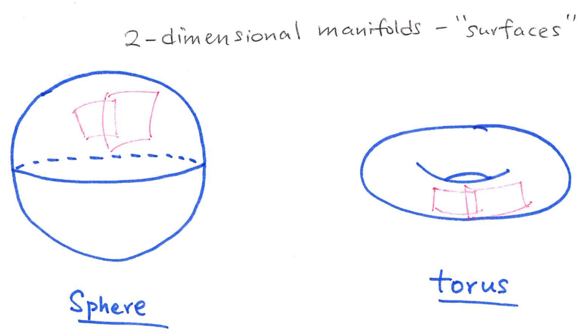
MATH 4030 Differential Geometry
Lecture Notes Part 2
last revised on October 14, 2015

After studying curves in $\mathbb{R}^n$, we now go up one dimension to study the theory of surfaces in $\mathbb{R}^n$ (we are mainly interested in the case $n = 3$ since the codimension is one). Unlike the theory of curves, where most theorems reduce to theorems about ODE systems, the situation gets drastically more difficult in dimension two since partial derivatives are involved. On the other hand, while there is no intrinsic geometry on a curve (i.e. every curve is locally "isomorphic" isometrically), there is both intrinsic and extrinsic geometry for a surface in $\mathbb{R}^n$. The intrinsic and extrinsic geometry are described by two quadratic forms, namely the *first* and *second fundamental forms*. Similar to curves, one can define various notions of *curvatures* to measure the deviation of the surface (both intrinsically and extrinsically) from the standard picture of a "flat" plane sitting inside $\mathbb{R}^3$.

Regular parametrized surfaces in $\mathbb{R}^3$

So what is a "surface"? It should be something two-dimensional in some sense, and just like curves, we want to give a surface some smooth parametrizations in order to do differential geometry. However, we should keep in mind that the parametrization is just a means of "coordinates" which allows us to do calculations but it should be completely artificial. Actual geometric quantities like area and curvatures should be independent of parametrizations. This property of parametrization invariance (think about Einstein's principle of equivalence in relativity) is a guiding principle to define many useful geometric quantities.

Therefore, a "surface" should locally look like $\mathbb{R}^2$, at least topologically but not isometrically. Image you have a piece of paper which is nice and flat originally, if the paper is rigid (i.e. it does not get torn or stretched or squeezed), you can only bent it into very restrictive shapes like a cylinder for example. But if the paper is made up of more flexible material which you can stretch and squeeze, then you can come up with all possible shapes you want. If we put these little pieces of "shapes" together, we would obtain a global surface like a sphere of a torus (i.e. surface of a donut). This gives the concept of a two dimensional manifold, *a.k.a.* a surface, which is an object locally is made up of pieces of $\mathbb{R}^2$ topologically.

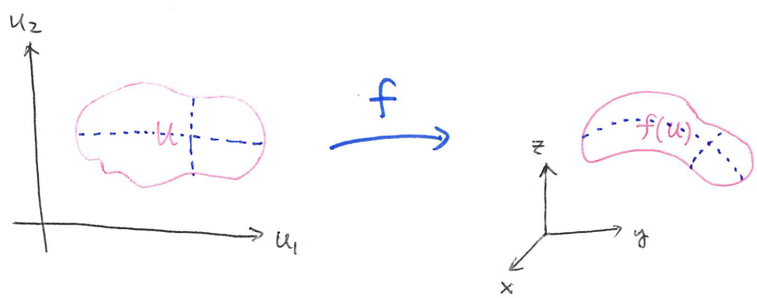


We will study the global geometry of surfaces later in the course. We first focus on studying the little pieces of surface elements which makes up the global surfaces.

Definition 1. *A (smooth) parametrized surface in* $\mathbb{R}^3$ *is a smooth map* $f : U \to \mathbb{R}^3$ *from an open set* $U \subset \mathbb{R}^2$*, which can be expressed in local coordinates by*

$$f(u_1, u_2) := (f_1(u_1, u_2), f_2(u_1, u_2), f_3(u_1, u_2)).$$

We say that a parametrized surface f *is regular if the differential* Df_u *has rank 2 at every* $u = (u_1, u_2) \in U$*. In this case, we also say that* f *is an immersion.*



In some sense we are just introducing a "coordinate system" on the image surface $f(U)$. Recall that the differential of f at $u \in U$ is the linear map $Df_u : \mathbb{R}^2 \to \mathbb{R}^3$ defined in local coordinates by the 3×2 matrix

$$Df_u = \left. \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} \end{pmatrix} \right|_u .$$

Therefore, the condition that Df_u has rank 2 is equivalent to the statement that the vectors $\left.\frac{\partial f}{\partial u_1}\right|_u := \left.(\frac{\partial f_1}{\partial u_1}, \frac{\partial f_2}{\partial u_1}, \frac{\partial f_3}{\partial u_1})\right|_u$ and $\left.\frac{\partial f}{\partial u_2}\right|_u := \left.(\frac{\partial f_1}{\partial u_2}, \frac{\partial f_2}{\partial u_2}, \frac{\partial f_3}{\partial u_2})\right|_u$ are linearly independent vectors in $\mathbb{R}^3$.

Remark 2. *The definition of a regular parametrized surface does not rule out self-intersections. However, since we are mainly concerned about local behaviors (except later in the course), this does not pose any problem.*

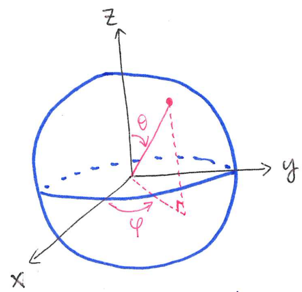
As in the case of curves, we can always reparametrize a given surface. Given a regular parametrized surface $f : U \to \mathbb{R}^3$, if $\varphi : \tilde{U} \to U$ is a smooth diffeomorphism between two open sets $U, \tilde{U}$ in $\mathbb{R}^2$, then the smooth map $\tilde{f} := f \circ \varphi : \tilde{U} \to \mathbb{R}^3$ is another regular parametrization of the "same" surface. We identify regular parametrized surfaces which differ only by a reparametrization.

Definition 3. *A* regular surface in $\mathbb{R}^3$ *is an equivalence class of regular parametrized surfaces up to reparametrizations.*

Therefore, one can think of a parametrization of a regular surface as putting a coordinate system (u_1, u_2) on the surface. Let us study some basic examples.

Example 4. *(1) The unit sphere* $\mathbb{S}^2 := \{p \in \mathbb{R}^3 : \|p\| = 1\} \subset \mathbb{R}^3$. *A well-known coordinate system on* $\mathbb{S}^2$ *is the spherical coordinates:*

$$f : (\varphi, \theta) \mapsto (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta).$$



If we take $U = \mathbb{R}^2 = \{(\varphi, \theta) : \varphi, \theta \in \mathbb{R}\}$, *then the image of* f *covers every point on* $\mathbb{S}^2$. *However, this is not regular everywhere (Exercise: Find a maximal subset of* $\mathbb{R}^2$ *on which the spherical coordinates* f *is regular.)*

On the other hand, one may also parametrize the sphere by the parametrization

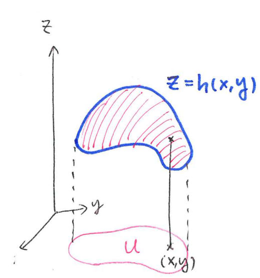
$$(x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2}).$$

However, this parametrization can at most be defined on the open set $\{(x, y) : x^2 + y^2 < 1\} \subset \mathbb{R}^2$ *and the image would only cover the open upper hemisphere. (Exercise: Is this parametrization regular everywhere?)*

(2) Graphical surfaces. Suppose we have a smooth function $h : U \to \mathbb{R}$ *defined on some open subset* $U \subset \mathbb{R}^2$, *then the map*

$$f : (u, v) \mapsto (u, v, h(u, v))$$

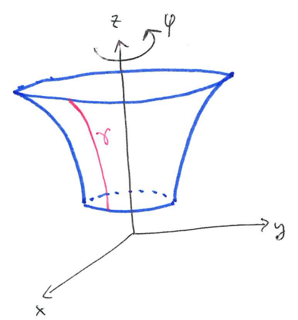
parametrizes the graph of $h = \{(x, y, z) \in \mathbb{R}^3 : z = h(x, y)\}$. *(Exercise: Is this parametrization regular?)*



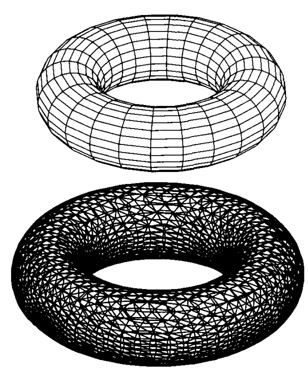
(3) *Surface of revolution. Suppose* $\gamma(t) = (x(t), 0, z(t)) : (a, b) \to \mathbb{R}^3$ *is a regular parametrized curve lying inside the open half plane* $\{(x, y, z) \in \mathbb{R}^3 : x > 0, y = 0\}$*. If we take the image curve and revolve it around the* z*-axis, we obtain a surface in* $\mathbb{R}^3$ *which is rotationally invariant about the* z*-axis. This surface of revolution can be parametrized by*

$$f : (t, \varphi) \mapsto (x(t) \cos\varphi, x(t) \sin\varphi, z(t)).$$

What is the domain of definition of f*? (Exercise: Is this parametrization regular?)*



