Differential Geometry

Jay Havaldar

1 Calculus on Euclidean Spaces

From Wikipedia:

> Differential geometry is a mathematical discipline that uses the techniques of differential calculus, integral calculus, linear algebra and multilinear algebra to study problems in geometry. The theory of plane and space curves and surfaces in the three-dimensional Euclidean space formed the basis for development of differential geometry during the 18th century and the 19th century.

In short, differential geometry tries to approximate smooth objects by linear approximations. These notes assume prior knowledge of multivariable calculus and linear algebra.

Definition: A smooth real-valued function f is one where all partial derivatives and are continuous.

1.1 Tangent Vectors

The first major concept in differential geometry is that of a tangent space for a given point on a manifold. Loosely, think of manifold as a space which locally looks like Euclidean space; for example, a sphere in $\mathbb{R}^3$. The tangent space of a manifold is a generalization of the idea of a tangent plane.

Tangent space for a point on a sphere. Image from Wikipedia.

Definition: A tangent vector v_p consists of a vector v and a point of application p.

There is a natural way to add tangent vectors at a point and multiply them by scalars.

Definition: If two tangent vectors have the same vector v but different points of application, they are said to be parallel.

The best explanation I've seen of tangent vectors is by analogy with the concept of a force in physics. A force applied at different areas of a rod will have different results. For example, applied at the midpoint we have translation; applied at the endpoint we have rotation.

Definition: The tangent space $T_p(\mathbb{R}^3)$ is the space consisting of all tangent vectors applied at p.

For convenience sake, we also define the **natural frame field** on $\mathbb{R}^{\mathbb{H}}$, which consists of the vector fields U_i which are defined so that at every point, U_i returns the ith coordinate of the point.

1.2 Directional Derivatives

We also define directional derivatives, which we recall from multivariable calculus. The directional derivative in the direction of v is the rate of change along a specific direction

v.

Definition: The directional derivative of f in the direction of v at a point p is denoted $v_p[f]$.

We can calculate it using the gradient:

$$v_p[f] = \nabla f(p) \cdot v$$

And as we would expect, $v_p[f]$ is linear in both v_p and f, and the Leibniz rule (product rule) applies as well. No matter what form differentiation takes, we can pretty much always count on linearity and the Leibniz property to hold. We can generalize the idea of a directional derivative to define the operation of a vector field on a function.

Definition: Define the operation $V[f]$ of a vector field V on a function f at each point as follows: $V[f] = V(p)[f]$. In other words, at each point p, this construction takes on the value of the directional direvative of f at p in the direction of $V(p)$.

It is clear that $V[f]$ is linear both in V and f and also follows the Leibniz rule, with scalar multiplication defined as usual.

1.3 Curves in $\mathbb{R}^3$

We define a curve as a map $\alpha : I \to \mathbb{R}^3$, where I is an open interval of the real line. We can also define a reparametrization $\beta : J \to \mathbb{R}^3$ if we define a suitable function $h : J \to I$.

Indeed, we can think of $\alpha'(t)$ as a collection of tangent vectors defined at each point $\alpha(t)$, and if we do so we can compute directional derivatives. For example, at a given point $\alpha'(t)$, we can define the following directional derivative:

$$\alpha'(t)[f] = \frac{d}{dt} f(\alpha(t))$$

We can prove the above statement using the chain rule. We are applying a tangent vector $\alpha'(t)$ at a point $\alpha(t)$; in other words, at each point of a curve, we compute the directional derivative at that point in the direction of the velocity of the curve. This is a way of computing the rate of change of f "along" the curve $\alpha(t)$.

1.4 Defining 1-Forms

Recall the definition of the total derivative from multivariable calculus.

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

We're going to unpack what this really means and finally get around to a rigorous definition of the dx and dy terms we work with everywhere in calculus. Intuitively, dx and dy refer to an infinitesimal quantity in the x and y directions, respectively. We will generalize this notion.

Definition: A 1-form is a linear real-valued function on a tangent space. Thus, a 1-form is an element of the dual space of $T_p(\mathbb{R})^3$.

So we can think of 1-forms as functions which send vectors to real numbers. Naturally, we can define the action of a 1-form on a vector field; the 1-form simply is applied to the vector field at each point. We now define the most important 1-form of them all:

Definition: The differential of a function df is a 1-form which acts as follows on vectors in a tangent space:

$$df(v_p) = v_p[f]$$

So, we can think of df as a 1-form which sends each tangent vector to the directional derivative in the direction of the tangent vector. Now we can finally rigorously define dx, dy, dz.

Example Let's look at the differential dx as a 1-form. From our definitions above, the action of dx on a tangent vector is as follows (here we omit the point of application):

$$dx = v[x] = \frac{dx}{dx}v_1 + \frac{dx}{dy}v_2 + \frac{dx}{dz}v3 = v_1$$

In terms of the gradient:

$$dx = v[x] = (1, 0, 0) \cdot v = v_1$$

In other words, the differential dx is a function which sends a vector to the directional derivative of x (whose gradient is defined as the unit vector $\hat{x}$ everywhere) in the direction of v.

Earlier, I mentioned that a good way to think about tangent vectors is to imagine forces at points on a manifold. Instead, perhaps a better way to think about forces is to imagine them as 1-forms. We can do this because, as we will later see, we can construct an explicit isomorphism between vector fields and their dual forms. Instead of recording the data of where the force points, we could instead record the amount of work which is done by the force when travelling in all possible directions. Thus, a force can be represented dually by either a tangent vector or a form which returns the dot product of any vector with a fixed tangent vector.

In this way, we can naturally define a "basis" for the set of all 1-forms as dx, dy, and dz. Any 1-form ϕ can be written $\phi = \sum_i f_i dx_i$, with f_i named the Euclidean coordinate functions of ϕ. In Euclidean 3 space, a 1-form is nothing more than a construction of the form:

$$fdx + gdy + hdz$$

And it is a function from tangent vectors to real numbers. It should also be clear that the differential operator d is linear and satisfies the Leibniz and chain rules. And finally, we can define the differential df in terms of partial derivatives:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

And indeed, applying this differential at a point returns the gradient's projection along that point.

Example Let's take a look at the function $f = (x^2 - 1)y + (y^2 + 2)z$. We could use the "partial derivative" definition of f, or instead use the product rule on its factors. In this example:

$$\begin{aligned} df &= (2x\ dx)y + (x^2 - 1)\ dy + (2y\ dy)z + (y^2 + 2)\ dz \\ &= 2xy\ dx + (x^2 + 2yz + 1)\ dy + (y^2 + 2)\ dz \end{aligned}$$

And we can evaluate this differential on a tangent vector v_p as follows:

$$df(v_p) = (2p_1p_2, p_1^2 + 2p_2p_3 + 1, p_2^2 + 2) \cdot (v_1, v_2, v_3)$$

In other words, we evaluated the coordinate functions at p, and calculated the projection along v.

1.5 Exterior Derivatives and the Wedge Product

We talked about differential 1-forms, but we can also have differential k-forms for any k. However, when computing products, we have to note that they are anticommutative. For example:

$$dxdy = -dydx$$

As an immediate consequence, $dxdx = 0$. Thus, 2-forms consist of terms of the form $fdxdy + gdxdz + hdydz$, and all 3-forms are in the form of $fdxdydz$. With the wedge product, we can multiply 1-forms together (thus obtaining a 2-form) term by term.

Definition: Suppose we have a 1-form ϕ on $\mathbb{R}^3$. We define its exterior derivative as the 2-form $d\phi = \sum df_i \wedge dx_i$.

Example Define the 1-form $\phi = xydx + x^2dz$. We can compute its exterior derivative term by term:

$$\begin{aligned} d\phi &= d(xy) \wedge dx + d(x^2) \wedge dz \\ &= (y\ dx + x\ dy) \wedge dx + (2x\ dx) \wedge dz \\ &= -x\ dxdy + 2x\ dxdz \end{aligned}$$

In fact, we did not even really need to calculate the $y\ dx$ term, as we knew beforehand that it would disappear with the wedge product.

The exterior derivative, much like the differential and the directional derivative, is linear and follows a modified Leibniz rule across the wedge product:

$$d(\phi \wedge \psi) = d\phi \wedge \psi - \phi \wedge d\psi$$

Which makes sense given the nature of the wedge product.

1.6 Div, Grad and Curl

Now that we have defined wedge products and exterior derivatives, we are ready to express the 3 main operations of differential calculus succinctly in our new notation! Note that the tangent space at $p \in \mathbb{R}^3$ is isomorphic to $\mathbb{R}^3$, so we can choose to define a basis dx, dy, dz as a basis for the tangent space, with the natural isomorphism $\hat{x} \to dx$ etcetera.

Gradient First, we define a real-valued function f as a 0-form. Then we take a look at the gradient (which is a 1-form):

$$\begin{aligned} \nabla f &= \frac{\partial f}{\partial x}\ dx + \frac{\partial f}{\partial y}\ dy + \frac{\partial f}{\partial z}\ dz \\ &= df \end{aligned}$$

Curl We move on to an understanding of the curl of a vector field $F = (U, V, W)$. We can read F as a 1-form, i.e. $F = U\ dx + V\ dy + W\ dz$. Then, we arrive at the following (using subscript notation for partial derivatives for convenience):

$$\begin{aligned} dF &= dU \wedge dx + dV \wedge dy + dW \wedge dz \\ &= (U_y - V_x)\ dxdy + (U_z - W_x)\ dxdz + (V_z - W_y)\ dydz \\ &= \text{curl}F \end{aligned}$$

So, we can think of the curl as the exterior derivative of a 1-Form, read as a two form.

Divergence We can similarly define the divergence for a vector field $F = (U, V, W)$. First, we read F as a 2-form, in other words:

$$F = U\ dydz - V\ dxdz + W\ dxdy$$

Then, we take a look at the exterior derivative dF:

$$\begin{aligned} dF &= \frac{\partial U}{\partial x}\ dxdydz + \frac{\partial V}{\partial y}\ dxdydz + \frac{\partial W}{\partial z}\ dxdydz \\ &= \text{div}F\ dxdydz \end{aligned}$$

You might be wondering about the minus sign in the second term of our construction of the divergence. That's a consequence of Hodge duality, which establishes a more explicit correspondence between 1-forms and 2-forms. That's outside the scope of this text. Loosely, however, we can say this: the gradient takes a 0-form to a one form (vector field); the divergence takes a 1-form to a 2-form (vector field); and the divergence takes a 2-form to a three form (function).

1.7 Mappings

Definition: Given a function $F : \mathbb{R}^n \to \mathbb{R}^m = (f_1(p), ..., f_m(p))$, we call f_i the Euclidean coordinate functions of F.

Definition: A mapping is a differentiable function from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Using mappings, we can send curves in one space to curves in another space.

Definition: Let $\alpha : I \to \mathbb{R}^n$ be a curve. Given a mapping F, we define the composite function $\beta = F(\alpha) : I \to \mathbb{R}^m$ as the image of α under F.

We will study one specific mapping in particular:

Definition: Given a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$, we define the tangent map to be the following tangent vector at $F(p)$:

$$F_\star(v) = (v[f_1], ..., v[f_m])$$

If we fix a point p, we can see that the tangent map $F\star$ sends tangent vectors at p to tangent vectors at $F(p) \in \mathbb{R}^m$. The tangent map is linear in v, and furthermore it turns out that The tangent map is the linear transformation that best approximates F at p. As a consequence of this, we can prove that tangent maps preserve velocities of curves:

Proof: Suppose $F = (f_1, f_2, f_3)$. Then we consider the image β of a curve α under F. $F_\star$ sends a tangent vector at $\alpha(t)$ to a tangent vector at $\beta(t)$:

$$\begin{aligned} F_\star(\alpha') &= (\alpha'[f_1], \alpha'[f_2], \alpha'[f_3]) \\ &= \frac{d}{dt}(f_1(\alpha), f_2(\alpha), f_3(\alpha)) \\ &= \beta'(t) \end{aligned}$$

Finally, we consider the tangent map applied to a particular type of vector field, the coordinate maps U_j. We use the definition of tangent maps to arrive at:

$$F_\star(U_j) = (U_j[f_1], U_j[f_2], ..., U_j[f_m])$$

And recall that $U_j[f_i] = \frac{\partial f_i}{x_j}$, so we finally arrive at:

$$F_\star(U_j) = \sum_{i=1}^{m} \frac{\partial f_i}{x_j} \hat{U_j}$$

Where $\hat{U_j}$ refer to the coordinate maps of the codomain $\mathbb{R}^m$. This construction may look familiar; indeed $F_\star(U_j)$ is the jth column of the Jacobian matrix of F.

In particular, we can think of $F_\star(U_j)$ in a special case. Suppose we are working with a mapping from $\mathbb{R}^n$ to M. Then, $x_\star(U_i)$ gives us the partial derivative of each component of x in terms of u.

Geometrically, we can think of this in the following way. We create a curve by taking $x(u, v)$ and holding $v = v_0$ constant while varying u and look at the image in M. This is called a **parameter curve** since only one parameter is varying. $x_\star(U_1)$ then returns for example, the derivative of the u-parameter curve. The u-parameter curve is a linear map, so its tangent or derivative is simply itself.

Definition: A mapping is regular if at every point in the domain, the tangent map is one-to-one.

We also know that tangent maps are linear transformations, so we can use results from matrices to analyze the above statement. For example, regular maps have tangent maps with a nullspace of exactly 0.

We can do a lot just with the idea of a tangent vector, which is defined as a pair consisting of a vector and a point of application. The duals of tangent vectors are 1-forms, and we

can generalize those to k-forms equipped with a natural addition and scalar multiplication. The exterior derivative works as a generalization of differentiation which maps 1-forms to 2-forms, etc; the velocity curve works as a generalization of the derivative to curves; and finally, the tangent map is a generalization of the derivative to mappings.

2 Frame Fields

This section is mostly concerned with constructing a "natural" frame for curves in $\mathbb{R}^3$, with the Frenet apparatus, and state some fundamental theorems using those constructions. This approach will be generalized when studying surfaces by looking at curves on the surface.

Definition: A set e_1, e_2, e_3 of mutually orthogonal vectors is a frame if $e_i \cdot e_j = \delta_{ij}$.

Definition: A regular curve $\alpha(t)$ has $\alpha'(t) = 0$. Basically, it has no cusps.

2.1 Curves

Definition: The arclength s of a curve $\alpha(t)$ is computed as $s = \int |\alpha'(t)|dt$. It is not hard to prove by the chain rule that the arclength is the same independent of parametrization.

Theorem There is a reparametrization of any regular curve $\alpha(t)$ with unit speed. In fact, this is the parametrization by arclength. Take a curve $\alpha(t)$ with $\beta(s) = \alpha(t(s))$. We define $t(s)$ to be the inverse function of the arclength function $s(t)$ (which exists by the inverse function theorem since by hypothesis the curve is regular). Then, $\beta'(s) = \alpha'(t(s)) \cdot \frac{dt}{ds} = \frac{ds}{dt}(t(s)) \cdot \frac{dt}{ds}(s) = 1$.

Definition: A reparametrization $\alpha(h)$ is called orientation preserving if $h' \geq 0$ and orientation reversing if $h' \leq 0$.

Definition: A vector field Y on a curve α assigns to each t a tangent vector $Y(t)$ at $\alpha(t)$; for example, velocity is a vector field defined on a curve.

We can differentiate vector fields, and operations on vector fields can give you more vector fields (addition, cross product, scalar multiplication by a real-valued function) or even a scalar field (dot product). We define the derivative of a vector field componentwise. Analogously for the definition for tangent vectors, Y is defined to be **parallel** if its vector components everywhere are the same; it is clear that this is equivalent to the condition that $Y' = 0$.

2.2 Frenet Formulas

Finally we get around to defining a local orthonormal nautral frame defined everywhere on a curve. Take a unit speed curve, then you can define its velocity vector ,T, which is

called the unit tangent vector field. T' measures the amount the curve is turning away from its trajectory and is called *curvature*. A short summary of the Frenet apparatus follows:

- $T = \alpha'(s)$ where α is a unit speed curve is called the **tangent** vector field.
- $\kappa(s) = |T'(s)|$ is the **curvature**.
- $N = \frac{T}{\kappa}$ is a unit vector field called the (principal) **normal** vector field. It is perpendicular to T.
- $B = T \times N$ is the **binormal** vector field, completing the Frenet frame.
- $\tau = -|B'(s)|$ is the **torsion**. The reason for the minus sign will be apparent later on.

Theorem The key here is that we can represent the derivatives of T, N, B in terms of this frame. We can write the matrix equation:

$$\begin{bmatrix} T \\ N \\ B \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} T \\ N \\ B \end{bmatrix}$$

Note that the matrix is skew symmetric! The spectral theorem tells us that a real symmetric matrix has only real eigenvalues, and similarly a real skew symmetric matrix has only pure imaginary eigenvalues.

Generally, geometric problems are solved using just the Frenet formulas and looking at derivatives. We can prove that a curve is a plane curve, that it is a circle, or place a lower bound on the curvature of a curve on a sphere. Here are some easy to prove results using the Frenet frame:

- $\kappa = 0$ iff the curve is a straight line.
- $\tau = 0$ iff the curve lies on a plane.
- κ is constant and $\tau = 0$ iff the curve is a circle of radius $\frac{1}{\kappa}$.
- $\frac{\tau}{\kappa}$ constant and nonzero iff the curve is a cylindrical helix.
- κ is constant and $\tau = 0$ iff the curve is a circular helix.

2.3 Covariant Derivatives

Recall that for a function, we were able to define the directional derivative of a real-valued function f with respect to a vector v:

$$v[f] = \frac{d}{dt} f(p + tv)_{t=0}$$

We will do a similar construction with vector fields instead of real-valued functions.

Definition: Let W be a vector field, v is a tangent vector field at p. Then the **covariant derivative** of W with respect to v is the tangent vector:

$$\nabla_v W = \frac{d}{dt} W(p + tv)_{t=0}$$

The definition is almost analogous, except in this case the result is a vector and not a scalar. The covariant derivative of a vector field with respect to a vector is clearly also a tangent vector, since it depends on a point of application p. The covariant derivative measures the initial rate of change of $W(p)$ along the v direction. Let's see an example.

Example Let $W = x^2 U_1 + yz U_3$, where U_i are the Euclidean coordinate functions. Pick a tangent vector $v = (-1, 0, 2)$ and a point $p = (2, 1, 0)$. Then we have the line:

$$p + tv = (2 - t, 1, 2t)$$

This gives rise to a vector field $W(p + tv) = (2 - t)^2 U_1(p + tv) + 2t U_3(p + tv)$. Finally we can compute:

$$\nabla_v W = W(p + tv)'(0) = -4U_1(p) + 2U_3(p)$$

And we can evaluate this vector field by computing the coordinates $U_1(p)$ and $U_2(p)$, which of course are x and y unit vectors everywhere. So indeed, the result lies in the tangent space at p.

Lemma The example above clues us into a simpler formula for the covariant derivative. Suppose $W = \sum w_i U_i$, and we take a tangent vector v at p. Then we extend the notion of the directional derivative to write:

$$\begin{aligned}
\nabla_v W &= W'(p + tv)(0) \\
&= \frac{d}{dt}\Big|_{t=0} \sum w_i(p + tv) U_i(p + tv) \\
&= \sum \frac{d}{dt}\Big|_{t=0} w_i(p + tv) U_i(p + tv) \\
&= \sum v[w_i] U_i(p)
\end{aligned}$$

To apply $\nabla_v W$ to a vector field, apply v to each of its Euclidean coordinate functions (take the directional derivative of each of its components in the direction of v). We can prove various properties similar to derivative properties using this definition.

