1 Curvature of Surfaces

1.1 Fundamental theorem for self adjoint linear operators

Theorem 1. *Let V be a two-dimensional vector space with inner product $\langle , \rangle$. Let $A : V \to V$ be a self-adjoint linear map. Then V admits an orthonormal basis consisting of eigenvectors of A. That is, there exists an orthonormal basis $\{e_1, e_2\}$ of V and real numbers $\lambda_1 \geq \lambda_2$ such that $A(e_1) = \lambda_1 e_1, A(e_2) = \lambda_2 e_2$ (i.e. λ_1, λ_2 are the corresponding eigenvalues of A). Moreover the eigenvalues are given by*

$$\lambda_1 = \max_{|v|=1} \langle A(v), v\rangle = \max_{|v|=1} B(v, v)$$
$$\lambda_2 = \min_{|v|=1} \langle A(v), v\rangle = \min_{|v|=1} B(v, v)$$

where $B(v, w) = \langle A(v), w\rangle$ is the bilinear, symmetric form associated to the self adjoint linear map A.

Remarks

- We apply this theorem to the case where $V = T_pS$ is the two-dimensional tangent space at $p \in S$, $A = L$ is the self-adjoint map on T_pS, and $\Pi(v, v)$ is the second fundamental form. Then the eigenvalues λ_1, λ_2 are the principal curvatures κ_1, κ_2.
- In the basis e_1, e_2 , A is a diagonal matrix with entries λ_1, λ_2. In our applications, we will work with a basis of T_pS that arises from a local chart $\mathbf{X}_1$, $\mathbf{X}_2$. Hence the matrix representing L might not appear diagonal; however, of course, the eigenvalues κ_i are basis-independent.

Proof

The proof follows the argument given in Do Carmo's text. The first part involves the following

Lemma 1. *Consider the function $f(x, y) = ax^2 + 2bxy + cy^2$, restricted to the circle $x^2 + y^2 = 1$. If $f(x, y)$ has a maximum at $(1, 0)$, then $b = 0$.*

Proof We can parametrize points on the circle by $(x, y) = (\cos t, \sin t)$. On the circle $f(t) \equiv f(x(t), y(t)) = a \cos^2 t + 2b \cos t \sin t + c \sin^2 t$. The point $(1, 0)$ corresponds to $t = 0$. Requiring $f'(0) = 0$ implies $b = 0$.

SInce V is a 2d vector space, we can think of elements $v \in V$ as points in $\mathbb{R}^2$ with coordinates (v^1, v^2) where v^1, v^2 are constants (the components of the vector). A symmetric bilinear form $B : V \times V \to \mathbb{R}$ can be thought of as a continuous function of the form $B(v^1, v^2) = a(v^1)^2 + 2bv^1v^2 + c(v^2)^2$, i.e. of the same form as $f(v^1, v^2)$ considered above.

Lemma 2. *Given the symmetric bilinear form B there is an orthonormal basis $\{e_1, e_2\}$ of V such that if we write a general element $v \in V$ as $v = v^1e_1 + v^2e_2$, then*

$$B(v, v) = \lambda_1(v^1)^2 + \lambda_2(v^2)^2$$

where λ_1 and λ_2 are the maximum and minimum respectively of $B(v, v)$ restricted to unit vectors v, i.e. $(v^1)^2 + (v^2)^2 = 1$.

Proof $B(v,v)$ restricted to unit-length vectors $|v|=1$ is a continuous function on a closed and bounded set. Hence from calculus we know it must achieve its maximum and minimum somewhere on this set. We call this maximum value λ_1 and the associated vector $v_{max}=(v^1_{max},v^2_{max})$. Define the first basis vector $e_1=v_{max}$. Then $B(e_1,e_1)=\lambda_1$. By standard methods (Gram-Schmidt orthogonalization) we can find a second linearly independent unit vector that is orthogonal to e_1 and we call it e_2. Hence $\{e_1,e_2\}$ forms an orthonormal basis. Now set $\lambda_2=B(e_2,e_2)$ and take a general $v=v^1e_1+v^2e_2$. Since $|v|^2=1$ we have $(v^1)^2+(v^2)^2=1$. Then by the bilinearity property of B,

$$\begin{aligned} B(v,v) &= B(v^1e_1+v^2e_2, v^1e_1+v^2e^2) \\ &= (v^1)^2B(e_1,e_2)+v^1v^2B(e_1,e_2)+v^2v^1B(e_2,e_1)+(v^2)^2B(e_2,e_2) \\ &= \lambda_1(v^1)^2+2v^1v^2B(e_1,e_2)+\lambda_2(v^2)^2 \end{aligned}$$

But we know that B has a maximum at e_1, which corresponds to $(v^1,v^2)=(1,0)$. So we are in the situation of the first Lemma with $b=B(e_1,e_2)$ and hence we must have $b=0$. So now we have

$$B(v,v)=\lambda_1(v^1)^2+\lambda_2(v^2)^2\geq\lambda_2((v^1)^2+(v^2)^2)=\lambda_2$$

since $|v|^2=1$ and $\lambda_2\leq\lambda_1$ and thus λ_2 is the minimum of $B(v,v)$. Hence we have established the lemma.

With these two Lemmas we can establish the theorem. We are given a self adjoint linear map A in V. We set $B(v,v)=\langle A(v),v\rangle$. Now by Lemma 2, we know there is an orthonormal basis $\{e_1,e_2\}$ of V with $B(e_1,e_1)=\lambda_1$ and $B(e_2,e_2)=\lambda_2\leq\lambda_1$ where λ_1,λ_2 are the maximum and minimum of $B(v,v)$ restricted to unit vectors, i.e. with $|v|^1=\langle v,v\rangle$. We have to show that e_1,e_2 are eigenvalues of A. Now we know $B(e_1,e_2)=0=\langle A(e_1),e_2\rangle$. If we expand in this basis $A(e_1)=A^1{}_1e_1+A^2{}_1e_2$, then

$$\langle A(e_1),e_2\rangle=A^2{}_1=0$$

since $\langle e_1,e_2\rangle=0$. Hence we gather $A(e_1)=A^1{}_1e_1$. Thus $\lambda_1=B(e_1,e_1)=\langle A(e_1),e_1\rangle=A^1{}_1$. (Note that we could have $\lambda_1=0$). We have thus shown that

$$A(e_1)=\lambda_1e_1$$

i.e. e_1 is a eigenvector of A with eigenvalue λ_1. Finally, since $B(e_1,e_2)=B(e_2,e_1)=\langle A(e_2),e_1\rangle=0$ and $\lambda_2=B(e_2,e_2)=\langle A(e_2),e_2\rangle$ we can run through the same argument as above to show that

$$A(e_2)=\lambda_2e_2$$

which thus establishes the Theorem.

1.2 Application to Curvature of Surfaces

As stated in the Remarks, we apply this result to self adjoint map L and its associated bilinear symmetric form $\Pi: T_pS\times T_pS\to\mathbb{R}$. The theorem then tells us that the *principal curvatures* are precisely the eigenvalues of L. The associated eigenvectors are called *principal directions*. Note that

$$\kappa_1=\kappa_1\langle e_1,e_1\rangle=\langle L(e_1),e_1\rangle=\Pi(e_1,e_1) \tag{1}$$

so that κ_1 is the *normal curvature* associated to the unit vector $e_1 \in T_pS$. A similar argument shows κ_2 is the normal curvature of the unit vector e_2. Now writing $L(e_j) = L^i{}_j e_i$ in this special basis , we find that L takes the form (the superscript counts the rows, and the lower subscript measures columns)

$$[L]^i{}_j = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}$$

The Gaussian curvature is then defined to be $K_g \equiv \det L = \kappa_1\kappa_2$. Note that again the determinant of a matrix is basis-independent, so obviously we can compute the determinant in any basis, in particular the canonical basis of T_pS associated to the chart $\mathbf{X} : U \to S$. Similarly $H = \mathrm{Tr}[L] = \kappa_1 + \kappa_2$ is called the mean curvature of S (at the point p).

Note that $L = -\mathrm{d}\,\mathbf{N}$. Hence if we switch our normal $\mathbf{N}$ by a sign (this corresponds to choosing either the 'outward' or 'inward' normal.), then $L \to -L$. This results in switching the signs of the eigenvalues, so $H \to -H$, but $K_g \to (-\kappa_1)(-\kappa_2) = K_g$ and so the Gaussian curvature remains invariant. This suggests that K_g is an intrinsic property of the surface (i.e. it does not depend on how the surface is embedded in $\mathbb{R}^3$.) This is indeed the case as shown by Gauss, and allows us to define curvature of Riemannian manifolds (without any reference to an ambient $\mathbb{R}^n$).

Example 1 Here we consider the saddle shaped surface $z = y^2 - x^2$ and compute the Gaussian curvature and mean curvature at the point $p = (0, 0, 0)$ by computing the eigenvalues of the Weingarten map (shape operator). Now the unit normal is can be found from considering it as a surface $f(x, y, z) = z - y^2 + x^2 = 0$. Hence

$$\mathbf{N} = \frac{(2x, -2y, 1)}{\sqrt{1 + 4x^2 + 4y^2}}$$

is the Gauss map (the unit normal vector at an arbitrary point (x, y, z). Note at p, $\mathbf{N}_p = (0, 0, 1)$. We could 'guess' the eigenvectors but let us take a systematic approach. Let $\mathbf{r}(t) = (x(t), y(t), y(t)^2 - x(t)^2)$ be an arbitrary differentiable curve on the surface such that $\mathbf{r}(0) = (0, 0, 0) = p$, i.e. $x(0) = y(0) = 0$. Now we have $\mathbf{r}'(t) = (x'(t), y'(t), 2y(t)y'(t) - 2x(t)x'(t))$, and so $\mathbf{r}'(0) = (x'(0), y'(0), 0)$. Now we compute

$$L(\mathbf{r}'(0)) = -\frac{\mathrm{d}}{\mathrm{d}t} N(\mathbf{r}(t))\bigg|_{t=0} = 2(-x'(0), y'(0), 0)$$

where we used $x(0) = y(0) = 0$ to simplify after taking the derivative. Now if we want to find the eigenvectors, we wish to solve

$$L(\mathbf{r}'(0)) = L((x'(0), y'(0), 0)) = 2(-x'(0), y'(0), 0) = \kappa(x'(0), y'(0), 0) = \kappa\,\mathbf{r}'(0)$$

One can easily check that the choice $e_1 = (0, 1, 0)$ (i.e. $x'(0) = 0, y'(0) = 1$) has eigenvalue $\kappa_1 = 2$ and $e_2 = (1, 0, 0)$ has eigenvalue $\kappa_2 = -2$. That is $L(e_1) = 2e_1, L(e_2) = -2e_2$. Note that e_1, e_2 form an orthonormal basis. We read off $K_g = \kappa_1\kappa_2 = -4$ and $H = 0$ at $p = (0, 0, 0)$.

Example 2 The sphere $x^2 + y^2 + z^2 = 1$ with outward unit normal $\mathbf{N} = (x, y, z)$. By symmetry it does not make a difference what specific point on S^2 we choose as all are clearly equivalent. Consider a general curve on the surface with $\mathbf{r}(t) = (x(), y(t), z(t))$. Then $\mathbf{r}'(t) = (x'(t), y'(t), z'(t))$ and it follows from the fact that $x(t)^2 + y(t)^2 + z(t)^2 = 1$ that $\mathbf{r}'(t) \cdot \mathbf{N} = 0$ as vectors in $\mathbb{R}^3$. This of course simply tells us that $\vec{\prime}(t)$ is tangent to the sphere at each t, i.e. $\mathbf{r}(t) \in T_{\mathbf{r}(t)}S^2$. Now it is easy to see that

$$L(\mathbf{r}'(t)) = -\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{N}(\mathbf{r}(t)) = -\mathbf{r}'(t)$$

since $\mathbf{N}(\mathbf{r}(t)) = \mathbf{r}(t)$. Hence *any* vector v is an eigenvector of L with eigenvalue -1; so L is just (-1) times the identity map. It follows that $\Pi(v,v) = \langle L(v), v\rangle = -|v|^2$ and so restricted to unit vectors, it gives -1. We conclude in either case that $\kappa_1 = \kappa_2 = -1$. Hence $K_g = 1, H = -2$ are the Gaussian and mean curvatures respectively. Positive Gaussian curvature on a sphere implies curves on the surface both bend in the same direction with respect to the normal (both 'away' or 'towards' depending on whether one chooses an outward or inward-pointing normal).

Example 3 The cylinder $x^2 + y^2 = a^2$ with unit normal $\mathbf{N} = \frac{1}{a}(x, y, 0)$. This example was covered in lectures. If we consider a general curve on the cylinder $\mathbf{r}(t) = (x(t), y(t), z(t))$ with tangent $\mathbf{r}'(t) = (x'(t), y'(t), z'(t))$ we see that L has eigenvalues $\kappa_1 = 0, \kappa_2 = -1/a$ with corresponding eigenvectors $e_1 = (0, 0, z'(t))$ and $e_2 = (x'(t), y'(t), 0)$ (we have not normalized them to have unit length, but this can be done). Note that $e_2 \cdot e_1 = 0$, i.e. the two eigenvectors, or principal directions, are orthogonal since they correspond to different eigenvalues. Also note that in e_2, the functions $x'(t), y'(t)$ are not independent, since of course $x(t)^2 + y(t)^2 = a^2$ so $x(t)x'(t) + y(t)y'(t) = 0$. We conclude $K_g = 0$ and $H = -1/a$. The Gaussian curvature of the cylinder and flat plane both vanish; we will see they are both intrinsically 'flat' (that is, their geometry is indistinguishable locally from the point of view of someone restricted to moving only on the surface). However, unlike in the plane case, $L \neq 0$ for a cylinder. These two principal directions are distinguished in the sense that one direction does not 'bend' with respect the normal (e_1, which is parallel to the $z-$axis), while the other bends away from it (e_2).

1.3 Classification of points on the surface by Gaussian curvature

In the examples of the plane, sphere and cylinder, K_g is actually a constant, that is it does not vary form point to point on these surfaces. On a general surface as we will see shortly, K_g will be some complicated function on the surface. However we can visualize the geometry near a point by comparing it to one of the 'model' cases with the same behaviour of L near that point.

Let p be a point on a regular surface S with Gaussian curvature $K_g = \det L = \kappa_1\kappa_2$. Then we have the following classification:

1. *Elliptic points* if $K_g > 0$ (so that κ_1, κ_2 have the same sign) the normal sections of the principal directions bend in the same direction 'away' or 'towards' the normal. Equivalently the normals $\mathbf{n}$ of curves passing through p point towards the same side of T_pS. The surface is then 'bowl-shaped' near p. Typical examples are the sphere of radius R with $\kappa_1 = \kappa_2 = \pm 1/R$ so $K_g = 1/R^2 > 0$ at all $p \in S^2$, and the elliptic paraboloid $z = x^2 + y^2$ at $p = (0, 0, 0)$.

2. *Hyperbolic points* if $K_g < 0$ (so that κ_1, κ_2 have opposite signs) then the normal sections of the principal directions bend in opposite directions (i.e. one towards, and the other away, from the normal $\mathbf{N}$ at p). The surface is then 'saddle shaped' near p. A typical example we have considered is the hyperbolic paraboloid $z = y^2 - x^2$ at $p = (0, 0, 0)$.

3. *Parabolic points* if $K_g = 0$ but $L = -\mathrm{d}\,\mathbf{N} \neq 0$ at p. The example we have seen is the cylinder, which as this property at all p. Hence one of κ_1, κ_2 vanishes indicating there is a 'flat' principal direction, such as curves parallel to the $z-$ axis on the cylinder $x^2 + y^2 = a^2$.

4. *Planar points* $L = -\mathrm{d}N = 0$ which obviously implies $K_g = 0$. The simplest example is a plane, such as $z = 0$. Curves passing through p neither bend way or towards the normal. $\mathbf{N}$.

The terminology used should be familiar to those studying 2nd order partial differential equations. Consider the two-dimensional cases: Laplace's equation, $\nabla^2 f = \partial_x^2 f + \partial_y^2 f = 0$ is the simplest elliptic equation because the signs of the 2nd derivative terms are the same whereas the wave equation $-\partial_t^2 f + \partial_x^2 f = 0$ is classified as a hyperbolic PDE. Finally the heat equation $\partial_x^2 f - \partial_t f = 0$ is a parabolic equation (the coefficient of the 2nd order derivative $\partial_t^2 f$ vanishes).

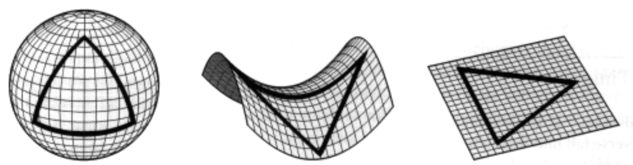

Figure 1: The area inside the triangles is related to the sign of the Gauss curvature, as we will see later.

1.4 Local Expressions for Curvature

So far we have studied simple, highly symmetric examples in which we could use the definition of the Weingarten amp to explicitly compute the principal curvatures of S. A spree, for example is *homogeneous* (loosely speaking, all points $p \in S^2$ are equivalent) and *isotropic* (again, loosely speaking all directions on the tangent space T_pS^2 are equivalent). Similarly, the cylinder $x^2 + y^2 = a^2$ is homogenous since we can always translate a point 'up' or 'down' on the $z-$axis and rotate about the xy plane) but not isotropic because directions on T_pS are 'distinguishable' . However, for more complicated surfaces, we will have to consider local charts $\mathbf{X}(u^i) \to S$ to cover patches of the surface. We can derive efficient formula for computing the Gaussian curvature with respect to coordinates in such a chart.

