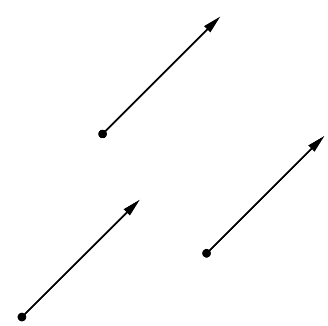
CALCULUS III, Third Semester
Table of Contents

XIII. Vector Algebra

13.1 Basic Concepts

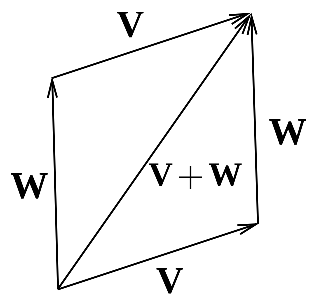
A *vector* $\mathbf{V}$ in the plane or in space is an arrow: it is determined by its length, denoted $|\mathbf{V}|$ and its direction. Two arrows represent the same vector if they have the same length and are parallel (see figure 13.1). We use vectors to represent entities which are described by magnitude and direction. For example, a force applied at a point is a vector: it is completely determined by the magnitude of the force and the direction in which it is applied. An object moving in space has, at any given time, a direction of motion, and a speed. This is represented by the velocity vector of the motion. More precisely, the velocity vector at a point is an arrow of length the speed (ds/dt), which lies on the tangent line to the trajectory. The success and importance of vector algebra derives from the interplay between geometric interpretation and algebraic calculation. In these notes, we will define the relevant concepts geometrically, and let this lead us to the algebraic formulation.

Figure 13.1



Newton did not write in terms of vectors, but through his diagrams we see that he clearly thought of forces in these terms. For example, he postulated that two forces acting simultaneously can be treated as acting sequentially. So suppose two forces, represented by vectors $\mathbf{V}$ and$\mathbf{W}$, act on an object at a particular point. What the object feels is the *resultant* of these two forces, which can be calculated by placing the vectors end to end (as in figure 13.2). Then the resultant is the vector from the initial point of the first vector to the end point of the second. Clearly, this is the same if we reverse the order of the vectors. We call this the *sum* of the two vectors, denoted $\mathbf{V} + \mathbf{W}$. For example, if an object is moving in a fluid in space with a velocity $\mathbf{V}$, while the fluid is moving with velocity $\mathbf{W}$, then the object moves (relative to a fixed point) with velocity $\mathbf{V} + \mathbf{W}$.

Figure 13.2



Definition 13.1.
a) A **vector** represents the length and direction of a line segment. The **length** is denoted $|\mathbf{V}|$. A **unit vector U** is a vector of length 1. The **direction** of a vctor **V** is the unit vector **U** parallel to **V**: $\mathbf{U} = \mathbf{V}/|\mathbf{V}|$.
b) Given two points P, Q, the vector from P to Q is denoted $\vec{PQ}$.
c) *Addition*. The **sum**, or **resultant**, $\mathbf{V} + \mathbf{W}$ of two vectors **V** and **W** is the diagonal of the parallelogram with sides **V**,**W**.
d) *Scalar Multiplication*. To distinguish them from vectors, real numbers are called **scalars**. If c is a positve real number, $c\mathbf{V}$ is the vector with the same direction as **V** and of length $c|\mathbf{V}|$. If c negative, it is the same, but directed in the opposite direction.

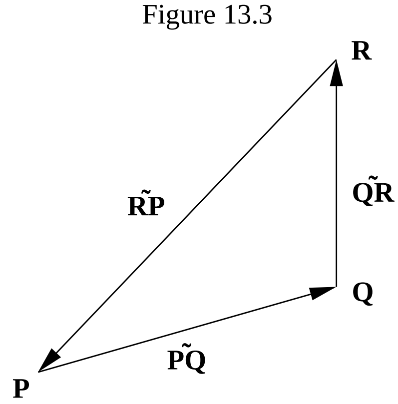
We note that the vectors **V**, $c\mathbf{V}$ are parallel, and conversely, if two vectors are parallel (that is, they have the same direction), then one is a scalar multiple of the other.

Example 13.1. Let $P,\ ,Q,\ R$ be three points in the plane not lying on a line. Then

$$\vec{PQ} + \vec{QR} + \vec{RP} = \mathbf{0}\ .$$

From figure 13.3, we see that the vector $\vec{RP}$ is the same line segment as $\vec{PQ} + \vec{QR}$, but points in the opposite direction. Thus $\vec{RP} = -(\vec{PQ} + \vec{QR})$.

Figure 13.3



The converse of this is also true: if three vectors $\mathbf{V},\ \mathbf{W},\ \mathbf{X}$ have sum $\mathbf{0}$: $\mathbf{V}+\mathbf{W}+\mathbf{X}=\mathbf{0}$, then if the are laid on the plane end to end, they close to form a triangle. For the definition of the sum of the vectors tellsl us that, since the sum is zero, the endpoint of $\mathbf{X}$ is the initial point of $\mathbf{V}$.

Example 13.2. Using vectors, show that if two triangles have corresponding sides parallel, that the lengths of corresponding sides are proportional.

Represent the sides of the two triangles by $\mathbf{U},\ ,\mathbf{V},\ \mathbf{W}$ and $\mathbf{U}',\ ,\mathbf{V}',\ \mathbf{W}'$ respectively. The hypothesis is that there are scalars $a,\ b,\ c$ such that $\mathbf{U}'=a\mathbf{U},\ \mathbf{V}'=b\mathbf{U},\ \mathbf{W}'=c\mathbf{W}$. The conclusion is that $a=b=c$. To show this, we start with the result of example 13.1; since these are the sides of a triangle, we have

$$\mathbf{U}+\mathbf{V}+\mathbf{W}=\mathbf{0}\ , \quad \mathbf{U}'+\mathbf{V}'+\mathbf{W}'=\mathbf{0}\ , \quad \text{or, what is the same,} \quad a\mathbf{U}+b\mathbf{V}+c\mathbf{W}=\mathbf{0}$$

The first equation gives us $\mathbf{U}=-\mathbf{V}-\mathbf{W}$, which, when substituted in the last equation gives

$$(b-a)\mathbf{V}+(c-a)\mathbf{W}=\mathbf{0}$$

Now, if $b\neq a$, this tells us that $\mathbf{V}$ and $\mathbf{W}$ are parallel, and so the triangle lies on a line: that is, there is no triangle. Thus we must have $b=a$, and by the same reasoning, $c=a$ also.

Problems 13.1

Give a vectorial demonstration of these geometric facts:

1. The line joining the midpoints of two sides of a triangle is parallel to the third side and half its length.

2. A median of a triangle is a line from a vertex to the midpoint of its opposing side. Show that the three medians of a triangle are themselves the sides of a triangle.

For the purposes of this and the next problem we make these definitions: a *quadrilateral* is a region enclosed by four connected line segments. a parallelogram is a quadrilateral with opposing pairs of sides parallel.

3. Show that a parallelogram has the property that opposing sides are of equal length.

4. Show that if one pair of opposing sides of a quadrilateral are parallel and of equal length, then the quadrilateral is a parallelogram.

13.2 Vectors in the Plane

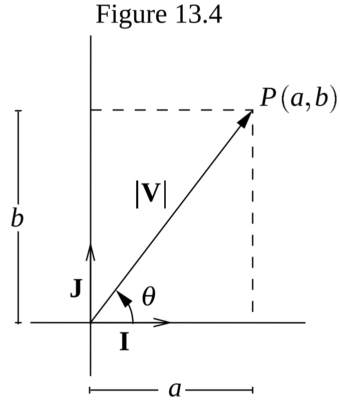
The advantage gained in using vectors is that they are moveable, and not tied to any particular coordinate system. As we have seen in the examples of the previous section, geometric facts can be easily derived using vectors; while working in coordinates may be cumbersome. However, it is often the case, that in working with vectors we must do calculations in a particular coordinate system. It is important to realize that it is the worker who gets to choose the coordinates; it is not necessarily inherent in the problem.

We now explain how to move back and forth between vectors and coordinates. Suppose, then, that a coordinate system has been chosen: a point O, the origin, and two perpendicular lines through the origin, the x- and y-axes. A vector $\mathbf{V}$ is determined by its length, $|\mathbf{V}|$ and its direction, which we can describe by the angle θ that $\mathbf{V}$ makes with the horizontal (see figure 13.4). In this figure, we have realized $\mathbf{V}$ as the vector $\vec{OP}$ from the origin to P. Let (a, b) be the cartesian coordinates of P. Note that $\mathbf{V}$ can be realized as the sum of a vector of length a along the x-axis, and a vector of length b along the y-axis. We express this as follows.

Definition 13.2. We let $\mathbf{I}$ represent the vector from the origin to the point (1,0), and $\mathbf{J}$ the vector from the origin to the point (0,1). These are the **basic** unit vectors (a unit vector is a vector of length 1). The unit vector in the direction θ is $\cos\theta\mathbf{I} + \sin\theta\mathbf{J}$.

If $\mathbf{V}$ is a vector of length r and angle θ, then $\mathbf{V} = r(\cos\theta\mathbf{I} + \sin\theta\mathbf{J})$. If $\mathbf{V}$ is the vector from the origin to the point (a, b); r is the length of $\mathbf{V}$, and $\cos\theta\mathbf{I} + \cos\theta\mathbf{J}$ is its direction. If $P(a, b)$ is the endpoint of $\mathbf{V}$, then $\mathbf{V} = \vec{OP} = a\mathbf{I} + b\mathbf{J}$. a and b are called the *components* of $\mathbf{V}$ (see figure 13.4).

Figure 13.4



Of course, r and θ are the usual polar coordinates, and we have these relations:

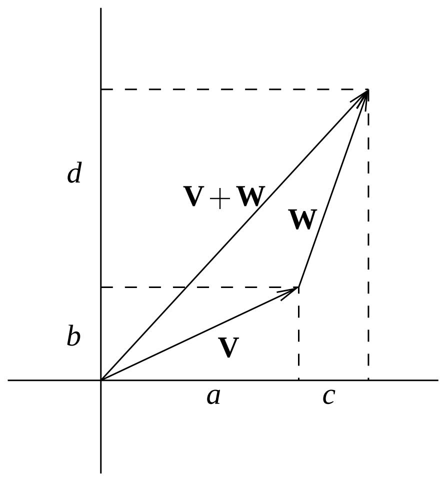
$$(13.1) \qquad |\mathbf{V}| = \sqrt{a^2 + b^2}\ , \quad \theta = \arctan\frac{b}{a}\ , \qquad a = |\mathbf{V}|\cos\theta\ , \quad b = |\mathbf{V}|\sin\theta\ .$$

We add vectors by adding their components, and multiply a vector by a scalar by multiplying the components by the scalar.

Proposition 13.1. If $\mathbf{V} = a\mathbf{I} + b\mathbf{J}$ and $\mathbf{W} = c\mathbf{I} + d\mathbf{J}$, then $\mathbf{V} + \mathbf{W} = (a + c)\mathbf{I} + (b + d)\mathbf{J}$.

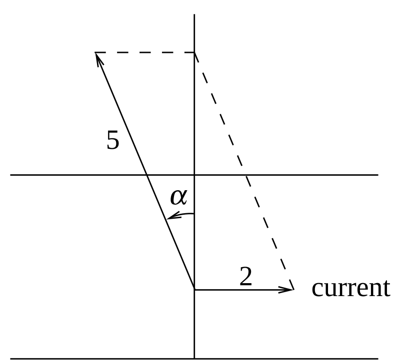
This is verified in figure 13.5.

Figure 13.5



Example 13.3. A boy can paddle a canoe at 5 mph. Suppose he wants to cross a river whose current moving at 2 mph. At what angle to the perpendicular from one bank to the other should he direct his canoe?

Figure 13.6



Draw a diagram so that the river is moving horizontally from left to right, and the direct crossing is vertical (see figure 13.6). Place the origin on the lower bank of the river, and choose the x-axis in the direction of flow, and the y-axis perpendicularly across the river. In these coordinates, the velocity vector of the current is $2\mathbf{I}$. Let $\mathbf{V}$ be the velocity vector of the canoe. We are given that $|\mathbf{V}| = 5$ and we want the sum of the two velocities to be vertical. If α is the desired angle, we see from the diagram that $\sin\alpha = 2/5$, so $\alpha = 23.5°$.

Example 13.4. An object on the plane is subject to the three forces $\mathbf{F} = 2\mathbf{I} + \mathbf{J}$, $\mathbf{G} = -8\mathbf{J}$, $\mathbf{H}$. Assuming the object doesn't move, find $\mathbf{H}$. At what angle to the horizontal is $\mathbf{H}$ directed?

By Newton's law, the sum of the forces must be zero. Thus

$$\mathbf{H} = -\mathbf{F} - \mathbf{G} = -2\mathbf{I} - \mathbf{J} + 8\mathbf{J} == -2\mathbf{I} + 7\mathbf{J} \ .$$

If α is the angle from the positive x-axis to $\mathbf{H}$, $\tan\alpha = -7/2$, so $\alpha = 105.95°$, since $\mathbf{H}$ points upward and to the left.

Since vectors represent magnitude and length, we need a computationally straightforward way of determining lengths and angles, given the components of a vector.

Definition 13.3. The **dot product** of two vectors $\mathbf{V_1}$ and $\mathbf{V_2}$ is defined by the equation

$$\mathbf{V_1} \cdot \mathbf{V_2} = |\mathbf{V_1}||\mathbf{V_2}|\cos\beta \ , \tag{13.2}$$

where β is the angle between the two vectors.

Note that since the cosine is an even function, it does not matter if we take β from $\mathbf{V_1}$ to $\mathbf{V_2}$, or in the opposite sense. In particular, we see that $\mathbf{V_1} \cdot \mathbf{V_2} = \mathbf{V_2} \cdot \mathbf{V_1}$. Now, we see how to write the dot product in terms of the components of the two vectors.

Proposition 13.2. Let $\mathbf{V_1} = a_1\mathbf{I} + b_1\mathbf{J}$ and $\mathbf{V_2} = a_2\mathbf{I} + b_2\mathbf{J}$. Then

$$\mathbf{V_1} \cdot \mathbf{V_2} = a_1a_2 + b_1b_2 \ . \tag{13.3}$$

To see this, we use the polar representation of the vectors:

$$\mathbf{V_1} = r_1(\cos\theta_1\mathbf{I} + \sin\theta_1\mathbf{J}) \ , \quad \mathbf{V_2} = r_2(\cos\theta_2\mathbf{I} + \sin\theta_2\mathbf{J}) \ .$$

Then

$$\mathbf{V_1} \cdot \mathbf{V_2} = r_1r_2\cos(\theta_1 - \theta_2) = r_1r_2\cos\theta_1\cos\theta_2 + r_1r_2\sin\theta_1\sin\theta_2$$

by the addition formula for the cosine. This is the same as

$$\mathbf{V_1} \cdot \mathbf{V_2} = (r_1\cos\theta_1)(r_2\cos\theta_2) + (r_1\sin\theta_1)(r_2\sin\theta_2)$$

which becomes (13.3) when we switch to Cartesian coordinates.

Proposition 13.3.
a) Two vectors $\mathbf{V}$ and $\mathbf{W}$ are orthogonal if and only if $\mathbf{V} \cdot \mathbf{W} = 0$.

b) If $\mathbf{L}$ and $\mathbf{M}$ are two unit vectors with $\mathbf{L}\cdot\mathbf{M}=0$, then for any vector $\mathbf{V}$, we can write

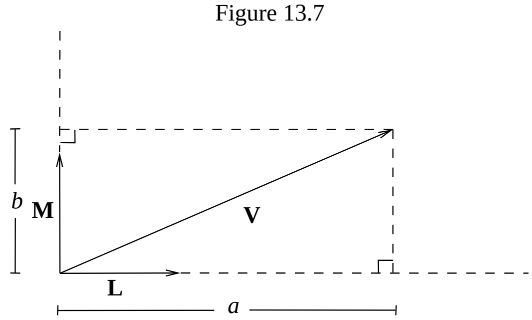
$$\mathbf{V}=a\mathbf{L}+b\mathbf{M}\ , \quad \text{with} \quad a=\mathbf{V}\cdot\mathbf{L},\ b=\mathbf{V}\cdot\mathbf{M}\ ,\ \text{and}\ |\mathbf{V}|=\sqrt{a^2+b^2}\ . \tag{13.4}$$

We shall say that a pair of unit vectors $\mathbf{L},\ \mathbf{M}$ with $\mathbf{L}\cdot\mathbf{M}=0$ form a **base** for the plane. This statement just reiterates that we can put cartesian coordinates on the plane with any point as origin and coordinate axes two orthogonal lines through the origin; for example, the lines in the directions of $\mathbf{L}$ and $\mathbf{M}$.

a) follows from the definition of the dot product, equation (13.2). For,if $\mathbf{V_1}$ and $\mathbf{V_2}$ are orthogonal, then $\beta=\pm\pi/2$, and $\mathbf{V}\cdot\mathbf{W}=0$. Conversely, if the dot product is zero, then $\beta=\pm\pi/2$, and the vectors are orthogonal.

To show part b) we start with figure 13.7.

Figure 13.7



From that figure, we see that we can write any vector as a sum $\mathbf{V}=a\mathbf{L}+b\mathbf{M}$ with (by the Pythagorean theorem) $|\mathbf{V}|=\sqrt{a^2+b^2}$. We now show that $a,\ b$ are as described;

$$\mathbf{V}\cdot\mathbf{L}=(a\mathbf{L}+b\mathbf{M})\cdot\mathbf{L}=a\mathbf{L}\cdot\mathbf{L}+b\mathbf{M}\cdot\mathbf{L}=a\ .$$

Similarly $\mathbf{V}\cdot\mathbf{M}=b$.

Example 13.5. Find the angle β between the vectors $\mathbf{V}=2\mathbf{I}-3\mathbf{J}$ and $\mathbf{W}=\mathbf{I}+2\mathbf{J}$.

We have $|\mathbf{V}|=\sqrt{2^2+3^2}=\sqrt{13}$, $|\mathbf{W}|=\sqrt{1^2+2^2}=\sqrt{5}$ and $\mathbf{V}\cdot\mathbf{W}=2(1)+(-3)(2)=-4$. Thus

$$\cos\beta=\frac{\mathbf{V}\cdot\mathbf{W}}{|\mathbf{V}||\mathbf{W}|}=\frac{-4}{\sqrt{65}}=-.496$$

so $\beta=-119.7°$.

Example 13.6. Suppose we have put cartesian coordinates on the plane, with $\mathbf{I},\ \mathbf{J}$ the standard base. Let

$$\mathbf{L}=\frac{\mathbf{I}+\mathbf{J}}{\sqrt{2}}\ , \quad \mathbf{M}=\frac{-\mathbf{I}+\mathbf{J}}{\sqrt{2}}$$

be a new base. Given the point $P(5,2)$, write $\vec{OP}$ in terms of $\mathbf{L}$ and $\mathbf{M}$.

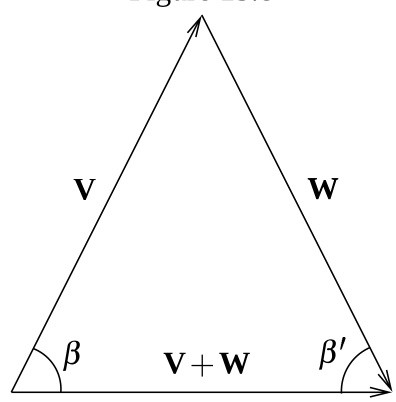
By the preceding proposition,

$$\vec{OP} \cdot \mathbf{L} = (5\mathbf{I} + 2\mathbf{J}) \cdot \left(\frac{\mathbf{I} + \mathbf{J}}{\sqrt{2}}\right) = \frac{7}{\sqrt{2}} \ , \quad \vec{OP} \cdot \mathbf{M} = (5\mathbf{I} + 2\mathbf{J}) \cdot \left(\frac{-\mathbf{I} + \mathbf{J}}{\sqrt{2}}\right) = -\frac{3}{\sqrt{2}} \ ,$$

so $\vec{OP} = (7\mathbf{L} - 3\mathbf{M})/\sqrt{2}$.

Example 13.7. Show, using vectors, that the interior angles of an isosceles triangle are equal.

Figure 13.8



In figure 13. 8 we have labelled the sides of equal length as $\mathbf{V}$ and $\mathbf{W}$. Thus, the base of the triangle is $\mathbf{V} + \mathbf{W}$. First of all, since $|\mathbf{V}| = |\mathbf{W}|$, we have $(\mathbf{V} + \mathbf{W}) \cdot \mathbf{V} = \mathbf{V} \cdot \mathbf{V} + \mathbf{W} \cdot \mathbf{V} = \mathbf{W} \cdot \mathbf{W} + \mathbf{V} \cdot \mathbf{W} = (\mathbf{V} + \mathbf{W}) \cdot \mathbf{W}$. Thus, by (13.2),

$$\cos\beta = \frac{(\mathbf{V} + \mathbf{W}) \cdot \mathbf{V}}{|\mathbf{V} + \mathbf{W}||\mathbf{V}|} = \frac{(\mathbf{V} + \mathbf{W}) \cdot \mathbf{W}}{|\mathbf{V} + \mathbf{W}||\mathbf{W}|} = \cos\beta' \ .$$

Since both angles are acute, $\beta = \beta'$.

Example 13.8. Find a vector orthogonal to $\mathbf{V} = 3\mathbf{I} + 4\mathbf{J}$ and of the same length.

The vectors $\mathbf{V} = a\mathbf{I} + b\mathbf{J}$, $\mathbf{W} = c\mathbf{I} + d\mathbf{J}$, are orthogonal precisely when $ac + bd = 0$. Thus, if we are given a, b, we take $c = -b$, $d = a$ to get an orthogonal vector. So for this example, we can take $\mathbf{W} = -4\mathbf{I} + 3\mathbf{J}$. Clearly, since the coefficients are the same but for sign, $|\mathbf{W}| = |\mathbf{V}|$. We could also take the vector in the opposite direction: $-\mathbf{W} = 4\mathbf{I} - 3\mathbf{J}$

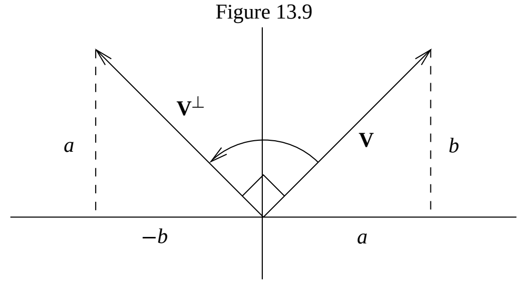
In general, if $\mathbf{V} = c\mathbf{I} + d\mathbf{J}$ then both $-d\mathbf{I} + c\mathbf{J}$ and $d\mathbf{I} - c\mathbf{J}$ are orthogonal to $\mathbf{V}$ and of the same length. The first is counterclockwise to $\mathbf{V}$, and the second, clockwise.

Definition 13.3. Given the vector $\mathbf{V}$, we shall denote by $\mathbf{V}^\perp$ that vector which is orthogonal to, of the same length as, and *counterclockwise* to $\mathbf{V}$. In components, we have:

(13.5) $$\text{If} \quad \mathbf{V} = a\mathbf{I} + b\mathbf{J} \ , \quad \text{then} \quad \mathbf{V}^\perp = -b\mathbf{I} + a\mathbf{J}$$

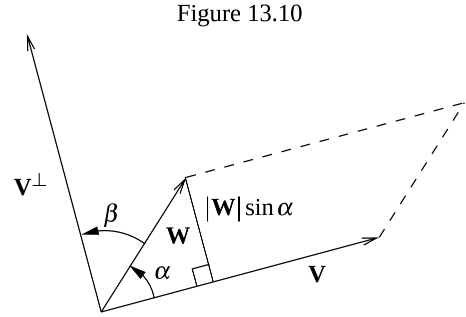
See figure 13.9 to see that $\mathbf{V}^\perp$ is counterclockwise to $\mathbf{V}$ (at least in the case where both a and b are positive).

Figure 13.9



Definition 13.4. Given two vectors $\mathbf{V}$ and $\mathbf{W}$, we define the **determinant** $\det(\mathbf{V},\mathbf{W})$ of the two vectors as the signed area of the parallelogram spanned by the two vectors. The sign is positive if $\mathbf{W}$ is counterclockwise from $\mathbf{V}$; otherwise negative.

Figure 13.10



In figure 13.10, α is the angle from $\mathbf{V}$ to $\mathbf{W}$. Thus

$$\det(\mathbf{V},\mathbf{W}) = |\mathbf{V}||\mathbf{W}|\sin\alpha\ . \tag{13.6}$$

Now, let β be the angle from $\mathbf{W}$ to $\mathbf{V}^\perp$ so that (in figure 13.10), $\alpha+\beta=\pi/2$, and we have $\sin\alpha=\cos\beta$. Since $|\mathbf{V}|=|\mathbf{V}^\perp|$, we can rewrite (13.6) as

$$\det(\mathbf{V},\mathbf{W}) = |\mathbf{V}^\perp||\mathbf{W}|\cos\beta = \mathbf{V}^\perp\cdot\mathbf{W} \tag{13.7}$$

This gives us the following.

Proposition 13.4. The determinant of the two vectors $\mathbf{V} = a\mathbf{I} + b\mathbf{J}$ and $\mathbf{W} = c\mathbf{I} + d\mathbf{J}$ is the determinant of the matrix whose rows are the vectors $\mathbf{V}$ and $\mathbf{W}$:

$$\det(\mathbf{V}, \mathbf{W}) = ad - bc \ . \tag{13.8}$$

For, $\mathbf{V}^\perp = -b\mathbf{I} + a\mathbf{J}$, and from (13.7), $\det(\mathbf{V}, \mathbf{W}) = \mathbf{V} \cdot \mathbf{W}^\perp = -bc + ad = ad - bc$.

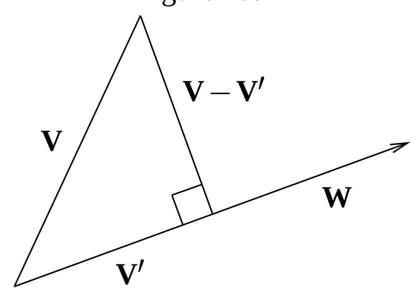
The vectors $\mathbf{V}$ and $\mathbf{W}$are parallel (or collinear) if and only if $\det(\mathbf{V}, \mathbf{W}) = 0$, for in this case there is no parallelogram. We also have the inequality

$$|\det(\mathbf{V}, \mathbf{W})| \leq |\mathbf{V}||\mathbf{W}| \ ,$$

with equality holding if and only if $\mathbf{V}$ and $\mathbf{W}$ are orthogonal.

Definition 13.5. Given two vectors $\mathbf{V}$ and $\mathbf{W}$, the **projection** of $\mathbf{V}$ in the direction of $\mathbf{W}$ is that vector $\mathbf{V}'$ parallel to $\mathbf{W}$ such that $\mathbf{V} - \mathbf{V}'$ is orthogonal to $\mathbf{V}'$ (see figure 13.11).

Figure 13.11



Proposition 13.5 The projection $\mathbf{V}'$ of $\mathbf{V}$ in the direction of $\mathbf{W}$ is given by the formula

$$\mathbf{V}' = pr_{\mathbf{W}}(\mathbf{V}) = \frac{\mathbf{V} \cdot \mathbf{W}}{\mathbf{W} \cdot \mathbf{W}}\mathbf{W} \ . \tag{13.9}$$

If $\mathbf{U}$ is a unit vector in the direction of $\mathbf{W}$, then

$$\mathbf{V}' = (\mathbf{V} \cdot \mathbf{U})\mathbf{U} \ , \quad \text{and} \quad \mathbf{V} = (\mathbf{V} \cdot \mathbf{U})\mathbf{U} + (\mathbf{V} \cdot \mathbf{U}^\perp)\mathbf{U}^\perp \ .$$

To show this we start with the equation $(\mathbf{V} - \mathbf{V}') \cdot \mathbf{V}' = \mathbf{0}$. Since $\mathbf{V}' = a\mathbf{W}$ for some a, this gives us

$$(\mathbf{V} - a\mathbf{W}) \cdot a\mathbf{W} = 0 \ , \quad \text{or} \quad a^2\mathbf{W} \cdot \mathbf{W} = a\mathbf{V} \cdot \mathbf{W}$$

If $a = 0$, then $\mathbf{V}' = \mathbf{0}$ and $\mathbf{V}$ and $\mathbf{W}$ are orthogonal. Otherwise

$$a = \frac{\mathbf{V} \cdot \mathbf{W}}{\mathbf{W} \cdot \mathbf{W}} \ ,$$

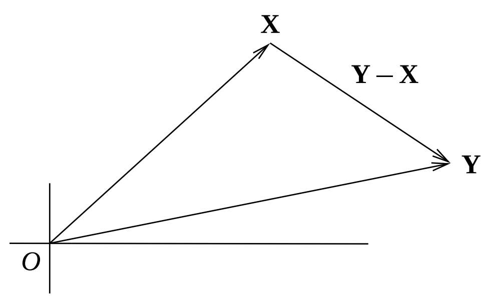
giving us (13.9). The rest of the proposition follows by replacing $\mathbf{W}$ by the unit vector $\mathbf{U}$, and should be viewed as a restatement of Proposition 13.6.

Example 13.9. Find the area of the parallelogram whose vertices are at $O(0,0), P(4,-2)$, $Q(5,8), R(9,6)$.

This is the parallelogram determined by the vectors from the origin O to the points P and Q: $\vec{OP} = 4\mathbf{I} - 2\mathbf{J},\ \vec{OQ} = 5\mathbf{I} - 8\mathbf{J}$, so has signed area $4(-8) - (-2)(5) = -22$. We verify these are the vertices of a parallelogram by calculating $\vec{OP} + \vec{OQ} = 9\mathbf{I} + 6\mathbf{J} = \vec{OR}$.

In order to discuss geometric objects in the coordinate plane, it is useful to represent a point $X(x,y)$ by the vector $\mathbf{X} = \vec{OX} = x\mathbf{I} + y\mathbf{J}$ from the origin to X. For Y another point, the vector from X to Y is thus represented by $\mathbf{Y} - \mathbf{X}$ (see figure 13.12).

Figure 13.12



A line L is determined by its direction and a point on the line. let $\mathbf{X_0}$ be a point on L, and $\mathbf{L}$ a vector parallel to the line L. Then, for any point $\mathbf{X}$, it is on the line if and only if $\mathbf{X} - \mathbf{X_0}$ is parallel to $\mathbf{L}$, or, what is the same, orthogonal to $\mathbf{L}^\perp$. This leads to these two equations, called the equation of the line:

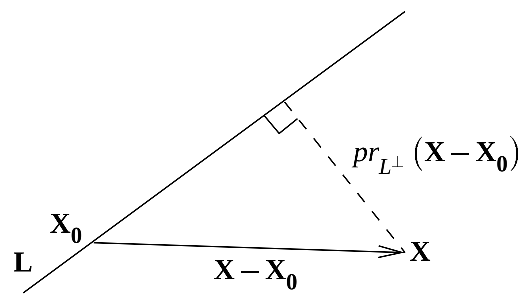
$$(\mathbf{X} - \mathbf{X_0}) \cdot \mathbf{L}^\perp = 0 \qquad \text{or} \qquad \det(\mathbf{X} - \mathbf{X_0}, \mathbf{L}) = 0\ . \tag{13.10}$$

Also, since $\mathbf{X} - \mathbf{X_0}$ is parallel to L if and only if $\mathbf{X} - \mathbf{X_0}$ is a scalar multiple of $\mathbf{L}$, we have the *parametric form* of the equation of the line:

$$L: \qquad \mathbf{X} = \mathbf{X_0} + t\mathbf{L}\ .$$

A line is also determined by two points $\mathbf{X_0},\ \mathbf{X_1}$ on the line. Given that information, we find the equations of the line by taking $\mathbf{L} = \mathbf{X_1} - \mathbf{X_0}$.

Figure 13.13



Now, suppose L is a line and $\mathbf{X}$ is a point not on the line. We seek a formula for the distance from the point $\mathbf{X}$ to the line. We see from figure 13.13 that this is the length of the projection in the direction perpendicular to L of a vector from $\mathbf{X}$ to any point $\mathbf{X_0}$ on L. This leads to the formula for the distance from $\mathbf{X}$ to L

$$d(\mathbf{X}, L) = |pr_{\mathbf{L}^\perp}(\mathbf{X} - \mathbf{X_0})| \tag{13.11}$$

Example 13.10. Let L be the line given by the equation $3x - y = 7$. Find the distance from (2,4) to L.

First of all, rewrite the equation defining the line vectorially. Let $\mathbf{X} = x\mathbf{I} + y\mathbf{J}$, so that the line is given as the set of points $\mathbf{X}$ satisfying $\mathbf{X}\cdot(3\mathbf{I} - \mathbf{J}) = 7$. If $\mathbf{X_0}$ is another point on the line we have $\mathbf{X_0}\cdot(3\mathbf{I} - \mathbf{J}) = 7$ as well, so

$$(\mathbf{X} - \mathbf{X_0})\cdot(3\mathbf{I} - \mathbf{J}) = 0 \ .$$

Comparing this with equation (13.10) we see that $\mathbf{L}^\perp = 3\mathbf{I} - \mathbf{J}$. To use (13.11) we need a point on the line; any solution of the equation $3x - y = 7$ will do. (3,2) is a solution, so we take $\mathbf{X_0} = 3\mathbf{I} + 2\mathbf{J}$. Thus, for our point, $\mathbf{X} = 2\mathbf{I} + 4\mathbf{J}$, the distance is

$$|pr_{\mathbf{L}^\perp}(\mathbf{X} - \mathbf{X_0})| = \frac{|(\mathbf{X} - \mathbf{X_0}) \cdot \mathbf{L}^\perp|}{|\mathbf{L}^\perp|} = \frac{|(-\mathbf{I} + 2\mathbf{J}) \cdot (3\mathbf{I} - \mathbf{J})|}{|3\mathbf{I} - \mathbf{J}|} = \frac{5}{\sqrt{10}}$$

Example 13.11. Find the distance from X(3,1) to the line through X_0(2,-3) and parallel to $\mathbf{V} = -\mathbf{I} + 4\mathbf{J}$.

The vector $\mathbf{L}^\perp = -4\mathbf{I} - \mathbf{J}$ is orthogonal to the line. The line segment from X_0 to X is given by the vector $(3 - 2)\mathbf{I} + (1 - (-3))\mathbf{J} = \mathbf{I} + 4\mathbf{J}$. Thus the distance is

$$|pr_{\mathbf{L}^\perp}(\mathbf{X} - \mathbf{X_0})| = \frac{|(\mathbf{I} + 4\mathbf{J}) \cdot (-4\mathbf{I} - \mathbf{J})|}{| - 4\mathbf{I} - \mathbf{J}|} = \frac{8}{\sqrt{17}}$$

Example 13.12. Find the point on the line $L:\ 2x - 3y = 17$ which is closest to the origin.

Let $\mathbf{X}$ be the vector from the origin to the desired point. Then $\mathbf{X}$ is orthogonal to the line, so is parallel to the vector $\mathbf{L}^\perp = 2\mathbf{I} - 3\mathbf{J}$. Writing $\mathbf{X} = t(2\mathbf{I} - 3\mathbf{J})$, since $\mathbf{X}$ ends on the line we have $\mathbf{X} \cdot \mathbf{L} = 2(2t) - 3(-3t) = 17$, so $t = 17/13$, and $\mathbf{X} = (34/13)\mathbf{I} - (51/13)\mathbf{J}$.

Problems 13.2

1. Find the components of a vector $\mathbf{V}$ making an angle of $45°$ with $\mathbf{W} = 2\mathbf{I} - \mathbf{J}$.

2. Show that the diagonals of a rhombus intersect at right angles. (A *rhombus* is a parallelogram with all sides of the same length).

3. A river ferry runs at a speed of 6 knots, across a river with a current of 2 knots. Assume that the river has shores which are parallel straight lines. In what direction should the barge head in order to cross the river perpendicular to the shores?

4. Find the area of the parallelogram determined by the vectors $\mathbf{V} = 6\mathbf{I} - 7\mathbf{J},\ \mathbf{W} = 3\mathbf{I} + 4\mathbf{J}$. What are the coordinates of the vertex of this parallelogram farthest from the origin?

5. A plane flies at a ground speed of 480 mph. The jet stream comes from $30°$ north of west at 60 mph. In what direction should the pilot direct the plane so as to be heading directly east? How many miles will the plane cover in 2 hours?

6. Let $\mathbf{V} = 2\mathbf{I} - 3\mathbf{J}$. Find the vector $\mathbf{W}$ to the left of $\mathbf{V}$, of the same length as $\mathbf{V}$, and making an angle of $60°$ with $\mathbf{V}$.

7. Find a base $\mathbf{L},\ \mathbf{M}$ (in a base, the vectors are of length 1 and orthogonal) so that $\mathbf{L}$ is on the line $y = 5x$ and $\mathbf{M}$ is left of $\mathbf{L}$. Any vector $\mathbf{X} = x\mathbf{I} + y\mathbf{J}$ can be written in the form $\mathbf{X} = u\mathbf{L} + v\mathbf{M}$. Find $u,\ v$ as functions of x, y.

8. What is the distance between the two parallel lines $L_1 : x + 2y = 7,\ L_2 : x + 2y = 11$?

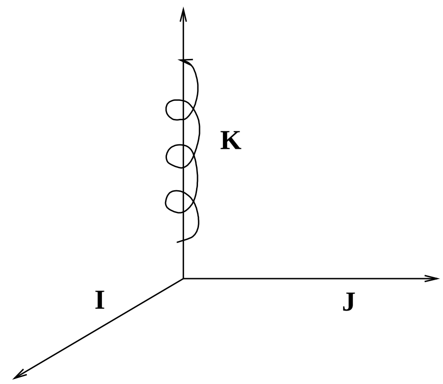
9. Find the distance from the point P(3,-2) to the line $L : 2x - y = 10$.

10. Find a point (x, y) on the line $2x - y = 10$ such that the triangle with vertices (0,0), (5,0), (x, y) has area equal to 15. How many such points are there?

13.3 Vectors in Space

In a Cartesian coordinate system for space, the vectors $\mathbf{I}$, $\mathbf{J}$, $\mathbf{K}$ are the vectors from the origin to the points $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ respectively. These are unit vectors, mutually orthogonal, and form the *standard base* for space. We always take a coordinatization so that $\{\mathbf{I},\ \mathbf{J},\ \mathbf{K}\}$ is a right-handed system. More precisely, if we situate $\mathbf{I}$ and $\mathbf{J}$ on the horizontal plane, then $\mathbf{I}$ is a unit vector, $\mathbf{J}$ is a unit vector perpendicular to $\mathbf{I}$ and counterclockwise from $\mathbf{I}$, and $\mathbf{K}$ is a unit vector orthogonal to the horizontal plane, pointing upwards (see figure 13.14).

Figure 13.14



Any vector $\mathbf{V}$ can be written uniquely as

$$\mathbf{V} = a\mathbf{I} + b\mathbf{J} + c\mathbf{K}, \tag{13.12}$$

where $a,\ b,\ c$ are called the **components** of $\mathbf{V}$. To add two vectors, add the components; to multiply a vector by a scalar, multiply the components by the scalar.

If $\mathbf{V}$ is given as in (12), its **length** is

$$|\mathbf{V}| = \sqrt{a^2 + b^2 + c^2}. \tag{13.13}$$

The *direction* of $\mathbf{V}$ is determined by the cosines of the angles between $\mathbf{V}$ and the coordinate axes. Thus, for any vector $\mathbf{V}$ we can write

$$\mathbf{V} = |\mathbf{V}|(\cos\alpha\mathbf{I} + \cos\beta\mathbf{J} + \cos\gamma\mathbf{K}) \tag{13.14}$$

where $\alpha,\ \beta,\ \gamma$ are those angles. The components of the unit vector in (13.14) are called the *direction cosines* of the vector $\mathbf{V}$. Note that, by (13.13), $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

Definition 13.6. The **dot product** of two vectors $\mathbf{V},\ \mathbf{W}$ is defined as

$$\mathbf{V}\cdot\mathbf{W} = |\mathbf{V}||\mathbf{W}|\cos\theta, \tag{13.15}$$

where θ is the angle between $\mathbf{V}$ and $\mathbf{W}$.

As for plane vectors, this has an easy formulation in terms of the components of the vectors.

Proposition 13.6. Let

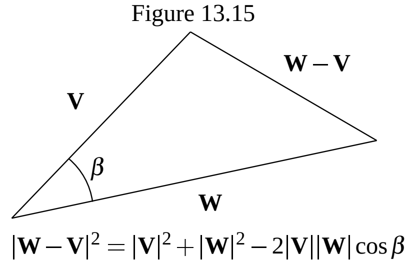
$$\mathbf{V} = a_1\mathbf{I} + b_1\mathbf{J} + c_1\mathbf{K}, \quad \mathbf{W} = a_2\mathbf{I} + b_2\mathbf{J} + c_2\mathbf{K} \tag{13.16}$$

in components. Then

$$\mathbf{V}\cdot\mathbf{W} = a_1a_2 + b_1b_2 + c_1c_2 \tag{13.17}$$

To see this, we start with the Law of Cosines for the triangle whose sides are the vectors $\mathbf{V},\ \mathbf{W},\ \mathbf{W}-\mathbf{V}$ (see figure 13.15):

Figure 13.15



$$|\mathbf{W}-\mathbf{V}|^2 = |\mathbf{W}|^2 + |\mathbf{V}|^2 - 2|\mathbf{V}||\mathbf{W}|\cos\beta = |\mathbf{W}|^2 + |\mathbf{V}|^2 - 2(\mathbf{W}\cdot\mathbf{V})\ ,$$

so that

$$\mathbf{W}\cdot\mathbf{V} = \frac{1}{2}(|\mathbf{W}|^2 + |\mathbf{V}|^2 - |\mathbf{W}-\mathbf{V}|^2)\ .$$

Now, writing the right hand side in terms of components, using (13.12) and (13.13), we get (13.17), after some cancellation.

In particular, just as in two dimensions, two vectors $\mathbf{V},\ \mathbf{W}$ are *orthogonal* if $\mathbf{V}\cdot\mathbf{W} = 0$.

Example 13.13. Find the angle between the vectors $\mathbf{V} = 2\mathbf{I} - 3\mathbf{J} + \mathbf{K},\ \mathbf{W} = 6\mathbf{I} + \mathbf{J} - 2\mathbf{K}$.

We have $\mathbf{V}\cdot\mathbf{W} = 12 - 3 - 2 = 7$ and $|\mathbf{V}| = \sqrt{2^2+3^2+1^2} = 3.74,\ |\mathbf{W}| = \sqrt{6^2+1^2+2^2} = 6.40$. Thus

$$\cos\alpha = \frac{7}{(3.74)(6.40)} = .2923$$

so $\alpha = 73°$.

