MATH 4030 Differential Geometry
Lecture Notes Part 3
last revised on November 25, 2015

We now move on to study the intrinsic geometry of surfaces. By *intrinsic* we mean all the geometric concepts and quantities that can be defined using the first fundamental form g alone. In other words, we just care about the tangential part but forget about the normal part (which is something extrinsic to the surface). The most important theorem here is the famous *Theorema Egregium* which says that the Gauss curvature K defined previously using the shape operator S, which is an extrinsic object, is indeed an intrinsic curvature which can be computed from the first fundamental form g and its derivatives only.

Curves on a surface

We will first begin with curves on surfaces and see that how the space curvature κ of a curve on the surface decompose into an intrinsic and extrinsic part. From our discussion of embedded submanifold, we know that all the concepts like the tangent space $T_p\Sigma$ and normal space $N_p\Sigma$ makes sense on an embedded submanifold $\Sigma \subset \mathbb{R}^3$, and we can identify $T_uU \cong T_uf \cong T_p\Sigma$ using the differential Df.

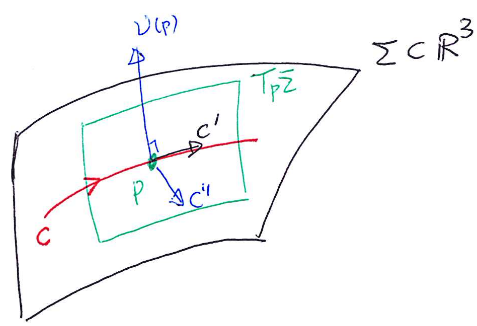
Let $\Sigma \subset \mathbb{R}^3$ be an embedded submanifold. Consider a space curve $c(s) : (a,b) \to \mathbb{R}^3$, parametrized by arc length, which lies completely in the surface Σ, i.e. $c(s) \in \Sigma$ for all $s \in (a,b)$. Let $p = c(s_0)$. Recall that the curvature of c at $s = s_0$ is defined as

$$\kappa := \|c''(s_0)\|.$$

Moreover, $c''(s_0)$ is a vector in $\mathbb{R}^3$ (with based point at p) and hence it decomposes into its tangential and normal components:

$$c''(s_0) = (c''(s_0))^T + (c''(s_0))^N,$$

with $(c''(s_0))^T \in T_p\Sigma$ and $(c''(s_0))^N \in N_p\Sigma$.



Definition 1. *We say that* $k_\nu := \|(c''(s_0))^N\|$ *is the* <u>*normal curvature*</u> *of* c *at* s_0 *and* $k_g := \|(c''(s_0))^T\|$ *is the* <u>*geodesic curvature*</u> *of* c *at* s_0*.*

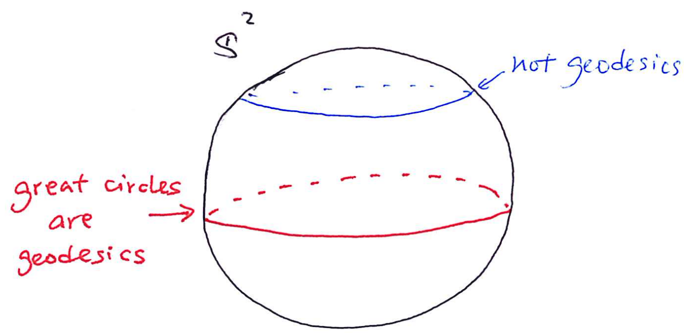
Since the tangent and normal spaces are orthogonal to each other, we have by Pythagorus theorem

$$\kappa^2 = k_g^2 + k_\nu^2.$$

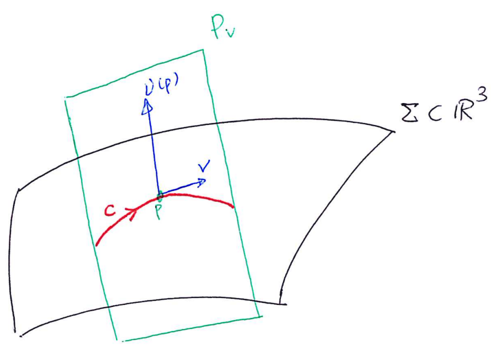
Therefore, the full curvature κ for the space curve c is a sum of the "tangential curvature" k_g and the "normal curvature" k_ν. When there is no "tangential curvature" on the entire curve, it is like a *straight line* on the surface.

Definition 2. *A curve* c *lying on an embedded submanifold* Σ *is a* <u>*geodesic*</u> *if* $k_g \equiv 0$ *on the curve* c*.*

For example, an equatorial circle on $\mathbb{S}^2$ is a geodesic while any other circle on it is not.



To obtain a curve on a given embedded submanifold $\Sigma \subset \mathbb{R}^3$, one can proceed as follows. Let $p \in \Sigma$ be a point on the surface whose unit normal is $\nu(p)$ at that point. Take any unit tangent vector $v \in T_p\Sigma$. The two vectors $\nu(p)$ and v span a plane P_v in $\mathbb{R}^3$ passing through the point p and contains $\nu(p)$ and v. Then, $P \cap \Sigma$ is (locally near p) a curve c on the surface Σ which passes through p and is tangent to v at p.



Let us calculate the various curvatures at p for this curve c. Assume $c(s)$ is an arc length parametrization of c with $c(0) = 0$ and $c'(0) = v$. Note that since $c(s) \in P$ for all s, the derivatives $c'(s)$ and $c''(s)$ must all parallel to P for all s. For simplicity, we just assume $p = 0$ is the origin. Then, $c'(0) = v$ and $c''(0) \in P$. Since $c''(0)$ is orthogonal to $c'(0)$ (since the curve is arc-length parametrized), we must have $c''(0)$ parallel to $\nu(p)$, which implies that $k_g(0) := \|(c''(0))^T\| = 0$. For the normal component, using differentiation by part,

$$k_\nu(0) := \|(c''(0))^N\| = |\langle c''(0), \nu(p)\rangle| = | - \langle v, D_v\nu(p)\rangle| = |h(v,v)|.$$

This gives another interpretation of the second fundamental form $h(v,v)$.

Note that the calculation above holds *only at the point* p. If $\nu \in P$ for a *fixed* plane P, then we get $k_g \equiv 0$ everywhere along the curve $c = P \cap \Sigma$ and hence c is a geodesic on the surface (think about the case for $\mathbb{S}^2$).

Einstein summation convention

We now introduce the widely adopted convention commonly used in tensor calculations. This has to do with *indices* (which appears everywhere in Riemannian Geometry). To make use of this convention, we have to separate indices into two types: *upper* and *lower* indices. In this convention, we always write the local coordinate functions as upper indices:

$$u^1, \cdots, u^n \qquad \text{or simply} \qquad u^i.$$

The coordinate vector fields generated by such a coordinate system will be denoted by lower indices:

$$\frac{\partial}{\partial u^1}, \cdots \frac{\partial}{\partial u^n} \qquad \text{or simply} \qquad \partial_i.$$

Hence, any (tangential) vector field X can be written as a linear combination of the form

$$X = \sum_{i=1}^n a^i \frac{\partial}{\partial u^i} = a^i \partial_i.$$

The last equality is what we meant by using the <u>Einstein summation convention</u>. The components are represented by upper indices a^i, and when we see the same index appearing as an upper and lower indices at the same time, it means that we are summing over the possible range of this index i (from 1 to n in our case). This convention saves us some ink on writing the summation signs which often become clumsy when you are summing over a lot of indices. Note that when we sum over an index, the "name" of the index is irrelevant as it is just a dummy index. For example, we can write

$$X = a^i \partial_i = a^j \partial_j = a^k \partial_k.$$

As an example, recall that the first fundamental form in local coordinates is given by $g_{ij} := g(\partial_i, \partial_j)$. Therefore, for any vector fields $X = X^i \partial_i$ and $Y = Y^j \partial_j$, then

$$g(X,Y) = g_{ij} X^i Y^j.$$

Note that the two indices i and j are being summed over by Einstein's summation convention. If we think of (g_{ij}) as a matrix, then we denote the inverse matrix using upper indices, i.e. $(g_{ij})^{-1} = (g^{ij})$. Hence we have $g^{ij} g_{jk} = \delta^i_k$, which just says that the product of a matrix with its inverse is the identity matrix. We will be using Einstein's summation convention throughout the course from now on.

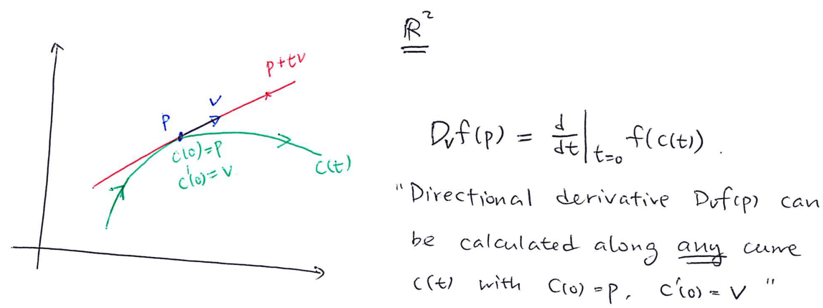
Calculus of vector fields in $\mathbb{R}^n$

We now revisit some of the concepts about vector fields in $\mathbb{R}^n$ and in the next section we will see how some of these concepts would be generalized to vector fields on surfaces.

First, we give a new perspective of vector fields in $\mathbb{R}^n$ as directional derivative. Recall that for a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, the directional derivative of f at p in the direction v is given by

$$D_v f(p) := \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t} = \left. \frac{d}{dt} \right|_{t=0} f(c(t)),$$

for *any* curve $c(t)$ in $\mathbb{R}^n$ such that $c(0) = p$ and $c'(0) = v$. Hence, to calculate the directional derivative $D_v f(p)$, we just need to know the vector v and the value of the function f restricted to *any* curve c passing through p and tangent to v at p. This is a useful observation when we talk about directional derivative on surfaces.



Note that we can interpret the directional derivative as the tangent vector $v \in T_p\mathbb{R}^n$ acts on functions by taking the directional derivative at that point. If we have a vector field $X : \mathbb{R}^n \to \mathbb{R}^n$, i.e. a smooth assignment to each point p a vector $X(p) \in T_p\mathbb{R}^n$, then the vector field X acts on smooth functions $f \in C^\infty(\mathbb{R}^n)$ and output another smooth function $X(f)$ by taking the directional derivative in the direction $X(p)$ at each point p, i.e.

$$X(f)(p) := D_{X(p)} f(p).$$

For example, if $X = (1, 0, \cdots, 0)$, then $X(f) = \frac{\partial f}{\partial u^1}$. This explains why we have the notation $X = \frac{\partial}{\partial u^i}$ since we are viewing the vector field X as a first order *partial differential operator*. This operator satisfies linearity and the product rule: let $f, g \in C^\infty(\mathbb{R}^n)$,

(i) $X(af+bg) = aX(f) + bX(g)$ for any constant $a, b \in \mathbb{R}$,

(ii) $X(fg) = fX(g) + gX(f)$.

