M435: INTRODUCTION TO DIFFERENTIAL GEOMETRY

MARK POWELL

Contents

4. Geometry of Surfaces

4.1. Smooth patches.

Definition 4.1. A *smooth patch* of a surface in $\mathbb{R}^3$ consists of an open subset $U \subseteq \mathbb{R}^2$ and a smooth map $\mathbf{r}\colon U \to \mathbb{R}^3$ which is a homeomorphism onto its image $\mathbf{r}(U)$, and which has injective derivative. The image $\mathbf{r}(U)$ is a smooth surface in $\mathbb{R}^3$.

Recall that the injective derivative condition can be checked by computing $\mathbf{r}_u \wedge \mathbf{r}_v \neq \mathbf{0}$ everywhere. A patch gives us coordinates (u, v) on our surface. The curves $\alpha(t) = \mathbf{r}(t, v)$ and $\beta(t) = \mathbf{r}(u, t)$, for fixed u, v, are the coordinate lines on our surface. The lines of constant longitude and latitude on the surface of the earth are examples.

Smooth means that all derivatives exist. In practice, derivatives up to third or fourth order at most are usually sufficient.

Example 4.2. The requirement that $\mathbf{r}$ be injective means that we sometimes have to restrict our domain, and we may not cover the whole of some compact surface with one patch.

(1) For example, recall a parametrisation of the sphere:

$$\mathbf{r}(u, v) := (\cos u \sin v, \sin u \sin v, \cos v).$$

We need to restrict to $(u, v) \in (0, 2\pi) \times (0, \pi)$ in order for this to be a 1-1 map. This misses out the Greenwich meridian, including the north and south poles of S^2.

(2) Similarly for the torus

$$\mathbf{r}(u, v) = \big((a + b \cos u) \cos v, (a + b \cos u) \sin v, b \sin u\big)$$

we let $U = (0, 2\pi)^2$. This misses out the union of the coordinate curves $\mathbf{r}(0, v)$ and $\mathbf{r}(u, 0)$, so that we have an injective function. In the version of the torus from identifying edges of the unit square, this means we parametrise the image of the interior of the square, and do not include the image of the edges, nor the corner points.

(3) We can actually parametrise the whole sphere apart from a single point, using what is called the *stereographic projection.* We define a map $\mathbf{r}\colon \mathbb{R}^2 \to S^2 \subset \mathbb{R}^3$. Imagine the unit sphere with centre $(0, 0, 0)$. For a point $(u, v) \in \mathbb{R}^2$, draw a straight line which passes through $(u, v, 0)$ in the $x - y$ plane and the north pole $(0, 0, 1)$ of S^2. This intersects the sphere in one other point (which is not the north pole). Let this be the point $\mathbf{r}(u, v)$.

Now let's figure out the formula. The straight line from $(0, 0, 1)$ to $(u, v, 0)$ has equation

$$\gamma(t) = (0, 0, 1) + t(u, v, -1) = (tu, tv, 1 - t).$$

This intersects the sphere when

$$\begin{aligned} & t^2u^2 + t^2v^2 + (1 - t)^2 = 1 \\ \Leftrightarrow \quad & t(tu^2 + tv^2 - 2 + t) = 0 \end{aligned}$$

The point $\gamma(0)$ is the north pole; we are looking for the other point. This is when

$$t = \frac{2}{u^2 + v^2 + 1}.$$

Plugging back into γ yields

$$\mathbf{r}(u, v) = \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1} \right).$$

This is the stereographic projection, which gives a bijection $\mathbb{R}^2 \to S^2 \setminus (0, 0, 1)$.

In general, we define a surface to be a subset of $\mathbb{R}^3$ such that every point is in the image of a smooth patch. A surface is constructed from multiple patches, with suitable change of coordinates on the overlaps of the patches. One smooth surface, which is a subset of $\mathbb{R}^3$, can have many parametrisations, or smooth patches.

Exercise 4.3. What is the smallest number of patches required to cover the torus?

4.2. **Tangent spaces.** Consider a point $p = \mathbf{r}(a,b)$ in a surface and consider the tangent vectors to the curve $\alpha(t) = (t,v)$. The derivative

$$\alpha'(t) = \lim_{h\to 0} \frac{\mathbf{r}(t+h,v) - \mathbf{r}(t,v)}{h} = \mathbf{r}_u(t,v).$$

Since the curve α lies in the surface, a tangent vector to α is a tangent vector to the surface. The partial derivative vectors give tangent vectors, which are tangents to the coordinate curves.

Definition 4.4. The *tangent plane* to a smooth patch of a surface at a point $p = \mathbf{r}(a,b)$ is the plane in $\mathbb{R}^3$ through $p \in \mathbf{r}(U)$ containing the vectors $\mathbf{r}_u(a,b)$ and $\mathbf{r}_v(a,b)$.

To write down the equation of the tangent plane compute the normal vector $\mathbf{n} = \mathbf{r}_u \wedge \mathbf{r}_v$ at (a,b). The equation $\mathbf{x}\cdot\mathbf{n} = \mathbf{r}(a,b)\cdot\mathbf{n}$ give the cartesian form of the tangent plane at $p = \mathbf{r}(a,b)$.

Example 4.5. Find the tangent plane to the sphere at the point $\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}\right)$.

The point in question occurs when $u = -\pi/4$ and $v = \pi/4$. Differentiating yields

$$\begin{aligned}\mathbf{r}_u &= (-\sin u \sin v, \cos u \sin v, 0)\\ \mathbf{r}_v &= (\cos u \cos v, \sin u \cos v, -\sin v)\end{aligned}$$

At the relevant point $\mathbf{r}_u(-1/2, 1/2) = (1/2, 1/2, 0)$ and $\mathbf{r}_v(-1/2,1/2) = (1/2, -1/2, -1/\sqrt{2})$. The tangent plane is the plane through $\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}\right)$ containing these two vectors. The normal vector is given by

$$\mathbf{r}_u \wedge \mathbf{r}_v = \left(-\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, -\frac{1}{2}\right).$$

Therefore the plane is given by the cartesian equation $-x + y - \sqrt{2}z = -2$.

Exercise 4.6. Find the cartesian equation for the tangent plane to the torus with $a = 5 - \sqrt{3}$, $b = 2$ at the point $(0, -5, 1)$.

Definition 4.7. A tangent vector at a point p in a patch of a surface Σ is a vector $v \in \mathbb{R}^3$ which is a linear combination of the vectors $\mathbf{r}_u(p)$ and $\mathbf{r}_v(p)$. The *tangent space* of Σ at p, $T_p\Sigma$, is the vector space consisting of all tangent vectors to Σ at p.

Definition 4.8. The *tangent bundle* of a smooth surface Σ is the set of pairs

$$T\Sigma := \{(p, \mathbf{v}) \mid \mathbf{v} \in T_p\Sigma\}$$

We should think of the tangent space at p as being a vector space whose origin lies at p, and a tangent vector as a vector whose tail is at p. For any surface and for any point p, $T_p\Sigma \cong \mathbb{R}^2$ abstractly. Recall any two 2-dimensional vector spaces are abstractly isomorphic. The map $\mathbf{r}$ enables us to define a specific isomorphism.

Example 4.9. Let $\mathbf{r}\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the inclusion as the x-y plane. For $p \in \mathbb{R}^2$, $T_p\mathbb{R}^2 \cong \mathbb{R}^2$. The isomorphism is canonical. Also there is a one to one correspondence $T\mathbb{R}^2 \cong \mathbb{R}^4$.

Let $\mathbf{r}\colon U \to \Sigma$ be a smooth patch and let $(p = (a,b), \mathbf{v}) \in T\mathbb{R}^2$, so that $\mathbf{v}$ is a vector in $T_{(a,b)}\mathbb{R}^2$. Define a map

$$\begin{array}{rcl} r_*\colon T_p\mathbb{R}^2 & \to & T_{\mathbf{r}(p)}\Sigma \\ (x,y) & \mapsto & x\mathbf{r}_u(a,b) + y\mathbf{r}_v(a,b) \end{array}$$

i.e. the Jacobian matrix acts on the tangent vectors. The parametrisation gives a basis of the tangent space.

Tangent planes are a linear approximation to a surface in the neighbourhood of a point. They enable us to use tools from linear algebra to make notions precise, and to measure distances, angles and areas, and to measure the curvature of a surface.

Exercise 4.10. Do Carmo Section 2-4, page 88. Exercises 2,3,4, 12, 16.

4.3. **Vector fields and covariant derivatives.** One useful notion is that of a vector field. This means a tangent vector at each point $p \in \Sigma$, which varies continuously. For example, one might have the velocity of the wind at each point, at an instance of time, as a vector field on the surface of a planet. For this definition, note that there is a canonical projection map $\pi\colon T\Sigma \to \Sigma$ which sends $(p, \mathbf{v}) \mapsto p$, forgetting the vector.