At this point to simplify notation we will have to use index notation. This allows us to write several equations in a compact form but takes time to get used to. For simplicity we will refer to the local coordinates (u, v) simply by $u^i = (u^1, u^2)$ so $i = 1, 2$. Derivatives with respect to these u^i are often referred to with a subscript, i.e.

$$\frac{\partial \mathbf{X}}{\partial u^i} = \mathbf{X}_i, \qquad \frac{\partial^2 \mathbf{X}}{\partial u^i \partial u^j} = \mathbf{X}_{ij}$$

Explicitly, we have

$$\mathbf{X}(u^1, u^2) = (x(u^1, u^2), y(u^1, u^2), z(u^1, u^2)) \tag{2}$$

$$\mathbf{X}_i = \left(\frac{\partial x}{\partial u^i}, \frac{\partial y}{\partial u^i}, \frac{\partial z}{\partial u^i}\right) \tag{3}$$

$$\mathbf{X}_{ij} = \left(\frac{\partial^2 x}{\partial u^i u^j}, \frac{\partial^2 y}{\partial u^i u^j}, \frac{\partial^2 z}{\partial u^i u^j}\right) \tag{4}$$

Given an arbitrary chart $\mathbf{X} : U \to S$ such that $p \in \mathbf{X}(U) \subset S$ we have seen $\mathbf{X}_i = (\mathbf{X}_1, \mathbf{X}_2)$ is a basis for T_pS. We wish to compute the action of $L : T_pS \to T_pS$ and $\Pi_p : T_pS \times T_pS \to \mathbb{R}$. We work out the components of L, Π on the basis elements $\mathbf{X}_i$; then their action on a general element $v = v^1 \mathbf{X}_1 + v^2 \mathbf{X}_2 = v^i \mathbf{X}_i \in T_pS$ can be found by using linearity. The components of L are defined as follows. Since $L(\mathbf{X}_j) \in T_pS$, we can expand it in again in terms of the basis vectors, so

$$L(\mathbf{X}_j) = L^i{}_j \mathbf{X}_i = L^1{}_j \mathbf{X}_1 + L^1{}_j \mathbf{X}_2$$

We may represent L therefore by a matrix $[L]$ with components $L^i{}_j$:

$$[L] = \begin{pmatrix} L^1{}_1 & L^1{}_2 \\ L^2{}_1 & L^2{}_2 \end{pmatrix}$$

(as you will show in the exercises, self adjointness of L implies $L^1{}_2 = L^2{}_1$). Similarly, the components of Π are the scalars $\Pi_{ij} = \Pi(\mathbf{X}_i, \mathbf{X}_j)$ (note the location of the indices is different to the components of L since they represent different types of maps). Once again, we may represent the operator Π in this basis as a matrix $[\Pi]$ with components[a] Π_{ij}:

$$[\Pi] = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12} & \Pi_{22} \end{pmatrix}$$

where, since Π is a *symmetric* map, $\Pi_{12} = \Pi_{21}$.

We will now work on he level of the chart with coordinates u^i. So when we say ‘ a curve on the surface $\mathbf{r}(t) = (u(t), v(t))$’ we understand this corresponds to the curve $\mathbf{X}(t) = \mathbf{X}(u(t), v(t))$. Similarly the tangent vector $\mathbf{r}'(t) = (u'(t), v'(t))$ corresponds to the tangent vector $\mathrm{d}\,\mathbf{X}(\mathbf{r}'(t)) = u'(t)\,\mathbf{X}_u + v'(t)\,\mathbf{X}_v = u^{i\prime}(t)\,\mathbf{X}_i$.

Proposition 1. $L(\boldsymbol{X}_j) = -\partial\,\boldsymbol{N}/\partial u^j$

Proof. From the definition, $L(\mathbf{X}_j) = -\mathrm{d}\,\mathbf{N}(\mathbf{X}_j)$. In our local chart, we have

$$\mathbf{N}(u^i) = \frac{\mathbf{X}_1 \times \mathbf{X}_2}{|\mathbf{X}_1 \times \mathbf{X}_2|}$$

where we have made the u^i-dependence explicit on the left hand side. Now consider a curve $u^i(t) = (u(t), v(t))$. By definition of the differential,

$$\mathrm{d}\,\mathbf{N}((u'(t), v'(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{N}(u(t), v(t)) = \frac{\partial\,\mathbf{N}}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}t} + \frac{\partial\,\mathbf{N}}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\partial\,\mathbf{N}}{\partial u^i}\frac{\mathrm{d}u^i}{\mathrm{d}t}$$

where we used the Chain rule in the second equality, and in the third equality we are using the summation convention to write the expression in a compact form. But of course the basis vector $\mathbf{X}_u$ corresponds to $(u'(t), v'(t)) = (1, 0)$ and $\mathbf{X}_v$ corresponds to $(u'(t), v'(t)) = (0, 1)$. Thus we read off

$$\mathrm{d}\,\mathbf{N}(\mathbf{X}_u) = \frac{\partial\,\mathbf{N}}{\partial u}, \quad \mathrm{d}\,\mathbf{N}(\mathbf{X}_v) = \frac{\partial\,\mathbf{N}}{\partial v} \quad \Rightarrow \quad \mathrm{d}\,\mathbf{N}(\mathbf{X}_j) = \frac{\partial\,\mathbf{N}}{\partial u^j}$$

and hence we arrive at

$$L(\mathbf{X}_j) = L^i{}_j\,\mathbf{X}_i = -\frac{\partial\,\mathbf{N}}{\partial u^i} \tag{5}$$

□

[a] For simplicity many authors will simply use the symbol Π_{ij} to stand for the operator Π with the understanding they the choice of basis is arbitrary.

This simple equation, however, does not allow us to compute the components ${L^i}_j$ immediately; one still has to expand the right hand side of the above equation in the basis $\mathbf{X}_i$.

Example Take a unit cylinder covered by the local chart $\mathbf{X}(u, v) = (\cos u, \sin u, v)$. The unit normal (Gauss map) is calculated easily to be

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

Thus

$$\frac{\partial \mathbf{N}}{\partial u} = (-\sin u, \cos u, 0) = \mathbf{X}_u, \qquad \frac{\partial \mathbf{N}}{\partial v} = (0, 0, 0) = 0\,\mathbf{X}_u + 0\,\mathbf{X}_v$$

where we have put the 0s in explicitly in the second expression. It follows that

$$L(\mathbf{X}_u) = -\mathbf{X}_u = {L^u}_u\,\mathbf{X}_u + {L^v}_u\,\mathbf{X}_v \Rightarrow {L^u}_u = -1, {L^v}_u = 0$$

and $L(\mathbf{X}_v) = 0 \Rightarrow {L^u}_v = {L^v}_v = 0$. So we can visualize the matrix representing L in this basis as

$$[L] = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

The coordinate chart for the cylinder is quite nice because we can immediately read off the eigenvalues $\kappa_2 = -1, \kappa_1 = 0$ with associated eigenvectors $(1, 0)$ and $(0, 1)$. Remember $(1, 0)$ corresponds to $\mathbf{X}_u$ and $(0, 1)$ corresponds to $\mathbf{X}_v$ - these are precisely the principal directions. This is of course completely consistent with the result we found earlier (see previous examples) when computing the Weingarten map using the first principles definition.

Computing $[L]$ can be tedious in practice because it involves differentiating the unit normal $\mathbf{N}$, which will invariably involve square roots. It is generally simpler to directly compute the Π_{ij} in a specific basis.

Proposition 2. *The components of the second fundamental form are given by* $\Pi_{ij} = \langle\, \boldsymbol{N}, \boldsymbol{X}_{ij}\rangle$

Proof. By definition $\langle\, \mathbf{N}, \mathbf{X}_i\rangle = 0$ since by definition the unit normal is orthogonal to $\mathbf{X}_i \in T_pS$. Differentiating with respect to u^i gives

$$\frac{\partial}{\partial u^j}\langle\, \mathbf{N}, \mathbf{X}_i\rangle = \langle \frac{\partial \mathbf{N}}{\partial u^j}, \mathbf{X}_i\rangle + \langle\, \mathbf{N}, \mathbf{X}_{ij}\rangle = 0$$

Thus

$$\langle \frac{\partial \mathbf{N}}{\partial u^j}, \mathbf{X}_i\rangle = -\langle\, \mathbf{N}, \mathbf{X}_{ij}\rangle = -\langle\, \mathbf{N}, \mathbf{X}_{ji}\rangle$$

with the second equality follows from the equality mixed partial derivatives $\mathbf{X}_{ij} = \mathbf{X}_{ji}$. Hence

$$\Pi_{ij} = \Pi(\mathbf{X}_i, \mathbf{X}_j) = \langle L(\mathbf{X}_i), \mathbf{X}_j\rangle = -\langle \frac{\partial \mathbf{N}}{\partial u^i}, \mathbf{X}_j\rangle = \langle \mathbf{N}, \mathbf{X}_{ij}\rangle = \Pi_{ji}$$

and as required, $\Pi_{ij} = \langle\, \mathbf{N}, \mathbf{X}_{ji}\rangle = \Pi_{ji}$. □

Example In the example above we considered the unit cylinder. Given $\mathbf{X}_{uu} = (-\cos u, -\sin u, 0)$, $\mathbf{X}_{uv} = \mathbf{X}_{vu} = (0,0,0)$, and $\mathbf{X}_{vv} = (0,0,0)$ it follows from $\mathbf{N} = (\cos u, \sin u, 0)$ from the above result that the only non-zero components of Π is $\Pi_{11} = -1$.

We can find a relationship between the components of L and Π:

Proposition 3. $\Pi_{ij} = g_{ik}L^k{}_j$, *that is as matrices* $[\Pi] = [g] \cdot [L]$

Proof. Recall that the components of the metric tensor g in the coordinate basis are $g_{ij} = \langle\, \mathbf{X}_i,\, \mathbf{X}_j\rangle$. But

$$\Pi_{ij} = \langle L(\,\mathbf{X}_i),\, \mathbf{X}_j\rangle = \langle L^k{}_i\,\mathbf{X}_k,\, \mathbf{X}_j\rangle = L^k{}_i\langle\,\mathbf{X}_k,\, \mathbf{X}_j\rangle = L^k{}_i g_{kj}$$

where we used linearity of the inner product to pull out the scalars $L^k{}_i$. Of course $\Pi_{ij} = \Pi_{ji}$ and simply switching the positions of i, j, we have $\Pi_{ji} = g_{ik}L^k{}_j$. This establishes the result. □

The operation of computing Π_{ij} from the components of g and L is known as 'lowering the index' because a map with both indices 'down' (i.e. Π) has been constructed form a map with one index 'up' and one 'down' (i.e. L). We also talk of 'contracting the first index of $L^i{}_j$ with the metric'. Such language is invaluable when performing calculations in Riemannian geometry and general relativity.

The expression we have proved is actually equivalent to 3 equations, one for each independent components of the matrix $[\Pi]$. For example for $(i,j) = (1,1)$ we have

$$\Pi_{11} = g_{ik}L^k{}_1 = g_{11}L^1{}_1 + g_{12}L^2{}_1$$

The equations can be written conveniently in form of a matrix equation:

$$[\Pi] = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12} & \Pi_{22} \end{pmatrix} = [g][L] = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix} \cdot \begin{pmatrix} L^1{}_1 & L^1{}_2 \\ L^1{}_2 & L^2{}_2 \end{pmatrix} \tag{6}$$

where as usual self-adjointness of L implies $L^1{}_2 = L^2{}_1$. The result (6) allows us to derive a simple expression for the Gaussian curvature directly without computing the components of L in a given basis.

Proposition 4. *The Gaussian curvature in a local chart is given by* $K_g = \dfrac{\det \Pi}{\det g}$.

Proof. Recall that $\det g$ is simply the determinant of the matrix $[g]$ representing the metric in a particular basis. Taking determinants of both sides of (6) gives

$$\det \Pi = \det g \det L = \det g K_g \tag{7}$$

where we are using the definition of the Gaussian curvature K_g (we have dropped the matrix notation [] as it is clear). By the fact that S is a regular surface, we must have $\det g \neq 0$ so it follows

$$K_g = \frac{\det \Pi}{\det g} = \frac{\Pi_{11}\Pi_{22} - \Pi_{12}^2}{g_{11}g_{22} - g_{12}^2} \tag{8}$$

□

This gives an efficient way to compute the Gaussian curvature in a local chart. In fact (6) also gives us a method to compute L in this coordinate basis simply by inverting:

$$[L] = [g]^{-1} \cdot [\Pi] \tag{9}$$

This single equation of course represents a separate equation for each component, known as the *equations of Weingarten*. We can be a bit more explicit by noting that

$$[g]^{-1} = \frac{1}{\det g}\begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \tag{10}$$

Thus

$$\begin{pmatrix} L^1{}_1 & L^1{}_2 \\ L^1{}_2 & L^2{}_2 \end{pmatrix} = \frac{1}{\det g}\begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \cdot \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12} & \Pi_{22} \end{pmatrix} \tag{11}$$

This could also be written in index notation as $L^i{}_j = g^{ik}\Pi_{kj}$ where g^{jk} stands for components of the inverse metric tensor g^{-1}. The fact that $[g] \cdot [g]^{-1} = \mathrm{Id}$ (the identity matrix) is written as $g_{ij}g^{jk} = \delta_i{}^j$ where $\delta_i{}^j = 1$ if $i = j$ and vanishes for $i \neq j$ (i.e. it just the components of the identity matrix).

Example Calculate the Gaussian curvature of a surface given by $z = f(x, y) = \frac{\kappa_1}{2}x^2 + \frac{\kappa_2}{2}y^2$ at $p = (0, 0, 0)$. We take as parameterization $\mathbf{X} = (x, y, f(x, y))$. The associated basis vectors are $\mathbf{X}_1 = (1, 0, \kappa_1 x)$ and $\mathbf{X}_2 = (0, 1, \kappa_2 y)$. We immediately read off the metric tensor $g_{11} = 1 + \kappa_1^2 x^2$, $g_{12} = g_{21} = \kappa_1\kappa_2 xy$, $g_{22} = 1 + \kappa_2^2 y^2$. This can also be written in line element form as

$$\mathrm{d}s^2 = (1 + \kappa_1^2 x^2)\mathrm{d}x^2 + 2\kappa_1\kappa_2 xy\mathrm{d}x\mathrm{d}y + (1 + \kappa_2^2 y^2)\mathrm{d}y^2 \ .$$

Note $\det g = 1 + \kappa_1^2 x^2 + \kappa_2^2 y^2$. We also have

$$\mathbf{N} = \frac{(-\kappa_1 x, -\kappa_2 y, 1)}{(1 + \kappa_1^2 x^2 + \kappa_2^2 y^2)^{1/2}}$$

where we have chosen the 'upward' pointing normal . We then can compute easily $\mathbf{X}_{11} = (0, 0, \kappa_1)$, $\mathbf{X}_{12} = (0, 0, 0)$, $\mathbf{X}_{22} = (0, 0, \kappa_2)$ so that at $x = y = 0$

$$\Pi_{ij} = \langle \mathbf{N}, \mathbf{X}_{ij} \rangle = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}$$

Then noting $\det[\Pi] = \kappa_1\kappa_2$ gives at $p = (0, 0, 0)$

$$K_g(p) = \kappa_1\kappa_2$$

This is of course consistent with previous computations of K_g we performed for the hyperbolic paraboloid $z = y^2 - x^2$ which had $K_g = -4$ at the origin. It is straightforward to find K_g for general values of (x, y).

Finally, it is worth remarking on the relationship between the principal directions and Gaussian and mean curvatures. If we denote these as v_1, v_2 with eigenvalues κ_1, κ_2 respectively, we have

$$L(v_1) = \kappa_1 v_1 \qquad L(v_2) = \kappa_2 v_2$$

so that $[L - \kappa\mathrm{Id}]\, v = 0$ for each κ, v. It follows this matrix has vanishing determinant:

$$\det\begin{pmatrix} L^1{}_1 - \kappa & L^1{}_2 \\ L^2{}_1 & L^2{}_2 - \kappa \end{pmatrix} = 0$$

giving

$$\det L - \kappa \mathrm{Tr} L + \kappa^2 = K_g - \kappa H + \kappa^2 = 0$$

We can use this relation to express the principal curvatures in terms of K_g, H:

$$\kappa = \frac{H}{2} \pm \sqrt{\frac{H^2}{4} - K_g^2}$$

with the upper (lower) sign corresponding to κ_1 (κ_2). Finally, note that from (11) we can read off

$$H = \mathrm{Tr} L = L^1{}_1 + L^2{}_2 = \frac{1}{\det g} \left[g_{22} \Pi_{11} - 2 g_{12} \Pi_{12} + g_{11} \Pi_{22} \right] \tag{12}$$

1.5 Gauss' Theorem Egregium and Riemann Curvature

One of the fundamental results of the differential geometry of surfaces is that Gaussian curvature is an *intrinsic* property, i.e. it depends only on the geometry of the surface (the metric tensor) and not on the details of how surface is embedded in the ambient $\mathbb{R}^3$. Concretely this means that although $L = -\mathrm{d}\,\mathbf{N}$ is defined in terms of the Gauss map which represents the direction of the normal to S, its determinant $K_g = \det L$ is can be computed directly in terms of g_{ij} and tis derivatives *alone*. Hence the name 'egregium' (Latin for 'remarkable'). On the way we will derive an expression for the **Riemann curvature tensor** which plays the key role in Riemannian geometry. The Riemann curvature is fairly simple in two dimensions, although tis form in a local coordinate system looks complicated, it has a nice geometric interpretation.

