10 Directional derivatives along surfaces

Goals

By the end of this activity, we should be able to do the following.

1. State the coordinate-free definition of the **directional derivative** of a function on a surface.
2. Prove that the directional derivative is **linear** in each of its arguments, and satisfies a **product rule** in its second, using either the definition or the coordinate expression.
3. State the definition of the **covariant derivative** and use this to discuss **parallel transport**.

Today marks the beginning of the third and final portion of our course. Instead of studying the geometry of curves in surfaces, we want to start thinking about the geometry of the surface itself — not necessarily from an "intrinsic" point of view. In spirit, we'll do this in much the same way we studied the geometry of curves in $\mathbb{R}^3$: by taking derivatives. The complication is that surfaces have more than one parameter, so we have to think about *directional* derivatives.

First, let's say what it means for a function to be differentiable at a point of our surface.

Definition. We say that a function $f : \operatorname{im}(\vec{x}) \to \mathbb{R}^n$ from a surface $\operatorname{im}(\vec{x})$ to $\mathbb{R}^n$ is **differentiable** at $p \in \operatorname{im}(\vec{x})$ if, for all possible C^1 surface curves $\vec{\alpha}\colon (-\epsilon,\epsilon) \to \operatorname{im}(\vec{x})$ with $\vec{\alpha}(0) = p$, the function

$$f \circ \vec{\alpha}\colon (-\epsilon,\epsilon) \to \mathbb{R}^n$$

is differentiable at 0.

We still need to think about why this is a good definition, but the next definition should not be too surprising, in light of the one we just wrote down.

Definition. Let $f : \operatorname{im}(\vec{x}) \to \mathbb{R}^n$ be a function on a surface, $p \in \operatorname{im}(\vec{x})$ a point in the surface at which f is differentiable, and $\vec{v} \in T_p\operatorname{im}(\vec{x})$ a vector tangent to $\operatorname{im}(\vec{x})$ at p. Then the **directional derivative** of f at p in the direction $\vec{v}$ is defined to be

$$D_{\vec{v}} f := (f \circ \vec{\alpha})'(0),$$

where $\vec{\alpha}\colon (-\epsilon,\epsilon) \to \operatorname{im}(\vec{x})$ is any C^1 surface curve with $\vec{\alpha}(0) = p$ and $\vec{\alpha}'(0) = \vec{v}$.

Remark. You might notice that the point p doesn't make an appearance in the notation $D_{\vec{v}} f$, which seems a bit strange. But the tangent vector $\vec{v}$ is based at p, so we can recover p from the notation.

Exercise 10.1. Consider the point $p = (1,0,0)$ on the unit sphere $S^2 \subset \mathbb{R}^3$ and the function $f : S^2 \to \mathbb{R}$ defined by $f(x,y,z) := z$. Then

$$D_{\vec{v}_1} f = 1 \quad \text{and} \quad D_{\vec{v}_2} f = 0,$$

where $\vec{v}_1 = (0,0,1)^T \in T_pS^2$ and $\vec{v}_2 = (0,1,0)^T \in T_pS^2$. Draw a picture and write some complete sentences explaining why these equations are true.

An important point to make about our directional derivative is that it doesn't just depend on the direction of $\vec{v}$; it also depends on the magnitude. But this relationship is linear, as you'll prove in the following exercise.

Exercise 10.2. Use the definition of the directional derivative to prove that

$$D_{c\,\vec{v}} f = c\, D_{\vec{v}} f,$$

for any tangent vector $\vec{v} \in T_p\operatorname{im}(\vec{x})$, differentiable function $f : \operatorname{im}(\vec{x}) \to \mathbb{R}^n$, and scalar $c \in \mathbb{R}$.
Hint: Start with a curve $\vec{\alpha}$ satisfying $\vec{\alpha}(0) = p$ and $\vec{\alpha}'(0) = \vec{v}$. Can you construct a curve $\vec{\beta}$ to use for computing $D_{c\,\vec{v}} f$?

While the definition we've given for $D_{\vec{v}} f$ makes good use of the geometry of $\vec{x}$, it's not the most useful thing for proving the properties we expect derivatives to enjoy, such as linearity and the product rule. However, the following lemma tells us that we can compute $D_{\vec{v}} f$ using the downstairs coordinates; the expected properties will be easier to prove in those coordinates.

Lemma 10.1. *Consider a simple surface $\vec{x}: U \to \mathbb{R}^3$, a point $p = \vec{x}(u_0^1, u_0^2)$ on the surface, and a tangent vector $\vec{v} \in T_p \mathrm{im}(\vec{x})$. If $f: \mathrm{im}(\vec{x}) \to \mathbb{R}^n$ is differentiable at p, then we have*

$$D_{\vec{v}} f = \left(V^1 \frac{\partial (f \circ \vec{x})}{\partial u^1} + V^2 \frac{\partial (f \circ \vec{x})}{\partial u^2} \right) \bigg|_{(u_0^1, u_0^2)},$$

where $\vec{v} = V^1 \vec{x}_1 + V^2 \vec{x}_2$.

Exercise 10.3. Prove Lemma 10.1.
Hint: Use the downstairs curve $\vec{\alpha}_U(t) = (u_0^1 + t\,V^1, u_0^2 + t\,V^2)$ when defining the curve $\vec{\alpha}$ you'll use in the definition of $D_{\vec{v}} f$. Then use the chain rule.