We can also define the covariant derivative of one vector field. We define $\nabla_V W$ to be at each point defined as $\nabla_{V(p)} W$. You can think of it as directional derivative of W along $V(p)$, then projecting it onto the tangent space given by V. In short, the covariant derivative measures the initial rate of change of $W(p)$ as p moves in the direction of $V(p)$. It's difficult to formulate a geometric meaning for what this is, since W has many components as does V; it is best to just think of it as some algebraically consistent way of extending the notion of the derivative consistently to vector fields. By consistently, what I mean is that we preserve linearity (with respect to both vector field arguments at a given point), as well as some suitable Leibniz-esque property for products. In the case of vector fields, we handle both scalar and vector (inner) products.

Properties of the Covariant Derivative We have the following properties, analogously with our other conceptions of the derivative.

- $\nabla_{av+bw} Y = a\nabla_v Y + b\nabla_w Y$
- $\nabla_v (aY + bZ) = a\nabla_v Y + b\nabla_v Z$
- $\nabla_v (fY) = \nabla_v f \; Y(p) + f(p)\nabla_v Y$
- $\nabla_v (Y \cdot Z) = \nabla_v Y \cdot V(p) + Y(p) \cdot \nabla_v Z$

And all these follow fairly straightforwardly from applying the definition. If instead we are taking the covariant derivative along a varying vector field, we have:

- $\nabla_{fV+gW} Y = f\nabla_V + g\nabla_W Y$
- $\nabla_V (aY + bZ) = a\nabla_V Y + b\nabla_V Z$
- $\nabla_V (fY) = \nabla_V \; fY + f\nabla_V Y$
- $\nabla_V (Y \cdot Z) = \nabla_V Y \cdot Z + Y \cdot \nabla_V Z$

Frame Fields **Definition:** Vector fields E_1, E_2, E_3 form a frame field on $\mathbb{R}^3$ if everywhere $E_i \cdot E_j = \delta_{ij}$.

With a frame field, we can define coordinate functions, i.e. for a vector field V and a frame field E_i, then we can write $V = \sum f_i E_i$ by orthonormal expansion, and $V \cdot E_i$ are called the coordinate functions. We can define dot products by multiplying componentwise with respect to these coordinate functions.

2.3.1 Connection Forms

We can, similarly to the construction with Frenet fields, express the covariant of coordinate fields in terms of the coordinate fields themselves. For example, take an arbitrary vector tangent v at a point p. Then we can indeed write:

$$\begin{bmatrix} \nabla_v E_1 \\ \nabla_v E_2 \\ \nabla_v E_3 \end{bmatrix} = \omega \cdot \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

Where $\omega_{ij}(v) = \nabla_v E_i \cdot E_j(p)$. In fact, each $\omega_{ij}(v)$ is a 1-form, called the **connection form**. Recall that a 1-form is a linear function in its arguments, which is clearly true here by the linear properties of the covariant derivative. Another way to look at it is that w_{ij} are elements of the dual space of the tangent space (hence, the cotangent space, or something). Furthermore, we can prove:

$$\begin{aligned} \nabla_v(E_i \cdot E_j) &= \nabla_v E_i \cdot E_j + \nabla_v E_j \cdot E_i \\ &= \omega_{ij}(v) + \omega_{ji}(v) \end{aligned}$$

This means that ω is a skew-symmetric matrix with 0 along the diagonals. By now, you should be reminded of the Frenet formulas, which had the same exact setup. The physical intuition is that $\omega_{ij}(v)$ is the initial rate at which E_i rotates towards E_j as p moves in the v direction.

We can now get to the connection equations of the frame field:

$$\nabla_V(E_i) = \sum_j \omega_{ij}(V) E_j$$

This bears a striking resemblance to the Frenet equations:

$$\begin{bmatrix} T \\ N \\ B \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} T \\ N \\ B \end{bmatrix}$$

Note that, the only key difference is that $\omega_{13}(V)$ is zero, since we define the derivative of the tangent vector such that it is entirely in the direction of N (in fact, that is how we pick N), so it makes sense that the rate at which T rotates towards B is always zero. The rates of change here are computed along T alone, but the coefficients can also apply to arbitrary vector fields. This is why the connection forms are 1-forms, as they are functions of vectors and not real-valued functions. In this sense, the connection forms allow us to encapsulate information about all the covariant derivatives at the same time.

We can find an explicit formula for the connection forms of an arbitrary frame field E_i. First, we express each E_i in terms of U_i, the natural frame field, to obtain a matrix:

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} = A \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix}$$

The matrix A is called the **attitude matrix** of the frame field. This matrix is orthogonal (its rows are orthogonal unit vectors). We define its differential dA to be the differential of each of the 1-forms in its entries, i.e. $dA_{ij} = d(a_{ij})$.

Theorem We can write the connection forms simply as:

$$\omega = dA \cdot A^T$$

Proof The proof is a little bit tedious. I'm warning you! First, recall that a_{ij} is the jth coordinate of E_i with respect to the natural coordinates U_i. We will then say that $a_{ij} = E_i^{(j)}$. Then, we write out an arbitrary w_{ij}:

$$\begin{aligned} \omega_{ij}(v) = (dA \cdot A^T)_{ij} &= \sum_k d(a_{ik})(v) \cdot A_{jk} \\ &= \sum_k v[a_{ik}] \cdot A_{jk} \\ &= \sum_k v[E_i^{(k)}] \cdot E_j^{(k)} \\ &= \nabla_v E_i \cdot E_j \end{aligned}$$

And we are done. This formulation will make some later proofs a lot easier.

2.4 The Structural Equations

Definition: The dual 1-forms of a frame field E_i are the 1-forms θ_i defined such that $\theta_i(v) = v \cdot E_i(p)$ for each tangent vector v.

For example, in the case of the natural frame field, $\theta_i = dx_i$ since $dx_i(v) = v_i$. Using dual 1-forms, we can write any vector field V in the tangent space in a different basis as: $V = \sum \theta_i(V) E_i$. In fact, θ_i is the dual basis of E_i.

Lemma We let θ_i be the dual forms of a frame field E_i. Then any 1-form ϕ has a unique representation $\phi = \sum \phi(E_i)\theta_i$; in this sense the dual forms form a basis for the 1-forms. This is the generalization of the earlier statement that any 1-form on $\mathbb{R}^3$ can be written in the form $f dx + g dy + h dz$.

Note that $\theta_i(U_j) = E_i \cdot U_j = A_{ij}$. So therefore, by the preceding lemma, we can concisely write the dual forms in the matrix equation:

$$\begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\end{bmatrix} = A \begin{bmatrix}dx_1\\ dx_2\\ dx_3\end{bmatrix}$$

Now we are faced with the question of finding exterior derivatives of our new objects, the connection forms and the dual forms. These are represented by Cartan's structural equations.

The Structural Equations The first structural equations are:

$$d\theta_i = \sum_j \omega_{ij} \wedge \theta_j$$

Compare to the connection equations:

$$\nabla_V(E_i) = \sum_j \omega_{ij}(V)E_j$$

And the second equations:

$$d\omega_{ij} = \sum_k \omega_{ik} \wedge \omega_{kj}$$

The first equations mirror closely the connection equations, since θ_i are the dual forms of E_i. To prove these two, first denote ξ to be the column vector consisting of the natural coordinates x_i of Euclidean space (in other words, 0-forms). Then we can write:

$$\theta = \begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\end{bmatrix} \text{ and } d\xi = \begin{bmatrix}dx_1\\ dx_2\\ dx_3\end{bmatrix}$$

So the formula for the dual forms can be written succinctly:

$$\theta = Ad\xi$$

Though matrix multiplication works as normal, recall that 1-forms have a separate multiplication operation.

Proof of First Structural Equation We have $d(d\xi) = 0$ by definition of the exterior derivative. Recall that A is orthogonal so:

$$d\theta = d(A\ d\xi) = dA \cdot d\xi + 0 = dA \cdot A^T \cdot A \cdot d\xi = \omega\theta$$

Which puts the first structural equations all in one matrix. ω and θ have 1-forms as their entries.

Proof of Second Structural Equation For functions f and g, we have:

$$d(g\ df) = -df \wedge dg$$

This is from the modified product rule for exterior derivatives. We apply this formula to ω:

$$\begin{aligned} d(\omega) &= d(dA \cdot A^T) = -dA \cdot d(A^T) \\ &= -dA \cdot A^T \cdot A \cdot (dA)^T \\ &= -\omega \cdot \omega^T \\ &= \omega^2 \end{aligned}$$

The last equality follows from the skew-symmetry of ω.

2.5 Summary

The first major idea here is that we can generalize our work with the triple of vector fields that is the Frenet frame to study arbitrary frames, and represent the covariant derivatives in terms of the frame itself with Cartan's equations. We can also exent these formulas to a "dual" notion by considering dual frames, which are not triples of vector fields but triples of cotangents, which are functionals on tangent vectors.

3 Euclidean Geometry

In this section, we discuss Euclidean geometry in more detail using the tools introduced in the previous two parts. The main ideas of this article revolve around isometries, which will stand in for a notion of congruence for geometrical objects. This will be a shorter chapter, since there's not much to do but apply definitions.

3.1 Isometries of $\mathbb{R}^3$

Definition An **isometry** (also called a rigid motion) of $\mathbb{R}^3$ is a mapping from $\mathbb{R}^3$ to itself so that $d(F(p), F(q)) = d(p, q)$ for all points $p, q \in \mathbb{R}^3$. $d(x, y)$ is Euclidean distance.

Translations and rotations are examples of isometries. As we will see, all isometries can be built from these two fundamental transformations.

Definition: An **orthogonal transformation** C in $\mathbb{R}^3$ preserves dot products, i.e. $C(x) \cdot C(y) = x \cdot y$

An example of an orthogonal transformation is rotation (easy to see if you consider the definition of dot products involving angles between vectors). It is hopefully easy to see that orthogonal transformations also then preserve norms, and therefore they are isometries.

Lemma Suppose F is an isometry such that $F(0) = 0$. Then F is orthogonal. The proof is fairly straightforward. First we show that F preserves norms:

$$d(0, F(p)) = d(F(0), F(p)) = d(0, p)$$

Since $F(p)$ is an isometry which preserves norms:

$$|F(p) - F(q)| = |p - q|$$
$$(F(p) - F(q)) \cdot (F(p) - F(q)) = (p - q) \cdot (p - q)$$

Multiplying out and cancelling the squared terms since they are simply norms:

$$F(p) \cdot F(q) = p \cdot q$$

And thus F is indeed orthogonal. It is not hard to prove that F is linear from what we have shown simply by taking dot products with a given orthonormal basis.

Theorem Let F be an isometry of $\mathbb{R}^3$. Then there is a unique translation T and a unique orthogonal transformation C so that $F = TC$.

This isn't hard to prove either. We let T^{-1} be the translation by $F(0)$. Then $T^{-1}F$ sends 0 to 0, and by our earlier lemma it is an orthogonal transformation. Proving uniqueness isn't too enlightening. This means we can write any isometry:

$$F(p) = a + Cp$$

Where C is an orthogonal matrix and a is a vector.

3.2 The Tangent Map of an Isometry

We now calculate the tangent map of an isometry. The main theorem here is that:

Theorem Let $F = TC$ be an isometry. Then:

$$F \star (v_p) = C(v)_{F(p)}$$

Recall that $F\star$ sends tangent vectors at p to tangent vectors at $F(p)$. We claim that the tangent map specifically sends any isometry to simply its orthogonal component. You can prove this from the definition:

$$\begin{aligned} F(p + tv) &= TC(p + tv) = a + C(p) + tC(v) \\ &= F(p) + tC(v) \end{aligned}$$

Taking the derivative at $t = 0$ gives us exactly the statement we need.

An immediate corollary is that isometries preserve dot products of tangent vectors, along with norms and orthogonality:

$$F_\star(v_p) \cdot F_\star(w_p) = C(v)_{F(p)} \cdot C(w)_{F(p)} = C(v) \cdot C(w) = v \cdot w$$

Where the last equality comes from the definition of orthogonal transformations as exactly those transformations which preserve dot products.

Now, we can also prove that isometries are uniquely determined by frames:

Theorem Given two frames of $\mathbb{R}^3$ e_i at p and f_i at q, there exists a unique isometry F so that $F_\star(e_i) = f_i$.

To prove this, we pick C to be the unique linear transformation from e_i to f_i; since both are orthogonal bases, it is not hard to see that C (a change of basis matrix) is indeed orthogonal. Finally, let T be the translation $T(p) = p + q - C(p)$. With these two transformations, define $F = TC$.

It is evident then that $F(p) = T(C(p)) = q$. By the earlier lemma, $F_\star(e_i) = f_{i_q}$.

We can compute this orthogonal component explicitly. It is not hard to check (just involves linear algebra) that if the attitude matrix of f_i is B, and the attitude matrix of e_i is A, then $C = B^{-1}A = B^T A$. The translation portion is as we saw earlier.

3.3 Orientation

We define the orientation of a frame to be the determinant of its attitude matrix, or equivalently the triple product of the frame's vectors. Since the attitude matrix is orthogonal, its determinant is ± 1, giving respectively **right** and **left handed** orientations.

We can also define the sign of an isometry F to be the determinant of its orthogonal component, i.e. $sgnF = \det C$. It is not hard to prove that the sign multiplies against orientation,

i.e. a positively oriented isometry preserves orientation while a negatively oriented isometry flips it; we call these **orientation-preserving** and **orientation-reversing** isometries. As a final theorem, we can check that the orientation of an isometry multiplies into cross products:

$$F_\star(v \times w) = (sgnF)(F_\star(v) \times F_\star(w))$$

3.4 Euclidean Geometry

This is more of a philosophical point than anything. Some consider to be "Euclidean geometry" all the properties which are preserved by isometries (dot products, cross products, etc). A more specific definition would be to consider properties which are preserved by isometries but not by arbitrary diffeomorphisms or mappings.

Note We define differentiation for vector fields componentwise, where the components are defined with respect to the natural coordinate functions U_i.

Velocity First, let's take a look at the velocity vector. We define $Y = \sum y_i U_i$. Then, using the definition of the tangent map, we can write:

$$\begin{aligned} F_\star(Y') &= F_\star(Y') = CY' \\ &= (CY)' = (F_\star(Y))' \end{aligned}$$

As an immediate result, isometries preserve acceleration.

Theorem: Isometries & The Frenet Frame Using the earlier proof, and the properties of the tangent map, we can prove that the Frenet apparatus is more or less preserved. Let β be a unit-speed curve in $\mathbb{R}^3$ with positive curvature, and let $\overline{\beta}$ be the image of the curve under an isometry F. Then:

- $\overline{\kappa} = \kappa$
- $\overline{T} = F_\star(T)$
- $\overline{N} = F_\star(N)$
- $\overline{B} = (sgnF)F_\star(B)$
- $\overline{\tau} = sgn(F)\tau$

The torsion and binormal formulas involve a sign because of the involvement of the cross product.

3.5 Congruence of Curves

Definition: Two curves α, β are **congruent** if there exists an isometry F such that $F(\alpha) = \beta$. Curves that are the same up to translation are called **parallel**.

This is the big theorem in this section. We hope that the "unique" identifying properties of curves relates to curvature and torsion. We first note that two curves are parallel if their velocities are everywhere parallel, by integration. We state our congruence theorem:

Theorem If $\alpha, \beta : I \to \mathbb{R}^3$ are unit speed curves such that $\kappa_\alpha = \kappa_\beta$ and $\tau_\alpha = \pm\tau_\beta$, then the two curves are congruent.

We can prove this theorem by construction. Fix an "initial point" in the interval I, and let F be the unique orientation-preserving isometry which sends the Frenet frame of α at $t = 0$ to the Frenet frame of β at $t = 0$. By our earlier theorem about isometries, it immediately follows that the new curve's Frenet apparatus (all three vectors, curvature, and torsion) is the same as that of β (if there is a sign error in torsion, apply instead an orienation-reversing isometry). It is not hard to show that this curve is parallel to β, and thus we are done since the two curves start at the same location.

4 Calculus on Surfaces

In this section, we discuss calculus on surfaces. Concretely, surfaces in $\mathbb{R}^3$ are comparable to sections of the Euclidean plane $\mathbb{R}^2$; with this correspondence, we can carry over much of the tools from calculus to surfaces (functions, vector fields, differential forms) and consider surfaces regardless of their ambient context.

4.1 Surfaces in $\mathbb{R}^3$

Definition: Suppose we have a set of points $M \subset \mathbb{R}^3$. A **coordinate patch** $x : D \subset \mathbb{R}^2 \to M$ is a one to one, smooth, regular mapping of an open set $D \subset \mathbb{R}^2$ into $\mathbb{R}^n$. A **proper patch** has a continuous inverse x^{-1}; in other words, a proper patch of M is a homemorphism from Euclidean space to a subset of M.

What we will do is construct an object as a union of (possibly overlapping) patches. Each mapping is one-to-one with a one-to-one tangent map everywhere, so we're sure this map is well-defined for our purposes. Now we can define a surface:

Definition: A **surface** is a $M \subset \mathbb{R}^3$ such that for each point $p \in M$, there is a proper patch from $\mathbb{R}^2$ to a neighborhood of p.

An important example is the graph of a differentiable function $f(x, y)$, which is evidently a surface with coordinate patch given by $x(u, v) = (u, v, f(u, v))$. We can extend this idea to prove that level sets are surfaces, given a reasonable criterion.

Theorem Let g be a differentiable real-valued function on $\mathbb{R}^3$. Let $M = \{(x, y, z) \mid g(x, y, z) = c\}$. Then, M is a surface if $dg \neq 0$ anywhere on m.

If you think about it, this is kind of a corollary of the Implicit Function Theorem (one of the only two useful theorems you learn in advanced calculus, along with the Inverse Function Theorem). We recall our definition of dg, which is a 1-form:

$$dg = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy + \frac{\partial g}{\partial z}dz$$

So the criterion that $dg \neq 0$ is equivalent to saying that the gradient of g is not zero anywhere. Let's pick a point $p = (p_1, p_2, p_3)$, and say WLOG that $g_z \neq 0$. By the implicit function theorem, there is a (unique) differentiable function h so that for all (u, v) in the neighorhood of (p_1, p_2), $g(u, v, h(u, v)) = c$. So for every point, we have constructed a proper patch $x(u, v) = (u, v, h(u, v))$; thus M is a surface.

For example, this means that a sphere is a surface, as are surfaces of revolution.

Definition: Let $x : \ D \to \mathbb{R}^3$ be a patch. Then at each point $u = (u_0, v_0)$, we define two separate curves $u \to x(u, v_0)$ and $v \to x(u_0, v)$. Their respective velocities are called **partial velocities** and are denoted x_u and x_v.

Note that x_u and x_v are tangent vectors applied at a given point. It isn't hard to see that if our patch is $x = (x_1, x_2, x_3)$, then:

$$x_u = (\frac{\partial x_1}{\partial u}, \frac{\partial x_2}{\partial u}, \frac{\partial x_3}{\partial u})$$

So this operation works very much like partial differentiation. We also define parametrization the usual way:

Definition: A regular mapping $x : \ D :\to M \subset \mathbb{R}^3$ is called a parametrization of M. With our definition of partial velocities from earlier, we can construct an easy test to see if x is a regular mapping. Let U_1, U_2, U_3 be the coordinate functions. Then we define the cross product:

$$x_u \times x_v = \begin{bmatrix} U_1 & U_2 & U_3 \\ \frac{\partial x_1}{\partial u} & \frac{\partial x_2}{\partial u} & \frac{\partial x_3}{\partial u} \\ \frac{\partial x_1}{\partial v} & \frac{\partial x_2}{\partial v} & \frac{\partial x_3}{\partial v} \end{bmatrix}$$

The last two rows of this matrix, notice, are the transpose of the Jacobian matrix. To say that a mapping is regular, then, is to say that its partial velocities are linearly independent, or equivalently that $x_u \times x_v \neq 0$ throughout D.

Now that we have defined a surface in multiple ways, we can begin the difficult task of defining differentiation.

4.2 Tangent Vectors, Vector Fields, and Directional Derivatives

From here on, we will mirror Part 1 of my notes, defining derivatives, 1-forms, and mappings. A lot of the proofs in this section are tedious and unenlightening, so I'll try my best to stick to the Cliffs Notes.

