

Lecture notes for Week 4

Lecturer: Mark Girard

September 24, 2019

4.1 Parametric surfaces in \mathbb{R}^3

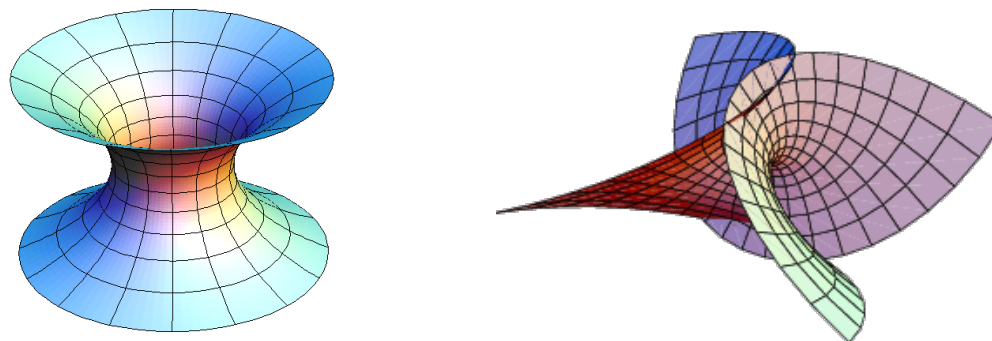
We'll now move on to studying *surfaces*, the two-dimensional analog of the one-dimensional curves that we had been studying up until now. Our consideration of surfaces will be limited to surfaces living in three dimensional space, as our primary objective will be the modeling of physical phenomena related to electromagnetism. Just as we had we had two types of line integrals over curves (integrals of scalar and vector fields), we will build up the machinery of *surface integrals* of scalar and vector fields.

Intuitively, a surface in three dimensional space is a “thin”, essentially “two-dimensional” object, such as a sheet of paper. Our first task is to make this somewhat vague notion mathematically precise in a clear definition.

Definition 4.1 (Surfaces in \mathbb{R}^3). Let $\Phi : D \rightarrow \mathbb{R}^3$ be a continuous function on a region $D \subseteq \mathbb{R}^2$. The (*parametric*) *surface* Σ in \mathbb{R}^3 that is traced out by the function Φ is the set of points

$$\Sigma = \{\Phi(s, t) \mid (s, t) \in D\}.$$

The function Φ is said to *trace out* the surface Σ . If Φ is a C^1 -function (i.e., if each of its component functions is C^1), then the resulting surface Σ is said to be C^1 .



(a) A surface in \mathbb{R}^3 is a “two-dimensional” object living in three dimensional space. (b) A non-simple surface in \mathbb{R}^3 could have self-intersections.

Figure 4.1: Some examples of surfaces in \mathbb{R}^3 .

Similar to curves, a function Φ that traces out the surface is typically defined in terms of its *component functions* $\Phi(s, t) = (f(s, t), g(s, t), h(s, t))$, where f , g , and h are some scalar fields on D . If each of the component functions is C^1 then the mapping Φ is C^1 .

In our study of curves, we were only interested in studying simple curves that can be parameterized by a one-to-one mapping γ . Such curves have no “self-intersections.” Similarly in our study of surfaces, we will ignore any surfaces that intersect themselves (as in Figure 4.1b).

Definition 4.2. A surface $\Sigma \subseteq \mathbb{R}^3$ is *simple* if there is a C^1 -mapping $\Phi : D \rightarrow \mathbb{R}^3$ tracing out the surface Σ such that Φ is one-to-one on D , except possibly on the boundary of D . The mapping Φ is said to *parameterize* the simple surface Σ .

As we've considered before, one type of surface that we can study are graphs. Given a continuous scalar field $f : D \rightarrow \mathbb{R}$ on a region $D \subseteq \mathbb{R}^2$, the *graph* of f is the surface in \mathbb{R}^3 defined by

$$\Sigma = \{(x, y, f(x, y)) \mid (x, y) \in D\},$$

which is parameterized by the function $\Phi(x, y) = (x, y, f(x, y))$. Note that in this case we've used the variable names x and y to parameterize the region D instead of s and t . It does really not matter what names we use for the variables. For example, we will use φ and θ for the variable names when using spherical coordinates.

Example 4.3. The surface of the “top half” of a sphere of radius $r > 0$ in \mathbb{R}^3 with center at the origin is the set of all points $(x, y, z) \in \mathbb{R}^3$ satisfying

$$x^2 + y^2 + z^2 = r^2 \quad \text{and} \quad z \geq 0.$$

This surface can also be described as the graph of a function as follows. Define the disc $D \subset \mathbb{R}^2$ of radius r as

$$D = \{(x, y) \mid x^2 + y^2 \leq r^2\},$$

and define the function $f : D \rightarrow \mathbb{R}$ as $f(x, y) = \sqrt{r^2 - x^2 - y^2}$ for all $(x, y) \in D$. It is clear that the points $(x, y, f(x, y))$ trace out the surface Σ .

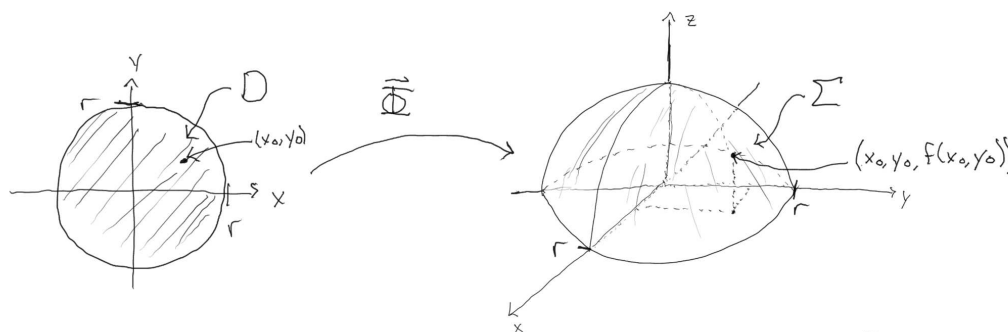


Figure 4.2: The top half of the sphere can be parameterized as the graph of a function.

Remark. Not all surfaces can be represented as graphs of scalar fields. Consider for example the surface depicted in Figure 4.3. Nonetheless, surfaces like this can still be parameterized by a function $\Phi : D \rightarrow \mathbb{R}^3$. We can think of a surface as a portion of a “deformed flat surface.” Just as it takes only two coordinates to specify a single point on the plane, it follows that one should likewise require only two “coordinates” to specify a point on a surface.