Example 13.14. Find a vector orthogonal to both the vectors $\mathbf{V}$ and $\mathbf{W}$ of example 13.13.

Let $\mathbf{X} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$ be the desired vector. We have the conditions

$$\mathbf{X}\cdot\mathbf{V} = 2x - 3y + z = 0\ , \qquad \mathbf{X}\cdot\mathbf{W} = 6x + y - 2z = 0\ .$$

We can solve these equations by replacing z by any nonzero value, say $z = 1$, and solving the resulting equations for x and y:

$$2x - 3y + 1 = 0\ , \qquad 6x + y - 2 = 0\ .$$

These have the solution $x = 1/4,\ y = 1/2$. Thus we can take

$$\mathbf{X_0} = \frac{1}{4}\mathbf{I} + \frac{1}{2}\mathbf{J} + \mathbf{K}$$

as our answer. Of course there is a line of such vectors, corresponding to all possible values for z. Thus the set of all vectors orthogonal to $\mathbf{V}$ and $\mathbf{W}$ is the set $\{t\mathbf{X_0}\}$.

Given vectors $\mathbf{V}$ and $\mathbf{W}$, the **projection** of $\mathbf{V}$ (denoted $pr_{\mathbf{W}}(\mathbf{V})$) in the direction of $\mathbf{W}$ is the vector $\mathbf{V}'$ parallel to $\mathbf{W}$ such that $\mathbf{V}'$ and $\mathbf{V} - \mathbf{V}'$ are orthogonal. If β is the angle between $\mathbf{V}$ and $\mathbf{W}$, this projection is the vector of length $|\mathbf{V}|\cos\beta$ in the direction of $\mathbf{W}$. The formula for the projection is (as in the plane):

$$pr_{\mathbf{W}}(\mathbf{V}) = (\frac{\mathbf{V}\cdot\mathbf{W}}{\mathbf{W}\cdot\mathbf{W}})\mathbf{W}\ . \tag{13.18}$$

Again, just as in the plane, if $\mathbf{U}$ is the unit vector in the direction of $\mathbf{W}$, then $pr_{\mathbf{W}}(\mathbf{V}) = (\mathbf{V}\cdot\mathbf{U})\mathbf{U}$. We note that for two vectors $\mathbf{V_1},\ \mathbf{V_2}$,

$$pr_{\mathbf{W}}(\mathbf{V_1}+\mathbf{V_2}) = pr_{\mathbf{W}}(\mathbf{V_1}) + pr_{\mathbf{W}}(\mathbf{V_2})\ .$$

Definition 13.7. The **cross product** of two vectors $\mathbf{V},\ \mathbf{W}$, denoted $\mathbf{V}\times\mathbf{W}$, is that vector

a) of length the area of the parallelogram spanned by $\mathbf{V},\ \mathbf{W}$,

b) perpendicular to the plane of $\mathbf{V},\ \mathbf{W}$ so that the system $\{\mathbf{V},\ \mathbf{W},\ \mathbf{V}\times\mathbf{W}\}$ is right-handed.

Now, since the area of the parallelogram spanned by the vectors $\mathbf{V},\ \mathbf{W}$ is $|\mathbf{V}||\mathbf{W}|\sin\beta$, where β is the angle between the two vectors, we have

$$|\mathbf{V}\times\mathbf{W}|^2 = |\mathbf{V}|^2|\mathbf{W}|^2 - (\mathbf{V}\cdot\mathbf{W})^2 \tag{13.19}$$

since

$$|\mathbf{V}\times\mathbf{W}|^2 = |\mathbf{V}|^2|\mathbf{W}|^2\sin^2\beta = |\mathbf{V}|^2|\mathbf{W}|^2(1-\cos^2\beta) = |\mathbf{V}|^2|\mathbf{W}|^2 - (|\mathbf{V_1}||\mathbf{V_2}|\cos\beta)^2$$

which is the right side of (13.19), from (13.15).

Note that interchanging $\mathbf{V}$ and $\mathbf{W}$ changes the sign of the cross product, for if the system $\{\mathbf{V},\ \mathbf{W},\ \mathbf{L}\}$ is right-handed, then the system $\{\mathbf{W},\ \mathbf{V},\ \mathbf{L}\}$ is left-handed, and thus $\{\mathbf{W},\ \mathbf{V},\ -\mathbf{L}\}$ is right handed. This gives us the first of the following identities:

$$\mathbf{V}\times\mathbf{W} = -\mathbf{W}\times\mathbf{V} \tag{13.20}$$

$$\mathbf{V}\times\mathbf{V} = \mathbf{0},$$

$$(a\mathbf{V})\times\mathbf{W}) = a(\mathbf{V}\times\mathbf{W})\ .$$

The second and third identities are clear from geometry. We now determine a formula for the cross product in components. It is useful to start with the determinant of three vectors in space, sometimes called the *triple scalar product.*

Definition 13.8. Given three vectors in space $\mathbf{U},\ \mathbf{V},\ \mathbf{W}$, we define the **determinant**, denoted $\det(\mathbf{U},\mathbf{V},\mathbf{W})$ as the signed volume of the parallelepiped spanned by the vectors. This is zero if the vectors all lie in the same plane. Otherwise, the sign is positive if the vectors $\{\mathbf{U},\mathbf{V},\mathbf{W}\}$ form a right-handed system, and negative if a left-handed system.

Proposition 13.7. Given two vectors $\mathbf{V}$, $\mathbf{W}$, then, for any third vector $\mathbf{U}$,

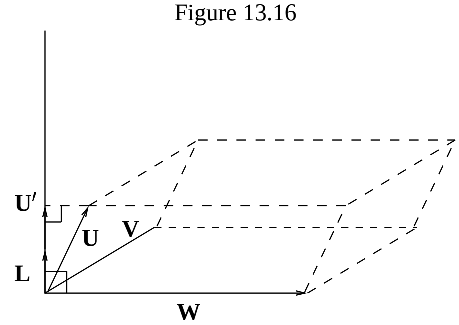
$$\det(\mathbf{U},\mathbf{V},\mathbf{W}) = \mathbf{U}\cdot(\mathbf{V}\times\mathbf{W}) \tag{13.21}$$

For any two vectors $\mathbf{U_1}$, $\mathbf{U_2}$

$$\det(\mathbf{U_1}+\mathbf{U_2},\mathbf{V},\mathbf{W}) = \det(\mathbf{U_1},\mathbf{V},\mathbf{W}) + \det(\mathbf{U_2},\mathbf{V},\mathbf{W}) \tag{13.22}$$

We now show (13.21) using a geometric argument similar to that used for proposition 13.4. If $\mathbf{V}$ and $\mathbf{W}$ lie on a line, then all terms are zero, and there is nothing to show. Otherwise, $\mathbf{V}$ and $\mathbf{W}$ determine a plane; let $\mathbf{L}$ be the unit vector orthogonal to that plane so that the triple $\mathbf{V},\mathbf{W},\mathbf{L}$ is right-handed. For any vector $\mathbf{U}$, let $\mathbf{U}'$ be the projection of $\mathbf{U}$ in the direction of $\mathbf{L}$. Then, we see geometrically that the volume of the parallelepiped spanned by $\mathbf{U},\mathbf{V},\mathbf{W}$ is the product of the area of the parallelogram spanned by $\mathbf{V},\mathbf{W}$ and the length of $\mathbf{U}'$ (see figure 13.16).

Figure 13.16



Since $\mathbf{V}\times\mathbf{W}$ has the same direction as $\mathbf{L}$, this volume is

$$|\mathbf{U}'||\mathbf{V}\times\mathbf{W}| = |\mathbf{U}||\mathbf{V}\times\mathbf{W}|\cos\beta = \mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})$$

where β is the angle between $\mathbf{U}$ and $\mathbf{L}$. The signs are right in (13.21), for on both sides they are determined by whether or not the system $\mathbf{U},\mathbf{V},\mathbf{W}$ is right-handed. (13.22) now follows directly from (13.21), since the right hand side is linear in $\mathbf{U}$:

$$\det(\mathbf{U_1}+\mathbf{U_2},\mathbf{V},\mathbf{W}) = (\mathbf{U_1}+\mathbf{U_2})\cdot(\mathbf{V}\times\mathbf{W})$$

$$= \mathbf{U_1}\cdot(\mathbf{V}\times\mathbf{W}) + \mathbf{U_2}\cdot(\mathbf{V}\times\mathbf{W}) = \det(\mathbf{U_1},\mathbf{V},\mathbf{W}) + \det(\mathbf{U_2},\mathbf{V},\mathbf{W})$$

Now, if we permute the three vectors $\mathbf{U}$. $\mathbf{V}$, $\mathbf{W}$, we just change the sign of the determinant, since it is always the parallelepiped spanned by the same vectors:

$$\det(\mathbf{U},\mathbf{V},\mathbf{W}) = -\det(\mathbf{V},\mathbf{U},\mathbf{W}) = \det(\mathbf{V},\mathbf{W},\mathbf{U})\ .$$

So, since, but for sign, we can move any of the vectors in $\det(\mathbf{U},\mathbf{V},\mathbf{W})$ to the first position, we conclude that the determinant is linear in all three variables. In particular, the cross product is linear in its variables. This allows us to calculate the determinant and cross product from the components of the given vectors. We first observe that the calculations for the basis vectors are immediate, since the area of the unit square is 1:

$$\mathbf{I}\times\mathbf{J}=\mathbf{K},\ \mathbf{J}\times\mathbf{K}=\mathbf{I},\ \mathbf{K}\times\mathbf{I}=\mathbf{J},\ \mathbf{I}\times\mathbf{I}=\mathbf{J}\times\mathbf{J}=\mathbf{K}\times\mathbf{K}=\mathbf{0}$$

Finally, from (13.21) $\mathbf{J}\times\mathbf{I}=-\mathbf{I}\times\mathbf{J}=-\mathbf{K}$, etc. After a long computation, we find:

Proposition 13.8. if $\mathbf{V_1}=a_1\mathbf{I}+b_1\mathbf{J}+c_1\mathbf{K},\ \mathbf{V_2}=a_2\mathbf{I}+b_2\mathbf{J}+c_2\mathbf{K}$, then

$$\mathbf{V_1}\times\mathbf{V_2}=(b_1c_2-c_1b_2)\mathbf{I}+(c_1a_2-a_1c_2)\mathbf{J}+(a_1b_2-b_1a_2)\mathbf{K}. \tag{13.23}$$

Now we see that the determinant of three vectors, or, what is the same, the triple scalar product: $\mathbf{V_1}\cdot(\mathbf{V_2}\times\mathbf{V_3})=(\mathbf{V_1}\times\mathbf{V_2})\cdot\mathbf{V_3}$ is, in fact, the determinant of the matrix whose rows are the components of the vectors $\mathbf{V_1},\mathbf{V_2},\mathbf{V_3}$, just by taking the dot product of $\mathbf{V_3}$ with the expression (13.23) for $\mathbf{V_1}\times\mathbf{V_2}$:

Proposition 13.9. If, in addition, $\mathbf{V_3}=a_3\mathbf{I}+b_3\mathbf{J}+c_3\mathbf{K}$, then

$$\det(\mathbf{V_1},\mathbf{V_2},\mathbf{V_3})=(\mathbf{V_1}\times\mathbf{V_2})\cdot\mathbf{V_3}=a_3(b_1c_2-c_1b_2)+b_3(c_1a_2-a_1c_2)+c_3(a_1b_2-b_1a_2)\ .$$

This is just the expansion of the determinant by minors of the third row. An easy way to remember the formula for the cross product is as this determinant:

$$\mathbf{V_1}\times\mathbf{V_2}=\begin{pmatrix}\mathbf{I}&\mathbf{J}&\mathbf{K}\\a_1&b_1&c_1\\a_2&b_2&c_2\end{pmatrix} \tag{13.24}$$

Example 13.15. Find $\mathbf{V_1}\cdot(\mathbf{V_2}\times\mathbf{V_3})$ where

$$\mathbf{V_1}=-\mathbf{I}+2\mathbf{J}+\mathbf{K}.\quad \mathbf{V_2}=2\mathbf{I}-2\mathbf{J}+3\mathbf{K},\quad \mathbf{V_3}=\mathbf{I}-2\mathbf{K}\ .$$

By proposition 13.9, this is the determinant

$$\mathbf{V_1}\cdot(\mathbf{V_2}\times\mathbf{V_3})=\det\begin{pmatrix}-1&2&1\\2&-2&3\\1&0&-2\end{pmatrix}=1(6+2)+0+(-2)(2-4)=12$$

where we calculate by minors of the third row.

Example 13.16. Find a vector $\mathbf{W}$ of length $|\mathbf{W}|=5$ which is orthogonal to both $\mathbf{V_1}$ and $\mathbf{V_2}$, so that the system $\{\mathbf{V_1},\ \mathbf{V_2},\ \mathbf{W}\}$ is right-handed.

$\mathbf{W}$ is a positive multiple of $\mathbf{V_1}\times\mathbf{V_2}$, which is

$$\mathbf{V_1}\times\mathbf{V_2}=\begin{pmatrix}\mathbf{I}&\mathbf{J}&\mathbf{K}\\-1&2&1\\2&-2&3\end{pmatrix}=8\mathbf{I}+5\mathbf{J}-2\mathbf{K}\ .$$

This vector has length $\sqrt{64+25+4} = \sqrt{93}$, so

$$\mathbf{W} = \frac{5}{\sqrt{93}}(8\mathbf{I} + 5\mathbf{J} - 2\mathbf{K}) \ .$$

Problems 13.3

1. Find a unit vector orthogonal to the vectors $\mathbf{V} = 6\mathbf{I} - 7\mathbf{J} + \mathbf{K},\ \mathbf{W} = -\mathbf{I} + 2\mathbf{J} - 3\mathbf{K}$. What is the volume of the parallelopiped determined by these three vectors?

2. Find a vector which makes an angle of 30° with the plane determined by the vectors $\mathbf{V} = 6\mathbf{I} - 7\mathbf{J} + \mathbf{K},\ \mathbf{W} = -\mathbf{I} + 2\mathbf{J} - 3\mathbf{K}$.

3. Given $\mathbf{V} = 2\mathbf{I} - \mathbf{J} + 3\mathbf{K},\ \mathbf{W} = -\mathbf{I} + 4\mathbf{J} - 2\mathbf{K}$, find the angle between $\mathbf{V}$ and $\mathbf{W}$.

4. Find $\mathbf{V_1} \cdot (\mathbf{V_2} \times \mathbf{V_3})$ where

$$\mathbf{V_1} = -\mathbf{I} + 2\mathbf{J} + \mathbf{K}. \quad \mathbf{V_2} = 2\mathbf{I} - 2\mathbf{J} + 3\mathbf{K}, \quad \mathbf{V_3} = \mathbf{I} - 2\mathbf{K} \ .$$

5. Find the volume of the parallelipiped determined by the three vectors $3\mathbf{I} - \mathbf{J} + \mathbf{K},\ \mathbf{I} + 2\mathbf{J},\ \mathbf{I} - 3\mathbf{K}$.

6. Show that if $\det(\mathbf{U}, \mathbf{V}, \mathbf{W}) = |\mathbf{U}||\mathbf{V}||\mathbf{W}|$, then the three vectors are mutually orthogonal.

7. Given vectors $\mathbf{V} = 2\mathbf{I} - \mathbf{J},\ \mathbf{W} = \mathbf{I} + 2\mathbf{J} + 3\mathbf{K}$, show that $\mathbf{V}$ and $\mathbf{W}$ are orthogonal. Find a vector orthogonal to both $\mathbf{V}$ and $\mathbf{W}$.

8. Suppose that $\mathbf{X}$ is orthogonal to both $\mathbf{V}$ and $\mathbf{W}$. Show that the magnitude of $\mathbf{X} \cdot (\mathbf{V} \times \mathbf{W})$ is the product of the length of $\mathbf{X}$ with the area of the parallelogram spanned by $\mathbf{V}$ and $\mathbf{W}$.

9. Given vectors $\mathbf{V} = \mathbf{I} + 2\mathbf{J},\ \mathbf{W} = -\mathbf{I} + 2\mathbf{J} + 3\mathbf{K}$, find $pr_{\mathbf{W}}\mathbf{V}$ and $pr_{\mathbf{V}}\mathbf{W}$

13.4 Lines and Planes in Space

A coordinate system in space consists of a choice of a particular point O as origin, and a right-handed system of mutually orthogonal unit vectors $\mathbf{I},\ \mathbf{J}, \mathbf{K}$. Once a coordinate system is selected, we can represent a point $P : (x, y, z)$ by the vector $\vec{OP} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$ from the origin to P. Given another point $Q = (x', y', z')$, the *vector* from P to Q is denoted $\vec{PQ} = (x' - x)\mathbf{I} + (y' - y)\mathbf{J} + (z' - z)\mathbf{K}$. We shall often write the vector $\vec{OP}$ as $\mathbf{P}$ for consistency of notation in formalas.

The **line** through a given point P and in the direction of a given vector $\mathbf{L}$ is the set of all points X of the form

$$\mathbf{X} = \mathbf{P} + t\mathbf{L} \tag{13.25}$$

where t runs over all real numbers. This is called the **parametric form** of the equation of the line. This says that the vector $\mathbf{X} - \mathbf{P}$ is collinear with the vector $\mathbf{L}$, and thus the components are proportional. In coordinates, writing $\mathbf{X} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$, $\mathbf{P} = x_0\mathbf{I} + y_0\mathbf{J} + z_0\mathbf{K}$, and $\mathbf{L} = a\mathbf{I} + b\mathbf{J} + c\mathbf{K}$, we get the **symmetric form** of the equation of a line:

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

Example 13.17. Find the symmetric equations of the line through the points P(2,-1,4) and Q(6, 2, -3).

The vector $\vec{PQ} = 4\mathbf{I} + 3\mathbf{J} - 7\mathbf{K}$ is on the line, so $\mathbf{X} = \mathbf{xI} + \mathbf{yJ} + \mathbf{zK}$ is on the line precisely when $\mathbf{X} - \mathbf{P}$ is parallel to $\vec{PQ}$. This gives us the symmetric equations

$$\frac{x-2}{4} = \frac{y+1}{3} = \frac{z-4}{-7} \ .$$

The **plane** through a point P, **spanned** by the vectors $\mathbf{V}$ and $\mathbf{W}$ is the set of all points X of the form

$$\mathbf{X} = \mathbf{P} + s\mathbf{V} + t\mathbf{W}$$

where s, t range over all real numbers. This is the **parametric form** of a plane. We note that a point X is on the plane if and only if the parallelipiped formed from $\mathbf{X} - \mathbf{P}$, $\mathbf{V}$, $\mathbf{W}$ has zero volume, that is

$$\det(\mathbf{X} - \mathbf{P}, \mathbf{V}, \mathbf{W}) = \mathbf{0} \ .$$

This is the equation of the plane. The vector $\mathbf{N} = \mathbf{V} \times \mathbf{W}$ is said to be **normal** to the plane, as it is orthogonal to all vectors lying on the plane. In terms of the normal, we have this as the equation of the plane:

$$(\mathbf{X} - \mathbf{P}) \cdot \mathbf{N} = 0 \ , \tag{13.26}$$

since $\det(\mathbf{X} - \mathbf{P}, \mathbf{V}, \mathbf{W}) = (\mathbf{X} - \mathbf{P}) \cdot \mathbf{N}$.

Turning to coordinates, let P be the point (x_0, y_0, z_0), and $\mathbf{N} = a\mathbf{I} + b\mathbf{J} + c\mathbf{K}$. Then for (x, y, z) the coordinates for the point X, (13.26) becomes

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \text{or} \quad ax + by + cz = d \ . \tag{13.27}$$

where

$$d = ax_0 + by_0 + cz_0 \ .$$

We can summarize this discussion with

Proposition 13.10. a) Given a point P and a vector $\mathbf{N}$, the plane through P and orthogonal to N is given by the equation

$$\mathbf{X} \cdot \mathbf{N} = \mathbf{P} \cdot \mathbf{N} \ .$$

b) The plane through P spanned by $\mathbf{V}$ and $\mathbf{W}$ has as normal $\mathbf{N} = \mathbf{V} \times \mathbf{W}$.
c) The coefficients of the cartesian equation (13.27) for a plane are the components of the normal vector.

Example 13.18. Find the equation of the plane through the point P(5,3,-1) perpendicular to the line in space whose symmetric equations are

$$\frac{x-2}{3} = \frac{y+1}{4} = \frac{z-1}{-2}$$

The vector $3\mathbf{I}+4\mathbf{J}-2\mathbf{K}$ has the direction of the line, so is normal to the plane, and can be taken to be $\mathbf{N}$. We know that the equation of the plane has the form $(\mathbf{X}-\mathbf{P})\cdot\mathbf{N}=0$, for $P:(5,3,-1)$ is a point on the plane. This gives the equation

$$\mathbf{X}\cdot\mathbf{N}=\mathbf{X_0}\cdot\mathbf{N} \quad \text{or} \quad 3x+4y-2z=15+12+2=29 \ .$$

Example 13.19. Find the equation of the plane containing the points $P(2,5,-1)$, $Q(6,-1,0)$, $R(3,1,4)$.

The vectors $\vec{PQ}=4\mathbf{I}-6\mathbf{J}+\mathbf{K}$, $\vec{PR}=\mathbf{I}-4\mathbf{J}+5\mathbf{K}$ lie on the plane, so the normal is

$$\mathbf{N}=\vec{PQ}\times\vec{PR}=(-30+4)\mathbf{I}+(1-20)\mathbf{J}+(-16+6)\mathbf{K}=-26\mathbf{I}-19\mathbf{J}-10\mathbf{K} \ .$$

The equation of the plane then is $\mathbf{X}\cdot\mathbf{N}=\mathbf{P}\cdot\mathbf{N}$, which comes to $26x+19y+10z=137$.

Example 13.20. Find the equation of the line through the origin and orthogonal to the plane $2x-y+3z=1$.

The vector $2\mathbf{I}-\mathbf{J}+3\mathbf{K}$ is normal to the plane, so lies in the direction of the line.Thus the symmetric equations of the line are

$$\frac{x}{2}=\frac{y}{-1}=\frac{z}{3}$$

Now, given two planes with equations $\mathbf{X}\cdot\mathbf{N_1}=d_1$, $\mathbf{X}\cdot\mathbf{N_2}=d_2$, the vector $\mathbf{N_1}\times\mathbf{N_2}$ has the direction of the line of intersection of the two planes. Thus if P is a point on that line (found by finding a simultaneous solution of the equations of the planes), the equation of the line is

$$\mathbf{X}=\mathbf{P}+t(\mathbf{N_1}\times\mathbf{N_2}) \ .$$

Example 13.21. Find the parametric form of the line given by the equations

$$2x-y+3z=1 \ , \quad x+5y-2z=0 \ .$$

To find a point P on the line we solve the simultaneous equations, taking $z=0$. This gives the equations for x and y: $2x-y=1$, $x+5y=0$. The solution is $x=5/11$, $y=-1/11$. Thus $P(5/11,-1/11,0)$ is on the line. The cross product of the two normals is

$$(2\mathbf{I}-1\mathbf{J}+3\mathbf{K})\times(\mathbf{I}+5\mathbf{J}-2\mathbf{K})=-13\mathbf{I}+7\mathbf{J}+12\mathbf{K}$$

giving the parametric equation of the line

$$\mathbf{X}=(\frac{5}{11}-13t)\mathbf{I}+(-\frac{1}{11}+7t)\mathbf{J}+12t\mathbf{K} \ .$$

Now, suppose we are given two lines in parametric form:

$$\mathbf{X}=\mathbf{P_1}+t\mathbf{L_1} \ , \quad \mathbf{X}=\mathbf{P_2}+t\mathbf{L_2} \ ,$$

and a point Q, and are asked to find the equation of the plane through Q and parallel to the lines. Then the normal to this plane is perpendicular to the two lines, so can be taken to be $\mathbf{L_1}\times\mathbf{L_2}$, and then the equation of the desired plane is

$$(\mathbf{X}-\mathbf{Q})\cdot(\mathbf{L_1}\times\mathbf{L_2})=0 \ .$$

Example 13.22. Find the equation of the plane through a (2,0,-1) parallel to the vectors $\mathbf{V} = 2\mathbf{I} - \mathbf{J},\ \mathbf{W} = 6\mathbf{I} + \mathbf{K}$.

$\mathbf{V} \times \mathbf{W}$ is perpendicular to the vectors $\mathbf{V}, \mathbf{W}$, so can be taken as the normal $\mathbf{N}$ to the plane. We get

$$\mathbf{N} = (2\mathbf{I} - \mathbf{J}) \times (6\mathbf{I} + \mathbf{K}) = 2\mathbf{I} \times \mathbf{K} - 6\mathbf{J} \times \mathbf{I} - \mathbf{J} \times \mathbf{K} = -2\mathbf{J} + 6\mathbf{K} - \mathbf{I}\ .$$

Taking $\mathbf{X_0} = 2\mathbf{I} - \mathbf{K}$ as a given point on the plane, the equation $\mathbf{X} \cdot \mathbf{N} = \mathbf{X_0} \cdot \mathbf{N}$ is

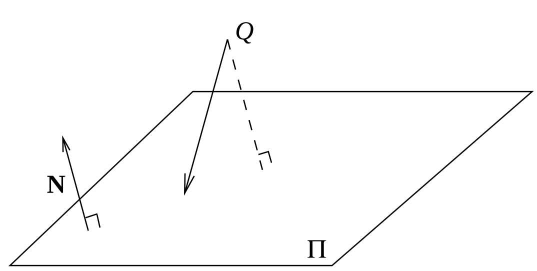
$$-x - 2y + 6z = 2(-1) + (-1)6 = -8$$

We can summarize this discussion in the form of two assertions.

Proposition 13.11
a) Given a line $\mathbf{X} = \mathbf{P} + t\mathbf{L}$, the plane through a given point Q and perpendicular to the line has the equation $(\mathbf{X} - \mathbf{Q}) \cdot \mathbf{L} = 0$.
b) Given the equation of a plane $\mathbf{X} \cdot \mathbf{N} = d$, a point P. the line through P and perpendicular to the plane has the equation $\mathbf{X} = \mathbf{P} + t\mathbf{N}$.

Now, suppose we want to find the distance of a point Q to a plane Π. We know from elementary geometry that the this distance is the length of the line segment from Q to Π which is perpendicular to Π. This line segment is thus in the direction of the normal to Π, and is seen (see figure 13.17) to be the projection of any vector from Q to Π in the normal direction.

Figure 13.17



This demonstrates the first part of

Proposition 13.12.
a) The distance from a point Q to a plane Π with normal $\mathbf{N}$ is

$$d(Q, \Pi) = \frac{|\vec{PQ} \cdot \mathbf{N}|}{|\mathbf{N}|}\ ,$$

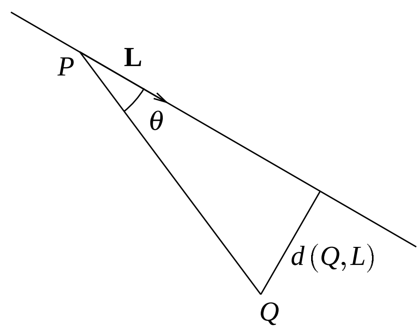
where P is any point on the plane.

b) The distance from a point Q to a line L in the direction $\mathbf{L}$ is

$$d(Q,L) = \frac{|\vec{PQ} \times \mathbf{L}|}{|\mathbf{L}|} \ ,$$

where P is any point on the line.

Figure 13.18



To show b), start with figure 13.18. We have

$$d(Q,L) = |\vec{PQ}| \sin\theta = \frac{|\vec{PQ}||\mathbf{L}|\sin\theta}{|\mathbf{L}|} = \frac{|\vec{PQ} \times \mathbf{L}|}{|\mathbf{L}|} \ .$$

Example 13.23. Find the distance of the point $(2,0,4)$ from the plane whose equation is $x+y-2z=0$.

Let $Q : (2,0,4)$. Pick a point P on the plane, for example, $P = (1,1,1)$. $\mathbf{N} = \mathbf{I} + \mathbf{J} - 2\mathbf{K}$ is normal to the plane, so the distance is the length of the projection of the vector from P to Q in the direction of $\mathbf{N}$:

$$\vec{PQ} \cdot \mathbf{N} = (\mathbf{I} - \mathbf{J} + 3\mathbf{K}) \cdot (\mathbf{I} + \mathbf{J} - 2\mathbf{K}) = -6 \ , \quad |\mathbf{N}| = \sqrt{6}$$

so the distance is $|\vec{PQ} \cdot \mathbf{N}|/|\mathbf{N}| = \sqrt{6}$.

Example 13.24. Find the distance of the point $(2,0,1)$ from the line whose symmetric equations are

$$\frac{x-2}{3} = \frac{y+1}{4} = \frac{z-1}{-2}$$

Let $\vec{OQ} = 2\mathbf{I} + \mathbf{K}$ be the vector to the given point, and $\vec{OP} = 2\mathbf{I} - \mathbf{J} + \mathbf{K}$ the vector to a point on the line, and $\mathbf{L} = 3\mathbf{I} + 4\mathbf{J} - 2\mathbf{K}$, a vector in the direction of the line. The distance is

$$\frac{|\vec{PQ} \times \mathbf{L}|}{|\mathbf{L}|} = \frac{|-\mathbf{J} \times (3\mathbf{I} + 4\mathbf{J} - 2\mathbf{K})|}{|3\mathbf{I} + 4\mathbf{J} - 2\mathbf{K}|} = \sqrt{\frac{13}{29}} \ .$$

Example 13.25. Find the distance between the two parallel planes

$$\Pi_1 \ : \quad x + 2y - 5z = 2 \ , \qquad \Pi_2 \ : \quad x + 2y - 5z = 11 \ .$$

The distance between the two planes is the length of any line segment perpendicular to both planes. Thus we need only find the length of the projection of $\vec{P_1P_2}$ on the common normal $\mathbf{N} = \mathbf{I} + 2\mathbf{J} - 5\mathbf{K}$, where P_1 is a point on Π_1 and P_2 is a point on Π_2. Since $\mathbf{P_1} \cdot \mathbf{N} = 2$, and $\mathbf{P_2} \cdot \mathbf{N} = 11$ for these points we get, for the distance:

$$\frac{|(\mathbf{P_2} - \mathbf{P_1}) \cdot \mathbf{N}|}{|\mathbf{N}|} = \frac{11 - 2}{\sqrt{1 + 4 + 25}} = \frac{9}{\sqrt{30}} \ .$$

Problems 13.4

1. Find a vector normal to the plane given by the equation

$$3x - 2y + z = 14 \ .$$

2. Find a vector normal to the plane through $P : (0, 0, 0),\ Q : (1, 0, -1),\ R : (0, 1, 1)$.

3. Find the equation of the plane through the origin which is normal to the line given parametrically by

$$\mathbf{X} = (3\mathbf{I} + 2\mathbf{J} - \mathbf{K}) + t(-\mathbf{I} + \mathbf{J} + 2\mathbf{K}) \ .$$

4. Find the symmetric equations of the line through the point (2,-1,3) which is perpendicular to the vectors $\mathbf{V} = 2\mathbf{I} - \mathbf{J} + 3\mathbf{K}$ and $\mathbf{W} = \mathbf{I} - \mathbf{J} + \mathbf{K}$.

5. Find the parametric equations of the line through the point (2,-1,3) which is parallel to the two planes given by the equations

$$3x + z = 4, \qquad x - 2y + 5z = 1 \ .$$

6. Find the equation of the plane through the point (2,-1,3) which is parallel to the vectors $\mathbf{I} - 2\mathbf{J} + 3\mathbf{K}$ and $3\mathbf{I} - 2\mathbf{J} + \mathbf{K}$.

7. Find the distance of the point $(2, 0, 1)$ from the plane given by the equation

$$\frac{x - 2}{3} + \frac{y + 1}{4} + \frac{z - 1}{2} = 0$$

8. Find the symmetric equations of the line of intersection of the planes of problems 1 and 2.

9. Find the equation of the line through $P(0, 2, 1)$ which is parallel to the line of intersection of the planes

$$\Pi_1 \ : \quad 3x - 5y + z = 0 \ , \qquad \Pi_2 \ : \quad 2x + y = 0 \ .$$

10. Let $L_1,\ L_2$ be two lines in three dimensions which do not intersect. There are points $P_1,\ P_2$ on $L_1,\ L_2$ respectively such that the line joining P_1 and P_2 intersects each line at a right angle. These are the points on the two lines which are closest together. Find a formula for the distance between P_1 and P_2, in terms of the equations of the lines

$$L_1 : \quad \mathbf{X} = \mathbf{Q_1} + t\mathbf{N_1} \qquad L_2 : \quad \mathbf{X} = \mathbf{Q_2} + s\mathbf{N_2} \ .$$

XIV. Particles in Motion; Kepler's Laws

14.1 Vector Functions

Vector notation is well suited to the representation of the motion of a particle. Fix a coordinate system with center **O**, and let the position of the object at time t, relative to **O** be represented by the vector function of t:

$$\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J} + z(t)\mathbf{K} \ . \tag{14.1}$$

The **velocity** of the object is the time rate of change of position:

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{dx}{dt}\mathbf{I} + \frac{dy}{dt}\mathbf{J} + \frac{dz}{dt}\mathbf{K},$$

and the **acceleration** is the time rate of change of velocity:

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2x}{dt^2}\mathbf{I} + \frac{d^2y}{dt^2}\mathbf{J} + \frac{d^2z}{dt^2}\mathbf{K}.$$

As shown, the differentiations are defined as they are for scalar functions, and are computed componentwise. Recall that we use the variable s to represent the length of the curve, as measured from a particular starting point. If we think of (14.1) as describing the curve parametrically, we find arc length by integrating

$$\frac{ds}{dt} = \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2} \tag{14.2}$$

with respect to t. The time rate of change of distance along the curve is the **speed** of the particle. From (14.2) we see that the speed of the particle is the magnitude of the velocity vector:

$$\frac{ds}{dt} = |\mathbf{V}| \ .$$

Example 14.1. A particle is rotating about the circle of radius R at constant angular velocity ω. To find the equations of motion, we start the clock with the particle on the positive x-axis; that is $\mathbf{X}(0) = R\mathbf{I}$. At time t the particle has moved through the angle ωt, so is at

$$\mathbf{X}(t) = R\cos\omega t\mathbf{I} + R\sin\omega t\mathbf{J} \ .$$

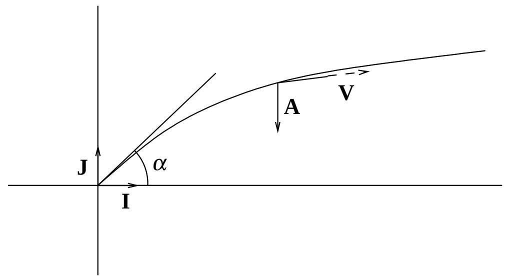
Differentiating, we find the velocity and acceleration:

$$\mathbf{V}(t) = -R\omega\sin\omega t\mathbf{I} + R\omega\cos\omega t\mathbf{J} \ , \quad \mathbf{A}(t) = -R\omega^2\cos\omega t\mathbf{I} - R\omega^2\sin\omega t\mathbf{J} \ .$$

Notice that $\mathbf{V}(t) = \omega\mathbf{X}(t)^\perp$, so $\mathbf{V}(t)$ is tangent to the circle at $\mathbf{X}(\mathbf{t})$ and has magnitude $\omega|\mathbf{X}(t)| = R\omega$, so the speed of the particle is $R\omega$. Also $\mathbf{A}(t) = -\omega^2\mathbf{X}(t)$, so the acceleration is directed toward the center of the circle and has magnitude $R\omega^2$.

Example 14.2. A missile is fired on the surface of the earth at an angle of elevation α and initial speed S ft/sec. Find the equations of motion.

Figure 14.1



The motion is shown in figure 14.1. We take the origin of coordinates to be the initial point of the trajectory of the missile. We know that the acceleration is that due to gravity, pointing downward and of magnitude 32 ft/sec^2. Thus $\mathbf{A}(t) = -32\mathbf{J}$. Integrating (componentwise), we have $\mathbf{V}(t) = -32t\mathbf{J} + \mathbf{C}$ where $\mathbf{C}$ is the constant (vector) of integration. Evaluating at $t = 0$, we have $\mathbf{C} = \mathbf{V}(0)$, which has magnitude S and direction α above the horizontal.

Thus

$$\mathbf{V}(t) = -32t\mathbf{J} + S(\cos\alpha\mathbf{I} + \sin\alpha\mathbf{J}) = S\cos\alpha\mathbf{I} + (S\sin\alpha - 32t)\mathbf{J}\ .$$

Integrating again

$$\mathbf{X}(t) = (S\cos\alpha)t\mathbf{I} + ((S\sin\alpha)t - 16t^2)\mathbf{J}\ ,$$

since $\mathbf{X}(0) = 0$. From this we can determine when the missile hits the ground again, and how far it has travelled. For $\mathbf{X}(t)$ is horizontal when its $\mathbf{J}$ component is zero; that is when $t = (S\sin\alpha)/16$. The distance travelled is the $\mathbf{I}$ component of $\mathbf{X}(t)$ for this t, so is

$$d = (S\cos\alpha)\frac{S\sin\alpha}{16} = \frac{S^2\sin 2\alpha}{32}\ .$$

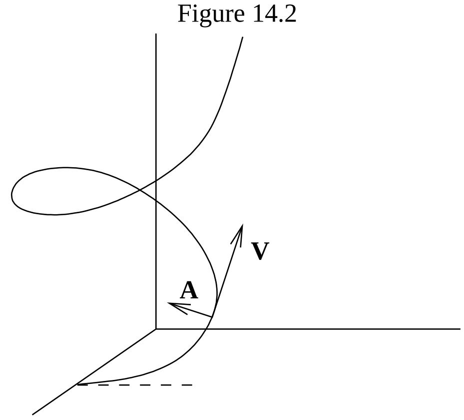
To choose the angle so that the horizontal distance has a maximum, we choose α so that this expression is largest. The maximum is $S^2/32$ when $\alpha = 45°$.

Example 14.3. Let $\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + t\mathbf{K}$. This particle is spiralling up the horizontal cylinder of radius 1 (see figure 14.2) at constant angular velocity. We have

$$\mathbf{V}(t) = -\sin t\mathbf{I} + \cos t\mathbf{J} + \mathbf{K}\ , \quad \mathbf{A}(t) = -\cos t\mathbf{I} - \sin t\mathbf{J}\ .$$

The speed of this particle is $|\mathbf{V}| = \sqrt{2}$, and its acceleration is of magnitude 1 and always pointed toward the z-axis. This is the same as circular motion in the xy-plane, except that the initial velocity has an elevation of 45° off the xy-plane.

Figure 14.2



In order to proceed we need to make some observations about the differentiation of vector functions.

Proposition 14.1. Suppose that $\mathbf{V}(t)$, $\mathbf{W}(t)$ are differentiable vector-valued functions of the variable t. Then

a)
$$\frac{d}{dt}(\mathbf{V}+\mathbf{W}) = \frac{d\mathbf{V}}{dt} + \frac{d\mathbf{W}}{dt}$$

If $w = w(t)$ is a differentiable scalar-valued function,

b)
$$\frac{d}{dt}(w\mathbf{V}) = \frac{dw}{dt}\mathbf{V} + w\frac{d\mathbf{V}}{dt}$$

c)
$$\frac{d}{dt}(\mathbf{V}\cdot\mathbf{W}) = \frac{d\mathbf{V}}{dt}\cdot\mathbf{W} + \mathbf{V}\cdot\frac{d\mathbf{W}}{dt}$$

d)
$$\frac{d}{dt}(\mathbf{V}\times\mathbf{W}) = \frac{d\mathbf{V}}{dt}\times\mathbf{W} + \mathbf{V}\times\frac{d\mathbf{W}}{dt}$$

e) If $\mathbf{U}(t)$ is a vector of length one for all t, then $\mathbf{U}$ and $d\mathbf{U}/dt$ are orthogonal.

The first four identities are all verified by writing all vectors in component form, and using the usual product rule for differentiation of scalar functions. e) follows from c) this way: since $\mathbf{U}\cdot\mathbf{U} = 1$, we have
$$\frac{d\mathbf{U}}{dt}\cdot\mathbf{U} + \mathbf{U}\cdot\frac{d\mathbf{U}}{dt} = 0$$
so that $(d\mathbf{U}/dt)\cdot\mathbf{U} = 0$.

Example 14.4. Let $\mathbf{L} = \mathbf{L}(t)$ be a unit-vector valued function whose values all lie in the xy-plane.Let $\theta(t)$ be the angle between $\mathbf{L}(t)$ and the horizontal. Then
$$\frac{d\mathbf{L}}{dt} = \frac{d\theta}{dt}\mathbf{L}^{\perp}\ ,$$

where $\mathbf{L}^\perp$ is the unit vector orthogonal and to the left of $\mathbf{L}$.

Writing $\mathbf{L}$ in components, we have

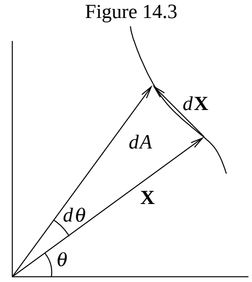
$$\mathbf{L}(t) = \cos\theta(t)\mathbf{I} + \sin\theta(t)\mathbf{J}, \quad \mathbf{L}^\perp(t) = -\sin\theta(t)\mathbf{I} + \cos\theta(t)\mathbf{J} \ .$$

Then

$$\frac{d\mathbf{L}}{dt} = -\sin\theta\frac{d\theta}{dt}\mathbf{I} + \cos\theta\frac{d\theta}{dt}\mathbf{J} = \frac{d\theta}{dt}\mathbf{L}^\perp \ .$$

Example 14.5. Let $\mathbf{X}(t)$ be a vector-valued function whose values always lie in the xy-plane. Let $A(t)$ be the area "swept out" by $\mathbf{X}$ from t_0 to t; that is, the area of the domain bounded by the vectors $\mathbf{X}(t_0)$, $\mathbf{X}(t)$ and the trajectory of the vector (see figure 14.3).

Figure 14.3



Show that

$$\frac{dA}{dt} = \frac{1}{2}\left|\frac{d\mathbf{X}}{dt} \times \mathbf{X}\right| \ .$$

If we argue by differentials, this is easy. In the diagram, dA is half the area of the parallelogram spanned by $\mathbf{X}$ and $d\mathbf{X}$, so $dA = (1/2)|d\mathbf{X} \times \mathbf{X}|$.