Definition: Suppose that f is a real-valued function defined on a surface M. Then f is differentiable if $f \circ x : \ D \to M$ is differentiable for any patch x of M.

Definition: Consider a function $F : S \subset \mathbb{R}^n \to M$. Then F is differentiable if for every patch $x : \ D \to M$, the function $x^{-1} \circ F$ is differentiable. We must first define $\mathcal{O} = (p \in \mathbb{R}^n \mid F(p) \in x(D))$ so that the composite function is well-defined; then, differentiability of $x^{-1}(F) : \ \mathcal{O} \to \mathbb{R}^3$ is defined in the usual way if $\mathcal{O}$ is an open set.

As a result, we can define coordinate functions of any function F with respect to a patch x. For example, look at a curve $\alpha : I \to M$. Then, $x^{-1} \circ \alpha : I \to D = (a_1, a_2)$ And so we have:

$$\alpha = x(a_1, a_2)$$

Since α is differentiable, you can then say that α_1, α_2 are differentiable functions. You might already see a problem here, which is that patches can overlap and give us disagreeing coordinate functions. We'll deal with that in a little bit.

There is a natural equivalence between our two concepts of differentiable maps. If we take a differentiable mapping $F : \ \mathbb{R}^n \to \mathbb{R}^3$ whose image lies in M, then considered as a function on manifolds, it is differentiable (as above). In other words, if the restriction of a mapping onto M is differentiable, then F works smoothly with all patches on M.

The proof is basically an exercise in set theory, so I'll spare you the details which you can read here[1] if you're curious. Also, I just realized that the term "atlas" is a pun, I guess, because in real life an atlas is a collection of maps, and a chart is a map. Wow. Anyway, it follows that since patches are differentiable mappings, they are differentiable on manifolds (meaning they agree with all other patches). More formally, we state the following corollary.

Corollary If x and y are overlapping patches for M, then $x^{-1}y$ and $y^{-1} \circ x$ are differentiable mappings defined on open sets of $\mathbb{R}^2$.

Indeed, since these composite functions are differentiable, their coordinate functions are differentiable, i.e., there exist unique differentiable functions $\bar{u}, \bar{v}$ so that:

$y = x(\bar{u}, \bar{v})$.

With this lemma, we don't have to check differentiability using all patches, only enough patches to cover M, because we can transition between different patches using a smooth

[1] http://www.maths.manchester.ac.uk/~tv/Teaching/Differentiable%20Manifolds/2014-2015/2-smooth-maps.pdf

map (albeit on a smaller domain). Now that we have differentiation, we're move on to define tangent vectors.

Definition: Let $p \in M$ be a point on a surface in $\mathbb{R}^3$. Then, a tangent vector v to $\mathbb{R}^3$ at p is **tangent to** M **at** p if v is the velocity of some curve in M. The set of all such tangent vectors is called the tangent plane of M at p, denoted $T_p(M)$.

So we now have a way of constructing tangent spaces of surfaces, where instead we only had the tangent space of $\mathbb{R}^3$. Now let's prove that a tangent plane is actually a plane (i.e. two dimensional).

Lemma Let p be a point of a surface M in $\mathbb{R}^3$, and let x be a patch so that $x(u_0, v_0) = p$. Then the partial velocities x_u, x_v at (u_0, v_0) form a basis for the tangent plane at M.

We get for free that the partial velocities are linearly independent; we must only show that they span the tangent plane. Firstly, as we saw before, x_u and x_v are velocities of curves in M through p. Suppose we have another vector v in the tangent plane; then there is some curve α with $\alpha(0) = p$ and $\alpha'(0) = v$.

Let (a_1, a_2) be the Euclidean coordinate functions for $x^{-1} \circ \alpha$. Then $\alpha = x(a_1, a_2)$. But then, by the chain rule:

$$\alpha' = x_u \frac{da_1}{dt} + x_v \frac{da_2}{dt}$$

Where x_u, x_v are evaluated at p. So we are done. It is not hard to prove that every linear combination of x_u, x_v is a tangent vector as well.

Now that we have tangent vectors, we define vector fields straightforwardly.

Definition: A Euclidean vector field Z on a surface $M \subset \mathbb{R}^3$ is a function that assigns to each $p \in M$ at tangent vector $Z(p)$ to $\mathbb{R}^3$ at p.

Note that vector fields do not necessarily give us vectors in the tangent plane. Such a vector field is instead called a **tangent vector field**. This is important because we can define things like a normal vector field. I will leave out the simple proof that for a level set which is a surface, the gradient vector field is a non-vanishing normal vector field. Finally, we are ready to describe derivatives, or how tangent vectors are applied to functions at points.

Definition: Let v be a tangent vector to M at p, and let f be a differentiable real-valued function on M. We define the derivative $v[f]$ as $\frac{d}{dt}(f\alpha(t))$ at $t = 0$ for any curve α with $\alpha'(0) = v$ and $\alpha(0) = p$.

This definition is pretty much the same as the one we had earlier for the directional derivative, except instead of taking a derivative along a line, we're taking a derivative along a curve through M at p. By the chain rule, you can see that the directional derivative does not depend on which curve α we pick. Finally, the usual properties of linearity and the Leibniz rule hold.

4.3 Differential Forms on a Surface

Forms on a manifold (as we will see later, a generalization of a surface) can tell us quite a bit about its geometry, as we will soon see. Our definitions normally.

- A 0-form is a differentiable real-valued function on M.
- A 1-form is a real-valued linear function on tangent vectors to M at p.
- A 2-form is a real-valued linear function on pairs of tangent vectors to M at p.

A 1-form can be evaluated as before on a vector field V, and 2-forms on pairs of vector fields V, W.

More concretely, a 2-form $\eta(v, w)$ is linear in both its arguments and antisymmetric, i.e. $\eta(v, w) = -\eta(w, v)$. As a result, we can determine the value of a 2-form just based on its evaluation on two linearly independent tangent vectors v and w. In fact, the result looks like the determinant:

$$\eta(av + bw, cv + dw) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \eta(v, w)$$

The only thing we're missing is a suitable definition for the wedge product of two 1-forms on a surface. We define:

Definition: If ϕ and ψ are 1-forms on M, then the wedge product $\phi \wedge \psi$ is the 2-form on M such that:

$$(\phi \wedge \psi)(v, w) = \phi(v)\psi(w) - \phi(w)\psi(v)$$

So we have the usual anticommutative property from our definition. In general, a wedge product of a p-form ξ and a q-form η has the law: $\xi \wedge \eta = (-1)^{pq}\eta \wedge \xi$. So on a surface, a minus sign only occurs when computing the wedge product of two 1-forms.

Now we move on to the exterior derivatives of forms on a surface. As before, we define the exterior derivative of a 0-form f to be the 1-form df so that $df(v) = v[f]$; we already defined directional derivatives so this makes sense. The only missing ingredient is the exterior derivative of a 1-form (which ought to be a 2-form).

Definition: Let ϕ be a 1-form on a surface M. Then the exterior derivative $d\phi$ is the 2-form such that for any patch x in M:

$$d\phi(x_u, x_v) = \frac{\partial}{\partial u}(\phi(x_v)) - \frac{\partial}{\partial v}(\phi(x_u))$$

You can check that this is indeed a 2-form. The issue here is that patches can overlap. We denote the above definition as $d_x\phi$. So our goal is to show when two patches x and y overlap at a point, then $d_x\phi = d_y\phi$.

Lemma We claim that definition above for the exterior derivative is well-defined, i.e. for any two overlapping patches x and y, $d_x\phi = d_y\phi$ on their overlap.

Consider any two linearly independent tangent vectors y_u and y_v. From the earlier discussion, we know that $d_x\phi(y_u, y_v)$ completely determines $d\phi$. We also know from earlier that we can find coordinate functions for y with respect to x, i.e. $y = x(\bar{u}, \bar{v})$. But then by the chain rule:

$$y_u = \frac{\partial \overline{u}}{\partial u} x_u + \frac{\partial \overline{v}}{\partial u} x_v$$
$$y_v = \frac{\partial \overline{u}}{\partial v} x_u + \frac{\partial \overline{v}}{\partial v} x_v$$

But now we can use the determinant fact from earlier to rewrite:

$$d_x\phi(y_u, y_v) = \det \begin{bmatrix} \frac{\partial \overline{u}}{\partial u} & \frac{\partial \overline{v}}{\partial u} \\ \frac{\partial \overline{u}}{\partial v} & \frac{\partial \overline{v}}{\partial v} \end{bmatrix} d_x\phi(x_u, x_v)$$

We can regard the matrix above as a sort of a Jacobian matrix (of the coordinate functions of y), or at least its transpose. Regardless, we denote J to be the determinant of that matrix and write succinctly:

$$d_x\phi(y_u, y_v) = J d_x\phi(x_u, x_v)$$

The rest of the proof proceeds by using the chain rule, but it is a tedious proof which I will omit here. Perhaps a more enlightening (and succinct) form of writing the exterior derivative of 1-forms on a surface is:

$$d\phi(V, W) = V[\phi(W)] - W[\phi[V]]$$

Where $\phi(W)$ is a function since ϕ is a one form, and so we can take directional derivatives. You can check that the two definitions are equivalent. The second definition shows us linearity and anti-commutativity far more clearly. Next we prove another property of exterior derivatives, which is that $d^2 = 0$. We only need to check this for 0-forms on a surface.

Theorem If f is a real-valued function on M, then $d(df) = 0$.

First, we write $\psi = df$; we claim that $d\psi = 0$. We have:

$$\psi(x_u) = df(x_u) = x_u[f] = \frac{\partial}{\partial u} f(x)$$
$$\psi(x_v) = df(x_v) = x_v[f] = \frac{\partial}{\partial v} f(x)$$

By the equality of mixed partial derivatives, you can check that the terms in $d\psi$ cancel.

Let's look at an example and examine differential forms on the plane.

Differential Forms on the Plane Define u_i to be the natural coordinate functions, U_i the natural frame field (in other words, du_i are the dual forms of U_i). Let f be a function on M. Then all 1-forms look like

$$\phi = f_1 du_1 + f_2 du_2$$

Where $\phi(U_i) = f_i$. The two forms are:

$$\eta = g du_1 du_2$$

Where $g = \eta(U_1, U_2)$. If we define $\psi = g_1 u_1 + g_2 u_2$, then we can take the wedge product:

$$\phi \wedge \psi = (f_1 g_2 - f_2 g_1) du_1 du_2$$

We can take the exterior derivative of the 0-form f:

$$df = \frac{\partial f}{\partial u_1} du_1 + \frac{\partial f}{\partial u_2} du_2$$

And the exterior derivative of the 1-form ϕ is:

$$d\phi = (\frac{\partial f_2}{\partial u_1} - \frac{\partial f_1}{\partial u_2}) du_1 du_2$$

We end this section with one last definition.

Definition: A differential form ϕ is closed if $d\phi = 0$. A differential form ϕ is exact if $\phi = d\xi$ for some ξ.

You may recognize the exact differential forms from differential equations; exact differential equations can be easily solved using Stokes Theorem.

4.4 Mappings of Surfaces

We now define mappings between surfaces in a straightforward generalization of the ideas we've seen so far for mappings.

Definition: A function $F : M \to N$ between surfaces is differentiable if for each patch x in M and y in N, the composite function $y^{-1} \circ F \circ x : D \to D'$ is differentiable. This is called a mapping between surfaces.

As before, we don't need to check every pair of patches, but only enough patches to cover the two surfaces since we can transition smoothly between patches.

We can prove, for example, that the stereographic projection is a differentiable map between surfaces. Now we define the tangent map.

Definition: Let $F : M \to N$ be a mapping between surfaces. Then, the tangent map $F_\star$ assigns to each tangent vector v to M at p a tangent vector $F_\star(v)$ to N at $F(p)$. We define $F_\star(v)$ as follows: if v is the initial velocity of a curve α in M, then $F_\star(v)$ is the initial velocity of the curve $F(\alpha)$ in N.

It is maybe not evident, but $F_\star$ is a linear map between $T_p(M)$ and $T_{F(p)}N$. By construction, this map preserves velocities of curves; notably, partial velocities get sent to partial velocities. We can rewrite succinctly for any curve α through $p \in M$:

$$F_\star(\alpha'(0)) = (F \circ \alpha)'(0)$$

We define a mapping to be regular if its tangent map (a square matrix) is one-to-one at each point; equivalently, if $F_\star$ is a linear isomorphism (invertible square matrix). A mapping with an inverse is called a **diffeomorphism**. This leads us to a special case of the inverse function theorem:

Theorem (Inverse Function Theorem) Let $F : M \to N$ be a mapping of surfaces such that $F_{\star,p}$ is a linear isomorphism for some $p \in M$. Then there is a neighborhood of p so that the restriction of F to that neighborhood yields a diffeomorphism onto a neighborhood of $F(p)$ in N.

We also have that a regular one-to-one mapping F of M onto N is a diffeomorphism, from definitions. Essentially, diffeomorphic objects are the same up to a smooth transformation.

Now, since mappings of surfaces map tangent vectors to tangent vectors in a well-defined way, we can imagine differential forms being easily transferrable between surfaces by a composition of functions. More concretely:

Definition: Let $F : M \to N$ be a mapping of surfaces, and ϕ a 1-form on N. Then, define $F^\star\phi$ to be the 1-form on M so that:

$$(F^\star\phi)(v) = \phi(F_\star v)$$

for all tangent vectors v to M.

Remark It may be useful to consider the terms "pushforward" and "pullback". A "pushforward", or differential, is the tangent map we saw earlier. It pushes "forward" because it sends tangent vectors in M to tangent vectors in N.

On the other hand, we also have $F^\star$, which pulls "back" because given a mapping F from M to N, it pulls back a differential form on N to a differential form on M, thus moving "backwards" through the map. I'll be using these terms from now on.

We can do the same for two forms:

Definition: Let $F : M \to N$ be a mapping of surfaces, and η a 2-form on N. Then, define $F_\star\eta$ to be the 2-form on M so that:

$$(F^\star\eta)(v, w) = \eta(F^\star(v), F^\star(w))$$

We will also denote $F^\star(f) = f(F)$ for 0-forms or functions on N, for consistency; $F^\star$ denotes a pullback, which make sense here as we are pulling back a function on N to a function on M. We finally end up with these elegant formulas for forms:

Theorem Let $F : M \to N$ be a mapping of surfaces, and η and ξ are forms on N. Then we have:

- $F^\star(\eta + \xi) = F^\star(\eta) + F^\star(\xi)$.
- $F^\star(\eta \wedge \xi) = F^\star(\eta) \wedge F^\star(\xi)$.
- $F^\star(d\xi) = d(F^\star\xi)$

The first two statements are pretty straightforward computations. The last one has a proof that's a little bit long, but not too important to the general understanding of the ideas, so I will omit it.

4.5 Integration on Surfaces

Recall how we defined calculus in Euclidean space; we integrate, for example, with respect to 1-forms dx and dy. But now that we have differential forms on surfaces, we can pull them back to Euclidean space and then integrate there. Thus we can formally define integration on surfaces.

For example, consider a curve on a surface $\alpha : I \to M$. Then what is the pullback of a 1-form on M? It is a 1-form on $\mathbb{R}$, so it is of the form $f(t)dt$ for some f. To find the coefficients of the ith dual form, we evaluate E_i; in this case we only need to evaluate $\alpha^\star\phi(U_i)$. So we have:

$$(\alpha^\star\phi)(U_1) = \phi(\alpha_\star(U_1))$$

Recall how we defined the tangent map:

$$F_\star(v) = \sum v[f_i]U_i$$

And in this case:

$$\alpha_\star(U_1) = U_1[\alpha] = \alpha'(t)$$

So for any 1-form ϕ on M, and a curve $\alpha(t)$, we have a pullback $\alpha^\star = \phi(\alpha'(t))dt$. This is a fairly flexible result and we can do a lot with it.

Definition: Let ϕ be a 1-form on M and let $\alpha : [a, b] \to M$ be a curve in M. Then we define the integral of ϕ over α as follows:

$$\int_\alpha \phi = \int_{[a,b]} \alpha^\star(\phi) = \int_a^b \phi(\alpha'(t))dt$$

This should look a lot like the definition of a line integral. And indeed, we can, for example calculate work.

Recall how earlier we said that we can think of a force as a 1-form. We take in every tangent vector on the curve (an infinitesimal portion of the curve), and for every such tangent vector calculate $\phi(\alpha)'$, the infinitesimal work done by the force in the direction of the velocity.

Suppose we have some vector field V, representing a force field, and consider the dual form ϕ to V, defined analogously to how we defined dual forms for frame fields in Part II. So we define $\phi(w) = w \cdot V$ at each point. Then, we can elegantly write the formula for work as a line integral:

$$W = \int_\alpha V(\alpha(t)) \cdot \alpha'(t)dt = \int_\alpha \phi(\alpha'(t))dt = \int_\alpha \phi$$

What if ϕ is an exact form, i.e. $\phi = df$ for some 0-form f? We get the corresponding generalization of the fundamental theorem of calculus.

Theorem Let f be a function on M, and let α be a curve where $\alpha(a) = p, \alpha(b) = q$. Then we claim that:

$$\int_\alpha df = f(q) - f(p)$$

As a corollary, this means that exact forms (related in physics to conservative vector fields) are independent of path, and vanish on closed paths. The proof comes from definitions:

$$\int_\alpha df = \int_a^b df(\alpha')dt$$

But we earlier defined the action of 1-forms on surfaces:

$$df(\alpha') = \alpha'[f] = \frac{d}{dt}(f \circ \alpha)$$

Where we omitted here the point of application determined by t. Returning to our integral, we have:

$$\int_\alpha df = \frac{d}{dt}(f \circ \alpha(t))dt = f(\alpha(b)) - f(\alpha(a)) = f(q) - f(p)$$

And we get the result by the fundamental theorem of calculus.

The amazing thing is that we can do something very similar if we pull back a 2-form ξ on M. Again, we know that the pullback will be of the form $h(u,v)dudv$, and we can compute h explicitly by calculating $h = x^\star\xi(U_1, U_2)$. But we have from our definitions:

$$h = (x^\star\eta)(U_1, U_2) = (\eta)(x_\star(U_1), x_\star(U_2))$$

We check that $x_\star(U_1) = \sum U_1[x_j]\bar{U_j} = x_u$, and proceed similarly for the other term; so our computation above simplifies to:

$$h = (x^\star\eta)(U_1, U_2) = \eta(x_u, x_v)$$

Now that we can pull back 2-forms, we can accordingly define double integrals over surfaces.

Definition: Let η be a 2-form on M, and let $x : R \to M$ be a map from a rectangle $[a,b] \times [c,d]$ in the plane into M. Then the integral of η over x is defined as:

$$\int\int_x \eta = \int\int_R x^\star\eta = \int_a^b \int_c^d \eta(x_u, x_v)dudv$$

The corresponding "conservation" theorem, wherein we compute the integral of an exact 2-form on M, gives us none other than Stokes' Theorem.

4.6 Theorem: Stokes' Theorem

We will compute the integral of a 2-form $\eta = d\phi$ over x as above. First, we define a boundary ∂x for x, which is a closed curve. We do this by defining its four sections:

- $\alpha(u) = x(u, c)$
- $\beta(v) = x(b, v)$
- $\gamma(u) = x(u, d)$
- $\delta(v) = x(a, v)$

Then we define the boundary as the counterclockwise traversal of these four segments, i.e. $\delta x = \alpha + \beta - (\gamma + \delta)$.