We begin with an arbitrary local chart $\mathbf{X}$. The metric tensor is given by the components $g_{ij} = g(\mathbf{X}_i, \mathbf{X}_j) = \langle \mathbf{X}_i, \mathbf{X}_j \rangle$ and has inverse g^{ij} given by (10). Note that since

$$[g][g]^{-1} = \mathrm{Id} \Rightarrow g_{ik} g^{kj} = \delta_i{}^j$$

We refer to the components of the identity matrix Id as the 'Kronecker delta'; that is, they are the components of the matrix

$$\delta_i{}^j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{13}$$

Now we know that the vectors $\mathbf{X}_i$ form a basis for T_pS. The set $\{\mathbf{X}_i, \mathbf{N}\}$ form a orthogonal basis for general vectors in $\mathbb{R}^3$ (the basis is not orthonormal of course since g_{11} need not be equal to one). The idea is to note that, although $\mathbf{X}_i$ are tangent to S, the vector $\mathbf{X}_{ij}$ need not be (indeed $\Pi_{ij} = \langle \mathbf{N}, \mathbf{X}_{ij} \rangle$ measures the component of $\mathbf{X}_{ij}$ orthogonal to S). This is done by expanding in the above basis: we write

$$\mathbf{X}_{ij} = \Gamma^1_{ij}\,\mathbf{X}_1 + \Gamma^2_{ij}\,\mathbf{X}_2 + \lambda_{ij}\,\mathbf{N} \tag{14}$$

$$= \Gamma^k_{ij}\,\mathbf{X}_k + \lambda_{ij}\,\mathbf{N} \tag{15}$$

where Γ^k_{ij} are called the *Christoffel symbols*. Now taking the scalar product of $\mathbf{X}_{ij}$ with $\mathbf{N}$ we read off $\lambda_{ij} = \Pi_{ij}$. Taking the scalar product with $\mathbf{X}_m$ (i.e. with $m = 1, 2$) we get

$$\langle \mathbf{X}_{ij}, \mathbf{X}_m \rangle = \Gamma^k_{ij} \langle \mathbf{X}_k, \mathbf{X}_m \rangle = \Gamma^k_{ij} g_{km} \tag{16}$$

Proposition 5. *We can express the components of $\boldsymbol{X}_{ij}$ in the $\boldsymbol{X}_i$ directions by*

$$\langle\, \boldsymbol{X}_{ij},\, \boldsymbol{X}_k\rangle = \frac{1}{2}\left[\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij}\right] \tag{17}$$

Proof.

$$\partial_k g_{ij} = \frac{\partial g_{ij}}{\partial u^k} = \partial_k\langle\, \mathbf{X}_i,\, \mathbf{X}_j\rangle = \langle\, \mathbf{X}_{ik},\, \mathbf{X}_j\rangle + \langle\, \mathbf{X}_i,\, \mathbf{X}_{jk}\rangle \tag{18}$$

We can also write the same equation by rearranging the indices

$$\partial_j g_{ik} = \langle\, \mathbf{X}_{ij},\, \mathbf{X}_k\rangle + \langle\, \mathbf{X}_i,\, \mathbf{X}_{jk}\rangle\,, \qquad \partial_i g_{jk} = \langle\, \mathbf{X}_{ij},\, \mathbf{X}_k\rangle + \langle\, \mathbf{X}_j,\, \mathbf{X}_{ik}\rangle \tag{19}$$

Adding these expressions give

$$\partial_j g_{ik} + \partial_i g_{jk} - \partial_k g_{ij} = 2\langle\, \mathbf{X}_{ij},\, \mathbf{X}_k\rangle \tag{20}$$

which establishes the result. □

Finally we can act on the right hand side of (16) with the inverse metric $[g]^{-1}$. This is equivalent to applying g^{mn} to both sides of the equation (this is known as 'contracting with the inverse metric' or 'raising an index'). This gives

$$\Gamma^k_{ij} g_{km} g^{mn} = \Gamma^k_{ij}\delta_k{}^n = \Gamma^n_{ij} = \langle\, \mathbf{X}_{ij},\, \mathbf{X}_m\rangle g^{mn}$$

Note how the free indices match on both sides of the equation. This gives us the fundamental formula for the Christoffel symbols:

$$\Gamma^n_{ij} = \frac{g^{mn}}{2}\left[\partial_j g_{im} + \partial_i g_{jm} - \partial_m g_{ij}\right] \tag{21}$$

Note that the Christoffel symbols are *symmetric* in the lower two components, i.e. $\Gamma^n_{ij} = \Gamma^n_{ji}$. In two dimensions, we have $i, j = 1, 2$ and in total this gives 6 independent components (there are $2^3 = 8$ components in total, but symmetry implies 2 of them are not independent). They can be computed in terms of the inverse metric and its first derivatives with respect to the local coordinates u^i. Computing these components can be a tedious task and for a specific chart, variational methods (i.e. computing the geodesic equation) is a far more efficient way of calculation. We will get to this later. We simply note here for those familiar with tensors that the Christoffel symbols are *not* tensors; they are a structure one can place on a manifold known as a *connection.* The particular connection we are using is induced from the metric and is known as the *Levi-Civita* connection. A connection gives rise to a notion of *parallel transport* (i.e. a way to determine if two vectors at different points on S are parallel).

The equations (14), known as *Gauss' formula*, together with the Weingarten's equation (11) (or equivalently $L(\,\mathbf{X}_j) = -\partial_j\,\mathbf{N}$) play the same role as the Frenet equations for curves, relating the first derivatives of the frame $(\,\mathbf{X}_1,\, \mathbf{X}_2,\, \mathbf{N})$ on the surface.

Example We compute the Christoffel symbols for the unit sphere S^2 in the parametrization $\mathbf{X}(\theta, \phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$. We set $u^i = (u^1, u^2) = (\theta, \phi)$. We have already computed the metric tensor in this basis:

$$[g] = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \quad\Rightarrow\quad [g]^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2\theta} \end{pmatrix} \tag{22}$$

The only non-vanishing components are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta \qquad \Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi} = \cot\theta \tag{23}$$

Exercise Compute the Christoffel symbols in the chart $\mathbf{X}(u,v)$ used to cover a portion of the unit cylinder.

We now proceed with the proof that K_g is an intrinsic quantity. The method follows Gauss. We begin by taking another derivative of Gauss' formula (14). This gives

$$\begin{aligned} \mathbf{X}_{ijk} &= \partial_k\Gamma^l_{ij}\,\mathbf{X}_l + \Gamma^l_{ij}\,\mathbf{X}_{lk} + \partial_k\Pi_{ij}\,\mathbf{N} + \Pi_{ij}\partial_k\,\mathbf{N} \\ &= \partial_k\Gamma^l_{ij}\,\mathbf{X}_l + \Gamma^l_{ij}\,[\Gamma^m_{lk}\,\mathbf{X}_m + \Pi_{lk}\,\mathbf{N}] + \partial_k\Pi_{ij}\,\mathbf{N} - \Pi_{ij}L^l{}_k\,\mathbf{X}_l \\ &= \left[\partial_k\Gamma^l_{ij} + \Gamma^m_{ij}\Gamma^l_{mk} - \Pi_{ij}L^l{}_k\right]\,\mathbf{X}_l + \left[\Gamma^l_{ij}\Pi_{lk} + \partial_k\Pi_{ij}\right]\,\mathbf{N} \end{aligned} \tag{24}$$

where in the last line we have again decomposed the vector into bits tangent to S and a component normal to S. Note carefully the position of the free indices (ijk) and the 'dummy indices' (those summed over). In doing these calculations it is important to realize one can relabel dummy indices - this was done in particular in the 2nd term of the 2nd line above (the roles of the dummy indices m and l were exchanged). Now we can rewrite (24), but with the relabellings $(j,k)\to(k,j)$:

$$\mathbf{X}_{ikj} = \left[\partial_j\Gamma^l_{ik} + \Gamma^m_{ik}\Gamma^l_{mj} - \Pi_{ik}L^l{}_j\right]\,\mathbf{X}_l + \left[\Gamma^l_{ik}\Pi_{lj} + \partial_j\Pi_{ik}\right]\,\mathbf{N} \tag{25}$$

Note that (24) and (25) represent the *same* set of equations. We have just rewritten them in order to exploit the fact that

$$\mathbf{X}_{ijk} = \frac{\partial^3\,\mathbf{X}}{\partial u^k\partial u^j\partial u^i} = \frac{\partial^3\,\mathbf{X}}{\partial u^j\partial u^k\partial u^i} = \mathbf{X}_{ikj} \tag{26}$$

again by equality of mixed partial derivatives. Note that as $\mathbf{X}_i$, $\mathbf{N}$ form a basis, it follows that two vectors are equal if and only if their components in the $\mathbf{N}$ direction and $\mathbf{X}_i$ directions are equal. Hence if we subtract (24) from (25) we must have

$$0 = \left[\partial_j\Gamma^l_{ik} + \Gamma^m_{ik}\Gamma^l_{mj} - \Pi_{ik}L^l{}_j - \left(\partial_k\Gamma^l_{ij} + \Gamma^m_{ij}\Gamma^l_{mk} - \Pi_{ij}L^l{}_k\right)\right]\,\mathbf{X}_l = 0 \tag{27}$$

and since $\mathbf{X}_l$ are non-vanishing, we must have

$$\partial_j\Gamma^l_{ik} - \partial_k\Gamma^l_{ij} + \Gamma^m_{ik}\Gamma^l_{mj} - \Gamma^m_{ij}\Gamma^l_{mk} = \Pi_{ik}L^l{}_j - \Pi_{ij}L^l{}_k \tag{28}$$

The quantity on the left hand side of this equation is one of the most important objects in differential geometry.

Definition. *The Riemann curvature tensor is defined by*

$$R^l{}_{ijk} \equiv \partial_j\Gamma^l_{ik} - \partial_k\Gamma^l_{ij} + \Gamma^m_{ik}\Gamma^l_{mj} - \Gamma^m_{ij}\Gamma^l_{mk} \tag{29}$$

Note that $R^l{}_{ijk} = -R^l{}_{ikj}$.

Remark. Note that the Riemann tensor can be computed *entirely from* g_{ij} *and its derivatives.*

We have been using the term *tensor* without a formal definition. We have already been introduced to tensors such as the metric tensor g and second fundamental form, or extrinsic curvature tensor, Π. These are bilinear maps from $T_pS\times T_pS\to\mathbb{R}$. As they have two arguments, they are easily represented by their components g_{ij},Π_{ij} respectively, and we may visualize them as matrices $[g],[\Pi]$. More generally, tensors are multilinear maps from n copies of T_pS and m copies of T_p^*S to $\mathbb{R}$. Here T_p^*S stands for the *cotangent space* , the vector space dual to T_pS. In this language, we refer to (m,n)-tensors. The metric tensor and second fundamental form are $(0,2)$ tensors since they take two elements of T_pS as arguments. The Riemann curvature tensor is naturally expressed as a $(1,3)$ tensor, as can most easily be seen from its index structure above (one index 'up' and three 'down'). Of course, we cannot represent it as a matrix - we would need a four-dimensional piece of paper to do so (!).

It can be shown that the number of independent components of the Riemann tensor in n dimensions is

$$\frac{n^2(n^2-1)}{12}$$

using its antisymmetry property and an additional identity known as the Bianchi identity (which is trivial in 2 dimensions). Substituting $n=2$ in the above formula one gets precisely only *one* independent component . Informally, this means that the intrinsic curvature of the surface at a point is completely specified by one number; as we now show, this number is the Gaussian curvature of S. Hence we will not have cause to undertake the formidable task of computing the Riemann tensor directly for two dimensional surfaces, as all we need to do is work out K_g. In higher dimensions, however, the computation is required and there are a number of techniques (including software) that allows one to do this efficiently.

Theorem 2. *(***Theorem Egregium***) The Gaussian curvature depends only on the intrinsic geometry of S*

Proof. From (28) we have

$$R^l_{ijk} = \Pi_{ik}L^l{}_j - \Pi_{ij}L^l{}_k \tag{30}$$

Lower the index on both sides by applying g_{ml}:

$$\begin{aligned} g_{ml}R^l{}_{ijk} &= \Pi_{ik}g_{ml}L^l{}_j - \Pi_{ij}g_{ml}L^l{}_k \\ &= \Pi_{ik}\Pi_{mj} - \Pi_{ij}\Pi_{mk} \end{aligned}$$

using Proposition 3. The above expression must be true for any chosen values for i,j,k,m. We take$(i,j,k,m) = (1,2,1,2)$ to find

$$g_{2l}R^l{}_{121} = \Pi_{11}\Pi_{22} - \Pi_{12}\Pi_{12} = \det\Pi = \det g \cdot \det L \tag{31}$$

and hence we arrive at

$$K_g = \frac{g_{2l}R^l{}_{121}}{\det g} = \frac{g_{21}R^1{}_{121} + g_{22}R^2{}_{121}}{\det g} \tag{32}$$

Since $R^l{}_{ijk}$ given by (29) is computable entirely in terms of the metric tensor g_{ij} and its derivatives (via the Christoffel symbols and its derivatives) the right hand side of Gauss' formula is computable entirely in terms of functions defined intrinsically on S. □

There is a final fundamental relation which relates the second fundamental form and the intrinsic geometry. Once again we subtract (25) from (24) and equating the $\mathbf{N}$ components gives

$$\partial_j\Pi_{ik} - \partial_k\Pi_{ij} = \Gamma^l_{ij}\Pi_{lk} - \Gamma^l_{ik}\Pi_{lj} \tag{33}$$

These are known as the Codazzi equations (or Codazzi-Mainardi); they express *compatibility* conditions between the intrinsic and extrinsic geometry. Indeed, the equation contains second derivatives of the Gauss map (i.e. the $\partial_j\Pi_{ik}$ terms) as well as first derivatives of $\mathbf{N}$ and g.

Naturally one would wonder if there are more relations between the metric, second fundamental form, and their derivatives. In fact the Guass equation and Codazzi equations exhaust all possibilities. This leads to the following theorem due to Bonnet, which is the analogue for regular surfaces of the Fundamental Theorem for Curves.

Theorem 3. Fundamental theorem of Surfaces *Let $g_{11}, g_{12}, g_{22}, \Pi_{11}, \Pi_{12}, \Pi_{22}$ be differentiable functions on an open set $V \subset \mathbb{R}^2$ (i.e. smooth functions of coordinates (u^1, u^2)) with $g_{11}, g_{22} > 0$ and $g_{11}g_{22} - g_{12}^2 > 0$ and assume that Gauss and Codazzi equations are satisfied. Then for each $p \in V$ there is an open set $U \subset V$ and a smooth map $\boldsymbol{X} : U \to \mathbb{R}^3$ such that the $\boldsymbol{X}(U)$ is a regular surface with metric tensor g_{ij} and second fundamental form Π_{ij}.Any other parametrization with the same properties will be related to $\boldsymbol{X}$ by a rigid motion.*

2 Maps between surfaces and isometries

Gauss' Theorem Egregium is often stated in terms of *isometries*, which are a certain type of maps between surfaces under which the geometry is invariant. We now briefly study maps between surfaces. Suppose S_1, S_2 are two regular surfaces and we have a continuous map $\Phi : V \subset S_1 \to S_2$. The map Φ is said to be differentiable at $p \in V \subset S_1$ if given parameterizations

$$\mathbf{X}_1 : U_1 \subset \mathbb{R}^2 \to S_1 \qquad \mathbf{X}_2 : U_2 \subset \mathbb{R}^2 \to S_2$$

with $\mathbf{X}_1(q) = p$ with $q \in U_1$ and $\Phi(\mathbf{X}_1(U_1)) \subset \mathbf{X}_2(U_2)$, the map from $U_1 \to U_2$ (which are two open sets of $\mathbb{R}^2$) given by

$$\hat{\Phi} \equiv \mathbf{X}_2^{-1} \circ \Phi \circ \mathbf{X}_1 : U_1 \to U_2$$

is differentiable at q. What this means in concrete examples is that if (u, v) are local coordinates for a part of the surface S_1, and (r, s) are local coordinates of a patch of S_2, then the map $\hat{\Phi}(u, v) = (r(u, v), s(u, v))$.

Example Suppose x, y are local coordinates for the $z = 0$ plane with $\mathbf{X}_1(x, y) = (x, y, 0)$ seen as a surface S_1 in $\mathbb{R}^3$. Now let S_2 be the unit cylinder aligned along the z-axis with chart $\mathbf{X}_2(u, v) = (\cos u, \sin u, v)$. Then a smooth map Φ from a portion of S_1 to S_2 is given by the map $u = x, v = y$.

Definition. *Two surfaces S_1 and S_2 are* diffeomorphic *if there is a differentiable map $\Phi : S_1 \to S_2$ with differentiable inverse $\Phi^{-1} : S_2 \to S_1$.*

Remark. Note that this means *all of* S_1 can be mapped to S_2 and vice versa, not just a portion of the surface. Intuitively, diffeormorphic surfaces are 'the same' from the point of view of differentiability. Of course their geometry (lengths, angles, curvature) might be different.

Example Let S_1 be the surface defined by the equation $x^2 + y^2 + z^2 = 1$. By previous results we know this is a regular surface. Now suppose we consider the map $\Phi(x, y, z) = (r, s, u) = (ax, by, cz)$. Then $(r/a)^2 + (s/b)^2 + (u/c)^2 = 1$ so we have a map from S^2 to an ellipsoid. Note that the map Φ 'stretches' or 'contracts' the coordinates (x, y, z), which is to say the sphere can be smoothly squashed and stretched into the ellipsoid. They are diffeormorphic.

In contrast, it may be the case that an open set U of S_1 can be mapped differentiably into into a open set V of S_1. We then call this is a *local diffeormophism* between S_1 and S_2. The example of the plane and the cylinder above is a local diffeormorphism (the u, v coordinates do not cover the whole cylinder).

Given a map Φ between surfaces, we may also define the differential $\mathrm{d}\Phi$ of the map. This is a map from the tangent spaces of S_1 to the tangent spaces of S_2 .That is, $\mathrm{d}\Phi : T_pS_1 \to T_{\Phi(p)}S_2$. The map $\mathrm{d}\Phi$ hanse takes vectors in T_pS_1 and 'pushes them forward' to vectors in $T_{\Phi(p)}S_2$. We can compute the explicit from of $\mathrm{d}\Phi$ in local charts just by using the usual rules. Suppose $\mathbf{X}_1(u, v)$ is a chart on S_1 and $\mathbf{X}_2(r, s)$ is a chart for S_2. The map Φ is expressed in local coordinates as $\hat{\Phi}(u, v) = (r(u, v), s(u, v))$. Now suppose we have a curve on S_1 described locally as $\alpha(t) = (u(t), v(t))$. This gets mapped to a curve in S_2 given locally by $\beta(t) = (r(u(t), v(t)), s(u(t), v(t)))$. So if p is the point on S_1 corresponding to $t = 0$ (i.e. $p = \mathbf{X}_1(u(0), v(0)))$ then the tangent to the curve at this point has components $(u'(0), v'(0))$ in the basis $\mathbf{X}_{1u}$, $\mathbf{X}_{1v}$. The curve in S_2 is given $\mathbf{X}_2(\beta(t)) = \mathbf{X}_2(r(u(t), v(t)), s(u(t), v(t)))$. Then by definition of the differential

$$\left.\frac{d\,\mathbf{X}(\beta(t))}{\mathrm{d}t}\right|_{t=0} = \left(\frac{\partial r}{\partial u}u'(0) + \frac{\partial r}{\partial v}v'(0)\right)\mathbf{X}_{2r} + \left(\frac{\partial s}{\partial u}u'(0) + \frac{\partial s}{\partial v}v'(0)\right)\mathbf{X}_{2s} \tag{34}$$

Hence we have obtained a vector in $T_{\Phi(p)}S_2$ (note $\mathbf{X}_{2r}$, $\mathbf{X}_{2s}$ is a basis for tangent vectors in S_2). In terms of coordinate components (which we are most often interested in) this can be written compactly as follows. if w is a vector in T_pS_1 with coordinates (w^1, w^2) in the usual basis associated to the chart $\mathbf{X}_1$, then it is mapped to the following vector with components in the basis associated with $\mathbf{X}_2$:

$$\mathrm{d}\hat{\Phi}(w) = \begin{pmatrix} \frac{\partial r}{\partial u} & \frac{\partial r}{\partial v} \\ \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \end{pmatrix} \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} \tag{35}$$

Since $\mathbf{X}_{1u}$ corresponds to $w = (1, 0)$ and $\mathbf{X}_{1v}$ to $(0, 1)$, we have

$$\mathrm{d}\Phi(\mathbf{X}_{1u}) = \frac{\partial r}{\partial u}\mathbf{X}_{2r} + \frac{\partial s}{\partial u}\mathbf{X}_{2s} \qquad \mathrm{d}\Phi(\mathbf{X}_{1v}) = \frac{\partial r}{\partial v}\mathbf{X}_{2r} + \frac{\partial s}{\partial v}\mathbf{X}_{2s} \tag{36}$$

The transformation matrix representing the change of the components in a particular basis

$$\mathrm{d}\hat{\Phi}^i_j = \begin{pmatrix} \frac{\partial r}{\partial u} & \frac{\partial r}{\partial v} \\ \frac{\partial s}{\partial u} & \frac{\partial s}{\partial v} \end{pmatrix} \tag{37}$$

plays an important role in the pushforward of vectors from TS_1 to TS_2. In particular, we want this linear map to be one-to one and onto (a bijection) . This way every vector in TS_1 has a unique image in TS_2 and vice versa. In the language of linear algebra we want the two vector spaces to be isomorphic. This just means that $\det \mathrm{d}\hat{\Phi} \neq 0$. This property is independent of the particular choice of chart used, so we often drop the '$\hat{\ }$' and say that the tangent spaces are isomorphic if $\det \mathrm{d}\Phi \neq 0$. This allows us to talk about a *local diffeormorphsim*.