For example, if γ *parametrizes a circle, then the surface of revolution obtained is a torus of revolution.*



Kuhnel p.53

Figure 3.3. Torus of revolution

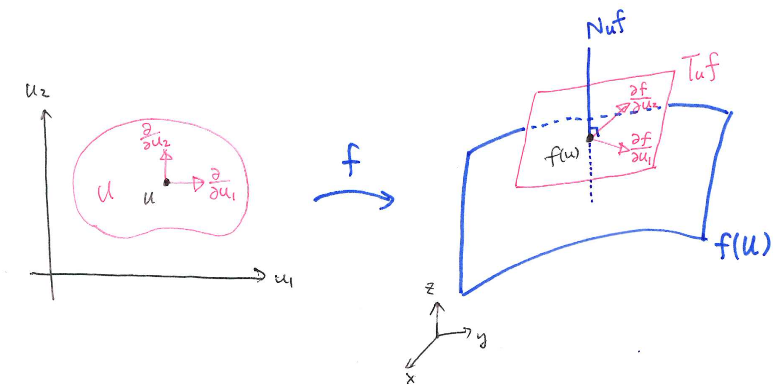
Tangent and normal spaces

Suppose we have a regular parametrized surface $f : U \to \mathbb{R}^3$. For each $u \in U$, we can define the tangent space and normal space at u as follows: let (u_1, u_2) be the standard rectangular coordinates on $U \subset \mathbb{R}^2$, the standard basis of $\mathbb{R}^2$ is given by $\{\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}\}$. Under the linear map $Df_u : \mathbb{R}^2 \to \mathbb{R}^3$, the standard basis $\{\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}\}$ goes to

$$\left.\frac{\partial f}{\partial u_1}\right|_u := Df_u\left(\frac{\partial}{\partial u_1}\right) = \left.\left(\frac{\partial f_1}{\partial u_1}, \frac{\partial f_2}{\partial u_1}, \frac{\partial f_3}{\partial u_1}\right)\right|_u,$$
$$\left.\frac{\partial f}{\partial u_2}\right|_u := Df_u\left(\frac{\partial}{\partial u_2}\right) = \left.\left(\frac{\partial f_1}{\partial u_2}, \frac{\partial f_2}{\partial u_2}, \frac{\partial f_3}{\partial u_2}\right)\right|_u,$$

which are linearly independent in $\mathbb{R}^3$ since f is regular. Therefore, one can define the tangent and normal space at u as

$$T_u f := \text{ span}\left\{\left.\frac{\partial f}{\partial u_1}\right|_u, \left.\frac{\partial f}{\partial u_2}\right|_u\right\} \subset \mathbb{R}^3,$$
$$N_u f := (T_u f)^\perp \subset \mathbb{R}^3.$$



We have made use of the parametrization f to define the tangent and normal spaces. In fact, they are well-defined for the equivalence class of regular surface $[f]$ because of the following lemma.

Lemma 5. *Let $\tilde{f} := f \circ \varphi : \tilde{U} \to \mathbb{R}^3$ be a reparametrization of the regular parametrized surface $f : U \to \mathbb{R}^3$ given some diffeomorphism $\varphi : \tilde{U} \to U$. Then, we have $T_{\tilde{u}}\tilde{f} = T_u f$ where $u = \varphi(\tilde{u})$.*

Proof. Let $\tilde{u}_1, \tilde{u}_2$ be the standard coordinates in $\tilde{U} \subset \mathbb{R}^2$, then φ can be written in coordinates as $\varphi(\tilde{u}_1, \tilde{u}_2) = (\varphi_1(\tilde{u}_1, \tilde{u}_2), \varphi_2(\tilde{u}_1, \tilde{u}_2))$. By chain rule, we have

$$\frac{\partial \tilde{f}}{\partial \tilde{u}_1} = Df_u \circ D\varphi_{\tilde{u}}\left(\frac{\partial}{\partial \tilde{u}_1}\right) = Df_u\left(\frac{\partial \varphi_1}{\partial \tilde{u}_1}\frac{\partial}{\partial u_1} + \frac{\partial \varphi_2}{\partial \tilde{u}_1}\frac{\partial}{\partial u_2}\right) = \frac{\partial \varphi_1}{\partial \tilde{u}_1}\frac{\partial f}{\partial u_1} + \frac{\partial \varphi_2}{\partial \tilde{u}_1}\frac{\partial f}{\partial u_2},$$
$$\frac{\partial \tilde{f}}{\partial \tilde{u}_2} = Df_u \circ D\varphi_{\tilde{u}}\left(\frac{\partial}{\partial \tilde{u}_2}\right) = Df_u\left(\frac{\partial \varphi_1}{\partial \tilde{u}_2}\frac{\partial}{\partial u_1} + \frac{\partial \varphi_2}{\partial \tilde{u}_2}\frac{\partial}{\partial u_2}\right) = \frac{\partial \varphi_1}{\partial \tilde{u}_2}\frac{\partial f}{\partial u_1} + \frac{\partial \varphi_2}{\partial \tilde{u}_2}\frac{\partial f}{\partial u_2}.$$

Since the linear map $D\varphi$ is invertible as φ is a diffeomorphism, we clearly have span$\left\{\frac{\partial \tilde{f}}{\partial \tilde{u}_1}, \frac{\partial \tilde{f}}{\partial \tilde{u}_2}\right\}$ = span$\left\{\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}\right\}$, i.e. $T_{\tilde{u}}\tilde{f} = T_u f$. □

Definition 6. *Let $f : U \to \mathbb{R}^3$ be a parametrization of a regular surface. For each $u \in U$, the* *tangent and normal spaces* *of f at u are defined by*

$$T_u f := Df_u(\mathbb{R}^2), \qquad N_u f := (T_u f)^{\perp}.$$

The *tangent and normal bundles* *are obtained by putting all these tangent and normal spaces together as a disjoint union:*

$$Tf := \bigsqcup_{u \in U} T_u f, \qquad Nf := \bigsqcup_{u \in U} N_u f.$$

By abuse of notation, one often uses $\{\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}\}$ also to denote the basis $\{\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}\}$ of $T_u f$ induced by the given parametrization f.

Recall that the dual of a real vector space V is the vector space of linear functionals on V, i.e. $V^* :=\text{Hom}(V, \mathbb{R})$. Therefore, one can define the cotangent space of f at u as

$$T_u^* f := (T_u f)^*.$$

We use $\{du_1, du_2\}$ to denote the dual basis of $\{\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}\}$, i.e. $du_i(\frac{\partial}{\partial u_j}) = \delta_{ij}$. The totality of all the cotangent space is called the cotangent bundle denoted by

$$T^* f := \bigcup_{u \in U} T_u^* f.$$

We will come back to these cotangent spaces later when we talk about forms (and tensors in general).

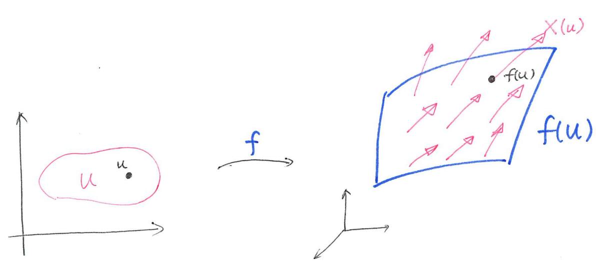
Vector fields along surfaces

We have learned from multivariable calculus the important concept of vector fields in $\mathbb{R}^n$. We would like to generalize this concept to vector fields defined on surfaces. Because a parametrized surface may intersect itself, we need to be slightly careful in our definitions.

Definition 7. *Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface. A* *vector field along f* *is a smooth map $X : U \to \mathbb{R}^3$. We say that X is a*

(i) *tangential vector field* *if $X(u) \in T_u f$ for all $u \in U$.*

(ii) *normal vector field* *if $X(u) \in N_u f$ for all $u \in U$.*



We think of the vector field X as a smooth association of the parameter u to a vector $X(u)$ in $\mathbb{R}^3$ based at the point $f(u)$ on the surface. Note that any vector field X along f can be decomposed into its tangential and normal components, i.e.

$$X = X^T + X^\perp,$$

where X^T is a tangent vector field and $X^\perp$ is a normal vector field. We can think of a tangential (normal) vector field as a "section" of the tangent (normal) bundle. Let $u = (u_1, u_2) \in U$. Since $\{\frac{\partial}{\partial u_1}\big|_u, \frac{\partial}{\partial u_2}\big|_u\}$ is a basis for $T_u f$, we can write a tangential vector field X as

$$X(u) = \alpha_1(u) \left.\frac{\partial}{\partial u_1}\right|_u + \alpha_2(u) \left.\frac{\partial}{\partial u_2}\right|_u,$$

where α_1, α_2 are some smooth functions of u. Similarly, since $\frac{\partial}{\partial u_1}\big|_u \times \frac{\partial}{\partial u_2}\big|_u$ is a non-zero vector in $N_u f$, any normal vector field X can be written as

$$X(u) = \alpha(u) \left.\frac{\partial}{\partial u_1}\right|_u \times \left.\frac{\partial}{\partial u_2}\right|_u,$$

for some smooth function $\alpha(u)$.

