Geometry of Curves and Surfaces

Weiyi Zhang

Mathematics Institute, University of Warwick

September 18, 2016

Contents

Chapter 1

Curves

1.1 Course description

Instructor: Weiyi Zhang
Email: weiyi.zhang@warwick.ac.uk
Webpage: http://homepages.warwick.ac.uk/staff/Weiyi.Zhang/
Lecture time/room:

Tuesday 10am - 11am MS.04
Thursday 12pm - 1pm MS.03
Monday 9am - 10am MS.05

Support class:

Thursday 2pm - 3pm B3.01

TA: Louis Bonthrone, L.Bonthrone@warwick.ac.uk

Reference books:

- John McCleary, "Geometry from a differentiable viewpoint", CUP 1994.
- Dirk J. Struik, "Lectures on classical differential geometry", Addison-Wesley 1950
- Manfredo P. do Carmo, "Differential geometry of curves and surfaces", Prentice-Hall 1976
- Barrett O'Neill, "Elementary differential geometry", Academic Press 1966
- Sebastian Montiel, Antonio Ros, "Curves and surfaces", American Mathematical Society 1998
- Alfred Gray, "Modern differential geometry of curves and surfaces", CRC Press 1993

- *Course Notes*, available on my webpage

I also make use of the following two excellence course notes:

- Brian Bowditch, "Geometry of curves and surfaces", University of Warwick, available at

  http://homepages.warwick.ac.uk/∼masgak/cas/course.html

- Nigel Hitchin, "The geometry of surfaces", University of Oxford, available at:

  http://people.maths.ox.ac.uk/∼hitchin/hitchinnotes/hitchinnotes.html

The following book has a lot of exercises with solutions available:

- Andrew Pressley, "Elementary Differential Geometry", 2nd Ed, Springer.

Prerequisites: MA 225 Differentiation, MA231 Vector Analysis and some basic notions from topology, namely open and closed sets, continuity etc. In practice these will only be applied to subsets of $\mathbb{R}^n$.

Contents: This course is about the analysis of curves and surfaces in 2- and 3-space using the tools of calculus and linear algebra. Emphasis will be placed on developing intuitions and learning to use calculations to verify and prove theorems We will cover

- local and global properties of curves: curvature, torsion, Frenet-Serret equations, and some global theorems;
- local and global theory of surfaces: local parameters, curves on surfaces, geodesic and normal curvature, first and second fundamental form, Gaussian and mean curvature, minimal surfaces, and Gauss-Bonnet theorem etc..

1.1.1 A bit preparation: Differentiation

Definition 1.1.1. *Let U be an open set in $\mathbb{R}^n$, and $f : U \to \mathbb{R}$ a continuous function. The function f is smooth (or C^∞) if it has derivatives of any order.*

Note that not all smooth functions are analytic. For example, the function

$$f(x) = \begin{cases} 0, & x \leq 0 \\ e^{-\frac{1}{x}}, & x > 0 \end{cases}$$

is a smooth function defined on $\mathbb{R}$ but is not analytic at $x = 0$. (Check this!)

Now let U be an open set in $\mathbb{R}^n$ and V be an open set in $\mathbb{R}^m$. Let $f = (f^1, \cdots, f^m) : U \to V$ be a continuous map. We say f is smooth if each component f^i, $1 \leq i \leq m$, is a smooth function.

Definition 1.1.2. *The* differential *of f, df, assigns to each point $x \in U$ a linear map $df_x : \mathbb{R}^n \to \mathbb{R}^m$ whose matrix is the* Jacobian matrix *of f at x,*

$$df_x = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x) & \cdots & \frac{\partial f^1}{\partial x^n}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial f^m}{\partial x^1}(x) & \cdots & \frac{\partial f^m}{\partial x^n}(x) \end{pmatrix}.$$

Now, we are ready to introduce the notion of *diffeomorphism*.

Definition 1.1.3. *A smooth map $f : U \to V$ is a* diffeomorphism *if f is one-to-one and onto, and $f^{-1} : V \to U$ is also smooth.*

Obviously

- If $f : U \to V$ is a diffeomorphism, so is f^{-1}.
- If $f : U \to V$ and $g : V \to W$ are diffeomorphisms, so is $g \circ f$.

As a consequence, we get

Theorem 1.1.4. *If $f : U \to V$ is a diffeomorphism, then at each point $x \in U$, the linear map df_x is an isomorphism. In particular,* $\dim U = \dim V$.

Proof. Applying the chain rule to $f^{-1} \circ f = id_U$, and notice that the differential of the identity map $id_U : U \to U$ is the identity transformation $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$, we get

$$df^{-1}_{f(x)} \circ df_x = \mathrm{Id}_{\mathbb{R}^n}.$$

The same argument applies to $f \circ f^{-1}$, which yields

$$df_x \circ df^{-1}_{f(x)} = \mathrm{Id}_{\mathbb{R}^m}.$$

By basic linear algebra, we conclude that $m = n$ and that df_x is an isomorphism. □

The inverse of the previous theorem is not true. For example, we consider the map

$$f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2 \setminus \{0\}, \quad (x^1, x^2) \mapsto ((x^1)^2 - (x^2)^2, 2x^1x^2).$$

Then at each point $x \in \mathbb{R}^2 \setminus \{0\}$, df_x is an isomorphism. However, f is not invertible since $f(x) = f(-x)$. (What is the map f if we identify $\mathbb{R}^2$ with $\mathbb{C}$?)

The inverse function theorem is a partial inverse of the previous theorem, which claims that an isomorphism in the linear category implies a *local* diffeomorphism in the differentiable category.

Theorem 1.1.5 (Inverse Function Theorem)**.** *Let $U \subset \mathbb{R}^n$ be an open set, $p \in U$ and $f : U \to \mathbb{R}^n$. If the Jacobian df_p is invertible at p, then there exists a neighbourhood U_p of p and a neighbourhood $V_{f(p)}$ of $f(p)$ such that*

$$f|U_p : U_p \to V_{f(p)}$$

is a diffeomorphism.

1.2 Methods of describing a curve

There are different ways to describe a curve.

1.2.1 Fixed coordinates

Here, the coordinates could be chosen as Cartesian, polar and spherical etc. **(a).** As a graph of explicitly given curves $y = f(x)$.

Example 1.2.1. *A parabola:* $y = x^2$*; A spiral:* $r = \theta$*.*

(b). Implicitly given curves

A plane curve (i.e. a curve in $\mathbb{R}^2$) could be given as $f(x, y) = 0$; A space curve (i.e. a curve in $\mathbb{R}^3$) could be given as $f_1(x, y, z) = 0, f_2(x, y, z) = 0$.

Example 1.2.2. *A unit circle could be given as* $x^2 + y^2 = 1$*. It could also be expressed as* $x^2 + y^2 + z^2 = 1, z = 0$*.*

1.2.2 Moving frames: parametrized curves

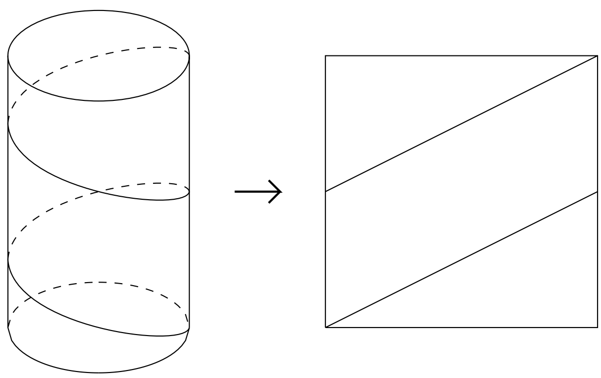
Definition 1.2.3. *A parametrized curve in* $\mathbb{R}^n$ *is a map* $\gamma : I \to \mathbb{R}^n$ *of an open interval* $I = (a, b)$*.*

Example 1.2.4. *Parabola:* $\gamma(t) = (t, t^2)$*,* $t \in (-\infty, \infty)$*;*

Circle: $\gamma(t) = (a \cos t, a \sin t)$*,* $-\epsilon < t < 2\pi + \epsilon, \epsilon > 0$*;*

Ellipse: $\gamma(t) = (a \cos t, b \sin t)$*,* $-\epsilon < t < 2\pi + \epsilon, \epsilon > 0$*;*

Helix: $\gamma(t) = (a \cos t, a \sin t, bt)$*,* $t \in (-\infty, \infty)$*.*



Why this description is called "moving frame" in the title? Roughly speaking, for a plane curve, (tangent vector= $\dot{\gamma}(t)$, normal vector) forms a coordinate, which changes as t varies. More precise explanation will be given in next section.

Like us, we could orient the world using (Front, Left) system and take ourselves as centres.

Remark 1.2.5. *1. Parametrizations are not unique. Parabola* $\gamma(t) = (t, t^2)$ *could also be parametrized as* (t^3, t^6), $(2t, 4t^2)$ *and other ways.*

2. All parametrized curves studied in this course are smooth.

3. If the tangent vector of a parametrized curve is constant, the image of the curve is part of a straight line. (proof: exercise)

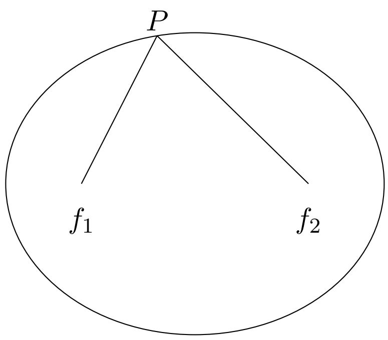
1.2.3 Intrinsic way(coordinate free)

The previous ways both involve coordinates. They are thus called extrinsic ways.

But for most cases, we only care the shapes of objects, but not the locations. Such a description will be called intrinsic.

Example 1.2.6. *Circle: set of points (on a plane) with given distance a to a point.*

Ellipse: Set of points such that the sum of the distances to two given points are fixed.



$$\overline{f_1P} + \overline{f_2P} \equiv C$$

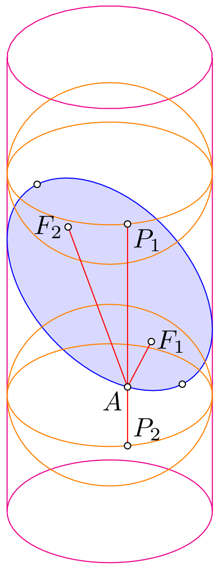
Similar definition for hyperbola and parabola.

Drawback: we do not have sufficient tools to further study our objects from intrinsic viewpoints. That is the actually reason there are no big breakthrough for geometry from Archimedes to Newton.

Goal of this course: use extrinsic ways (mainly parametrized way) to prove intrinsic results. Thus we could use tools we learnt in the last two years: analysis, linear algebra, $\cdots$.

However, there are some elegant results could be proved intrinsically! The next example is from "geometry and imagination" by Hilbert and Cohn-Vossen.

We want to prove the intersection of a cylinder and a plane is an ellipse if the curve is closed. To achieve that, we use two identical balls with radii equal to that of the cylinder. Then move them: one from top and the other from bottom, until they hit the plane at F_1 and F_2. We claim these two points are foci of the ellipse. Let A be any point on the intersection of the plane and the cylinder. Draw a vertical line through A such that it intersects the top ball at P_1 and the bottom ball at P_2. Since AP_1 and AF_1 and two tangents of the same ball from the same point, they have equal length. Similarly $AP_2 = AF_2$. But $AP_1 + AP_2$ is the distance between the centres of the two balls, thus a constant. So $AF_1 + AF_2$ is a constant, which proves our claim.



Exercise: prove similar results for a cone $x^2 + y^2 = z^2$. (Similar arguments, but now you have to deal with different cases: ellipse, parabola, hyperbola and pair of straight lines.)

1.3 Curves in $\mathbb{R}^n$: Arclength Parametrization

To use parametrization to study curves, we should start with building moving frames. An orthonormal basis is usually easier to play with.

First component of the moving frame is the tangent direction $\dot{\gamma}(t)$. This step works for any curve in $\mathbb{R}^n$. So the first step to build an orthonormal frame would be

$$||\dot{\gamma}(t)|| \equiv 1$$

for some parametrization and for all $t \in (\alpha, \beta)$. These curves are called it unit-speed curves.

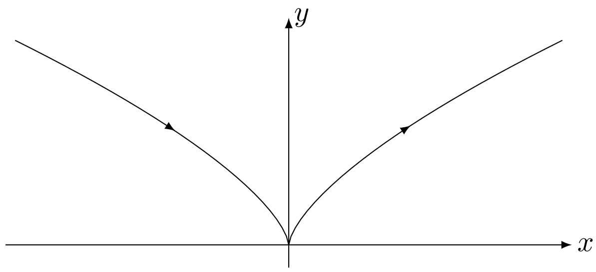
Idea: reparametrization or change variables.

Definition 1.3.1. *A parametrized curve $\tilde{\gamma} : J \to \mathbb{R}^n$ is a reparametrization of $\gamma : I \to \mathbb{R}^n$ if there is a diffeomorphism $\phi : J \to I$ such that $\tilde{\gamma}(\tilde{t}) = \gamma \circ \phi(\tilde{t})$ for all $\tilde{t} \in J$.*

The first obtacle to find unit-speed parametrization is $\dot{\gamma}(t) = 0$ at some point.

Definition 1.3.2. *A point $\gamma(t)$ is called a regular point if $\dot{\gamma}(t) \neq 0$; otherwise $\gamma(t)$ is a singular point of γ. A curve is regular if all points are regular.*

Example 1.3.3. *$\gamma : \mathbb{R} \to \mathbb{R}^2$ given by $\gamma(t) = (t^3, t^2), t \in \mathbb{R}$ is not regular since $\dot{\gamma}(0) = (0, 0)$.*



Remark 1.3.4. *Any reparametrization of a regular curve is regular (exercise of Chain rule).*

Then why is the next example?

Example 1.3.5. *$\gamma(t) = (t, t^2)$ is regular, but another parametrization $\tilde{\gamma}(t) = (t^3, t^6)$ is not regular.*

Answer: $\tilde{\gamma}$ is not a reparametrization of γ.
Reason: The bijection $\phi(t) = t^3$ is not a diffeomorphism (ϕ^{-1} is not smooth at 0).

Later on, a curve is a parametrized smooth regular curve.

To find reparametrization s such that $||\tilde{\gamma}_s(s)|| = 1$, we only need $\frac{ds}{dt} = ||\gamma_t(t)||$. In other word,

$$s(t) = \int_{t_0}^{t} ||\gamma_t(t)||dt = \int_{t_0}^{t} \sqrt{\dot{x}_1^2(t) + \cdots + \dot{x}_n^2(t)}dt.$$

Hence $||\tilde{\gamma}_s|| = ||\gamma_t \cdot \frac{dt}{ds}|| = 1$.

It need a proof that s is indeed a reparametrization for a regular curve.Only thing need to verify is the following:

Proposition 1.3.6. *If $\gamma(t)$ is a regular curve, then $s(t)$ is a diffeomorphism.*

Proof. We first show that s is a smooth function of t.

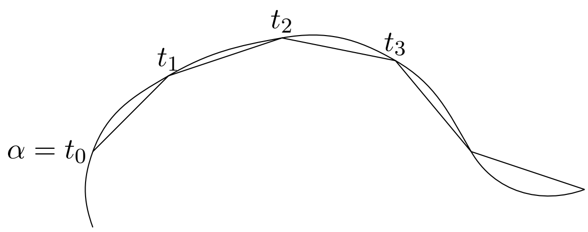
$\frac{ds}{dt} = f(\dot{x}_1^2(t) + \cdots + \dot{x}_n^2(t))$ if $\gamma(t) = (x_1(t), \cdots, x_n(t))$ and $f = \sqrt{x}$. Hence $f^{(n)}(x) = c_n x^{\frac{1}{2}-n}$ for some positive constant c_n. Since $\dot{\gamma} \neq 0$, $\frac{ds}{dt}$ is smooth. Hence s is smooth.

The conclusion s is a diffeomorphism follows from IFT (inverse function theorem) and the fact $\frac{ds}{dt} > 0$ by above calculation. Or more directly,every monotone function has its inverse. □

Definition 1.3.7. *We refer to s as arc length and to $\tilde{\gamma}$ as the arc-length (or unit-speed) reparametrization of γ.*

Remark 1.3.8. *Let $\mathcal{P} = \{\alpha = t_0 < t_1 < \cdots < t_k = \beta$ be a partition of (α, β), and $l(\gamma, \mathcal{P}) = \sum_{i=1}^{k} ||\gamma(t_i) - \gamma(t_{i-1})||$. Then*

$$s(t) = \sup\{l(\gamma, \mathcal{P}) : a \ partition \ \mathcal{P}\}.$$



The next lemma ends this section.

Lemma 1.3.9. *Suppose $\overrightarrow{f(t)} : I = (\alpha, \beta) \to \mathbb{R}^n$ are differentiable. Then $||\overrightarrow{f(t)}|| = const$ if and only if $\overrightarrow{f(t)} \cdot \overrightarrow{f'(t)} = 0$ for all t.*

Proof. The function $\overrightarrow{f(t)} \cdot \overrightarrow{f(t)}$ is a constant if and only if $\overrightarrow{f(t)} \cdot \overrightarrow{f'(t)} + \overrightarrow{f'(t)} \cdot \overrightarrow{f(t)} = 2\overrightarrow{f(t)} \cdot \overrightarrow{f'(t)} = 0$. □

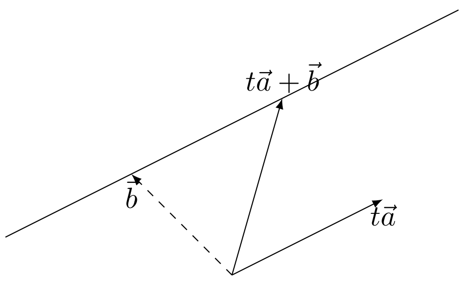
Corollary 1.3.10. *If γ is unit-speed, then $\ddot{\gamma}$ is zero or perpendicular to the tangent vector $\dot{\gamma}$.*

1.4 Curvature

From now on, let us focus on space curves, i.e. curves $\gamma(t) : I \to \mathbb{R}^3$.

Curvature measures how far a curve is different from a straight line (how far it bends from a straight line).

Recall a straight line has parametrized form $\gamma(t) = t\vec{a} + \vec{b}$. Here $\vec{a}$ is the direction and $\vec{b} = \gamma(0)$.



It has vanishing higher derivatives. Its Taylor expansion

$$\gamma(t + \Delta t) = \gamma(t) + \dot{\gamma}(t)\Delta t + \cdots = \gamma(t) + \vec{a} \cdot \Delta t$$

only has 2 terms.

For a general curve, using unit-speed parametrization,

$$\gamma(t+\Delta t) = \gamma(t) + \dot{\gamma}(t)\Delta t + \frac{1}{2}\ddot{\gamma}(t)(\Delta t)^2 + R(t)$$

with $\lim_{\Delta t \to 0} \frac{R(t)}{(\Delta t)^2} = 0$. Since $\ddot{\gamma}(t) \perp \dot{\gamma}(t)$ when $\ddot{\gamma}(t)$ is nonzero, $||\ddot{\gamma}(t)||$ measures how far γ is deviated from its tangent line at $\gamma(t)$.

1^{st} definition of curvature: extrinsic one.

Definition 1.4.1. *If γ is a unit-speed curve with parameter t, its curvature $\kappa(t)$ at $\gamma(t)$ is defined to be $||\ddot{\gamma}(t)||$.*

2^{nd} definition of curvature: more extrinsic.

Proposition 1.4.2. *Let $\gamma(t)$ be a regular curve. Then its curvature is*

$$\kappa = \frac{||\gamma_{tt} \times \gamma_t||}{||\gamma_t||^3},$$

where $\times$ is the vector (or cross) product.

Proof. $\gamma_s = \gamma_t \cdot \frac{dt}{ds} = \frac{\gamma_t}{||\gamma_t||}$, since $\frac{ds}{dt} = ||\gamma_t||$. The second derivative

$$\gamma_{ss} = \gamma_{tt} \cdot (\frac{dt}{ds})^2 + \gamma_t \cdot \frac{d^2t}{ds^2} = \frac{\gamma_{tt}}{||\gamma_t||^2} - \gamma_t \frac{\gamma_{tt} \cdot \gamma_t}{||\gamma_t||^4},$$

since

$$\frac{d^2t}{ds^2} = \frac{d}{ds}\frac{1}{||\gamma_t||} = \frac{d\dfrac{1}{(\gamma_t \cdot \gamma_t)^{\frac{1}{2}}}}{dt} \cdot \frac{dt}{ds} = (-\frac{1}{2(\gamma_t \cdot \gamma_t)^{\frac{3}{2}}} \cdot 2\gamma_t \cdot \gamma_{tt})\frac{1}{||\gamma_t||} = -\frac{\gamma_{tt} \cdot \gamma_t}{||\gamma_t||^4}.$$

So

$$\kappa^2 = ||\gamma_{ss}||^2 = \frac{||\gamma_{tt}||^2 \dot{|}|\gamma_t||^2 - (\gamma_{tt} \cdot \gamma_t)^2}{||\gamma_t||^6} = \frac{||\gamma_{tt} \times \gamma_t||^2}{||\gamma_t||^6}$$

Thus $\kappa = \frac{||\gamma_{tt} \times \gamma_t||}{||\gamma_t||^3}$.

Here we make use of the relation

$$||\vec{a}||^2 \cdot ||\vec{b}||^2 = (\vec{a} \cdot \vec{b})^2 + ||\vec{a} \times \vec{b}||^2$$

because

$$\vec{a} \cdot \vec{b} = ||\vec{a}|| \cdot ||\vec{b}|| \cos\theta, ||\vec{a} \times \vec{b}|| = ||\vec{a}|| \cdot ||\vec{b}|| \sin\theta.$$

□

Corollary 1.4.3. *1. If γ is a plane curve, i.e. $\gamma(t) = (x(t), y(t))$, then*

$$\kappa = \frac{|\ddot{y}\dot{x} - \ddot{x}\dot{y}|}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}}.$$

2. If γ is a graph $y = f(x)$,

$$\kappa = \frac{|f''(x)|}{(1 + (f'(x))^2)^{\frac{3}{2}}}.$$

Example 1.4.4. *1. A curve is a (part of) straightline if and only if its curvature is everywhere zero.*

2. Look at a circle in $\mathbb{R}^2$: centred at (x_0, y_0) and of radius R.

A unit-speed parametrization is $\gamma(t) = (x_0 + R \cdot \cos \frac{t}{R}, y_0 + R \sin \frac{t}{R})$. We calculate $\ddot{\gamma}(t) = (-\frac{1}{R} \cos \frac{t}{R}, -\frac{1}{R} \sin \frac{t}{R})$. Hence $\kappa = ||\ddot{\gamma}(t)|| = \frac{1}{R}$.