The Clairaut's theorem says that mixed partial derivatives are equal (as long as the function being differentiated is smooth enough), i.e.

$$\frac{\partial^2 f}{\partial u^i \partial u^j} = \frac{\partial^2 f}{\partial u^j \partial u^i}.$$

Hence, if we write $X = \frac{\partial}{\partial u_i}$ and $Y = \frac{\partial}{\partial u^j}$, then it becomes $X(Y(f)) = Y(X(f))$, i.e. the two vector fields X and Y *commute* as operators acting on functions. Just like matrices, this is not true in general and the failure of commutativity is measured by the *Lie bracket.*

Definition 3. *The Lie bracket of two vector fields X and Y in $\mathbb{R}^n$ is defined by*

$$[X, Y] := XY - YX.$$

Note that we interpret the right hand side as an operator acting on functions. So, $[X, Y](f) = X(Y(f)) - Y(X(f))$.

Lemma 4. *$[X, Y]$ is a vector field.*

Proof. Let $X = X^i \partial_i$ and $Y = Y^j \partial_j$. Then for any $f \in C^\infty(\mathbb{R}^n)$, we have

$$\begin{array}{rcl} [X, Y] & = & X(Y(f)) - Y(X(f)) \\ & = & (X^i \partial_i)(Y^j \partial_j f) - (Y^j \partial_j)(X^i \partial_i f) \\ & = & X^i(\partial_i Y^j)(\partial_j f) + X^i Y^i (\partial_i \partial_j f) - Y^j(\partial_j X^i)(\partial_i f) - X^i Y^j (\partial_j \partial_i f) \\ & = & X^i(\partial_i Y^j)(\partial_j f) - Y^i(\partial_i X^j)(\partial_j f) + X^i Y^j (\partial_i \partial_j f - \partial_j \partial_i f) \\ & = & (X^i(\partial_i Y^j) - Y^i(\partial_i X^j))\partial_j f \end{array}.$$

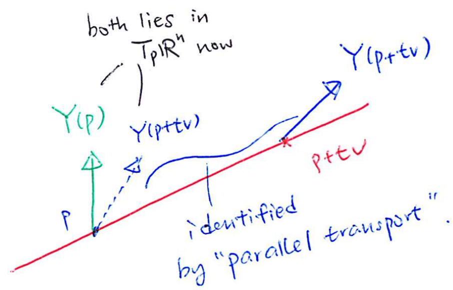
Note that we have applied Clairaut's theorem and switched the dummy indices. Therefore, we have $[X, Y] = (X^i(\partial_i Y^j) - Y^i(\partial_i X^j))\partial_j$, which is a vector field. □

Note that we can restate Clairaut's theorem as $[\partial_i, \partial_j] = 0$ for coordinate vector fields. Now, the next question we want to address is "*how do we differentiate vector fields*?" In $\mathbb{R}^n$, it is easy since any vector field Y can be written in components as $Y = (Y^1, \cdots, Y^n)$ where each Y^i is the component functions of Y. Therefore, we can define the derivative of Y in the direction given by a vector field X as

$$D_X Y := (X(Y^1), \cdots, X(Y^n)),$$

which makes sense since each Y^i is a function which can be act on by the vector field X as directional derivatives. Equivalently, $D_X Y$ is a vector field which at each point p is given by

$$(D_X Y)(p) = \lim_{t \to 0} \frac{Y(p + tX(p)) - Y(p)}{t}.$$



Note that the two vectors in the numerator in fact lie in different tangent spaces. The Euclidean structure (that we can translate vectors around) of $\mathbb{R}^n$ gives a canonical identification of each tangent space $T_p\mathbb{R}^n \cong \mathbb{R}^n$. The situation will be drastically different for surfaces since there is no canonical identification between tangent spaces at different points. We therefore need a more refined notion for differentiating (tangential) vector fields on a surface. We make one remark that the "differentiation" D is compatible with both the Euclidean metric $\langle \cdot, \cdot \rangle$ and the differentiable structure of $\mathbb{R}^n$:

Proposition 5. *For any vector fields X, Y, Z in $\mathbb{R}^n$, we have*

(i) $X\langle Y, Z\rangle = \langle D_X Y, Z\rangle + \langle Y, D_X Z\rangle$,

(ii) $D_X Y - D_Y X = [X, Y]$.

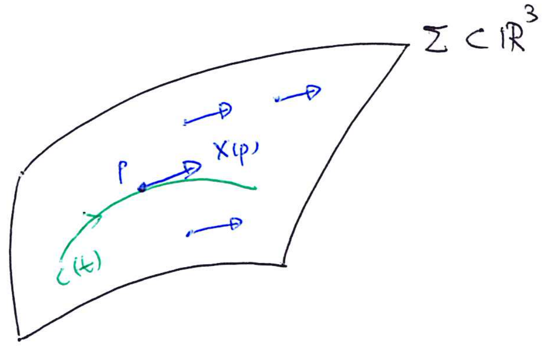
Proof. Direct from the definitions. □

Covariant derivatives

We now carry over our previous discussions to functions and (tangential) vector fields defined on a surface $\Sigma \subset \mathbb{R}^3$. Let $f \in C^\infty(\Sigma)$ be a smooth function on Σ and X, Y are tangential vector fields defined on Σ. We want to make sense of $X(f)$ and $D_X Y$ as in the case for $\mathbb{R}^n$. The first observation is that it is easy to define $X(f) \in C^\infty(\Sigma)$ as

$$X(f)(p) = D_{X(p)} f(p).$$

Although f is only defined on the surface Σ, and so is X, but since X is tangential, we can take a curve $c(t)$ *lying completely inside* Σ such that $c(0) = p$ and $c'(0) = X(p)$. By our previous discussion, it is sufficient to calculate the directional derivative $D_{X(p)} f(p)$. Therefore, the right hand side makes perfect sense and thus tangential vector fields X on Σ act to smooth functions f on Σ by taking directional derivatives as usual.



Now, we ask whether we can generalize the derivative $D_X Y$ in $\mathbb{R}^n$ to tangential vector fields X, Y defined on a surface. If we try to generalize directly as above, we immediately encounter a problem: we could still define $D_X Y$ be taking derivatives componentwise, the resulting vector, however, may no longer stay *tangential* to Σ anymore! In fact, it decomposes into two parts: a *tangential* component and a *normal* component using the orthogonal splitting $T_p\mathbb{R}^3 = T_p\Sigma \oplus N_p\Sigma$ with respect to the metric $\langle \cdot, \cdot \rangle$:

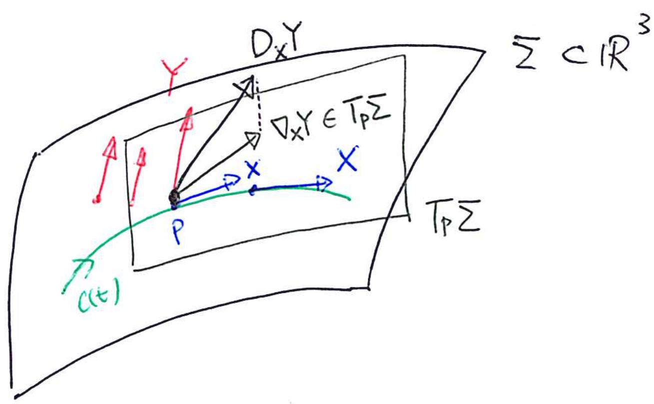
$$D_X Y = (D_X Y)^T + (D_X Y)^N \in T\Sigma \oplus N\Sigma.$$

The tangential part is which remains *intrinsic* on the surface Σ. Therefore, we define

Definition 6. *Let X, Y be tangential vector fields on a surface $\Sigma \subset \mathbb{R}^3$. The covariant derivative of Y along X is defined as*

$$\nabla_X Y := (D_X Y)^T.$$

We also call ∇ a connection.



The following properties can be checked easily from the definitions.

Proposition 7. *Let X, Y, Z be tangential vector fields on Σ.*

(i) $\nabla_X Y$ *is a tangential vector field.*

(ii) The covariant derivative is linear in each variable: for any constants $a, b \in \mathbb{R}$,

$$\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z,$$

$$\nabla_{aX+bY} Z = a\nabla_X Z + b\nabla_Y Z.$$

(iii) For any $f \in C^\infty(\Sigma)$, *we have*

$$\text{(Leibniz rule)} \qquad \nabla_X(fY) = X(f)\,Y + f\,\nabla_X Y,$$

$$\text{(Tensorial property)} \qquad \nabla_{fX} Y = f\,\nabla_X Y.$$

(iv) $(D_X Y)^N = -h(X, Y)\nu$ *where h is the second fundamental form with respect to the normal ν.*

Proof. (i) is by definition. (ii) and (iii) follows from the properties of directional derivatives. For (iv), note that

$$(D_X Y)^N = \langle D_X Y, \nu\rangle\nu = -\langle Y, D_X \nu\rangle\nu = -h(Y, X)\nu = -h(X, Y)\nu,$$

where we have used that $\langle Y, \nu\rangle = 0$ and the symmetry of the second fundamental form h. □

In summary, the directional derivative $D_X Y$ can be decomposed into an *intrinsic* part given by the covariant derivative $\nabla_X Y$ and an *extrinsic* part described by the second fundamental form of the surface Σ. Note that to calculate $\nabla_X Y(p)$ at a point $p \in \Sigma$, we just need to know the vector $X(p)$ and the vector field Y along some curve $c(t)$ with $c(0) = 0$ and $c'(0) = X(p)$. Similar to Proposition 5, we have the following properties for the covariant derivative.

Theorem 8. *The connection ∇ satisfies the following: for any tangential vector fields X, Y, Z on Σ, we have*

(i) (metric compatible) $X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$.

(ii) (torsion free) $\nabla_X Y - \nabla_Y X = [X, Y]$.

Any connection ∇ satisfying both (i) and (ii) is said to be a <u>Levi-Civita</u> or <u>Riemannian connection</u>.