We can also argue using the polar representation of vectors. Let r be the length of $\mathbf{X}$ and $\mathbf{L}$ the unit vector in the direction of $\mathbf{X}$, so that $\mathbf{X} = r\mathbf{L}$. We know (the polar form for area):

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} \ .$$

Now, using example 14.4,

$$\frac{d\mathbf{X}}{dt} = \frac{d}{dt}(r\mathbf{L}) = \frac{dr}{dt}\mathbf{L} + r\frac{d\mathbf{L}}{dt} = \frac{dr}{dt}\mathbf{L} + r\frac{d\theta}{dt}\mathbf{L}^\perp \ ,$$

so that

$$\left|\frac{d\mathbf{X}}{dt} \times \mathbf{X}\right| = \left|\left(\frac{dr}{dt}\mathbf{L} + r\frac{d\theta}{dt}\mathbf{L}^\perp\right) \times r\mathbf{L}\right| = r^2\frac{d\theta}{dt} \ ,$$

since $\mathbf{X}$ and $\mathbf{L}$ are colinear, and $\mathbf{L}$ and $\mathbf{L}^\perp$ are orthogonal unit vectors.

Problems 14.1

1. Let $\mathbf{L} = \cos\theta(t)\mathbf{I} + \sin\theta(t)\mathbf{J}$ be a unit vector -valued function of t in the xy-plane. Show that

$$\frac{d}{dt}(\mathbf{L} \times \mathbf{K}) = \frac{d\theta}{dt}\mathbf{L}.$$

2. An object of mass m is moving in the plane subject to a force directed toward the origin and of magnitude k times the distance from the origin. Suppose that its initial position, (at time $t = 0$) is at (1,0) and its initial velocity is $b\mathbf{J}$. Show that the trajectory is an ellipse of eccentricity

$$e = \sqrt{1 - \frac{k}{mb^2}} \quad \text{or} \quad e = \sqrt{1 - \frac{mb^2}{k}} \ ,$$

whichever is real.

3. On a level field, a baseball is thrown with a speed of 96 feet/sec. at an angle of 30° with the horizontal. How far has the ball traveled when it hits the ground? For convenience, assume that the ball leaves the thrower's hand at ground level).

4. The same situation as in problem 3, but this time the thrower wants the ball to land 80 feet away. At what angle should he throw the ball?

5. A particle moves along the ellipse

$$x^2 + \frac{y^2}{4} = 1$$

according to the equation

$$\mathbf{X}(t) = \cos(\frac{\pi}{2}t)\mathbf{I} + 2\sin(\frac{\pi}{2}t)\mathbf{J} \ .$$

Find the area swept out as a function of time.

6. Let $\mathbf{M}$ be a vector in space. Suppose that a particle moves in such a way that its acceleration is always orthogonal to $\mathbf{M}$ and its initial velocity is orthogonal to $\mathbf{M}$. Then the particle moves in a plane orthogonal to $\mathbf{M}$.

7. Let $\mathbf{M}$ be a vector in space. Suppose that a particle moves in such a way that for all t, $\mathbf{V}(t) \times \mathbf{M} = 0$. Show that the particle motion is in a plane.

14.2 Planar Particle Motion

As the above examples demonstrate, a good understanding of the motion is achieved when it is described in terms of the position of the particle, rather than relative to a fixed origin. For this reason we want to represent the velocity and acceleration in terms which relate directly to the particle motion. We start with motion in the plane, with $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J}$ representing the position of the particle relative to a given coordinate system. As above, we have the velocity and acceleration of the particle given by

$$\mathbf{V} = \frac{d\mathbf{X}}{dt} = \frac{dx}{dt}\mathbf{I} + \frac{dy}{dt}\mathbf{J} \ , \quad \mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2x}{dt^2}\mathbf{I} + \frac{d^2y}{dt^2}\mathbf{J} \ . \tag{14.3}$$

The *speed* of the moving point is

$$\frac{ds}{dt} = |\mathbf{V}| = \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2} \ . \tag{14.4}$$

The unit vector in the direction of motion, called the *tangent*, is denoted $\mathbf{T}$. Thus

$$\mathbf{V} = \frac{ds}{dt}\mathbf{T} \ . \tag{14.5}$$

Now the acceleration vector represents two aspects of the motion, describing both the way the direction of motion is turning, and the way the speed is changing. Differentiating (14.5), we have

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}\frac{d\mathbf{T}}{dt} \ , \tag{14.6}$$

The first term gives the change of speed, and the second, the change in direction. By example 14.4, the second term is orthogonal to the first. We say that it is *normal* to the curve of motion (called the *trajectory*). Now, by example 14.4,

$$\frac{d\mathbf{T}}{ds} = \frac{d\theta}{ds}\mathbf{T}^{\perp}$$

where θ is the angle between $\mathbf{T}$ and the horizontal. Since both $\mathbf{T}$ and arc length are independent of time, this equation has to do only with the trajectory. We define the *curvature*, κ, of the curve as the magnitude of $d\mathbf{T}/ds$, and the *unit normal*, $\mathbf{N}$, to the curve as the direction of $d\mathbf{T}/ds$. Thus

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N} \ , \tag{14.7}$$

so that $\kappa = |d\mathbf{T}/ds|$ and $\mathbf{N} = \pm\mathbf{T}^{\perp}$, and equation (14.6) becomes

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2s}{dt^2}\mathbf{T} + (\frac{ds}{dt})^2\kappa\mathbf{N} \ , \tag{14.8}$$

since

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds}\frac{ds}{dt}$$

by the chain rule. The component of $\mathbf{A}$ in the direction of $\mathbf{T}$ is called the *tangential acceleration* and denoted a_T, and the component in the normal direction is the *normal acceleration* denoted a_N. Thus (14.8) can be rewritten as the set of equations

$$\mathbf{A} = a_T\mathbf{T} + a_N\mathbf{N} \ , \quad \text{where} \quad a_T = \mathbf{A}\cdot\mathbf{T} = \frac{d^2s}{dt^2} \ , \quad a_N = \mathbf{A}\cdot\mathbf{N} = \kappa(\frac{ds}{dt})^2 \ . \tag{14.9}$$

Example 14.6. In example 14.1 we considered a particle moving around a circle of radius R at constant angular velocity ω and found

$$\mathbf{V}(t) = -R\omega\sin\omega t\mathbf{I} + R\omega\cos\omega t\mathbf{J} = R\omega(-\ \sin\omega t\mathbf{I} + \cos\omega t\mathbf{J}) \ ,$$

$$\mathbf{A}(t) = -R\omega^2\cos\omega t\mathbf{I} - R\omega^2\sin\omega t\mathbf{J} = R\omega^2(-\cos\omega t\mathbf{I} - \sin\omega t\mathbf{J}) \ .$$

Thus

$$\frac{ds}{dt} = |\mathbf{V}| = R\omega \ , \quad \mathbf{T} = \frac{\mathbf{V}}{|\mathbf{V}|} = -\ \sin\omega t\mathbf{I} + \cos\omega t\mathbf{J} \ , \quad \mathbf{N} = \mathbf{T}^{\perp} = -\cos\omega t\mathbf{I} - \sin\omega t\mathbf{J} \ ,$$

so that $\mathbf{A} = R\omega^2\mathbf{N}$ and, by (14.8):

$$a_T = \frac{d^2s}{dt^2} = 0 \ , \quad a_N = (\frac{ds}{dt})^2\frac{1}{R} = R\omega^2 \ , \quad \kappa = \frac{1}{R} \ .$$

Example 14.7. In example 14.2, we considered the trajectory of a missile fired at an initial speed of S ft/sec, and at an angle α to the horizontal. We found

$$\mathbf{V}(t) = S\cos\alpha\mathbf{I} + (S\sin\alpha - 32t)\mathbf{J} \ , \quad \mathbf{A} = -32\mathbf{J} \ .$$

Thus

$$\frac{ds}{dt} = \sqrt{(S\cos\alpha)^2 + (S\sin\alpha - 32t)^2} = S\sqrt{\cos^2\alpha + (\sin\alpha - 32t/S)^2} \ ,$$

$$\mathbf{T} = \frac{\cos\alpha\mathbf{I} + (\sin\alpha - 32t/S)\mathbf{J}}{\sqrt{\cos^2\alpha + (\sin\alpha - 32t/S(\sin\alpha - 32t/S)^2}} \ .$$

Now, since the acceleration is clockwise to $\mathbf{T}$, so is $\mathbf{N}$, thus

$$\mathbf{N} = -\mathbf{T}^{\perp} = \frac{(\sin\alpha - 32t/S)\mathbf{I} - \cos\alpha\mathbf{J}}{\sqrt{\cos^2\alpha + (\sin\alpha - 32t/S(\sin\alpha - 32t/S)^2}} \ .$$

Finally,

$$a_T = \mathbf{A}\cdot\mathbf{T} = \frac{-32(\sin\alpha - 32t/S)}{\sqrt{\cos^2\alpha + (\sin\alpha - 32t/S(\sin\alpha - 32t/S)^2}} \ ,$$

$$a_N = \mathbf{A}\cdot\mathbf{N} = \frac{32\cos\alpha}{\sqrt{\cos^2\alpha + (\sin\alpha - 32t/S(\sin\alpha - 32t/S)^2}} \ ,$$

and the curvature is

$$\kappa = \frac{a_N}{(ds/dt)^2} = \frac{32\cos\alpha}{(\cos^2\alpha + (\sin\alpha - 32t/S(\sin\alpha - 32t/S)^2)^{3/2}} \ .$$

An effective procedure to follow to make these calculations is this:
1. Given the formula $\mathbf{X}(t)$ of motion, first differentiate it twice to obtain $\mathbf{V}$ and $\mathbf{A}$.
2. The magnitude of $\mathbf{V}$ is the speed, its direction vector is $\mathbf{T}$. Calculate $a_T = \mathbf{A}\cdot\mathbf{T}$.
3. The normal is $\pm\mathbf{T}^{\perp}$. Calculate $\mathbf{A}\cdot\mathbf{T}^{\perp}$; if it is positive it is a_N, and $\mathbf{N} = \mathbf{T}^{\perp}$; otherwise change the signs.
4. Finally, the curvature is $\kappa = a_N/(ds/dt)^2$.

Example 14.8. A particle moves in the plane according to the equation

$$\mathbf{R}(t) = t^{-1}\mathbf{I} + \ln t\mathbf{J}$$

Find the tangential and normal accelerations and the curvature of the trajectory at the time $t = 1$.

First we calculate the velocity and acceleration:

$$\mathbf{V} = \frac{-1}{t^2}\mathbf{I} + \frac{1}{t}\mathbf{J} \ , \quad \mathbf{A} = \frac{2}{t^3}\mathbf{I} - \frac{1}{t^2}\mathbf{J}$$

Evaluating at $t = 1$, we have $\mathbf{V} = -\mathbf{I} + \mathbf{J}$, $\mathbf{A} = 2\mathbf{I} - \mathbf{J}$. Then

$$\frac{ds}{dt} = |\mathbf{V}| = \sqrt{2} \ , \quad \text{and} \quad \mathbf{T} = \frac{-\mathbf{I} + \mathbf{J}}{\sqrt{2}} \ .$$

Since $\mathbf{A}$ is clockwise to $\mathbf{T}$, we must take

$$\mathbf{N} = \frac{\mathbf{I} + \mathbf{J}}{\sqrt{2}} \ ,$$

and finally

$$a_T = \mathbf{A} \cdot \mathbf{T} = \frac{-3}{\sqrt{2}} \ , \quad a_N = \mathbf{A} \cdot \mathbf{N} = \frac{1}{\sqrt{2}} \ , \quad \kappa = \frac{a_N}{(\frac{ds}{dt})^2} = \frac{1}{2\sqrt{2}} \ .$$

Problems 14.2

1. A particle moves along the Archimedean spiral ($r = a\theta$ in polar coordinates) so that its angular speed is constant. Find its speed.

2. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = t \sin t\mathbf{I} + \cos t\mathbf{J}$$

Find the speed, tangential and normal accelerations and the curvature of the trajectory at the time $t = 2\pi$.

3. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = t\mathbf{I} - t^3\mathbf{J}$$

Find the velocity, speed, acceleration, tangent and normal vectors, and normal acceleration of the particle at any time t.

4. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = (t^2 + t + 1)\mathbf{I} + t^3\mathbf{J}$$

Find the velocity, speed, acceleration, tangent and normal vectors of the particle at any time $t > 0$.

5. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = t\mathbf{I} + \ln t\mathbf{J}$$

Find the velocity, speed, acceleration, tangent and normal vectors, and tangential and normal acceleration of the particle at any time $t > 0$.

6. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = \ln t\mathbf{I} + \frac{1}{t}\mathbf{J}$$

Find the velocity, speed, acceleration, tangent and normal vectors, and normal acceleration of the particle at any time t.

7. A particle moves in the plane according to the equation

$$\mathbf{X}(t) = e^{at}(\cos t\mathbf{I} + \sin t\mathbf{J})$$

Show that the angle between the position vector and the tangent line to the trajectory is constant.

14.3 Particle Motion in Space

For motion in space, the ideas are precisely the same; however, a little more work is needed to find the normal direction. Again we start with the motion described by a vector function $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J} + z(t)\mathbf{K}$ in a given coordinate system. The *speed* of the moving point is

$$\frac{ds}{dt} = |\mathbf{V}| = \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2} \ . \tag{14.10}$$

The unit vector in the direction of motion (called the *tangent*) is denoted $\mathbf{T}$. Thus

$$\mathbf{V} = \frac{ds}{dt}\mathbf{T} \ . \tag{14.11}$$

Since $\mathbf{T}$ is a unit vector, its derivative is orthogonal to it. We define the *normal*, $\mathbf{N}$ as the unit vector in the direction of the derivative of $\mathbf{T}$. The plane determined by the tangent and normal vectors is called the *tangent plane* of the motion. We introduce the *curvature* of the trajectory of the motion as the magnitude of the deriviative of $\mathbf{T}$ with respect to arc length; thus

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N} \ . \tag{14.12}$$

Note that, by the chain rule, the derivative of $\mathbf{T}$ with respect to time is

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds}\frac{ds}{dt} = \kappa\frac{ds}{dt}\mathbf{N} \ . \tag{14.13}$$

Now the acceleration vector lies in the plane of $\mathbf{T}$, $\mathbf{N}$, which we see by differentiating (14.11):

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}\frac{d\mathbf{T}}{dt} = \frac{d^2s}{dt^2}\mathbf{T} + (\frac{ds}{dt})^2\kappa\mathbf{N} \ , \tag{14.14}$$

using (14.4). The component of $\mathbf{A}$ in the direction of $\mathbf{T}$ is called the *tangential acceleration* and denoted a_T, and the component in the normal direction is the *normal acceleration* denoted a_N. Thus (14.14) can be rewritten as the set of equations

$$\mathbf{A} = a_T\mathbf{T} + a_N\mathbf{N} \quad \text{where} \quad a_T = \mathbf{A}\cdot\mathbf{T} = \frac{d^2s}{dt^2} \ , \quad a_N = \mathbf{A}\cdot\mathbf{N} = \kappa(\frac{ds}{dt})^2 \ . \tag{14.15}$$

Notice that, since the normal acceleration is always positive, $\mathbf{N}$ lies on the same side of $\mathbf{T}$ as the acceleration $\mathbf{A}$.

In order to find the components of the acceleration in any particular example, it is best to use the geometry as a guide, rather than to do the calculations by direct applications of these formulae. The procedure to follow to make these calculations is this:

1. Given the formula $\mathbf{X}(t)$ of motion, first differentiate it twice to obtain $\mathbf{V}$ and $\mathbf{A}$.
2. The magnitude of $\mathbf{V}$ is the speed, its direction vector is $\mathbf{T}$.
3. Calculate $a_T = \mathbf{A} \cdot \mathbf{T}$.
4. From (14.15) we have

$$a_N \mathbf{N} = \mathbf{A} - (\mathbf{A} \cdot \mathbf{T})\mathbf{T}$$

so calculate the vector on the right. Its magnitude is a_N and its direction vector is $\mathbf{N}$.
5. The curvature is $\kappa = a_N/(ds/dt)^2$.

Formulas for the normal acceleration and curvature which are sometimes useful are

$$a_N = \frac{|\mathbf{V} \times \mathbf{A}|}{|\mathbf{V}|}\ , \quad \kappa = \frac{|\mathbf{V} \times \mathbf{A}|}{|\mathbf{V}|^3}. \tag{14.16}$$

To see this, start with (14.15): $\mathbf{A} = a_T\mathbf{T} + a_N\mathbf{N}$, and take the vector product with $\mathbf{V}$. Since $\mathbf{V} \times \mathbf{T} = \mathbf{0}$,

$$\mathbf{V} \times \mathbf{A} = a_N \mathbf{V} \times \mathbf{N}$$

Now, take lengths: $|\mathbf{V} \times \mathbf{A}| = a_N|\mathbf{V}|$, since $\mathbf{V}$ is orthogonal to $\mathbf{N}$, and $\mathbf{N}$ has length one.

One final remark: if the problem is to find the components of the acceleration at a particular value of t, then, immediately after differentiating, evaluate at t. Then the rest of the problem is just vector algebra with specific vectors.

Example 14.9. In example 14.3 we considered a particle moving counterclockwise up a spiral at constant speed. There we found

$$\mathbf{V} = \cos t\mathbf{I} - \sin t\mathbf{J} + \mathbf{K}\ , \quad \mathbf{A} = -\sin t\mathbf{I} - \cos t\mathbf{J}\ , \quad \frac{ds}{dt} = |\mathbf{V}| = \sqrt{2}\ .$$

Note that $\mathbf{A} \cdot \mathbf{T} = 0, |\mathbf{A}| = 1$. Thus, since $\mathbf{A}$ is a unit vector orthogonal to $\mathbf{T}$ on the same side of $\mathbf{T}$ as $\mathbf{A}$, it *is* the normal vector: $\mathbf{A} = \mathbf{N}$. Finally,

$$a_T = 0\ , \quad a_N = 1\ , \quad \kappa = \frac{a_N}{(ds/dt)^2} = \frac{1}{2}\ .$$

Example 14.10. A particle moves in space according to the formula

$$\mathbf{R}(t) = t^2\mathbf{I} + \ln t\mathbf{J} + \frac{1}{t}\mathbf{K}\ .$$

Find $\mathbf{V}, \mathbf{A}, \mathbf{T}, \mathbf{N}, a_T, a_N, \kappa$ at the point $t = 1$.

First we differentiate to find the velocity and acceleration.

$$\mathbf{V} = 2t\mathbf{I} + \frac{1}{t}\mathbf{J} - \frac{1}{t^2}\mathbf{K}\ , \quad \mathbf{A} = 2\mathbf{I} - \frac{1}{t^2}\mathbf{J} + \frac{2}{t^3}\mathbf{K}\ .$$

At $t = 1$, $\mathbf{V} = \mathbf{I} + \mathbf{J} - \mathbf{K}$, $\mathbf{A} = 2\mathbf{I} - \mathbf{J} + 2\mathbf{K}$, and $ds/dt = |\mathbf{V}| = \sqrt{6}$. Thus

$$\mathbf{T} = \frac{2\mathbf{I} + \mathbf{J} - \mathbf{K}}{\sqrt{6}}\ ,$$

$$a_T = \mathbf{A} \cdot \mathbf{T} = \frac{1}{\sqrt{6}} \ .$$

For the normal acceleration, we calculate $a_N\mathbf{N}$ in terms of what we already know:

$$a_N\mathbf{N} = \mathbf{A} - a_T\mathbf{T} = 2\mathbf{I} - \mathbf{J} + 2\mathbf{K} - \frac{1}{\sqrt{6}}(\frac{2\mathbf{I} + \mathbf{J} - \mathbf{K}}{\sqrt{6}}) = \frac{10\mathbf{I} - 7\mathbf{J} + 13\mathbf{K}}{6} \ .$$

a_N is the magnitude of this vector, and $\mathbf{N}$ its direction, so

$$a_N = \frac{\sqrt{318}}{6} \ , \quad \mathbf{N} = \frac{10\mathbf{I} - 7\mathbf{J} + 13\mathbf{K}}{\sqrt{318}}, \quad \kappa = \frac{a_N}{(\frac{ds}{dt})^2} = \frac{\sqrt{318}}{6} \ .$$

Problems 14.3

1. A particle moves in space according to the formula

$$\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + \cos(2t)\mathbf{K} \ .$$

Find the tangential and normal accelerations and the curvature at $t = \pi/4$.

2. A particle moves in space according to the formula

$$\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + e^t\mathbf{K} \ .$$

Find the tangential and normal accelerations for this motion.

3. A particle moves in space according to the formula $\mathbf{X}(t) = \mathbf{I} + t\mathbf{J} - t^2\mathbf{K}$. Find the tangential and normal accelerations as functions of t.

4. A particle moves in space according to the formula $\mathbf{X}(t) = t\mathbf{I} + e^{2t}\mathbf{J} - 2e^t\mathbf{K}$. Find the normal acceleration at the point $t = 0$.

5. A particle moves in space according to the formula $\mathbf{X}(t) = \ln t\mathbf{I} + t^2\mathbf{J} - t\mathbf{K}$.

   a) Find the values of t at which the velocity and acceleration vectors are orthogonal.

   b) What is the tangential acceleration of the particle at these points?

6. A particle moves in space according to the formula

$$\mathbf{X}(t) = \frac{1}{2}t^2\mathbf{I} + \frac{1}{t}\mathbf{J} - t\mathbf{K} \ .$$

Find a_N, κ at the point $t = 1$.

7. Let $\mathbf{U}(t)$, $\mathbf{V}(t)$ be unit vectors which are always orthogonal. Let $\mathbf{W} = \mathbf{U} \times \mathbf{V}$. Show that there are scalar functions α, β, γ such that

$$\frac{d\mathbf{U}}{dt} = \alpha\mathbf{V} + \beta\mathbf{W} \ , \quad \frac{d\mathbf{V}}{dt} = -\alpha\mathbf{U} + \gamma\mathbf{W} \ , \quad \frac{d\mathbf{W}}{dt} = -\beta\mathbf{U} - \gamma\mathbf{V}$$

8. Let $\mathbf{X} = \mathbf{X}(t)$ represent the motion of a particle in space.

a) Show that if $|\mathbf{V}|$ is constant, then $\mathbf{A}$ is orthogonal to $\mathbf{T}$.

b) Show that if $\mathbf{A}$ is constant, the motion lies in a plane.

c) Show that if $\mathbf{V} \cdot \mathbf{X} = 0$, then particle moves on a fixed sphere centered at the origin.

9. Show that a space curve with no curvature ($\kappa = 0$) must be a line.

10. Let $\mathbf{X} = \mathbf{X}(t)$ parametrize the motion of a particle in space. Let $\mathbf{T}$, $\mathbf{N}$ be the unit tangent and normal vectors as in the text above. Define $\mathbf{B} = \mathbf{T} \times \mathbf{N}$ (this is called the *binormal* of the curve. Using problem 7, show the Frenet-Serret formulas:

$$\frac{d\mathbf{T}}{ds} = \kappa \mathbf{N} \ ,$$

$$\frac{d\mathbf{T}}{ds} = -\kappa \mathbf{T} + \tau \mathbf{B} \ ,$$

$$\frac{d\mathbf{B}}{ds} = -\tau \mathbf{N} \ ,$$

where κ is the curvature and τ is another scalar function, called the *torsion* .

14.4 Derivation of Kepler's Laws of Planetary Motion from Newton's Laws

Historical Background

Newton's *Principia Mathematica*, published in 1684, is the fundamental text for rational mechanics, especially the dynamics of bodies in motion. It forms a bridge between late medieval rational philosophy and modern physics; written as it is in the style of its day, but in conception and execution as modern as any present day text on the subject.

In Book I of the *Principia* Newton develops the fundamentals of the theory of motion based on his methods of the Calculus and three laws of motion:

Law 1. Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.

Law 2. The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed.

Law 3. To every action there is always opposed an equal reaction: or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

In Book II Newton shows how the laws of Kepler on planetary motion and Galileo's laws of falling bodies lead, in the context of the above laws, to the concept of a universal force of attraction between massive bodies. In particular, he shows that Kepler's laws imply that this force is inversely proportional to the square of the distance between the bodies. This is the hypothesis to which he is headed, his law of universal gravitation.

Newton's Law of Universal Gravitation. Given two massive bodies, they exert a force on one another which is proportional to their masses and inversely proportional to the square of the distance between them.

In those days this was called an "occult" force: action at a distance. Descartes' physics, based on strict ratonalism, denied the existence of mysterious, or occult causes, and thus with it, forces that act at a distance. This led Descartes to a description of planetary motion based on collisions within hypothetical vortices. In order for Newton to establish the validity of his theory, it was necessary to repudiate these Cartesian constructs, show that the concept of gravitation is forced by the work of Galileo and Kepler, and conversely, to deduce their Laws from his. These tasks he completely accomplished in Book II. It is worth noting that Newton remarks at length that, although there was no (concrete) physical evidence for gravitation, the mathematical arguments were forcing: the postulate of Universal Gravitation was forced by, and in turn predicted correctly, the accumulated data. Finally in Book III Newton showed how his Laws lead to accurate predictions of the motions of planets, and he gave rational explanations for the tides, recession and many other astronomical phenomena.

Kepler's Laws

Kepler's laws are descriptive of the motion of planets around the Sun. For this discussion we shall locate the Sun at the point $\mathbf{O}$, and let $\mathbf{X}(t)$ represent the position of the planet (relative to $\mathbf{O}$) at time t after the initial observation ($t = 0$).

Kepler I. a) A planet revolves around the Sun in an elliptical orbit.
b) The Sun is located at a focus of the ellipse.

Kepler II. The line joining the Sun to a planet sweeps out equal areas in equal times. More expicitly, let $A(t)$ be the area swept out by the vector $\mathbf{X}(t)$. Then for any t, s, $A(t+s)-A(t) = hs$, where h is constant. Equivalently, $dA/dt = h$ is constant.

Kepler III. The square of the period of revolution of a planet is proportional to the cube of the length of the major axis of its orbit.

Newton's second law of motion is

$$\mathbf{F} = m\mathbf{A}$$

where $\mathbf{F}$ is the force applied to the object, m is its mass, and $\mathbf{A}$ is the acceleration of the object. In particular, for planetary motion, this and Newton's law of gravitation imply that the acceleration of the planet is directed toward the sun; that is, it is **centripetal**. Newton's first observation is that this is equivalent to Kepler's second law.

Proposition 14.2. Suppose the only external force exerted on an object is centripetal. Then its motion is planar and Kepler II holds.

Let $\mathbf{X}$ be the position vector to the object, $\mathbf{V}$ its velocity, and $\mathbf{A}$ its acceleration. Since $\mathbf{A}$ is colinear with $\mathbf{X}$, $\mathbf{X} \times \mathbf{A} = \mathbf{0}$. Now

$$\frac{d}{dt}(\mathbf{X} \times \mathbf{V}) = \frac{d\mathbf{X}}{dt} \times \mathbf{V} + \mathbf{X} \times \frac{d\mathbf{V}}{dt} = \mathbf{V} \times \mathbf{V} + \mathbf{X} \times \mathbf{A} = \mathbf{0}.$$

Thus $\mathbf{X} \times \mathbf{V}$is a constant vector $\mathbf{H}$. If $\mathbf{H} = \mathbf{0}$, then $\mathbf{X}$ and $\mathbf{V}$ are colinear, and the motion is on a line directed toward $\mathbf{O}$. Otherwise $\mathbf{X}$ and $\mathbf{V}$ always lie on the plane perpendicular to the constant vector $\mathbf{H}$, so the motion is planar.

Finally, let $h = |\mathbf{H}|$; h is constant, and we have

$$h = |\mathbf{X} \times \mathbf{V}| = |\mathbf{X} \times \frac{d\mathbf{X}}{dt}| = r^2 \frac{d\theta}{dt} = 2\frac{dA}{dt} \tag{14.17}$$

(by example 14.5), giving us Kepler's second law.

We now show that the assumption of planar motion and Kepler's second law implies the force is centripetal. The assumption that the motion is planar implies that $\mathbf{X} \times \mathbf{V}$ has a fixed direction, and Kepler's second law implies (see example 5) that the length of that vector is constant. Thus $\mathbf{X} \times \mathbf{V}$ is a constant vector. Differentiating, we obtain

$$\mathbf{V} \times \mathbf{V} + \mathbf{X} \times \mathbf{A} = \mathbf{0},$$

so $\mathbf{X} \times \mathbf{A} = \mathbf{0}$, and $\mathbf{A}$ and $\mathbf{X}$ are colinear, that is, $\mathbf{A}$ is centripetal. As an aside, we observe that the fact that $\mathbf{X} \times \mathbf{V}$ is constant can be interpreted as the conservation of angular momentum.

The crux of Newton's argument is that the first Kepler Law follows from the inverse-square law. We shall go straight to this argument, although Newton did take the trouble to also show that Kepler's first law forces the inverse-square law.

Proposition 14.3. Suppose that an object moves in a plane subject to a force directed toward a fixed point $\mathbf{O}$ of magnitude inversely proportional to the square of the distance from $\mathbf{O}$. Then the orbit of the object is a conic section.

We continue the argument of Proposition 14.2, continuing with that notation. The additional hypothesis is that

$$\frac{d\mathbf{V}}{dt} = -\frac{k}{r^2}\mathbf{L}, \tag{14.18}$$

where k is a positive constant. The trick now is to take the cross product with $\mathbf{K}$. Even though the action takes place completely in the plane perpendicular to $\mathbf{K}$, this allows us to introduce the area information in vectorial form.

$$\frac{d\mathbf{V}}{dt} \times \mathbf{K} = -\frac{k}{r^2}\mathbf{L} \times \mathbf{K} = \frac{k}{r^2}\mathbf{L}^\perp \ ,$$

since $\mathbf{L}$ and $\mathbf{K}$ are orthogonal unit vectors and the system $\mathbf{L},\ \mathbf{K},\ \mathbf{L}^\perp$ is left-handed. But, as we saw in example 14.4,

$$\frac{d\mathbf{L}}{dt} = \frac{d\theta}{dt}\mathbf{L}^\perp \quad \text{or} \quad \mathbf{L}^\perp = \frac{d\mathbf{L}}{dt} \Big/ \frac{d\theta}{dt} \ .$$

By (14.17), $r^2 d\theta/dt = h$, so

$$\frac{d\mathbf{V}}{dt} \times \mathbf{K} = \frac{k}{h}\frac{d\mathbf{L}}{dt}, \tag{14.19}$$

Now, integrating (14.19), (since $\mathbf{K}$, k, h are constant),

$$\mathbf{V} \times \mathbf{K} = \frac{k}{h}(\mathbf{L} + \mathbf{E}), \tag{14.20}$$

where $\mathbf{E}$ is a constant vector in the plane of action. Take the dot product with $\mathbf{X}$:

$$\mathbf{X} \cdot \mathbf{V} \times \mathbf{K} = \frac{k}{h}\mathbf{X} \cdot (\mathbf{L} + \mathbf{E}).$$

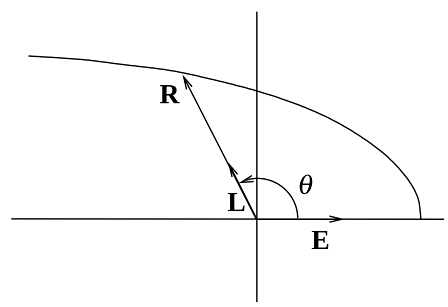
Observe that

$$\mathbf{X} \cdot \mathbf{V} \times \mathbf{K} = (\mathbf{X} \times \mathbf{V}) \cdot \mathbf{K} = h\mathbf{K} \cdot \mathbf{K} = h \ ,$$

since $\mathbf{X} \times \mathbf{V} = h\mathbf{K}$. This gives

$$\mathbf{X} \cdot (\mathbf{L} + \mathbf{E}) = \frac{h^2}{k}. \tag{14.21}$$

The motion therefore always satisfies this equation. We now show that it describes an ellipse. We move to polar coordinates, taking the origin at the Sun, and the ray $\theta = 0$ in the direction of $\mathbf{E}$.



We have: $\mathbf{E} = e\mathbf{I}$, $\mathbf{L} = \cos\theta\mathbf{I} + \sin\theta\mathbf{J}$ and $\mathbf{X} = r\mathbf{L}$. Then $\mathbf{X} \cdot \mathbf{L} = r$, $\mathbf{X} \cdot \mathbf{E} = r\cos\theta$ and equation (14.21) becomes

$$r(1 + e\cos\theta) = \frac{h^2}{k} \tag{14.22}$$

which is the equation of a conic with eccentricity e, the x-axis as major axis and a focus at the origin. If $e < 1$ this is an ellipse, if $e = 1$ a parabola, and if $e > 1$ a hyperbola. Since the planets have recurring orbits, the only possibility is that of an ellipse, $e < 1$.

Finally, we can derive Kepler's third law as a consequence, either of his first and second laws, or Newton's Law of Gravitational Attraction.

Proposition 14.4. Let T be the length of the planetary year, a and b the lengths of the half-major and half-minor axes of the orbit of the planet. Then

$$T^2 = \frac{4\pi^2}{k}a^3 \ . \tag{14.23}$$

The important thing here is that the equation which asserts Kepler II,

$$\frac{dA}{dt} = \frac{1}{2}h \ ,$$

allows us to relate the area (πab) of the ellipse with the time it takes to sweep out one full orbit, the planetary year T. Integrate with respect to t over a full cycle:

$$\pi ab = \frac{1}{2}hT \ . \tag{14.24}$$

The form of the equation of the orbit (14.22) allows us to replace h in this equation by the gravitational constant k relating the planet to the sun. We know that the orbit crosses the major axis at the points $\theta = 0$ and $\theta = \pi$, and the distance between these two points is the major axis. This distance (the length of the major axis) is the sum of the values of r at $\theta = 0, \ \pi$:

$$2a = \frac{h^2}{k(1+e)} + \frac{h^2}{k(1-e)} \ ,$$

so $a = h^2/k(1-e^2)$. Using the relation $b^2 = a^2(1-e^2)$, this gives us

$$\frac{b^2}{a} = \frac{h^2}{k} \ . \tag{14.25}$$

Combining (6.9) and (6.10) gives us the desired result:

$$T^2 = \Big(\frac{2\pi ab}{h}\Big)^2 = (2\pi ab)^2 \frac{a}{kb^2} = \frac{4\pi^2}{k} a^3 \ .$$

Note (in (14.22)) that since $h/2$ is the rate of change of the swept-out area, and e, the eccentricity of the ellipse, is related to the major and minor axes, these can be computed by terrestrial measurements for any planet (and in fact, had been by Kepler). Using his law of universal gravitiation and other known cosmological computations, Newton was able to estimate the constant k as well, so he was able to give good estimates for the actual orbits of the planets. These estimates agreed quite strongly with observations. Since Newton could not exclude the possibility of parabolic or hyperbolic orbits, he hypothesized that these were the orbits for some comets. Using the observations of the comet of 1681-2, (and Kepler's third law) he discovered its orbit to be elliptical and fairly accurately predicted the date of its return.

XV. Coordinates and Surfaces

15.1 Change of Coordinates in Two Dimensions

Suppose that E is an ellipse centered at the origin. If the major and minor axes are horizontal and vertical, as in figure 15.1, then the equation of the ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \ ,$$

where a and b are the lengths of the major and minor radii.

Figure 15.1

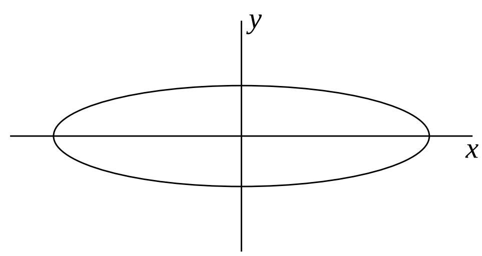
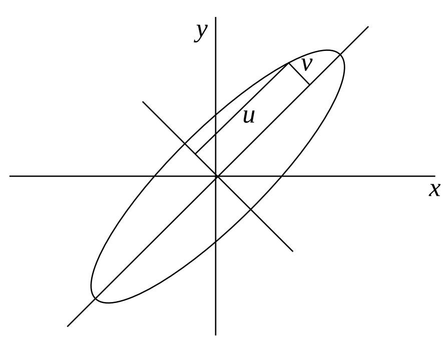


Figure 15.2



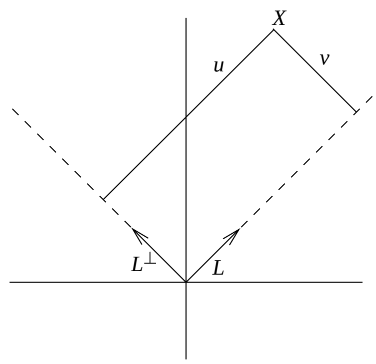
However, if the axes of E are neither horizontal nor vertical, as in figure 15.2, then we do not have this simple form of the equation. What we do know, since the curves in figures 15.1 and 15.2 are the same but for a rotation, is this: for u and v as shown in figure 15.2,

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

Now, the point here is that u and v can be expressed in terms of the cartesian coordinates, and in turn, x and y can be determined from u and v. This replacement of one pair of variables which determine a point by another is called a **change of coordinates**. We now show how to do this in the context of the ellipse of figure 15.2. First, recall Proposition 13.3, and in particular example 13.6.

Introduce a unit vector $\mathbf{L}$ in the positive direction of the major axis, so that $\mathbf{L}^\perp$ is in the direction of the minor axis. Now, any point X can be represented by the vector $\mathbf{X} = u\mathbf{L} + v\mathbf{L}^\perp$ (see figure 15.3).

Figure 15.3



The following proposition tells us how to find u and v.

Proposition 15.1. Let $\mathbf{L}$ be a unit vector in the plane. Then any vector $\mathbf{X}$ can be written as

$$\mathbf{X} = u\mathbf{L} + v\mathbf{L}^\perp \quad \text{where} \quad u = \mathbf{X} \cdot \mathbf{L} \ , \quad v = \mathbf{X} \cdot \mathbf{L}^\perp \ .$$

To see this, we calculate the dot products;

$$\mathbf{X} \cdot \mathbf{L} = (u\mathbf{L} + v\mathbf{L}^\perp) \cdot \mathbf{L} = u\mathbf{L} \cdot \mathbf{L} + v\mathbf{L}^\perp \cdot \mathbf{L} = u \ ,$$

$$\mathbf{X} \cdot \mathbf{L}^\perp = (u\mathbf{L} + v\mathbf{L}^\perp) \cdot \mathbf{L}^\perp = u\mathbf{L} \cdot \mathbf{L}^\perp + v\mathbf{L}^\perp \cdot \mathbf{L}^\perp = v \ ,$$

since $\mathbf{L} \cdot \mathbf{L} = 1,\ \mathbf{L}^\perp \cdot \mathbf{L}^\perp = 1,\ \mathbf{L} \cdot \mathbf{L}^\perp = 0$. In this way, if we are given a geometric description of the ellipse, we can find its equation in the cartesian coordinates $x,\ y$.

Example 15.1. Let E be the ellipse centered at the origin, with major radius of length 5, major axis the line $3x - 4y = 0$, and minor radius of length 2. Find the equation of the ellipse.

The point (4,3) is on the given line, so $4\mathbf{I} + 3\mathbf{J}$ lies in the direction of the major axis. The length of this vector is 5, so we can take the unit vector in the direction of the major axis to be $\mathbf{L} = (4\mathbf{I} + 3\mathbf{J})/5$. Thus, in the context of the above discussion,

$$\mathbf{L} = \frac{4}{5}\mathbf{I} + \frac{3}{5}\mathbf{J} \ , \quad \mathbf{L}^\perp = -\frac{3}{5}\mathbf{I} + \frac{4}{5}\mathbf{J} \ ,$$

and a point $\mathbf{X} = x\mathbf{I} + y\mathbf{J}$ is on the ellipse if and only if

$$\frac{u^2}{5^2} + \frac{v^2}{2^2} = 1 \quad \text{where} \quad u = \mathbf{X} \cdot \mathbf{L} \ , \quad v = \mathbf{X} \cdot \mathbf{L}^\perp \ .$$

Now, $\mathbf{X}\cdot\mathbf{L} = (4/5)x + (3/5)y,\ \mathbf{X}\cdot\mathbf{L}^{\perp} = -(3/5)x + (4/5)y$, so the equation of the ellipse is

$$\frac{[(4/5)x + (3/5)y]^2}{25} + \frac{[-(3/5)x + (4/5)y]^2}{4} = 1\ .$$

This simplifies to $.1156x^2 - .2016xy + .1744y^2 = 1$, or $289x^2 - 504xy + 436y^2 = 2500$.

In general, it happens that, in solving a particular problem, the situation can be easily realized in variables adapted to the problem, but the solution requires presentation in terms of an initial cartesian coordinate system. For example, the above ellipse is easily described in terms of the variables u and v adapted to its axes, but to realize the ellipse by an equation, we had to represent u and v in terms of x and y. We now state the proposition which gives the procedure used in example 15.1.

Proposition 15.2. Given a unit vector $\mathbf{L} = \cos\theta\mathbf{I} + \sin\theta\mathbf{J}$ we can write any vector $\mathbf{X} = x\mathbf{I} + y\mathbf{J}$ as $\mathbf{X} = u\mathbf{L} + v\mathbf{L}^{\perp}$, where $u = \mathbf{X}\cdot\mathbf{L}$ and $v = \mathbf{X}\cdot\mathbf{L}^{\perp}$; that is

$$u = x\cos\theta + y\sin\theta\ , \quad v = -x\sin\theta + y\cos\theta\ . \tag{15.1}$$

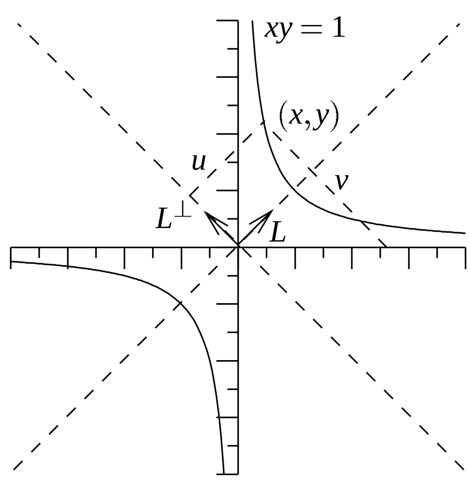
We refer to equations (15.1) as a **change of coordinate** by **rotation** through the angle θ. We can also reverse the roles of these variables and return to $x,\ y$ from $u,\ v$ just by rotating back through an angle $-\theta$. This gives us the equations

$$x = u\cos\theta - v\sin\theta\ , \quad y = u\sin\theta + v\cos\theta\ . \tag{15.2}$$

Of course, this is just what we get by solving equations (1) for x and y in terms of u and v.

Example 15.2. The curve $xy = 1$ is symmetric about the axes $x = y,\ x = -y$. Write the curve in coordinates $u,\ v$ relative to these axes, as in figure 15.4.

Figure 15.4



Since the line $x = y$ makes an angle of 45° with the horizontal, the change of coordinates is accomplished by a rotation through 45°. Since $\cos(45°) = \sin(45°) = 1/\sqrt{2}$, we have the relations

$$u = \frac{x+y}{\sqrt{2}} \ , \quad v = \frac{-x+y}{\sqrt{2}} \ , \quad x = \frac{u-v}{\sqrt{2}} \ , \quad y = \frac{u+v}{\sqrt{2}} \ .$$

Substituting for x and y in terms of u and v in $xy = 1$, we get

$$xy = [\frac{u-v}{\sqrt{2}}][\frac{u+v}{\sqrt{2}}] = 1 \quad \text{leading to} \quad u^2 - v^2 = 2 \ ,$$

the equation of a hyperbola in the $u,\ v$ coordinates.