Example 4.4. Although the top-half of a sphere can be considered as the graph of a function, the surface of the *entire* sphere cannot be the graph of some function. So we will need some other way to parameterize this surface. One way to specify a point \mathbf{r} on the sphere of radius r is to consider the *angles* that the line connecting the origin to \mathbf{r} makes with the x - and z -axes (see Figure 4.4). First set φ to be the angle that the line connecting the origin to \mathbf{r} makes with the z -axis, where $0 \leq \varphi \leq \pi$. Then drop a line from the point

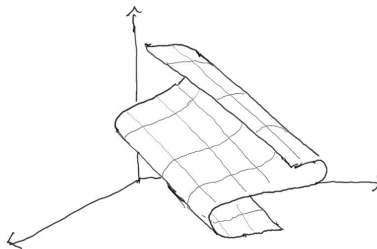


Figure 4.3: A surface in \mathbb{R}^3 with a “fold” cannot be represented as the graph of some function.

r straight down to the xy -plane. The line connecting this point to the origin makes an angle of θ with the x -axis, where θ is taken between 0 and 2π . The coordinates of this point as functions of φ and θ are

$$x(\theta, \varphi) = r \sin \varphi \cos \theta, \quad y(\theta, \varphi) = r \sin \varphi \sin \theta, \quad \text{and} \quad z(\theta, \varphi) = r \cos \varphi.$$

This yields the parameterization $\Phi : D \rightarrow \mathbb{R}^3$ of the sphere of radius r , where we take D to be the region

$$\begin{aligned} D &= \{(\theta, \varphi) \mid 0 \leq \theta \leq 2\pi \text{ and } 0 \leq \varphi \leq \pi\} \\ &= [0, 2\pi] \times [0, \pi], \end{aligned}$$

and define $\Phi(\theta, \varphi) = (r \sin \varphi \cos \theta, r \sin \varphi \sin \theta, r \cos \varphi)$.

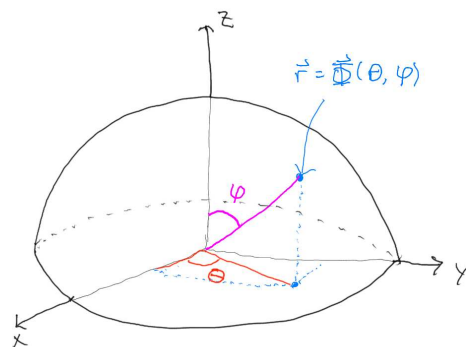


Figure 4.4: Parameterizing the surface that is the sphere of radius r .

Example 4.5. In this example we write (r, θ) for the parametric variables names, since we want to regard the variable r as the “radius” and the variable θ as the “angle”. Define

$$\Phi(r, \theta) = (r \cos \theta, r \sin \theta, \theta)$$

for all $r \in [0, 1]$ and $\theta \in [0, 4\pi]$. We have a parametric mapping $\Phi : D \rightarrow \mathbb{R}^3$ where the region D is the rectangle in the “ $r\theta$ -plane” given by

$$\begin{aligned} D &= \{(r, \theta) \mid 0 \leq r \leq 1 \text{ and } 0 \leq \theta \leq 2\pi\} \\ &= [0, 1] \times [0, 4\pi]. \end{aligned}$$

The surface in \mathbb{R}^3 traced out by $\Phi(r, \theta)$ as (r, θ) traverses the rectangle D is called a *helicoid* and is shown in Figure 4.5.

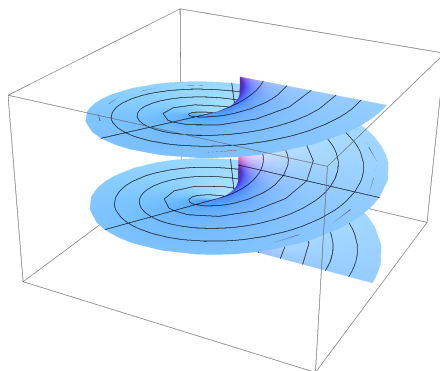


Figure 4.5: The helicoid in Example 4.5.

4.1.1 Grid curves

Suppose we are given a parameterization $\Phi : D \rightarrow \mathbb{R}^3$ of a surface for which the region $D \subset \mathbb{R}^2$ is a rectangle of the form

$$D = [a, b] \times [c, d]$$

for some numbers $a < b$ and $c < d$. For some fixed value $t_0 \in [c, d]$, we can define a path by holding the second parameter constant at $t = t_0$ and vary over the first parameter by defining

$$\gamma_{t=t_0} : [a, b] \rightarrow \mathbb{R}^3 \quad \text{by} \quad \gamma_{t=t_0}(s) = \Phi(s, t_0) \quad \text{for all } s \in [a, b].$$

The resulting curve $\Gamma_{t=t_0} = \{\Phi(s, t_0) \mid s \in [a, b]\}$ is called a *grid curve* of the parameterization for this surface. Analogously, we can define grid curves where we hold the first parameter constant and vary over the second parameter. For a fixed value $s_0 \in [a, b]$, we can define a path

$$\gamma_{s=s_0} : [c, d] \rightarrow \mathbb{R}^3 \quad \text{by} \quad \gamma_{s=s_0}(t) = \Phi(s_0, t) \quad \text{for all } t \in [c, d].$$

By picking a number of values $s_0, s_1, \dots, s_M \in [a, b]$ and $t_0, t_1, \dots, t_N \in [c, d]$, we construct a “grid” on the rectangle D in the st -plane. Sketching the resulting grid curves in \mathbb{R}^3 gives us an idea of what the surface looks like.

Example 4.6. Consider the region $D = [0, 1] \times [0, 2\pi]$ and let $\Phi : D \rightarrow \mathbb{R}^3$ be the parameterization defined by

$$\Phi(s, t) = (s \cos t, s \sin t, s).$$

Sketching out some of the grid curves allows us to get an understanding of what this surface looks like. If we hold t constant at the value $t = 0$, we get the curve traced out by the path

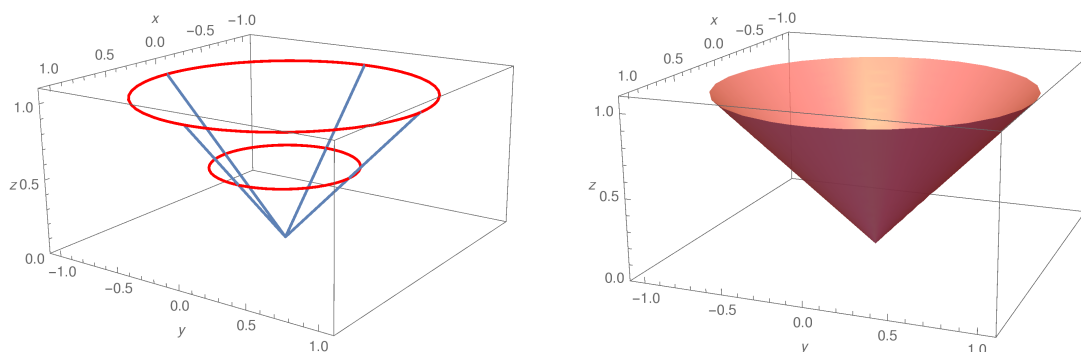
$$\gamma_{t=0}(s) = (s, 0, s)$$

for $0 \leq s \leq 1$. The resulting curve is exactly the straight line segment connecting the origin to the point $(1, 0, 1)$. Similarly, holding t at the constant values $t = \pi/2$, $t = \pi$, and $t = 3\pi/2$ yields the paths

$$\gamma_{t=\frac{\pi}{2}}(s) = (0, s, s), \quad \gamma_{t=\pi}(s) = (-s, 0, s), \quad \text{and} \quad \gamma_{t=\frac{3\pi}{2}}(s) = (0, -s, s),$$

which trace out the lines connecting the origin to $(0, 1, 1)$, $(-1, 0, 1)$, and $(0, -1, 1)$ respectively. To get the grid curves of this surface where we hold the first variable constant, we see that setting $s = 0$ the grid curve is just a point at the origin. Meanwhile setting $s = 1/2$ and $s = 1$ gives us the grid curves

$$\gamma_{s=\frac{1}{2}} = \frac{1}{2}(\cos t, \sin t, 1) \quad \text{and} \quad \gamma_{s=1} = (\cos t, \sin t, 1),$$



(a) Some grid curves of the surface in Example 4.6. (b) The full surface of the cone from Example 4.6.