Proof. Using Proposition 5 (i), we have

$$\begin{aligned} X(g(Y, Z)) &= X\langle Y, Z\rangle \\ &= \langle D_X Y, Z\rangle + \langle Y, D_X Z\rangle \\ &= \langle (D_X Y)^T + (D_X Y)^N, Z\rangle + \langle Y, (D_X Z)^T + (D_X Z)^N\rangle \\ &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \end{aligned}$$

Similarly, using Proposition 5 (ii) and Proposition 7 (iv), we have

$$[X,Y] = D_X Y - D_Y X = \nabla_X Y - \nabla_Y X + (-h(X,Y) + h(Y,X))\nu = \nabla_X Y - \nabla_Y X,$$

since $h(X,Y) = h(Y,X)$ by the symmetry of the second fundamental form. □

Christoffel symbols

In this section, we write down explicit how to calculate the covariant derivative $\nabla_X Y$ using a local coordinate system (i.e. a parametrization). Suppose we have coordinate system u^i on the surface and the associated coordinate vector fields are $\partial_i = \frac{\partial}{\partial u^i}$. Since $\nabla_{\partial_i}\partial_j$ is a tangential vector fields so we can express it as linear combinations of ∂_k:

$$\nabla_{\partial_i}\partial_j = \Gamma_{ij}^k \partial_k,$$

the coefficients Γ_{ij}^k are called the Christoffel symbols associated with this coordinate system.

We remark that the Christoffel symbols depend heavily on the chosen coordinate system. In fact Γ_{ij}^k is not a tensor (we can choose coordinates such that all of them vanishes at a given point). Therefore, the Christoffel symbols themselves do not carry any geometric meaning. It is describing the distortions of the coordinate system from the rectangular coordinate system in flat Euclidean space.

Theorem 9. *The Christoffel symbols are symmetric in the lower indices, i.e.* $\Gamma_{ij}^k = \Gamma_{ji}^k$*, and it can be calculated by the formula*

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell}(\partial_i g_{\ell j} + \partial_j g_{i\ell} - \partial_\ell g_{ij}).$$

Proof. The symmetry is an easy consequence of Clairaut's theorem $[\partial_i, \partial_j] = 0$ and that the connection ∇ is torsion free:

$$0 = [\partial_i, \partial_j] = \nabla_{\partial_i}\partial_j - \nabla_{\partial_j}\partial_i = (\Gamma_{ij}^k - \Gamma_{ji}^k)\partial_k.$$

To show that Γ_{ij}^k is computed by the given formula, first note that by metric compatibility of ∇:

$$g_{k\ell}\Gamma_{ij}^k = g(\nabla_{\partial_i}\partial_j, \partial_\ell) = \partial_i g_{j\ell} - g(\partial_j, \nabla_{\partial_i}\partial_\ell) = \partial_i g_{j\ell} - g(\partial_j, \nabla_{\partial_\ell}\partial_i).$$

where we have used $\nabla_{\partial_i}\partial_\ell = \nabla_{\partial_\ell}\partial_i$ from the symmetry of Christoffel symbols. Switch the role of i and j in the equation and add them together, using that $\Gamma_{ij}^k = \Gamma_{ji}^k$, we have

$$2g_{k\ell}\Gamma_{ij}^k = \partial_i g_{j\ell} + \partial_j g_{i\ell} - (g(\partial_j, \nabla_{\partial_\ell}\partial_i) + g(\partial_i, \nabla_{\partial_\ell}\partial_j)) = \partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij},$$

where we have used the metric compatibility again in the last equality. To get the desired formula, we multiply $\frac{1}{2}g^{p\ell}$ on both sides:

$$\Gamma_{ij}^p = \delta_k^p \Gamma_{ij}^k = g^{p\ell} g_{k\ell}\Gamma_{ij}^k = \frac{1}{2} g^{p\ell}(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}).$$

This is the formula we want by changing the name of the index p to k. □

As a trivial example, if we use the standard rectangular coordinates on the flat plane $\mathbb{R}^2$, then we have $\Gamma_{ij}^k \equiv 0$ since the (g_{ij}) is the constant identity matrix (so the right hand side in the formula vanishes upon differentiation). *Exercise: What about polar coordinates?*

Proposition 10. *Let $f : U \to \mathbb{R}^3$ be a parametrization of a regular parametrized surface. Then, we have*

$$\text{Gauss equation:} \qquad \partial_i \partial_j f = \Gamma_{ij}^k \partial_k f - h_{ij}\nu,$$

$$\text{Weingarten equation:} \qquad \partial_i \nu = g^{jk} h_{ij} \partial_k f.$$

Proof. Homework problem. □

Geodesics and parallel transport

Now, with the notion of covariant derivative, we can define the concept of *parallel transport* and *geodesics* on a surface $\Sigma \subset \mathbb{R}^3$.

Definition 11. *(i) A tangential vector field Y on Σ is said to be parallel if $\nabla_X Y \equiv 0$ for all tangential vector field X on Σ.*

(ii) Let Y be a tangential vector field defined on a curve $c(t)$ lying inside Σ. We say that Y is parallel along $c(t)$ if $\nabla_{c'(t)} Y \equiv 0$ for all t.

(iii) A curve $c(t)$ lying inside Σ is a geodesic if $\nabla_{c'(t)} c'(t) \equiv 0$ for all t. In other words, c' is parallel along c.

The following proposition follows easily from the definition.

Proposition 12. *Let X, Y be two tangential vector fields which are parallel along a curve $c(t)$ on Σ. Then, $g(X,Y) \equiv$constant (i.e. is independent of t).*

Proof. By metric compatibility, we have

$$\frac{d}{dt} g(X,Y) = g(\nabla_{c'(t)} X, Y) + g(X, \nabla_{c'(t)} Y).$$

However, both terms on the right hand side vanishes since X and Y are parallel along c. □

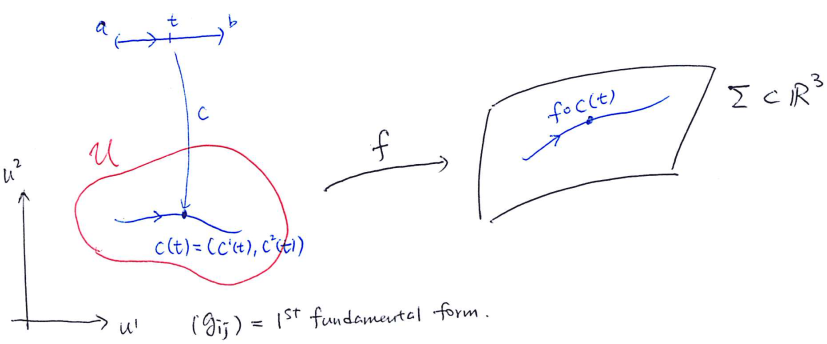
Note that previously we said that an arc-length parametrized curve $c(s)$ on Σ is a geodesic if the geodesic curvature $k_g := \|c''(s)^T\| \equiv 0$ for all s. This is in fact compatible with Definition 11 (iii) above in the following manner. Recall that one could identify regular parametrized curves up to orientation preserving reparametrizations, let $\mathcal{C} = [c]$ denote the equivalence class of a regular parametrized curve $c(t)$. In each equivalence class $\mathcal{C} = [c]$, there exists an (essentially unique) arc-length parametrized representative $\tilde{c}(s) \in \mathcal{C}$. Thus, we say that the equivalence class $\mathcal{C}$ is a geodesic if $\tilde{c}(s)$ is a geodesic in the sense that $k_g \equiv 0$ (see Definition 2). These two definitions are related in the following way.

Proposition 13. *If a regular parametrized curve $c(t)$ is a geodesic in the sense of Definition 11, then $\mathcal{C} = [c]$ is a geodesic in the sense that its arc-length parametrized representative $\tilde{c}(s)$ is a geodesic in the sense of Definition 2. Conversely, if $\mathcal{C}$ is a geodesic, then $\tilde{c}(s)$ is a geodesic in the sense of Definition 11.*

Proof. Homework problem. □

Note that a corollary of Proposition 12 is that if $c(t)$ is a geodesic in the sense of Definition 11, then $\|c'(t)\|$ is constant and thus t is up to affine transformation the arc length parameter s. Therefore, a "trajectory" $c(t)$ is a geodesic on Σ if (i) the image is a "straight line" as seen from the surface Σ, and (ii) the straight line is being traced out at constant speed. Physically, this is the trajectory of an object subject to no external force by Newton's first law of motion.

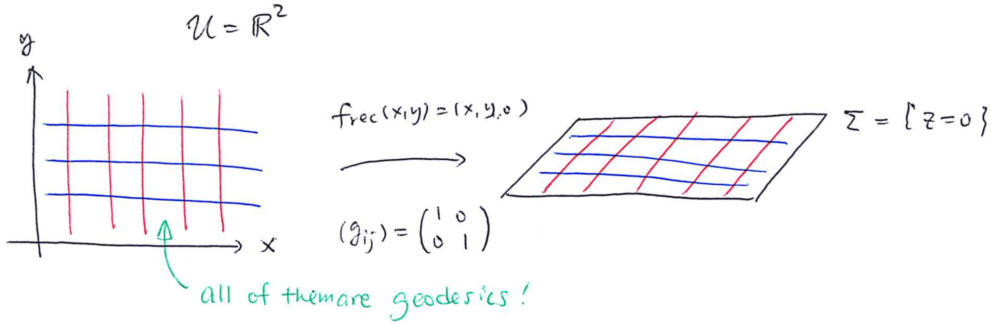
Let us now study the concept of parallel vector fields and geodesics using local coordinates (i.e. a parametrization $f : U \to \Sigma \subset \mathbb{R}^3$). Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface. If we have a parametrized curve $c(t) : (a, b) \to U \subset \mathbb{R}^2$, then $f \circ c(t)$ gives a curve on the surface Σ. The converse is also true locally, i.e. any curve on Σ locally can be expressed under a local coordinate system in this form. This is useful as instead of having a parametrized curve $c(t)$ lying on the surface $\Sigma \subset \mathbb{R}^3$, which needs three components to describe it (together with one constraint equation that it lies on the surface Σ), we only need two components to describe a curve in $U \subset \mathbb{R}^2$ (without any constraints!). From this point of view, we can understand the (local) behavior of geodesics on a surface $\Sigma \subset \mathbb{R}^3$ by looking at the corresponding curves in U under a parametrization (i.e. local coordinate system) $f : U \to \Sigma \subset \mathbb{R}^3$.