Example 8. *(1) To find a unit normal vector field on the unit sphere $\mathbb{S}^2$, recall that we have the spherical coordinates given by $f(\varphi, \theta) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. Therefore,*

$$\frac{\partial}{\partial \varphi} = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0),$$

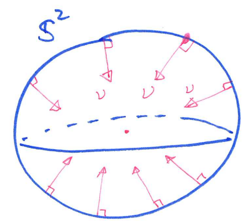
$$\frac{\partial}{\partial \theta} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta),$$

$$\frac{\partial}{\partial \varphi} \times \frac{\partial}{\partial \theta} = (-\sin^2\theta\cos\varphi, -\sin^2\theta\sin\varphi, -\sin\theta\cos\theta).$$

Notice that $\left\|\frac{\partial}{\partial\varphi} \times \frac{\partial}{\partial\theta}\right\| = |\sin\theta|$. Therefore, for $\varphi \in \mathbb{R}$ and $\theta \in (0, \pi)$, the unit normal vector field is given by

$$\nu = \frac{\frac{\partial}{\partial\varphi} \times \frac{\partial}{\partial\theta}}{\left\|\frac{\partial}{\partial\varphi} \times \frac{\partial}{\partial\theta}\right\|} = (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, -\cos\theta).$$

Notice that $\nu(\varphi, \theta) = -f(\varphi, \theta)$ is the "inward" pointing unit normal on $\mathbb{S}^2$ (excluding the north and south poles).



(2) *Consider the following parametrization of the cylinder* $\{(x, y, z) : x^2 + y^2 = 1\}$*:*

$$f(\varphi, z) = (\cos\varphi, \sin\varphi, z).$$

Let $z_0 \in \mathbb{R}$ *be a fixed positive number, we claim that*

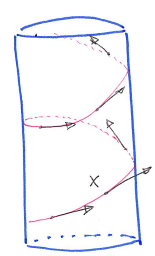
$$X(\varphi, z) = (-\sin\varphi, \cos\varphi, z_0),$$

is a tangential vector field along f*. To see this, we compute the standard basis of the tangent space* $T_{(\varphi,z)}f$*:*

$$\frac{\partial}{\partial\varphi} = (-\sin\varphi, \cos\varphi, 0), \qquad \frac{\partial}{\partial z} = (0, 0, 1).$$

Therefore,

$$X(\varphi, z) = \frac{\partial}{\partial\varphi} + z_0\frac{\partial}{\partial z} \in T_{(\varphi,z)}f.$$



Observe that the helix $c(t) = (\cos t, \sin t, tz_0 + c)$*, where* c *is a constant, lies on the cylinder and that is tangent to* X *at any point on the curve. We call such curves the* <u>*integral curves*</u> *of the vector field* X*. It can be shown that any tangential vector field on a surface generates a unique set of integral curves which foliates the surface (except at points where* X *vanishes). One can then define a family of diffeomorphisms (called the* <u>*flow associated to* X</u>*) on the surface by moving points along the integral curves. Therefore, one can think of a vector field as an infinitesimal diffeomorphism.*

First fundamental form

The definition of regular surfaces allows us to do calculus on surfaces. In order to do *geometry*, we need the notion of a *metric* which can measure length of a vector and the angle between two vectors. What we need in linear algebra terms is an *inner product*. In the 19th century, the German mathematician Riemann (who was a student of Gauss) proposed that a differentiable manifold together with a Riemannian metric on it is all we need to study its (intrinsic) geometry. Here, "Riemannian" means that the metric (i.e. inner product) has to be positive definite. Soon after this groundbreaking discovery of Riemann, people realized that much of the theory survives as long as the inner product is only non-degenerate. In 1915, Einstein proposed that our universe consists of a spacetime which can be described as a four dimensional manifold equipped with a Lorentzian metric and gravity can be interpreted as the geometry of this underlying Lorentzian manifold. This is a triumph for science in the early 20th century showing how closely physics and mathematics are related to each other.

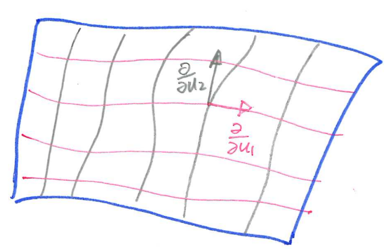
We now go back to our study of the geometry of surfaces in $\mathbb{R}^3$. Recall that the Euclidean space $\mathbb{R}^3$ is equipped with the standard inner product $\langle \cdot, \cdot \rangle$. This of course induces an inner product on any linear subspace of $\mathbb{R}^3$.

Definition 9. *Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface. The first fundamental form g is the positive definite, symmetric bilinear form defined on each tangent space of f by*

$$g(X, Y) := \langle X, Y \rangle,$$

for any $X, Y \in T_u f \subset \mathbb{R}^3$.

One can think of g as a smooth family of inner products defined on the smooth family of tangent spaces $T_u f$ parametrized by u. For surfaces in $\mathbb{R}^3$, the situation is simpler since all the tangent spaces $T_u f$ are two dimensional subspaces of the underlying space $(\mathbb{R}^3, \langle \cdot, \cdot \rangle)$. The first fundamental form g is then just the restriction of $\langle \cdot, \cdot \rangle$ to each of these two dimensional subspaces.



For a local expression of g, recall that $\{\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}\}$ is the standard basis of $T_u f$ associated to the parametrization f. Under this basis, we can express the inner product g as a 2×2 matrix given by

$g_{ij} = \langle \frac{\partial}{\partial u_i}, \frac{\partial}{\partial u_j} \rangle$, $i, j = 1, 2$, i.e. if $X = X_1 \frac{\partial}{\partial u_1} + X_2 \frac{\partial}{\partial u_2}$ and $Y = Y_1 \frac{\partial}{\partial u_1} + Y_2 \frac{\partial}{\partial u_2}$ are two tangential vector fields along f, then

$$g(X,Y) = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}.$$

Note that the 2×2 matrix on the right hand side is symmetric and positive definite. Equivalently, one can express the first fundamental form as a symmetric $(0,2)$-tensor (i.e. $du_1 du_2 = du_2 du_1$):

$$ds^2 := g_{11}\, du_1^2 + 2g_{12}\, du_1 du_2 + g_{22}\, du_2^2.$$

The fact that this expression really defines a symmetric $(0,2)$-tensor because of the following invariance under coordinate transformations.

Lemma 10. *Let $\tilde{f} : \tilde{U} \to \mathbb{R}^3$ be a reparametrization of a regular parametrized surface $f : U \to \mathbb{R}^3$, where $\varphi : \tilde{U} \to U$ is a diffeomorphism. Suppose g_{ij} and $\tilde{g}_{ij}$ are the local expressions of the first fundamental form g associated with the parametrizations f and $\tilde{f}$ respectively. Then, we have*

$$(\tilde{g}_{ij}) = (D\varphi)^T (g_{ij}) (D\varphi),$$

where $D\varphi$ is the 2×2 Jacobian matrix for φ, i.e. $(D\varphi)_{ij} = \frac{\partial \varphi_i}{\partial \tilde{u}_j}$, $i, j = 1, 2$. Here, we have $\varphi(\tilde{u}_1, \tilde{u}_2) = (\varphi_1(\tilde{u}_1, \tilde{u}_2), \varphi_2(\tilde{u}_1, \tilde{u}_2))$.

Proof. Observe that

$$(g_{ij}) = \begin{pmatrix} \langle \frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_1} \rangle & \langle \frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2} \rangle \\ \langle \frac{\partial f}{\partial u_2}, \frac{\partial f}{\partial u_1} \rangle & \langle \frac{\partial f}{\partial u_2}, \frac{\partial f}{\partial u_2} \rangle \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_2}{\partial u_1} & \frac{\partial f_3}{\partial u_1} \\ \frac{\partial f_1}{\partial u_2} & \frac{\partial f_2}{\partial u_2} & \frac{\partial f_3}{\partial u_2} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} \end{pmatrix}.$$

In matrix form, we have $(g_{ij}) = (Df)^T (Df)$. Similarly, we have $(\tilde{g}_{ij}) = (D\tilde{f})^T (D\tilde{f})$. On the other hand, $D\tilde{f} = (Df)(D\varphi)$ by chain rule. Combining these, we have

$$(\tilde{g}_{ij}) = (D\tilde{f})^T (D\tilde{f}) = (D\varphi)^T (Df)^T (Df)(D\varphi) = (D\varphi)^T (g_{ij})(D\varphi).$$

□

Example 11. *(1) Flat surfaces. Consider the parametrized surface $f(u_1, u_2) = (u_1, u_2, 0)$ into the xy-plane of $\mathbb{R}^3$, then it is trivial to see that*

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

However, if we parametrize the xy-plane instead by $\tilde{f}(\tilde{u}_1, \tilde{u}_2) = (\tilde{u}_1 + \tilde{u}_2, \tilde{u}_1 - 2\tilde{u}_2, 0)$, then

$$(\tilde{g}_{ij}) = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}.$$

(2) For the spherical coordinate system $f(\varphi, \theta) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, the first fundamental form is given by

$$(g_{ij}) = \begin{pmatrix} \sin^2\theta & 0 \\ 0 & 1 \end{pmatrix}.$$

We will see later that the first fundamental form, which can be viewed as a family of positive definite symmetric 2×2 matrices depending on $u \in U$, is all we need to describe the *intrinsic geometry* of the surface. In other words, we do not need to know that (g_{ij}) comes from the inner product of some tangent vectors in $\mathbb{R}^3$ but the data (U, g_{ij}) alone should be enough to describe the intrinsic geometry of some abstract surface, which is not sitting as a surface in $\mathbb{R}^3$.