As we have seen in the above examples, a hyperbola or ellipse leads to a quadratic equation, which will have a nonzero xy term if the axes are not horizontal and vertical. This is always true; as well as the reverse: any quadratic equation is the equation of a conic curve. We now see how to find the standard description of the conic from the equation (first with an example).

Example 15.3. Let C be the curve given by the equation

$$x^2 - xy = 12 \ . \tag{15.3}$$

Find coordinates which put this in standard form.

We want to make a substitution of the form (15.2) so that the coefficient of the uv term is 0. Making the substitution gives us

$$(u\cos\theta - v\sin\theta)^2 - (u\cos\theta - v\sin\theta)(u\sin\theta + v\cos\theta) = 12 \ . \tag{15.4}$$

The coefficient of uv is

$$-2\cos\theta\sin\theta - \cos^2\theta + \sin^2\theta \quad \text{or} \quad -\sin(2\theta) - \cos(2\theta) \ .$$

This is zero when $\tan(2\theta) = -1$, or $\theta = -\pi/8$. For this value of θ we now compute the coefficients A and B of u^2 and v^2:

$$A = \cos^2\theta - \cos\theta\sin\theta = \frac{1}{2}(1 + \cos(2\theta) - \sin(2\theta)) = \frac{1+\sqrt{2}}{2} \ ,$$

$$B = \sin^2\theta + \cos\theta\sin\theta = \frac{1}{2}(1 - \cos(2\theta) + \sin(2\theta)) = \frac{1-\sqrt{2}}{2} \ ,$$

so that the equation for the curve in the $u,\ v$ coordinates is

$$(\sqrt{2}+1)u^2 - (\sqrt{2}-1)v^2 = 24 \ ,$$

the equation of a hyperbola.

Following this example, given any quadratic equation in x and y:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{15.5}$$

we can find a rotation which eliminates the cross term. The resulting equation in the new variables $u,\ v$ is that of a conic section. Of course, there will be exceptional cases; for example the equation $x^2+y^2+1=0$ has no solutions, and the equation $x^2-y^2=0$ is a pair of straight lines. But, if (5) defines a curve, it must be an ellipse, hyperbola or parabola. If we introduce the new variables u and v by a rotation through an angle θ, the equation in the new coordinates is still quadratic in u and v; that is, the equation is of the form

$$A'u^2+B'uv+C'v^2+D'u+E'v+F'=0 \tag{15.6}$$

where the new coefficients are expressed in terms of θ and the old ones. By setting $B'=0$, we see how to choose θ. So, let's make the substitution (15.2) in the equation (15.6). The part which is purely quadratic is

$$A(u\cos\theta-v\sin\theta)^2+B(u\cos\theta-v\sin\theta)(u\sin\theta+v\cos\theta)+C(u\sin\theta+v\cos\theta)^2 \ .$$

The coefficient of uv in this expression is

$$B'=-2A\cos\theta\sin\theta+B(\cos^2\theta-\sin^2\theta)+2C\cos\theta\sin\theta \ .$$

Set this to zero and solve for θ. Using double angle formulas, the equation is

$$(-A+C)\sin(2\theta)+B\cos(2\theta)=0\ , \quad \text{or} \quad \tan(2\theta)=\frac{B}{A-C}\ . \tag{15.7}$$

If $A=C$, the denominator is zero, so we take $2\theta=\pi/2$, or $\theta=\pi/4$. Now, the equation (15.6) (with $B'=0$) is of the form considered in chapter 11, and can be put in standard form with center at some other point.

Example 15.4. Let C be the curve given by the equation

$$x^2-2\sqrt{3}xy-3y^2+6\sqrt{3}x+6y=16\ . \tag{15.8}$$

Find coordinates which put C in standard form.

First, we use (15.7) to find the angle of rotation:

$$\tan(2\theta)=\frac{-2\sqrt{3}}{1-3}=\sqrt{3}\ , \quad \text{so} \quad 2\theta=\frac{\pi}{3}\ , \quad \theta=\frac{\pi}{6}\ ,$$

and $(\sin\theta=1/2,\ \cos\theta=\sqrt{3}/2)$, the substitution (15.2) is

$$x=\frac{\sqrt{3}u-v}{2}\ , \quad y=\frac{u+\sqrt{3}}{2}\ .$$

We do this in two steps. First, the quadratic terms of (8) are, in the coordinates $u,\ v$:

$$\frac{1}{4}[(3u^2-2\sqrt{3}uv+v^2)-2\sqrt{3}(\sqrt{3}u^2+2uv+\sqrt{3}v^2)+3(u^2+2\sqrt{3}uv+3v^2]$$

which reduces to v^2. Now incorporate the linear terms of (8) in terms of $u,\ v$:

$$v^2+\frac{1}{2}[6\sqrt{3}(\sqrt{3}u-v)+6(u+\sqrt{3}v)=16\ ,$$

which can be put in the standard form

$$v^2 = -\frac{15}{2}(u - \frac{32}{15}) \ .$$

Thus the curve is a parabola, with axis at an angle of $\pi/3$ with the x-axis, which opens downward.

We summarize this discussion as follows.

Proposition 15.3. A curve given by the equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{15.9}$$

is a conic section. Rotate coordinates by the angle θ given by

$$\tan(2\theta) = \frac{B}{A - C} \ ; \tag{15.10}$$

that is, make the substitution in (9):

$$x = u\cos\theta - v\sin\theta \ , \quad y = u\sin\theta + v\cos\theta \ . \tag{15.11}$$

There is no uv term, so after completing the squares, the equation is in standard form. In particular, the axes of the conic are at an angle θ with the coordinate axes.

Proposition 15.4. For a curve given by equation (15.9),
If $B^2 - 4AC < 0$, the curve is an ellipse.
If $B^2 - 4AC > 0$, the curve is a hyperbola.
If $B^2 - 4AC = 0$, the curve is a parabola.

To indicate why this is true, let us consider just the quadratic terms and start with an equation of the form

$$Ax^2 + Bxy + Cy^2 = 1 \ . \tag{15.12}$$

Supposing $A > 0$, we complete the square for the first two terms, rewriting (12) as

$$A(x^2 + 2\frac{By}{2A}x + (\frac{By}{2A})^2) - \frac{B^2}{2A}y^2 + Cy^2 = 1 \ ,$$

or

$$A(x + \frac{By}{2A}y)^2 + \frac{4AC - B^2}{4A}y^2 = 1 \ .$$

If the coefficients of the squared terms are both positive, then there are no solutions for large x and y, so the curve is an ellipse. On the other hand, if the signs of the coefficients are different, there are always solutions for large x and y, so the curve must be a hyperbola. Thus the shape is determined by the sign of $4AC - B^2$, and if we carefully follow through the argument, we arrive at proposition 15.4.

Example 15.5. Describe the curve $x^2 + xy + y^2 = 1$.

Since $B^2 - 4AC = -1 < 0$, this is an ellipse. Since $A = C$, we need to rotate coordinates by $\pi/4$. We make the subsititution

$$x = \frac{u - v}{\sqrt{2}} \ , \quad y = \frac{u + v}{\sqrt{2}}$$

getting

$$\frac{u^2 - 2uv + v^2}{2} + \frac{u^2 - v^2}{2} + \frac{u^2 + 2uv + v^2}{2} = 1 \ ,$$

which reduces to $3u^2 + 2v^2 = 2$.

Problems 15.1

1. Consider the line L in the plane given by the equation $2x + 5y + 10 = 0$. Find a base $\{\mathbf{U}, \mathbf{V}\}$ with $\mathbf{U}$ parallel to L, and $\mathbf{V}$ counterclockwise to $\mathbf{U}$. Find the equation of the line in coordinates $\{u, v\}$ relative to the base $\{\mathbf{U}, \mathbf{V}\}$.

2. A conic in the plane is given by the equation

$$x^2 - 2xy + y^2 + 2x - y = 0 \ .$$

a) What conic is it?

b) At what angle to the x-axis are the axes of the conic?

3. A conic in the plane is given by the equation $5x^2 - xy + y^2 = 50$.

a) What conic is it?

b) At what angle to the x-axis are the axes of the conic?

4. A parabola has its vertex at the origin, and its focus at the point (3,4). Give the equation of the parabola.

5. An ellipse has center at the point (2,1), and its major axis is the line $x+y = 3$. Its major radius is 3 and its minor radius is 1. What is the equation of the ellipse?

6. a) Write down an equation of a hyperbola whose major axis is the line $y = x$.

b) Write down an equation of an ellipse (not a circle) whose major axis is the line $y = x$.

15.2. Special Coordinate Systems

Often a problem can be seen as that of understanding the motion of a particle relative to a fixed point or a fixed axis. In these cases it is useful to express everything in coordinates which emphasize positions relative to the fixed point or axis.

Polar coordinates

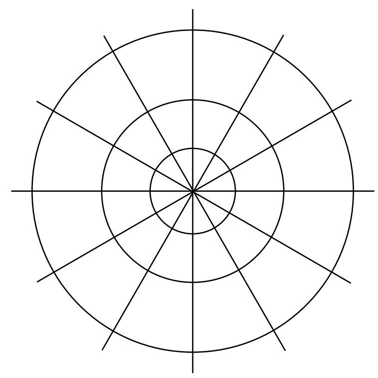
First, we recall, from Chapter 11, polar coordinates in the plane. We consider the fixed point as the origin of these coordinates, and take the positive x-axis as the "zero" direction. Then any other direction is described by the angle between it and the positive x axis, which we denote as θ. The distance of a point on this line from the origin is denoted r. These equations relate the cartesian coordinates (x, y) with the polar coordinates r, θ:

(15.9) $$x = r\cos\theta \ , \quad y = r\sin\theta \ , \qquad r = \sqrt{x^2 + y^2} \ , \quad \theta = \arctan\frac{y}{x}$$

Polar coordinates have some ambiguities. Every value of (r, θ) determines a point in the plane. However, if $r = 0$, the point is the origin, and θ doesn't make sense. Secondly, the values (r, θ)

and $(r, \theta + 2\pi)$, and in fact, $(r, \theta + 2n\pi)$ for any n give the same point. The curve $\theta = a$ is the ray of angle a emanating from the origin, and the curve $r = a$ is the circle of radius a centered at the origin (see figure 15.5).

Figure 15.5

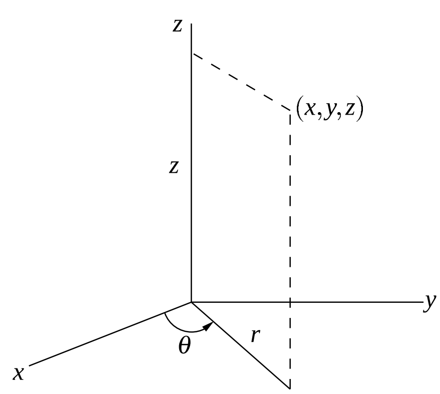


In three dimensions, we introduce two new coordinate systems, the first oriented toward a fixed axis, the z-axis, and the second oriented toward the origin.

Cylindrical coordinates

Here a point is described by its z-coordinate and its polar coordinates in the plane (see figure 15.6).

Figure 15.6



The formulas for the change from cartesian coordinates are

$$x = r\cos\theta\ , \quad y = r\sin\theta\ , \quad z = z\ , \quad r = \sqrt{x^2 + y^2}\ , \quad \theta = \arctan\frac{y}{x} \tag{15.10}$$

The equation $r = at$ is a circular cylinder of radius a centered along $z = 0$; $\theta = a$ describes the half plane with its edge along $z = 0$ making an angle a with the xz-plane, and $z = a$ is a horizontal plane . (See figures 15.7 and 15.8)

Figure 15.7

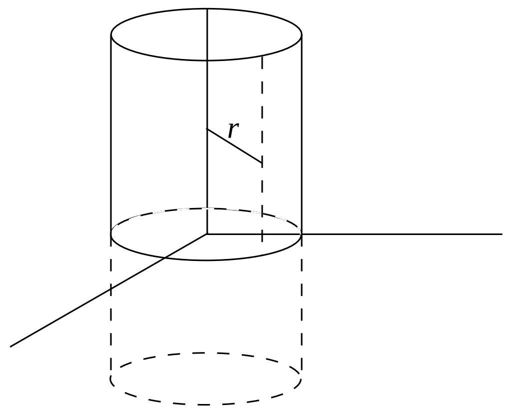
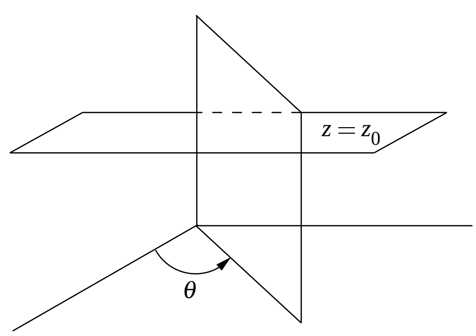


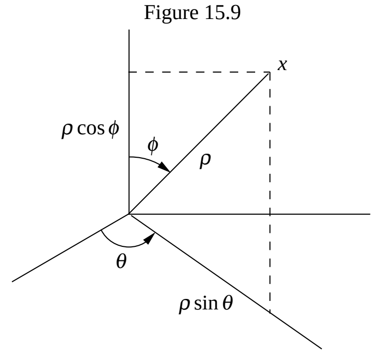
Figure 15.8



Spherical coordinates

These coordinates are oriented toward the origin, so that a point is described by its distance ρ from the origin and the ray from the origin on which it lies. We describe the ray by the angle ϕ it makes with the z-axis and the angle θ it makes with the xz-plane (see figure 15.9).

Figure 15.9



We can read off from figure 15.9 the equations relating spherical coordinates with polar coordinates;

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

Note that, although θ ranges through a whole circle, ϕ ranges from 0 to π. The curve $\rho = a$ is the sphere of radius a centered at the origin, $\theta = a$ a half-plane, and $\phi = a$ the half-cone with axis the z-axis, making the angle a with its axis (see figure 15.10=15.12).

Figure 15.10

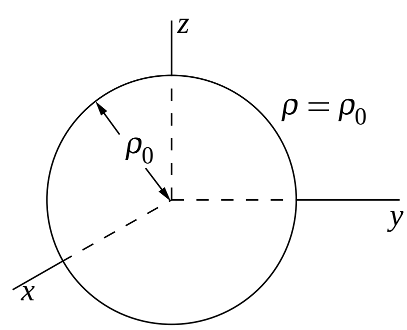


Figure 15.11

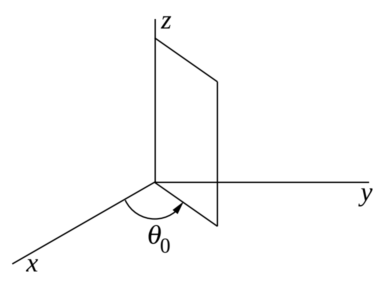
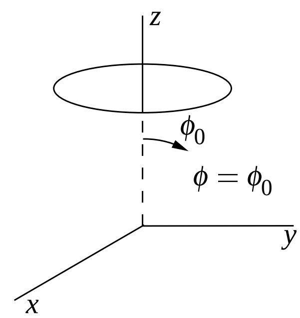


Figure 15.12



Example 15.6. Describe the curves $C_1 : \phi = \phi_0,\ \rho = 1$, $C_2 : \theta = \theta_0,\ \rho = R$. Give their equations in cylindrical and cartesian coordinates.

The curve C_1 is the intersection of the cone $\phi = \phi_0$ with the sphere of radius R. If we think of this sphere as the globe, C_1 is a circle of latitude. The radius of this circle is $R \sin \phi_0$, and its center is on the z-axis, at a distance $R \cos \phi_0$ from the origin. In cylindrical coordinates, this curve lies on the plane $z = R \cos \phi_0$, and the cylinder $r = R \sin \phi_0$; these are then the equations of C_1 in cylindrical coordinates. In rectangular coordinates the equations are $x^2 + y^2 = (R \sin \phi_0)^2,\ z = R \cos \phi_0$.

The curve C_2 is the intersection of the plane $\theta = \theta_0$ with the sphere of radius R. If we think of this sphere as the globe, C_2 is a circle of longitude. Its center is the origin and its radius is R. In cylindrical coordinates, C_2 is given by the equations $r^2 + z^2 = R^2,\ \theta = \theta_0$. In cartesian coordinates, the equations are $x^2 + y^2 + z^2 = R^2, y = x \tan \theta_0$. (When $\theta_0 = \pi/2$ the second equation is $x = 0$.)

Problems 15.2

1. Show that the intersection of a plane with a sphere is a circle.

2. Consider the set of all points P in space such that the vector from O to P has length 2 and makes an angle of 45 degrees with $\mathbf{I} + \mathbf{J}$.
   a) What kind of geometric object is this set?
   b) Give equations in cartesian coordinates for this set.

3. Write down the equations of the paraboloid of revolution $z = x^2 + y^2$ in cylindrical and spherical coordinates.

4. Describe the surfaces given in cylindrical coordinates by the equations a) $r = z$; b) $\theta = z$.

5. Show that the equation
$$\mathbf{X}(t) = t \cos t\mathbf{I} + t \sin t\mathbf{J} + t\mathbf{K}$$
parametrizes the curve of intersection of the two surfaces of problem 4. Show that this curve lies on the surface $r = \theta$ (cylindrical coordinates).

6. Describe the curve in the plane given in polar coordinates by $r^2 \sin(2\theta) = 1$. Write its equation in cartesian coordinates.

15.3 Surfaces: Graphs and Level curves

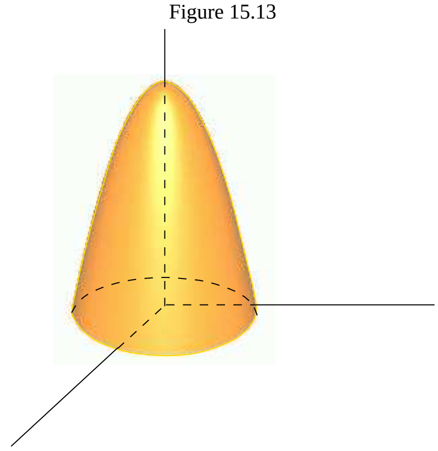
A relation among the variables x, y, z defines a *surface* in three dimensions: the set of all x, y, z which satisfy the equation. For example, we have seen that a linear relation $ax + by + cz + d = 0$ is the equation of a plane; that is, the set of all points (x, y, z) which satisfy that relation is a plane. Similarly, the sphere of radius R has the equation $x^2 + y^2 + z^2 = R^2$. As we have observed already, it is sometimes difficult to visualize a surface given by an equation in three variables. In this section we shall discuss various ways of visualizing surfaces.

In the case that the relation can be solved for z in terms of x and y, then the surface is the *graph* of the function $z = f(x, y)$. In this case it is a good idea to try to sweep out the surfaces by the curves of intersection of the curve with the planes $z =$ const. These are the *level curves* of the surface. Then we can imagine the surface as a stack of these level sets. In order to understand how the level sets stack, we may want to look at representative *profiles*: these are the curves of intersection of the surface with planes perpendicular to the xy-plane.

Example 15.7. Draw the level curves of the surface $z = 4 - x^2 - y^2$, and sketch the surface.

We see first of all, since $x^2 + y^2$ is never negative, that $z \le 4$. The level surface $z = 4$ is just the origin $(0, 0)$, but as z decreases from 4 we get a family of circles centered at the origin of ever increasing radius (the radius is $\sqrt{4 - z}$). Our surface then is a stack of circles. To see the shape of the stack, we look at a representative profile: the intersection of the surface with a plane through the z-axis. For example, for $y = 0$ we get the parabola $z = 4 - x^2$, and now can safely sketch the graph (figure 15.13).

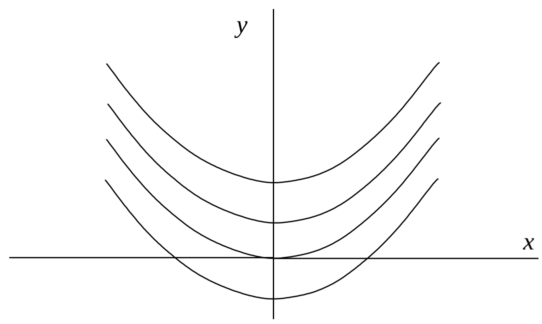
Figure 15.13



Example 15.8. Do the same for the surface $z = y - x^2$.

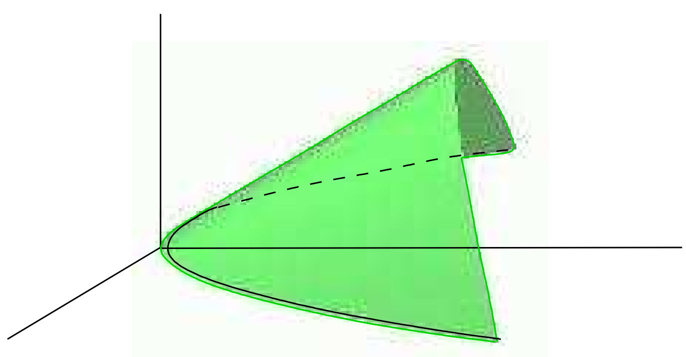
Setting z equal to the constant z_0, we get the parabola $x^2 = y - z_0$, so the level curves are the family of parabolas with axis the y-axis, opening upward, with vertex at $(0, z_0)$. We have shown typical level curves in figure 15.14.

Figure 15.14



Thus the surface is is a stack of parallel parabolas with vertex moving linearly up the y-axis; that is, the vertices lie on the line $y = z,\ x = 0$. To get a further idea of the shape of the surface we look at a profile $y =$ constant , say $y = 0$. There the surface is given by the parabola $z = -x^2$, a parabola opening downward. Putting this information together we get figure 15.15.

Figure 15.15



As is clear from these examples, sketching surfaces is an imprecise science, and the configuration of level sets gives an idea of the shape, but not very precise. If we draw a large number of level sets on the xy plane, we can observe that at points where the level sets are close together, the surface is steep, and where they are far apart, the surface is quite flat. This is illustrated in figures 15.16 and 15.17: figure 15.16 is that of the surface, and figure 15,17 the configuration of its level sets in the xy-plane.

Figure 15.16

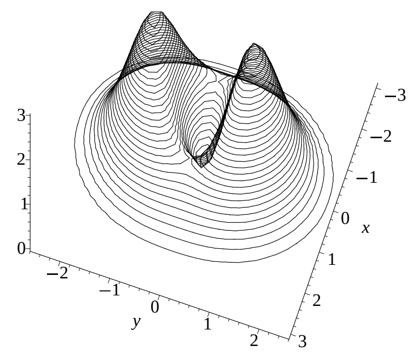
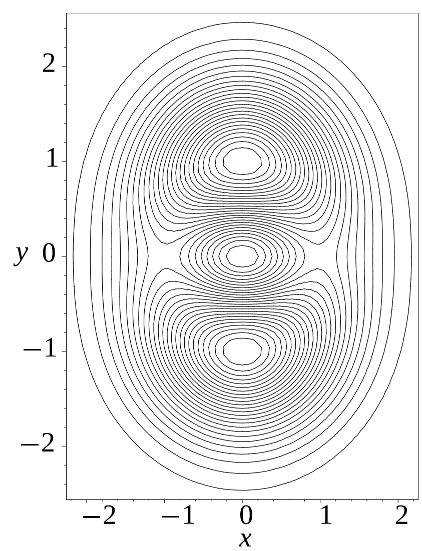


Figure 15.17



Problems 15.3.

1. a) Draw some typical level curves in the (x, y)-plane for the function

$$f(x, y) = (1 + x^2 + y^2)^{-1}.$$

b) Sketch the surface $z = f(x, y)$.

2. Consider the helix $\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + t\mathbf{K}$. For each t, let L_t be the line perpendicular to the z-axis intersecting the helix at $\mathbf{X}(t)$. Find the equation of the surface swept out by the lines L_t. What are its level sets? (See problems 4 and 5 of the preceding section.)

3. Consider the graph $z = f(x, y)$, where $f(x, y) = x^2 - y^2$. Show that the level sets $f(x, y) = a$ for $a > 0$ are conics centered at the origin. Find the foci of the level sets. How would you describe the level sets $f(x, y) = a$ for $a < 0$?

4. Consider the graph $z^2 = x^2 + 2y^2$. Show that the level sets $z = a,\ a \neq 0$ are conics centered at the origin. Find the foci of the level sets.

5. Consider the graph $z = f(x, y)$, where $f(x, y) = \exp(x^2+y^2)$. Describe the level sets $f(x, y) = a$.

6. Describe the surface $xyz = 1$.

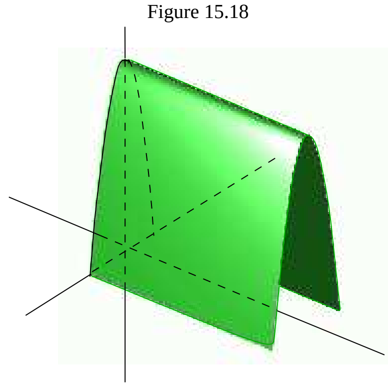
15.4 Cylinders and Surfaces of Revolution

Starting with a curve C in a plane Π, the surface swept out by the translates of this curve in the direction perpendicular to Π is the **cylinder** over the curve C. This is the case when the relation defining the surface S is independent of one of the variables. For example, the surface S given by the relation $x^2 - y^2 = 1$ is independent of z, so if $(x,\ y)$ is a point satisfying this relation, then all points $(x,\ y,\ z)$ are on the surface, so S is the cylinder over the hyperbola $C:\ x^2 - y^2 = 1$ in the xy-plane.

Example 15.9. Sketch the surface $z = 9 - x^2$.

Since the relation is independent of x we just draw the parabola given by this equation on the xz-plane, and extend it by lines parallel to the y-axis.

Figure 15.18



If a surface has the property that, for a particular line L, the intersection of the surface with a plane perpendicular to L is a circle centered on L, then the surface is a **surface of revolution** about the **axis** L. Suppose that S is a surface of revolution about the z axis. Then the intersection of S with the half-plane $y = 0,\ x > 0$ completely determines the surface. Let C be that curve of intersection, and suppose it is given by the equation $z = f(x)$. If $(x_0, 0, z_0)$ is a point on C, then every point on the plane $z = z_0$ whose distance from the z axis is x_0 is on S. That is, if $\sqrt{x^2 + y^2} = x_0$, then $z_0 = f(\sqrt{x^2 + y^2})$, so (x, y, z_0) is on the surface S. Thus the equation of a surface $S$of revolution is given by the equation $z = f(x)$ defining its profile just by replacing x by $\sqrt{x^2 + y^2}$: S is the surface $z = f(\sqrt{x^2 + y^2})$.

Example 15.10. Sketch the surface $z^2 = x^2 + y^2$.

This is the surface obtained by revolving the curve $z^2 = x^2$ about the z axis. Since that curve consists of the two lines $z = \pm x$, we get the cone in figure 15.24 (see the end of the chapter).

15.5 Quadric surfaces

These are the surfaces which are given by a quadratic relation among the variables. By completing the square, and - if necessary - rotating the axes, we can reduce every quadric surface to one of the surfaces in this section. It is a good exercise to trace out these surfaces using the technique of level sets and profiles from the preceding sections. The figures are collected together at the end of the text. It is essential to become familiar with these surfaces, for they are the fundamental examples for the rest of the course.

Sphere of radius R $\qquad x^2 + y^2 + z^2 = R^2 \qquad$ (Figure 15.19)

The sphere is symmetric about all axes and planes through its center. The intersection of the sphere with any plane is a circle. The intersection of the sphere with a plane through its center is a **great circle**.

Ellipsoid $\qquad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1 \qquad$ (Figure 15.20)

This equation is just that of the sphere, but with the coordinates $x,\ y,\ z$ replaced by $x/a,\ y/b,\ z/c$. The effect is that the sphere has been dilated in the $x-,\ y-,\ z-$directions by the factors $a,\ b,\ c$ respectively, producing the ellipsoid with vertices $(\pm a, 0, 0),\ (0, \pm b, 0),\ (0, 0, \pm c)$ along the coordinate axes.

Hyperboloid of one sheet $\qquad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1 \qquad$ (Figure 15.21)

First, let's consider the case $a = b = c = 1$. Then the equation can be written as

(15.11) $$x^2 + y^2 = 1 + z^2$$

so the intersection of this surface with the plane $z = z_0$ is a circle centered at the origin of radius $\sqrt{1 + z_0^2}$. This is a stack of circles of ever increasing radius as we move away from the xy-plane. If we set $y = 0$, we get the profile $x^2 - z^2 = 1$: a hyperbola, and thus figure 15.21. We could also have come to this figure by observing that this is the surface of revolution of this hyperbola.

Note that if we set $x = 1,\ y = z$ in equation (15.11) we get an identity. Thus this line lies on the surface. More importantly, since this is a surface of revolution, if we revolve this line about the z-axis, we generate the surface (see figure 15.21). The line $x = 1,\ y = -z$ also lies on the surface and generates it by rotation.

Now, for general a, b, c, the level set of the surface at $z = z_0$ is the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{z_0^2}{c^2}$$

Thus, the surface is a stack of similar ellipses of size increasing as we move away from the xy-plane, with, again, a hyperbolic profile along each plane through the z-axis. We observe that this surface has the same shape as that in the case $a = b = c = 1$, except for dilations along the coordinate axes.

Again, this surface is generated by lines as in the first case, for figure 15.21 still describes the surface, but for a change in scale in the coordinate directions. In particular, the pair of lines lying on the surface that go through the point $(a, 0, 0)$ are the lines

$$x = a\ , \qquad \frac{y}{b} = \pm\frac{z}{c}$$

Hyperboloid of two sheets $\qquad \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1 \qquad$ (Figure 15.23)

The level sets $z = z_0$ of this surface are the hyperbolas

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 + \frac{z_0^2}{c^2}$$

whose vertices lie on the x-axis, but move further and further from the origin as z_0 moves away from the origin. To get a better view of this surface, we look at the level curves $x = x_0$:

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = \frac{x_0^2}{a^2} - 1$$

There is no curve for $|x| < a$, and for larger values of $|x|$, we get a family of ever increasing ellipses. This leads quickly to figure 15.22.

Elliptical Cone $\qquad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0 \qquad$ (Figure 15.24)

Here the level curves are the ever-widening ellipses

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

and the profile in the plane $x = 0$ is given by

$$\frac{y^2}{b^2} = \frac{z^2}{c^2} \quad \text{or} \quad \frac{y}{b} = \pm\frac{z}{c}\ ,$$

a pair of lines. The profile in the plane $y = 0$ is the pair of lines

$$\frac{x}{a} = \pm\frac{z}{c} \ . \tag{15.12}$$

This gives us figure 15.24. If $a = b$, the level sets are circles, and the profiles are the same, for the surface is the surface of revolution of the curves given by (15.12).

Now, the above list exhausts all possibilities for surfaces given by quadratic equations of the form

$$Ax^2 + By^2 + Cz^2 + Dx + Ey + Fz + G = 0 \ , \tag{15.13}$$

with none of $A,\ B,\ C,$ equal to zero. Each of the surfaces illustrated in the figures has the origin as center, and has a particular axis (the z-axis, except in the case of the hyperboloid of two sheets, in which case it is the x-axis). For the general equation (15.13), the center might be any point, and the axis could be one of the other coordinate axes.

For, if we complete the square in each of the variables in equation (15.13) we end up with an equation of the form

$$A(x - x_0)^2 + B(y - y_0)^2 + C(z - z_0)^2 = H$$

If $H = 0$ we have two cases: $A,\ B,\ C,$ all of the same sign, in which case there is no surface. Otherwise we get a cone. The axis is identified by the coefficient which is of a sign different from the other two.

If $H \neq 0$, we can divide by H, leading to one of the previous cases. The number of negative signs determines the surface; no negatives: ellipsoid; one negative: hyperboloid of one sheet; two negative: hyperboloid of two sheets.

Finally, there are two more surfaces; corresonding to the cases where one or more of the coefficients of the quadratic term is zero

Elliptical Paraboloid $\qquad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = z \qquad$ (Figure 15.25)

We look at the level curves $z = z_0$. If $z_0 < 0$ we get no curve. For $z_0 > 0$ we get a family of ellipses of ever increasing size. The profile on the plane $x = 0$ is the parabola $y^2 = b^2 z$, and on the $y = 0$ the parabola $x^2 = a^2 z$. This gives us enough information for figure 15.25.

Hyperbolic Paraboloid $\qquad \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = z \qquad$ (Figure 15.26)

We look at the level curves $z = z_0$. We get the hyperbolas

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z_0 \ .$$

If $z_0 > 0$, the axis of the hyperbola is the x axis, but if $z_0 < 0$, the axis is the y-axis. As $|z_0|$ increases, the vertices move away from the z-axis. The level curve for $z_0 = 0$ consists of the two lines $x/a = \pm y/b$. To get a better grip on the surface, we look at its profiles. The profile for $y = 0$ is the parabola $z = x^2/a^2$, which has the z-axis as axis, and opens upwards. The profile for $x = 0$

is the parabola $z = -y^2/b^2$, which opens downward. putting this information together gives us figure 24.

We can get a different (and perhaps more readable) view of this surface by interchanging the x and y coordinates. Figure 15.26 is that of the hyperbolic paraboloid

$$-\frac{x^2}{a^2} + \frac{y^2}{b^2} = z \ .$$

Problems 15.5

1. Let L be the line $x = 1,\ z = 3y$. If we rotate the line about the z-axis, it describes a surface. Find the equation of that surface.

2. Sketch or describe the surface given by the equation

$$\frac{y^2}{4} - \frac{z^2}{9} = x^2.$$

3. Sketch or describe the surface given by the equation

$$\frac{x^2}{9} + \frac{z^2}{4} = y.$$

4. Sketch the surface $z^2 - x - y = 0$,

5. Sketch the surface $(x + z)^2 = (y + z)^2$,

6. Sketch the surface

$$z^2 + \frac{x^2}{4} - \frac{y^2}{16} = 1 \ .$$

Find a pair of intersecting lines that lie on the surface.

7 The surface

$$z^2 - \frac{x^2}{9} - \frac{y^2}{16} = 1$$

has two pieces $S^+,\ S^-$. Find the point p_+ on S^+ and the point p_- on S^- which minimize the distance between S^+ andS^-.

Figure 15.19: Sphere.

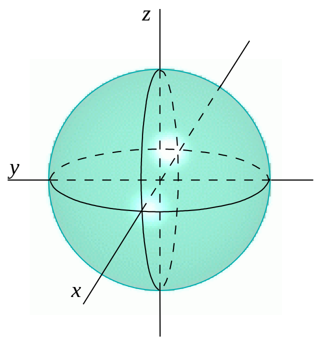


Figure 15.20: Ellipsoid

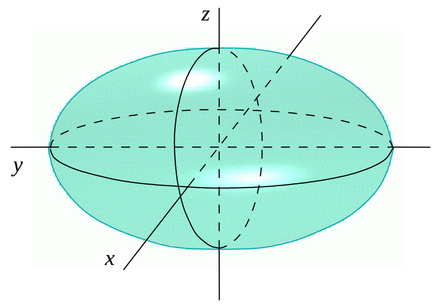


Figure 15.22: Hyperboloid of one sheet.

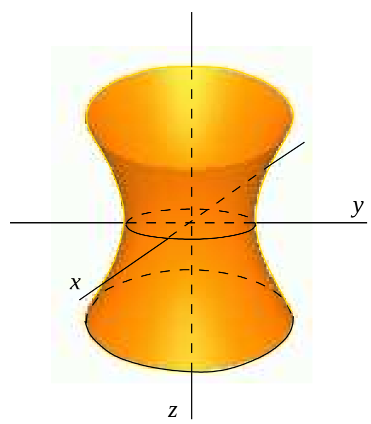


Figure 15.23: Hyperboloid of two sheets

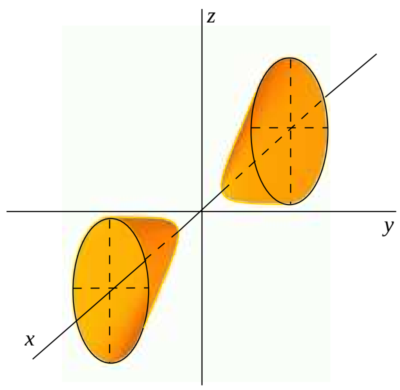


Figure 15.24: Elliptic cone

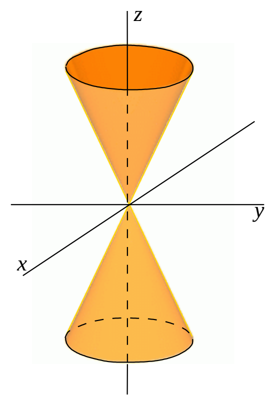


Figure 15.25: Elliptic paraboloid

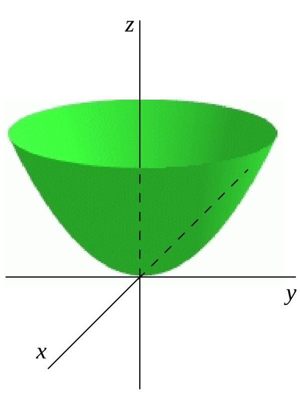
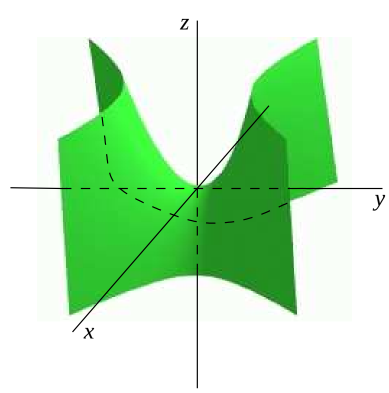


Figure 15.26: Hyperboloid paraboloid



XVI. Differentiable Functions of Several Variables

16.1 The Differential and Partial Derivatives

Let $w = f(x, y, z)$ be a function of the three variables x, y, z. In this chapter we shall explore how to evaluate the way w changes near a point (x_0, y_0, z_0), and make use of that evaluation. For functions of one variable, this led to the derivative: dw/dx is the rate of change of w with respect to x. But in more than one variable, the existence of more than one direction in which w changes makes this more complicated. In particular, we must see how to formulate the rates of change along each direction in which we move away from the base point (x_0, y_0, z_0). We start by using the one variable theory to define change in w with respect to one variable at a time. So, for each variable, we differentiate the function with respect to that variable, considering the other variables as held constant. These are the *partial derivatives* of the function.

Definition 16.1. Suppose we are given a function $w = f(x, y, z)$ defined near the point (x_0, y_0, z_0). The **partial derivative** of f **with respect to** x is

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0) = \frac{d}{dx} f(x, y_0, z_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0, z_0) - f(x_0, y_0, z_0)}{h} .$$

Similarly, if we keep x and z constant, we define the **partial derivative** of f **with respect to** y by

$$\frac{\partial f}{\partial y} = \frac{d}{dy} f(x_0, y, z_0) ,$$

and by keeping x and y constant, we define the **partial derivative** of f **with respect to** z by

$$\frac{\partial f}{\partial z} = \frac{d}{dz} f(x_0, y_0, z) .$$

Example 16.1. Find the partial derivatives of $f(x, y) = x(1 + xy)^2$.

Thinking of y as a constant, we have

$$\frac{\partial f}{\partial x} = (1 + xy)^2 + x(2(1 + xy)y) = (1 + xy)(1 + 3xy) .$$

Now, we think of x as constant and differentiate with respect to y:

$$\frac{\partial f}{\partial y} = x(2(1 + xy)x) = 2x^2(1 + xy) .$$

Example 16.2. The partial derivatives of $f(x, y, z) = xyz$ are

$$\frac{\partial f}{\partial x} = yz , \quad \frac{\partial f}{\partial y} = xz , \quad \frac{\partial f}{\partial z} = zy .$$

Of course, the partial derivatives are themselves functions, and when it is possible to differentiate the partial derivatives, we do so, obtaining higher order derivatives. More precisely, the partial derivatives are found by differentiating the formula for f with respect to the relevant variable, treating the other variable as a constant. Apply this procedure to the functions so obtained to get the **second partial derivatives**:

$$(16.1) \qquad \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial x}) \ , \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}), \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial y}) \ , \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(\frac{\partial f}{\partial y})$$

Example 16.3. Calculate the second partial derivatives of the function in example 1.

We have $f(x, y) = x(1 + xy)^2$, and have found

$$\frac{\partial f}{\partial x} = (1 + xy)(1 + 3xy) \ , \quad \frac{\partial f}{\partial y} = 2x^2(1 + xy) \ .$$

Differentiating these expressions, we obtain

$$\frac{\partial^2 f}{\partial x^2} = (1 + xy)(3y) + y(1 + 3xy) = 4y + 6xy^2$$

$$\frac{\partial^2 f}{\partial y \partial x} = (1 + xy)(3x) + x(1 + 3xy) = 4x + 6x^2 y$$

$$\frac{\partial^2 f}{\partial x \partial y} = 4x(1 + xy) + 2x^2 y = 4x + 6x^2 y$$

$$\frac{\partial^2 f}{\partial y^2} = 2x^2(x) = 2x^3 \ .$$

Notice that the second and third lines are equal. This is an important general fact: the mixed partials (the middle terms above) are equal when the second partials are continuous:

$$(16.2) \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

This is not easily proven, but is easily verified by many examples. Thus $\partial^2 f / \partial x \partial y$ can be calculated in whatever is the most convenient order. Finally, we note an alternative notation for partial derivatives;

$$f_x = \frac{\partial f}{\partial x}, \quad f_y = \frac{\partial f}{\partial y}, \quad f_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f_{xy} = f_{yx} = \frac{\partial^2 f}{\partial x \partial y} \ , \quad f_{yy} = \frac{\partial^2 f}{\partial y^2} \ , \text{ etc.}$$

Example 16.4. Let $f(x, y) = y \tan x + x \sec y$. Show that $f_{xy} = f_{yx}$.

We calculate the first partial derivatives and then the mixed partials in both orders:

$$f_x = y \sec^2 x + \sec y \ , \qquad f_y = \tan x + x \sec y \tan y$$

$$f_{yx} = \sec^2 x + \sec y \tan y \qquad f_{xy} = \sec^2 x + \sec y \tan y \ .$$

The partial derivatives of a function $w = f(x, y, z)$ tell us the rates of change of w in the coordinate directions. But there are many directions at a point on the plane or in space: how do we find these rates in other directions? More generally, if two or three variables are changing, how do we explore the corresponding change in w? The answer to these questions starts with the generalization of the idea of the differential as linear approximation.

For a function of one variable, a function $w = f(x)$ is differentiable if it is can be locally approximated by a linear function

$$w = w_0 + m(x - x_0) \tag{16.3}$$

or, what is the same, the graph of $w = f(x)$ at a point (x_0, y_0) is more and more like a straight line, the closer we look. The line is determined by its slope $m = f'(x_0)$. For functions of more than one variable, the idea is the same, but takes a little more explanation and notation.