Let us study two simple examples. Let Σ be the xy-plane in $\mathbb{R}^3$. One can parametrize it by the rectangular coordinate system $f_{rec} : \mathbb{R}^2 \to \mathbb{R}^3$ as

$$f_{rec}(x, y) := (x, y, 0).$$

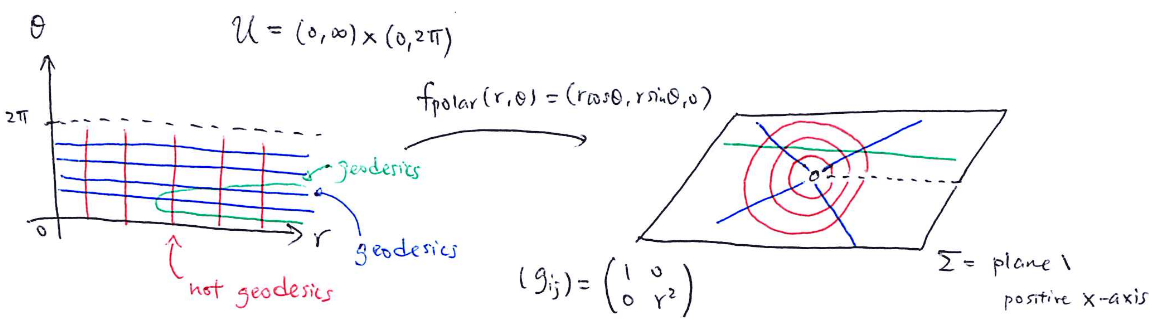
Under this coordinate system, the geodesics in Σ are just ordinary straight lines in the parameter space $U = \mathbb{R}^2$:



On the other hand, if we consider the polar coordinates instead, i.e. $f_{polar} : (0,\infty) \times (0,2\pi) \to \mathbb{R}^3$, where

$$f_{polar}(r,\theta) := (r\cos\theta, r\sin\theta, 0).$$

Then, under this coordinate system, the straight lines in $U = (0,\infty) \times (0,2\pi)$ are not necessarily geodesics (with respect to the metric described by the first fundamental form (g_{ij})):



In other words, what we learned from this example is that the geodesic would look different in the parameter space U if we use different parametrizations for even the same surface! What really makes the difference is that the first fundamental forms for these two coordinate systems are different

$$(g_{ij}^{rec}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (g_{ij}^{polar}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.$$

We will study this intrinsic way of looking at surfaces in the next section.

Let us turn to study the concept of parallel transport and geodesics using a local coordinate system. As we said, a curve $c(t)$ in the parameter space $U \subset \mathbb{R}^2$ can be represented in components as $c(t) = (c^1(t), c^2(t))$. A vector field Y along $c(t)$ can thus be expressed in local coordinates as $Y(t) = Y^i(t)\partial_i$.

Note that $c'(t) = \frac{dc^j}{dt}\partial_j$ and therefore

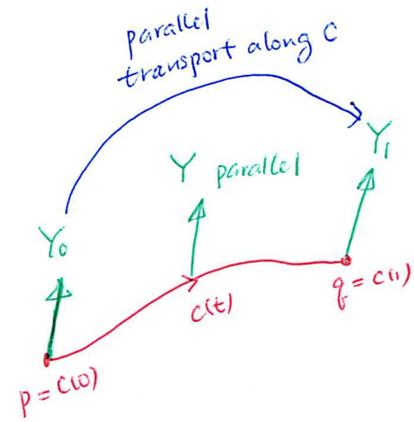
$$\nabla_{c'(t)}Y(t) = \nabla_{c'(t)}(Y^i(t)\partial_i) = \frac{dY^i}{dt}\partial_i + \Gamma^k_{ji}(c(t))\frac{dc^j}{dt}Y^i(t)\partial_k = \left(\frac{dY^k}{dt} + \Gamma^k_{ij}(c(t))\frac{dc^i}{dt}Y^j(t)\right)\partial_k.$$

Therefore, Y is parallel along $c(t)$ if and only if the components Y^k satisfy the linear system of first order ODE:

$$\frac{dY^k}{dt} + \Gamma^k_{ij}(c(t))\frac{dc^i}{dt}Y^j(t) = 0, \qquad k = 1, 2.$$

Note that the matrix coefficients $\Gamma^k_{ij}(c(t))\frac{dc^i}{dt}$ in the linear system depend on t and $c(t)$ but not on Y^k. Therefore, by the existence and uniqueness of linear first order ODEs, the system is uniquely solvable once an initial condition $Y(0)$ is given. Therefore, we have the following theorem.

Theorem 14. *Let $c : [0,1] \to \Sigma \subset \mathbb{R}^3$ be a regular curve on a surface Σ. Let $p = c(0)$ and $q = c(1)$. For each fixed $Y_0 \in T_p\Sigma$, there exists a unique vector field Y defined on c which is parallel along $c(t)$ and that $Y(c(0)) = Y_0$. For such a parallel vector field Y, the vector $Y(c(1)) \in T_q\Sigma$ is said to be the parallel transport of Y_0 from p to q along the curve c.*



For the special case of $Y = c'(t)$, we get the $c(t)$ is a geodesic if and only if its components in local coordinates satisfy the geodesic equations:

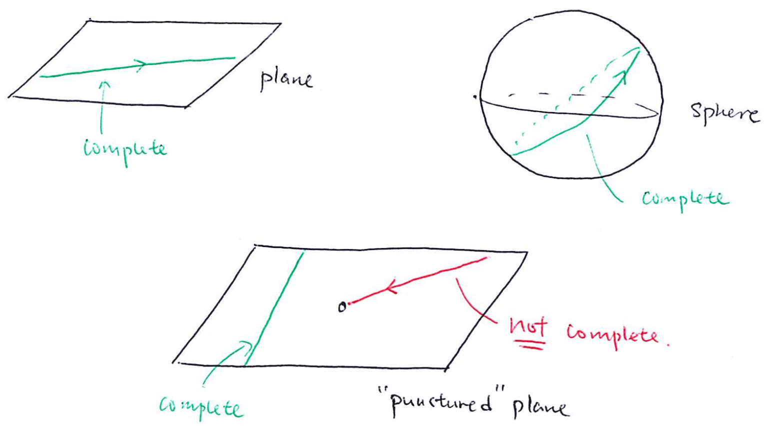
$$\frac{d^2c^k}{dt^2} + \Gamma^k_{ij}(c(t))\frac{dc^i}{dt}\frac{dc^j}{dt} = 0, \qquad k = 1, 2.$$

Note that the equations are quadratic in the first order terms involving $c'(t)$. Therefore, it is indeed a second order nonlinear ODE system. The general local existence theory for second ODE system implies the following:

Theorem 15. *For each fixed $p \in \Sigma$ and $v \in T_p\Sigma$, there exists some $\epsilon > 0$ and a unique geodesic $c : (-\epsilon, \epsilon) \to \Sigma \subset \mathbb{R}^3$ such that $c(0) = p$ and $c'(0) = v$. Moreover, the geodesic c depends smoothly on the initial conditions p and v.*

Note that the ϵ above depends on p and v as well, if ϵ can be taken as infinity, then we can that the geodesic is complete.

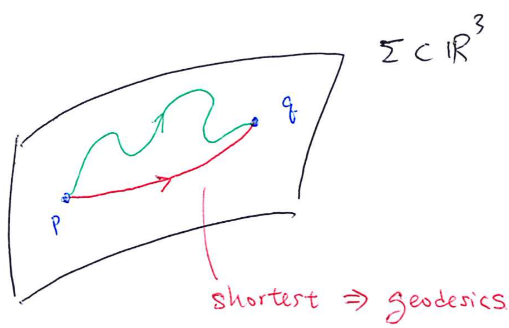
Definition 16. *A geodesic $c(t) : (a,b) \to \Sigma\mathbb{R}^3$ is said to be* <u>complete</u> *if it can be extended to a geodesic $c(t)$ such that t is defined for all real numbers.*



We like to think of geodesics as "straight lines" on the surface Σ. Therefore, it should share similar geometric properties that ordinary straight lines enjoy. For example, in a plane, any two distinct point can be connected by a unique straight line segment, and vice versa any curve of shortest length joining two points must be a straight line. For geodesics we have the following:

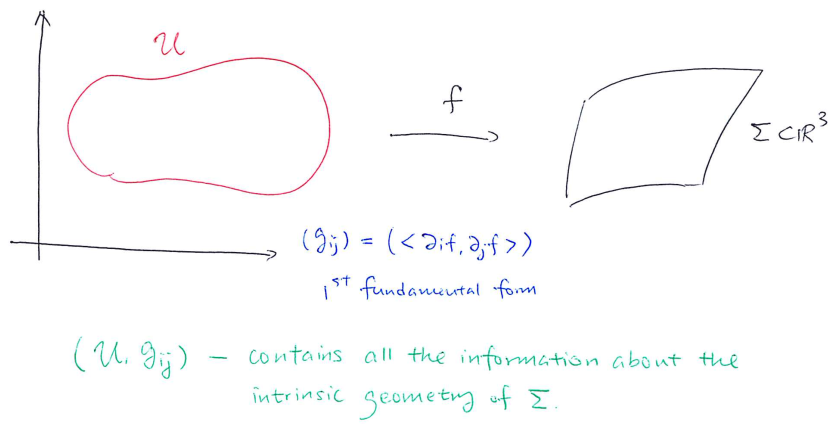
Proposition 17. *If $c(t) : [a,b] \to \Sigma \subset \mathbb{R}^3$ is a curve of shortest length joining two points p, q on Σ, i.e. $c(a) = p$ and $c(b) = q$, then $c(t) : (a,b) \to \Sigma$ is a geodesic after reparametrization.*

Proof. We will postpone the proof until later when we discuss first variation formula. $\square$



Abstract Riemannian surfaces

Given a regular parametrized surface $f : U \to \Sigma \subset \mathbb{R}^3$, we can define the first fundamental form (g_{ij}) which is a smooth family of 2×2 positive definite symmetric matrices on U. Anything that can be computed *only* with these (g_{ij}) are said to be intrinsic in the surface Σ.



Since (U, g_{ij}) is all we need to study the intrinsic geometry of a surface, it motivates the following definition.

Definition 18. *An abstract Riemannian surface is a pair (U, g_{ij}) where $U \subset \mathbb{R}^2$ is a connected open set and (g_{ij}) is a smooth family of 2×2 positive definite symmetric matrices defined on U.*

Recall the example of rectangular and polar coordinates on a (subset of) plane. In rectangular coordinates, we have an abstract Riemannian surface given by

$$U = \mathbb{R}^2, \qquad (g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

On the other hand, for polar coordinates, we have

$$U = (0, \infty) \times (0, 2\pi), \qquad (g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.$$

Even though the two abstract Riemannian surfaces look different, they are both describing the same surface (i.e. the flat plane) but just in different coordinates. Therefore, we would like to say that these two abstract Riemannian surfaces are "the same" (i.e. *isomorphic*). The following definition states when we regard two abstract Riemannian surfaces as "the same".

