2 Surfaces (and Volumes)

The main purpose of this chapter is to understand how to generalise the idea of an integral. Rather than integrating over a line, we will instead look at how to integrate over a 2d surface. We'll then see how to generalise to the integration over a 3d volume or, more generally, an n-dimensional space.

2.1 Multiple Integrals

We'll start by explaining what it means to integrate over a region in $\mathbb{R}^2$ or over a region in $\mathbb{R}^3$. The former are called area, or surface, integrals; the latter volume integrals. By the time we've understood volume integrals, the extension to $\mathbb{R}^n$ will be obvious.

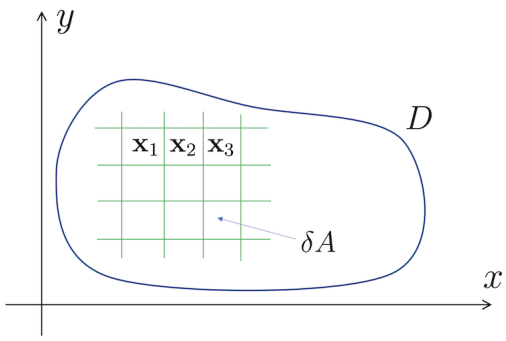
2.1.1 Area Integrals

Consider a region $D \subset \mathbb{R}^2$. Given a scalar function $\phi : \mathbb{R}^2 \to \mathbb{R}$, we want to find a way to integrate ϕ over D. We write this as

$$\int_D \phi(\mathbf{x})\, dA \tag{2.1}$$

You should think of the area element dA as representing an infinitesimally small area, with the $\int$ sign telling us that we're summing over many such small areas, in much the same was as $\int dx$ should be thought of as summing over infinitesimally small line elements dx. The area element is also written as $dA = dx\, dy$.

The rough idea underlying the integral is straightforward. First, we find a way to tesselate D with some simple shape, say a rectangle or other polygon. Each shape has common area δA. Admittedly, there might be some difficulty in making this work around the edge, but we'll ignore this for now. Then we might approximate the integral as

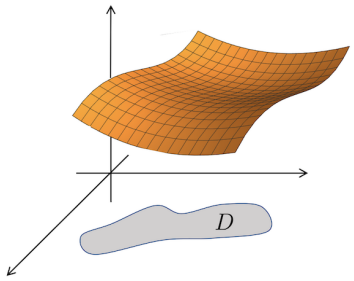


$$\int_D \phi(\mathbf{x})\, dA \;\approx\; \sum_n \phi(\mathbf{x}_n)\, \delta A$$

where $\mathbf{x}_n$ is a point in the middle of each shape. We can then consider making δA smaller and smaller, so that we tesselate the region D with finer and finer shapes. Intuitively, we might expect that as $\delta A \to 0$, the sum converges on an answer and, moreover, this answer is independent on any choices that we made along the way, such

as what shape we use and how we deal with the edges. When the limit exists – as it will for any sensible choice of function ϕ and region D – then it converges to the integral. If the function in the integrand is simply $\phi = 1$ then the integral (2.1) calculates the area of the region D.

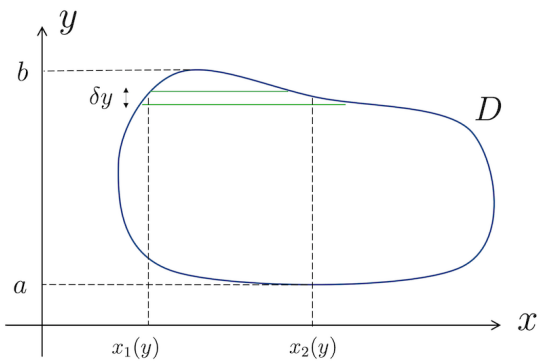
Just as an ordinary, one-dimensional integral can be viewed as the area under a curve, so too can an area integral be viewed as the volume under a function. This interpretation follows simply by plotting $z = \phi(x, y)$ in 3d as shown to the right.



Evaluating Area Integrals

In practice, we evaluate area integrals (or, indeed, higher dimensional integrals) by reducing them to multiple ordinary integrals.

There are a number of different ways to do this, and some may be more convenient than others, although all will give the same answer. For example, we could parcel our region D into narrow horizontal strips of width δy like so:

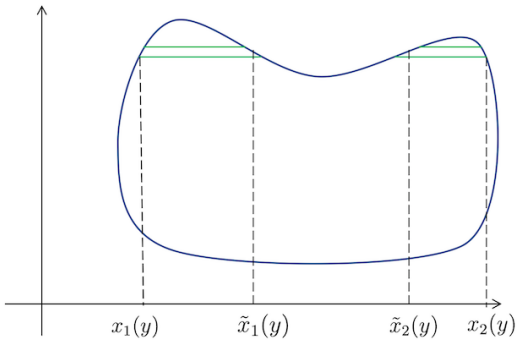


For each value of y, we then do the x integral between the two limits $x_1(y)$ and $x_2(y)$. We then subsequently sum over all such strips by doing the y integral between the two outer limits of the shape which we call a and b. The net result is

$$\int_D \phi(x, y)\, dA = \int_a^b dy \int_{x_1(y)}^{x_2(y)} dx\ \phi(x, y) \tag{2.2}$$

In this approach, the information about the shape D appears in the limits of the integral $x_1(y)$ and $x_2(y)$ which trace the outline of D as y changes.

If your shape is suitably annoying, then one or more of the integrals may have to be decomposed into disjoint sets. An example is shown on the right. In this case, for some values of y we need two further functions $\tilde{x}_1(y)$ and $\tilde{x}_2(y)$ to trace the outline of D and the integral in (2.2) is defined to be



$$\int_{x_1(y)}^{x_2(y)} dx = \int_{x_1(y)}^{\tilde{x}_1(y)} dx + \int_{\tilde{x}_2(y)}^{x_2(y)} dx$$

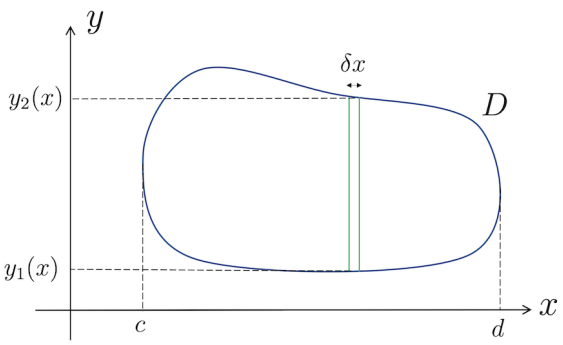
with the obvious generalisation if more disjoint intervals are needed.

We should pause at this point to make a comment on notation. You may be used to writing integrals as $\int$(integrand) dx, with the thing you're integrating sandwiched between the $\int$ sign and the dx. Indeed, that's the convention that we've been using up until now. But, as you progress through mathematics, there is a time to dump this notation and we have now reached that time. When performing multiple integrals, it becomes annoying to remember where you should place all those dx's and dy's, not least because they're not conveying any further information. So we instead write integrals as $\int dx$ (integrand), with the dx placed next to the integral sign. There's nothing deep in this. It's just a different convention, albeit one that holds your hand a little less. Think of it like that time you took the training wheels off your bike.