Definition. *We say that a map $\Phi : U \subset S_1 \to S_2$ is a* local diffeomorphism *at $p \in U$ if there is a neighbourhood $V \subset U$ of p such that Φ restricted to V is a diffeomorphism onto an open set $\Phi(V) \subset S_2$.*

Informally this means that locally near p, the two surfaces 'look the same' from the point of view of differentiability. Of course they might not look the same globally because a local diffeomorphism might not exist for every $p \in S_1$.

Theorem 4. *(***Inverse function theorem***) Let $\Phi : U \subset S_1 \to S_2$ be a map between two regular surfaces such that the differential $d\Phi_p : T_pS_1 \to T_{\Phi(p)}S_2$ is an isomorphism, i.e. $\det d\Phi_p \neq 0$. Then Φ is a local diffeomorphism near p.*

Example This may seem abstract but it is all quite concrete in specific examples. Return to the previous example where S_1 is the $z = 0$ plane and S_2 is the cylinder with the given charts $\mathbf{X}_1$ and $\mathbf{X}_2$. The map $\hat{\Phi}(x, y)$ is simply given by $u = x, v = y$. Hence the differential of the map $\mathrm{d}\hat{\Phi}$ is just the identity matrix, which has determinant 1, that is

$$\mathrm{d}\hat{\Phi}^i_j = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{38}$$

Hence the plane and the cylinder are locally diffeormorphic according to the above theorem. Of course they are not globally diffeormorphic, because if you move on a straight path around the circle of the cylinder, you will eventually return to where you started, whereas this cannot happen on the plane. The plane and cylinder have different *topology*, or equivalently they are not *homeomorphic*.

We can now define the meaning of a map between surfaces to be an *isometry*. As the name suggests this means that the metric tensor g is preserved under the map. This implies that the lengths and angles between vectors in TS_1 is unchanged after they are 'pushed forward' into TS_2. We write this as follows. To avoid too many subscripts, we refer to S_2 as $\hat{S}$.

Definition. *Let $S, \hat{S}$ be two surfaces with metric tensors $g, \hat{g}$ respectively and Φ is a diffeomorphism $\Phi : S \to \hat{S}$. Then $S, \hat{S}$ are* isometric *if for all $p \in S$, and all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in T_pS$, we have*

$$g(\mathbf{v}_1, \mathbf{v}_2) = \langle v_1, v_2 \rangle = \hat{g}(\mathrm{d}\Phi_p(\mathbf{v}_1), \mathrm{d}\Phi_p(\mathbf{v}_2)) = \langle \mathrm{d}\Phi(\mathbf{v}_1), \mathrm{d}\Phi(\mathbf{v}_2) \rangle \tag{39}$$

Remark. The definition extends in the obvious way for a *local isometry.*

To emphasize, two diffeomorphic surfaces may not be isometric; the latter condition indicates a 'rigidity' of the smooth map. As we have seen, the intrinsic geometry (and indeed the Gaussian curvature) is totally determined in terms of the metric. So if the metric is preserved, then $S, \hat{S}$ actually have the *same* geometry. We will state this more precisely shortly. Firstly though we want to have a concrete way in terms of local coordinates of determining whether a map Φ is an isometry.

Take charts $\mathbf{X}_1(u^i)$, $\hat{\mathbf{X}}(\hat{u}^i)$ on $S, \hat{S}$ respectively (to avoid a lot of clutter, we will just refer to $u^i = (u, v)$ and $\hat{u}^i = (r, s)$). As usual to check if the map is a local isometry, we need only check the above condition on the basis vectors $\mathbf{X}_i$; the result for general $\mathbf{v}_1, \mathbf{v}_2$ will follow by linearity. In terms of local coordinates, the map Φ will be given by $\hat{u}^i = \hat{\Phi}^i(u^1, u^2)$. According to (36) the basis vectors $\mathbf{X}_i$ transform to:

$$\mathrm{d}\Phi(\mathbf{X}_i) = \frac{\partial \hat{\Phi}^j}{\partial u^i} \hat{\mathbf{X}}_j \tag{40}$$

So in terms of charts, the condition that the map generates an isometry is

$$g_{ij} = g(\mathbf{X}_i, \mathbf{X}_j) = \hat{g}(\mathrm{d}\Phi(\mathbf{X}_i), \mathrm{d}\Phi(\mathbf{X}_j)) = \hat{g}\left(\frac{\partial \hat{\Phi}^k}{\partial u^i} \hat{\mathbf{X}}_k, \frac{\partial \hat{\Phi}^l}{\partial u^j} \hat{\mathbf{X}}_l \right) = \hat{g}_{kl} \frac{\partial \hat{\Phi}^k}{\partial u^i} \frac{\partial \hat{\Phi}^l}{\partial u^j} \tag{41}$$

This single equation represents three equations (one for each choice of i, j, taking into account the symmetry $g_{ij} = g_{ji}$.)

Example Let us take the previous example of the local diffeomorphism between the plane and the cylinder. Here (x, y) are coordinates on the plane and (u, v) are coordinates on the cylinder. In terms of our general language above, the unhatted coordinates $u^i = (x, y)$ and the hatted coordinates $\hat{u}^i = (u, v)$. Our map is given by $(u, v) = (\hat{\Phi}^1(x, y), \hat{\Phi}^2(x, y)) = (x, y),$. Now the metric tensor on the plane is clearly just

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{42}$$

in the standard basis associated to the coordinates (x, y). We have already computed the metric tensor on the cylinder in the coordinates (u, v). For convenience the reader is reminded that $\hat{\mathbf{X}}_u = (-\sin u, \cos u, 0)$ and $\hat{\mathbf{X}}_v = (0, 0, 1)$. Taking the dot products simply gives

$$\hat{g}_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{43}$$

To check the map is an isometry, we need to compute

$$\hat{g}_{kl} \frac{\partial \hat{\Phi}^k}{\partial u^i} \frac{\partial \hat{\Phi}^l}{\partial u^j} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{44}$$

This is easy to see because the matrices $\partial \hat{\Phi}^i / \partial u^i = \delta_j{}^i$, i.e. the identity, and $\hat{g}$ is just the identity in this chart. Hence the cylinder and plane are locally isometric - they have the same geometry, and locally measurements of lengths and angles would give the same answer (recall that a creature living on the surface only has access to the local coordinate on the chart it is using). We already have noted that $K_g = 0$ for both the cylinder and plane; this is no accident, because we have seen they have the same metric tensor and the Gaussian curvature is intrinsic, that is, it is completely determined in terms of the metric and its derivatives. In summary we can state Gauss' Theorem Egregium in a different way, which in fact is the way he originally described it:

Theorem 5. Theorema Egregium (ii) *The Gauss curvature K_g is invariant under local isometries of a surface*

Remark. We have already proved this fact. K_g can be determined entirely by measuring angles and distances (and their rates of change) on S itself, without further reference to the way in which the surface is embedded in the ambient $\mathbb{R}^3$. For example, by performing local measurements, we can compute the Gaussian curvature of the Earth - roughly $1/R_e^2$ up to ellipsoidal deformations - without ever looking at the Earth from space.

Example Take the parameterization of the unit S^2 given by

$$\mathbf{X}(\theta, \phi) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)$$

and a parametrization of the ellipse: $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ given by

$$\hat{\mathbf{X}}(\alpha, \beta) = (a \cos\alpha \sin\beta, b \sin\alpha \cos\beta, c \cos\beta)$$

along with the map $(\alpha, \beta) = \hat{\Phi}(\theta, \phi) = (\theta, \phi)$. Is this an isometry?

3 Intrinsic geometry

We now focus on the geometry of the surface without reference to how it is embedded. In particular, we want to understand the notion of 'straightness' on the surface which leads naturally to the definition of geodesics, parallel transport, and the covariant derivative. All these ideas generalize naturally to Riemannian geometry. Finally, we will arrive at a sketch of a proof of the deepest theorem in differential geometry of surfaces, the *Gauss-Bonnet* theorem. This establishes a remarkable relationship between *local* geometry (the Gaussian curvature K_g) of S with the *topology* of S (the Euler characteristic χ).

3.1 Geodesics and Covariant Derivatives

Preliminaries

Our starting point is to extend the notion of 'straightness' from $\mathbb{R}^n$ to curved surfaces. Consider therefore first a straight line in Euclidean space $\mathbb{R}^n$. A straight line can be parametrized as $\mathbf{r}(s) = \mathbf{v}s + \mathbf{c}$ where $|\mathbf{v}| = 1$ is a unit vector, s is the arc length parameter, and $\mathbf{c}$ is a constant vector. Clearly $\mathbf{r}''(s) = 0$ and so we see the curvature $\kappa = |\mathbf{r}''(s)| = 0$ of the straight line is unsurprisingly, zero. Alternatively, straight lines can be described as the curves of shortest length joining two points $p, q \in \mathbb{R}^n$. Finding the 'extremal curve' is a problem of variational calculus; we consider all possibly paths joining p, q and compare their lengths, trying to find the curve of minima length (assuming it even exists). That is we wish to minimize the functional

$$L(\mathbf{r}) = \int_a^b |\mathbf{r}(s)| \mathrm{d}s \tag{45}$$

over all curves with $\mathbf{r}(a) = p$, $\mathbf{r}(b) = q$. The Euler-Lagrange equations for this problem can be solved to show the straight lines are indeed the minimizers. Straight lines are called *geodesics*.

Now we turn to surfaces. Consider a sphere. Clearly the 'straight lines' of the sphere will still be curved in the sense they are restricted to lie on the surface. The preceding discussion suggest two ways to extend the definition of geodesics to surfaces:

1. Demand that geodesics on S satisfy the condition $\kappa = 0$. Recall that on surfaces we distinguish two components of curvature on S; a component normal to $\mathbf{N}$, which we called the normal curvature $\kappa_N = \kappa n \cdot \mathbf{N} = \kappa \cos\theta$ where$n$ is the usual normal vector of the curve considered as a curve in $\mathbb{R}^3$, and the tangential curvature, which we define as the component of the curvature κ tangent to S, say $\kappa_T = \kappa \sin\theta$. Then the natural analogue of a straight line in $\mathbb{R}^n$ is a curve that is 'straight' from the point of the surface: thus $\kappa_T = 0$. For a general curve lying on the surface $\mathbf{r}(s) \subset S$ this translates into a second order differential equation, just like $\mathbf{r}''(s) = 0$ in Euclidean space.

2. We could return to the length functional given above, replacing the integrand $|\mathbf{r}'(s)|$ with its natural analogue on the surface. This gives the functional on curves lying on the surface:

$$L(\mathbf{r}) = \int_a^b [g(\mathbf{r}'(s), \mathbf{r}'(s))]^{1/2} \, \mathrm{d}s \tag{46}$$

   The Euler-Lagrange equations for this problem also yield a second order differential equation.

Remarkably these two methods of approaching geodesics coincide. The second approach is in a sense intuitively more natural, but it turns out one runs into difficulties if the points $p, q \in S$ are sufficiently far apart - a geodesic may fail to be a minimizer if there are 'conjugate points' between p and q. On the level of computation however, using the Euler-Lagrange equations for the functional above is the most efficient way to finding geodesics. and the Christoffel symbols.

Example We consider an example of a curve whose curvature as measured on S vanishes, which by the suggestion above, is what we mean by a 'straight line'. Consider the unit sphere $x^2 + y^2 + z^2 = 1$ and the great circle on the $z = 0$ plane given by $\mathbf{r}(s) = (\cos s, \sin s, 0)$. Remember the normal vector is $\mathbf{N} = (x, y, z)$. Now note that $\mathbf{r}''(s) = \kappa n = (-\cos s, -\sin s, 0) = -\mathbf{N}$. Hence the normal curvature $\kappa_N = -\kappa = -1$, that is $n = -\mathbf{N}$ so $\cos\theta = -1$. Thus $\kappa_T = 0$ (i.e. $\mathbf{r}''(s)$ has no components tangent to the sphere). So the curve is not bending from the point of view of S^2: it is therefore a 'straightest line' or geodesic of S^2. We now turn to more precise definitions.

Geodesic curvature

Intuitively, a geodesic is a curve $\gamma(t)$ on S whose acceleration $\gamma''(t)$ is either zero or parallel to $\mathbf{N}$ (i.e. it is not accelerating relative to S). This definition is somewhat unsatisfactory because it depends on the extrinsic geometry and so cannot be generalized to Riemannian manifolds. However, it does give good intuition before proceeding to the intrinsic definition.

Proposition 6. *A geodesic $\gamma(t)$ has constant speed.*

Proof. We know

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \gamma'(t), \gamma'(t)\rangle = 2\langle \gamma''(t), \gamma'(t)\rangle \tag{47}$$

But if $\gamma(t)$ is a geodesic, then the acceleration $\gamma''(t)$ is normal to the surface, i.e. orthogonal to vectors tangent to S at each value of t. But since $\gamma(t)$ lies on the surface, by definition $\gamma'(t) \in T_{\gamma(t)}S$. Hence $\gamma'' \cdot \gamma' = 0$ and so the speed $|\gamma'(t)|$ is constant. □

We now turn to defining the geodesic curvature. Let $\mathbf{r}(s)$ be a curve on S (we parametrize by arc length for simplicity). The curve has unit tangent $T = \mathbf{r}'$ and at each point we have a normal vector to the surface $\mathbf{N}$. Recall that from our definition of curvature for space curves,

$$T' = \frac{\mathrm{d}T}{\mathrm{d}s} = \kappa\mathbf{n} \tag{48}$$

where $\mathbf{n}$ is the normal vector to the curve and κ is the curvature. We now define another unit vector $\hat{T} = T \times \mathbf{N}$. Notice that $\hat{T}$ is orthogonal to $\mathbf{N}$ so it lies in T_pS; further it is orthogonal to T, so the set $(T, \hat{T})$ can be thought of as an orthonormal basis for T_pS. We now expand T' into the basis $(T, \hat{T}, \mathbf{N})$. Obviously $T' \cdot T = 0$ since $|T|^2 = 1$. Hence we must have

$$T' = \kappa\mathbf{n} = \alpha\,\mathbf{N} + \beta\hat{T} \tag{49}$$

and taking the dot product with $\mathbf{N}$ gives $\alpha = \kappa_N = \kappa(\mathbf{n} \cdot \mathbf{N})$, the normal curvature of the curve which we have already discussed. The scalar β measures the component of the curvature that lies tangent to S and we call this the geodesic curvature κ_g.

Definition. *The geodesic curvature of a curve lying on the surface is given by*

$$\kappa_g = T' \cdot \hat{T} = T' \cdot (T \times \boldsymbol{N}) \tag{50}$$

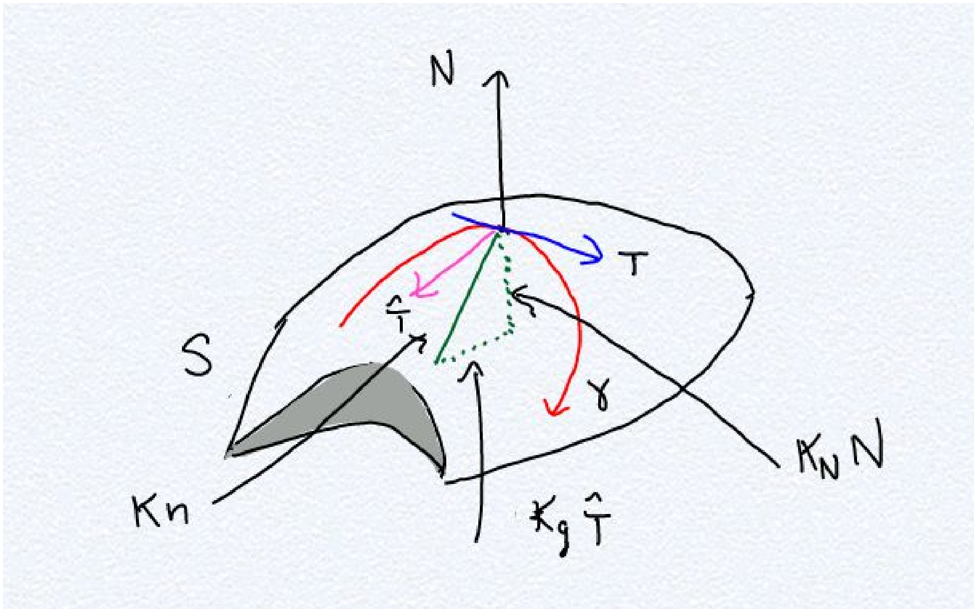

Figure 2: Geodesic curvature. Note that $\kappa^2 = \kappa_g^2 + \kappa_N^2$. Note that $(T, \hat{T})$ span T_pS.

Notice that $T' \cdot T' = \kappa^2(\mathbf{n} \cdot \mathbf{n}) = \kappa^2 = \kappa_N^2 + \kappa_g^2$. We may therefore say that the curvature of a curve lying on the surface splits into two pieces, one parallel $\mathbf{N}$ and the other tangent to S.