Now we can state the Stokes' Theorem in full:

$$\int\int_x d\eta = \int_{\partial x} \phi$$

Proof We first apply our definition above to rearrange the integral on the left hand side:

$$\int\int_x d\phi = \int_a^b \int_c^d d\phi(x_u, x_v) du dv$$

And from our definition of the exterior derivative of a 1-form earlier, we can rewrite:

$$\int\int_R d\phi(x_u, x_v)\ du dv = \int\int_R \left[\frac{\partial}{\partial u}\phi(x_v) - \frac{\partial}{\partial v}\phi(x_u)\right]\ du dv$$

For convenience sake, we write $f = \phi(x_u)$ and $g = \phi(x_v)$, to arrive at:

$$\int\int_R d\phi(x_u, x_v)\ du dv = \int\int_R \frac{\partial g}{\partial u}\ du dv - \int\int_R \frac{\partial f}{\partial v}\ du dv$$

Let's take a look at the integral on the left hand side. We treat it as an interated integral, i.e. define $I(v) = \int_a^b \frac{\partial g}{\partial u}\ du dv$:

$$\int\int_R \frac{\partial g}{\partial u}\ du dv = \int_c^d I(v) dv$$

Now note that for a given fixed value of v, the integrand of $I(v)$ is not a partial derivative but just a derivative. We then apply the fundamental theorem of calculus to get:

$$I(v) = \int_a^b \frac{dg}{du} du = g(b,v) - g(a,v)$$
$$\int\int_R \frac{\partial g}{\partial u} du dv = \int_c^d g(b,v)dv - \int_c^d g(a,v)dv$$

Now, we look at the left integral, remembering that we defined earlier $g = \phi(x_v)$. Using our boundary from earlier, we have that $x_v(b,v) = \beta'(v)$. So finally we write:

$$\int_c^d g(b,v)dv = \int_c^d \phi(\beta'(v))dv = \int_\beta \phi$$

Repeating this argument for the other integral above, we have:

$$\int\int_R \frac{\partial g}{\partial u} du dv = \left(\int_\beta \phi - \int_\delta \phi\right)$$

Continuing onwards, we finally arrive at:

$$\int\int_x d\phi = \left(\int_\beta \phi - \int_\delta \phi\right) - \left(\int_\gamma \phi - \int_\delta \alpha\right) = \int_{\partial x} \phi$$

And we are done. We also note that reparametrizations can be orientation-preserving or orienation reversing, which affect the line integral by a sign.

4.7 Topological Properties of Surfaces

I will assume here a basic knowledge of concepts in topology such as connectedness and compactness. We'll look at how these properties can be defined on a surface.

Definition: A surface is connected if for each $p, q \in M$, there is a curve from p to q in M.

Lemma A surface M is compact iff it can be covered by the images of a finite number of rectangles in M.

First, we prove the backwards direction. Suppose M is compact. For each $p \in M$, p lies in the image of some rectangle under a patch x. A finite number of such patches covers M, since M is compact; therefore the number of corresponding rectangles is finite.

Conversely, assume that M is covered by the images of a finite number of rectangles.

We first prove a lemma, that the image of a rectangle under a patch x is compact. Since x is differentiable, we assume it can be extended to an open set containing the rectangle.

Now, let (U_α) be an open covering of M,and $r \in D$ be a point in the rectangle. Then $x(r) \in U_r$ for some $U_r \in (U_\alpha)$. Since x is continuous, the preimage of U_r is a neighborhood N_r in the rectangle. For all r, these neighborhoods form an open covering of the rectangle. Since there are finitely many N_r, there are finitely many U_r; thus, we have constructed a finite covering of $x(R)$.

Returning to the original problem, assume that M is covered by the images of a finite number of rectangles. Suppose that we have some open covering for M. We just showed that the image of each rectangle is covered by a finite subcovering; therefore, taking a finite union of finite sets, we arrive at a finite subcovering for M.

Lemma A continuous function on a compact region in a surface M takes on a maximum somewhere in the region.

This lemma follows straightforwardly from the similar theorem that a function in a compact region in Euclidean space (a rectangle) attains a maximum.

This can show us that many objects are not compact. For example, a cylinder is not compact since the coordinate z is unbounded. Loosely, you expect a compact object to be closed with a smooth boundary.

Definition: A surface M is orientable if there exists a differentiable 2-form μ on M which does not vanish on M.

This definition makes a little more sense if we connect it to unit normal vector fields.

Proposition A surface M in $\mathbb{R}^3$ is orientable iff there is a unit vector field U on M. If M is connected, then $\pm U$ are the only two unit normal vector fields.

Proof First, we prove the backwards direction. Say that there is a unit normal vector field U on M. Then we can define:

$$\eta(v, w) = U \cdot (v \times w)$$

It is clear that this is a 2-form, since the determinant is linear in its rows, and the cross product is anti-symmetric. The 2-form is determined by its value on a linearly independent pair of vectors, in which case the three vectors U, v, w are linearly independent, and so the triple product is non-zero.

Conversely, suppose that M is orientable and there is a non-vanishing 2-form μ. Then define the following vector field:

$$Z(p) = \frac{v \times w}{\mu(v, w)}$$

Where v, w are any two linearly independent tangent vectors. Since they are two tangent vectors, their cross product is a normal vector; if you check its magnitude, it is the determinant of the coordinates of v, w, which is a multiple of $mu(v, w)$. So this vector field is normal to each tangent plane, does not vanish and does not depend on the choice of v, w. To make it a unit normal vector field, all we have to do is divide by the magnitude.

On a connected surface, a non-vanishing differential function cannot change sign (intermediate value theorem), so that up to sign, there is only one unit normal vector field.

For example, level sets have a unit normal vector field (the gradient), so they are all orientable. Similarly, parametrizations are orientable.

Definition: A closed curve α in M is homotopic to a constant if there is a rectangle R and a patch $x : R \to M$, called a **homotopy**, defined on $[a, b] \times [0, 1]$, such that $x(u, 0) = \alpha(u)$ and the other three edge curves $x(u, 1), x(b, v), x(u, 1) = \alpha(a) = \alpha(b) = p$ are constant.

Intuitively, a homotopy is a continuous deformation. This motivates our definition:

Definition: A surface M is simply connected if it is connected and every loop in M is homotopic to a constant.

This means, intuitively, that there are no "holes"; each loop in M can be shrunk down to a single point. We can come up with a fairly easy test for simple connectedness:

Lemma Let ϕ be a closed 1-form on M. If a loop α is homotopic to a constant, then $\int_\alpha \phi = 0$.

The proof is straightforward. $d\phi = 0$ by definition. By Stokes' Theorem, $\int_{\partial x} \phi = 0$ for any boundary ∂x. In particular, there is a boundary on which three edges are constant, so that $\int_{\partial x} \phi = \int_\alpha \phi = 0$.

Next, we consider that earlier we defined exact forms such that every exact form is closed (since $d^2 = 0$). But under what conditions is a closed form exact? Well, turns out, on any simply connected surface, this is true for 1-forms.

Lemma (Poincaré) This nice lemma says that on a simply connected surface, every closed 1-form is exact; so that if $d\phi = 0$, then $\phi = df$ for some f.

Proof First, we show that the integral of a closed 1-form is path independent on a simply connected surface. Suppose we have two curves α and β from p to q. Then $\alpha - \beta$ is a closed loop. Every loop on a simply connected surface is homotopic to a constant. Thus, by our lemma, $\int_{\alpha-\beta} \phi = 0$, or equivalently $\int_\alpha \phi = \int_\beta \phi$.

Now, suppose that ϕ is a 1-form on a simply connected surface. We define $f(p) = \int_\delta \phi$, where δ is any curve from a fixed point p_0 to p. From earlier, we know that f is well-defined

everywhere. We claim then, that $df = \phi$, or in other words that $df(v) = \phi(v)$ for every tangent vector at p.

Let $\alpha : [a, b] \to M$ be any curve with initial velocity $\alpha'(a) = v$ and initial position $\alpha(a) = p$. Then we can append this curve to δ to obtain a curve from p_0 to $\alpha(t)$. Let's call this curve γ. By the definition of f:

$$f(\alpha(t)) = \int_\gamma \phi = f(p) + \int_a^t \phi(\alpha'(u))du$$

Taking the derivative and applying the fundamental theorem of calculus, we have $\alpha'[f] = (f \circ \alpha)'(t) = \phi(\alpha'(t))$. In particular, at $t = 0$, we get $v[f] = \phi(v)$. Indeed, $df(v) = \phi(v)$. This construction is more or less what we would expect, and mirrors the proof of the fundamental theorem of calculus.

I leave the last two theorems in this section without proof, as we'll get to them later.

Theorem A compact surface in $\mathbb{R}^3$ is orientable. (Corollary of the Jordan Curve Theorem)

Theorem A simply connected surface is orientable.

4.8 Manifolds

Now we will generalize the basic concepts about surfaces to consider examples beyond $\mathbb{R}^3$. First we start with patches. We may have an **abstract patch**, which is a one-to-one map from $D \subset \mathbb{R}^2$ into the set M, which does not necessarily sit inside an ambient Euclidean space. By gluing together abstract patches, we get a manifold.

Definition: A surface is a set M with a collection $\mathcal{P}$ of abstract patches satisfying the following axioms. - The images of the patches $\mathcal{P}$ cover M. - Patches smoothly overlap. For any patches $x, y \in \mathcal{P}$, the composite functions $x^{-1}y, y^{-1}x$ are differentiable in the Euclidean sense, defined on open sets in $\mathbb{R}^2$.

We also wish to define a topology on the surface. So, for any patch x, the image of an open set in D is an open set in M, as well as their unions. We must add one last axiom for open sets to work nicely.

- The Hausdorff axiom. For any points $p \neq$ in M, there are disjoint patches x, y with p in $x(D)$ and q in $y(E)$, i.e. a surface is a Hausdorff space.

Mostly everything else carries over except for velocities.

Definition: Let $\alpha : I \to M$ be a curve in an abstract surface M. For each $t \in I$ the velocity vector $\alpha'(t)$ is the function so that:

$$\alpha'[f] = \frac{d}{dt}(f \circ \alpha)$$

So, $\alpha'(t)$ is a function whose domain is the set of real-valued functions on M. Finally, we generalize further to manifolds.

Definition: An n-dimensional manifold M is a set with a collection $\mathcal{P}$ of abstract patches (one to one functions $x : D \to M, D \subset \mathbb{R}^n$ open) satisfying:

- The covering property. The images of the patches in $\mathcal{P}$ cover M.
- The smooth overlap property. For any patches $x, y \in \mathcal{P}$, the composite functions $x^{-1}y, y^{-1}x$ are differentiable in the Euclidean sense, defined on open sets in $\mathbb{R}^2$.
- The Hausdorff axiom. For any points $p \neq$ in M, there are disjoint patches x, y with p in $x(D)$ and q in $y(E)$, i.e. a surface is a Hausdorff space.

Thus, an abstract surface is simply a 2-manifold.

4.9 Example

An example of a 4-dimensional manifold is the tangent bundle of a surface. For a surface M, let $T(M)$ be the set of all tangent vectors to M at all points of M.

To get patches in $T(M)$, we convert our patches on M to abstract patches. Given a patch $x : D \to M$, let $\tilde{D}$ be the open set in $\mathbb{R}^4$ consisting of all points $(p_1, p_2, p_3, p4)$ for which (p_1, p_2) inD. Then we define an abstract patch $\tilde{x} : \tilde{D} \to T(M)$ given by:

$$\tilde{x}(p_1, p_2, p_3, p_4) = p_3 x_u(p_1, p_2) + p_4 x_v(p_1, p_2)$$

It is not difficult to check that each $\tilde{x}$ is one-to-one and that the collection of patches satisfies the properties of a manifold. $T(M)$ is called the tangent bundle of M.

5 Shape Operators

We can now study the shape of surfaces through the lens of linear algebra, namely a symmetric operator called the shape operator. We draw comparison with the Frenet formulas and their generalizations using the connection forms[2], which describe curves in terms of their curvature and torsion, to develop similar machinery describing surfaces. Finally, we generalize the uniqueness theorem for curves[3] to a congruence theorem for surfaces, although that will have to wait until Part VI.

[2] /notes/2017/07/08/dffgeo2.html
[3] /notes/2017/07/10/diffgeo-notes3.html

5.1 The Shape Operator of $M \subset \mathbb{R}^3$

Recall how we defined the covariant derivative of a vector field Z with respect to a vector v, denoted as $\nabla_v Z$. By an earlier lemma, we were able to show that in Euclidean space:

$$\begin{aligned}\nabla_v Z &= \frac{d}{dt} Z(p+tv) \mid_t = 0 \\ &= \sum v[w_i] U_i\end{aligned}$$

This definition still more or less makes sense for any vector field on a surface if we pick v tangent to a surface M so that the derivative is defined. In addition to the earlier method of computation, we can also refer to the definition of derivatives on surfaces to obtain:

$$\begin{aligned}\nabla_v Z &= (Z_\alpha)'(0) \\ &= \sum v[w_i] U_i\end{aligned}$$

Where Z_α is the vector field Z restricted to a curve α through a point p where $\alpha'(0) = v$.

Using the covariant derivative, we now define the shape operator.

Note that if a surface is orientable, it has a (differentiable) unit normal vector field U everywhere. If it is connected, there are exactly two (in opposite directions). For a non-orientable surface, we still have local parameter curves, which allow us to create a local unit normal vector field. All this tells us that the following definition makes sense for surfaces.

Definition: If $p \in M$, then for each tangent vector v to M at p, define the **shape operator**:

$$S_p(v) = -\nabla_v U$$

Where U is a normal vector field to M defined in a neighborhood of p.

The shape operator tells us how the normal to the tangent plane (and thus the tangent plane itself) is moving in every direction; so it tells us the way M is curving in all directions at p.

Lemma Suppose $p \in M \subset \mathbb{R}^3$. Then S_p is a linear map from the tangent space $T_p(M)$ to itself.

The linearity of the shape operator is clear from the properties of covariant derivatives. Furthermore, since U has unit length everywhere, the derivative of $|U| = 0$ everywhere. So we have:

$$0 = v[U \cdot U] = (\nabla_v U) \cdot U(p)$$

So indeed the shape operator is orthogonal to the unit normal, and thus lies in the tangent plane.

Example For example, take a sphere of radius r. It is evident that the unit normal is pointing out radially from the center everywhere. So we have:

$$\begin{aligned} S_p(v) = -\nabla_v U &= -\frac{1}{r}\sum v[x_i]U_i \\ &= -\frac{v}{r} \end{aligned}$$

This has some geometric meaning; the shape operator simply is scalar multiplication, and this reflects in the uniformity of the sphere itself. The sphere bends in the same exact way at every point.

Lemma The shape operator is symmetric, i.e.:

$$S(v) \cdot w = S(w) \cdot v$$

This proof appears later on the chapter.

5.2 Normal Curvature

The first application of the shape operator is in the discussion of normal curvature.

Definition: Suppose u is a unit vector tangent to M at p. Then $k(u) = S(u) \cdot u$ is called the normal curvature of M in the u direction.

So what does the normal curvature mean? First, we prove a brief lemma.

Lemma Suppose α is a curve in M. Then:

$$\begin{aligned} (\alpha \cdot U) &= 0 \\ \alpha'' \cdot U + \alpha' \cdot U' &= 0 \\ \alpha'' \cdot U &= -\alpha' \cdot U' = S(\alpha') \cdot \alpha' \end{aligned}$$

First, we differentiated both sides. In the second step, we note that $S(\alpha) = -U'$, perhaps abusing notation to really get across the point that covariant derivatives are derivatives, and in a fairly straightforward way, too. In fact, using the chain rule we can define $S(\alpha') = -\nabla_{\alpha'} U = -U(\alpha(t))'(0)$.

The interpretation of this lemma is that all curves with a given velocity v have the same normal component of acceleration; namely a multiple of the normal curvature of $\hat{v}$.

So we parametrize α to be in particular a unit speed curve in M so that $\alpha'(0) = u$. Then we can compute:

$$\begin{aligned} k(u) &= S(u) \cdot u \\ &= \alpha'' \cdot U \\ &= \kappa N \cdot U \\ &= \kappa \cos(\theta) \end{aligned}$$

Where θ is the angle between the normal vector to the curve (N) and the unit normal to the surface (U), and κ is our familiar curvature. So the maximum possible value the normal curvature can take is in fact, our regular notion of curvature.

To make more geometric sense of this, we define the normal section of M through the u direction, which is the plane containing U as well as u; this plane slices out a curve σ inside of M. If we choose to make σ unit speed, then $\sigma'(0) = u$, and indeed the normal vector $N = \pm U$, the unit normal vector for the surface. So we can say that the normal curvature of M in the direction of u is just the curvature of the unit-speed curve determined by the corresponding normal section.

Definition: The maximum and minimum values of $k(u)$ at M are called the **principal curvatures** of M at p, denoted k_1, k_2. The corresponding directions are called **principal directions**.

It should maybe become clear to you now that these are precisely the eigenvalues and eigenvectors of the shape operator, for these are precisely the vectors u such that $S(u) = \pm cu$.

Definition: A point p is unmbilic if $k(u)$ is constant everywhere.

At an umbilic point, the shape operator is just multiplication on tangent vectors. So the following computations only apply to a non-umbilic point. Due to the fact that S is linear, its action is completely determined by its evaluation in principal directions. And furthermore due to the fact that S is symmetric, it has a set of orthogonal eigenvectors[4]. So for any unit vector u, we can write $u = \cos\theta e_1 + \sin\theta e_2$, where e_1, e_2 are unit vectors in the principal directions. Then, by linear algebra:

$$\begin{aligned} k(u) = S(u) \cdot u &= S(\cos\theta e_1 + \sin\theta e_2) \cdot (\cos\theta e_1 + \sin\theta e_2) \\ &= k_1 \cos^2\theta + k_2 \sin^2\theta \end{aligned}$$

Fun fact, this is called the Euler curvature formula. He really got around, didn't he.

5.3 Gaussian Curvature

The following two measures have some geometric significance, but their use won't become totally obvious until a bit later. We define:

[4] post/2017/07/02/symmetric-matrices.html

Definition: The Gaussian curvature of a surface $K = k_1 k_2$ is the product of its principal values (the determinant of the shape operator).

Definition: The Mean curvature of a surface $H = \frac{k_1+k_2}{2}$ is the mean of its principal values (half the trace of the shape operator).

Interestingly, the Gaussian curvature does not change if we change the sign of U. The sign of the Gaussian curvature can actually tell us quite a lot about the surface. For example:

- If $K > 0$, the principal directions have the same signs, so M is bending away from the tangent plane in all directions (like a bowl), though it could be upwards or downwards.
- If $K < 0$, the principal directions have opposite signs, so M is "saddle"-shaped.
- If $K = 0$, one or both of the principal values is zero. If it's just one, M is shaped like a "U". If it's both, then M looks like a plane near the point.

Using K and H, you can determine computational formulas for k_1, k_2. I won't get too much into those, as they all really have the same upshot. Let's take for example the algebraic identity:

$$k_1, k_2 = H \pm \sqrt{H^2 - K}$$

An immediate corollary of this is that k_1, k_2 are continuous, and away from umbilic points where the quadratic discriminant is zero, they are also differentiable.

Definition: A surface M is **flat** if $K = 0$, and **minimal** if $H = 0$.

5.4 Computational Techniques

Now we get into the every-day process of computing shape operators. We first define for any patch x of a surface M:

$$\begin{aligned} E &= x_u \cdot x_u \\ F &= x_u \cdot x_v \\ G &= x_v \cdot x_v \end{aligned}$$

These tell us more or less what the inner product looks like on $T_p(M)$, since we know that x_u, x_v form a basis for the tangent space at a point. For example, if $v = ax_u + bx_v$ and $w = cx_u + dx_v$, then we can easily check that:

$$v \cdot w = Eac + F(ac + bd) + G(bd)$$

Furthermore, $|x_u \times x_v|^2 = EG - F^2$ by distributivity. You may recognize this as a quadratic form.