Definition 19. *(i) Two embedded submanifolds $\Sigma_1, \Sigma_2 \subset \mathbb{R}^3$ are* <u>*isometric*</u> *if there exists a differmorphism (i.e. a bijective smooth map with smooth inverse) $F : \Sigma_1 \to \Sigma_2$ such that*

$$\langle dF_p(X), dF_p(Y) \rangle = \langle X, Y \rangle, \qquad \text{for all } X, Y \in T_p\Sigma_1,\ p \in \Sigma_1.$$

(ii) Two abstract Riemannian surfaces (U, g_{ij}) and $(\tilde{U}, \tilde{g}_{ij})$ are isometric if there exists a diffeomorphism $\varphi : \tilde{U} \to U$ such that

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

The diffeomorphisms F and φ above are called isometries.

By Lemma 10 in Part 2 of the lecture notes, we know that (ii) is satisfied if we simply do a reparametrization of a regular parametrized surface. In other words, a change of coordinates do not affect the intrinsic geometry.

According to Definition 19, to establish that two surfaces are isometric, one needs to exhibit an explicit map which is an isometry. For example, show that the cylinder $\{x^2 + y^2 = 1\}$ is locally isometric to a piece of the flat plane. On the other hand, it is in principle much harder to show that two surfaces are *not* isometric, since we have to say that there is *no* isometry between them. We cannot check out of infinitely many maps that none is an isometry. We need some other ways instead to distinguish two surfaces if they are different. We need to come up with geometric quantities that remain unchanged under isometries. These are called isometric invariants. Since an isometry preserves distance, quantities like length and area should be preserved as well. These are examples of isometric invariants. Indeed, in the next section, we will see that the Gauss curvature K for a surface in $\mathbb{R}^3$ is also an isometric invariant. This is the famous Gauss' *Theorema Egregium.*

Gauss and Codazzi equations and Theorema Egregium

Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface. Recall that the Gauss map $\nu : U \to \mathbb{S}^2 \subset \mathbb{R}^3$ is well-defined and we have the first and second fundamental forms (g_{ij}) and (h_{ij}) associated with f. The Gauss curvature K is defined to be the determinant of the shape operator S or equivalently by $K = \det(h_{ij})/\det(g_{ij})$. From this it seems that K is an extrinsic quantity since it depends on the second fundamental form (h_{ij}) which describes the extrinsic geometry of the surface. Surprisingly, Gauss discovered that this is just a disguise and in fact K is an intrinsic quantity of the surface.

Theorem 20 (Gauss' Theorema Egregium)**.** *The Gauss curvature K is an intrinsic invariant, i.e. it depends only on the first fundamental form (g_{ij}) (and their derivatives).*

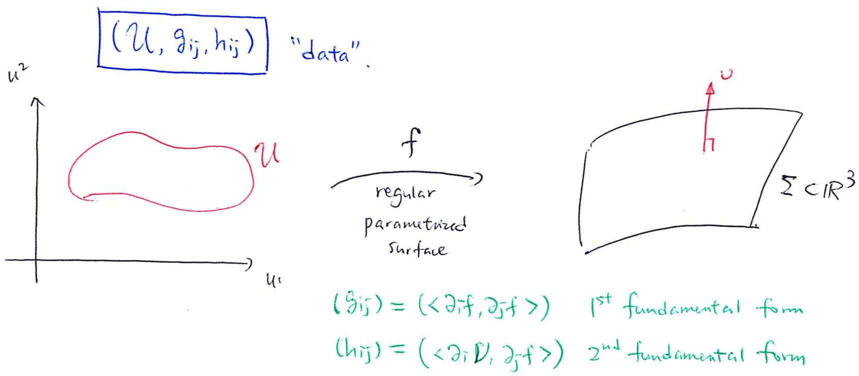
Since an isometry preserves the metric, i.e. the first fundamental form (g_{ij}), therefore, the Gauss curvature is an isometric invariant.

Corollary 21. *If $\varphi : (\tilde{U}, \tilde{g}_{ij}) \to (U, g_{ij})$ is an isometry as in Definition 19, then $K \circ \varphi = \tilde{K}$.*

Note that Gauss Theorema Egregium implies that the above corollary makes sense as (U, g_{ij}) is the only data we need to calculate K (the explicit formula is complicated though). In particular, Corollary 21 is effective in proving that two surfaces are *not* isometric. For example, the sphere $\mathbb{S}^2$ and the plane $\mathbb{R}^2$ can never be isometric (even locally) because $K_{\mathbb{S}^2} \equiv 1 \neq 0 \equiv K_{\mathbb{R}^2}$. On the other hand, the plane is locally isometric to the round cylinder since both of them has $K \equiv 0$. In fact, it is a theorem that two surfaces with the same constant Gauss curvature are necessarily locally isometric.

The proof of Gauss' Theorema Egregium comes from the "constraint equations" or "integrability conditions" for surfaces Σ in $\mathbb{R}^3$. For each regular parametrized surface $f : U \to \mathbb{R}^3$, we have the associated data (U, g_{ij}, h_{ij}) where (g_{ij}) and (h_{ij}) are the first and second fundamental forms associated to the parametrized surface $f : U \to \mathbb{R}^3$. Therefore, it is natural to ask for the reverse process:

Realization Problem: Given (U, g_{ij}, h_{ij}), can be find a regular parametrized surface $f : U \to \mathbb{R}^3$ such that (g_{ij}) and (h_{ij}) are exactly the first and second fundamental forms?



The answer turns out to be no in general. In contrast with the case of curve (recall the fundamental theorem for curves), the data (g_{ij}) and (h_{ij}) have to satisfy some compatibility conditions called *constraint equations* or *integrability conditions*. They are precisely the *Gauss* and *Codazzi equations*. It is a remarkable theorem that these necessary conditions turns out to be sufficient!

Theorem 22 (Bonnet)**.** *Let $U \subset \mathbb{R}^2$ be a connected open set, (g_{ij}) be a smooth family of 2×2 positive definite symmetric matrices defined on U and (h_{ij}) be a smooth family of 2×2 symmetric matrices defined on U. Suppose furthermore that the following integrability conditions are satisfied: for all $i, j, k, \ell = 1, 2$,*

$$(\textit{Gauss}) \qquad \partial_k \Gamma^\ell_{ij} - \partial_j \Gamma^\ell_{ik} + \Gamma^p_{ij}\Gamma^\ell_{pk} - \Gamma^p_{ik}\Gamma^\ell_{pj} = g^{\ell p}(h_{ij}h_{kp} - h_{ik}h_{jp}),$$

$$(\textit{Codazzi}) \qquad \partial_k h_{ij} - \partial_j h_{ik} + \Gamma^p_{ij} h_{pk} - \Gamma^p_{ik} h_{pj} = 0.$$

Then, there exists a regular parametrized surface $f : U \to \mathbb{R}^3$ with first and second fundamental forms (g_{ij}) and (h_{ij}) respectively. Moreover, f is unique up to rigid motions of $\mathbb{R}^3$ (i.e. a composition of rotations and translations).

Proof. The proof is beyond the scope of this class. Interested readers can refer to a proof in Kühnel (Theorem 4.24) $\square$

We now try to understand why the Gauss and Codazzi equations are necessary conditions for the existence of a regular parametrized surface f which realize the data (U, g_{ij}, h_{ij}). Note that Bonnet's theorem is in the same spirit as the fundamental theory of curves which says that once all the curvatures κ_i are prescribed, we can find a (unique up to rigid motions) curve in $\mathbb{R}^n$ which has all the curvatures as prescribed. However, there is *no* constraints that have to be satisfied by the prescribed curvatures κ_i. Why is there such a difference between the local theory of curves and surfaces?

This phenomenon can be understood from a calculus point of view. This is related to the existence of a potential function to a vector field. Consider a smooth vector field X in $\mathbb{R}^n$, we ask whether there exists a function f on $\mathbb{R}^n$ such that $X = \nabla f$. When $n = 1$ (which is analogous to the curve case), this is equivalent to finding the primitive function of a given function - which can *always* be done by integration! However, the situation gets more complicated already when $n = 2$. Let $X = (X_1, X_2)$ be the components of the vector field in $\mathbb{R}^2$. A potential function f for X is a function which satisfies the following system of partial differential equations:

$$\frac{\partial f}{\partial x} = X_1, \qquad \frac{\partial f}{\partial y} = X_2.$$

When does such an f exist? Note that first of all we should not expect this to be solvable for *all* X_1, X_2. Since mixed partial derivatives commute (as long as f is smooth enough), we must have

$$\frac{\partial X_1}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial X_2}{\partial x}.$$

Therefore, $\frac{\partial X_1}{\partial y} = \frac{\partial X_2}{\partial x}$ is a necessary condition for the solvability of f. From advanced calculus, we know that as long as the domain of definition of (X and f) under consideration is simply connected, then this necessary condition is also sufficient. The thing we learned from this simpler problem is that when dimension is at least two, the prescribed data (the vector field X in this case) has to satisfy some constraints as necessary conditions, and they are also sufficient conditions when the domain topology is simple enough.

We will now show that a very similar argument naturally leads us to the Gauss and Codazzi equations. Note that since a surface is *curved*, the constraint equations we get are *nonlinear* equations. Nonetheless, this is still a first order system of PDEs on our initial data g_{ij} and h_{ij}. (Recall that Γ_{ij}^k can be written as some combinations of the first derivatives of g_{ij}.)

Lemma 23. *Let $f : U \to \mathbb{R}^3$ be a regular parametrized surface in $\mathbb{R}^3$ with (g_{ij}) and (h_{ij}) be its first and second fundamental forms. Then, we have the following equations:*

$$(Gauss) \qquad \partial_k \Gamma_{ij}^\ell - \partial_j \Gamma_{ik}^\ell + \Gamma_{ij}^p \Gamma_{pk}^\ell - \Gamma_{ik}^p \Gamma_{pj}^\ell = h_{ij} h_k^\ell - h_{ik} h_j^\ell,$$

$$(Codazzi) \qquad \partial_k h_{ij} - \partial_j h_{ik} + \Gamma_{ij}^p h_{pk} - \Gamma_{ik}^p h_{pj} = 0,$$

where $\Gamma_{ij}^k := \frac{1}{2} g^{k\ell}(\partial_i g_{\ell j} + \partial_j g_{i\ell} - \partial_\ell g_{ij})$ are the Christoffel symbols and $h_i^j := g^{jk} h_{ik}$ (we say that we have raised the index j of h_{ij} using the metric g).