Proposition 7. *A curve on a surface is a geodesic if and only if its geodesic curvature is zero*

Proof. If $\gamma(s)$ is a geodesic, then by the above definition we have

$$\kappa_g = T' \cdot (T \times \mathbf{N}) = \gamma'' \cdot (\gamma' \times \mathbf{N}) \tag{51}$$

and further $\gamma''(s)$ is parallel to $\mathbf{N}$; hence it must be orthogonal to $\gamma' \times \mathbf{N}$and so $\kappa_g = 0$. On the other hand if $\kappa_g = 0$ then γ'' must be orthogonal to $\gamma' \times \mathbf{N}$, i.e. $\gamma'' \cdot (\gamma' \times \mathbf{N}) = 0$ (assuming of course $\gamma'' \neq 0$). But since $\gamma' \cdot \gamma' = 1$ (unit speed) it follows $\gamma'' \cdot \gamma' = 0$. Then it follows that γ'' must be parallel to $\gamma' \times (\gamma' \times \mathbf{N})$ (since it is orthogonal to each vector appearing on the left and right of the first cross product). Therefore γ'' is parallel to $\mathbf{N}$. This can be made more transparent by using vector identity

$$\gamma'' \propto \gamma' \times (\gamma' \times \mathbf{N}) = (\gamma' \cdot \mathbf{N})\gamma' - (\gamma' \cdot \gamma')\,\mathbf{N} = -\,\mathbf{N} \tag{52}$$

□

A simpler way to see this actually is to expand $\gamma'' = c_1 T + c_2 \hat{T} + c_3\,\mathbf{N}$ using our orthonormal basis. If $\kappa_g = 0$ then $c_2 = 0$ and further $c_1 = 0$ because $\gamma'' \cdot \gamma' = \gamma'' \cdot T = 0$. Thus we immediately see γ'' is parallel to $\mathbf{N}$. The notion of geodesic curvature gives a simple way to identity geodesics in cases with high symmetry (the condition is sufficient although not necessary).

Proposition 8. *Suppose P is a plane that intersects S orthogonally at every point of intersection. Then the curve of intersection is a geodesic.*

Proof. The curve is a plane curve (i.e. it is tangent to P everywhere). So $T = \gamma'$ is tangent to P. Likewise $T' = \gamma''$ is tangent to P (otherwise the curve would be leaving, or 'accelerating' off, P). Since the plane intersects S orthogonally at each point of intersection, this means that $\mathbf{N}$ lies on P. Thus $(T, \mathbf{N})$ form an orthonormal basis for vectors tangent to P at each point of intersection. But $\gamma'' \cdot T = T' \cdot T = 0$. Thus $\gamma'' \propto \mathbf{N}$ (i.e. $\kappa_g = 0$) and so $\gamma(s)$ is a geodesic. □

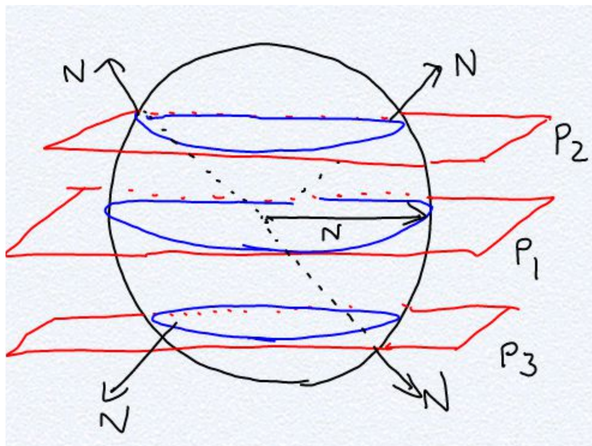

Figure 3: Curves of intersection of the planes P_i with S^2 are shown in blue. Only the middle curve is a geodesic, since P_1 contains $\mathbf{N}$

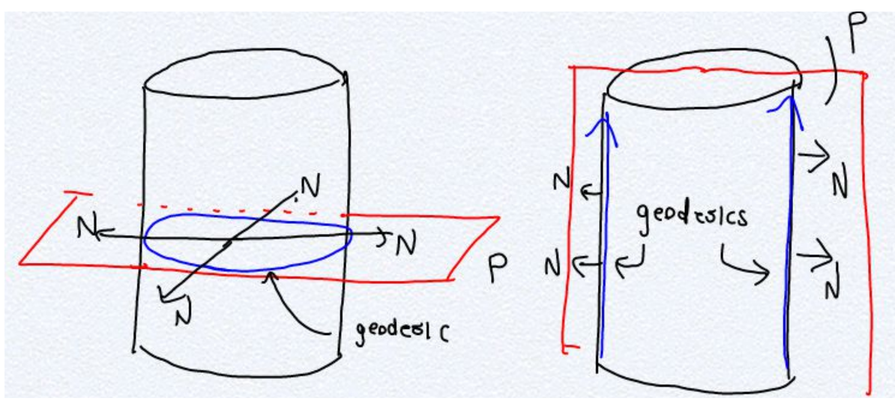

Figure 4: Planes parallel to xy plane intersect the cylinder along geodesics, as well as vertical planes.

Covariant Derivatives

We first formulae a notion of 'differentiation of vector fields' on S. Then the geodesic condition will be equivalent to requiring the 'covariant second derivative' of the curve vanishes, much like $\mathbf{r}''(s) = 0$ defines a straight line for a curve in $\mathbb{R}^n$.

Definition. *A* vector field *on S is a map $V : p \in S \to T_pS$ that assigns a vector in T_pS to each point $p \in S$.*

Remark. Examples abound in vector calculus of vector fields in $\mathbb{R}^3$- for example gravitational or electromagnetic force fields, or the velocity vector field of the flow of water. More importantly for us, the vector field $\mathbf{T}$ which assigns to each point p on a curve $\mathbf{r}$ the tangent vector at p is a vector field, known as the 'tangent vector field' to the curve.

In a local chart $\mathbf{X}(u^i)$ a vector field will be given in the form $\mathbf{V} = V^1(u^1, u^2)\,\mathbf{X}_1 + V^2(u^1, u^2)\,\mathbf{X}_2$ or more compactly, $\mathbf{V} = V^i\,\mathbf{X}_i$. The $V^i(u^1, j^2)$ are the *component functions* of the vector field in this basis. The vector field is differentiable (smooth) if each function V^1, V^2 are smooth functions of (u^1, u^2).

When differentiating a vector field, just as we do with differentiating functions, we must take a directional derivative along some curve on S (so we refer to the 'covariant derivative of $\mathbf{V}$ along the curve γ'). The idea is quite simple: we compute the derivative of the vector field, but then project down to T_pS by throwing away the component along the $\mathbf{N}$ direction. Thus we are left with a quantity which describes how the vector field changes 'on S'.

As usual we first give the chart-independent definition, and then give an expression in terms of an arbitrary chart on S.

Definition. *Let $p \in S$ and $\boldsymbol{V}$ be a differentiable vector field defined in a neighbourhood $U \subset S$ of p. Let $\alpha(t)$ be a curve on S, such that $\alpha(0) = p$ and $\alpha'(0) = T \in T_pS$ is the tangent to the curve at p. Let $\boldsymbol{V}(t) = \boldsymbol{V} \circ \alpha(t)$ be the restriction of $\boldsymbol{V}$ to α. The projection of the vector*

$$\nabla_T \mathbf{V}\bigg|_p \equiv \frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\bigg|_{t=0} \tag{53}$$

onto T_pS *is the* covariant derivative of $\mathbf{V}$ with respect to T at p. *Explicitly,*

$$\nabla_T \mathbf{V}\bigg|_p = \left[\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} - \langle\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}, \mathbf{N}\rangle\,\mathbf{N}\right]\bigg|_{t=0} \tag{54}$$

Note that the definition is independent of the curve $\alpha(t)$; it only depends on $T = \alpha'(0)$.

Remark. Many authors will use the notation $\dfrac{D\mathbf{V}}{\mathrm{d}t}$ to denote a covariant derivative along the curve. The notation used here, however, has the advantage that it emphasizes the derivative is in the direction T.

Remark. The covariant derivative can also be defined for scalar functions f on the surface S. We define it simply be the usual directional derivative along the curve:

$$\nabla_T f \equiv \frac{\mathrm{d}(f \circ \alpha(t))}{\mathrm{d}t} = \frac{\mathrm{d}f(t)}{\mathrm{d}t} \tag{55}$$

Remark. The covariant derivative ∇_T along the curve with tangent T has the properties of a derivation (i.e. it is linear and satisfies the Leibniz rule). For example, from (54)

$$\nabla_T(f\mathbf{V}) = \left[\frac{\mathrm{d}(f\mathbf{V})}{\mathrm{d}t} - \langle\frac{\mathrm{d}(f\mathbf{V})}{\mathrm{d}t}, \mathbf{N}\rangle\,\mathbf{N}\right] = f\nabla_T\mathbf{V} + (\nabla_T f)\mathbf{V} \tag{56}$$

where $\nabla_T f$ is just the usual directional derivative of f along T.

Remark. The covariant derivative along T also acts naturally on inner products:

$$\nabla_T g(\mathbf{V},\mathbf{W}) = g(\nabla_T \mathbf{V},\mathbf{W}) + g(\mathbf{V},\nabla_T \mathbf{W}) \tag{57}$$

as follows easily from (54), where $\mathbf{V},\mathbf{W}$ are two vector fields in S. To see this, remember $g(\mathbf{V},\mathbf{W})$ is just a scalar function on S and so the covariant derivative by the above remark is simply the usual directional derivative along the curve $\alpha(t)$:

$$\begin{aligned}
\nabla_T g(\mathbf{V},\mathbf{W}) &= \frac{\mathrm{d}g(\mathbf{V},\mathbf{W})}{\mathrm{d}t} = g\left(\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t},\mathbf{W}\right) + g\left(\mathbf{V},\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t}\right)\\
&= \langle\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t},\mathbf{W}\rangle + \langle\mathbf{V},\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t}\rangle\\
&= \langle\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} - \langle\frac{d\mathbf{V}}{\mathrm{d}t},\mathbf{N}\rangle\,\mathbf{N},\mathbf{W}\rangle + \langle\mathbf{V},\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t} - \langle\frac{\mathrm{d}\mathbf{W}}{\mathrm{d}t},\mathbf{N}\rangle\,\mathbf{N}\rangle\\
&= g(\nabla_T\mathbf{V},\mathbf{W}) + g(\mathbf{V},\nabla_T\mathbf{W})
\end{aligned}$$

where in the third equality we used the fact that $\langle\,\mathbf{N},\mathbf{V}\rangle = \langle\,\mathbf{N},\mathbf{W}\rangle = 0$.

Remark. The above expressions defines the covariant derivative $\nabla_T\mathbf{V}$ at $\alpha(0) = p$. More generally, if $T(t)$ is the tangent vector field to the curve $\alpha(t)$, we refer to the vector field $\nabla_T\mathbf{V} = \nabla_{\alpha'(t)}[\mathbf{V}\circ\alpha(t)]$ as the *covariant derivative of* $\mathbf{V}$ *along the curve* α .

Local Expression for the covariant derivative

We work in an arbitrary chart $\mathbf{X}(u^i)$. We are given a curve which is described by in coordinates by $\alpha(t) = (u^1(t), u^2(t))$ (so that the image on the S is $\mathbf{X}(u^i(t))$). Its tangent vector is $T = (\mathrm{d}u^1/\mathrm{d}t, \mathrm{d}u^2/\mathrm{d}t)$. Consider the vector field $\mathbf{V} = V^i\,\mathbf{X}_i$. The restriction of $\mathbf{V}$ to the curve is $\mathbf{V}(t) = V^i(t)\,\mathbf{X}_i(t)$ (the right hand side is a function of the u^i, which are themselves functions of t). Then applying the chain rule and using (14)

$$\begin{aligned}
\nabla_T\mathbf{V} &= \frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} - \langle\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t},\mathbf{N}\rangle\,\mathbf{N} &(58)\\
&= \frac{\mathrm{d}V^i}{\mathrm{d}t}\,\mathbf{X}_i + V^i(t)\frac{\mathrm{d}\,\mathbf{X}_i}{\mathrm{d}t} - \text{normal component} &(59)\\
&= \frac{\mathrm{d}u^j}{\mathrm{d}t}\left[\frac{\partial V^i}{\partial u^j}\,\mathbf{X}_i + V^i\,\mathbf{X}_{ij}\right] - \text{normal component} &(60)\\
&= \frac{\mathrm{d}u^j}{\mathrm{d}t}\left[\frac{\partial V^i}{\partial u^j}\,\mathbf{X}_i + V^i\left(\Gamma^k_{ij}\,\mathbf{X}_k + \Pi_{ij}\,\mathbf{N}\right)\right] - \text{normal component} &(61)\\
&= \frac{\mathrm{d}u^j}{\mathrm{d}t}\left[\frac{\partial V^i}{\partial u^j} + V^k\Gamma^i_{kj}\right]\,\mathbf{X}_i = T^j\left[\frac{\partial V^i}{\partial u^j} + V^k\Gamma^i_{kj}\right]\,\mathbf{X}_i &(62)
\end{aligned}$$

Thus we have the result that $\nabla_T\mathbf{V} = (\nabla_T\mathbf{V})^i\,\mathbf{X}_i$ in the coordinate basis with components

$$(\nabla_T\mathbf{V})^i = T^j\left[\frac{\partial V^i}{\partial u^j} + V^k\Gamma^i_{kj}\right] \tag{63}$$

Inspecting this equation, we see is the usual directional derivative of the components of $\mathbf{V}$ (i.e. a $T\cdot\nabla\mathbf{V}$ term) plus a second term which arises because the basis vectors themselves are varying on S. In standard vector calculus in $\mathbb{R}^3$, one usually works in the Cartesian basis i, j, k of vectors, which are obviously constant everywhere and hence when differentiating vectors (e.g. taking divergence and curl of a vector field) these additional terms vanish. However, if one were to work with the basis vectors of spherical or cylindrical coordinates, one will see these extra terms.

Parallel Transport

We can now give a definition of what it means for a vector field $\mathbf{V}$ to be 'parallel' at different points. Obviously, vectors in different vector spaces (i.e. T_pS and T_qS) cannot be compared. However, one can imagine a curve $\alpha(t)$ on S with tangent vector field T which passes through both points. The covariant derivative along the curve gives us a wa to say that the vector field is 'constant' along α.

Definition. *The vector field $\boldsymbol{V}$ is said to be parallelly transported along the curve $\alpha(t)$ if*

$$\nabla_T \mathbf{V} = 0 \tag{64}$$

for all t along the curve.

This is the generalization of the idea of a 'constant vector field ' in $\mathbb{R}^3$ to surfaces. In general however, the notion of two vectors being parallel is path dependent; it depends on the curve $\alpha(t)$ that joins the points p, q.

Remark. We could use this definition to construct a parallel vector field $\mathbf{V}$ as follows. Suppose we are given a curve $\alpha(t) \subset S$ with tangent vector field $T(t)$ and at the point p, choose a given vector $\mathbf{V}_0$ lying in T_pS. We then solve the differential equation $\nabla_T\mathbf{V} = 0$ (with the initial condition $V(0) = V_0$). The theory of ordinary differential equations implies we can always set the left-hand side of (63) to zero at least for a small interval $t \in (t_1, t_2)$. The resulting vector field $\mathbf{V}(t)$ is defined on an interval of t along the curve and is called the *parallel transport of* $\mathbf{V}_0$ along the curve α.

Proposition 9. *Suppose $\boldsymbol{V}$ and $\boldsymbol{W}$ are two vector fields are parallel vector fields along the curve $\alpha(t)$ with tangent vector field $T = \alpha'(t)$. THen the angle between them is constant on the curve. In particular the length $g(\boldsymbol{V}, \boldsymbol{V})$ is constant.*

Proof. We compute

$$\nabla_T g(\mathbf{V}, \mathbf{W}) = g(\nabla_T \mathbf{V}, \mathbf{W}) + g(\mathbf{V}, \nabla_T W) = 0 \tag{65}$$

since $\mathbf{V}, \mathbf{W}$ are parallel vector fields on the curve. □

We are now in a position to define why we mean be a geodesic on a surface. Intuitively, we want a geodesic curve to be the 'straightest' on the surface. This means that (form the point of the ambient $\mathbb{R}^3$) it bends 'just enough' to stay on S. For a straight curve in $\mathbb{R}^3$, this means in the arc length parametrization $\mathbf{r}''(s) = 0$. Alternatively we can say that the tangent vector to the curve T is constant along the curve i.e $T'(s) = 0$. For a surface, this means that the derivative of the tangent vector field is always normal (i.e. it points in the $\mathbf{N}$ direction) on the curve. This gives rise to

Definition. *A curve $\alpha(t)$ for $t \in I = (t_1, t_2)$ on a regular surface S with tangent vector field $T = \alpha'(t)$. is said to be a* geodesic *if*

$$\nabla_T T = 0 \quad \Rightarrow \quad \frac{\mathrm{d}^2\alpha(t)}{\mathrm{d}t^2} \text{ is normal to } S \tag{66}$$

for all $t \in I$, i.e. the tangent vector field T is parallel transported along the curve.