We also generally compute the unit normal vector. Luckily, for a patch x, we can always use:

$$U = \frac{x_u \times x_v}{|x_u \times x_v|}$$

And finally, we can define the shape operator just by looking at the following values:

$$\begin{aligned} L &= S(x_u) \cdot x_u \\ M &= S(x_u) \cdot x_v = S(x_v) \cdot x_u \\ N &= S(x_v) \cdot x_v \end{aligned}$$

Similarly to the earlier case, these coefficients allow us to compute $S(v) \cdot w$ where v, w are expressed with respect to the x_u, x_v basis, and indeed we get a quadratic form again. As promised before, we have essentially proven that the shape operator is symmetric. So how do we compute L, M, N? Well, by the Leibniz rule we have:

$$\begin{aligned} 0 &= \frac{\partial}{\partial v}(U \cdot x_u) \\ 0 &= \nabla_v U \cdot x_u + U \cdot x_{uv} \\ S(x_u) \cdot x_v &= U \cdot x_{uv} \end{aligned}$$

So this gives us easier formulas to compute L, M, N:

$$\begin{aligned} L &= S(x_u) \cdot x_u = U \cdot x_{uu} \\ M &= S(x_u) \cdot x_v = U \cdot x_{uv} \\ N &= S(x_v) \cdot x_v = U \cdot x_{vv} \end{aligned}$$

So we merely need to compute the normal components of the second derivatives of x.

Using vector formulas, we can get the following convenient formula for the Gaussian curvature:

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

5.5 Special Curves in a Surface

Now we look at some important classes of curves on a surface, the principal, normal, and asymptotic curves.

Definition: A regular curve α is a principal curve if α' always points in a principal direction.

Lemma If α is a regular curve, and U is a normal vector field, then α is a principal curve iff α' is collinear with U' at each point. Furthermore, the principal curvature of M in the direction of α' is $k = \frac{\alpha'' \cdot U}{\alpha' \cdot \alpha'}$.

The first part of the theorem we get basically for free as a consequence of how we defined principal curves (they maximize the covariant derivative of U). The second theorem follows from a few steps of vector algebra:

$$\begin{aligned} k_1 &= k(\frac{\alpha'}{|\alpha'|}) \\ &= S(\frac{\alpha'}{|\alpha'|}) \cdot \frac{\alpha'}{|\alpha'|} \\ &= \frac{S(\alpha') \cdot \alpha'}{\alpha' \cdot \alpha'} \\ &= \frac{\alpha'' \cdot U}{\alpha' \cdot \alpha'} \end{aligned}$$

The second equality comes from the linearity of the shape operator. The final equality comes from a lemma in the normal curvature section of this chapter.

Lemma Let α be a curve cut from a surface M by a plane P. If the angle between M and P is constant along α, then α is a principal curve of M.

Let U, V be the normal vector fields to M, P respectively. Since the angle between them is constant, we have:

$$0 = (U \cdot V)' = U \cdot V' + U' \cdot V = U' \cdot V$$

Where the last equality follows from the fact that V is constant. So we know that U' is orthogonal to U and to V. Similarly, α' is orthogonal to U since it is a tangent vector, and orthogonal to V since it lies on the plane P. So the two are collinear. By our earlier lemma, α is a principal curve. We omit the case where $U = \pm V$, in which case α is trivially principal.

The geometric meaning of this theorem is that for a surface of revolution, the meridians and parallels (which are cut from the surface with the "right" planes) are principal curves.

So we covered curves for which the normal curvature is k_1 or k_2. On the other extreme, what about points where the normal curvature $k = 0$?

Definition: A direction in which the normal curvature is zero is called **asymptotic**. An asymptotic curve α is one where α' always points in an asymptotic direction.

The asymptotic directions at a point are the ones where M does not bend away from the tangent plane at all. Let's take a look at the possibilities depending on the Gaussian curvature.

- $K(p) > 0$. In this case k_1, k_2 share the same sign. By the intermediate value theorem, there is no asymptotic direction.
- $K(p) < 0$. There are exactly two asymptotic directions at p. They are bisected by the principal curves. From Euler's curvature formula, it should be easy to see that $\tan^2\theta = -\frac{k_1}{k_2}$ defines the angle between the asymptotic and a principal direction.
- $K(p) = 0$. There is either one asymptotic (also principal) direction; or p is a planar point, and all directions are asymptotic trivially.

As a quick corollary, we can check the following fact. Suppose α is on M; then $U \cdot \alpha' = 0$. By the product rule, this means that $U' \cdot \alpha' = -U \cdot \alpha''$. Thus, a curve is asymptotic if and only if its acceleration has no normal component.

By the Euler curvature formula, we also have an important result for minimal surfaces. We find that minimal surfaces have two orthogonal asymptotic curves at each point which is non-planar.

Finally, we discuss special curves called geodesics.

Definition: A curve $\alpha \in M$ is **geodesic** if α'' is always normal to M.

So, whereas an asymptotic curve has no normal component of acceleration, a geodesic has only a normal component of acceleration. For inhabitants of the surface, a geodesic is a curve where there appears to be no acceleration at all, so these are the "straight lines".

It isn't hard to show as well that geodesics have constant speed, since α'' is normal to α'.

In particular, if a unit-speed curve α lies in a plane P everywhere orthogonal to M along the curve, the α is a geodesic. The two orthogonal vectors α' and α'' are both orthogonal to M, and α' is tangent to M, so it must be the case that α'' is normal to M. As a consequence, all meridians are geodesics on a surface of revolution.

We summarize these three special kinds of curves:

- A principal curve: $k(\alpha) = k_1, k_2$. $S(\alpha')$ is collinear to α'.
- An asymptotic curve: $k(\alpha) = 0$. $S(\alpha')$ is orthogonal to α'. α'' is tangent to M.
- An geodesic curve: α'' is normal to M.

Since the shape operator is a linear map on the tangent space, think of principal curves as "eigencurves". A geodesic curve is, as we mentioned before, kind of a generalization of a straight line. One interesting application of geodesics is in finding shortest paths on surfaces, but we'll get to that later.

5.6 Surfaces of Revolution

This section is pretty much optional and exists solely to show the apparatus in action.

We discuss a particular class of surfaces to get a better sense of these concepts: surfaces of revolution. These have a "profile" curve:

$$x(u) = (g(u), h(u))$$

And we can extend this to a surface of revolution if $h(u) > 0$:

$$x(u, v) = (g(u), h(u)\cos(v), h(u)\sin v)$$

We'll compute the whole shape operator apparatus for a surface of revolution. First, we compute a basis for the tangent space:

$$\begin{aligned} x_u &= (g', h'\cos v, h'\sin v) \\ x_v &= (0, -h\sin v, h\cos v) \end{aligned}$$

Take their dot products to get E, F, G:

Take their cross product to get a unit vector field:

$$\begin{aligned} U &= \frac{x_u \times x_v}{|x_u \times x_v|} \\ &= \frac{(h', -g'\cos v, -g'\sin v)}{\sqrt{g'^2 + h'^2}} \end{aligned}$$

Note that the denominator here is simply the speed of our profile curve, or equivalently, $\sqrt{E}$. We continue onwards, taking second derivatives:

$$\begin{aligned} x_{uu} &= (g'', h''\cos v, h''\sin v) \\ x_{uv} &= (0, -h'\sin v, h'\cos v) \\ x_{vv} &= (0, -h\cos v, -h\sin v) \end{aligned}$$

Now, computing dot products we can compute L, M, N:

$$\begin{aligned} L &= \frac{-g'h'' + g''h'}{\sqrt{g'^2 + h'^2}} \\ M &= 0 \\ N &= \frac{g'h}{\sqrt{g'^2 + h'^2}} \end{aligned}$$

So we are finally ready to compute the shape operators. We know that x_u, x_v are principal directions (well, we don't know, but I don't want to prove it here; essentially all you need

to know is that the matrices of the quadratic forms we discussed in the last section are diagonal here, so this gives us a good idea that we picked the "right basis").

Given the lemma in the "normal curvature" section, we can write:

$$S(x_u) \cdot x_u = \frac{U \cdot x_u u}{x_u \cdot x_u} = \frac{L}{E}$$
$$S(x_v) \cdot x_v = \frac{U \cdot x_v}{x_v \cdot x_v} = \frac{N}{G}$$

And by plugging in, we can compute the principal curvatures, and their product, the Gaussian curvature.

Before going on to do all the ugly computations, let me go back and direct your attention to the term appearing everywhere relating to the speed of the profile curve. If the curve is one-to-one (does not cross back on itself), then by the implicit function theorem we can pick $g(u) = u$ without losing generality. Then, most of the terms above disappear. Note that $g'h'' + g''h' = (g'h')' = h''$, and that the speed term becomes $(1 + h'^2)$. So we now have the principal directions:

$$k_\mu = -\frac{h''}{(1+h'^2)^{3/2}}$$
$$k_\pi = -\frac{1}{h(1+h'^2)^{3/2}}$$
$$K = k_\mu k_\pi = -\frac{h''}{h(1+h'^2)^2}$$

We can further simplify this by picking a unit-speed profile curve, called the **canonical parametrization** of the surface. In this case, it is easy to see that:

$$E = g'^2 + h'^2 = 1$$
$$F = 0$$
$$G = h^2$$

And our principal curvatures and Gausssian curvature are now:

$$k_\mu = -h''$$
$$k_\pi = -\frac{1}{h}$$
$$K = k_\mu k_\pi = -\frac{h''}{h}$$

Theorem Suppose M is a minimal surface of revolution. Then M is a section of either a plane of a catenoid.

To prove this theorem, we look at $x = (g, h\cos v, h\sin v)$. We consider three cases.

Case 1 In this case, $g' = 0$ everywhere. Then, g is constant, so M is planar.

Case 2 In this case, $g' = 0$ nowhere. Then, we can repeat the earlier trick of picking another parametrization $y(u, v) = (u, h\cos v, h\sin v)$. However, looking at the earlier formulas, if M is minimal, then

$$h''h = 1 + h'^2$$

You can solve this differential equation by setting $v = y'$ and using separation of variables to arrive at:

$$1 + v^2 = ah^2$$

And by inspection, we can see that $v = \sinh(u/a + b)$ gives us a solution. This is precisely a catenoid.

Case 2 In this case, $g' = 0$ somewhere, but not everywhere. This is a contradiction. If $g' \neq 0$ at a point, we showed earlier that it locally looks like a catenoid. We cannot preserve continuity and yet have h'/g' tend to infinity.

So indeed, the only complete minimal surface is a catenoid (this means something in physics! We'll get to that later)

5.7 Summary

The shape operator S measures the instantaneous change of the unit normal. The unit normal is a sort of "first derivative" (telling us about the tangent plane), and so S is something like a second derivative. By looking at S as an algebraic object, we get some invariants of a surface which will become helpful in defining, as we did for curves, the notion of congruence: principal curvatures and directions, the Gaussian curvature, and the mean curvature (algebraically: the eigenvalues of the shape operator).

6 Geometry of Surfaces in $\mathbb{R}^3$

We continue by analogy. First, we set up the essential machinery for discussing curves in $\mathbb{R}^3$, such as frame fields, connection forms, and the structural equations in Part II. Then in Part III, we discussed the geometry of $\mathbb{R}^3$, or more specifically, the unique properties which are preserved by Euclidean isometries. Similarly, in Part V we discussed the machinery for discussing surfaces in $\mathbb{R}^3$, namely the shape operator. And in Part VI we will discuss the intrinsic geometry of a surface in $\mathbb{R}^3$, concluding as we did Part III with a congruence theorem for surfaces. Man, it's like poetry. Makes me tear up.

6.1 The Fundamental Equations

Once more, the Cartan equations from Part II find their use. As with the Frenet frame approach, we first need to set up a system of frames on M.

Definition: An **adapted frame field** E_i on a region $\mathcal{O}$ in $M \subset \mathbb{R}^3$ is a Euclidean frame field (three orthonormal vector fields) such that E_3 is always normal to M.

Lemma On a region $\mathcal{O}$ in $M \subset \mathbb{R}^3$, there exists an adapted frame field if and only if $\mathcal{O}$ is orientable and there exists a non-vanishing tangent vector field on $\mathcal{O}$.

This condition is clearly necessary, since the orientable condition guarantees the existence of a unit normal vector field (see: Part IV), and the tangent space condition guarantees us the remaining two vector fields. From a unit vector field U and a tangent vector V, we can define a frame:

$$\begin{aligned} E_1 &= \frac{V}{|V|} \\ E_2 &= U \times E_1 \\ E_3 &= U \end{aligned}$$

This lemma immediately implies that on the image of any patch in M, there is an adapted frame field on M in an open region. Thus adaptable frame fields exist on surfaces, at least locally.

Suppose E_i are an adapted frame field on M. We can extend the frame field so that it is defined on an open set in $\mathbb{R}^3$. Now, we look at the connection equations:

$$\nabla_v E_i = \sum \omega_{ij}(v) E_j$$

This definition really only makes sense if we apply them to tangent vectors v to M. Now, we have a set of 1-forms defined on M.

Again, $\omega_{ij}(v)$ defines the initial rate at which E_i rotates towards E_j as p moves in the direction v.

Corollary As a corollary, we can define the shape operator in terms of the connection forms, since $E_3 = U$ is a unit normal vector field. Then:

$$S_p(v) = \omega_{13}(v) E_1(p) + \omega_{23} E_2(p)$$

This follows directly from the connection equations.

We also can carry over directly the **dual** 1 **-forms**, $\theta_i(v) = v \cdot E_i$, as long as we restrict v to tangent vectors to M. This means that in particular, $\theta_3(v) = 0$, since E_3 does not lie in the tangent plane. Thus, due to skew-symmetry, we essentially only have five 1-forms:

- θ_1, θ_2 are dual to the tangent vector fields E_1, E_2.
- ω_{12} gives the rate of rotation of E_1 towards E_2.
- ω_{13}, ω_{23} completely describe the shape operator.

Theorem The big one! If E_i is an adapted frame field on $M \subset \mathbb{R}^3$, then its dual forms and connection forms satisfy:

First Structural Equations

$$d\theta_1 = \omega_{12} \wedge \theta_2$$
$$d\theta_2 = \omega_{21} \wedge \theta_1$$

Gauss Equation

$$d\omega_{12} = \omega_{13} \wedge \omega_{32}$$

Codazzi Equations

$$d\omega_{13} = \omega_{13} \wedge \omega_{23}$$
$$d\omega_{23} = \omega_{21} \wedge \omega_{13}$$

The first structural equations are the same as in Part II. The second structural equations imply the Gauss equations and the Codazzi equations, the former of which describes the rotation between E_1, E_2 and the latter of which describes the shape operator. It is important to note that these equations depend on the choice of frame field.

Since the corresponding connection forms describe the shape operator, the Codazzi equations describe the rate at which the shape of M is changing.

Lemma We can describe geodesics using connection forms. Let α be a unit speed curve in M. If E_1, E_2, E_3 is an adapted frame field where $E_1 = T$ along α (T being the unit tangent vector from the Frenet frame), we claim that α is a geodesic if and only if $\omega_{12}(T) = 0$.

First, denote $E_i' = E_i(\alpha(t))$ at $t = 0$. But this is the same as the covariant derivative: $E_i' = \nabla_{\alpha'} E_i$. In particular for this curve, $\alpha' = T$.

Since E_i is an adapted frame field and T is a unit vector, E_i' must be entirely in the diretction of E_2. This means that: $\omega_{12}(T) = (\nabla_T E_1) \cdot E_2$. But since α is a geodesic, α'' is normal to M; so there is no such component in the E_2 direction, so indeed $\omega_{12}(T) = 0$.

So, the connection forms already find some use in making sense of curves on surfaces.

6.2 Form Computations

If E_1, E_2, E_3 is an adapted frame field on $M \subset \mathbb{R}^3$, we say that E_1, E_2 constitute a tangent frame field on M. So any tangent vector field V on M can be decomposed into its components $V \cdot E_1, V \cdot E_2$ along each of the tangent frames.

Thus, to show that two forms are equal, we just have to compare them on basis vector fields E_1, E_2. So now we can equivalently form a "basis" of forms on M.

Lemma Let θ_1, θ_2 be the dual 1-forms of E_1, E_2 on M. If ϕ is a 1-form and ψ a 2-form, then:

$$\phi = \phi(E_1)\theta_1 + \phi(E_2)\theta_2$$

$$\mu = \mu(E_1, E_2)\theta_1 \wedge \theta_2$$

The first statement comes easily from the criteria for equality of frame fields; the second comes from our definition of the wedge product:

$$\begin{aligned}\theta_1 \wedge \theta_2(E_1, E_2) &= \theta_1(E_1)\theta_2(E_2) - \theta_1(E_2)\theta_2(E_1) \\ &= 1\end{aligned}$$

Lemma

$$\omega_{13} \wedge \omega_{23} = K\theta_1 \wedge \theta_2$$

$$\omega_{13} \wedge \theta_2 + \theta_1 \wedge \omega_{23} = 2H\theta_1 \wedge \theta_2$$

To prove these equations, first we need to write a matrix representation of the shape operator S. From the connection equations:

$$\begin{aligned}S(E_1) &= -\nabla_{E_1} E_3 = \omega_{13}(E_1)E_1 + \omega_{23}(E_1)E_2 \\ S(E_2) &= -\nabla_{E_2} E_3 = \omega_{13}(E_2)E_1 + \omega_{23}(E_2)E_2\end{aligned}$$

So the matrix of S is then:

$$\begin{bmatrix}\omega_{13}(E_1) & \omega_{13}(E_2) \\ \omega_{23}(E_1) & \omega_{23}(E_2)\end{bmatrix}$$

Now, we test equality of forms in the usual way. For the first form:

$$\begin{aligned}(\omega_{13} \wedge \omega_{23})(E_1, E_2) &= \omega_{13}(E_1)\omega_{23}(E_2) - \omega_{13}(E_2)\omega_{23}(E_1) \\ &= \det S \\ &= K\end{aligned}$$

And similarly for the second form:

$$\begin{aligned}(\omega_{13} \wedge \theta_2 + \theta_1 \wedge \omega_{23})(E_1, E_2) &= \omega_{13}(E_1) + \omega_{23}(E_2) \\ &= \mathrm{tr} S \\ &= 2H\end{aligned}$$

And from the first half of this lemma and the Gauss equation, we get:

$$d\omega_{12} = -K\theta_1 \wedge \theta_2$$

Which we will call the second structural equation. Since ω_{12} is something like the rate of rotation of the tangent frame field E_1, E_2, its derivative K measures something like a second derivative of E_1, E_2.

Definition: A **principal frame field** on M is an adapted frame field so that E_1, E_2 are principal vectors of M.

Where there are no umbilic points, there is a unique principal frame on M up to changes in sign. We leave this proof out, but it's not hard to prove using linear algebra and some analysis.

Suppose we have a principal frame field on M, then we can rename the principal directions so that $S(E_1) = k_1 E_1$ and $S(E_2) = k_2 E_2$. Thus by the basis formula from earlier, we have:

$$\begin{aligned}\omega_{13} &= k_1 \theta_1 \\ \omega_{23} &= k_2 \theta_2\end{aligned}$$

This leads us to an interesting corollary of the Codazzi equations.

Theorem If E_1, E_2, E_3 is a principal frame field on M, then:

$$\begin{aligned}E_1[k_2] &= (k_1 - k_2)\omega_{12}(E_2) \\ E_2[k_1] &= (k_1 - k_2)\omega_{12}(E_1)\end{aligned}$$

We can prove this just with differential forms algebra. First, we apply the Codazzi equations:

$$\begin{aligned}d(\omega_{13}) = d(k_1\theta_1) &= \omega_{12} \wedge \omega_{23} \\ dk_1 \wedge \theta_1 + k_1 d\theta_1 &= k_2 \omega_{12} \wedge \theta_2\end{aligned}$$

But we also know from the First Structural Equations that $d\theta_1 = \omega_{12} \wedge \theta_2$. So finally we have:

$$dk_1 \wedge \theta_2 = (k_2 - k_1)\omega_{12} \wedge \theta_2$$

To prove this statement, we simply apply each of the 2-forms to (E_1, E_2):

$$\begin{aligned}(dk_1 \wedge \theta_2)(E_1, E_2) &= (k_2 - k_1)\omega_{12} \wedge \theta_2(E_1, E_2) \\ -dk_1(E_2) &= (k_2 - k_1)\omega_{12}(E_1)\end{aligned}$$

Therefore:

$$E_2[k_1] = dk_1(E_2) = (k_1 - k_2)\omega_{12}(E_1)$$

And similarly for the second statement.