Definition 16.2. Let $w = f(x, y, z)$ be a function defined near the point (x_0, y_0, z_0). We say that f is **differentiable** if it can be well- approximated near (x_0, y_0, z_0) by a linear function

$$w - w_0 = a(x - x_0) + b(y - y_0) + c(z - z_0) \ . \tag{16.4}$$

In this case, we call the linear function the **differential** of f at (x_0, y_0, z_0), denoted $df((x_0, y_0, z_0)$. It is important to keep in mind that the differential is a function of a vector at the point; that is, of the increments $(x - x_0, y - y_0, z - z_0)$.

If $f(x, y)$ is a function of two variables, we can consider the **graph** of the function as the set of points (x, y, z) such that $z = f(x, y)$. To say that f is differentiable is to say that this graph is more and more like a plane, the closer we look. This plane, called the **tangent plane** to the graph, is the graph of the approximating linear function, the differential.

For a precise definition of what we mean by "well" approximated, see the discussion in section 16.3. The following example illustrates this meaning.

Example 16.5. Let $f(x, y) = x^2 + y$. Find the differential of f at the point (1,3). Find the equation of the tangent plane to the graph of $z = f(x)$ at the point.

We have $(x_0, y_0) = (1, 3)$, and $z_0 = f(x_0, y_0) = 4$. Express $z - 4$ in terms of $x - 1$ and $y - 3$:

$$z - 4 = x^2 + y - 4 = (1 + (x - 1))^2 + (3 + (y - 3)) - 4$$

$$= 1 + 2(x - 1) + (x - 1)^2 + 3 + (y - 3) - 4 \ , \quad \text{simplifing to}$$

$$z - 4 = 2(x - 1) + y - 3 + (x - 1)^2 \ .$$

Comparing with (16.4), the first two terms give the differential. $(x - 1)^2$ is the error in the approximation. The equation of the tangent plane is

$$z - 4 = 2(x - 1) + y - 3 \quad \text{or} \quad z = 2x + y - 1 \ .$$

If we just follow the function along the line where $y = y_0,\ z = z_0$, then the linear approximation (16.4) along this line is just $w - w_0 = a(x - x_0)$. Comparing this with the one-variable theory, using definition 16.1, we conclude that a is the derivative of w in the x-direction, that is $a = \partial w/\partial x$. Similarly $b = \partial w/\partial y$ and $c = \partial w/\partial z$. Finally, since the variables $x,\ y,\ z$ are themselves linear, we have that dx *is* $x - x_0$,and so forth. This leads to the following reformulation of the concept of differentiability in several variables:

Proposition 16.1. Suppose that $w = f(x, y, z)$ is differentiable at (x_0, y_0, z_0). Then

$$dw = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \ . \tag{16.5}$$

That is, if the function $w = f(x, y, z)$ can be well-approximated near the base point by a linear function, then that linear function must be given by (16.5). There are a variety of ways to use this formula, which we now illustrate.

Example 16.6. Let

$$z = f(x, y) = x^2 - xy + y^3 \tag{16.6}$$

Find the equation of the tangent plane to the graph at the point (2,-1).

At $(x_0, y_0) = (2, -1)$, we have $z_0 = f(x_0, y_0) = 6$. We calculate

$$\frac{\partial f}{\partial x} = 2x - y \ , \quad \frac{\partial f}{\partial y} = -x + 3y^2 \ ,$$

so the differential of z is

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = (2x - y)dx + (-x + 3y^2)dy \ . \tag{16.7}$$

Now we turn to the point (2,-1), evaluating the partial derivatives: $\partial f/\partial x = 5$, $\partial f/\partial y = 1$. Substituting these values in (16.7), and replacing dz by $z - 6$, dx by $x - 2$ and dy by $y + 1$, we obtain

$$z - 6 = 5(x - 2) + (y + 1) \quad \text{or} \quad z = 5x + y - 3 \ , \tag{16.8}$$

as the equation of the tangent plane at the point (2,-1).

An alternative approach is to differentiate equation (16.6) implicitly:

$$dz = 2xdx - xdy - ydx + 3y^2dy \ .$$

Evaluating at (2,-1), we have $z_0 = 6$, and

$$dz = 4dx - 2dy + dx + 3dy = 5dx + dy \ .$$

This is the equation of the tangent plane, Now we can replace the differentials dx, dy, dz replaced by the increments $x - 2, y + 1, z - 5$ to get the equation of the tangent plane (16.8).

Example 16.7. Find the equation of the tangent plane to the graph of the function $z = x^2 + xy - y$ at (2,-1, 1).

First, we calculate the differential

$$dz = 2xdx + xdy + ydx - dy$$

and then evaluate it at the point:

$$dz = 4dx + 2dy - dx - dy = 3dx + dy \ .$$

We now get the equation of the tangent plane by replacing the differentials by the increments:

$$z - 1 = 3(x-2) + (y+1) \quad \text{or} \quad z = 3x + y - 4 \ .$$

Example 16.8. Find the points at which the graph of $z = f(x,y) = x^2 - 2xy + y$ has a horizontal tangent plane.

The horizontal plane through the point (x_0, y_0, z_0) has the equation $z - z_0 = 0$. Thus our points are those where $df = 0$; i.e., solutions of the pair of equations

$$\frac{\partial f}{\partial x} = 0 \qquad \frac{\partial f}{\partial y} = 0 \ .$$

Calculating, we get $2x - 2y = 0, \ -2x + 1 = 0$. Now, solve this pair of simultaneous equations to get $x = 1/2, \ y = 1/2$ and our point is $(1/2, 1/2)$.

Example 16.9. Given the function $z = x^2 - xy + y^3$, in what direction, at the point (1,1,1) is the rate of change of z equal to zero?

The differential of z is $dz = (2x - y)dx + (-x + 3y^2)dy$, so at (1,1,1), we have $dz = dx + 2dy$. This is zero for the direction in which $dx = -2dy$; that is along the line of slope -1/2. Thus the answer is given by a vector in that direction, for example: $-2\mathbf{I} + \mathbf{J}$.

Example 16.10. Suppose that we have designed a cylindrical silo of base radius 6 meters and height 10 meters, and we are asked to increase the radius by .25 m and the height by .2 m. By (approximately) how much do we increase the volume?

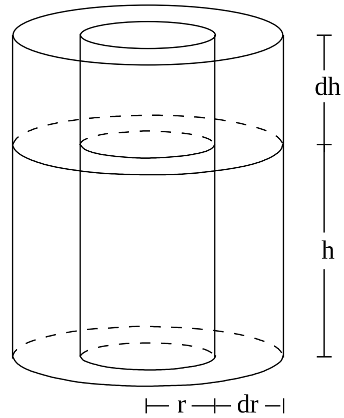
The volume of a cylinder of radius r and height h is $V = \pi r^2 h$. To answer this question, we consider the linear approximation of volume, so we take the differential of V:

$$dV = 2\pi r h dr + \pi r^2 dh \ .$$

Now, in our case $r = 5, \ h = 10, \ dr = .25, \ dh = .2$, so we calculate

$$dV = 2\pi(5)(10)(.25) + \pi(5)^2(.2) = \pi(25 + 5) = 30\pi \quad \text{cubic meters} \ .$$

Figure 16.1



By looking at figure 16.1, we can identify the two terms in the increment of volume: the first is the volume of the shell of width dr around the cylinder, and the second is the volume of the cap of height dh. The negligible part is the volume $2\pi dr dh$ of the washer at the top of width dr and height dh.

Proposition 16.2 (The Chain Rule). Let $w = f(x, y, z)$ be a function defined in a region R in space. Suppose that γ is a curve in R given parametrically by $x = x(t),\ y = y(t).\ z = z(t)$, with $t = 0$ corresponding to (x_0, y_0, z_0). Then, considering $w = f(x(t), y(t), z(t))$ as a function of t along γ, we have

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \ . \tag{16.9}$$

That is, the rate of change of w with respect to t along γ is given by (16.9). We shall give an explanation of this formula in section 16.3

Example 16.11. Let $w = f(x, y, z) = xy + y^2 z$. Consider the curve given parametrically by $x = t,\ y = t^2,\ z = \ln t$. Find dw/dt at $t = 2$.

Differentiating,

$$\frac{\partial w}{\partial x} = y \ , \quad \frac{\partial w}{\partial y} = x + 2yz \ , \quad \frac{\partial w}{\partial z} = y^2 \ ,$$

$$\frac{dx}{dt} = 1 \ , \quad \frac{dy}{dt} = 2t \ , \quad \frac{dz}{dt} = \frac{1}{t} \ ,$$

so

$$\frac{dw}{dt} = y(1) + (x + 2yz)(2t) + y^2(\frac{1}{t}) \ .$$

At $t = 2$ we calculate $x = 2,\ y = 4$ and $z = \ln 2$, giving

$$\frac{dw}{dt} = 4 + (2 + 8\ln 2)(4) + 16/2 = 20 + 32\ln 2 \ .$$

Example 16.12. Let $z = x^2+y^2$. Find the maximum value of z on the ellipse given parametrically by $x = \cos t,\ y = 2\sin t$.

We need to find the points at which $dz/dt = 0$. Now

$$\frac{\partial z}{\partial x} = 2x \ , \quad \frac{\partial z}{\partial y} = 2y \ , \quad \frac{dx}{dt} = -\sin t \ , \quad \frac{dy}{dt} = 2\cos t \ ,$$

and thus

$$\frac{dz}{dt} = -2x\sin t + 4y\cos t \ .$$

Since $x = \cos t$ and $y = 2\sin t$, this gives $dz/dt = -2\sin t\cos t + 8\sin t\cos t$. Set this equal to zero to obtain $6\sin t\cos t = 0$, or $3\sin(2t) = 0$, which has the solutions $t = 0,\ \pm\pi/2,\ \pi,$. For these values of t, the corresponding values of x and y are $x = \pm 1,\ y = 0$ or $x = 0,\ y = \pm 2$. These are the candidates for the maximum point; to determine the actual maximum we compare the values of z at these points. These turn out to be 1 and 4, so 1 is the minimum value and 4 is the maximum value of z on the ellipse.

We can think of an equation of the form $f(x,y,z) = 0$ as defining z **implicitly** as a function of x and y, in the sense that we could solve for z, given specific values of x and y. However, just as in one dimension, we need not solve for z to find the partial derivatives. If we take the differential of the defining equation $f(x,y,z) = 0$ we get

$$f_x dx + f_y dy + f_z dz = 0 \quad \text{so that} \quad dz = -\frac{f_x}{f_z}dx - \frac{f_y}{f_z}dy \ .$$

The coefficient of dx is thus $\partial z/\partial x$, and the coefficient of dy is $\partial z/\partial y$. Of course if $f_z = 0$, these are not defined. But if $f_z \neq 0$, then this method works.

Proposition 16.3. Suppose that f is a differentiable function of (x,y,z) near the point (x_0,y_0,z_0), and that $f_z((x_0,y_0,z_0) \neq 0$. Then the equation $f(x,y,z) = 0$ defines z implicitly as a function of x,y and

$$\frac{\partial z}{\partial x} = -\frac{f_x}{f_z} \quad \text{and} \quad \frac{\partial z}{\partial y} = -\frac{f_y}{f_z} \ . \tag{16.10}$$

Example 16.13. Let $f(x,y,z) = z^3 + 3xz^2 + y^2z$. The relation $f(x,y,z) = 5$ defines z implicitly as a function of (x,y). Find expressions for $\partial z/\partial x$ and $\partial z/\partial y$. Evaluate these at the point (1,1,1).

First we calculate the partial derivatives:

$$\frac{\partial f}{\partial x} = 3z^2 \ , \quad \frac{\partial f}{\partial y} = 2yz \ , \quad \frac{\partial f}{\partial z} = 3z^2 + 6zx + y^2,$$

so that

$$\frac{\partial z}{\partial x} = -\frac{3z^2}{3z^2+6zx+y^2} \ , \quad \frac{\partial z}{\partial y} = -\frac{2yz}{3z^2+6zx+y^2} \ .$$

The values at (1,1,1) are $\partial z/\partial x = -3/10,\ \partial z/\partial x = -1/10$.

Problems 16.1

1. A particle moves in space according to the equation

$$\mathbf{X}(t) = t^2\mathbf{I} + (1-t^2)\mathbf{J} + (1-t)\mathbf{K} \ .$$

For $w = xy + yz^2 + xz^2$, find dw/dt along the trajectory. What is dw/dt when $t = 2$?

2. Let $w = xyz$ and let γ be the helix given by

$$\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + t\mathbf{k} \ .$$

Find dw/dt at $t = 2\pi/3$.

3. Find the second derivatives of $f(x, y) = \ln(x^2 + y^2)$. Show that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \quad \text{and} \quad \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \ .$$

4. Let $f(x, y, z) = \ln(x^2 + y^2 + z^2)$. Calculate

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \ .$$

5. Let $f(x, y, z) = (x^2 + y^2 + z^2)^{-1}$. Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 \ .$$

6. Find the minimum of the absolute value of $u(x, y) = 2x^2 + 3y^3$ on the circle $x = \cos t,\ y = \sin t$.

7. Find the point(s) on the curve $(x - y)^2 + 4(x + y)^2 = 1$ which is closest to the origin.

16.2 Gradients and Vector Methods

Let $w = f(x, y, z)$, where f is a differentiable function. To put the formula for the differential,(16.5), in vector form, we introduce the **gradient** of the function f:

(16.11) $$\nabla f = \frac{\partial f}{\partial x}\mathbf{I} + \frac{\partial f}{\partial y}\mathbf{J} + \frac{\partial f}{\partial z}\mathbf{K} \ ,$$

and the vector differential $d\mathbf{X} = dx\mathbf{I} + dy\mathbf{J} + dz\mathbf{K}$. We interpret $d\mathbf{X}$ as a small change in the vector $\mathbf{X}$/ Then (16.5) can be rewritten as

(16.12) $$dw = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = (\nabla f) \cdot d\mathbf{X} \ .$$

This leads to the following vectorial form of the chain rule.

Proposition 16.4 (The Gradient Form of the Chain Rule). Let $w = f(x, y, z)$ be a function defined in a region R in space. Suppose that γ is a curve in R given parametrically by $\mathbf{X} = \mathbf{X}(t)$. Then, considering $w = f(\mathbf{X}(t))$ as a function of t along γ, we have

$$\frac{dw}{dt} = (\nabla f) \cdot \frac{d\mathbf{X}}{dt} \ . \tag{16.13}$$

The partial derivatives tell us the rate of change of the function f in the coordinate directions. Using the gradient, we can calculate the rate of change in any direction.

Definition 16.3. Let $w = f(x, y, z)$ be differentiable in a neighborhood of $\mathbf{X_0}$. For any vector $\mathbf{V}$, let $\mathbf{X}(t) = \mathbf{X_0} + t\mathbf{V}$ parametrize the line through $\mathbf{X_0}$ in the direction $\mathbf{V}$. The **derivative of f along V** is

$$D_{\mathbf{V}} f(\mathbf{X_0}) = \lim_{t \to 0} \frac{f(\mathbf{X_0} + t\mathbf{V}) - f(\mathbf{X_0})}{t} \ . \tag{16.14}$$

Propostion 16.5. Given the differentiable function f and a vector $\mathbf{V}$, we have

$$D_{\mathbf{V}} f(\mathbf{X_0}) = \nabla f \cdot \mathbf{V} \ . \tag{16.15}$$

The right hand side of (16.14) is the derivative of f along the line in the direction of $\mathbf{V}$. That line is parametrized by $\mathbf{X}(t) = \mathbf{X_0} + t\mathbf{V}$, so $d\mathbf{X}/dt = \mathbf{V}$. Now, by the chain rule

$$D_{\mathbf{V}} f(\mathbf{X_0}) = \frac{d}{dt} f(\mathbf{X}(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = \nabla f \cdot \frac{d\mathbf{X}}{dt} = \nabla f \cdot \mathbf{V} \ .$$

If we replace $\mathbf{V}$ by a unit vector $\mathbf{U}$, then the parameter t represents distance along the line, since $|\mathbf{X}(t) - \mathbf{X_0}| = t|\mathbf{U}| = t$. We say that the line is parametrized by arc length, and refer to $D_{\mathbf{U}} f$ as the **directional derivative** of f in the direction $\mathbf{U}$.

Example 16.14. Let $f(x, y) = x^3 - 3x^2 + xy + 7$ and $\mathbf{U} = 0.6\mathbf{I} - 0.8\mathbf{J}$. Find $D_{\mathbf{U}} f(1, -2)$.

We have $f_x = 3x^2 - 6x + y, f_y = x$. Evaluating at (1,-2), we have $\nabla f(1, -2) = -5\mathbf{I} + \mathbf{J}$. Thus

$$D_{\mathbf{U}} f(1, -2) = \nabla f(1, -2) \cdot \mathbf{U} = -3 - 0.8 = -3.8 \ .$$

Example 16.15. For f as above, find the direction $\mathbf{U}$ at (1,-2) in which $D_{\mathbf{U}} f = 0$.

Let $\mathbf{U} = a\mathbf{I} + b\mathbf{J}$. We must solve

$$\nabla f(1, -2) \cdot \mathbf{U} = (-5\mathbf{I} + \mathbf{J} \cdot (a\mathbf{I} + b\mathbf{J}) = -5a + b = 0 \ .$$

This gives $b = 5a$. Since $\mathbf{U}$ is a unit vector, we have $a^2 + b^2 = 26a^2 = 1$, so $a = 1/\sqrt{26},\ b = 5/\sqrt{26}$ will do Thus

$$\mathbf{U} = \frac{\mathbf{I} + 5\mathbf{J}}{\sqrt{26}} \ .$$

We also have the answer $-\mathbf{U}$. Notice that both these vectors are unit vectors in the direction of $\nabla f^{\perp}$.

Example 16.16. Let γ be parametrized by $\mathbf{X}(t) = t^2\mathbf{I} + \ln t\mathbf{J} + t\mathbf{K}$, and let $w = f(x, y) = xyz$. Find dw/dt along γ. What is the rate of change of w with respect to t at the point $t = 2$?

To use (16.15), we calculate

$$\nabla f = yz\mathbf{I} + xz\mathbf{J} + xy\mathbf{K} \ , \qquad \frac{d\mathbf{X}}{dt} = 2t\mathbf{I} + \frac{1}{t}\mathbf{J} + \mathbf{K} \ ,$$

so that

$$\frac{dw}{dt} = (\nabla f) \cdot \frac{d\mathbf{X}}{dt} = 2tyz + \frac{xz}{t} + xy = 3t^2(\ln t + 1) \ ,$$

since $x = t^2$, $y = \ln t$ and $z = t$. At $t = 2$, we get $dz/dt = 12(\ln 2 + 1)$.

Example 16.17. Let $\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J}$ parametrize the unit circle, and let $f(x, y) = x^2 + 2xy$. Find the maximum value of f on the unit circle.

The function $z = f(\mathbf{X}(t))$ has a maximum when $dz/dt = 0$. We calculate:

$$\nabla f = (2x + 2y)\mathbf{I} + 2x\mathbf{J} = 2((\cos t + \sin t)\mathbf{I} + 2\cos t\mathbf{J} \ ,$$

$$\frac{d\mathbf{X}}{dt} = -\sin t\mathbf{I} + \cos t\mathbf{J} \ ,$$

so that

$$\frac{dz}{dt} = (\nabla f) \cdot \frac{d\mathbf{X}}{dt} = 2((\cos t + \sin t)(-\sin t) + 2\cos^2 t \ .$$

To solve $dz/dt = 0$ we use double angle formulas:

$$2((\cos t + \sin t)(-\sin t) + 2\cos^2 t = -2\cos t \sin t + 2(\cos^2 t - \sin^2 t) = -\sin(2t) + 2\cos(2t)$$

which is zero when $\tan(2t) = 2$, or $t = 31.7°, 211.7°$. The corresponding values of $x = \cos t$, $y = \sin t$ are $x = \pm.526$, $y = \pm.851$. Calculating the values of z at these points gives the maximum 1.172.

For a function $w = f(x, y, z)$ of three variables defined near the point $\mathbf{X_0} : \ (x_0, y_0, z_0)$, let $w_0 = f(x_0, y_0, z_0)$. The equation $w = w_0$ is the level surface S of w at (x_0, y_0, z_0). For f differentiable at a point $\mathbf{X_0}$, the fact that f can be approximated by a linear function implies that the surface S looks more and more like a plane, the closer we look. This plane, given by the equation $df(\mathbf{X_0}) = 0$, is the *tangent plane* to S at $\mathbf{X_0}$. We now note that the gradient of f is the normal to this surface, and points in the direction of maximum increase of f.

Propostion 16.6. Let f be a function differentiable in a neighborhood of the point $\mathbf{X_0}$.
a) $\nabla f(\mathbf{X_0})$ points in the direction of maximum increase of the function f at $\mathbf{X_0}$.
b) $\nabla f(\mathbf{X_0})$ is the normal to the tangent plane of the level set of f through $\mathbf{X_0}$.

To show a), start with a unit vector $\mathbf{U}$. From (16.15) we have

$$D_{\mathbf{U}}(\mathbf{X_0}) = \nabla f \cdot \mathbf{U} = |\nabla f| \cos \beta$$

where β is the angle between ∇f and $\mathbf{U}$ (since $|\mathbf{U}| = 1$). This takes its greatest value when $\cos \beta = 0$, that is $\mathbf{U} = \nabla f$.

To show b), let $\mathbf{V}$ be a vector on the tangent plane. By definition, $df(\mathbf{X_0})(\mathbf{V}) = 0$, so

$$\nabla f(\mathbf{X_0}) \cdot \mathbf{V} = df(\mathbf{X_0})(\mathbf{V}) = 0 \ .$$

Thus $\nabla f(\mathbf{X_0})$ is orthogonal to every vector in the tangent plane, so can be taken to be its normal. Now, a point $\mathbf{X}$ lies in the tangent plane if and only if the vector $\mathbf{X} - \mathbf{X_0}$ lies on the tangent plane, or

$$\nabla f(\mathbf{X_0}) \cdot (\mathbf{X} - \mathbf{X_0}) = 0 \ , \tag{16.16}$$

which is thus the equation of the tangent plane.

Example 16.18. Let $f(x,y) = x^3 + 3x^2y^2 + 2y$. Find the equation of the line tangent to the curve $f(x,y) = 9$ at (2,-1).

The above discussion for three dimensions holds just as well in two dimensions. Thus, by proposition 16.6, the normal to the tangent line to the curve is ∇f.We calculate $\nabla f = (3x^2 + 6xy^2)\mathbf{I} + (6x^2y + 2)\mathbf{J}$; which at $x = 2,\ y = -1$ is the vector $24\mathbf{I} - 22\mathbf{J}$. Now, the equation of the tangent line is given by (15), where $\mathbf{X_0} = 2\mathbf{I} - \mathbf{J}$ is the vector to the point (2,-1):

$$(24\mathbf{I} - 22\mathbf{J}) \cdot ((x-2)\mathbf{I} + (y+1)\mathbf{J}) = 0 \quad \text{or} \quad 24(x-2) - 22(y+1) = 0 \ ,$$

which simplifies to $24x - 22y = 70$.

Example 16.19. Let $f(x,y,z) = xyz$. Find the gradient of f. Find the equation of the tangent plane to the level surface $f(x,y,z) = 2$ at the point $\mathbf{X_0} : (1,2,1)$.

We calculate:

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{I} + \frac{\partial f}{\partial y}\mathbf{J} + \frac{\partial f}{\partial z}\mathbf{K} = yz\mathbf{I} + xz\mathbf{J} + xy\mathbf{K} \ . \tag{16}$$

At $\mathbf{X_0}$, $\nabla f = 2\mathbf{I} + \mathbf{J} + 2\mathbf{K}$, so the equation of the tangent plane is $\nabla f \cdot (\mathbf{X} - \mathbf{X_0}) = 0$:

$$2(x-1) + (y-2) + 2(z-1) = 0 \quad \text{or} \quad 2x + y + 2z = 5 \ .$$

Exampe 16.20. Let $w = x + xy - yz^2$. Find the equation of the tangent plane to the surface $w = 2$ at (3,1,2).

We calculate $\nabla w = (1+y)\mathbf{I} + (x - z^2)\mathbf{J} - 2zy\mathbf{K}$. At the given point (3,1,2), $\nabla w = 2\mathbf{I} - \mathbf{J} - 6\mathbf{K}$. This is the normal to the tangent plane, at $\mathbf{X_0} = 3\mathbf{I} + \mathbf{J} + 2\mathbf{K}$, so the equation of that plane is

$$\nabla w \cdot (\mathbf{X} - \mathbf{X_0}) = 2(x-3) - (y-1) - 6(z-2) = 0 \ ,$$

or $2x - y - 6z + 7 = 0$.

Example 16.21. Let S be the sphere $x^2 + y^2 + z^2 = a^2,\ a > 0$. Show that at any point $\mathbf{X}$ on the sphere, the vector $\mathbf{X}$ is orthogonal to the sphere.

Let $w = x^2 + y^2 + z^2$, so that S is the level set $w = a^2$. Then ∇w is normal to S at $\mathbf{X} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$. But

$$\nabla w = 2x\mathbf{I} + 2y\mathbf{J} + 2z\mathbf{K} = 2\mathbf{X} \ .$$

Example 16.22. Let S_1 be the sphere $x^2 + y^2 + z^2 = 4$ and S_2 the cylinder $x^2 + y^2 = 1$. Let $\mathbf{X} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$ be a point on the curve γ of intersection of the surfaces S_1 and S_2. Find a vector tangent to γ at $\mathbf{X}$.

Let $w_1 = x^2 + y^2 + z^2$, $w_2 = x^2 + y^2$, so that γ is the intersection of the level sets $w_1 = 4$ and $w_2 = 1$. Then ∇w_1 and ∇w_2 are both orthogonal to the tangent to γ, so $\nabla w_1 \times \nabla w_2$ points in the direction of the tangent to γ. We calculate:

$$\nabla w_1 \times \nabla w_2 = 2(x\mathbf{I} + y\mathbf{J} + z\mathbf{K}) \times 2(x\mathbf{I} + y\mathbf{J}) = 4(-yz\mathbf{I} + zx\mathbf{J}) \ .$$

In the above, we have considered a surface as a graph or as a level set of a function. Surfaces can also be given *parametrically.* Let u and v be the variables in a region R of the plane, and let $\mathbf{X}(u,v) = x(u,v)\mathbf{I} + y(u,v)\mathbf{J} + z(u,v)\mathbf{K}$ be a vector-valued function on R. Then the set of values of $\mathbf{X}(u,v)$, as (u,v) ranges over R describes a surface in space.

Example 16.23. Consider the function $\mathbf{X}(u,v) = (u-v)\mathbf{I} + (u+v)\mathbf{J} + uv\mathbf{K}$ defined in (u,v) space. In coordinates, this is given by the equations

$$x = u - v \quad y = u + v \quad z = uv \ .$$

We can solve for u and v in terms of x and y;

$$u = \frac{x+y}{2} \quad v = \frac{-x+y}{2} \ ; \tag{16.17}$$

putting these in the formula for z we have

$$z = uv = \frac{x+y}{2}\frac{-x+y}{2} = \frac{-x^2+y^2}{4} \ ,$$

so the surface is the hyperbolic paraboloid $4z = y^2 - x^2$.

Now, in general it may not be so easy (or simple) to realize a parametric surface as a level set; however, we can use the parametric equations to , for example, find the tangent plane to the surface at a point.

Proposition 16.7. Let $\mathbf{X}(u,v) = x(u,v)\mathbf{I}+y(u,v)\mathbf{J}+z(u,v)\mathbf{K}$ be a vector-valued function defined on a region in R-space. Define

$$\mathbf{X_u} = \frac{\partial x}{\partial u}\mathbf{I} + \frac{\partial y}{\partial u}\mathbf{J} + \frac{\partial z}{\partial u}\mathbf{K} \ ,$$

$$\mathbf{X_v} = \frac{\partial x}{\partial v}\mathbf{I} + \frac{\partial y}{\partial v}\mathbf{J} + \frac{\partial z}{\partial v}\mathbf{K} \ .$$

a) The vector $\mathbf{X_u} \times \mathbf{X_v}$ is normal to the surface.

b) If $w = f(x,y,z)$ is a function defined near the surface, we can consider it as a function of u and v by writing $w = f(x(u,v), y(u,v), z(u,v))$. Then

$$\frac{\partial w}{\partial u} = \nabla w \cdot \mathbf{X_u} \quad \frac{\partial w}{\partial v} = \nabla w \cdot \mathbf{X_v} \ .$$

To see this, fix a point (u_0, v_0). If we set $v = v_0$ and let u vary, we get the curve C given paramterically by

$$\mathbf{X}(u) = x(u, v_0)\mathbf{I} + y(u, v_0)\mathbf{J} + z(u, v_0)\mathbf{K} \ .$$

The tangent vector to this curve is $\mathbf{X_u}$, and since the curve lies in the surface, its tangent vector lies in the tangent plane. Similarly, considering the curve $u = u_0$, we see that the vector $\mathbf{X_v}$ also lies in the tangent plane. Thus $\mathbf{X_u} \times \mathbf{X_v}$ is normal to the tangent plane. Part b) follows directly from the chain rule, applied to the curves $u = u_0$ and $v = v_0$.

Example 16.24. Find the equation of the tangent plane to the surface of example 16.23 at the point (-2,4,3).

From (16.17), this point corresponds to the values $u = 1,\ v = 3$. Now, we differentiate the function defining the surface, obtaining

$$\mathbf{X_u} = \mathbf{I} + \mathbf{J} + v\mathbf{K} \ , \quad \mathbf{X_v} = -\mathbf{I} + \mathbf{J} + u\mathbf{K} \ .$$

The values at $u = 1,\ v = 3$ are $\mathbf{X_u} = \mathbf{I} + \mathbf{J} + 3\mathbf{K},\ \mathbf{X_v} = -\mathbf{I} + \mathbf{J} + \mathbf{K}$. Thus, a normal to the tangent plane is $\mathbf{N} = \mathbf{X_u} \times \mathbf{X_v} = -2\mathbf{I} - 4\mathbf{J} + 2\mathbf{K}$, and the equation of the tangent plane is

$$-2(x + 2) - 4(y - 4) + 2(z - 3) = 0 \quad \text{or} \quad z = x + 2y - 3$$

Example 16.25. Consider the surface given parametrically by

$$\mathbf{X}(u, v) = 3\cos u \cos v\mathbf{I} + 4\cos u \sin v\mathbf{J} + 5\sin u\mathbf{K} \ .$$

Find the normal to the tangent plane ant the point corresponding to $u = \pi/3,\ v = \pi/6$.

Differentiate:

$$\mathbf{X_u} = -3\sin u \cos v\mathbf{I} - 4\sin u \sin v\mathbf{J} + 5\cos u\mathbf{K} \ ,$$

$$\mathbf{X_v} = -3\cos u \sin v\mathbf{I} + 4\cos u \cos v\mathbf{J} \ .$$

Evaluating at the given point, we have

$$\mathbf{X_u} = -3\frac{\sqrt{3}}{2}\frac{\sqrt{3}}{2}\mathbf{I} - 4\frac{\sqrt{3}}{2}\frac{1}{2}\mathbf{J} + 5\frac{1}{2}\mathbf{K} = -\frac{9}{4}\mathbf{I} - \sqrt{3}\mathbf{J} + \frac{5}{2}\mathbf{K} \ ,$$

$$\mathbf{X_v} = -3\frac{1}{2}\frac{1}{2}\mathbf{I} + 4\frac{1}{2}\frac{\sqrt{3}}{2}\mathbf{J} = -\frac{3}{4}\mathbf{I} + \sqrt{3}\mathbf{J} \ .$$

A normal to the plane is $\mathbf{N} = \mathbf{X_u} \times \mathbf{X_v} = -(5/2)\sqrt{3}\mathbf{I} + (15/8)\mathbf{J} + 3\sqrt{3}\mathbf{K}$.

Problems 16.2

1. Let $f(x, y, z) = x \ln z + 2yz$. a) What is ∇f?
   b) Show that

$$\frac{\partial^2 f}{\partial z \partial x} = \frac{\partial^2 f}{\partial x \partial z} \ .$$

2. Find the direction $\mathbf{U}$ of maximal change for the function $w = x^3y^2z + xyz^2$ at the point (2,-1,2). What is $D_{\mathbf{U}}w$ at this point?

3. Suppose that the function $w = f(x, y)$ is differentiable at the point P in the plane. Let $\mathbf{V} = \mathbf{I} + 2\mathbf{J}$, $\mathbf{W} = \mathbf{I} - \mathbf{J}$, and that $D_{\mathbf{V}}w = 2$, $D_{\mathbf{W}}w = 3$ at P. What is $\nabla w(P)$?

4. Suppose $z = f(x, y)$ is differentiable at (1,1), and suppose that

$$\frac{d}{dt}f(1+t, 1+t^2)|_{t=0} = 3, \quad \frac{d}{dt}f(1, 1+t)|_{t=0} = 2\ .$$

What is $\nabla f(1, 1)$?

5. Let $f(x, y) = x^2 + 2xy + 2y^2 + 4x - y$. Find the point(s) at which $\nabla f = 0$.

6. Let S be the level surface $f(x, y, z) = 0$ through the point $\mathbf{X_0}$. In proposition 16.6. we saw that the tangent plane to S at $\mathbf{X_0}$ is the set of vectors orthogonal to $\nabla f(\mathbf{X_0})$. Suppose also that S is a graph, given by the equation $z = g(x, y)$. In section 16.1 we claimed that the tangent plane is given by the equation $dz - dg = 0$ evaluated at $\mathbf{X_0}$. Show that these two conditiions define the same plane.

7. Find the equation of the tangent plane to the surface

$$x^{1/2} + y^{1/2} - z^{1/2} = 0$$

at the point (4,9,25).

8. Find the equation of the tangent plane to the surface

$$x \ln z + 2yz = 0$$

at the point $(-e^2, 1, e^2)$.

9. Find the equation of the tangent plane to the surface given parametrically by

$$\mathbf{X}(u, v) = u^3\mathbf{I} + 2uv\mathbf{J} + v^2\mathbf{K}$$

at the point where $u = 1,\ v = 2$.

10. Let S be a surface which goes through the origin, and whose normal is the z-axis. Let Π be a plane containing the z-axis, and γ the curve of intersection of the surface S and the plane Π. Show that the principal normal to γ is $\pm\mathbf{K}$.

16.3 Theoretical considerations

In order to make the intuitive concept of linear approximation, as used above, more precise we start with the idea of closeness in the space itself. We measure the "nearness" of two points by the length of the line segment joining the points. Thus, in vectorial terms, the **distance** between $\mathbf{X}$ and $\mathbf{X_0}$ is $|\mathbf{X} - \mathbf{X_0}|$, that is, the square root of the sum of the squares of the components. We define limits in terms of this distance.

Definition 16.3. The **ball** of radius c centered at $\mathbf{X_0}$ (denoted $B(\mathbf{X_0}, c)$) is the set of all points of distance less than c from $\mathbf{X_0}$. A **neighborhood** of $\mathbf{X_0}$ is any set which contains some ball centered at $\mathbf{X_0}$.

Definition 16.4. Suppose that f is a function defined in a neighborhood of $\mathbf{X_0}$. We say that

$$\lim_{\mathbf{X}\to\mathbf{X_0}} f(\mathbf{X}) = L$$

if we can insure that $|f(\mathbf{X}) - L|$ can be made as small as we please by taking $\mathbf{X}$ close enough to $\mathbf{X_0}$. We say that f is **continuous** at $\mathbf{X_0}$ if

$$\lim_{\mathbf{X}\to\mathbf{X_0}} f(\mathbf{X}) = f(\mathbf{X_0}) \ .$$

Just as in one variable, we are assured that all functions which can be expressed by polynomials in the coordinates are continuous.

Definition 16.5. A **linear function** is a function of the form $L(\mathbf{X}) = \mathbf{C} \cdot \mathbf{X}$, for some vector $\mathbf{C}$. In coordinates we have $L(x, y, z) = ax + by + cz$, where we have written $\mathbf{C} = a\mathbf{I} + b\mathbf{J} + c\mathbf{K}$ and $\mathbf{X} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$. Its level surface through $\mathbf{X_0}$ is the plane $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$, or $\mathbf{C} \cdot (\mathbf{X} - \mathbf{X_0}) = 0$.

Now we define differentiability at $\mathbf{X_0}$ of a function f: that it can be well- approximated by a linear function. This is the direct generalization of the definition of the derivative in one dimension.

Definition 16.6. Suppose that f is a function defined in a neighborhood of $\mathbf{X_0}$. We say that f is **differentiable** at $\mathbf{X_0}$ if there is a linear function L such that

$$\lim_{\mathbf{X}\to\mathbf{X_0}} \frac{|f(\mathbf{X}) - f(\mathbf{X_0}) - L(\mathbf{X} - \mathbf{X_0})|}{|\mathbf{X} - \mathbf{X_0}|} = 0 \ . \tag{16.18}$$

In this case, we call L the **differential** of f at $\mathbf{X_0}$, denoted $df(\mathbf{X_0})$.

We must verify that this definition of the differential leads to the formula given in (16.5). It will suffice to look at the situation in two variables. Suppose that f is differentiable at $\mathbf{X_0}$, and its differential there is $L(x - x_0, y - y_0) = a(x - x_0) + b(y - y_0)$.

First we see what happens to equation (16.18) along the line $y = y_0$. Then $\mathbf{X} - \mathbf{X_0} = (x - x_0)\mathbf{I}$, and we get

$$\lim_{x\to x_0} \left|\frac{f(x, y_0) - f(x_0, y_0) - a(x - x_0)}{x - x_0}\right| = \lim_{x\to x_0} \left|\frac{f(x, y_0) - f(x_0, y_0}{x - x_0}) - a\right| = 0 \ ,$$

or

$$a = \lim_{x\to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = \frac{\partial f}{\partial x}(x_0, y_0) \ .$$

In the same way, we see that $b = \partial f/\partial y(x_0, y_0)$.

Now we turn to an argument for the chain rule in two dimensions.

The Chain Rule (Propositions 16.2 and 16.4). Let $w = f(x, y)$ be a differentiable function defined in a region R. Suppose that γ is a differentiable curve in R given parametrically by $\mathbf{X} = \mathbf{X}(t)$, with $t = 0$ corresponding to $\mathbf{X_0}$. Then, considering $w = f(\mathbf{X}(t))$ as a function of t along γ, we have

$$\frac{dw}{dt} = (\nabla f) \cdot \frac{d\mathbf{X}}{dt} \ .$$

evaluated at $\mathbf{X_0}$.

We start with the definition of differentiability. Let

$$\eta(t) = f(\mathbf{X}(t)) - f(\mathbf{X_0}) - \mathbf{L} \cdot (\mathbf{X}(\mathbf{t}) - \mathbf{X_0}) \ .$$

where we have written $\mathbf{L}$ for the gradient of w evaluated at $\mathbf{X_0}$. By (16.18),

$$(16.19) \qquad \lim_{\mathbf{X}(\mathbf{t}) \to \mathbf{X_0}} \frac{|\eta(t)|}{|\mathbf{X}(\mathbf{t}) - \mathbf{X_0}|} = 0 \ .$$

Now, by continuity $\mathbf{X}(t) \to \mathbf{X_0}$ as $t \to 0$, and thus

$$(16.20) \qquad \lim_{t \to 0} \left| \frac{\eta(t)}{t} \right| = \lim_{\mathbf{X}(t) \to \mathbf{X_0}} \frac{|\eta(t)|}{|\mathbf{X}(t) - \mathbf{X_0}|} \lim_{t \to 0} \frac{|\mathbf{X}(t) - \mathbf{X_0}|}{|t|} = 0 \ ,$$

by (16.19), and the assumption of differentiability of $\mathbf{X}(t)$, which assure that the second limit on the right exists. Now, by the definition of η:

$$f(\mathbf{X}(t)) - f(\mathbf{X_0}) = \mathbf{L} \cdot (\mathbf{X}(\mathbf{t}) - \mathbf{X_0}) + \eta(t) \ .$$

Divide by t, and take the limit as $t \to 0$:

$$\frac{dw}{dt} = \lim_{t \to 0} \frac{f(\mathbf{X}(t)) - f(\mathbf{X_0})}{t} = \lim_{t \to 0} \mathbf{L} \cdot \frac{\mathbf{X}(\mathbf{t}) - \mathbf{X_0}}{t} + \lim_{t \to 0} \frac{\eta(t)}{t} = \mathbf{L} \cdot \frac{d\mathbf{X}}{dt} \ .$$

16.4 Optimization

In this section we will develop techniques for finding maxima and minima of a function of several variables. First, we consider a function $z = f(x, y)$ of two variables.

Definition 16.7. If $f(x_0, y_0) \geq f(x, y)$ for all $(x, y$ in a neighborhood of (x_0, y_0), we say that f has a **local maximum** at (x_0, y_0). More precisely, if, for some $a > 0$ we have

$$f(x_0, y_0) \geq f(x, y)$$

for all (x, y) within a distance a of (x_0, y_0), then (x_0, y_0) is a local maximum point for f. Similarly, if instead we have

$$f(x_0, y_0) \leq f(x, y)$$

for all (x, y) sufficiently close to (x_0, y_0), then (x_0, y_0) is a **local minimum** point for f.

The first derivative test for functions of one variable gives us the following criterion:

Proposition 16.8. Suppose that $\mathbf{X_0}$ is a local maximum (or minimum) for f. Then $\nabla f = 0$.

To see this, pick a vector $\mathbf{V}$ and consider the line given by the equation $\mathbf{X}(t) = \mathbf{X_0} + t\mathbf{V}$. Then $f(\mathbf{X}(t))$ has a maximum at $t = 0$, so

$$\nabla f \cdot \mathbf{V} = \frac{d}{dt} f(\mathbf{X}(t))\big|_0 = 0 \ .$$

This can only be true for all $\mathbf{V}$ if $\nabla f = 0$.

Definition 16.8. If $\nabla f(x_0, y_0) = 0$ we say that (x_0, y_0) is a **critical point**.

Thus, to find the local maxima or minima of a function in a given region, one must look among the critical points.

Example 16.26. Find the critical points of the function $f(x, y) = x^3 + xy + y^2 - x$.

We calculate the components of the gradient:

$$\frac{\partial f}{\partial x} = 3x^2 + y - 1 \ , \quad \frac{\partial f}{\partial y} = x + 2y \ . \tag{16.21}$$

Now, we set these equal to zero and solve. The second equation gives $x = -2y$; substituting that in the first gives $12y^2 + y - 1 = 0$, which has the roots

$$y = \frac{-1 \pm \sqrt{1 + 48}}{24} \ , \quad \text{or} \quad y = \frac{1}{4}, \ -\frac{1}{3} \ .$$

Thus, the critical points are $(-1/2, 1/4)$, $(2/3, -1/3)$.