Proof. The two main ingredients of the proof is that (i) partial derivatives of functions in $\mathbb{R}^n$ commutes, and (ii) we can decompose any vector on the surface into its tangential and normal components. Recall that any regular parametrized surface $f : U \to \mathbb{R}^3$ satisfies the Gauss and Weingarten equations:

$$(Gauss) \qquad \partial_i \partial_j f = \Gamma_{ij}^k \partial_k f - h_{ij}\nu,$$

$$(Weingarten) \qquad \partial_i \nu = h_i^j \partial_j f.$$

Notice that the above equations are interpreted as taking partial derivatives of each of the components of f and ν. We can actually view Gauss and Weingarten equations as a "Frenet equations" for surfaces. To see this, note that $\{\partial_1 f, \partial_2 f, \nu\}$ together form a (not orthonormal!) moving basis of $\mathbb{R}^3$. The left hand side of Gauss and Weingarten equations are derivatives of this moving basis along tangential directions, and the right hand side simply express the derivatives in this moving basis. On the other hand, the situation is more complicated here as ordinary derivatives become partial derivatives and that there is no natural way to associate an *orthonormal* moving frame along a surface.

Since partial derivatives of functions commute in $\mathbb{R}^n$, we have two compatibility conditions derived from the Gauss and Weingarten equations respectively:

$$\partial_k(\partial_i \partial_j f) = \partial_j(\partial_i \partial_k f), \tag{1}$$

$$\partial_\ell \partial_i \nu = \partial_i \partial_\ell \nu. \tag{2}$$

Notice that (1) and (2) are vector equations and therefore the corresponding tangential and normal components must equal. We will see shortly that the tangential component of (1) will give the Gauss equation while the normal component will give the Codazzi equation. On the other hand, the tangential part of of (2) again gives the Codazzi equation and the normal component gives an equation which is always trivially satisfied. Therefore, all together the Gauss and Codazzi equations form a complete set of constraint equations.

Let us now compute the left hand side of (1). Note that the right hand side can be obtained similarly by just switching the indices j and k. Using Gauss and Weingarten equations, and decomposing everything in tangential and normal components, we have

$$\begin{aligned}
\partial_k(\partial_i \partial_j f) &= \partial_k(\Gamma_{ij}^\ell \partial_\ell f - h_{ij}\nu) \\
&= (\partial_k \Gamma_{ij}^\ell)\partial_\ell f + \Gamma_{ij}^\ell(\partial_k \partial_\ell f) - (\partial_k h_{ij})\nu - h_{ij}(\partial_k \nu) \\
&= (\partial_k \Gamma_{ij}^\ell)\partial_\ell f + \Gamma_{ij}^\ell(\Gamma_{k\ell}^p \partial_p f - h_{k\ell}\nu) - (\partial_k h_{ij})\nu - h_{ij}(h_k^\ell \partial_\ell f) \\
&= (\partial_k \Gamma_{ij}^\ell + \Gamma_{ij}^p \Gamma_{kp}^\ell - h_{ij} h_k^\ell)\partial_\ell f - (\partial_k h_{ij} + \Gamma_{ij}^p h_{kp})\nu,
\end{aligned}$$

where we have renamed some of the dummy indices in the last equality in order to help us group the terms together. By switching the indices j and k, we obtain the right hand side of (1) automatically,

$$\partial_j(\partial_i \partial_k f) = (\partial_j \Gamma_{ik}^\ell + \Gamma_{ik}^p \Gamma_{jp}^\ell - h_{ik} h_j^\ell)\partial_\ell f - (\partial_j h_{ik} + \Gamma_{ik}^p h_{jp})\nu.$$

By comparing the tangential and normal components separately, we get the Gauss and Codazzi equations respectively. (Recall that we have the symmetries $\Gamma_{ij}^k = \Gamma_{ji}^k$ and $h_{ij} = h_{ji}$.)

We can do a similar calculation for (2). The left hand side of (2) becomes

$$\begin{aligned}
\partial_\ell \partial_i \nu &= \partial_\ell (h_i^k \partial_k f) \\
&= (\partial_\ell h_i^k)\partial_k f + h_i^k (\partial_\ell \partial_k f) \\
&= (\partial_\ell h_i^k)\partial_k f + h_i^k (\Gamma_{\ell k}^p \partial_p f - h_{\ell k}\nu) \\
&= (\partial_\ell h_i^k + \Gamma_{\ell p}^k h_i^p)\partial_k f - (h_{\ell p} h_i^p)\nu,
\end{aligned}$$

again we have renamed some of the indices in the last equality to group similar terms together. The right hand side of (2) is easily obtained by switching i and ℓ:

$$\partial_i \partial_\ell \nu = (\partial_i h_\ell^k + \Gamma_{ip}^k h_\ell^p)\partial_k f - (h_{ip} h_\ell^p)\nu.$$

Let's first look at the normal component. By the definition of *raising an index*, we see that

$$h_{\ell p} h_i^p = h_{\ell p}(g^{pq} h_{iq}) = (g^{pq} h_{\ell p}) h_{iq} = h_\ell^q h_{iq} = h_{ip} h_\ell^p.$$

Therefore, the normal component has to match *automatically.* We now argue that the tangential component would give again the Codazzi equation. We will need to use the following identity, whose proof we postpone until the very end:

$$\partial_\ell g^{kq} + g^{pq}\Gamma_{\ell p}^k = -g^{kj}\Gamma_{j\ell}^q. \tag{3}$$

Let's us now assume (3) and use it to show that the tangential component of (2) gives the Codazzi equation. Using (3) and $h_i^j = g^{jk} h_{ik}$, the tangential component of $\partial_\ell \partial_i \nu$ is

$$\begin{aligned}
\partial_\ell h_i^k + \Gamma_{\ell p}^k h_i^p &= \partial_\ell (g^{kq} h_{qi}) + \Gamma_{\ell p}^k g^{pq} h_{qi} \\
&= h_{qi}(\partial_\ell g^{kq} + g^{pq}\Gamma_{\ell p}^k) + g^{kq}\partial_\ell h_{qi} \\
&= -g^{kj} h_{qi}\Gamma_{j\ell}^q + g^{kq}\partial_\ell h_{qi} \\
&= g^{kq}(-h_{ji}\Gamma_{q\ell}^j + \partial_\ell h_{qi}).
\end{aligned}$$

Switching i and ℓ, we obtain

$$g^{kq}(-h_{ji}\Gamma_{q\ell}^j + \partial_\ell h_{qi}) = g^{kq}(-h_{j\ell}\Gamma_{qi}^j + \partial_i h_{q\ell}).$$

Multiplying g_{kp} on both sides and using $g_{ij}g^{jk} = \delta_i^k$, we get

$$-h_{ji}\Gamma_{k\ell}^j + \partial_\ell h_{ki} = -h_{j\ell}\Gamma_{ki}^j + \partial_i h_{k\ell},$$

which is easily seen to be equivalent to the Codazzi equation.

It now remains to prove the identity (3). Since $g^{kq}g_{qi} = \delta_i^k$ holds identically *at every point*, we can differentiate this identity to get

$$0 = \partial_\ell (g^{kq} g_{qi}) = (\partial_\ell g^{kq}) g_{qi} + g^{kq}\partial_\ell g_{qi}.$$

Therefore, $\partial_\ell g^{kq} = -g^{qi}g^{kp}\partial_\ell g_{pi}$. Using the formula $\Gamma^k_{\ell j} = \frac{1}{2}g^{kp}(\partial_\ell g_{pj} + \partial_j g_{p\ell} - \partial_p g_{\ell j})$, the left hand side of (3) is

$$\begin{aligned}\partial_\ell g^{kq} + g^{pq}\Gamma^k_{\ell p} &= -g^{qi}g^{kp}\partial_\ell g_{pi} + \frac{1}{2}g^{jq}g^{kp}(\partial_\ell g_{pj} + \partial_j g_{p\ell} - \partial_p g_{\ell j}) \\ &= -g^{kp}\cdot\frac{1}{2}g^{jq}(\partial_\ell g_{pj} + \partial_p g_{\ell j} - \partial j g_{p\ell}) = -g^{kp}\Gamma^q_{p\ell},\end{aligned}$$

which is equivalent to (3). This finishes the proof of the lemma. □

We now observe that Gauss' Theorema Egregium is simply a special case of the Gauss equation.

Proof of Gauss' Theorema Egregium: First, we recall the formula $K = \det(h_{ij})/\det(g_{ij})$. Therefore, to show that K is an intrinsic invariant, we just need to show that $\det(h_{ij})$ is an intrinsic invariant, i.e. can be expressed solely in terms of g_{ij} and its derivatives. Note that Γ^k_{ij} can be expressed in terms of g_{ij} and its first derivatives. On the other hand, the Gauss equation implies that

$$g_{q\ell}(\partial_k\Gamma^q_{ij} - \partial_j\Gamma^q_{ik} + \Gamma^p_{ij}\Gamma^q_{pk} - \Gamma^p_{ik}\Gamma^q_{pj}) = h_{ij}h_{k\ell} - h_{ik}h_{j\ell}.$$

In particular, if we take $i = j = 1$ and $k = \ell = 2$, the right hand side is $\det(h_{ij})$ and the left hand side is completely determined by g_{ij} and its first and second derivatives. This proves the theorem.

Riemann curvature tensor

In the last section, we have derived the Gauss and Codazzi equations in terms of local coordinates induced by the parametrization $f : U \to \mathbb{R}^3$. One may wonder if there is a more coordinate-free way to get such equations. Now, recall that we have the covariant derivative ∇ intrinsically associated with a (abstract) surface. Out of this we can define the important concept of *Riemann curvature tensor*:

Definition 24. *The Riemann curvature tensor R is defined as follows: for any tangential vector field X, Y, Z on a (abstract) surface,*

$$R(X,Y)Z := \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z.$$

Note that the output $R(X,Y)Z$ is itself a tangential vector field. The Riemann curvature tensor satisfies a number of nice properties.

Proposition 25. *For any tangential vector fields X, Y, Z, W and any function f on the surface, we have*

(i) $R(X,Y)Z = -R(Y,X)Z$.

(ii) $g(R(X,Y)Z, W) = -g(R(X,Y)W, Z)$.