Area of a surface

Just like the length of a curve does not depend on how the the curve is put into $\mathbb{R}^n$, the area of a surface should not depend on how it curves inside $\mathbb{R}^3$. In other words, *area* should be an intrinsic quantity that can be computed just from the intrinsic information (U, g_{ij}). Therefore, we have the following definition:

Definition 12. *Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface. The area of the parametrized surface f is defined as*

$$A(f) := \int_U \sqrt{\det(g_{ij})}\, du_1 du_2,$$

where (g_{ij}) is the first fundamental form in the parametrization f.

To show that $A(f)$ is a geometric quantity that really calculates the area of the (image) surface, we need the following invariance under reparametrization. After all, the area of a surface should not depend on what coordinates we use to describe the surface.

Lemma 13. *Let $\tilde{f} : \tilde{U} \xrightarrow{\varphi} U \to \mathbb{R}^3$ be a reparametrization of a regular parametrized surface $f : U \to \mathbb{R}^3$. Then, $A(\tilde{f}) = A(f)$.*

Proof. Recall from Lemma 10 that the first fundamental form transforms in the way $(\tilde{g}_{ij}) = (D\varphi)^T (g_{ij})(D\varphi)$. Taking determinant on both sides, using that $\det(AB) = \det(A)\det(B)$ and $\det A^T = \det A$, we have

$$\det(\tilde{g}_{ij}) = (\det(D\varphi))^2 \det(g_{ij}).$$

Therefore, using the change of variable formula for double integrals, we have

$$A(\tilde{f}) = \int_{\tilde{U}} \sqrt{\det(\tilde{g}_{ij})}\, d\tilde{u}_1 d\tilde{u}_2 = \int_{\tilde{U}} \sqrt{\det(g_{ij})} |\det(D\varphi)|\, d\tilde{u}_1 d\tilde{u}_2 = \int_U \sqrt{\det(g_{ij})}\, du_1 du_2 = A(f).$$

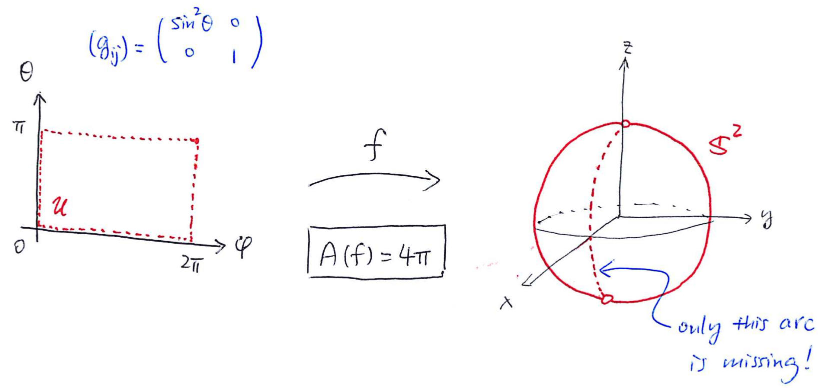
□

Note that the area is the same even for an orientation-reversing reparametrization $\tilde{f}$ with $\det(D\varphi) < 0$. This is because the Jacobian has an absolute value in it so the formula is insensitive to orientation. Later, we will introduce the concept of *differential forms* which are the natural objects to be integrated on a surface. These concept of forms will require a fixed orientation so that a general Stokes' theorem would hold. Think about the single-variable fundamental theorem of calculus which says $\int_a^b f'(x)\,dx = f(b) - f(a)$. The right hand side is evaluating the function at the end points with *different signs* because we are viewing the interval as an *oriented* curve, i.e. it starts from a and ends at b. This orientation issue is essential which allows cancellations of terms in the interior.

As an example, let's calculate the area of a unit sphere given in spherical coordinates by $f(\varphi, \theta) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, where $(\varphi, \theta) \in (0, 2\pi) \times (0, \pi) =: U$. Note that the image $f(U)$ of this parametrization covers almost the whole sphere $\mathbb{S}^2$ except that it misses a great arc joining the north and south poles, which has zero area. So $A(f)$ should in fact give the area of the unit sphere (which we know should be 4π). Recall from Example 12 (2) that we have $\sqrt{\det(g_{ij})} = \sin\theta > 0$. Therefore, we have

$$A(f) = \int_U \sqrt{\det(g_{ij})}\,du_1du_2 = \int_0^\pi \int_0^{2\pi} \sin\theta\,d\varphi d\theta = 4\pi.$$

One should compare this formula with the familiar formula of the volume element in $\mathbb{R}^3$ in spherical coordinates $dV = \rho^2 \sin\theta\,d\rho d\varphi d\theta$.



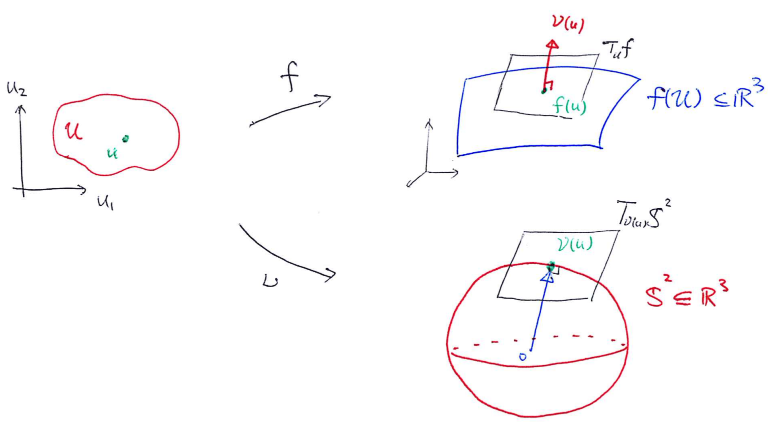
The Gauss map

We now turn to the study of the *extrinsic* geometry of surfaces. In contrast with the intrinsic geometry, we are interested in how the surface actually sits inside $\mathbb{R}^3$. In fact all the extrinsic geometry is contained in the normal bundle Nf of the surface. The *twisting* of the normal bundle tells us how the surface is curved inside $\mathbb{R}^3$. For surfaces in $\mathbb{R}^3$, the normal bundle is just a family of normal lines parametrized by the surface itself. Hence, the normal line is determined (up to a sign) by a unit normal vector at a point on the surface. This is the main result why we restrict to surfaces in $\mathbb{R}^3$ rather than in $\mathbb{R}^n$, because the normal bundle can be equivalently described (at least locally) by a unit normal vector field on the surface, instead of an $(n-2)$-dimensional sphere of unit normals at each point. It is similar to the reason why plane curves are simpler to study than space curves since the *codimension* is one.

Definition 14. *For any regular parametrized surface $f: U \to \mathbb{R}^3$, the Gauss map is defined to be the smooth map $\nu: U \to \mathbb{S}^2 \subset \mathbb{R}^3$ where*

$$\nu = \frac{\frac{\partial}{\partial u_1} \times \frac{\partial}{\partial u_2}}{\|\frac{\partial}{\partial u_1} \times \frac{\partial}{\partial u_2}\|}.$$

Note that the regularity of f implies that ν is a well-defined unit vector in $\mathbb{R}^3$, which depends smoothly on u. We remark that there are two ways to visualize the Gauss map ν. One is to view ν as a normal vector field along f where each unit vector $\nu(u)$ is based at the point $f(u)$ on the surface. Another way is to forget about the surface but simply think of ν as a smooth map that takes a point $u \in U$ to a point $\nu(u)$ on the unit sphere $\mathbb{S}^2$. We will use both point of views interchangeably from time to time.



