[ (D\sigma \cdot \alpha')' \cdot N = \left[ \left( \frac{\partial^2 \sigma}{\partial u \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial u \partial v} \frac{\partial v}{\partial s} \right) u'(s) + \left( \frac{\partial^2 \sigma}{\partial v \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial v \partial v} \frac{\partial v}{\partial s} \right) v'(s) \right] \cdot N. \]

Therefore

\[ (D\sigma \cdot \alpha')' \cdot N = \left( \frac{\partial^2 \sigma}{\partial u \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial u \partial v} \frac{\partial v}{\partial s} \right) u'(s) + \left( \frac{\partial^2 \sigma}{\partial v \partial u} \frac{\partial u}{\partial s} + \frac{\partial^2 \sigma}{\partial v \partial v} \frac{\partial v}{\partial s} \right) v'(s) \]

\[ = a_{11} \frac{\partial u}{\partial s} + a_{12} \frac{\partial v}{\partial s} + a_{21} \frac{\partial u}{\partial s} + a_{22} \frac{\partial v}{\partial s} = \Pi(\alpha', \alpha'). \]

Compute at \( s = 0 \) we see

\[ \kappa n \cdot N = \gamma''(0) \cdot N = (D\sigma \cdot \alpha')' \cdot N = \Pi(\alpha'(0), \alpha'(0)) = \Pi(W, W). \]

Assume \( X \) is a tangent vector of \( \Sigma \). Take some local coordinate we can find \( W \) so that \( X = \sigma_* W = D\sigma(W) \). Therefore we can define

\[ \Pi(X, X) = \Pi(W, W). \]

The above lemma implies that \( \Pi(X, X) \) does not depend on the coordinate system. In fact, we have the following

**Theorem 6.20.** Assume \( X \) is a tangent vector of \( \Sigma \) at \( p \in \Sigma \) with length \( |X| = 1 \). Let \( \gamma \) be any curve that lies in \( \Sigma \), with arclength parameter \( s \), so that

\[ \gamma(0) = p, \quad \gamma'(0) = X. \]

Then

\[ \Pi(X, X) = \kappa n \cdot N, \]

where \( k \) is the curvature of \( \gamma \) at \( p \), \( n \) is the normal vector of \( \gamma \), and \( \gamma'' = \kappa n \).
Proof. Take any local coordinate system. We can find a curve \( \alpha \) so that \( \sigma(\alpha) = \gamma \), then let \( W = \alpha'(0) \). Then

\[
X = \gamma'(0) = D\sigma \cdot \alpha'(0) = D\sigma(W).
\]

Observe

\[
\langle W, W \rangle = (D\sigma W) \cdot (D\sigma W) = X \cdot X = 1.
\]

the proof now follows as in the previous lemma.

Exercise 45. Let \( X \) be a unit tangent vector to a ruled surface so that \( X \) is in the line direction. Compute \( \Pi(X, X) \) using two methods: 1. the above theorem, and 2. direct computation.

In particular \( \Pi(X, X) \) does not depend on the choice of \( \gamma \). Therefore we also view the second fundamental form as a bilinear form on the tangent space of \( \Sigma \), i.e. we can compute \( \Pi(X, X) \) with \( X \) a tangent vector of \( \Sigma \).

Assume we use arclength parameter for the curve \( \gamma \) in the surface. Denote by \( T \) its tangent vector. So \( T \perp N \). So \( N, T, N \times T \) are unit vector that are mutually perpendicular to each other. Notice \( \kappa n \perp T \). So we can write

\[
\kappa n = \kappa_n N + \kappa_g N \times T,
\]

here

\[
\kappa_n = \langle \kappa n, N \rangle
\]

is called the normal curvature of \( \gamma \),

\[
\kappa_g = \langle \kappa n, N \times T \rangle
\]

is called the geodesic curvature of \( \gamma \). We are especially interested in the geodesic curvature. Of course the geodesic curvature is not quite well defined: if we use \(-N\) as the normal vector, we will get the value \(-\kappa_g\).

Lemma 6.28. Let \( \Sigma \) be a surface, \( p \) is a point in \( \Sigma \). Let \( P \) be the plane that passes through \( p \) and contains \( N \). Then \( P \) intersects \( \Sigma \) at a curve \( \beta \) (say near \( p \)).

Then

\[
\Pi(\beta'(0), \beta'(0)) = \pm \kappa,
\]

where \( \kappa \) is the curvature of \( \beta \) at \( p \).

Proof. \( \beta \) is a plane curve (stays in \( P \)) that passes through \( p \). For \( \beta \), the vector \( \kappa n \) is in the plane \( P \); since both \( \kappa n \) and \( N \) are in \( P \) and both are perpendicular to \( T = \beta'(0) \), we see

\[
n = \pm N.
\]

So

\[
\Pi(\beta'(0), \beta'(0)) = \kappa n \cdot N = \pm \kappa N \cdot N = \pm \kappa |N|^2 = \pm \kappa.
\]

\(\square\)
Of course this is true only for \( \beta \) constructed in this way. Now among all such \( \beta \), we want to find the maximum and minimum of \( k \). From the above discussion, this is equivalent to the following question:

**Among all 2-dimensional vectors \( A \) with \( \langle A, A \rangle = 1 \), find the maximum and minimum of \( \|A\| \).**

It is not hard to see that the set of vectors \( A \) with \( \langle A, A \rangle = 1 \) forms an ellipse. In fact, view \( g \) as a symmetric matrix,

\[
(6.32) \quad g = \begin{bmatrix} E & F \\ F & G \end{bmatrix}.
\]

By linear algebra, because \( g \) is a symmetric matrix, one can find an orthogonal matrix \( Q \) so that

\[
(6.33) \quad Q^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} Q = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix},
\]

where \( \mu_1, \mu_2 \) are the two eigenvalues of \( g \) - they are both real! In fact, since \( g \) is positive definite, \( \mu_1, \mu_2 \) are both positive. So we can use the new coordinate \( y = Q^T x \), we see the locus \( \langle A, A \rangle = 1 \) is the set of points \( y^T = (y_1, y_2) \) satisfying

\[
(6.34) \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 = 1,
\]

that is an ellipse.

So there is a parameter \( \theta \), so that there is a differentiable vector function \( A = A(\theta) \) so that \( \langle A(\theta), A(\theta) \rangle = 1 \). So the above problem becomes an **unconstrained** max/min problem for \( \Pi(A(\theta), A(\theta)) \) as a function of \( \theta \). View \( \Pi \) as a symmetric matrix,

\[
(6.35) \quad \Pi = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.
\]

So at maximum/minimum,

\[
(6.36) \quad 0 = \frac{d}{d\theta} \Pi(A, A) = 2\Pi(A_\theta, A) = 2(A_\theta)^T \Pi A.
\]

In the last two equal signs we view \( \Pi \) as a \( 2 \times 2 \) matrix. On the other hand, for every \( \theta \),

\[
(6.37) \quad 0 = \frac{d}{d\theta} \langle A, A \rangle = \frac{d}{d\theta} (A^T g A) = 2(A_\theta^T g A).
\]

Note we view \( g \) as a \( 2 \times 2 \) matrix. Now at maximum/minimum both \( \Pi A, gA \) are perpendicular (in \( \mathbb{R}^2 \)) to \( A_\theta \), so \( \Pi A \) is a multiple of \( gA \), so we can write

\[
(6.38) \quad \Pi A - \lambda gA = 0
\]

at maximum/minimum. Observe here we used the fact \( A \neq 0 \) and \( gA \neq 0 \) because \( g(A, A) = A^T g A = 1 \). Therefore

\[
(6.39) \quad (\Pi - \lambda g)A = 0.
\]

In order that this has a solution \( A \neq 0 \), it must be

\[
(6.40) \quad \det (\Pi - \lambda g) = 0;
\]
in this case, we get

\[(6.41) \quad 0 = A^T (\Pi - \lambda g) A = \Pi(A, A) - \lambda g(A, A) = \Pi(A, A) - \lambda; \]

so the solutions \( \lambda \) to \( \det(\Pi - \lambda g) = 0 \) are the maximum and minimum we are looking for.

Compute

\[
\begin{align*}
\det(\Pi - \lambda g) &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} - \lambda \begin{vmatrix} E & F \\ F & G \end{vmatrix} \\
&= \begin{vmatrix} a_{11} - E\lambda & a_{12} - F\lambda \\ a_{21} - E\lambda & a_{22} - G\lambda \end{vmatrix} \\
&= (a_{11} - E\lambda)(a_{22} - G\lambda) - (a_{12} - F\lambda)^2 \\
&= (EG - F^2)\lambda^2 + \ldots
\end{align*}
\]

recall that \( g \) is positive definite, so

\[(6.43) \quad |g| = EG - F^2 > 0.\]

So \( \det(\Pi - \lambda g) \) is a second order polynomial in \( \lambda \) - in particular it is not a first or zero order polynomial. Since \( EG - F^2 \neq 0 \), either \( E \neq 0 \) or \( G \neq 0 \) (or both). So we can choose some real \( \lambda' \) so that

\[(6.44) \quad (a_{11} - E\lambda')(a_{22} - G\lambda') = 0, \]

so

\[
\begin{align*}
\det(\Pi - \lambda' g) &= (a_{11} - E\lambda')(a_{22} - G\lambda') - (a_{12} - F\lambda')^2 \\
&= - (a_{12} - F\lambda')^2 \leq 0.
\end{align*}
\]

Recall that for a second order polynomial with positive leading coefficient (i.e. \( EG - F^2 \)), if it can take \( \leq 0 \) value then it has two (distinct or identical) real roots.

In conclusion, \( \det(\Pi - \lambda g) \) has two real roots \( \lambda_1, \lambda_2 \). Their product

\[(6.46) \quad K = \lambda_1 \lambda_2 \]

is called the Gauss curvature. Their sum

\[(6.47) \quad H = \lambda_1 + \lambda_2 \]

is called mean curvature.

If \( \lambda_1 \neq \lambda_2 \), they are called principal curvatures.

**Lemma 6.48.** The Gauss curvature is well defined, i.e. does not depend on the choice of \( N \).

**Proof.** A different choice of \( N \) changes \( \Pi \) into \(-\Pi\). If

\[(6.49) \quad \det (\Pi - \lambda g) = 0, \]

then

\[(6.50) \quad \det (-\Pi - (-\lambda) g) = 0, \]

the roots will be \(-\lambda_1, -\lambda_2\), then we calculate

\[(6.51) \quad K = (-\lambda_1)(-\lambda_2) = \lambda_1 \lambda_2 \]

as before. \( \Box \)
From the example of plane and cylinder we see the principal curvatures and mean curvatures are not intrinsic, i.e. they can not be decided by the first fundamental form alone.

**Lemma 6.52.** $K = \det II / \det g$.

**Proof.** Notice $II - \lambda g = (IIg^{-1} - \lambda I)g$. Solve
\begin{equation}
0 = \det (II - \lambda g) = \det [(IIg^{-1} - \lambda I)g] = \det (IIg^{-1} - \lambda I) \det g.
\end{equation}
Notice $\det g \neq 0$, we see $\lambda_1, \lambda_2$ are just the eigenvalues of $IIg^{-1}$. So
\begin{equation}
K = \lambda_1 \lambda_2 = \det(IIg^{-1}) = \det II \cdot \det(g^{-1}) = \det II / \det g.
\end{equation}

□

**Exercise 46.** What happens to the mean curvature $H$ if we use the normal vector $-N$?

**7. Theorema Egregium**

**Lemma 7.1.** Near any point on the surface, there is a local coordinate $(\hat{u}, \hat{v})$ under which the first fundamental form is diagonalized, i.e.
\begin{equation}
g = \begin{bmatrix} \hat{E} & 0 \\ 0 & \hat{G} \end{bmatrix}.
\end{equation}

**Proof.** Take any local coordinate system $(u, v)$. Take a 2-dimensional vector field, say
\begin{equation}
A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\end{equation}
Let
\begin{equation}
X = \begin{bmatrix} a \\ b \end{bmatrix}.
\end{equation}
be the vector field so that
\begin{equation}
g(A, X) = 0, \quad a > 0, \quad a^2 + b^2 = 1
\end{equation}
everywhere. That is,
\begin{equation}
0 = g(A, X) = [0, 1] \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = [F, G] \begin{bmatrix} a \\ b \end{bmatrix} = aF + bG.
\end{equation}

At each $(u, v)$, let $\hat{u}(u, v) = u$. Passing through $(u, v)$ there is a unique curve that is tangent to $X$ everywhere, this curve intersects the $v$-axis at $(0, c(u, v))$, define
\begin{equation}
\hat{v}(u, v) = c(u, v).
\end{equation}
Observe that the following directional derivative is 0:

\[(7.8) \quad D_X \hat{v} = 0.\]

That is,

\[(7.9) \quad \frac{\partial \hat{v}}{\partial u} a + \frac{\partial \hat{v}}{\partial v} b = 0.\]

And \(\hat{u} = u\). Compute

\[(7.10) \quad \begin{bmatrix} \frac{\partial u}{\partial \hat{u}} & \frac{\partial u}{\partial \hat{v}} \\ \frac{\partial v}{\partial \hat{u}} & \frac{\partial v}{\partial \hat{v}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{\partial \hat{v}}{\partial u} & \frac{\partial \hat{v}}{\partial v} \end{bmatrix}^{-1};\]

by the implicit mapping theorem,

\[(7.11) \quad \begin{bmatrix} \frac{\partial u}{\partial \hat{u}} & \frac{\partial u}{\partial \hat{v}} \\ \frac{\partial v}{\partial \hat{u}} & \frac{\partial v}{\partial \hat{v}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{u}}{\partial u} & \frac{\partial \hat{u}}{\partial v} \\ \frac{\partial \hat{v}}{\partial u} & \frac{\partial \hat{v}}{\partial v} \end{bmatrix}^{-1},\]

therefore

\[(7.12) \quad \begin{bmatrix} \frac{\partial u}{\partial \hat{u}} & \frac{\partial u}{\partial \hat{v}} \\ \frac{\partial v}{\partial \hat{u}} & \frac{\partial v}{\partial \hat{v}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{\partial \hat{v}}{\partial u} & \frac{\partial \hat{v}}{\partial v} \end{bmatrix}^{-1} = \frac{1}{\partial \hat{u} / \partial v} \begin{bmatrix} \frac{\partial \hat{v}}{\partial v} & 0 \\ -\frac{\partial \hat{v}}{\partial u} & 1 \end{bmatrix}.\]

We obviously have

\[(7.13) \quad \frac{\partial \hat{v}}{\partial u} (\frac{\partial \hat{v}}{\partial v}) + \frac{\partial \hat{v}}{\partial v} \left( -\frac{\partial \hat{v}}{\partial u} \right) = 0,\]

Combine with (7.9), we see both \([\frac{\partial \hat{v}}{\partial u}, -\frac{\partial \hat{v}}{\partial u}]^T\) and \([a, b]^T\) are perpendicular to \([\frac{\partial \hat{v}}{\partial u}, -\frac{\partial \hat{v}}{\partial u}]^T\), and this last vector is not 0 (why?), so we see \([\frac{\partial \hat{v}}{\partial u}, -\frac{\partial \hat{v}}{\partial u}]^T\) is a multiple of \([a, b]^T\), so for some \(\mu\) we have

\[(7.14) \quad \frac{\partial \hat{v}}{\partial v} = \mu a, \quad -\frac{\partial \hat{v}}{\partial u} = \mu b.\]

By (7.12), we have

\[(7.15) \quad \frac{\partial u}{\partial \hat{u}} = \mu a, \quad \frac{\partial v}{\partial \hat{u}} = \mu b.\]
We now show that \((\hat{u}, \hat{v})\) is the desired coordinate: compute

\[
\hat{F} = \sigma_u \cdot \sigma_v = \left( \frac{\partial \sigma}{\partial u} \frac{\partial u}{\partial \hat{u}} + \frac{\partial \sigma}{\partial v} \frac{\partial v}{\partial \hat{u}} \right) \cdot \left( \frac{\partial \sigma}{\partial u} \frac{\partial u}{\partial \hat{v}} + \frac{\partial \sigma}{\partial v} \frac{\partial v}{\partial \hat{v}} \right)
\]

\[
= E \frac{\partial u}{\partial \hat{u}} \frac{\partial u}{\partial \hat{v}} + F \frac{\partial u}{\partial \hat{v}} \frac{\partial u}{\partial \hat{u}} + F \frac{\partial u}{\partial \hat{v}} \frac{\partial u}{\partial \hat{v}} + G \frac{\partial v}{\partial \hat{u}} \frac{\partial v}{\partial \hat{v}}
\]

\[
= E \mu a \frac{\partial u}{\partial \hat{u}} \frac{\partial u}{\partial \hat{v}} + F \frac{\partial u}{\partial \hat{v}} \frac{\partial v}{\partial \hat{u}} + F \frac{\partial u}{\partial \hat{u}} \frac{\partial v}{\partial \hat{v}} + G \frac{\partial v}{\partial \hat{u}} \frac{\partial v}{\partial \hat{v}}
\]

\[
= (E \mu_a + F \mu_b) \frac{\partial u}{\partial \hat{v}} \frac{\partial v}{\partial \hat{u}}.
\]

By (7.12) \(\partial u/\partial \hat{v} = 0\); by (7.6) \(Fa + Gb = 0\). So \(\hat{F} = 0\). \(\Box\)

So from now on we will assume \(F = 0\) in the first fundamental form.