Remark. A brief comment is required here concerning parametrized curves. Even in $\mathbb{R}^3$ the curve $\mathbf{r}(t) = \mathbf{v}e^t$ where $|\mathbf{v}| = 1$, does not satisfy $\mathbf{r}''(t) = 0$, although it is a straight line; just change to the arc length parameter $s = e^t$ to see that $\kappa = 0$. So in a given parametrization, a geodesic might not satisfy the above condition. For this reason we focus on *unit speed geodesics.*

Proposition 10. *A geodesic has constant speed, i.e.* $g(T, T) = c$ *where* T *is the tangent vector field.*

Proof. This immediately follows form Proposition 9. Since $\nabla_T T = 0$, we have $g(T, T) = \langle T, T\rangle = c$ on the curve. □

Proposition 11. *A unit speed curve* $\alpha(s)$ *is a geodesic if and only if* $\nabla_T T = 0$

Proof. First suppose we have a geodesic $\alpha(s)$ with unit speed. Then by definition $\nabla_T T = 0$ where $T = \alpha'(s)$. On the other hand suppose we have a curve $\alpha(t)$ satisfying $\nabla_T T = 0$ where $T = \alpha'(t)$. Then $g(T, T) = c^2$ is a constant. So then since $\mathrm{d}s/\mathrm{d}t = |T| = \sqrt{g(T, T)} = c$ we can integrate this to find $s = ct$. Then $T = c\alpha'(s)$, so

$$\nabla_{\alpha'(s)}\alpha'(s) = \nabla_{c^{-1}T}(c^{-1}T) = c^{-2}\nabla_T T = 0$$

Another way of seeing it is to note that since by assumption

$$\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} \propto \mathbf{N}$$

then because

$$\frac{\mathrm{d}\alpha}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{1}{c}\frac{\mathrm{d}\alpha}{\mathrm{d}t} \quad \text{and} \quad \frac{\mathrm{d}^2\alpha}{\mathrm{d}s^2} = \frac{1}{c}\frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} = \frac{1}{c^2}\frac{\mathrm{d}^2\alpha}{\mathrm{d}t^2} \propto \mathbf{N}$$

Thus we see that that $\alpha(s)$ is also a geodesic. □

In a local chart $\mathbf{X}(u^i)$, the equation satisfied by a geodesic $\alpha(t)$ can be read off directly from (63) by substituting $\mathbf{V} = \alpha'(t) = T$. It is easiest to just derive it directly. So let $\alpha(t)$ be a curve with tangent $T = \alpha'(t)$. In local coordinates, the curve is expressed by $u^i(t) = (u^1(t), u^2(t))$ with tangent $\mathrm{d}u^i/\mathrm{d}t$ and hence

$$\alpha(t) = \mathbf{X}(u^i(t)) \subset S\ , \qquad T = \frac{\mathrm{d}u^i}{\mathrm{d}t}\mathbf{X}_i \tag{67}$$

and thus

$$\begin{aligned}
\nabla_T T &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}u^i}{\mathrm{d}t}\mathbf{X}_i\right) - \langle\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}u^i}{\mathrm{d}t}\mathbf{X}_i\right), \mathbf{N}\rangle\,\mathbf{N} && (68)\\
&= \frac{\mathrm{d}^2u^i}{\mathrm{d}t^2}\mathbf{X}_i + \frac{\mathrm{d}u^i}{\mathrm{d}t}\frac{\mathrm{d}u^j}{\mathrm{d}t}\mathbf{X}_{ij} - \frac{\mathrm{d}u^i}{\mathrm{d}t}\frac{\mathrm{d}u^j}{\mathrm{d}t}\langle\mathbf{X}_{ij}, \mathbf{N}\rangle\,\mathbf{N} && (69)\\
&= \frac{\mathrm{d}^2u^i}{\mathrm{d}t^2}\mathbf{X}_i + \frac{\mathrm{d}u^i}{\mathrm{d}t}\frac{\mathrm{d}u^j}{\mathrm{d}t}\Gamma^k_{ij}\,\mathbf{X}_k && (70)\\
&= \left[\frac{\mathrm{d}^2u^i}{\mathrm{d}t^2} + \frac{\mathrm{d}u^j}{\mathrm{d}t}\frac{\mathrm{d}u^k}{\mathrm{d}t}\Gamma^i_{jk}\right]\mathbf{X}_i \quad \text{exchanging dummy indices i and k} && (71)
\end{aligned}$$

Thus we have $\nabla_T T = 0$ if and only if the $u^i(t)$ satisfy the following second order ordinary differential equation:

$$\frac{\mathrm{d}^2u^i}{\mathrm{d}t^2} + \Gamma^i_{jk}\frac{\mathrm{d}u^j}{\mathrm{d}t}\frac{\mathrm{d}u^k}{\mathrm{d}t} = 0 \tag{72}$$

This is the geodesic equation in local coordinates for a curve $u^i(t)$. It is a second order ordinary differential equation. The basic existence and uniqueness theorem for ordinary differential equations imply

Proposition 12. *Given $p \in S$ and $v \in T_pS$, there exists a unique geodesic $\gamma(t)$ for $t \in (t_1, t_2)$ satisfying $\gamma(0) = p$ and $\gamma'(0) = v$.*

In other words, given an initial point and initial velocity (direction in the tangent space) there is a unique geodesic, defined in a neighbourhood of the point (it may not exist of course for all t). We now consider some important examples.

Example For the plane in Cartesian coordinates, the metric tensor is given by $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$ and so $g_{xx} = g_{yy} = 1, g_{xy} = 0$. Thus all $\Gamma^i_{jk} = 0$ and the geodesic equation reduces to $\ddot{u}^i = 0$ with $u^i = (x, y)$ and the dot refers to a derivative with respect to the parameter t. Thus the geodesics are simply the straight lines $\gamma(t) = (x(t), y(t)) = (at+c, bt+d) = (a, b)t+(c, d)$ where a, b, c, d are constants. This can be written as well in the usual form $\gamma(t) = vt + p$ where v is a constant vector and p represents the position of the curve at $t = 0$.

Geodesics on a sphere

This example is more complicated. Recall we have already shown $\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, \Gamma^\phi_{\phi\theta} = \Gamma^\phi_{\theta\phi} = \cot\theta$ are the only non-zero Christoffel symbols in the standard parametrization $\mathbf{X}(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ for which the metric tensor is given by (22). We seek a geodesic $\gamma(t) = \mathbf{X}(\theta(t), \phi(t))$ and we will restrict to unit-speed curves without loss of generality, so we require $g(\gamma'(t), \gamma'(t)) = 1$. Hence

$$g(\gamma'(t), \gamma'(t)) = \theta'(t)^2 + \sin^2\theta(t)\phi'(t)^2 = 1 \tag{73}$$

This is equivalent to choosing our parameter t to be the arclength parameter s. We have the geodesic equations

$$\phi'' + 2\cos\theta\theta'\phi' = 0 \tag{74}$$
$$\theta'' - \sin\theta\cos\theta\phi'^2 = 0 \tag{75}$$

These form a coupled set of 2nd order differential equations. An analytic solution can be found as follows. First multiply (74) by $\sin^2\theta$ to find

$$\sin^2\theta\phi'' + 2\sin\theta\cos\theta\theta'\phi' = \frac{\mathrm{d}}{\mathrm{d}t}\left[\sin^2\theta\phi'\right] = 0 \tag{76}$$

which gives us the first integral $\sin^2\theta\phi' = C$ for some constant C. Suppose first $C = 0$. Then $\phi' = 0$ so $\phi = \phi_0$ a constant. Then (75) gives $\theta'^2 = 1$ so that $\theta' = \pm 1$ and $\theta = t + t_0$. Hence our geodesic is given by the coordinate curve $\mathbf{X}(t + t_0, \phi_0)$. This curve represents a meridian (a line of longitude) on S^2 (note that it is a great circle). Next, suppose $C \neq 0$. Subbing

$$\phi'^2 = \frac{C^2}{\sin^4\theta} \tag{77}$$

into (73) gives

$$\theta'^2 = 1 - \frac{C^2}{\sin^2\theta} \tag{78}$$

To solve these two (non-linear) equations, one can eliminate the t variable, by writing $\theta(\phi(t))$ and using the Chain rule to give

$$\left(\frac{\mathrm{d}\theta}{\mathrm{d}\phi}\right)^2 = \frac{\theta'^2}{\phi'^2} = \frac{\sin^2\theta}{C^2}\left(\sin^2\theta - C^2\right) \tag{79}$$

This can be immediately integrated to give

$$\phi - \phi_0 = \pm \int \frac{C \mathrm{d}\theta}{\sin\theta\sqrt{\sin\theta - C^2}} \tag{80}$$

We will take the upper sign without loss of generality. This can be solved by setting $u = \cot\theta$ and $\mathrm{d}u = -\csc^2\theta \mathrm{d}\theta$ to give

$$\phi - \phi_0 \quad = \quad \int \frac{-C \mathrm{d}u}{\sqrt{1 - C^2 \csc^2\theta}} \tag{81}$$

since $\sin\theta\csc\theta = 1$. But $1 + \cot^2\theta = 1 + u^2 = \csc^2\theta$ so

$$\begin{aligned} \phi - \phi_0 &= \int \frac{-C \mathrm{d}u}{\sqrt{1 - C^2(1+u^2)}} = \int \frac{-\mathrm{d}u}{\sqrt{\alpha^2 - u^2}} \qquad \text{where } \alpha = \frac{\sqrt{1-C^2}}{C} && (82) \\ &= \arccos(u/\alpha) && (83) \end{aligned}$$

So the end result is that the geodesics satisfy

$$\cot\theta = \alpha\cos(\phi - \phi_0) = \alpha(\cos\phi\cos\phi_0 + \sin\phi\sin\phi_0) \tag{84}$$

and this can be rewritten

$$\cos\theta = \alpha(\cos\phi\cos\phi_0 + \sin\phi\sin\phi_0)\sin\theta \Leftrightarrow z = ax + by \tag{85}$$

for constants a, b satisfying $a^2 + b^2 = \alpha^2$ and we have used our parametrization $\mathbf{X}(\theta, \phi)$ to rewrite the equation in terms of the (x, y, z) coordinates in $\mathbb{R}^3$ which lie on the sphere. So one can easily see that the geodesics are precisely the intersection of S^2 with planes that pass through the origin $(0, 0, 0)$, i.e. the centre of the sphere - these are, of course, the great circles, as we obtained from a more geometric argument. Note that these geodesics depend on two arbitrary constants (α, ϕ_0))(or alternatively (a, b, α) with one constraint between them) which determine the initial position and velocity.

Surfaces of Revolution

Imagine one has the curve $(x, z) = (f(v), h(v))$ is a curve in the xz plane. If we rotate this surface about the $z-$axis, we obtain a surface of revolution, with parametrization

$$\mathbf{X}(u, v) = (f(v)\cos u, f(v)\sin u, h(v)) \tag{86}$$

with $0 < u < 2\pi$ and $f(v) > 0$. Further since the original curve is assumed to be regular, we take $f'^2 + h'^2 > 0$. The surface obtained is rotationally symmetric about the z axis, and this class of surfaces contains many familiar examples (for example the sphere corresponds to $(f, h) = (\sin v \cos v)$). A further simplification occurs if we choose v to be an arclength parameter on the original curve, i.e. $f'^2 + h'^2 = 1$. The basis vectors associated to this chart are

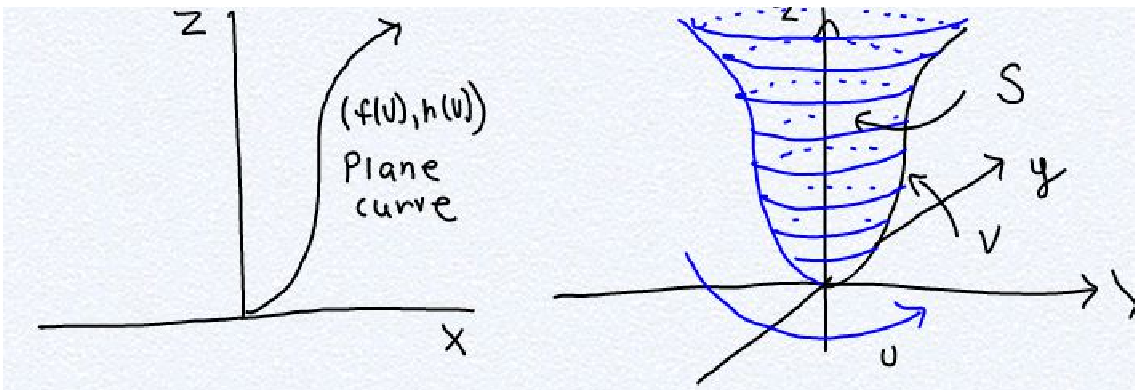

Figure 5: Surface of revolution

$$\mathbf{X}_v = (f'\cos u, f'\sin u, h') \qquad \mathbf{X}_u = (-f\sin u, f\cos u, 0) \tag{87}$$

The metric can be found to be

$$\mathrm{d}s^2 = (f'^2 + h'^2)\mathrm{d}v^2 + f^2\mathrm{d}u^2 \tag{88}$$

with Christoffel symbols

$$\Gamma^u_{uv} = \Gamma^u_{vu} = \frac{f'}{f}\ , \quad \Gamma^v_{uu} = -\frac{f'f}{f'^2 + h'^2}\ , \quad \Gamma^v_{vv} = \frac{f'f'' + h'h''}{f'^2 + h'^2} \tag{89}$$

and so the equation satisfied by the geodesics $(u(t), v(t))$ is

$$\ddot{u} + \frac{2f'}{f}\dot{u}\dot{v} = 0\ , \qquad (f'^2 + g'^2)\ddot{v} + (f'f'' + h'h'')\dot{v}^2 - ff'\dot{u}^2 = 0 \tag{90}$$

where the overdot refers to a total derivative with respect to the parameter t. We can identity some general results about these geodesics, which will agree with results we found directly for the sphere and cylinder. First consider 'meridians' of the form u =constant, so $u^i(t) = (u(t), v(t)) = (u_0, v(t)$ are geodesics. On these curves $\dot{u} = 0$ so the first geodesic equation is satisfied. Further by the unit speed condition, we must have

$$g_{ij}\frac{\mathrm{d}u^i(t)}{\mathrm{d}t}\frac{\mathrm{d}u^j(t)}{\mathrm{d}t} = g_{vv}\dot{v}^2 = \dot{v}^2(f'^2 + h'^2) = 1 \tag{91}$$

Thus differentiating this equation with respect to t we have

$$\dot{v}\ddot{v}(f'^2 + h'^2) + \dot{v}^3(f'f'' + h'h'') = 0 \tag{92}$$

and assuming $\dot{v} \neq 0$ (otherwise the curve is simply a point) so then see that the second geodesic equation (90) is automatically satisfied.

We can also identity geodesics that are *parallels* (i.e. fixed v). We consider paths with $v = v_0$ =constant, so $u^i(t) = (u(t), v_0)$. The unit speed condition is just $f^2(v_0)\dot{u}^2 = 1$. Now the first equation of (90) implies $\ddot{u} = 0$ so $\dot{u} = \alpha \neq 0$ is constant. The second requires

$$ff'\alpha^2 = 0 \tag{93}$$

and since $f(v) > 0$ we must have $f'(v_0) = 0$. So we must have that the parallel of a surface of revolution is a geodesic if it is generated by the rotation of a 'vertical' part of the original curve (i.e. the tangent line is parallel to the axis of rotation, in this case the $z-$axis. Note that at this point the normal $\mathbf{N}$ to the surface is also normal to the curve, so by Proposition 8 this is sufficient to show that the curve is a geodesic. Finally, returning to a general unit speed geodesic $(u(t), v(t))$ if we multiply the first equation

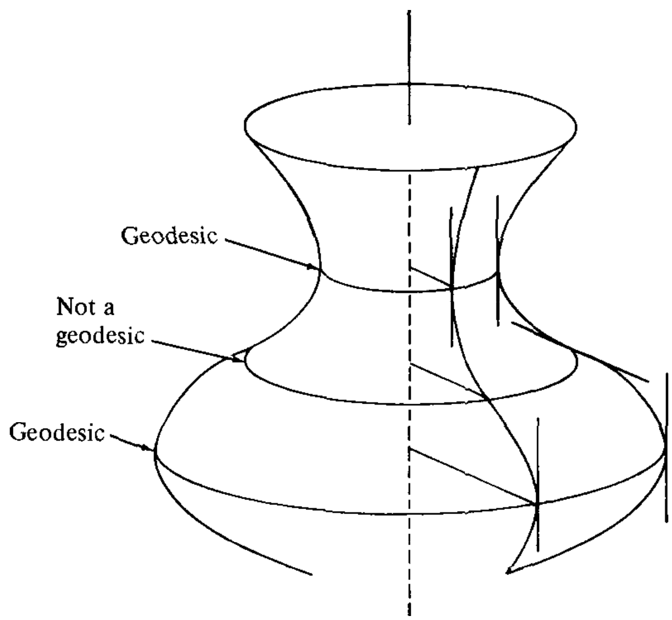

Figure 6: Parallels are geodesics when their normals agree with $\mathbf{N}$ (from do Carmo)

in (90) by f^2 we can write it as a total derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[f^2\dot{u}\right] = 0 \Rightarrow f^2\dot{u} = c \tag{94}$$

which must be satisfied by geodesics. The angle $0 \leq \theta < \pi/2$ this unit speed geodesic makes with a parallel (i.e. a curve with tangent $\mathbf{X}_u$) is given by

$$\mathbf{X}_u \cdot (\dot{u}\,\mathbf{X}_u + \dot{v}\,\mathbf{X}_v) = \dot{u}f^2 = \cos\theta|\,\mathbf{X}_u| = f\cos\theta \tag{95}$$

so that $\cos\theta = f\dot{u} = c$. Now $f^2 = r^2 = x^2 + y^2$ is the radius of the parallel at the intersection points (the distance from the point on the surface to axis of rotation) so we have

$$r\cos\theta = f^2\dot{u} = c \tag{96}$$

which is known as Clairaut's relation.

4 The Gauss-Bonnet Theorem

One of the deepest results in theory of surfaces is the Gauss-Bonnet theorem. The theorem shows that the Gaussian curvature can be used to determine the topology of a closed orientable surface (closed means the surface is compact without boundary, such as a sphere or torus). The topology of the surface is determined by a topological invariant of the surface, the Euler characteristic $\chi(S)$. The Gauss-Bonnet theorem related χ to the integral of K over S. This is quite a powerful and beautiful result that relates the *geometry* of S and its *topology*. We have developed the necessary tools to understand the geometrical part of the theorem, although we will state various topological facts without proof. These are intuitively clear to grasp, however. The Gauss-Bonnet theory is part of a large body of research in the 20th century on *index theory*, which relates certain analytical data on a manifold with topological invariants. This theory involves ideas from functional analysis, differential geometry, and topology.

Before proceeding we should record an important identity. Recall $\nabla_T\mathbf{V}$ is the vector field that represents the covariant derivative of the vector field X with respect to the curve with tangent vector field X. We may therefore differentiate it again, say with respect to another curve with tangent vector field Y to form the vector field $\nabla_Y\nabla_X\mathbf{V}$. Conversely we could perform the differentiation in the reverse order to form $\nabla_X\nabla_Y\mathbf{V}$. Let us work in a coordinate chart $\mathbf{X}(u^1,u^2)$ with basis vector fields $\mathbf{X}_i$ and our vector field $\mathbf{V} = V^i\,\mathbf{X}_i$. Let us choose our vector fields X and Y to be the basis vectors $\mathbf{X}_i$ and $\mathbf{X}_j$ (that is, we are taking covariant derivatives along the coordinate curves of u^1 and u^2). It is straightforward to verity the identity (check it)

$$(\nabla_i\nabla_j - \nabla_j\nabla_i)\mathbf{V} = [(\nabla_i\nabla_j - \nabla_j\nabla_i)\mathbf{V}]^k\,\mathbf{X}_k = R^k{}_{nij}V^n\,\mathbf{X}_k \tag{97}$$

where $R^l{}_{nij}$ are the components of the Riemann curvature tensor that we defined previously (29). Note that in this identity, *no derivatives* of $\mathbf{V}$ appear on the right hand side. In fact the operator $\nabla_i\nabla_j - \nabla_j\nabla_i$ is a linear map on $\mathbf{V}$ which can be thought of as a rotation of $\mathbf{V}$ followed by a scalar multiplication (this 'stretches' or 'contracts' the vector). In particular , if we take $i = 2$ and $j = 1$, we get

$$[(\nabla_2\nabla_1 - \nabla_1\nabla_2)\mathbf{V}]^k = R^k{}_{n21}V^n \tag{98}$$

which represents two equations, one for each value of the free index k.