3. Helix $\gamma(\theta) = (a \cos \theta, a \sin \theta, b\theta), \theta \in \mathbb{R}$.

$\dot{\gamma}(\theta) = (-a \sin \theta, a \cos \theta, b)$, so $||\dot{\gamma}(\theta)|| = \sqrt{a^2 + b^2}$.

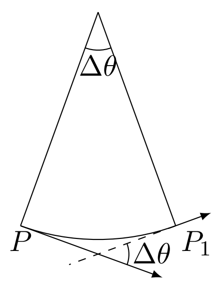
$\ddot{\gamma} = (-a \cos \theta, -a \sin \theta, 0)$, $\dot{\gamma} \times \ddot{\gamma} = (ab \sin \theta, -ab \cos \theta, a^2)$. Hence

$$\kappa = \frac{(a^2 b^2 + a^4)^{\frac{1}{2}}}{(a^2 + b^2)^{\frac{3}{2}}} = \frac{|a|}{a^2 + b^2}.$$

When $b = 0$, this is a circle of radius $|a|$, and $\kappa = \frac{1}{|a|}$ (coincides with previous calculation).

Intrinsic Viewpoint:

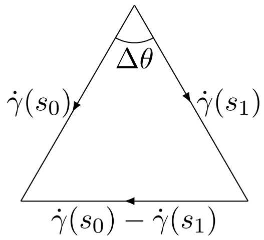
For circle, $\kappa = \frac{1}{R} = \frac{\Delta\theta}{\Delta s}$, where $\Delta\theta$ could be understood as the difference of angles between tangent vectors at s and $s + \Delta s$.



In general, it may not be true, since different Δs give different values. We assume $\gamma(s)$ is a unit-speed curve. Let $P = \gamma(s_0)$ and $P_1 = \gamma(s_1)$, then $\Delta s = |s_1 - s_0|$. And $\Delta\theta$ is the angle between tangent vectors $\dot{\gamma}(s_0)$ and $\dot{\gamma}(s_1)$.

Theorem 1.4.5. $\kappa(s_0) = \lim_{\Delta s \to 0} \frac{\Delta\theta}{\Delta s} = \lim_{P_1 \to P} \frac{\Delta\theta}{\Delta s}$.

Proof. $||\dot\gamma(s_0)|| = ||\dot\gamma(s_1)|| = 1$ implies $2\sin\frac{\Delta\theta}{2} = ||\dot\gamma(s_0) - \dot\gamma(s_1)||$.



Thus

$$\lim_{\Delta s \to 0} \frac{\Delta\theta}{\Delta s} = \lim_{\Delta\theta \to 0} \frac{\Delta\theta}{2\sin\frac{\Delta\theta}{2}} \cdot \lim_{\Delta s \to 0} \frac{||\dot\gamma(s_0) - \dot\gamma(s_1)||}{\Delta s} = ||\ddot\gamma(s_0)|| = \kappa(s_0).$$

□

1.5 Orthonormal frame: Frenet-Serret equations

Now, let $\gamma(s)$ be a unit-speed curve in $\mathbb{R}^3$. We are ready to build orthonormal moving frame for it.

Denote $\mathbf{t} = \dot\gamma(s)$ be the unit tangent vector.

If $\kappa(s) \neq 0$, by Corollary 1.3.10, $\ddot\gamma(s) \perp \dot\gamma(s)$. We define the *principal normal* at $\gamma(s)$ be

$$\mathbf{n}(s) = \frac{1}{\kappa(s)}\mathbf{t}'(s).$$

We have $||\mathbf{n}|| = 1$ and $\mathbf{t} \cdot \mathbf{n} = 0$.

Finally, we define bi-normal vector

$$\mathbf{b}(s) = \mathbf{t} \times \mathbf{n}.$$

It is a unit vector perpendicular to both $\mathbf{t}$ and $\mathbf{n}$.

To summarize: $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ is an orthonormal basis of $\mathbb{R}^3$ and it is right-handed (i.e. $\mathbf{b} = \mathbf{t} \times \mathbf{n}, \mathbf{n} = \mathbf{b} \times \mathbf{t}, \mathbf{t} = \mathbf{n} \times \mathbf{b}$).

There are standard names for planes spanned by any two of them:

- osculating plane: by $\mathbf{t}$ and $\mathbf{n}$;
- rectifying plane: by $\mathbf{b}$ and $\mathbf{t}$;
- normal plane: by $\mathbf{n}$ and $\mathbf{b}$.

$||\mathbf{b}'(s)||$ measures the rate of change of angles Θ between osculating planes, which is the same as the changing of binormal vector as well. Similar reasoning as in Theorem 1.4.5,

$$||\mathbf{b}'(s)|| = \lim_{\Delta s \to 0} \frac{\Delta\Theta}{\Delta s}.$$

Notice

$$\mathbf{b}'(s) = \mathbf{t}' \times \mathbf{n} + \mathbf{t} \times \mathbf{n}' = \mathbf{t} \times \mathbf{n}'$$

since $\mathbf{t}' \times \mathbf{n} = \kappa\mathbf{n} \times \mathbf{n} = 0$. This implies $\mathbf{b}' \perp \mathbf{t}$.

On the other hand, $||\mathbf{b}|| = 1$ implies $\mathbf{b}' \perp \mathbf{b}$. Hence $\mathbf{b}' \parallel \mathbf{n}$. We define

$$\mathbf{b}' = -\tau\mathbf{n}.$$

Here τ is called *torsion*[1] of the curve.

Remark 1.5.1. *Please notice that since* $\mathbf{n}(s)$ *is defined only when* $\kappa(s) \neq 0$, $\tau(s)$ *is so as well.*

There is a formula of τ for an arbitrary parametrization as Proposition 1.4.2 for κ.

Proposition 1.5.2. *Let* $\gamma(t)$ *be a regular curve in* $\mathbb{R}^3$ *with* $\kappa(t) \neq 0$, *then*

$$\tau(t) = \frac{(\dot\gamma(t) \times \ddot\gamma(t)) \cdot \dddot\gamma(t)}{||\dot\gamma(t) \times \ddot\gamma(t)||^2}$$

The proof is a tedious calculation and left as an exercise.

Example 1.5.3. *Planar curve* $\gamma(t)$. *There is a constant vector* $\mathbf{a}$ *such that* $\gamma(t) \cdot \mathbf{a}$ *is a constant. Without loss, we could assume the parametrization is unit-speed. So* $\mathbf{t} \cdot \mathbf{a} = 0$ *and* $\mathbf{n} \cdot \mathbf{a} = 0$. *Hence* $\mathbf{t}$ *and* $\mathbf{n}$ *is perpendicular to* $\mathbf{a}$, *thus parallel to the plane. Finally,* $\mathbf{b} = \mathbf{t} \times \mathbf{n}$ *is a unit vector orthogonal to the plane and thus a constant vector. Hence* $\tau = 0$.

Example 1.5.4. *Let* $\gamma(\theta) = (a\cos\theta, a\sin\theta, b\theta)$.

$$\dot\gamma(\theta) = (-a\sin\theta, a\cos\theta, b), \ddot\gamma(\theta) = (-a\cos\theta, -a\sin\theta, 0), \dddot\gamma(\theta) = (a\sin\theta, -a\cos\theta, 0).$$

So $\dot\gamma(\theta) \times \ddot\gamma(\theta) = (ab\sin\theta, -ab\cos\theta, a^2)$ *and*

$$||\dot\gamma \times \ddot\gamma||^2 = a^2(a^2 + b^2), (\dot\gamma \times \ddot\gamma) \cdot \dddot\gamma = a^2 b.$$

Hence, $\tau = \frac{b}{a^2+b^2}$.

[1] In some books, torsion is defined as $-\tau$ in our notation.

Exercise: Use unit-speed parametrization to calculate the above example again.

Let us come back to the unit-speed parametrization. We have $\mathbf{n} = \mathbf{b} \times \mathbf{t}$, so

$$\mathbf{n}'(s) = \mathbf{b}' \times \mathbf{t} + \mathbf{b} \times \mathbf{t}' = -\tau\mathbf{n} \times \mathbf{t} + \kappa\mathbf{b} \times \mathbf{n} = -\kappa\mathbf{t} + \tau\mathbf{b}.$$

To summarize, we have the following set of *Frenet-Serret equations* (when $\kappa(s) \neq 0$!!):

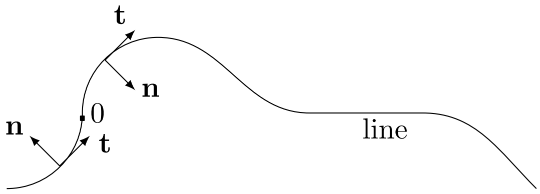
$$\begin{cases} \mathbf{t}' = & \kappa\mathbf{n} & \\ \mathbf{n}' = -\kappa\mathbf{t} & & +\tau\mathbf{b} \\ \mathbf{b}' = & -\tau\mathbf{n} & \end{cases}$$

In other writing,

$$\begin{pmatrix} \mathbf{t} \\ \mathbf{n} \\ \mathbf{b} \end{pmatrix}' = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{t} \\ \mathbf{n} \\ \mathbf{b} \end{pmatrix}$$

The matrix is skew-symmetric.

What if $\kappa = 0$ at some points? $\mathbf{n}$ is not well-defined at these points.



But there is a way to resolve this issue for plane curves.

In summary, curvature measures how far the curve is from a line; torsion measures how far it is from its osculating plane, or how far the curve is from a plane curve.

1.6 Plane curves

The trick to resolve the issue mentioned in the last section for plane curve is to define the *signed unit normal* $\mathbf{n}_s$ as the unit vector obtained by rotating $\mathbf{t}$ counter-clockwise $\frac{\pi}{2}$.

The signed curvature κ_s is defined as

$$\ddot{\gamma} = \mathbf{t}' = \kappa_s \mathbf{n}_s.$$

The relation with curvature is $\kappa = ||\ddot{\gamma}|| = |\kappa_s|$.

Look at Figure 1.1 and 1.2. The intrinsic viewpoint of κ_s: change of angle for tangent vectors. How to define the angle?

Let $\gamma(s) = (x(s), y(s))$ be a unit-speed plane curve. Let us first define it locally. Let $\phi(s) \in (0, 2\pi)$ be the angle that $\mathbf{t}(s)$ makes with x-axis.

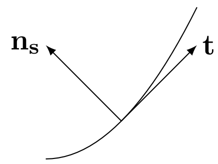


Figure 1.1: $\kappa_s > 0$, angle increases

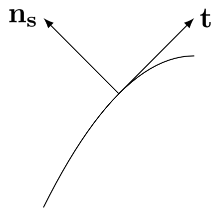


Figure 1.2: $\kappa_s < 0$, angle decreases

So $\phi(s) = \arctan \frac{y'(s)}{x'(s)}$ and $(x'(s), y'(s)) = (\cos\phi(s), \sin\phi(s))$. It is locally well-defined.

$$\ddot{\gamma} = \frac{d\mathbf{t}}{ds} = \frac{d\phi}{ds}(-\sin\phi, \cos\phi) = \frac{d\phi}{ds}\mathbf{n}_s.$$

So

$$\kappa_s(s) = \frac{d\phi}{ds} = x'y'' - x''y'.$$

Motivated by last calculation, we define it globally:

$$\phi(s) = \int_{s_0}^{s} \kappa_s(s)ds.$$

This is called the *turning angle.* Up to constant, it is just the previous defined local version.

It is particularly interesting to study the total turning angle for a closed curve.

Definition 1.6.1. *A smooth curve $\gamma : \mathbb{R} \to \mathbb{R}^n$ is called a closed curve if there is $T \neq 0$ such that $\gamma(t + T) = \gamma(t)$ for all $t \in \mathbb{R}$.*

The minimal such T is called *period.* Later we may write $\gamma : [0, T] \to \mathbb{R}^n$ with $\gamma(0) = \gamma(T)$ to represent a closed curve.

A *simple closed curve* is a closed curve with no self-intersection, i.e if $|t_1 - t_2| < T$, then $\gamma(t_1) \neq \gamma(t_2)$. A simple closed curve is also called Jordan curve in some literature. We have the following intuitively clear but hard to prove theorem.

Theorem 1.6.2 (Smooth Schoenflies)**.** *For any simple closed curve γ, there is a diffeomorphism $f : \mathbb{R}^2 \to \mathbb{R}^2$ sending the unit circle to γ.*

Hence, we can define the interior (resp. exterior) of the Jordan curve γ as the bounded (resp. unbounded) region with boundary γ.

For closed curves, the total signed curvature

$$\int_0^T \kappa_s(s)ds = \phi(T) - \phi(0) = 2\pi I,$$

where I is an integer called *rotation index.*

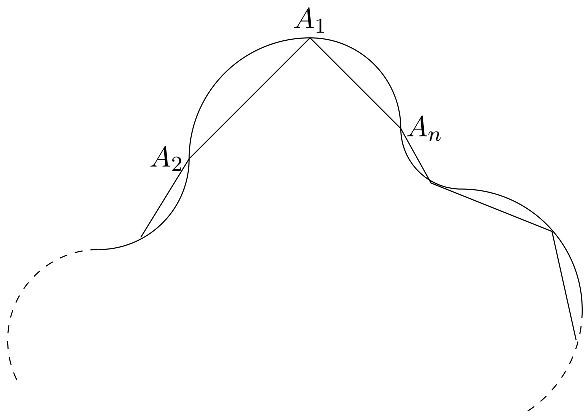
Example 1.6.3. *A counter clockwise circle has rotation index 1. A clockwise ellipse has rotation index* -1*.*

What is the rotation index of figure 8?

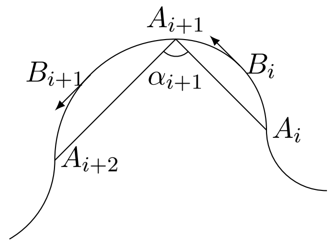
Theorem 1.6.4 (Hopf's Umlaufsatz)**.** *The rotation index of a simple closed curve is* ± 1*. (sign depends on the orientation)*

This will be a corollary of Gauss-Bonnet theorem. But there is simple proof which also motivates the proof of Gauss-Bonnet.

Proof. We denote $\gamma_i = \widetilde{A_iA_{i+1}}$ as part of the curve γ. Assume the total integral curvature of every arc γ_i is less than π, and no self-intersection for the polygon $A_1 \cdots A_n A_{n+1}$ with $A_{n+1} = A_1$.



On each γ_i, we choose B_i such that $\mathbf{t}(B_i) \parallel A_iA_{i+1}$.



Then by the local definition of turning angle

$$\int_{\widetilde{B_iB_{i+1}}} \kappa_s(s)ds = \phi(B_{i+1}) - \phi(B_i) = \pi - \alpha_{i+1}.$$

So

$$\int_0^T \kappa_s(s)ds = \int_\gamma \kappa_s ds = \sum_{i=1}^n \int_{\widetilde{B_iB_{i+1}}} \kappa_s ds = \sum_{i=1}^n (\pi - \alpha_{i+1}) = n\pi - \sum_{i=1}^n \alpha_i = 2\pi$$

□

One could compare this result to

Theorem 1.6.5 (Fenchel). *The total curvature of any closed space curve is at least 2π, i.e $\int \kappa ds \geq 2\pi$. The equality holds if and only if the curve is a convex planar curve.*

1.7 More results for space curves

We have shown a space curve is a straight line if and only if its curvature is everywhere 0.

Proposition 1.7.1. *A space curve with nowhere vanishing curvature is planar if and only if its torsion is everywhere* 0.

Proof. Take unit-speed parametrization. We have shown the "only if" part.

On the other hand, if $\tau = 0$, then $\mathbf{b}' = 0$ and so $\mathbf{b}$ is a constant vector. By calculation

$$\frac{d}{ds}(\gamma \cdot \mathbf{b}) = \dot{\gamma} \cdot \mathbf{b} = \mathbf{t} \cdot \mathbf{b} = 0.$$

So $\gamma \cdot \mathbf{b}$ is a constant C, which implies γ is contained in the plane $(\cdot) \cdot \mathbf{b} = C$. □

Proposition 1.7.2. *The only planar curves with non-zero constant curvature are (part of) circles.*

Proof. We have shown a circle of radius R has constant curvature $\kappa = \frac{1}{R}$. Now suppose a planar curve γ (thus $\tau = 0$) has constant curvature κ.

$$\frac{d}{ds}(\gamma(s) + \frac{1}{\kappa}\mathbf{n}) = \mathbf{t} + \frac{1}{\kappa}\mathbf{n}' = \mathbf{t} - \mathbf{t} + \frac{\tau}{\kappa}\mathbf{b} = 0.$$

Hence $\gamma + \frac{1}{\kappa}\mathbf{n}$ is a constant vector $\mathbf{a}$. So $||\gamma - \mathbf{a}|| = \frac{1}{\kappa}$. This is a circle with centre $\mathbf{a}$ and radius $\frac{1}{\kappa}$. □

Especially, a space curve with constant κ and $\tau = 0$ is a part of circle.

1.7.1 Taylor expansion of a curve

Frenet-Serret equations gives the local picture of space curves. Let us look at the Taylor expansion of a space curve

$$\gamma(s) = \gamma(0) + s\dot{\gamma}(0) + \frac{s^2}{2}\ddot{\gamma}(0) + \frac{s^3}{6}\dddot{\gamma}(0) + R$$

where $\lim_{s\to 0} \frac{R}{s^3} = 0$. We know

$$\dot{\gamma}(0) = \mathbf{t}(0), \ddot{\gamma}(0) = \kappa(0)\mathbf{n}(0)$$

and

$$\dddot{\gamma} = (\kappa \mathbf{n})'(0) = \kappa_0' \mathbf{n}(0) + \kappa_0(-\kappa_0 \mathbf{t}(0) + \tau_0 \mathbf{b}(0)).$$

So

$$\gamma(s)-\gamma(0) = (s-\frac{\kappa_0^2}{6}s^3+\cdots)\mathbf{t}(0)+(\frac{s^2\kappa_0}{2}+\frac{s^3\kappa_0'}{6}+\cdots)\mathbf{n}(0)+(\frac{1}{6}\kappa_0\tau_0 s^3+\cdots)\mathbf{b}(0).$$

Here $\lim_{s\to 0} \frac{\cdots}{s^3} = 0$.

See the local pictures in next page. Notice the sign of τ will affect the projection in rectifying and normal planes, thus the whole local picture.

Exercise: Draw the local pictures when $\tau < 0$.

1.7.2 Fundamental Theorem of the local theory of curves

Theorem 1.7.3. *Given two smooth functions κ and τ with $\kappa > 0$ everywhere, there is a unit-speed curve in $\mathbb{R}^3$ whose curvature is κ and torsion is τ. The curve is unique up to a rigid motion.*

Here two curves are related by a rigid motion if $\tilde{\gamma} = A \circ \gamma + \mathbf{c}$ where A is a orthogonal linear map of $\mathbb{R}^3$ with $\det A > 0$ and $\mathbf{c}$ is a vector. In other words, they are related by a composition of a translation and a rotation.

We omit the proof, which is an application of the existence and uniqueness theorem of linear systems. The 2D version is problem 4 in Example sheet 1.

1.8 Isoperimetric Inequality

We prove the famous Isoperimetric inequality here.

Theorem 1.8.1. *If a simple closed plane curve γ has length L and encloses area A, then*

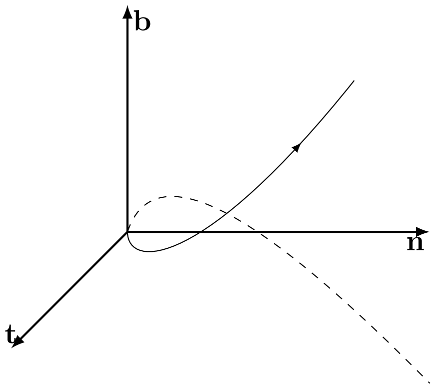
$$L^2 \geq 4\pi A,$$

and the equality holds if and only if γ is a circle.

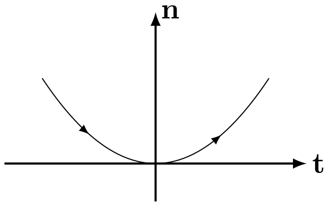
Without loss, we could assume γ is unit-speed. So our simple closed curve $\gamma(s) = (x(s), y(s))$ where $s \in [0, L]$. We first derive formulae for area A.

Lemma 1.8.2. *For any parametrization of the curve γ,*

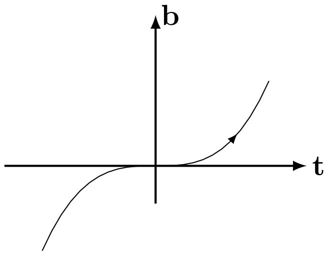
$$A = -\int_0^{L'} y(t)x'(t)dt = \int_0^{L'} x(t)y'(t)dt = \frac{1}{2}\int_0^{L'} (x(t)y'(t) - y(t)x'(t))dt$$



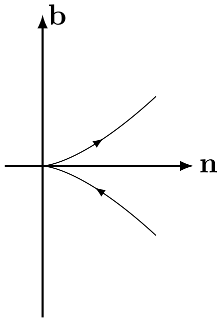
local picture of a space curve when $\tau > 0$



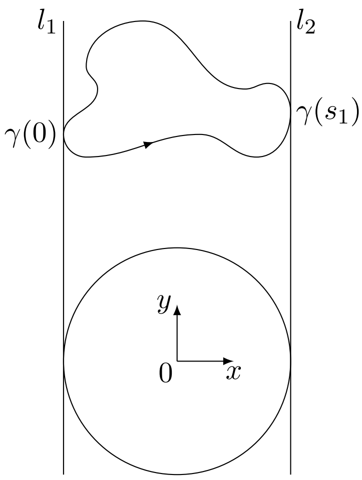
Osculating plane: $(u, \frac{\kappa_0}{2}u^2 + \frac{\kappa_0'}{6}u^3 + \cdots)$



Rectifying plane: $(u, (\frac{\kappa_0\tau_0}{6})u^3 + \cdots)$



Normal plane: $(u^2, (\frac{\sqrt{2}\tau_0}{3\sqrt{\kappa_0}})u^3 + \cdots)$



Project the curve to a circle

Proof. It is a corollary of Green's theorem:

$$\int_{int(\gamma)} (\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}) dxdy = \int_\gamma f(x,y)dx + g(x,y)dy$$

The three formulae correspond to $f = -y, g = 0$; $f = 0, g = x$; and $f = -\frac{1}{2}y, g = \frac{1}{2}x$ respectively. □

Notice in the proof, we make use of smooth Schoenflies implicitly to talk about the interior of a simple closed curve.