The following is the Theorema Egregium of Gauss:

**Theorem 7.17.** The Gauss curvature \(K\) is intrinsic.

**Proof.** Write

\[
\sigma_{uu} = \frac{\partial^2 \sigma}{\partial u^2} = \begin{bmatrix} \frac{\partial^2 x}{\partial u^2} & \frac{\partial^2 y}{\partial u^2} & \frac{\partial^2 z}{\partial u^2} \end{bmatrix},
\]

similarly \(\sigma_{uv}, \sigma_{vv}\).

Observe, by assuming \(F = 0\), the three vectors \(\sigma_u/|\sigma_u|, \sigma_v/|\sigma_v|\) and \(N\) are orthonormal, i.e. mutually perpendicular and all has length 1. Compute

\[
\text{det } \Pi = (\sigma_{uu} \cdot N)(\sigma_{vv} \cdot N) - (\sigma_{uv} \cdot N)^2
\]

\[
= \sigma_{uu} \cdot \sigma_{uv} - \left( \sigma_{uu} \cdot \frac{\sigma_u}{|\sigma_u|} \right) \left( \sigma_{uv} \cdot \frac{\sigma_u}{|\sigma_u|} \right) - \left( \sigma_{uu} \cdot \frac{\sigma_v}{|\sigma_v|} \right) \left( \sigma_{uv} \cdot \frac{\sigma_v}{|\sigma_v|} \right)
\]

\[
- \sigma_{uv} \cdot \sigma_{uv} + \left( \sigma_{uv} \cdot \frac{\sigma_u}{|\sigma_u|} \right) \left( \sigma_{uv} \cdot \frac{\sigma_u}{|\sigma_u|} \right) + \left( \sigma_{uv} \cdot \frac{\sigma_v}{|\sigma_v|} \right) \left( \sigma_{uv} \cdot \frac{\sigma_v}{|\sigma_v|} \right),
\]

this is essentially the Pythagorean theorem. So

\[
\text{det } \Pi = \sigma_{uu} \cdot \sigma_{vv} - \frac{1}{|\sigma_u|^2} (\sigma_{uu} \cdot \sigma_u)(\sigma_{uv} \cdot \sigma_u) - \frac{1}{|\sigma_v|^2} (\sigma_{uu} \cdot \sigma_v)(\sigma_{uv} \cdot \sigma_v)
\]

\[
- (\sigma_{uv} \cdot \sigma_{uv}) + \frac{1}{|\sigma_u|^2} (\sigma_{uv} \cdot \sigma_u)(\sigma_{uv} \cdot \sigma_u) + \frac{1}{|\sigma_v|^2} (\sigma_{uv} \cdot \sigma_v)(\sigma_{uv} \cdot \sigma_v).
\]

Notice

\[
|\sigma_u|^2 = E, \quad |\sigma_v|^2 = G,
\]
and
\begin{align*}
\sigma_{uu} \cdot \sigma_u &= \frac{1}{2} (\sigma_u \cdot \sigma_u)'_u = \frac{1}{2} E_u, \\
\sigma_{uu} \cdot \sigma_v &= (\sigma_u \cdot \sigma_v)'_u - (\sigma_u \cdot \sigma_{uv}) = (\sigma_u \cdot \sigma_v)'_u - \frac{1}{2} (\sigma_u \cdot \sigma_u)'_v = - \frac{1}{2} E_v, \\
\sigma_{uv} \cdot \sigma_u &= \frac{1}{2} (\sigma_u \cdot \sigma_u)'_v = \frac{1}{2} E_v, \\
\sigma_{uv} \cdot \sigma_v &= \frac{1}{2} (\sigma_v \cdot \sigma_v)'_u = \frac{1}{2} G_u, \\
\sigma_{vv} \cdot \sigma_u &= (\sigma_v \cdot \sigma_u)'_v - (\sigma_v \cdot \sigma_{uv}) = (\sigma_v \cdot \sigma_u)'_v - \frac{1}{2} (\sigma_v \cdot \sigma_v)'_u = - \frac{1}{2} G_u, \\
\sigma_{vv} \cdot \sigma_v &= \frac{1}{2} (\sigma_v \cdot \sigma_v)'_v = \frac{1}{2} G_v.
\end{align*}

(7.22)

So
\begin{align*}
\text{det } II &= \sigma_{uu} \cdot \sigma_{vv} + \frac{E_u G_u}{4E} + \frac{E_v G_v}{4G} - \sigma_{uv} \cdot \sigma_{uv} + \frac{E_u E_v}{4E} + \frac{G_u G_v}{4G}.
\end{align*}

Then
\begin{align*}
\sigma_{uu} \cdot \sigma_{vv} &= (\sigma_{uu} \cdot \sigma_{vv})'_v - \sigma_{uv} \cdot \sigma_{vv} \\
&= (\sigma_{uu} \cdot \sigma_{vv})'_v - (\sigma_{uv} \cdot \sigma_{uv})'_u - \sigma_{uv} \cdot \sigma_{uv} \\
&= - \frac{1}{2} E_{vv} - \frac{1}{2} G_{uu} + \sigma_{uv} \cdot \sigma_{uv}.
\end{align*}

(7.24)

Plug this in,
\begin{align*}
\text{det } II &= - \frac{1}{2} E_{vv} - \frac{1}{2} G_{uu} + \frac{E_u G_u}{4E} + \frac{E_v G_v}{4G} + \frac{E_u E_v}{4E} + \frac{G_u G_v}{4G}.
\end{align*}

(7.25)

Since \( \text{det } g = EG \), we get
\begin{align*}
K &= \frac{1}{EG} \left( - \frac{1}{2} E_{vv} - \frac{1}{2} G_{uu} + \frac{E_u G_u}{4E} + \frac{E_v G_v}{4G} + \frac{E_u E_v}{4E} + \frac{G_u G_v}{4G} \right).
\end{align*}

(7.26)

So \( K \) can be written in term of the first fundamental form.

\textbf{Corollary 7.27.}

\begin{align*}
K &= - \frac{1}{\sqrt{EG}} \left[ \left( \frac{\sqrt{E}}{\sqrt{G}} \right)_v + \left( \frac{\sqrt{G}}{\sqrt{E}} \right)_u \right].
\end{align*}

(7.28)

\textit{Proof.} Direct computation, using (7.26).

\textbf{Exercise 47.} Compute the Gauss curvature for the sphere with radius \( R \).

\textbf{Exercise 48.} Compute the Gauss curvature for the plane.

\textbf{Exercise 49.} Assume the first fundamental form is
\begin{align*}
\frac{du^2 + dv^2}{v^2}.
\end{align*}

This is called the upper space model of hyperbolic plane.

Compute the Gauss curvature.
How do you visualize this surface?

**Exercise 50.** Assume the first fundamental form is

\[(7.30) \quad 4 \frac{du^2 + dv^2}{(1 - u^2 - v^2)^2}. \]

This is called the disc model of hyperbolic plane.

Compute the Gauss curvature.

How do you visualize this surface?

**Exercise 51.** Compute the Gauss curvature for the surface \(z = x^2 - y^2\).

**Exercise 52.** Compute the Gauss curvature for the surface \(x^2 + y^2 - z^2 = 1\).

**Exercise 53.** Compute the Gauss curvature for the surface \(x^2 + y^2 - z^2 = 0\).

**Exercise 54.** Compute the Gauss curvature for the surface \(x^2 + y^2 - z^2 = -1\).

**Exercise 55.** Assume \(\Sigma\) is a surface in \(\mathbb{R}^3\). Assume for \(p \in \Sigma\), there is a straight line passing through \(p\) that is entirely contained in \(\Sigma\). Prove that the Gauss curvature of \(\Sigma\) at \(p\) is \(\leq 0\).

Can you conclude this curvature must be 0?

**Exercise 56.** Assume the first fundamental form is

\[(7.31) \quad du^2 + f(u)^2 dv^2. \]

How do you visualize this surface?

Compute the Gauss curvature.

**Exercise 57.** In the above exercise, find some function \(f\) so that the Gauss curvature is constant 1. Can you be convinced that you get a sphere?

In the above exercise, find some function \(f\) so that the Gauss curvature is constant 0. Can you be convinced that you get a plane?

In the above exercise, find some function \(f\) so that the Gauss curvature is constant \(-1\).

**Exercise 58.** Rotate the graph of \(y = f(x)\) around the \(x\)-axis and get a surface. Compute its Gauss curvature.

**Exercise 59.** i. Compute the curvature of the curve \(y = f(x)\).

ii. Assume \(f(0) = 0, f'(0) = 0\), and we have

\[(7.32) \quad f(x) \geq 1 - \sqrt{1 - x^2} \]
when \(|x|\) is small. Prove that the curvature for \(y = f(x)\) at \(x = 0\) is \(\geq 1\). What is the geometric meaning of this result?

**Exercise 60.** Assume \(\Sigma\) is a compact (i.e. bounded, closed as a set, without boundary) surface in \(\mathbb{R}^3\).

Prove that there exists at least one point \(p \in \Sigma\) so that

\[(7.33) \quad K(p) > 0. \]
Exercise 61. Given a curve \( \gamma(t) = (x(t), y(t), z(t)) \) and \( R > 0 \), define
\[
(7.34) \quad R\gamma(t) = (Rx(t), Ry(t), Rz(t)).
\]

Given a surface \( \sigma(u, v) = (x(u, v), y(u, v), z(u, v)) \) and \( R > 0 \), define
\[
(7.35) \quad R\sigma(u, v) = (Rx(u, v), Ry(u, v), Rz(u, v)).
\]

The procedure of changing \( \gamma \) (or \( \sigma \)) into \( R\gamma \) (or \( R\sigma \)) is called rescale. It is a very important notion in geometry.

Discuss what happens to curvature (for curves), first and second fundamental forms, mean curvature and Gauss curvature (for surfaces), after doing a rescale.

8. Triangle Gauss-Bonnet theorem

Theorem 8.1. For a curve on a surface, its geodesic curvature \( \kappa_g \) is intrinsic.

Proof. As before, take a coordinate system with \( F = 0 \).

Assume \( \gamma = \sigma(u(s), v(s)) \) is a curve with arclength parameter \( s \). So \( T = \sigma_u u' + \sigma_v v' \);
so
\[
(8.2) \quad 1 = T \cdot T = E \cdot (u')^2 + G \cdot (v')^2.
\]

Compute
\[
(8.3) \quad \kappa_g = \gamma'' \cdot (N \times T) = (\sigma_u u' + \sigma_v v')' \cdot [N \times (\sigma_u u' + \sigma_v v')].
\]

Since \( \sigma_u \perp \sigma_v \), we see \( \sigma_u, \sigma_v, N = \sigma_u \times \sigma_v/|\sigma_u \times \sigma_v| \) are mutually perpendicular, and right handed. So
\[
(8.4) \quad \kappa_g = (\sigma_{uu}(u')^2 + \sigma_{vv}(v')^2 + 2\sigma_{uv}u'v' + \sigma_u u'' + \sigma_v v'') \cdot \left( \frac{\sigma_u}{|\sigma_v|} u' - \frac{\sigma_v}{|\sigma_u|} v' \right)
\
- (\sigma_{uu} \cdot \sigma_v) (u')^3 \frac{|\sigma_u|}{|\sigma_v|} - (\sigma_{uv} \cdot \sigma_u) (u')^2 v' \frac{|\sigma_v|}{|\sigma_u|} + 2(\sigma_{uv} \cdot \sigma_v) (v')^2 v' \frac{|\sigma_u|}{|\sigma_v|}
\
- 2(\sigma_{uv} \cdot \sigma_u) u'(v')^2 \frac{|\sigma_v|}{|\sigma_u|} + (\sigma_{vv} \cdot \sigma_v) (v')^2 u' \frac{|\sigma_u|}{|\sigma_v|} - (\sigma_{vv} \cdot \sigma_u) (v')^3 \frac{|\sigma_v|}{|\sigma_u|}
\
- E \frac{|\sigma_u|}{|\sigma_v|} u'' + G \frac{|\sigma_u|}{|\sigma_v|} v'' + u'.
\]
Apply (7.22), we have

\(\kappa_g = \frac{E_v u'^2 E}{2 \sqrt{EG}} - \frac{E_u (u')^2 v' \sqrt{G}}{2 \sqrt{E}} + G_u (u')^2 v' \sqrt{G} - E_v v'^2 u' \sqrt{G} \)

\(\frac{G_v (v')^2}{\sqrt{G}} + \frac{G_u (v')^3}{\sqrt{E}} \sqrt{G} - \sqrt{EG} u'' v' + \sqrt{EG} v'' u' \)

\(\frac{E_v u'}{2 \sqrt{EG}} + (v')^2 u' \left(\frac{G_v \sqrt{E}}{2 \sqrt{EG}} - \frac{E_u \sqrt{G}}{2 \sqrt{E}}\right) \)

\(- \sqrt{EG} u'' v' + \sqrt{EG} v'' u'.\)

We used (8.2) in the second equal sign. So

\(\kappa_g = - \frac{(\sqrt{E})_v}{\sqrt{G}} u' + \frac{(\sqrt{G})_u}{\sqrt{E}} v' - G \cdot (v')^2 \left(\frac{\sqrt{E} u'}{\sqrt{G} v'}\right)'\).