Before proceeding to a proof of the Local Gauss Bonnet theorem, it is helpful to consider some simple examples, to serve as motivation. First, let us consider the plane $z = 0$, which is parameterized by $\mathbf{X}(x,y) = (x,y,0)$. Then $\mathbf{N} = (0,0,1)$ points in the $z-$direction. Let us consider a circle C on the plane, given by $\mathbf{r}(s) = (\cos s, \sin s, 0)$. We consider this circle as a simple curve (it is travelled round once). A simple calculation shows $T = (-\sin s, \cos s, 0)$ and $T' = (-\cos s, -\sin s, 0)$ and the curvature is obviously $\kappa = 1$ (it is a unit circle on the plane). We also have $\mathbf{N}\times T = (-\cos s, -\sin s, 0)$ and thus $\kappa_g = T'\cdot(\,\mathbf{N}\times T) = 1$. Thus in this case we find

$$\int_C \kappa_g\,\mathrm{d}s = \int_0^{2\pi}\mathrm{d}s = 2\pi \tag{99}$$

The Gaussian curvature of the plane is of course zero, and so we have trivially

$$\int_C \kappa_g \, \mathrm{d}s = 2\pi - \iint_A K \, \mathrm{d}A \tag{100}$$

where we have taken A to be the area enclosed within C.

Now let us take a second example. Consider a unit sphere S^2, which as we know has Gaussian curvature $K = 1$. We now take our curve C to be a great circle, for simplicity the equatorial circle, which in our standard parameterization of a sphere, is given by $(\theta(s), \phi(s)) = (\pi/2, s)$ with $s \in [0, 2\pi]$. We have already seen that C is a geodesic, and thus $\kappa_g = 0$. On the other hand, consider the area enclosed by C (we take this to be the area enclosed to the 'left' of C). This is the hemisphere region defined by $0 < \theta < \pi/2, 0 < \phi < 2\pi$. Noting that $\mathrm{d}A = \sqrt{\det g} \mathrm{d}\theta \mathrm{d}\phi$ is the area element with $\det g = \sin^2\theta$, we have

$$\iint_A K \mathrm{d}A = \int_0^{\pi/2} \mathrm{d}\theta \int_0^{2\pi} \sin\theta \, \mathrm{d}\phi = 2\pi \tag{101}$$

by a simple integration. Once again we find

$$0 = \int_C \kappa_g \, \mathrm{d}s = 2\pi - \iint_A K \, \mathrm{d}A \tag{102}$$

Hence the same relation (100) that we found for a circle on the flat plane holds, although different terms in the expression are non-vanishing. We consider a final non-trivial example in which both $K, \kappa_g \neq 0$.

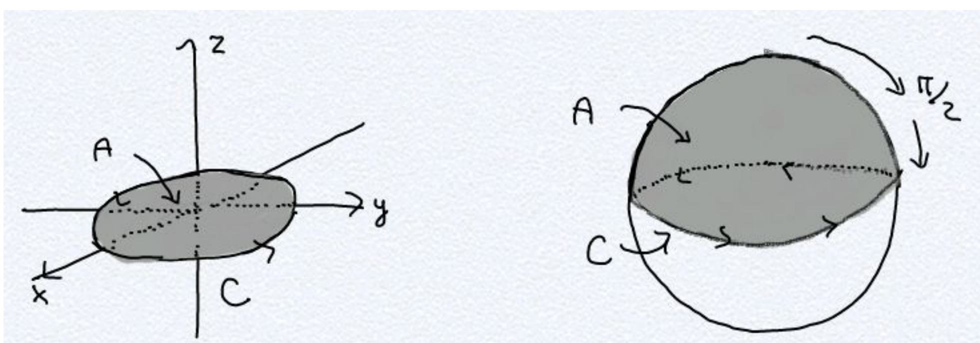

Figure 7: Simple closed curves on the plane and S^2 with enclosed area shaded.

Consider the S^2 above but now take C to be a parallel (meridian) with fixed polar angle $\theta = \theta_0$ and traversed once around. Such a curve will therefore take the form

$$\mathbf{r}(t) = (\cos t \sin\theta_0, \sin t \sin\theta_0, \cos\theta_0) \tag{103}$$

where $t \in [0, 2\pi]$. It is straightforward to check that the arclength parameter satisfies $s = \sin\theta_0 t$ and $s \in [0, 2\pi \sin\theta_0]$. A short calculation (exercise) reveals that $\kappa_g = \cot\theta_0$. It is then easily seen that

$$\int_C \kappa_g \mathrm{d}s = \cot\theta_0 \int_C \mathrm{d}s = 2\pi\cos\theta_0 \tag{104}$$

whereas

$$2\pi - \iint_A K \mathrm{d}A = 2\pi - \int_0^{\theta_0} \int_0^{2\pi} \sin\theta \, \mathrm{d}\phi = 2\pi - 2\pi(1 - \cos\theta_0) = 2\pi\cos\theta_0 \tag{105}$$

Once again the relation (100) is verified to hold. We will now show this result holds quite generally.

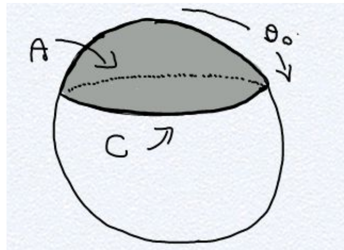

Figure 8: Parallel on S^2 with $\theta = \theta_0$.

Local Gauss Bonnet theorem

Recall that the geodesic curvature of a curve $\gamma(s)$ parameterized by arclength is defined by[b] $\kappa_g = \gamma'' \cdot (\mathbf{N} \times \gamma'(s))$. The following result is the first version of the Gauss-Bonnet theorem:

Theorem 6. *Let γ be a smooth, simple closed curve that lies in a single chart of the surface S that encloses a region R. Then*

$$\int_\gamma \kappa_g \; ds = 2\pi - \iint_R K_g \, dA \tag{106}$$

where ds is the element of arclength of γ and $dA = \sqrt{g} du^1 du^2$ is the element of area on S.

Proof. The proof is involved and can be omitted on first reading. First note the following identity from (31). For any choice of fee index (j, k)

$$g_{1m} R^m{}_{1jk} = 0 \;, \qquad g_{2m} R^m{}_{2jk} = 0 \tag{107}$$

This can easily be seen by setting $i = m$ on the right-hand side of (31). Because $g_{2m} R^m{}_{121} = K_g \det g$, we have

$$g_{21} R^1{}_{121} + g_{22} R^2{}_{121} = g_{22} R^2{}_{121} - \frac{g_{12}^2}{g_{11}} R^2{}_{121} = \frac{\det g}{g_{11}} R^2{}_{121} = K \det g \tag{108}$$

where we are using the fact

$$g_{12} R^2{}_{121} = -g_{11} R^1{}_{121} \tag{109}$$

(set $j = 2, k = 1$ in first equation in the identity above, and $g_{12} = g_{21}$). Putting this together gives the simple result

$$\frac{R^2{}_{121}}{g_{11}} = K_g \tag{110}$$

which is another way of writing the Theorem Egregium; the Gauss curvature is equal to functions built solely from the intrinsic geometry. It turns out that it is this particular relation between K_g and g (as opposed to (31)) which is needed in the proof.

The main idea of the proof is to apply Green's theorem, which is an integral relation *on the plane* between line integrals of vector fields over curves and a integral over the two-dimensional region which has the curve as its boundary. Recall that in the statement of the theorem, the curve γ is required to lie within a coordinate chart and enclose a bounded region within it. This means that rather than work on S, we can perform all calculations on the open set U on which the chart $\mathbf{X}$ is based: recall $\mathbf{X} : U \to S$

[b]Note the choice $\mathbf{N} \times \gamma'$ is chosen so that $(T, \mathbf{N} \times T, \mathbf{N})$ form a right-handed basis for vectors in $\mathbb{R}^3$

and U is a subset of the plane $\mathbb{R}^2$. Hence for curves and regions in U, we can apply Green's theorem. This will be then be equivalent to a statement for vector fields and curves on S. Let us state Green's theorem, as follows. If $\mathbf{V} = (P, Q)$ is a vector field on the plane (in the standard basis of vectors in $\mathbb{R}^2$ with coordinates (u^1, u^2)) and γ is a curve $\gamma(s) = (u^1(s), u^2(s))$ which encloses the region R, then

$$\int_\gamma \left(P\frac{\mathrm{d}u^1}{\mathrm{d}s} + Q\frac{\mathrm{d}u^2}{\mathrm{d}s} \right) \mathrm{d}s = \iint_R (Q_1 - P_2)\mathrm{d}A \tag{111}$$

where $Q_1 = \partial Q/\partial u^1$, etc. The basic strategy is to rewrite this expression in terms of intrinsic quantities on S to arrive at (133). To do this, we must find a vector field $\mathbf{V}$ on S such that the line integral over γ gives the geodesic curvature κ_g. We will now introduce an orthonormal basis for the tangent space. Let $\mathbf{e}_1 = \mathbf{X}_1/\sqrt{g_{11}}$ and $\mathbf{e}_2 = \mathbf{N} \times \mathbf{e}_1$. By construction they are orthogonal, have unit length, and are both tangent vector fields to S and form a right-handed coordinate system[c]. We can write $\mathbf{e}_2$ in terms of our usual basis:

$$\mathbf{e}_2 = \left(\frac{\mathbf{X}_1 \times \mathbf{X}_2}{|\,\mathbf{X}_1 \times \mathbf{X}_2|} \right) \times \frac{\mathbf{X}_1}{\sqrt{g_{11}}} = \frac{1}{\sqrt{g_{11}}\sqrt{\det g}} \left[-g_{12}\,\mathbf{X}_1 + g_{11}\,\mathbf{X}_2 \right] \tag{112}$$

where we used an identity to simplify the tuple cross product and note that $|\,\mathbf{X}_1 \times \mathbf{X}_2| = \sqrt{\det g}$.

Now since $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 1$, it follows that $\langle \nabla_i \mathbf{e}_1, \mathbf{e}_1 \rangle = 0$ where $i = 1, 2$ represent the usual covariant derivatives in the direction $\mathbf{X}_1$, $\mathbf{X}_2$. So we can write $\nabla_1 \mathbf{e}_1 = P\mathbf{e}_2, \nabla_2 \mathbf{e}_1 = Q\mathbf{e}_2$ for some functions P, Q on S (equivalently, on our chart they will be functions of u^1, u^2). Notice that

$$P\frac{\mathrm{d}u^1}{\mathrm{d}s} + Q\frac{\mathrm{d}u^2}{\mathrm{d}s} = \left(\frac{\mathrm{d}u^1}{\mathrm{d}s}\nabla_1 \mathbf{e}_1 + \frac{\mathrm{d}u^2}{\mathrm{d}s}\nabla_2 \mathbf{e}_2 \right) \cdot \mathbf{e}_2 = \frac{\mathrm{d}u^i}{\mathrm{d}s}\nabla_i \mathbf{e}_1 \cdot \mathbf{e}_2 = \mathbf{e}_1' \cdot \mathbf{e}_2 \tag{113}$$

and we can also write

$$\frac{\mathrm{d}u^i}{\mathrm{d}s}\nabla_i \mathbf{e}_1 \cdot \mathbf{e}_2 = \nabla_T \mathbf{e}_1 \cdot \mathbf{e}_2 \tag{114}$$

where T is the unit tangent vector to γ. Geometrically, we are measuring how the unit vector field $\mathbf{e}_1$ 'rotates' towards $\mathbf{e}_2$ as it travels around the closed path. Obviously, $\mathbf{e}_1 \cdot \mathbf{e}_2 = 0$ everywhere, but remember that each vector field is itself rotating as it moves around S (see Figure 9). Integrating this quantity over the whole curve gives the total rotation of $\mathbf{e}_1$. This is known as the *holonomy* of $\mathbf{e}_1$ along the closed path γ. Let us express T in terms of our orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2\}$:

$$T(s) = \cos\theta \mathbf{e}_1 + \sin\theta \mathbf{e}_2 \tag{115}$$

(note this always has unit length) where $\theta = \theta(s)$ is the angle that T makes with the basis vectors, measured from $\mathbf{e}_1$ towards $\mathbf{e}_2$, as it travels around the curve γ. The path is a simple closed, smooth curve, and T must return to its original value once a complete circuit is made. The basis vectors themselves are changing along γ. After one circuit has been made, θ will have changed by 2π. Now we have

$$T' = \frac{\mathrm{d}T}{\mathrm{d}s} = \cos\theta \mathbf{e}_1' - \sin\theta\theta' \mathbf{e}_1 + \cos\theta\theta' \mathbf{e}_2 + \sin\theta \mathbf{e}_2' \tag{116}$$

It follows that

$$T' \cdot \mathbf{e}_2 = \cos\theta(\mathbf{e}_1' \cdot \mathbf{e}_2) + \cos\theta\theta' \tag{117}$$

[c] this is needed to apply Green's theorem with the correct sign

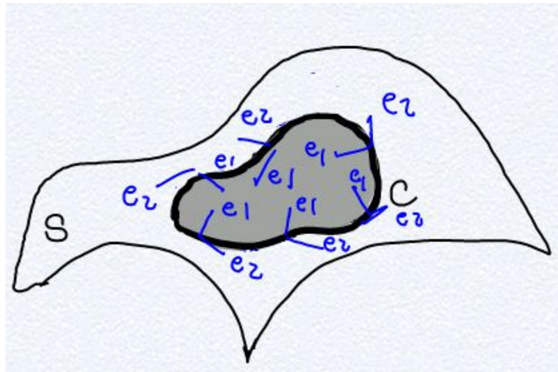

Figure 9: The orthonormal frame $(\mathbf{e}_1, \mathbf{e}_2)$ rotates as it turns around a closed curve C.

Now we have defined the geodesic curvature to by $\kappa_g = T' \cdot (\mathbf{N} \times T)$, i.e.

$$T' = \kappa_N \mathbf{N} + \kappa_g(\mathbf{N} \times T) \tag{118}$$

Inserting the above expression for T, and using the fact $\mathbf{e}_2 \times \mathbf{N} = (\mathbf{N} \times \mathbf{e}_1) \times \mathbf{N} = \mathbf{e}_1$, gives

$$T' = \kappa_N \mathbf{N} + \kappa_g(\cos\theta \mathbf{e}_2 - \sin\theta \mathbf{e}_1) \tag{119}$$

Thus

$$T' \cdot \mathbf{e}_2 = \kappa_g \cos\theta = \cos\theta \mathbf{e}_1' \cdot \mathbf{e}_2 + \cos\theta\theta' \tag{120}$$

which gives finally

$$\mathbf{e}_1' \cdot \mathbf{e}_2 = \kappa_g - \theta' \tag{121}$$

Hence the left hand side of (111) becomes, using (113)

$$\int_\gamma \mathbf{e}_1' \cdot \mathbf{e}_2 \,\mathrm{d}s = \int_\gamma (\kappa_g - \theta')\mathrm{d}s = \int_\gamma \kappa_g \,\mathrm{d}s - 2\pi \tag{122}$$

Remark. Note that (121) is equivalent to our previous definition of curvature of a plane curve. κ measures the rate at which the unit tangent vector turns relative to a fixed reference direction (the direction of the unit tangent T). In the present case, the geodesic curvature κ_g of γ lying in S measures the rate at which T is turning relative to a parallel vector field on the curve: θ' measures how fast T turns relative to $\mathbf{e}_1$, and $\nabla_T \mathbf{e}_1 \cdot \mathbf{e}_2$ measures how fast $\mathbf{e}_1$ is itself turning, and κ_g is the sum of these two contributions.