Figure 4.6: The cone with and some of its grid curves.

which trace out circles around the z -axis of radius $1/2$ and 1 , each having centre at $(0, 0, 1/2)$ and $(0, 0, 1)$ respectively. Sketching these curves out in \mathbb{R}^3 , we see that the resulting surface is a cone along the z -axis whose vertex is at the origin and whose base is the unit circle on the $z = 1$ plane centered at $(0, 0, 1)$.

4.1.2 Review: Equations of lines and planes in \mathbb{R}^3

Here we review the equations that define lines and planes in \mathbb{R}^3 . A *line* L in space is determined when we know a point \mathbf{r}_0 on the line L and the direction of L . If \mathbf{v} is a vector in \mathbb{R}^3 that is *parallel* to the line L , then the line can be described as the set

$$L = \{\mathbf{r}_0 + t\mathbf{v} \mid t \in \mathbb{R}\}.$$

(This is a parametric representation of the line L .)

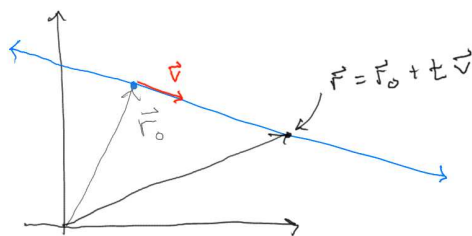


Figure 4.7: A line L in \mathbb{R}^3 specified by a point \mathbf{r}_0 on the line and a direction \mathbf{v} .

A *plane* is a bit more difficult to describe, since a single vector parallel to the plane in \mathbb{R}^3 is not enough to specify the “direction” of the plane. However, a vector *perpendicular* to the plane does completely specify the direction. Thus, a plane E in \mathbb{R}^3 can be completely specified by a single point \mathbf{r}_0 on the plane and a *normal vector* \mathbf{n} that is orthogonal to the plane. If $\mathbf{r} \in \mathbb{R}^3$ is any other point on the plane, the normal vector \mathbf{n} is *orthogonal* to the vector $\mathbf{r} - \mathbf{r}_0$. Mathematically, this is written as

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0 \quad \text{or} \quad \mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0. \quad (4.1)$$

So we may describe the plane as the set of all points \mathbf{r} satisfying the equation in (4.1),

$$E = \{\mathbf{r} \in \mathbb{R}^3 \mid \mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0\}.$$

The equation in (4.1) is the *vector equation* for the plane. Note that this representation is not unique, since we could have chosen *any* point \mathbf{r}_0 on the plane to start from. Moreover, any nonzero scalar multiple of the normal vector (i.e., $a\mathbf{n}$ for any $a \neq 0$) will yield the same plane, since multiplying the equation in (4.1) by a nonzero scalar a doesn't change it. It is perhaps more common to see the equation for a plane written as something like

$$ax + by + cz = d \quad (4.2)$$

for some constants $a, b, c, d \in \mathbb{R}$. The value of these constants can be found from the normal vector. An arbitrary point in \mathbb{R}^3 is described in coordinate form as $\mathbf{r} = (x, y, z)$, the constants in (4.2) correspond to the normal vector $\mathbf{n} = (a, b, c)$, and we may take $d = \mathbf{n} \cdot \mathbf{r}_0$ for some point \mathbf{r}_0 on the plane. Then (4.1) is equivalent to the *equation for the plane* in (4.2), since $\mathbf{n} \cdot \mathbf{r} = (a, b, c) \cdot (x, y, z) = ax + by + cz$.

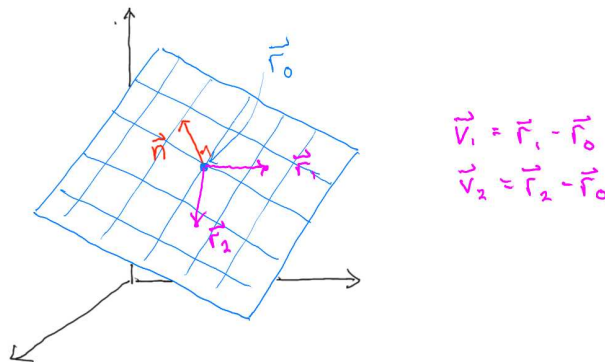


Figure 4.8: A plane E in \mathbb{R}^3 specified by a point \mathbf{r}_0 on the plane and a normal vector \mathbf{n} . A normal vector of a plane can be found from three non-co-linear points \mathbf{r}_0 , \mathbf{r}_1 , and \mathbf{r}_2 on the plane.

A plane in \mathbb{R}^3 is clearly a surface, so it makes sense that we should be able to find a parameterization for a plane as per Definition 4.1. We first show how we can find the normal vector of a plane. Let E be a plane and let \mathbf{r}_0 , \mathbf{r}_1 , and \mathbf{r}_2 be three points on the plane that do not lie on the same *line* (i.e., not *co-linear*), and define the vectors

$$\mathbf{v}_1 = \mathbf{r}_1 - \mathbf{r}_0 \quad \text{and} \quad \mathbf{v}_2 = \mathbf{r}_2 - \mathbf{r}_0.$$

The vectors \mathbf{v}_1 and \mathbf{v}_2 both lie *tangent* to the plane and they do not point along the same line (since the points \mathbf{r}_0 , \mathbf{r}_1 , and \mathbf{r}_2 were taken to be not co-linear). Taking the cross product of \mathbf{v}_1 and \mathbf{v}_2 yields a vector that is orthogonal to both of them. This is the normal vector that we are looking for:

$$\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 = (\mathbf{r}_1 - \mathbf{r}_0) \times (\mathbf{r}_2 - \mathbf{r}_0).$$

Since the equation defining a plane is not changed when multiplying the normal vector \mathbf{n} by a nonzero scalar, it is often useful to pick the normal vector to be a *unit normal vector*. To get a normal vector that has length equal to one, we simply divide by its norm to get

$$\hat{\mathbf{n}} = \frac{\mathbf{n}}{\|\mathbf{n}\|},$$

where the ‘hat’ symbol $\hat{}$ over a vector will always be used to specify that the vector is a unit vector.

To find a parametric representation of the plane, first note that the vectors \mathbf{v}_1 and \mathbf{v}_2 span a *subspace* in \mathbb{R}^3 by taking all of the possible linear combinations:

$$V = \{s\mathbf{v}_1 + t\mathbf{v}_2 \mid s, t \in \mathbb{R}\}.$$

This is a parametric representation for the plane going through the origin that is “parallel” to the tangent plane we are looking for. To find a parametric representation for the desired tangent plane, we only have to shift each point in the subspace V by \mathbf{r}_0 :

$$E = \{\mathbf{r}_0 + s\mathbf{v}_1 + t\mathbf{v}_2 \mid s, t \in \mathbb{R}\}.$$

This is the desired parametric representation, since any other point \mathbf{r} on the plane E can be described as $\mathbf{r} = \mathbf{r}_0 + s\mathbf{v}_1 + t\mathbf{v}_2$ for some values of s and t .