You can check that the last term is indeed correct. \(\square\)

**Corollary 8.7.**

\(\kappa_g = - \frac{(\sqrt{E})_v}{\sqrt{G}} u' + \frac{(\sqrt{G})_u}{\sqrt{E}} v' + E \cdot (u')^2 \left(\frac{\sqrt{G} v'}{\sqrt{E} u'}\right)'.\)

**Proof.** The last term in the \(\kappa_g\) formula in the previous theorem is

\(G \cdot (v')^2 \left(\frac{\sqrt{E} u'}{\sqrt{G} v'}\right)' = G \cdot (v')^2 \left[\left(\frac{\sqrt{G} v'}{\sqrt{E} u'}\right)^{-1}\right]'\)

\(= - G \cdot (v')^2 \left(\frac{\sqrt{G} v'}{\sqrt{E} u'}\right)^{-2} \left(\frac{\sqrt{G} v'}{\sqrt{E} u'}\right)' = - E \cdot (u')^2 \left(\frac{\sqrt{G} v'}{\sqrt{E} u'}\right)'.\) \(\square\)

Assume \(\Omega\) is a (curved) **triangle** on the surface \(\Sigma\), i.e. \(\Omega\) is a region on \(\Sigma\) that is **homeomorphic** to a triangle, and the boundary \(\partial\Omega\) of \(\Omega\) is the union of three smooth segments of curves. Assume \(\Omega = \sigma(D)\), so \(D\) is a (curved) triangle, with boundary \(\partial D\), in the \(u, v\) plane.

**Theorem 8.10** (Triangle Gauss-Bonnet theorem).

\(\sum_{i=1}^{3} (\pi - \phi_i) + \int_{\Omega} K + \int_{\partial \Omega} \kappa_g = 2\pi.\)

Here \(\phi_1, \phi_2, \phi_3\) are the three angles of the triangle \(\Omega\). And

\(\kappa_g = \kappa_n \cdot (N \times T)\)
is the geodesic curvature on the sides, with the convention that $N \times T$ is pointing inward to the triangle $\Omega$.

**Proof.** As before assume $F = 0$.

$$
\int_{\partial \Omega} K = \int_{D} -\frac{1}{\sqrt{EG}} \left[ \left( \frac{\sqrt{E}}{\sqrt{G}} \right)_v + \left( \frac{\sqrt{G}}{\sqrt{E}} \right)_u \right] \sqrt{EG} \, du \, dv
$$

(8.13)

$$
= \int_{\partial D} \frac{(\sqrt{E})_v}{\sqrt{G}} du - \frac{(\sqrt{G})_u}{\sqrt{E}} dv.
$$

Notice in the second equal sign we applied Green’s formula. Observed that we used the assumption on the orientation, i.e. looking from $N$, we see the boundary $\partial \Omega$ goes counterclockwise.

By Corollary 8.7,

$$
\int_{\partial \Omega} \kappa_g = \int_{\partial \Omega} \left( \frac{\sqrt{E}}{\sqrt{G}} \right)_v \cdot u' + \left( \frac{\sqrt{G}}{\sqrt{E}} \right)_u \cdot v' + E \cdot (u')^2 \left( \frac{\sqrt{G}v'}{E^{1/2}} \right)_s
$$

(8.14)

In order to compute this integral, recall we use the arclength parameter $s$ on the boundary $\partial D$. So write the boundary as $\gamma(s) = \sigma(u(s), v(s))$, where $s$ runs on $[0, L]$ - i.e. $L$ is the length of $\partial \Omega$,

$$
\int_{\partial \Omega} \kappa_g = \int_{0}^{L} \left[ -\frac{(\sqrt{E})_v}{\sqrt{G}} u' + \frac{(\sqrt{G})_u}{\sqrt{E}} v' + E \cdot (u')^2 \left( \frac{\sqrt{G}v'}{E^{1/2}} \right)_s \right] ds
$$

(8.15)

$$
= \int_{\partial D} \frac{(\sqrt{E})_v}{\sqrt{G}} du + \frac{(\sqrt{G})_u}{\sqrt{E}} dv + \int_{0}^{L} E \cdot (u')^2 \left( \frac{\sqrt{G}v'}{E^{1/2}} \right)_s ds.
$$

So

$$
\int_{\Omega} K = -\int_{\partial \Omega} \kappa_g + \int_{0}^{L} E \cdot (u')^2 \left( \frac{\sqrt{G}v'}{E^{1/2}} \right)_s ds.
$$

(8.16)

Consider one of the three pieces of $\partial \Omega$. Remember the tangent vector of $\partial \Omega$ is

$$
T = \sigma_u u'(s) + \sigma_v v'(s).
$$

Therefore (because $F = \langle \sigma_u, \sigma_v \rangle = 0$),

$$
\cos \angle(\sigma_u, T) = \frac{\langle T, \sigma_u \rangle}{|T| \cdot |\sigma_u|} = \frac{E u'}{\sqrt{E}} = \sqrt{E} u',
$$

(8.18)

$$
\cos \angle(\sigma_v, T) = \frac{\langle T, \sigma_v \rangle}{|T| \cdot |\sigma_v|} = \frac{G v'}{\sqrt{G}} = \sqrt{G} v'.
$$

(8.19)

We get

$$
E \cdot (u')^2 \left( \frac{\sqrt{G}v'}{E^{1/2}} \right)_s = (\cos \angle(\sigma_u, T))^2 \left( \frac{\cos \angle(\sigma_u, T)}{\cos \angle(\sigma_u, T)} \right)_s
$$

(8.20)

$$
= (\cos \angle(\sigma_u, T))^2 \left( \tan \angle(\sigma_u, T) \right)_s = (\angle(\sigma_u, T))_s.
$$

Here $\angle(\sigma_u, T)$ is defined in the following way. Since $F = 0$, we have $\sigma_u \perp \sigma_v$. So there is a direction, in which we can reach $\sigma_v$ by rotating $\sigma_u$ for an angle of $\pi/2$; then if we
rotate from $\sigma_u$ by angle $\angle(\sigma_u, T)$ in this direction we reach $T$. Notice that $\angle(\sigma_u, T)$ is still not well defined - it is defined up to adding an integer multiple of $2\pi$. On the other hand, $\cos \angle(\sigma_u, T)$ and $\angle(\sigma_u, T)'_s$ are well defined.

The orientation plays the following role. With the choice of orientation in $\partial \Omega$, e.g. in the following picture $N$ is pointing away from us and $\partial \Omega$ goes “clockwise”.

Now the tangent vector $T$ of $\partial \Omega$ is not continuous. We can smooth the three vertices in the triangle $\Omega$, so that we get a domain $\omega_1$ with smooth boundary $\partial \Omega_1$, that can be as close to $\Omega$ as we want; in fact, $\Omega_1$ differ from $\Omega$ only in the three very small discs around the three vertices of $\Omega$.

Let $T_1$ be the tangent vector (field) of $\partial \Omega_1$, the orientation is the same as $\partial \Omega$. Let $L_1$ be the length of $\partial \Omega_1$ and use the arclength parameter on $\partial \Omega$. Recall that $[\angle(\sigma_u, T_1)]'_s$ is well defined. We call the integral

$$
(8.21) \quad \frac{1}{2\pi} \int_0^{L_1} [\angle(\sigma_u, T_1)]'_s ds
$$

the rotating number of $\partial \Omega_1$. We can imaging that we travel along $\partial \Omega_1$, sitting on a “car” that always pointing to the $\sigma_u$ direction. When we finish a full round, $T_1$ returns to the original position (but perhaps after some rounds of rotation). Therefore we conclude that the total change in $\angle(\sigma_u, T_1)$ is an integer multiple of $2\pi$. In particular the rotating number is an integer.

We can deform $\partial \Omega_1$ smoothly to a small disc $C$ inside $\Omega$, while keeping the choice of orientation continuous. Observe that the rotating number depends continuously on loops,
and it takes only integer value, we see

\[(8.22) \text{ rotating number } (\partial \Omega_1) = \text{ rotating number } (C) \]

Now clearly the rotating number of \(C\) is 1.

Now remember

\[(8.23) \quad \int_0^L E \cdot (u')^2 \left( \frac{\sqrt{G'}v'}{\sqrt{E'u'}} \right)' \frac{ds}{s} = \int_0^L \left( \angle (\sigma_u, T) \right)' \frac{ds}{s}. \]

Observe near the vertices of \(\partial \Omega\), \(\partial \Omega_1\) makes a very sharp turn and quickly increases \(\angle (\sigma_u, T_1)\) by \(\approx \pi - \phi\). Let \(\omega_1 \to \Omega\), we get

\[(8.24) \quad \int_0^L \left( \angle (\sigma_u, T) \right)' ds + (\pi - \phi_1) + (\pi - \phi_2) + (\pi - \phi_3) = 2\pi \cdot \text{rotating number } (\partial \Omega_1) = 2\pi. \]

The conclusion follows. \(\square\)

**Exercise 62.** A geodesic triangle is a triangle all whose three sides are geodesics, i.e. shortest path connecting the vertices.\(^6\)

On the standard sphere of radius \(R\), find a relation between the total angle of a geodesic triangle and the area it bounds. Prove this relation by Gauss-Bonnet. Also prove this relation using elementary geometry.

From an intrinsic point of view, we have to view geodesics as “straight”. Thus on a curved space, we no longer have 180\(^\circ\) as the total angle of a triangle.

**9. Global Gauss-Bonnet Theorem**

It is known that any closed surface \(\Sigma\) has a triangulation, i.e. \(\Sigma\) is the union of finitely many triangles, and any edge is the common boundary of two triangles. Let \(\mathcal{F}\) be the number of faces, \(\mathcal{E}\) the number of edges, \(\mathcal{V}\) the number of vertices. The Euler characteristic of \(\Sigma\) is defined to be

\[(9.1) \quad \chi(\Sigma) = \mathcal{V} - \mathcal{E} + \mathcal{F}. \]

For example, we can view tetrahedron as a triangulation of the sphere \(S^2\); it has 4 faces, 4 vertices and 6 edges; therefore

\[(9.2) \quad \chi(S^2) = 4 - 6 + 4 = 2. \]

It is known that \(\chi\) is well defined - it does not depend on particular triangulation of \(\Sigma\).

**Theorem 9.3** (Global Gauss-Bonnet theorem). For a closed surface \(\Sigma\),

\[(9.4) \quad \frac{1}{2\pi} \int_{\Sigma} K = \chi(\Sigma). \]

\(^6\)In the future we shall relax the definition, allowing just three geodesics \(AB, BC\) and \(CA\), while the three sides \(AB, BC, CA\) may not bound a “triangle”. For example on the torus it is easy to find such “triangles”. We can even allow one side to be not distance minimizing.
Proof. Assume $\Sigma$ is the union of $\Omega_1, \ldots, \Omega_F$. Assume $\Sigma$ is oriented, i.e. there is a consistent way of assigning “counterclockwise” to each $\Omega_j$: i.e. if two triangles $\Omega_{i_1}, \Omega_{i_2}$ have a common side $L$, then travel counterclockwise along $\partial \Omega_{i_1}, \partial \Omega_{i_2}$ will yield opposite directions on $L$.

Apply triangle Gauss-Bonnet to each $\Omega_j$,

\begin{equation}
\sum_{j=1}^{F} (\pi - \phi^{(j)}_i) + \int_{\Omega_j} K + \int_{\partial \Omega_j} \kappa_g = 2\pi.
\end{equation}

Here $\phi^{(j)}_1, \phi^{(j)}_2, \phi^{(j)}_3$ are the three angles of the triangle $\Omega_j$. Add all these equations, we get

\begin{equation}
\sum_{i=1}^{3} \sum_{j=1}^{F} (\pi - \phi^{(j)}_i) + \int_{\Sigma} K = 2F\pi.
\end{equation}

Here notice all geodesic curvature terms canceled in pairs, because $\Sigma$ is orientable. In fact, $\kappa_g = \langle \kappa n, N \times T \rangle$, changing the orientation of $\partial \Omega$ is just replace the arclength parameter by $\hat{s} = -s$, then

\begin{equation}
\frac{d\gamma}{d\hat{s}} = \frac{d\gamma}{ds} \frac{ds}{d\hat{s}} = -\frac{d\gamma}{ds} = -T, \quad \hat{k}n = \frac{d\hat{T}}{d\hat{s}} = -\frac{dT}{ds} \frac{ds}{d\hat{s}} = -\kappa n(-1) = \kappa n;
\end{equation}

so

\begin{equation}
\kappa_g = \hat{k} n \cdot (N \times \hat{T}) = \kappa n \cdot (-N \times T) = -\kappa_g.
\end{equation}

Notice

\begin{equation}
\sum_{j=1}^{F} \sum_{i=1}^{3} (\pi - \phi^{(j)}_i) = \sum_{j=1}^{F} \sum_{i=1}^{3} \pi - \sum_{j=1}^{F} \sum_{i=1}^{3} \phi^{(j)}_i = 3F\pi - 2\pi V.
\end{equation}

So

\begin{equation}
\int_{\Sigma} K = 2F\pi - 3F\pi + 2\pi V.
\end{equation}

Finally, notice each triangle has three edges, each edge is the edge of exactly two triangles, so

\begin{equation}
3F = 2E.
\end{equation}

In the nonorientable case, we just need to observe that although there is no globally defined $N$, the orientation assumption in the triangle Gauss-Bonnet demands that $N \times T$ is pointing inward to a triangle, so if we change $N$ into $-N$, $T$ must be changed into $-T$. As we have seen this does not change $\kappa_n$. And $\kappa_g$ is changed into

\begin{equation}
\kappa n \cdot [(N \times T)] = \kappa n \cdot (-N \times T) = \kappa_g.
\end{equation}

So we can pick any $N$ on a triangle, and this does change the boundary integral, so as before we see neighboring boundary integrals cancel each other. Thus the proof goes the same way. \qed

Consider two compact surfaces (without boundary) $\Sigma_1, \Sigma_2$, remove a disc (i.e. a subset in the surface that is homeomorphic to a disc) from each surface, we get two surfaces $\Sigma'_1, \Sigma'_2$ each with a boundary piece that is a circle (i.e. a subset in the surface that is
homeomorphic to a circle). Now glue $\Sigma'_1, \Sigma'_2$ together along the boundary we get a new surface, $\Sigma_1 \# \Sigma_2$, which is called the connected sum of $\Sigma_1, \Sigma_2$.

**Exercise 63.** Convince yourself that $\Sigma \# S^2$ is homeomorphic to $\Sigma$. Here $S^2$ is the 2-sphere.

We say the sphere $S^2$ is of genus 0. We say the torus $T^2$ is of genus 1. In general, we say the surface $T^2 \# T^2 \# \ldots \# T^2$, where there are $k$ many $T^2$‘s, is of genus $k$.

So genus just counts the number of “holes” in a compact surface (without boundary).

**Exercise 64.** Convince yourself that
\[
\chi(\Sigma_1 \# \Sigma_2) = \chi(\Sigma_1) + \chi(\Sigma_2) - 2.
\]
(Actually you should prove this, assuming the fact that $\chi$ does not depend on how you triangulate a surface).