We now turn to the right hand side of (111) with the above identifications for P and Q. Note that

$$\nabla_2\nabla_1\mathbf{e}_1 = \nabla_2 (P\mathbf{e}_2) = P_2\mathbf{e}_2 + P\nabla_2(\mathbf{N} \times \mathbf{e}_1) = P_2\mathbf{e}_2 + P(\mathbf{N} \times \nabla_2\mathbf{e}_1) = PQ\,\mathbf{N} \times (\mathbf{N} \times \mathbf{e}_1). \tag{123}$$

Note that $\mathbf{e}_1 \times \nabla_2 \mathbf{N}$ is actually *normal* to the surface, and hence is thrown away when taking the covariant derivative. To see this, just note that it is obviously orthogonal to $\mathbf{e}_1$, and since $\nabla_2 \mathbf{N}$ is *tangential* to S, writing $\nabla_2 \mathbf{N} = \alpha\mathbf{e}_1 + \beta\mathbf{e}_2$, we see that $\mathbf{e}_1 \times \nabla_2 \mathbf{N} \propto \mathbf{e}_1 \times \mathbf{e}_2 \propto \mathbf{N}$. Likewise $\nabla_1\nabla_2\mathbf{e}_1 = Q_1\mathbf{e}_2 + PQ\,\mathbf{N} \times (\mathbf{e}_1 \times \mathbf{N})$. Therefore

$$(\nabla_2\nabla_1 - \nabla_1\nabla_2)\mathbf{e}_1 = (P_2 - Q_1)\mathbf{e}_2 \tag{124}$$

But from (98), we know this is equal to

$$(\nabla_2\nabla_1 - \nabla_1\nabla_2)\mathbf{e}_1 = R^l{}_{n21}e_1^n\,\mathbf{X}_l = R^l{}_{121}\frac{\mathbf{X}_l}{\sqrt{g_{11}}} \tag{125}$$

We wish to compare these two expressions. To do note that we can write the $\mathbf{X}_i$ in terms of our orthonormal frame by inverting (112):

$$\mathbf{X}_1 = \sqrt{g_{11}}\mathbf{e}_1\ , \qquad \mathbf{X}_2 = \frac{1}{\sqrt{g_{11}}}\left[g_{12}\mathbf{e}_1 + \sqrt{\det g}\mathbf{e}_2\right] \tag{126}$$

Thus

$$\begin{aligned} R^l{}_{121}\frac{\mathbf{X}_l}{\sqrt{g_{11}}} &= R^1{}_{121}\frac{\mathbf{X}_1}{\sqrt{g_{11}}} + R^2{}_{121}\frac{\mathbf{X}_2}{\sqrt{g_{11}}} = R^1{}_{121}\mathbf{e}_1 + R^2{}_{121}\frac{g_{12}\mathbf{e}_1 + \sqrt{\det g}\mathbf{e}_2}{g_{11}} \\ &= \frac{1}{g_{11}}\left[g_{1m}R^m{}_{121}\right]\mathbf{e}_1 + R^2{}_{121}\frac{\sqrt{\det g}}{g_{11}}\mathbf{e}_2 \end{aligned} \tag{127}$$

and the first term in the final line proportional to $\mathbf{e}_1$ vanishes by (107). This of course had to happen, since the right hand sides of (124) and (125) must be equal and a necessary condition is that they both are proportional to $\mathbf{e}_2$. Finally putting this altogether and using (110) to rewrite the curvature tensor in terms of the Gaussian curvature, we have

$$Q_1 - P_2 = -K_g\sqrt{\det g} \tag{128}$$

and inserting this into (111) and noting $\mathrm{d}A = \sqrt{\det g}\mathrm{d}u^1\mathrm{d}u^2$ is the area element for the surface, we arrive at

$$\int_\gamma \kappa_g\ \mathrm{d}s - 2\pi = -\iint_R K_g\ \mathrm{d}A \tag{129}$$

which establishes the result. □

We are using the fact that on a closed curve, the change in the angle θ made between the tangent to the curve and some fixed basis vector as one moves around the curve is

$$\int_\gamma \frac{\mathrm{d}\theta(s)}{\mathrm{d}s}\ \mathrm{d}s = 2\pi \tag{130}$$

where you can think of θ as the angle measured from the tangent vector to a fixed direction (say the $x-$axis in the plane). For example consider a circle, $\mathbf{r}(s) = (\sin s, -\cos s)$, $s \in [0, 2\pi]$, then $T = \mathbf{r}'(s) = (\cos s, \sin s)$ and here s is measuring precisely the angle between T and the $x-$axis. Hence $\theta = s$, and the above integral is clearly 2π. If the curve γ is pieecwise smooth rather than smooth (e.g. a curved polygon, or a triangle with curved sides), then we must replace the 2π that appears in the above theorem by 2π minus the amount the angle 'jumps' as one goes around the segment of the curve (measuring the angle from some fixed direction, say $\mathbf{X}_1$, the first basis vector associated to the chart $\mathbf{X}$) . This result is actually a topological fact which is known as the *Theorem of turning tangents*:

$$\int_\gamma \frac{\mathrm{d}\theta}{\mathrm{d}s}\mathrm{d}s = \Delta\theta = 2\pi - \sum_i \delta_i \tag{131}$$

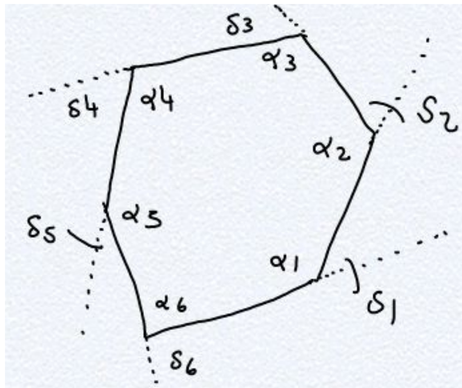

Figure 10: Internal angles α_i and external angles δ_i for a piecewise smooth curve with $n = 6$.

where δ_i is the *external angle* at each vertex (see Figure). We can write this in terms of the 'internal angles' $\alpha_i = \pi - \delta_i$

$$\Delta\theta = \sum_i \alpha_i - (n-2)\pi \tag{132}$$

where n is the number of indices, $i = 1 \ldots n$. For such a piecewise curve the local Gauss-Bonnet theorem reads

$$\int_\gamma \kappa_g \, \mathrm{d}s = \sum_i \alpha_i - (n-2)\pi - \iint_R K_g \mathrm{d}A \tag{133}$$

Let us consider an example. Suppose we consider a triangle with constructed by straight line segments on a flat plane P. Then $K_g = 0$ on P and $n = 3$ in the above formula. The line segments are geodesics and so $\kappa_g = 0$ for each line integral. Then we get $\alpha_1 + \alpha_2 + \alpha_3 = \pi$, the well known result from Euclidean geometry.

On the other hand, suppose we consider a geodesic triangle on S^2. The great circles are geodesics, so we form the sides of our triangle with segments of great circles. Again we have $\kappa_g = 0$ and $n = 3$. But now $K_g = 1$ as we have previously computed, and so

$$\text{Sum of internal angles} = \sum_i^3 \alpha_i = \pi + \iint_R \mathrm{d}A = \pi + \text{Area enclosed} \tag{134}$$

Thus on a sphere, the sum of the internal angles of a triangle is *greater* than π. On the other hand on a surface with negative Gaussian curvature $K_g < 0$, the sum of internal angles of a triangle will be *less than* π.

Example Let us take an explicit example on S^2. We know the great circles are geodesics. Consider the triangle formed by (i) a meridian starting from the North Pole (say with $\phi = 0$ in the standard chart) to the equator at $\theta = \pi/2$; (ii) the segment on the equatorial circle to the meridian $\phi = \pi/2$; and (iii) travelling back up the meridian $\phi = \pi/2$ to the North Pole. By construction each internal angle of the triangle is $\pi/2$ so the sum of the internal angles is $3\pi/2$ and $\kappa_g = 0$ everywhere on the triangle. On the unit circle, $K_g = 1$ and if we compute the area of this triangular region $0 < \phi < \pi/2, 0 < \theta < \pi/2$ (and

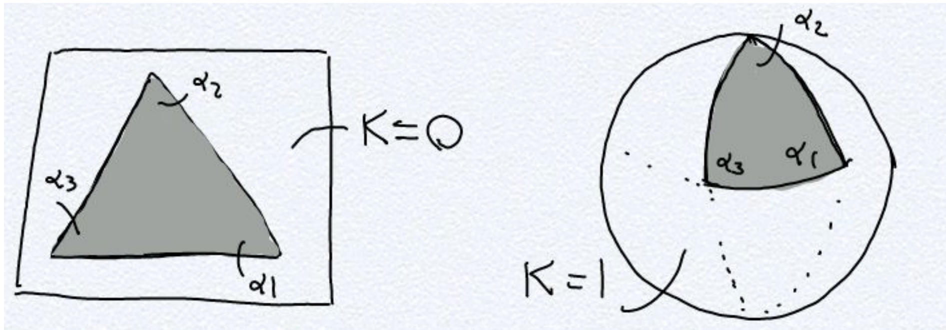

Figure 11: The sum of the internal angles on a triangle on the plane is less than that of S^2

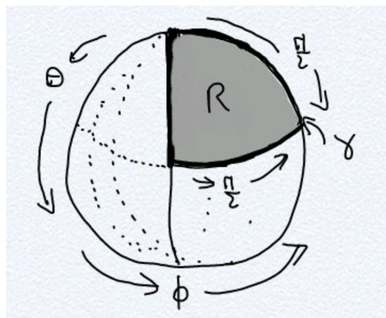

Figure 12: The sum of the internal angles of the triangle lying on S^2 is $3\pi/2$.

note the area element is $\mathrm{d}A = \sin\theta \mathrm{d}\theta \mathrm{d}\phi$), we find

$$\pi + \iint_R K \mathrm{d}A = \pi + \int_0^{\pi/2} \mathrm{d}\phi \int_0^{\pi/2} \sin\theta \mathrm{d}\theta = \pi + \frac{\pi}{2} \int_0^{\pi/2} \sin\theta \mathrm{d}\theta = \frac{3\pi}{2} \tag{135}$$

in accordance with the Local Gauss-Bonnet theorem.

Global Version of Gauss-Bonnet theorem

The above result applies to (piecewise)-smooth curves on S that lie within the image of a coordinate chart $\mathbf{X}$ for S. The *global* version of the theorem requires us to apply this result to the whole surface. We restrict firstly to *oriented* surfaces - that is, we can a continuous unit normal vector field defined everywhere on S. We only consider compact, oriented surfaces with a piecewise-smooth boundary (we are particularly interested in closed surfaces, which are compact without boundary, like a sphere or torus). It is a theorem of Rado that such a S can be *tringulated*, that is, it can be covered with a finite number of triangles.

Definition. *A triangulation is a finite family $\mathbb{T}$ of triangles $T_i, i = 1 \dots n$ such that*

1. $\cup_{i=1}^{n} T_i = S$
2. *if $T_i \cap T_j \neq \phi$ then the intersection is either a common edge of each triangle or a common vertex.*
3. *At most one edge of T_i is contained in the boundary ∂S of S.*

It is an important result of topology due to Rado that

Note that by choosing an orientation of the triangles compatibly with our choice of normal vector field for S, we get an orientation on the boundary of S. This is just the same as what happens when dealing with Green's theorem or Stokes' theorem.

Proposition 13. *Every regular surface admits a triangulation*

We now define an important topological quantity:

Definition. *Given a triangulation of S with V vertices, E edges, and F faces, the* Euler-Poincare characteristic $\chi(M, \mathcal{T}) = V - E + F$.

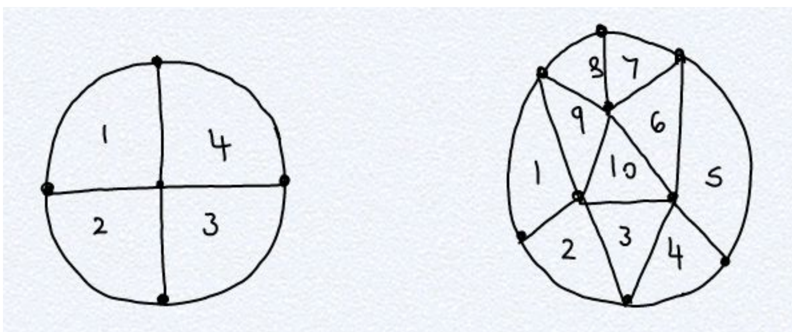

Figure 13: Two triangulations of a closed disc with $\chi = 1$

Example The Euler-Poincare characteristic can be calculated for two triangulations of a closed disc. One can imagine that each triangle is the image of a 'straight-line triangle' in $U \subset \mathbb{R}^2$ which is mapped to the above 'curved' triangle in S. In the diagram, the triangulation on the right has $F = 4, V = 5, E = 8$ so $\chi = 5-8+4 = 1$. In the triangulation on the right, $F = 10, V = 9, E = 18$ so again $\chi = 9-18+10 = 1$.

Proposition 14. *The Euler-Poincare characteristic does not depend on the triangulation of S and we can just write $\chi(S)$.*

This shows that $\chi(S)$ is a *topological invariant* of the surface (it is unchanged if we perform a continuous deformation of the surface). It turns out that it completely classifies two-dimensional compact surfaces. Note that χ is a combinatorial quantity (and defined for a much wider class of objects than smooth surfaces, like polyhedra). For example in Figure 14 a triangulation of S^2 is given demonstrating that $\chi = 2$. We could perform the same calculation on a cube, which is homeomorphic to S^2. There are 6 sides on the cube, which we turn into triangles by drawing a diagonal across each, giving 12 faces. We have $V = 8$ as there were originally 8 vertices in the cube and we have not added any; and $E = 18$ since the original cube had 12 edges, and we added 6 edges when we added the diagonals across each face of the cube. So $\chi = 8 - 18 + 12 = 2$ again. A more visual quantity is the *genus* of a connected orientable surface (see Figure 4). This is an integer which counts, informally, the number of 'handles' in the surface (intuitively it is the number of 'holes' in a closed surface). For a sphere $\chi(S^2) = 2$, for a torus $\chi(\mathbb{T}^2) = 0$ and for a $n-$torus (a sphere with n handles) one has $\chi(S) = -2(n-1)$. It is possible to prove the following:

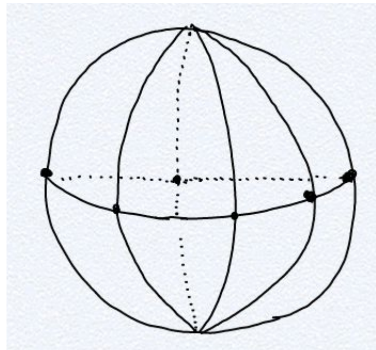

Figure 14: In this triangulation of S^2, $F = 12, V = 8, E = 18$, so $\chi(S^2) = 2$.

Theorem 7. *The genus of a closed surface is given in terms of* $\chi(S)$ *by*

$$g = \frac{2 - \chi(S)}{2} \tag{136}$$

The following important theorem classifies closed, orientable two-dimensional surfaces:

Theorem 8. *Let* S *be a closed connected surface. Then one of the values* $2, 0, -2 \ldots -2n \ldots$ *is assumed by the Euler-Poincare characteristic* $\chi(S)$. *Any other such surface* S' *with* $\chi(S') = \chi(S')$ *is homeomorphic (continuously deformable to)* S.

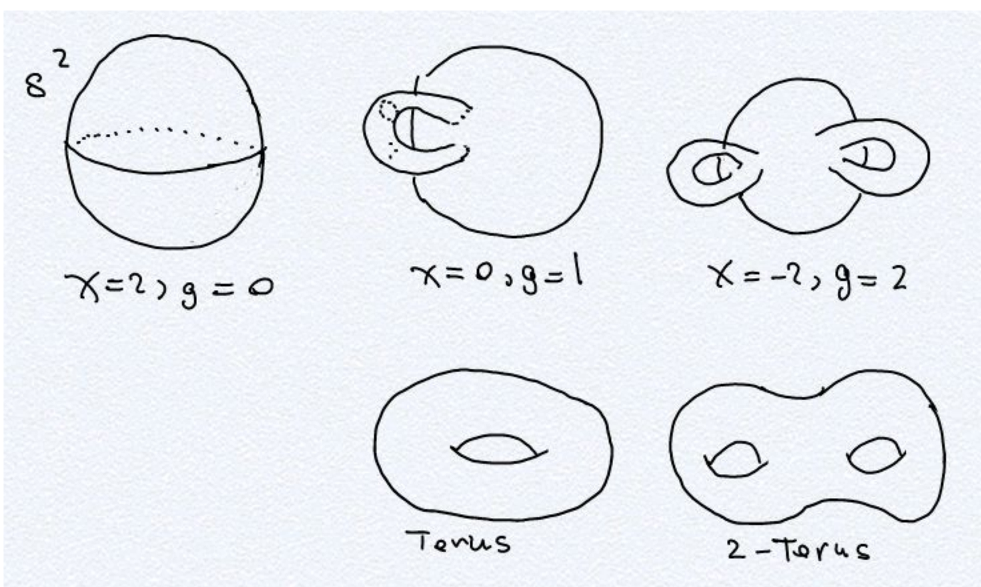

Figure 15: Classification of closed orientable surfaces

Summarizing, any closed orientable surface up to homeomorphism can be obtained by adding handles to S^2. We now state

Theorem 9. (Global Gauss Bonnet) *Let S be a compact orientable surface with piecewise smooth boundary. Then*

$$\int_{\partial S} \kappa_g ds + \iint_S K_g \; dA + \sum_{k=1}^{n} \delta_i = 2\pi\chi(S) \tag{137}$$

where δ_i are the external angles of ∂S.

An important implication of this result occurs if S is *closed* (i.e. has no boundary). Then

Theorem 10. *If S is a closed surface, then*

$$\iint_S K \; dA = 2\pi\chi(S) = 4\pi(1-g) \tag{138}$$

Proof. We take a given triangulation of S, so that each triangle lies inside the image of some chart $\mathbf{X} : U \to S$, using (133) with $n = 3$ (for triangles) and add each result to find the sum of the internal angles. The integrals of κ_g over the edges bounding each triangle must cancel because of the orientation of the edges of adjacent triangles is opposite (we are using an anticlockwise orientation for the edges of each triangle). We get a contribution of $\pi + \int_{T_i} K\mathrm{d}A$ for each triangle in $\mathcal{T}$. The total sum of internal angles is given by

$$\text{Sum of internal angles} = \pi F + \iint_S K \; \mathrm{d}A \tag{139}$$

where F is the number of faces of the triangulation (there is one face for each triangle). The sum of the internal angles may be written simply as $2\pi V$, since about each vertex in the triangulation, the internal angles must sum to 2π. This gives

$$2\pi V - \pi F = \iint_S K \; \mathrm{d}A \tag{140}$$

Finally, we must count how many edges E occur in the triangulation. Foe each face of a triangle, there are of course 3 edges. However, enumerating the total number of edges in this way will lead to an overcounting by a factor of two; indeed each edge borders two faces (this is true because S is *closed* - there are no 'boundary triangles' whose edges to not intersect another triangle). Hence $E = \frac{3}{2}F$. Thus we get

$$2\pi\chi(S) = 2\pi(V - E + F) = 2\pi V - \pi F = \iint_S K \; \mathrm{d}A \tag{141}$$

□

Remarkably, the integral of K (the total curvature of S) does not change if we smoothly deform the geometry. This would mean that the Gaussian curvature (which is a local invariant) would have to distribute itself in such a way that the integral over S does not change - further, this integral is an integer modulo 2π (!) By the classification of closed orientable surfaces, the total curvature completely determines the topology of such surfaces. For example, suppose we have an arbitrary closed orientable surface with $K > 0$ everywhere on S. Then S must be homeomorphic to S^2 because this is the only such surface with $\chi > 0$.

Example We have seen that the Gaussian curvature of the torus with outer R and inner radius r embedded in $\mathbb{R}^3$ using the standard parametrization (see Assignment 3) with $0 < \theta, \phi < 2\pi$ is

$$K = \frac{\cos\phi}{r(R + r\cos\phi)} \tag{142}$$

and the metric is given by

$$\mathrm{d}s^2 = (r\cos\phi + R)^2 \mathrm{d}\theta^2 + r^2 \mathrm{d}\phi^2 \tag{143}$$

Noting that $\sqrt{\det g} = r(r\cos\phi + R)$, it is easily seen that

$$\iint_{\mathbb{T}^2} K \ \mathrm{d}A = \int_0^{2\pi} \mathrm{d}\theta \int_0^{2\pi} \cos\phi \ \mathrm{d}\phi = 0 \tag{144}$$

which again gives $\chi(\mathbb{T}^2) = 0$.