Definition 4.11. A vector field on Σ is a map $s\colon \Sigma \to T\Sigma$ such that $\pi \circ s = \mathrm{Id}_\Sigma$. (Such a map is also called a section of a bundle).

Example 4.12. On S^2, suppose there is a westerly wind, which is strong near the equator and light near the poles, we could describe this in terms of the first parametrisation with the vector field.

$$\mathbf{r}(u,v) \mapsto (\sin v, 0) \in T_{\mathbf{r}(u,v)}S^2.$$

Here the coordinates in the tangent space are given as usual in terms of the basis $\{\mathbf{r}_u, \mathbf{r}_v\}$.

Exercise 4.13. A zero of a vector field is a point where $s(p) = \mathbf{0} \in T_p\Sigma$. Define a vector field on the torus $(a = 2, b = 1)$ which has no zeroes. Define a vector field on the torus which has 4 zeroes. A good way to do this is to project the torus to the x coordinate and take the gradient of this function. Make a sketch of your vector field. What is the smallest number of zeroes you can achieve for a vector field on S^2?

Covariant differentiation is a fancy word for directional derivative, which was in Calculus 3. Let $\alpha\colon I \to \mathbb{R}^3$ be a curve and let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be a function. We want to differentiate f along α. For all $t \in I$ we have values $f(\alpha(t)) \in \mathbb{R}$, and we want to see how these change as we travel along α. Suppose that $\alpha(t) = (x(t), y(t), z(t))$. Then by the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}(f\circ\alpha)(t) = \frac{\partial f}{\partial x}(\alpha(t))x'(t) + \frac{\partial f}{\partial y}(\alpha(t))y'(t) + \frac{\partial f}{\partial z}(\alpha(t))z'(t) = \alpha'(t)\cdot\nabla f(\alpha(t)).$$

Note that this just depends on the tangent vector of α at t. We therefore define the directional derivative, or *covariant derivative* of f along a vector $\mathbf{v} \in \mathbb{R}^3$ by

$$\nabla_{\mathbf{v}} f = \mathbf{v}\cdot\nabla f.$$

This measures the rate of change of f along $\mathbf{v}$. If $|\mathbf{v}| = 1$ then this measures the rate of change of f in the direction defined by $\mathbf{v}$, whence the nomenclature directional derivative.

We can characterise the covariant derivative of f at p along $\mathbf{v} \in T_p\mathbb{R}^3$ as: $\nabla_{\mathbf{v}_p} f = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \alpha)(t)|_{t=t_0}$ where $\alpha\colon I \to \mathbb{R}^3$ is a curve with $\alpha(t_0) = p$ and $\alpha'(t_0) = \mathbf{v}_p$.

Now we can define the covariant derivative more generally of a vector field $\mathbf{w}\colon \mathbb{R}^3 \to T\mathbb{R}^3$ along another vector field $\mathbf{v}\colon \mathbb{R}^3 \to T\mathbb{R}^3$. Recall that for any p $T_p\mathbb{R}^3 \cong \mathbb{R}^3$ canonically, however it is preferably to think of them as separate, vectors starting at p and points p, respectively.

Define a new vector field $\nabla_{\mathbf{v}}\mathbf{w}$ by:

$$\big(\nabla_{\mathbf{v}}\mathbf{w}\big)(p) = \Big(\big(\nabla_{\mathbf{v_p}} w_1\big)(p), \big(\nabla_{\mathbf{v_p}} w_2\big)(p), \big(\nabla_{\mathbf{v_p}} w_3\big)\Big)$$

where $\mathbf{w} = (w_1, w_2, w_3)$ are the coordinates of $\mathbf{w}$.

Example 4.14. Find the covariant derivative of $\mathbf{w} = (x^2, y^2, z^2)$ along the tangent vector to the helix at time t. The helix is given by $\alpha(t) = (\cos t, \sin t, t)$. Its tangent vector is $\mathbf{v} = \alpha'(t) = (-\sin t, \cos t, 1)$. We have $w_1 = x^2$ so $\nabla w_1 = (2x, 0, 0)$ and

$$\nabla_{\alpha'(t)} w_1 = (2x, 0, 0) \cdot (-\sin t, \cos t, 1) = -2x \sin t = -2 \cos t \sin t.$$

Similarly

$$\nabla_{\alpha'(t)} w_2 = (0, 2y, 0) \cdot (-\sin t, \cos t, 1) = 2y \cos t = 2 \sin t \cos t$$

and

$$\nabla_{\alpha'(t)} w_3 = (0, 0, 2z) \cdot (-\sin t, \cos t, 1) = 2z = 2t.$$

Thus $(\nabla_{\mathbf{v}}\mathbf{w})(t) = (-2 \cos t \sin t, 2 \cos t \sin t, 2t)$.

Example 4.15. A nice example is given by taking $\mathbf{w} = (x, y, z)$, the coordinate vector field. Then

$$\nabla \mathbf{w} = (e_1, e_2, e_3) \cdot \mathbf{v} = (v_1, v_2, v_3) = \mathbf{v}.$$

The covariant derivative of $\mathbf{w}$ along $\mathbf{v}$ is again $\mathbf{v}$.

To take the covariant derivative of vector fields $a(u, v)\mathbf{r}_u + b(u, v)\mathbf{r}_v \in T_p\Sigma$, we will consider them as elements of $T\mathbb{R}^3$.

This does not pose too much difficulty. Let $\mathbf{w}\colon U \to T\mathbb{R}^3$ be a map where $\mathbf{w}(u, v) \in T_{\mathbf{r}(u,v)}\mathbb{R}^3$. A vector field $s\colon \Sigma \to T\Sigma$ determines such a map. Let's compute the covariant derivatives of $\mathbf{w}$ along the coordinate directions $\mathbf{r}_u$ and $\mathbf{r}_v$. We can use the coordinate curves

$$\alpha_u(t) = \mathbf{r}(u_0 + t, v_0)$$

and

$$\alpha_v(t) = \mathbf{r}(u_0, v_0 + t),$$

where $\mathbf{r}(u_0, v_0) = p$ is the point at which we want to compute covariant derivatives. We obtain:

$$\big(\nabla_{\mathbf{r}_u}\mathbf{w}\big)(u_0, v_0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \mathbf{w}(u_0 + t, v_0) = \frac{\partial \mathbf{w}}{\partial u}\big|(u_0, v_0)$$

and

$$\big(\nabla_{\mathbf{r}_v}\mathbf{w}\big)(u_0,v_0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\mathbf{w}(u_0,v_0+t) = \frac{\partial \mathbf{w}}{\partial v}\big|(u_0,v_0).$$

Thus as long as we always covariantly differentiate along tangent vectors we can just take the partial derivatives with respect to the smooth patch coordinates to compute. We need the following linearity formula, which shows how to take covariant derivative along any tangent vector $a\mathbf{r}_u + b\mathbf{r}_v$.

$$\nabla_{a\mathbf{r}_u+b\mathbf{r}_v}\mathbf{w} = a\nabla_{\mathbf{r}_u}\mathbf{w} + b\nabla_{\mathbf{r}_v}\mathbf{w}.$$

To prove this use the curve

$$\alpha_{(a,b)}(t) = \mathbf{r}(u_0+at, v_0+bt).$$

Then

$$\big(\nabla_{a\mathbf{r}_u+b\mathbf{r}_v}\mathbf{w}\big)(u_0,v_0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\mathbf{w}(u_0+at,v_0+bt) = a\cdot\frac{\partial \mathbf{w}}{\partial u}+b\cdot\frac{\partial \mathbf{w}}{\partial v} = a\cdot\nabla_{\mathbf{r}_u}\mathbf{w}+b\cdot\nabla_{\mathbf{r}_v}\mathbf{w}.$$

In this entire argument, there was no requirement for $\mathbf{w}$ to be tangent to Σ. For example, crucially, we can apply the argument to covariantly differentiate the normal vector.

Definition 4.16. The unit normal vector to a patch of a surface $\mathbf{r}\colon U \to \mathbb{R}^3$ is

$$\frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|}$$

Example 4.17. Consider the saddle surface $z = xy$. We let the smooth patch be $\mathbf{r}(u,v) = (u,v,uv)$. Compute that the unit normal

$$\mathbf{n}(u,v) = \frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|} = \frac{(-v,u,1)}{\sqrt{1+u^2+v^2}}.$$

Then the covariant derivatives of $\mathbf{n}$ along the coordinate curves are

$$\nabla_{\mathbf{r}_u}\mathbf{n} = \frac{\partial \mathbf{n}}{\partial u} = \frac{(uv,-1-v^2,-u)}{(1+u^2+v^2)^{3/2}}$$

and

$$\nabla_{\mathbf{r}_v}\mathbf{n} = \frac{\partial \mathbf{n}}{\partial v} = \frac{(-1-u^2,uv,-v)}{(1+u^2+v^2)^{3/2}}.$$

At the origin $\mathbf{r}(0,0) = (0,0,0)$ we have

$$\nabla_{\mathbf{r}_u}\mathbf{n}\big|_{(0,0)} = \frac{\partial \mathbf{n}}{\partial u}\big|_{(0,0)} = (0,-1,0) = -\mathbf{r}_v|_{(0,0)}$$

and

$$\nabla_{\mathbf{r}_v}\mathbf{n}\big|_{(0,0)} = \frac{\partial \mathbf{n}}{\partial v}\big|_{(0,0)} = (-1,0,0) = -\mathbf{r}_u|_{(0,0)}$$

This can be interpreted by understanding how the normal vector tilts when one starts at the origin and moves a small amount in the positive x and y directions.