Let's look at an explicit example. Consider again our favorite parametrization of the unit sphere $\mathbb{S}^2$ by spherical coordinates $f(\varphi, \theta) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, direct calculation shows that

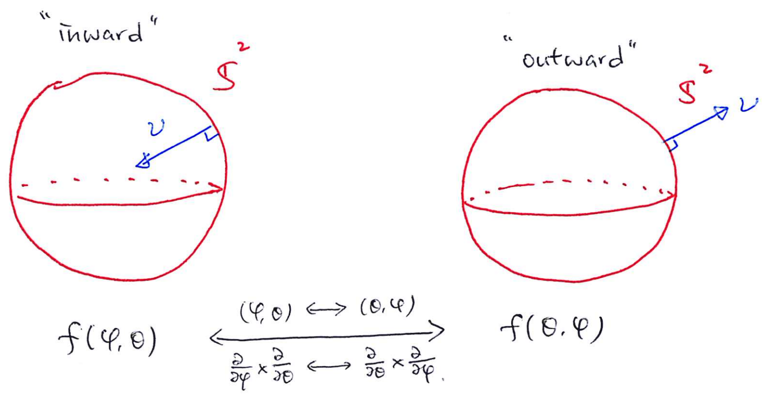
$$\frac{\partial}{\partial \varphi} = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0), \quad \frac{\partial}{\partial \theta} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta),$$

$$\frac{\partial}{\partial \varphi} \times \frac{\partial}{\partial \theta} = (-\sin^2\theta\cos\varphi, -\sin^2\theta\sin\varphi, -\sin\theta\cos\theta).$$

Therefore, $\|\frac{\partial}{\partial \varphi} \times \frac{\partial}{\partial \theta}\| = \sin\theta$ and thus we have

$$\nu = \frac{\frac{\partial}{\partial u_1} \times \frac{\partial}{\partial u_2}}{\|\frac{\partial}{\partial u_1} \times \frac{\partial}{\partial u_2}\|} = (-\sin\theta\cos\varphi, -\sin\theta\sin\varphi, -\cos\theta) = -f(\varphi, \theta).$$

Therefore, ν is the inward pointing unit normal on $\mathbb{S}^2$. To get the outward unit normal instead, we could simply flip the order of φ and θ, i.e. take $f(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, since the cross product changes sign if we interchange the order.



Although the definition of the Gauss map ν seemingly depends on the choice of the parametrization f, in fact it should not because the normal bundle Nf is well-defined independent of parametrizations. So the unit normal should be determined uniquely up to a sign.

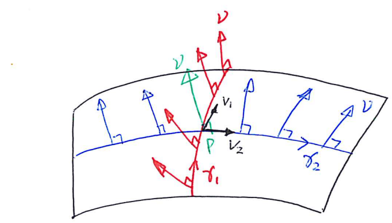
Lemma 15. *The Gauss map is well-defined, i.e. invariant under orientation-preserving reparametrizations. In other words, if* $\tilde{f} : \tilde{U} \xrightarrow{\varphi} U \xrightarrow{f} \mathbb{R}^3$ *is an orientation-preserving reparametrization (i.e.* $\det(D\varphi) > 0$*), then* $\tilde{\nu} = \nu \circ \varphi$*.*

Proof. Exercise for the reader. □

A quest for the notion of curvatures on surfaces

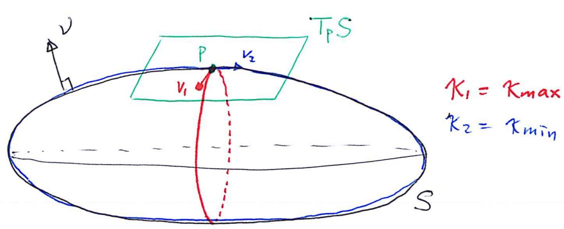
We now have the notion of Gauss map ν which encodes how the unit normal is changing along the surface. It will helps us define the notion of *curvatures* on a surface. Note that a flat plane has constant Gauss map, and any sensible definition of its curvatures should be zero. In the case of plane curves, we saw that curvature is the rate of change of the unit tangent vector e_1 in the normal direction e_2 when we move along the curve, i.e. $\kappa := \langle e_1', e_2 \rangle$. By the Frenet equations, it is equivalent to the rate of change of the unit normal e_2 in the tangential $-e_1$ direction, i.e. $\kappa = \langle e_2', -e_1 \rangle$. We will take this second point of view to generalize it to a notion of curvatures for a surface in $\mathbb{R}^3$.

Given our understanding of the concept of curvatures for space curves, we now attempt to define a sensible notion of curvatures for surfaces. The theory of curves is considerably simpler because curves are one-dimension, in other words, there is only one parameter (or degree of freedom) to move along a curve. That's why single variable calculus and ODE techniques are sufficient for a well-developed theory of curves in $\mathbb{R}^n$. However, the story is different for surfaces since there are two parameters (or degrees of freedom) to move along a surface. Suppose we are originally standing at a point p on a surface. After picking a direction v_1 tangential to the surface at p, we can move along a curve γ_1 on the surface towards the direction v_1 (note that we cannot stay in a straight line in $\mathbb{R}^3$ as we are constrained to move only on the surface.) One can then say that the rate of change of the unit normal ν along this curve γ_1 gives the "curvature" of the surface at p *along the direction* v_1. If we pick a different direction v_2 to begin with, we will get a different curve γ_2 on the surface. Presumably the rate of change of the normal ν would be different along this curve γ_2. Therefore, it seems that we will have a family of curvatures for each tangential direction at p, which can be parametrized by a circle. (Note that we do not care about the length of the "direction" since we always normalize our rate of change relative to unit-speed parametrizations.)



"The normal ν changes differently along different directions!"

Apparently this sounds like a nightmare. There are infinitely many "curvatures" at each point p, one for each tangential direction! Fortunately, the linearity of differentiation saves us from the nightmare. There are indeed only two distinctive curvatures κ_1 and κ_2, called *principal curvatures*, at each point p. One is the maximum curvature and the other one is the minimum. Each principal curvature corresponds to a *principal curvature direction*, called v_1 and v_2. The linearity property says that if we take a direction $v = av_1 + bv_2$, such that $a^2 + b^2 = 1$, then the curvature along the direction v should be given by the same linear combination $\kappa_v = a\kappa_1 + b\kappa_2$ of the principal curvatures. This is somewhat not too surprising as this is we expect from our knowledge of directional derivatives.



"an elongated ellipsoid"

Consider an ellipsoid with p being the "north pole", it is not hard to see that moving in the v_1 direction gives you the maximum rate of change of the unit normal ν and moving in the v_2 direction gives you the minimum. Note that the two principal directions v_1 and v_2 are *orthogonal* to each other. We will see shortly that this is indeed always the case!

Some useful facts from calculus and linear algebra

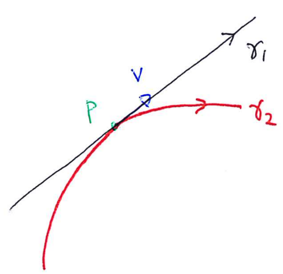
We now digress a little bit to recall some basic concepts and theorems in multivariable calculus and linear algebra. First, we discuss directional derivatives of a function $f : \mathbb{R}^3 \to \mathbb{R}$. Recall that the directional derivative of f at p in the direction v, where $\|v\| = 1$, is defined as

$$D_v f(p) := \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

Note that the definition still makes sense even when f is a vector-valued function (we just differentiate component-wise). Therefore, $D_v f(p)$ is just the derivative of the function f along the straight line parallel to v and passes through p at $t = 0$. In fact, as long as f is smooth enough, it is not necessary to take a straight line.

Lemma 16. *Let $c(t) : (-\epsilon, \epsilon) \to \mathbb{R}^3$ be a smooth curve such that $c(0) = p$ and $c'(0) = v$. Then, we have*

$$D_v f(p) = \left.\frac{d}{dt}\right|_{t=0} f(c(t)).$$



" $D_v f(p)$ can be calculated by differentiating along either γ_1 or γ_2, as long as they have the same tangent vector at p "

Note that when $c(t) = p + tv$, this reduces to the original definition of direction derivative. However, an important observation from the lemma is that to calculate $D_v f(p)$ we do not need f is be defined everywhere in $\mathbb{R}^3$, we just need to know the values of f along a curve through the point p and tangent to v at this point. In particular, if f is just defined on a surface $S \subset \mathbb{R}^3$, and v is a vector tangential to the surface at a point p, then we can calculate $D_v f(p)$ by differentiating it along a curve *lying on the surface* through p and is tangent to v there.

Second, we recall the spectral theorem for self-adjoint operators. Let V be an n-dimensional real vector space equipped with an inner product $\langle \cdot, \cdot \rangle$. A linear operator $S : V \to V$ is said to be self-adjoint if $\langle S(u), v \rangle = \langle u, S(v) \rangle$ for all $u, v \in V$. For any self-adjoint operator $S : V \to V$, one can associated (canonically through the inner product $\langle \cdot, \cdot \rangle$) to it a symmetric bilinear form $h : V \times V \to \mathbb{R}$ defined by $h(u, v) = \langle S(u), v \rangle$ and vice versa. Therefore, for an inner product space $(V, \langle \cdot, \cdot \rangle)$, self-adjoint operators S and symmetric bilinear forms h are isomorphic objects. The *spectral theorem* for self adjoint operators is as follows:

Theorem 17. *Any self-adjoint operator on a finite dimensional inner product space (over $\mathbb{R}$) is diagonalizable with an orthonormal basis of eigenvectors whose eigenvalues are all real.*

One proof of the spectral theorem above is to use the variational characterization of eigenvalues using *Rayleigh quotients*. The eigenvalues (and eigenvectors) can be obtained through an extremal problem. For example, when $\dim(V) = 2$, then the eigenvalues κ_1, κ_2 of a self-adjoint operator S (with associated bilinear form h) is given by

$$\kappa_1 = \min_{v \neq 0} \frac{h(v, v)}{\|v\|^2}, \quad \kappa_2 = \max_{v \neq 0} \frac{h(v, v)}{\|v\|^2}.$$

We emphasize that the important conclusion of the spectral theorem is that not only is the self-adjoint operator diagonalizable. We can find an eigenbasis consisting of *orthonormal* eigenvectors. This will be relevant to our previous observation that the principal curvature directions have to be orthogonal to each other.