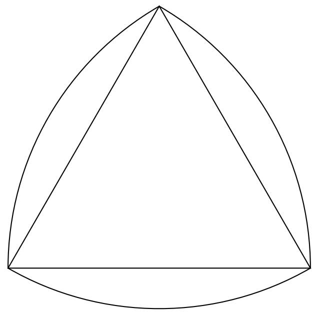
Now we prove the theorem. The idea is to "project" the curve to a circle. We choose parallel lines l_1 and l_2 tangent and enclosing γ. Draw a circle α tangent to both lines but does not meet γ. Let O be the centre of the circle. Take $\gamma(0) \in l_1$ and $\gamma(s_1) \in l_2$.

Assume the equation of $\alpha = (x(s), \bar{y}(s))$.

$$A = A(\gamma) = \int_0^L xy'ds, A(\alpha) = \pi R^2 = -\int_0^L \bar{y}x'ds$$

So

$$\begin{aligned} A + \pi R^2 &= \int_0^L (xy' - \bar{y}x')ds \\ &= \int_0^L (x, \bar{y}) \cdot (y', -x')ds \\ &\leq \int_0^L \sqrt{x^2 + \bar{y}^2}\sqrt{(x')^2 + (y')^2}ds \\ &= LR \end{aligned}$$



Reuleaux triangle

Hence

$$2\sqrt{A}\sqrt{\pi R^2} \le A + \pi R^2 \le LR.$$

Thus the isoperimetric inequality

$$L^2 \ge 4\pi A.$$

If the equality holds, $A = \pi R^2$ and $L = 2\pi R$. Especially, R is independent of the direction of l_1, l_2. Hence $(x, \bar{y}) = R(y', x')$. So $x = Ry'$. Rotate l_i for 90 degrees, we have y for x and $-x$ for y, so $y = -Rx'$. Thus

$$x^2 + y^2 = R^2((x')^2 + (y')^2) = R^2,$$

and γ is a circle.

Remark 1.8.3. *R is independent of the direction does not ensure α is a circle. We have curves of constant width which are not circle. A Reuleaux triangle is the simplest example but only piecewise smooth. One can construct smooth ones by move it outwards along the normal direction with a fixed distance for example.*

But actually smooth examples are ubiquitous: our 20p and 50p coins.

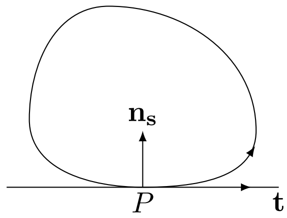
1.9 The Four Vertex Theorem

This is about a plane curve $\gamma(t) = (x(t), y(t))$, and its vertex:

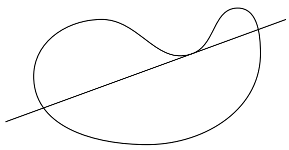
Definition 1.9.1. *A vertex of a plane curve $\gamma(t)$ is a point where its signed curvature κ_s has a critical point, i.e. where $\frac{d\kappa_s}{dt} = 0$.*

Exercise: check the definition is independent of the parametrization.

Recall the definition of κ_s: Assume s is unit-speed parametrization then $\mathbf{t}' = \kappa_s \mathbf{n}_s$, where $\mathbf{n}_s$ is a 90 degree rotation of $\mathbf{t}$. The curvature $\kappa = |\kappa_s|$.



Convex curve. It has $\kappa_s \geq 0$ if the parameter increases counter-clockwise around its interior.



Non-convex curve

Definition 1.9.2. *A simple closed plane curve γ is convex if it lies on one side of its tangent line at each point.*

Equivalent definitions:

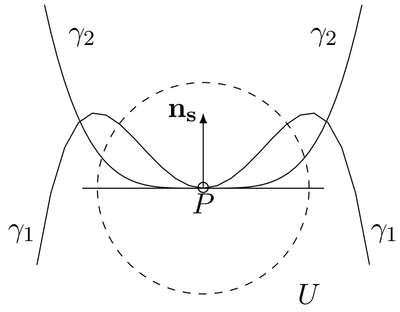
- If the interior D is convex: if $A \in D, B \in D$, then the segment $AB \subset D$.

- If a simple closed curve has a non-negative signed curvature at each of its points.

It is easy to see that the first is an equivalent definition. Leave as exercise. To prove the second equivalence, we should use the relation with turning angle and signed curvature. We do not provide the proof here. Instead we mention the following result which is more general and implies one side of the equivalence.

Proposition 1.9.3. *Let γ_1 and γ_2 tangent to each other at P, and the signed curvature $\kappa_1 > \kappa_2$. Then there is a neighbourhood U of P in which $\gamma_1 \cap U$ is located in one side of $\gamma_2 \cap U$ defined by $\mathbf{n}_s$.*

Proof. We choose P at origin. And express γ_1 and γ_2 locally as graphs of functions f_1, f_2. So $f_1(0) = f_2(0) = 0$ and $f_1'(0) = f_2'(0)$. By the formula of signed curvature in section 1.6, $\kappa_s = x'y'' - x''y' = f''$ for graph of function f. Hence by Taylor expansion $f_1(x) - f_2(x) = \frac{x^2}{2}(\kappa_1 - \kappa_2) + o(x^2)$ which is greater than 0 in a neighbourhood of $P = (0, 0)$. This finishes the proof. $\square$



Any simple closed curve has at least two vertices: maximum and minimum of κ_s. Actually we have more

Theorem 1.9.4 (Four Vertex Theorem)**.** *Every convex simple closed curve in $\mathbb{R}^2$ has at least four vertices.*

Remark 1.9.5. *The conclusion holds for simple closed curves, but we only prove it for convex ones.*

Suppose γ has fewer than 4 vertices. Then κ_s must have 2 or 3 critical points. Under this circumstance, we have the following

Lemma 1.9.6. *There is a straight line L that divides γ into 2 segments, in one of which $\kappa_s' > 0$ and in the other $\kappa_s' \leq 0$. (or possibly $\kappa_s' \geq 0$ and $\kappa_s' < 0$ respectively)*

Proof. Let the max/min points of κ_s be P and Q.

If P and Q are the only vertices, $\kappa_s' > 0$ on one of the segments and $\kappa_s' < 0$ on the other. This is because a closed curve will keep the same value of κ_s after one turn.

If there is one more vertex R. Then P, Q, R determined 3 segments in γ, each of them $\kappa_s' > 0$ or $\kappa_s' < 0$. Then there are two adjacent ones on which κ_s' has the same sign. □

A slightly more careful argument along this line will actually show that it is impossible to have exactly 3 critical points.

Proof. (of the theorem) Let the equation of L be $\mathbf{a} \cdot \mathbf{x} = 0$, where we choose the unit vector $\mathbf{a}$ such that $\kappa_s' \geq 0$ precisely when $\mathbf{a} \cdot \gamma(s) \geq 0$. Then

$$\int_0^T \kappa_s'(\mathbf{a} \cdot \gamma(s))ds > 0.$$

Integration by parts

$$\begin{aligned}\int_0^T \kappa_s'(\mathbf{a}\cdot\gamma(s))ds =&\kappa_s(\mathbf{a}\cdot\gamma(s))|_0^T - \int_0^T \kappa_s(\mathbf{a}\cdot\mathbf{t}(s))ds\\ =&0+\int_0^T \mathbf{a}\cdot\mathbf{n}_s'(s)ds\\ =&\mathbf{a}\cdot\int_0^T \mathbf{n}_s'(s)ds\\ =&0\end{aligned}$$

This is a contradiction!

Notice for the second equality. The first term is 0 because the curve is simple closed. The second term is because of $\mathbf{n}_s' = -\kappa_s\mathbf{t}$. ($\mathbf{t}\cdot\mathbf{n}_s = 0$, so $\mathbf{t}'\cdot\mathbf{n}_s + \mathbf{t}\cdot\mathbf{n}_s' = 0$. Since $\mathbf{t}' = \kappa_s\mathbf{n}_s$ by definition, $\mathbf{n}_s' = -\kappa_s\mathbf{t}$) □

Exercise: Find all the vertices for the ellipse $\gamma(t) = (p\cos t, q\sin t)$ when $p \neq q$.

Chapter 2

Surfaces in $\mathbb{R}^3$

2.1 Definitions and Examples

First, we assume you know the definition of open sets and continuous maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Definition 2.1.1. *If $f : X \to Y$ is continuous and bijective, and if its inverse map $f^{-1} : Y \to X$ is also continuous, then f is called a homeomorphism and X and Y are said to be homeomorphic.*

Theorem 2.1.2 (Invariance of domain)**.** *If $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ is an injective continuous map, then $V = f(U)$ is open and f is a homeomorphism between U and V.*

Definition 2.1.3. *A subset $S \subset \mathbb{R}^3$ is a regular surface if for each $p \in S$, there exists a neighbourhood $W \subset \mathbb{R}^3$ and a map $\sigma : U \to W \cap S$ of an open set $U \subset \mathbb{R}^2$ onto $W \cap S \subset \mathbb{R}^3$, such that*

- *σ is smooth*
- *σ is homeomorphism*
- *at all points $(u, v) \in U$, $\sigma_u \times \sigma_v \neq 0$.*

The mapping σ is called a *(regular) parametrization* or a *chart.* We will call its image a *coordinate patch.* A collection of charts such that every point of S is contained in a coordinate patch is called an *atlas.* The condition 3 above means σ_u and σ_v are linearly independent, or $d\sigma_q : \mathbb{R}^2 \to \mathbb{R}^3$ is one to one.

For any point of a regular surface S, there might be more than one charts.

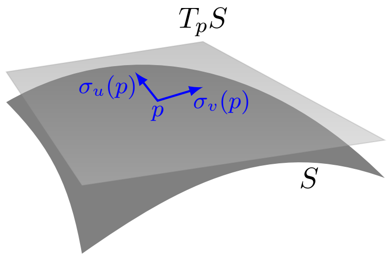
Proposition 2.1.4. *Let $\sigma : U \to S$, $\tilde{\sigma} : V \to S$ be two charts of S such that $p \in \sigma(U) \cap \tilde{\sigma}(V) = W$. Then the transition map $h = \sigma^{-1} \circ \tilde{\sigma} : \tilde{\sigma}^{-1}(W) \to \sigma^{-1}(W)$ is a diffeomorphism.*

Proof is omitted. It is another application of Inverse Function Theorem. The diffeomorphism h gives a *reparametrization*.

Definition 2.1.5. *A reparametrization of surface is a composition* $\sigma \circ f : V \to \mathbb{R}^3$ *where* $f : V \to U$ *is a diffeomorphism.*

Since the Jacobian *df* is invertible, let $f(x, y) = (u(x, y), v(x, y))$, $(\sigma \circ f)_x$ and $(\sigma \circ f)_y$ are linearly independent if and only if σ_u and σ_v are. So the following is well defined.

Definition 2.1.6. *The tangent plane* T_pS *of a surface* S *at the point* p *is the vector space spanned by* $\sigma_u(p)$ *and* $\sigma_v(p)$.



This space is independent of parametrization. One should think of the origin of the vector space as the point p.

Definition 2.1.7. *The unit vector*

$$\mathbf{N}_\sigma(u, v) = \frac{\sigma_u \times \sigma_v}{||\sigma_u \times \sigma_v||}$$

is the standard normal to the surface at point $p = \sigma(u, v)$.

Here are examples of parametrized surfaces. For the pictures of these, look at Hichin's notes.

Example:

1. A plane:

   $$\sigma(u, v) = \mathbf{a} + u\mathbf{b} + v\mathbf{c}$$

   for constant vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{b} \times \mathbf{c} \neq 0$. The normal vector

   $$\mathbf{N} = \frac{\mathbf{b} \times \mathbf{c}}{||\mathbf{b} \times \mathbf{c}||}$$

2. A cylinder:

   $$\sigma(u, v) = (a \cos u, a \sin u, v), a > 0$$

   $$\mathbf{N} = (\cos u, \sin u, 0)$$

3. A cone (without cone point):

$$\sigma(u,v) = (au\cos v, au\sin v, u)$$

4. A helicoid:

$$\sigma(u,v) = (au\cos v, au\sin v, v)$$

5. A sphere (minus a half circle connecting poles) in spherical coordinates: $U = (-\frac{\pi}{2}, \frac{\pi}{2}) \times (0, 2\pi)$.

$$\sigma(u,v) = (a\cos v\cos u, a\sin v\cos u, a\sin u)$$

$$\mathbf{N} = -\frac{1}{a}\sigma$$

6. A torus

$$\sigma(u,v) = ((a + b\cos u)\cos v, (a + b\cos u)\sin v, b\sin u),$$

$a > b$ are constants.

7. A surface of revolution

$$\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u))$$

is obtained by rotating a plane curve (called profile curve) $\gamma(u) = (f(u), 0, g(u))$ around z-axis. We assume $f(u) > 0$ for all u. We have

$$\sigma_u = (f_u\cos v, f_u\sin v, g_u), \sigma_v = (-f\sin v, f\cos v, 0).$$

So

$$\sigma_u \times \sigma_v = (-f\dot{g}\cos v, -f\dot{g}\sin v, f\dot{f}), ||\sigma_u \times \sigma_v||^2 = f^2(\dot{f}^2 + \dot{g}^2) \neq 0$$

8. A generalized cylinder

$$\sigma(u,v) = \gamma(u) + v\mathbf{a}.$$

$$\sigma_u = \dot{\gamma}, \sigma_v = \mathbf{a}$$

σ is regular if γ is never tangent to the ruling $\mathbf{a}$.

But usually, a surface has more than one patches. That is the reason why we need more preparation of surfaces local theory than that of curves. For curves, only one patch is enough since the topology is simpler. The following example shows how a closed (i.e. compact without boundary) surface is different from a closed curve, where we can use a periodic one patch parametrization.

Example 2.1.8. *The unit sphere* $S^2 = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 + z^2 = 1\}$ *is a regular surface. We let* $\sigma_1 : U \subset \mathbb{R}^2 \to \mathbb{R}^3$.

$$\sigma_1(x, y) = (x, y, \sqrt{1 - x^2 - y^2}), (x, y) \in U = B_1(0)$$

is a parametrization. Please check the 3 conditions (general statement is Proposition 2.1.9).

Similarly $\sigma_2(x, y) = (x, y, -\sqrt{1 - x^2 - y^2}), (x, y) \in U = B_1(0)$ *is also a parametrization. And* $\sigma_1(U) \cup \sigma_2(U)$ *covers* S^2 *minus equator* $z = 0$.

With 4 more parametrizations

$$\sigma_3(x, z) = (x, \sqrt{1 - x^2 - z^2}, z)$$

$$\sigma_4(x, z) = (x, -\sqrt{1 - x^2 - z^2}, z)$$

$$\sigma_5(y, z) = (\sqrt{1 - y^2 - z^2}, y, z)$$

$$\sigma_6(y, z) = (-\sqrt{1 - y^2 - z^2}, y, z)$$

they cover S^2. *So* S^2 *is a regular surface.*

To check each σ_i is a parametrization, one could prove the following more general result, whose proof is left as an exercise.

Proposition 2.1.9. *If* $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ *is a smooth function in an open set* U *of* $\mathbb{R}^2$, *then the graph of* f, *i.e.* $\sigma(x, y) = (x, y, f(x, y))$ *for* $(x, y) \in U$ *is a regular surface.*

2.1.1 Compact surfaces

A subset X of $\mathbb{R}^3$ is *compact* if it is closed and bounded (i.e. X is contained in some open ball).

Non-examples: A plane is not compact. The open disc $\{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 < 1, z = 0\}$ is not compact.

There are very few compact surfaces:

Example 2.1.10. *Any sphere is compact. Let us consider the unit sphere* S^2.

It is bounded because it is contained in the open ball $D_2(0)$.

To show S^2 *is closed, i.e. the complement is open: if* $||p|| \neq 1$, *say* $||p|| > 1$. *Let* $\epsilon = ||p|| - 1$, $D_\epsilon(p)$ *does not intersect* S^2. *This is because if* $q \in D_\epsilon(p)$, *then* $||q|| \geq ||p|| - ||p - q|| > ||p|| - \epsilon = 1$.

Other examples are torus $\Sigma_1 = T^2$, and surface of higher "genus" $\Sigma_{g \geq 2}$.

Theorem 2.1.11. *For any* $g \geq 0$, Σ_g *has an atlas such that it is a smooth surface. Moreover, every compact surface is diffeomorphic to one of* Σ_g.

2.1.2 Level sets

There is another family of regular surfaces: the level sets. Suppose that $f : U \subset \mathbb{R}^n \to \mathbb{R}$ is smooth. For each $p \in U$, we have Jacobian

$$df_p = \nabla f(p) = (f_{x_1}, f_{x_2}, \cdots, f_{x_n})(p).$$

Definition 2.1.12. *We say $p \in U$ is a critical point if $df_p = 0$. Otherwise it is regular.*

The image $f(p)$ of a critical point is called a critical value. $t \in \mathbb{R}$ is a regular value if every point of the level set $f^{-1}(t)$ is regular.

The following shows the notions of regular surface and regular value coincide in some sense.

Theorem 2.1.13. *If $f : U \subset \mathbb{R}^3 \to \mathbb{R}$ is a smooth function and $t \in f(U)$ is a regular value of f, then $f^{-1}(t)$ is a regular surface in $\mathbb{R}^3$.*

Proof. Let p be a point of $f^{-1}(t)$. Without loss, we assume $f_z(p) \neq 0$. Define $F : U \subset \mathbb{R}^3 \to \mathbb{R}^3$ by

$$F(x, y, z) = (x, y, f(x, y, z)).$$

Its Jacobian is

$$dF_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ f_x & f_y & f_z \end{pmatrix}.$$

$det(dF_p) = f_z \neq 0$.

Thus by Inverse Function Theorem, we have a neighbourhood $V \subset \mathbb{R}^3$ of p and $W \subset \mathbb{R}^3$ of $F(p)$ such that $F : V \to W$ is invertible and $F^{-1} : W \to V$ is smooth, i.e. $F^{-1}(u, v, w) = (u, v, g(u, v, w))$ with $(u, v, w) \in W$ and g smooth. Especially $g(u, v, t) = h(u, v)$ is smooth, where h takes value from $W' = \{(u, v)|(u, v, t) \in W\} \subset \mathbb{R}^2$. Since $F(f^{-1}(t) \cap V) = \{(u, v, t)\} \cap W$, the graph of $h(u, v)$ is $F^{-1}(u, v, t) = f^{-1}(t) \cap V$. Hence $h : W' \to f^{-1}(t) \cap V$ is a parametrization containing p. Hence by Proposition 2.1.9 $f^{-1}(t)$ is a regular surface. □

Example 2.1.14.

- *$f(x, y, z) = x^2 + y^2 + z^2$. $\nabla f = (2x, 2y, 2z)$. Thus $f^{-1}(t)$ is an embedded surface for all $t > 0$. It is a sphere of radius t.*

- *$f(x, y, z) = x^2 + y^2 - z^2$. $\nabla f = (2x, 2y, -2z)$. $f^{-1}(0)$ is a cone which is singular at the origin. $f^{-1}(t)$ is a regular surface for $t \neq 0$. It is a hyperboloid — 1-sheeted for $t > 0$ and 2-sheeted for $t < 0$.*

2.2 The First Fundamental Form

Choose a parametrization $\sigma : U \to \mathbb{R}^3$ of S, such that $p \in \sigma(U)$ and $\sigma(u_0, v_0) = p$. A curve γ lies on S and passes through p when $t = t_0$ if $\gamma(t) = \sigma(u(t), v(t))$ with $u(t_0) = u_0$ and $v(t_0) = v_0$. By Inverse Function Theorem, both u and v are smooth.

Since $||\dot{\gamma}||^2 = <\dot{\gamma}, \dot{\gamma}> = E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2$, where

$$E = \sigma_u \cdot \sigma_u, F = \sigma_u \cdot \sigma_v, G = \sigma_v \cdot \sigma_v,$$

the arc length of such a curve from $t = a$ to $t = b$ is

$$\int_a^b ||\dot{\gamma}(t)||dt = \int_a^b \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2}dt.$$

Definition 2.2.1. *The first fundamental form of a surface in* $\mathbb{R}^3$ *is the expression*

$$I = Edu^2 + 2Fdudv + Gdv^2.$$

This is just the quadratic form

$$Q(\mathbf{v}, \mathbf{v}) = \mathbf{v} \cdot \mathbf{v}$$

on the tangent plane written in terms of the basis σ_u and σ_v. (And we assume the formal computations $du(\sigma_u) = dv(\sigma_v) = 1, du(\sigma_v) = dv(\sigma_u) = 0$.) So it tells us how the surface S inherits the inner product of $\mathbb{R}^3$. It is represented in this basis by the symmetric matrix

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

It is clear that the first fundamental form only depends on S and p. Especially, it does not depend on the parametrization. A reparametrization $\tilde{\sigma} = \sigma \circ f$ will change it to the same form $\bar{E}dx^2 + 2\bar{F}dxdy + \bar{G}dy^2$ which is identical to the one calculated from coordinate change

$$du = u_x dx + u_y dy, dv = v_x dx + v_y dy,$$

where $f(x, y) = (u(x, y), v(x, y))$. It helps us to make measurement (e.g. Length of curves, angles, areas) on the surface directly, so we say a property of S is *intrinsic* if it can be expressed in terms of the first fundamental form.