(iii) *R is "tensorial" in every slot, i.e.* $R(fX,Y)Z = R(X,fY)Z = R(X,Y)(fZ) = fR(X,Y)Z$.

Proof. Exercise for the reader. □

Note that the last property in the Proposition above is very important which makes the Riemann curvature tensor a "*tensor*". In particular, for each p on the surface, $R(X,Y)Z|_p$ only depends on $X|_p$ and $Y|_p$ and $Z|_p$. So it makes sense to write $R(X|_p, Y|_p)Z|_p$ as $R(X,Y)Z|_p$ is independent of how we extend the vector fields X, Y, Z to a neighborhood of p.

The Riemann curvature tensor is a fundamental geometric invariant which generalize the concept of Gauss curvature K to higher dimensions. From the definition it suggests that R measure the failure of the commutativity of the mixed second covariant derivatives of vector fields. Note that if we were to replace Z by a function f, then the expression would always vanish by the definition of the Lie bracket $[X,Y] := XY - YX$. Therefore, mixed partials on functions always commute but mixed covariant derivatives of vectors do not and in fact it reflects the intrinsic geometry of the surface.

We will now use the Riemann curvature tensor to re-derive the Gauss and Codazzi equations in covariant (i.e. coordinate-free) form. Let D be the usual derivative of vector fields in $\mathbb{R}^n$ (by componentwise taking partial derivatives). Then we have

$$D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z = 0 \tag{4}$$

for any vector fields X, Y, Z in $\mathbb{R}^n$. Note that this is just a covariant way of saying that mixed partial derivatives of functions (and hence vector fields in $\mathbb{R}^n$) commute. Now, if we have a surface $\Sigma \subset \mathbb{R}^3$ with global unit normal ν (we are only concerned with local properties here so we do not lose generality here), the shape operator (or Weingarten map) $S = D\nu$ is the linear operator on $T\Sigma$, which can be expressed equivalently by the second fundamental form $h(X,Y) = \langle S(X), Y \rangle$. The Gauss and Weingarten equations can be expressed covariantly as:

$$D_X Y = \nabla_X Y - h(X,Y)\nu,$$
$$D_X \nu = S(X),$$

for any tangential vector fields X, Y. Using (4) and the Gauss and Weingarten equations, we can derive the covariant form of Gauss and Codazzi equations:

Theorem 26. *For any tangential vector fields X, Y, Z, we have*

$$(Gauss) \qquad R(X,Y)Z = h(Y,Z)S(X) - h(X,Z)S(Y),$$
$$(Codazzi) \qquad \nabla_X(S(Y)) - \nabla_Y(S(X)) = S([X,Y]).$$

Proof. The proof is in principle the same as in the last section. Remember that the two ingredients needed are the commutativity of mixed partials (which is expressed by (4) now) and the Gauss and Weingarten equations. Using the Gauss and Weingarten equations, we obtain

$$\begin{aligned} D_X D_Y Z &= D_X(\nabla_Y Z - h(Y,Z)\nu) \\ &= \nabla_X \nabla_Y Z - h(X, \nabla_Y Z)\nu - D_X(h(Y,Z))\nu - h(Y,Z) D_X \nu \\ &= (\nabla_X \nabla_Y Z - h(Y,Z)S(X)) - (\langle S(X), \nabla_Y Z \rangle + D_X \langle S(Y), Z \rangle)\nu. \end{aligned}$$

By switching X and Y, we easily obtain

$$D_Y D_X Z = (\nabla_Y \nabla_X Z - h(X,Z)S(Y)) - (\langle S(Y), \nabla_X Z\rangle + D_Y \langle S(X), Z\rangle)\nu.$$

Moreover, by Gauss equation (note that $[X,Y]$ is a tangential vector field)

$$D_{[X,Y]} Z = \nabla_{[X,Y]} Z - h([X,Y],Z)\nu.$$

Therefore, combining all these and using (4) and the definition of Riemann curvature tensor, we have

$$\begin{aligned} 0 &= D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z \\ &= (R(X,Y)Z - h(Y,Z)S(X) + h(X,Z)S(Y)) \\ &\quad -(\langle S(X), \nabla_Y Z\rangle + D_X\langle S(Y), Z\rangle - \langle S(Y), \nabla_X Z\rangle - D_Y\langle S(X), Z\rangle - h([X,Y],Z))\nu. \end{aligned}$$

Note that the vanishing of the tangential component gives the Gauss equation. To see that the vanishing of the normal component gives the Codazzi equation, notice that by metric compatibility of ∇, we can simplify the normal component to the equation

$$\langle \nabla_Y (S(X)) - \nabla_X (S(Y)) + S([X,Y]), Z\rangle = 0.$$

Since the above equation holds for *any* Z, we get the Codazzi equation as desired. □

Why is the Riemann curvature tensor a generalization of the Gauss curvature K? Recall that R is tensorial in every slots, therefore, if we take an orthonormal basis $\{X, Y\}$ of any tangent plane $T_p\Sigma$ at $p \in \Sigma$, the Gauss equation says

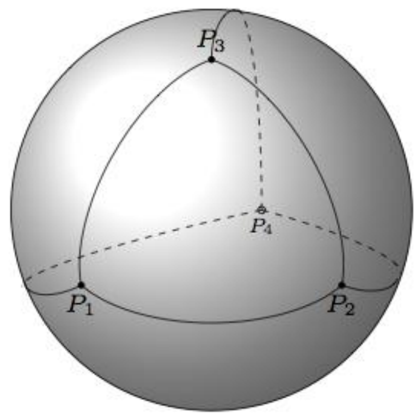
$$g(R(X,Y)Y, X) = h(Y,Y)h(X,X) - h(X,Y)^2.$$

Note that the right hand side is equal to $K(p)$ since X, Y forms an orthonormal basis at p. Therefore, $g(R(X,Y),Y,X)$ is equal to $K(p)$ when $\{X,Y\}$ forms an orthonormal basis at p. In general, if we consider Riemannian manifolds in higher dimension, we can take an two orthogonal unit tangent vectors X, Y inside a tangent space $T_p\Sigma$, the quantity $g(R(X,Y)Y,X)$ is called the sectional curvature of the place spanned by X and Y. This gives us a way to understand curvatures in higher dimensions through our understanding of curvatures of surfaces. This is important for General Relativity as we will be confronted with a $3 + 1 = 4$ dimensional spacetime whose curvature represents the gravity and other matter fields in our universe.

Gauss-Bonnet Theorems

We now turn to the global theory of surfaces and give the most important theorem in the theory - the *Gauss-Bonnet Theorem.* This is a formula which relates the geometry and the global topology of a surface. We first recall a fundamental topological invariant called the *Euler characteristic.* There are many equivalent definitions for this but we will choose the one using triangulation which is natural in the proof of the Gauss-Bonnet Theorem.

Recall that a triangulation of a surface Σ is a finite collection of closed "triangles" $\{T_i\}_i$ on the surface Σ such that $\Sigma = \cup_i T_i$ and that any two T_i's are either disjoint or intersect each other along a common "edge" or a common "vertex". An example of a triangulation of the sphere is shown below:



Definition 27. *The Euler characteristic of a surface Σ is defined as*

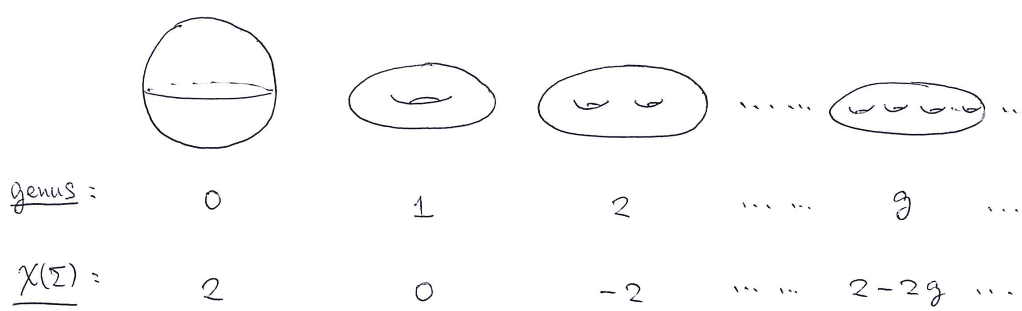
$$\chi(\Sigma) := v - e + f,$$

where v, e and f are the number of vertices, edges and faces respectively for a triangulation of Σ.

It is a fact that the expression $v - e + f$ is independent of the choice of triangulation on Σ and hence $\chi(\Sigma)$ is a well-defined topological invariant of Σ. For example, taking the triangulation of the sphere above, we have $v = 4$, $e = 6$ and $f = 4$ and hence $\chi(\mathbb{S}^2) = 4 - 6 + 4 = 2$.

There is a complete classification of all the closed (i.e. compact without boundary) orientable surfaces: their topological type is totally determined by the genus, which is just the "number of holes", of the surface. By putting in explicit triangulation, one can prove the following proposition.

Proposition 28. *For a closed orientable surface Σ, we have $\chi(\Sigma) = 2(1 - genus(\Sigma))$.*



We are now ready to state the global Gauss-Bonnet Theorem.

Theorem 29 (Global Gauss-Bonnet Theorem)**.** *Let $\Sigma \subset \mathbb{R}^3$ be a closed orientable surface. Then we have*

$$\int_\Sigma K\, dA = 2\pi\chi(\Sigma).$$

This theorem is remarkable since the left hand side is computed in terms of the Gauss curvature K which measures the local geometry of the surface while the right hand side is purely a topological invariant which is insensitive to the local geometry. Therefore, the Gauss-Bonnet Theorem provides a bridge between geometry and topology. Because of this, this formula has profound applications in geometry and topology. We also note that the global Gauss-Bonnet theorem still holds if Σ is a closed orientable abstract Riemannian surface (not necessarily inside $\mathbb{R}^3$) since everything in the formula is intrinsically defined. (Recall from Gauss' Theorema Egregium that K is an *intrinsic* quantity!)

As an illustration, let us consider the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ which has $K \equiv 1$. Therefore, the Gauss-Bonnet formula reads

$$\int_{\mathbb{S}^2} K\, dA = \text{Area}(\mathbb{S}^2) = 4\pi = 2\pi\chi(\mathbb{S}^2).$$

The Gauss-Bonnet theorem can be used to constrain the topology of a surface under certain curvature assumptions. For example, if $K > 0$ everywhere on a surface Σ, then the Gauss-Bonnet theorem implies that $\chi(\Sigma) > 0$, which by the classification theorem for surfaces, Σ must be homeomorphic to the sphere.