Exercise 4.18. Check the computations in the previous example.

Exercise 4.19. Consider the sphere with the stereographic projection from the north pole. Compute the covariant derivative of the normal vector with respect to the coordinate vectors $\mathbf{r}_u$ and $\mathbf{r}_v$ at the south pole. Interpret your answer geometrically.

4.4. **The normal vector and the Gauss map.** As above:

Definition 4.20. The unit normal vector to a patch of a surface $\mathbf{r}\colon U \to \mathbb{R}^3$ is

$$\frac{\mathbf{r}_u \wedge \mathbf{r}_v}{|\mathbf{r}_u \wedge \mathbf{r}_v|}$$

There are two possible choices of unit normal vector; the parametrisation makes the choice for us. From now on when we talk about the normal vector we mean the unit normal and write $\mathbf{n}$.

The unit normal defines a map $G\colon \Sigma \to S^2$, called the *Gauss map*. By definition the Gauss map sends a point in Σ to its unit normal vector. By considering this as a vector based at the origin of $\mathbb{R}^3$, we obtain a point in S^2.

There are really two Gauss maps, one being the negative of the other. For a patch of a surface, the parametrisation fixes one. So we have a well defined map $G\colon U \to S^2$. Since there are no closed loops in the surface, there can be no Mobius bands, so this is well-defined. For non-orientable surfaces, the Gauss map can only be made well-defined as a map to $\mathbb{P}$. We will restrict to orientable surfaces.

Definition 4.21. An *orientation* of a surface Σ is a consistent choice of unit normal vector $\mathbf{n}_p$ for all $p \in \Sigma$. Here consistent means that the Gauss map defined by this choice is continuous. We call an orientable surface together with an orientation an *oriented* surface.

We will only need local properties of the Gauss map; we want to use it to measure how the surface changes near a point by comparing it to S^2. So as long as one fixes a choice of a Gauss map, that is fixes a choice of orientation, that will be fine. We claim that the Gauss map is well-defined i.e. independent of the choice of parametrisation, so is a function purely of the subset of $\mathbb{R}^3$. In other words, up to sign, all parametrisations yield the same unit normal vector.

4.5. **Change of parametrisation.** Suppose we have smooth patches $\mathbf{x}\colon U \to \mathbb{R}^3$ and $\mathbf{y}\colon V \to \mathbb{R}^3$ such that $W := \mathbf{x}(U) \cap \mathbf{y}(V) \neq \emptyset$ is an open subset of a surface Σ. Here U, V are open sets in $\mathbb{R}^2$. Since $\mathbf{x}$ and $\mathbf{y}$ are homeomorphisms onto their images, they have continuous inverses. We want to show that these inverses are smooth.

Proposition 4.22. *The homeomorphism $h := \mathbf{x}^{-1} \circ \mathbf{y}\colon \mathbf{y}^{-1}(W) \to \mathbf{x}^{-1}(W)$ is a diffeomorphism.*

A diffeomorphism is a homeomorphism which is differentiable and has differentiable inverse. We will need the *inverse function theorem*. Recall that a function $f\colon \mathbb{R} \to \mathbb{R}$ has an inverse where it is a monomorphism (1-1, injective). A function with nonzero derivative is *locally* a monomorphism. The following theorem is the higher dimensional version of this fact.

Theorem 4.23 (Inverse function theorem)**.** *Let $f\colon U \to \mathbb{R}^n$ be a continuously differentiable function, where U is an open subset of $\mathbb{R}^n$, and suppose $|Df|(a) \neq 0$ for some $a \in U$. Then there exists an open neighbourhood $V \ni f(a) =: b$ and a function $g\colon V \to U$ such that $g(b) = a$, $f \circ g(v) = v$ for all $v \in V$, g is continuously differentiable and $g(V) \subseteq U$ is an open neighbourhood of a.*

The conclusions imply that the restriction of f is a bijection between $g(V) \subseteq U$ and V. In the proof we frequently use the following argument. Suppose the derivative of a function is bounded above by $1/2$ in some region. Then by the mean value theorem, the difference in the value of the function at a and b in that region is bounded above by $\frac{1}{2}\|b-a\|$. Thus being able to control the derivative allows us to control a differentiable function.

Also recall the concept of a Cauchy sequence.

Before the proof we also define the *norm* of a linear transformation $A\colon \mathbb{R}^n \to \mathbb{R}^m$. Let

$$\|A\| := \sup\{\|A\xi\| \in \mathbb{R}^m \mid \xi \in \mathbb{R}^n,\ \|\xi\| = 1\}.$$

Intuitively, the norm $\|A\|$ is the farthest away from the origin that any point of the unit sphere gets stretched by the map A.

Theorem 4.24. *Let $f\colon U \to \mathbb{R}^n$ be a continuously differentiable function from a convex open subset $U \subset \mathbb{R}^n$ to $\mathbb{R}^n$ and let $x, y \in \mathbb{R}^n$. There exists $c \in U$ such that*

$$\|f(x) - f(y)\| = \|x - y\| \cdot \|Df(c)\|.$$

The following proof is quite sophisticated, and non-examinable.

Proof of Inverse function theorem. We may assume WLOG that $a = b = 0$ and that $Df(0) = I$, the $n \times n$ identity matrix. The assumption that $|Df(0)| = \det(Df(0)) \neq 0$ implies that $Df(0)$ is invertible. To see this, note that we may replace f by $\widetilde{f}\colon \mathbb{R}^n \to \mathbb{R}^n$ which sends

$$x \mapsto A^{-1}f(x+a) - b$$

where $A := Df(a)$. This replacement is valid for the following reasons.

First note that the functions $x \mapsto x+a$ and $x \mapsto A^{-1}x-b$ are invertible functions. Thus if the conclusion holds for the new function $\widetilde{f}$ it holds for the old one f too. Secondly, the hypothesis that $|Df(a)| \neq 0$ implies that $|D\widetilde{f}(0)| \neq 0$, because:

$$D\widetilde{f}(0) = D(A^{-1})(f(a)) \cdot Df(a) = A^{-1} \cdot Df(a) = A^{-1} \cdot A = I.$$

Here we used the fact that $D(A^{-1})(x) = A^{-1}$ for all x, on which we now elaborate. For any $n \times n$ matrix B we have a linear map $B\colon \mathbb{R}^n \to \mathbb{R}^n$. In coordinates this is given by

$$(x_1, \dots, x_n) \mapsto \left(\sum_{j=1}^n B_{1j}x_j, \dots, \sum_{j=1}^n B_{nj}x_j\right).$$

Therefore the derivative is given by the matrix

$$(DB)_{ik} = \left(\frac{\partial\left(\sum_{j=1}^n B_{ij}x_j\right)}{\partial x_k}\right)_{ik} = B_{ik}.$$

So we now assume that $a = b = 0$ and that $Df(0) = I$. Since Df is continuous, we may choose $\varepsilon > 0$ such that

$$\|Df(x) - I\| < \frac{1}{2}$$

whenever $\|x\| < \varepsilon$. To see this, think of Df as a function from $\mathbb{R}^n$ to the space of linear transformations on $\mathbb{R}^n$, and use the norm on linear transformations to define open balls. Once we have open balls we can define continuous, just like the beginning of chapter 3. Df is continuous at 0 means that for all $\tau > 0$ (choose $\tau = 1/2$ here), there exists $\varepsilon > 0$ such that

$$\|Df(x) - Df(0)\| = \|Df(x) - I\| < \tau = \frac{1}{2}$$

whenever $\|x\| < \varepsilon$.

Define the sets

$$X := \{x \in \mathbb{R}^n \mid \|x\| < \varepsilon\}$$

and

$$V := \{x \in \mathbb{R}^n \mid \|x\| < \frac{\varepsilon}{2}\}.$$

These sets will be important for the rest of the proof.