Example:

1. Plane $\sigma(u, v) = \mathbf{a} + u\mathbf{b} + v\mathbf{c}$ with $\mathbf{b} \perp \mathbf{c}$ and

$$||\mathbf{b}|| = ||\mathbf{c}|| = 1.$$

$\sigma_u = \mathbf{b}, \sigma_v = \mathbf{c}$, so

$$E = ||\mathbf{b}||^2 = 1, F = \mathbf{b} \cdot \mathbf{c} = 0, G = ||\mathbf{c}||^2 = 1.$$

The first fundamental form is

$$I = du^2 + dv^2.$$

2. Surface of revolution

$$\sigma(u, v) = (f(u) \cos v, f(u) \sin v, g(u)).$$

We could assume the profile curve $\gamma(u) = (f(u), 0, g(u))$ is unit-speed, i.e. $f_u^2 + g_u^2 = 1$, and $f > 0$. We have

$$\sigma_u = (f_u \cos v, f_u \sin v, g_u), \sigma_v = (-f \sin v, f \cos v, 0).$$

So

$$E = f_u^2 + g_u^2 = 1, F = 0, G = f^2.$$

Hence

$$I = du^2 + f(u)^2 dv^2$$

The unit sphere S^2 is a special case where $u = \theta, v = \phi, f(\theta) = \cos \theta, g(\theta) = \sin \theta$. We have

$$I = d\theta^2 + \cos^2 \theta d\phi^2$$

3. Generalized cylinder $\sigma(u, v) = \gamma(u) + v\mathbf{a}$. We assume γ is unit-speed, $\mathbf{a}$ is a unit vector, and $\dot{\gamma} \perp \mathbf{a}$. Since $\sigma_u = \dot{\gamma}, \sigma_v = \mathbf{a}$,

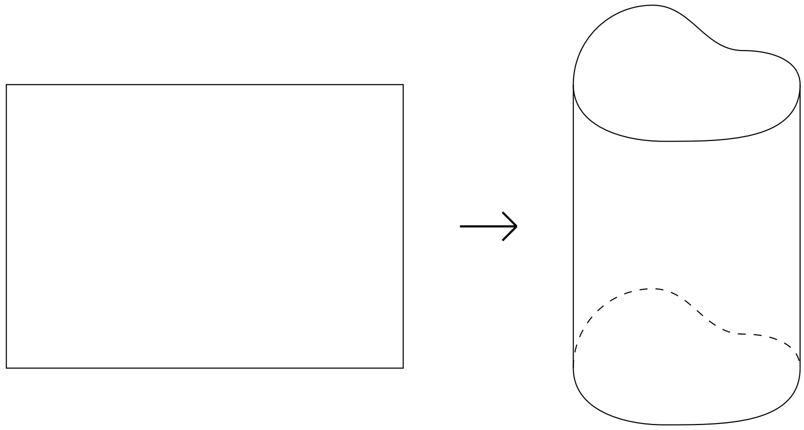
$$I = du^2 + dv^2.$$

Exercise: Calculate the first fundamental form for all other examples in previous section.

Observe that the first fundamental form of a generalized cylinder is the same as that of a plane! This is not a coincidence. The reason is the generalized cylinder is obtained from bending a piece of paper. Or it could be cut through one of its ruling to a flat paper. This is called a local isometry.

2.3 Length, Angle, Area: Isometric, Conformal, Equiareal

In this section, we explore several intrinsic properties.



2.3.1 Length: Isometry

Definition 2.3.1. *Two surfaces S_1 and S_2 are isometric if there is a diffeomorphism $f : S_1 \to S_2$ which maps curves in S_1 to curves in S_2 of the same length. The map f is called an isometry.*

The map from a plane to a cylinder is not an isometry since it is not a diffeomorphism. But indeed it has the second property. A smooth map like this is called a *local isometry*. This suggests us to look at this definition for a coordinate patch.

Theorem 2.3.2. *The coordinate patches U_1 and U_2 are isometric if and only if there exist parametrizations $\sigma_1 : V \to \mathbb{R}^3$ and $\sigma_2 : V \to \mathbb{R}^3$ with the same first fundamental form, and $\sigma_1(V) = U_1, \sigma_2(V) = U_2$.*

Proof. Suppose such parametrizations exist, then the identity map is an isometry since the first fundamental form determines the length of curves.

Conversely, assume U_1, U_2 are isometric. And let the charts be $\sigma_1 : V_1 \to \mathbb{R}^3$ and $\sigma_2 : V_2 \to \mathbb{R}^3$. So we could assume the diffeomorphism is realized by $f : V_1 \to V_2$. Then

$$\sigma_2 \circ f, \sigma_1 : V_1 \to \mathbb{R}^3$$

are parametrizations from the same open set $V = V_1$. So the fundamental forms are defined using same coordinate (u, v) as

$$E_1 du^2 + 2F_1 dudv + G_1 dv^2, E_2 du^2 + 2F_2 dudv + G_2 dv^2.$$

We have

$$\int_I \sqrt{E_1 \dot{u}^2 + 2F_1 \dot{u}\dot{v} + G_1 \dot{v}^2} dt = \int_I \sqrt{E_2 \dot{u}^2 + 2F_2 \dot{u}\dot{v} + G_2 \dot{v}^2} dt$$

for all curves and all intervals. Take derivative, we have

$$\sqrt{E_1 \dot{u}^2 + 2F_1 \dot{u}\dot{v} + G_1 \dot{v}^2} = \sqrt{E_2 \dot{u}^2 + 2F_2 \dot{u}\dot{v} + G_2 \dot{v}^2}$$

for all $u(t)$ and $v(t)$. Hence $E_1 = E_2, F_1 = F_2, G_1 = G_2$. □

2.3.2 Angle: conformal

One notices that the dot product inherited from $\mathbb{R}^3$ is also preserved under isometry (and *vice versa*), since it is determined by the first fundamental form:

$$\mathbf{v} \cdot \mathbf{w} = \frac{1}{2}(||\mathbf{v} + \mathbf{w}||^2 - ||\mathbf{v}||^2 - ||\mathbf{w}||^2).$$

Hence, the angle is also an intrinsic invariant. Let us define it.

Look at two curves α, β on the surface S intersecting at $t = 0$. The angle between them at $t = 0$ is given by

$$\cos\theta = \frac{\dot{\alpha} \cdot \dot{\beta}}{||\dot{\alpha}||||\dot{\beta}||}, 0 \leq \theta \leq \pi.$$

Everything is expressed in terms of the coefficients of the first fundamental form.

Definition 2.3.3. *Two surfaces S_1 and S_2 are conformal if there is a diffeomorphism f which preserves the angle for any pair of curves.*

Notice the invariance of the expression of $\cos\theta$ if we scale the first fundamental form by a positive function λ^2. Hence we have a similar characterization as for isometry.

Theorem 2.3.4. *The coordinate patches U_1 and U_2 are conformal if and only if there exist parametrizations $\sigma_1 : V \to \mathbb{R}^3$ and $\sigma_2 : V \to \mathbb{R}^3$ with $\sigma_1(V) = U_1, \sigma_2(V) = U_2$, and $E_2 = \lambda^2 E_1, F_2 = \lambda^2 F_1, G_2 = \lambda^2 G_1$ in V, where λ^2 is a nowhere zero differentiable function in V.*

We call them locally conformal. The most important property of conformal maps is the following.

Theorem 2.3.5. *Any two regular surfaces are locally conformal.*

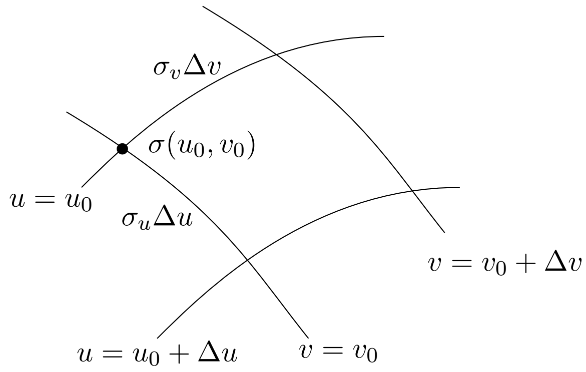
To prove the theorem, we need to choose a special parametrization. For a neighborhood of any point of a regular surface by *Isothermal parametrization*, in which the first fundamental form is $\lambda^2(u, v)(du^2 + dv^2)$.

2.3.3 Area: equiareal

Let us focus on a parametrized surface $\sigma : U \to \mathbb{R}^3$. There are two families of curves $u = const$ and $v = const$. Fix $(u_0, v_0) \in U$, we have the following picture.

The area of the "parallelogram" is

$$||\sigma_u \Delta u \times \sigma_v \Delta v|| = ||\sigma_u \times \sigma_v||\Delta u \Delta v.$$



Local "parallelogram"

Definition 2.3.6. *The area $A_\sigma(R)$ of the part $\sigma(R)$ of $\sigma : U \to \mathbb{R}^3$ for region $R \subset U$ is*

$$A_\sigma(R) = \int_R ||\sigma_u \times \sigma_v|| dudv = \int_R \sqrt{EG - F^2} dudv.$$

The second equality follows from

$$||\sigma_u \times \sigma_v||^2 = ||\sigma_u||^2 ||\sigma_v||^2 - (\sigma_u \cdot \sigma_v)^2 = EG - F^2.$$

As a corollary, we know the area of a surface patch is unchanged by reparametrization.

There is a characterization for equiareal map.

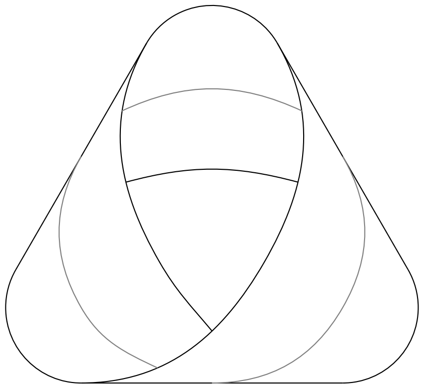
Theorem 2.3.7. *A diffeomrphism $f : U_1 \to U_2$ is equiareal, i.e. it takes any region in S_1 to a region of same area in S_2, if and only if for any surface patch σ on S_1, the first fundamental forms of charts σ and $f \circ \sigma$ satisfy*

$$E_1 G_1 - F_1^2 = E_2 G_2 - F_2^2.$$

We summarize that being isometric is a stronger condition than being conformal or equiareal.

2.4 The Second Fundamental Form

The first fundamental form describe the intrinsic geometry of a surface, namely independent of the choice of its sitting in $\mathbb{R}^3$. The second fundamental form describes how the surface is bent in $\mathbb{R}^3$.



Möbius band

2.4.1 Normals and orientability

A unit normal to surface S at p, up to sign, is a unit vector perpendicular to T_pS. Recall that we define standard unit normal for a parametrization $\sigma : U \to \mathbb{R}^3$ as

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

However, we do not always have a smooth choice of the unit normal at any point of S. For instance, the Möbius band is such an example. Intuitively, if we walk along the middle circle of it, after one turn, the normal vector $\mathbf{N}$ will come back as $-\mathbf{N}$. In other words, we cannot make a consistent choice of a definite "side" on Möbius band. But apparently, $\mathbf{N}_\sigma$ is a smooth choice on one surface patch. Actually, the reason of this phenomenon is $\mathbf{N}_\sigma$ depends on the choice of patches.

Let $\tilde{\sigma} : \tilde{U} \to \mathbb{R}^3$ be another. Then

$$\tilde{\sigma}_{\tilde{u}} \times \tilde{\sigma}_{\tilde{v}} = (\frac{\partial u}{\partial \tilde{u}}\frac{\partial v}{\partial \tilde{v}} - \frac{\partial u}{\partial \tilde{v}}\frac{\partial v}{\partial \tilde{u}})\sigma_u \times \sigma_v = \det J(\Phi)\sigma_u \times \sigma_v,$$

where $J(\Phi)$ is the Jacobian of the transition map $\Phi = \tilde{\sigma}^{-1} \circ \sigma$. So $\mathbf{N}_{\tilde{\sigma}} = \pm\mathbf{N}_\sigma$. The sign is that of $\det J(\Phi)$.

Definition 2.4.1. *A surface S is orientable if we have a smooth choice of unit normal at any point of S. Such a choice of unit normal vector field is called an orientation of S.*

A surface with a chosen orientation is called oriented.

Example 2.4.2. *Every compact surface in $\mathbb{R}^3$ is orientable. This is because every compact surface is diffeomorphic to one of Σ_g.*

The next follows from the above discussion.

Proposition 2.4.3. *A surface S is orientable if there exists an atlas $\mathcal{A}$ of S such that for transition map Φ between any two charts in $\mathcal{A}$, we have* $\det J(\Phi) > 0$.

After on, without particular mentioning, our surface will be orientable.

2.4.2 Gauss map and second fundamental form

Let $S \subset \mathbb{R}^3$ be a surface with an orientation $\mathbf{N}$, we have the *Gauss map*

$$\mathcal{G} : S \to S^2, p \mapsto \mathbf{N}_p,$$

where $\mathbf{N}_p$ is the unit normal of S at p. The rate at which $\mathbf{N}$ varies across S is measured by the derivative. It is denoted as $D_p\mathcal{G} : T_pS \to T_{\mathcal{G}(p)}S^2$. But as planes in $\mathbb{R}^3$, $T_{\mathcal{G}(p)}S^2$ and T_pS are parallel since both are perpendicular to $\mathbf{N}$. So we actually look at the *Weingarten map*

$$\mathcal{W}_{p,S} = -D_p\mathcal{G} : T_pS \to T_pS.$$

It is defined as the unique linear map determined by

$$\mathcal{W}(\sigma_u) = -\mathbf{N}_u, \mathcal{W}(\sigma_v) = -\mathbf{N}_v$$

for any parametrization σ.

Exercise: Prove $\mathcal{W}$ is independent of the choice of surface parametrization.

Parallel to the discussion of first fundamental form, we have

Definition 2.4.4. *The second fundamental form of an oriented surface is the expression*

$$II = Ldu^2 + 2Mdudv + Ndv^2$$

where $L = \sigma_{uu} \cdot \mathbf{N}, M = \sigma_{uv} \cdot \mathbf{N}, N = \sigma_{vv} \cdot \mathbf{N}$.

There is another expression. Note that $\sigma_u \cdot \mathbf{N} = 0$, we have

$$(\sigma_u \cdot \mathbf{N})_u = \sigma_{uu} \cdot \mathbf{N} + \sigma_u \cdot \mathbf{N}_u = 0$$

and similarly

$$\sigma_{vu} \cdot \mathbf{N} + \sigma_v \cdot \mathbf{N}_u = 0, \sigma_{uv} \cdot \mathbf{N} + \sigma_u \cdot \mathbf{N}_v = 0, \sigma_{vv} \cdot \mathbf{N} + \sigma_v \cdot \mathbf{N}_v = 0.$$

Hence we also have

$$L = -\sigma_u \cdot \mathbf{N}_u$$

$$M = -\sigma_u \cdot \mathbf{N}_v = -\sigma_v \cdot \mathbf{N}_u$$

$$N = -\sigma_v \cdot \mathbf{N}_v$$

Hence the second fundamental form is the symmetric bilinear form

$$II(\mathbf{w}) = \mathcal{W}_{p,S}(\mathbf{w}) \cdot \mathbf{w} =< \mathcal{W}_{p,S}(\mathbf{w}), \mathbf{w} > .$$

It is represented by

$$\begin{pmatrix} L & M \\ M & N \end{pmatrix}.$$

in terms of basis σ_u and σ_v.

There is a third interpretation. Recall that the curvature of a curve could be understand as $\mathbf{t}' \cdot \mathbf{n}$, or the second term of Taylor expansion of $\gamma(s)$. We could understand the second fundamental form in a similar way. We look at Taylor expression

$$\sigma(u+\Delta u, v+\Delta v)-\sigma(u,v) = \sigma_u \Delta u+\sigma_v \Delta v+\frac{1}{2}(\sigma_{uu}(\Delta u)^2+2\sigma_{uv}\Delta u \Delta v+\sigma_{vv}(\Delta v)^2)+R$$

where $\lim_{\Delta u, \Delta v \to 0} \frac{R}{(\Delta u)^2+(\Delta v)^2} = 0$. Since $\sigma_u \cdot \mathbf{N} = \sigma_v \cdot \mathbf{N} = 0$,

$$(\sigma(u+\Delta u, v+\Delta v) - \sigma(u,v)) \cdot \mathbf{N} = \frac{1}{2}(L(\Delta u)^2 + 2M\Delta u \Delta v + N(\Delta v)^2) + R'.$$

The fourth interpretation is more geometric: we take surface $\sigma(u,v)$ and push it inwards a distance t along its normal to get a family of surfaces

$$R(u,v,t) = \sigma(u,v) - t\mathbf{N}(u,v).$$

We calculate the first fundamental form $Edu^2 + 2Fdudv + Gdv^2$ of R which depends on t, then the derivative

$$\frac{1}{2}\frac{\partial}{\partial t}(Edu^2 + 2Fdudv + Gdv^2)|_{t=0} = Ldu^2 + 2Mdudv + Ndv^2$$

where $Ldu^2 + 2Mdudv + Ndv^2$ is the second fundamental form of σ. So it describes how the first fundamental form varies along the unit normal direction.

Example:

1. Plane $\sigma(u,v) = \mathbf{a} + u\mathbf{b} + v\mathbf{c}$ has $\sigma_{uu} = \sigma_{uv} = \sigma_{vv} = 0$. So the second fundamental form vanishes.

2. Surface of revolution

$$\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u)).$$

We again assume $f_u^2 + g_u^2 = 1$ and $f > 0$. We have

$$\sigma_u = (f_u \cos v, f_u \sin v, g_u), \sigma_v = (-f \sin v, f \cos v, 0).$$

So

$$\sigma_u \times \sigma_v = (-fg_u \cos v, -fg_u \sin v, ff_u), ||\sigma_u \times \sigma_v|| = f.$$

Hence

$$\begin{aligned} \mathbf{N} &= (-g_u \cos v, -g_u \sin v, f_u), \\ \sigma_{uu} &= (f_{uu} \cos v, f_{uu} \sin v, g_{uu}), \\ \sigma_{uv} &= (-f_u \sin v, f_u \cos v, 0), \\ \sigma_{vv} &= (-f \cos v, -f \sin v, 0). \end{aligned}$$

So the second fundamental form

$$II = (f_u g_{uu} - f_{uu} g_u) du^2 + f g_u dv^2.$$

There are two special cases:

(a) Unit sphere: $u = \theta$, $v = \phi$, $f(\theta) = \cos\theta$, $g(\theta) = \sin\theta$.

$$II = d\theta^2 + \cos^2\theta d\phi^2$$

the same as its first fundamental form.

(b) Unit cylinder: $f(u) = 1$, $g(u) = u$. So

$$II = dv^2.$$

This is different from that of a plane, although their first fundamental forms are the same.

These examples tells us second fundamental form is an extrinsic concept, although it is not independent of the first fundamental form.

Exercise: Prove the converse of Example 1: If the second fundamental form vanishes, it is part of a plane.

2.5 Curvatures

2.5.1 Definitions and first properties

The shape of a surface influences the curvature of curves on the surface.

Let $\gamma(t) = \sigma(u(t), v(t))$ be a unit-speed curve on an oriented surface S. Hence $\dot\gamma = \dot u \sigma_u + \dot v \sigma_v \in T_{\gamma(t)}S$, which means $\dot\gamma \perp \mathbf{N}$. So $\mathbf{N}$, $\dot\gamma$ and $\mathbf{N} \times \dot\gamma$ is a right handed orthonormal basis of $\mathbb{R}^3$. Since $\ddot\gamma \perp \dot\gamma$,

$$\ddot\gamma = \kappa_n \mathbf{N} + \kappa_g \mathbf{N} \times \dot\gamma \tag{2.1}$$

Hence κ_n is called the *normal curvature* and κ_g is called the *geodesic curvature* of γ. Notice when σ is a plane and γ a plane curve, the geodesic curvature is just the signed curvature κ_s.

On a general (non-oriented) surface, only magnitudes of κ_n and κ_g are well defined.

Proposition 2.5.1. *1.* $\kappa_n = \ddot{\gamma} \cdot \mathbf{N}, \kappa_g = \ddot{\gamma} \cdot (\mathbf{N} \times \dot{\gamma})$.

2. $\kappa^2 = \kappa_n^2 + \kappa_g^2$.

3. $\kappa_n = \kappa \cos\psi, \kappa_g = \kappa \sin\psi$, *where* κ *is the curvature of* γ *and* ψ *is the angle between* $\mathbf{N}$ *and* $\mathbf{n}$ *of* γ.

Proof. The first is obtained by multiplying $\mathbf{N}$ and $\mathbf{N} \times \dot{\gamma}$ respectively to (2.1).

The second is by multiplying $\ddot{\gamma}$ to it.

For the last notice $\ddot{\gamma} = \kappa\mathbf{n}$. Comparing the coefficients of (2.1) and

$$\kappa\mathbf{n} = \kappa \cos\psi \mathbf{N} + \kappa \sin\psi \mathbf{N} \times \dot{\gamma}$$

gives us the equalities. □

Proposition 2.5.2. *If* γ *is a unit-speed curve on* S,

$$\kappa_n = II(\dot{\gamma}).$$

In other words, for $\gamma(t) = \sigma(u(t), v(t))$,

$$\kappa_n = L\dot{u}^2 + 2M\dot{u}\dot{v} + N\dot{v}^2$$

Proof. Since $\mathbf{N} \cdot \dot{\gamma} = 0$, $\mathbf{N} \cdot \ddot{\gamma} = -\dot{\mathbf{N}} \cdot \dot{\gamma}$. So

$$\kappa_n = \mathbf{N} \cdot \ddot{\gamma} = -\dot{\mathbf{N}} \cdot \dot{\gamma} = < \mathcal{W}(\dot{\gamma}), \dot{\gamma} > = II(\dot{\gamma}).$$

□

So κ_n only depends on the point p and the tangent vector $\dot{\gamma}(p)$, but not the curve γ.

Theorem 2.5.3 (Meusnier's Theorem)**.** *Let* $p \in S$, $\mathbf{v} \in T_pS$ *a unit vector. Let* Π_θ *be the plane containing* $\mathbf{v}$ *and making angle* $\theta \neq 0$ *with* T_pS*. Suppose* Π_θ *intersects* S *in a curve with curvature* κ_θ*. Then* $\kappa_\theta \sin\theta$ *is independent of* θ.

Proof. Let $\gamma_\theta = \Pi_\theta \cap S$, and parametrize it by arc length.