Compute $\chi$ of a surface with genus $k$, where $k = 0, 1, 2, \ldots$.

It turns out that for any compact orientable surface (without boundary) $\Sigma$, there exists an integer $g \geq 0$ so that $\Sigma$ is diffeomorphic to a surface of of genus $g$. In the differentiable category, i.e. assume the surface is smooth, we can use Morse theory to prove it [See Simon Donaldson’s book *Riemann Surfaces*, Chapter 2]. One can also prove this in the topological category, but the proof is harder [See Moise’s book *Geometric topology in dimensions 2 and 3*].

In the following, by saying a “sphere” we mean any surface that is homeomorphic (actually, diffeomorphic) to $S^2$. So a “sphere” may not be “round”. Similar remarks for torus and all surfaces of higher genus.

**Exercise 65.** Prove that a sphere cannot have a metric with nonpositive curvature.

**Exercise 66.** Prove that if a torus has a metric with nonnegative curvature, then the curvature must be constant 0.

**Exercise 67.** Prove that a surface of genus $> 1$ cannot have a metric with nonnegative curvature.

**Lemma 9.14.** The torus admits a metric with zero curvature.

**Proof.** (sketch) One can view a torus as obtained by gluing opposite sides of a square together. Thus it “inherits” the flat metric of the (flat) square. Alternatively, view the torus as a circle $S^1$ times another circle $S^1$. Let $u$ be a parameter on the first circle, $v$ be a parameter on the second circle, then we can put the metric
\[
(9.15) \quad du^2 + dv^2.
\]

By Exercise 60, the above “flat” metric cannot be achieved by any surface in $\mathbb{R}^3$ that is diffeomorphic to a torus.

---

7In general, for a set $X$ and a set $Y$, their product is $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$. 
Exercise 68. Consider the four dimensional Euclidean space $\mathbb{R}^4$ with coordinate $x, y, z, t$.

Consider following the subset of $\mathbb{R}^4$:

$$\Sigma = \{(x, y, z, t) \mid x^2 + y^2 = z^2 + t^2 = 1\}.$$  

Prove that $\Sigma$ is diffeomorphic to a torus.

Prove that the curvature of $\Sigma$ is constant 0. Hint: take a local coordinate chart $\sigma$ on some domain in the $u, v$ plane, you can define the first fundamental form as before: $E = \langle \sigma_u, \sigma_u \rangle$, etc.

It turns out a surface of genus $> 1$ admits a metric with constant curvature $-1$. One can use hyperbolic geometry, or use the Ricci flow to prove this. We may return to this later.

We can start with a Möbius band (do an internet search if necessary), its boundary is a single circle, so we can glue a disc to it and make a closed surface, which is $RP^2$. The connected sum $RP^2 \# RP^2$ is called a Klein bottle.

Exercise 69. Compute $\chi$ for $RP^2$ and the Klein bottle.

It can be proved that any compact nonorientable surfaces (without boundary) is diffeomorphic to the connected sum of $k$ many $RP^2$’s, where $k = 1, 2, 3, \ldots$ Again the proof can be done using Morse theory.

10. Higher dimensions: manifolds and tangent vectors

Roughly speaking, differential geometry concerns curved spaces of many dimensions. First of all, what is “space”? Then how do we know it is “curved”?

It is easier to understand 2-dimensional spaces, these are just surfaces in the ordinary sense. Examples are the sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$, the torus (i.e. a surface of one “hole”), surfaces of two “holes”; the following is a picture of a surface of three holes:

---

8Quote from J. Cheeger.
There are also surfaces with boundary (e.g. north hemisphere), but for simplicity we will not consider these. In general, a compact (i.e. bounded, without boundary) surface of \( k \) holes is called a surface of \textbf{genus} \( k \).

There are also noncompact surfaces, for example the hyperboloid of one sheet
\[
x^2 + y^2 - z^2 = 1,
\]
and the paraboloid \( z = x^2 + y^2 \). There are also surface look like this:

![Image of a surface]

However in this course we will not study the surface
\[
x^2 + y^2 - z^2 = 0,
\]
because there is a singularity at \((0, 0, 0)\); in fact this surface contains two cones, whose tips touch at \((0, 0, 0)\).

Although many of these surfaces cannot be written as the graph of a function (i.e. they are not globally of the form \( z = f(x, y) \)), there is one common feature of all these surfaces: everywhere there is a \textbf{local coordinate system}. A coordinate system, in dimension 2, is a map \( \sigma \) from some bounded domain \( \Omega \subset \mathbb{R}^2 \) to a set \( \sigma(\Omega) \) in the surface, so that the map is one-to-one, and as before we need \( \sigma_1, \sigma_2 \) to be linearly independent.

For example, for the surface of genus as in the above picture, near the left extreme we see the normal vector is almost in the negative \( x \) direction (we can assume the \( x \)-axis goes from left to right), thus locally the surface can be written as a graph:
\[
x = f(y, z).
\]
Thus locally the surface has the parametric equation
\[
\begin{align*}
x &= f(u, v) \\
y &= u \\
z &= v.
\end{align*}
\]
In particular, we get a coordinate system \((u, v)\). However, it is clear that this can only be done locally.

In general, the “space” we want to study are \textbf{smooth manifolds} of dimension \( n \). A manifold of dimension \( n \) is a topological space \( M \) so that for any \( p \in M \), there is a bounded domain \( \Omega \subset \mathbb{R}^n \), a map \( \varphi : \Omega \to M \) that is a \textbf{homeomorphism} from \( \Omega \) to \( \varphi(\Omega) \). To be precise, there are more conditions in the definition of a smooth manifold (e.g. what happen when two coordinate systems overlap?), however the above discussion is enough for our course.
Thus in an \( n \)-dimensional manifold \( M \), locally we have some coordinate system \((x_1, \ldots, x_n)\). In general, a coordinate system covers only a small portion of \( M \); however this portion can be a very long, for example a “neighborhood” of a curve.

**Example 10.1.** Our “\( n \)-surface”, defined earlier, is a special case of \( n \)-manifolds.

**Example 10.2.** The most basic example are the Euclidean spaces: \( \mathbb{R}, \mathbb{R}^2, \ldots, \mathbb{R}^n, \ldots \).

For such examples, one can cover the whole space with only one coordinate system.

**Example 10.3.** The \( n \)-dimensional sphere, is the following subset in \( \mathbb{R}^{n+1} \):

\[
S^n = \{(x_1, x_2, \ldots, x_{n+1}) \mid x_1^2 + x_2^2 + \ldots + x_{n+1}^2 = 1\}.
\]

When \( n = 1 \), this is the circle \( S^1 \); when \( n = 2 \) this is the ordinary sphere.

For such examples, one can cover the whole space with only two coordinate systems (one slightly larger than the “north hemisphere”, the other slightly larger than the “south hemisphere”).

Cartesian products of manifolds are also manifold. For example, the torus \( S^1 \times S^1 \), the cylinder \( S^1 \times \mathbb{R} \), and \( S^1 \times S^2, S^2 \times S^2 \), etc. More examples, e.g. projective spaces, can be found in textbooks of topology.

Now we try to define “tangent vectors” of a manifold. Recall a tangent vector on a 2-surface \( \Sigma \) is traditionally defined as a vector in \( \mathbb{R}^3 \), where \( \Sigma \) stays in. The problems in the general situation is, manifold is an independent existence, there is no need to view a manifold as a subset of a Euclidean space \(^9\). Therefore the definition has to be more complicated.

Assume \( x^1, x^2, \ldots, x^n \) is a local coordinate system of an \( n \)-dimensional space \( M \). So \( x^1, x^2, \ldots, x^n \) are functions on \( M \). We can talk about curves in \( M \), i.e. maps from \([0, L]\) to \( M \) in the following form:

\[
\gamma : \begin{cases} 
  x^1 = u^1(t), \\
  x^2 = u^2(t), \\
  \ldots \ldots \\
  x^n = u^n(t).
\end{cases}
\]

One writes \( x^j \) instead of \( x_j \) for historical reasons. This does have some advantage when we “raise or lower indices” (in the future...). In the physics literature this rule is strictly followed. For mathematicians this is not so strictly followed, as many tensors are written in a “coordinate-free” manner. See Chapter 1 of the book *Einstein manifolds* by A. Besse.

\(^9\)Whitney’s theorem, nevertheless, says that any smooth compact \( n \)-manifold can be **smoothly** embedded in \( \mathbb{R}^{2n} \). It is quite easy to show that they can be smoothly embedded in \( \mathbb{R}^{2n+1} \); see Hirsch’s book *Differential topology*. 
Pick any point \( p = (a^1, a^2, ..., a^n) \). We will introduce an equivalence relation on all curves passing through \( p \). Assume \( \gamma_1, \gamma_2 \) are two curves,

\[
\begin{align*}
\gamma_1 : & \quad \begin{cases} 
  x^1 = u^1(t), \\
  x^2 = u^2(t), \\
  \vdots \\
  x^n = u^n(t), 
\end{cases} \\
\gamma_2 : & \quad \begin{cases} 
  x^1 = v^1(t), \\
  x^2 = v^2(t), \\
  \vdots \\
  x^n = v^n(t). 
\end{cases}
\end{align*}
\]

Assume

\[(10.6) \quad \gamma_1(0) = \gamma_2(0) = p.\]

Then we say \( \gamma_1 \sim \gamma_2 \) if \( \dot{u}^j(0) = \dot{v}^j(0) \) for \( j = 1, 2, ..., n \). Here \( \dot{u} \), etc. means derivative.

A tangent vector is an equivalence class \([\gamma]\). The space of all tangent vectors at \( p \) is called the tangent space \( T_pM \). One can show that \( T_pM \) has the structure of an \( n \)-dimensional vector space; for example if

\[
\begin{align*}
\gamma_1 : & \quad \begin{cases} 
  x^1 = a^1 + c^1 t, \\
  x^2 = a^2 + c^2 t, \\
  \vdots \\
  x^n = a^n + c^n t 
\end{cases} \\
\gamma_2 : & \quad \begin{cases} 
  x^1 = a^1 + d^1 t, \\
  x^2 = a^2 + d^2 t, \\
  \vdots \\
  x^n = a^n + d^n t, 
\end{cases}
\end{align*}
\]

then \([\gamma_1] + [\gamma_2]\) is represented by the curve

\[(10.8) \quad \gamma : \quad \begin{cases} 
  x^1 = a^1 + (c^1 + d^1) t, \\
  x^2 = a^2 + (c^2 + d^2) t, \\
  \vdots \\
  x^n = a^n + (c^n + d^n) t. 
\end{cases}\]

For any function \( f : M \to \mathbb{R} \) and any vector \( X \), define

\[(10.9) \quad Xf(p) = \frac{d}{dt} f(u^1(t), u^2(t), ..., u^n(t)).\]

Here \( \gamma(t) = (u^1(t), u^2(t), ..., u^n(t)) \) is any representative of \( X = [\gamma] \).

**Exercise 70.** Prove that \( Xf(p) \) does not depend on the choice of representative \( \gamma \).

Pick any \( i \) so that \( 1 \leq i \leq n \). Let

\[(10.10) \quad \frac{\partial}{\partial x^i} \]

be the unique vector so that

\[(10.11) \quad \frac{\partial}{\partial x^i} x^i = 1, \quad \frac{\partial}{\partial x^j} x^j = 0 \text{ if } j \neq i.\]

In another word,

\[\text{In some way this is similar to the definition of real numbers using Cauchy sequences - a real number } x \text{ is an equivalent class of Cauchy sequences, which turns out to converge to } x. \text{ We don’t need to know what a real number “really” is, we just need to know its laws: arithmetic, order, supremum, etc.}\]
Lemma 10.12. At $p = (a^1, a^2, ..., a^n)$, $\frac{\partial}{\partial x^i}$ is the class $[\gamma]$, where

\[
\gamma : \begin{cases} 
  x^1 = a^1, \\
  \ldots \ldots, \\
  x^i = a^i + t, \\
  \ldots \ldots, \\
  x^n = a^n.
\end{cases}
\]

(10.13)

Lemma 10.14. If $f : M \to \mathbb{R}$ is a function, then

\[
\frac{\partial}{\partial x^i} f = \frac{\partial f}{\partial x^i}.
\]

(10.15)

Lemma 10.16. If $\gamma = (u^1(t), u^2(t), ..., u^n(t))$, $\gamma(0) = p = (a^1, a^2, ..., a^n)$, then

\[
[\gamma] = \dot{u}^1(0) \frac{\partial}{\partial x^1} + \dot{u}^2(0) \frac{\partial}{\partial x^2} + \ldots + \dot{u}^n(0) \frac{\partial}{\partial x^n}.
\]

(10.17)

Having a local coordinate system does not tell us if the space is “curved” or not. We must be able to do geometry, that is, at least we want to measure length and angle. For a curve $\gamma$, the arclength formula is the integration of $|\gamma'|$, therefore we must have a way to measure the length of a tangent vector (e.g. $\gamma'(t)$). Similar to the theory of $n$-surfaces, we define the Riemannian metric $g$, now in an abstract sense:

Let $\text{GL}(n)$ be the set of $n \times n$ matrix. A Riemannian metric on $M$ is a smooth function $g : M \to \text{GL}(n)$ so that for all $p \in M$,

\[
g(p) = \begin{bmatrix} 
  g_{11}(p) & g_{12}(p) & \ldots & g_{1n}(p) \\
  g_{21}(p) & g_{22}(p) & \ldots & g_{2n}(p) \\
  \vdots & \vdots & \ddots & \vdots \\
  g_{n1}(p) & g_{n2}(p) & \ldots & g_{nn}(p)
\end{bmatrix}.
\]

(10.18)

is a positive definite (so by definition it is symmetric) matrix. Recall in local coordinate we can write

\[
p = (x^1, x^2, ..., x^n).
\]

(10.19)

Assume

\[
A = a^1 \frac{\partial}{\partial x^1} + a^2 \frac{\partial}{\partial x^2} + \ldots + a^n \frac{\partial}{\partial x^n}, \quad B = b^1 \frac{\partial}{\partial x^1} + b^2 \frac{\partial}{\partial x^2} + \ldots + b^n \frac{\partial}{\partial x^n},
\]

(10.20)

Define

\[
\langle A, B \rangle = \sum_{i,j=1}^{n} a^i g_{ij} b^j
\]

(10.21)

\[
=[a^1, a^2, ..., a^n] \begin{bmatrix} 
  g_{11}(p) & g_{12}(p) & \ldots & g_{1n}(p) \\
  g_{21}(p) & g_{22}(p) & \ldots & g_{2n}(p) \\
  \vdots & \vdots & \ddots & \vdots \\
  g_{n1}(p) & g_{n2}(p) & \ldots & g_{nn}(p)
\end{bmatrix} \begin{bmatrix} 
  b^1 \\
  b^2 \\
  \vdots \\
  b^n
\end{bmatrix}.
\]
Remark 10.22. From now on, we will use the Einstein summation convention. This means, when an index is being repeated twice, it should be understood that we sum this index from 1 to $n$. Examples:

\begin{equation}
(10.23) \quad g^{ij}\phi_{ij} \text{ means } \sum_{i=1}^{n} \sum_{j=1}^{n} g^{ij}\phi_{ij}.
\end{equation}

\begin{equation}
(10.24) \quad g_{ij}u^{i}u^{j} \text{ means } \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}u^{i}u^{j}.
\end{equation}

\begin{equation}
(10.25) \quad g^{ij}\psi_{j} \text{ means } \sum_{j=1}^{n} g^{ij}\psi_{j}, \text{ so this result depends on } i.
\end{equation}

Remark 10.26. One usually denote by $(g^{ij})$ the inverse of $g$, i.e. $g^{ij}$ is the $(i,j)$ entry in the inverse matrix of $g$.