Fix $y \in V$, and define a function $\phi\colon U \to \mathbb{R}^n$ by

$$\phi(x) = x + y - f(x).$$

Note that $f(x) = y$ if and only if $\phi(x) = x$. So to define an inverse function we are looking for fixed points of ϕ. Note that $D\phi = I - Df$ so

$$\|D\phi(x)\| = \|(I - Df)(x)\| < \frac{1}{2}$$

for all $x \in X$. This will be a useful bound. Bounding the derivative enables bounds on a function in a neighbourhood, using the Mean Value theorem, as we will see. Intuitively, if the derivative is bounded a function can only change by so much in a small region.

We *claim* that $\phi(X) \subseteq X$ i.e. that $\|\phi(x)\| < \varepsilon$ for all $x \in X$. So see this let $x \in X$. Note that $\phi(0) = y - f(0) = y \in V$. The mean value theorem for multivariable functions implies that there exists $c \in X$ such that

$$\|\phi(x) - \phi(0)\| = \|x - 0\| \cdot \|D\phi(c)\|,$$

which implies that

$$\|\phi(x) - y\| = \|x\| \cdot \|D\phi(c)\| < \varepsilon \cdot \frac{1}{2}.$$

Therefore

$$\|\phi(x)\| = \|\phi(x) - 0\| \leq \|\phi(x) - y\| + \|y - 0\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $\phi(x) \in X$ as claimed.

Now define a sequence $\{x_k\}_{k=0}^{\infty}$ by $x_0 = 0$ and $x_k := \phi(x_{k-1})$. We have that

$$\|x_k - x_{k-1}\| = \|\phi(x_{k-1}) - \phi(x_{k-2})\| = \|x_{k-1} - x_{k-2}\| \cdot \|D\phi(c)\|$$

for some $c \in X$ by the Mean Value Theorem. Then $\|D\phi(c)\| < \frac{1}{2}$ so that

$$\|x_k - x_{k-1}\| \leq \frac{1}{2}\|x_{k-1} - x_{k-2}\|.$$

This says that $\{x_k\}$ is a *Cauchy sequence* (a type of convergent sequence where you don't have to know what the limit is to define convergence). Let

$$g(y) := \lim_{k\to\infty} x_k.$$

Since $x_k \in X$ for all X, $g(y) \in X$. We therefore have a map $g\colon V \to X$. Note that $g(y) = \phi(g(y))$ since $g(y)$ is the limit, so $f(g(y)) = y$. Also $g(0) = 0$: looking at the sequence with $y = 0$, we have $x_0 = 0$ and $\phi(x) = x - f(x)$, so $\phi(0) = 0 - f(0) = f(0) = 0$. Therefore $x_k = 0$ for all $k \in \mathbb{N} \cup \{0\}$ and $g(0) = 0$ as claimed.

Next we want to prove that g as we have defined it is continuously differentiable. Write $g(y) = x$. Given $k \in \mathbb{R}^n$ (small) let $h \in \mathbb{R}^n$ be such that $g(y + k) = x + h$. Therefore

$$f(x + h) = f \circ g(y + k) = y + k = f(x) + k.$$

Thus $f(x + h) - f(x) = k$, which implies that

$$f(x + h) - (x + h) - f(x) + x = k - h.$$

Taking norms

$$\begin{aligned}\|k - h\| &= \|f(x + h) - (x + h) - f(x) + x\| = \|(f - \mathrm{Id})(x + h) - (f - \mathrm{Id})(x)\| \\ &= \|(x + h) - x\| \cdot \|D(f - \mathrm{Id})(c)\|\end{aligned}$$

for some c by the mean value theorem. Thus $\|k - h\| \leq \|h\| \cdot \frac{1}{2}$.

Claim.

$$\frac{\|g(y + k) - g(y) - (Df(x))^{-1}(k)\|}{\|k\|} \to 0$$

as $k \to 0$.

(Think of $\frac{f(x+h)-f(x)}{h} - a = \frac{f(x+h)-f(x)-ha}{h} \to 0$ as $h \to 0$ if and only if $f'(x) = a$.)

Assume that the claim holds, and let $k = te_i$, where e_i is the ith standard basis vector of $\mathbb{R}^n$. Then $Df(x)^{-1}k = t(Df(x)^{-1})_i$, the ith column of $Df(x)^{-1}$. Since $t \to 0$ as $k \to 0$, we obtain the $\frac{\partial}{\partial x_i}$ partial derivative of the components of g. So g is continuously differentiable with $Dg(y) = Df(x)^{-1}$.

Proof of claim. We have

$$\begin{aligned}&g(y + k) - g(y) - Df(x)^{-1}k \\ =&x + h - x - Df(x)^{-1}(f(x + h) - f(x)) \\ =&- Dfx^{-1} \cdot \big(f(x + h) - f(x) - Df(x) \cdot h\big).\end{aligned}$$

Consider the function

$$t \mapsto f(x + th) - tDf(x) \cdot h$$

on $t \in [0, 1]$. The mean value theorem implies that there exists $c \in [0, 1]$ such that

$$\|f(x + h) - Df(x) \cdot h - f(x)\| = 1 \cdot \|Df(x + ch) \cdot h - Df(x) \cdot h\| \leq R(h)\|h\|$$

where

$$R(h) := \sup\{\|Df(x+th) - Df(x)\| \,|\, t \in [0,1]\}.$$

Therefore

$$\|g(y+k) - g(y) - Df(x)^{-1}k\| \leq R(h) \cdot \|Df(x)^{-1}\| \cdot \|h\| \leq R(h) \cdot \|Df(x)^{-1}\| \cdot 2\|k\|.$$

The last inequality uses that $\|h\| \leq 2\|k\|$ whose proof follows. Since $\|k - h\| \leq \frac{\|h\|}{2}$ we have

$$\|h\| = \|h - k + k\| \leq \|h - k\| + \|k\| \leq \frac{1}{2}\|h\| + \|k\|.$$

Rearranging yields $\|h\| \leq 2\|k\|$ as desired.

Now rearranging gives

$$\frac{\|g(y+k) - g(y) - Df(x)^{-1}k\|}{\|k\|} \leq 2R(h)\|Df(x)^{-1}\| \to 0$$

as $k \to 0$ since $R(h) \to 0$ as $h \to 0$ by continuity of Df, and since $h \to 0$ as $k \to 0$. This completes the proof of the claim. □

We check that $f|_X$ is injective. Suppose that $f(x) = f(x')$. Then $(f(x) - x) - f'(x) - x) = -(x - x')$. But $\|Df - I\| < \frac{1}{2}$ implies that

$$\|(f - \mathrm{Id})(x - x')\| \leq \|x - x'\|\frac{1}{2}.$$

Then $\|x - x'\| \leq \|x - x'\|\frac{1}{2}$ which implies that $\|x - x'\| = 0$ so $x = x'$ and f is injective.

Therefore $g(V) = X \cap f^{-1}(V)$. To see this, $g(V) \in X$ by definition of g. Also $g(V) \subseteq f^{-1}(V)$ since $f \circ g(v) = v$ for all $v \in V$. So $g(V) \subseteq X \cap f^{-1}(V)$; this we already knew. Suppose $x \in X$ and $x \in f^{-1}(V)$ i.e. $f(x) \in V$. We know $g \circ f(x) = y$ is such that $f(y) = f \circ g \circ f(x) = f(x)$. Therefore $y = x$ as f is injective. So $g(f(x)) = x$ and $x \in g(V)$. So $g(V) \subseteq X \cap f^{-1}(X)$, and indeed $g(V) = X \cap f^{-1}(V)$ as claimed.

Both X and $f^{-1}(V)$ are open (since f is continuous) so $g(V)$ is open, and $g(V) \ni 0$ since $g(0) = 0$, therefore $g(V)$ is an open neighbourhood of $a = 0 \in \mathbb{R}^n$ as required. This completes the proof of the inverse function theorem.

□

We now show that the coordinate change maps are diffeomorphisms.