With Lemma 10.1 in hand, our directional derivative on surfaces is able to inherit some helpful properties from the directional derivative on $\mathbb{R}^2$.

Exercise 10.4. Use Lemma 10.1 to show that

$$D_{\vec{v}+\vec{w}} f = D_{\vec{v}} f + D_{\vec{w}} f,$$

for any $\vec{v}, \vec{w} \in T_p \mathrm{im}(\vec{x})$ and differentiable function $f: \mathrm{im}(\vec{x}) \to \mathbb{R}^n$. Combined with Exercise 10.2, this tells us that the directional derivative is **linear** in the vector argument. Why is this fact harder than Exercise 10.2 to prove from the definition?

If we fix a point $p \in \mathrm{im}(\vec{x})$ and a vector $\vec{v} \in T_p \mathrm{im}(\vec{x})$, we can think of $D_{\vec{v}}$ as a "derivative operator" which can be applied to any differentiable function $f: \mathrm{im}(\vec{x}) \to \mathbb{R}^n$. We generally expect our derivatives to be linear, and indeed this is the case.

Exercise 10.5. Fix a point $p \in \mathrm{im}(\vec{x})$ and a vector $\vec{v} \in T_p \mathrm{im}(\vec{x})$. Use Lemma 10.1 to show that

$$D_{\vec{v}}(c\,f + d\,g) = c\,D_{\vec{v}} f + d\,D_{\vec{v}} g,$$

for any differentiable functions $f, g: \mathrm{im}(\vec{x}) \to \mathbb{R}^n$ and scalars $c, d \in \mathbb{R}$.

Another property we expect of our derivatives is that they have some version of a **product rule**. Symbolically, we want

$$D_{\vec{v}}(f \cdot g) = (D_{\vec{v}} f) \cdot g + f \cdot (D_{\vec{v}} g), \tag{10.1}$$

where $\cdot$ represents the relevant notion of product. For instance, a vector field $\vec{X}$ on our surface can be thought of as a function $\vec{X}: \mathrm{im}(\vec{x}) \to \mathbb{R}^3$, and thus it makes sense to think about the directional derivative $D_{\vec{v}} \vec{X}$. Here, the relevant notion of product is the inner product.

Exercise 10.6. Fix a point $p \in \mathrm{im}(\vec{x})$ and a vector $\vec{v} \in T_p \mathrm{im}(\vec{x})$. Use Lemma 10.1 to show that

$$D_{\vec{v}} \langle \vec{X}, \vec{Y} \rangle = \left\langle D_{\vec{v}} \vec{X}, \vec{Y}(p) \right\rangle + \left\langle \vec{X}(p), D_{\vec{v}} \vec{Y} \right\rangle, \tag{10.2}$$

for any differentiable vector fields $\vec{X}, \vec{Y}: \mathrm{im}(\vec{x}) \to \mathbb{R}^3$.
Note: This works for any differentiable vector fields $\vec{X}, \vec{Y}: \mathrm{im}(\vec{x}) \to \mathbb{R}^3$ — we are not assuming that these are tangent *vector fields.*
Hint: There are a few ways to do this, and they'll probably all become a bit messy. I suggest using the fact that

$$\langle \vec{X}, \vec{Y} \rangle \circ \vec{x} = \langle \vec{X} \circ \vec{x}, \vec{Y} \circ \vec{x} \rangle,$$

plus the fact that partial differentiation has a product rule for inner products.

Remark. The more general product rule in Equation 10.1 is also true, but we're proving the concrete case that we'll use most often.

Finally, there's a notion of directional derivative that we've hinted at before. This directional derivative only records what an inhabitant of the surface can see, and helps us to make sense of things like parallel transport.

Definition. Let $\vec{X} : \mathrm{im}(\vec{x}) \to \mathbb{R}^3$ be a differentiable vector field on a surface, $p \in \mathrm{im}(\vec{x})$, and $\vec{v} \in T_p\mathrm{im}(\vec{x})$ a vector tangent to $\mathrm{im}(\vec{x})$ at p. Then the **covariant derivative** of $\vec{X}$ at p in the direction $\vec{v}$ is defined to be

$$\nabla_{\vec{v}}\vec{X} := \mathrm{proj}_{T_p\mathrm{im}(\vec{x})} D_{\vec{v}}\vec{X} = D_{\vec{v}}\vec{X} - \langle D_{\vec{v}}\vec{X}, \vec{n}\rangle\, \vec{n}.$$

Because our focus is shifting from the intrinsic geometry of surfaces to the *extrinsic* geometry of surfaces, we will use the covariant derivative only sparingly going forward. But it's useful to notice that this is the correct general way to talk about parallelism. (Remember that our previous discussion of parallel transport only worked for circles of latitude on a sphere.)

Exercise 10.7. Consider a surface curve $\vec{\alpha} = \vec{x} \circ \vec{\alpha}_U$ and a vector field $\vec{V}$ defined along $\vec{\alpha}$. Remember that we call a vector field $\vec{V}$ **parallel** along $\vec{\alpha}$ if it doesn't appear to be changing as we move along $\vec{\alpha}$ — from the perspective of an inhabitant of the surface. Write down an equation to determine whether or not $\vec{V}$ is parallel using the covariant derivative. Write a sentence or two explaining why you wrote the equation you did.
Hint: Today we thought about how to measure changes as we move in a given direction. Is there a direction associated to movement along $\vec{\alpha}$?