Our new notation does, however, retain the idea of ordering. You should work from right to left, first performing the $\int dx$ integration in (2.2) to get a function of y, and subsequently performing the $\int dy$ integration.

Note also that the number of $\int$ signs is not conserved in (2.2). On the left, $\int dA$ is an area integral and so requires us to do two normal integrals which are then written explicitly on the right. Shortly we will meet volume integrals and denote them as $\int dV$. Some texts prefer a convention in which there is a conservation of integral signs and so write area integrals as $\int\int dA$ and volume integrals as $\int\int\int dV$. The authors of these texts aren't string theorists and have never had to perform an integral in ten dimensions. Here we refuse to adopt this notation on the grounds that it looks silly.

There is a different way to do the integral (2.2). We could just as well divide our formula D into vertical strips of width δx, so that it looks like this:



For each value of x, we do the y integral between the two limits $y_1(x)$ and $y_2(x)$. As before, these functions trace the shape of the region D. We then subsequently sum over all strips by doing the x integral between the two outer limits of the shape which we now call c and d. Now the result is

$$\int_D \phi(x,y)\, dA = \int_c^d dx \int_{y_1(x)}^{y_2(x)} dy\ \phi(x,y) \tag{2.3}$$

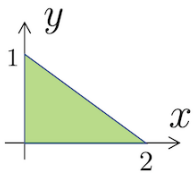
There are other ways to divide up the region D, some of which we will meet below when we discuss different coordinate choices. *Fubini's theorem*, proven in 1907, states that, for suitably well behaved functions $\phi(x,y)$ and regions D, all different ways of decomposing the integral agree. We won't prove this theorem here but it guarantees that the result that you get from doing the integrals in (2.2) coincides with the result from (2.3).

An Example

As a simple example, consider the function

$$\phi(x,y) = x^2 y$$

integrated over the triangle D shown in the figure.



We'll do the area integral in two different ways. If we first do the $\int dx$ integration, as in (2.2), then we have

$$\int_D \phi\, dA = \int_0^1 dy \int_0^{2-2y} dx\ x^2 y = \int_0^1 dy\ y \left[\frac{x^3}{3}\right]_0^{2-2y} = \frac{8}{3}\int_0^1 dy\ y(1-y)^3 = \frac{2}{15}$$

Meanwhile, doing the $\int dy$ integration first, as in (2.3), we have

$$\int_D \phi\, dA = \int_0^2 dx \int_0^{1-x/2} dy\ x^2 y = \int_0^2 dx\ x^2 \left[\frac{y^2}{2}\right]_0^{1-x/2} = \frac{1}{2}\int_0^2 dx\ x^2 \left(1-\frac{x}{2}\right)^2 = \frac{2}{15}$$

The two calculations give the same answer as advertised.

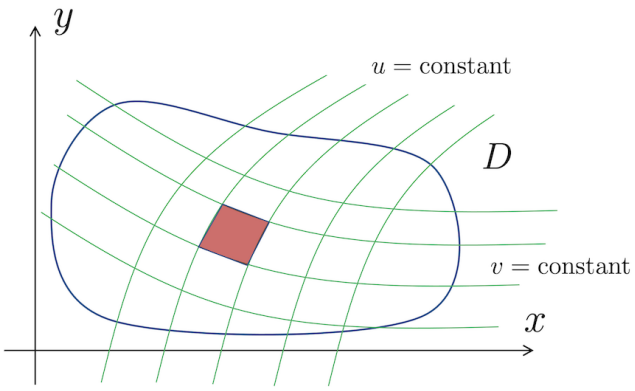


Figure 4. A change of coordinates from (x, y) to (u, v).

2.1.2 Changing Coordinates

Our discussion above was very much rooted in Cartesian coordinates. What if we choose to work with a different set of coordinates on $\mathbb{R}^2$?

Consider a change of variables $(x, y) \to (u, v)$. To be a good change of coordinates, the map should be smooth and invertible and we will assume that this is the case. The region D can then equally well be parameterised by coordinates (u, v). An example is shown in Figure 4, with lines of constant u and constant v plotted in green. We want to know how to do the area integral in the (u, v) coordinates.

Claim: The area integral can be written as

$$\int_D dx\, dy\ \phi(x, y) = \int_{D'} du\, dv\ |J(u, v)|\, \phi(u, v) \tag{2.4}$$

The region D in the (x, y) plane is mapped into a different region D' in the (u, v) plane. Here $\phi(u, v)$ is slightly sloppy shorthand: it means the function $\phi(x(u, v), y(u, v))$. The additional term $J(u, v)$ is called the *Jacobian* and is given by the determinant

$$J(u, v) = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix}$$

The Jacobian is an important enough object that it also gets its own notation and is sometimes written as

$$J = \frac{\partial(x, y)}{\partial(u, v)}$$

Proof(ish): Here is a sketch of the proof to give you some intuition for why this is the right thing to do. We evaluate the integral by summing over small areas δA, formed by lines of constant u and v as shown by the red shaded region in Figure 4. The sides of this small region have length δu and δv respectively, but what is its area? It's not simply $\delta u\,\delta v$ because the sides aren't at necessarily right angles. Instead, the small shaded region is approximately a parallelogram.

We think of the original coordinates as functions of the new, so $x = x(u, v)$ and $y = y(u, v)$. If we make vary u and v slightly, then the change in the original x and y coordinates is

$$\delta x = \frac{\partial x}{\partial u}\delta u + \frac{\partial x}{\partial v}\delta v + \ldots \quad \text{and} \quad \delta y = \frac{\partial y}{\partial u}\delta u + \frac{\partial y}{\partial v}\delta v + \ldots$$

where the $+\ldots$ hide second order terms $\mathcal{O}(\delta u^2)$, $\mathcal{O}(\delta v^2)$ and $\mathcal{O}(\delta u\,\delta v)$. This means that we have

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix}$$

The small parallelogram is then spanned by the two vectors $\mathbf{a} = (\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u})\delta u$ and $\mathbf{b} = (\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v})\delta v$. Recall that the area of a parallelogram is $|\mathbf{a} \times \mathbf{b}|$, so we have

$$\delta A = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| \delta u\,\delta v = |J|\,\delta u\,\delta v$$

which is the promised result □

An Example: 2d Polar Coordinates

There is one particular choice of coordinates that vies with Cartesian coordinates in their usefulness. This is plane polar coordinates, defined by

$$x = \rho\cos\phi \quad \text{and} \quad y = \rho\sin\phi$$

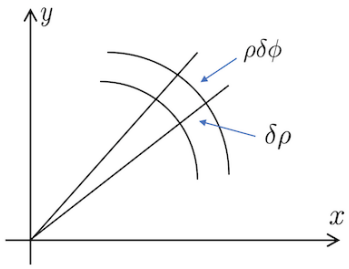
where the radial coordinate $\rho \geq 0$ and the angular coordinate takes values in $\phi \in [0, 2\pi)$. (Note: we used $\phi(x, y)$ to describe a general scalar field earlier in this section. This shouldn't be confused with the coordinate ϕ that we've introduced here.) We can easily compute the Jacobian to find

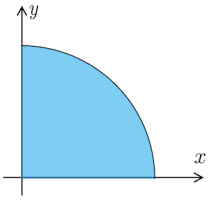
$$J = \frac{\partial(x, y)}{\partial(\rho, \theta)} = \begin{vmatrix} \cos\phi & -\rho\sin\phi \\ \sin\phi & \rho\cos\phi \end{vmatrix} = \rho$$

So we learn that the area element is given by

$$dA = \rho\, d\rho d\phi$$

There is also a simple graphical explanation of this result: it follows by looking at the area of the rounded square shape in the figure to the right (again, ignoring terms second order in $\delta\phi$ and $\delta\rho$).