The Weingarten map and second fundamental form

We are now ready to define the notions of curvatures for surfaces. We will make use of our intuition that curvature should be some kind of derivative of the Gauss map ν. Remember we can think of ν as a unit normal vector field that associates to each point on a surface with its unit normal vector, so it is a vector-valued function defined on the surface. Therefore, we can take the directional derivative $D_v \nu$ of the Gauss map in the direction of a tangent vector v at a point p on the surface. We first observe that $D_v \nu$ is always a vector tangent to the surface at p.

Lemma 18. *Let $\nu : U \to \mathbb{S}^2 \subset \mathbb{R}^3$ be the Gauss map of a regular parametrized surface $f : U \to \mathbb{R}^3$. Then, $\frac{\partial \nu}{\partial u_i} \in T_u f$ for $i = 1, 2$. Therefore, for any $v \in T_u f$, we have $D_v \nu \in T_u f$.*

Proof. Since ν is a unit vector, $\langle \nu, \nu \rangle \equiv 1$. Differentiating with respect to u_i gives $\langle \frac{\partial \nu}{\partial u_i}, \nu \rangle = 0$. Therefore, $\frac{\partial \nu}{\partial u_i} \in T_u f$ since $\nu \perp T_u f$. □

Definition 19. *We define the Weingarten map or shape operator $S : T_u f \to T_u f$ as the linear operator*

$$S(v) = D_v \nu,$$

i.e. $S = D\nu \circ (Df)^{-1}$.

In the homework exercise, you will be asked to prove that S is well-defined independent of parametrizations. Therefore, S is a well-defined linear operator on the tangent space of the surface. Remember that the tangent space $T_u f$ comes with an inner product g, the first fundamental form, which is just the restriction of the Euclidean inner product $\langle \cdot, \cdot \rangle$. The following lemma is important.

Lemma 20. *The shape operator S is a self adjoint operator on the tangent space $T_u f$ with respect to the first fundamental form g.*

Proof. By definition of self adjoint operators, we have to show that $g(S(X), Y) = g(X, S(Y))$ for all $X, Y \in T_u f$. By linearity, it suffices to check it for $X = \frac{\partial f}{\partial u_1}$ and $Y = \frac{\partial f}{\partial u_2}$. Note that we try to be more careful to distinguish the two vectors $\frac{\partial}{\partial u_i} \in T_u U$ and $\frac{\partial f}{\partial u_i} \in T_u f$ to avoid confusion. By the definition of $S = D\nu \circ (Df)^{-1}$, we have

$$g\left(S(\frac{\partial f}{\partial u_1}), \frac{\partial f}{\partial u_2}\right) = g\left(D\nu(\frac{\partial}{\partial u_1}), \frac{\partial f}{\partial u_2}\right) = \langle \frac{\partial \nu}{\partial u_1}, \frac{\partial f}{\partial u_2} \rangle = -\langle \nu, \frac{\partial^2 f}{\partial u_1 \partial u_2} \rangle,$$

where in the last equality we have used "differentiation by part" and the fact that $\langle \nu, \frac{\partial f}{\partial u_2} \rangle \equiv 0$. Similarly, we find that

$$g\left(S(\frac{\partial f}{\partial u_2}), \frac{\partial f}{\partial u_1}\right) = -\langle \nu, \frac{\partial^2 f}{\partial u_2 \partial u_1} \rangle.$$

By the equality of mixed partial derivatives, $\frac{\partial^2 f}{\partial u_1 \partial u_2} = \frac{\partial^2 f}{\partial u_2 \partial u_1}$, and hence the right hand side are the same. The assertion follows. □

Since S is a self adjoint operator on the inner product space $(T_u f, g)$, by the identification of self adjoint operators and symmetric bilinear forms. We can define the following:

Definition 21. *The second fundamental form $h : T_u f \times T_u f \to \mathbb{R}$ is the symmetric bilinear form on the tangent space $T_u f$ associated to the self-adjoint shape operator S, i.e.*

$$h(X, Y) := g(S(X), Y).$$

Note that h is a symmetric bilinear form, i.e. a $(0, 2)$-tensor, but it is not necessarily positive definite like the first fundamental form g. As for the first fundamental form, we can write h in local coordinates as a 2×2 symmetric matrix:

$$(h_{ij}) = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix},$$

where $h_{ij} := h(\frac{\partial}{\partial u_i}, \frac{\partial}{\partial u_j})$. We note that h also transforms similar to g does by $(\tilde{h}_{ij}) = (D\varphi)^T (h_{ij})(D\varphi)$ (c.f. Lemma 10). (Exercise: Prove this!) Note that from the proof of Lemma 20, we have

$$h_{ij} = \langle \frac{\partial \nu}{\partial u_i}, \frac{\partial f}{\partial u_j} \rangle = -\langle \nu, \frac{\partial^2 f}{\partial u_i \partial u_j} \rangle.$$

Gauss and mean curvatures

With the shape operator S and second fundamental form h defined, we can now define our notions of curvatures.

Definition 22. *The Gauss curvature K and mean curvature H of a regular parametrized surface $f : U \to \mathbb{R}^3$ is defined respectively by:*

$$K := \det(S), \qquad H := \operatorname{tr}(S).$$

Here, det *and* tr *are the determinant and trace of an operator (which is defined independent of a chosen basis).*

Before we study more closely these two notions of curvatures, let us look at our favorite example of the parametrization of the unit sphere by spherical coordinates $f(\theta, \varphi)$, recall that in this order of θ, φ, we get $\nu = f$. Therefore, $D\nu = Df$ and thus $S = \mathrm{id}$ is just the identity operator on the tangent space. So we have $H = 2$ and $K = 1$. So the unit sphere $\mathbb{S}^2$ has constant (positive) Gauss and mean curvatures.

Applying the spectral theorem (Theorem 17) to the self adjoint shape operator S on $T_u f$, we know that S is diagonalizable by an orthonormal basis of eigenvectors.

Definition 23. *An eigenvector of the shape operator S is called a principal curvature direction and the corresponding eigenvalue is called the principal curvature. Therefore, in an orthonormal basis of eigenvectors for S, we can express the operator S as a matrix*

$$S \sim \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix},$$

where κ_i are the principal curvatures. Hence, $H = \kappa_1 + \kappa_2$ and $K = \kappa_1 \kappa_2$.

Beware that in a basis, the matrix of S is in general not the same as the matrix (h_{ij}), except in the case that the basis is an orthonormal basis with respect to g. Although S and h essentially represent the same information, they transform differently under change of basis. In the language of tensors, we say that the $(1,1)$-tensor S can be obtained from the $(0,2)$-tensor h by raising an index using the metric g. However, we still have the following formula which is often a useful way to calculate K and H.

Proposition 24. *Let (g_{ij}) and (h_{ij}) be the matrices representing the first and second fundamental form in the coordinate basis induced by some parametrization $f : U \to \mathbb{R}^3$. Then, we have*

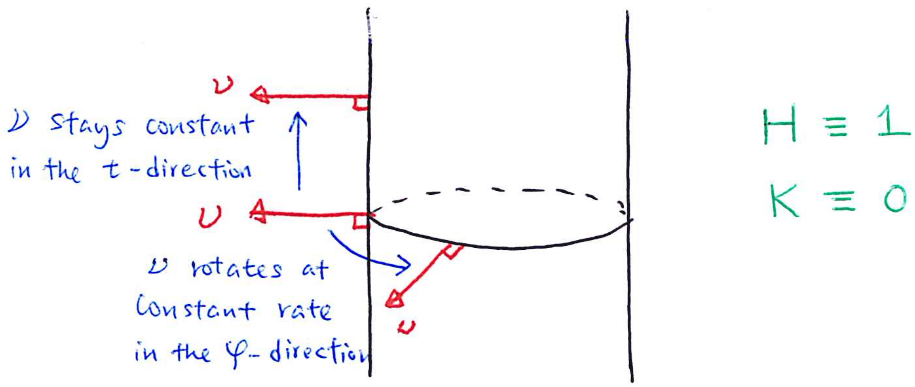
$$K = \frac{\det(h_{ij})}{\det(g_{ij})}, \qquad H = \operatorname{tr}((g_{ij})^{-1}(h_{ij})).$$

Proof. First of all, using the transformation laws for the first and second fundamental forms, i.e. $(\tilde{g}_{ij}) = (D\varphi)^T(g_{ij})(D\varphi)$ and $(\tilde{h}_{ij}) = (D\varphi)^T(h_{ij})(D\varphi)$, we see that the formula on the right hand side is independent of reparametrization. Therefore, by taking an orthonormal basis of $T_u f$ with respect to g, we get the desired result. $\square$

Let us make use of the formulas in Proposition 24 for an explicit example. Consider a parametrization of the cylinder given by $f(\varphi, t) = (\cos\varphi, \sin\varphi, t)$, in this coordinates the first and second fundamental forms are

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (h_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Therefore, we have $K \equiv 0$ and $H \equiv 1$. Note that the principal curvature directions are given by $\frac{\partial}{\partial\varphi}$ and $\frac{\partial}{\partial t}$ with principal curvatures 1 and 0 respectively.