6.3 Some Global Theorems

We know apply our knowledge of the shape operator to draw some conclusions of (connected) surfaces.

Theorem If S is identically zero, then M is part of the plane in $\mathbb{R}^3$.

This means that the unit vector field E_3 is the same vector everywhere. So, we pick a fixed point p in M and consider any other point q in M, and draw a curve α in M so that $\alpha(0) = p, \alpha(1) = q$. So we consider the function:

$$\begin{aligned}f(t) &= (\alpha(t) - p) \cdot E_3 \\ \frac{df}{dt} &= \alpha' \cdot E_3 = 0\end{aligned}$$

So indeed, since $f(0) = 0$, f is identically zero. By checking $f(1)$, we see that indeed $(q - p) \cdot E_3 = 0$, so that indeed, q lies in the plane normal to E_3.

Theorem If M is umbilic everywhere, then M has constant Gaussian curvature.

To prove this, for each region in M we pick an adapted frame field E_1, E_2, E_3. Since every direction is a principal direction, indeed this is a principal frame field. So our earlier lemma from the Codazzi equations in the previous section, we have:

$$dk(E_1) = dk(E_2) = 0$$

By the basis formulas, $dk = 0$ in the region. But then we know that $dK = d(k_1^2) = 2k_1 dk = 0$. So indeed, we have shown that K is constant everywhere since $dK = 0$ in every region of M.

Theorem If M is umbilic everywhere, and $K > 0$, then M is part of a sphere in $\mathbb{R}^3$ of radius $\frac{1}{\sqrt{K}}$.

To prove this theorem, We will construct a point which is equidistant from every point in M. Pick any point p of M and denote the unit normal E_3. Then we claim that the center is:

$$c = p + \frac{1}{k(p)} E_3(p)$$

Now, we pick any point q of M, and draw a curve α in M as in the preceding proof with $\alpha(0) = p, \alpha(1) = q$. We extend E_3 to a unit normal vector field all along α. Now all we have to do is consider the curve:

$$\gamma = \alpha + \frac{1}{k} E_3$$

How do we know k is a constant? Well, since E_3 was extended continuously, $k(p)$ is also continuous; but our previous theorem tells us that the Gaussian curvature $K = k^2$ is constant. Taking the derivative of γ, we get:

$$\gamma' = \alpha' + \frac{1}{k} E_3'$$

But by the definition of the shape operator (which we recall is just scalar multiplication at umbilic points):

$$E_3' = -S(\alpha') = -k\alpha'$$

So indeed $\gamma' = 0$, and γ is constant. So in particular:

$$\gamma(0) = p + \frac{1}{k} E_3 = \gamma(1) = q + \frac{1}{k} E_3$$

And in particular, $|c - p| = |c - q|$. Furthermore, since $K = k^2$, the distance is indeed $\frac{1}{\sqrt{K}}$.

Combining the three previous theorems: A surface is all-umbilic if and only if M is a part of a plane or sphere.

Corollary A compact all-umbilic surface is a sphere.

First, we note from topology that a connected space has only two clopen sets: the empty set, and the whole space. In particular, both planes and spheres are connected.

Now, if M is compact, then it is automatically closed and bounded. Furthermore, for each patch in M, since x is continuous with continuous inverse, open sets map to open sets. Suppose that $M \subset N$ for some other surface N; then this means M, being the union of a collection of patches, is indeed a collection of open sets and thus open. Since M is open and closed in a connected surface N, it is either empty, or all of N itself.

If M is compact, it can't be a plane; therefore it must be a sphere.

Before proceeding into the next theorem, we first prove a useful lemma which relates the Gaussian curvature to the connection forms. First, by the basis theorem, we can write:

$$\omega_{12} = \omega_{12}(E_1)\theta_1 + \omega_{12}(E_2)\theta_2$$

And we also have from an earlier lemma:

$$d\omega_{12} = -K\theta_1 \wedge \theta_2$$

So, we differentiate each term in the earlier expansion of ω_{12}:

$$\begin{aligned} d\omega_{12} = d\omega_{12}(E_1) \wedge \theta_1 + \omega_{12}(E_1)d\theta_1 \\ + d\omega_{12}(E_2) \wedge \theta_2 + \omega_{12}(E_2)d\theta_2 \end{aligned}$$

But the first structural equations allow us to write $d\theta_i$. And indeed if we apply the 2-form above to the pair (E_1, E_2), we arrive at:

$$K = -d\omega_{12}(E_1, E_2) = E_2[\omega_{12}(E_1)] - E_1[\omega_{12}(E_2)] - \omega_{12}(E_1)^2 - \omega_{12}(E_2)^2$$

This puts Gaussian curvature entirely in terms of covariant derivatives and connection forms.

Theorem On every compact surface M in $\mathbb{R}^3$, there is a point at which the Gaussian curvature K is strictly positive.

We consider the real-valued (differentiable) function $f(p) = p \cdot p$ on a compact surface. Since M is compact, f attains a maximum at some point m on M\$

Take any unit tangent vector u to M at m, and pick a unit curve α on M so that $\alpha(0) = m, \alpha'(0) = u$. Then indeed, $f(\alpha(t))$ also has a maximum there, i.e.:

$$\frac{d}{dt}(f \circ \alpha)(0) = 0$$
$$\frac{d^2}{dt^2}(f \circ \alpha)(0) \leq 0$$

But we note that by our definition, $\frac{d}{dt}(f \circ \alpha) = 2\alpha \cdot \alpha$. So indeed:

$$\frac{d}{dt}(f \circ \alpha)(0) = 0 = 2(m \cdot u)$$

So this means that m is normal to M, since it is perpendicular to any tangent vector, and this means that locally, M looks like a sphere. Differentiating again:

$$\begin{aligned}\frac{d^2}{dt^2}(f \circ \alpha)(0) &= 2\alpha' \cdot \alpha' + 2\alpha \cdot \alpha'' \\ &= 2(1 + m \cdot \alpha''(0)) \leq 0\end{aligned}$$

Where the second equality comes from evaluating at $t = 0$.

Now, earlier we proved that m is a normal to M, so all we have to do is divide by its length r to get a unit normal vector. Thus, $m/r \cdot \alpha''$ is exactly the normal curvature $k(u)$ (See: Part V, Normal Curvature). So our inequality then becomes:

$$k(u) \leq \frac{1}{r}$$

Since both principal curvatures satisfy this inequality, we have:

$$K(m) \geq \frac{1}{r^2} > 0$$

So we have shown there is a point where the Gaussian curvature is positive.

Corollary As an immediate corollary, there is no compact surface in $\mathbb{R}^3$ with $K \leq 0$. So if we have a compact surface with constant Gaussian curvature, it must be positive. A sphere is an example of such a surface – and, as it turns out, it is the only example. To prove this, we begin wth a lemma.

Lemma Let m be a point in M so that: - k_1 achieves its local maximum at m. - k_2 achieves its local minimum at m. - $k_1(m) > k_2(m)$

Then, $K(m) \leq 0$.

Proof First, if f is a differentiable function on M and V a vector field, then we have also a first derivative $V[f]$ which yields a second derivative $VV[f] = V[V[f]]$. At a local maximum, it is not hard to prove that $V[f] = 0$ and $VV[f] \leq 0$, just as in calculus.

Now since $k_1 > k_2$ at m, and this is a strict inequality, then m is not umbilic; so there exists a principal frame field E_1, E_2, E_3 on a neighborhood of m, and indeed by calculus:

$$\begin{aligned} E_1[m] = E_2[m] = 0 \\ E_1E_1[k_2] \geq 0 \\ E_2E_2[k_1] \leq 0 \end{aligned}$$

Where the inequality follows from the fact that k_2 achieves a minimum, and k_2 a maximum. Now we apply the Codazzi equations in the way they appear on a principal frame field, and discover that at m:

$$\omega_{12}(E_1) = \omega_{12}(E_2) = 0$$

Now, by an earlier lemma:

$$\begin{aligned} K &= E_2[\omega_{12}(E_1)] - E_1[\omega_{12}(E_2)] - \omega_{12}(E_1)^2 - \omega_{12}(E_2)^2 \\ &= E_2[\omega_{12}(E_1)] - E_1[\omega_{12}(E_2)] \end{aligned}$$

To understand these terms, we apply E_i to the Codazzi equations to derive the inequalities:

$$\begin{aligned} E_1[\omega_{12}(E_2)] \geq 0 \\ E_2[\omega_{12}(E_1)] \leq 0 \end{aligned}$$

And thus our earlier expression for K reveals that indeed, $K \leq 0$.

Theorem Now we're finally ready to prove the theorem that a compact surface in $\mathbb{R}^3$ with constant Gaussian curvature is indeed a sphere of radius $\frac{1}{\sqrt{K}}$

Since the surface is compact, it is orientable, giving us a continuous shape operator and continuous principal curvatures k_1, k_2. Since M has constant Guassian curvature, we earlier proved that $K > 0$. Since we're working on a compact surface, k_1 has a maximum at some point p. Since K is constant, that also means k_2 is at a maximum.

From the immediately preceding lemma, the only missing condition is that $k_1 > k_2$ at p; if that were true, then $K \leq 0$, which is a contradiction. So we know that $k_1 = k_2$, not just at this point, but everywhere. So M is all-umbilic and it is therefore a sphere.

Note that compactness is essential for this proof.

6.4 Isometries and Local Isometries

Definition: If p, q are points in M, we consider all the curves α from p to q. The **intrinsic distance** $\rho(p, q)$ is the greatest lower bound on the lengths $L(\alpha)$ of the curves.

Note that this is just a greatest lower bound; there is not necessarily an actual curve with that length (due to limits). Now that we have a "metric", we can define an isometry.

Definition: An isometry $F : M \to \bar{M}$ of surfaces in $\mathbb{R}^3$ is a one-to-one mapping from M onto $\bar{M}$ that preserves dot products of tangent vectors. Explicitly:

$$F_\star(v) \cdot F_\star(w) = v \cdot w$$

Note that this was a theorem in Part III; the tangent map of an isometry in Euclidean space is simply its orthogonal component, evaluated at $F(p)$, and therefore dot products are preserved. Along with dot products, of course, we get lengths and orthogonality. It follows that isometries are regular mappings (its tangent map is one-to-one), because the tangent map sends zero vectors to zero vectors; and indeed this means that by the inverse function theorem, F is a diffeomorphism, i.e. has an inverse mapping (which is also an isometry).

Theorem Isometries preserve intrinsic distance.

We would hope that this holds. To prove this, consider any curve and recall how we defined its arclength by integrating its speed. Well, for any curve α, α' is a tangent vector; thus isometries preserve the norm of α' (speed), and therefore preserve arclength.

If there is an isometry between two surfaces, we call them **isometric**; for example, bending a piece of paper without stretching, folding, or tearing is an isometry.

Definition: A **local isometry** $F : M \to N$ of surfaces is a mapping that preserves dot products of tangent vectors.

Thus a local isometry need not be one-to-one and onto, and thus F is not necessarily a diffeomorphism. However, since F is regular on some neigborhood of any point, F carries that neigborhood diffeomorphically to a neighborhood of $F(p)$. Indeed, a local isometry works like an isometry, but only locally.

Lemma Let $F : M \to N$ be a mapping. For each patch $x : D \to M$, we consider the composite mapping:

$$\bar{x} = F(x) : D \to N$$

Then, F is a local isometry if and only if for each patch x we have:

$$E = \bar{E}$$
$$F = \bar{F}$$
$$G = \bar{G}$$

We prove both directions simultaneously.

First, recognize that we only need to show that $F_\star$ preserves dot products between x_u, x_v. The curve $\alpha = x(u, v_0)$ for which $\alpha' = x_u$ gets sent by F to $F \circ \alpha$, a curve in N for which $(F \circ \alpha)' = \bar{x_u}$, So indeed:

$$F_\star(x_u) = \bar{x_u}$$
$$F_\star(x_v) = \bar{x_v}$$

And since F is a local isometry, it therefore preserves dot products and therefore E, F, G. Reversing this argument, we deduce the converse statement; the tangent map preserves dot products, therefore it is a local isometry.

We can use this theorem to construct local isometries. Suppose we have $x : D \to M$ and $y : D \to N$. Then we can construct a mapping:

$$F(x(u, v)) = y(u, v)$$

And if $E = \bar{E}, F = \bar{F}, G = \bar{G}$, then F is a local isometry.

Definition: A mapping of surfaces $F : M \to N$ is **conformal** provided there exists a real-valued function $\lambda > 0$ on M such that:

$$|F_\star(v_p)| = \lambda(p)|v(p)|$$

The function λ is called a scale factor for F. This is a generalization of a local isometry where dot products are not preserved, but at a certain point the tangent vectors are all stretched by the same factor.

6.5 Intrinsic Geometry of Surfaces in $\mathbb{R}^3$

As before, we wish to study the properties of surfaces preserved by isometries, but not by arbitrary mappings. First note that since dot products are preserved, we still have tangent frames. From an adapted frame field we still have E_1, E_2 and their dual forms. What else can we say?

Lemma The connection form ω_{12} is the only 1-form that satisfies the first structural equations:

$$d\theta_1 = \omega_{12} \wedge \theta_2$$
$$d\theta_2 = \omega_{21} \wedge \theta_1$$

To prove this, we apply the wedge products on the RHS to the pair of tangent vector fields E_1, E_2:

$$d\theta_1(E_1, E_2) = \omega_{12}(E_1)$$
$$d\theta_2(E_1, E_2) = \omega_{21}(E_2) = -\omega_{12}(E_2)$$

Thus, by the basis lemma, the connection form ω_{12} is completely determined by θ_1, θ_2, which as we said before, is preserved by isometries. In fact, we can regard the above lemma as the definition of the connection forms, and it has nothing to do with the shape operator or covariant derivatives, just differential forms.

If $F : M \to \bar{M}$ is an isometry, we transfer the tangent frame E_1, E_2 at p to the pair $\bar{E}_1, \bar{E}_2 = F_\star(E_1), F_\star(E_2)$ at $F(p) = q$ (which is unique). This is also a frame field. Succinctly, we write:

$$F_\star(E_1) = \bar{E}_1$$
$$F_\star(E_2) = \bar{E}_2$$

Lemma Let $F : M \to \bar{M}$ be an isometry, and let E_1, E_2 be a tangent frame field on M. Then if $\bar{E}_1, \bar{E}_2$ is the transferred frame field on $\bar{M}$, then:

$$\theta_1 = F^\star(\bar{\theta_1})$$
$$\theta_2 = F^\star(\bar{\theta_2})$$
$$\omega_{12} = F^\star(\bar{\omega_{12}})$$

Proof First, we check the dual forms on a basis. By definition:

$$F^\star(\bar{\theta_i})(E_j) = \bar{\theta_i}(F_\star(E_j)) = \bar{\theta_i}(\bar{E_j}) = \delta_{ij}$$

So indeed, the pullback of the dual forms work exactly the same way. Next, we check the connection forms, remembering the first structural equation:

$$d\theta_1 = \omega_{12} \wedge \theta_2$$

Recall that $F_\star$ preserves wedge products and "commutes" with exterior derivatives (Part IV). So we have:

$$d(F^\star(\bar{\theta_1})) = F^\star(d\bar{\theta_1}) = F^\star(\bar{\omega_{12}}) \wedge F^\star(\bar{\theta_2})$$

By the first part of our lemma we can rewrite $\theta_1 = F^\star(\bar{\theta_1})$ and similarly for θ_2 to obtain:

$$d\theta_1 = F^\star(\bar{\omega_{12}}) \wedge \theta_2$$

And similarly:

$$d\theta_2 = F^\star(\bar{\omega_{21}}) \wedge \theta_1$$

But now by the uniqueness lemma, we have uniquely determined $F^\star(\bar{\omega_{12}}) = \omega_{12}$, as this is there is only one connection form satisfying the first structural equations.

Theorem We now arrive at Gauss' *Theorema Egregium* (Remarkable Theorem), which is really pretty remarkable. Simply, the Gaussian curvature is invariant under local isometry. More explicitly, if $F : M \to \bar{M}$ is an isometry, then:

$$K(p) = \bar{K}(F(p))$$

Proof For an arbitrary point $p \in M$, we pick a tangent frame field E_1, E_2 in a neighborhood of p and then transfer via $F_\star$ to a tangent frame field $\bar{E_1}, \bar{E_2}$. By the previous lemma, $\omega_{12} = F^\star(\bar{\omega_{12}})$ as well. According to an earlier theorem, we can compute:

$$d\bar{\omega_{12}} = -\bar{K}\bar{\theta_1} \wedge \bar{\theta_2}$$

Now we push forward this equation by applying $F_\star$ to both sides, and arrive at:

$$d\omega_{12} = -\bar{K}(F)\theta_1 \wedge \theta_2$$

Where $\bar{K}(F) = F_\star(\bar{K})$.

Thus, by comparison, indeed $K = \bar{K}(F)$.

The key point of this proof was the second structural equation, $d\omega_{12} = -K\theta_1 \wedge \theta_2$.

Note Philosophical note: the inhabitants of M don't have any knowledge of the shape operator or really even the shape of M; but they can determine the Gaussian curvature of their surface, just by constructing a local frame. An isometry may change the principal curvature, but it does not change their product. Local isometries also preserve Gaussian curvature, albeit only locally.

A plane and a cylinder are both called flat, even though a cylinder is "curved", because they have the same Gaussian curvature by local isometries.

6.6 Orthogonal Coordinates

So we can completely describe the intrinsic geometry of a surface with three forms $\theta_1, \theta_2, \omega_{12}$ derived from a frame field E_1, E_2. These forms are completely determined by:

**First Structural Equations:

$$d\theta_1 = \omega_{12} \wedge \theta_2$$
$$d\theta_2 = \omega_{21} \wedge \theta_1$$

Second Structrual Equations:

$$d\omega_{12} = -K\theta_1 \wedge \theta_2$$

Now, we will come up with a practical way to compute all these forms and thus the Gaussian curvature.

Definition: An **orthogonal coordinate patch** $x : D \to M$ is a patch for which $F = x_u \cdot x_v = 0$.

And indeed, dividing by $\sqrt{E}$ and $\sqrt{G}$, we can turn x_u, x_v into a frame field.

Definition: The associated frame field E_1, E_2 of an orthogonal patch $x : D \to M$ consists of the orthogonal unit vector fields:

$$E_1 = \frac{x_u}{\sqrt{E}}$$
$$E_2 = \frac{x_v}{\sqrt{G}}$$

At the point $x(u, v)$.

With each patch we can come up with coordinate functions $\tilde{u}, \tilde{v}$ which give us for each point $x(u, v)$ the cooresponding coordinates via the inverse function. We refer to these as u, v from now on. It is not hard to prove that:

$$du(x_u) = 1$$
$$du(x_v) = 0$$
$$dv(x_u) = 0$$
$$dv(x_v) = 1$$

So we can concisely write the dual forms for our associated frame field as:

$$\theta_1 = \sqrt{E}du$$
$$\theta_2 = \sqrt{G}du$$

And by using the structural equations (I leave the differential forms algebra to you), you can also solve for ω_{12} and K the way we did before.

6.7 Integration and Orientation

We start with a patch $x : D \to M$ and ask what its area should be. We start with a rectangle, with sides $\Delta u, \Delta v$.

Under the map, it gets distorted into a rectangle with sides $x_u \Delta u$ and $x_v \Delta v$. So, to find the length of the parallelogram, we take the cross product and arrive at $|x_u \times x_v| \Delta u \Delta v \approx \sqrt{EG - F^2} du dv$. Thus, we define something like an area by integrating this quantity. There's one small problem: patches have open domains, and integrals are defined for closed and bounded sets. So we define:

Definition: The interior R^o of a rectangle $[a, b] \times [c, d]$ is the open set $(a, b) \times (c, d)$. A two-cell $x : R \to M$ is called **patchlike** if the mapping $x : R^o \to M$ is a patch in M.