Let's now use this to do an integral. Let D be a the wedge in the (x, y) plane defined by $x \geq 0$, $y \geq 0$ and $x^2 + y^2 \leq R^2$. This is shown to the left. In polar coordinates, this region is given by

$$0 \leq \rho \leq R \quad \text{and} \quad 0 \leq \phi \leq \frac{\pi}{2}$$

We'll integrate the function $f = e^{-(x^2+y^2)/2} = e^{-\rho^2/2}$ over the region D. In polar coordinates, we have

$$\int_D f\, dA = \int_0^{\pi/2} d\phi \int_0^R d\rho\ \rho e^{-\rho^2/2}$$

where the extra power of ρ in the integrand comes from the Jacobian. The $\int d\phi$ integral just gives us $\pi/2$, while the $\int d\rho$ integral is easily done. We have

$$\int_D f\, dA = \frac{\pi}{2}\Big[- e^{-\rho^2/2}\Big]_0^R = \frac{\pi}{2}\left(1 - e^{-R^2/2}\right)$$

As a final application, consider taking the limit $R \to \infty$, so that we're integrating over the quadrant $x, y \geq 0$. Clearly the answer is $\int_D f\, dA = \pi/2$. Back in Cartesian coordinates, this calculation becomes

$$\int_D f\, dA = \int_0^\infty dx \int_0^\infty dy\ e^{-(x^2+y^2)/2} = \left(\int_0^\infty dx\ e^{-x^2/2}\right)\left(\int_0^\infty dy\ e^{-y^2/2}\right)$$

Comparing to our previous result, we find the well-known expression for a Gaussian integral

$$\int_0^\infty dx\ e^{-x^2/2} = \sqrt{\frac{\pi}{2}}$$

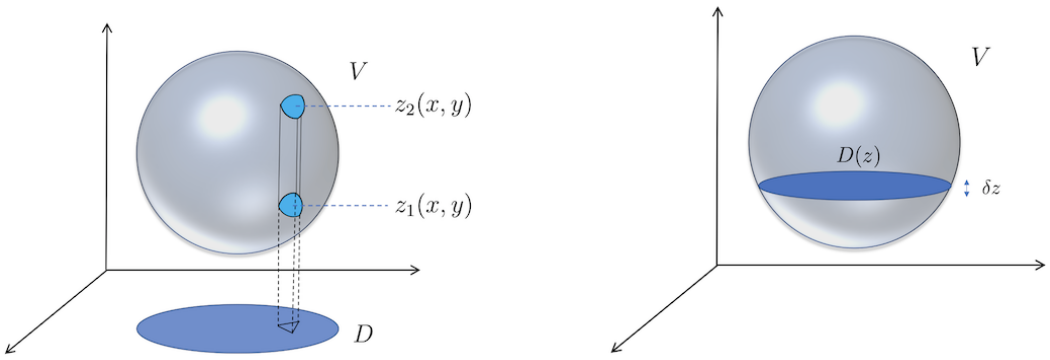


Figure 5. Two different ways to do a volume integral. On the left: perform the $\int dz$ integral first; on the right, perform the $\int_{D(z)} dA$ area integral first.

2.1.3 Volume Integrals

Most of this chapter will be devoted to discussing surfaces, but this is as good a place as any to introduce volume integrals because they are a straightforward generalisation of area integrals.

The basic idea should by now be familiar. The integration of a scalar function $\phi : \mathbb{R}^3 \to \mathbb{R}$ over a three-dimensional region V can be approximated by dividing the region into many small 3d pieces, each with volume δV and located at some position $\mathbf{x}_n$. You then find a way to take the limit

$$\int_V \phi(\mathbf{x})\, dV \;= \lim_{\delta V \to 0} \sum_n \phi(\mathbf{x}_n)\, \delta V$$

In practice, we evaluate volume integrals in the same way as we evaluate area integrals: by performing successive integrations. If we use Cartesian coordinates (x, y, z) we have a number of ways to proceed. For example, we could choose to first do the $\int dz$ integral, subsequently leaving us with an area integral over the (x, y) plane.

$$\int_V \phi(x,y,z)\, dV = \int dA \int_{z_1(x,y)}^{z_2(x,y)} dz\ \phi(x,y,z)$$

This approach is shown on the left-hand side of Figure 5. Alternatively, we could first do an area integral over some sliver of the region V and subsequently integrate over all slivers. This is illustrated on the right-hand side of Figure 5 and results in an integral of the form

$$\int_V \phi(x,y,z)\, dV = \int dz \int_{D(z)} dxdy\ \phi(x,y,z)$$

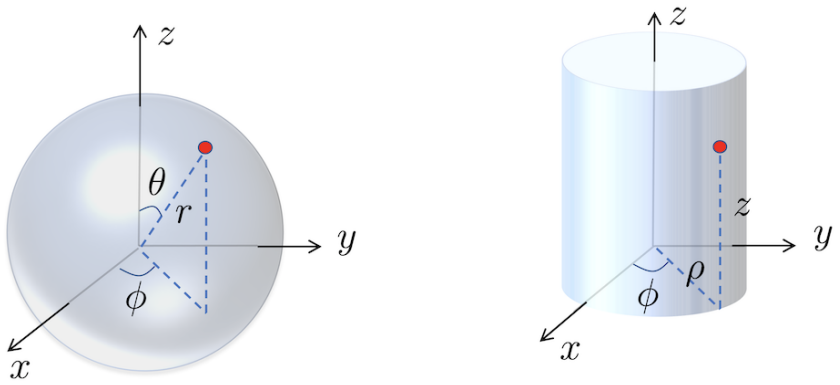


Figure 6. Spherical polar coordinates on the left, and cylindrical polar coordinates on the right.

As before, for suitably nice functions ϕ and regions V, the order of integration is unimportant.

There are many reasons to do a volume integral. You might, for example, want to know the volume of some object, in which case you just integrate the function $\phi = 1$. Alternatively, it's common to integrate a density of something, which means stuff per unit volume. Integrating the density over the region V tells you the amount of stuff in V. Examples of stuff that we will meet in other courses include mass, electric charge and probability.

2.1.4 Spherical Polar and Cylindrical Polar Coordinates

If your region V is some blocky shape, then Cartesian coordinates are probably the right way forward. However, for many applications it is more convenient to use a different choice of coordinates.