Higher fundamental forms

We have now defined two fundamental forms g and h for a parametrized surface and used them to define the Gauss and mean curvatures K and H. We may ask if we need to compute the third, fourth etc. fundamental forms to know the surface better. Fortunately, the answer is no because all the higher fundamental forms can be derived from the first and second fundamental forms. Therefore, in principle, the first two fundamental forms are all we need to determine the surface.

Recall that $g(X, Y) := \langle X, Y\rangle$ and $h(X, Y) := \langle S(X), Y\rangle$ where S is the shape operator defined in Definition 19. Thus, we can define similarly the higher fundamental forms by taking higher powers of S:

Definition 25. *The third fundamental form III is the bilinear form on the tangent space defined by*

$$III(X, Y) := \langle S^2(X), Y\rangle.$$

By self-adjointness of the shape operator S, III is a symmetric bilinear form. The proposition below tells us that III is completely determined by g and h. Note that the Gauss and mean curvatures K and H are determined by g and h (Proposition 24).

Proposition 26. *$III = Hh - Kg$.*

Proof. Since both sides of the formula are bilinear forms, it suffices to check that they agree on *any* basis. In particular, if we take an eigenbasis $\{v_1, v_2\}$ for the self-adjoint operator S, then we have

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (h_{ij}) = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix},$$

and also $H = \kappa_1 + \kappa_2$, $K = \kappa_1 \kappa_2$. Since v_i are eigenvectors for S, we have $Sv_i = \kappa_i v_i$. Therefore,

$$III(v_i, v_j) = \langle S^2(v_i), v_j \rangle = g(S(\kappa_i v_i), v_j) = \kappa_i^2 g(v_i, v_j) = \kappa_i^2 \delta_{ij},$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$ is the Kronecker delta. Therefore, in this basis we have

$$III = \begin{pmatrix} \kappa_1^2 & 0 \\ 0 & \kappa_2^2 \end{pmatrix}.$$

On the other hand,

$$Hh - Kg = (\kappa_1 + \kappa_2) \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} - \kappa_1 \kappa_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \kappa_1^2 & 0 \\ 0 & \kappa_2^2 \end{pmatrix} = III.$$

□

Geometric meaning of curvatures

Now we have defined two notions of curvatures for surfaces: Gauss curvature K and mean curvature H. So what do these curvatures represent about the surface geometrically? You will see in the homework problems that K and H scales accordingly when we dilate or shrink a surface by a factor. Therefore, modulo similarity, we would like to understand what geometric meaning is carried by the signs of K and H.

First of all, we will argue that the sign of H do not have much geometric meaning. The reason is that one can always flip the sign of H by changing the orientation of the surface (i.e. by switching "inward" and "outward" unit normal ν). This can be observed as follows. If we switch ν to $-\nu$, then the shape operator S changes sign and so are its eigenvalues. Therefore, the mean curvature H, which is defined to be the *sum* of eigenvalues, changes sign. However, the Gauss curvature K, which is defined to be the *product* of eigenvalues, does not change sign! This insensitivity of K to the choice of normal ν hints that the Gauss curvature K may be something intrinsic to the surface (in fact this is true and it is the famous *Theorema Egregium* of Gauss - we will talk about this when we study the intrinsic geometry of surfaces).

Recall that for curves we define the curvature $\kappa := \|c''\|$ as some kind of second derivatives of c. By putting a surface locally into a *canonical form*, we can also see that our notions of curvatures for surfaces are some kind of second derivatives. Here is how to see this. Suppose we have a regular parametrized surface $f(x,y) : U \to \mathbb{R}^3$ such that $0 \in U$, and that it is a graphical surface

$$f(x,y) := (x, y, u(x,y)),$$

for some function $u : U \to \mathbb{R}$. Assume that the graph is put into "standard position", i.e. $u(0) = 0$ and $\nabla u(0) = 0$. Then, a direct calculation shows that the second fundamental form at $(0,0)$ is given by

$$(h_{ij})_{(0,0)} = \begin{pmatrix} -u_{xx} & -u_{xy} \\ -u_{yx} & -u_{yy} \end{pmatrix}\Bigg|_{(0,0)} = -\text{Hess}_{(0,0)}(u).$$

Therefore, using the formula in Proposition 24, the Gauss and mean curvatures at $(0,0)$ will be given by

$$K(0,0) = u_{xx}u_{yy} - u_{xy}^2, \qquad H(0,0) = -u_{xx} - u_{yy},$$

which are the second derivatives of the defining function u at $(0,0)$.

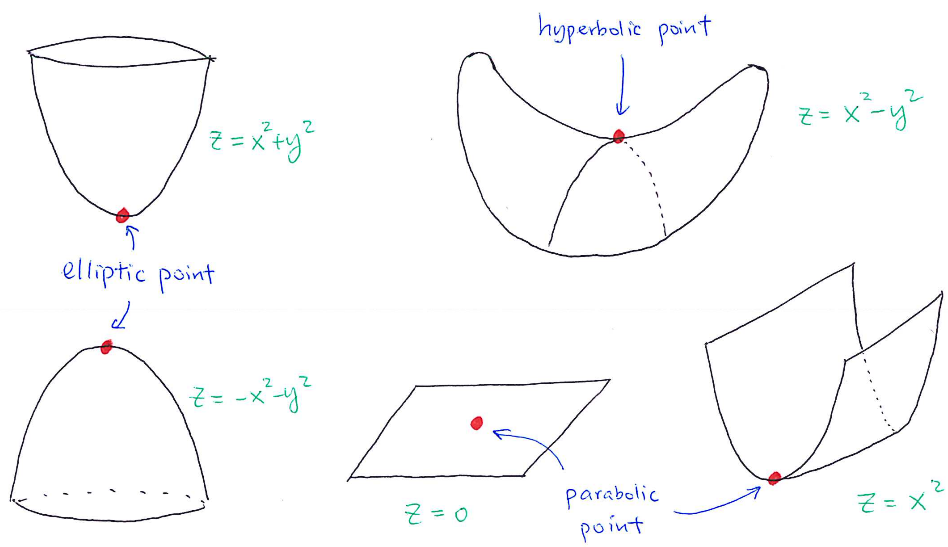
Definition 27. *Given a regular parametrized surface $f : U \to \mathbb{R}^3$, a point $p = f(u)$ is said to be a*

(a) elliptic point if the principal curvatures have the same signs, i.e. either $\kappa_1, \kappa_2 > 0$ or $\kappa_1, \kappa_2 < 0$;

(b) hyperbolic point if the principal curvatures have different signs, i.e. either $\kappa_1 < 0 < \kappa_2$ or $\kappa_2 < 0 < \kappa_1$;

(c) parabolic point if either one of the principal curvatures is zero or both.

The picture below shows some typical cases of these scenarios:



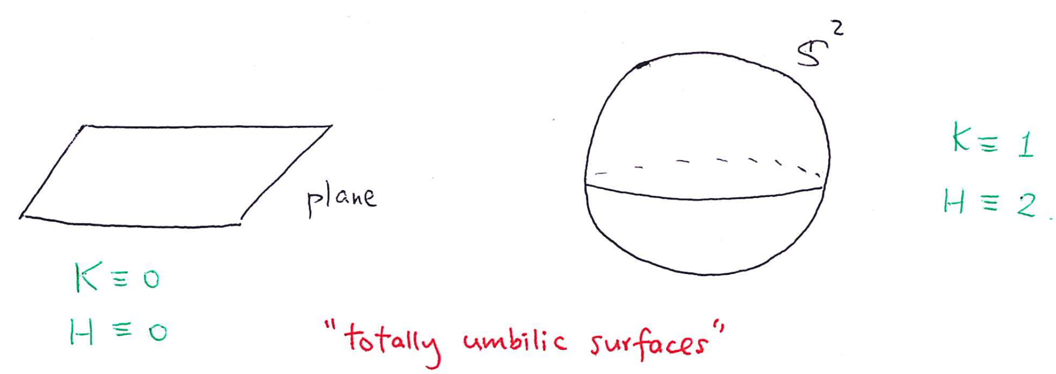
Here is a summary of what we can say about each case:

type	K	H
elliptic	> 0	> 0 or < 0
hyperbolic	< 0	undetermined
parabolic	$= 0$	undetermined

From the table we see that the type of points is completely classified by the sign of the Gauss curvature at that point. The mean curvature does not say anything much on this. From the pictures above, we sort of see that the Gauss curvature K measures how the surface deviates (intrinsically) from the flat plane. The mean curvature H, on the other hand, measures whether the positive or the negative curvature is dominating. For example, a mean convex surface (i.e. $H > 0$) means that the bigger positive curvature always dominate the other (possibly negative) curvature. Note that surfaces with $H < 0$ indeed have the same meaning since one can flip the sign by changing orientation. The sign of H in fact says that whether the surface is curving towards or away from the chosen unit normal ν.

Totally umbilic surfaces

In our study of plane curves, we saw that only the straight line and the circles have constant curvature. For surfaces, we have studied two examples which also have constant curvatures:



Notice that these two examples not only have constant K and H, but all the principal curvatures are the same and constant across the whole surface. We give a name to situations like this.