Then at p, $\dot{\gamma}_\theta = \pm\mathbf{v}$, so $\ddot{\gamma}_\theta \perp \mathbf{v}$ and $\parallel \Pi_\theta$ since γ_θ is a plane curve. Thus $\psi = \frac{\pi}{2} - \theta$ and $\kappa_\theta \sin\theta = \kappa_n$, independent of θ. □

The Weingarten map is a linear map. It could be viewed as a symmetric 2×2 matrix after fixing basis, say σ_u, σ_v, since the second fundamental form is a symmetric bilinear form. Its determinant and trace are two invariant associate with it, which is independent of the choice of basis.

Definition 2.5.4. *Let $\mathcal{W}_p$ be the Weingarten map at $p \in S$. Then the Gaussian curvature*

$$K = \det(\mathcal{W}_p),$$

and mean curvature

$$H = \frac{1}{2} trace(\mathcal{W}_p).$$

For a linear map/matrix, we also look at their eigenvalues and eigenvectors. For $\mathcal{W}$, the eigenvalues are real numbers since it is symmetric.

So at $p \in S$, there are κ_1, κ_2 and a basis $\{\mathbf{t}_1, \mathbf{t}_2\}$ of T_pS such that

$$\mathcal{W}(\mathbf{t}_1) = \kappa_1 \mathbf{t}_1, \mathcal{W}(\mathbf{t}_2) = \kappa_2 \mathbf{t}_2.$$

Moreover, if $\kappa_1 \neq \kappa_2$, then $< \mathbf{t}_1, \mathbf{t}_2 >= 0$. We call κ_1, κ_2 *principal curvatures*, and $\mathbf{t}_1, \mathbf{t}_2$ *principal vectors*. Points of the surface with $\kappa_1 = \kappa_2$ is called *umbilical points*, where $\mathcal{W}_p$ is $\kappa_1 \cdot I_{2\times 2}$ and every direction is a principal direction.

Hence, for any points, there is an orthonormal basis of T_pS consisting of principal vectors. We also know that

$$H = \frac{1}{2}(\kappa_1 + \kappa_2), K = \kappa_1 \cdot \kappa_2.$$

Theorem 2.5.5 (Euler's Theorem)**.** *Let γ be a curve on an oriented surface S, and let κ_1, κ_2 be the principal curvatures with principal vectors $\mathbf{t}_1, \mathbf{t}_2$. Then the normal curvature of γ is*

$$\kappa_n = \kappa_1 \cos^2\theta + \kappa_2 \sin^2\theta,$$

where θ is the angle from $\mathbf{t}_1$ to $\dot\gamma$ in the orientation of T_pS (which is denoted as $\widehat{\mathbf{t}_1\dot\gamma}$).

Proof. We assume $\{\mathbf{t}_1, \mathbf{t}_2\}$ is an orthonormal basis and $\widehat{\mathbf{t}_1\mathbf{t}_2} = \frac{\pi}{2}$. So

$$\dot\gamma = \cos\theta \mathbf{t}_1 + \sin\theta \mathbf{t}_2.$$

Then

$$\kappa_n = II(\dot\gamma) = \cos^2\theta \cdot II(\mathbf{t}_1) + 2\sin\theta\cos\theta < \mathcal{W}(\mathbf{t}_1), \mathbf{t}_2 > + \sin^2\theta \cdot II(\mathbf{t}_2).$$

Here, recall $II(\mathbf{v}) =< \mathcal{W}(\mathbf{v}), \mathbf{v} >$.

Finally, the conclusion follows since

$$< \mathcal{W}(\mathbf{t}_i), \mathbf{t}_j >=< \kappa_i \mathbf{t}_i, \mathbf{t}_j >= \begin{cases} \kappa_i & i = j \\ 0 & i \neq j \end{cases}$$

□

We want to remark that Meusnier's Theorem and Euler's Theorem are most ancient results on the theory of surfaces.

Corollary 2.5.6. *The principal curvatures at a point of a surface are maximum and minimum of the normal curvature of all curves on the surface that pass through this point.*

Proof. If $\kappa_1 \geq \kappa_2$, then $\kappa_1 \geq \kappa_n \geq \kappa_2$. □

2.5.2 Calculation of Gaussian and mean curvatures

Now we want to calculate Gaussian curvature K and mean curvature H in terms of first and second fundamental forms. Let $\sigma(u, v)$ be a chart, and

$$I = Edu^2 + 2Fdudv + Gdv^2, II = Ldu^2 + 2Mdudv + Ndv^2.$$

We denote

$$\mathcal{F}_I = \begin{pmatrix} E & F \\ F & G \end{pmatrix}, \mathcal{F}_{II} = \begin{pmatrix} L & M \\ M & N \end{pmatrix}.$$

Proposition 2.5.7. *Let σ be a parametrization. Then the matrix $\mathcal{W}_p$ with respect to the basis $\{\sigma_u, \sigma_v\}$ of T_pS is $\mathcal{F}_{II}\mathcal{F}_I^{-1} = (\mathcal{F}_I^{-1}\mathcal{F}_{II})^T$.*

Proof. We know that $\mathcal{W}(\sigma_u) = -\mathbf{N}_u, \mathcal{W}(\sigma_v) = -\mathbf{N}_v$. So the matrix of $\mathcal{W}$ is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

where

$$-\mathbf{N}_u = a\sigma_u + b\sigma_v, -\mathbf{N}_v = c\sigma_u + d\sigma_v.$$

Paring each with σ_u, σ_v, we have

$$L = aE + bF, M = aF + bG, M = cE + dF, N = cF + dG,$$

i.e.

$$\mathcal{F}_{II} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mathcal{F}_I$$

□

Corollary 2.5.8.

$$H = \frac{LG - 2MF + NE}{2(EG - F^2)}, K = \frac{LN - M^2}{EG - F^2}.$$

Proof.

$$K = \det(\mathcal{F}_I^{-1}\mathcal{F}_{II}) = \frac{\det \mathcal{F}_{II}}{\det \mathcal{F}_I} = \frac{LN - M^2}{EG - F^2}.$$

$$\begin{aligned} \mathcal{F}_I^{-1}\mathcal{F}_{II} =& \frac{1}{EG - F^2}\begin{pmatrix} G & -F \\ -F & E \end{pmatrix}\begin{pmatrix} L & M \\ M & N \end{pmatrix} \\ =& \frac{1}{EG - F^2}\begin{pmatrix} LG - MF & MG - NF \\ ME - LF & NE - MF \end{pmatrix} \end{aligned}$$

So

$$H = \frac{1}{2} trace(\mathcal{F}_I^{-1}\mathcal{F}_{II}) = \frac{LG - 2MF + NE}{2(EG - F^2)}.$$

□

Example 2.5.9. *Surface of revolution*

$$\sigma(u, v) = (f(u)\cos v, f(u)\sin v, g(u)).$$

We again assume $f_u^2 + g_u^2 = 1$ *and* $f > 0$*.*

$$I = du^2 + f^2 dv^2, II = (f_u g_{uu} - f_{uu} g_u)du^2 + f g_u dv^2.$$

Hence

$$K = \frac{(f_u g_{uu} - f_{uu} g_u) f g_u}{f^2}.$$

Taking derivative on $f_u^2 + g_u^2 = 1$*, we have*

$$f_u f_{uu} + g_u g_{uu} = 0.$$

So

$$(f_u g_{uu} - f_{uu} g_u) g_u = -f_{uu}(f_u^2 + g_u^2) = -f_{uu},$$

and

$$K = -\frac{f_{uu} f}{f^2} = -\frac{f_{uu}}{f}.$$

Especially, for a unit sphere $u = \theta, v = \phi, f(\theta) = \cos\theta, g(\theta) = \sin\theta$*. We thus have* $K = 1$*.*

Gauss uses another way to define K, roughly speaking it is the ratio of the area changed under Gaussian map $\mathcal{G}$, or

$$\lim_{R \to p} \frac{Area(\mathcal{G}(R))}{Area R}.$$

Next theorem makes it precisely.

Theorem 2.5.10. *Let $\sigma : U \to \mathbb{R}^3$ be a parametrization, with $(u_0, v_0) \in U$. Let $R_\delta = \{(u, v) \in \mathbb{R}^2 | (u - u_0)^2 + (v - v_0)^2 \leq \delta^2\}$. Then*

$$\lim_{\delta \to 0} \frac{A_{\mathbf{N}}(R_\delta)}{A_\sigma(R_\delta)} = |K|,$$

where K is the Gaussian at $\sigma(u_0, v_0)$.

Proof. Recall that

$$\frac{A_{\mathbf{N}}(R_\delta)}{A_\sigma(R_\delta)} = \frac{\int_{R_\delta} ||\mathbf{N}_u \times \mathbf{N}_v|| dudv}{\int_{R_\delta} ||\sigma_u \times \sigma_v|| dudv}.$$

$$\begin{aligned}\mathbf{N}_u \times \mathbf{N}_v =& (a\sigma_u + b\sigma_v) \times (c\sigma_u + d\sigma_v) \\ =& (ad - bc)\sigma_u \times \sigma_v \\ =& \det(\mathcal{F}_I^{-1}\mathcal{F}_{II})\sigma_u \times \sigma_v \\ =& K\sigma_u \times \sigma_v\end{aligned}$$

So we could choose δ small, such that $|K(u, v) - K(u_0, v_0)| < \epsilon$ if $(u, v) \in R_\delta$. So

$$|K(u_0, v_0)| - \epsilon < \frac{A_{\mathbf{N}}(R_\delta)}{A_\sigma(R_\delta)} < |K(u_0, v_0)| + \epsilon.$$

This finishes the proof. □

2.5.3 Principal curvatures

Let us come back to principal curvatures. They are the roots κ of $\det(\mathcal{F}_I^{-1}\mathcal{F}_{II} - \kappa I) = 0$, which is

$$\det(\mathcal{F}_{II} - \kappa \mathcal{F}_I) = 0.$$

$\mathbf{t} = \xi\sigma_u + \eta\sigma_v$ is a principal vector if

$$(\mathcal{F}_{II} - \kappa\mathcal{F}_I)\begin{pmatrix}\xi \\ \eta\end{pmatrix} = \begin{pmatrix}0 \\ 0\end{pmatrix}.$$

Example 2.5.11. *For unit sphere*

$$I = II = d\theta^2 + \cos^2\theta d\phi^2.$$

So principal curvatures are repeated roots $\kappa = 1$ and thus every tangent vector is principal, every point is umbilical.

Example 2.5.12. *For cylinder*

$$\sigma(u, v) = (\cos v, \sin v, u).$$

$$I = du^2 + dv^2, II = dv^2.$$

Principal curvatures are solutions of

$$\det\begin{pmatrix} 0-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} = 0.$$

So $\kappa = 0, 1$, and no point is umbilical.

The principal vector

$$\mathbf{t}_1 = \sigma_u = (0,0,1), \mathbf{t}_2 = \sigma_v = (-\sin v, \cos v, 0).$$

Proposition 2.5.13. *Let S be a connected surface of which every point is umbilical. Then S is an open subset of a plane or a sphere.*

Proof. For every tangent vector $\mathbf{t}$, $\mathcal{W}(\mathbf{t}) = \kappa\mathbf{t}$ where κ is the principal curvature. Since $\mathcal{W}(\sigma_u) = -\mathbf{N}_u, \mathcal{W}(\sigma_v) = -\mathbf{N}_v$, then

$$\mathbf{N}_u = -\kappa\sigma_u, \mathbf{N}_v = -\kappa\sigma_v.$$

Hence by taking derivatives,

$$\kappa_v\sigma_u = \kappa_u\sigma_v.$$

Since σ_u and σ_v are linearly independent, $\kappa_u = \kappa_v = 0$. Thus $\kappa \equiv C$.

If $\kappa = 0$, $\mathbf{N}$ is constant. Then $(\mathbf{N}\cdot\sigma)_u = (\mathbf{N}\cdot\sigma)_v = 0$, so $\mathbf{N}\cdot\sigma \equiv C$. Thus $\sigma(U)$ is an open subset of the plane $P\cdot\mathbf{N} \equiv C$.

If $\kappa \neq 0$, $\mathbf{N} = -\kappa\sigma + \mathbf{a}$. Hence

$$||\sigma - \frac{1}{\kappa}\mathbf{a}||^2 = ||-\frac{1}{\kappa}\mathbf{N}||^2 = \frac{1}{\kappa^2}.$$

So $\sigma(U)$ is an open subset of the sphere with centre $\kappa^{-1}\mathbf{a}$ and radius $|\kappa|^{-1}$.

To complete the proof, notice that each patch is contained in a plane or a sphere. But if the images of two patches intersect, they must clearly be part of the same plane or same sphere. So complete the proof. □

Principal curvature at $p \in S$ provides the information about shape. We choose the coordinates as following: p is the origin, T_pS is the xy-plane in $\mathbb{R}^3$, principal vectors $\mathbf{t}_1 = (1,0,0)$ and $\mathbf{t}_2 = (0,1,0)$ and $\mathbf{N} = (0,0,1)$. We could always choose such a coordinate up to an isometry, i.e. rotation and translation, of $\mathbb{R}^3$.

Let σ be a parametrization with $\sigma(0,0) = 0$ (point p). The tangent plane is $\{(x,y,0)\} = s\sigma_u(0,0) + t\sigma_v(0,0)$. Taylor expansion gives us

$$\sigma(s,t) = \sigma(0,0)+s\sigma_u(0,0)+t\sigma_v(0,0)+\frac{1}{2}(s^2\sigma_{uu}(0,0)+2st\sigma_{uv}(0,0)+t^2\sigma_{vv}(0,0))+\cdots$$

If x, y (hence s, t) are small, we have $\sigma(s,t) \approx (x,y,z)$ where

$$z \approx \frac{1}{2}(s^2\sigma_{uu}(0,0) + 2st\sigma_{uv}(0,0) + t^2\sigma_{vv}(0,0))\cdot\mathbf{N} = \frac{1}{2}(Ls^2 + 2Mst + Nt^2).$$

Since

$$\mathcal{W}(\mathbf{t}) = x\mathcal{W}(\mathbf{t}_1) + y\mathcal{W}(\mathbf{t}_2) = \kappa_1 x\mathbf{t}_1 + \kappa_2 y\mathbf{t}_2 = (\kappa_1 x, \kappa_2 y, 0)$$

for $\mathbf{t} = (x, y, 0)$, hence

$$Ls^2 + 2Mst + Nt^2 =< \mathcal{W}(\mathbf{t}), \mathbf{t} >= \kappa_1 x^2 + \kappa_2 y^2.$$

So near p, S is approximated by $z = \frac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2)$.

There are 4 cases of local behaviour:

1. *Elliptic* if $K_p > 0$, so $z = \frac{1}{2}(\kappa_1 x^2 + \kappa_2 y^2)$ is an elliptic paraboloid.
2. *Hyperbolic* if $K_p < 0$, it is a hyperbolic paraboloid.
3. *Parabolic* if one of κ_1, κ_2 is zero, and the other is non-zero. It is a parabolic cylinder.
4. *Planar* if both $\kappa_1 = \kappa_2 = 0$ (or $\mathcal{W}_p \equiv 0$). We need higher derivatives to know the shape.

2.6 Gauss's Theorema Egregium

Since the definitions of curvatures involve the second fundamental form, they are usually not intrinsic. But actually Gaussian curvature K is an intrinsic invariant.

Theorem 2.6.1 (Gauss's Theorema Egregium)**.** *The Gaussian curvature K of a surface is invariant of the first fundamental form.*

In this section, we prove it by detailed calculations.

For regular surface S, and a chart $\sigma : U \to S$, $\sigma_u, \sigma_v, \mathbf{N}$ would be a basis. We express $\sigma_{uu}, \sigma_{uv}, \sigma_{vv}$ by

$$\sigma_{uu} = \Gamma^1_{11}\sigma_u + \Gamma^2_{11}\sigma_v + L_1 \cdot \mathbf{N} \tag{2.2}$$

$$\sigma_{uv} = \Gamma^1_{12}\sigma_u + \Gamma^2_{12}\sigma_v + L_2 \cdot \mathbf{N} \tag{2.3}$$

$$\sigma_{vv} = \Gamma^1_{22}\sigma_u + \Gamma^2_{22}\sigma_v + L_3 \cdot \mathbf{N} \tag{2.4}$$

Here Γ^k_{ij} are called *Christoffel symbols.*

First, by taking dot product with $\mathbf{N}$, $L_1 = L, L_2 = M, L_3 = N$.

Next, we claim Γ^k_{ij} only depends on the first fundamental form. More precisely,

$$\begin{aligned}\Gamma^1_{11} = \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}, \Gamma^2_{11} = \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)} \\ \Gamma^1_{12} = \frac{GE_v - FG_u}{2(EG - F^2)}, \Gamma^2_{12} = \frac{EG_u - FE_v}{2(EG - F^2)} \\ \Gamma^1_{22} = \frac{2GF_v - GG_u - FG_v}{2(EG - F^2)}, \Gamma^2_{22} = \frac{EG_v - 2FF_v + FG_u}{2(EG - F^2)}\end{aligned} \tag{2.5}$$

They are determined by following set of equations

$$\begin{cases} (2.2)\cdot\sigma_u: & \Gamma_{11}^1\cdot E+\Gamma_{11}^2\cdot F=\sigma_{uu}\cdot\sigma_u=\frac{1}{2}E_u \\ (2.2)\cdot\sigma_v: & \Gamma_{11}^1\cdot F+\Gamma_{11}^2\cdot G=\sigma_{uu}\cdot\sigma_v=(\sigma_u\cdot\sigma_v)_u-\sigma_u\cdot\sigma_{uv}=F_u-\frac{1}{2}E_v \\ (2.3)\cdot\sigma_u: & \Gamma_{12}^1\cdot E+\Gamma_{12}^2\cdot F=\sigma_{uv}\cdot\sigma_u=\frac{1}{2}E_v \\ (2.3)\cdot\sigma_v: & \Gamma_{12}^1\cdot F+\Gamma_{12}^2\cdot G=\sigma_{uv}\cdot\sigma_v=\frac{1}{2}G_u \\ (2.4)\cdot\sigma_u: & \Gamma_{22}^1\cdot E+\Gamma_{22}^2\cdot F=\sigma_{vv}\cdot\sigma_u=F_v-\frac{1}{2}G_u \\ (2.4)\cdot\sigma_v: & \Gamma_{22}^1\cdot F+\Gamma_{22}^2\cdot G=\sigma_{vv}\cdot\sigma_v=\frac{1}{2}G_v \end{cases}$$

Finally, recall that we have determined

$$\mathbf{N}_u=a_{11}\sigma_u+a_{21}\sigma_v$$

$$\mathbf{N}_v=a_{12}\sigma_u+a_{22}\sigma_v$$

in Proposition 2.5.7, where

$$(a_{ij})_{2\times 2}=-\mathcal{F}_I^{-1}\mathcal{F}_{II}=-\frac{1}{EG-F^2}\begin{pmatrix}LG-MF & MG-NF\\ ME-LF & NE-MF\end{pmatrix}.$$

Especially $K=\frac{LN-M^2}{EG-F^2}$.

Now Gauss's Theorema Egregium follows from any of the following equations.

Proposition 2.6.2 (Gauss equations)**.**

$$\begin{aligned} EK=&(\Gamma_{11}^2)_v-(\Gamma_{12}^2)_u+\Gamma_{11}^1\Gamma_{12}^2+\Gamma_{11}^2\Gamma_{22}^2-\Gamma_{12}^1\Gamma_{11}^2-(\Gamma_{12}^2)^2\\ FK=&(\Gamma_{12}^1)_u-(\Gamma_{11}^1)_v+\Gamma_{12}^2\Gamma_{12}^1-\Gamma_{11}^2\Gamma_{22}^1\\ =&(\Gamma_{12}^2)_v-(\Gamma_{22}^2)_u+\Gamma_{12}^1\Gamma_{12}^2-\Gamma_{22}^1\Gamma_{11}^2\\ GK=&(\Gamma_{22}^1)_u-(\Gamma_{12}^1)_v+\Gamma_{22}^1\Gamma_{11}^1+\Gamma_{22}^2\Gamma_{12}^1-(\Gamma_{12}^1)^2-\Gamma_{12}^2\Gamma_{22}^1 \end{aligned}$$

Proof. For first "FK": $(\sigma_{uu})_v=(\sigma_{uv})_u$. Then

$$(\Gamma_{11}^1\sigma_u+\Gamma_{11}^2\sigma_v+L\mathbf{N})_v=(\Gamma_{12}^1\sigma_u+\Gamma_{12}^2\sigma_v+M\mathbf{N})_u,$$

which is

$$\begin{aligned} &((\Gamma_{11}^1)_v-(\Gamma_{12}^1)_u)\sigma_u+((\Gamma_{11}^2)_v-(\Gamma_{12}^2)_u)\sigma_v+(L_v-M_u)\mathbf{N}\\ =&\Gamma_{12}^1\sigma_{uu}+(\Gamma_{12}^2-\Gamma_{11}^1)\sigma_{uv}-\Gamma_{11}^2\sigma_{vv}-L\mathbf{N}_v+M\mathbf{N}_u\\ =&\Gamma_{12}^1(\Gamma_{11}^1\sigma_u+\Gamma_{11}^2\sigma_v+L\mathbf{N})+(\Gamma_{12}^2-\Gamma_{11}^1)(\Gamma_{12}^1\sigma_u+\Gamma_{12}^2\sigma_v+M\mathbf{N})\\ &-\Gamma_{11}^2(\Gamma_{22}^1\sigma_u+\Gamma_{22}^2\sigma_v+N\mathbf{N})-L(a_{12}\sigma_u+a_{22}\sigma_v)+M(a_{11}\sigma_u+a_{21}\sigma_v) \end{aligned}$$

Comparing the coefficients of σ_u:

$$(\Gamma_{11}^1)_v-(\Gamma_{12}^1)_u=\Gamma_{12}^2\Gamma_{12}^1-\Gamma_{11}^2\Gamma_{22}^1-La_{12}+Ma_{11}$$

where $a_{11}=\frac{MF-LG}{EG-F^2}, a_{12}=\frac{NF-MG}{EG-F^2}$.

Hence

$$\begin{aligned} La_{12} - Ma_{11} =& \frac{1}{EG - F^2}(L(NF - MG) - M(MF - LG)) \\ =& \frac{F(LN - M^2)}{EG - F^2} = FK \end{aligned}$$

This completes the proof of first "FK".