Since $g$ is symmetric, we have $g^{ji} = g^{ij}$.

The symbol $\delta_{\mu}^{\nu}$ (more serious physicists write this as $\delta_{\mu\nu}$) takes value 1 if $\mu = \nu$ and take the value 0 if $\mu \neq \nu$.

So we have

\begin{equation}
(10.27) \quad g^{\mu i}g_{\nu j} = \delta_{\nu}^{\mu},
\end{equation}

Exercise 71. Prove that

\begin{equation}
(10.28) \quad \dot{g}^{ij} = -g^{ip}\dot{g}_{pq}g^{qj}.
\end{equation}

Here $\dot{g}$ means any directional derivative, e.g. by restricting $g$ on a curve $\gamma(t)$ we view $g$ as a function of $t$, so we take the $t$ derivative.

Given a tangent vector $A \in T_{p}M$, define its length to be

\begin{equation}
(10.29) \quad |A| = \sqrt{\langle A, A \rangle}.
\end{equation}

In local coordinate this is

\begin{equation}
(10.30) \quad \sqrt{g_{ij}a^{i}a^{j}}.
\end{equation}

Given two nonzero tangent vectors $A, B \in T_{p}M$, define the angle between them to be

\begin{equation}
(10.31) \quad \angle(A, B) = \arccos \frac{\langle A, B \rangle}{|A| \cdot |B|}.
\end{equation}

Let $\gamma : [0, C] \to M$ be a curve. Its length is

\begin{equation}
(10.32) \quad L(\gamma) = \int_{0}^{C} |\gamma'(t)| dt = \int_{0}^{C} \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt.
\end{equation}

in local coordinate this is

\begin{equation}
(10.33) \quad \int_{0}^{C} \sqrt{g_{ij}\dot{x}^{i}\dot{x}^{j}} dt.
\end{equation}
One often write

\[ g = g_{ij}dx^i dx^j \]

if \( g_{ij} \) is a Riemannian metric.

**Exercise 72.** Assume Riemannian metric

\[ g_{ij} = \frac{4}{(1 - (x^1)^2 - \ldots - (x^n)^2)^2}. \]

This is called the **disc model of hyperbolic space**.

Assume \( x^1, \ldots, x^n \) is a point with \( (x^1)^2 + \ldots + (x^n)^2 < 1 \). Compute the length of the curve

\[ \gamma(t) = (tx^1, \ldots, tx^n), \quad 0 \leq t \leq 1. \]

**Exercise 73.** Take a local coordinate for the **n-sphere**

\[ S^n = \{(x_1, \ldots, x_n) \in \mathbb{R}^{n+1} | x_1^2 + \ldots + x_n^2 + x_{n+1}^2 = 1\}, \]

then find \( g_{ij} \).

11. Geodesics: First Variation Formula

Assume \( \gamma : [0, C] \to M \) is a **minimal geodesic** connects two points \( p, q \), i.e. \( \gamma(0) = p \), \( \gamma(C) = q \), its length is minimal among all curves that connect \( p, q \). Assume \( \gamma \) has arclength parameter.

Let \( \gamma_r(s) \), with \( -\epsilon < r < \epsilon, 0 \leq s \leq C \) be a family of curves. Assume \( s \) is the arclength parameter for \( \gamma = \gamma_0 \), i.e.

\[ |\gamma'(s)| = 1. \]

Assume \( \gamma_r(0) = p \) for all \( r \), \( \gamma_r(C) = q \) for all \( r \). Because \( \gamma_0 = \gamma \) is minimal, we see

\[ 0 = \frac{d}{dr} \Big|_{r=0} L(\gamma_r) = \frac{d}{dr} \Big|_{r=0} \int_0^C \sqrt{\langle \gamma_\prime_r(s), \gamma_\prime_r(s) \rangle} \, ds. \]

Locally, we can write

\[ \gamma_r : \begin{cases} x^1 = x^1(r, s), \\ x^2 = x^2(r, s), \\ \ldots \\ x^n = x^n(r, s). \end{cases} \]

So

\[ \langle \gamma_\prime_r(s), \gamma_\prime_r(s) \rangle = g_{ij} \dot{x}^i \dot{x}^j. \]

Here \( \dot{x} = x' \) means the \( s \)-partial derivative of \( x \), \( x'' \) means second \( s \)-partial derivative of \( x \). So

\[ \frac{d}{dr} \Big|_{r=0} L(\gamma_r) = \int_0^C \frac{(g_{ij} \dot{x}^i \dot{x}^j)'_r}{2\sqrt{g_{ij} \dot{x}^i \dot{x}^j}} \, ds = \frac{1}{2} \int_0^C (g_{ij} \dot{x}^i \dot{x}^j)' \, ds. \]
Here notice we take derivative at \( r = 0 \), and \(|\gamma'(s)| = 1 \) when \( r = 0 \).

\[
\frac{d}{dr} \bigg|_{r=0} L(\gamma_r) = \frac{1}{2} \int_0^C \frac{\partial g_{ij}}{\partial x^k} \frac{\partial x^k}{\partial r} \hat{x}^i \hat{x}^j + 2 g_{ij} \frac{\partial^2 x^i}{\partial s \partial r} \frac{\partial x^j}{\partial s} \, ds
\]

\[
= \frac{1}{2} \int_0^C \frac{\partial g_{ij}}{\partial x^k} \frac{\partial x^k}{\partial r} \hat{x}^i \hat{x}^j \, ds + 2 \int_0^C g_{ij} \frac{\partial^2 x^i}{\partial s \partial r} \frac{\partial x^j}{\partial s} \, ds
\]

(11.6)

\[
= \frac{1}{2} \int_0^C \frac{\partial g_{ij}}{\partial x^k} \frac{\partial x^k}{\partial r} \hat{x}^i \hat{x}^j + g_{ij} \frac{\partial x^i}{\partial r} \frac{\partial x^j}{\partial s} \, ds - \int_0^C \frac{\partial x^i}{\partial r} \frac{\partial}{\partial s} \left( g_{ij} \frac{\partial x^j}{\partial s} \right) \, ds
\]

In the last equal sign, notice \( \gamma_r(0) = p, \gamma_r(C) = q \), they do not depend on \( r \), so \( \partial x^i / \partial r = 0 \) when \( s = 0 \) or \( C \).

Since \( \partial x^k / \partial r \) can be arbitrary, we see that on a geodesic,

\[
\frac{\partial}{\partial s} (g_{ij} \frac{\partial x^j}{\partial s}) = 0.
\]

So

\[
0 = g_{kj} \frac{\partial^2 x^j}{\partial s^2} + \frac{\partial g_{kj}}{\partial x^m} \frac{\partial x^m}{\partial s} \frac{\partial x^j}{\partial s} - \frac{1}{2} g_{ij} \frac{\partial^2}{\partial x^k \partial x^j} \hat{x}^i \hat{x}^j
\]

(11.7)

\[
= g_{kj} \frac{\partial^2 x^j}{\partial s^2} + \frac{1}{2} g_{kj} \frac{\partial x^m}{\partial s} \frac{\partial x^j}{\partial s} + \frac{1}{2} g_{kj} \frac{\partial x^m}{\partial s} \frac{\partial x^j}{\partial s} - \frac{1}{2} g_{ij} \frac{\partial^2}{\partial x^k \partial x^j} \hat{x}^i \hat{x}^j.
\]

Let \((g^{ij})\) be the inverse matrix of \( g \), so

\[
g^{ik} g_{kj} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}
\]

Multiply by \( g^{ak} \) (then sum over \( k \)),

\[
0 = g^{ak} g_{kj} \frac{\partial^2 x^j}{\partial s^2} + \frac{1}{2} g^{ak} \frac{\partial g_{kj}}{\partial x^i} \frac{\partial x^i}{\partial s} + \frac{1}{2} g^{ak} \frac{\partial g_{kj}}{\partial x^j} \frac{\partial x^j}{\partial s} - \frac{1}{2} g^{ak} \frac{\partial g_{ij}}{\partial x^k} \hat{x}^i \hat{x}^j.
\]

(11.10)

So

\[
\frac{\partial^2 x^a}{\partial s^2} + \frac{1}{2} g^{ak} \left( \frac{\partial g_{kj}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right) \hat{x}^i \hat{x}^j = 0.
\]

Write

\[
\Gamma^a_{ij} = \frac{1}{2} g^{ak} \left( \frac{\partial g_{kj}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right),
\]

(11.12)

(this is called the Christoffel symbol) we get

\[
\frac{\partial^2 x^a}{\partial s^2} + \Gamma^a_{ij} \hat{x}^i \hat{x}^j = 0, \quad a = 1, 2, \ldots, n.
\]

(11.13)

This is called the \textbf{geodesic equation}. Any curve that satisfies the geodesic equation is called a \textbf{geodesic}.

\textbf{Exercise 74.} Prove that the Christoffel symbols are symmetric, i.e. \( \Gamma^a_{ij} = \Gamma^a_{ji} \).

\textbf{Exercise 75.} Find all geodesics on \( \mathbb{R}^n \).
Exercise 76. The upper space model of hyperbolic plane is the domain \( \{(u,v) \mid v > 0\} \) with the Riemannian metric

\[
(11.14) \quad du^2 + dv^2 \over v^2.
\]

Find the Christoffel symbol.

Can you find all the geodesics? What do they look like?

Exercise 77. Use the spherical polar coordinate \((r,\theta)\) on the 2-dimensional sphere \(S^2\), i.e.

\[
(11.15) \quad x = \sin r \cos \theta, \quad y = \sin r \sin \theta, \quad z = \cos r.
\]

Find the Riemannian metric. Find the Christoffel symbol.

Can you find all the geodesics? What do they look like?

Theorem 11.16. Given any unit tangent vector \(T\) at \(p \in M\), there is a unique geodesic starting at \(p\) with initial tangent vector \(T\).

Proof. This a result in ODE. Assume in local coordinate \(p = (p^1,\ldots,p^n)\), and \(T = (b^1,\ldots,b^n)\), then the geodesic is decided by the ODE system

\[
(11.17) \begin{cases}
\frac{\partial^2 x^k}{\partial s^2} + \Gamma^k_{ij} \dot{x}^i \dot{x}^j = 0, \\
x^k(0) = p^k \\
\dot{x}^k = b^k
\end{cases} \quad k = 1, 2, \ldots, n.
\]

The conclusion follows from the existence and uniqueness theorem of ODE systems. \(\square\)

Exercise 78. Prove that in the sphere \(S^2\), geodesics are the “big circles”.

Hint: choose a suitable coordinate system!

At this moment, actually we have not proved that a geodesic is a path connecting \(p,q\) that minimizes distance, even locally. You can show the existence of geodesic using for example Arzela-Ascoli theorem, but the geodesic obtained that way is only Lipschitz (not sure about smooth). It turns out that geodesics are indeed smooth; the proof starts with defining the exponential map and normal coordinate, then prove the Gauss lemma. Moreover, one can prove that a geodesic is truly distance minimizing if its length is sufficiently small. These will be done later.

Example 11.18. On the 2-sphere \(S^2\) of radius \(R\) (therefore of constant curvature \(K = R^{-2}\)), geodesics are big circles. Notice that a piece of big circle is minimal only when its length is at most

\[
(11.19) \quad D = \pi R = \frac{\pi}{\sqrt{K}}.
\]

12. Connection

Assume \(M\) is a curved space with a Riemannian metric \(g\) (so \(M\) is a “Riemannian manifold”). Let \(X\) be a vector at \(p \in M\), and let \(Y\) be a vector field on \(M\).
Think about the case when \( M \) is the standard Euclidean plane \( \mathbb{R}^2 \) with standard coordinate \((u,v)\), let

\[
X = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ at } p = (1,0), \quad Y = \begin{bmatrix} -v \\ u \end{bmatrix}.
\]

Then we have

\[
"D_XY" = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\]

However, if we use the polar coordinate, we see

\[
X = \frac{\partial}{\partial r} \text{ at } p = (1,0), \quad Y = \frac{\partial}{\partial \theta},
\]

and "\(D_XY\)" shall be 0 here.

So it is not straightforward to define the "directional derivative" of \( Y \) in the \( X \) direction. Notice we have not used the metric \( g \).

However, on the other hand, if \( X \) is a vector at \( p \in M \) and \( f : M \to \mathbb{R} \) is a function, then the direction derivative \( Xf \) is well defined: take any curve \( \gamma \) representing \( X \), then

\[
Xf = \left. \frac{d}{dt} f(\gamma(t)) \right|_{t=0}.
\]

In local coordinate, if

\[
X = a^i \frac{\partial}{\partial x^i},
\]

then

\[
Xf = a^i \frac{\partial f}{\partial x^i}.
\]

**Exercise 79.** Prove that \( Xf \) does not depend on the choice of coordinate systems (The proof is just the chain rule).

We want to define a "derivative" for vector fields, called a connection, so that

\[
\nabla_{A_1+A_2}B = \nabla_{A_1}B + \nabla_{A_2}B, \\
\nabla_{A}(B_1+B_2) = \nabla_{A}B_1 + \nabla_{A}B_2, \\
\n\nabla_{fA}(B) = f \nabla_{A}B, \\
\n\nabla_{A}(fB) = (Af)B + f \nabla_{A}B.
\]

These are just the properties of the standard Euclidean derivative \( D \) on \( \mathbb{R}^n \) associated with the "linear" structure of \( \mathbb{R}^n \).

Without further restriction, we can define many (different) connections on \( M \).

If \( X, Y \) are two vector fields,

\[
X = a^i \frac{\partial}{\partial x^i}, \quad Y = b^i \frac{\partial}{\partial x^i},
\]

define their Lie bracket \([X,Y]\) to be another vector field, so that if \( f \) is a function,

\[
[X,Y]f = X(Yf) - Y(Xf).
\]
in local coordinate,

\[(12.10) \quad [X, Y] = a^i \frac{\partial b^j}{\partial x^i} \frac{\partial}{\partial x^j} - b^j \frac{\partial a^i}{\partial x^j} \frac{\partial}{\partial x^i}.\]

**Theorem 12.11.** Given a Riemannian metric, then there exists a unique connection so that

\[(12.12) \quad X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,\]

and

\[(12.13) \quad \nabla_X Y - \nabla_Y X - [X, Y] = 0.\]

The condition (12.12) says the connection is **compatible with the metric**; this looks like the product rule of derivatives on \(\mathbb{R}^n\) - but we have taken the metric on \(\mathbb{R}^n\) for granted.

The condition (12.13) says the connection is **torsion-free**.

Such a connection is called the **Levi-Civita connection**. Some people also call it **covariant derivative**. Connection seems to have a long history; it seems the modern notation is due to Koszul (1921-).