Definition 28. *A point p on a regular parametrized surface is said to be an umbilic point if the principal curvatures at p are the same, i.e. $\kappa_1 = \kappa_2$ at p. A regular parametrized surface is said to be totally umbilic if all the points are umbilic.*

Note that a totally umbilic surface just have the two principal curvatures matching at each point p, but the value of the principal curvatures could change from point to point on the surface. However, by the theorem below, in fact this cannot happen and the plane and the spheres are the only examples.

Theorem 29. *A totally umbilic surface in $\mathbb{R}^3$ is either a piece of a plane or a sphere.*

Proof. Let $f : U \to \mathbb{R}^3$ be a totally umbilic regular parametrized surface. Since it is totally umbilic, we have $\kappa_1 = \kappa_2 = \lambda(u)$ for some function $\lambda : U \to \mathbb{R}$. We will first show that $\lambda(u)$ is in fact constant (assuming U is connected) and then we use this to show that the surface must be part of a plane or sphere.

Since $\kappa_1 = \kappa_2 = \lambda(u)$, by definition, we have $S(v) = \lambda(u)v$ for all $v \in T_u f$. In particular, taking $v = \frac{\partial f}{\partial u_i}$ we get

$$\frac{\partial \nu}{\partial u_1} = \lambda(u)\frac{\partial f}{\partial u_1} \qquad \text{and} \qquad \frac{\partial \nu}{\partial u_2} = \lambda(u)\frac{\partial f}{\partial u_2}.$$

Differentiating the first equation with respect to u_2:

$$\frac{\partial^2 \nu}{\partial u_2 \partial u_1} = \frac{\partial^2 f}{\partial u_2 \partial u_1} + \frac{\partial \lambda}{\partial u_2}\frac{\partial f}{\partial u_1}.$$

Similarly, if we differentiate the second equation with respect to u_2, we obtain

$$\frac{\partial^2 \nu}{\partial u_1 \partial u_2} = \frac{\partial^2 f}{\partial u_1 \partial u_2} + \frac{\partial \lambda}{\partial u_1}\frac{\partial f}{\partial u_2}.$$

Now, comparing the last two equations and using that mixed partial derivatives are equal, we conclude that

$$\frac{\partial \lambda}{\partial u_2}\frac{\partial f}{\partial u_1} - \frac{\partial \lambda}{\partial u_1}\frac{\partial f}{\partial u_2} = 0.$$

Note that this is a vector equation and since $\{\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}\}$ is a basis for $T_u f$ (hence is linearly independent), the coefficients must all vanish. Hence, we get

$$\frac{\partial \lambda}{\partial u_1} = \frac{\partial \lambda}{\partial u_2} = 0,$$

which implies that $\lambda(u) \equiv$ constant since U is connected.

To prove that the surface is in fact a plane or a sphere, we divide into two cases: $\lambda = 0$ or $\lambda \neq 0$. When $\lambda = 0$, we have $S = 0$ and hence ν is constant across the whole surface. This can only happen for a plane (Exercise: can you proof this rigorously?). In the second case, $\lambda \neq 0$ and since $D\nu = \lambda Df$, we have $f - \frac{1}{\lambda}\nu \equiv p_0$ where p_0 is a constant vector (or point) in $\mathbb{R}^3$. This point p_0 is supposed to be the center of the sphere. On the other hand, note that $\|\frac{1}{\lambda}\nu\| \equiv \frac{1}{\lambda}$ as ν is a unit vector. This shows that the surface f lies in the sphere of radius $\frac{1}{\lambda}$ centered at p_0. □

Embedded submanifolds and orientability

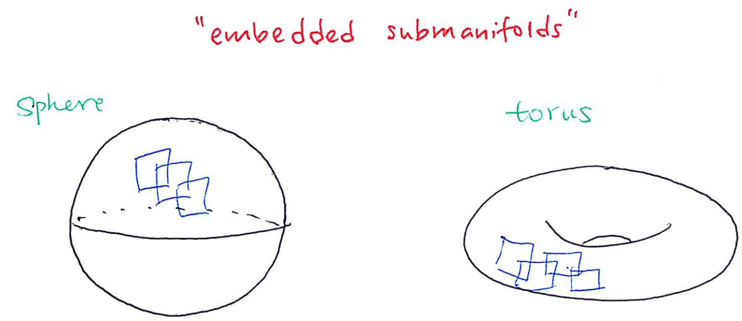
So far we have studied surfaces as regular parametrized surfaces $f : U \to \mathbb{R}^3$. In the course of our study, we have seen that many of the concepts we defined are actually independent of parametrizations. So these are indeed geometric concepts defined on the image surface $f(U)$ inside $\mathbb{R}^3$ which has nothing to do with the chosen parametrization f. This motivates our definition below.

Definition 30. *A subset $\Sigma \subset \mathbb{R}^3$ is said to be a (two dimensional) embedded submanifold in $\mathbb{R}^3$ if locally it is the image of a regular parametrized surface, i.e. for all $p \in \Sigma$, there exists some $\epsilon > 0$ such that for the open ball $B_\epsilon(p)$ of radius ϵ centered at p, we have*

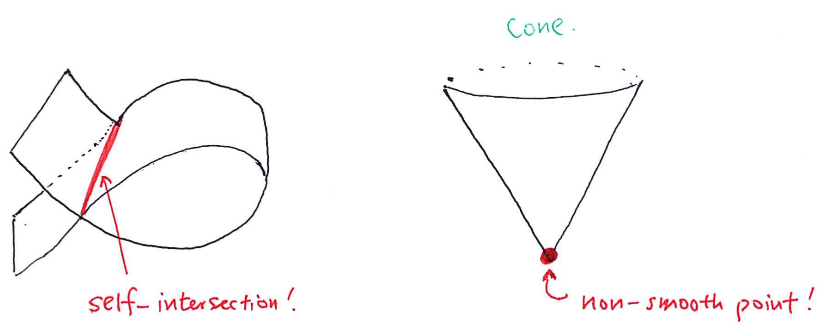
$$\Sigma \cap B_\epsilon(p) = image(f)$$

for some regular parametrized surface $f : U \to \mathbb{R}^3$.

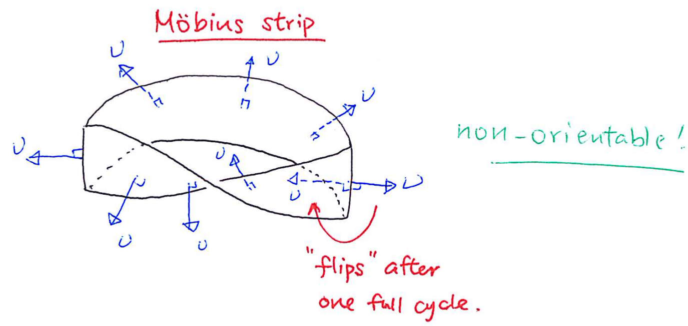
Therefore, an embedded submanifold is just pieces of $\mathbb{R}^2$ patched together to form the entire "surface". We often called the parametrizations f "*charts*" where we put coordinates near that point through the choice of chart f.



Note that the definition of embedded submanifolds rules out the case of self-intersection (which is allowed for a regular parametrized surface!) and non-smooth points.



Note that by the parametrization invariance property, the concept of tangent space $T_p\Sigma$ and normal space $N_p\Sigma$ make perfect sense at a point $p \in \Sigma$ of an embedded submanifold. Moreover, one can visualize the Gauss map as a map $\nu : \Sigma \to \mathbb{S}^2$, where we associate each $p \in \Sigma$ to a unit normal vector $\nu(p) \in N_p\Sigma$. However, for an embedded submanifold Σ, a *continuous* globally defined unit normal $\nu : \Sigma \to \mathbb{S}^2$ may not exist. A famous example is given by the Möbius band. If you fix a unit normal a point on the band and goes around the band once, varying the normal in a continuous way, we see that it will not match with the original choice of normal!



Definition 31. *An embedded submanifold $\Sigma \subset \mathbb{R}^3$ is* <u>*orientable*</u> *if there exists a globally defined unit normal $\nu : \Sigma \to \mathbb{S}^2$ which is continuous. Otherwise, it is said to be* <u>*non-orientable*</u>.

Note that orientability is a *global* property. Any small piece of an embedded submanifold is orientable (since it can be oriented by the choice of a parametrization f).

We end with a remark that orientability is an important notion which has already (secretly) made its debut when you learn single variable calculus. Remember the fundamental theorem of calculus says that for a continuously differentiable function $f : [a, b] \to \mathbb{R}$, we have

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

The key point is that integration of derivative over the interval $I = [a, b]$ becomes evaluations at the boundary $\partial I = \{a, b\}$. However, the end points a and b are assigned with different signs! This implicitly says that a is the *starting point* and b is the *end point*. Therefore, we are in fact thinking of the interval $I = [a, b]$ as an *oriented* interval from a to b. This seemingly over-careful distinction is indeed the whole idea why fundamental theorem of calculus works. All the interior contributions are cancelled out because they have different signs! This is also the heart of the proofs of all the fundamental theorem of calculus type results you have learned in advanced calculus, for example the Green's, Stokes' and Divergence Theorems. Later on, we will see that how one can take into account of orientation by the introduction of *differential forms*, these are in fact the actual objects that one is integrating. We will postpone this theory until we have studied the intrinsic geometry of surfaces.