4.1.3 Tangent plane and normal vector

In this section we will introduce the idea of a *tangent plane* of a surface at a point on that surface. We first recall the idea of a *tangent line* on a curve. Consider a simple C^1 -curve $\Gamma \subseteq \mathbb{R}^3$ that is parameterized by a C^1 -path $\gamma : [a, b] \rightarrow \mathbb{R}^3$. At a point $\mathbf{r}_0 = \gamma(t_0)$ on the curve for some $t_0 \in [a, b]$, the vector velocity of the path $\gamma'(t_0)$ is tangent to the curve. The *tangent line* of Γ at \mathbf{r}_0 (which is the set of all points on the line that lies tangent to the curve at the point) is described as

$$L = \{\gamma(t_0) + t\gamma'(t_0) \mid t \in \mathbb{R}\}.$$

See Figure 4.9 for a sketch.

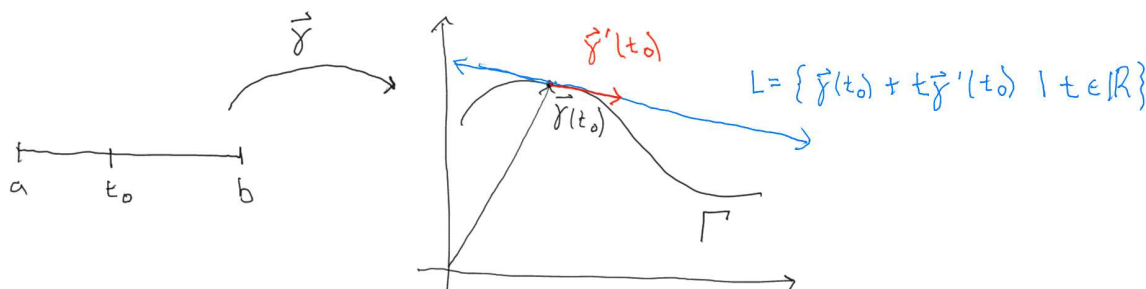


Figure 4.9: The tangent line of a curve at a point $\mathbf{r}_0 = \gamma(t_0)$.

Using similar ideas, we can define the *tangent plane* of a surface. Consider a simple C^1 -surface $\Sigma \subseteq \mathbb{R}^3$ and let $\Phi : D \rightarrow \mathbb{R}^3$ be a C^1 -parametric representation of this surface for some region $D \subseteq \mathbb{R}^2$. Let (s_0, t_0) be a point in the interior (i.e., not on the boundary) of D . Then there must be some small rectangle inside D ,

$$[a, b] \times [c, d] \subset D \quad \text{such that} \quad a < s_0 < b \quad \text{and} \quad c < t_0 < d.$$

As before, we can construct “grid curves” corresponding to this parameterization of the surface. We take the paths

$$\gamma_{t=t_0}(s) = \Phi(s, t_0) \quad \text{for } s \in [a, b] \quad \text{and} \quad \gamma_{s=s_0}(t) = \Phi(s_0, t) \quad \text{for } t \in [c, d], \quad (4.3)$$

and consider the resulting grid curves

$$\Gamma_{t=t_0} = \{\Phi(s, t_0) \mid a \leq s \leq b\} \quad \text{and} \quad \Gamma_{s=s_0} = \{\Phi(s_0, t) \mid c \leq t \leq d\} \quad (4.4)$$

which intersect at the point on the surface $\mathbf{r}_0 = \Phi(s_0, t_0)$. Since the curves resulting from these paths lie *on* the surface, the tangent lines to these curves will *also* be tangent to the surface at the point \mathbf{r}_0 . (See Figure 4.11.)

The velocities of the paths in (4.3) defining the grid curves in (4.4) at the values of the parameters $s = s_0$ and $t = t_0$ are

$$\gamma'_{t=t_0}(s_0) = \left. \frac{\partial \Phi}{\partial s} \right|_{(s_0, t_0)} \quad \text{and} \quad \gamma'_{s=s_0}(t_0) = \left. \frac{\partial \Phi}{\partial t} \right|_{(s_0, t_0)}$$

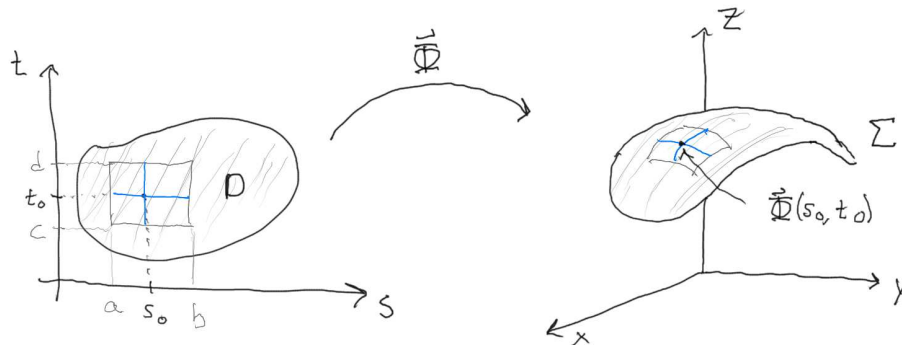


Figure 4.10: Even if D is not itself a rectangle, we can always consider a smaller region inside D that is a rectangle around the point (s_0, t_0) inside.

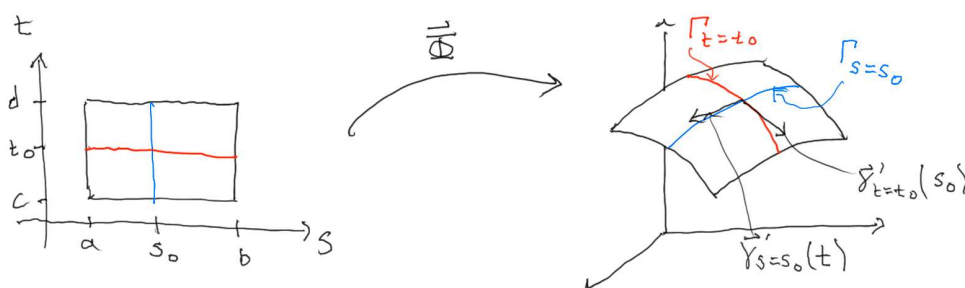


Figure 4.11: The grid curves $\Gamma_{t=t_0}$ and $\Gamma_{s=s_0}$ intersect at the point $\mathbf{r}_0 = \Phi(s_0, t_0)$. The velocity vectors of the paths are tangent to the surface at this point.

respectively, where, to take the partial derivatives of the parameterization Φ with respect to each of its parametric variables, we take the partial derivatives of its components $\Phi(s, t) = (x(s, t), y(s, t), z(s, t))$,

$$\Phi_s(s_0, t_0) = \frac{\partial \Phi}{\partial s} \Big|_{(s_0, t_0)} = \left(\frac{\partial x}{\partial s} \Big|_{(s_0, t_0)}, \frac{\partial y}{\partial s} \Big|_{(s_0, t_0)}, \frac{\partial z}{\partial s} \Big|_{(s_0, t_0)} \right)$$

and

$$\Phi_t(s_0, t_0) = \frac{\partial \Phi}{\partial t} \Big|_{(s_0, t_0)} = \left(\frac{\partial x}{\partial t} \Big|_{(s_0, t_0)}, \frac{\partial y}{\partial t} \Big|_{(s_0, t_0)}, \frac{\partial z}{\partial t} \Big|_{(s_0, t_0)} \right),$$

where, for simplicity, we often instead use the notation $\Phi_s(s_0, t_0)$ and $\Phi_t(s_0, t_0)$ to denote the partial derivatives of Φ evaluated at (s_0, t_0) .