But now, how can we tell whether or not we have a local maximum or a local minimum at either of these points? In fact, we may have neither; there is a third possibility: that along certain lines through the critical point, the value is a local maximum, and along other lines, the value is a local minimum. Such a point is a *saddle point*.

Example 16.27. Let $z = x^2 - y^2$. Then the origin is a critical point for z. Since $z = x^2$ along the line $y = 0$, z has a minimum at the origin on this line, but on the line $x = 0$, we have $z = -y^2$ which has a maximum at the origin along this line.

We distinguish local maxima, local minima and saddle points among these critical points by using the second derivative test in one variable. In order to make clear what the criterion is, we first consider the case of a quadratic function.

Example 16.28. Let z be a quadratic function of the variables u, v: $z = au^2 + 2buv + cv^2$. The origin is a critical point. By completing the square we can discover what kind of critical point it is:

$$z = a(u^2 + 2\frac{b}{a}uv + \frac{b^2}{a^2}v^2) + (c - \frac{b^2}{a})v^2 = a(u + \frac{b}{a}v)^2 + \frac{ac - b^2}{a}v^2 \ . \tag{16.22}$$

Thus if both terms have positive coefficients, the origin is a minimum; if both terms have negative coefficients, the origin is a maximum, and if the signs differ, the origin is a saddle point. We call the expression $D = ac - b^2$ the **discriminant** of the quadratic function defining z. Notice that if $D > 0$, that the coefficients of (16.22) have the same sign, and if also $a > 0$, we have a minimum. and if $a < 0$, a maximum. If $D < 0$, the coefficients have different signs and we have a saddle point.

This example leads us directly to the general criterion by applying the second derivative test along each line through the critical point. For the function $z = f(x, y)$, let $f_{xx},\ f_{x,y},\ f_{yy}$ represent the second partial derivatives of f.

Proposition 16.9. Suppose that $\nabla f(x_0, y_0) = 0$, that is (x_0, y_0) is a critical point. Then (evaluating at (x_0, y_0)):

If $D = f_{xx}f_{yy} - (f_{xy})^2 < 0$ at a point (x_0, y_0), the f has a saddle point there.

If $D = f_{xx}f_{yy} - (f_{xy})^2 > 0$ and $f_{xx} > 0$, at a point (x_0, y_0), then f has a local minimum there.

If $D = f_{xx}f_{yy} - (f_{xy})^2 > 0$ and $f_{xx} < 0$, at a point (x_0, y_0), then f has a local maximum there.

If $D = 0$, we can conclude nothing. We note that when $D > 0$ the second derivative along all lines has the same sign, so we could check whether f_{yy} is greater or less than 0 instead, if that were easier.

To see this, choose a vector $\mathbf{V} = u\mathbf{I} + v\mathbf{J}$ and consider the function $f_{\mathbf{V}}(t) = f(\mathbf{X_0} + t\mathbf{V}) = f(x_0 + tu, y_0 + tv)$. Differentiating we find, by the chain rule,

$$\frac{d}{dt}f_{\mathbf{V}} = uf_x + vf_y \ , \quad \frac{d^2}{dt^2}f_{\mathbf{V}} = \frac{d}{dt}(uf_x + vf_y) = u\frac{df_x}{dt} + v\frac{df_y}{dt} \ .$$

We compute this by applying the chain rule to the functions $f_x,\ f_y$:

$$\frac{d^2}{dt^2}f_{\mathbf{V}} = u(uf_{xx} + vf_{yx}) + v(uf_{xy} + vf_{yy}) = u^2 f_{xx} + 2uvf_{xy} + v^2 f_{yy} \ . \tag{16.23}$$

By the second derivative test of one variable, if this is positive, then the function $f_{\mathbf{V}}$ has a minimum; that is the function f has a minimum along the line in the direction of $\mathbf{V}$. If this holds for all directions $\mathbf{V}$; that is, for all values of $u,\ v$, then f has a local minimum at (x_0, y_0). But, referring back to example 16.28, this is true if $D > 0,\ f_{xx} > 0$. Similarly, if $D < 0,\ f_{xx} < 0$, then f has a maximum along all lines through (x_0, y_0), so f has a local maximum there. However, if D and f_{xx} have different signs, then f has a local maximum in some directions, and a local minimum in others, so we have a saddle point.

Example 16.29. We continue with example 16.26. We found critical points at $P(-1/2, 1/4)$ and $Q(2/3, -1/3)$. Differentiating the first partials (see (16.21)), we get

$$f_{xx} = 6x \ , \quad f_{xy} = 1 \ , \quad f_{yy} = 2 \ .$$

Thus

$$\text{at } P : \quad D = 6(-\frac{1}{2})(2) - 1^2 = -7 \ , \quad \text{and at } Q : \quad D = 6(\frac{2}{3})(2) - 1^2 = 7 \ ,$$

so P is a saddle point, and since $f_{xx} = 4 > 0$, Q is a local minimum.

Example 16.30. Let $f(x,y) = x^2 + 2y^4 + xy + 4x + 2y$. Find the local maxima and minima of z. Does f have a global maximum or minimum?

First we find the critical points:

$$f_x = 2x + y + 4\ , \quad f_y = 8y^3 + x + 2\ .$$

To find the points where both are zero, we obtain $x = -8y^3 - 2$ from the second equation. Putting this in the first,we get

$$2(-8y^3 - 2) + y + 4 = 0\ , \quad \text{or} \quad -16y^3 + y = 0\ .$$

This has the solutions $y = 0, \pm\frac{1}{4}$, so the critical points are $P(-2, 0),\ Q(-\frac{17}{8}, \frac{1}{4})$ and $R(-\frac{15}{8}, \frac{1}{4})$. We now calculate the second derivatives:

$$f_{xx} = 2\ , \quad f_{xy} = 1\ , \quad f_{yy} = 24y^2\ .$$

Then $D = 48y^2 - 1$, which is positive at all of these points. Since f_{xx} is everywhere positive, these are all local minima. To determine the global minimum, we evaluate: $f(P) = -4,\ f(Q) = -4.0078,\ f(R) = -4.0078$. Thus the global minimum is -4.0078, attained at both Q and R. Everywhere else the function has a direction in which it is increasing, so it has no global maximum.

Notice, in these problems we have to solve several equations simultaneously, and usually they are not linear. There are no universsal algorithms for solving such systems of equations, and we have to follow our intuition. Usually the technique of substitution works (although in the above problem, with other constants the cubic equation in (16.24) would be much more difficult). So, in general the procedure to follow is to look at the given equations to see if, in one of the equations one of the variables can be easily written in terms of the other. If so, substitute that expression in the other equation.

The Method of Lagrange Multipliers.

Let C be a curve in the plane, not going through the origin. Let's find the point on C which is closest to the origin. This amounts to finding the minimum value of $f(x,y) = x^2 + y^2$ on the curve C. If C is given parametrically by the equations $x = x(t),\ y = y(t)$, we know what to do: differentiate $f(x(t), y(t))$ and set the derivative equal to zero. But, if the curve is given implicitly by an equation $g(x,y) = c$, we don't want to solve the equation explicitly, and we don't have to. Looking at the condition

$$\frac{d}{dt} f(x(t), y(t)) = 0 \quad \text{as} \quad \nabla f \cdot \frac{d\mathbf{X}}{dt} = 0$$

we see that the requirement is that ∇f is orthogonal to the tangent to the curve at the minimizing point. But ∇g is orthogonal to its level set C everywhere, so at the minimizing point we have that ∇f and ∇g are collinear; that is, they are multiples of each other. Thus, we can solve the problem by fiinding the solution of the system

$$\nabla f = \lambda \nabla g\ , \quad g(x,y) = c\ .$$

This gives three scalar equations in three unknowns, which, in principle, can be solved. Of course the value of λ is not of interest, but is useful as an auxiliary to finding the values of $x,\ y$. Since this method was discovered by Lagrange, we call λ the *Lagrange multiplier.*

Example 16.31. Find the point on the line $3x - 2y = 1$ which is closest to the point (4,7).

Given the constraint $g(x,y) = 3x - 2y = 1$, we want to minimize $f(x,y) = (x-4)^2 + (y-7)^2$. The gradients are

$$\nabla f = 2(x-4)\mathbf{I} + 2(y-7)\mathbf{J} \quad \text{and} \quad \nabla g = 3\mathbf{I} - 2\mathbf{J}\ .$$

These gradients are collinear at the minimizing point, so we have to solve the equations

$$2(x-4) = 3\lambda\ , \quad 2(y-7) = 2\lambda \quad \text{and} \quad 3x - 2y = 1\ .$$

We can eliminate λ from the first two equations:

$$4(x-4) = 6\lambda = 6(y-7) \quad \text{so that} \quad 4x - 6y = -26\ .$$

Now we have simultaneous linear equations in x and y which we can solve, getting the point (16,47/2). We note that the Lagrangian equations just say that the line from this point to (4,7) has to be orthogonal to the given line; something we knew from geometry.

Example 16.32. Find the maximum value of $f(x,y) = xy$ on the ellipse $x^2 + 4y^2 = 1$.

Let $g(x,y) = x^2 + 4y^2$. We calculate the gradients: $\nabla f = y\mathbf{I} + x\mathbf{J}$ and $\nabla g = 2x\mathbf{I} + 8y\mathbf{J}$. At the point on the ellipse at which we have the maximum, we have ∇f orthogonal to the tangent to the ellipse, so is collinear with∇g. Thus we have the equation $\nabla f = \lambda \nabla g$ for some λ. This gives the scalar equations

$$y = 2\lambda x\ , \quad x = 8\lambda y\ , x^2 + 8y^2 = 1\ .$$

We can eliminate λ by dividing the first equation by the second:

$$\frac{y}{x} = \frac{2\lambda x}{8\lambda y} = \frac{x}{4y} \quad \text{giving} \quad x^2 = 4y^2\ .$$

Substituting that in the last equation gives $4y^2 + 4y^2 = 1$, so that $y = \pm 1/(2\sqrt{2})$. Then

$$x^2 = 4y^2 = \frac{4}{8} \quad \text{so that} \quad x = \pm\frac{1}{\sqrt{2}}\ .$$

The possible values of $f(x,y) = xy$ at these points are $\pm 1/4$, so the maximum value of f is 1/4, and its minimum is $-1/4$.

The Lagrange multiplier λ serves the purpose of finding a relation between x and y which is a consequence of the optimization. The value of λ is not important, and if it can be eliminated from the conditions, do so first.However, in some cases it may make the problem easier to first determine λ.

To summarize: given the problem: minimize (or maximize) a function $f(x,y)$ subject to a constraint $g(x,y) = c$. We observe that the chain rule tells us that, at the optimizing point, ∇f is orthogonal to the tangent to the level set of g. But so is ∇g, so we must have $\nabla f = \lambda \nabla g$ for some λ. Solve this equation in conjunction with $g(x,y)=c$ to find the point. This method (of Lagrange multipliers) works in three dimensions as well.

Proposition 16.10. Suppose that $w=f(x,y,z)$ is a differentiable function, and we wish to find its maxima and minima subject to a constraint $g(x,y,z)=c$. At an optimizing point P there is a λ such that

$$\nabla f(P)=\lambda\nabla g(P)\ ,\quad g(x,y,z)=c\ .$$

These equations give a system of four equations in four unknowns which, in typical circumstances, has only a finite number of solutions. The maximum (minimum) of the function must occur at one of these points.

To see why this is true, we follow the two dimensional argument. Let S be the level surface $g(x,y,z)=c$. Let C be a curve through P lying in the surface S. Then f is optimized along C, so that the derivative of f along the curve is zero at P. But this just says that $\nabla f(P)$ is orthogonal to the tangent to the curve. Since every vector in the tangent plane to S is the tangent vector to such a curve, $\nabla f(P)$ is orthogonal to the tangent plane to S. But so is $\nabla g(P)$, so $\nabla f(P)$ and $\nabla g(P)$ must be colinear.

Example 16.33. Find the point on the plane $2x+3y+z=1$ closest to the point $(1,-1,0)$.

Here the constraint is $g(x,y,z)=2x+3y+z=1$ and the function to be minimized is $f(x,y,z)=(x-1)^2+(y+1)2+z^2$. Taking the gradients and introducing the Lagrange multipier, we are led to the equations

$$2(x-1)=2\lambda\ ,\quad 2(y+1)=3\lambda\ , 2z=\lambda\ ,\quad 2x+3y+z=1\ .$$

We use the first three equations to express the variables in terms of λ, and then use the last to solve for λ:

$$x=\lambda+1\ ,\quad y=\frac{3\lambda-2}{2}\ ,\quad z=\frac{\lambda}{2}\ , \tag{16.24}$$

so that

$$2(\lambda+1)+3\frac{3\lambda-2}{2}+\frac{\lambda}{2}=1\ .$$

This gives $\lambda=1/7$. Substituting into equations 16.24, we find the desired point to be (1/7, -11/14, 1/14).

Example 16.34. Farmer Brown wishes to enclose a rectangular coop of 1000 square feet. He will build three sides of brick, costing \$25 per linear foot, and the fourth of chain link fence, at \$ 15 per linear foot. What should the dimensions be to minimize the cost?

Let x and y be the dimensions of the coop, where x represents the sides, both of which are to be of brick. The constraint is $g(x,y)=xy=1000$, and the cost function is $C=25(2x+y)+15y$. We have $\nabla C=50\mathbf{I}+40\mathbf{J}$, and $\nabla g=y\mathbf{I}+x\mathbf{J}$. The equations to solve are:

$$50=\lambda y\ ,\quad 40=\lambda x\ , xy=1000\ ,$$

so

$$1000=xy=(\frac{50}{\lambda})(\frac{40}{\lambda})\ ,$$

or $\lambda^2 = (50)(40)/1000 = 2$, giving $\lambda = \sqrt{2}$. Then $x = 40/\sqrt{2} = 20\sqrt{2},\ y = 50/\lambda = 25\sqrt{2}$.

Many problems involve finding the maximum or minimum of a function of many variables subject to many constraints The technique of Lagrange multipliers works in this general context, but - of course - is much more difficult to employ. To give a sense of the general procedure, we state the proposition in the case of a function of three variables with two constraints.

Proposition 16.11. To find the extreme values of a function $f(x, y, z)$ subject to two constraints (say along a curve), $g(x, y, z) = c,\ h(x, y, z) = d$, we have to solve the five equations in the five unknowns x, y, z, λ, μ:

$$\nabla f(P) = \lambda \nabla g(P) + \mu \nabla h\ , \quad g(x, y, z) = c, \quad h(x, y, z) = d\ .$$

Problems16.4

1. Let $f(x, y, z) = xyz - x^3 + x^2 + yz$. Find the critical points of f.

2. Let

$$f(x, y) = x^3 - 4y^3 + 3x^2y - 18x + 6.$$

Find all critical points and classify as maxima, minima, saddle points.

3. Find all the critical points of the function $f(x, y) = x^3 - y^2 + z^3 + 2x^2 + xy$.

4. Let

$$g(x, y, z) = x^2y^2z.$$

Find the point on the surface $g(x, y, z) = 1$ which is closest to the origin.

5. (This is a Calculus I problem. You are to do it using the methods of Lagrange multipliers). John and Mary work part-time at the Widget factory, and are willing to work as much as 40 hours a week. John gets paid \$27/hour, and Mary gets \$45/hour. If John works x hours and Mary works y hours, they produce $3xy + (1/2)y^2$ widgets. The company has allocated \$1600/week for compensation to John and Mary (together). How many hours should they each work in order to produce the maximum number of widgets?

6. The material for the bottom of a box costs three times as much per square foot as the material for the sides and top. We wish to know the greatest volume such a box can have if the total maount of money available for material is \$12, and the material for the bottom costs \$0.60 per square foot. Find the system of equations which must be solved to get the answer.

7. Let $f(x, y) = x^2 + 2y^2 + 2x$. Find the minimum and maximum of f on the ball $x^2 + y^2 \leq 16$.

8. Do problem 7 of section 16.1 using Lagrange multipliers: Minimize $x^2 + y^2$ on the curve $(x - y)^2 + 4(x + y)^2 = 1$.

9. Maximize the area of a rectangle with sides parallel to the coordinate axis, that lies above the x-axis and below the curve $x^2 + 4y^2 = 16$.

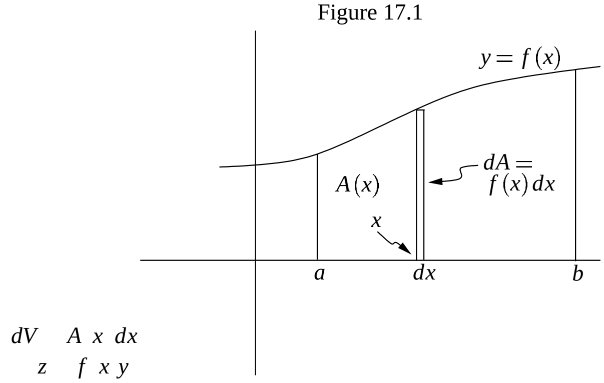
10. Maximize the volume of a rectangular parallelipiped with sides parallel to the coordinate planes, and lying inside the ellipsoid $x^2 + 2y^2 + z^2 = 9$.

Chapter 17. Multiple Integration

17.1 Integration on Planar Regions

Integration of functions in several variables is done following the ideas of "accumulation" introduced in Chapter 4. There, for example, we calculated the area under a curve $y = f(x)$ as x ranges from $x = a$ to $x = b$ by accumulating the area as we swept the region out along the x-axis from a to b. If we define the function $A(x)$ to be the area swept out up to the value x, then we calculated, using figure 16.1, $dA = f(x)dx$: the increment in area is equal to the increment in x times the height of the rectangle at x.

Figure 17.1



We will show that the same idea, now for calculating volumes, works in two dimensions, leading to what is called an "iterated integral". In section 17.3 we shall give a more formal definition of the double integral, and then see that its computation uses the technique of iteration introduced in this section.

Definition 17.1. Let $f(x, y)$ be a function defined on a region R in the plane.

a) If $f(x, y)$ is positive for all (x, y) in the region R, then the volume of the solid lying over the region R and under the graph $z = f(x, y)$ is the **double integral** of f over R, denoted

$$\int\int_R f(x, y)dA \ . \tag{17.1}$$

b) For a general f, the double integral (17.1) is the signed volume bounded by the graph $z = f(x, y)$ over the region; that is, the volume of the part of the solid below the xy-plane is taken to be negative.

Proposition 17.1 (Iterated Integrals). We can compute $\int\int_R f dA$ on a region R in the following way.

Figure 17.2

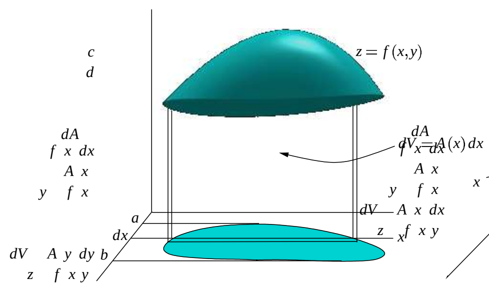
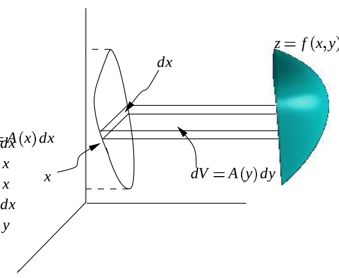


Figure 17.3



a) Suppose R lies between the lines $x = a$ and $x = b$. For each x between a and b, let $A(x)$ be the signed area of the region defined by the graph of $z = f(x, y)$ over R, with x held constant (see figure 17.2). Then

$$\int\int_R f(x,y)dA = \int_a^b A(x)dx = \int_a^b [\int f(x,y)dy]dx \ . \tag{17.2}$$

b) Suppose R lies between the lines $y = c$ and $x = d$. For each y between c and d, let $A(y)$ be the signed area of the region defined by the graph of $z = f(x, y)$ over R, with y held constant (see figure 17.3). Then

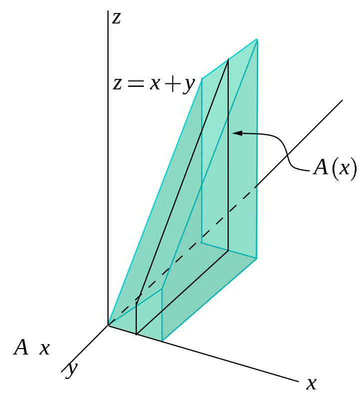
$$\int\int_R f(x,y)dA = \int_c^d A(y)dy = \int_c^d [\int f(x,y)dx]dy \ . \tag{17.3}$$

For (17.2), we sweep out the volume along the x-axis, letting $V(x)$ be the volume accumulated from a to x. Now, an approximation to the increment ΔV in V by moving a small distance Δx is the product of the cross-sectional area $A(x)$ with Δx (see figure 17.2). This leads to the the differential equation $dV = A(x)dx$, which is just the differential form of the first equality of (17.2). But, the area $A(x)$ is just $\int f(x,y)dy$. (17.3) is demonstrated in the same way by sweeping out in the direction of the y-axis.

We left out the limits of integration in the inner integrals of equations (17.2) and (17.3) for simplicity of notation. Determining them could be quite complicated. We start with some simple cases.

Example 17.1. Find the volume of the solid over the rectangle $0 \le x \le 1,\ 0 \le y \le 3$ and bounded by the xy-plane and the plane $z = x + y$ (see figure 17.4).

Figure 17.4



For x between 0 and 1, we calculate the area of the section of the solid by the plane with x fixed;

$$A(x) = \int_0^3 z dy = \int_0^3 (x+y)dy = (xy + \frac{y^2}{2})\Big|_0^3 = 3x + \frac{9}{2} \ .$$

Then the volume is

$$\int_0^1 A(x)dx = \int_0^1 (3x + \frac{9}{2})dx = (\frac{3}{2}x^2 + \frac{9}{2}x)\Big|_0^1 = 6 \ .$$

Example 17.2. Suppose that a house is situated in one corner of a rectangular plot of land 300 feet by 200 feet. The contour of this plot of land is given by the equation $E(x,y) = 10^{-4}(x^2 - xy/2)$ where the house is situated at the origin and the x-axis is the 300 foot length. It is desired to level this property at the level of the house. How much fill has to be removed (or brought in) to accomplish this?

We want to know the difference between the volume of land above house level and that below house level; that is, we want to find the signed volume determined by the graph of $z = E(x,y)$ over the plot of land R. We have

$$\int\int_R E(x,y)dA = \int_0^{300} A(x)dx \ ,$$

where $A(x)$ is the cross-sectional (signed) area of the terrain profile on a section perpendicular to the x axis at a distance x from the house. This is

$$\int_0^{200} E(x,y)dy = 10^{-4}\int_0^{200}(x^2 - \frac{xy}{2})dy = 10^{-4}(x^2y - \frac{xy^2}{4})\Big|_0^{200} = .02x^2 - x \ .$$

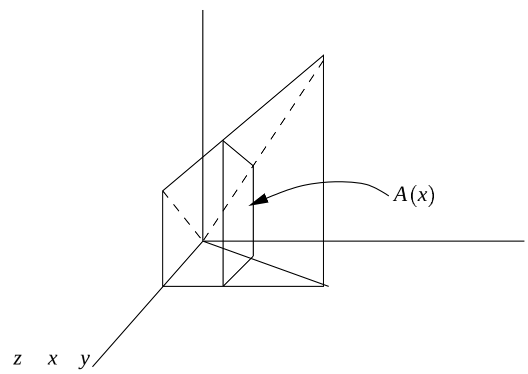
Then

$$\int\int_R E(x,y)dA = \int_0^{300}[\int_0^{200} E(x,y)dy]dx = \int_0^{300}(.02x^2 - x)dx = (.02\frac{x^3}{3} - \frac{x^2}{2})\Big|_0^{300} \ .$$

This evaluates to 13.5×10^4 cubic feet which, because it is positive, have to be removed. Were we to ask for the average elevation of the property above house level, we divide this by the area of the property, getting $13.5 \times 10^4/(6 \times 10^4) = 2.25$ feet.

Example 17.3. Find the volume under the plane $z = x + 2y + 1$ over the triangle bounded by the lines $y = 0,\ x = 1,\ y = 2x$ (see figure 17.5).

Figure 17.5



If we sweep out along the x-axis, we can calculate the volume as $\int_0^1 A(x)dx$, where, for fixed $x,\ A(x)$ is the area under the curve $z = x + 2y + 1$ over the line segment at x in the triangle. This is the line from $y = 0$ to $y = 2x$. Thus

$$A(x) = \int_0^{2x} (x + 2y + 1)dy = (xy + y^2 + y)\Big|_0^{2x} = 2x^2 + 4x^2 + 2x = 6x^2 + 2x \ .$$

Then the volume is

$$Volume \ = \int_0^1 [\int_0^{2x} (x + 2y + 1)dy]dx = \int_0^1 (6x^2 + 2x)dx = (2x^3 + x^2)\Big|_0^1 = 3 \ .$$

For confirmation that we can calculate integrals by iterating in either order, we'll calculate the volume by sweeping out along the y axis first. Now, y ranges from 0 to 2, and for fixed $y,\ x$ ranges from $y/2$ to 1. This computation leads to

$$Volume \ = \int_0^2 [\int_{y/2}^1 (x + 2y + 1)dx]dy \ .$$

The inner integral is

$$\int_{y/2}^1 (x + 2y + 1)dy = (\frac{x^2}{2} + 2xy + x)\Big|_{y/2}^1 = \frac{3}{2} + \frac{3}{2}y - \frac{9}{8}y^2 \ .$$

We then get

$$Volume \ = \int_0^2 (\frac{3}{2} + \frac{3}{2}y - \frac{9}{8}y^2)dy = (\frac{3}{2}y + \frac{3}{4}y^2 - \frac{3}{8}y^3\Big|_0^2 = 3 + \frac{3}{4} - \frac{3}{4} = 3 \ .$$

This last example illustrates two important considerations:

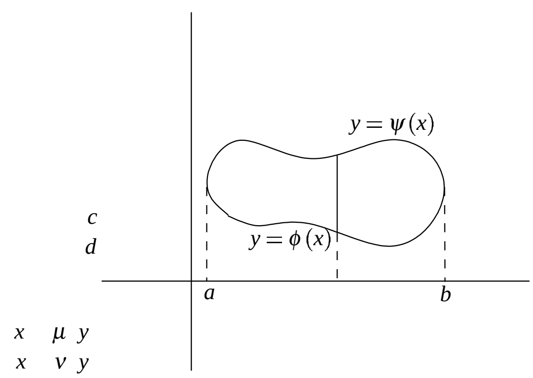
1. We can try to integrate by sweeping out along either coordinate axis, and one order of integration may be simpler than the other.

2. To integrate a function f over a domain, first draw a diagram of the domain to determine the preferred (sometimes, only possible) order of integration. If we try sweeping out along the x-axis, first determine the range in the variable x, and then assure yourself that any line $x =$ constant intersects the region in an interval. If not, sweep in the other direction, so that any line $y =$ constant is an interval. Of course, both attempts may fail; we'll look into this in the next section. If one or the other criterion holds, we say the region is regular.

Definition 17.2. A region in the plane is **regular** if it can be described in either of these two ways:

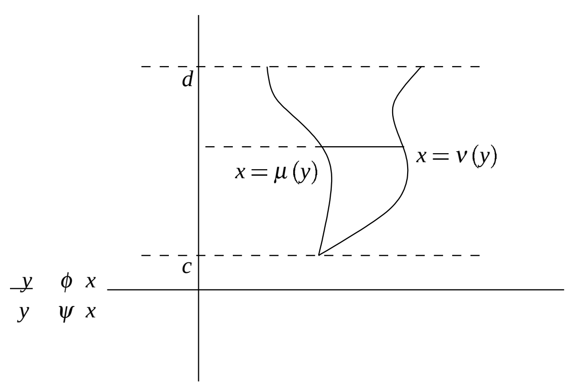
(type 1): as the set of (x, y) where x runs from a to b, and for each such x, y lies between $\phi(x$ and $\psi(x)$ (see figure 17.6);

Figure 17.6: Type 1



(type 2): as the set of (x, y) where y runs from c to d, and for each such y, x lies between $\mu(y)$ and $\nu(y)$ (see figure 17.7).

Figure 17.7: Type 2



Proposition 17.2. Suppose that f is defined over a regular region R. Then we can calculate the double integral $\int\int_R f dA$ as an **iterated** integral:

$$\int\int_R f dA = \int_a^b [\int_{\phi(x)}^{\psi(x)} f(x, y) dy] dx \qquad \text{(type 1)} ,$$

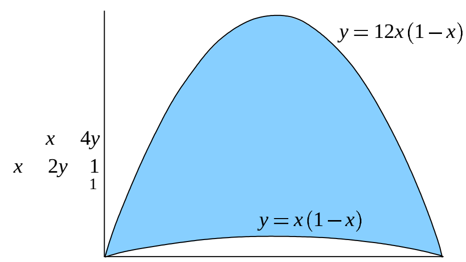
$$\int\int_R f dA = \int_c^d [\int_{\mu(y)}^{\nu(y)} f(x,y)dx]dy \qquad \text{(type 2)} \ .$$

If R is regular of both types 1 and 2, then we can calculate either way, whichever is more convenient.

Example 17.4 . Let R be the region in the first quadrant between the curves $y = 12x(1-x)$ and $y = x(1-x)$. Find $\int\int_R xy dA$.

Draw the figure (see figure 17.8).

Figure 17.8



This is a type 1 region, where for any x, y lies on the interval from $x(1-x)$ to $12x(1-x)$. This interval shrinks to a point when $x = 0$ or 1, so the range of x is $0 \le x \le 1$. Thus

$$\int\int_R xy dA = \int_0^1 [\int_{x(1-x)}^{12x(1-x)} xy dy]dx \ .$$

The inner integral is

$$x\frac{y^2}{2}\Big|_{x(1-x)}^{12x(1-x)} = \frac{x}{2}(144x^2(1-x)^2 - x^2(1-x)^2) = \frac{143}{2}(x^2 - 2x^3 + x^4) \ .$$

Thus

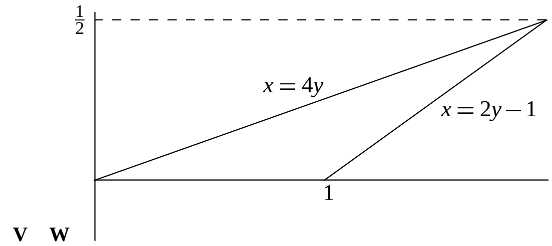
$$\int\int_R xy dA = \frac{143}{2}\int_0^1 (x^2 - 2x^3 + x^4)dx = \frac{143}{2}(\frac{1}{3} - 2\frac{1}{4} + \frac{1}{5}) = \frac{143}{60} \ .$$

Example 17.5 . Let R be the region in the first quadrant between the lines $x - 4y = 0$ and $x - 2y = 1$. Evaluate

$$\int\int_R \frac{x}{1+y^2} dA \ .$$

Draw the figure (see figure 17.9).

Figure 17.9



This is not a type 1 region, but it is type 2. The point of intersection of the lines is the simultaneous solution of the two equations" $x = 2,\ y = 1/2$. Thus the region can be described as $0 \leq y \leq 1/2$, and for any such y, $4y \leq x \leq 1 + 2y$. Thus

$$\int\int_R \frac{x}{1+y^2} dA = \int_0^{\frac{1}{2}} [\int_{4y}^{1+2y} \frac{x}{1+y^2} dx] dy \ .$$

The inner integral is

$$\frac{1}{1+y^2}\frac{x^2}{2}\Big|_{4y}^{1+2y} = \frac{(1+2y)^2 - 16y^2}{2(1+y^2)} = \frac{1+4y-12y^2}{2(1+y^2)} \ .$$

We now want to integrate this from 0 to1/2. To do so, we will need to do the long division:

$$\frac{1+4y-12y^2}{2(1+y^2)} = -6 + 2\frac{y}{1+y^2} + \frac{13}{2(1+y^2)} \ .$$

Finally, the value of the double integral we seek is

$$\int_0^{\frac{1}{2}} \big(-6 + \frac{2y}{1+y^2} + \frac{13}{2(1+y^2)}\big) dy = \big[-6y + \ln(1+y^2) + \frac{13}{2}\arctan y\big]_0^{\frac{1}{2}}$$

which evaluates to $-3 + \ln(1.25) + 6.5\arctan(.5) = .2368$.

Problems 17.1

1. What is the volume of the region under the surface $z = \ln x + y$ and over the rectangle in the xy-plane with vertices (1,0), (1,5), (3,0), (3,5)?

2. What is the volume of the solid bounded by the surfaces $z = x^3$ and $z = x^2 + 2y^2$ lying directly over the rectangle $0 \leq x \leq 1,\ 0 \leq y \leq 3$?

3. What is the volume of the solid bounded above by the surface $z = y^2 - x^2$ lying directly over the triangle $T:\ 0 \leq y \leq 2, -y \leq x \leq y$?

4. Let D be the triangle bounded by the x-axis, the line $x = 1$ and the line $x = y$. Find

$$I = \int\int_D xye^{x^2+y^2} dA \ .$$

5. Find the volume of the region in the first octant under the plane $x + y + z = 1$.

6. What is the volume of the region under the surface $z = e^{x+y}$ and over the triangle in the xy-plane with vertices (0,0), (1,0), (0,2).

7. Let R be the region in the plane bounded by the curves $x = y^2$, $x = 3 - 2y^2$. Calculate

$$I = \int\int_R (y^2 - x)dA \ .$$

8. Find the volume of the region lying over the domain bounded by the y-axis and the curve $y = \sin x$ and the line$y = 1$, and under the surface $z = \sin x + y$.

17.2 Applications

Just as in one dimension (as in Chapter 5), we can apply integration to any calculation that is "accumulative": that is, its value over a whole is the sum of its value of the parts.

Average

Let R be a region in the plane. Its area is $\int\int_R dA$. If f is a function defined on R, its *average value* is

$$f_{ave} = \frac{\int\int_R f dA}{\int\int_R dA} \ .$$

We already calculated an average, in example 2. Here is another example.

Example 17.6. We can model the state of Kansas as the rectangle $K : 0 \le x \le 500,\ 0 \le y \le 300$, where the units are miles. On a cold winter's day the temperature at (x, y) was $T(x, y) = .01x(1 - 10^{-5}xy)$ degrees Celsius. What was the average temperature in the state?

To answer the question we consider $T(x, y)$ as a measure of the amount of "heat" at (x, y). Then the total heat in the state is

$$H = \int\int_K T(x, y)dA = \int_0^{500}\int_0^{300} (10^{-2}x - 10^{-7}x^2y)dydx \ .$$

The inner integral is

$$\int_0^{300} (10^{-2}x - 10^{-7}x^2y)dy = (10^{-2}xy - 10^{-7}x^2\frac{y^2}{2})_0^{300} = 3x - 4.5 \times 10^{-3}x^2 \ .$$

Finally

$$H = \int_0^{500(} 3x - 4.5 \times 10^{-3}x^2)dx = \frac{3}{2}x^2 - 1.5 \times 10^{-3}x^3)_0^{500} = 18.75 \times 10^4 \ .$$

Since the area of Kansas is 15×10^4 square miles, the average is

$$\frac{H}{A} = \frac{18.75 \times 10^4}{15 \times 10^4} = 1.25^\circ\mathrm{C} \ .$$

Mass

A thin plate covering a region R in the plane is called a *lamina*. We suppose the lamina is filled with some inhomogeneous material, whose density over the point (x, y) is $\delta(x, y)$. Then the *total mass* of the plate is

$$Mass = \int\int_R \delta dA \ .$$

Example 17.7. Suppose the rectangle R: $0 \leq x \leq 3,\ 0 \leq y \leq 4$ is filled with an inhomogeneous fluid whose density at the point (x, y) is $\delta(x, y) = xy/6$. Find the total mass.

$$Mass = \int\int_R \frac{xy}{6} dA = \frac{1}{6}\int_0^3 [\int_0^4 xydy]dx \ .$$

The inner integral is

$$\int_0^4 xydy = x\frac{y^2}{2}|_0^4 = 8x \ ,$$

so

$$Mass = \frac{1}{6}\int_0^3 8xdx = \frac{1}{6}(4x^2)_0^3 = 6 \ .$$

Moment

Given a mass m at a point P, its *moment* about an axis L is $d \cdot m$, where d is the distance of P to the axis L. It is Archimedes' principle that moments are additive, so if we have a lamina covering a region R, we can approximate its total moment about an axis L by covering it with a grid of small rectangles, and adding up the moments of each of the rectangles. In the limit this becomes a double integral: the *moment of the lamina R about the axis L* is

$$Mom_L = \int\int_R d_L \delta dA \ ,$$

where δ is the density function and d_L is the signed distance from the axis L. We take the signed distance because, when the axis passes through the region the moments on one side have the opposite effect of the moments on the other. The lamina is *balanced* along the axis L if $Mom_L = 0$. The *center of mass* of R is a point $\bar{P} = (\bar{x}, \bar{y})$ such that R is balanced along every axis through $\bar{P}$. To find this point, we first write down the fact that the lamina is balanced along the lines $x = \bar{x},\ y = \bar{y}$. First, note that the signed distance of x to $\bar{x}$ is just $x - \bar{x}$. Thus

$$Mom_{x=\bar{x}} = \int\int_R (x - \bar{x})\delta dA \quad \text{or} \quad \bar{x}\int\int_R \delta dA = \int\int_R x\delta dA \ ,$$

$$Mom_{y=\bar{y}} = \int\int_R (y - \bar{y})\delta dA \quad \text{or} \quad \bar{y}\int\int_R \delta dA = \int\int_R y\delta dA \ .$$

Thus

$$\bar{x} = \frac{Mom_{x=0}}{Mass}\ , \qquad \bar{y} = \frac{Mom_{y=0}}{Mass}\ .$$

Example 17.8. Find the center of mass of the rectangle in example 17.7.

$$Mom_{x=0} = \int\int_R x\delta dA = \frac{1}{6}\int_0^3 [\int_0^4 x^2 y dy]dx = \frac{1}{6}\int_0^3 8x^2 dx = 12\ .$$

$$Mom_{y=0} = \int\int_R x\delta dA = \frac{1}{6}\int_0^3 [\int_0^4 xy^2 dy]dx = \frac{1}{6}\int_0^3 \frac{64x}{3} dx = 16\ .$$

In example 17.7, we found the mass to be 6, so the center of mass is at (12/6,16/6) = (2,2.67).

The *moment of inertia* of a lamina$R$ about an axis L is $\int\int_R d_L^2 \delta dA$. This is a concept used when considering the rotation of the lamina. In particular, for the corrdinate axes we have

$$I_{x=0} = \int\int_R x^2 \delta dA\ ,$$

$$I_{y=0} = \int\int_R y^2 \delta dA\ .$$

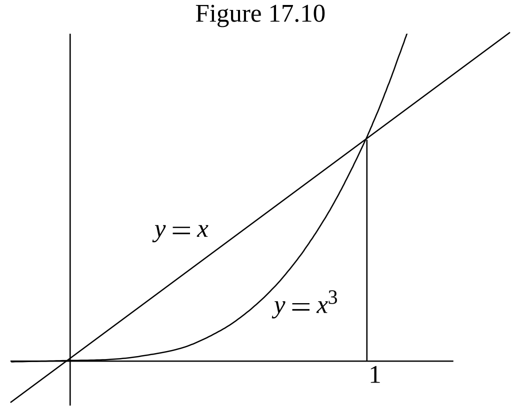
If we are interested in rotation of the region in its plane about the origin, we take as the squared term the distance to the origin (which can also be considered as rotation about the z-axis). Thus the moment of inertia of the region about the origin is

$$I_O = \int\int_R (x^2 + y^2)\delta dA$$

Example 17.9. Consider a lamina over the region R in the first quadrant bounded by the curves $y = x$ and $y = x^3$. The density of the material in the lamina is $\delta(x, y) = y$. Find the center of mass of R and the moment of inertia about the origin.

First, draw a diagram of the region R (see figure 17.10):

Figure 17.10



The figure is the key to performing these integrations:

$$Mass = \int\int_R \delta dA = \int_0^1 [\int_{x^3}^x ydy]dx = \int_0^1 \frac{1}{2}(x^2 - x^6)dx = \frac{1}{2}[\frac{1}{3} - \frac{1}{7}] = \frac{2}{21} \ .$$

$$Mom_{x=0} = \int\int_R x\delta dA = \int_0^1 [\int_{x^3}^x xydy]dx = \int_0^1 \frac{1}{2}(x^3 - x^7)dx = \frac{1}{2}[\frac{1}{4} - \frac{1}{8}] = \frac{1}{16} \ .$$

$$Mom_{y=0} = \int\int_R y\delta dA = \int_0^1 [\int_{x^3}^x y^2dy]dx = \int_0^1 \frac{1}{3}(x^3 - x^9)dx = \frac{1}{3}[\frac{1}{4} - \frac{1}{10}] = \frac{1}{20} \ .$$

Thus the center of mass is $(\overline{x}, \overline{y})$, where

$$\overline{x} = \frac{1/16}{2/21} = \frac{21}{32} \ , \qquad \overline{y} = \frac{1/20}{2/21} = \frac{21}{40} \ .$$

Finally, the moment of inertia about the origin is

$$I_O = \int\int_R (x^2 + y^2)\delta dA = \int_0^1 [\int_{x^3}^x (x^2 + y^2)ydy]dx \ .$$

The inner integral is

$$\int_{x^3}^x (x^2y + y^3)dy = (x^2\frac{y^2}{2} + \frac{y^4}{4})|_{x^3}^x = \frac{3x^4}{4} - \frac{x^8}{2} - \frac{x^{12}}{4} \ ,$$

so, finally

$$I_O = \int_0^1 (\frac{3x^4}{4} - \frac{x^8}{2} - \frac{x^{12}}{4}) = \frac{44}{585} \ .$$

Problems 17.2

1. What is the mass of the lamina bounded by the curves $y = 1 + x$ and $y = 1 - x^3$ and the x-axis, where the density function is $\delta(x, y) = x^2$?

2. A lamina filled with a homogeneous material (the density is identicaly equal to 1) is in the shape of the region R bounded by the curves $y = 1$ and $y = x^2$. What is its center of mass?

3. Find the mass of the solid bounded by the surface $z = \sqrt{x^2 + y}$, the coordinate planes and the planes $x = 1$, $y = 2$, where the density is $\delta(x, y, z) = x$.

4. What is the mass of the lamina bounded by the curves $y = 3x$ and $y = 6x - x^2$ where the density function is $\delta(x, y) = xy$?

5. A lamina filled with a homogeneous material (the density is identicaly equal to 1) is in the shape of the region $0 \leq x \leq \pi, \ 0 \leq y \leq \sin x$. Find its center of mass.