The function $\sqrt{EG - F^2}$ is bounded in a rectangle, so the area is well-defined and finite. Note that a patchlike segment is not necessarily one-to-one. To define something similar for a more complex region, we add up the areas of a bunch of smaller regions.

Definition: A **paving** of a region $\mathcal{P}$ in a surface M is a finite number of patchlike 2-segments $x_1, x_2, ...x_k$ whose images fill M in such a way that each point of M is in at most one set $x_i(R_i^o)$.

Not all regions are pavable, but it turns out that an entire compact surface is pavable. To find its area, we sum together the area of all its regions.

To get a more rigorous definition in terms of differential forms, we first remember how we defined the integration of a 2-form μ. We defined:

$$\int \int_x \mu = \int \int_R \mu(x_u, x_v) du dv$$

So in particular, we want to pick a $\mu(x_u, x_v) = \sqrt{EG - F^2}$. We now define such a form.

Definition: An **area form** on a surface M is a differentiable 2-form μ so that its value on any pair of tangent vectors is:

$$\mu(v, w) = \pm|v \times w|$$

Due to the linearity of forms, equivalently $\mu(E_1, E_2) = \pm 1$ for every frame E_1, E_2 on M. The sign ambiguity cannot be avoided if we want to keep, for example, Stokes' Theorem.

Lemma A surface M has an area form iff it is orientable. On a connected orientable surface, there are precisely two area forms, denoted $\pm dM$.

First, we know that a surface is orientable iff there is a non-vanishing 2-form on it. So if M has an area form, it is orientable. Given an orientable surface with a unit normal U, we can construct its area form:

$$dM(v, w) = \pm U \cdot v \times w$$

To **orient** a surface, we pick a unit normal. Let x be a patchlike segment, then we can then define:

$$\int\int_x dM = \int\int_R dM(x_u, x_v)dudv$$

There are three cases here depending on the sign of dM.

- If $dM(x_u, x_v)$ is positive, we say x is positively oriented; $dM = \sqrt{EG - F^2}$
- If $dM(x_u, x_v)$ is negative, we say x is negatively oriented; $dM = -\sqrt{EG - F^2}$

To find the area of a pavable region $\mathcal{P}$, we must use a paving that is positively oriented (each of its patches is positively oriented). Then we can construct the area by simply summing the integrals over each patchlike segment.

Definition: Let v be a 2-form on a pavable oriented region $\mathcal{P}$ in a surface. Then, the integral of v over $\mathcal{P}$ is defined:

$$\int\int_{\mathcal{P}} v = \sum_i \int\int_{x_i} v$$

Where $x_1, x_2, ...x_k$ is a positively oriented paving of $\mathcal{P}$.

Definition: Let f be a continuous function on a pavable region $\mathcal{P}$, then we define its integral over $\mathcal{P}$ to be:

$$\int\int_{\mathcal{P}} f dM$$

We can rigorously define the integration of any strictly positive function on an arbitrary region by taking the least upper bound of all of its integrals over all pavable regions.

6.8 Total Curvature

We define another interesting invariant of a surface, its total curvature.

Definition: Let K be the Gaussian curvature of a compact surface M oriented by an area form dM. Then we define the **total Gaussian curvature** of M as the integral:

$$\int\int_{M} K dM$$

Definition: Let M, N be surfaces oriented by area forms dM, dN. Then the **Jacobian** of the mapping $F : M \to N$ is the real-valued function J_F so that:

$$F^\star(dN) = J_F dM$$

So essentially, if we pull back an area form on N, we want to quantify by how much we shrink to get an area form on M:

$$J_F dM(v, w) = F^\star(dN)(v, w) = dN(F_\star(v), F_\star(w))$$

In particular, this means that F is regular iff J_F is nonzero at a point. The above equation also sufficiently motivates the following definitions:

Definition: If $J_F(p) > 0$, F is **orientation preserving** at p. If $J_F(p) < 0$, it is **orientation reversing**.

This also motivates a new way of calculating a signed area:

Definition: Let M, N be surfaces with a mapping $F : M \to N$. Then we define the algebraic area of $F(M)$ to be:

$$\int\int_{M} J_F dM = \int\int_{M} F^\star(dN)$$

If F switches orientation on a region, then that region's area makes a negative contribution, and similarly for preserving orientations.

In particular, we look at the **Gauss map**, which sends each point on a surface to the corresponding unit normal on the unit sphere. From the geometry of a sphere (the unit

normal is always pointing outwards), we know that the unit normal $\bar{U}(G(p))$ on the sphere is parallel to to $G(p)$, and thus to $U(p)$ on the original surface itself.

Theorem The Gaussian curvature K of an oriented surface M is the Jacobian of its Gauss map.

First, note that:

$$-S(v) = \sum v[g_i]U_i(p)$$
$$G_\star(v) = \sum v[g_i]U_i(G(p))$$

And indeed this means that these two vectors are parallel.

To prove this theorem, we must show the equivalence of the 2-forms: KdM and $G^\star(d\sum)$, where $\sum$ is the unit shere S^2. As always, we test this out by applying both forms to the (linearly independent) pair (v, w).

$$\begin{aligned} KdM(v,w) &= K(p)[U(p) \cdot v \times w] \\ &= U(p) \cdot S(v) \times S(w) \end{aligned}$$

To prove that last equality, simply write the matrix of S with respect to the basis v, w, take the cross product of an arbitrary pair $(av + bw, cv + dw)$, and note that the result is $\det Sv \times w$.

On the other hand, for $G^\star(d\sum)$:

$$G^\star(d\sum) = d\sum(G_\star(v), G_\star(w)) = \bar{U}(G(p)) \cdot G_\star(v) \times G_\star(w)$$

First, we noted earlier that $U(p)$ and $\bar{U}(G(p))$ are parallel. Furthermore, $G_\star(v)$ and $-S(v)$ are parallel and the minus signs cancel. So indeed the triple products above are equal.

Corollary The total Gaussian curvature of an oriented surface M is precisely the algebraic area of the image of its Gauss map $G : M \to S^2$.

This follows from the earlier theorem simply by integrating over M.

Corollary Let $\mathcal{R}$ be an oriented region in M on which: - The Gauss map is one-to-one - K does not change sign

Then the total curvature of $\mathcal{R}$ is, up to a sign, the area of $\mathcal{R}$; the sign is determined by the sign of K.

Evidently, this also means that the total curvature is bounded by the area of $\sum$, which is 4π.

On an oriented surface, we can define rotation rigorously as well.

Definition: On an oriented surface, the **rotation operator** is $J(v) = U \times v$.

This operator rotates tangent vectors by 90 degrees counterclockwise.

Definition: Let v, w be unit tangent vectors at a point p on an an oriented surface M. A number φ is called an **oriented angle** from v to w if:

$$w = \cos\varphi v + \sin\varphi J(v)$$

So indeed, an oriented angle measures rotation in the (well-defined) plane determined by v and $J(v)$.

Lemma Let α be a curve on an oriented surface M. If V, W are nonvanishing tangent vector fields on α, then there is a differentiable function φ on I such that φ is the oriented angle between $V(t)$ and $W(t)$.

WLOG, we turn V and W into unit frame fields and define the frame field $V, J(V)$ on α. By orthonormal expansion, we can define $W = fV + gJ(V)$. And finally, since W is of unit length, we can define $f = \cos\theta$ and $g = \sin\theta$. φ is then called an **angle function** from V to W.

Indeed, on an oriented surface, if we pick dM on positively oriented patches, with positively oriented frame fields, then we have $dM = \theta_1 \wedge \theta_2$.

In particular, given any nonvanishing vector field V on M, by the above discussion we can generate a positively oriented frame field. Namely, we take $V, J(V)$ and make them unit length; the resulting frame E_1, E_2 is called the **associated frame field** of V.

6.9 Congruence of Surfaces

Finally, we get to the moment we've been waiting for: establishing congruence of surfaces. We assume for simplicity we're working with connected, orientable surfaces.

Definition: Two surfaces are congruent if there is an isometry that carries one surface directly onto the other.

Theorem If F is a Euclidean isometry such that $F(M) = \bar{M}$, then its restriction to M is an isometry of surfaces, and F preserves shape operators, i.e.:

$$F_\star(S(v)) = \bar{S}(F_\star(v))$$

The first part of this theorem is fairly obvious. The restriction of $F_\star$ should still preserve dot products of tangent vectors, and the restriction of F is by hypothesis one-to-one and onto, and therefore we have an isometry.

We look at the shape operators. Let U be a unit normal in M. Since $F_\star$ preserves dot products and therefore orthogonality, $F_\star(U(p))$ is indeed orthogonal to all the tangent vectors at $F(p)$. Thus:

$$F_\star(U(p)) = \bar{U}(F(p))$$

Where $\bar{U}$ is one of the two unit normals on $\bar{M}$. So we define the two shape operators S derived from U and $\bar{S}$ defined from $\bar{U}$. So we pick a curve α with initial velocity v, and map it to $F(\alpha)$, a curve with initial velocity $F_\star(v)$. But we know that the tangent map of an isometry preserves velocities, so:

$$F_\star(S(v)) = -F_\star(U') = -[F_\star(U)]' = -\bar{U}' = \bar{S}(F_\star(v))$$

And indeed this is well-defined since v and $S(v)$ both lie in the tangent space of M.

So we know that by definition, congruent surfaces are isometric. But not all isometric surfaces are congruent. To make the converse relation true, we need to talk about the shape operator. Just as before, we defined congruence by looking at κ and τ for curves, we define congruence for surfaces by looking at the shape operator.

To prove this statement, we first prove a lemma about the congruence of curves theorem, which is instrumental to this generalization.

Lemma Let α, β be two curves defined on the same interval with frame fields E_1, E_2, E_3 and F_1, F_2, F_3 respectively. For any $t_0 \in I$, if F is the unique Euclidean isometry that sends each $E_i(t_0)$ to $F_i(t_0)$, or equivalently if: - $\alpha' \cdot E_i = \beta' \cdot F_i$ - $E_i' \cdot E_j = F_i' \cdot F_j$

Then the two curves are congruent, with $F(\alpha) = \beta$.

We won't consider the full proof here but the main takeaway is this: if you have two curves for which the velocities have the same expansion (relative to the respective frames) and the derivatives of the frames have the same expansion (relative to the respective frames), then indeed the two curves are congruent.

Theorem Let M and $\bar{M}$ be oriented surfaces in $\mathbb{R}^3$, and let $F : M \to \bar{M}$ be an isometry of oriented surfaces between them, such that the shape operators are preserved, i.e.:

$$F_\star(S(v)) = \bar{S}(F_\star(v))$$

Then, M and $\bar{M}$ are congruent; there is a Euclidean isometry whose restriction is precisely F.

I'll just proceed with the sketch of a proof here. First, we take the unique Euclidean isometry which pushes forward a tangent frame at p in M to precisely the pushforward of that tangent frame under F, and which pushes forward the unit normal E_3 at p determining the shape operator of M to the unit normal $\bar{E}_3$ at $F(p)$ determining the shape operator of $\bar{M}$. In other words, if we have a Euclidean isometry whose restriction to M is F, then certainly its tangent map agrees with that of F. Fix a point p_0 and for any point p, let α be any curve in M between the two points.

Next, transfer a tangent frame field on M via the isometry F to a tangent frame field on $\bar{M}$, and append the respective unit normals to get adapted frame fields for $M, \bar{M}$.

Finally, we confirm the conditions of our previous lemma using the properties of isometries (they preserve velocities and derivatives). The expansion of α' in the first frame is the same as the expansion of $F(\alpha')$ in the second frame. Using the pullback of the connection forms as well as the fact that isometries preserve shape operators, we prove that E_1, E_2 has the same expansion in M as $\bar{E}_1, \bar{E}_2$. Thus, we have found a Euclidean isometry which restricts to F; and indeed, this means every curve in M is congruent to a curve in $\bar{M}$, so we are done.

To draw analogy with the congruence theorem for curves: isometries between (orientable, connected) surfaces function sort of like reparametrizations of unit-speed curves on the same interval. And matching curvatures and torsions loosely correspond to matching shape operators. Putting the two kinds of hypotheses together yields a sufficient condition for congruence.

7 Riemannian Geometry

We now generalize the concepts built up in the previous discussion of geometry on surface in $\mathbb{R}^3$. The most important element of our earlier discussion was the Euclidean dot product. The simple generalization, then, is to define a general inner product on a vector space (namely, the tangent space). From the inner product alone we can construct a rich intrinsic geometry on surfaces which may be quite different from the Euclidean case.

7.1 Geometric Surfaces

Definition: An **inner product** on a vector space V is a real-valued, bilinear, symmetric, and positive definite (non-negative for all inputs) function $\langle v, w \rangle$ on pairs of vectors.

Now that we have something like a dot product, we can define the generalization of a surface.

Definition: A geometric surface is an abstract surface with an inner product on each tangent space.

In particular this means that if V, W are differentiable vector fields on a surface M, then $\langle V, W \rangle$ is a differentiable scalar function on M.

Definition: A metric tensor is a function on all pairs of vectors such that:

$$g_p(v, w) = \langle v, w \rangle_p$$

A metric tensor is thus similar to a 2-form, but instead of being skew-symmetric, it is symmetric. In its arguments. Loosely, we have:

$$\text{surface} + \text{metric tensor} = \text{geometric surface}$$

Now, we define a **frame field** on a surface M, as before, as a pair of unit vector fields E_1, E_2 so that:

$$\langle E_i, E_j \rangle = \delta_{ij}$$

And from this we have the dual 1-forms:

$$\theta_i(E_j) = \delta_{ij}$$

Finally, we carry over the connection form ω_{12}, which uniquely satisfied the structural equations:

$$d\theta_1 = \omega_{12} \wedge \theta_2$$
$$d\theta_2 = \omega_{21} \wedge \theta_1$$

In a neighborhood p of M, we get a differentiable angle function θ so that for any unit tangent vector field $\bar{E_1}$:

$$\bar{E_1} = \cos\varphi E_1 + \sin\varphi E_2$$

So indeed we can pick:

$$\bar{E_2} = -\sin\varphi E_1 + \cos\varphi E_2$$

And we get another (positively oriented) frame field. We can indeed pick $-\bar{E_2}$ to get a negatively oriented frame.

Lemma Let E_1, E_2 and $\bar{E_1}, \bar{E_2}$ be frame fields on some region in M:

$$\bar{\omega_{12}} = \pm(\omega_{12} + d\varphi)$$
$$\bar{\theta_1} \wedge \bar{\theta_2} = \pm(\theta_1 \wedge \theta_2)$$

Where we pick the sign depending on whether the frames have the same or opposite orientations. This proof follows by simple algebra to write θ_1, θ_2 in terms of the barred versions; then, take exterior derivatives and apply the first structural equations. By the uniqueness of solutions to the structural equations, we get our first result; by applying both wedge products to (E_1, E_2), we get the second result.

Definition: A manifold with a metric tensor is called a **Riemannian manifold**.

So indeed, a geometric surface is simply a 2-dimensional Riemmanian manifold.

7.2 Gaussian Curvature

Though we initially defined the Gaussian curvature in terms of the shape operator, the Theorema Egregium showed us we need not consider the ambient context; instead, we can calculate K from purely properties of the surface which are intrinsic to the surface itself. So we claim:

Theorem On a geometric surface M, there is a unique real-valued function K so that for every frame field on M, the second structural equation:

$$d\omega_{12} = -K\theta_1 \wedge \theta_2$$

The proof follows from the basis formulas, which tell us that there is a unique K satisfying the above (the second structural equation). But why does this hold for all frame fields? Say, that for example:

$$d\bar{\omega_{12}} = -\bar{K}\bar{\theta_1} \wedge \bar{\theta_2}$$

We wish to show that $K = \bar{K}$ on the overlap of the domains of the two frame fields. By our earlier theorem, we have:

$$\bar{\omega_{12}} = \omega_{12} \pm d\varphi$$

Where the sign is determiend depending on orientation. Taking the exterior derivative of both sides and noting that $d^2 = 0$, we indeed get that:

$$d\bar{\omega_{12}} = d\omega_{12}$$
$$\bar{K}\bar{\theta_1} \wedge \bar{\theta_2} = K\theta_1 \wedge \theta_2$$

Since both of the wedge products are nonzero, indeed $\bar{K} = K$. So indeed, the first structural equations uniquely determine the connection form ω_{12}, and the second structural equations uniquely determine K everywhere.

Example Let T be a torus of revolution with the usual parametrization:

$$x(u, v) = ([R + r\cos u]\cos v, [R + r\cos u]\sin v, r\sin u)$$

We define the metric with:

$$\begin{aligned}\langle x_u, x_u\rangle &= 1\\ \langle x_u, x_v\rangle &= 0\\ \langle x_v, x_v\rangle &= 1\end{aligned}$$

This determines a unique inner product at each point and consequently at each tangent plane. But this torus is flat! Recall:

$$\langle v, w\rangle_M = \langle F_\star(v), F_\star(w)\rangle_N$$

For any isometry F between surfaces M, N. And indeed x is a regular mapping, so we can indeed push forward an inner product on $\mathbb{R}^2$ to one on M:

$$\langle U_1, U_2\rangle_{\mathbb{R}^2} = \langle x_\star(U_1), x_\star(U_2)\rangle_M$$

And this is indeed the inner product we just defined, since $x_\star(U_1) = x_u$. What we have done is pull back this torus under x to get the plane. By Gauss' Theorema Egregium, K is preserved, so we have defined a flat torus. However, this torus is compact, and we know from earlier discussion that in $\mathbb{R}^3$ that a compact surface has positive curvature. Therefore, the flat torus does not exist in $\mathbb{R}^3$.

Corollary For the plane $\mathbb{R}^2$ with the metric tensor $\langle v, w\rangle = \frac{v\cdot w}{h^2(p)}$ for some function $h(u, v) > 0$, the Gaussian curvature is:

$$K = h(h_{uu} + h_{vv}) - (h_u^2 + h_v^2)$$

The proof is simple. The identity map $\mathbb{R}^2 \to M$ is a patch with $E = G = \frac{1}{h^2}, F = 0$. Recall earlier that we proved we can compute K purely in terms of E, F, G; using this formula the result follows by algebra.

Example For example, we can define the so-called hyperbolic plane with $h(u,v) = 1 - \frac{u^2+v^2}{4}$ defined in the disk $u^2+v^2 < 4$. By our earlier lemma, $K = h(h_{uu}+h_{vv})-(h_u^2+h_v^2) = -1$.

Note that any regular mapping $F : M \to N$ pulls back a metric tensor on N to one on M, but in general you can only push forward a metric via a diffeomorphism (which has a smooth inverse). The essential problem is that if $F(p_1) = F(p_2)$, then inner products might transfer differently at those two points. We address this issue with the following proposition.

Proposition Suppose F is a regular mapping from a geometric surface M to a surface N. Suppose when $F(p_1) = F(p_2)$, that there is an isometry G_{12} from a neighborhood of p_1 to a neighborhood of p_2 such that:

$$FG_{12} = F$$
$$G_{12}(p_1) = p_2$$

Then, there exists a unique metric tensor on N which makes F a local isometry.

So indeed, this gives us a sufficient condition for the pushforward of a metric under a regular mapping to be well-defined.

Proof Since F is a local isometry, the inner product is completely determined on N by the isometry. F is regular, so for every p_1 in M with $F(p_1) = q$, there are unique vectors v_1, w_1 so that:

$$F_\star(v_1) = v$$
$$F_\star(w_1) = w$$

And we write in this case that $\langle v, w\rangle_N = \langle v_1, w_1\rangle_M$. What if $F(p_2) = q$ and $p_1 \neq p_2$? The same argument holds, and we have a unique pair v_2, w_2 at p_2 which are pushed forward by the tangent map of the isometry to v, w. We want to show:

$$\langle v_1, w_1\rangle_M = \langle v_2, w_2\rangle_M$$

Let $G = G_{12}$ be an isometry. Then, $FG = F$, so by the chain rule, $F_\star G_\star = F_\star$. Therefore, $G_\star$ carries v_1, w_1 to v_2, w_2, and since G is an isometry it indeed preserves the inner product.