Theorem 30. *A closed orientable surface Σ with $K > 0$ everywhere must be homeomorphic to the sphere.*

Suppose $\Sigma \subset \mathbb{R}^3$ is a compact orientable surface *with boundary.* Then we still have a version of the Gauss-Bonnet theorem. Note that the left hand side would have an extra term involving the curvature of the boundary $\partial\Sigma$.

Theorem 31 (Local Gauss-Bonnet Theorem)**.** *Let $\Sigma \subset \mathbb{R}^3$ be a compact oriented surface with smooth boundary $\partial\Sigma$ such that Σ is homeomorphic to a disk. Then we have*

$$\int_\Sigma K\, dA + \int_{\partial\Sigma} k_g\, ds = 2\pi,$$

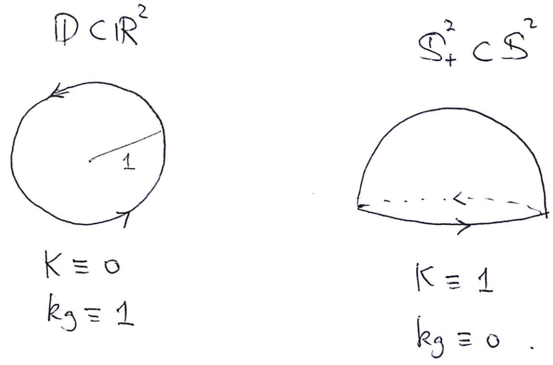
where $k_g := \langle \nabla_{e_1} e_1, e_2 \rangle$ is the geodesic curvature of $\partial\Sigma$ defined as follows: let ν be the global unit normal on Σ giving its orientation, e_1 is the unit tangent vector along $\partial\Sigma$ such that $e_2 := \nu \times e_1$ is the inward unit normal of $\partial\Sigma$ with respect to Σ.

For example, if Σ is the flat unit disk $\mathbb{D}$ of radius 1, then $K = 0$ and $k_g = 1$ and hence

$$\int_{\mathbb{D}} K\, dA + \int_{\partial\mathbb{D}} k_g\, ds = 0 + \text{length}(\partial\mathbb{D}) = 2\pi.$$

On the other hand, if Σ is the upper hemisphere $\mathbb{S}^2_+$ of the unit sphere $\mathbb{S}^2$, then $K = 1$ and $k_g = 0$ and hence

$$\int_{\mathbb{S}^2_+} K\, dA + \int_{\partial \mathbb{S}^2_+} k_g\, ds = \text{area}(\mathbb{S}^2_+) = 2\pi.$$

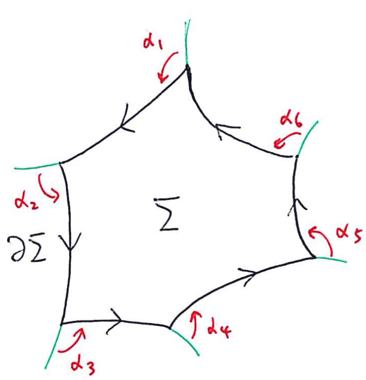


In fact, we have a version of the local Gauss-Bonnet theorem when the boundary $\partial\Sigma$ is only piecewise smooth.

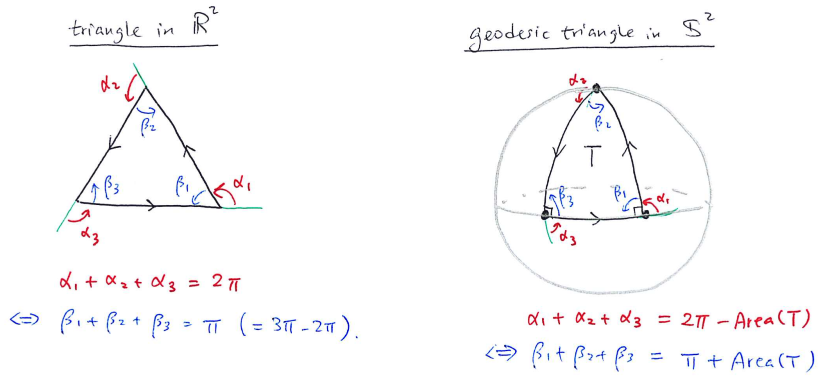
Theorem 32 (Local Gauss-Bonnet Theorem - piecewise smooth version)**.** *Let $\Sigma \subset \mathbb{R}^3$ be a compact oriented surface with piecewise smooth boundary $\partial\Sigma$ such that Σ is homeomorphic to a disk. Then we have*

$$\int_{\Sigma} K\, dA + \int_{\partial\Sigma} k_g\, ds + \sum_j \alpha_j = 2\pi,$$

where k_g is the geodesic curvature of $\partial\Sigma$ as defined in Theorem 31 and that α_j is the exterior angle at the j-th vertex.



For example, if we consider Σ be a flat triangle inside $\mathbb{R}^2$ with straight lines as boundaries, then $K = k_g = 0$ and that the Gauss-Bonnet theorem simply says that the total sum of exterior angles is equal to 2π. Equivalently, the total interior angle sum of a triangle in $\mathbb{R}^2$ is equal to π.



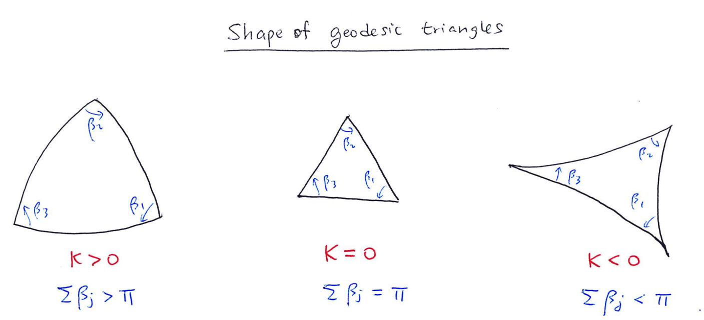
On the other hand, if we consider a geodesic triangle $T \subset \mathbb{S}^2$ by taking one of the vertices at the north pole and the other two on the equator, connected by geodesics (i.e. arcs of great circles), then by Gauss-Bonnet we have

$$\int_T K\,dA + \int_{\partial T} k_g\,ds + \sum_j \alpha_j = \text{area}(T) + \sum_{j=1}^{3} \alpha_j = 2\pi.$$

Equivalently, the interior angle sum is

$$\sum_{j=1}^{3} \beta_j = \sum_{j=1}^{3} (\pi - \alpha_j) = 3\pi - \sum_{j=1}^{3} \alpha_j = \pi + \text{area}(T) > \pi.$$

Therefore, the total interior angle sum is greater than π! In fact, positive Gauss curvature makes the geodesic triangles "fatter" and thus have bigger interior angle sum, while negative Gauss curvature makes the geodesic triangles "slimmer" with interior angle sum less than π.



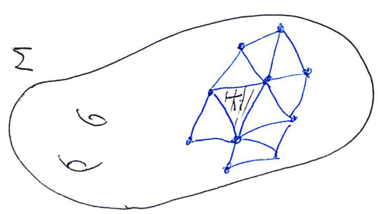
We now proceed to prove the Gauss-Bonnet Theorems. First of all, by an approximation argument, the piecewise smooth version of the local Gauss-Bonnet theorem follows from the smooth version. We will now show that the piecewise smooth local Gauss-Bonnet implies the global Gauss-Bonnet.

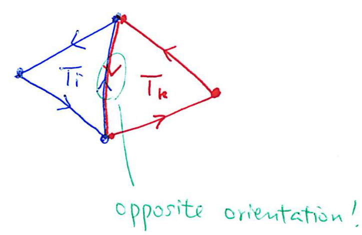
Let $\Sigma \subset \mathbb{R}^3$ be a closed orientable surface. Let $\{T_i\}_i$ be a triangulation on Σ with v vertices, e edges and f faces. On each T_i, we can apply the piecewise smooth version of the local Gauss-Bonnet theorem to get

$$\int_{T_i} K\,dA + \int_{\partial T_i} k_g\,ds + \sum_{j=1}^{3} \alpha_j^i = 2\pi,$$

where $\alpha_1^i, \alpha_2^i, \alpha_3^i$ are the exterior angles of the triangle T_i. If we let $\beta_j^i := \pi - \alpha_j^i$ be the corresponding interior angles of T_i, we have

$$\int_{T_i} K\,dA + \int_{\partial T_i} k_g\,ds + 3\pi - \sum_{j=1}^{3} \beta_j^i = 2\pi.$$





If we sum over all the T_i's, the first term becomes the total curvature on Σ (since the interior of T_i's are pairwise disjoint and that there union is the whole surface), the second term all cancels out since each edge is shared by two adjacent triangles which must induce different orientation on the edge, and the last term on the left hand side become the total sum of all the interior angles of all the T_i's hence is equal to $2\pi v$ since each vertex would contribute 2π. Therefore, we have

$$\int_{\Sigma} K\,dA + 3\pi f - 2\pi v = 2\pi f.$$

On the other hand, since each edge is shared by exactly two T_i's, therefore the total number of edges $e = 3f/2$, i.e. $3f = 2e$. Using this we have

$$\int_{\Sigma} K\,dA = 2\pi(v - e + f) = 2\pi\chi(\Sigma),$$

which is the desired Gauss-Bonnet formula. Therefore, it remain to prove the smooth version of the local Gauss-Bonnet theorem, which we will do later after we introduce the concept of differential forms. At last, we point out that the local Gauss-Bonnet theorem actually has a version for compact surfaces which are not necessarily homeomorphic to a disk.

Theorem 33 (General Gauss-Bonnet Theorem with boundary)**.** *Let Σ be a compact orientable surface with smooth boundary $\partial\Sigma$. Then we have*

$$\int_\Sigma K\,dA + \int_{\partial\Sigma} k_g\,ds = 2\pi\chi(\Sigma).$$

Note that $\chi(\Sigma)$ can be defined for surfaces with boundary exactly the same way through any triangulation, but now we allow some of the edges to lie completely inside $\partial\Sigma$. In fact, we have the formula

$$\chi(\Sigma) = 2 - 2g - k,$$

where $g = \text{genus}(\Sigma)$ and k is the number of boundary curves of $\partial\Sigma$. In particular, when Σ is homeomorphic to a disk, then $g = 0$ and $k = 1$ and thus $\chi(\Sigma) = 2 - 1 = 1$, which gives Theorem 31. More examples and applications will be discussed in the tutorial (see Tutorial notes).