The vectors $\Phi_s(s_0, t_0)$ and $\Phi_t(s_0, t_0)$ lie tangent to the surface at the point $\mathbf{r}_0 = \Phi(s_0, t_0)$. A representation for the *tangent plane* to the surface Σ at this point can therefore be given by

$$\text{tangent plane at } \mathbf{r}_0 = \left\{ \mathbf{r}_0 + s \frac{\partial \Phi}{\partial s} \Big|_{(s_0, t_0)} + t \frac{\partial \Phi}{\partial t} \Big|_{(s_0, t_0)} \mid s, t \in \mathbb{R} \right\}.$$

We can find a *normal vector* to the tangent plane at the point $\mathbf{r}_0 = \Phi(s_0, t_0)$ corresponding to the parameterization Φ as

$$\mathbf{n}_\Phi(s_0, t_0) = \frac{\partial \Phi}{\partial s} \Big|_{(s_0, t_0)} \times \frac{\partial \Phi}{\partial t} \Big|_{(s_0, t_0)}. \quad (4.5)$$

Since this vector is normal to the tangent plane of Σ at \mathbf{r}_0 , it is normal to the surface at that point as well! Given a parameterization $\Phi : D \rightarrow \mathbb{R}^3$ of a surface Σ , we can define another vector-valued function

$\mathbf{n}_{\Phi} : D \rightarrow \mathbb{R}^3$ defined for all $(s, t) \in D$ by

$$\begin{aligned}\mathbf{n}_{\Phi}(s, t) &= \left. \frac{\partial \Phi}{\partial s} \right|_{(s,t)} \times \left. \frac{\partial \Phi}{\partial t} \right|_{(s,t)} \\ &= \Phi_s(s, t) \times \Phi_t(s, t)\end{aligned}\tag{4.6}$$

which gives us a normal vector at $\Phi(s, t)$ corresponding to this parameterization. Note that this normal vector depends on which parameterization we choose, since choosing a different parameterization might result in different partial derivatives. The resulting normal vector from a different parameterization will necessarily point along the same direction, but might have different length.

Example 4.7. Consider a cylinder of radius r and height h in \mathbb{R}^3 that is parameterized by $\Phi : D \rightarrow \mathbb{R}^3$ where

$$\Phi(\theta, t) = (r \cos \theta, r \sin \theta, t) \quad \text{and} \quad D = [0, 2\pi] \times [0, h].$$

We can find an expression for a normal vector at each point on the cylinder by taking the partial derivatives

$$\left. \frac{\partial \Phi}{\partial \theta} \right|_{(\theta,t)} = (-r \sin \theta, r \cos \theta, 0) \quad \text{and} \quad \left. \frac{\partial \Phi}{\partial t} \right|_{(\theta,t)} = (0, 0, 1).$$

The normal vector at each point on the surface corresponding to this parameterization is therefore

$$\begin{aligned}\mathbf{n}_{\Phi}(\theta, t) &= \left(\frac{\partial \Phi}{\partial \theta} \times \frac{\partial \Phi}{\partial t} \right) \Big|_{(\theta,t)} \\ &= (-r \sin \theta, r \cos \theta, 0) \times (0, 0, 1) \\ &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -r \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= r \cos \theta \hat{\mathbf{i}} + r \sin \theta \hat{\mathbf{j}}.\end{aligned}$$

This is the *outward facing* normal vector, since we can see in Figure 4.12 that this normal vector points outward (away from the z -axis) at any point. Consider for example the point $\Phi(0, h/2)$ on the cylinder that is half way up. The normal vector (determined by this parameterization) at this point is

$$\mathbf{n}_{\Phi}\left(0, \frac{h}{2}\right) = (r, 0, 0),$$

which points parallel to the x -axis.

If we want to find the *unit* normal vector corresponding to a parameterization of a surface, we divide by the norm:

$$\hat{\mathbf{n}}_{\Phi}(s, t) = \frac{\mathbf{n}_{\Phi}(s, t)}{\|\mathbf{n}_{\Phi}(s, t)\|} = \frac{\left. \frac{\partial \Phi}{\partial s} \right|_{(s,t)} \times \left. \frac{\partial \Phi}{\partial t} \right|_{(s,t)}}{\left\| \left. \frac{\partial \Phi}{\partial s} \right|_{(s,t)} \times \left. \frac{\partial \Phi}{\partial t} \right|_{(s,t)} \right\|}.$$

Notice, however, that we cannot divide by zero, so the unit normal vector can only be found if the normal vector is nonzero ($\mathbf{n}_{\Phi}(s, t) \neq \mathbf{0}$). This can occur if either

$$\frac{\partial \Phi}{\partial s} = \mathbf{0} \quad \text{or} \quad \frac{\partial \Phi}{\partial t} = \mathbf{0} \quad \text{or} \quad \text{if the partial derivatives } \frac{\partial \Phi}{\partial s} \text{ and } \frac{\partial \Phi}{\partial t} \text{ are parallel.}$$

If either of these things happen, the resulting parametric surface might have a “sharp edge” at the point where $\mathbf{n}_{\Phi}(s, t) = \mathbf{0}$. This will cause troubles when we get to integrating over surfaces. In order to eliminate this possibility, we restrict our attention only to simple surfaces that are *smooth*. In this course we will only be interested in smooth surfaces.

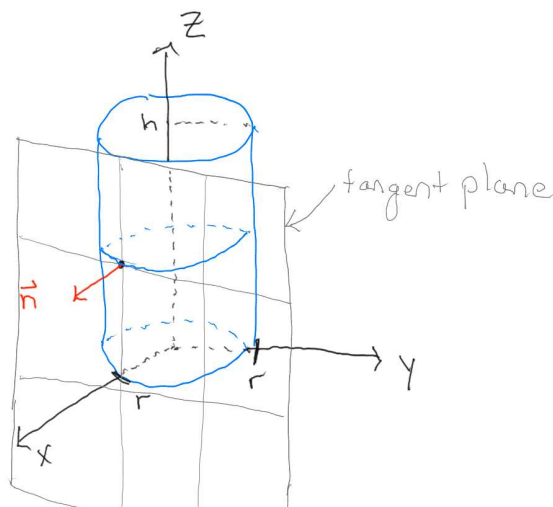


Figure 4.12: The tangent plane of the cylinder from Example 4.7.

Definition 4.8 (Smooth surface). Let $\Sigma \subseteq \mathbb{R}^3$ be a simple surface and let $\mathbf{r}_0 \in \Sigma$ be a point on the surface. The surface Σ is said to be *smooth at the point* \mathbf{r}_0 if there is a C^1 -parameterization $\Phi : D \rightarrow \mathbb{R}^3$ of Σ such that $\mathbf{r}_0 = \Phi(s_0, t_0)$ for some $(s_0, t_0) \in D$ such that

$$\frac{\partial \Phi}{\partial s} \Big|_{(s_0, t_0)} \times \frac{\partial \Phi}{\partial t} \Big|_{(s_0, t_0)} \neq \mathbf{0}.$$

A simple surface Σ is called *smooth* if it has a parameterization $\Phi : D \rightarrow \mathbb{R}^3$ such that Σ is smooth at $\Phi(s_0, t_0)$ for each point $(s_0, t_0) \in D$.