For the others:

"EK": equating coefficients of σ_v in $(\sigma_{uu})_v = (\sigma_{uv})_u$.
Second 'FK": equating coefficients of σ_u in $(\sigma_{uv})_v = (\sigma_{vv})_u$.
"GK": equating coefficients of σ_v in $(\sigma_{uv})_v = (\sigma_{vv})_u$. □

Notice if we substitute Equations (2.5) of Christoffel symbols into the 4 equations, we have the same Gaussian curvature:

$$\begin{aligned} K = \frac{-1}{2\sqrt{EG - F^2}}((\frac{E_v - F_u}{\sqrt{EG - F^2}})_v - (\frac{F_v - G_u}{\sqrt{EG - F^2}})_u) \\ - \frac{1}{4(EG - F^2)^2} \cdot \det \begin{pmatrix} E & E_u & E_v \\ F & F_u & F_v \\ G & G_u & G_v \end{pmatrix} \end{aligned}$$

Abstractly, the Gaussian curvature has the following form

$$K = -\frac{1}{2(EG - F^2)}(E_{vv} - 2F_{uv} + G_{uu}) + \mathcal{R}$$

where the remainder $\mathcal{R}$ is quadratic in the first derivatives of E, F, G.

What if we compare coefficients of $\mathbf{N}$ in both equations? We would have the following

Proposition 2.6.3 (Codazzi-Mainardi Equations)**.**

$$L_v - M_u = L\Gamma_{12}^1 + M(\Gamma_{12}^2 - \Gamma_{11}^1) - N\Gamma_{11}^2$$

$$M_v - N_u = L\Gamma_{22}^1 + M(\Gamma_{22}^2 - \Gamma_{12}^1) - N\Gamma_{12}^2$$

Proof. The first equation is obtained by comparing the coefficients of $\mathbf{N}$ in $(\sigma_{uu})_v = (\sigma_{uv})_u$. The second is from $(\sigma_{uv})_v = (\sigma_{vv})_u$. □

Along with Equations (2.5) of Christoffel symbols, Gauss and Codazzi-Mainardi equations are the equations of the coefficients of the first and second fundamental forms. They are called the *compatible equations* of surfaces. Actually, these are only relations of compatibility between the first and the second fundamental forms by a theorem of Bonnet. This should be compared with the fundamental theorem of curves Theorem 1.7.3.

Theorem 2.6.4 (Bonnet)**.** *Let $Edu^2+2Fdudv+Gdv^2$ and $Ldu^2+2Mdudv+Ndv^2$ be two arbitrary fundamental forms in an open set $U \subset \mathbb{R}^2$. And $E>0, G>0, EG-F^2>0$. If the coefficients of these fundamental forms satisfy the Gauss and Codazzi-Mainardi equations, then there is a unique, up to a rigid motion of the space $\mathbb{R}^3$, surface $\sigma : U \to \mathbb{R}^3$ for which these forms are the first and the second fundamental forms, respectively.*

2.6.1 Gaussian curvature for special cases

There are several specially cases we have most interests.

Corollary 2.6.5. *Let the first fundamental form be $Edu^2+2Fdudv+Gdv^2$.*

1. *If $F=0$, then*
$$K = -\frac{1}{2\sqrt{EG}}\left(\frac{\partial}{\partial u}\left(\frac{G_u}{\sqrt{EG}}\right) + \frac{\partial}{\partial v}\left(\frac{E_v}{\sqrt{EG}}\right)\right).$$
2. *If $E=1, F=0$,*
$$K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial u^2}.$$

Proof. 1 implies 2 because $K = -\frac{1}{2\sqrt{G}}\frac{\partial}{\partial u}(\frac{G_u}{\sqrt{G}}) = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial u^2}$.

To prove 1, we calculate the Christoffel symbols:
$$\Gamma^1_{11} = \frac{E_u}{2E}, \Gamma^2_{11} = -\frac{E_v}{2G}, \Gamma^1_{12} = \frac{E_v}{2E}, \Gamma^2_{12} = \frac{G_u}{2G}, \Gamma^1_{22} = -\frac{G_u}{2E}, \Gamma^2_{22} = \frac{G_v}{2G}.$$
So
$$EK = -\left(\frac{E_v}{2G}\right)_v - \left(\frac{G_u}{2G}\right)_u + \frac{E_uG_u}{4EG} - \frac{E_vG_v}{4G^2} + \frac{E_v^2}{4EG} - \frac{G_u^2}{4G^2}.$$
i.e.
$$\begin{aligned}
-2K\sqrt{EG} =& \frac{E_{vv}+G_{uu}}{\sqrt{EG}} - \frac{E_v(EG_v+E_vG)}{2\sqrt{(EG)^3}} - \frac{G_u(E_uG+EG_u)}{2\sqrt{(EG)^3}} \\
=& \frac{E_{vv}}{\sqrt{EG}} - \frac{1}{2}\frac{E_v(EG)_v}{\sqrt{(EG)^3}} + \frac{G_{uu}}{\sqrt{EG}} - \frac{1}{2}\frac{G_u(EG)_u}{\sqrt{(EG)^3}} \\
=& \left(\frac{E_v}{\sqrt{EG}}\right)_v + \left(\frac{G_u}{\sqrt{EG}}\right)_u
\end{aligned}$$
□

Example 2.6.6. *Surface of revolution*
$$\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u)).$$
We again assume $f_u^2+g_u^2=1$ and $f>0$. We have $E=1, F=0, G=f^2(u)$. So
$$K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial u^2} = -\frac{f_{uu}}{f}.$$
This is the same as our previous calculation.

2.7 Surfaces of constant Gaussian curvature

Let us look at surface of revolution

$$\sigma(u,v) = (f(u)\cos v, f(u)\sin v, g(u))$$

with $f_u^2 + g_u^2 = 1$ and $f > 0$. We have shown

$$K = -\frac{f_{uu}}{f}.$$

We want to look at these surfaces with constant Gaussian curvature. There are three cases.

Before that, let us first summarize the effect of dilation

$$(x,y,z) \mapsto (ax, ay, az), a \neq 0.$$

E, F, G are multiplied by a^2; L, M, N are multiplied by a; H is multiplied by a^{-1}; K is multiplied by a^{-2}. So the constant Gaussian curvatures are reduced to the cases of $K = 1, 0, -1$.

1. $K = 0$ everywhere.

So $f_{uu} = 0$ or $f(u) = au + b$. Since $f_u^2 + g_u^2 = 1$, we know $|a| \leq 1$ and $g_u = \pm\sqrt{1-a^2}$. Hence

$$g(u) = \pm\sqrt{1-a^2}u + C.$$

We could assume $C = 0$ by translating along z-axis, and the sign is + by rotating degree π for the profile curve if necessary. Then we have the surface

$$\sigma(u,v) = (b\cos v, \sin v, 0) + u(a\cos v, a\sin v, \sqrt{1-a^2}).$$

This is a *ruled surface* in the sense that this is a union of straight lines (the v-curves $v \equiv c$ gives straight lines).

2. $K > 0$ and $K = -\frac{1}{R^2}$.

We have the differential equation

$$f_{uu} + \frac{f}{R^2} = 0.$$

So $f(u) = a\cos(\frac{u}{R} + b)$. By reparametrization $\tilde{u} = u + Rb$, we could assume $b = 0$ in the expression. Then $g(u) = \int \sqrt{1 - \frac{a^2}{R^2}\sin^2\frac{u}{R}}du$. If $a = R$, we have $f(u) = R\cos\frac{u}{R}$, $g(u) = R\sin\frac{u}{R}$. It is a sphere of radius R.

3. $K < 0$. Up to a dilation of $\mathbb{R}^3$, we could assume $K = -1$.

We have $f_{uu} - f = 0$ whose solutions are

$$f(u) = ae^u + be^{-u}.$$

$g(u)$ could be expressed in terms of elementary functions only if one of a or b is zero. If $b=0$, we could assume $a=1$ by translation $\tilde{u}=u+c$. (It is similar for $a=0$.) In this case, $f(u)=e^u$, and

$$g(u)=\int\sqrt{1-e^{2u}}du.$$

We must have $u\leq 0$, so we assume $\sin\theta=e^u>0$.

$$\begin{aligned}\int\sqrt{1-e^{2u}}du &= \int\frac{\cos^2\theta}{\sin\theta}d\theta=-\int\frac{\cos^2\theta}{\sin^2\theta}d\cos\theta\\ &=\cos\theta-\int\frac{1}{1-\cos^2\theta}d\cos\theta\\ &=\cos\theta-\frac{1}{2}\ln(\cos\theta+1)+\frac{1}{2}\ln(1-\cos\theta)\\ &=\cos\theta-\frac{1}{2}\ln\frac{(\cos\theta+1)^2}{\sin^2\theta}\\ &=\cos\theta-\ln(\csc\theta+\cot\theta)\\ &=\sqrt{1-e^{2u}}-\ln(e^{-u}+\sqrt{e^{-2u}-1})\end{aligned}$$

Recall that $\cosh^{-1}(v)=\ln(v+\sqrt{v^2-1})$, the profile curve in xz-plane is

$$z=\sqrt{1-x^2}-\cosh^{-1}(\frac{1}{x}).$$

This is the *tractrix* we met in Example Sheet 1. The condition $u\leq 0$ read as $0<x\leq 1$. This surface of revolution is called a *pseudo-sphere*. It is amusing to calculate the Gaussian curvature in terms of principal curvatures. One could calculate that the curvature of the tractrix is $\tan\theta$. Since $\mathbf{n}=-\mathbf{N}$ we know $\mathbf{N}_u$ is parallel to σ_u which is one of the principal directions, and we have $\kappa_1=\kappa_n=\kappa\cos\psi=-\tan\theta$. For the circle, we know the radius is $\sin\theta$ thus the curvature is $\csc\theta$. But the angle between $\mathbf{n}$ and $\mathbf{N}$ is θ. So $\kappa_2=\kappa_n=\kappa\cos\psi=\tan\theta$. Thus $K=\kappa_1\kappa_2=-1$.

Notice that in above calculation, the surface of constant curvatures are more than one might expect. However locally, plane, sphere, pseudo-sphere are the only possibilities.

Proposition 2.7.1. *Any point of a surface of constant Gaussian curvature is contained in a patch that is isometric to an open subset of a plane, a sphere, or a pseudo-sphere.*

We will prove this theorem in a later stage.

Notice a plane or a pseudo-sphere is not compact. For compact surface, we have the following stronger result.

Theorem 2.7.2. *Every connected compact surface whose Gaussian curvature is constant is a sphere.*

The proof consists the following parts.

Proposition 2.7.3. *Suppose $S \subset \mathbb{R}^3$ is a compact surface. Then there is a point $p \in S$ with $K(p) > 0$.*

Proof. Because S is compact, the continuous function $f(\mathbf{x}) = ||\mathbf{x}||$ achieves maximum at $p \in S$. Let $f(p) = R$. So any curve $\alpha \subset S$ at p has curvature at least $\frac{1}{R}$ (Proposition 1.9.3). So every normal curvature is no less than $\frac{1}{R}$, and so $K(p) \geq \frac{1}{R^2} > 0$. □

Proposition 2.7.4. *Suppose p is not an umbilical point of S. Then there is a "principal coordinate" near p, so that u-curves (v-curves respectively) are lines of curvature with principal curvature κ_1 (κ_2 resp.). Especially,*

$$I = Edu^2 + Gdv^2, II = Ldu^2 + Ndv^2.$$

Proof. Need the fundamental theorem of ODE. But the idea is the following: principal vectors $\mathbf{t}_1$ of κ_1 and $\mathbf{t}_2$ of κ_2 form well-defined vector fields X and Y near p, since p is not umbilical. These are the direction fields of some ODE $\frac{d\mathbf{x}}{dt} = \mathbf{t}_i$ and the integral curves are u-curves and v-curves. □

Proposition 2.7.5. *p is not umbilical and $\kappa_1(p) > \kappa_2(p)$. Suppose κ_1 has a local maximum at p and κ_2 has a local minimum at p. Then $K(p) \leq 0$.*

Proof. Choose principal coordinate, by switching u and v if necessary, we could assume $\kappa_1 = \frac{L}{E}$ and $\kappa_2 = \frac{N}{G}$. We know $(\kappa_1)_v = (\kappa_2)_u = 0$, so

$$E_v = -\frac{2E}{\kappa_1 - \kappa_2}(\kappa_1)_v = 0, G_u = \frac{2G}{\kappa_1 - \kappa_2}(\kappa_2)_u = 0.$$

The equality follows from an exercise in Example Sheet 4.

Under our coordinates,

$$\begin{aligned} K &= -\frac{1}{2\sqrt{EG}}(\frac{\partial}{\partial u}(\frac{G_u}{\sqrt{EG}}) + \frac{\partial}{\partial v}(\frac{E_v}{\sqrt{EG}})) \\ &= -\frac{1}{2EG}(G_{uu} + E_{vv}) \\ &= -\frac{1}{2EG}(\frac{2G}{\kappa_1 - \kappa_2}(\kappa_2)_{uu} - \frac{2E}{\kappa_1 - \kappa_2}(\kappa_1)_{vv}) \end{aligned}$$

Since κ_1 is a local maximum at p, we know $(\kappa_1)_{vv} \leq 0$. And κ_2 is a local minimum, so $(\kappa_2)_{uu} \geq 0$. Hence $K \leq 0$. □

Proof. (of the theorem 2.7.2) Since S is compact and K is constant, we have $K > 0$ by Proposition 2.7.3. We assume $\kappa_1 \geq \kappa_2$. By compactness, κ_1 reaches maximum at some $p \in S$. Since $K = \kappa_1\kappa_2$ is a positive constant, κ_2 reaches minimum at the same point p. By Proposition 2.7.5, p is umbilical, i.e $\kappa_1(p) = \kappa_2(p)$.

Then for any other point $q \in S$. we have

$$\kappa_1(p) \geq \kappa_1(q) \geq \kappa_2(q) \geq \kappa_2(p) = \kappa_1(p).$$

This implies $\kappa_1(q) = \kappa_2(q)$, i.e. every point of S is umbilical. By Proposition 2.5.13, we know this is a sphere. □

2.8 Parallel transport and covariant derivative

Running with constant velocity on the earth is actually not a constant move in $\mathbb{R}^3$, since the directions keep changing. But it should be a constant move in certain sense since that is what we feel as the runner. To achieve this, we need to compare velocity at each tangent plane.

The velocity vectors form a tangent vector field along its trajectory, i.e a curve γ on a surface S. Here, by a *tangent vector field* $\mathbf{v}$ along γ, we mean a smooth map from the interval (a, b) to $\mathbb{R}^3$ such that $\mathbf{v}(t) \in T_{\gamma(t)}S$ for all $t \in (a, b)$.

Denote the rate of change of $\mathbf{v}$ in $\mathbb{R}^3$ by $\dot{\mathbf{v}}$. In our previous example, the tangent vectors of a great circle on a sphere, $\dot{\mathbf{v}}$ is the normal direction. So, we are interested in its tangent component

$$\nabla_\gamma \mathbf{v} = \dot{\mathbf{v}} - (\dot{\mathbf{v}} \cdot \mathbf{N})\mathbf{N}.$$

One can check that $\nabla_\gamma \mathbf{v} \cdot \mathbf{N} = 0$. It is defined on any surface, orientable or not, since this is unchanged if $\mathbf{N}$ is replaced by $-\mathbf{N}$.

This $\nabla_\gamma \mathbf{v}$ is called the *covariant derivative* of $\mathbf{v}$ along γ.

Definition 2.8.1. $\mathbf{v}$ *is said to be parallel along* γ *if* $\nabla_\gamma \mathbf{v} = 0$ *at every point of* γ.

By definition, $\mathbf{v}$ is parallel along γ if and only if $\dot{\mathbf{v}} \perp T_{\gamma(t)}S$.

Here is a local calculation of this condition.

Proposition 2.8.2. *Let* $\gamma(t) = \sigma(u(t), v(t))$ *be a curve, and* $\mathbf{w}(t) = \alpha(t)\sigma_u + \beta(t)\sigma_v$ *be a tangent vector field along* γ. *Then* $\mathbf{w}$ *is parallel long* γ *if and only if*

$$\dot{\alpha} + (\Gamma^1_{11}\dot{u} + \Gamma^1_{12}\dot{v})\alpha + (\Gamma^1_{12}\dot{u} + \Gamma^1_{22}\dot{v})\beta = 0$$
$$\dot{\beta} + (\Gamma^2_{11}\dot{u} + \Gamma^2_{12}\dot{v})\alpha + (\Gamma^2_{12}\dot{u} + \Gamma^2_{22}\dot{v})\beta = 0$$

Proof.

$$\begin{aligned}\dot{\mathbf{w}} =& \dot{\alpha}\sigma_u + \dot{\beta}\sigma_v + \alpha(\dot{u}\sigma_{uu} + \dot{v}\sigma_{uv}) + \beta(\dot{u}\sigma_{uv} + \dot{v}\sigma_{vv}) \\ =& \dot{\alpha}\sigma_u + \dot{\beta}\sigma_v + \alpha\dot{u}(\Gamma^1_{11}\sigma_u + \Gamma^2_{11}\sigma_v + L\mathbf{N}) \\ &+ (\alpha\dot{v} + \beta\dot{u})(\Gamma^1_{12}\sigma_u + \Gamma^2_{12}\sigma_v + M\mathbf{N}) + \beta\dot{v}(\Gamma^1_{22}\sigma_u + \Gamma^2_{22}\sigma_v + N\mathbf{N})\end{aligned}$$

Hence

$$\begin{aligned}\nabla_\gamma \mathbf{w} =& (\dot\alpha + (\Gamma_{11}^1 \dot u + \Gamma_{12}^1 \dot v)\alpha + (\Gamma_{12}^1 \dot u + \Gamma_{22}^1 \dot v)\beta)\sigma_u \\ &+ (\dot\beta + (\Gamma_{11}^2 \dot u + \Gamma_{12}^2 \dot v)\alpha + (\Gamma_{12}^2 \dot u + \Gamma_{22}^2 \dot v)\beta)\sigma_v\end{aligned}$$

And the result follows. □

Examples:

1. When in the plane, $E = G = 1, F = 0$. So $\Gamma_{ij}^k = 0$. And $\nabla_\gamma \mathbf{v} = \dot\alpha\sigma_u + \dot\beta\sigma_v$ is the usual vector derivative.

2. S is a unit sphere with latitude-longitude parametrization

$$\sigma(\theta, \phi) = (\cos\phi\cos\theta, \sin\phi\cos\theta, \sin\theta).$$

So

$$I = d\theta^2 + \cos^2\theta d\phi^2$$

$$\Gamma_{11}^1 = \Gamma_{11}^2 = \Gamma_{22}^2 = \Gamma_{12}^1 = 0, \Gamma_{12}^2 = -\tan\theta, \Gamma_{22}^1 = \sin\theta\cos\theta.$$

Let the curve γ be a circle of latitude $\theta = \theta_0$, i.e

$$\gamma(\phi) = \sigma(\theta_0, \phi), -\frac{\pi}{2} < \theta_0 < \frac{\pi}{2}.$$

So the equations become

$$\dot\alpha = -\beta\sin\theta_0\cos\theta_0, \dot\beta = \alpha\tan\theta_0.$$

If $\theta_0 = 0$, α, β are constants. Especially, the tangent field σ_ϕ is parallel along $\theta_0 = 0$.

If $\theta_0 \neq 0$,

$$\ddot\alpha + \alpha\sin^2\theta_0 = 0,$$

so

$$\begin{aligned}\alpha(\phi) &= A\cos(\phi\sin\theta_0) + B\sin(\phi\sin\theta_0),\\ \beta(\phi) &= A\frac{\sin(\phi\sin\theta_0)}{\cos\theta_0} - B\frac{\cos(\phi\sin\theta_0)}{\sin\theta_0}.\end{aligned}$$

We consider $\mathbf{w}(0) = \sigma_\phi$ which is tangent to $\gamma(0)$. Then $\alpha(0) = 0, \beta(0) = 1$ and $A = 0, B = -\sin\theta_0$. Hence

$$\mathbf{w}(\phi) = -\sin\theta_0\sin(\phi\sin\theta_0)\sigma_\theta + \cos(\phi\sin\theta_0)\sigma_\phi.$$

In general $\mathbf{w}(\phi)$ is not the tangent vector σ_ϕ of γ.

A natural question motivated from the previous example is: when the tangent vector $\dot\gamma$ is parallel along γ? Or when we drive along a "straight" road on a surface?

It is measured by the tangent component of the acceleration $\ddot\gamma$.

2.9 Geodesics

Definition 2.9.1. *A curve γ on a surface S is called a geodesic if $\ddot{\gamma}(t)$ is zero or perpendicular to the tangent plane of surface at $\gamma(t)$, i.e. parallel to its unit normal.*

By definition, γ is a geodesic if and only if $\dot{\gamma}$ is parallel along γ, i.e. when $\nabla_\gamma \dot{\gamma} = 0$.

Proposition 2.9.2. *Let $\mathbf{w}(t)$ and $\mathbf{v}(t)$ be parallel vector fields along $\gamma : I \to S$. Then $\mathbf{w}(t) \cdot \mathbf{v}(t)$ is constant, in particular, $||\mathbf{w}(t)||$ and $||\mathbf{v}(t)||$ are constant, and the angle between them is constant.*

Proof. We have $\frac{d}{dt}\mathbf{w}(t) \cdot \mathbf{v}(t) = \dot{\mathbf{w}}(t) \cdot \mathbf{v}(t) + \mathbf{w}(t) \cdot \dot{\mathbf{v}}(t)$.

$\mathbf{w}(t)$ is parallel along γ means $\dot{\mathbf{w}}(t) \perp T_{\gamma(t)}S$. So $\dot{\mathbf{w}}(t) \cdot \mathbf{v}(t) = 0$. Similarly, $\mathbf{w}(t) \cdot \dot{\mathbf{v}}(t) = 0$.

So $\mathbf{w}(t) \cdot \mathbf{v}(t)$ is a constant. □

By taking $\mathbf{v}(t) = \mathbf{w}(t) = \dot{\gamma}(t)$, we have the following.

Corollary 2.9.3. *Any geodesic has constant speed.*

Thus we could always choose unit-speed geodesic if we want, because if $||\dot{\gamma}|| = \lambda$, $\tilde{\gamma}(t) = \gamma(\frac{t}{\lambda})$ does so.