**Proof.** The compatibility with the metric requires that

\[
X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,
\]

\[
Y\langle X, Z \rangle = \langle \nabla_Y X, Z \rangle + \langle X, \nabla_Y Z \rangle,
\]

\[
Z\langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle.
\]

Add the first two equations then subtract the third,

\[
X\langle Y, Z \rangle + Y\langle X, Z \rangle - Z\langle X, Y \rangle = \langle \nabla_X Y + \nabla_Y X, Z \rangle + \langle Y, \nabla_X Z - \nabla_Z X \rangle
+ \langle X, \nabla_Y Z - \nabla_Z Y \rangle.
\]

Use the torsion-free condition, we get \(\nabla_Y X = \nabla_X Y - [X, Y]\) for example, and

\[
\nabla_X Y + \nabla_Y X = 2\nabla_X Y - [X, Y].
\]

Overall we get

\[
X\langle Y, Z \rangle + Y\langle X, Z \rangle - Z\langle X, Y \rangle = \langle 2\nabla_X Y - [X, Y], Z \rangle + \langle Y, [X, Z] \rangle + \langle X, [Y, Z] \rangle.
\]

Thus

\[
\langle \nabla_X Y, Z \rangle = \frac{X\langle Y, Z \rangle + Y\langle X, Z \rangle - Z\langle X, Y \rangle + \langle [X, Y], Z \rangle - \langle Y, [X, Z] \rangle - \langle X, [Y, Z] \rangle}{2}.
\]

This uniquely decides a connection. For example, set

\[
X = \frac{\partial}{\partial x^i}, \quad Y = \frac{\partial}{\partial x^j}, \quad Z = \frac{\partial}{\partial x^m}.
\]
Thus all the Lie bracket are 0. So
\[
2\left\langle \nabla \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^m} \right\rangle = \frac{\partial}{\partial x^i} \left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^m} \right\rangle + \frac{\partial}{\partial x^j} \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^m} \right\rangle - \frac{\partial}{\partial x^m} \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = \frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m}.
\]

We can write, for some coefficients $C^k_{ij}$,
\[
(12.14) \quad \nabla \frac{\partial}{\partial x^i} = C^k_{ij} \frac{\partial}{\partial x^k}.
\]

Then
\[
\frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m} = 2\left\langle C^a_{ij} \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^m} \right\rangle = 2C^a_{ij} g_{am}.
\]

Fix an index $k$. Multiply the above by $g^{km}$ then sum over $m$,
\[
g^{mk} \left( \frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m} \right) = 2C^a_{ij} g_{am} g^{mk} = 2C^k_{ij}.
\]

Thus
\[
C^k_{ij} = \frac{g^{mk}}{2} \left( \frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m} \right).
\]

But this is just the Christoffel symbol $\Gamma^k_{ij}$ in the geodesic equation! Thus
\[
(12.15) \quad \nabla \frac{\partial}{\partial x^i} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}.
\]

Now we check that $\nabla$ defined like this is indeed a connection satisfying (12.12), (12.13). First, in fact for any choice of $C^k_{ij}$, not necessarily $\Gamma^k_{ij}$, we can simply impose (12.7):
\[
\nabla_{\frac{\partial}{\partial x^i}} \left( b_j \frac{\partial}{\partial x^j} \right) = a_i \frac{\partial b_j}{\partial x^i} \frac{\partial}{\partial x^j} + a_i b_j C^k_{ij} \frac{\partial}{\partial x^k}.
\]

Now it is routine to check (12.7) is satisfied.

Next check compatibility to metric. Assume
\[
(12.16) \quad X = a^i \frac{\partial}{\partial x^i}, \quad Y = b^j \frac{\partial}{\partial x^j}, \quad Z = c^k \frac{\partial}{\partial x^k}.
\]

So
\[
X \langle Y, Z \rangle = a^i \frac{\partial}{\partial x^i} \left( b^j \frac{\partial}{\partial x^j}, \ c^k \frac{\partial}{\partial x^k} \right) = a^i \frac{\partial}{\partial x^i} \left( b^j c^k g_{jk} \right) = a^i \frac{\partial b^j}{\partial x^i} c^k g_{jk} + a^i \frac{\partial c^k}{\partial x^i} b^j g_{jk} + a^i \frac{\partial g_{jk}}{\partial x^i} b^j c^k.
\]
if you can view torsion-free as a technical condition. It does lead to simpler computations, Lemma 12.21. Assume \( \gamma(s) = (x^1(s), ..., x^n(s)) \) is a curve with arclength parameter \( s \). Let \( \gamma' = T \), then

\[
\nabla_{\gamma'} T = ((x^k)^{''} + \Gamma^m_{ij} \dot{x}^i \dot{x}^j) \frac{\partial}{\partial x^k}.
\]

Finally, check torsion free:

\[
\nabla_X Y - \nabla_Y X - [X, Y] = \nabla_{a^i \frac{\partial}{\partial x^i}} b^j \frac{\partial}{\partial x^j} - \nabla_{b^j \frac{\partial}{\partial x^j}} a^i \frac{\partial}{\partial x^i} - \left[ a^i \frac{\partial}{\partial x^i}, b^j \frac{\partial}{\partial x^j} \right]
\]

\[
= a^i \frac{\partial b^j}{\partial x^j} \frac{\partial}{\partial x^i} + a^i b^j \Gamma^m_{ij} \frac{\partial}{\partial x^m} - b^j a^i \frac{\partial}{\partial x^i} - a^i b^j \Gamma^m_{ij} \frac{\partial}{\partial x^m}
\]

\[
- \left( a^i \frac{\partial b^j}{\partial x^i \partial x^j} - b^j a^i \frac{\partial}{\partial x^i} \right)
\]

\[
= a^i b^j (\Gamma^m_{ij} - \Gamma^m_{ji}) \frac{\partial}{\partial x^m}.
\]

This is zero because \( \Gamma^m_{ij} = \Gamma^m_{ji} \).

Example 12.20. On \( \mathbb{R}^n \), use the standard coordinate. Thus \( g_{ij} = 0 \) if \( i \neq j \) and \( g_{ij} = 1 \) if \( i = j \). In particular \( g_{ij} \) are constants, so \( \Gamma_{ij} = 0 \).

Thus the Levi-Civita connection on \( \mathbb{R}^n \) (with the standard Euclidean metric) is just the ordinary directional derivative \( D \) on vector fields.

A more detailed discussion of Lie brackets exceeds the scope of this course. Meanwhile you can view torsion-free as a technical condition. It does lead to simpler computations, moreover, the close relation between the Levi-Civita \( \Gamma^k_{ij} \) and the geodesic equation makes it worthwhile.

Lemma 12.21. Assume \( \gamma(s) = (x^1(s), ..., x^n(s)) \) is a curve with arclength parameter \( s \). Let \( \gamma' = T \), then

\[
\nabla_{\gamma'} T = ((x^k)^{''} + \Gamma^m_{ij} \dot{x}^i \dot{x}^j) \frac{\partial}{\partial x^k}.
\]
In particular, $\gamma$ is a geodesic if and only if 

$$\nabla_T T = 0.$$  

Proof. Recall 

$$T = \dot{x}^k \frac{\partial}{\partial x^k}.$$ 

\[\square\] 

Exercise 80. Given a surface $\Sigma$, and (locally defined) vector fields $A$ and $B$ on $\Sigma$. 

Is there a (local) coordinate system $(u,v)$ on $\Sigma$ so that 

$$\frac{\partial}{\partial u} = A, \quad \frac{\partial}{\partial v} = B ?$$ 

Show that a necessary condition is, $A,B$ are linearly independent and the Lie bracket $[A,B] = 0$.  

It turns out that the above is also a sufficient condition - this is part of the Frobenius theorem. 

13. Theorem of Clairaut 

Assume a radial symmetric surface has a metric 

$$du^2 + f(u)^2 dv^2,$$  

i.e. 

$$g = \begin{bmatrix} 1 & 0 \\ 0 & f(u)^2 \end{bmatrix}.$$ 

This surface is also called a warped product. 

Observe 

$$\left< \frac{\partial}{\partial v}, \frac{\partial}{\partial v} \right> = g_{12} = 0.$$ 

Assume $\gamma(s) = (u(s), v(s))$ is a geodesic whose tangent is 

$$T = \dot{u} \frac{\partial}{\partial u} + \dot{v} \frac{\partial}{\partial v}.$$ 

Theorem 13.5 (Clairaut, 1713-1765). Assume $\gamma$ is a geodesic on a radial symmetric surface. Then along $\gamma$, the quantity 

$$f \cdot \sin \angle \left( T, \frac{\partial}{\partial u} \right)$$ 

is a constant (depend only on $\gamma$).
Proof. We now study

\[(13.7) \quad \langle T, \frac{\partial}{\partial v} \rangle.\]

This is a function of \(s\). Take derivative of this function,

\[(13.8) \quad \frac{d}{ds} \langle T, \frac{\partial}{\partial v} \rangle = T \langle T, \frac{\partial}{\partial v} \rangle = \langle \nabla_T T, \frac{\partial}{\partial v} \rangle + \langle T, \nabla_T \frac{\partial}{\partial v} \rangle = \langle T, \nabla_T \frac{\partial}{\partial v} \rangle.\]

Compute

\[(13.9) \quad \Gamma^1_{11} = 0, \quad \Gamma^1_{12} = \Gamma^1_{21} = 0, \quad \Gamma^1_{22} = -ff';\]

\[(13.10) \quad \Gamma^2_{11} = 0, \quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{f'}{f}, \quad \Gamma^2_{22} = 0.\]

So

\[(13.11) \quad \frac{d}{ds} \langle T, \frac{\partial}{\partial v} \rangle = \langle \dot{u} \frac{\partial}{\partial u} + \dot{v} \frac{\partial}{\partial v}, \dot{u} \nabla_{\frac{\partial}{\partial u}} \frac{\partial}{\partial v} + \dot{v} \nabla_{\frac{\partial}{\partial v}} \frac{\partial}{\partial v} \rangle = \langle \dot{u} \frac{\partial}{\partial u} + \dot{v} \frac{\partial}{\partial v}, \dot{u} \frac{f'}{f} \frac{\partial}{\partial v} - \dot{v} \frac{f'}{f} \frac{\partial}{\partial u} \rangle = -\dot{u} \dot{v} ff' + \dot{v} \frac{f'}{f} f^2 = 0.\]

Notice

\[(13.12) \quad \langle T, \frac{\partial}{\partial v} \rangle = \cos \angle \left( T, \frac{\partial}{\partial v} \right) \cdot \left| \frac{\partial}{\partial v} \right| = f \sin \angle \left( T, \frac{\partial}{\partial u} \right).\]

□

Exercise 81. Rotate the curve \(y = e^{-x}\) around the \(x\)-axis, get a surface \(\Sigma\).

Show that passing through every point \(p \in \Sigma\), there exists only one geodesic that can reach \(x = +\infty\); except this one all other geodesic has two ends that goes to \(x = -\infty\).

Draw a picture to show this conclusion.

14. Levi-Civita as induced connection

In general, assume \(D\) is a connection on a manifold \(M\), i.e. \(D\) is a “derivative” satisfying (12.7). Let \(\Sigma\) be a “submanifold” in \(M\). This means \(X\) is a smooth “surface” in \(M\) - an important example is \(M\) is \(\mathbb{R}^3\) and \(\Sigma\) is a smooth surface in it.

Assume \(X, Y\) be vector fields on \(\Sigma\). For any point \(p \in \Sigma\), notice the tangent space of \(\Sigma\) at \(p\), denoted by \(T_p^\Sigma\), is a subspace in the tangent space of \(M\) at \(p\), denoted by \(T_p^M\):

\[T_p^\Sigma \subset T_p^M.\]

Assume \(M\) has a Riemannian metric, then we can talk about two vectors in \(T_p^M\) being orthogonal.

Then define the induced connection \(\nabla\) on \(\Sigma\) (associated with the connection \(D\) on \(M\)) by

\[\nabla_X Y = \text{proj}_{T_p^\Sigma} D_X Y.\]
Here \( \text{proj}_{\Sigma} \) is the orthogonal projection from \( T_p^M \) to \( T_p^\Sigma \).

**Lemma 14.1.** \( \nabla \) is a connection on \( \Sigma \).

*Proof.* It is straightforward to check that (12.7) is satisfied. \( \square \)

**Lemma 14.2.** If \( D \) is the Levi-Civita connection on \( M \), then \( \nabla \) is the Levi-Civita connection on \( \Sigma \).

*Proof.* We first check compatibility with metric. Since \( D \) is Levi-Civita,

\[
X\langle Y, Z \rangle = \langle D_X Y, Z \rangle + \langle Y, D_X Z \rangle.
\]

Notice, by definition, \( D_X Y = \nabla_X Y + W \), where \( W \) is a vector in \( T_p^M \) so that \( W \perp T_p^\Sigma \).

Since \( Z \in T_p^\Sigma \), we get \( \langle W, Z \rangle = 0 \). So

\[
\langle D_X Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.
\]

Similarly, \( \langle Y, D_X Z \rangle = \langle Y, \nabla_X Z \rangle \). Thus

\[
X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.
\]

Next, we check the torsion free condition: it is satisfied by \( D \),

\[
D_X Y - D_Y X - [X, Y] = 0.
\]

Project this to \( T_p^\Sigma \),

\[
\nabla_X Y - \nabla_Y X - \text{proj}_{\Sigma}[X, Y] = 0.
\]

Since \( X, Y \) are tangent to \( \Sigma \), we have \( [X, Y] \in T_p^\Sigma \). Therefore \( \text{proj}_{\Sigma}[X, Y] = [X, Y] \).

Thus by the uniqueness theorem 12.11, \( \nabla \) is the Levi-Civita connection on \( \Sigma \). \( \square \)

There is a difference between the Levi-Civita connection on \( \Sigma \) and the flat connection (i.e. Levi-Civita connection on \( \mathbb{R}^3 \)):

\[
\Pi(X, Y) = D_X Y - \nabla_X Y.
\]

Therefore \( \Pi(X, Y) \) is perpendicular to \( \Sigma \). \( \Pi(X, Y) \) is called the second fundamental form. In particular,

**Lemma 14.4.** \( \Pi(X, Y) \) is perpendicular to \( \Sigma \).

Now assume \( M = \mathbb{R}^3 \) and \( \Sigma \) is a surface. Then \( D \) is the standard directional derivative on vector fields, and by projection this induces the Levi-Civita connection on \( \Sigma \).

In this situation, the second fundamental form defined before is essentially the same as the one defined earlier. First, we shall identify a 2-dimensional vector on the \((u, v)\) plane with a “real” tangent vector of \( \Sigma \) by the differential \( \sigma_* = D\sigma \): say

\[
X = \sigma_* \frac{\partial}{\partial u} = \sigma_u, \quad Y = \sigma_* \frac{\partial}{\partial v} = \sigma_v.
\]

Then the second fundamental form defined earlier is just

\[
\begin{bmatrix}
\langle \Pi(X, X), N \rangle & \langle \Pi(X, Y), N \rangle \\
\langle \Pi(X, Y), N \rangle & \langle \Pi(Y, Y), N \rangle
\end{bmatrix}.
\]
In fact, for example the \((1,1)\) entry should be \(\langle \sigma_{uu}, N \rangle\), but
\[
\langle \sigma_{uu}, N \rangle = \left\langle \frac{\partial}{\partial u} \frac{\partial \sigma}{\partial u}, N \right\rangle = \langle D_{\sigma_u} \sigma_u, N \rangle = \langle \nabla_{\sigma_u} \sigma_u + \Pi(\sigma_u, \sigma_u), N \rangle = \langle \Pi(\sigma_u, \sigma_u), N \rangle;
\]
here notice \(\nabla_{\sigma_u} \sigma_u\) is in the tangent plane \(T_p^\Sigma\), so perpendicular to \(N\).