17.3 Theoretical Considerations

Let $f(x, y)$ be a function defined for all (x, y) in a region R in the plane. The integral of f over R is defined, as in one variable, as a limit of approximating sums. Recall that in one variable, the interval of integration was partitioned into small interval, and for each interval we formed the product $f(\bar{x})\Delta x$, where $\bar{x}$ is a point in the interval, and Δx is the length of the interval. The sum $\sum f(\bar{x})\Delta x$ over all intervals is the approximation to the integral for this partition. Now we do the same thing, taking small rectangles instead of intervals.

Select a grid G of horizontal and vertical lines, and let $|G|$ represent the maximal distance between any two successive lines. Given such a grid, let $G(R)$ be the set of rectangles wholly contained inside the region R. Then form the sum

$$\sum f(\bar{x},\bar{y})\Delta A \tag{17.4}$$

over all rectangles in $G(R)$. Here, ΔA is the area of one such rectangle, and $(\bar{x},\bar{y})$ is any point in the rectangle. This sum is called the bf Riemann sum for f on R over the grid G. If the grid is very fine (that is, $|G|$ is very small), this is an approximation to the integral.

Definition 17.3. f is said to be **integrable** over the region R if

$$\lim_{|G|\to 0}\sum f(x,y)\Delta A \tag{17.5}$$

exists. In this case, the limit is the **definite integral** of f over R, denoted

$$\int\int_R f(x,y)dA\ .$$

The limit is taken as the grid becomes exceedingly fine; that is, as the dimensions of all the rectangles goes to zero. It is very hard to find comprehensive general conditions on the function f and the region R so that the intregral $\int\int_R f(x,y)dA$ exists. However, for our present purposes the following fact suffices.

Theorem. Suppose that R is a rectangle ($a \le x \le b,\ c \le y \le d$), and f is bounded and continuous on R except along a finte set of differentiable curves, then f is integrable on R; that is, the limit (16.5) exists.

To evaluate the integral, we reduce the problem to the one variable calculus, by adding the terms in (16.5) first in the vertical columns, and then adding the sums corresponding to the columns. Suppose then, that R is the rectangle $a \le x \le b,\ c \le y \le d$. For each column between two grid lines on the x-axis, form the sum $\sum f(x,y)\Delta y$. Now take these expressions, multiply by the width of the column (Δx) and sum over all columns, obtaining

$$\sum \Big[\sum f(x,y)\Delta y\Big]\Delta x\ .$$

Taking the limits (as $|G|\to 0$, (which is the same as $\Delta x\to 0, \Delta y\to 0$) we obtain

$$\int\int f(x,y)dA = \int_a^b \Big[\int_c^d f(x,y)dy\Big]dx\ ,$$

where the inner integral is taken with x treated as a constant. Of course, we can sum first over the rows, and then add these sums up the y-axis:

$$\int\int f(x,y)dA = \int_c^d \Big[\int_a^b f(x,y)dx\Big]dy\ .$$

Now, what do we do if the region R is more general than a rectangle? There is no general procedure; but for regular domains, we can deduce proposition 17.2 from the above theorem. Suppose, for example that R is a type 1 domain, bounded on the left and right by lines $x = a$, $x = b$, and below by a curve $y = \phi(x)$, and above by a curve $y = \psi(x)$. Then we can enclose the region R in a rectangle R_0 bounded by the lines $x = a,\ x = b,\ y = c,\ y = d$. If we extend f to all of R_0 by defining it to be zero outside R, this extension is continuous on R_0 except along the curve $y = \phi(x),\ y = \psi(x)$, so the theorem applies. But now

$$\int_c^d f(x,y)dy = \int_{\phi(x)}^{\psi(x)} f(x,y)dy$$

since $f(x,y) = 0$ for $y < \phi(x)$ or $y > \psi(x)$. Thus we have

$$\int\int_R f(x,y)dA = \int_a^b \Big[\int_{\phi(x)}^{\psi(x)} f(x,y)dy\Big]dx \ .$$

Similarly, for a type 2 region R, bounded below and above by lines $y = c$, $y = d$, and on the left by a curve $x = \mu(y)$, and on the right by a curve $x = \nu(y)$, then the result is

$$\int\int_R f(x,y)dA = \int_c^d \Big[\int_{\mu(y)}^{\nu(y)} f(x,y)dx\Big]dy \ .$$

These are the **iterated integrals**, and reduce the problem of evaluating double integrals to one-variable methods..

For a general region, we try to break it up into a finite set of nonoverlapping pieces, each of which is either type 1 or type 2. Then we evaluate over each piece by the iterated integral, and add them all together. This uses part (a) of the following proposition, all of which can be easily verified as for integration in one variable.

Proposition 17.3.
If R consists of n nonoverlapping regions $R_1, R_2, \ldots, R_n$, then

b)
$$\int\int_R fdA = \int\int_{R_1} fdA + \cdots + \int\int_{R_n} fdA.$$

b)
$$\int\int_R (f+g)dA = \int\int_R fdA + \int\int_R gdA \ .$$

c) If C is a constant

$$\int\int_R CfdA = C\int\int_R fdA \ .$$

d) If $f \geq g$ then

$$\int\int_R fdA \geq \int\int_R gdA \ .$$

e) The **area** of the region R is

$$A(R) = \int\int_R dA \ .$$

f) If f is non-negative, the **volume** of the region lying over the region R in the xy-plane and under the surface $z = f(x,y)$ is

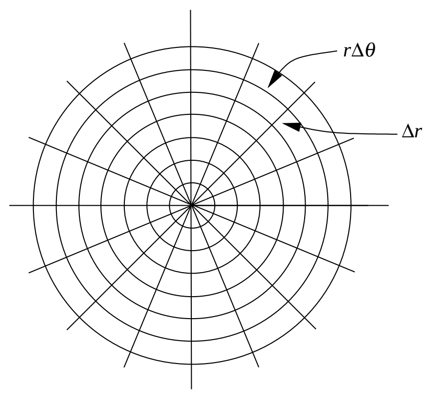
$$\int\int_R fdA \ .$$

17.4 Integration in Other Coordinates

Polar coordinates

Calculations of double integrals are often simplified by turning to appropriate coordinates. If, for example, the problem setup is suggestive of polar coordinates, the change can be made as follows. Cover the region with a grid this time made up of the curves $r =$ constant, $\theta =$ constant (see figure 17.11).

Figure 17.11



Then form the sum

$$\sum f(r,\theta)\Delta A\ ,$$

where now ΔA is the area of one of the figures cut out by this grid, and f is evaluated at a point in the grid. If the grid is very fine, we can take its area to be the product of the lengths of its sides, respectively Δr and $r\Delta\theta$. Using this approximation, in the limit, we get $dA = rdrd\theta$, so that

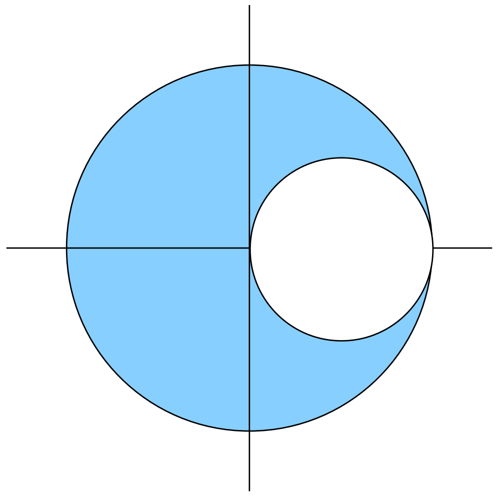
$$\int\int_R f dA = \int\int f(r,\theta) r dr d\theta\ , \qquad (17.6)$$

which we now can evaluate by appropriate iteration. So, if the region R is of the form $\alpha \le \theta \le \beta,\ u(\theta) \le r \le v(\theta)$, we calculate the double integral by the iterated integral:

$$\int\int_R f(r,\theta) dA = \int_\alpha^\beta \Big[\int_{u(\theta)}^{v(\theta)} f(r,\theta) r dr\Big] d\theta\ .$$

Example 17.10. Find $\int\int_R x dA$ where R is the region bounded by the circle of radius 1 centered at the origin and the circle of radius 1/2 centered at the point (1/2,0) (see figure 17.12).

Figure 17.12



If we move to polar coordinates we have $x = r\cos\theta$, the outer boundary of R is $r = 1$, and the inner boundary is given by $r = \cos\theta$. To proceed, we first notice that the integral is twice that of the integral over the part of R in the upper half plane, by symmetry. To reduce the integral to iterated integrals, we have to consider the pieces in the first and second quadrants separately. Denote these by I and II.

$$\int\int_I xdA = \int_0^{\pi/2}\int_{\cos\theta}^1 (r\cos\theta)rdrd\theta = \int_0^{\pi/2}\int_{\cos\theta}^1 r^2 cos\theta dr]d\theta\ .$$

The inner integral is

$$\frac{r^3}{3}cos\theta|_{\cos\theta}^1 = \frac{1}{3}(\cos\theta - \cos^4\theta)\ .$$

Using the double angle formula twice, we find

$$\cos^4\theta = \frac{3}{8} + \frac{\cos(2\theta)}{2} + \frac{\cos(4\theta)}{4}\ .$$

Then

$$\int\int_I xdA = \frac{1}{3}\int_0^{\pi/2}(\cos\theta - \cos^4\theta)d\theta = \frac{1}{3}(1 - \frac{3\pi}{16})\ .$$

Now an easier computation gives

$$\int\int_{II} xdA = \int_{\pi/2}^{\pi}\int_0^1 (r\cos\theta)rdrd\theta = -\frac{1}{3}\ ,$$

so that finally

$$\int\int_R xdA = 2(\frac{1}{3}(1-\frac{3\pi}{16})-\frac{1}{3}) = -\pi/8\ .$$

Example 17.11. Find the center of mass of the region described in example 17.10 (refer to figure 17.12).

That region is bounded on the outside by a circle of radius 1, and on the inside by a circle of radius 1/2, so has area $\pi - \pi/4 = 3\pi/4$. In example 6, we calculated the moment about $x = 0$ to be

$$Mom_{x=0} = \int\int_R xdA = -\frac{\pi}{8}\ ,$$

so the x coordinate of the center of mass is $x = -(\pi/8)/(3\pi/4) = -1/6$. Since the region is symmetric about the x-axis, the center of mass is on that axis, and thus is at $(-1/6, 0)$.

Our next example is a trick calculation, obtained by working backwards from the iterated integral to the double integral.

Example 17.12. $\int_0^\infty e^{-x^2}dx = \sqrt{\pi}/2$.

Observe first that

$$(\int_0^\infty e^{-x^2}dx)^2 = (\int_0^\infty e^{-x^2}dx)(\int_0^\infty e^{-y^2}dy)\ ,$$

just by renaming the variable of integration in the second factor. But now, this last can be viewed as an iterated integral, and then as a double integral:

$$(\int_0^\infty e^{-x^2}dx)(\int_0^\infty e^{-y^2}dy) = \int_0^\infty [\int_0^\infty e^{-(x^2+y^2)}dy]dx = \int\int_Q e^{-(x^2+y^2)}dA\ ,$$

where Q is the first quadrant. Now, this is the integral of e^{-r^2} over Q, which in polar coordinates is given by $0 \le \theta \le \pi,\ 0 \le r \le \infty$. Making the change to polar coordinates, we get :

$$(\int_0^\infty e^{-x^2}dx)^2 = \int_0^{\pi/2}[\int_0^\infty e^{-r^2}rdr]d\theta = \frac{\pi}{2}(\frac{1}{2}) = \frac{\pi}{4}\ .$$

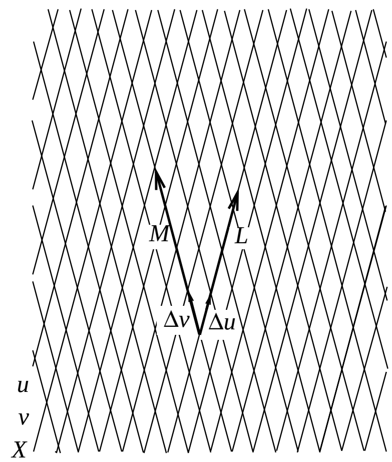
Taking square roots, we get the result.

General Coordinate Changes

Let **L** and **M** be two vectors in the plane which are not parallel, so that the area of the parallelogram they span, $|\det(\mathbf{L},\mathbf{M})|$, is not zero. Just as in the case of Cartesian coordinates, we can represent any point **X** in the plane by a vector sum of the form $u\mathbf{L} + v\mathbf{M}$ (recall section 15.1). We will call $u,\ v$ the **coordinates** of the point **X** relative to the **basis** $\mathbf{L}, \mathbf{M}$.

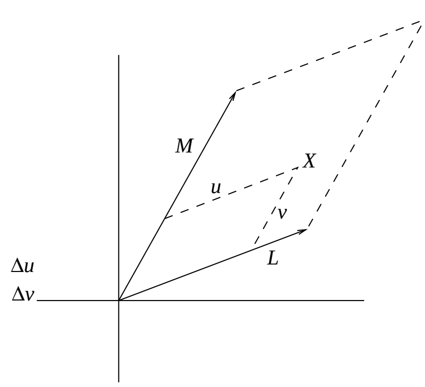
Now cover the plane with a grid of lines parallel to **L** and **M** (see figure 17.13).

Figure 17.13



If we take the grid sufficiently fine, we can approximate the area of a region R by the sum of the areas of the grid parallelograms wholly contained inside the region; that is, the limit of these sums, as the grid becomes infinitely fine, is the area of the region. If the side lengths of a typical parallelogram (in u, v coordinates) are $\Delta u,\ \Delta v$, the area of the parallelogram is $|\det(\mathbf{L}, \mathbf{M})|\Delta u \Delta v$ (see figure 17.14).

Figure 17.14



Thus, in the limit:

$$\int\int_R dA = \lim \sum |\det(\mathbf{L}, \mathbf{M})|\Delta u \Delta v = \int\int_S |\det(\mathbf{L}, \mathbf{M})| du dv\ .$$

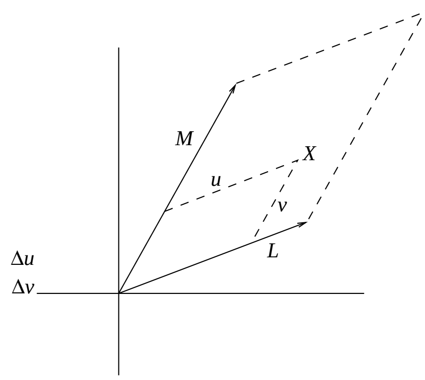
where S is the region in u, v coordinates corresponding to R. Thus the area of R is $|\det(\mathbf{L}, \mathbf{M})|$ times the area of S. This same argument works for the integral of a function over the region R:

$$\int\int_R f(x, y) dA = \lim \sum f(\overline{x}, \overline{y}) |\det(\mathbf{L}, \mathbf{M})|\Delta u \Delta v$$

$$= \int\int_S f(x(u,v), y(u,v)) |\det(\mathbf{L}, \mathbf{M})| du dv \ .$$

Example 17.13. Find $\int\int_R y dA$, where R is the parallelogram with vertices at (0,0), (5,3), (2,6), (8,9) (see figure 17.15).

Figure 17.14



R can be described as the parallelogram with sides $\mathbf{L} = 5\mathbf{I} + 3\mathbf{J}, \ \mathbf{M} = 2\mathbf{I} + 6\mathbf{J}$. We see (using figure 17.14) that a point $\mathbf{X} = x\mathbf{I} + y\mathbf{J}$ is in R if we can write $\mathbf{X} = u\mathbf{L} + v\mathbf{M}$ with $0 \le u \le 1, \ 0 \le v \le 1$. Now, we solve for x and y in terms of u and v:

$$\mathbf{X} = u\mathbf{L} + v\mathbf{M} = u(5\mathbf{I} + 3\mathbf{J}) + v(2\mathbf{I} + 6\mathbf{J}) = (5u + 2v)\mathbf{I} + (3u + 6v)\mathbf{J} \ ,$$

so $x = 5u + 2v, \ y = 3u + 6v$. That change of coordinates realizes R as the image of the unit square S in (u, v) space. We have

$$\det(\mathbf{L}, \mathbf{M}) = \det \begin{pmatrix} 5 & 3 \\ 2 & 6 \end{pmatrix} = 24 \ .$$

Then

$$\int\int_R y dA = \int\int_S (3u + 6v) |\det(\mathbf{L}, \mathbf{M})| du dv = 24 \int_0^1 [\int_0^1 (3u + 6v) du] dv = 4.5 \ .$$

Now, suppose that we make a (not necessarily linear) change of variables in a region $R: \ x = x(u, v), \ y = y(u, v)$. To say that this is a change of variables is to say that the values of u and v are determined by the point (x, y) in R, that is, in principle, we can solve these equations for u and v in terms of x and y.

In terms of vectors we can write this as $\mathbf{X}(u, v) = x(u, v)\mathbf{I} + y(u.v)\mathbf{J}$. We can see how to calculate integrals in the (u, v) plane by following the above argument for a linear change of variables. However, now the factor by which we multiply is not constant, and depends at each point upon the relationship between differential rectangles at that point. This relationship is given by the linear approximation at that point. Specifically, If we select, at a point in (u, v) space, a rectangle with sides in the coordinate directions, and of side lengths du, dv, the image is, in the linear approximation, the parallelogram spanned by the vectors $\mathbf{X}_u du$ and $\mathbf{X}_v dv$ (see figures 17.15 and 17.16).

Figure 17.15

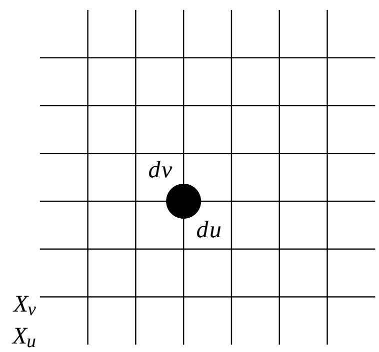
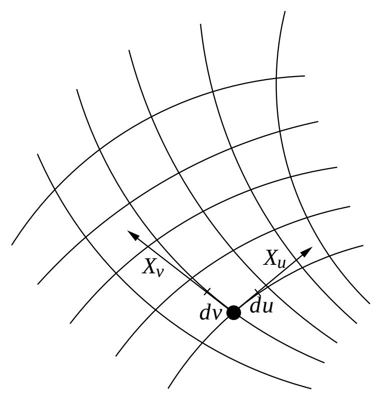


Figure 17.16



In this figure, the vectors $\mathbf{X_u}$ and $\mathbf{X_v}$ play the role of the vectors **L** and **M** in the linear case, since they represent a move of one unit in the respective coordinate directions. This parallellogram then has area $|\det(\mathbf{X}_u, \mathbf{X}_v)|dudv = |\mathbf{X}_u du \times \mathbf{X}_v dv|$.

Definition 17.4. Given a change of coordinates $x = x(u, v),\ y = y(u, v)$, the determinant

$$\det(\mathbf{X}_u, \mathbf{X}_v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}$$

is called the **Jacobian** of the change of variables and is denoted by

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

Proposition 17.4. If $x = x(u, v),\ y = y(u, v)$ is a change of coordinates in the region R, then we have this equation for the differential of area:

$$dA = |\mathbf{X}_u \times \mathbf{X}_v| dudv = \left| \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix} \right| dudv . \tag{17.7}$$

If R is the image of the region S in (u,v)-space, and f is an integrable function on R, then

$$(17.8) \qquad \int_R f(x,y)dxdy = \int_S f(x(u,v),y(u,v))\Big|\det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}\Big|dudv \ .$$

As we have seen in Chapter 13, the determinant $\det(\mathbf{X}_u, \mathbf{X}_v)$ is the determinant of the two by two matrix whose rows are the components of the vectors; that is, the determinant in (17.7). (17.8) follows from (17.7) intuitively: we can cover the region R by a grid formed of the level curves of the functions u and v, and get a good approximation by the right hand side, for a grid fine enough.

To illustrate this, we return to polar coordinates as the change of variables $x = r\cos\theta,\ y = r\sin\theta$, or, vectorially

$$\mathbf{X}(r.\theta) = r\cos\theta\mathbf{I} + r\sin\theta\mathbf{J} \ .$$

Now, for a differential rectangle in r,θ-space of side lengths $dr,\ d\theta$, the corresponding figure in x,y-space is the parallelogram bounded by the vectors $\mathbf{X}_r dr$ and $\mathbf{X}_\theta d\theta$ (see figure 17.11). We have

$$\mathbf{X}_r \times \mathbf{X}_\theta = \det\begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} = r \ ,$$

and the element of area is

$$dA = |\mathbf{X}_r dr \times \mathbf{X}_\theta d\theta| = rdrd\theta \ .$$

Example 17.14. Show that the area of an ellipse of major radius a and minor radius b is πab.

We consider the ellipse as the region E bounded by the curve

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \ .$$

The change of variables $x = au\ y = bv$ realizes this ellipse as the image of the unit disk D in (u,v) space. Thus

$$Area = \int\int_E dxdy = \int\int_D \Big|\frac{\partial(x,y)}{\partial(u,v)}\Big|dudv = ab\int\int_D dxdy = \pi ab \ ,$$

since the last integral is the area of the disc of radius 1, which is π.

Example 17.15. Find the area of the ellipse E bounded by the curve $x^2 + 4xy + 13y^2 = 16$.

If we complete the square the equation becomes

$$(x+2y)^2 + 9y^2 = 16 \ .$$

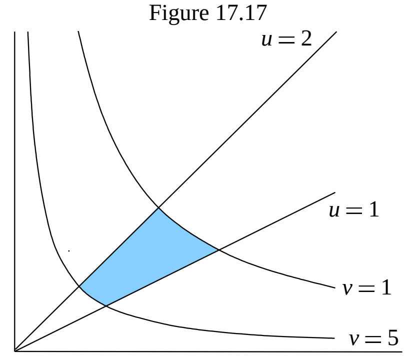
If we let $u = x+2y,\ v = 3y$, this is the image of the disk D in (u,v) space bounded by $u^2+v^2 = 16$. The radius is 4, and the area is 16π. Solving for x,y in terms of u,v, we have $x = u-2v/3,\ y = v/3$. The Jacobian of the change of variables is

$$\det\begin{pmatrix} 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix} = \frac{1}{3} \ .$$

Thus
$$Area(E) = \int\int_E dxdy = \frac{1}{3}\int\int_D dudv = \frac{1}{3}(Area(D)) = \frac{16\pi}{3}\ .$$

Example 17.16. Find the area of the region R in the first quadrant bounded by the curves $x = y,\ x = 2y,\ xy = 1,\ xy = 5$ (see figure 17.17).

Figure 17.17



Make the change of variables $u = x/y,\ v = xy$. Then the region can be described by the inequalities $1 \le u \le 2,\ 1 \le v \le 5$. Let S represent this region in u, v space. To integrate in u, v space, we have to find the Jacobian. First we solve the equations for x and y in terms of u and v, obtaining:

$$x = \sqrt{uv}\ , \quad y = \sqrt{\frac{v}{u}}\ ,$$

so that
$$\det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix} = \det\begin{pmatrix} \frac{1}{2}\sqrt{\frac{v}{u}} & -\frac{1}{2u}\sqrt{\frac{v}{u}} \\ \frac{1}{2}\sqrt{\frac{u}{v}} & \frac{1}{2}\sqrt{\frac{1}{uv}} \end{pmatrix} = \frac{1}{4}\sqrt{\frac{v}{u^2v}} + \frac{1}{4u}\sqrt{\frac{vu}{uv}} = \frac{1}{2u}\ .$$

Thus

$$Area(R) = \int\int_R dxdy = \int\int_S \frac{1}{2u}dudv = \frac{1}{2}\int_1^2 [\int_1^5 \frac{1}{u}dv]du = 2\int_1^2 \frac{du}{u} = 2\ln 2\ .$$

Surface Area

Consider a surface $\mathbf{S}$ in three dimensions given parametrically by

$$\mathbf{X}(u, v) = x(u, v)\mathbf{I} + y(u, v)\mathbf{J} + z(u, v)\mathbf{K}$$

where (u, v) lies in a region R. To approximate the area of $\mathbf{S}$ we cut R up into small rectangles given by a fine grid, calculate the area of the parallelogram (in the tangent plane at a point in R) approximating the image of the rectangle on $\mathbf{S}$, and add this up over all rectangles, as we did for the plane (see figure 17.15 and 17.16; now figure 17.17 is a grid on a surface in space). Turning to the language of differentials, let dS represent the differential of the surface area. This is the area of the rectangle in space spanned by the vectors $\mathbf{X}_u du$ and $\mathbf{X}_v dv$. Thus

$$dS = \left|\mathbf{X}_u du \times \mathbf{X}_v dv\right| , \tag{17.9}$$

and thus the area is

$$Area(\mathbf{S}) = \int\int_R dS = \int\int_R \left|\mathbf{X}_u \times \mathbf{X}_v\right| dudv \quad \text{(Surface given parametrically)} . \tag{17.10}$$

Example 17.17. Find the area of the piece of the paraboloid $z = x^2 + y^2$ cut off by the plane $z = 1$.

Parametrize the surface using cylindrical coordinates:

$$\mathbf{X}(x, y) = r\cos\theta\mathbf{I} + r\sin\theta\mathbf{J} + r^2\mathbf{K} , \quad 0 \le \theta \le 2\pi,\ 0 \le r \le 1 .$$

Then

$$\mathbf{X}_\mathbf{r} = \cos\theta\mathbf{I} + \sin\theta\mathbf{J} + 2r\mathbf{K} , \quad \mathbf{X}_\theta = -r\sin\theta\mathbf{I} + r\cos\theta\mathbf{J} .$$

$$dS = \left|\mathbf{X}_r \times \mathbf{X}_\theta\right| drd\theta = r\sqrt{1 + 4r^2}dr .$$

$$Area = \int_0^{2\pi}\int_0^1 r\sqrt{1 + 4r^2}drd\theta = \frac{\pi}{6}(\sqrt{125} - 1) .$$

If the surface is the graph of a function $z = f(x, y)$, we can look at this as given parametrically by

$$\mathbf{X}(x, y) = x\mathbf{I} + y\mathbf{J} + f(x, y)\mathbf{K} ,$$

and we find $\mathbf{X}_x = \mathbf{I} + f_x\mathbf{K}, \mathbf{X}_y = \mathbf{J} + f_y\mathbf{K}$ and so, from (17.9): $dS = \sqrt{1 + f_x^2 + f_y^2}dxdy$ and we have

$$Area(\mathbf{S}) = \int\int_R dS = \int\int_R \sqrt{1 + f_x^2 + f_y^2}dudv \quad \text{(Surface given as a graph)} . \tag{17.11}$$

Example 17.18. Find the area of the cone $z^2 = x^2 + y^2$ lying over the triangle in the first quadrant bounded by the line $2x + y = 2$.

We represent the surface as the graph of the function $f(x, y) = \sqrt{x^2 + y^2}$ over the region $) \le x \le 1,\ 0 \le y \le 2 - 2x$. Differentiating:

$$f_x = \frac{x}{\sqrt{x^2 + y^2}} , \quad f_y = \frac{y}{\sqrt{x^2 + y^2}} ,$$

so

$$dS = \sqrt{1 + \frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}}dxdy = \sqrt{2}dxdy .$$

Thus

$$Area = \sqrt{2}\int_0^1 \int_0^{2-2x} dydx = \sqrt{2}\ .$$

Example 17.19. If the function is given in polar coordinates: $z = f(r,\theta)$. we write

$$\mathbf{X}(r,\theta) = r\cos\theta\mathbf{I} + r\sin\theta\mathbf{J} + f(r,\theta)\mathbf{K}\ .$$

Then $\mathbf{X}_r = \cos\theta\mathbf{I} + \sin\theta\mathbf{J} + f_r\mathbf{K},\ \mathbf{X}_\theta = -r\sin\theta\mathbf{I} + r\cos\theta\mathbf{J} + f_\theta\mathbf{K}$ and the calculation gives

$$dS = \sqrt{r^2 + r^2 f_r^2 + f_\theta^2}drd\theta\ .$$

Notice, if the surface is the plane $z = 0$, then dS is just the element of area in polar coordinates: $rdrd\theta$.

Problems 17.4

1. The surface H, given in cylindrical coordinates by $z = 2\theta$ is a helicoid. What is the volume of the region R bounded above by $H, 0 \le \theta \le 2\pi$, below by the plane $z = 0$ and lying over the disc $r \le 1$?

2. A beach B is shaped in the form of a crescent. We model this on the region in the right half plane between the circle of radius 1, centered at the origin, and the circle of radius 3/4 centered at the point (3/4,0), where the units are in miles. Suppose that the human density σ decreases as we move from the beach according to $\sigma(x,y) = 1000(x^2+y^2)^{-2}$ people per square mile.What is the population on that beach?

3. The curve $z = (x-1)^2, 0 \le z \le 1$ is rotated about the z-axis, enclosing, together with the xy-plane, a 3-dimensional region R. R is filled with a substance whose density is inversely proportional to the distance from the z-axis. Find the total mass of this object.

4. As (u,v) runs through the region $u^2+v^2 \le 1$, the vector function

$$\mathbf{X}(u,v) = (u^2+v^2)\mathbf{I} + (u^2-v^2)\mathbf{J} + uv\mathbf{K}$$

describes a surface S in three space. Write down the double integral which must be calculated to find the surface area of S.

5. Find the volume of the region lying above the disc $x^2+y^2 \le 1$ in the xy-plane, and below the surface $z = \sin(\pi\sqrt{x^2+y^2}/2)$.

6. Find the mass of the lamina of the region R lying between the ellipses $x^2+4y^2 = 1$ and $x^2+4y^2 = 4$, where the density function is $\delta(x,y) = x^2+y^2$.

7. Find the area of the region R in the first quadrant bounded by the curves $y^2 = 2x,\ y^2 = 5x,\ x^2 = 4y,\ x^2 = 10y$.

8. Find the surface area of the part of the hyperbolic paraboloid $z = xy$ that lies inside the cylinder $x^2+y^2 \le 4$.

9. Find the surface area of the part of the hyperbolic paraboloid $z = y^2 - x^2$ that lies between the cylinders $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.

10. Find the surface area of the part of the surface $z = (2/3)(x^{3/2} + y^{3/2})$ that lies above the triangle in the first quadrant bounded by the line $x + y = 1$.

17.5 Triple Integrals

These are just like double integrals, but with another integration to perform. Although this is conceptually a simple extension of the idea, in practice it can get very complicated. For example, in two variables, there are just two different ways to integrate by iteration, depending on how we order the variables. But since there are six different orderings of three variables, there are now six different types of iterated integrals. In this section we shall illustrate the basic ideas, and not go into very much detail or complexity.

We begin with the definition of the integral of a function $w = f(x, y, z)$ over a region R in three dimensions. Cover R with a grid formed of the coordinate planes. and form the sum

$$\sum f(x, y, z)\Delta V$$

over all the cubes of the grid in R, where f is evaluated at a point of the cube, and ΔV is the volume of the cube. Taking the limit as the grid becomes infinitely fine, we obtain

$$\iiint_R f(x, y, z)dV \ . \tag{17.12}$$

As remarked above, triple integrals can be evaluated as iterated integrals.

Proposition 17.5. Suppose that $w = f(x, y, z)$ is a continuous function on the rectangular parallelipiped $R:\ a \le x \le b,\ c \le y \le d,\ p \le z \le q$. Then the triple integral (12) can be evaluated by iteration in any of six ways, depending upon which variable is chosen first. For example, if the variables are chosen in the order x, y, z, we have

$$\iiint_R f(x, y, z)dV = \int_a^b \Big[\int_c^d \Big[\int_p^q f(x, y, z)dz\Big]dy\Big]dx \ . \tag{17.13}$$

Example 17.20. Find the integral of $x^2 + yz$ over the region $R:\ 0 \le x \le 2,\ 1 \le y \le 4, 0 \le z \le 5$. If we iterate in the order x, y, z we get

$$\iiint_R (x^2 + yz)dV = \int_0^2 \Big[\int_1^4 \Big[\int_0^5 (x^2 + yz)dz\Big]dy\Big]dx \ .$$

The first integration gives

$$\int_0^5 (x^2 + yz)dz = (x^2 z + \frac{yz^2}{2})\Big|_0^5 = 5x^2 + \frac{25}{2}y \ .$$

The next integral is

$$\int_1^4 (5x^2 + \frac{25}{2}y)dy = (5x^2 y + \frac{25}{4}y^2)\Big|_1^4 = 20x^2 + 100 - 5x^2 - \frac{25}{4} = 15x^2 + \frac{375}{4} \ .$$

Finally, the last integration gives the desired result:

$$\int_0^2 (15x^2 + \frac{375}{4})dx = \frac{15x^3}{3} + \frac{375}{4}x\Big|_0^2 = \frac{455}{2} \ .$$

If we iterate in another order, the calculations are different (and could be easier or more difficult), but lead to the same answer. For example, let us iterate in the order y, z, x:

$$\iiint_R (x^2 + yz)dV = \int_1^4 \Big[\int_0^5 \Big[\int_0^2 (x^2 + yz)dx\Big]dz\Big]dy \ .$$

The first integral is

$$\frac{x^3}{3} + yzx\Big|_0^2 = 2yz + \frac{8}{3} \ .$$

The next integration leads to

$$yz^2 + \frac{8}{3}z\Big|_0^5 = 25y + \frac{40}{3} \ ,$$

and the final integration is

$$\int_1^4 (25y + \frac{40}{3})dy = \frac{25y^2}{2} + \frac{40}{3}y\Big|_1^4 = \frac{455}{2} \ .$$

Now, for the integral over a general region R, the situation can easily become quite complicated. Here is one of the possibilities;

Proposition 17.6 Suppose that $w = f(x, y, z)$ is a continuous function on the region R. If R can be described as the set of points (x, y, z) such that $a \le x \le b,\ u(x) \le y \le v(x)\ , \psi(x, y) \le z \le \phi(x, y)$, then

$$\iiint f(x,y,z)dV = \int_a^b \Big[\int_{u(x)}^{v(x)} \Big[\int_{\psi(x,y)}^{\phi(x,y)} f(x,y,z)dz\Big]dy\Big]dx \ . \tag{17.14}$$

Each ordering of the variables leads to a different possibility of calculating the triple integral by iteration; some of which may not work, and some of which may be easier than others.

Example 17.21 Find the volume of the region in the first octant bounded by the plane $3x+y+2z = 12$.

Since the values of the variables must be nonnegative, we see that $0 \le x \le 4,\ 0 \le y \le 12$ and $0 \le z \le 6$. For any x between 0 and 4, y ranges between 0 and $12 - 3y$. For such an x and y, z ranges between 0 and $(12 - 3x - y)/2$. Thus we can describe the region by the inequalities $0 \le x \le 4,\ 0 \le y \le 12 - 3x,\ 0 \le z \le (12 - 3x - y)/2$ which direct us to the iterated integral

$$Volume = \int_0^4 \Big[\int_0^{12-3x} \Big[\int_0^{\frac{12-3x-y}{2}} dz\Big]dy\Big]dx \ .$$

The first integration gives $(12 - 3x - y)/2$, and the second is

$$\frac{1}{2}\int_0^{12-3x} (12 - 3x - y)dy = \frac{12-3x}{2}y - \frac{y^2}{4}\Big|_0^{12-3x} = \frac{(12-3x)^2}{4} \ ,$$

and finally

$$Volume = \int_0^4 (\frac{(12-3x)^2}{4})dx = -\frac{(12-3x)^3}{36}|_0^4 = 48 \ .$$

Example 17.22. Find the z-coordinate of the centroid of the region of example 17. 21.

To do that, we find the moment of the region about the plane $z = 0$:

$$Mom_{\{z=0\}} = \iiint_R z dV \ .$$

In general, if we have a choice, it is best to leave the most involved integration to the last; thus we may want to take the variables in the order z, x, y. This gives us the iterated integral

$$Mom_{\{z=0\}} = \int_0^6 [\int_0^{12-2z} [\int_0^{(12-2z-y)/3} z dx] dy] dz \ .$$

The first integral leads to $z(12-2z-y)/3$, and the second is

$$\frac{1}{3}\int_0^{12-2z} z(12-2z-y)dy = \frac{z}{3}((12-2z)y - \frac{y^2}{2})|_0^{12-2z} = \frac{z}{6}(12-2z)^2 \ .$$

After a little algebra, we obtain

$$Mom_{\{z=0\}} = \frac{2}{3}\int_0^6 (36z - 12z^2 + z^3)dz = 72 \ .$$

Thus

$$\bar{z} = \frac{Mom_{\{z=0\}}}{Mass} = \frac{72}{48} = 1.5 \ .$$

Example 17.23. Let R be the region bounded by the xy-plane, the plane $x = 1$ and under the surface $z = x^2 - y^2$. Find the centroid of R.

First we draw the diagram(see figure 17.19).

We see that we can describe the region as: $0 \le x \le 1, \ -x \le y \le x, \ 0 \le z \le x^2 - y^2$. Thus

$$Volume = \int_0^1 [\int_{-x}^x [\int_0^{x^2-y^2} dz] dy] dx = \int_0^1 [\int_{-x}^x (x^2-y^2)dy = \frac{4}{3}\int_0^1 x^3 dx = \frac{1}{3} \ ,$$

$$Mom_{\{x=0\}} = \int_0^1 [\int_{-x}^x [\int_0^{x^2-y^2} x dz] dy] dx = \frac{4}{3}\int_0^1 x^4 dx = \frac{4}{15} \ .$$

The moment about the plane $y = 0$ is zero because of symmetry. Finally,

$$Mom_{\{z=0\}} = \int_0^1 [\int_{-x}^x [\int_0^{x^2-y^2} z dz] dy] dx = \frac{1}{2}\int_0^1 \int_{-x}^x (x^2-y^2)^2 dy dx \ .$$

Working out the square, we find the inner integral to be

$$\int_{-x}^{x}(x^4 - x^2y^2 + y^4)dy = (x^4y - \frac{2}{3}x^2y^3 + \frac{y^5}{5})\Big|_{-x}^{x} = 2x^5(1 - \frac{2}{3} + \frac{1}{5}) = \frac{16}{15}x^5 \ .$$

Thus

$$Mom_{\{z=0\}} = \frac{8}{15}\int_0^1 x^5 dx = \frac{4}{45} \ ,$$

and

$$\bar{x} = \frac{4/15}{1/3} = \frac{4}{5} \ , \quad \bar{y} = 0 \ . \quad \bar{z} = \frac{4/45}{1/3} = \frac{4}{15} \ .$$

Integration in other coordinates.

First, recall that the volume of the parallelipiped spanned by the three vectors $\mathbf{U}$, $\mathbf{V}, \mathbf{W}$ is $|\mathbf{U} \times (\mathbf{V} \cdot \mathbf{W})| = |\det(\mathbf{U}, \mathbf{V}, \mathbf{W})|$, where the determinant is that of the matrix whose rows are the components of $\mathbf{U}, \mathbf{V}, \mathbf{W}$. Now, suppose that we want to evaluate the triple integral

$$\iiint_R f dxdydz \tag{17.15}$$

where R corresponds to a region S in u, v, w space under a change of variables

$$\mathbf{X} = \mathbf{X}(u, v, w) = x(u, v, w)\mathbf{I} + y(u, v, w)\mathbf{J} + z(u, v, w)\mathbf{K} \ . \tag{17.16}$$

By a change of variables, we mean that the coordinates (u, v, w) in S uniquely determine a point in R; that is, in principle, we could solve the equations (16) for u, v, w in terms of x, y, z. Now, we can perform the integration in u, v, w space by selecting a fine grid in S, covering R by the image of the grid, and forming the sum $\sum f(X)\Delta V$, where ΔV is the volume of the figure in R corresponding to a typical cube in the grid on S. This will be an approximation to (17.15), which gets better as the grid gets finer.

If the grid is sufficiently fine, we can, at each cube, approximate by the linear approximation to the change of variables:

$$d\mathbf{X} = \mathbf{X}_u du + \mathbf{X}_v dv + \mathbf{X}_w dw \ .$$

A cube in S of side lengths du, dv, dw corresponds to the parallelipiped in R spanned by the vectors $\mathbf{X}_u du$, $\mathbf{X}_v dv$, $\mathbf{X}_w dw$, and the volume of that parallelipiped is

$$dV = |\det(\mathbf{X}_u, \mathbf{X}_v, \mathbf{X}_w)| dudvdw \ . \tag{17.17}$$

The factor in (17.17) is called the **Jacobian** of the variable change, and is calculated as the (absolute value of the) determinant of the matrix whose rows are $\mathbf{X}_u$, $\mathbf{X}_v$, $\mathbf{X}_w$. This matrix is denoted by

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

Proposition 17.7. Suppose that we are given the change of variables (16) so that the region R in (x, y, z) space corresponds to the region S in (u, v, w) space. Then we can calculate triple integrals by integrating over S as follows:

$$\iiint_R f(\mathbf{X})dxdydz = \iiint_S f(\mathbf{X}(u, v, w))\Big|\frac{\partial(x, y, z)}{\partial(u, v, w)}\Big|dudvdw \ .$$

Example 17.24. Find the volume of the region R given by the inequalities

$$0 \leq z \leq 4 \ , \quad 0 \leq y + z \leq 3 \ , \quad 0 \leq x + y + z \leq 5 \ .$$

This region is a parallelipiped, so by the appropriate change of coordinates, can be made to correspond to a rectangular parallelipiped. That is, we make the change of variables

$$u = x + y + z \ , \quad v = y + z \ , \quad w = z$$

so that R corresponds to the region S given by the inequalities $0 \leq u \leq 5,\ 0 \leq v \leq 3,\ 0 \leq w \leq 4$. Thus

$$Volume = \int\int\int_R dxdydz = \int\int\int_S \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| dudvdw \ .$$

Now, to calculate the Jacobian, we solve for x, y, z in terms of u, v, w:

$$x = u - v \ , \quad y = v - w \ , \quad z = w \ ,$$

so that

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \det\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \ .$$

Thus

$$Volume = \int_0^5 \int_0^3 \int_0^4 dudvdw = 60 \ .$$

When we make a linear change of variables the Jacobian is a constant (but not always 1, as above). By using the following fact about determinants, we need not actually solve for x, y, z in terms of u, v, w:

$$\frac{\partial(x,y,z)}{\partial(u,v,w)}\frac{\partial(u,v,w)}{\partial(x,y,z)} = 1 \ . \tag{17.18}$$

Example 17.25. Let $u = 3x - y + z,\ v = x + y + 2z,\ w = x - z$. Find the volume of the parallelipiped given by the inequalities $-2 \leq u \leq 3,\ 0 \leq v \leq 4,\ 2 \leq u \leq 7$.

We calculate

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \det\begin{pmatrix} 3 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & 2 & -1 \end{pmatrix} = -7 \ .$$

Thus, by (17.18),

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = -\frac{1}{7} \ ,$$

so that

$$Volume = \int_{-2}^3 \int_0^4 \int_2^7 \frac{1}{7} dudvdw = \frac{100}{7} \ .$$

Usually, we change coordinates only when the statement of the problem strongly suggests a different set of coordinates. Thus, if the problem has symmetry about the z-axis, we may change to spherical coordinates, and if there is symmetry about the origin, an change to spherical coordinates is indicated. The calculation of theJacobians for cylindrical and spherical coordinates leads to these expressions:

Cylindrical coordinates: $dV = rdrd\theta dz$.