Proof of Proposition 4.22. Let $r \in \mathbf{y}^{-1}(W)$ be a point, and define $q := h(r) \in \mathbf{x}^{-1}(W)$. Write $\mathbf{x}$ in the form

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

Since the derivative of $\mathbf{x}$ is injective at q, the wedge product of the columns does not vanish, so at least one of the determinants of the 2 by 2 minors are nonzero. These minors are Jacobians; the determinant of the minors are the coordinates of $\mathbf{x}_u \wedge \mathbf{x}_v$. Assume without loss of generality (WLOG) that it is:

$$\frac{\partial(x,y)}{\partial(u,v)}(q) \neq 0.$$

Now define a function

$$\begin{aligned} F\colon U\times\mathbb{R} &\to \mathbb{R}^3 \\ F(u,v,t) &= (x(u,v),y(u,v),z(u,v)+t). \end{aligned}$$

We make a map of the cylinder over U to the cylinder over $\mathbf{x}(U)$. Note that $F|_{U\times\{0\}}$ coincides with $\mathbf{x}$. The derivative of F is given by

$$DF = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & 0 \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & 0 \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & 1 \end{pmatrix}$$

The determinant of this at q is $\frac{\partial(x,y)}{\partial(u,v)}(q)\neq 0$, so by the inverse function theorem there is a local inverse F^{-1} to F in a neighbourhood of $q\in\mathbf{x}^{-1}(W)$. Let $M\subseteq W$, $M\ni\mathbf{x}(q)=\mathbf{y}(r)$ be the intersection of the neighbourhood given by the inverse function theorem with $W\subseteq\Sigma$. Choose $N\subseteq V$ with $N\ni r$ such that $\mathbf{y}(N)\subseteq M$. This is possible by continuity of $\mathbf{y}$. Since $F^{-1}|_M$ is an inverse to $\mathbf{x}$ we have

$$h|_N = F^{-1}\circ\mathbf{y},$$

and this is a composition of differentiable functions and is therefore differentiable. To see that h is differentiable it is enough to see that it is in small neighbourhoods. We may switch the rôles of $\mathbf{x}$ and $\mathbf{y}$ in this proof to see that h^{-1} is also differentiable, so that h is a diffeomorphism as claimed. □

The key ingredients for smooth changes in coordinates are the inverse function theorem and that we had regular parametrisations, so the Jacobian in question was nonzero. Note that we also used that our surface is a subset of $\mathbb{R}^3$. For abstract surfaces, smooth changes in coordinates is an assumption as part of the definition, for embedded surfaces it is a theorem.

Next, a change of coordinates acts on the tangent space in a specific way.

Proposition 4.25. *A change of coordinates $(\widetilde{u},\widetilde{v})=(\widetilde{u}(u,v),\widetilde{v}(u,v))$ induces a change of basis on the tangent space as follows. The vector $\mathbf{a}\in\mathbb{R}^2$ representing an element of $T_p\Sigma$ with respect to the basis $\{\mathbf{r}_u,\mathbf{r}_v\}$ is represented by the vector $J^T\mathbf{a}$ with respect to the basis $\{\widetilde{\mathbf{r}}_{\widetilde{u}},\widetilde{\mathbf{r}}_{\widetilde{v}}\}$ where*

$$J = \begin{pmatrix} \frac{\partial\widetilde{u}}{\partial u} & \frac{\partial\widetilde{v}}{\partial u} \\ \frac{\partial\widetilde{u}}{\partial v} & \frac{\partial\widetilde{v}}{\partial v} \end{pmatrix}$$

is the Jacobian matrix of the change of coordinates.

Proof. Suppose $\mathbf{r}(u,v)=\widetilde{\mathbf{r}}(\widetilde{u}(u,v),\widetilde{v}(u,v))$. Then by the chain rule

$$\mathbf{r}_u = \widetilde{\mathbf{r}}_{\widetilde{u}}\frac{\partial\widetilde{u}}{\partial u}+\widetilde{\mathbf{r}}_{\widetilde{v}}\frac{\partial\widetilde{v}}{\partial u}$$

and

$$\mathbf{r}_v = \widetilde{\mathbf{r}}_{\widetilde{u}}\frac{\partial\widetilde{u}}{\partial v}+\widetilde{\mathbf{r}}_{\widetilde{v}}\frac{\partial\widetilde{v}}{\partial v}.$$

□

This is a great illustration of the connection between calculus and linear algebra. The derivative matrix is a linear map which acts on tangent spaces. This proposition will be very useful when we are investigating the effect of a change of coordinates. Only change of coordinates whose Jacobian matrix has nonzero determinant are permitted: otherwise one of the patches would not be regular.

Exercise 4.26. Compute the Jacobian matrix for the change of coordinates on $\mathbb{R}^2 \setminus \{0\}$ given by $\mathbf{r}(x,y) = (x,y)$ and $\mathbf{r}(r,\theta) = (r\cos\theta, r\sin\theta)$, with $u = r, v = \theta, \widetilde{u} = x, \widetilde{v} = y$

Corollary 4.27. *The derivatives satisfy:*

$$D\mathbf{r} = (D\widetilde{\mathbf{r}}) \cdot J^T.$$

Since the Gauss map gives a normal vector to the plane determined by the columns of the derivative matrix, we see that the plane is unchanged by change of coordinates if $\det(J^T) = \det(J) \neq 0$. In addition recall that the Gauss map requires a choice of normal vector, which is determined by the order of the basis vectors $\mathbf{r}_u$ and $\mathbf{r}_v$. If $\det(J) > 0$, the normal vector determined by this order is unchanged. For an *oriented* surface, only changes of coordinates whose Jacobian has positive determinant are allowed.

This is a common useful trick in differential geometry: the fewer change of coordinate functions there are, the easier it is for a function on a surface to be well-defined.

To elaborate:

Definition 4.28. A function on a surface $f\colon \Sigma \to \mathbb{R}$ is a function $f_{\mathbf{r}}\colon U \to \mathbb{R}$ for each smooth patch $\mathbf{r}$, such that for any two coordinate patches $\mathbf{r}$ and $\widetilde{\mathbf{r}}$, if we define

$$h = \widetilde{\mathbf{r}}^{-1} \circ \mathbf{r}\colon \mathbf{r}^{-1}(\mathbf{r}(U) \cap \widetilde{\mathbf{r}}(\widetilde{U})) \to \widetilde{\mathbf{r}}^{-1}(\mathbf{r}(U) \cap \widetilde{\mathbf{r}}(\widetilde{U}))$$

then we have

$$f_{\mathbf{r}}(u,v) = f_{\widetilde{\mathbf{r}}} \circ h(u,v) \in \mathbb{R}$$

for all $(u,v) \in \mathbf{r}^{-1}(\mathbf{r}(U) \cap \widetilde{\mathbf{r}}(\widetilde{U})) \subset U$.

The idea is that we can define a function by defining it on all smooth patches, provided the definitions agree on the overlaps. Actually there was no need for the codomain to be $\mathbb{R}$ here. It could even be another smooth surface.

Let Σ_1, Σ_2 be surfaces and let $\mathbf{r}_1\colon U_1 \to \Sigma_1$, $\mathbf{r}_2\colon U_2 \to \Sigma_2$ be smooth patches for the surfaces. Define $W := \mathbf{r}_1(U_1)$ and $V := \mathbf{r}_2(U_2)$ be the overlap of the smooth patches.

A function $f\colon \Sigma_1 \to \Sigma_2$ is smooth if the function

$$g := \mathbf{r}_2^{-1} \circ f \circ \mathbf{r}_1\colon \mathbf{r}_1^{-1}(W \cap f^{-1}(V)) \to \mathbf{r}_2^{-1}(f(W) \cap V)$$

is smooth for each pair of smooth patches $\mathbf{r}_1$, $\mathbf{r}_2$.

We can talk about differentiability of functions $\mathbb{R}^2 \to \mathbb{R}^2$, so we use this to define differentiability of functions on surfaces. A *diffeomorphism* between surfaces is a homeomorphism which is smooth and has smooth inverse. We say that two surfaces are diffeomorphic if there is a diffeomorphism between them. This is the natural equivalence relation to consider on smooth surfaces, just as homeomorphism was the natural equivalence relation on topological surfaces. Every smooth surface is a topological surface and every diffeomorphism is a homeomorphism, but the converses of these statements are false.

Exercise 4.29. Do Carmo Section 2–3. Questions 1,3,4,7,8.

4.6. **Linear Algebra review.**

Definition 4.30. A symmetric bilinear form on a vector space V is a function $B\colon V \times V \to \mathbb{R}$ such that for all $u, v, w \in V$ and for all $\lambda, \mu \in \mathbb{R}$, we have:

(i) $B(\lambda v + \mu w, u) = \lambda B(v, u) + \mu B(w, u)$; and
(ii) $B(v, w) = B(w, v)$.

A form is *nondegenerate* if $B(v, v) = 0$ if and only if $v = 0$. A form is *(positive/negative) definite* if $(B(v, v) > 0/B(v, v) < 0)$ for all $v \neq 0$. A positive definite symmetric bilinear form (which in particular is nondegenerate) is called an inner product on V. We write $B(v, w) = \langle v, w\rangle$.

Exercise 4.31. Prove from the axioms that a bilinear form is also linear in the second variable.

The length of a vector $v \in V$ is $\|v\| := \langle v, v\rangle^{\frac{1}{2}}$. The angle between two nonzero vectors $v, w \in V$ is

$$\cos^{-1}\left(\frac{\langle v, w\rangle}{\|v\|\|w\|}\right).$$

The area of a parallelogram defined by linearly independent vectors v, w in $\mathbb{R}^3$ is $|v \wedge w|$. By putting the structure of an inner product on each tangent space of a surface, we will be able to compute lengths, angles and areas in a surface.