Given an invertible, smooth transformation $(x, y, z) \to (u, v, w)$ then the volume elements are mapped to

$$dV = dx\, dy\, dz = |J|\, du\, dv\, dw$$

with the Jacobian given by

$$J = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

The sketch of the proof is identical to the 2d case: the volume of the appropriate parallelepiped is $\delta V = |J|\delta u \delta v \delta w$.

Two sets of coordinates are particularly useful. The first is *spherical polar coordinates*, related to Cartesian coordinates by the map

$$\begin{aligned} x &= r\sin\theta\,\cos\phi \\ y &= r\sin\theta\,\sin\phi \\ z &= r\cos\theta \end{aligned} \tag{2.5}$$

The range of the coordinates is $r \in [0,\infty)$, $\theta \in [0,\pi]$ and $\phi \in [0,2\pi)$. The Jacobian is

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = r^2\sin\theta \quad \Rightarrow \quad dV = r^2\sin\theta\, dr\, d\theta\, d\phi \tag{2.6}$$

The second is *cylindrical polar coordinates*, which coincides with plane polar coordinates in the (x,y) plane, leaving z untouched

$$\begin{aligned} x &= \rho\cos\phi \\ y &= \rho\sin\phi \\ z &= z \end{aligned} \tag{2.7}$$

with $\rho \in [0,\infty)$ and $\phi \in [0,2\pi)$ and, of course, $z \in (-\infty,+\infty)$. (Later in the course, we will sometimes denote the radial coordinate in cylindrical polar coordinates as r instead of ρ.) This time the Jacobian is

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \rho \quad \Rightarrow \quad dV = \rho\, d\rho\, d\phi\, dz$$

We can do some dimensional analysis to check that these results make sense. In spherical polars we have one coordinate, r, with dimensions of length and two dimensionless angular coordinates. Correspondingly, the Jacobian has dimension length^2 to ensure that dV has the dimension of volume. In cylindrical polars, we have two coordinates with dimension of length, ρ and z, and just a single angular coordinate. This is the reason that the Jacobian now has dimension of length rather than length^2.

Example 1: The Volume of a Sphere

Consider a spherically symmetric function $f(r)$. We can integrate it over a ball of radius R using spherical polar coordinates, with $dV = r^2 \sin\theta \, dr d\theta d\phi$ to get

$$\begin{aligned}\int_V f \, dV &= \int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} d\phi \; r^2 f(r) \sin\theta \\ &= 2\pi \Big[-\cos\theta \Big]_0^\pi \int_0^R dr \; r^2 f(r) \\ &= 4\pi \int_0^R dr \; r^2 f(r)\end{aligned}$$

In particular, if we take $f(r) = 1$ then we get the volume of a sphere $\text{Vol} = 4\pi R^3/3$.

Example 2: A Cylinder Cut Out of a Sphere

Next consider a more convoluted example: we want the volume of a sphere of radius R, with a cylinder of radius $s < R$ removed from the middle. The region V is then $x^2 + y^2 + z^2 \leq R^2$, together with $x^2 + y^2 \geq s^2$. Note that we don't just subtract the volume of a cylinder from the that of a sphere because the top of the cylinder isn't flat: it stops where it intersects the sphere.

In cylindrical coordinates, the region V spans $s \leq \rho \leq R$ and $-\sqrt{R^2 - \rho^2} \leq z \leq \sqrt{R^2 - \rho^2}$. And, of course, $0 \leq \phi < 2\pi$. We have $dV = \rho \, d\rho dz d\phi$ and

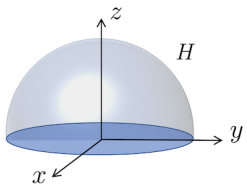
$$\text{Vol} = \int_V dV = \int_0^{2\pi} d\phi \int_s^R d\rho \; \rho \int_{-\sqrt{R^2-\rho^2}}^{\sqrt{R^2-\rho^2}} dz = 4\pi \int_s^R d\rho \; \rho \sqrt{R^2 - \rho^2}$$

It is now straightforward to do the integral to find the volume

$$\text{Vol} = \frac{4\pi}{3}(R^2 - s^2)^{3/2}$$

Example 3: Electric Charge On a Hemisphere

Consider a density of electric charge that increases linearly in the z-direction, with $f(z) = f_0 z/R$, in a hemisphere H of radius R, with $z \geq 0$ and f_0 a constant. What is the total charge in H?



In spherical polar coordinates, the coordinates for the hemisphere H are $0 \leq r \leq R$ and $0 \leq \phi < 2\pi$ and, finally, $0 \leq \theta \leq \pi/2$, which restricts us to the hemisphere with

$z \geq 0$. We integrate the function $f = f_0 r\cos\theta/R$ over H with $dV = r^2\sin\theta\, dr d\theta d\phi$ to find

$$\int_H f\, dV = \frac{f_0}{R}\int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta \int_0^R dr\ r^2\sin\theta\, r\cos\theta = \frac{2\pi f_0}{R}\left[\frac{r^4}{4}\right]_0^R \left[\frac{1}{2}\sin^2\theta\right]_0^{\pi/2} = \frac{1}{4}\pi R^3 f_0$$

As a quick check on our answer, note that f_0 is the charge density so the dimensions of the final answer are correct: the total charge is equal to the charge density times a volume.

Vector Valued Integrals

We can also integrate vector valued fields $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ over a volume V. There's nothing subtle here: we just do the integral component by component and the final answer is also a vector.

A common example arises when we compute the centre of mass. Let $\rho(\mathbf{x})$ be the density of an object. (Note that this isn't a great choice of notation if we're working in cylindrical polar coordinates.) The total mass is

$$M = \int_V \rho(\mathbf{x})\, dV$$

and the *centre of mass* is

$$\mathbf{X} = \frac{1}{M}\int_V \mathbf{x}\rho(\mathbf{x})\, dV$$

For example, consider again the solid hemisphere H from the previous example, covering $0 \leq r \leq R$ and $z \geq 0$. We'll take this object to have constant density ρ. The total mass is

$$M = \int_H \rho\, dV = \frac{2\pi}{3}\rho R^3$$

Writing $\mathbf{X} = (X, Y, Z)$ for the centre of mass, we need to compute the three components individually. We have

$$\begin{aligned} X = \frac{\rho}{M}\int_H x\, dV &= \frac{\rho}{M}\int_0^{2\pi} d\phi \int_0^R dr \int_0^{\pi/2} d\theta\ xr^2\sin\theta \\ &= \frac{\rho}{M}\int_0^{2\pi} d\phi \int_0^R dr \int_0^{\pi/2} d\theta\ r^3\sin^2\theta\cos\phi = 0 \end{aligned}$$

where the integral $\int d\phi \cos\phi = 0$. A similar calculation shows that $Y = 0$. Indeed, the fact that the centre of mass lies at $(X, Y) = (0, 0)$ follows on symmetry grounds. We're left only computing the centre of mass in the z-direction. This is

$$Z = \frac{\rho}{M}\int_H z\, dV = \frac{\rho}{M}\int_0^{2\pi} d\phi \int_0^R dr \int_0^{\pi/2} d\theta\ r^3 \cos\theta \sin\theta = \frac{3R}{8}$$

We learn that the centre of mass sits at $\mathbf{X} = (0, 0, 3R/8)$.