4.2 Surface integration

Previously, we studied how to extend the notion of integration of real-valued functions on the real line to line integrals over curves in space. We will now study how to take integrals over two-dimensional surfaces in three-dimensional space. We'll introduce this idea by first using parameterizations to determine the area of a surface by integration.

4.2.1 Area of surfaces

Consider a smooth surface Σ that is parameterized by some function $\Phi : D \rightarrow \mathbb{R}^3$. For now, suppose that the region $D \subseteq \mathbb{R}^2$ is a rectangle $D = [a, b] \times [c, d]$ for some values $a < b$ and $c < d$. We can obtain an approximation for the area of Σ by segmenting D into many smaller rectangles D_1, D_2, \dots, D_N , where each of the smaller rectangles has sides of length Δs and Δt for some small values of Δs and Δt , and the area of each of these smaller rectangles is $\text{area}(D_i) = \Delta s \Delta t$. (See Figure 4.13.)

This segments the surface Σ into N smaller surfaces $\Sigma_1, \Sigma_2, \dots, \Sigma_N$, each of which corresponds to one of the pieces D_1, \dots, D_N of the region D . We can obtain an approximation for the total area of the surface Σ by approximating the areas of each of the pieces $\Sigma_1, \dots, \Sigma_N$ and adding all of the approximate areas together. To obtain an approximation for the area of each of the pieces $\Sigma_1, \dots, \Sigma_N$, for each $i \in \{1, 2, \dots, N\}$ we may

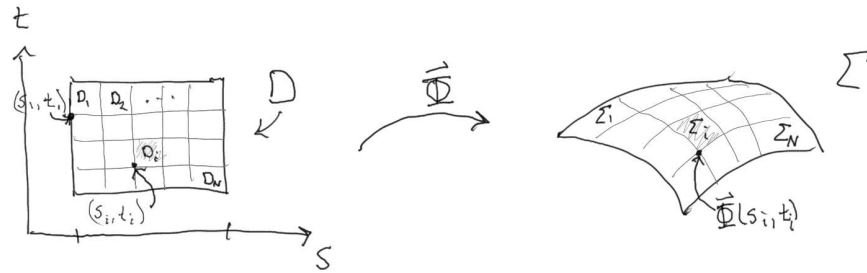


Figure 4.13

think of the region D_i as being a small rectangle

$$\begin{aligned} D_i &= [s_i, s_i + \Delta s] \times [t_i, t_i + \Delta t] \\ &= \{(s, t) \in \mathbb{R}^2 \mid s_i \leq s \leq s_i + \Delta s \text{ and } t_i \leq t \leq t_i + \Delta t\} \end{aligned}$$

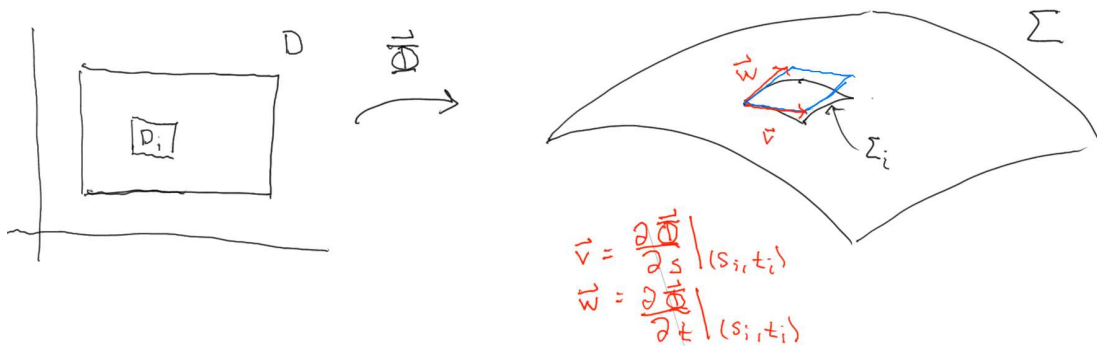
and the piece Σ_i as the small piece of surface mapped to from the piece D_i ,

$$\Sigma_i = \{\Phi(s, t) \mid (s, t) \in D_i\}.$$

For small enough values of Δs and Δt , each piece Σ_i is “approximately” a flat parallelogram, as seen in Figure 4.14, and the edges of Σ_i can approximately be thought of as straight edges given by the vectors

$$\mathbf{v} = \Phi(s_i + \Delta s, t_i) - \Phi(s_i) \quad \text{and} \quad \mathbf{w} = \Phi(s_i, t_i + \Delta t) - \Phi(s_i)$$

that connect the corners of the small piece of the surface Σ_i .

Figure 4.14: A small piece of surface Σ_i can be approximated by a small parallelogram.

Since the values Δs and Δt are small, these vectors are approximately

$$\mathbf{v} \approx \frac{\partial \Phi}{\partial s} \Big|_{(s_i, t_i)} \Delta s \quad \text{and} \quad \mathbf{w} \approx \frac{\partial \Phi}{\partial t} \Big|_{(s_i, t_i)} \Delta t.$$

Recall that the area of the parallelogram that approximates Σ_i with edges given by the vectors \mathbf{v} and \mathbf{w} is

given by the norm of the cross-product $\mathbf{v} \times \mathbf{w}$, namely

$$\begin{aligned} \text{area}(\Sigma_i) &\approx \|\mathbf{v} \times \mathbf{w}\| \\ &\approx \left\| \left(\frac{\partial \Phi}{\partial s} \Big|_{(s_i, t_i)} \Delta s \right) \times \left(\frac{\partial \Phi}{\partial t} \Big|_{(s_i, t_i)} \Delta t \right) \right\| \\ &= \left\| \left(\frac{\partial \Phi}{\partial s} \times \frac{\partial \Phi}{\partial t} \right) \Big|_{(s_i, t_i)} \right\| \Delta s \Delta t \\ &= \|\mathbf{n}_\Phi(s_i, t_i)\| \Delta s \Delta t \end{aligned}$$

where in the last line we recall the definition in (4.5) of the normal vector determined by the parameterization Φ . The area of the surface can therefore be approximated as

$$\begin{aligned} \text{area}(\Sigma) &= \sum_{i=1}^N \text{area}(\Sigma_i) \\ &\approx \sum_{i=1}^N \|\mathbf{n}_\Phi(s_i, t_i)\| \Delta s \Delta t \end{aligned}$$

where we recall that $\Delta s \Delta t$ is the area of each piece D_i of the region D . In the limit as $N \rightarrow \infty$ and $\Delta s, \Delta t \rightarrow 0$, where we segment D into infinitely many smaller segments, this sum approaches a limiting value, which we denote by the integral

$$\begin{aligned} \text{area}(\Sigma) &= \iint_D \|\mathbf{n}_\Phi(s, t)\| \, ds \, dt. \\ &= \iint_D \left\| \left(\frac{\partial \Phi}{\partial s} \times \frac{\partial \Phi}{\partial t} \right) \Big|_{(s, t)} \right\| \, ds \, dt. \end{aligned} \tag{4.7}$$

That is, we can compute the surface area of the surface in \mathbb{R}^3 as a two-dimensional integral over the parameterizing region D in \mathbb{R}^2 !