4.7. **First Fundamental Form.** A fundamental form on a surface Σ is an inner product $B : T_p\Sigma \times T_p\Sigma \to \mathbb{R}$ for each $p \in \Sigma$, such that the form varies smoothly as p varies. With respect to the basis given by $\mathbf{r}_u, \mathbf{r}_v$, we can write B as a matrix whose coordinates are functions of u and v. We require that these are smooth functions.

One has to be careful when one defines a notion, how much that notion depends on the choice of coordinate system. A closely related anxiety is whether a notion associated to a vector space depends on a choice of basis for that vector space.

Let $(a, b) \in \mathbb{R}^2$ and let $r_*\colon T_{(a,b)}\mathbb{R}^2 \to T_{\mathbf{r}(a,b)}\Sigma$ be the map defined above. We require, for all pairs of vectors $x, y \in \mathbb{R}^2$ (which defines vectors $x, y \in T_{(a,b)}\mathbb{R}^2$ since each tangent space is canonically identified with $\mathbb{R}^2$), that:

$$(a, b) \mapsto B(r_*(x), r_*(y))$$

is a smooth function $\mathbb{R}^2 \to \mathbb{R}$. It can be checked that if this holds for one smooth parametrisation of a surface, then it holds for all of them.

Now we define the first fundamental form. This is one of the most important objects of the course. (It is a special case of a Riemannian metric, for those interested in Riemannian geometry.)

Suppose I want to compute the length of a curve in a surface. Let $\gamma(t) = (u(t), v(t))$ be a curve in $\mathbb{R}^2$. Then $\alpha(t) = \mathbf{r}(\gamma(t)) = \mathbf{r}(u(t), v(t))$ is a curve in Σ.

Differentiating yields

$$\alpha'(t) = \mathbf{r}_u(u(t), v(t))u'(t) + \mathbf{r}_v(u(t), v(t))v'(t)$$

so

$$|\alpha'(t)| = \sqrt{\alpha' \cdot \alpha'} = \sqrt{\mathbf{r}_u \cdot \mathbf{r}_u (u')^2 + 2\mathbf{r}_u \cdot \mathbf{r}_v u'v' + \mathbf{r}_v \cdot \mathbf{r}_v (v')^2}.$$

One integrates this function of t to compute the length of α. Define functions $U \to \mathbb{R}$

$$E(u, v) := \mathbf{r}_u \cdot \mathbf{r}_u$$
$$F(u, v) := \mathbf{r}_u \cdot \mathbf{r}_v$$
$$G(u, v) := \mathbf{r}_v \cdot \mathbf{r}_v.$$

So to compute the length of a curve in a surface α over the interval I, we integrate

$$L(\alpha) = \int_{t \in I} \sqrt{E(u')^2 + 2Fu'v' + G(v')^2} \mathrm{d}t$$

Here E, F and G depend on the surface, and u', v' came from the original curve γ in $\mathbb{R}^2$.

The functions E, F and G are called the *coefficients of the first fundamental form.* With respect to the basis $\mathbf{r}_u$, $\mathbf{r}_v$ for a tangent space, the first fundamental form is represented by the matrix

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

This means that for vectors $x = a\mathbf{r}_u + b\mathbf{r}_v$ and $y = c\mathbf{r}_u + d\mathbf{r}_v$ we have

$$B(x, y) = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = Eac + F(ad + bc) + Gbd$$

If $a = c$ and $b = d$ we have

$$B(x, x) = Ea^2 + 2Fab + Gb^2.$$

We record the first fundamental form *formally* by the notation

$$E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2.$$

Example 4.32. (a) Let's compute the first fundamental form of the sphere.

$$\mathbf{r}_u = (-\sin u \sin v, \cos u \sin v, 0)$$
$$\mathbf{r}_v = (\cos u \cos v, \sin u \cos v, -\sin v)$$

Then

$$E = \mathbf{r}_u \cdot \mathbf{r}_u = \sin^2 u \sin^2 v + \cos^2 u \sin^2 v = \sin^2 v$$
$$F = \mathbf{r}_u \cdot \mathbf{r}_v = -\sin u \sin v \cos u \cos v + \cos u \sin v \sin u \cos v = 0$$
$$G = \mathbf{r}_v \cdot \mathbf{r}_v = \cos^2 u \cos^2 v + \sin^2 u \cos^2 v + \sin^2 v = 1.$$

Therefore the first fundamental form is

$$\sin^2 v \mathrm{d}u^2 + \mathrm{d}v^2.$$

(b) For the plane we have $\mathbf{r}(u, v) = (u, v, 0)$ and $\mathbf{r}_u = (1, 0, 0)$, $\mathbf{r}_v = (0, 1, 0)$. So $E = G = 1$ and $F = 0$. The first fundamental form is

$$\mathrm{d}u^2 + \mathrm{d}v^2.$$

(c) For the plane in polar coordinates $\mathbf{r}(r,\theta) = (r\cos\theta, r\sin\theta, 0)$, we have

$$\mathbf{r}_r = (\cos\theta, \sin\theta, 0)$$
$$\mathbf{r}_\theta = (-r\sin\theta, r\cos\theta, 0)$$

so that $E = 1$, $G = r^2$ and $F = 0$ so that the first fundamental form is

$$\mathrm{d}r^2 + r^2\mathrm{d}\theta.$$

(d) The cone $\mathbf{r}(u,v) = (au\cos v, au\sin v, u)$. $u > 0, 0 < v < 2\pi$. We have

$$\mathbf{r}_u = (a\cos v, a\sin v, 1)$$
$$\mathbf{r}_v = (-au\sin v, au\cos v, 0)$$

so that $E = 1 + a^2$, $F = 0$ and $G = a^2u^2$ so we have

$$(1+a^2)\mathrm{d}u^2 + a^2u^2\mathrm{d}v^2$$

for the first fundamental form.

Exercise 4.33. Compute the first fundamental forms of the following surfaces:

(i) The hemisphere with parametrisation $(u,v) \mapsto (u, v, \sqrt{1-u^2-v^2})$, $\|(u,v)\| < 1$.

(ii) The surface of revolution

$$\mathbf{r}(u,v) = (f(u)\cos v, f(u)\sin v, u).$$

For some differentiable function $f\colon \mathbb{R} \to \mathbb{R}_{>0}$. $(u,v) \in \mathbb{R} \times (0, 2\pi)$.

(iii) The torus with $a = 3, b = 1$.

We have seen how to use the first fundamental form to compute lengths of curves on a surface. Next we will use it to compute angles between intersecting curves and areas of subsets of a surface.

4.8. **Angles.** Let $\alpha\colon I \to \mathbb{R}^2$ and $\gamma\colon J \to \mathbb{R}^2$ define curves on a surface patch $\mathbf{r}\colon U \to \mathbb{R}^3$. Suppose that $\alpha(0) = \gamma(0) = p$, so that the curves in Σ intersect at $\mathbf{r}(p) = q \in \Sigma$. The curves $\mathbf{r}(\alpha(t))$ and $\mathbf{r}(\gamma(t))$ in Σ can be expressed as

$$\mathbf{r}(u_1(t), v_1(t)) \text{ and } \mathbf{r}(u_2(t), v_2(t))$$

respectively. The tangent vectors are

$$\mathbf{r}_u u_1' + \mathbf{r}_v v_1' \text{ and } \mathbf{r}_u u_2' + \mathbf{r}_v v_2'.$$

Therefore the angle between the curves at their point of intersection is

$$\begin{aligned}
\theta &= \cos^{-1}\left(\frac{(\mathbf{r}_u u_1' + \mathbf{r}_v v_1')\cdot(\mathbf{r}_u u_2' + \mathbf{r}_v v_2')}{\|\mathbf{r}_u u_1' + \mathbf{r}_v v_1'\|\,\|\mathbf{r}_u u_2' + \mathbf{r}_v v_2'\|}\right)\\
&= \cos^{-1}\left(\frac{Eu_1'u_2' + F(u_1'v_2 + u_2'v_1') + Gv_1'v_2'}{\sqrt{E(u_1')^2 + 2Fu_1'v_1' + G(v_1')^2}\sqrt{E(u_2')^2 + 2Fu_2'v_2' + G(v_2')^2}}\right)\\
&= \cos^{-1}\left(\frac{\begin{pmatrix} u_1' & v_1'\end{pmatrix}\begin{pmatrix} E & F\\ F & G\end{pmatrix}\begin{pmatrix} u_2'\\ v_2'\end{pmatrix}}{\sqrt{\begin{pmatrix} u_1' & v_1'\end{pmatrix}\begin{pmatrix} E & F\\ F & G\end{pmatrix}\begin{pmatrix} u_1'\\ v_1'\end{pmatrix}}\sqrt{\begin{pmatrix} u_2' & v_2'\end{pmatrix}\begin{pmatrix} E & F\\ F & G\end{pmatrix}\begin{pmatrix} u_2'\\ v_2'\end{pmatrix}}}\right)
\end{aligned}$$

The purpose of the last formula is so that we will be able to see that the angle is independent of the choice of coordinate system i.e. is the same for all smooth patches representing the same surface in $\mathbb{R}^3$.