Recall we have defined geodesic curvature $\kappa_g = \ddot{\gamma} \cdot (\mathbf{N} \times \dot{\gamma})$ (if γ is unit-speed), along with the normal curvature $\kappa_n = \ddot{\gamma} \cdot \mathbf{N}$. It gives another equivalent definition of geodesic.

Proposition 2.9.4. *A unit-speed curve on S is a geodesic if and only if $\kappa_g \equiv 0$.*

Proof. "$\Rightarrow$": It is clear since $\mathbf{N} \times \dot{\gamma} \in T_{\gamma(t)}S$.

"$\Leftarrow$": $\ddot{\gamma} \perp \mathbf{N} \times \dot{\gamma}$ since $\kappa_g = 0$; $\ddot{\gamma} \perp \dot{\gamma}$ since γ is unit-speed. So $\ddot{\gamma} \perp T_{\gamma(t)}S$ or $\ddot{\gamma} \parallel \mathbf{N}$. □

2.9.1 General facts for geodesics

Here comes several simple examples of geodesics.

Fact 1: Any part of straight line on a surface is a geodesic.

Proof. $\gamma(t) = \mathbf{a} + \mathbf{b}t$, so $\ddot{\gamma} = 0$. □

Example 2.9.5. *Lines in the plane; rulings of any ruled surface, such as those of a cylinder or a cone.*

Fact 2: Any normal section of a surface is a geodesic. Here a *normal section* means the intersection C of S with a plane Π, such that $\Pi \perp T_pS$ for any $p \in C$.

Proof. By Meusnier's Theorem 2.5.3, $\kappa_n = \kappa_\theta \sin\theta = \pm\kappa(C)$. So $\kappa_g = 0$ since $\kappa^2 = \kappa_g^2 + \kappa_n^2$. □

Example 2.9.6. *1. All great circles on a sphere.*

2. The intersection of a generalized cylinder with a plane which is perpendicular to the ruling.

Question 2.9.7. *Do we have other geodesics on a sphere other than (part of) great circles?*

We need the following geodesic equations.

Theorem 2.9.8. *A curve γ on a surface S is a geodesic if and only if for $\gamma(t) = \sigma(u(t), v(t))$ and $I_\sigma = Edu^2 + 2Fdudv + Gdv^2$, the following two equations hold*

$$\frac{d}{dt}(E\dot{u} + F\dot{v}) = \frac{1}{2}(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2)$$

$$\frac{d}{dt}(F\dot{u} + G\dot{v}) = \frac{1}{2}(E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2)$$

There is another equivalent form.

Theorem 2.9.9. *A curve γ on a surface S is a geodesic if and only if for $\gamma(t) = \sigma(u(t), v(t))$, the following equations hold*

$$\ddot{u} + \Gamma^1_{11}\dot{u}^2 + 2\Gamma^1_{12}\dot{u}\dot{v} + \Gamma^1_{22}\dot{v}^2 = 0$$

$$\ddot{v} + \Gamma^2_{11}\dot{u}^2 + 2\Gamma^2_{12}\dot{u}\dot{v} + \Gamma^2_{22}\dot{v}^2 = 0$$

Proof. We have shown as Proposition 2.8.2 that $\mathbf{w}(t) = \alpha(t)\sigma_u + \beta(t)\sigma_v$ is parallel long γ if and only if

$$\dot{\alpha} + (\Gamma^1_{11}\dot{u} + \Gamma^1_{12}\dot{v})\alpha + (\Gamma^1_{12}\dot{u} + \Gamma^1_{22}\dot{v})\beta = 0$$

$$\dot{\beta} + (\Gamma^2_{11}\dot{u} + \Gamma^2_{12}\dot{v})\alpha + (\Gamma^2_{12}\dot{u} + \Gamma^2_{22}\dot{v})\beta = 0$$

And by definition, γ is a geodesic if and only if $\dot{\gamma}$ is parallel along γ. Since $\dot{\gamma} = \dot{u}\sigma_u + \dot{v}\sigma_v$, we have the two equations in the theorem. □

Now we prove Theorem 2.9.8 is equivalent to Theorem 2.9.9

Proof. Two equations in Theorem 2.9.8 is equivalent to

$$E\ddot{u} + F\ddot{v} + \frac{1}{2}E_u\dot{u}^2 + E_v\dot{u}\dot{v} + (F_v - \frac{1}{2}G_u)\dot{v}^2 = 0$$

$$F\ddot{u} + G\ddot{v} + (F_u - \frac{1}{2}E_v)\dot{u}^2 + G_u\dot{u}\dot{v} + \frac{1}{2}G_v\dot{v}^2 = 0$$

Solve $\ddot{u}$ and $\ddot{v}$, we have the following set of equations which are equivalent to the above two.

$$(EG-F^2)\ddot{u}+\frac{1}{2}(E_uG+FE_v-2FF_u)\dot{u}^2+(E_vG-G_uF)\dot{u}\dot{v}+(GF_v-\frac{1}{2}G_uG-\frac{1}{2}FG_v)\dot{v}^2=0$$

$$(EG-F^2)\ddot{v}+\frac{1}{2}(2EF_u-E_vE-E_uF)\dot{u}^2+(G_uE-E_vF)\dot{u}\dot{v}+\frac{1}{2}(G_vE-2F_vF+G_uF)\dot{v}^2=0$$

Comparing with (2.5), we know these are exactly the equations in Theorem 2.9.9. □

Example 2.9.10. *Let us determine all geodesics on S^2. We have*

$$\sigma(\theta,\phi)=(\cos\theta\cos\phi,\cos\theta\sin\phi,\sin\theta)$$

$$I=d\theta^2+\cos^2\theta d\phi^2$$

Assume $\gamma(t)=\sigma(\theta(t),\phi(t))$ is unit-speed, i.e.

$$\dot{\theta}^2+\dot{\phi}^2\cos^2\theta=1$$

If γ is a geodesic, second equation in Theorem 2.9.8 tells us

$$\frac{d}{dt}(\dot{\phi}\cos^2\theta)=0.$$

So $\dot{\phi}\cos^2\theta=\Omega$ is a constant. If $\Omega=0$, $\dot{\phi}=0$ so ϕ is constant and γ is a great circle passing through north and south poles.

If $\Omega\neq 0$, by unit-speed condition

$$\dot{\theta}^2=1-\frac{\Omega^2}{\cos^2\theta}.$$

So along γ,

$$(\frac{d\theta}{d\phi})^2=\frac{\dot{\theta}^2}{\dot{\phi}^2}=\cos^2\theta(\Omega^{-2}\cos^2\theta-1).$$

Hence

$$\begin{aligned}\pm(\phi-\phi_0)&=\int\frac{d\theta}{\cos\theta\sqrt{\Omega^{-2}\cos^2\theta-1}}\\&=\int\frac{d\theta}{\cos^2\theta}\cdot\frac{1}{\sqrt{\Omega^{-2}-1-\tan^2\theta}}\\&=\sin^{-1}(\frac{\tan\theta}{\sqrt{\Omega^{-2}-1}})\end{aligned}$$

So

$$\begin{aligned}\tan\theta&=\pm\sqrt{\Omega^{-2}-1}\sin(\phi-\phi_0)\\&=\mp\sqrt{\Omega^{-2}-1}\sin\phi_0\cos\phi\pm\sqrt{\Omega^{-2}-1}\cos\phi_0\sin\phi\end{aligned}$$

Hence $\gamma(t)$ satisfies $z=ax+by$ where $a=\mp\sqrt{\Omega^{-2}-1}\sin\phi_0$, $b=\pm\sqrt{\Omega^{-2}-1}\cos\phi_0$. In other words, $\gamma(t)$ is the intersection of the sphere with a plane passing through origin, i.e. a great circle.

Actually, the above example also follows from the following
Fact 3: There is a unique geodesic through any given point of a surface in any given tangent direction.

More precisely,

Proposition 2.9.11. *Let $p \in S$, $\mathbf{t} \in T_pS$ with $||\mathbf{t}|| = 1$. Then there exists a unique unit-speed geodesic γ on S which passes through p and tangent to $\mathbf{t}$.*

Proof. Let $p = \sigma(a, b)$, $\mathbf{t} = c\sigma_u + d\sigma_v$.

A unit-speed curve $\gamma(t) = \sigma(u(t), v(t))$ passes through p if $u(t_0) = a, v(t_0) = b$; is tangent to $\mathbf{t}$ if $\dot{u}(t_0) = c, \dot{v}(t_0) = d$.

The geodesic equations in Theorem 2.9.9 read abstractly as

$$\begin{cases} \ddot{u} = f(u, v, \dot{u}, \dot{v}) \\ \ddot{v} = g(u, v, \dot{u}, \dot{v}) \end{cases}$$

where f, g are smooth.

So by existence and uniqueness theorem of ODE, for any initial value problem $u(t_0) = a, v(t_0) = b, \dot{u}(t_0) = c, \dot{v}(t_0) = d$ associated to the system, there is a unique solution. This is the geodesic passing through p and tangent to $\mathbf{t}$. □

Example 2.9.12. *1. Plane. There is a unique straight line passing through any point with given slope $y - y_0 = k(x - x_0)$.*

2. Sphere. There is a unique great circle passing through any point and tangent to any given direction at this point. It means starting from any given place on the earth, choose a direction to go, there is a unique route without turns.

Fact 4: Any local isometry $f : S_1 \to S_2$ takes the geodesics of S_1 to the geodesics of S_2.

Reason: For any surface patch σ of S_1, σ and $f \circ \sigma$ have the same first fundamental form. Notice geodesic equations only involve first fundamental form. So if $\gamma_1(t) = \sigma(u(t), v(t))$ is a geodesic on S_1, $\gamma_2 = f \circ \sigma(u(t), v(t))$ is a geodesic on S_2.

Example 2.9.13. *Unit cylinder $x^2 + y^2 = 1$. It is obtained from the plane by*

$$(u, v) \mapsto (\cos u, \sin u, v).$$

We have learnt that all straight lines on cylinder, i.e. those parallel to z-axis, are geodesics. These correspond to the lines parallel to y-axis in the plane.

We also know circles obtained by intersecting with plane $z = c$ are geodesics by Fact 2. These correspond to the lines parallel to x-axis on the plane.

What else? Straight lines on the plane $y = mx + c$ (if not parallel to y-axis) is mapped to

$$\gamma(u) = (\cos u, \sin u, mu + c)$$

on cylinder. These are helix we have studied in curve theory. By Fact 4, they are geodesics.

These are all the geodesics since now for any point and any direction, there is a unique such curve.

2.9.2 Geodesics on surfaces of revolution

We look at the surface of revolution

$$\sigma(u, v) = (f(u) \cos v, f(u) \sin v, g(u)),$$

with $f_u^2 + g_u^2 = 1$ and $f > 0$. So

$$I = du^2 + f^2(u)dv^2.$$

Hence two geodesic equations

$$\frac{d}{dt}(E\dot{u} + F\dot{v}) = \frac{1}{2}(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2)$$

$$\frac{d}{dt}(F\dot{u} + G\dot{v}) = \frac{1}{2}(E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2)$$

read as

$$(r1): \quad \ddot{u} = f(u)\frac{df}{du}\dot{v}^2$$

$$(r2): \ \frac{d}{dt}(f^2(u)\dot{v}) = 0$$

We may consider unit-speed geodesics, so we have

$$(r3): \dot{u}^2 + f^2(u)\dot{v}^2 = 1.$$

We have the followings:

1. *Every meridian $v = v_0$ is a geodesic.*

   $(r2)$ is obviously satisfied for v-curves. $(r3)$ implies $\dot{u} = \pm 1$, so $(r1)$ holds as well.

2. *A parallel $u = u_0$ is a geodesic if and only if $\frac{df}{du} = 0$ when $u = u_0$.*

   If $u = u_0$, $(r3)$ implies $\dot{v} \neq 0$ is a constant, so $(r2)$ holds. $(r1)$ holds if and only if $\frac{df}{du} = 0$ at $u = u_0$.

3. *Let $\theta \in [0, \frac{\pi}{2}]$ be the angle of a geodesic with a parallel that intersects it, R be the radius of the parallel at the intersection. Then we have Clairaut's relation:*

$$R\cos\theta = const.$$

Here $(r2)$ implies $f^2\dot{v} = const$. On the other hand,

$$\cos\theta = \frac{<\sigma_v, \sigma_u\dot{u} + \sigma_v\dot{v}>}{||\sigma_v||} = f\dot{v}.$$

Since $f = R$, so $R\cos\theta = const$.

The converse is also true: If $R\cos\theta$ is a constant and no point has parallel tangent vector, then it is a geodesic.

4. $(r2) + (r3) \Rightarrow (r1)$ *when $\dot{v}, \dot{u} \neq 0$.*

   $(r2) \Rightarrow f^2\dot{v} = C \neq 0$. *And $(r3)$ becomes $\dot{u}^2 + C\dot{v} = 1$. Take derivatives for the above two equations, we have*

$$\ddot{v} = -\frac{\dot{v} \cdot 2f\frac{df}{du} \cdot \dot{u}}{f^2},$$

$$2\dot{u}\ddot{u} + C\ddot{v} = 0.$$

So

$$2\dot{u}\ddot{u} = \dot{v}^2 \cdot 2f\frac{df}{du} \cdot \dot{u}.$$

This is just $(r1)$.

5. *Finally, we could solve the geodesic*

$$(r3) \times (\frac{dt}{dv})^2 : (\frac{dt}{dv})^2 = f^2 + (\frac{dt}{dv} \cdot \frac{du}{dt})^2 = f^2 + (\frac{du}{dv})^2.$$

Since $(r2)$ implies $f^2\frac{dv}{dt} = C$, we have

$$\frac{f^4}{C^2} - f^2 = (\frac{du}{dv})^2.$$

So

$$\frac{du}{dv} = \frac{1}{C}f\sqrt{f^2 - C^2},$$

or

$$v = C\int \frac{1}{f} \cdot \frac{1}{\sqrt{f^2 - C^2}}du + C'.$$

2.9.3 Geodesics and shortest paths

On a plane or a unit sphere, shortest paths are geodesics. But not every geodesic is the shortest path.

A natural question is : for $p, q \in S$, is a shortest path between p and q always a geodesic?

We need to compare with nearby curves. Assume $\gamma(t)$ is a unit-speed curve in a surface patch σ. Look at a family of curves γ^τ on σ with $-\delta < \tau < \delta$, such that

1. There exists $\epsilon > 0$, such that $\gamma^\tau(t)$ is defined for $t \in (-\epsilon, \epsilon)$ and all $\tau \in (-\delta, \delta)$.

2. For some a, b, with $-\epsilon < a < b < \epsilon$, we have $\gamma^\tau(a) = p, \gamma^\tau(b) = q$ for all $\tau \in (-\delta, \delta)$.

3. The map from $(-\delta, \delta) \times (-\epsilon, \epsilon)$ into $\mathbb{R}^3$
$$(\tau, t) \mapsto \gamma^\tau(t)$$
is smooth.

4. $\gamma^0 = \gamma$.

Recall the length between p and q is

$$L(\tau) = \int_a^b ||\dot{\gamma}^\tau|| dt.$$

Theorem 2.9.14. *The unit speed curve γ is a geodesic if and only if $\frac{d}{d\tau}L(0) = 0$ for all families of curves γ^τ with $\gamma^0 = \gamma$.*

Note that we cannot assume γ^τ be unit speed as well for $\tau \neq 0$. Otherwise $L(\tau)$ is fixed.

Let us only sketch the proof. Basically, it follows from choosing a vector field $V(t) = f(t)A(t)$ where $f \geq 0$ is a smooth function with $f(a) = f(b) = 0$, and $A(s) = \nabla_\gamma \dot{\gamma}$. Then the result follows from the calculation

$$L'(0) = -\int_a^b < A(t), V(t) > dt = -\int_a^b f(t)|A(t)|^2 dt.$$

Corollary 2.9.15. *A shortest path is always a geodesic.*

Proof. If γ is a shortest path, then $L(\tau)$ must have absolute minimum when $\tau = 0$. So $\frac{d}{d\tau}L(\tau) = 0$ when $\tau = 0$, which implies τ is a geodesic by above theorem. ☐

But the converse need not to be true. On a non-compact surface, there might not exist shortest curve for a given pair of points. Think about $\mathbb{R}^2 \backslash 0$ and $p, q = (\pm 1, 0)$. But for compact surfaces, there is always one shortest curve.

2.9.4 Geodesic coordinates

Let $p \in S$, γ a unit speed geodesic with $\gamma(0) = p$. For any v, let $\tilde{\gamma}^v(u)$ be a unit speed geodesic such that $\tilde{\gamma}^v(0) = \gamma(v)$ and perpendicular to γ at $\gamma(0)$. Define $\sigma(u, v) = \tilde{\gamma}^v(u)$.

Proposition 2.9.16. *There is an open subset $U \subset \mathbb{R}^2$ containing $(0,0)$ such that $\sigma : U \to \mathbb{R}^3$ is a parametrization of S. Moreover,*

$$I_\sigma = du^2 + G(u, v)dv^2,$$

where G is a smooth function with

$$G(0, v) = 1, G_u(0, v) = 0$$

whenever $(0, v) \in U$.

Proof. $\gamma(v) = \sigma(0, v)$. $\sigma_u(0, v) = \frac{d}{du}\tilde{\gamma}^v(u)|_{u=0}$, $\sigma_v(0, v) = \frac{d}{dv}\gamma(v)$, and they are perpendicular to each other.

So if $\sigma(u, v) = (f, g, h)$, the Jacobian

$$\begin{pmatrix} f_u & f_v \\ g_u & g_v \\ h_u & h_v \end{pmatrix}$$

has rank 2 at $u = v = 0$. Hence at least one of 2×2 block is invertible at $(0, 0)$, say

$$\begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix}$$

By inverse function theorem, there is an open subset U of $\mathbb{R}^2$ such that the map $F(u, v) = (f(u, v), g(u, v))$ is bijection from U to an open set $F(U) \subset \mathbb{R}^2$, and its inverse $F(U) \to U$ is also smooth (i.e F is a diffeomorphism onto its image). So the matrix is invertible for all U and σ_u, σ_v are linearly independent. So σ is a parametrization.

We calculate

$$E = ||\sigma_u||^2 = ||\frac{d}{du}\tilde{\gamma}^v(u)||^2 = 1$$

since $\tilde{\gamma}^v(u)$ is unit speed.

Also $F|_{u=0} = 0$ since $\sigma_u(0, v) \perp \sigma_v(0, v)$. Apply the second equation in Proposition 2.9.8

$$\frac{d}{dt}(F\dot{u} + G\dot{v}) = \frac{1}{2}(E_v\dot{u}^2 + 2F_v\dot{u}\dot{v} + G_v\dot{v}^2)$$

to $\tilde{\gamma}^v$, where v is a constant implies $\dot{v} = 0$, $E = 1$ implies $E_v = 0$ and $t = u$. So we have $F_u = 0$, and hence $F = 0$ everywhere. All together,

$$I_\sigma = du^2 + G(u, v)dv^2.$$

$$G(0,v) = ||\sigma_v(0,v)||^2 = ||\frac{d\gamma}{dv}||^2 = 1$$

since γ is unit-speed. Apply the first equation in Proposition 2.9.8

$$\frac{d}{dt}(E\dot{u} + F\dot{v}) = \frac{1}{2}(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2)$$

to $\gamma(v) = \tilde{\gamma}^v(0) = \sigma(0,v)$, where $u = 0$ implies $\dot{u} = 0$ and $t = v$. So $G_u(0,v) = 2F_v(0,v) = 0$. □

Actually, we a partial converse statement: If $I = du^2 + g^2(u,v)dv^2$, then $v = const = c$ is a geodesic. We could definitely apply geodesic equations to check this fact. However, there is an alternative way. We could actually prove that they are shortest path among all curves with $\gamma(t_1) = \sigma(a,c)$ and $\gamma(t_2) = \sigma(b,c)$ with $u(t_1) = a, u(t_2) = b$.

$$L(\gamma) = \int_{t_1}^{t_2} \sqrt{\dot{u}^2 + g^2\dot{v}^2}dt \geq \int_a^b du = b - a.$$

An application of the geodesic coordinates is the following

Theorem 2.9.17. *Any point of a surface of constant Gaussian curvature is contained in a patch that is isometric to an open subset of a plane, a sphere or a pseudosphere.*

Proof. Only need to consider the cases $K = -1, 0, 1$. Take a geodesic patch $\sigma(u,v)$ with $\sigma(0,0) = p$. Write $g = \sqrt{G}$, then

$$I = du^2 + g^2(u,v)dv^2.$$

Since when $E = 1$ and $F = 0$, $K = -\frac{1}{\sqrt{G}}\frac{\partial^2\sqrt{G}}{\partial u^2}$. So

$$\frac{\partial^2 g}{\partial u^2} + Kg = 0$$

with $g(0,v) = 1, g_u(0,v) = 0$.

1. $K = 0$. $g(u,v) = \alpha(v)u + \beta(v)$ and the initial conditions imply $\alpha = 0, \beta = 1$, and
$$I = du^2 + dv^2.$$
So σ is isometric to an open subset of the plane.

2. $K = 1$. $g(u,v) = \alpha(v)\cos u + \beta(v)\sin u$. Initial conditions imply $\alpha = 1, \beta = 0$, and
$$I = du^2 + \cos^2 u dv^2,$$
the same as unit sphere. So σ is isometric to an open subset of S^2.

3. $K = -1$. $g(u,v) = \alpha(v)e^{-u} + \beta(v)e^{u}$. Initial condition tells us $\alpha = \frac{1}{2} = \beta$. So
$$I = du^2 + \cosh^2 u dv^2.$$
This indeed tell us for every surface with $K = -1$, including pseudosphere, the geodesic coordinate has such first fundamental form. Hence, all are local isometric to each other and especially to pseudosphere.

□

Exercise: If we reparametrize the fundamental form $I = du^2 + \cosh^2 u dv^2$ by $V = e^v \tanh u, W = \frac{e^v}{\cosh u}$, then we have
$$I = \frac{dV^2 + dW^2}{W^2}.$$

2.9.5 Half plane model of hyperbolic plane

Let us look at the half-plane model $\mathcal{H}$ of pseudosphere. This means the last parametrizaton appeared above:
$$I_{\mathcal{H}} = \frac{dv^2 + dw^2}{w^2}$$
where v takes any real value and $w > 0$. So it basically gives a metric on the half space and $K = -1$.