Remark. The following important question arises in connection with the area formula given by (4.7), namely the right side appears to depend on the particular parametric representation Φ that we have chosen for the surface Σ . However, the area of a surface is intrinsic, and should not depend on the particular parametric representation we have chosen for the surface. This means that if we choose to represent the same surface Σ by two different parametric representations

$$\Phi_1 : D_1 \rightarrow \mathbb{R}^3 \quad \text{and} \quad \Phi_2 : D_2 \rightarrow \mathbb{R}^3,$$

for regions $D_1 \subseteq \mathbb{R}^2$ and $D_2 \subseteq \mathbb{R}^2$, each of which parameterize the same surface Σ ,

$$\{\Phi_1(s, t) \mid (s, t) \in D_1\} = \Sigma = \{\Phi_2(u, v) \mid (u, v) \in D_2\}$$

for $(s, t) \in D_1$ and $(u, v) \in D_2$, then it must be the case that

$$\iint_{D_1} \left\| \left(\frac{\partial \Phi_1}{\partial s} \times \frac{\partial \Phi_1}{\partial t} \right) \Big|_{(s, t)} \right\| \, ds \, dt = \iint_{D_2} \left\| \left(\frac{\partial \Phi_2}{\partial u} \times \frac{\partial \Phi_2}{\partial v} \right) \Big|_{(u, v)} \right\| \, du \, dv.$$

This is analogous to our analysis for lengths of curves; the value we get for computing the length of a curve is independent of which parameterization we choose. This means, in particular, that if we have several parametric representations of a surface Σ then we should use that particular representation which involves the least amount of work in the integration for calculating the area. This will become clear in later examples.

Example 4.9. In this example we will consider the two different parametric representations for the top half of the sphere from Examples 4.3 and 4.3. Let Σ be the surface

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = r^2 \text{ and } z \geq 0\}$$

that is the top half of the sphere which we can parameterize in two different ways.

- For the first parameterization, we will consider Σ as the graph of the function $f : D_1 \rightarrow \mathbb{R}^3$ defined as

$$f(x, y) = \sqrt{r^2 - x^2 - y^2} \quad (4.8)$$

over the region that is the disc of radius r in the plane $D_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}$. This gives us the parameterization $\Phi_1 : D_1 \rightarrow \mathbb{R}^3$ defined by

$$\Phi_1(x, y) = (x, y, f(x, y)) = (x, y, \sqrt{r^2 - x^2 - y^2}). \quad (4.9)$$

The partial derivatives of this parameterization are

$$\frac{\partial \Phi_1}{\partial x} \Big|_{(x,y)} = \left(1, 0, \frac{x}{\sqrt{r^2 - x^2 - y^2}} \right) \quad \text{and} \quad \frac{\partial \Phi_1}{\partial y} \Big|_{(x,y)} = \left(0, 1, \frac{y}{\sqrt{r^2 - x^2 - y^2}} \right),$$

such that the normal vector on the top half of the sphere at the point $\Phi_1(x, y)$ that is obtained from this parameterization is

$$\begin{aligned} \mathbf{n}_{\Phi_1}(x, y) &= \frac{\partial \Phi_1}{\partial y} \Big|_{(x,y)} \times \frac{\partial \Phi_1}{\partial x} \Big|_{(x,y)} \\ &= \left(0, 1, \frac{y}{\sqrt{r^2 - x^2 - y^2}} \right) \times \left(1, 0, \frac{x}{\sqrt{r^2 - x^2 - y^2}} \right) \\ &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 0 & 1 & \frac{y}{\sqrt{r^2 - x^2 - y^2}} \\ 1 & 0 & \frac{x}{\sqrt{r^2 - x^2 - y^2}} \end{vmatrix} = \frac{x}{\sqrt{r^2 - x^2 - y^2}} \hat{\mathbf{i}} + \frac{y}{\sqrt{r^2 - x^2 - y^2}} \hat{\mathbf{j}} + \hat{\mathbf{k}}, \end{aligned}$$

which has norm equal to

$$\|\mathbf{n}_{\Phi_1}(x, y)\| = \sqrt{\frac{x^2 + y^2}{r^2 - x^2 - y^2} + 1} = \sqrt{\frac{r^2}{r^2 - x^2 - y^2}} = \frac{r}{\sqrt{r^2 - x^2 - y^2}}.$$

The area of the top half of the sphere can therefore be computed as

$$\text{area}(\Sigma) = \iint_{D_1} \frac{r}{\sqrt{r^2 - x^2 - y^2}} dx dy. \quad (4.10)$$

Evaluation of this integral is quite laborious and complicated (although not impossible) because the integrand involves the reciprocal of a square root (which is usually quite awkward to deal with) and because the integration is over the disc $D_1 = \{(x, y) \mid x^2 + y^2 \leq r^2\}$ rather than over a nice simple region such as a rectangle.

- As an alternative to the parameterization above, we can use the *spherical coordinates* parameterization from Example 4.4. In that example we have parameterized the entire surface of the sphere, but we can restrict the region to

$$D_2 = [0, 2\pi] \times [0, \pi/2] = \left\{ (\theta, \varphi) \mid 0 \leq \theta \leq 2\pi \text{ and } 0 \leq \varphi \leq \frac{\pi}{2} \right\}$$

(where we only consider points that make an angle of less than $90^\circ = \pi/2$ with the z -axis), and take the parameterization $\Phi_2 : D_2 \rightarrow \mathbb{R}^3$ defined as

$$\Phi_2(\theta, \varphi) = (r \sin \varphi \cos \theta, r \sin \varphi \sin \theta, r \cos \varphi). \quad (4.11)$$

This also parameterizes the top half of the sphere. The relevant partial derivatives are

$$\frac{\partial \Phi_2}{\partial \theta} = (-r \sin \varphi \sin \theta, r \sin \varphi \cos \theta, 0) \quad \text{and} \quad \frac{\partial \Phi_2}{\partial \varphi} = (r \cos \varphi \cos \theta, r \cos \varphi \sin \theta, -r \sin \varphi)$$

such that the normal vector on the top half of the sphere at the point $\Phi_2(\theta, \varphi)$ that is obtained from this parameterization is

$$\begin{aligned} \mathbf{n}_{\Phi_2}(\theta, \varphi) &= \left. \frac{\partial \Phi_2}{\partial \theta} \right|_{(\theta, \varphi)} \times \left. \frac{\partial \Phi_2}{\partial \varphi} \right|_{(\theta, \varphi)} \\ &= (-r \sin \varphi \sin \theta, r \sin \varphi \cos \theta, 0) \times (r \cos \varphi \cos \theta, r \cos \varphi \sin \theta, -r \sin \varphi) \\ &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -r \sin \varphi \sin \theta & r \sin \varphi \cos \theta & 0 \\ r \cos \varphi \cos \theta & r \cos \varphi \sin \theta & -r \sin \varphi \end{vmatrix} \\ &= -r^2 (\sin^2 \varphi \cos \theta \hat{\mathbf{i}} + \sin^2 \varphi \sin \theta \hat{\mathbf{j}} + \sin \varphi \cos \varphi \hat{\mathbf{k}}), \end{aligned}$$

which has norm

$$\begin{aligned} \|\mathbf{n}_{\Phi_2}(\theta, \varphi)\| &= r^2 \sqrt{\sin^4 \varphi \cos^2 \theta + \sin^4 \varphi \sin^2 \theta + \sin^2 \varphi \cos^2 \varphi} \\ &= r^2 \sqrt{\sin^2 \theta (\sin^2(\varphi \cos^2 \theta + \sin^2 \theta) + \cos^2 \varphi)} \\ &= r^2 |\sin \varphi|. \end{aligned}$$

We may now compute the area of the top half of the sphere as

$$\begin{aligned} \text{area}(\Sigma) &= \iint_{D_2} \|\mathbf{n}_{\Phi_2}(\theta, \varphi)\| \, d\theta \, d\varphi \\ &= r^2 \int_0^{2\pi} \left(\int_0^{\pi/2} |\sin \varphi| \, d\varphi \right) d\theta \\ &= r^2 \int_0^{2\pi} 1 \, d\theta \\ &= 2\pi r^2. \end{aligned}$$

Note that here we used the fact that D_2 is a rectangle, and thus the integral can easily be computed as a double integral.