4.9. **Areas.** Recall that the area of the parallelogram defined by linearly independent vectors $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a}\wedge\mathbf{b}|$. This holds since $|\mathbf{a}\wedge\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta$ where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

Thus the size of a small area element at a point $\mathbf{r}((a,b))$ of a smooth patch $\mathbf{r}\colon U \to \mathbb{R}^3$ of a surface is given by $|\mathbf{r}_u \wedge \mathbf{r}_v|(a,b)\mathrm{d}u\mathrm{d}v$. Let C be a region of a smooth patch $\mathbf{r}$. Then the area of C is given by

$$\int_{\mathbf{r}^{-1}(C)} |\mathbf{r}_u \wedge \mathbf{r}_v| \,\mathrm{d}u\mathrm{d}v.$$

Recall the vector identity

$$|\mathbf{r}_u \wedge \mathbf{r}_v|^2 = (\mathbf{r}_u\cdot\mathbf{r}_u)(\mathbf{r}_v\cdot\mathbf{r}_v) - (\mathbf{r}_u\cdot\mathbf{r}_v)(\mathbf{r}_u\cdot\mathbf{r}_v) = EG - F^2.$$

Therefore the area of C can be computed in terms of the first fundamental form as

$$\int_{\mathbf{r}^{-1}(C)} \sqrt{EG-F^2}\,\mathrm{d}u\mathrm{d}v = \int_{\mathbf{r}^{-1}(C)} \sqrt{\det\begin{pmatrix} E & F \\ F & G\end{pmatrix}}\,\mathrm{d}u\mathrm{d}v.$$

Exercise 4.34. (i) Compute the area of the torus with $a = 3$ and $b = 1$.
(ii) Compute the area of the sphere with radius r using the first fundamental form.
(iii) Give an expression for the area of a surface of revolution for $u \in [a,b]$.

Exercise 4.35. Do Carmo Section 2–5, questions 1,2,5,6,10,11,12.

4.10. **Effect of change of coordinates on first fundamental form.**

Exercise 4.36. Suppose that a bilinear form is represented with respect to the standard basis of $\mathbf{R}^2$ by the matrix $\begin{pmatrix} 2 & 1 \\ 1 & 3\end{pmatrix}$. Find the matrix representing the form with respect to the basis $\{(1,2),(-1,1)\}$ for $\mathbb{R}^2$.

Since

$$\mathbf{r}_u = \widetilde{\mathbf{r}}_{\widetilde{u}}\frac{\partial\widetilde{u}}{\partial u} + \widetilde{\mathbf{r}}_{\widetilde{v}}\frac{\partial\widetilde{v}}{\partial u}$$

and

$$\mathbf{r}_v = \widetilde{\mathbf{r}}_{\widetilde{u}}\frac{\partial\widetilde{u}}{\partial v} + \widetilde{\mathbf{r}}_{\widetilde{v}}\frac{\partial\widetilde{v}}{\partial v}.$$

we have that

$$\begin{aligned} E = \mathbf{r}_u\cdot\mathbf{r}_u &= \left(\frac{\partial\widetilde{u}}{\partial u}\right)^2 \widetilde{\mathbf{r}}_{\widetilde{u}}\cdot\widetilde{\mathbf{r}}_{\widetilde{u}} + 2\frac{\partial\widetilde{v}}{\partial u}\frac{\partial\widetilde{u}}{\partial v}\widetilde{\mathbf{r}}_{\widetilde{u}}\cdot\widetilde{\mathbf{r}}_{\widetilde{v}} + \left(\frac{\partial\widetilde{v}}{\partial v}\right)^2 \widetilde{\mathbf{r}}_{\widetilde{v}}\cdot\widetilde{\mathbf{r}}_{\widetilde{v}} \\ &= \left(\frac{\partial\widetilde{u}}{\partial u}\right)^2 \widetilde{E} + 2\frac{\partial\widetilde{v}}{\partial u}\frac{\partial\widetilde{u}}{\partial v}\widetilde{F} + \left(\frac{\partial\widetilde{v}}{\partial v}\right)^2 \widetilde{G}. \end{aligned}$$

Similarly for F and G. We eventually obtain the simple matrix formula:

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} = \begin{pmatrix} \frac{\partial \widetilde{u}}{\partial u} & \frac{\partial \widetilde{v}}{\partial u} \\ \frac{\partial \widetilde{u}}{\partial v} & \frac{\partial \widetilde{v}}{\partial v} \end{pmatrix} \begin{pmatrix} \widetilde{E} & \widetilde{F} \\ \widetilde{F} & \widetilde{G} \end{pmatrix} \begin{pmatrix} \frac{\partial \widetilde{u}}{\partial u} & \frac{\partial \widetilde{v}}{\partial u} \\ \frac{\partial \widetilde{u}}{\partial v} & \frac{\partial \widetilde{v}}{\partial v} \end{pmatrix}^T = J \begin{pmatrix} \widetilde{E} & \widetilde{F} \\ \widetilde{F} & \widetilde{G} \end{pmatrix} J^T$$

This is seen to be consistent with the change of basis matrix from Proposition 4.25 and the way that bases change for bilinear forms, as in the exercise above.

We can use this change of coordinate formula to see that the length, angle and area formulas given above, in terms of the coefficients of the first fundamental form, are independent of the coordinate system. That is, if we change coordinates to a different smooth patch for the same underlying surface, the length of a curve, the angle between curves, and the area of a region are measured to be the same in the new and old coordinate systems.

To see this, replace $\begin{pmatrix} E & F \\ F & G \end{pmatrix}$ in the formulae with $J \begin{pmatrix} \widetilde{E} & \widetilde{F} \\ \widetilde{F} & \widetilde{G} \end{pmatrix} J^T$ and observe for the length and angle formulae that

$$J^T \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} \widetilde{u}' \\ \widetilde{v}' \end{pmatrix}.$$

For the area formula,

$$\det \left(J \begin{pmatrix} \widetilde{E} & \widetilde{F} \\ \widetilde{F} & \widetilde{G} \end{pmatrix} J^T \right) = (\det J)^2 (\widetilde{E}\widetilde{G} - \widetilde{F}^2),$$

so that we have

$$\int_{\widetilde{\mathbf{r}}^{-1}(C)} \sqrt{\widetilde{E}\widetilde{G} - \widetilde{F}^2} \det(J) \mathrm{d}u \mathrm{d}v = \int_{\widetilde{\mathbf{r}}^{-1}(C)} \sqrt{\widetilde{E}\widetilde{G} - \widetilde{F}^2} \mathrm{d}\widetilde{u} \mathrm{d}\widetilde{v}$$

by the change of variable formula for double integration.

4.11. **Isometries and local isometries.** We have seen that the first fundamental form transforms in a certain way under change of variables, and that this governs lengths, angles and areas. In geometry, we consider two surfaces to be "the same" if they are diffeomorphic; that is there is a smooth homeomorphism between them which has a smooth inverse, and if the diffeomorphism preserves lengths of curves.

Definition 4.37. An *isometry* between two surfaces Σ_1, Σ_2 is a *diffeomorphism* $f \colon \Sigma_1 \to \Sigma_2$ which sends smooth paths in Σ_1 to smooth paths in Σ_2 of the same length.

So if $\gamma \colon I \to \Sigma_1$ is a smooth path then $L(\gamma) = L(f \circ \gamma)$.

Definition 4.38. Let $p \in \Sigma_1$ be a point with an open neighbourhood $V_1 \subseteq \Sigma_1$. A map $f \colon V_1 \to \Sigma_2$ is a local isometry $f \colon V_1 \to f(V_1)$ is an isometry.

If f is a diffeomorphism and a local isometry for all $p \in \Sigma_1$, then f is an isometry.

Next we show that it is equivalent that a local isometry f sends the first fundamental form of Σ_1 to the first fundamental form of Σ_2. Thus the mapping should preserve lengths, angles and areas. The first fundamental form is a local property, so it is preserved under local isometries.