7.3 The Covariant Derivative

We want to generalize the concept of the covariant derivative, so that: - $\nabla_V W$ at p is the rate of change of W in the direction of $V(p)$. - $\omega_{12}(V) = \langle \nabla_V E_1, E_2\rangle$.

Lemma Assume there is a covariant derivative ∇ defined on M with linearity and the Leibniz property, satisfying the connection form equation as well. Then, ∇ satisfies the **connection equations**:

$$\nabla_V E_1 = \omega_{12}(V)E_2$$
$$\nabla_V E_2 = \omega_{21}(V)E_1$$

And for a vector field $W = f_1E_1 + f_2E_2$:

$$\nabla_V W = (V[f_1] + f_2\omega_{21}(V))E_1 + (V[f_2] + f_1\omega_{12}(V))E_2$$

This is called the **covariant derivative formula**.

Proof Let $\nabla_V = \langle \nabla_V E_1, E_1\rangle E_1 + \langle \nabla_V E_1, E_2\rangle E_2$, by orthonormal expansion. Then, we have:

$$0 = V[\langle E_1, E_1\rangle] = 2\langle \nabla_V E_1, E_1\rangle$$

So indeed, $\nabla_V E_1 = \omega_{12}(V)E_2$, since the covariant derivative has no component in the direction of E_1. The covariant derivative formula then follows by siply applying the chain rule to $W = f_1E_1 + f_2E_2$.

Theorem This theorem is often called the fundamental theorem of Riemannian Geometry. On each surface M, there exists a unique covariant derivative ∇ satisfying: - The linearity and Leibniz properties. - $\omega_{12}(V) = \langle \nabla_V E_1, E_2\rangle$. For every frame field E_1, E_2.

The surprising upshot of this theorem is that there is a single covariant derivative which satisfies the connection form condition for *all* frame fields. The preceding lemma shows that for each frame, there is at most one covariant derivative (the one specified by the covariant derivative formula above). So we wish to prove that there is at least one covariant derivative.

Proof First, we consider the local definition. For a frame field E_1, E_2 on a region in M, we know how to define the covariant derivative by the covariant derivative formula; we can check that the product rule is satisfied and linearity holds usig that formula. Then, to prove the condition regarding the connection forms, set $W = E_1$.

Secondly, we deal with the issue of consistency. Do our local definitions agree for two different frame fields? We only need to check that two frame fields defined on the same region yields covariant derivatives which agree on $W = E_1, W = E_2$. Assume WLOG that the two frames have the same orientation. Then:

$$\begin{aligned}\bar{E_1} &= \cos\varphi E_1 + \sin\varphi E_2 \\ \nabla_V \bar{E_1} &= \sin\varphi(-V[\varphi] + \omega_{21}(V))E_1 + \cos\varphi(V[\varphi] + \omega_{12}(V))E_2\end{aligned}$$

But we have $\bar{\omega_{12}} = \omega_{12} + d\varphi$, and so we our above equation reduces to:

$$\begin{aligned}\nabla_V \bar{E_1} &= \bar{\omega_{12}}(V)(-\sin\varphi E_1 + \cos\varphi E_2) \\ &= \bar{\omega_{12}}(V)\bar{E_2} \\ &= \bar{\nabla}_V \bar{E_1}\end{aligned}$$

And similarly for E_2.

In particular, for $\mathbb{R}^2$, we get the covariant derivative defined earlier: if $W = \sum f_i U_i$, then $\nabla_V W = \sum V[f_i]U_i$.

For curves, we want to define something like $\nabla_{\alpha'} Y$ as we did before. But the quantities involved are not necessarily defined on open sets. So we simply decide:

$$Y' = (f_1' + f_2\omega_{21}(\alpha'))E_1 + Y' = (f_2' + f_1\omega_{12}(\alpha'))E_2$$

What we have done is simply adapt the covariant derivative by replacing $V[f_i]$ with f_i'.

Definition: A vector field V on α in a geometric surface is parallel if $V' = 0$ along α.

In other words, V is paralle if the covariant derivative vanishes. Note that in this case, $\langle V, V\rangle' = 2\langle V', V\rangle = 0$, so a parallel vector field has perfect length.

LEmma If *alpha* is a curve, and v a tangent vector at $p = \alpha(t_0)$, then there is a unique parallel vector field V so that $V(t_0) = v$.

This follows from the uniqueness and existence theorems for differential equations. IN this case, the conditions are: $V'(t) = 0$ and $V(t_0) = v$. It is not hard to see by the previous lemma that the angle function has the corresponding conditions $\varphi' + \omega_{12}(\alpha') = 0$, and $\varphi(t_0)$ is the angle between $E_1(p)$ and v. This allows us to get an explicit formula for φ by integrating. This is called the parallel translation of v at p along α.

If v is taken by parallel translation around a closed curve, it might not end up as v when it returns to the same spot, a phenomenon called holonomy. We know, however, that in particular:

$$\varphi(b) - \varphi(a) = -\int_\alpha \omega_{12}$$

So we call this quantity ψ_α, the holonomy angle of α.

What is the relation between our new concept of the covariant derivative and the one we earlier defined in Euclidean space? This is illuminated in the following lemma.

Lemma If V, W are tangent vector fields to M, and if $\tilde{\nabla}$ denotes the Euclidean covariant derivative, then: - ∇_V, W is the component of $\tilde{\nabla}_V W$ which is tangent to M. - $\tilde{\nabla}_V W = \nabla_V W + (S(V) \cdot W)U$ Where $S(V)$ is the shape operator and U the associated unit normal.

To prove this, we need only prove the second condition, which implies the first. But we have by definition using the connection forms:

$$\begin{aligned}\tilde{\nabla}_V E_1 &= \sum \omega_{ij}(V) E_j \\ &= \omega_{12}(V) E_2 + \omega_{13}(V) E_3\end{aligned}$$

And the definition of the covariant derivative on surfaces gives:

$$\nabla_V E_1 = \omega_{12}(V) E_2$$

And, since $E_3 = U$, we have:

$$\begin{aligned}\omega_{13}(V) E_3 &= (\tilde{\nabla}_V E_1 \cdot E_3) E_3 \\ &= (\tilde{\nabla}_V E_3 \cdot E_1) E_3 \\ &= (S(V) \cdot E_1) U\end{aligned}$$

We can do the same thing and instead take the covariant derivative of E_2 to get out result, and then prove the result for an arbitrary $W = f_1 E_1 + f_2 E_2$ by calculus.

7.4 Geodesics

We now can define geodesics without using the unit normal.

Definition: A curve γ in a geometric surface is a geodesic if $\gamma'' = 0$; equivalently, γ' is parallel; equivlaently, γ' has constant length.

Since acceleration is preserved by isometries, geodesics are isometric invariants under our new definition. Indeed, that means if γ is a geodesic, then $F \circ \gamma$ is also a geodesic. We use the notation:

$$\begin{aligned}\alpha' &= v_1 E_1 + v_2 E_2 \\ \alpha'' &= A_1 E_1 + A_2 E_2\end{aligned}$$

By definition, α is a geodesic iff $A_1 = A_2 = 0$. In particular, if we apply the above definition for the covariant derivative of α, we arrive at the conditions:

$$A_1 = v_1' + v_2\omega_{21}(\alpha') = 0$$
$$A_2 = v_2' + v_1\omega_{12}(\alpha') = 0$$

We can indeed use an associated frame field, which gives the following equivalent conditions (I will not go through the proof here):

$$A_1 = a_1'' + \frac{1}{2E}(E_u a_1'^2 + 2E_v a_1' a_2' - G_u a_2'^2) = 0$$
$$A_2 = a_2'' + \frac{1}{2G}(-E_v a_1'^2 + 2G_u a_1' a_2' + G_v a_2'^2) = 0$$

We now arrive at a theorem which gives us the ability to construct geodesics easily.

Theorem Given a tangent vector v to M, there exists a unique geodesic γ, defined around 0, so that $\gamma(0) = p, \gamma'(0) = v$.

To prove this, let x be an orthogonal patch around $p = x(u_0, v_0)$ (orthogonal meaning $F = 0$). Let $v = cx_u + dx_v$. Then our above conditions become:

$$a_1'' = f_1(a_1, a_2, a_1', a_2')$$
$$a_2'' = f_2(a_1, a_2, a_1', a_2')$$

And with the initial conditions $(a_1, a_2)(0) = (u_0, v_0)$ and $(a_1', a_2')(0) = (c, d)$. By the existence and uniqueness of differential equations, we get our result.

Note that this allows us to extend geodesics infinitely. Let us fix a tangent vector v. Then at every point, we have a unique geodesic γ pointing in the direction of v defined in a neighborhood of 0. If we pick a point near the edge of that neighborhood, we define another unique geodesic, and so on. By gluing together all these geodesics, we get what's called a **maximal geodesic**.

Definition: A geometric surface M is complete if every maximal geodesic γ_v in M is defined on the whole real line $\mathbb{R}$.

Here, we only need to parametrize the maximal geodesics by their direction v. In a complete geometric surface, a maximal geodesic runs infinitely. For example, on a sphere, a great circle is a geodesic which repeats itself periodically and thus runs forever.

As it turns out, all compact geometric surfaces are complete.

Lemma Let E_1, E_2 be a frame field with constant speec so that α' is never orthogonal to E_2. If $A_1 = 0$, then α is a geodesic.

The proof is straightforward. Since α' is constant length, we have $\langle \alpha', \alpha' \rangle' = 2\langle \alpha'', \alpha' \rangle = 0$. So indeed:

$$\langle A_1 E_1 + A_2 E_2, \alpha' \rangle = A_1 \langle E_1, \alpha' \rangle + A_2 \langle E_2, \alpha' \rangle = 0$$

By hypothesis, $A_1 = 0$. Since E_2 is not orthogonal to α', we must have $A_2 = 0$, and thus indeed α is a geodesic.

In particular, if you pick E_1, E_2 to be frame fields so that $\langle x_u, x_v \rangle = \delta_{ij}$, then our above conditions on A_1, A_2 can be simplified. Adding and integrating, we get:

$$Ea_1'^2 + Ga_2'^2 = \text{const}$$

We now introduce a critical notion for the next section. If $\beta : I \to M$ is aunit-speed curve on an oriented surface, then we can define $T = \beta', N = J(T)$ to get an analogue of the Frenet frame field. We then get an analogue for curvature, the **geodesic curvature** k_g:

$$T' = k_g N$$

As a direct corollary, if φ is the angle between E_1 and β', then we have:

$$k_g = \frac{d\varphi}{ds} + \omega_{12}(\beta')$$

Note that in the case of Euclidean space, this confirms our notion that curvature measures the "rate of turning" of the tangent vector.

We get this by setting $Y = T$, $\alpha = \beta$, and applying the definition of the covariant derivative for vector fields with respect to curves on geometric surfaces. Another interesting consequence of the definition of geometric curvature is that:

$$\alpha' = vT$$
$$\alpha'' = \frac{dv}{dt}T + k_g v^2 N$$

Where v is the speed of the curve α. This formual gives is an easy condition to see if a curve is a geodesic in terms of its geodesic curvature:

Lemma A regular curve α is a geodesic iff it has constant speed and $k_g = 0$.

Since the speed $v > 0$ by regularity, by the above formula $\alpha''0$ is equivalent to the condition that $\frac{dv}{dt} = k_g = 0$. This means that α'' is collinear to the tangent vector, as in our earlier definition of geodesics (curves whose accelerations were normal to the surface). We can easily turn any curve with $k_g = 0$ (called a **pre-geodesic**) into a geodesic, then, by reparametrizing and setting speed to be constant.

7.5 The Gauss-Bonnet Theorem

Definition: $\alpha : [a, b] \to M$ is a regular curve segment in an oriented geometric surface. We define the **total geodesic curvature** to be the quantity:

$$\int_\alpha k_g ds = \int_{s(a)}^{s(b)} k_g(s(t)) \frac{ds}{dt} dt$$

The total geodesic curvature is the analogue of the total Gaussian curvature, and we will soon see there is a link between them.

Lemma Let α be a regular curve in M which has an oriented frame field E_1, E_2. Then we have:

$$\int_\alpha k_g ds = \varphi(b) - \varphi(a) + \int_\alpha \omega_{12}$$

Where φ is the angle from E_1 to α' along α. This lemma follows from simply integrating our earlier formula for the geodesic curvature.

We now begin the road to proving the famous Gauss-Bonnet theoerem, which tells us that the total curvature is intricately connected to topological invariants of a surface.

Definition: Let $x : R \to M$ be a one-to-one regular patch from a 2-segment. We define the exterior angle ϵ_j of x at the vertice p_j to be the angle between the relevant tangent vectors of the boundary curves. The interior angle i_j is similarly $\pi - \epsilon_j$; the interior and exterior angles add up to π.

Theorem Let $x : R \to M$ be a one-to-one, regular mapping to a 2-segment in a geometric surfcae M. If dM is an area form, then:

$$\int\int_x K dM + \int_{\partial x} k_g ds + \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 = 2\pi$$

This is often called the **Gauss-Bonnet** formula.

Proof We pick the associated frame field $E_1 = \frac{x_u}{\sqrt{E}}, E_2 = J(E_1)$, where J is the rotation operator. Then we have $dM(E_1, E_2) = 1$; we picked the positive orientation. By the second structural equation:

$$\begin{aligned} d\omega_{12} &= -K\theta_1 \wedge \theta_2 \\ &= -KdM \end{aligned}$$

Now, by Stokes Theorem, we have:

$$\int_{\partial x} \omega_{12} = \int\int_x d\omega_{12} = \int\int_x -KdM$$
$$\int_{\partial x} \omega_{12} + \int\int_x KdM = 0$$

But recall earlier we defined:

$$\int_\alpha k_g ds = \varphi(b) - \varphi(a) + \int_\alpha \omega_{12}$$

Adding up the integrals for each of the four boundary curves, we get the required result.

Definition: A regular decomposition D of a surface M is a finite collection x_i of rectangles which intersect at a common vertex or edge.

This is essentiall what we did earlier with the concept of a paving, but slightly modified.

Theorem Every compact surface M has a rectangular decomposition. We won't prove this here, as it is a result from topology.

Theorem In dimension 2, two surfaces are diffeomorphic iff they are homemomorphic. We won't prove this here.

What the above theorem implies is that topological invariants (those preserved by homeomorphisms) are now in fact preserved by diffeomorphisms. This reduces our work considerably, since diffeomorphisms give us differentiability.

Theorem Let D be a rectangular decomposition of a compact surface. let v, e, f denote the number of vertices, edges, faces in D. Then $v - e + f$, called the **Euler characteristic** $\chi(M)$, is the same for each rectangular decomposition.

This theorem from topology tells us that the Euler characteristic does not depend on our rectangularization. So, if we have two separate rectangular decompositions, we can construct a diffeomorphism by varying smoothly between them, and thus we know that diffeomorphic surfaces have the same Euler characteristic.

Theorem If M is a comapct, connected, orientable surface, there is a unique positive integer h such that M is diffeomorphic to $\Sigma[h]$, which is the sphere with h handles added on.

A handle can be constructed as follows. Take any rectangular decomposition of M, and glue together one "tile" from M to a tile from a rectangular decomposition of a torus. By looking at the case of the sphere which has Euler characteristic 2, when we add on a handle we get a surface diffeomorphic to a torus, which has Euler characteristic 0. In general, adding h handles reduces the Euler characteristic of the sphere by $2h$.

As a direct corollary, compact orientable surfaces M, N have the same Euler characteristic if and only if they are diffeomorphic.

To prove this statement, consider that $\chi(M) = \chi(N)$ means that both surfaces are diffeomorphic to a sphere with h handles; so they are diffeomorphic to each other.

We can now state the incredible Gauss-Bonnet Theorem.

Theorem (Gauss-Bonnet) For a compact, orientable surface, we have:

$$\int\int_M KdM = 2\pi\chi(M)$$

This theorem connects a topological invariant (the Euler characteristic) to a geometric/isometric invariant (the Gaussian curvature). In fact, this says that the Gaussian curvature is a topological invariant.

Proof I will sketch the proof here, avoiding some details. First, we orient M by an area form dM, and construct a rectangular decomposition D which is positively oriented (we can simply reorder the tangent vectors until we get positive orientations on each rectangle). Then, D is an oriented paving. We can now integrate to obtain:

$$\int\int_M KdM = \sum \int\int_{x_i} KdM$$

But by the ealrier Gauss-Bonnet formula, we have:

$$\sum \int\int_{x_i} KdM = -\int_{\partial x_i} k_g ds - 2\pi + i_1 + i_2 + i_3 + i_4$$

But on the intersection, we note that since all the rectangles have the same orientaton, the Gaussian curvatures on boundary curves cancel out! This is similar to a common trick you see in integration. If we split up a rectangle into four smaller rectangles and integrate over each, the only terms that do not cancel are associated with the boundary curves. So we are left with:

$$\sum \int \int_{x_i} K dM = \sum -2\pi + i_1 + i_2 + i_3 + i_4$$

Where we are summing over the number of tiles in our paving (the number of faces). So we can rewrite the sum above as:

$$\sum -2\pi + i_1 + i_2 + i_3 + i_4 = -2\pi f + \sum i_1 + i_2 + i_3 + i_4$$

Where f is the number of faces. Note that around every vertex, we have four interior angles which add up to 2π. Therefore, we can reduce the sum on the right to obtain:

$$\sum \int \int_{x_i} K dM = 2\pi(v - f)$$

And finally, we try to count the number of edges in a surface. If we simply multiply the amount of faces by four, we are double counting, since every edge appears on two faces. So we have $4f = 2e$. Equivalently above:

$$\sum \int \int_{x_i} K dM = 2\pi\chi(M)$$

And we are done.

7.6 Applications of Gauss-Bonnet

We now run through some simple consequences of the Gauss-Bonnet Theorem.

First, there is no metric on a sphere $\sum$ with $K < 0$. If there was, then the total Gaussian curvature would be negative. But by Gauss-Bonnet, the total Gaussian curvature is $2\pi\chi(M) = 4\pi$, which is positive.

Second, every compact, orientable geometric surface M which has $K > 0$ is diffeomorphic to a sphere. If the geometric surface has positive total Gaussian curvature, then $\chi(M)$ is positive by Gauss-Bonnet. But if we add even one handle to a sphere, the Euler characteristic drops to 0. Therefore, any geometric surface is diffeomorphic to a sphere.

Theorem The following are equivalent on a compact, orientable surface. - There exists a non-vanishing tangent vector field on M. - The Euler characteristic of M is 0. - M is diffeomorphic to a torus.

First, we prove that the first statement implies the second. We construct an associated frame field from the non-vanishing vector field:

$$E_1 = \frac{V}{|V|}$$
$$E_2 = J(E_1)$$

By the second structural equation, $d\omega_{12} = -KdM$.

By Gauss-Bonnet:

$$\int\int_M KdM = 2\pi\chi(M)$$

By Stokes' Theorem:

$$\int\int_M KdM = -\int\int_M d\omega_{12}$$

But for the associated frame field, $\omega_{12} = 0$, so indeed this means the integral is 0 and thus so is the Euler characteristic.

Next, we prove the second statement implies the third. If the Euler characteristic is 0, then M is diffeomorphic to a sphere with one handle, which is a torus.

Finally, we prove the third statement implies the first. We use x_u, x_v from the usual parametrization of the torus; either is a non-vanishing tangent vector field.

7.7 Summary

Much of our discussion of surfaces revolved around the Euclidean dot product. However, in this chapter we showed that we can create arbitrary metric spaces on surfaces to obtain a Riemmanian manifold. On a Riemannian manifold, many concepts and ideas see analogous results: differential forms, frame fields, the structural equations, covariant derivatives, curvature, and geodesics. An interesting result we can prove is that the total curvature is indeed a topological quantity; thus, we have established a fundamental link between geometry and topology.