This example illustrates a very important aspect of the area formula (4.7), namely the choice of parametric representation of the surface Σ can substantially influence the amount of work involved in using this formula. Indeed, we saw that the parametric representation of Σ as the graph of the function in (4.8) over the region D_1 given by (4.9) leads to the rather complicated integral in (4.10), whereas the area formula is quite easy to use for the parametric representation of Σ given by (4.11).

4.2.2 Surface integral of a scalar field

In Section 4.2.1 we obtained a formula for the area of a surface Σ with a parametric representation $\Phi : D \rightarrow \mathbb{R}^3$. In this section our goal is to generalize this idea to construct the integral of a given scalar field over

the surface Σ . For a concrete instance of how this type of integral could be useful, suppose that the surface Σ describes an infinitesimally thin sheet of plastic, and for each point $(x, y, z) \in \Sigma$ on the surface function value $\sigma(x, y, z)$ gives the charge density (in units Coulombs/m²) concentrated on the surface Σ at the point (x, y, z) . The integral that we are going to define will enable us to determine the *total charge* on the surface.

Consider a simple C^1 surface Σ and choose a parameterization $\Phi : D \rightarrow \mathbb{R}^3$, for which we will suppose for simplicity that $D \subseteq \mathbb{R}^2$ is a rectangle

$$D = [a, b] \times [c, d],$$

and let $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ be a continuous scalar field. As in 4.2.1, we will divide the region D into tiny rectangles D_1, \dots, D_N , each with area $\Delta s \Delta t$ for some small values Δs and Δt . Then Φ maps D_i onto a small piece of the surface Σ_i which is approximately a parallelogram with area

$$\Delta A_i = \text{area}(\Sigma_i) \approx \|\mathbf{n}_\Phi(s_i, t_i)\| \Delta s \Delta t,$$

where we recall that the normal vector from this parameterization is

$$\mathbf{n}_\Phi(s, t) = \left. \frac{\partial \Phi}{\partial s} \right|_{(s,t)} \times \left. \frac{\partial \Phi}{\partial t} \right|_{(s,t)}.$$

We now multiply the area ΔA_i of Σ_i by the value of the scalar field at the point $\Phi(s_i, t_i)$ corresponding to the corner of Σ_i to get

$$f(\Phi(s_i, t_i)) \Delta A_i \approx f(\Phi(s_i, t_i)) \|\mathbf{n}_\Phi(s_i, t_i)\| \Delta s \Delta t. \quad (4.12)$$

To get some idea of what this quantity means, suppose that at each point (x, y, z) on the surface Σ_i the value $f(x, y, z)$ gives the density of charge per unit area concentrated on the surface around the point (x, y, z) . Then the *total amount of charge* contained on the *small piece* of surface Σ_i is given by the expression in (4.12). Summing over all of the pieces gives us the total amount of charge on the entire surface. If we take the limit as $N \rightarrow \infty$ and $\Delta s, \Delta t \rightarrow 0$, the value of this limit is equal to

$$\lim_{\substack{N \rightarrow \infty \\ \Delta s, \Delta t \rightarrow 0}} \sum_{i=1}^N f(\Phi(s_i, t_i)) \Delta s \Delta t = \iint_D f(\Phi(s, t)) ds dt.$$

As before, it turns out that the value of this limit does not depend on which parameterization we choose for the surface. We can therefore define the value of the surface integral of a scalar field f over a surface Σ as

$$\boxed{\iint_\Sigma f dA = \iint_D f(\Phi(s, t)) ds dt.} \quad (4.13)$$

As before, the value of the integral on the left-hand side is *independent of parameterization*, so we may choose whichever parameterization we choose that allows us to compute the right-hand side of (4.13) most efficiently.

Example 4.10. Consider the conic surface determined by the equations $x^2 + y^2 = z^2$ for $0 \leq z \leq 1$. Suppose that the surface charge density at a point (x, y, z) on the surface is equal to $\sigma(x, y, z) = 1 + xyz$. To find the total charge contained on the surface of this cone, we must set up a surface integral. To do so we must first find a parameterization. We may choose $\Phi : D \rightarrow \mathbb{R}^3$ defined by

$$\Phi(r, \theta) = (r \cos \theta, r \sin \theta, r)$$

over the rectangular region $D = [0, 1] \times [0, 2\pi]$, which has partial derivatives given by

$$\frac{\partial \Phi}{\partial r} = (\cos \theta, \sin \theta, 1) \quad \text{and} \quad \frac{\partial \Phi}{\partial \theta} = (-r \sin \theta, r \cos \theta, 0).$$

The normal vector to the surface that is determined by this parameterization is

$$\begin{aligned}\mathbf{n}_{\Phi}(r, \theta) &= \left(\frac{\partial \Phi}{\partial \theta} \times \frac{\partial \Phi}{\partial r} \right) \Big|_{(r, \theta)} \\ &= (-r \sin \theta, r \cos \theta, 0) \times (\cos \theta, \sin \theta, 1) \\ &= (r \cos \theta, r \sin \theta, r),\end{aligned}$$

which has norm equal to

$$\|\mathbf{n}_{\Phi}(r, \theta)\| = \sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta + r^2} = \sqrt{2}r.$$

The value of the function f at the points on the surface determined by the parameterization is

$$f(\Phi(r, \theta)) = 1 + r^3 \sin \theta \cos \theta.$$

The total charge can now be computed by setting up a surface integral

$$\begin{aligned}\iint_{\Sigma} f \, dA &= \iint_D f(\Phi(r, \theta)) \|\mathbf{n}_{\Phi}(r, \theta)\| \, dr \, d\theta \\ &= \iint_D (1 + r^3 \sin \theta \cos \theta)(\sqrt{2}r) \, dr \, d\theta \\ &= \sqrt{2} \int_0^1 \left(\int_0^{2\pi} (r + r^4 \sin \theta \cos \theta) \, d\theta \right) \, dr \\ &= \sqrt{2} \int_0^1 \left(r \underbrace{\int_0^{2\pi} d\theta}_{2\pi} + r^4 \underbrace{\int_0^{2\pi} \cos \theta \sin \theta \, d\theta}_0 \right) \, dr \\ &= 2\sqrt{2}\pi \int_0^1 r \, dr = \sqrt{2}\pi.\end{aligned}$$

Remark. We note that the area of a smooth surface can be expressed using this notation as an integral over a scalar field,

$$\text{area}(\Sigma) = \iint_{\Sigma} dA,$$

where we take the constant scalar field $f = 1$.