When we compute \(\nabla_X Y\) at a point \(p\), for \(X\) we only need its value at \(p\), but for \(Y\) the value at \(p\) is not enough; we can take any function \(f\) so that \(f = 1\) at \(p\), then
\[
\nabla_X (fY) = (Xf)Y + f\nabla_X Y = (Xf)Y + \nabla_X Y \text{ at } p,
\]
this result is different from \(\nabla_X Y\) while \(fY = Y\) at \(p\). However, when compute \(\Pi(X,Y)\) at \(p\), the values of \(X,Y\) at \(p\) is sufficient - that makes \(\Pi\) a tensor. In fact, if \(Y_1 = Y_2\) at \(p\), write in local coordinate
\[
Y_1 = a_i \frac{\partial}{\partial x^i}, \quad Y_2 = b_j \frac{\partial}{\partial x^j},
\]
then
\[
\Pi(X, Y_1) - \Pi(X, Y_2) = D_X Y_1 - \nabla_X Y_1 - (D_X Y_2 - \nabla_X Y_2)
\]
\[
= Xa_i \frac{\partial}{\partial x^i} + a_i D_X \frac{\partial}{\partial x^i} - \left( Xa_i \frac{\partial}{\partial x^i} + a_i \nabla_X \frac{\partial}{\partial x^i} \right)
\]
\[
- \left[ Xb_i \frac{\partial}{\partial x^i} + b_i D_X \frac{\partial}{\partial x^i} - \left( Xb_i \frac{\partial}{\partial x^i} + b_i \nabla_X \frac{\partial}{\partial x^i} \right) \right]
\]
\[
= a_i D_X \frac{\partial}{\partial x^i} - a_i \nabla_X \frac{\partial}{\partial x^i} - b_i D_X \frac{\partial}{\partial x^i} + b_i \nabla_X \frac{\partial}{\partial x^i} = 0 \text{ at } p,
\]
because \(a_i(p) = b_i(p)\).

**Lemma 14.5.** Assume \(\gamma\) is a curve with arclength parameter in a surface \(\Sigma\). \(\kappa_g\) is its geodesic curvature and \(T\) is its tangent vector. Then
\[
|\nabla_T T| = |\kappa_g|.
\]
In particular, \(\gamma\) is a geodesic if and only if \(\kappa_g = 0\).

**Proof.** Recall
\[
\kappa_g = \langle \gamma'', N \times T \rangle,
\]
and \(\gamma'' = D_T (\gamma') = D_T T\). So
\[
\kappa_g = \langle D_T T, N \times T \rangle = \langle \nabla_T T + \Pi(T,T), N \times T \rangle = \langle \nabla_T T, N \times T \rangle,
\]
here recall \(\Pi(T,T)\) is perpendicular to \(\Sigma\) while \(N \times T\) is tangent to \(\Sigma\), so their product is 0.

Next, since \(\langle T, T \rangle = 1\), we see
\[
0 = T \langle T, T \rangle = 2 \langle \nabla_T T, T \rangle,
\]
so \(\nabla_T T \perp T\). Now the tangent plane \(T_p^\Sigma\) is 2-dimensional and \(N \times T \perp T\), so \(\nabla_T T\) must be in the \(\pm N \times T\) direction. We see directly \(|N \times T| = 1\), thus
\[
\langle \nabla_T T, N \times T \rangle = |\nabla_T T| \cos(0 \text{ or } \pi) = \pm |\nabla_T T|.
\]
\(\square\)
In particular, $\kappa_g$ is intrinsic. Notice our proof here is much easier - the price we pay is the more abstract notions of connection and induced connection.

15. **Space of constant negative curvature**

The Poincare disc, $\mathbb{H}$, is the domain $u^2 + v^2 < 1$ on the $u, v$ plane, with a Riemannian metric

$$4 \frac{du^2 + dv^2}{(1 - u^2 - v^2)^2}.\tag{15.1}$$

That is, the metric ("first fundamental form") is

$$g = \begin{bmatrix} \frac{4}{(1-u^2-v^2)^2} & 0 \\ 0 & \frac{4}{(1-u^2-v^2)^2} \end{bmatrix}.\tag{15.2}$$

So curvature is

$$K = -\frac{1}{\sqrt{EG}} \left[ \left( \frac{\sqrt{E}}{\sqrt{G}} \right)_v + \left( \frac{\sqrt{G}}{\sqrt{E}} \right)_u \right]\tag{15.3}$$

$$= -\frac{1}{4} \left( 1 - u^2 - v^2 \right)^2 \left[ \left( \frac{2v}{1 - u^2 - v^2} \right)_v + \left( \frac{2u}{1 - u^2 - v^2} \right)_u \right]$$

$$= -\frac{1}{4} \left[ 2(1 - u^2 - v^2) + 4v^2 + 2(1 - u^2 - v^2) + 4u^2 \right] = -1.$$

So it is of constant curvature $-1$.

Solving the geodesic equation is a daunting job. So we look for other ways to find geodesic.

In general, assume $M$ is a Riemannian manifold, then we can make it into a metric space, i.e. for $p, q \in M$, define $d(p, q)$ be the length of a minimal geodesic $\gamma$ linking $p$ and $q$.

As an example, on the unit sphere, the distance from the north pole $p$ to the south pole $q$ is $\pi$, and there are infinitely many minimal geodesics linking $p$ and $q$.

A map $\tau : M \to M$ is called an isometry, if for any $p, q \in M$,

$$d(\tau(p), \tau(q)) = d(p, q).\tag{15.4}$$

**Lemma 15.5.** Assume $\tau : M \to M$ is an isometry, and the fixed point set

$$\{ p | \tau(p) = p \}\tag{15.6}$$

is a curve $\gamma$. Then $\gamma$ is a geodesic.

**Proof.** Take a point $p \in \gamma$. Let $T$ be the tangent vector of $\gamma$ at $p$. Let $\alpha$ be the geodesic starting at $p$ with initial tangent vector $T$.

Observe an isometry moves a geodesic to a geodesic.
Since \( \tau \) moves \( p \) to \( p \) and \( T \) to \( T \), it must move \( \alpha \) to a geodesic at \( a \) with initial tangent \( T \). By the uniqueness of geodesic, Theorem 11.16, we see \( \tau \alpha = \alpha \). So \( \alpha \) is contained in \( \gamma \). So \( \alpha = \gamma \).

\[ \square \]

**Exercise 82.** Prove that in the sphere \( S^2 \), geodesics are the “big circles”.

*Hint:* use the above lemma.

Especially, any “diameter”, i.e. the curve \( u = kv \) in \( \mathbb{H} \) is a geodesic.

The distance to 0 is

\[
(15.7) \quad s = \int_0^r \sqrt{\frac{4}{(1-t^2)^2}} dt = \ln(1+t) - \ln(1-t) \bigg|_0^r = \ln \frac{1+r}{1-r}.
\]

We can then solve

\[
(15.8) \quad r = \frac{e^s - 1}{e^s + 1}.
\]

With the origin removed, we can think the Poincare disc as a warped product

\[
(15.9) \quad ds^2 + \frac{4r^2}{(1-r^2)^2} d\theta^2 = ds^2 + \sinh^2 s d\theta^2
\]

For any geodesic \( \gamma \), let \( \beta \) be the angle between \( \gamma' \) and the outward radius. The Clairaut theorem tells us

\[
(15.10) \quad \sinh s \sin \beta = c,
\]

that is,

\[
(15.11) \quad 2r \sin \beta = c(1-r^2).
\]

**Remark 15.12.** Notice that the metric here is conformal to the Euclidean metric on the \( u,v \) disc, i.e. the metric \( g \) (as a matrix) is a multiple of the identity matrix \( I \), although this multiple \( f \) (which equals to \( 4(1-u^2-v^2)^{-2} \)) changes from point to point, thus for two 2-dimensional vectors \( A,B \), we see

\[
\langle A, B \rangle = f A \cdot B,
\]

here the left is the metric, the right is the ordinary dot product on the \( u,v \) space. Thus

\[
\frac{\langle A, B \rangle}{\sqrt{\langle A, A \rangle \langle B, B \rangle}} = \frac{A \cdot B}{\sqrt{(A \cdot A)(B \cdot B)}};
\]

Thus we can compute angle on such coordinate system directly using the \( u,v \)-coordinate in the Euclidean way (without using the metric).

**Lemma 15.13.** Geodesics in \( \mathbb{H} \) are diameters or circles that intersects \( u^2 + v^2 = 1 \) at right angle.

*Proof.* Elementary geometry. Let \( c \) be the unit circle \( u^2 + v^2 = 1 \). Let \( x \) be a circle intersects \( c \) at right angle. We prove that the part of \( x \) inside \( c \), which we denote by \( x_{\text{in}} \),
is a geodesic. In fact, take a point $P \in x_{in}$, write $r = OP$. Assume the line $OP$ intersects $x$ at $P, B$. Let $d$ be the radius of $x$, and $O'$ be the center of $x$, then $\angle O'PB = \frac{\pi}{2} - \theta$, thus

$$2d \sin \theta = PB = OB - OP = \frac{1}{r} - r = \frac{1 - r^2}{r}.$$ \hfill \Box

**Theorem 15.14.** Any compact surface admits a metric of constant curvature.

**Proof.** For convenience, we only consider orientable surfaces. A sphere admits a metric with constant curvature 1, and we have seen that a torus admits a metric of curvature 0.

So we only need to consider surfaces with genus $g > 1$. We follow the notes of McMullen.

First, observe that on the hyperbolic plane, there is a regular pentagon with angle $\pi/2$, that is, a pentagon all 5 of its sides are geodesics of equal length, and all angles are $\pi/2$. In fact, we can work on the disc model of hyperbolic plane; from the origin we can travel in directions $\theta = 0, \frac{2\pi}{5}, \frac{4\pi}{5}, \frac{6\pi}{5}, \frac{8\pi}{5}$ with the same distance $L$, end up at five points; use these five points as vertices we build a pentagon $P_L$ with equal geodesic sides. If $L$ is small, then the angle of $P$ is almost the Euclidean value $3\pi/5$; and when $L \to \infty$ we see the angle $\to 0$. Thus by the intermediate value theorem, there is some $L$ so that the angle of $P_L$ is $\pi/2$.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{pentagon.png}
\caption{Find a regular tiling pentagon with angle $\pi/2$.}
\end{figure}

Next, we can take a disc with $g$ many holes, and take another copy (replica) of it then glue together along boundary, we get a surface of genus $g$.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{disc_with_holes.png}
\caption{A disc with 3 holes.}
\end{figure}
Then a surface of genus $g$ can be tiled by $8g - 8$ many pentagons; for such a tiling we can compute the Euler characteristic:

$$\frac{5 \cdot (8g - 8)}{4} - \frac{5 \cdot (8g - 8)}{2} + 8g - 8 = 2 - 2g.$$ 

Thus we can use the $\pi/2$ angle pentagon to tile a surface of genus $g$, the result is smooth, with constant curvature $-1$. □

As we have seen, surface with constant curvature and genus at least 1 cannot be embedded into $\mathbb{R}^3$.

A stronger result is the uniformization of surfaces, which says that given any metric $g$ on a compact surface $\Sigma$, one can conformally deform the metric into one with constant curvature, i.e. there is a function $f$ so that the metric $e^f g$ is of constant curvature. The proof uses nonlinear PDE technique.

**Theorem 15.15.** Any geodesic triangle in $\mathbb{H}$ has area $\pi - \phi_1 - \phi_2 - \phi_3$. In particular, any geodesic triangle has total angle $< \pi$.

Assume $p$ is a point that is not on a geodesic line $\gamma$, then passing through $p$ there are infinitely many geodesic lines that does not intersect (“parallel to”) $\gamma$. The tangents of these lines at $p$ are all within one angle.

This is the basic theorem (axiom ?) of the Lobachevsky geometry.

**Exercise 83.** Decide all geodesics in the upper half plane model of the hyperbolic plane.

**Exercise 84.** Decide all geodesics in the upper half plane model of the hyperbolic plane.

**Exercise 85.** Consider the curve ("tractrix"): for $t \geq 0$,

$$x = \cos t, \quad y = \ln(\sec t + \tan t) - \sin t$$

(15.16)

Sketch this curve.

Rotate this curve around the $y$-axis, get a surface (called a “pseudo-sphere”) $\Sigma$.

Compute the curvature of $\Sigma$. 

[Figure. A surface of genus 3 tiled by 16 pentagons - we show 8 of them.]
Exercise 86. Prove the law of cosine on sphere:

Assume $ABC$ is a “geodesic triangle” on the sphere $S^2$, i.e. the three sides $AB$, $BC$, $CA$ are all minimal geodesics. Let $AB$ denote the length of side $AB$. Prove

$$
(15.17) \quad \cos BC = \cos AB \cos AC + \sin AB \sin AC \cos \angle A.
$$

Exercise 87. Prove the law of sine on sphere:

Assume $ABC$ is a “geodesic triangle” on the sphere $S^2$, i.e. the three sides $AB$, $BC$, $CA$ are all minimal geodesics. Let $AB$ denote the length of side $AB$. Prove

$$
(15.18) \quad \frac{\sin \angle A}{\sin BC} = \frac{\sin \angle B}{\sin CA} = \frac{\sin \angle C}{\sin AB}.
$$

Exercise 88. Prove the law of cosine on $\mathbb{H}$:

Assume $ABC$ is a “geodesic triangle” on the $\mathbb{H}$, i.e. the three sides $AB$, $BC$, $CA$ are all minimal geodesics. Let $AB$ denote the length of side $AB$. Prove

$$
(15.19) \quad \cosh BC = \cosh AB \cosh AC - \sinh AB \sinh AC \cos \angle A.
$$

Exercise 89. Prove the law of sine on $\mathbb{H}$:

Assume $ABC$ is a “geodesic triangle” on the $\mathbb{H}$, i.e. the three sides $AB$, $BC$, $CA$ are all minimal geodesics. Let $AB$ denote the length of side $AB$. Prove

$$
(15.20) \quad \frac{\sin \angle A}{\sinh BC} = \frac{\sin \angle B}{\sinh CA} = \frac{\sin \angle C}{\sinh AB}.
$$

16. CURVATURE TENSOR

The curvature tensor is defined to be

$$
(16.1) \quad R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \frac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}} \nabla_{\frac{\partial}{\partial x^j}} Z - \nabla_{\frac{\partial}{\partial x^j}} \nabla_{\frac{\partial}{\partial x^i}} Z.
$$

So it is a vector that depends on three indices. Why do we call it a “curvature”? In Euclidean space, we have for any vector fields $V$,

$$
(16.2) \quad \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} V - \frac{\partial}{\partial x^j} \frac{\partial}{\partial x^i} V = 0,
$$

that is, we can exchange the order of $i,j$ derivatives. One cannot hope this to hold on a curved space, with the derivative replaced by the Levi-civita connection. So one uses the above curvature, which records the extent of non-commutative of derivatives, as a measure of how curved a space is.