Generalisation to $\mathbb{R}^n$

Finally, it is straightforward to generalise multiple integrals to $\mathbb{R}^n$. If we make a smooth, invertible change of coordinates from Cartesian $x^1, \ldots, x^n$ to some other coordinates $u^1, \ldots, u^n$ then the integral over some n-dimensional region M is

$$\int_M f(x^i)\, dx^1\ \ldots dx^n = \int_{M'} f(x(u^i))\, |J|\, du^1 \ldots du^n$$

where the Jacobian

$$J = \frac{\partial(x^1, \ldots, x^n)}{\partial(u^1, \ldots, u^n)} = \det\left(\frac{\partial x^i}{\partial u^a}\right)$$

is the obvious generalisation of our previous results.

2.2 Surface Integrals

Our next task is to understand how to integrate over a surface that doesn't lie flat in $\mathbb{R}^2$, but is instead curved in some way in $\mathbb{R}^3$. We will start by looking at how we define such surfaces in the first place.

2.2.1 Surfaces

There are (at least) two different ways to describe a surface in $\mathbb{R}^3$.

- A surface can be viewed as the level set of a function,

  $$F(x, y, z) = 0$$

  This is one condition on three variables, so results in a two-dimensional surface in $\mathbb{R}^3$. (In general, a single constraint like this results in an $(n-1)$-dimensional space in $\mathbb{R}^n$. Alternatively we say that the space has *codimension* one.)

- We can consider a *parameterised surface*, defined by the map

  $$\mathbf{x} : \mathbb{R}^2 \to \mathbb{R}^3$$

  This is the extension of the parameterised curve that we discussed in Section 1. This now defines a dimension two surface in any space $\mathbb{R}^3$.

At each point on the surface, we can define a *normal vector*, $\mathbf{n}$ which points perpendicularly away from the surface. When the surface is defined as the level set of function $F(\mathbf{x}) = 0$, the normal vector lies in the direction

$$\mathbf{n} \sim \nabla F$$

To see this, note that $\mathbf{m} \cdot \nabla F$ describes the rate of change of F in the direction $\mathbf{m}$. If $\mathbf{m}$ lies tangent to the surface then we have, by definition, $\mathbf{m} \cdot \nabla F = 0$. Conversely, the normal to the surface $\mathbf{n}$ lies in the direction in which the function F changes most quickly, and this is ∇F.

It's traditional to normalise the normal vector, so we usually define

$$\mathbf{n} = \pm \frac{1}{|\nabla F|} \nabla F$$

where we'll say more about the choice of minus sign below.

Meanwhile, for the parameterised surface $\mathbf{x}(u, v) \in \mathbb{R}^3$, we can construct two tangent vectors to the surface, namely

$$\frac{\partial \mathbf{x}}{\partial u} \quad \text{and} \quad \frac{\partial \mathbf{x}}{\partial v}$$

where each partial derivative is taken holding the other coordinate fixed. Each of these lies within the surface, so the normal direction is

$$\mathbf{n} \sim \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}$$

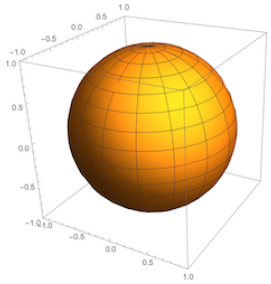
If $\mathbf{n} \neq 0$ anywhere on the surface then the parameterisation is said to be *regular*. Note that, although a parameterised surface can be defined in any $\mathbb{R}^n$, the normal direction is only unique in $\mathbb{R}^3$ where we have the cross product at our disposal.

Examples

Here are a number of examples using the definition involving a level set. A sphere of radius R is defined by

$$F(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0$$

the normal direction is given by $\nabla F = 2(x, y, z)$ and points radially outwards.



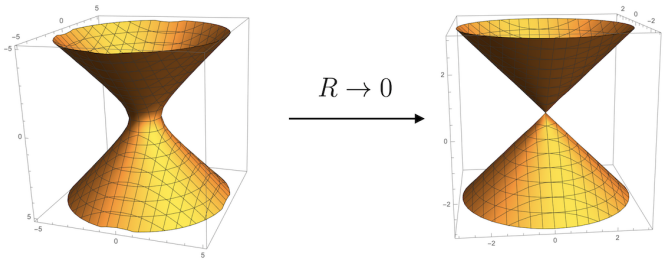


Figure 7. When the hyperboloid degenerates to a cone, there is no well defined normal at the origin.

A hyperboloid is defined by

$$F(x, y, z) = x^2 + y^2 - z^2 - R^2 = 0$$

with normal direction given by $\nabla F = 2(x, y, -z)$. Note that for both the sphere and hyperboloid, the normal vector is nowhere vanishing because the origin $\mathbf{x} = 0$ doesn't lie on the surface. However, if we take the limit $R \to 0$ then the hyperboloid degenerates to two cones, meeting at the origin. In this case, $\nabla F = 0$ at the origin, reflecting the fact there is no unique direction away from the surface at this point. This is shown in Figure 7

2.2.2 Surfaces with Boundaries

A surface S can have a boundary. This is a piecewise smooth closed curve. If there are several boundaries, then this curve should be thought of as having several disconnected pieces.

For example, we could define the surfaces above now restricted to the region $z \geq 0$. In this case both the sphere and hyperboloid are truncated and their boundary is the circle $x^2 + y^2 = R^2$ in the $z = 0$ plane.

The boundary of a surface S is denoted ∂S with ∂ the standard notation to denote the boundary of any object. For example, later in the lectures we will denote the boundary of a 3d volume V as ∂V. You might reasonably wonder why we use the partial derivative symbol ∂ to denote the boundary of something. There are some deep and beautiful reasons behind this that will only become apparent in later courses. But there is also a simple, intuitive reason. Consider a collection of 3d objects, all the same shape but each bigger than the last. We'll denote these volumes as V_r. Then, roughly

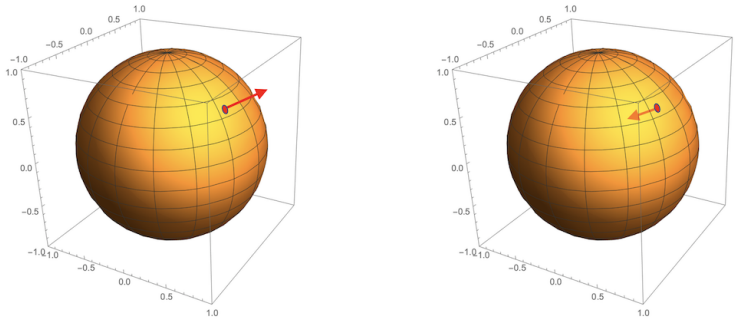


Figure 8. Two orientations of a sphere, with the unit normal pointing outwards or inwards.

speaking, you can view the boundary surface as

$$\partial V_r = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\Big(V_{r+\epsilon} \setminus V_r\Big)$$

where $\setminus$ means that you remove the 3d object V_r from inside the slightly larger object $V_{r+\epsilon}$. This, of course, looks very much like the formula for a derivative.

This "derivative equals boundary" idea also shows up when we calculate volumes, areas and lengths. For example, a disc of radius r has area πr^2. The length of the boundary is $\frac{d}{dr}(\pi r^2) = 2\pi r$. This relation continues to higher dimensional balls and spheres.

There is something important lurking in the idea of a boundary. The boundary is necessarily a *closed* curve C, meaning that it has no end points. Another way of saying this is that a closed curve C itself has no boundary, or $\partial C = 0$. We see that if a curve arises as the boundary of a surface, then the curve itself has no boundary. This is captured in the slogan "the boundary of a boundary vanishes" or, in equation form, $\partial^2 S = 0$. It is a general and powerful principle that extends to higher dimensional objects where $\partial^2(\text{anything}) = 0$. The idea that the boundary of a boundary vanishes is usually expressed simply as $\partial^2 = 0$.

A couple of quick definitions. A surface is said to be *bounded* if it doesn't stretch off to infinity. More precisely, a bounded surface can be contained within some solid sphere of fixed radius. A surface that does stretch off to infinity is said to be *unbounded*. Obviously, the sphere is a bounded surface, while the hyperboloid is unbounded.

Finally, a bounded surface with no boundary is said to be *closed*.

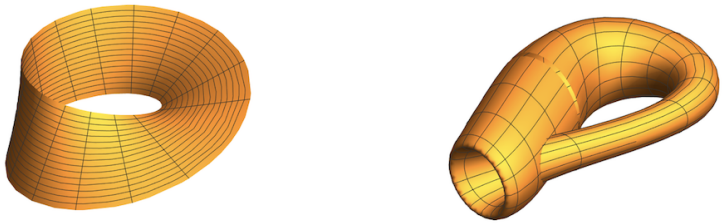


Figure 9. Two unorientable surfaces: the Möbius strip on the left, and the Klein bottle on the right.

2.2.3 Orientability

As long as the normal vector $\mathbf{n} \neq 0$, we can always normalise it so that it has unit length. But in general, there is no canonical way to fix the sign. This is a matter of convention and determines what we mean by "outside the surface" and what we mean by "inside".

A surface is said to be *orientable* if there is a consistent choice of unit normal $\mathbf{n}$ which varies smoothly over the surface. The sphere and hyperboloid above are both orientable, with the two choices of an orientation for the sphere shown in Figure 8. Throughout these lectures we will work only with orientable surfaces. For such surfaces, a choice of sign fixes the unit normal everywhere and is said to determine the *orientation* of the surface.

We note in passing that unorientable surfaces exist. You can easily make one of these in the comfort of your own home. Take a strip of paper and glue the two ends together. You've got two different ways to glue them as shown by the arrows below:

If you glue by aligning the arrows shown on the left, then you're just left with a boring strip of paper. But if you glue with the arrows aligned on the right, then you end up with something new and exciting: an unorientable surface, known as a *Möbius strip*. You can see one that I made earlier on the left of Figure 9. If you pick a normal vector and evolve it smoothly around the strip then you'll find that it comes back pointing in the other direction. Relatedly, the Möbius strip has a single boundary, rather than two.

We can also make closed, unorientable surfaces using a similar construction. This time we glue together two edges, like so

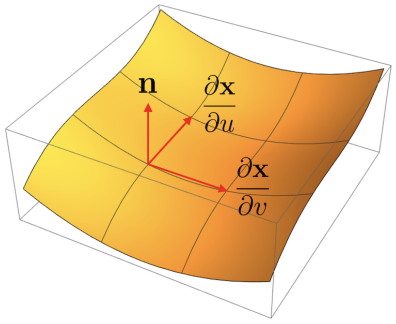
Again, you have various choices. If you glue with the arrows aligned as shown on the left, then you'll end up with a torus which is very much orientable. If you glue with the arrows aligned as shown on the right then you get an unorientable surface called a *Klein bottle*. It's a little tricky to draw embedded in 3d space (and, indeed, tricky to make with paper and glue) as it appears to intersect itself, but an attempt is shown on the right of Figure 9.

2.2.4 Scalar Fields

We're now in a position to start integrating objects over surfaces. For this, we work with parameterised surfaces $\mathbf{x}(u, v)$.

Sit at some point (u, v) on the surface, and move in both directions by some small amount δu and δv. This defines an approximate parallelogram on the surface, as shown in the figure. The area of this parallelogram is

$$\delta S = \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| \delta u \, \delta v$$



where, as usual, we've dropped higher order terms. This is called the *scalar area*. (We'll see the need for the adjective "scalar" below when we introduce a variant known as the vector area.)

Now we're in a position to define the surface integral of a scalar field $\phi(\mathbf{x})$. Given a parameterised surface S, the surface integral is given by

$$\int_S \phi(\mathbf{x}) \, dS = \int_D du dv \, \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| \phi(\mathbf{x}(u, v)) \tag{2.8}$$

where D is the appropriate region in the (u, v) plane. This is now the same kind of area integral that we learned to do in Section 2.1.

The area integral of a scalar field does not depend on the orientation of the surface. It doesn't matter what you choose as the inside of the surface S and what you choose as the outside, the integral of a scalar field over S always gives the same answer. In particular, if we integrate $\phi = 1$ over a surface S then we get the area of that surface, and this is always positive. This is entirely analogous to the line integral of a scalar field that we met in Section 1.2 that was independent of the orientation of the curve.

Reparameterisation Invariance

Importantly, the surface integral (2.8) is independent of the choice of parameterisation of the surface. To see this, suppose that we replace our original parameterisation $\mathbf{x}(u, v)$ with an alternative parameterisation $\mathbf{x}(\tilde{u}, \tilde{v})$, both of which are assumed to be regular. We then have

$$\frac{\partial \mathbf{x}}{\partial u} = \frac{\partial \mathbf{x}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial u} + \frac{\partial \mathbf{x}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial u} \quad \text{and} \quad \frac{\partial \mathbf{x}}{\partial v} = \frac{\partial \mathbf{x}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial v} + \frac{\partial \mathbf{x}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial v}$$

Taking the cross-product, we have

$$\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} = \frac{\partial(\tilde{u}, \tilde{v})}{\partial(u, v)} \frac{\partial \mathbf{x}}{\partial \tilde{u}} \times \frac{\partial \mathbf{x}}{\partial \tilde{v}}$$

This means that the scalar area element can equally well be written as

$$dS = \left| \frac{\partial \mathbf{x}}{\partial \tilde{u}} \times \frac{\partial \mathbf{x}}{\partial \tilde{v}} \right| d\tilde{u}\, d\tilde{v}$$

where we've used the result (2.4) which, in the current context, is $d\tilde{u}\, d\tilde{v} = \frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}\, du\, dv$.

The essence of this calculation is the same as we saw for line integrals: the two derivatives $\partial/\partial u$ and $\partial/\partial v$ in the integrand cancel the Jacobian factor under a change of variables. The upshot is that we can write the surface integral (2.8) using any parameterisation that we wish: the answer will be the same.

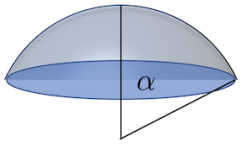
An Example

Consider a sphere of radius R. Let S be the subregion that sits at an angle $\theta \leq \alpha$ from the vertical. This is the grey region shown in the figure. We want to compute the area of this cap.

We start by constructing a parameterisation of the sphere. This is straightforward if we use the spherical polar angles θ and ϕ defined in (2.5) as parameters. We have

$$\mathbf{x}(\theta, \phi) = R(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) := R\,\mathbf{e}_r$$

Here $\mathbf{e}_r$ is the unit vector that points radially outwards. (We will also use the notation $\mathbf{e}_r = \hat{\mathbf{r}}$ later in these lectures.) We can then easily calculate



$$\frac{\partial \mathbf{x}}{\partial \theta} = R(\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta) := R\,\mathbf{e}_\theta$$
$$\frac{\partial \mathbf{x}}{\partial \phi} = R(-\sin\theta\sin\phi, \sin\theta\cos\phi, 0) := R\sin\theta\,\mathbf{e}_\phi$$

Here, by construction, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ are unit vectors pointing in the direction of increasing θ and ϕ respectively. We'll have more to say about the triplet of vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ in Section 3.3. For now, we can compute

$$\frac{\partial \mathbf{x}}{\partial \theta} \times \frac{\partial \mathbf{x}}{\partial \phi} = R^2 \sin\theta\,\mathbf{e}_r$$

From this, we have the scalar area element

$$dS = R^2 \sin\theta\, d\theta\, d\phi \tag{2.9}$$

We've seen a result very similar to this before. The volume element in spherical polar coordinates (2.6) is $dV = r^2\sin\theta\, dr\, d\theta\, d\phi$. Our area element over a sphere simply comes from setting $r = R$ and ignoring the dr piece of the volume element.

It is now straightforward to compute the area. We have

$$A = \int_0^{2\pi} d\phi \int_0^{\alpha} d\theta\ R^2\sin\theta = 2\pi R^2(1-\cos\alpha)$$

Note that if we set $\alpha = \pi$ then we get the area of a full sphere: $A = 4\pi R^2$.

2.2.5 Vector Fields and Flux

Now we turn to vector fields. There is a particularly interesting and useful way to integrate a vector field $\mathbf{F}(\mathbf{x})$ over a surface S so that we end up with a number. We do this by taking the inner product of the vector field with the normal to the surface, $\mathbf{n}$, so that

$$\int_S \mathbf{F}(\mathbf{x})\cdot\mathbf{n}\, dS = \int_D du dv\ \left(\frac{\partial \mathbf{x}}{\partial u}\times\frac{\partial \mathbf{x}}{\partial v}\right)\cdot\mathbf{F}(\mathbf{x}(u,v)) \tag{2.10}$$

This is called the *flux* of $\mathbf{F}$ through S.

The definition of the flux is independent of our choice of parameterisation: the argument is identical to the one we saw above for a scalar field.

It's convenient to introduce some new notation. The *vector area element* is defined as

$$d\mathbf{S} = \mathbf{n}\, dS = \left(\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right) du\, dv$$

This has magnitude dS and points in the normal direction $\mathbf{n}$.

The flux of a vector field depends on the orientation of the surface S. This can be seen in the presence of the normal vector in (2.10). In the parameterised surface $\mathbf{x}(u, v)$, the choice of orientation can be traced to the parameterisation (u, v) and, in particular, the order in which they appear in the cross product. Changing the orientation of the surface flips the sign of the flux.

The physical importance of the flux can be seen by thinking about a fluid. Let $\mathbf{F}(\mathbf{x})$ be the *velocity field* of a fluid. (Usually we would denote this as $\mathbf{u}(\mathbf{x})$ or $\mathbf{v}(\mathbf{x})$, but we've already used u and v as the parameters of the surface so we'll adopt the non-standard name $\mathbf{F}$ for the velocity to avoid confusion.) In a small time δt, the amount of fluid flowing through a small surface element δS is given by

$$\text{Fluid Flow} = \mathbf{F}\, \delta t \cdot \mathbf{n}\, \delta S$$

where the dot product ensures that we don't include the component of fluid that flows parallel to the surface. Integrating over the whole surface, we see that the flux of fluid

$$\text{Flux} = \int_S \mathbf{F} \cdot d\mathbf{S}$$

is the amount of fluid crossing S *per unit time*. In other words, the flux is the rate of fluid flow.

We also talk of "flux" in other contexts, where there's no underlying flow. For example, in our course on Electromagnetism, we will spend some time computing the flux of the electric field through various surfaces, $\int_S \mathbf{E} \cdot d\mathbf{S}$.

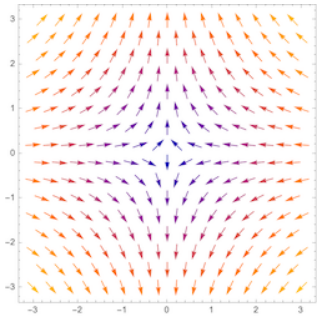
An Example

Consider the vector field

$$\mathbf{F} = (-x, 0, z)$$

This is plotted in the $y =$ constant plane in the figure.

We want to integrate this vector field over the hemispherical cap, subtended by the angle α that we used as an example in Section 2.2.4. This is the region of a sphere of radius R, spanned by the polar coordinates



$$0 \leq \theta \leq \alpha \quad \text{and} \quad 0 \leq \phi < 2\pi$$

We know from our previous work that

$$d\mathbf{S} = R^2 \sin\theta\, \mathbf{e}_r\, d\theta\, d\phi \quad \text{with} \quad \mathbf{e}_r = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

In particular, we have

$$\mathbf{F}\cdot\mathbf{e}_r = -x\sin\theta\cos\phi + z\cos\theta = R(-\sin^2\theta\cos^2\phi + \cos^2\theta)$$

The flux through the hemispherical cap is then

$$\int \mathbf{F}\cdot d\mathbf{S} = \int_0^\alpha d\theta \int_0^{2\pi} d\phi\ R^3\sin\theta\left[-\sin^2\theta\cos^2\phi + \cos^2\theta\right]$$

We use $\int_0^{2\pi} d\phi\ \cos^2\phi = \pi$ to get

$$\begin{aligned}\int \mathbf{F}\cdot d\mathbf{S} &= \pi R^3 \int_0^\alpha d\theta\ \sin\theta\left[-\sin^2\theta + 2\cos^2\theta\right] \\ &= \pi R^3\Big[\cos\theta\sin^2\theta\Big]_0^\alpha = \pi R^3\cos\alpha\sin^2\alpha \end{aligned} \tag{2.11}$$

2.2.6 A Sniff of the Gauss-Bonnet Theorem

The methods described in this section have many interesting applications to geometry. Here we sketch two important ideas. We prove neither.

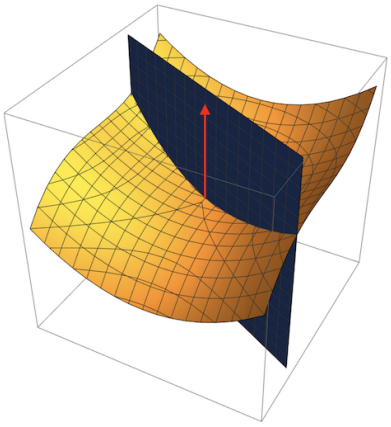
Consider a surface S and pick a point with normal $\mathbf{n}$. We can construct a plane containing $\mathbf{n}$, as shown in the figure. The intersection of the original surface and the plane describes a curve C that lies in S. Associated to this curve is a curvature κ, defined in (1.6).

Now, we rotate the plane about $\mathbf{n}$. As we do so, the curve C changes and so too does the curvature. Of particular interest are the maximum and minimum curvatures

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}$$

These are referred to as *principal curvatures*. The *Gaussian curvature* of the surface S at our chosen point is then defined to be

$$K = \kappa_{\min}\kappa_{\max}$$



As defined, the curvature K would appear to have as much to do with the embedding of the surface in $\mathbb{R}^3$ as the surface itself. The *theorem egregium* (or "remarkable" theorem) due to Gauss, is the statement that this is misleading: the curvature K is a property of the surface alone, irrespective of any choice of embedding. We say that K is *intrinsic* to the surface.

The idea that curved surfaces have a life of their own, independent of their embedding, is an important one. It generalises to higher dimensional spaces, known as *manifolds*, which are the subject of differential geometry. In physics, curved space (or, more precisely, curved spacetime) provides the framework for our understanding of gravity. Both Riemannian geometry and its application to gravity will be covered in lectures on General Relativity.

The Gaussian curvature K has a number of interesting properties. Here's one. Consider a *geodesic triangle* drawn on the surface as shown in Figure 10. This means that we connect three points with *geodesics*, which are lines of the shortest distance as measured using the arc length (1.5). Let θ_1, θ_2 and θ_3 be the interior angles of the triangle, defined by the inner product of tangent vectors of the geodesic curves. Then it turns out that

$$\theta_1 + \theta_2 + \theta_3 = \pi + \int_D K\, dS \tag{2.12}$$

where D is the interior region of the triangle. If the triangle is drawn on flat $\mathbb{R}^2$, then $K = 0$ and this theorem reduces to the well known statement that the angles of a triangle add up to π.

We can check this formula for the simple case of a triangle drawn on a sphere. If the sphere has radius R then the geodesics are great circles and, as we saw in Section 1.1, they all have curvature $\kappa = 1/R$. Correspondingly, the Gaussian curvature for a sphere is $K = 1/R^2$. A geodesic triangle is shown in the figure to the below: it has two right-angles $\pi/2$ sitting at the equator, and an angle α at the top.

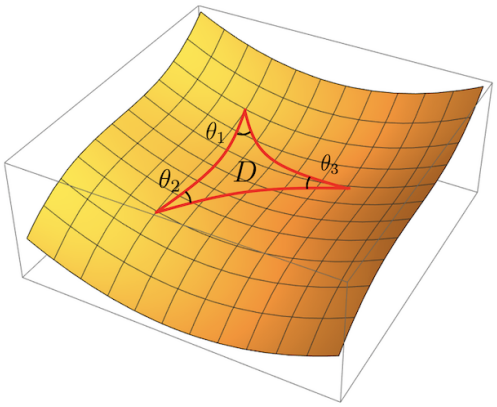
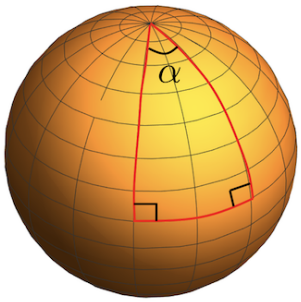


Figure 10. A geodesic triangle inscribed on a surface.

The area of the region inside the triangle is $A = \alpha R^2$ (so that $A = 2\pi R^2$ when $\alpha = 2\pi$ which is the area of the upper hemisphere.). We then have



$$\int_D K\, dS = \frac{A}{R^2} = \alpha$$

which agrees with the result (2.12).

Here's another beautiful application of the Gaussian curvature. Consider a closed surface S. Any such surface can be characterised by the number of holes that it has. This number of holes is known as the *genus*. Three examples are given in Figure 11: a sphere with $g = 0$, a torus with $g = 1$ and some kind of baked-good with genus $g = 3$. It turns out that if you integrate the Gaussian curvature over the entire surface then you get

$$\int_S K\, dS = 4\pi(1 - g) \tag{2.13}$$

This result is all kinds of wonderful. The genus g tells us about the topology of the surface. It's a number that only makes sense when you stand back and look at the object as a whole. In contrast, the Gaussian curvature is a locally defined object: at any given point it depends only on the neighbourhood of that point. But this result tells us that integrating something local can result in something global.

The round sphere provides a particularly simple example of this result. As we've seen above, the Gaussian curvature is $K = 1/R^2$ which, when integrated over the

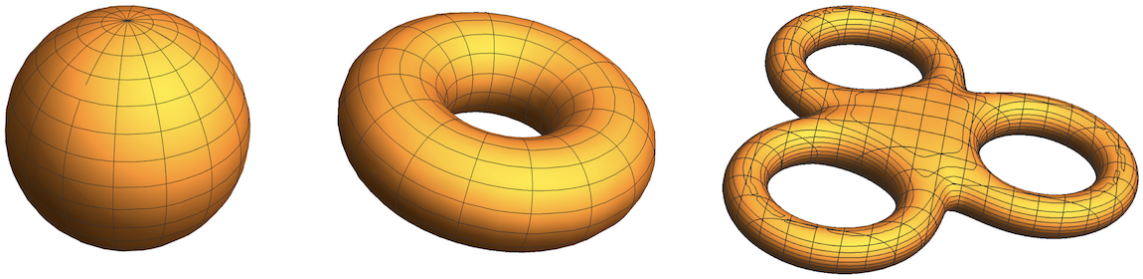


Figure 11. Three closed surfaces with different topologies. The sphere has genus $g = 0$, the torus has genus $g = 1$ and the surface on the right has $g = 3$.

whole sphere, does indeed give 4π as befits a surface of genus $g = 0$. However, this simple calculation hides the magic of the formula (2.13). Suppose that we start to deform the sphere. We might choose to pull it out in some places, push it inwards in others. We could try to mould some likeness of our face in some part of it. Everything that we do changes the local Gaussian curvature. It will increase in some parts and decrease in others. But the formula (2.13) tells us that this must, at the end of the day, cancel out. As long as we don't tear the surface, so its topology remains that of a sphere, the integral of K will always give 4π.

The results (2.12) and (2.13) are two sides of the wondrous Gauss-Bonnet theorem. A proof of this theorem will have to wait for later courses. (You can find a somewhat unconventional proof using methods from physics in the lectures on Supersymmetric Quantum Mechanics. This proof also works for a more powerful generalisation to higher dimensional spaces, known as the Chern-Gauss-Bonnet theorem.)