Theorem 4.39. *Two* Σ_1 *and* Σ_2 *are locally isometric if and only if for each point* $p \in \Sigma_1$ *there is an open set* $V \in \mathbb{R}^2$ *and smooth patches* $\mathbf{r}\colon V \to \Sigma_1$, $\widetilde{\mathbf{r}}\colon V \to \Sigma_2$ *with* $p \in \mathbf{r}(V)$ *such that the first fundamental forms of* Σ_1 *and* Σ_2*, with respect to* $\mathbf{r}$ *and* $\widetilde{\mathbf{r}}$ *respectively, are the same.*

Proof. The if direction. We claim that the local isometry is given by $f = \widetilde{\mathbf{r}} \circ \mathbf{r}^{-1}\colon \Sigma_1 \to \Sigma_2$. Let $\gamma\colon [a,b] \to \Sigma_1$ be a smooth curve in Σ_1 with $\gamma(t) = \mathbf{r}(u(t), v(t))$. Then $f \circ \gamma(t) = \widetilde{\mathbf{r}}(u(t), v(t))$. Therefore

$$L(\gamma) = \int_a^b \sqrt{E(u')^2 + 2Fu'v' + G(v')^2}\, \mathrm{d}t = L(f \circ \gamma).$$

So the length of curves is preserved and f is a local isometry as claimed.

The only if direction. This direction is a little harder. Suppose $f\colon \Sigma_1 \to \Sigma_2$ is a local isometry and let $\mathbf{r}\colon V \to \Sigma_1$ be a smooth patch, with $p \in V$. Define $\widetilde{\mathbf{r}}\colon V \to \Sigma_2$ by $\widetilde{\mathbf{r}} = f \circ \mathbf{r}$. Also define

$$\begin{aligned} E &= \mathbf{r}_u \cdot \mathbf{r}_u \\ F &= \mathbf{r}_u \cdot \mathbf{r}_v \\ G &= \mathbf{r}_v \cdot \mathbf{r}_v \\ \widetilde{E} &= \widetilde{\mathbf{r}}_u \cdot \widetilde{\mathbf{r}}_u \\ \widetilde{F} &= \widetilde{\mathbf{r}}_u \cdot \widetilde{\mathbf{r}}_v \\ \widetilde{G} &= \widetilde{\mathbf{r}}_v \cdot \widetilde{\mathbf{r}}_v \end{aligned}$$

Since f is a local isometry by assumption, we have

$$\int_a^b \sqrt{E(u')^2 + 2Fu'v' + G(v')^2}\, \mathrm{d}t = \int_a^b \sqrt{\widetilde{E}(u')^2 + 2\widetilde{F}u'v' + \widetilde{G}(v')^2}\, \mathrm{d}t$$

for all smooth paths $t \mapsto (u(t), v(t))$ in V. Choose

$$(u(t), v(t)) = (u_0 + t, v_0).$$

By the fundamental theorem of calculus,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^\varepsilon \sqrt{E(u_0 + t, v_0)}\mathrm{d}t = \sqrt{E(u_0, v_0)}$$

and

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^\varepsilon \sqrt{\widetilde{E}(u_0 + t, v_0)}\mathrm{d}t = \sqrt{\widetilde{E}(u_0, v_0)}.$$

As above, the two integrals are equal for all $\varepsilon \neq 0$, so we have $\sqrt{E(u_0, v_0)} = \sqrt{\widetilde{E}(u_0, v_0)}$, which implies that $\widetilde{E} = E$. Similarly taking the path $(u(t), v(t)) = (u_0, v_0 + t)$ yields $\widetilde{G} = G$. Then taking $(u(t), v(t)) = (u_0 + t, v_0 + t)$ implies that

$$E + 2F + G = \widetilde{E} + 2\widetilde{F} + \widetilde{G},$$

so that $\widetilde{F} = F$ as well. □

Example 4.40. The cylinder and the plane are locally isometric but not isometric. To see that the cylinder and the plane are locally isometric compute the first fundamental form of the cylinder with parametrisation $\mathbf{r}(u,v) = (\cos u, \sin u, v)$, $(u,v) \in (0,2\pi) \times \mathbb{R}$ to be $\mathrm{d}u^2 + \mathrm{d}v^2$, which is the same as that of the subset of the plane with the parametrisation $\widetilde{\mathbf{r}}(\widetilde{u},\widetilde{v}) = (\widetilde{u},\widetilde{v},0)$, $(\widetilde{u},\widetilde{v}) \in (0,2\pi) \times \mathbb{R}$. They are not homeomorphic, so cannot be diffeomorphic and therefore cannot be isometric. To see this we'll consider the compact surfaces: for the plane use the domain $[0,2\pi] \times [0,1]$ and for the cylinder use the domain $[[0,2\pi) \times [0,1]$. We can compute the Euler characteristic of the plane to be 1 and the Euler characteristic of the cylinder to be 0. Thus they cannot be homeomorphic.

Example 4.41. The cone is given by

$$\mathbf{r}(u,v) = (au\cos v, au\sin v, u)$$

We compute

$$\mathbf{r}_u = (a\cos v, a\sin v, 1)$$
$$\mathbf{r}_v = (-au\sin v, au\cos v, 0)$$

so that

$$E = 1 + a^2;\ F = 0 \text{ and } G = a^2u^2.$$

The first fundamental form is

$$(1+a^2)\mathrm{d}u^2 + a^2u^2\mathrm{d}v^2.$$

Make the change of coordinates

$$(u,v) \mapsto \left(\sqrt{1+a^2}u, \sqrt{\frac{a^2}{1+a^2}}v\right) = (r,\theta).$$

Thus $u = \frac{r}{\sqrt{1+a^2}}$ and $v = \sqrt{\frac{1+a^2}{a^2}}\theta$.

We can compute a formal substitution for the first fundamental form using the chain rule:

$$\mathrm{d}u = \frac{\partial u}{\partial r}\mathrm{d}r + \frac{\partial u}{\partial \theta}\mathrm{d}\theta = \frac{1}{\sqrt{1+a^2}}\mathrm{d}r$$

and

$$\mathrm{d}v = \frac{\partial v}{\partial r}\mathrm{d}r + \frac{\partial v}{\partial \theta}\mathrm{d}\theta = \sqrt{\frac{1+a^2}{a^2}}\mathrm{d}\theta.$$

After making these substitutions the first fundamental form becomes:

$$\mathrm{d}r^2 + r^2\mathrm{d}\theta^2$$

which is the first fundamental form of the plane in polar coordinates.

Note that when $v \in (0,2\pi)$, the variable $\theta \in \left(0, 2\pi\sqrt{\frac{a^2}{1+a^2}}\right)$. Thus there is an isometry between this part of the plane and the cone with $u > 0, v \in (0,2\pi)$.

The cone and the plane should be isometric, since we can make the cone from a sheet of paper. You couldn't make a sphere or a torus from paper unless it was very stretchy paper. The stretching would not constitute an isometry.

Exercise 4.42. The catenoid has a smooth patch (which misses out one meridian)

$$\mathbf{r}(u, v) = (a \cosh v \cos u, a \cosh v \sin u, av)$$

$(u, v) \in (0, 2\pi) \times \mathbb{R}$. The helicoid has a smooth patch

$$\mathbf{x}(r, t) = (t \cos r, t \sin r, ar)$$

$(r, t) \in (0, 2\pi) \times \mathbb{R}$.

(i) Compute the first fundamental forms of these surfaces.
(ii) Find the first fundamental form of the helicoid after the change of coordinates $(r, s) = (u, a \sinh v)$.
(iii) Show that the smooth patches are isometric.

(Since the helicoid is really for all $r \in \mathbb{R}$, this is a local isometry between the helicoid and the catenoid. But restricted to the given patches, it's an isometry.

Exercise 4.43. A smooth patch is said to be conformal if it preserves angles from $\mathbb{R}^2$, and is equiareal if it preserves areas.

(a) Show that a smooth patch is conformal if and only if $E = G$ and $F = 0$.
(b) Show that a smooth patch is equiareal if and only if $EG - F^2 = 1$.
(c) Consider the smooth patch on the sphere

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

$(\theta, \phi) \in (-\pi, \pi) \times (-\pi/2, \pi/2)$. Show that this is neither conformal nor equiareal.
(d) The Mercator projection of the sphere with the Date Line removed takes a point $\mathbf{r}(\theta, \phi)$ to

$$\Big(\theta, \log\Big(\tan\Big(\frac{\phi}{2} + \frac{\pi}{4}\Big)\Big)\Big) \in (-\pi, \pi) \times \mathbb{R}.$$

Show that the Mercator projection is conformal but not equiareal. Hint: change coordinates from (c) to

$$(u, v) = (\theta, \log\Big(\tan\Big(\frac{\phi}{2} + \frac{\pi}{4}\Big)\Big)).$$

(e) Lambert's cylindrical projection (projecting a sphere outwards to a cylinder) takes a point $\mathbf{r}(\theta, \phi)$ to $(\theta, \sin\phi)$. Show that this is equiareal.

References

[DoCa] Manfredo P. DoCarmo *Differential Geometry of Curves and Surfaces*
[Hit] Nigel Hitchens *Lecture notes for B3: Geometry of Surfaces* http://www.maths.ox.ac.uk/system/files/coursematerial/2013/2377/9/notes2013.pdf

Department of Mathematics, Indiana University, Bloomington, IN 47405, USA
E-mail address: macp@indiana.edu