The definition we give here is a rather oversimplified one. In general, one defines the curvature tensor $R(X,Y)Z$ to be

$$
(16.3) \quad R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.
$$

However, many authors (especially those doing general relativity) use the definition

$$
(16.4) \quad R(X,Y)Z = -\nabla_X \nabla_Y Z + \nabla_Y \nabla_X Z + \nabla_{[X,Y]} Z.
$$

Unfortunately, there is no general agreement on which one is best, even among workers in differential geometry. A small inconvenience one must get used to.
Remark 16.5. We omitted a lot of fundamental properties here. One need to read Kobayashi-Nomizu to gain a working knowledge of curvature tensor.

We write
\begin{equation}
R_{i,j,k,l} = \left\langle \nabla \frac{\partial}{\partial x^k} \nabla \frac{\partial}{\partial x^l}, \nabla \frac{\partial}{\partial x^j} \nabla \frac{\partial}{\partial x^i} - \nabla \frac{\partial}{\partial x^j} \nabla \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l} \right\rangle.
\end{equation}

Similarly, this is a tensor. What is a “tensor”? You can think this one is a function from a small domain (a coordinate system) in $M$ to the $n^4$ dimensional space $\mathbb{R}^{n^4}$; for example, if $n = 5$ there are five base vectors,
\begin{equation}
\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}, \frac{\partial}{\partial x^4}, \frac{\partial}{\partial x^5},
\end{equation}
so there are $5^4 = 625$ many components:
\begin{equation}
R_{1111}, R_{1112}, R_{1113}, R_{1114}, R_{1115}, R_{1121}, R_{1122}, R_{1123}, R_{1124}, \ldots, R_{5555}.
\end{equation}
On the other hand, if $n = 2$ there are only two vectors,
\begin{equation}
\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2},
\end{equation}
so there are $2^4 = 16$ many components:
\begin{equation}
R_{1111}, R_{1112}, R_{1121}, R_{1122}, R_{1211}, R_{1212}, R_{1221}, R_{1222}, R_{2111}, R_{2112}, R_{2121}, R_{2211}, R_{2212}, R_{2221}, R_{2222}.
\end{equation}
Fortunately, many components are always 0:

**Theorem 16.11.** We have $R_{iijk} = R_{jkii} = 0$ (here repeated index does not mean summation!)

(Try to prove it; the first identity is easy.)

**Theorem 16.12.** We have $R_{ijkl} = -R_{jikl}, R_{ijkl} = -R_{ijlk}, R_{ijkl} = -R_{klij}$.

(Try to prove it; the first identity is easy.)

So accepting the above results, for $n = 2$ we only care for
\begin{equation}
R_{1221} = \left\langle \nabla \frac{\partial}{\partial u} \nabla \frac{\partial}{\partial v}, \nabla \frac{\partial}{\partial v} \nabla \frac{\partial}{\partial u}, \frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right\rangle.
\end{equation}

**Remark 16.14.** You can view $\frac{\partial}{\partial u}$ is the vector $\sigma_u$ on the surface, $\frac{\partial}{\partial v}$ is the vector $\sigma_v$ on the surface; they are not exactly the same but there is no harm thinking this way.

**Theorem 16.15** (Gauss equation). Assume $\Sigma$ is a submanifold in $M$. Assume $x_1 = u, x_2 = v, \ldots$ is a coordinate system on $\Sigma$.

Then
\begin{equation}
R^M_{1221} = R^\Sigma_{1221} - \left[ II(\frac{\partial}{\partial u}, \frac{\partial}{\partial u}) II(\frac{\partial}{\partial v}, \frac{\partial}{\partial v}) - II(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}) II(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}) \right].
\end{equation}

\[11\] There are many different types of tensors.
Recall on $\sigma$, the Levi-Civita connection is just the induced connection associated with the Levi-Civita connection on $\Sigma$.

**Proof.** Compute

(16.17)  
\[ R_{1221}^M = \left\langle \nabla_{\frac{\partial}{\partial u}} \nabla_{\frac{\partial}{\partial v}}, \frac{\partial}{\partial u} \right\rangle. \]

We see

(16.18)

\[
\left\langle \nabla_{\frac{\partial}{\partial u}} \nabla_{\frac{\partial}{\partial v}}, \frac{\partial}{\partial u} \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial u}} \left[ \nabla_{\frac{\partial}{\partial v}} \frac{\partial}{\partial u} + \Pi \left( \frac{\partial}{\partial v}, \frac{\partial}{\partial u} \right) \right], \frac{\partial}{\partial u} \right\rangle \\
= \left\langle \nabla_{\frac{\partial}{\partial u}} \nabla_{\frac{\partial}{\partial v}} \frac{\partial}{\partial u}, \frac{\partial}{\partial u} \right\rangle + \left\langle \Pi \left( \frac{\partial}{\partial u}, \nabla_{\frac{\partial}{\partial v}} \frac{\partial}{\partial u} \right), \frac{\partial}{\partial u} \right\rangle \\
+ \left\langle \nabla_{\frac{\partial}{\partial u}} \Pi \left( \frac{\partial}{\partial v}, \frac{\partial}{\partial u} \right), \frac{\partial}{\partial u} \right\rangle.
\]

Since the second fundamental form $\Pi$ is normal, the second term is 0, while

(16.19)

\[
\left\langle \nabla_{\frac{\partial}{\partial u}} \Pi \left( \frac{\partial}{\partial v}, \frac{\partial}{\partial v} \right), \frac{\partial}{\partial u} \right\rangle = -\left\langle \Pi \left( \frac{\partial}{\partial v}, \frac{\partial}{\partial v} \right), \nabla_{\frac{\partial}{\partial u}} \frac{\partial}{\partial u} \right\rangle - \left\langle \Pi \left( \frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right), \frac{\partial}{\partial u} \right\rangle.
\]

Similarly we get

(16.20)

\[
\left\langle \nabla_{\frac{\partial}{\partial v}} \nabla_{\frac{\partial}{\partial u}}, \frac{\partial}{\partial u} \right\rangle = \left\langle \nabla_{\frac{\partial}{\partial u}} \nabla_{\frac{\partial}{\partial v}} \frac{\partial}{\partial u}, \frac{\partial}{\partial u} \right\rangle - \left\langle \Pi \left( \frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right), \Pi \left( \frac{\partial}{\partial u}, \frac{\partial}{\partial v} \right) \right\rangle.
\]

Assume $M = \mathbb{R}^3$, then $R_{1221}^M = 0$. Assume at the point $p \in \Sigma$ where we perform this computation, we have

(16.21)

\[
\left| \frac{\partial}{\partial u} \right| = \left| \frac{\partial}{\partial v} \right| = 1, \quad \frac{\partial}{\partial u} \perp \frac{\partial}{\partial v}.
\]

Then the quantity involving $\Pi$ is exactly the Gauss curvature defined extrinsically, and $R_{1221}^\Sigma$ is the Gauss curvature!

**Definition 16.22.** The Ricci curvature is defined by

(16.23)  
\[ \text{Ric}_{ij} = g^{pq}R_{pijq}. \]

So the Ricci curvature is a tensor with two indices.

17. **Jacobi field and meaning of curvature**

Let $p \in M$, and $V \in T_p M$ is a vector at $p$. Find a curve $\gamma$ so that

(17.1)  
\[ \gamma(0) = p, \quad \gamma'(0) = V/|V|, \quad \nabla_T T = 0. \]
So $\gamma$ is a geodesic parametrized by arclength $s$. Define the exponential map $\exp_p : T_pM \to M$ to be

$$\exp_p(V) = \gamma(|V|).$$

**Example.** On $\mathbb{R}^3$, $\exp_0$ is the identity map.

**Example.** On sphere $S^2$, the exponential map is easy to understand.

Let $F(s, \theta)$ be a map $[0, L] \times [-\epsilon, \epsilon] \to M$. Assume for each fixed $\theta$, the curve

$$\gamma_\theta(s) = F(s, \theta)$$

is a geodesic and $s$ the arclength parameter. Then the vector

$$\frac{\partial F}{\partial \theta} \bigg|_{\theta=0}$$

which is the tangent vector of $\theta \mapsto F(s, \theta)$, is called a Jacobi field. Now the $s$ derivative of $F$ gives a vector field, $T$. Since all $\gamma_s$ are geodesic,

$$\nabla_T T = 0.$$

**Lemma 17.6.**

$$\nabla_T J = \nabla_J T.$$

**Proof.** For any test function $\varphi$,

$$T\varphi = \frac{\partial \varphi \circ F}{\partial s}, \quad JT\varphi = \frac{\partial^2 \varphi \circ F}{\partial \theta \partial s},$$

$$J\varphi = \frac{\partial \varphi \circ F}{\partial \theta}, \quad TJ\varphi = \frac{\partial^2 \varphi \circ F}{\partial s \partial \theta},$$

so $TJ\varphi = JT\varphi$, so $[T, J] = 0$.

By (12.13), $\nabla_T J = \nabla_J T$. \qed

So

$$\nabla_T \nabla_T J = \nabla_T \nabla_J T = \nabla_T \nabla_J T - \nabla_J \nabla_T T = R(T, J)T.$$

This is the Jacobi field equation.

**Exercise 90.** Solve the Jacobi field equation on sphere, on plane, and on $\mathbb{H}$.

**Lemma 17.11.** If

$$J(0) \perp T, \quad \nabla_T J(0) \perp T,$$

then $J \perp T$ for all $s$.

**Proof.**

$$\langle \nabla_T \nabla_T J, T \rangle = 0.$$

This implies

$$\langle J, T \rangle''_s = TT \langle J, T \rangle = \langle \nabla_T \nabla_T J, T \rangle = 0.$$
So \( \langle J, T \rangle \) is linear in \( s \). \( J(0) \perp T \) implies
\[
\langle J, T \rangle(0) = 0,
\]
\( \nabla_T J(0) \perp T \) implies
\[
\frac{d}{ds} \langle J, T \rangle(0) = T \langle J, T \rangle(0) = \langle \nabla_T J, T \rangle(0) = 0.
\]
\[\square\]

A vector field \( P \) along \( \gamma \) is called \textbf{parallel}, if
\[
\nabla_T P = 0.
\]
Note then
\[
T \langle P, P \rangle = 2 \langle \nabla_T P, P \rangle = 0,
\]
so \( |P| \) is constant along \( \gamma \). Moreover,
\[
T \langle T, P \rangle = \langle \nabla_T T, P \rangle + \langle T, \nabla_T P \rangle = 0,
\]
so \( \angle(T, P) \) remain constant along \( \gamma \). In particular, if at \( \gamma(0) \), we have \( |P(0)| = 1 \), \( P(0) \perp T(0) \), then for all \( s \),
\[
|P(s)| = 1 \quad P(s) \perp T(s).
\]
Let \( J \) be the Jacobi field associated with
\[
F(s, \theta) = \exp_p(s \cos \theta, s \sin \theta)^T.
\]
So \( J(0) = 0 \). Then \( T(0) = (1, 0)^T \), and
\[
\nabla_T J(0) = \nabla_T J = \frac{d}{d\theta} \bigg|_{\theta=0} (\cos \theta, \sin \theta)^T = (-\sin \theta, \cos \theta)^T \bigg|_{\theta=0} = (0, 1)^T,
\]
so \( \nabla_T J(0) \perp T(0) \). Now assume \( n = 2 \). Take \( P(0) = (0, 1)^T \). So there is a function \( J = fP \), we can assume \( f \geq 0 \) for small \( s \). So
\[
\nabla_T \nabla_T J = f''(s)P = R(TJ)T
\]
So
\[
f'' = \langle f''(s)P, P \rangle = \langle R(T, J)T, P \rangle = -f \langle R(P, T)T, P \rangle = -Kf.
\]

18. \textbf{Bonnet-Myers theorem and meaning of curvature}

Here is the theorem of Bonnet (1819-1892):

\textbf{Theorem 18.1.} Assume the curvature satisfies \( K \geq c > 0 \), then the space is bounded and the diameter of the space is at most
\[
\frac{\pi}{\sqrt{c}}.
\]
Proof. We have

\[ f'' \leq -cf, \quad f(0) = 0, \quad f'(0) = 1. \]  

Now let

\[ h(s) = \frac{1}{\sqrt{c}} \sin s\sqrt{c}. \]

Then

\[ h(0) = 0, \quad h'(0) = 1, \quad h'' = -ch. \]

Now by l'Hospital's rule,

\[ \lim_{s \to 0} \frac{f}{h} = 1. \]

Now

\[ \left( \frac{f}{h} \right)' = \frac{f'h - fh'}{h^2}. \]

Note

\[ (f'h - fh')(0) = 0, \quad (f'h - fh')' = f''h - fh'' \leq -cfh - cfh = 0. \]

so \( f'h - fh' \leq 0. \) So

\[ \left( \frac{f}{h} \right)' = \frac{f'h - fh'}{h^2} \leq 0. \]

Since \( f > 0 \) for small \( s \), we see \( f \) will have the second zero before \( s = \frac{\pi}{\sqrt{c}} \).

By a geometric argument (give in class) we see that \( \gamma \) cannot be minimal beyond distance \( \frac{\pi}{\sqrt{c}} \). \( \square \)

Bonnet's theorem was later improved by Myers, who reaches the same conclusion but using weaker assumptions (assuming only the Ricci curvature bound). So people mention this result by the Bonnet-Myers theorem.

We try to find a formal power series solution \( f \): assume

\[ f = s + a_2 s^2 + a_3 s^3 + \ldots \]

so

\[ f'' = 2a_2 + 6a_3 s + \ldots = -Kf = -K(s + a_2 s^2 + \ldots) \]

so

\[ a_2 = 0, \quad a_3 = -\frac{K}{6}, \ldots \]

So the perimeter of curve \( \theta \mapsto \exp_p(s \cos \theta, s \sin \theta)^T \) is

\[ 2\pi s - 2\pi \frac{Ks^3}{6} + O(s^4). \]

Area enclosed by this curve is

\[ \pi s^2 - \frac{\pi s^4}{12} + O(s^5). \]
So here comes the geometric meaning of curvature: if $K > 0$, then the nearby geodesics tend to converge (relatively; compared with $\mathbb{R}^2$); if $K < 0$, then the nearby geodesic tend to diverge (relatively; compared with $\mathbb{R}^2$). This has an effect on area. So the area of positive curvature surface "shrinks" compare with $\mathbb{R}^2$.

**Exercise 91.** On $S^2$, the plane, and $\mathbb{H}$, compute the length of

(18.15) $\partial B_R(p) = \{x \mid d(x, p) = R\}$.

**Exercise 92.** On $S^2$, the plane, and $\mathbb{H}$, compute the area of

(18.16) $B_R(p) = \{x \mid d(x, p) \leq R\}$.

19. **General relativity**
FURTHER readings

More on differential geometry of 2-surfaces:


This is a classical book that is at a level similar to our course.

The following are more advanced:

Analysis:


J. Dieudonné: Foundations of Modern Analysis, I.

After that you can read some real analysis, e.g. Tao’s book, and some functional analysis, e.g. the book by Lax.

Manifolds, differential forms, Stokes formula:


Topology:


Riemannian geometry:


After that you can read Gilbarg-Trudinger and Schoen-Yau.

Geometry of fiber bundles: