

Calculus III

Lukas Zamora

November 28, 2017

Contents

Contents	1
1 Vectors	5
1.1 The Basics	5
1.2 Vector Arithmetic	6
1.3 Dot Product	8
1.4 Cross Product	9
2 Three Dimensional Space	13
2.1 The 3-D Coordinate System	13
2.2 Equations of Lines	14
2.3 Equations of Planes	16
2.4 Quadric Surfaces	16
2.5 Vector Functions	18
2.6 Calculus with Vector Functions	19
2.7 Tangent, Normal, and Binormal Vectors	20
2.8 Arc Length with Vector Functions	20
2.9 Curvature	21
2.10 Velocity and Acceleration	22
2.11 Cylindrical Coordinates	22
2.12 Spherical Coordinates	23
3 Partial Derivatives	25
3.1 Limits	25
3.2 Partial Derivatives	27
3.3 Higher Order Partial Derivatives	28

3.4	Differentials	29
3.5	Chain Rule	29
3.6	Directional Derivatives	31
4	Applications of Partial Derivatives	33
4.1	Tangent Planes and Linear Approximations	33
4.2	Gradient Vector, Tangent Planes and Normal Lines	35
4.3	Relative Minimums and Maximums	36
4.4	Absolute Minimums and Maximums	37
4.5	Lagrange Multipliers	39
5	Multiple Integrals	41
5.1	Double Integrals	41
5.2	Iterated Integrals	43
5.3	Double Integrals Over General Regions	44
5.4	Double Integrals in Polar Coordinates	47
5.5	Triple Integrals	49
5.6	Triple Integrals in Cylindrical Coordinates	52
5.7	Triple Integrals in Spherical Coordinates	52
5.8	Change of Variables	53
5.9	Surface Area	54
5.10	Physical Applications of Double Integrals	55
6	Line Integrals	57
6.1	Vector Fields	57
6.2	Line Integrals - Part I	58
6.3	Line Integrals - Part II	60
6.4	Line Integrals of Vector Fields	61
6.5	Fundamental Theorem for Line Integrals	61
6.6	Conservative Vector Fields	63
6.7	Green's Theorem	64
6.8	Curl and Divergence	66
7	Surface Integrals	69
7.1	Parametric Surfaces	69

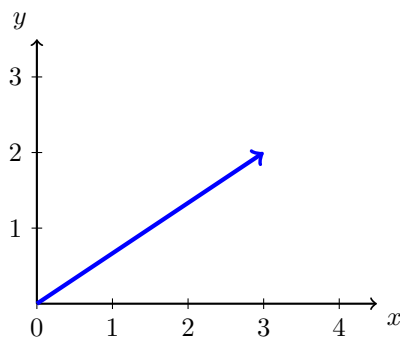
<i>CONTENTS</i>	3
7.2 Surface Integrals	70
7.3 Surface Integrals of Vector Fields	71
7.4 Stokes' Theorem	72
7.5 Divergence Theorem	73
	73

Chapter 1

Vectors

1.1 The Basics

Vectors are used to represent quantities that have both a magnitude and a direction. Good examples of quantities that can be represented by vectors are force and velocity. Both of these have a direction and a magnitude. Consider the sketch below



In this case the vector starts at a specific point then moves 3 units to the right and 2 units up. The notation that we'll use for this vector is,

$$\mathbf{v} = \langle 3, 2 \rangle$$

and each of the directed line segments in the sketch are called *representations* of the vector. Be careful to distinguish vector notation, $\langle 3, 2 \rangle$, from the notation we use to represent coordinates of points, $(3, 2)$. The vector denotes a magnitude and a direction of a quantity while the point denotes a location in space. So don't mix the notations up!

A representation of the vector $\mathbf{v} = \langle a_1, a_2 \rangle$ in two dimensional space is any directed line segment, \overrightarrow{AB} , from the point $A = (x, y)$ to the point $B = (x + a_1, y + a_2)$. Likewise a representation of the vector $\mathbf{v} = \langle a_1, a_2, a_3 \rangle$ in three dimensional space is any directed line segment, \overrightarrow{AB} , from the point $A = (x, y, z)$ to the point $B = (x + a_1, y + a_2, z + a_3)$. Note that there is very little difference between the two dimensional and three dimensional formulas. To get from the three dimensional

formula to the two dimensional formula all we did is take out the third component/coordinate. Because of this most of the formulas here are given only in their three dimensional version. If we need them in their two dimensional form we can easily modify the three dimensional form.

The representation of the vector $\mathbf{v} = \langle a_1, a_2, a_3 \rangle$ that starts at the point $A = (0, 0, 0)$ and ends at the point $B = (a_1, a_2, a_3)$ is called the *position vector* of the point (a_1, a_2, a_3) . Position vectors are useful if we ever need to represent a point as a vector.

The *magnitude*, or length of the vector $\mathbf{v} = \langle a_1, a_2, a_3 \rangle$ is given by,

$$\|\mathbf{v}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

We also have the fact that

$$\text{If } \|\mathbf{a}\| = 0 \text{ then } \mathbf{a} = \mathbf{0}$$

Which makes sense. Because we square all the components the only way we can get zero out of the formula was for the components to be zero in the first place.

Any vector with magnitude of 1, is called the *unit vector*. Given a vector \mathbf{v} , $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ will be a unit vector that points in the same direction as \mathbf{v} . The vector whose components are all zeros, is called the *zero vector*, denoted by $\mathbf{0}$.

In \mathbb{R}^3 , there are unit vectors that are called the *standard basis vectors*. They are denoted by

$$\hat{\mathbf{i}} = \langle 1, 0, 0 \rangle \quad \hat{\mathbf{j}} = \langle 0, 1, 0 \rangle \quad \hat{\mathbf{k}} = \langle 0, 0, 1 \rangle$$

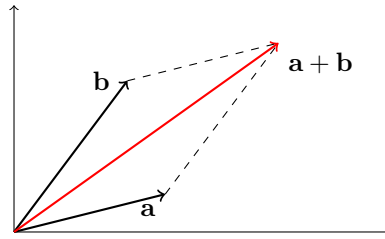
which are also unit vectors.

1.2 Vector Arithmetic

Given the vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, the addition of the two vectors is given by

$$\mathbf{a} + \mathbf{b} = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle$$

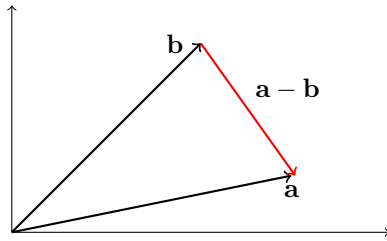
The following figure gives the geometric interpretation of the addition of two vectors



Subtraction is very similar. Given the vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, the difference of the two vectors is given by

$$\mathbf{a} - \mathbf{b} = \langle a_1 - b_1, a_2 - b_2, a_3 - b_3 \rangle$$

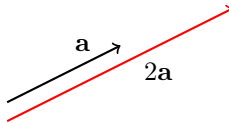
Here is the geometric interpretation of the difference of two vectors



The next arithmetic operation that we want to look at is *scalar multiplication*. Given the vector $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and any number r the scalar multiplication is

$$r\mathbf{a} = \langle ra_1, ra_2, ra_3 \rangle$$

Here is its geometric interpretation. Suppose $r = 2$, then



Two vectors are parallel if they have the same direction or are in exactly opposite directions. Let's suppose that \mathbf{a} and \mathbf{b} are parallel vectors. If they are parallel then there must be a number r so that,

$$\mathbf{a} = r\mathbf{b}$$

So, two vectors are parallel if one is a scalar multiple of the other.

In the previous section we introduced the idea of standard basis vectors without discussing why they were important, so let's do that. Start with the vector

$$\mathbf{a} = \langle a_1, a_2, a_3 \rangle$$

We can use the addition of vectors to break this up

$$\begin{aligned} \mathbf{a} &= \langle a_1, a_2, a_3 \rangle \\ &= \langle a_1, 0, 0 \rangle + \langle 0, a_2, 0 \rangle + \langle 0, 0, a_3 \rangle \end{aligned}$$

and use scalar multiplication to break this up as follows

$$\begin{aligned} \mathbf{a} &= \langle a_1, 0, 0 \rangle + \langle 0, a_2, 0 \rangle + \langle 0, 0, a_3 \rangle \\ &= a_1 \langle 1, 0, 0 \rangle + a_2 \langle 0, 1, 0 \rangle + a_3 \langle 0, 0, 1 \rangle \end{aligned}$$

Notice how these are just the three standard basis vectors for \mathbb{R}^3

$$\mathbf{a} = a_1 \hat{\mathbf{i}} + a_2 \hat{\mathbf{j}} + a_3 \hat{\mathbf{k}}$$

So we can take any vector and write it in terms of the standard basis vectors.

Properties

If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors (each with the same number of components) and a and b are two numbers, then we have the following properties

$$\begin{array}{ll} \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} & \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \mathbf{u} + \mathbf{0} = \mathbf{u} & 1 \cdot \mathbf{u} = \mathbf{u} \\ a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v} & (a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u} \end{array}$$

1.3 Dot Product

Given two vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ the dot product is,

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 \quad (1.1)$$

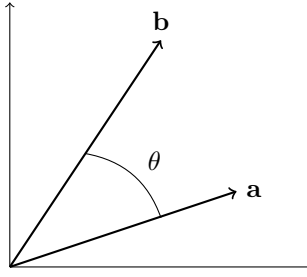
Notice how our answer is just a scalar, not a vector. Here are some properties of the dot product.

Properties

If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors and c is a number then,

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) &= \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} & (c\mathbf{u}) \cdot \mathbf{v} &= \mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v}) \\ \mathbf{u} \cdot \mathbf{v} &= \mathbf{v} \cdot \mathbf{u} & \mathbf{u} \cdot \mathbf{0} &= 0 \\ \mathbf{u} \cdot \mathbf{u} &= \|\mathbf{u}\|^2 & \text{If } \mathbf{u} \cdot \mathbf{u} = 0 &\text{ then } \mathbf{u} = \mathbf{0} \end{aligned}$$

There is also a nice geometric interpretation to the dot product. First suppose that θ is the angle between \mathbf{a} and \mathbf{b} such that $0 \leq \theta \leq \pi$ as shown in the figure below.



We can then have the following theorem

Theorem 1.3.1.

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos(\theta)$$

The formula from this theorem is often used not to compute a dot product but instead to find the angle between two vectors.

The dot product gives us a very nice method for determining if two vectors are perpendicular and it will give another method for determining when two vectors are parallel. Note as well that often we will use the term *orthogonal* in place of perpendicular. If two vectors are orthogonal then we know that the angle between them is 90 degrees. From (1.3.1) this tells us that if two vectors are orthogonal then,

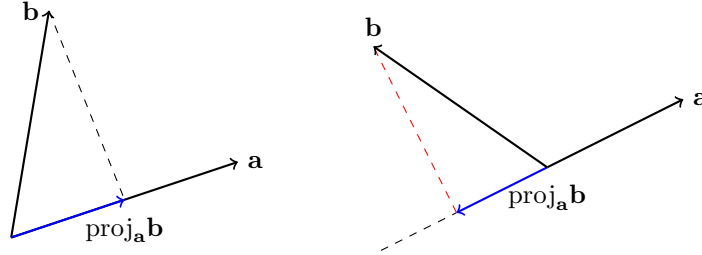
$$\mathbf{a} \cdot \mathbf{b} = 0$$

Likewise, if two vectors are parallel then the angle between them is either 0 degrees (pointing in the same direction) or 180 degrees (pointing in the opposite direction). Once again using (1.3.1) this would mean that one of the following would have to be true

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| (\theta = 0^\circ) \quad \text{or} \quad \mathbf{a} \cdot \mathbf{b} = -\|\mathbf{a}\| \|\mathbf{b}\| (\theta = 180^\circ)$$

Projections

Given two vectors \mathbf{a} and \mathbf{b} we want to determine the projection of \mathbf{b} onto \mathbf{a} . The projection is denoted by $\text{proj}_{\mathbf{a}}\mathbf{b}$. Here are a couple of sketches illustrating the projection



So, to get the projection of \mathbf{b} onto \mathbf{a} we drop straight down from the end of \mathbf{b} until we hit (and form a right angle) with the line that is parallel to \mathbf{a} . The projection is then the vector that is parallel to \mathbf{a} , starts at the same point both of the original vectors started at and ends where the dashed line hits the line parallel to \mathbf{a} . Here is a nice formula for finding the projection of \mathbf{b} onto \mathbf{a} ,

$$\text{proj}_{\mathbf{a}}\mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|^2} \mathbf{a}$$

Note that we also need to be very careful with notation here. The projection of \mathbf{a} onto \mathbf{b} is given by

$$\text{proj}_{\mathbf{b}}\mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2} \mathbf{b}$$

1.4 Cross Product

Given two vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, the cross product is given by the formula,

$$\mathbf{a} \times \mathbf{b} = \langle a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1 \rangle$$

This is not an easy formula to remember! So we use the fact that we can write the cross product as the determinant of a 3x3 matrix, which looks like this,

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

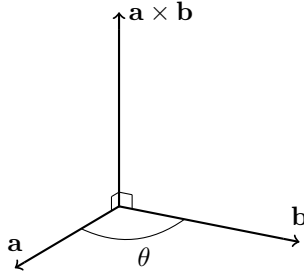
The first row is the standard basis vectors and must appear in the order given above. The second row is the components of our first vector \mathbf{a} and the third row is the components of our second vector \mathbf{b} . We can now use the Method of Cofactors to expand

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \hat{\mathbf{i}} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \hat{\mathbf{j}} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \hat{\mathbf{k}} \\ &= (a_2b_3 - a_3b_2)\hat{\mathbf{i}} - (a_1b_3 - a_3b_1)\hat{\mathbf{j}} + (a_1b_2 - a_2b_1)\hat{\mathbf{k}} \\ &= \langle a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1 \rangle \end{aligned}$$

There is also a geometric interpretation of the cross product. First we will let θ be the angle between the two vectors \mathbf{a} and \mathbf{b} and assume that $0 \leq \theta \leq \pi$, then we have the following fact,

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin(\theta) \quad (1.2)$$

and the following figure,



But how do we know that the cross product pointed in the direction that we're given here? First, the figure implies that the cross product is orthogonal to both of the original vectors. This will always be the case with one exception that we'll get to in a second.

Second, we knew that it pointed in the upward direction (in this case) by the "right hand rule". This says that if we take our right hand, start at \mathbf{a} and rotate our fingers toward \mathbf{b} our thumb will point in the direction of $\mathbf{a} \times \mathbf{b}$. Therefore, if we'd sketched $\mathbf{b} \times \mathbf{a}$ we should have gotten a vector in the downward direction.

Now, let's address the one times where the cross product will not be orthogonal to the original vectors. If the two vectors, \mathbf{a} and \mathbf{b} , are parallel then the angle between them is either 0 or 180 degrees. From (1.3) this implies that

$$\|\mathbf{a} \times \mathbf{b}\| = 0$$

From a fact about the magnitude we saw in the first section we know that this implies

$$\mathbf{a} \times \mathbf{b} = \mathbf{0}$$

In other words, it won't be orthogonal to the original vectors since we have the zero vector. This does give us another test for parallel vectors however.

Fact

If $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ then \mathbf{a} and \mathbf{b} will be parallel vectors.

Let's also formalize up the fact about the cross product being orthogonal to the original vectors.

Fact

Provided $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$ then $\mathbf{a} \times \mathbf{b}$ is orthogonal to both \mathbf{a} and \mathbf{b} . Here are some nice properties about the cross product.

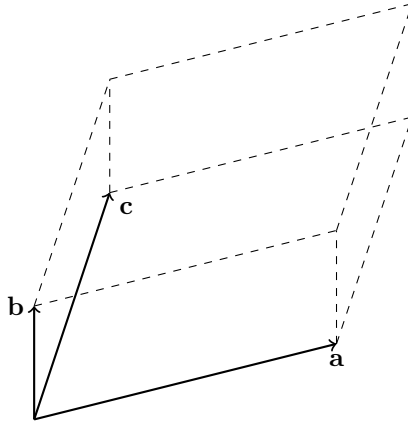
Properties

If \mathbf{u} , \mathbf{v} , and \mathbf{w} are vectors and c is a number then,

$$\begin{aligned}\mathbf{u} \times \mathbf{v} &= -\mathbf{v} \times \mathbf{u} & (c\mathbf{u}) \times \mathbf{v} &= \mathbf{u} \times (c\mathbf{v}) = c(\mathbf{u} \times \mathbf{v}) \\ \mathbf{u} \times (\mathbf{v} + \mathbf{w}) &= \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} & \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}\end{aligned}$$

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

There are a couple of geometric applications to the cross product as well. Suppose we have three vectors \mathbf{a} , \mathbf{b} , and \mathbf{c} and we form the three dimensional figure shown below



The area of the parallelogram (two dimensional front of this object) is given by,

$$\text{Area} = \|\mathbf{a} \times \mathbf{b}\|$$

and the volume of the parallelepiped (the whole three dimensional object) is given by,

$$\text{Volume} = |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$$

Note that the absolute value bars are required since the quantity could be negative and volume isn't negative.

Chapter 2

Three Dimensional Space

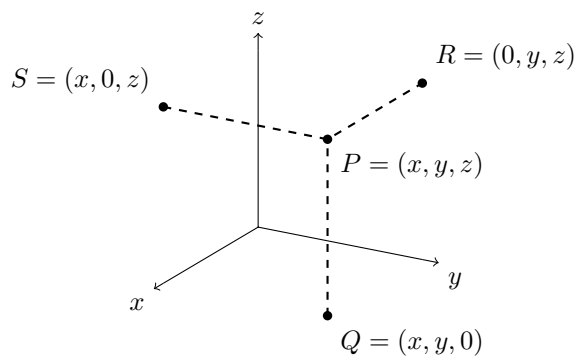
In this chapter we will start to take a more detailed look at three dimensional space (3-D space or \mathbb{R}^3). This is a very important topic in Calculus III since a good portion of Calculus III is done in three (or higher) dimensional space.

We will be looking at the equations of graphs in 3-D space as well as vector valued functions and how we do calculus with them. We will also be taking a look at a couple of new coordinate systems for 3-D space.

2.1 The 3-D Coordinate System

Let's first get some basic notation out of the way. The 3-D coordinate system is often denoted by \mathbb{R}^3 . Likewise the 2-D coordinate system is often denoted by \mathbb{R}^2 and the 1-D coordinate system is denoted by \mathbb{R} . Also, as you might have guessed then a general n dimensional coordinate system is often denoted by \mathbb{R}^n .

Next, let's take a quick look at the basic coordinate system.



This is the standard placement of the axes in this class. It is assumed that only the positive directions are shown by the axes. If we need the negative axes for any reason we will put them in as

needed. Also note the various points on this sketch. The point P is the general point sitting out in 3-D space. If we start at P and drop straight down until we reach a z -coordinate of zero we arrive at the point Q . We say that Q sits in the xy -plane. The xy -plane corresponds to all the points which have a zero z -coordinate. We can also start at P and move in the other two directions as shown to get points in the xz -plane (this is S with a y -coordinate of zero) and the yz -plane (this is R with an x -coordinate of zero).

Collectively, the xy , xz , and yz -planes are sometimes called the coordinate planes. Also, the point Q is often referred to as the projection of P in the xy -plane. Likewise, R is the projection of P in the yz -plane and S is the projection of P in the xz -plane.

Many of the formulas that you are used to working with in \mathbb{R}^2 have natural extensions in \mathbb{R}^3 . For instance the distance between two points in \mathbb{R}^2 is given by,

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

While the distance between any two points in \mathbb{R}^3 is given by,

$$d(P_1, P_2, P_3) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

2.2 Equations of Lines

So, before we get into the equations of lines we first need to briefly look at vector functions. We're going to take a more in depth look at vector functions later. At this point all that we need to worry about is notational issues and how they can be used to give the equation of a curve.

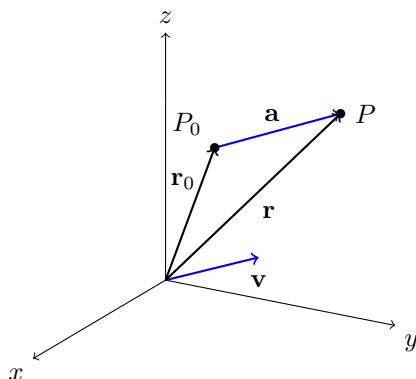
A vector function is a function that takes one or more variables, one in this case, and returns a vector. Note as well that a vector function can be a function of two or more variables. However, in those cases the graph may no longer be a curve in space. We represent vectors with a bold face.

We want to write down the equation of a line in \mathbb{R}^3 and we will need a vector function to do this. To see how we're going to do this let's think about what we need to write down the equation of a line in \mathbb{R}^2 . In two dimensions we need the slope (m) and a point that was on the line in order to write down the equation.

Suppose that we know a point that is on the line, $P_0 = (x_0, y_0, z_0)$, and that $\mathbf{v} = \langle a, b, c \rangle$ is some vector that is parallel to the line. Note, in all likelihood, will not be on the line itself. We only need to be parallel to the line. Finally, let $P = (x, y, z)$ be any point on the line.

Now, since our "slope" is a vector let's also represent the two points on the line as vectors. We'll do this with position vectors. So, let \mathbf{r}_0 and \mathbf{r} be the position vectors for P_0 and P respectively. Also, let's define \mathbf{a} to be the vector with representation $\mathbf{P}_0\mathbf{P}$.

We now have the following sketch with all these points and vectors on it.



Now, we've shown the parallel vector, \mathbf{v} , as a position vector but it doesn't need to be a position vector. It can be anywhere, a position vector, on the line or off the line, it just needs to be parallel to the line. Next, notice that we can write \mathbf{r} as follows,

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{a}$$

Now, notice that the vectors \mathbf{a} and \mathbf{v} are parallel. Therefore there is a number, t , such that

$$\mathbf{a} = t\mathbf{v}$$

Combining these, we now have

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} = \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle$$

This is called the *vector form of the equation of a line*. Notice that $t\mathbf{v}$ will be a vector that lies along the line and tells us how far from the original point that we should we move.

There are several other forms of the equation of a line. To get the first alternate form let's start with the vector form and do a slight rewrite.

$$\begin{aligned}\mathbf{r} &= \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle \\ \langle x, y, z \rangle &= \langle x_0 + ta, y_0 + tb, z_0 + tc \rangle\end{aligned}$$

The only way for two vectors to be equal is for the components to be equal. In other words,

$$\begin{aligned}x &= x_0 + ta \\ y &= y_0 + tb \\ z &= z_0 + tc\end{aligned}$$

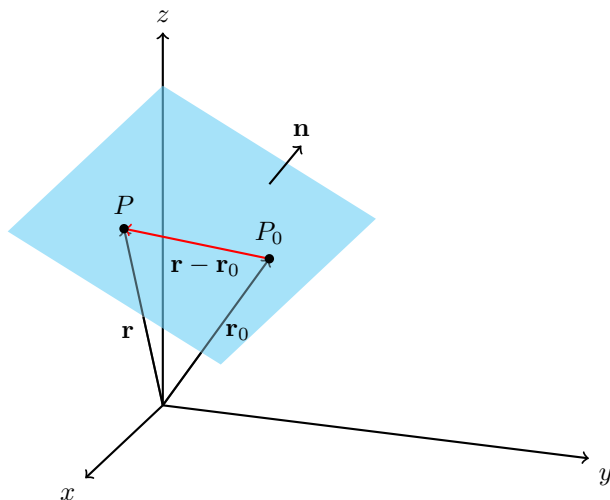
This set of equations are called *parametric form of the equation of a line*. To get a point on the line all we do is pick a t and plug into either form of the line. In the vector form of the line we get a position vector for the point and in the parametric form we get the actual coordinates of the point. If we assume that a , b , and c are all non-zero numbers we can solve each of the equations in the parametric form of the line for t . We can then set all of them equal to each other since t will be the same number in each. Doing this gives the following,

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

This is called the *symmetric equation of the line*.

2.3 Equations of Planes

Let's start assuming that we know a point on the plane, $P_0 = (x_0, y_0, z_0)$. Let's also suppose that we have a vector that is orthogonal (perpendicular) to the plane, $\mathbf{n} = \langle a, b, c \rangle$. This is called the *normal vector*. Now, assume that $P = (x, y, z)$ is any point in the plane. Finally, we'll let \mathbf{r} and \mathbf{r}_0 be the position vectors for P and P_0 respectively. Here's the sketch,



Notice that the vector $\mathbf{r} - \mathbf{r}_0$ lies completely on the plane. Now, because \mathbf{n} is orthogonal to the plane, it's also orthogonal to any vector that lies in the plane. In particular, it's orthogonal to $\mathbf{r} - \mathbf{r}_0$. Recall from the Dot Product that two orthogonal vectors will have a dot product of zero. In other words,

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0 \quad \Rightarrow \quad \mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0$$

This is called the *vector equation of the plane*.

A slightly more useful form of the equation is as follows. Start with the first form of the vector equation and write down a vector for the difference.

$$\begin{aligned} \langle a, b, c \rangle \cdot (\langle x, y, z \rangle - \langle x_0, y_0, z_0 \rangle) &= 0 \\ \langle a, b, c \rangle \cdot \langle x - x_0, y - y_0, z - z_0 \rangle &= 0 \end{aligned}$$

Now compute the dot product to get,

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

This is called the *scalar equation of the plane*. Often this will be written as,

$$ax + by + cz = d$$

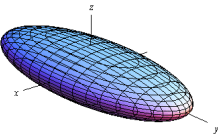
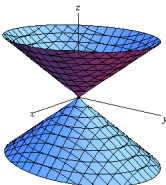
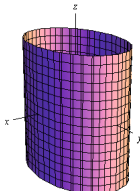
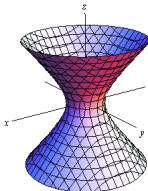
where $d = ax_0 + by_0 + cz_0$

2.4 Quadric Surfaces

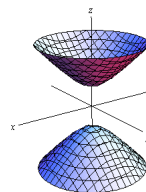
Quadric surfaces are the graphs of any equation that can be put into the general form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0$$

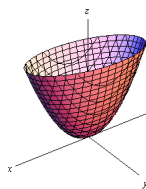
where A, \dots, J are constants. Now, there is no way we can possibly list all of them, but there are some standard equations so here is a table of some of the more common quadric surfaces.

Name	Equation	Graph
Ellipsoid	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$	
Cone	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$	
Cylinder	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	
Hyperboloid of One Sheet	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$	

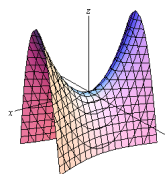
Hyperboloid of Two Sheets $-\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$



Elliptic Paraboloid $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z}{c}$



Hyperbolic Paraboloid $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}$



2.5 Vector Functions

A vector function is a function that takes one or more variables and returns a vector. A vector function of a single variable in \mathbb{R}^2 and \mathbb{R}^3 have the form,

$$\mathbf{r}(t) = \langle f(t), g(t) \rangle \quad \mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$$

where $f(t)$, $g(t)$, and $h(t)$ are called the *component functions*.

The domain of a vector function is the set of all t 's for which all the component functions are defined.

Any vector function can be broken down into a set of parametric equations that represent the same graph. In general, the two dimensional vector function, $\mathbf{r}(t) = \langle f(t), g(t) \rangle$, can be broken down into the parametric equations,

$$x = f(t) \quad y = g(t)$$

Likewise, a three dimensional vector function, $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$, can be broken down into the parametric equations,

$$x = f(t) \quad y = g(t) \quad z = h(t)$$

2.6 Calculus with Vector Functions

Given the vector function $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$, the limit of $\mathbf{r}(t)$ as it approaches a value α , is

$$\begin{aligned}\lim_{t \rightarrow \alpha} \mathbf{r}(t) &= \lim_{t \rightarrow \alpha} \langle f(t), g(t), h(t) \rangle \\ &= \left\langle \lim_{t \rightarrow \alpha} f(t), \lim_{t \rightarrow \alpha} g(t), \lim_{t \rightarrow \alpha} h(t) \right\rangle \\ &= \lim_{t \rightarrow \alpha} f(t) \hat{\mathbf{i}} + \lim_{t \rightarrow \alpha} g(t) \hat{\mathbf{j}} + \lim_{t \rightarrow \alpha} h(t) \hat{\mathbf{k}}\end{aligned}$$

So, all that we do is take the limit of each of the component's functions and leave it as a vector.

Derivatives are very similar,

$$\begin{aligned}\mathbf{r}'(t) &= \langle f'(t), g'(t), h'(t) \rangle \\ &= f'(t) \hat{\mathbf{i}} + g'(t) \hat{\mathbf{j}} + h'(t) \hat{\mathbf{k}}\end{aligned}$$

Most of the basic facts that we know about derivatives still hold. Given two vectors \mathbf{u} , \mathbf{v} , a real valued function $f(t)$, and a real number c , we have the following properties

$$\begin{aligned}\frac{d}{dt}(\mathbf{u} + \mathbf{v}) &= \mathbf{u}' + \mathbf{v}' \\ \frac{d}{dt}(c\mathbf{u}) &= c\mathbf{u}' \\ \frac{d}{dt}(f(t)\mathbf{u}) &= f(t)\mathbf{u}' + f'(t)\mathbf{u} \\ \frac{d}{dt}(\mathbf{u} \cdot \mathbf{v}) &= \mathbf{u} \cdot \mathbf{v}' + \mathbf{u}' \cdot \mathbf{v} \\ \frac{d}{dt}(\mathbf{u} \times \mathbf{v}) &= \mathbf{u} \times \mathbf{v}' + \mathbf{u}' \times \mathbf{v} \\ \frac{d}{dt}(\mathbf{u}(f(t))) &= f'(t)\mathbf{u}'(f(t))\end{aligned}$$

Before getting to integration we need to define what a smooth curve is. A *smooth curve* is any curve for which $\mathbf{r}(t)$ is continuous and $\mathbf{r}'(t) \neq 0$ for any t except possibly at the endpoints. A helix is a smooth curve, for example.

Finally, the integral of a vector function is defined as

$$\begin{aligned}\int \mathbf{r}(t) &= \left\langle \int f(t) dt, \int g(t) dt, \int h(t) dt \right\rangle + \mathbf{c} \\ &= \int f(t) dt \hat{\mathbf{i}} + \int g(t) dt \hat{\mathbf{j}} + \int h(t) dt \hat{\mathbf{k}} + \mathbf{c}\end{aligned}$$

and the following for definite integrals

$$\begin{aligned}\int_{\alpha}^{\beta} \mathbf{r}(t) dt &= \left\langle \int_{\alpha}^{\beta} f(t) dt, \int_{\alpha}^{\beta} g(t) dt, \int_{\alpha}^{\beta} h(t) dt \right\rangle \\ &= \int_{\alpha}^{\beta} f(t) dt \hat{\mathbf{i}} + \int_{\alpha}^{\beta} g(t) dt \hat{\mathbf{j}} + \int_{\alpha}^{\beta} h(t) dt \hat{\mathbf{k}}\end{aligned}$$

2.7 Tangent, Normal, and Binormal Vectors

In the past we've used the fact that the derivative of a function was the slope of the tangent line. With vector functions we get exactly the same result, with one exception.

Given the vector function, $\mathbf{r}(t)$, we call $\mathbf{r}'(t)$ the *tangent vector* provided it exists and provided $\mathbf{r}(t) \neq \mathbf{0}$. The tangent line to $\mathbf{r}(t)$ at P is then the line that passes through the point P and is parallel to the tangent vector, $\mathbf{r}'(t)$. Also, provided $\mathbf{r}'(t) \neq \mathbf{0}$, the *unit tangent vector* to the curve is given by,

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

While, the components of the unit tangent vector can be somewhat messy on occasion there are times when we will need to use the unit tangent vector instead of the tangent vector.

The unit normal is orthogonal (or normal, or perpendicular) to the unit tangent vector and hence to the curve as well. The *unit normal vector* is given by,

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}$$

Next is the binormal vector. The binormal vector is defined to be,

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t)$$

Because the binormal vector is defined to be the cross product of the unit tangent and unit normal vector we then know that the binormal vector is orthogonal to both the tangent vector and the normal vector.

2.8 Arc Length with Vector Functions

In this section we'll recast an old formula into terms of vector functions. We want to determine the length of a vector function,

$$\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$$

on the interval $\alpha \leq t \leq \beta$. We actually already know how to do this. Recall that we can write the vector function into the parametric form,

$$x = f(t) \quad y = g(t) \quad z = h(t)$$

Also, recall that with two dimensional parametric curves the arc length is given by,

$$L = \int_{\alpha}^{\beta} \sqrt{(f'(t))^2 + (g'(t))^2} dt$$

There is a natural extension of this to three dimensions. So, the length of the curve $\mathbf{r}(t)$ on the interval $\alpha \leq t \leq \beta$ is,

$$L = \int_{\alpha}^{\beta} \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2} dt$$

There is a nice simplification that we can make for this. Notice that the integrand is nothing more than the magnitude of the tangent vector,

$$\|\mathbf{r}'(t)\| = \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2}$$

Therefore, the arc length can be written as,

$$L = \int_{\alpha}^{\beta} \|\mathbf{r}'(t)\| dt$$

2.9 Curvature

Curvature is the measure of how fast a curve is changing direction at a given point. The formal definition of curvature is,

$$\kappa = \left| \frac{d\mathbf{T}}{ds} \right|$$

where \mathbf{T} is the unit tangent vector and s is the arc length. In general the formal definition of the curvature is not easy to use so there are two alternate formulas that we can use. Here they are.

$$\kappa = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|} \quad \kappa = \frac{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}{\|\mathbf{r}'(t)\|^3}$$

These may not be particularly easy to deal with either, but at least we don't need to reparameterize the unit tangent.

There is a special case that we can look at. Suppose we have a curve given by $y = f(x)$ and we want to find its curvature. We can write it as a standard vector function in the form

$$\mathbf{r}(t) = x \hat{\mathbf{i}} + f(x) \hat{\mathbf{j}}$$

If we then use the second formula for the curvature we will arrive at the following formula for the curvature.

$$\kappa = \frac{|f'(x)|}{(1 + (f'(x))^2)^{3/2}}$$

2.10 Velocity and Acceleration

From Calculus I we know that given the position function of an object that the velocity of the object is the first derivative of the position function and the acceleration of the object is the second derivative of the position function. So given an object's position vector $\mathbf{r}(t)$ the object's velocity and acceleration are given by

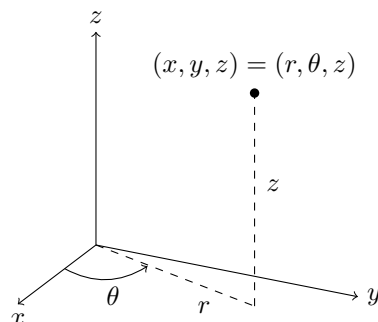
$$\mathbf{v}(t) = \mathbf{r}'(t) \quad \mathbf{a}(t) = \mathbf{r}''(t)$$

Notice that the velocity and acceleration are going to be vector functions as well.

2.11 Cylindrical Coordinates

The last two sections of this chapter will be looking at some alternate coordinate systems for three dimensional space. As with two dimensional space the standard (x, y, z) coordinate system is called the Cartesian coordinate system.

Cylindrical coordinates are fairly simple as it is nothing more than extending polar coordinates into \mathbb{R}^3 . Not only is it an extension of polar coordinates, but we extend it into the third dimension just as we extend Cartesian coordinates into the third dimension. All that we do is add a z on as the third coordinate. The r and θ are the same as with polar coordinates. Here is a sketch.



The conversions for x and y are the same conversions that we used back when we were looking at polar coordinates. So, if we have a point in cylindrical coordinates the Cartesian coordinates can be found by using the following conversions

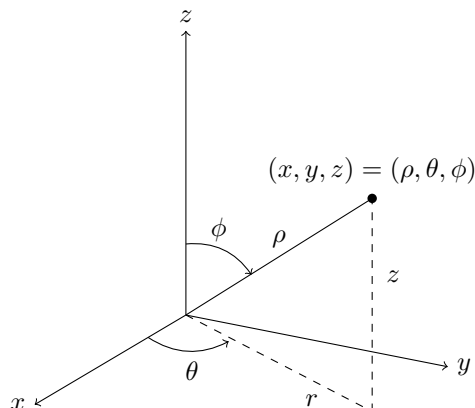
$$\begin{aligned} x &= r \cos(\theta) \\ y &= r \sin(\theta) \\ z &= z \end{aligned}$$

The third equation is just an acknowledgement that the z -coordinate of a point in Cartesian and polar coordinates is the same. Likewise, if we have a point in Cartesian coordinates the cylindrical coordinates can be found by using the following conversions

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \quad \text{or} \quad r^2 = x^2 + y^2 \\ \theta &= \tan^{-1} \left(\frac{y}{x} \right) \\ z &= z \end{aligned}$$

2.12 Spherical Coordinates

Let's start with a sketch



Spherical coordinates consist of the following three quantities.

First there is ρ . This is the distance from the origin to the point, we require $\rho \geq 0$.

Next there is θ . This is the same angle from polar/cylindrical coordinates. It is the angle between the positive x -axis and the line denoted by r (which is also the same r in polar/cylindrical coordinates).

Finally there is ϕ . This is the angle between the positive z -axis and the line from the origin to the point. We will require $0 \leq \phi \leq \pi$.

So, given a point in spherical coordinates the cylindrical coordinates of the point will be,

$$r = \rho \sin(\phi)$$

$$\theta = \theta$$

$$z = \rho \cos(\phi)$$

Note as well that,

$$\begin{aligned} r^2 + z^2 &= \rho^2 \cos^2(\phi) + \rho^2 \sin^2(\phi) \\ &= \rho^2 (\cos^2(\phi) + \sin^2(\phi)) \\ &= \rho^2 \end{aligned}$$

Chapter 3

Partial Derivatives

In Calculus III we will extend our knowledge of calculus into functions of two or more variables. Despite the fact that this chapter is about derivatives we will start out the chapter with a section on limits of functions of more than one variable. In the remainder of this chapter we will be looking at differentiating functions of more than one variable. As we will see, while there are differences with derivatives of functions of one variable, if you can do derivatives of functions of one variable you shouldn't have any problems differentiating functions of more than one variable.

3.1 Limits

In this section we will take a look at limits involving functions of more than one variable. In fact, we will concentrate mostly on limits of functions of two variables, but the ideas can be extended out to functions with more than two variables. Before getting into this let's briefly recall how limits of functions of one variable work. We say that,

$$\lim_{x \rightarrow \alpha} f(x) = L$$

provided that

$$\lim_{x \rightarrow \alpha^+} f(x) = \lim_{x \rightarrow \alpha^-} f(x) = L$$

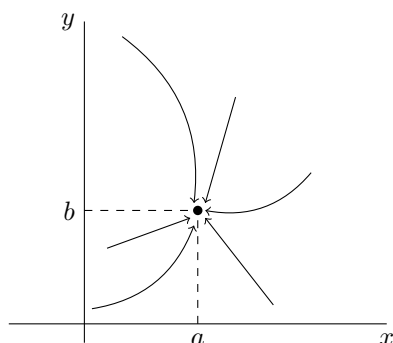
In other words, we will have $\lim_{x \rightarrow \alpha} f(x) = L$ provided $f(x)$ approaches L as we move in towards $x = \alpha$ (without letting $x = \alpha$) from both sides. Now, notice that in this case there are only two paths that we can take as we move in towards $x = \alpha$. We can either move in from the left or we can move in from the right. Then in order for the limit of a function of one variable to exist the function must be approaching the same value as we take each of these paths in towards $x = \alpha$.

With functions of two variables we will have to do something similar, except this time there is (potentially) going to be a lot more work involved. Let's first address the notation and get a feel for just what we're going to be asking for in these kinds of limits. We will be asking to take the limit of the function $f(x, y)$ as x approaches a and as y approaches b . Here is the standard notation,

$$\lim_{(x, y) \rightarrow (a, b)} f(x, y)$$

In taking a limit of a function of two variables we are really asking what the value of $f(x, y)$ is doing as we move the point (x, y) in closer and closer to the point (a, b) without actually letting it be (a, b) .

Just like with limits of functions of one variable, in order for this limit to exist, the function must be approaching the same value regardless of the path that we take as we move in towards (a, b) . The problem that we are immediately faced with is that there are literally an infinite number of paths that we can take as we move in towards (a, b) . Here are a few examples of paths that we could take.



In other words, to show that a limit exists we would technically need to check an infinite number of paths and verify that the function is approaching the same value regardless of the path we are using to approach the point. The way we approach this is to use one of the main ideas from Calculus I limits.

Definition

A function $f(x, y)$ is continuous at the point (a, b) if,

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b)$$

From a graphical standpoint this definition means the same thing as it did when we first saw continuity in Calculus I. A function will be continuous at a point if the graph doesn't have any holes or breaks at that point.

Example

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y^2}{x^4 + 3y^4}$$

Now, in this case the function is not continuous at the point in question and so we can't just plug in the point. So, since the function is not continuous at the point there is at least a chance that the limit doesn't exist. If we could find two different paths to approach the point that gave different values for the limit then we would know that the limit didn't exist. Two of the more common paths to check are the x and y -axis so let's try those.

So, let's see what happens along the x -axis. If we are going to approach $(0, 0)$ along the x -axis we can take advantage of the fact that along the x -axis we know that $y = 0$. This means that, along the x -axis, we will plug in $y = 0$ into the function and then take the limit as x approaches

zero.

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y^2}{x^4 + 3y^4} = \lim_{(x,0) \rightarrow (0,0)} \frac{x^2(0)^2}{x^4 + 3(0)^4} = \lim_{(x,0) \rightarrow (0,0)} 0 = 0$$

So, along the x -axis the function will approach zero as we move in towards the origin. Now, let's try the y -axis. Along this axis we have $x = 0$ and so the limit becomes,

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y^2}{x^4 + 3y^4} = \lim_{(0,y) \rightarrow (0,0)} \frac{(0)^2 y^2}{(0)^4 + 3y^4} = \lim_{(0,y) \rightarrow (0,0)} 0 = 0$$

So, the same limit along two paths. Don't misread this. This does NOT say that the limit exists and has a value of zero. This only means that the limit happens to have the same value along two paths. Let's take a look at a third fairly common path to take a look at. In this case we'll move in towards the origin along the path $y = mx$, where m is just a number. This is what we meant previously about relating x and y through a function. To do this we will replace all the y 's with mx 's and then let mx approach zero. Let's take a look at this limit.

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y^2}{x^4 + 3y^4} = \lim_{(x,mx) \rightarrow (0,0)} \frac{x^2 (mx)^2}{x^4 + 3(mx)^4} = \lim_{(x,mx) \rightarrow (0,0)} \frac{m^2 x^4}{x^4 + 3m^4 x^4} = \lim_{(x,mx) \rightarrow (0,0)} \frac{m^2}{1 + 3m^4}$$

Since we're just left with m 's and we know that m can be any number, this means that that we will get a different value with each m we pick. So since the value of the limit is different from the two previous paths we chose, the limit does not exist.

3.2 Partial Derivatives

Recall that given a function of one variable, $f(x)$, the derivative, $f'(x)$, represents the rate of change of the function as x changes. This is an important interpretation of derivatives and we are not going to want to lose it with functions of more than one variable. The problem with functions of more than one variable is that there is more than one variable. In other words, what do we do if we only want one of the variables to change, or if we want more than one of them to change? In fact, if we're going to allow more than one of the variables to change there are then going to be an infinite amount of ways for them to change. For instance, one variable could be changing faster than the other variable(s) in the function. Notice as well that it will be completely possible for the function to be changing differently depending on how we allow one or more of the variables to change.

Let's start with the function $f(x, y) = 2x^2 y^3$ and let's determine the rate at which the function is changing at a point, (a, b) , if we hold y fixed and allow x to vary and if we hold x fixed and allow y to vary.

We'll start by looking at the case of holding y fixed and allowing x to vary. Since we are interested in the rate of change of the function at (a, b) and are holding y fixed this means that we are going to always have $y = b$ (if we didn't have this then eventually y would have to change in order to get to the point). Doing this will give us a function involving only x s and we can define a new function as follows,

$$g(x) = f(x, b) = 2x^2 b^3$$

Now, this is a function of a single variable and at this point all that we are asking is to determine the rate of change of $g(x)$ at $x = a$. In other words, we want to compute $g'(a)$ and since this is

a function of a single variable we already know how to do that. Here is the rate of change of the function at (a, b) if we hold y fixed and allow x to vary.

$$g'(a) = 4ab^3$$

We will call $g'(a)$ the *partial derivative* of $f(x, y)$ with respect to x at (a, b) and we will denote it in the following way,

$$f_x(a, b) = 4ab^3$$

Now, let's do it the other way. We will now hold x fixed and allow y to vary. We can do this in a similar way. Since we are holding x fixed it must be fixed at $x = a$ and so we can define a new function of y and then differentiate this as we've always done with functions of one variable.

$$h(y) = f(a, y) = 2a^2y^3 \quad \Rightarrow \quad h'(y) = 6a^2y^2$$

In this case we call $h'(y)$ the *partial derivative* of $f(x, y)$ with respect to y at (a, b) and we denote it as follows,

$$f_y(a, b) = 6a^2b^2$$

Now, as this quick example has shown taking derivatives of functions of more than one variable is done in pretty much the same manner as taking derivatives of a single variable. To compute $f_x(x, y)$ all we need to do is treat all the y 's as constants (or numbers) and then differentiate the x 's as we've always done. Likewise, to compute $f_y(x, y)$ we will treat all the x 's as constants and then differentiate the y 's as we are used to doing.

Since we can think of the two partial derivatives above as derivatives of single variable functions it shouldn't be too surprising that the definition of each is very similar to the definition of the derivative for single variable functions. Here are the formal definitions of the two partial derivatives we looked at above.

$$f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h} \quad f_y(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h}$$

Now let's take a quick look at some of the possible alternate notations for partial derivatives. Given the function $z = f(x, y)$ the following are all equivalent notations,

$$\begin{aligned} f_x(x, y) = f_x &= \frac{\partial f}{\partial x} = z_x = \frac{\partial z}{\partial x} = D_x f \\ f_y(x, y) = f_y &= \frac{\partial f}{\partial y} = z_y = \frac{\partial z}{\partial y} = D_y f \end{aligned}$$

3.3 Higher Order Partial Derivatives

Just as we had higher order derivatives with functions of one variable we will also have higher order derivatives of functions of more than one variable. However, this time we will have more options since we do have more than one variable.

Consider the case of a function of two variables, $f(x, y)$ since both of the first order partial derivatives are also functions of x and y we could in turn differentiate each with respect to x or y . This means that for the case of a function of two variables there will be a total of four possible second

order derivatives. Here they are and the notations that we'll use to denote them.

$$\begin{aligned}(f_x)_x &= f_{xx} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} \\(f_x)_y &= f_{xy} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x} \\(f_y)_x &= f_{yx} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} \\(f_y)_y &= f_{yy} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2}\end{aligned}$$

The second and third second order partial derivatives are often called *mixed partial derivatives* since we are taking derivatives with respect to more than one variable. Note as well that the order that we take the derivatives in is given by the notation for each these. If we are using the subscripting notation, e.g. f_{xy} , then we will differentiate from left to right. In other words, in this case, we will differentiate first with respect to x and then with respect to y . With the fractional notation, e.g. $\partial^2 f / \partial y \partial x$, it is the opposite. In these cases we differentiate moving along the denominator from right to left. So, again, in this case we differentiate with respect to x first and then y .

3.4 Differentials

Just like we had differentials for functions of one variable we also have them for functions of more than one variable. Also, as we've already seen in previous sections, when we move up to more than one variable things work pretty much the same, but there are some small differences. Given the function $z = f(x, y)$ the differential dz or df is given by,

$$dz = f_x dx + f_y dy \quad \text{or} \quad df = f_x dx + f_y dy$$

There is a natural extension to functions of three or more variables. For instance, given the function $w = g(x, y, z)$ the differential is given by,

$$dw = g_x dx + g_y dy + g_z dz$$

3.5 Chain Rule

Recall the chain rule for a single variable function $F = f(g(x))$,

$$F' = f'(g(x))g'(x)$$

There is an alternate notation however that while probably not used much in Calculus I is more convenient at this point because it will match up with the notation that we are going to be using in this section. If $y = f(x)$ and $x = g(t)$ then

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

Notice that the derivative dy/dt really does make sense here since if we were to plug in for x then y really would be a function of t . One way to remember this form of the chain rule is to note that

if we think of the two derivatives on the right side as fractions the dx 's will cancel to get the same derivative on both sides.

We will be looking at two distinct cases prior to generalizing the whole idea out. Let's look at the first case

Case I. $z = f(x, y)$, $x = g(t)$, $y = h(t)$. Compute dz/dt .

This case is analogous to the standard chain rule from Calculus I that we looked at above. In this case we are going to compute an ordinary derivative since z really would be a function of t only if we were to substitute in for x and y . The chain rule for this case is,

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

So, basically what we're doing here is differentiating f with respect to each variable in it and then multiplying each of these by the derivative of that variable with respect to t . The final step is to then add all this up.

Case II. $z = f(x, y)$, $x = g(s, t)$, $y = h(s, t)$. Compute $\partial z/\partial s$ and $\partial z/\partial t$.

In this case if we were to substitute in for x and y we would get that z is a function of s and t and so it makes sense that we would be computing partial derivatives here and that there would be two of them. Here is the chain rule for both of these cases.

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial s} \quad \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t}$$

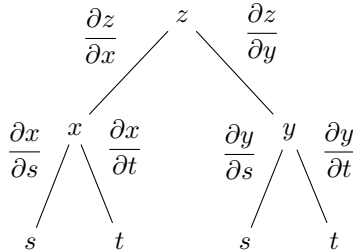
Now for the general version of the chain rule. Suppose that z is a function of n variables, x_1, x_2, \dots, x_n , and that each of these variables are in turn functions of m variables, t_1, t_2, \dots, t_m . Then for any variable t_i , $i = 1, 2, \dots, m$ we have the following,

$$\frac{\partial z}{\partial t_i} = \frac{\partial z}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial z}{\partial x_2} \frac{\partial x_2}{\partial t_i} + \dots + \frac{\partial z}{\partial x_n} \frac{\partial x_n}{\partial t_i}$$

Which is a lot to remember! There is actually an easier way to construct all the chain rules that we've discussed in the section. We can build up a tree diagram that will give us the chain rule for any situation. To see how these work let's go back and take a look at the chain rule for $\partial z/\partial s$ given that $z = f(x, y)$, $x = g(s, t)$, $y = h(s, t)$. We already know what this is, but it may help to illustrate the tree diagram if we already know the answer. For reference here is the chain rule for this case,

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial s}$$

Here is the tree diagram for this case.



We start at the top with the function itself and the branch out from that point. The first set of branches is for the variables in the function. From each of these endpoints we put down a further set of branches that gives the variables that both x and y are a function of. We connect each letter with a line and each line represents a partial derivative as shown. Note that the letter in the numerator of the partial derivative is the upper “node” of the tree and the letter in the denominator of the partial derivative is the lower “node” of the tree. Note that we don’t usually put the derivatives in the tree. They are always an assumed part of the tree.

3.6 Directional Derivatives

To this point we’ve only looked at the two partial derivatives $f_x(x, y)$ and $f_y(x, y)$. Recall that these derivatives represent the rate of change of f as we vary x (holding y fixed) and as we vary y (holding x fixed) respectively. We now need to discuss how to find the rate of change of f if we allow both x and y to change simultaneously. The problem here is that there are many ways to allow both x and y to change. For instance one could be changing faster than the other and then there is also the issue of whether or not each is increasing or decreasing.

The rate of change of $f(x, y)$ in the direction of the unit vector $\mathbf{u} = \langle a, b \rangle$ is called the *directional derivative* and is denoted by $D_{\mathbf{u}}f(x, y)$. The definition of the directional derivative is

$$D_{\mathbf{u}}f(x, y) = \lim_{h \rightarrow 0} \frac{f(x + ah, y + bh) - f(x, y)}{h}$$

So, the definition of the directional derivative is very similar to the definition of partial derivatives. However, in practice this can be a very difficult limit to compute so we need an easier way of taking directional derivatives. It’s actually fairly simple to derive an equivalent formula for taking directional derivatives. To see how we can do this let’s define a new function of a single variable,

$$g(z) = f(x_0 + az, y_0 + bz)$$

where x_0, y_0, a, b are fixed numbers. Then by the definition of the derivative for functions of a single variable we have,

$$g'(z) = \lim_{h \rightarrow 0} \frac{g(z + h) - g(z)}{h}$$

and the derivative at $z = 0$ is given by,

$$g'(0) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h}$$

If we now substitute in for $g(z)$ we get

$$g'(0) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = \lim_{h \rightarrow 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = D_{\mathbf{u}}f(x_0, y_0)$$

So it looks like we have the following relationship

$$g'(0) = D_{\mathbf{u}}f(x_0, y_0) \tag{3.1}$$

Now let’s look at this from another perspective. Let’s rewrite $g(z)$ as follows,

$$g(z) = f(x, y) \text{ where } x = x_0 + az, y = y_0 + bz$$

We can now use the chain rule from the previous section to compute,

$$g'(z) = \frac{dg}{dz} = \frac{\partial f}{\partial x} \frac{dx}{dz} + \frac{\partial f}{\partial y} \frac{dy}{dz} = f_x(x, y)a + f_y(x, y)b$$

So, from the chain rule we get the following relationship.

$$g'(z) = f_x(x, y)a + f_y(x, y)b \quad (3.2)$$

If we now take $z = 0$ we will get that $x = x_0$ and $y = y_0$ (from how we defined x and y above) and plug these into (3.2) we get,

$$g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b \quad (3.3)$$

Now, simply equate (3.1) and (3.3) to get that,

$$D_{\mathbf{u}}f(x_0, y_0) = g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$$

If we now go back to allowing x and y to be any number we get the following formula for computing directional derivatives.

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b$$

Which is much simpler than the limit definition. Also note that this definition assumed that we were working with functions of two variables. There are similar formulas that can be derived by the same type of argument for functions with more than two variables. For instance, the directional derivative of $f(x, y, z)$ in the direction of the unit vector $\mathbf{u} = \langle a, b, c \rangle$ is given by,

$$D_{\mathbf{u}}f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c$$

There is another form of the formula that we used to get the directional derivative that is a little nicer and somewhat more compact. It is also a much more general formula that will encompass both of the formulas above. Let's start with the second one and notice that we can write it as follows,

$$\begin{aligned} D_{\mathbf{u}}f(x, y, z) &= f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c \\ &= \langle f_x, f_y, f_z \rangle \cdot \langle a, b, c \rangle \end{aligned}$$

Now let's give a name and notation to the first vector in the dot product since this vector will show up fairly regularly throughout this course (and in other courses). The *gradient* of f or *gradient vector* of f is defined to be,

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

Or, if we want to use the standard basis vectors the gradient is,

$$\nabla f = f_x \hat{\mathbf{i}} + f_y \hat{\mathbf{j}} + f_z \hat{\mathbf{k}}$$

With the definition of the gradient we can now say that the directional derivative is given by,

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}$$

Chapter 4

Applications of Partial Derivatives

In this section we will take a look at a couple of applications of partial derivatives. Most of the applications will be extensions to applications to ordinary derivatives that we saw back in Calculus I. For instance, we will be looking at finding the absolute and relative extrema of a function and we will also be looking at optimization. Both of these subjects were major applications back in Calculus I. They will, however, be a little more work here because we now have more than one variable.

4.1 Tangent Planes and Linear Approximations

The graph of a function $z = f(x, y)$ is a surface in \mathbb{R}^3 and we can now start thinking of a plane that is “tangent” to the surface at a point. Let’s start out with a point (x_0, y_0) and let’s let C_1 represent the trace of $f(x, y)$ for the plane $y = y_0$ (i.e. allowing x to vary with y held fixed) and we’ll let C_2 represent the trace of $F(x, y)$ for the plane $x = x_0$ (i.e. allowing y to vary with x held fixed). Now, we know that $f_x(x_0, y_0)$ is the slope of the tangent line to the trace C_1 and $f_y(x_0, y_0)$ is the slope of the tangent line to the trace C_2 . So, let L_1 be the tangent line to the trace C_1 and let L_2 be the tangent line to the trace C_2 .

The tangent plane will then be the plane that contains the two lines L_1 and L_2 . Geometrically this plane will serve the same purpose that a tangent line did in Calculus I. A tangent line to a curve was a line that just touched the curve at that point and was “parallel” to the curve at the point in question. Well tangent planes to a surface are planes that just touch the surface at the point and are “parallel” to the surface at the point. Note that this gives us a point that is on the plane. Since the tangent plane and the surface touch at (x_0, y_0) the following point will be on both the surface and the plane.

$$(x_0, y_0, z_0) = (x_0, y_0, f(x_0, y_0))$$

What we need to do now is determine the equation of the tangent plane. We know that the general equation of a plane is given by,

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

where (x_0, y_0, z_0) is a point that is on the plane, which we know already. Let's rewrite this a little. We'll move the x terms and y terms to the other side and divide both sides by c . Doing this gives,

$$z - z_0 = -\frac{a}{c}(x - x_0) - \frac{b}{c}(y - y_0)$$

Now, let's rename the constants to simplify up the notation a little. Let's rename them as follows,

$$A = -\frac{a}{c} \quad B = -\frac{b}{c}$$

With this renaming the equation of the tangent plane becomes,

$$z - z_0 = A(x - x_0) + B(y - y_0)$$

and we need to determine values for A and B .

Let's first think about what happens if we hold y fixed, i.e. if we assume that $y = y_0$. In this case the equation of the tangent plane becomes,

$$z - z_0 = A(x - x_0)$$

This is the equation of a line and this line must be tangent to the surface at (x_0, y_0) (since it's part of the tangent plane). In addition, this line assumes that $y = y_0$ (i.e. fixed) and A is the slope of this line. But if we think about it this is exactly what the tangent to C_1 is, a line tangent to the surface at (x_0, y_0) assuming that $y = y_0$. In other words,

$$z - z_0 = A(x - x_0)$$

is the equation for L_1 and we know that the slope of L_1 is given by $f_x(x_0, y_0)$. Therefore we have the following,

$$A = f_x(x_0, y_0)$$

If we hold x fixed at $x = x_0$ the equation of the tangent plane becomes,

$$z - z_0 = B(y - y_0)$$

However, by a similar argument to the one above we can see that this is nothing more than the equation for L_2 and that its slope is B or $f_y(x_0, y_0)$. So,

$$B = f_y(x_0, y_0)$$

The equation of the tangent plane to the surface given by $z = f(x, y)$ at (x_0, y_0) is then,

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

One nice use of tangent planes is they give us a way to approximate a surface near a point. As long as we are near to the point (x_0, y_0) then the tangent plane should nearly approximate the function at that point. Because of this we define the linear approximation to be,

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

and as long as we are "near" (x_0, y_0) then we should have that,

$$f(x, y) \approx L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

4.2 Gradient Vector, Tangent Planes and Normal Lines

Let's first recall the equation of a plane that contains the point (x_0, y_0, z_0) with normal vector $\mathbf{n} = \langle a, b, c \rangle$ is given by,

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

When we introduced the gradient vector we gave the following fact.

Fact

The gradient vector $\nabla f(x_0, y_0)$ is orthogonal to the level curve $f(x, y) = k$ at the point (x_0, y_0) . Likewise, the gradient vector $\nabla f(x_0, y_0, z_0)$ is orthogonal to the level surface $f(x, y, z) = k$ at the point (x_0, y_0, z_0) .

Also recall that the gradient vector is,

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

So, the tangent plane to the surface given by $f(x, y, z) = k$ at (x_0, y_0, z_0) has the equation,

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0$$

This is a much more general form of the equation of a tangent plane than the one that we derived in the previous section.

Note however, that we can also get the equation from the previous section using this more general formula. To see this let's start with the equation $z = f(x, y)$ and we want to find the tangent plane to the surface given by $z = f(x, y)$ at the point (x_0, y_0, z_0) where $z_0 = f(x_0, y_0)$. In order to use the formula above we need to have all the variables on one side. This is easy enough to do. All we need to do is subtract a z from both sides to get,

$$f(x, y) - z = 0$$

Now, if we define a new function

$$F(x, y, z) = f(x, y) - z$$

we can see that the surface given by $z = f(x, y)$ is identical to the surface given by $F(x, y, z) = 0$ and this new equivalent equation is in the correct form for the equation of the tangent plane that we derived in this section. So, the first thing that we need to do is find the gradient vector for F .

$$\nabla F = \langle F_x, F_y, F_z \rangle = \langle f_x, f_y, -1 \rangle$$

The equation of the tangent plane is then,

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0$$

Or, upon solving for z , we get,

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

which is identical to the equation that we derived in the previous section. We can get another nice piece of information out of the gradient vector as well. We might on occasion want a line that is orthogonal to a surface at a point, sometimes called the *normal line*. This is easy enough to get if we recall that the equation of a line only requires that we have a point and a parallel vector. Since we want a line that is at the point (x_0, y_0, z_0) we know that this point must also be on the line and we know that $\nabla f(x_0, y_0, z_0)$ is a vector that is normal to the surface and hence will be parallel to the line. Therefore the equation of the normal line is,

$$\mathbf{r}(t) = \langle x_0, y_0, z_0 \rangle + t \nabla f(x_0, y_0, z_0)$$

4.3 Relative Minimums and Maximums

We are going to be looking at identifying relative minimums and relative maximums. Recall as well that we will often use the word *extrema* to refer to both minimums and maximums.

Definition

1. A function $f(x, y)$ has a *relative minimum* at the point (a, b) if $f(x, y) \geq f(a, b)$ for all points (x, y) in some region around (a, b) .
2. A function $f(x, y)$ has a *relative maximum* at the point (a, b) if $f(x, y) \leq f(a, b)$ for all points (x, y) in some region around (a, b) .

Next we need to extend the idea of critical points up to functions of two variables. Recall that a critical point of the function $f(x)$ was a number $x = c$ so that either $f'(c) = 0$ or $f'(c)$ doesn't exist. We have a similar definition for critical points of functions of two variables.

Definition

The point (a, b) is a *critical point* of $f(x, y)$ provided one of the following is true,

1. $\nabla f(a, b) = \mathbf{0}$ (this is equivalent to saying that $f_x(a, b) = 0$ and $f_y(a, b) = 0$).
2. $f_x(a, b)$ and/or $f_y(a, b)$ does not exist.

To see the equivalence in the first part let's start off with $\nabla f = \mathbf{0}$ and put in the definition of each part.

$$\begin{aligned}\nabla f(a, b) &= \mathbf{0} \\ \langle f_x(a, b), f_y(a, b) \rangle &= \langle 0, 0 \rangle\end{aligned}$$

The only way that these two vectors can be equal is to have $f_x(a, b) = 0$ and $f_y(a, b) = 0$. In fact, we will use this definition of the critical point more than the gradient definition since it will be easier to find the critical points if we start with the partial derivative definition. We now have the following fact that, at least partially, relates critical points to relative extrema.

Fact

If the point (a, b) is a relative extrema of the function $f(x, y)$ and the first order derivatives of $f(x, y)$ exist at (a, b) then (a, b) is also a critical point of $f(x, y)$ and in fact we'll have $\nabla f(a, b) = \mathbf{0}$.

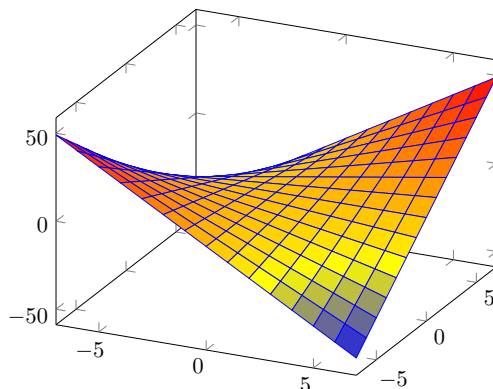
Note that this does NOT say that all critical points are relative extrema. It only says that relative extrema will be critical points of the function. To see this let's consider the function

$$f(x, y) = xy$$

The two first order partial derivatives are

$$f_x = y \quad f_y = x$$

The only point that will make both of these derivatives zero at the same time is $(0, 0)$ and so $(0, 0)$ is a critical point for the function. Here is a graph of the function.



Note that the axes are not in the standard orientation here so that we can see more clearly what is happening at the origin, i.e. at $(0,0)$. If we start at the origin and move into either of the quadrants where both x and y are the same sign the function increases. However, if we start at the origin and move into either of the quadrants where x and y have the opposite sign then the function decreases. In other words, no matter what region you take about the origin there will be points larger than $f(0,0) = 0$ and points smaller than $f(0,0) = 0$. Therefore, there is no way that $(0,0)$ can be a relative extrema. Critical points that exhibit this kind of behavior are called *saddle points*.

So, once we have all the critical points in hand all we will need to do is test these points to see if they are relative extrema or not. To determine if a critical point is a relative extrema (and in fact to determine if it is a minimum or a maximum) we can use the following fact.

Fact

Suppose that (a,b) is a critical point of $f(x,y)$ and that the second order partial derivatives are continuous in some region that contains (a,b) . Next define,

$$D(a,b) = f_{xx}(a,b)f_{yy}(a,b) - [f_{xy}(a,b)]^2$$

We then have the following classifications of the critical point

1. If $D > 0$ and $f_{xx}(a,b) > 0$ then there is a relative minimum at (a,b) .
2. If $D > 0$ and $f_{xx}(a,b) < 0$ then there is a relative maximum at (a,b) .
3. If $D < 0$ then (a,b) is a saddle point.
4. If $D = 0$ then the point (a,b) may be a relative minimum, relative maximum or a saddle point.

4.4 Absolute Minimums and Maximums

In this section we want to optimize a function, that is identify the absolute minimum and/or the absolute maximum of the function, on a given region in \mathbb{R}^2 . Note that when we say we are going to be working on a region in \mathbb{R}^2 we mean that we're going to be looking at some region in the xy -plane. In order to optimize a function in a region we are going to need to get a couple of definitions out of the way and a fact. Let's first get the definitions out of the way.

Definitions

1. A region in \mathbb{R}^2 is called *closed* if it includes its boundary. A region is called *open* if it doesn't include any of its boundary points.
2. A region in \mathbb{R}^2 is called *bounded* if it can be completely contained in a disk. In other words, a region will be bounded if it is finite.

Let's think a little more about the definition of closed. We said a region is closed if it includes its boundary. Just what does this mean? Let's think of a rectangle. Below are two definitions of a rectangle, one is closed and the other is open.

Open	Closed
$-5 < x < 3$	$-5 \leq x \leq 3$
$1 < y < 6$	$1 \leq y \leq 6$

In this first case we don't allow the ranges to include the endpoints (i.e. we aren't including the edges of the rectangle) and so we aren't allowing the region to include any points on the edge of the rectangle. In other words, we aren't allowing the region to include its boundary and so it's open. In the second case we are allowing the region to contain points on the edges and so will contain its entire boundary and hence will be closed. This is an important idea because of the following fact.

Extreme Value Theorem

If $f(x, y)$ is continuous in some closed, bounded set D in \mathbb{R}^2 then there are points in D , (x_1, y_1) and (x_2, y_2) so that $f(x_1, y_1)$ is the absolute maximum and $f(x_2, y_2)$ is the absolute minimum of the function in D .

Note that this theorem does NOT tell us where the absolute minimum or absolute maximum will occur. It only tells us that they will exist. Note as well that the absolute minimum and/or absolute maximum may occur in the interior of the region or it may occur on the boundary of the region.

The basic process for finding absolute maximums is pretty much identical to the process that we used in Calculus I when we looked at finding absolute extrema of functions of single variables. There will however, be some procedural changes to account for the fact that we now are dealing with functions of two variables. Here is the process.

Finding Absolute Extrema

1. Find all the critical points of the function that lie in the region D and determine the function value at each of these points.
2. Find all extrema of the function on the boundary. This usually involves the Calculus I approach for this work.
3. The largest and smallest values found in the first two steps are the absolute minimum and the absolute maximum of the function.

The main difference between this process and the process that we used in Calculus I is that the "boundary" in Calculus I was just two points and so there really wasn't a lot to do in the second

step. For these problems the majority of the work is often in the second step as we will often end up doing a Calculus I absolute extrema problem one or more times.

4.5 Lagrange Multipliers

In this section we are going to take a look at another way of optimizing a function subject to given constraint(s). The constraint(s) may be the equation(s) that describe the boundary of a region although in this section we won't concentrate on those types of problems since this method just requires a general constraint and doesn't really care where the constraint came from.

We want to optimize (i.e. find the minimum and maximum value of) a function, $f(x, y, z)$, subject to the constraint $g(x, y, z) = k$. Again, the constraint may be the equation that describes the boundary of a region or it may not be. The process is actually fairly simple, although the work can still be a little overwhelming at times.

Method of Lagrange Multipliers

1. Solve the following system of equations

$$\begin{aligned}\nabla f(x, y, z) &= \lambda \nabla g(x, y, z) \\ g(x, y, z) &= k\end{aligned}$$

2. Plug in all solutions, (x, y, z) , from the first step into $f(x, y, z)$ and identify the minimum and maximum values, provided they exist.

The constant, λ , is called the *Lagrange Multiplier*.

Notice that the system of equations actually has four equations, we just wrote the system in a simpler form. To see this let's take the first equation and put in the definition of the gradient vector to see what we get.

$$\begin{aligned}\langle f_x, f_y, f_z \rangle &= \lambda \langle g_x, g_y, g_z \rangle \\ &= \langle \lambda g_x, \lambda g_y, \lambda g_z \rangle\end{aligned}$$

In order for these two vectors to be equal the individual components must also be equal. So, we actually have three equations here.

$$f_x = \lambda g_x \quad f_y = \lambda g_y \quad f_z = \lambda g_z$$

These three equations along with the constraint, $g(x, y, z) = k$, give four equations with four unknowns x , y , z , and λ .

Naturally, we can extend this to multiple constraints. Suppose we want to optimize $f(x, y, z)$ subject to the constraints $g(x, y, z) = c$ and $h(x, y, z) = k$. The system that we need to solve in this case is,

$$\begin{aligned}\nabla f(x, y, z) &= \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \\ g(x, y, z) &= c \\ h(x, y, z) &= k\end{aligned}$$

Chapter 5

Multiple Integrals

In Calculus I we moved on to the subject of integrals once we had finished the discussion of derivatives. The same is true in this course. Now that we have finished our discussion of derivatives of functions of more than one variable we need to move on to integrals of functions of two or three variables.

Most of the derivatives topics extended somewhat naturally from their Calculus I counterparts and that will be the same here. However, because we are now involving functions of two or three variables there will be some differences as well. here will be new notation and some new issues that simply don't arise when dealing with functions of a single variable.

5.1 Double Integrals

Before starting on double integrals let's do a quick review of the definition of a definite integrals for functions of single variables. First, when working with the integral,

$$\int_a^b f(x) dx$$

we think of x 's as coming from the interval $a \leq x \leq b$. For these integrals we can say that we are integrating over the interval $a \leq x \leq b$. Note that this does assume that $a < b$, however, if we have $b > a$ then we can just use the interval $b \leq x \leq a$.

Now, when we derived the definition of the definite integral we first thought of this as an area problem. We first asked what the area under the curve was and to do this we broke up the interval $a \leq x \leq b$ into n subintervals of width Δx and choose a point, x_i^* , from each interval. Each of the rectangles has height of $f(x_i^*)$ and we could then use the area of each of these rectangles to approximate the area as follows.

$$A \approx f(x_1^*)\Delta x + f(x_2^*)\Delta x + \dots + f(x_i^*)\Delta x + \dots + f(x_n^*)\Delta x$$

To get the exact area we then took the limit as n goes to infinity and this was also the definition of the definite integral.

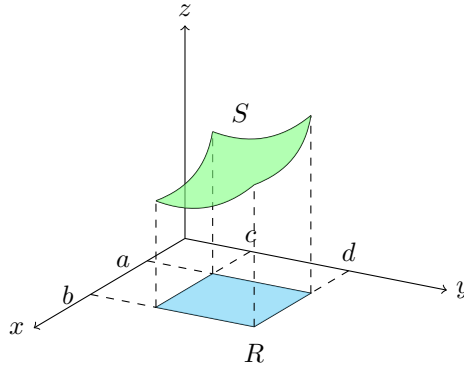
$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x$$

In this section we want to integrate a function of two variables, $f(x, y)$. With functions of one variable we integrated over an interval (i.e. a one-dimensional space) and so it makes some sense then that when integrating a function of two variables we will integrate over a region of \mathbb{R}^2 .

We will start out by assuming that the region in \mathbb{R}^2 is a rectangle which we will denote as follows,

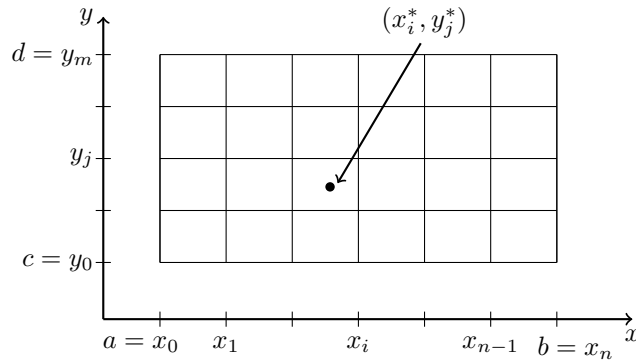
$$R = [a, b] \times [c, d]$$

This means that the ranges for x and y are $a \leq x \leq b$ and $c \leq y \leq d$. Also, we will initially assume that $f(x, y) > 0$ although this doesn't really have to be the case. Let's start out with the graph of the surface S given by graphing $f(x, y)$ over the rectangle R .



Now, just like with functions of one variable let's not worry about integrals quite yet. Let's first ask what the volume of the region under S (and above the xy -plane of course) is.

We will first approximate the volume much as we approximated the area above. We will first divide up $a \leq x \leq b$ into n subintervals and divide up $c \leq y \leq d$ into m subintervals. This will divide up R into a series of smaller rectangles and from each of these we will choose a point (x_i^*, y_j^*) . Here is a sketch of this set up.



Now, over each of these smaller rectangles we will construct a box whose height is given by $f(x_i^*, y_j^*)$. Each of the rectangles has a base area of ΔA and a height of $f(x_i^*, y_j^*)$ so the volume of each of these boxes is $f(x_i^*, y_j^*)\Delta A$. The volume under the surface S is then approximately,

$$V \approx \sum_{i=1}^n \sum_{j=1}^m f(x_i^*, y_j^*)\Delta A$$

We will have a double sum since we will need to add up volumes in both the x and y directions. To get a better estimation of the volume we will take n and m larger and larger and to get the exact volume we will need to take the limit as both n and m go to infinity. In other words,

$$V = \lim_{n, m \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^m f(x_i^*, y_j^*)\Delta A$$

Now, this should look familiar. This looks a lot like the definition of the integral of a function of single variable. In fact this is also the definition of a double integral, or more exactly an integral of a function of two variables over a rectangle.

Here is the official definition of a double integral of a function of two variables over a rectangular region R as well as the notation that we'll use for it.

$$\iint_R f(x, y) dA = \lim_{n, m \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^m f(x_i^*, y_j^*)\Delta A$$

Note that one interpretation of the double integral of $f(x, y)$ over the rectangle R is the volume under the function $f(x, y)$ (and above the xy -plane). Or,

$$\text{Volume} = \iint_R f(x, y) dA$$

5.2 Iterated Integrals

The following theorem tells us how to compute a double integral over a rectangle.

Fubini's Theorem

If $f(x, y)$ is continuous on $R = [a, b] \times [c, d]$ then,

$$\iint_R f(x, y) dA = \int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy$$

These integrals are called *iterated integrals*

Note that there are in fact two ways of computing a double integral and also notice that the inner differential matches up with the limits on the inner integral and similarly for the outer differential and limits. In other words, if the inner differential is dy then the limits on the inner integral must be y limits of integration and if the outer differential is dx then the limits on the outer integral must be x limits of integration.

Now, on some level this is just notation and doesn't really tell us how to compute the double integral. Let's just take the first possibility above and change the notation a little.

$$\iint_R f(x, y) dA = \int_a^b \left[\int_c^d f(x, y) dy \right] dx$$

We will compute the double integral by first computing

$$\int_c^d f(x, y) dy$$

and we compute this by holding x constant and integrating with respect to y as if this were a single integral. This will give a function involving only x 's which we can in turn integrate.

We've done a similar process with partial derivatives. To take the derivative of a function with respect to y we treated the x 's as constants and differentiated with respect to y as if it was a function of a single variable. Double integrals work in the same manner. We think of all the x 's as constants and integrate with respect to y or we think of all y 's as constants and integrate with respect to x .

The next topic of this section is a quick fact that can be used to make some iterated integrals somewhat easier to compute on occasion.

Fact

If $f(x, y) = g(x)h(y)$ and we are integrating over the rectangle $R = [a, b] \times [c, d]$ then

$$\iint_R f(x, y) dA = \iint_R g(x)h(y) dA = \left(\int_a^b g(x) dx \right) \left(\int_c^d h(y) dy \right)$$

So, if we can break up the function into a function only of x times a function of y then we can do the two integrals individually and multiply them together.

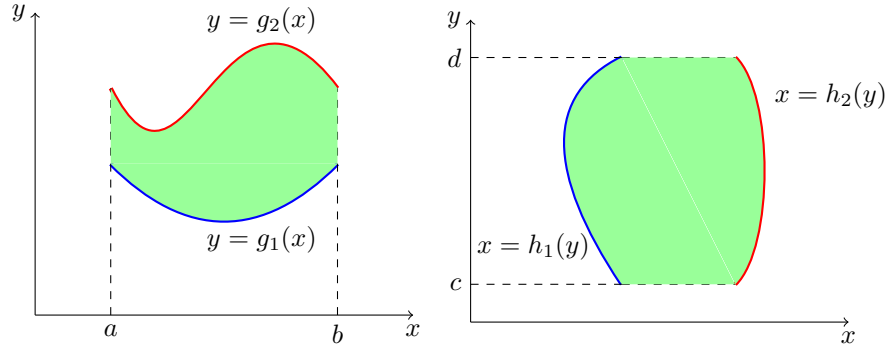
5.3 Double Integrals Over General Regions

In the previous section we looked at double integrals over rectangular regions. The problem with this is that most of the regions are not rectangular so we need to now look at the following double integral,

$$\iint_D f(x, y) dA$$

where D is any region.

There are two types of regions that we need to look at. Here is a sketch of both of them.



We will often use set builder notation to describe these regions. Here is the definition for the region in Case I

$$D = \{(x, y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

and here is the definition for the region in Case II

$$D = \{(x, y) \mid h_1(y) \leq x \leq h_2(y), c \leq y \leq d\}$$

The double integral for both of these cases are defined in terms of iterated integrals as follows.

In Case I where $D = \{(x, y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$ the integral is defined to be,

$$\iint_D f(x, y) dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx$$

In Case 2 where $D = \{(x, y) \mid h_1(y) \leq x \leq h_2(y), c \leq y \leq d\}$ the integral is defined to be,

$$\iint_D f(x, y) dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) dx dy$$

Here are some properties of the double integral. Note that all three of these properties are really just extensions of properties of single integrals that have been extended to double integrals.

Properties

1. $\iint_D f(x, y) + g(x, y) dA = \iint_D f(x, y) dA + \iint_D g(x, y) dA$
2. $\iint_D cf(x, y) dA = c \iint_D f(x, y) dA$, where c is any constant.
3. If the region D can be split into two separate regions D_1 and D_2 then the integral can be written as

$$\iint_D f(x, y) dA = \iint_{D_1} f(x, y) dA + \iint_{D_2} f(x, y) dA$$

The final topic of this section is two geometric interpretations of a double integral. The first interpretation is an extension of the idea that we used to develop the idea of a double integral in the first section of this chapter. We did this by looking at the volume of the solid that was below the surface of the function $z = f(x, y)$ and over the rectangle R in the xy -plane. This idea can be extended to more general regions.

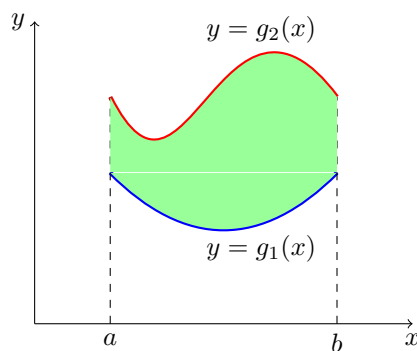
The volume of the solid that lies below the surface given by $z = f(x, y)$ and above the region D in the xy -plane is given by,

$$V = \iint_D f(x, y) dA$$

The second geometric interpretation of a double integral is the following.

$$\text{Area of } D = \iint_D dA$$

This is easy to see why this is true in general. Lets suppose that we want to find the area of the region shown below.



From Calculus I we know that this area can be found by the integral,

$$A = \int_a^b g_2(x) - g_1(x) dx$$

Or in terms of a double integral we have,

$$\begin{aligned}
 \text{Area of } D &= \iint_D dA \\
 &= \int_a^b \int_{g_1(x)}^{g_2(x)} dy \, dx \\
 &= \int_a^b y \Big|_{g_1(x)}^{g_2(x)} dx \\
 &= \int_a^b g_2(x) - g_1(x) \, dx
 \end{aligned}$$

This is exactly the same formula we had in Calculus I.

5.4 Double Integrals in Polar Coordinates

To this point we've seen quite a few double integrals. However, in every case we've seen to this point the region D could be easily described in terms of simple functions in Cartesian coordinates. In this section we want to look at some regions that are much easier to describe in terms of polar coordinates. For instance, we might have a region that is a disk, ring, or a portion of a disk or ring. In these cases using Cartesian coordinates could be somewhat cumbersome. For instance let's suppose we wanted to do the following integral,

$$\iint_D f(x, y) \, dA, \quad D \text{ is the disk of radius } 2$$

To this we would have to determine a set of inequalities for x and y that describe this region. These would be,

$$\begin{aligned}
 -2 &\leq x \leq 2 \\
 -\sqrt{4-x^2} &\leq y \leq \sqrt{4-x^2}
 \end{aligned}$$

With these limits the integral would become,

$$\iint_D f(x, y) \, dA = \int_{-2}^2 \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} f(x, y) \, dy \, dx$$

Due to the limits on the inner integral this is liable to be an unpleasant integral to compute.

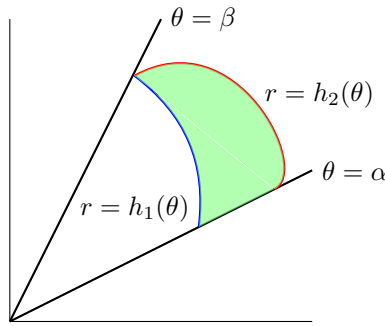
However, a disk of radius 2 can be defined in polar coordinates by the following inequalities,

$$\begin{aligned}
 0 &\leq r \leq 2 \\
 0 &\leq \theta \leq 2\pi
 \end{aligned}$$

These are very simple limits and, in fact, are constant limits of integration which almost always makes integrals somewhat easier.

So, if we could convert our double integral formula into one involving polar coordinates we would be in pretty good shape. The problem is that we can't just convert the dx and the dy into a dr and a $d\theta$. In computing double integrals to this point we have been using the fact that $dA = dx dy$ and this really does require Cartesian coordinates to use. Once we've moved into polar coordinates $dA \neq dr d\theta$ and so we're going to need to determine just what dA is under polar coordinates.

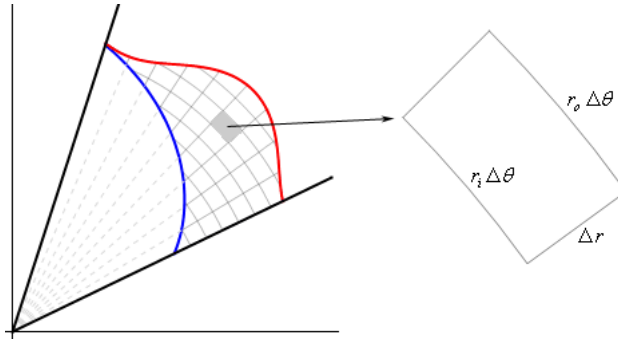
So, let's step back a little bit and start off with a general region in terms of polar coordinates and see what we can do with that. Here is a sketch of some region using polar coordinates.



So, our general region will be defined by inequalities,

$$\begin{aligned}\alpha &\leq \theta \leq \beta \\ h_1(\theta) &\leq r \leq h_2(\theta)\end{aligned}$$

Now, to find dA let's redo the figure above as follows,



As shown, we'll break up the region into a mesh of radial lines and arcs. Now, if we pull one of the pieces of the mesh out as shown we have something that is almost, but not quite a rectangle. The area of this piece is ΔA . The two sides of this piece both have length $\Delta r = r_o - r_i$ where r_o is the radius of the outer arc and r_i is the radius of the inner arc. Basic geometry then tells us that the length of the inner edge is $r_i \Delta \theta$ while the length of the out edge is $r_o \Delta \theta$ where $\Delta \theta$ is the angle between the two radial lines that form the sides of this piece.

Now, let's assume that we've taken the mesh so small that we can assume that $r_i \approx r_o = r$

and with this assumption we can also assume that our piece is close enough to a rectangle that we can also then assume that,

$$\Delta A \approx r \Delta \theta \Delta r$$

Also, if we assume that the mesh is small enough then we can also assume that,

$$dA \approx \Delta A \quad d\theta \approx \Delta \theta \quad dr \approx \Delta r$$

With these assumptions we then get $dA \approx r dr d\theta$.

In order to arrive at this we had to make the assumption that the mesh was very small. This is not an unreasonable assumption. Recall that the definition of a double integral is in terms of two limits and as limits go to infinity the mesh size of the region will get smaller and smaller. In fact, as the mesh size gets smaller and smaller the formula above becomes more and more accurate and so we can say that,

$$dA = r dr d\theta$$

Now, if we're going to be converting an integral in Cartesian coordinates into an integral in polar coordinates we are going to have to make sure that we've also converted all the x 's and y 's into polar coordinates as well. To do this we'll need to remember the following conversion formulas,

$$x = r \cos \theta \quad y = r \sin \theta \quad r^2 = x^2 + y^2$$

Putting this all together we have,

$$\iint_D f(x, y) dA = \int_{\alpha}^{\beta} \int_{h_1(\theta)}^{h_2(\theta)} f(r \cos \theta, r \sin \theta) r dr d\theta$$

It is important to not forget the added r and don't forget to convert the Cartesian coordinates in the function over to polar coordinates.

5.5 Triple Integrals

Now that we know how to integrate over a two-dimensional region we need to move on to integrating over a three-dimensional region. We used a double integral to integrate over a two-dimensional region and so it shouldn't be too surprising that we'll use a triple integral to integrate over a three dimensional region. The notation for the general triple integrals is,

$$\iiint_E f(x, y, z) dV$$

Let's start simple by integrating over the box,

$$B = [a, b] \times [c, d] \times [u, v]$$

Note that when using this notation we list the x 's first, the y 's second and the z 's third.

The triple integral in this case is,

$$\iiint_E f(x, y, z) dV = \int_u^v \int_c^d \int_a^b f(x, y, z) dx dy dz$$

Note that we integrated with respect to x first, then y , and finally z here, but in fact there is no reason to the integrals in this order. There are 6 different possible orders to do the integral in and which order you do the integral in will depend upon the function and the order that you feel will be the easiest. We will get the same answer regardless of the order however.

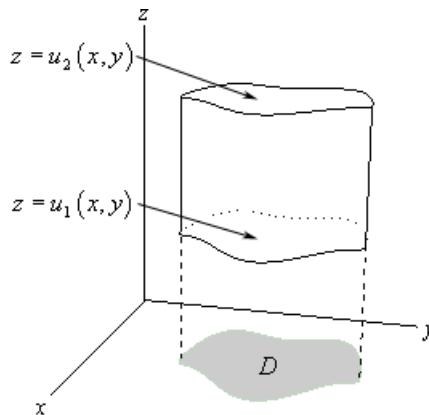
Before moving on to more general regions let's get a nice geometric interpretation about the triple integral out of the way so we can use it in some of the examples to follow.

Fact

The volume of the three-dimensional region E is given by the integral,

$$V = \iiint_E dV$$

Let's now move on to the more general three-dimensional regions. We have three different possibilities for a general region. Here is a sketch of the first possibility.



In this case we define the region E as follows,

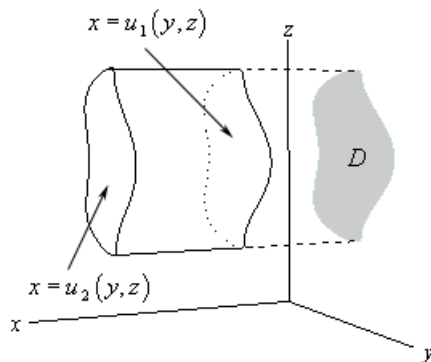
$$E = \{(x, y, z) \mid (x, y) \in D, u_1(x, y) \leq z \leq u_2(x, y)\}$$

In this case we will evaluate the triple integral as follows,

$$\iiint_E f(x, y, z) dV = \iint_D \left[\int_{u_1(x, y)}^{u_2(x, y)} f(x, y, z) dz \right] dA$$

where the double integral can be evaluated in any of the methods that we saw in the previous couple of sections. In other words, we can integrate first with respect to x , we can integrate first with respect to y , or we can use polar coordinates as needed.

Let's now move onto the second possible three-dimensional region we may run into for triple integrals. Here is a sketch of this region.



For this possibility we define the region E as follows,

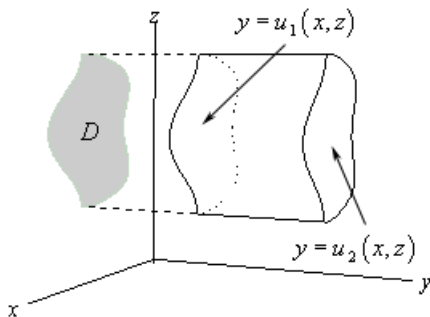
$$E = \{(x, y, z) \mid (y, z) \in D, u_1(y, z) \leq x \leq u_2(y, z)\}$$

So, the region D will be a region in the yz -plane. Here is how we will evaluate these integrals.

$$\iiint_E f(x, y, z) dV = \iint_D \left[\int_{u_1(y, z)}^{u_2(y, z)} f(x, y, z) dx \right] dA$$

As with the first possibility we will have two options for doing the double integral in the yz -plane as well as the option of using polar coordinates if needed.

We now need to look at the third (and final) possible three-dimensional region we may run into for triple integrals. Here is a sketch of this region.



In this final case E is defined as,

$$E = \{(x, y, z) \mid (x, z) \in D, u_1(x, z) \leq y \leq u_2(x, z)\}$$

and here the region D will be a region in the xz -plane. Here is how we will evaluate these integrals.

$$\iiint_E f(x, y, z) dV = \iint_D \left[\int_{u_1(x, z)}^{u_2(x, z)} f(x, y, z) dy \right] dA$$

where we will can use either of the two possible orders for integrating D in the xz -plane or we can use polar coordinates if needed.

5.6 Triple Integrals in Cylindrical Coordinates

In this section we want to take a look at triple integrals done completely in Cylindrical Coordinates. Recall that cylindrical coordinates are really nothing more than an extension of polar coordinates into three dimensions. The following are the conversion formulas for cylindrical coordinates.

$$x = r \cos \theta \quad y = r \sin \theta \quad z = z$$

In order to do the integral in cylindrical coordinates we will need to know what dV will become in terms of cylindrical coordinates. We will be able to show in the Change of Variables section of this chapter that,

$$dV = r \, dz \, dr \, d\theta$$

The region, E , over which we are integrating becomes,

$$\begin{aligned} E &= \{(x, y, z) \mid (x, y) \in D, u_1(x, y) \leq z \leq u_2(x, y)\} \\ &= \{(r, \theta, z) \mid \alpha \leq \theta \leq \beta, h_1(\theta) \leq r \leq h_2(\theta), u_1(r \cos \theta, r \sin \theta) \leq z \leq u_2(r \cos \theta, r \sin \theta)\} \end{aligned}$$

Note that we've only given this for E 's in which D is in the xy -plane. We can modify this accordingly if D is in the yz -plane or the xz -plane as needed.

In terms of cylindrical coordinates a triple integral is,

$$\iiint_E f(x, y, z) \, dV = \int_{\alpha}^{\beta} \int_{h_1(\theta)}^{h_2(\theta)} \int_{u_1(r \cos \theta, r \sin \theta)}^{u_2(r \cos \theta, r \sin \theta)} f(r \cos \theta, r \sin \theta, z) \, r \, dz \, dr \, d\theta$$

5.7 Triple Integrals in Spherical Coordinates

In the previous section we looked at doing integrals in terms of cylindrical coordinates and we now need to take a quick look at doing integrals in terms of spherical coordinates. Here are the conversion formulas for spherical coordinates.

$$\begin{aligned} x &= \rho \sin \phi \cos \theta & y &= \rho \sin \phi \sin \theta & z &= \rho \cos \phi \\ x^2 + y^2 + z^2 &= \rho^2 \end{aligned}$$

We also have the following restrictions on the coordinates.

$$\rho \geq 0 \quad 0 \leq \phi \leq \pi$$

In the next section we will show that

$$dV = \rho^2 \sin \phi \, d\rho \, d\theta \, d\phi$$

If $a \leq \rho \leq b$, $\alpha \leq \theta \leq \beta$, $\delta \leq \phi \leq \gamma$, then the triple integral becomes

$$\iiint_E f(x, y, z) \, dV = \int_{\delta}^{\gamma} \int_{\alpha}^{\beta} \int_a^b \rho^2 \sin \phi \, f(\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi) \, d\rho \, d\theta \, d\phi$$

This looks bad, but given that the limits are all constants the integrals here tend to not be too bad.

5.8 Change of Variables

Back in Calculus I we had the substitution rule that told us that,

$$\int_a^b f(g(x))g'(x) dx = \int_c^d f(u) du \quad \text{where } u = g(x)$$

In essence this is taking an integral in terms of x 's and changing it into terms of u 's. We want to do something similar for double and triple integrals. In fact we've already done this to a certain extent when we converted double integrals to polar coordinates and when we converted triple integrals to cylindrical or spherical coordinates. The main difference is that we didn't actually go through the details of where the formulas came from. If you recall, in each of those cases we commented that we would justify the formulas for dA and dV eventually. Now is the time to do that justification.

While often the reason for changing variables is to get us an integral that we can do with the new variables, another reason for changing variables is to convert the region into a nicer region to work with. When we were converting the polar, cylindrical or spherical coordinates we didn't worry about this change since it was easy enough to determine the new limits based on the given region. That is not always the case however.

We will start with double integrals. In order to change variables in a double integral we will need the Jacobian of the transformation. Here is the definition of the Jacobian.

Definition

The *Jacobian* of the transformation $x = g(u, v), y = h(u, v)$ is

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

The Jacobian is defined as a determinant of a 2×2 matrix. Now that we have the Jacobian out of the way we can give the formula for change of variables for a double integral.

Change of Variables for Double Integrals

Suppose that we want to integrate $f(x, y)$ over the region R . Under the transformation $x = g(u, v)$, $y = h(u, v)$, the region becomes S and the integral becomes,

$$\iint_R f(x, y) dA = \iint_S f(g(u, v), h(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$$

If we look just at the differentials in the above formula we can also say that

$$dA = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$$

Let's now briefly look at triple integrals. In this case we will again start with a region R and use the transformation $x = g(u, v, w)$, $y = h(u, v, w)$, and $z = k(u, v, w)$ to transform the region into the new region S . To do the integral we will need a Jacobian, just as we did with double integrals. Here is the definition of the Jacobian for this kind of transformation.

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

In this case the Jacobian is defined in terms of the determinant of a 3×3 matrix. The integral under this transformation is,

$$\iiint_E f(x, y, z) dV = \iiint_S f(g(u, v, w), h(u, v, w), k(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du dv dw$$

Let's verify the formula for dV for spherical coordinates. Here the transformation is just the standard conversion formulas.

$$x = \rho \sin \phi \cos \theta \quad y = \rho \sin \phi \sin \theta \quad z = \rho \cos \phi$$

The Jacobian is,

$$\begin{aligned} \frac{\partial(x, y, z)}{\partial(u, v, w)} &= \begin{vmatrix} \sin \cos \theta & -\rho \sin \phi \sin \theta & \rho \cos \phi \cos \theta \\ \sin \phi \sin \theta & \rho \sin \phi \cos \theta & \rho \cos \phi \sin \theta \\ \cos \phi & 0 & -\rho \sin \phi \end{vmatrix} \\ &= -\rho^2 \sin^3 \phi \cos^2 \theta - \rho^2 \sin \phi \cos^2 \phi \sin^2 \theta + 0 - \rho^2 \sin^3 \phi \sin^2 \theta - 0 \\ &\quad - \rho^2 \sin \phi \cos^2 \phi \cos^2 \theta \\ &= -\rho^2 \sin^3 \phi (\cos^2 \theta + \sin^2 \theta) - \rho^2 \sin \phi \cos^2 \phi (\sin^2 \theta + \cos^2 \theta) \\ &= -\rho^2 \sin^3 \phi - \rho^2 \sin \phi \cos^2 \phi \\ &= -\rho^2 \sin \phi (\sin^2 \phi + \cos^2 \phi) \\ &= -\rho^2 \sin \phi \end{aligned}$$

Finally, dV becomes,

$$dV = |-\rho^2 \sin \phi| d\rho d\theta d\phi = \rho^2 \sin \phi d\rho d\theta d\phi$$

Recall that we restricted ϕ to the range $0 \leq \phi \leq \pi$ for spherical coordinates and so we know that $\sin \phi \geq 0$ and so we don't need the absolute value bars on the sine.

5.9 Surface Area

In this section we will look at the lone application (aside from the area and volume interpretations) of multiple integrals in this material. This is not the first time that we've looked at surface area. We first saw surface area in Calculus II, however, in that setting we were looking at the surface area of a solid of revolution. In other words we were looking at the surface area of a solid obtained by rotating a function about the x or y axis. In this section we want to look at a much more

general setting although you will note that the formula here is very similar to the formula we saw back in Calculus II.

Here we want to find the surface area of the surface given by $z = f(x, y)$ where (x, y) is a point from the region D in the xy -plane. In this case the surface area is given by,

$$\mathcal{S} = \iint_D \sqrt{(f_x)^2 + (f_y)^2 + 1} \, dA$$

5.10 Physical Applications of Double Integrals

There is a ton of applications for double integrals. Here are a few.

Mass and Static Moments of a Lamina

Suppose we have a lamina which occupies a region R in the xy -plane and is made of non-homogeneous material. Its density at a point (x, y) in the region R is $\rho(x, y)$. The total *mass of the lamina* is expressed through the double integral as follows:

$$m = \iint_R \rho(x, y) \, dA$$

The *static moment of the lamina about the x -axis* is given by the formula

$$M_x = \iint_R y\rho(x, y) \, dA$$

Similarly, the *static moment of the lamina about the y -axis* is

$$M_y = \iint_R x\rho(x, y) \, dA$$

The coordinates of the center of mass of a lamina occupying the region R in the xy -plane with density function $\rho(x, y)$ are described by the formulas

$$\begin{aligned}\bar{x} &= \frac{M_y}{m} = \frac{1}{m} \iint_R x\rho(x, y) \, dA \\ \bar{y} &= \frac{M_x}{m} = \frac{1}{m} \iint_R y\rho(x, y) \, dA\end{aligned}$$

When the mass density of the lamina is $\rho(x, y) = 1$ for all (x, y) in the region R , the center of mass is defined only by the shape of the region and is called the *centroid* of R .

Moments of Inertia of a Lamina

The *moment of inertia of a lamina about the x -axis* is defined by the formula

$$I_x = \iint_R y^2 \rho(x, y) \, dA$$

Similarly, the *moment of inertia of a lamina about the y -axis* is given by

$$I_y = \iint_R x^2 \rho(x, y) dA$$

The *polar moment of inertia* is

$$I_0 = \iint_R (x^2 + y^2) \rho(x, y) dA$$

Charge of a Plate

Suppose electrical charge is distributed over a region which has area R in the xy -plane and its charge density is defined by the function $\sigma(x, y)$. Then the total *charge* Q of the plate is defined by the expression

$$Q = \iint_R \sigma(x, y) dA$$

Average of a Function

Let $f(x, y)$ be a continuous function over a closed region R in the xy -plane. The average value μ of the function $f(x, y)$ in the region R is given by the formula

$$\mu = \frac{1}{S} \iint_R f(x, y) dA$$

where $S = \iint_R dA$ is the area of the region of integration R .

Chapter 6

Line Integrals

In this chapter we are going to start looking at Calculus with vector fields. In particular we will be looking at a new type of integral, the line integral and some of the interpretations of the line integral. We will also take a look at one of the more important theorems involving line integrals, Green's Theorem.

6.1 Vector Fields

We need to start this chapter off with the definition of a vector field as they will be a major component of both this chapter and the next. Let's start off with the formal definition of a vector field.

Definition

A *vector field* on two (or three) dimensional space is a function \mathbf{F} that assigns to each point (x, y) (or (x, y, z)) a two (or three dimensional) vector given by $\mathbf{F}(x, y)$ (or $\mathbf{F}(x, y, z)$).

The standard notation for the function \mathbf{F} is,

$$\begin{aligned}\mathbf{F}(x, y) &= P(x, y)\hat{\mathbf{i}} + Q(x, y)\hat{\mathbf{j}} \\ \mathbf{F}(x, y, z) &= P(x, y, z)\hat{\mathbf{i}} + Q(x, y, z)\hat{\mathbf{j}} + R(x, y, z)\hat{\mathbf{k}}\end{aligned}$$

depending on whether or not we're in two or three dimensions. The function P , Q , R (if it is present) are sometimes called *scalar functions*.

In the third chapter we looked at the gradient vector. Recall that given a function $f(x, y, z)$ the gradient vector is defined by,

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

This is a vector field and is often called a *gradient vector field*.

The final topic of this section is that of conservative vector fields. A vector field \mathbf{F} is called a *conservative vector field* if there exists a function f such that $\mathbf{F} = \nabla f$. If \mathbf{F} is a conservative vector field then the function, f , is called a *potential function* for \mathbf{F} .

6.2 Line Integrals - Part I

In this section we are now going to introduce a new kind of integral. However, before we do that it is important to note that you will need to remember how to parameterize equations, or put another way, you will need to be able to write down a set of parametric equations for a given curve.

In Calculus I we integrated $f(x)$, a function of a single variable, over an interval $[a, b]$. In this case we were thinking of x as taking all the values in this interval starting at a and ending at b . With line integrals we will start with integrating the function $f(x, y)$, a function of two variables, and the values of x and y that we're going to use will be the points, (x, y) , that lie on a curve γ . Note that this is different from the double integrals that we were working with in the previous chapter where the points came out of some two-dimensional region.

Let's start with the curve γ that the points come from. We will assume that the curve is *smooth* (defined shortly) and is given by the parametric equations,

$$x = h(t) \quad y = g(t) \quad a \leq t \leq b$$

We will often want to write the parameterization of the curve as a vector function. In this case the curve is given by,

$$\mathbf{r}(t) = h(t)\mathbf{i} + g(t)\mathbf{j} \quad a \leq t \leq b$$

The curve is called smooth if $\mathbf{r}'(t)$ is continuous and $\mathbf{r}'(t) \neq 0$ for all t . The line integral of $f(x, y)$ along γ is denoted by,

$$\int_{\gamma} f(x, y) ds$$

We use a ds here to acknowledge the fact that we are moving along the curve, γ , instead of the x -axis (denoted by dx) or the y -axis (denoted by dy). Because of the ds this is sometimes called the *line integral of f with respect to arc length*. If you recall from Calculus II when we looked at the arc length of a curve given by parametric equations we found it to be,

$$L = \int_a^b ds, \quad \text{where } ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt$$

It is no coincidence that we use ds for both of these problems. The ds is the same for both the arc length integral and the notation for the line integral. So, to compute a line integral we will convert everything over to the parametric equations. The line integral is then,

$$\int_{\gamma} f(x, y) ds = \int_a^b f(h(t), g(t)) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt$$

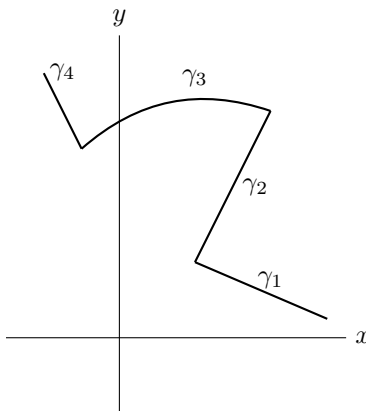
If we use the vector form of the parameterization we can simplify the notation up somewhat by noticing that,

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} = \|\mathbf{r}'(t)\|$$

Using this notation the line integral becomes,

$$\int_{\gamma} f(x, y) ds = \int_a^b f(h(t), g(t)) \|\mathbf{r}'(t)\| dt$$

Next we need to talk about line integrals over piecewise smooth curves. A *piecewise smooth curve* is any curve that can be written as the union of a finite number of smooth curves, $\gamma_1, \dots, \gamma_n$ where the end point of γ_i is the starting point of γ_{i+1} . Below is an illustration of a piecewise smooth curve.



Evaluation of line integrals over piecewise smooth curves is a relatively simple thing to do. All we do is evaluate the line integral over each of the pieces and then add them up. The line integral for some function over the above piecewise curve would be,

$$\int_{\gamma} f(x, y) ds = \int_{\gamma_1} f(x, y) ds + \int_{\gamma_2} f(x, y) ds + \int_{\gamma_3} f(x, y) ds + \int_{\gamma_4} f(x, y) ds$$

Let's suppose that the curve γ has the parameterization $x = h(t)$, $y = g(t)$. Let's also suppose that the initial point on the curve is A and the final point on the curve is B . The parameterization $x = h(t)$, $y = g(t)$ will then determine an *orientation* for the curve where the positive direction is the direction that is traced out as t increases. Finally, let $-\gamma$ be the curve with the same points as γ , however in this case the curve has B as the initial point and A as the final point, again t is increasing as we traverse this curve. In other words, given a curve γ , the curve $-\gamma$ is the same curve as γ except the direction has been reversed. We then have the following fact about line integrals with respect to arc length.

Fact

$$\int_{\gamma} f(x, y) ds = \int_{-\gamma} f(x, y) ds$$

So, for a line integral with respect to arc length we can change the direction of the curve and not change the value of the integral. This is a useful fact to remember as some line integrals will be easier in one direction than the other.

We can do line integrals over three-dimensional curves as well. Let's suppose that the three-dimensional curve γ is given by the parameterization,

$$x = x(t) \quad y = y(t) \quad z = z(t) \quad a \leq t \leq b$$

then the line integral is given by,

$$\int_{\gamma} f(x, y, z) ds = \int_a^b f(x(t), y(t), z(t)) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt$$

Note that often when dealing with three-dimensional space the parameterization will be given as a vector function.

$$\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$$

Notice that we changed up the notation for the parameterization a little. Since we rarely use the function names we simply kept the x , y , and z and added on the (t) part to denote that they may be functions of the parameter. Also notice that, as with two-dimensional curves, we have,

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} = \|\mathbf{r}'(t)\|$$

and the line integral can again be written as,

$$\int_{\gamma} f(x, y, z) ds = \int_a^b f(x(t), y(t), z(t)) \|\mathbf{r}'(t)\| dt$$

6.3 Line Integrals - Part II

In this section we want to look at line integrals with respect to x and/or y . As with the last section we will start with a two-dimensional curve γ with parameterization,

$$x = x(t) \quad y = y(t) \quad a \leq t \leq b$$

The line integral of f with respect to x is,

$$\int_{\gamma} f(x, y) dx = \int_a^b f(x(t), y(t)) x'(t) dt$$

The line integral of f with respect to y is,

$$\int_{\gamma} f(x, y) dy = \int_a^b f(x(t), y(t)) y'(t) dt$$

Note that the only notational difference between these two and the line integral with respect to arc length (from the previous section) is the differential. These have a dx or dy while the line integral with respect to arc length has a ds . So when evaluating line integrals be careful to first note which differential you've got so you don't work the wrong kind of line integral.

These two integral often appear together and so we have the following shorthand notation for these cases.

$$\int_{\gamma} P dx + Q dy = \int_{\gamma} P(x, y) dx + \int_{\gamma} Q(x, y) dy$$

Fact

If γ is any curve then,

$$\int_{-\gamma} P dx + Q dy = \int_{\gamma} P dx + Q dy$$

6.4 Line Integrals of Vector Fields

In this section we are going to evaluate line integrals of vector fields. We'll start with the vector field,

$$\mathbf{F}(x, y, z) = P(x, y, z)\hat{\mathbf{i}} + Q(x, y, z)\hat{\mathbf{j}} + R(x, y, z)\hat{\mathbf{k}}$$

and the three-dimensional, smooth curve given by

$$\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}} \quad a \leq t \leq b$$

The line integral of \mathbf{F} along γ is,

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) dt$$

We can also write line integrals of vector fields as a line integral with respect to arc length as follows,

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = \int_{\gamma} \mathbf{F} \cdot \mathbf{T} ds$$

where $\mathbf{T}(t)$ is the unit tangent vector given by,

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

6.5 Fundamental Theorem for Line Integrals

In Calculus I we had the Fundamental Theorem of Calculus that told us how to evaluate definite integrals. This told us,

$$\int_a^b F'(x) dx = F(b) - F(a)$$

It turns out that there is a version of this for line integrals over certain kinds of vector fields. Here it is.

Theorem 6.5.1. Suppose that γ is a smooth curve given by $\mathbf{r}(t)$, $a \leq t \leq b$. Also suppose that f is a function whose gradient vector, ∇f , is continuous on γ . Then,

$$\int_{\gamma} \nabla f \cdot d\mathbf{r} = f(\mathbf{r}(b)) - f(\mathbf{r}(a))$$

Proof. For the purposes of the proof we'll assume that we're working in three dimensions, but it can be done in any dimension. Let's start by just computing the line integral.

$$\begin{aligned}\int_{\gamma} \nabla f \cdot d\mathbf{r} &= \int_a^b \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t) dt \\ &= \int_a^b \left(\frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt} \right) dt\end{aligned}$$

Now, at this point we can use the Chain Rule to simplify the integrand as follows,

$$\begin{aligned}\int_{\gamma} \nabla f \cdot d\mathbf{r} &= \int_a^b \left(\frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt} \right) dt \\ &= \int_a^b \frac{d}{dt} [f(\mathbf{r}(t))] dt\end{aligned}$$

To finish this off we just need to use the Fundamental Theorem of Calculus for single integrals.

$$\int_{\gamma} \nabla f \cdot d\mathbf{r} = f(\mathbf{r}(b)) - f(\mathbf{r}(a))$$

■

The important idea from this proof (and hence about the Fundamental Theorem of Calculus) is that, for these kinds of line integrals, we didn't really need to know the path to get the answer. In other words, we could use any path we want and we'll always get the same results. Here are some definitions. The first one we've already seen before, but it's been a while and it's important in this section so we'll give it again. The remaining definitions are new.

Definitions

First suppose that \mathbf{F} is a continuous vector field in some domain D .

1. \mathbf{F} is a *conservative* vector field if there is a function f such that $\mathbf{F} = \nabla f$.
2. $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is *independent of path* if $\int_{\gamma_1} \mathbf{F} \cdot d\mathbf{r} = \int_{\gamma_2} \mathbf{F} \cdot d\mathbf{r}$ for any two paths γ_1 and γ_2 in D with the same initial and final points.
3. A path γ is called *closed* if its initial and final points are the same point. For example a circle is a closed path.
4. A path γ is *simple* if it doesn't cross itself. A circle is a simple curve.
5. A region D is *open* if it doesn't contain any of its boundary points.
6. A region D is *connected* if we can connect any two points in the region with a path that lies completely in D .

7. A region D is *simply-connected* if it is connected and it contains no holes.

With these definitions we can now give some nice facts.

Facts

1. $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is independent of path.
2. If \mathbf{F} is a conservative vector field then $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is independent of path.
3. If \mathbf{F} is a continuous vector field on an open connected region D and if $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is independent of path (for any path in D) then \mathbf{F} is a conservative vector field on D .
4. If $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is independent of path then $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed path γ .
5. If $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed path γ then $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r}$ is independent of path.

These are some nice facts to remember as we work with line integrals over vector fields. Also notice that 2 & 3 and 4 & 5 are converses of each other.

6.6 Conservative Vector Fields

In this section we want to look at two questions. First, given a vector field \mathbf{F} is there any way of determining if it is a conservative vector field? Secondly, if we know that \mathbf{F} is a conservative vector field how do we go about finding a potential function for the vector field?

The first question is easy to answer at this point if we have a two-dimensional vector field. For higher dimensional vector fields we'll need to wait until the final section in this chapter to answer this question. With that being said let's see how we do it for two-dimensional vector fields.

Theorem 6.6.1. Let $\mathbf{F} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a vector field on an open and simply-connected region D . Then if P and Q have continuous first order partial derivatives in D and

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

then \mathbf{F} is a conservative vector field.

Now that we know how to identify if a two-dimensional vector field is conservative we need to address how to find a potential function for the vector field. This is actually a fairly simple process. First, let's assume that the vector field is conservative and so we know that a potential function, $f(x, y)$ exists. We can then say that,

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} = P \mathbf{i} + Q \mathbf{j} = \mathbf{F}$$

Or by setting components equal we have,

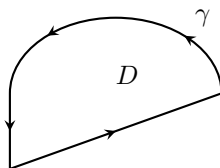
$$\frac{\partial f}{\partial x} = P \quad \text{and} \quad \frac{\partial f}{\partial y} = Q$$

By integrating each of these with respect to the appropriate variable we can arrive at the following two equations.

$$f(x, y) = \int P(x, y) dx \quad \text{or} \quad f(x, y) = \int Q(x, y) dy$$

6.7 Green's Theorem

In this section we are going to investigate the relationship between certain kinds of line integrals (on closed paths) and double integrals. Let's start off with a simple closed curve γ and let D be the region enclosed by the curve. Here is a sketch of such a curve and region.



First, notice that because the curve is simple and closed there are no holes in the region D . Also notice that a direction has been put on the curve. We will use the convention here that the curve γ has a *positive orientation* if it is traced out in a counter-clockwise direction. Another way to think of a positive orientation (that will cover much more general curves as well see later) is that as we traverse the path following the positive orientation the region D must always be on the left.

Given curves/regions such as this we have the following theorem.

Theorem 6.7.1 (Green's Theorem). Let γ be a positively oriented, piecewise smooth, simple, closed curve and let D be the region enclosed by the curve. If $P(x, y)$ and $Q(x, y)$ have continuous first order partial derivatives on D then,

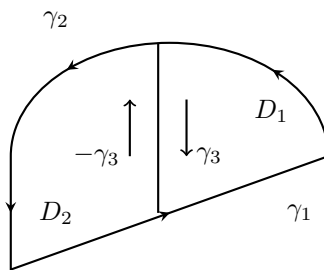
$$\int_{\gamma} P(x, y) dx + Q(x, y) dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

When working with a line integral in which the path satisfies the condition of Green's Theorem we will often denote the line integral as,

$$\oint_{\gamma} P(x, y) dx + Q(x, y) dy$$

So, Green's theorem will not work on regions that have holes in them. However, many regions do have holes in them. So, let's see how we can deal with those kinds of regions.

Let's start with the following region. Even though this region doesn't have any holes in it the arguments that we're going to go through will be similar to those that we'd need for regions with holes in them, except it will be a little easier to deal with and write down.



The region D will be $D_1 \cup D_2$. The boundary of D_1 is $\gamma_1 \cup \gamma_3$ while the boundary of D_2 is $\gamma_2 \cup (-\gamma_3)$ and notice that both of these boundaries are positively oriented. As we traverse each boundary the corresponding region is always on the left. Finally, also note that we can think of the whole boundary, γ , as,

$$\gamma = (\gamma_1 \cup \gamma_3) \cup (\gamma_2 \cup (-\gamma_3)) = \gamma_1 \cup \gamma_2$$

Now, let's start with the following double integral and use a basic property of double integrals to break it up.

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_{D_1 \cup D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_{D_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA + \iint_{D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

Next, use Green's theorem on each of these and again use the fact that we can break up line integrals into separate line integrals for each portion of the boundary.

$$\begin{aligned} \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA &= \iint_{D_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA + \iint_{D_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \\ &= \oint_{\gamma_1 \cup \gamma_3} P dx + Q dy + \oint_{\gamma_2 \cup (-\gamma_3)} P dx + Q dy \\ &= \oint_{\gamma_1} P dx + Q dy + \oint_{\gamma_3} P dx + Q dy + \oint_{\gamma_2} P dx + Q dy + \oint_{-\gamma_3} P dx + Q dy \end{aligned}$$

Recall that changing the orientation of a curve with line integrals with respect to x and/or y will simply change the sign on the integral. Using this fact we get,

$$\begin{aligned} \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA &= \oint_{\gamma_1} P dx + Q dy + \oint_{\gamma_3} P dx + Q dy + \oint_{\gamma_2} P dx + Q dy - \oint_{\gamma_3} P dx + Q dy \\ &= \oint_{\gamma_1} P dx + Q dy + \oint_{\gamma_2} P dx + Q dy \\ &= \oint_{\gamma_1 \cup \gamma_2} P dx + Q dy \\ &= \oint_{\gamma} P dx + Q dy \end{aligned}$$

So, what did we learn from this? If you think about it this was just a lot of work and all we got out of it was the result from Green's Theorem which we already knew to be true. What this exercise

has shown us is that if we break a region up as we did above then the portion of the line integral on the pieces of the curve that are in the middle of the region (each of which are in the opposite direction) will cancel out. This idea will help us in dealing with regions that have holes in them.

We will close out this section with an interesting application of Green's Theorem. Recall that we can determine the area of a region D with the following double integral.

$$A = \iint_D dA$$

Let's think of this double integral as the result of using Green's Theorem. In other words, let's assume that

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$$

and see if we can get some functions P and Q that will satisfy this. There are many functions that will satisfy this. Here are some of the more common functions.

$$\begin{array}{c|c|c} P = 0 & P = -y & P = -\frac{y}{2} \\ Q = x & Q = 0 & Q = \frac{x}{2} \end{array}$$

Then, if we use Green's Theorem in reverse we see that the area of the region D can also be computed by evaluating any of the following line integrals.

$$A = \oint_{\gamma} x dy = - \oint_{\gamma} y dx = \frac{1}{2} \oint_{\gamma} x dy - y dx$$

where γ is the boundary of the region D .

6.8 Curl and Divergence

In this section we will introducing new vector operations, the curl and the divergence of a vector.

Given the vector field $\mathbf{F} = P(x, y, z)\hat{\mathbf{i}} + Q(x, y, z)\hat{\mathbf{j}} + R(x, y, z)\hat{\mathbf{k}}$ the curl of \mathbf{F} is defined to be,

$$\text{curl } \mathbf{F} = (R_y - Q_z)\hat{\mathbf{i}} + (P_z - R_x)\hat{\mathbf{j}} + (Q_x - P_y)\hat{\mathbf{k}}$$

There is actually an easier definition of the curl of a vector field. To use it we will first need to define the ∇ operator. This is defined to be,

$$\nabla = \frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}$$

Using the ∇ we can define the curl as the following cross product,

$$\text{curl } \mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

We have a couple of nice facts that use the curl of a vector field.

Facts

1. If $f(x, y, z)$ has continuous second order partial derivatives then $\text{curl}(\nabla f) = \mathbf{0}$.
2. If \mathbf{F} is a conservative vector field then $\text{curl } \mathbf{F} = \mathbf{0}$. This is a direct result of what it means to be a conservative vector field and the previous fact.
3. If \mathbf{F} is defined on all of \mathbb{R}^3 whose components have continuous first order partial derivative and $\text{curl } \mathbf{F} = \mathbf{0}$ then \mathbf{F} is a conservative vector field.

Next we should talk about a physical interpretation of the curl. Suppose that \mathbf{F} is the velocity field of a flowing fluid. Then $\text{curl } \mathbf{F}$ represents the tendency of particles at the point (x, y, z) to rotate about the axis that points in the direction of \mathbf{F} . If $\text{curl } \mathbf{F} = \mathbf{0}$ then the fluid is called irrotational.

Given the vector field $\mathbf{F} = P(x, y, z)\hat{\mathbf{i}} + Q(x, y, z)\hat{\mathbf{j}} + R(x, y, z)\hat{\mathbf{k}}$ the divergence of \mathbf{F} is defined to be,

$$\text{div } \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

There is also a definition of the divergence in terms of the ∇ operator. The divergence can be defined in terms of the following dot product.

$$\text{div } \mathbf{F} = \nabla \cdot \mathbf{F}$$

We also have the following fact about the relationship between the curl and the divergence.

$$\text{div}(\text{curl } \mathbf{F}) = 0$$

Note that the result is a scalar.

We also have a physical interpretation of the divergence. If we again think of \mathbf{F} as the velocity field of a flowing fluid then $\text{div } \mathbf{F}$ represents the net rate of change of the mass of the fluid flowing from the point (x, y, z) per unit volume. This can also be thought of as the tendency of a fluid to diverge from a point. If $\text{div } \mathbf{F} = 0$ then the \mathbf{F} is called incompressible.

The final topic in this section is to give two vector forms of Green's Theorem. The first form uses the curl of the vector field and is,

$$\oint_{\gamma} \mathbf{F} \cdot d\mathbf{r} = \iint_D (\text{curl } \mathbf{F}) \cdot \hat{\mathbf{k}} dA$$

The second form uses the divergence. In this case we also need the outward unit normal to the curve γ . If the curve is parameterized by

$$\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}}$$

then the outward unit normal is given by,

$$\mathbf{n} = \frac{y'(t)}{\|\mathbf{r}'(t)\|}\hat{\mathbf{i}} - \frac{x'(t)}{\|\mathbf{r}'(t)\|}\hat{\mathbf{j}}$$

So the vector form of Green's Theorem that uses the divergence is given by,

$$\oint_{\gamma} \mathbf{F} \cdot \mathbf{n} ds = \iint_D \text{div } \mathbf{F} dA$$

Chapter 7

Surface Integrals

In the previous chapter we looked at evaluating integrals of functions or vector fields where the points came from a curve in two or three-dimensional space. We now want to extend this idea and integrate functions and vector fields where the points come from a surface in three-dimensional space. These integrals are called surface integrals.

7.1 Parametric Surfaces

Before we get into surface integrals we first need to talk about how to parameterize a surface. When we parameterized a curve we took values of t from some interval $[a, b]$ and plugged them into

$$\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}$$

and the resulting set of vectors would be the position vectors for the points on the curve.

With surfaces we'll do something similar. We will take points, (u, v) , out of some two-dimensional space D and plug them into

$$\mathbf{r}(u, v) = x(u, v)\hat{\mathbf{i}} + y(u, v)\hat{\mathbf{j}} + z(u, v)\hat{\mathbf{k}}$$

and the resulting set of vectors will be the position vectors for the points on the surface S that we are trying to parameterize. This is often called the *parametric representation* of the *parametric surface* S .

Let's take a quick look at a couple of applications. Let's look at finding the tangent plane to the parametric surface S given by,

$$\mathbf{r}(u, v) = x(u, v)\hat{\mathbf{i}} + y(u, v)\hat{\mathbf{j}} + z(u, v)\hat{\mathbf{k}}$$

First, define

$$\begin{aligned}\mathbf{r}_u(u, v) &= \frac{\partial x}{\partial u}(u, v)\hat{\mathbf{i}} + \frac{\partial y}{\partial u}(u, v)\hat{\mathbf{j}} + \frac{\partial z}{\partial u}(u, v)\hat{\mathbf{k}} \\ \mathbf{r}_v(u, v) &= \frac{\partial x}{\partial v}(u, v)\hat{\mathbf{i}} + \frac{\partial y}{\partial v}(u, v)\hat{\mathbf{j}} + \frac{\partial z}{\partial v}(u, v)\hat{\mathbf{k}}\end{aligned}$$

Now, provided $\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$ it can be shown that the vector $\mathbf{r}_u \times \mathbf{r}_v$ will be orthogonal to the surface S . This means that it can be used for the normal vector that we need in order to write down the equation of a tangent plane. This is an important idea that will be used many times throughout the next couple of sections.

The second application that we want to take a quick look at is the surface area of the parametric surface S given by,

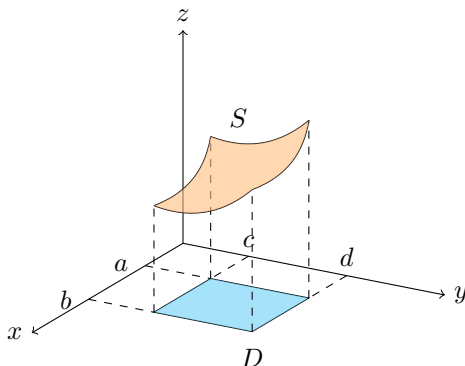
$$\mathbf{r}(u, v) = x(u, v)\hat{\mathbf{i}} + y(u, v)\hat{\mathbf{j}} + z(u, v)\hat{\mathbf{k}}$$

Provided S is traced out exactly once as (u, v) ranges over the points in D the surface area of S is given by,

$$A = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\| dA$$

7.2 Surface Integrals

It is now time to think about integrating functions over some surface, S , in three-dimensional space. Let's start off with a sketch of the surface S since the notation can get a little confusing once we get into it. Here is a sketch of some surface S .



The region S will lie above (in this case) some region D that lies in the xy -plane. We used a rectangle here, but it doesn't have to be of course. Also note that we could just as easily look at a surface S that was in front of some region D in the yz -plane or the xz -plane. Do not get so locked into the xy -plane that you can't do problems that have regions in the other two planes.

Now, how we evaluate the surface integral will depend upon how the surface is given to us. There are essentially two separate methods here, although as we will see they are really the same.

First, let's look at the surface integral in which the surface S is given by $z = g(x, y)$. In this case the surface integral is,

$$\iint_S f(x, y, z) dS = \iint_D f(x, y, g(x, y)) \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1} dA$$

Now, we need to be careful here as both of these look like standard double integrals. In fact the integral on the right is a standard double integral. The integral on the left however is a surface

integral. The way to tell them apart is by looking at the differentials. The surface integral will have a dS while the standard double integral will have a dA .

The second method for evaluating a surface integral is for those surfaces that are given by the parameterization,

$$\mathbf{r}(u, v) = x(u, v)\hat{\mathbf{i}} + y(u, v)\hat{\mathbf{j}} + z(u, v)\hat{\mathbf{k}}$$

In these cases the surface integral is,

$$\iint_S f(x, y, z) dS = \iint_D f(\mathbf{r}(u, v)) \|\mathbf{r}_u \times \mathbf{r}_v\| dA$$

where D is the range of the parameters that trace out the surface S .

7.3 Surface Integrals of Vector Fields

Just as we did with line integrals we now need to move on to surface integrals of vector fields. Recall that in line integrals the orientation of the curve we were integrating along could change the answer. The same thing will hold true with surface integrals. So, before we really get into doing surface integrals of vector fields we first need to introduce the idea of an *oriented surface*.

Let's start off with a surface that has two sides that has a tangent plane at every point (except possibly along the boundary). Making this assumption means that every point will have two unit normal vectors, \mathbf{n}_1 and $\mathbf{n}_2 = -\mathbf{n}_1$. This means that every surface will have two sets of normal vectors. The set that we choose will give the surface an orientation.

There is one convention that we will make in regards to certain kinds of oriented surfaces. First we need to define a *closed surface*. A surface S is closed if it is the boundary of some solid region E . A good example of a closed surface is the surface of a sphere. We say that the closed surface S has a *positive* orientation if we choose the set of unit normal vectors that point outward from the region E while the negative orientation will be the set of unit normal vectors that point in towards the region E .

In order to work with surface integrals of vector fields we will need to be able to write down a formula for the unit normal vector corresponding to the orientation that we've chosen to work with. We have two ways of doing this depending on how the surface has been given to us. First, let's suppose that the function is given by $z = g(x, y)$. In this case we first define a new function,

$$f(x, y, z) = z - g(x, y)$$

In terms of our new function the surface is then given by the equation $f(x, y, z) = 0$. Now, recall that ∇f will be orthogonal (or normal) to the surface given by $f(x, y, z) = 0$. This means that we have a normal vector to the surface. The only potential problem is that it might not be a unit normal vector. That isn't a problem since we also know that we can turn any vector into a unit vector by dividing the vector by its length. In our case this is,

$$\mathbf{n} = \frac{\nabla f}{\|\nabla f\|}$$

In this case it will be convenient to actually compute the gradient vector and plug this into the formula for the normal vector. Doing this gives,

$$\mathbf{n} = \frac{\nabla f}{\|\nabla f\|} = \frac{-g_x \hat{\mathbf{i}} - g_y \hat{\mathbf{j}} + g_z \hat{\mathbf{k}}}{\sqrt{(g_x)^2 + (g_y)^2 + 1}}$$

Notice that the component of the normal vector in the z -direction is always positive and so this normal vector will generally point upwards. It may not point directly up, but it will have an upwards component to it.

This will be important when we are working with a closed surface and we want the positive orientation. If we know that we can then look at the normal vector and determine if the positive orientation should point upwards or downwards. Remember that the “positive” orientation must point out of the region and this may mean downwards in places. Of course if it turns out that we need the downward orientation we can always take the negative of this unit vector and we’ll get the one that we need. Again, remember that we always have that option when choosing the unit normal vector.

Now, we need to discuss how to find the unit normal vector if the surface is given parametrically as,

$$\mathbf{r}(u, v) = x(u, v) \hat{\mathbf{i}} + y(u, v) \hat{\mathbf{j}} + z(u, v) \hat{\mathbf{k}}$$

In this case recall that the vector $\mathbf{r}_u \times \mathbf{r}_v$ will be normal to the tangent plane at a particular point. But if the vector is normal to the tangent plane at a point then it will also be normal to the surface at that point. So, this is a normal vector. In order to guarantee that it is a unit normal vector we will also need to divide it by its magnitude. So, in the case of parametric surfaces one of the unit normal vectors will be,

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}$$

So, given a vector field \mathbf{F} with unit normal vector \mathbf{n} then the surface integral of \mathbf{F} over the surface S is given by,

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_S \mathbf{F} \cdot \mathbf{n} dS$$

where the right hand integral is a standard surface integral. This is sometimes called *the flux of \mathbf{F} across S* .

We will leave this section with a quick interpretation of a surface integral over a vector field. If \mathbf{v} is the velocity field of a fluid then the surface integral

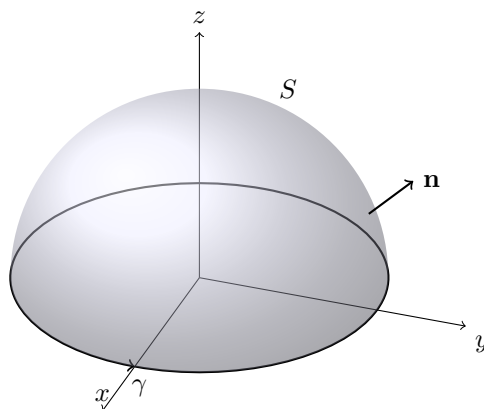
$$\iint_S \mathbf{v} \cdot d\mathbf{S}$$

represents the volume of fluid flowing through S per unit time.

7.4 Stokes’ Theorem

In this section we are going to take a look at a theorem that is a higher dimensional version of Green’s Theorem. In Green’s Theorem we related a line integral to a double integral over some

region. In this section we are going to relate a line integral to a surface integral. However, before we give the theorem we first need to define the curve that we're going to use in the line integral. Let's start off with the following surface with the indicated orientation.



Around the edge of this surface we have a curve γ . This curve is called the boundary curve. The orientation of the surface S will induce the positive orientation of γ . To get the positive orientation of γ think of yourself as walking along the curve. While you are walking along the curve if your head is pointing in the same direction as the unit normal vectors while the surface is on the left then you are walking in the positive direction on γ . Now that we have this curve definition out of the way we can give Stokes' Theorem.

Theorem 7.4.1 (Stokes' Theorem). Let S be an oriented smooth surface that is bounded by a simple, closed, smooth boundary curve γ with positive orientation. Also let \mathbf{F} be a vector field then,

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S}$$

In this theorem note that the surface S can actually be any surface so long as its boundary curve is given by γ . This is something that can be used to our advantage to simplify the surface integral on occasion.

7.5 Divergence Theorem

In the last section we relate surface integrals to triple integrals. We do this using the Divergence Theorem

Theorem 7.5.1 (Divergence Theorem). Let E be a simple solid region and S is the boundary surface of E with positive orientation. Let \mathbf{F} be a vector field whose components have continuous first order partial derivatives. Then,

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iiint_E \text{div } \mathbf{F} dV$$