Fact 1: Hyperbolic angles in $\mathcal{H}$ is the same as Euclidean angles.
This is because $I_{\mathcal{H}}$ is conformal to the Euclidean metric $I = dv^2 + dw^2$.

Fact 2: Geodesics are half-lines orthogonal to $w = 0$ and the semi-circles with centres on $w = 0$.
We show it by
Step 1: They are geodesics.
It is easy to see $v \equiv C$ is a geodesic since it is a shortest path: $\int \frac{dv^2+dw^2}{w^2} > \int \frac{dw^2}{w^2}$.
We show the circles $(v - v_0)^2 + w^2 = R^2$ are geodesics. We use polar coordinates
$$v - v_0 = r\cos\theta, w = r\sin\theta$$
Then
$$\sigma_r = \sigma_v \cdot \frac{\partial v}{\partial r} + \sigma_w \cdot \frac{\partial w}{\partial r} = \sigma_v \cos\theta + \sigma_w \sin\theta,$$
$$\sigma_\theta = -\sigma_v r \sin\theta + \sigma_w r \cos\theta.$$
So
$$||\sigma_r||^2 = \frac{1}{w^2}(\cos^2\theta + \sin^2\theta) = \frac{1}{r^2 \sin^2\theta}$$

$$\sigma_r \cdot \sigma_\theta = 0$$
$$||\sigma_\theta||^2 = \frac{1}{w^2} r^2 = \frac{1}{\sin^2\theta}$$

Let $\tilde{\theta} = \int_0^\theta \frac{1}{\sin\theta} d\theta, \tilde{r} = r$, then

$$I = d\tilde{\theta}^2 + \frac{1}{\tilde{r}^2 \sin^2\theta} d\tilde{r}^2.$$

So $\tilde{r} \equiv R$ is a geodesic, which is $(v - v_0)^2 + w^2 = R^2$.

Step 2: They are all the geodesics.

This is because for any point, any direction, there is a unique circle with centre at $w = 0$. (How to construct it?)

2.10 Gauss-Bonnet Theorem

Recall Hopf's Umlaufsatz tells us for a simple closed curve in the plane

$$\int \kappa_s ds = 2\pi.$$

Now we want to generalize it to a simple closed curve C in a surface $S \subset \mathbb{R}^3$. We assume C is the boundary of a set $\Delta \subset S$ which is homeomorphic to a disc. Then we have

Theorem 2.10.1 (Local Gauss-Bonnet)**.**

$$\int_C \kappa_g ds = 2\pi - \int_Y K dA$$

The following proof is given by Donaldson.

Let us first introduce an algebraic lemma whose proof is left as an exercise. Let $P \subset \mathbb{R}^3$ be a plane through the origin and $\mathbf{N}$ be a unit normal to P. For $\mathbf{x}, \mathbf{y} \in P$, set

$$\mathbf{x} \wedge \mathbf{y} = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{N}.$$

Then

Lemma 2.10.2. *For any three vectors* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in P$*, we have*

$$(\mathbf{x} \wedge \mathbf{y})\mathbf{z} + (\mathbf{y} \wedge \mathbf{z})\mathbf{x} + (\mathbf{z} \wedge \mathbf{x})\mathbf{y} = 0.$$

Now let us prove local Gauss-Bonnet.

Proof. We use polar coordinates (r, θ) in the plane. We suppose a local parametrization σ of S maps the unit disc to Δ and the unit circle to C. For $r \leq 1$ we let C_r be the closed curve $\sigma(r, \theta)$ in S. We set

$$I(r) = \int_{C_r} \kappa_g ds.$$

Let Δ_r be the image of the disc of radius r. We compute

$$\frac{d}{dr}(I(r) + \int_{\Delta_r} K dA).$$

We parametrize C_r by arc length and let $\mathbf{t}$ be the tangent to curve C_r and $\mathbf{N}$ be the normal to S. Thus $\mathbf{t} \cdot \mathbf{N} = 0$ everywhere. Then

$$I(r) = \int_{C_r} \kappa_g ds = \int_{C_r} (\mathbf{N}\times\dot{\gamma})\cdot\ddot{\gamma} ds = -\int_{C_r} (\mathbf{t}_s\times\mathbf{t})\cdot\mathbf{N} ds = -\int_0^{2\pi} (\mathbf{t}_\theta\times\mathbf{t})\cdot\mathbf{N} d\theta.$$

Thus

$$\frac{dI}{dr} = -\int_0^{2\pi} \frac{\partial}{\partial r}((\mathbf{t}_\theta \times \mathbf{t}) \cdot \mathbf{N}) d\theta.$$

Now consider

$$S = \frac{\partial}{\partial r}((\mathbf{t}_\theta \times \mathbf{t}) \cdot \mathbf{N}) - \frac{\partial}{\partial \theta}((\mathbf{t}_r \times \mathbf{t}) \cdot \mathbf{N})$$

Then

$$S = ((\mathbf{t}_\theta\times\mathbf{t}_r)\cdot\mathbf{N} + (\mathbf{t}_{\theta r}\times\mathbf{t})\cdot\mathbf{N} + (\mathbf{t}_\theta\times\mathbf{t})\cdot\mathbf{N}_r) - ((\mathbf{t}_r\times\mathbf{t}_\theta)\cdot\mathbf{N} + (\mathbf{t}_{r\theta}\times\mathbf{t})\cdot\mathbf{N} + (\mathbf{t}_r\times\mathbf{t})\cdot\mathbf{N}_\theta)$$

which is

$$S = 2(\mathbf{t}_\theta \times \mathbf{t}_r) \cdot \mathbf{N} + (\mathbf{t}_\theta \times \mathbf{t}) \cdot \mathbf{N}_r - (\mathbf{t}_r \times \mathbf{t}) \cdot \mathbf{N}_\theta.$$

Since $\mathbf{t}$ is a unit vector, so the vectors $\mathbf{t}_r, \mathbf{t}_\theta, \mathbf{N}$ are orthogonal to it. Thus the three vectors are on the same plane and

$$\begin{aligned}
S =&(\mathbf{t}_\theta \times \mathbf{t}) \cdot \mathbf{N}_r - (\mathbf{t}_r \times \mathbf{t}) \cdot \mathbf{N}_\theta \\
=&\mathbf{t}_\theta \cdot (\mathbf{t} \times \mathbf{N}_r) - \mathbf{t}_r \cdot (\mathbf{t} \times \mathbf{N}_\theta) \\
=&(\mathbf{t}_\theta \cdot \mathbf{N})(\mathbf{t} \wedge \mathbf{N}_r) - (\mathbf{t}_r \cdot \mathbf{N})(\mathbf{t} \wedge \mathbf{N}_\theta) \\
=&\mathbf{t} \cdot ((\mathbf{t} \wedge \mathbf{N}_\theta)\mathbf{N}_r - (\mathbf{t} \wedge \mathbf{N}_r)\mathbf{N}_\theta) \\
=&-\mathbf{t} \cdot ((\mathbf{N}_\theta \wedge \mathbf{N}_r)\mathbf{t}) \\
=&\mathbf{N}_r \wedge \mathbf{N}_\theta \\
=&K\sigma_r \wedge \sigma_\theta
\end{aligned}$$

This third equality is because of $\mathbf{t}, \mathbf{N}_r, \mathbf{N}_\theta$ are orthogonal to $\mathbf{N}$ and thus the cross products are parallel to it. The fourth is because of $\mathbf{t} \cdot \mathbf{N} = 0$. The fifth is because of the lemma. The last is because the calculation in Theorem 2.5.10.

Hence

$$\frac{dI}{dr} = -\int_0^{2\pi} K\sigma_r \wedge \sigma_\theta d\theta.$$

Let E, F, G be the components of the first fundamental form in the r, θ coordinates. Then

$$\frac{dI}{dr} = -\int_0^{2\pi} K\sqrt{EG - F^2} d\theta.$$

But

$$\int_{\Delta_r} KdA = \int_0^r \int_0^{2\pi} K\sqrt{EG - F^2}d\theta d\rho,$$

so

$$\frac{d}{dr}\int_{\Delta_r} KdA = \int_0^{2\pi} K\sqrt{EG - F^2}d\theta d\rho.$$

Hence the two derivatives cancel and we conclude that

$$I(r) + \int_{\Delta_r} KdA$$

is a constant, independent of r.

Finally, $\lim_{r\to 0} I(r) = 2\pi$ since it approaches that of the plane over very small regions. Hence the proof is complete. □

2.10.1 Geodesic polygons

The *geodesic polygons* are regions formed by the intersection of geodesics. For spheres, we know all geodesics are great circles; for plane, all geodesics are straight lines; for hyperbolic plane, we also classify the geodesics. So we see geodesic polygons are generalizations of polygons in a plane. How to calculate the areas?

Now, let us digress on the version of local Gauss-Bonnet for piecewise smooth curves. It γ is such a curve with Γ_i as each smooth component. Let θ_i be the internal angle. Since locally, especially at the corners, the κ_g approaches to signed curvature of the plane curve which measures the turning angle. Hence the integration of geodesic curvature would read as

$$\int_\gamma \kappa_g ds = \sum \int_{\Gamma_i} \kappa_g ds + \sum(\pi - \theta_j).$$

It could be proven rigorously by approximating γ by smooth curves.

Now, let Δ be a geodesic polygon, homeomorphic to the disc, whose boundary is made up of smooth geodesics Γ_i meeting at corners. By local Gauss-Bonnet, let θ_j be the internal angle at corner j, we have

$$\begin{aligned}\int_\Delta KdA =& 2\pi - \sum \int_{\Gamma_i} \kappa_g ds - \sum(\pi - \theta_j)\\ =& 2\pi - (n\pi - \sum \theta_j)\\ =& -(n-2)\pi + \sum \theta_j\end{aligned}$$

For sphere, $K = 1$, so

$$area(\Delta) = \sum \theta_j - (n-2)\pi.$$

Especially, it implies angle sum of n-gon is greater than $(n-2)\pi$. So for example, a half unit sphere has area 2π. A triangle obtained by the 3 great circles cut by x, y, z planes has area $\frac{\pi}{2}$.

For pseudosphere, $K = -1$.

$$area(\Delta) = (n-2)\pi - \sum \theta_j.$$

So the angle sum of n-gon is less than $(n-2)\pi$. So especially any ideal triangle in the hyperbolic plane (i.e. a triangle with all vertices on the x-axis) has area π.

This circle of ideas applies to shortest path geodesics. First we define a surface S is *simply connected* if any simple closed curve divides S into two regions, one of them is homeomorphic to a disc.

Theorem 2.10.3. *If S is a simply connected surface of $K \leq 0$, then there exists no more than one geodesics through any 2 points on S.*

Proof. If there are two geodesics γ_1 and γ_2 pass through P, Q, then they form a 2-gon since S is simply connected. Let α, β be the two angles at P and Q. Apply local Gauss-Bonnet to this 2-gon,

$$\int K dA + \pi - \alpha + \pi - \beta = 2\pi$$

which implies

$$\alpha + \beta = \int K dA \leq 0$$

Hence $\alpha = \beta = 0$, contradicting to the assumption that there are two geodesics. □

We have two immediate corollaries.

Corollary 2.10.4. *If S is a simply connected surface of $K \leq 0$, then*

1. *There are no closed geodesic on S.*
2. *Any arc of a geodesic on S is a shortest path.*

2.10.2 Global Gauss-Bonnet

A triangulation of a region $R \subset S$ is a finite family of triangles $T_i, i = 1, \cdots, n$, such that

1. $\cup_{i=1}^{n} T_i = R$.
2. If $T_i \cap T_j \neq \emptyset$, then $T_i \cap T_j$ is either a common edge or a common vertex.

Theorem 2.10.5. *Every compact surface has a triangulation.*

We denote

- V =total number of vertices;
- E =total number of edges;
- F =total number of faces (triangles).

Definition 2.10.6. *The Euler numer χ of a triangulation is*

$$\chi = V - E + F.$$

Example 2.10.7. *Tetrahedron, Octahedron and Icosahedron give triangulations of a sphere. Actually, we could divide sphere into polygonal region and make the same definitions of V, E, F, χ.*

	V	E	F	χ
Tetrahedron	*4*	*6*	*4*	*2*
Octahedron	*6*	*12*	*8*	*2*
Icosahedron	*12*	*30*	*20*	*2*
Texahedron	*8*	*12*	*6*	*2*
Dodecahedron	*20*	*30*	*12*	*2*

Especially, we observe that $\chi = 2$ in all the above examples.

Theorem 2.10.8 (Global Gauss-Bonnet)**.** *For any compact surface S and any triangulation on it*

$$\int_S K dA = 2\pi\chi(S)$$

Corollary 2.10.9.

1. *$\chi(S)$ depends only on S and not on the choice of the triangulations.*
2. *$\int_S K dA$ does not depend on the way surface is embedded in $\mathbb{R}^3$.*

Recall the classification of compact surfaces: $S^2, T^2, \cdots, \Sigma_g, \cdots$ Since we only need one triangulation to determine the Euler number, we know there is a triangulation of T^2 with $V = 1, E = 3, F = 2$, then $\chi(T^2) = 0$.

For Σ_g, we use induction to claim

Theorem 2.10.10. *$\chi(\Sigma_g) = 2 - 2g$.*

Proof. We prove it by induction.

We understand Σ_{g+1} as the gluing of Σ_g and T^2 along a triangle in the triangulations $\mathcal{T}'$ and $\mathcal{T}''$ of Σ_g and T^2. It induces a triangulation of Σ_{g+1}. So

$$V = V' + V'' - 3, E = E' + E'' - 3, F = F' - 1 + F'' - 1.$$

Hence

$$\begin{aligned}\chi(\Sigma_{g+1}) =& V - E + F \\ =& \chi(\Sigma_g) + \chi(T^2) - 2 \\ =& 2 - 2g + 0 - 2 \\ =& 2 - 2(g+1)\end{aligned}$$

□

Corollary 2.10.11. $\int_{\Sigma_g} KdA = 4\pi(1-g)$

Now, let us prove global Gauss-Bonnet theorem.

Proof. For any triangular region, we have

$$\int_{\Delta} KdA = 2\pi - \sum \int_{\Gamma_i} \kappa_g ds - \sum(\pi - \theta_j).$$

Add them, we have

$$\sum_{\alpha} \int_{\Delta_\alpha} KdA = \int_S KdA.$$

Let us look at each component. For $\sum \int_{\Gamma_i} \kappa_g ds$, each edge is shared by two triangles which induce different orientations on the edge. So in the sum $\sum_\alpha \int_{\Delta_\alpha} KdA$, these terms add up to 0.

For the rest, 2π adds up to $2\pi F$. The angle $\sum \theta_j$ adds up to $2\pi V$ since each vertex has degree 2π. And finally, -3π in the last component adds up to

$$-3\pi F = -3\pi \cdot \frac{2E}{3} = -2\pi E$$

We use $3F = 2E$ since each face has 3 edges and each edge is counted twice in the sum.

Add all these up to $2\pi(F - E + V) = 2\pi\chi(S)$. □

2.11 Vector fields and Euler number

On a regular surface S, let V be a smooth tangent vector field on S, i.e $V = \alpha(u,v)\sigma_u + \beta(u,v)\sigma_v$ locally.

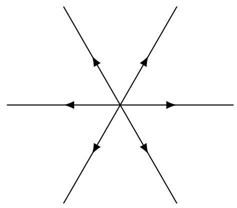
Definition 2.11.1. *A point $p \in S$ at which $V = 0$ is called a singular (or stationary) point.*

Now we introduce the index of singular points. Take a surface patch σ containing p, let γ be a simple closed (positively oriented) regular curve with period l, which contains p as the unique singular point in its interior. Let $\phi(t)$ be the angle from σ_u to $V(t)$, so $\phi(l) - \phi(0) = 2\pi j$ where j is an integer. We call j the index of the singular point. We could write it in terms of formula

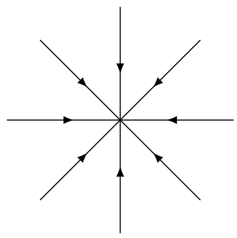
$$j(p) = \frac{1}{2\pi}\int_0^l \frac{d\phi}{dt}dt$$

Notice this j is finite and independent of γ and initial point $\gamma(0)$.

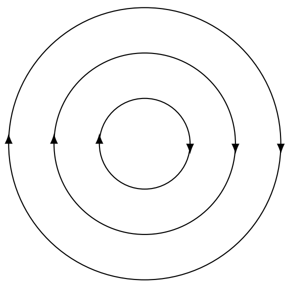
Let us see several examples.



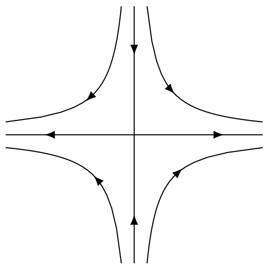
source: $V(x,y) = (x,y)$, $j = 1$



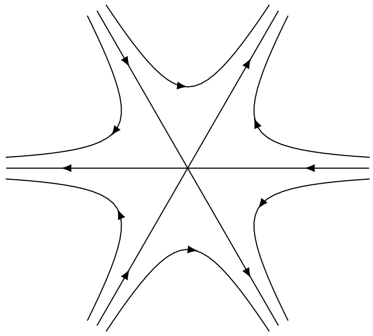
sink: $V(x,y) = (-x,-y)$, $j = 1$



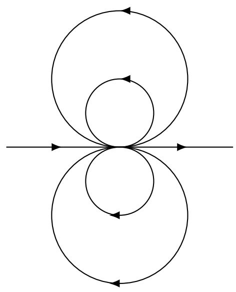
center (vortex): $V(x,y) = (y,-x)$, $j = 1$



simple saddle: $V(x, y) = (x, -y)$, $j = -1$



monkey saddle: $j = -2$



dipole: $j = 2$

Theorem 2.11.2 (Poincaré-Hopf)**.** *Let V be a smooth vector field on a compact surface S with finitely many singular points $p_1, \cdots, p_n$, then*

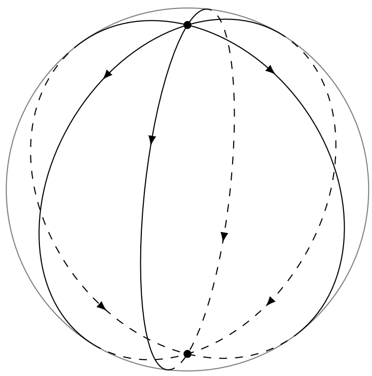
$$\sum_r j(p_r) = \chi(S).$$

Since $\chi(S^2) = 2$, we have

Corollary 2.11.3 (Hairy-Ball theorem)**.** *Any tangent vector field on S^2 has singular point.*

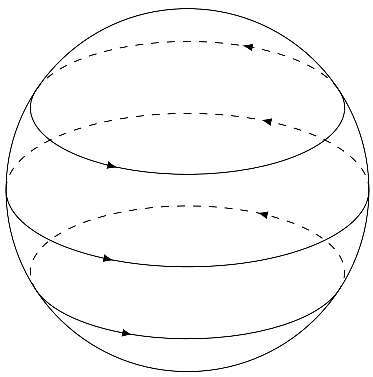
This means you cannot comb a hairy ball without cow-lick.

source

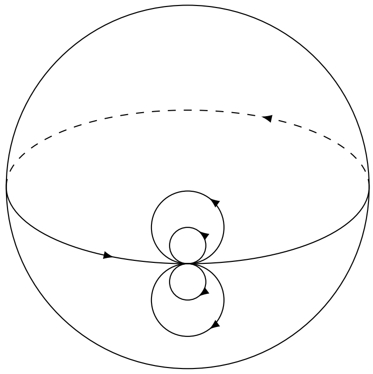


sink

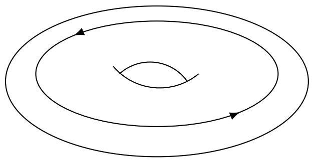
$j_1 + j_2 = 2$



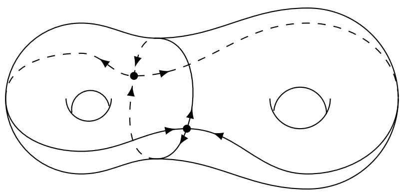
$j_1 + j_2 = 2$



$j = 2$



No singularity



two simple saddles: $j_1 + j_2 = -2$

Now we prove Poincaré-Hopf theorem.

Proof. We prove our result in 2 steps.

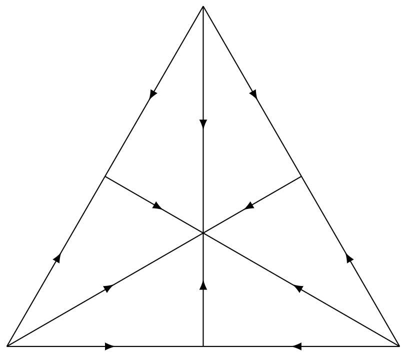
1. $\sum j(p_r)$ has the same value for all vector fields. Consider two vector fields F and F'.

Divide S, such that every 2-cell at most has one singularity of F or F'. Calculate $(j_F - j_{F'})(p)$. Notice if p is not singular for F', $j_{F'}(p) = 0$.

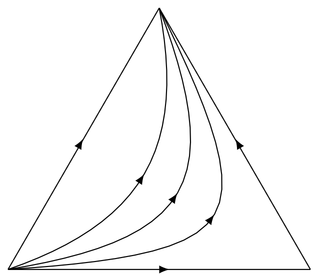
$(j_F - j_{F'})$ is $\frac{1}{2\pi}$ of the change in the difference of directions of F and F'. But for $\sum(j_F - j_{F'})$, every edge bounds two 2-cells, and will be calculated twice in opposite signs. So $\sum(j_F - j_{F'}) = 0$ or $\sum j_F = \sum j_{F'}$.

2. Construct a special vector field and calculate the sum $\sum_r j(p_r)$.

So we do triangulation. For each triangle, introduce 4 additional points: 1 in the centre and 3 on each side. Add vectors such that the centre is a



Add 4 points and subdivide triangles



vector fields in a new triangle

sink, each original vertex is a source and the ones on sides are saddle. We could extend the vector into each small triangle. No other singularities other than these. So now, each vertex of the original triangulation corresponds to a source, which has index 1. Each face corresponds to a sink, which has index 1 and each edge corresponds to a saddle, which has index -1. Hence

$$\sum_r j(p_r) = V + F - E = \chi(S).$$

□