Spherical coordinates: $dV = \rho^2 \sin\phi d\rho d\theta d\phi$.

Example 17.26. Find the volume and the centroid of the region bounded by the hyperboloid $x^2 + y^2 - z^2 = 1$ and the planes $z = 0,\ z = 2$.

Here, because of the symmetry about the z-axis, we are led to cylindrical coordinates. In these coordinates, the region is given by $0 \le \theta \le 2\pi,\ 0 \le z \le 2,\ 0 \le r \le \sqrt{1+z^2}$. Thus

$$Volume = \int\int\int_R dxdydz = \int\int\int_R rdrdzd\theta = \int_0^{2\pi}[\int_0^2[\int_0^{\sqrt{1+z^2}} rdr]dz]d\theta\ .$$

The inner integral is

$$\int_0^{\sqrt{1+z^2}} rdr = \frac{1+z^2}{2}\ ,$$

and the next integral is

$$\int_0^2 \frac{1+z^2}{2}dz = \frac{z}{2} + \frac{z^3}{6}\Big|_0^2 = \frac{7}{3}\ ,$$

and the final integration gives the volume as $14\pi/3$. Now, because of the symmetry, $\bar{x} = 0,\ \bar{y} = 0$. To calculate $\bar{z}$, we need

$$Mom_{\{z=0\}} = \int\int\int_R zrdrdzd\theta = \int_0^{2\pi}[\int_0^2[\int_0^{\sqrt{1+z^2}} zrdr]dz]d\theta\ .$$

This time the first integral gives us $z(1+2^2)/2$, and the next is

$$\int_0^2 \frac{z+z^3}{2}dz = \frac{z^2}{4} + \frac{z^4}{8}\Big|_0^2 = 3\ ,$$

so that $Mom_{\{z=0\}} = 6\pi$. Thus

$$\bar{z} = \frac{6\pi}{14\pi/3} = 9/7\ .$$

Example 17.27. The region R between the spheres of radius 4 and 5 is filled with a material whose density is given by $\delta(x,y,z) = 1 + x^2 + y^2$. Find the mass of this region.

In spherical coordinates, the region is given by the inequalities $0 \le \theta \le 2\pi,\ 0 \le \phi \le \pi,\ 4 \le \rho \le 5$, and $\delta = 1 + \rho^2 \sin^2\phi$. Thus

$$Mass = \int_0^{2\pi}[\int_0^{\pi}[\int_4^5 (1 + \rho^2 \sin^2\phi)\rho^2 \sin\phi d\rho]d\phi]d\theta\ .$$

The innermost integral is

$$\int_4^5 (1+\rho^2\sin^2\phi)\rho^2\sin\phi d\rho = \frac{61}{3}\sin\phi + \frac{2101}{5}\sin^3\phi \ .$$

Now, to integrate with respect to ϕ, we calculate

$$\int_0^\pi \sin\phi d\phi = 2 \ , \quad \int_0^\pi \sin^3\phi d\phi = \int_0^\pi (1-\cos^2\phi)\sin\phi d\phi = \frac{4}{3} \ ,$$

and the integration with respect to θ introduces a factor of 2π. Thus

$$Mass = 2\pi(\frac{61}{3}(2) + \frac{2101}{5}\frac{4}{3}) = 1201.87\pi = 3775.67 \ .$$

Problems 17.5

1. Find the mass and the x-coordinate of the center of mass of the solid bounded by the planes $x=0,\ y=0,\ z=0,\ x+y+z=1$ with the density function $\rho(x,y,z)=y$.

2. Find the center of mass of the piece of the solid parabolic shell $z \leq 16-(x^2+y^2)$ lying above the xy-plane.

3. Find the average value of $f(x,y,z)=x+y+z$ over the region R in the first octant (the region where all the coordinates are positive) under the plane $x+y+z=1$.

4. The curve $z=(x-1)^2, 0 \leq z \leq 1$ is rotated about the z-axis, enclosing, together with the xy-plane, a 3-dimensional region R. R is filled with a substance whose density is inversely proportional to the distance from the z-axis. Find the total mass of this object.

5. Evaluate

$$\int\int\int_R (x^2+y^2+z^2)dxdydz$$

where R is the ball $x^2+y^2+z^2 \leq 4$.

6. Find the centroid of the region R described in example 17.24.

Chapter XVIII. Vector Calculus

In this chapter we develop the fundamental theorem of the Calculus in two and three dimensions. This begins with a slight reinterpretation of that theorem. Consider the endpoints a, b of the interval $[a, b]$ from a to b as the boundary of that interval. Then the fundamental theorem, in this form:

$$f(b) - f(a) = \int_a^b \frac{df}{dx}(x)dx \ ,$$

relates the values of a function at the boundary with the values of its derivative in the interior. Stated this way, the fundamental theorems of the Vector Calculus (Green's, Stokes' and Gauss' theorems) are higher dimensional versions of the same idea. However, in higher dimensions, things are far more complex: regions in the plane have curves as boundaries, and for regions in space, the boundary is a surface, and surfaces in space have curves as boundaries. This requires a reinterpretation of the term $f(b) - f(a)$, as a *signed* sum of the values of f on the boundary, the sign being determined by the side on which the interval lies (it is to the right of a and to the left of b). This leads to the understanding that in higher dimensions both sides will be integrals; for example, for a region R in the plane with C as its boundary, the term $f(b) - f(a)$ becomes an integral over the curve C. And in three dimensions, we will have two versions of the fundamental theorem, one relating integrals over a region with integrals over the bounding surface, and another relating integrals over surfaces with integrals over the bounding curve (and with the relation involving some form of differentiation).

We will not give derivations, or even intuitive arguments for the proofs of these theorems. First of all, the idea of the proof is to reduce the theorem to the one-variable fundamental theorem; in this process, the notational complexity is constantly threatening to get out of hand. The proofs then become masterful displays of technical control, and provide little insight. The insight comes from the physical interpretation of these theorems (indeed, so also did the first proofs), particularly in terms of fluid flows. For example, Gauss' theorem simply says that, for a fluid in flow we can measure the rate of change of the amount of fluid in a given region in two ways: directly over the region, or instead, by measuring the rate of passage through the boundary.

18.1 Vector Fields

Definition 18.1. A **vector field** is an association of a vector to each point $\mathbf{X}$ of a region R:

$$\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{I} + Q(x, y, z)\mathbf{J} + R(x, y, z)\mathbf{K} \ .$$

For example, the vector field

$$\mathbf{X}(x, y, z) = x\mathbf{I} + y\mathbf{J} + z\mathbf{K} \tag{18.1}$$

is the field of vectors pointing outward from the origin, whose length is equal to the distance from the origin. The field $\mathbf{U} = (1/r)\mathbf{X}$ (where $r(x, y, z) = (x^2 + y^2 + z^2)^{1/2}$) is the unit vector field with the same direction.

Example 18.1 (Gravitation). According to Newton's Law of gravitation, two bodies attract each other with a force proportional to the product of the masses, and inversely proportional to the square of the distance between them. Suppose one body, of mass M is situated at the origin.

Then another body of mass m, situated at the point $\mathbf{X}$ experiences the gravitational force due to M:

$$\mathbf{F} = -\frac{GMm}{r^2}\mathbf{U} \ ,$$

where G is Newton's universal constant of gravitation, and $\mathbf{U}$ is the unit vector pointing the direction of $\mathbf{X}$. If we want to concentrate on the effect of the mass M on bodies in its vicinity, we introduce the **gravitational field** of M:

$$\mathbf{G}(\mathbf{X}) = -\frac{GM}{r^2}\mathbf{U} = -\frac{GM}{r^3}\mathbf{X} \ .$$

Since $\mathbf{F} = m\mathbf{A}$, a body of mass m at $\mathbf{X}$ accelerates toward the origin with acceleration $\mathbf{G}(\mathbf{X})$.

Definition 18.2. Suppose the region R is filled with a fluid which is in motion. We can describe the motion by following the individual particles. Let $\mathbf{X}(\mathbf{X_0}, t)$ be the position at time t of the particle which was at $\mathbf{X_0}$ at time $t = 0$. The **velocity field** of the motion is the velocity of the particle at position $\mathbf{X}$ at time t, represented by $\mathbf{V}(\mathbf{X}, t)$ .This is a time-dependent vector field in the region R. We say that the flow is **steady** if its velocity field is independent of time.

In studying a fluid in motion, we are not interested in the history of particular particles, but in the fluid as a whole. Thus, it is the velocity field of the fluid that is the object of study, rather than the equations of motion. It can be shown that the velocity field completely determines the motion.

Example 18.2. Suppose a fluid is flowing on the plane radially away from the origin. In this case the origin is called a *source*; if the fluid were flowing toward the origin, we call it a *sink*. The equation of motion is given by

$$\mathbf{X}(\mathbf{X_0}, t) = f(t)\mathbf{X_0} \quad \text{for some scalar function } f \text{ with } f(0) = 1 \ .$$

Let's look at the case $f(t) = e^{at}$. We find the velocity field as follows. First, the velocity of the particle originally at $\mathbf{X_0}$ is

$$\frac{\partial}{\partial t}\mathbf{X}(\mathbf{X_0}, t) = \frac{d}{dt}(e^{at})\mathbf{X_0} = ae^{at}\mathbf{X_0} \ .$$

But this is $a\mathbf{X}$, so the velocity field is $\mathbf{V}(\mathbf{X}) = a\mathbf{X}$, and the flow is steady.

However, if, say $f(t) = 1 + t$ so that $\mathbf{X}(\mathbf{X_0}, t) = (1 + t)\mathbf{X_0}$,we have

$$\frac{\partial}{\partial t}\mathbf{X}(\mathbf{X_0}, t) = \mathbf{X_0} = (1 + t)^{-1}\mathbf{X} \ ,$$

so the flow is time-dependent.

The terminology may seem confusing: in the first case, the particle's speed is increasing exponentially, while in the second case the particle's speed is constant. But, if we look at a particular point $\mathbf{X}$ in space, then in the first case, the fluid is always moving with the same velocity through that point, while in the second case, the fluid slows down at that point over time.

Example 18.3. Suppose a fluid is rotating on the plane about the origin in the counterclockwise direction at constant angular velocity ω. From the description, this is a steady flow; let's find its velocity field. At a point $\mathbf{X}$, particles move through $\mathbf{X}$ along the circle of radius $|\mathbf{X}|$ at angular velocity ω.Thus the velocity of the fluid at $\mathbf{X}$ is of magnitude $\omega|\mathbf{X}|$ and in the direction tangent to to the circle through $\mathbf{X}$, so $\mathbf{V}(\mathbf{X}) = \omega\mathbf{X}^\perp$.

Definition 18.3. A differentiable function $w = f(x, y, z)$ has associated to it its **gradient field**

$$\nabla w = \frac{\partial f}{\partial x}\mathbf{I} + \frac{\partial f}{\partial y}\mathbf{J} + \frac{\partial f}{\partial z}\mathbf{K} \ . \tag{18.2}$$

The surfaces $f(x, y, z)$ =const. are orthogonal to the vector field (18.2), and are called the **equipotentials**, and the function f, a **potential** for the field.

So, the flow associated to a gradient field is easily visualized as being in the direction perpendicular to these equipotential surfaces. A natural question is: when is a vector field $\mathbf{F}$ the gradient of a function; that is, when does a vector field have a potential function? If the vector field with the components $\mathbf{F} = P\mathbf{I} + Q\mathbf{J} + R\mathbf{K}$ is a gradient, so looks like (18.2), then, because of the equality of mixed derivatives, we must have

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \qquad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \qquad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \ . \tag{18.3}$$

If these conditions are satisfied, then we can try to find the potential function by integrating one variable at a time.

Example 18.4. Let $\mathbf{F} = (2xy + x)\mathbf{I} + x^2 - y\mathbf{J}$. Is $\mathbf{F}$ a gradient field? If so, find the potential function.

First, we check that the condition (18.3) is satisfied:

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y}(2xy + x) = 2x \qquad \frac{\partial Q}{\partial x} = \frac{\partial}{\partial y}(x^2 - y) = 2x \ .$$

So, we have a chance of finding a function f such that $\nabla f = \mathbf{F}$. To find f we have to solve the equations

$$\frac{\partial f}{\partial x} = 2xy + x, \ \frac{\partial f}{\partial y} = x^2 - y \ .$$

We can find a function satisfying the first equation by integrating with respect to x; so we try $f(x, y) = x^2y - x^2/2$. Now we see if this f satisfies the second equation:

$$\frac{\partial f}{\partial y} = x^2 \ , \tag{18.4}$$

which unfortunately is not $x^2 - y$. However, since the derivative with respect to x of any function of y is zero, we could also have tried

$$f(x, y) = x^2y + x^2/2 + \phi(y)$$

for some yet-to-be-determined $\phi(y)$. Now, we have, instead of (18.4),

$$\frac{\partial f}{\partial y} = x^2 + \phi'(y) \ ;$$

setting that equal to Q gives the equation $\phi'(y) = -y$, so we can take $\phi(y) = -y^2/2$. We conclude that our solution is

$$f(x, y) = x^2 y + \frac{x^2}{2} - \frac{y^2}{2} + C \ ,$$

for any constant C. The reason that the terms involving x disappear in equation (18.4) is precisely that the condition $\partial P/\partial y = \partial Q/\partial x$ is satisfied; if it were not, this procedure would break down at this point.

Example 18.5. The procedure in three dimensions is the same, but longer. Suppose we are given the vector field $\mathbf{F} = (y^2 z + 1)\mathbf{I} + (2xyz + z)\mathbf{J} + (xy^2 + y + 1)\mathbf{K}$, and we are told that it is the differential of a function f. Find f.

Since we are told that there is a potential function, we need not verify conditions (18.3). We start with

$$\frac{\partial f}{\partial x} = y^2 z + 1 \ .$$

Integrating both sides with respect to x, (thinking of y and z as constants), we obtain

$$f(x, y, z) = xy^2 z + x + \phi(y, z) \tag{18.5}$$

where ϕ is an unknown function of y and z alone. Now, differentiating this equation, since $\partial f/\partial y = 2xyz + z$, we obtain

$$2xyz + z = 2xyz + \frac{\partial \phi}{\partial y} \ , \tag{18.6}$$

or

$$\frac{\partial \phi}{\partial y} = z \ . \tag{18.7}$$

Now we do the same, integrating both sides with respect to y:

$$\phi(y, z) = yz + \psi(z) \ , \tag{18.8}$$

for some unknown function $\psi(z)$. Thus (18.5) now becomes

$$f(x, y, z) = xy^2 z + x + yz + \psi(y, z) \ . \tag{18.9}$$

Differentiating now with respect to z:

$$xy^2 + y + 1 = xy^2 + y + \frac{\partial \psi}{\partial z} \tag{18.10}$$

so $\partial\psi/\partial z = 1$, and thus $\psi(z) = z + C$. Putting this back in (18.9), we have found

$$f(x, y, z) = xy^2 z + x + yz + z + C \ .$$

The reason that the variable x disappeared from (18.6) and x and y from (18.10) is precisely because of the conditions (18.3); if they did not hold there would be no such function f, and we could not have solved equations (18.7) and (18.10).

Example 18.6. We point out at this time that these methods make sense only in the domain in which the solution function f is well-defined, even if the given vector field is well-defined in a bigger region. Take, for example, the polar function

$$\theta = \arctan \frac{y}{x} \ . \tag{18.11}$$

Since θ is periodic, it is only well-defined (single-valued) in the plane outside of a ray from the origin, say the ray $x \geq 0$. However,

$$\nabla\theta = -\frac{y}{x^2+y^2}\mathbf{I} + \frac{x}{x^2+y^2}\mathbf{J} \ , \tag{18.12}$$

and this is well-defined in the whole plane, except for the origin. Thus, if we apply the above procedure to the vector field (18.12), we get (18.11), and we have to pick a particular branch of the arc tangent.

Two important concepts associated to a vector fields are its *divergence* and *curl*.

Definition 18.4. Let $\mathbf{F}$ be a vector field given by

$$\mathbf{F} = P\mathbf{I} + Q\mathbf{J} + R\mathbf{K} \ ,$$

where P, Q, R are scalar functions. The **divergence** of $\mathbf{F}$ is

$$\text{div } \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \ , \tag{18.13}$$

and the **curl** of $\mathbf{F}$ is

$$\text{curl } \mathbf{F} = (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})\mathbf{I} + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})\mathbf{J} + (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\mathbf{K} \ . \tag{18.14}$$

These are best interpreted in terms of the velocity field of a fluid flow. The divergence is the rate of expansion of the fluid at a point. The curl is a vector describing the rotation of the fluid near the point (the direction of the curl is the axis of rotation and the magnitude is a measure of the rate of rotation). The flow is called *incompressible* if its divergence is zero, and *irrotational* if its curl is zero. We note that the condition (18.3) for a vector field to be a gradient can be expressed as follows:

Propositon 18.1. Given a differentiable function f, its gradient field is irrotational; that is: curl $\nabla f = 0$. In order for a vector field to be a gradient field, it must be irrotational.

There is a notation which is very convenient in representing the gradient, div and curl. We consider ∇ as an operator on functions:

$$\nabla = \frac{\partial}{\partial x}\mathbf{I} + \frac{\partial}{\partial y}\mathbf{J} + \frac{\partial}{\partial z}\mathbf{K} \ . \tag{18.15}$$

Then, we have

$$\text{(18.16)} \qquad \text{div } \mathbf{F} = \nabla \cdot \mathbf{F} \ , \qquad \text{curl } \mathbf{F} = \nabla \times \mathbf{F} \ .$$

Two useful formulas are:

$$\nabla \cdot (\nabla \times \mathbf{F}) = 0 \ , \quad \text{or} \quad \text{div (curl } \mathbf{F}) = 0 \ .$$

$$\nabla \times \nabla f = 0 \ , \quad \text{or} \quad \text{curl } (\nabla f) = 0 \ .$$

If we are discussing vector fields in two dimensions, we have, for

$$\mathbf{F} = P(x,y)\mathbf{I} + Q(x,y)\mathbf{J} \ ,$$

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

$$\text{curl } \mathbf{F} = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\mathbf{K} \ .$$

Example 18.7. Find the divergence and curl of the velocity fields a) associated to a source (see example 18.2), and for rotation about a point (see example 18.3).

In example 18.2 we had $\mathbf{V} = a\mathbf{X} = a(x\mathbf{I} + y\mathbf{J})$. Then

$$\text{div } \mathbf{V} = 2a \ , \quad \text{curl } \mathbf{V} = 0 \ .$$

Note that in this case $\mathbf{V} = \nabla r^2/2$, so the field has the circles centered at the origin as equipotentials.

In example 18.3, $\mathbf{V} == \omega\mathbf{X}^{\perp} = \omega(-y\mathbf{I} + x\mathbf{J}))$, so that

$$\text{div } \mathbf{V} = 0 \ , \quad \text{curl } \mathbf{V} = 2\omega\mathbf{K} \ ,$$

and the vector field is not a gradient.

Problems 18.1

1. Suppose that a fluid is rotating about the z-axis with constant angular speed ω. Let $\mathbf{V}$ be the velocity field of the motion.

a) Show that $\mathbf{V} = \omega\mathbf{K} \times \mathbf{X}$.

b) Calculate div $\mathbf{V}$, curl $\mathbf{V}$.

c) Calculate $\int_C \mathbf{V} \cdot d\mathbf{X}$, where C is the helicoid given parametrically by $\mathbf{X}(t) = 2\cos t\mathbf{I} + 2\sin t\mathbf{J} + t\mathbf{K}$.

2. Let $\mathbf{F}(\mathbf{X})$ be the vector field

$$\cos x \cos y(cosz + 2)\mathbf{I} - (\sin x \sin y(2 + \cos z) + \cos y \cos z)\mathbf{J} + \sin z(\sin y - \sin x \cos y)\mathbf{K} \ .$$

If it exists, find $w = f(x,y,z)$ such that $\nabla w = \mathbf{F}$.

3. A *radial* field is a field of the form $\mathbf{F}(\mathbf{X}) = g(\rho)\mathbf{X}$, where $\rho = |\mathbf{X}|$. Show that a radial field has a potential function; that is, there is a function $w = G(\mathbf{X})$ such that $\mathbf{F} = \nabla G$.

4. Here are two vector fields:

$$(A) \qquad \mathbf{F}(x,y,z) = (2xy^2z + yz^2)\mathbf{I} + (2x^2yz + xz^2)\mathbf{J} + (x^2y^2 + 2xyz)\mathbf{K}$$

$$(B) \qquad \mathbf{G}(x,y,z) = (2xy^2z + yz^2)\mathbf{I} + (x^2z + xz)\mathbf{J} + (x^2y^2 + xy)\mathbf{K}$$

a) Which of the vector fields $\mathbf{F}$ and $\mathbf{G}$ has a chance of being a gradient? Why?

b) Pick one of your answers to a) and find the function f whose gradient is that field.

5. Describe the equipotentials of these vector fields $\mathbf{F}$:

$$a)\ y\mathbf{I} + x\mathbf{J}\ , \quad b)\ y\mathbf{I} - x\mathbf{J}\ , \quad c)\ x\mathbf{I} + 2y\mathbf{J} + z\mathbf{K}\ , \quad d)\ y\mathbf{I} + x\mathbf{J} - 2z\mathbf{K}\ .$$

6. Show that

$$a) \qquad \nabla \cdot (\nabla \times \mathbf{F}) = 0\ ,$$

$$b) \qquad \nabla \times \nabla f = 0\ .$$

18.2 Line Integrals and Work.

Suppose $\mathbf{F}$ is a vector field defined on a region R, and C is a curve lying in R. We define the line integral of $\mathbf{F}$ along C, by analogy with other integrals as follows.

Definition 18.5. Let $\mathbf{X}_i, 0 \le i \le n$ be a sequence of points on the curve, with $\mathbf{X}_0, \mathbf{X}_n$ the endpoints. Form the sum

$$\sum_{i=1}^{n} \mathbf{F}(\mathbf{X}_i) \cdot (\mathbf{X}_i - \mathbf{X}_{i-1})\ .$$

If the limit of this sum exists (as the maximum distance between successive points approaches zero), it is the **line integral** of $\mathbf{F}$ along C:

$$(18.17) \qquad \int_C \mathbf{F} \cdot d\mathbf{X} = \lim_{\max|\Delta\mathbf{X}_i| \to 0} \sum_{i=1}^{n} \mathbf{F}(\mathbf{X}_i) \cdot \Delta\mathbf{X}_i\ ,$$

where $\Delta\mathbf{X}_i$ represents the vector increment between successive points.

If we have a parametric representation of the curve: $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J} + z(t)\mathbf{K}$, for $a \le t \le b$, where the functions $x(t), y(t), z(t)$ are differentiable, then we can compute the line integral by integration with respect to t. For, as successive points become arbitrarily close, we can replace each $\Delta\mathbf{X}_i$ by its linear approximation, and in the limit, we obtain

$$\lim \sum_{i=1}^{n} \mathbf{F}(\mathbf{X}_i) \cdot \Delta\mathbf{X} = \lim_{\Delta t_i \to 0} \sum_{i=1}^{n} \mathbf{F}(\mathbf{X}(t_i)) \cdot \frac{d\mathbf{X}}{dt}(t_i)\Delta t_i = \int_a^b \mathbf{F} \cdot \frac{d\mathbf{X}}{dt} dt\ .$$

Proposition 18.2. If C is a curve parametrized by $\mathbf{X} = \mathbf{X}(t)$ for $a \leq t \leq b$, and $\mathbf{F}$ is a vector field defined on C, then

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_a^b \mathbf{F}(\mathbf{X}(t)) \cdot \frac{d\mathbf{X}}{dt} dt \ . \tag{18.18}$$

Example 18.8. Find $\int_C \mathbf{F} \cdot d\mathbf{X}$ where C is the curve $\mathbf{X}(t) = t^2\mathbf{I} + (t+1)\mathbf{J},\ 0 \leq t \leq 3$, and $\mathbf{F}(x,y) = x^2\mathbf{I} + xy\mathbf{J}$.

Here

$$\frac{d\mathbf{X}}{dt} = 2t\mathbf{I} + \mathbf{J}$$

and, along C,

$$\mathbf{F}(x,y) = x^2\mathbf{I} + xy\mathbf{J} = (t^2)^2\mathbf{I} + t^2(t+1)\mathbf{J}$$

so

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_0^3 \mathbf{F} \cdot \frac{d\mathbf{X}}{dt} dt = \int_0^3 ((t^2)^2(2t) + t^2(t+1))dt$$

$$= \int_0^3 (2t^5 + t^3 + t)dt = \Big(\frac{t^6}{3} + \frac{t^4}{4} + \frac{t^2}{2}\Big)_0^3 = 9(27 + \frac{9}{4} + \frac{1}{2}) = 267.75 \ .$$

To summarize, line integrals are computed this way. Let $\mathbf{F} = P\mathbf{I} + Q\mathbf{J} + R\mathbf{K}$ be a vector field in three dimensions, and suppose that C is given parametrically by the equation $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J} + z(t)\mathbf{K}$, for $a \leq t \leq b$, where the functions $x(t), y(t), z(t)$ are differentiable. Then

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_a^b \mathbf{F} \cdot \frac{d\mathbf{X}}{dt} dt = \int_a^b \big(P\frac{dx}{dt} + Q\frac{dy}{dt} + R\frac{dz}{dt}\big)dt \ . \tag{18.19}$$

If the curve is given as the graph $y = y(x),\ z = z(x)$, then we still use the same formula, thinking of the parameter as x and the trajectory given by $\mathbf{X}(x) = x\mathbf{I} + y(x)\mathbf{J} + z(x)\mathbf{K}$. Of course, as we have defined the line integral, it is independent of the parametrization of the curve, and depends only on the direction along the curve in which we integrate.

The line integral (18.19) may appear in several different forms. First, if we want to interpret the line integral as the integral of a differential (as in all cases of integration), we write (18.19) as

$$\int_a^b Pdx + Qdy + Rdz \ , \tag{18.20}$$

as the integral of the differential $Pdx+Qdy+Rdz$. To calculate the integral, we choose a convenient parametrization and calculate as in (18.19). it is also useful to refer to the parametrization by arc length. Since $d\mathbf{X}/ds = \mathbf{T}$ where $\mathbf{T}$ is the unit tangent to the curve. we can write $d\mathbf{X} = \mathbf{T}ds$ and the line integral is

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_C \mathbf{F} \cdot \mathbf{T}ds \ .$$

This expresses the line integral as the integral with respect to arc length of the component of the field in the direction of the curve. Finally we note that the integral is additive over curves.

Proposition 18.3. If the curve C can be written as a finite succession of curves $C_1, \ldots, C_n$ such that the initial point of each C_i is the same as the terminal point of its predecessor, then, for any vector fied $\mathbf{F}$ defined on C:

$$\int_C \mathbf{F}\cdot d\mathbf{X} = \int_{C_1} \mathbf{F}\cdot d\mathbf{X} + \cdots + \int_{C_n} \mathbf{F}\cdot d\mathbf{X}\ . \tag{18.21}$$

Example 18.9. Find $\int_C \mathbf{F}\cdot d\mathbf{X}$, where $\mathbf{F}(x,y) = xy\mathbf{I} + y^2\mathbf{J}$, and C is the triangle from (0,0) to (2,0) to (3,0) and back to (0,0).

C consists of three line segments:

$$C_1 : 0 \le x \le 2,\ y = 0 \quad C_2 : 0 \le y \le 3,\ x = 2 - \frac{2}{3}y \quad C_3 : 3 \ge y \ge 0,\ x = 0\ .$$

We calculate the three integrals separately, and then, by (21), take their sum. On C_1, we take x as the parameter, and $dy = 0$.

$$\int_{C_1} \mathbf{F}\cdot d\mathbf{X} = \int_{C_1} xydx + y^2dy = \int_0^2 0dx = 0\ .$$

On C_2 we take y as the parameter, and we have $dx = -(2/3)dy$.

$$\int_{C_2} \mathbf{F}\cdot d\mathbf{X} = \int_{C_2} xydx + y^2dy = \int_0^3 (2 - \frac{2}{3}y)(-\frac{2}{3})dy + y^2dy$$

$$= \int_0^3 (-\frac{4}{3} + \frac{4}{3}y + y^2)dy = (-\frac{4}{3}y + \frac{2}{3}y^2 + \frac{y^3}{3}|_0^3 = 11\ .$$

Finally, since $x = 0$ on C_3:

$$\int_{C_3} \mathbf{F}\cdot d\mathbf{X} = \int_{C_3} y^2dy = \int_3^0 y^2dy = -9\ ,$$

and

$$\int_C \mathbf{F}\cdot d\mathbf{X} = \int_{C_1} \mathbf{F}\cdot d\mathbf{X} + \int_{C_2} \mathbf{F}\cdot d\mathbf{X} + \int_{C_3} \mathbf{F}\cdot d\mathbf{X} = 0 + 11 - 9 = 2\ .$$

Example 18.10. Find $\int_C \mathbf{F}\cdot d\mathbf{X}$, where $\mathbf{F}(x,y) = y\mathbf{I} + x\mathbf{J}$, and C is the curve given parametrically as $x = 1 + 3\cos t,\ y = 3\sin(2t)$.

We first calculate the differentials $dx = -3\sin tdt,\ dy = 6\cos(2t)dt$, so

$$\mathbf{F}\cdot d\mathbf{X} = 3\sin(2t)(-3\sin tdt) + (1 + 3\cos t)(6\cos(2t)dt)$$

$$= (-9\sin(2t)\sin t + 6\cos(2t) + 18\cos(2t)\sin(t))dt\ .$$

Performing the integration, we get $\int_C \mathbf{F}\cdot d\mathbf{X} = -.00111$.

If $\mathbf{F}$ is a force field in the plane in space, then the work done in moving from one point $\mathbf{X}_0$ to another point $\mathbf{X}_1$ is $W = \mathbf{F}\cdot(\mathbf{X}_1 - \mathbf{X}_0)$, since the action of the force is only in the direction from $\mathbf{X}_0$ to $\mathbf{X}_1$. Now, if $\mathbf{X}(t)$ represents a curve C then the contribution to work along a small piece of the curve $d\mathbf{X}$ is $dW = F \cdot d\mathbf{X}$. We find the **total work** done by the force along the trajectory as the integral:

$$Work = \int_C F \cdot d\mathbf{X} \ . \tag{18.22}$$

Example 18.11. Let $\mathbf{F} = -z\mathbf{I} + x\mathbf{J} + \mathbf{K}$ be a force field in space. How much work is done by this force in moving an object from the origin to the point (1,1,1) along the path $C: \ y = x^2, \ z = x^3$?

First we express C parametrically by $\mathbf{X} = x\mathbf{I} + x^2\mathbf{J} + x^3\mathbf{K}, \ 0 \leq x \leq 1$, so that $d\mathbf{X}/dx = \mathbf{I} + 2x\mathbf{J} + 3x^2\mathbf{K}$. The force along C is, in terms of the parameter x: $\mathbf{F} = -x^3\mathbf{I} + x\mathbf{J} + \mathbf{K}$. Then, the work done by this force is

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_0^1 (-x^3 + 2x(x) + 3x^2)dx = \int_0^1 (5x^2 - x^3)dx = \frac{17}{12} \ .$$

Recall that the kinetic energy of a particle of mass m in motion is $(1/2)m|\mathbf{V}|^2$, where $\mathbf{V}$ is its velocity. If we differentiate this with respect to t, and use Newton's Second law $\mathbf{F} = m\mathbf{A}$, we have:

$$\frac{d}{dt}\big(\frac{1}{2}m\mathbf{V} \cdot \mathbf{V}\big) = m\mathbf{A} \cdot \mathbf{V} = \mathbf{F} \cdot \frac{d\mathbf{X}}{dt} \ . \tag{18.23}$$

This expresses the law of conservation of energy for a particle in motion in the presence of a force field: the change in the kinetic energy along the trajectory is equal to the work done to the particle. For suppose that the particle travels along the path C from time $t = a$ to $t = b$. We integrate (18.23) along the path, getting

$$\frac{m}{2}|\mathbf{V}(b)|^2 - \frac{m}{2}|\mathbf{V}(a)|^2 = \int_C \mathbf{F} \cdot d\mathbf{X} \ . \tag{18.24}$$

Example 18.12. A particle of mass 2 g. moves around the circle of radius 1 on the plane in the presence of a centripetal force field (keeping it on the circle) and of the force field $\mathbf{F}(x, y) = (1 + y)\mathbf{I} + y^2\mathbf{J}$ (where the magnitude is in dynes). Suppose that at time $t = 0$ the particle is at the point (1,0) travelling at a speed of 3 cm/sec. What is its speed the next time it passes through the point (1,0)?

We parametrize the path using polar coordinates $C: \ x = \cos\theta, \ y = \sin\theta$ for $0 \leq \theta \leq 2\pi$. In terms of this parametrization,

$$\mathbf{F} = (1 + \sin\theta)\mathbf{I} + (\sin^2\theta)\mathbf{J} \ , \quad \frac{d\mathbf{X}}{d\theta} = -\sin\theta\mathbf{I} + \cos\theta\mathbf{J} \ .$$

Since the centripetal force is orthogonal to $d\mathbf{X}/d\theta$, the work done in this motion is

$$\int_C \mathbf{F} \cdot d\mathbf{X} = \int_0^{2\pi} (-\sin\theta - \sin^2\theta + \sin^2\theta\cos\theta)d\theta = -\pi \ .$$

Since $m = 2$, letting b be the time the particle next passes through (1,0), (18.24) gives us

$$|\mathbf{V}(b)|^2 = -\pi + (3)^2 = 5.8584 \ ,$$

so $|\mathbf{V}(b)| = 2.420$ cm/sec.

Problems 18.2

1. Let C be the curve $x = y = z$ going from the point (1,1,1) to the point $(4, 4, 4)$. Find

$$\int_C \mathbf{F} \cdot d\mathbf{X}, \qquad \int_C \mathbf{G} \cdot d\mathbf{X}$$

for the vector fields given in (A) and (B) of problem 4 of section 18.1.

2. Consider the force field in three dimensions $\mathbf{F}(x, y, z) = y\mathbf{I} + z\mathbf{K}$. Let C be the curve given parametrically by $\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + t\mathbf{K}$. What is the work required to move a particle along C from $(1, 0, 0)$ to $(1, 0, 10\pi)$?

3. Consider the vector field defined in the firt quadrant by

$$\mathbf{F}(x, y) = (\frac{y}{x} + \ln y)\mathbf{I} + (\frac{x}{y} + \ln x)\mathbf{J} \ .$$

Find $\int_C \mathbf{F} \cdot d\mathbf{X}$ where C is the straight line from (1,1) to (e, e^2).

4. Calculate $\int_C \mathbf{V} \cdot d\mathbf{X}$, where C is the helicoid given parametrically by

$$\mathbf{X}(t) = 2\cos t\mathbf{I} + 2\sin t\mathbf{J} + t\mathbf{K},\ 0 \le t \le a \ ,$$

and $\mathbf{V} = -y\mathbf{I} + x\mathbf{J}$.

5. Consider the vector field in the plane described in example 18.6:

$$\mathbf{V} = \frac{-y}{x^2 + y^2}\mathbf{I} + \frac{x}{x^2 + y^2}\mathbf{J} \ .$$

a) Show that $\nabla \times \mathbf{V} = 0$.

b) Calculate $\int_C \mathbf{V} \cdot d\mathbf{X}$, where C is the unit circle.

18.3 Independence of Path

In this and the next section, we shall restrict attention to two dimensions. First, let us summarize the preceding sections. A vector field defined in a region D is of the form $\mathbf{F} = P\mathbf{I} + Q\mathbf{J}$ where P and Q are scalar functions on R. If C is a curve in R parametrized by $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J},\ a \le t \le b$, then

$$\int_C \mathbf{F} \cdot \mathbf{X} = \int \mathbf{F} \cdot \mathbf{T} ds = \int_C P dx + Q dy = \int_a^b (P\frac{dx}{dt} + Q\frac{dy}{dt}) dt \ .$$

This is the integral with respect to arc length of the component of $\mathbf{F}$ in the direction of the curve.

If $\mathbf{F}$ is a force field, this is the work done by the force along the curve C.

If $\mathbf{F}$ is interpreted as the velocity field of a flow, this is the total flow of fluid in the direction of the curve.

We might also be interested in the flow of the fluid across the curve; this is the integral of the component of $\mathbf{F}$ orthogonal to the curve; that is, $\int \mathbf{F}\cdot\mathbf{N}ds$ where $\mathbf{N}$ is the normal to the curve. Since there are two unit normals to the curve, we must specify the direction in which the curve is crossed. For this discussion we shall take the normal pointing to the right of the direction in which the curve is traversed. Since $\mathbf{T}ds = dx\mathbf{I} + dy\mathbf{J}$, we are taking $\mathbf{N}ds = dy\mathbf{I} - dx\mathbf{J}$, so that $\mathbf{F}\cdot\mathbf{N}ds = \det(\mathbf{F}, d\mathbf{X})$.

Definition 18.5, Let $\mathbf{F} = P\mathbf{I} + Q\mathbf{J}$ be a vector field defined in a region R, and C a curve in R. The **circulation** of $\mathbf{F}$ along C is

$$\int \mathbf{F}\cdot\mathbf{T}ds = \int_C \mathbf{F}\cdot d\mathbf{X} = \int_C Pdx + Qdy\ .$$

The **flux** of $\mathbf{F}$ across C from left to right is

$$\int \mathbf{F}\cdot\mathbf{N}ds = \int_C \det(\mathbf{F}, d\mathbf{X}) = \int_C -Qdx + Pdy\ .$$

Example 18.13 Calculate the circulation and flux of $\mathbf{F} = x^2\mathbf{I} - xy\mathbf{J}$ across the line from (0,0) to (3,4).

The line is easily parametrized by $x = 3t,\ y = 4t,\ 0 \le t \le 1$, so that $dx = 3dt,\ dy = 4dt$. Then

$$Circulation = \int_C x^2dx - xydy = \int_0^1 (3t)^2(3dt) - (3t)(4t)(4dt) = \int_0^1 (27-48)t^2dt = -7\ .$$

$$Flux = \int_C xydx + x^2dy = \int_0^1 (3t)(4t)(3dt) - (3t)^2(4dt) = \int_0^1 (36+36)t^2dt = 24\ .$$

Proposition 18.4. If the vector field $\mathbf{F}$ is the gradient of a function in R, then, for any path C,

$$\int_C \mathbf{F}\cdot d\mathbf{X} = f(\mathbf{X_1}) - f(\mathbf{X_0})$$

where $\mathbf{X_0}$ is the initial point of the path, and $\mathbf{X_1}$ is its endpoint.

To see this, let C have the parametrization $\mathbf{X}(t) = x(t)\mathbf{I} + y(t)\mathbf{J}$ for $a \le t \le b$, so that $\mathbf{X_0} = \mathbf{X}(a)$ and $\mathbf{X_1} = \mathbf{X}(b)$. We have

$$\mathbf{F} = \frac{\partial f}{\partial x}\mathbf{I} + \frac{\partial f}{\partial y}\mathbf{J}$$

so that

$$\int_C \mathbf{F}\cdot d\mathbf{X} = \int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\right)dt = \int_a^b \frac{d}{dt}f(\mathbf{X}(t))dt = f(\mathbf{X}(b)) - f(\mathbf{X}(a))\ ,$$

by the fundamental theorem of the Calculus.

Definition 18.6. A region D is called **connected** if, for any two points P and Q in D, there is a curve C with endpoints P and Q. A differential $Pdx + Qdy + Rdz$ is said to be **independent**

of path in D if the integral $\int_C Pdx + Qdy + Rdz$ is the same for all curves C with the same endpoints. A differential is said to be **exact** if it is the differential of a function; that is, there is a function f such that $df = Pdx + Qdy + Rdz$. A vector field $\mathbf{F}$ is called **conservative** if $\int_C \mathbf{F} \cdot d\mathbf{X}$ is independent of path.

Proposition 18.5. A differential form $Pdx + Qdy + Rdz$ defined on a connected region D is independent of path there if and only if it is exact. Equivalently, given a vector field $\mathbf{F}$, the line integral $\int_C \mathbf{F} \cdot d\mathbf{X}$ is independent of path if and only if $\mathbf{F} = \nabla f$ for some function f (called its *potential*).

The above proposition tells us that gradient fields are independent of path. Now, we must show that if the differential form $Pdx + Qdy$ is independent of path in D, then it is a gradient. Fix a point (x_0, y_0) in D, and define the function f by $f(x, y) = \int_C Pdx + Qdy$ where C is any path joining (x_0, y_0) to (x, y). To show that $\partial f/\partial x = P$, we take a point $(x + h, y)$ near (x, y), and consider the path C' which is C followed by the line segment L from (x, y) to $(x + h, y)$ (see the figure). Then

$$f(x+h,y) = \int_{C'} Pdx + Qdy = \int_C Pdx + Qdy + \int_L Pdx + Qdy = f(x,y) + \int_L Pdx + Qdy \ .$$

Now, we can parametrize L by $(x(t), y(t)) = (x + t, y),\ 0 \le t \le h$. Since $dy = 0$ along L, we have

$$\frac{f(x+h,y) - f(x,y)}{h} = \frac{1}{h}\int_0^h P(x+t,y)dx \ ,$$

which converges to $P(x, y)$. Similarly, $\partial f/\partial y = Q$.

Definition 18.7 A curve C is said to be **closed** if its endpoints are the same (under any parametrization). The integral over a closed curve is denoted $\oint_C$.

Proposition 18.5 can be restated this way: we have $\mathbf{F} = \nabla f$ if and only if the line integral $\oint_C \mathbf{F} \cdot d\mathbf{X}$ over every closed curve is zero.

Example 18.14. Let $\mathbf{F} = -y\mathbf{I} + x\mathbf{J}$ and C be the boundary of the ellipse $x = 2\cos t,\ y = \sin t,\ 0 \le t \le 2\pi$. Then

$$\oint_C \mathbf{F} \cdot d\mathbf{X} = \oint_C -ydx + xdy = \int_0^{2\pi} -\sin t(-2\sin t)dt + 2\cos t\cos tdt =$$

$$2\int_0^{2\pi} (\sin^2 t + \cos^2 t)dt = 4\pi \ .$$

Problems 18.3

1. A particle of mass 4 grams moves in the plane along a track subject to an external force $\mathbf{F} = 2xy\mathbf{I} + (x^2 + y)\mathbf{J}$. When the particle is at the origin its speed is 5 cm/sec. Some time later it arrives at the point (3,2).What is its velocity at that point?

2. Let $\mathbf{F} = yx\mathbf{I} + xz\mathbf{J} + xy\mathbf{K}$. Calculate $\int_C \mathbf{F} \cdot d\mathbf{X}$ where $C = C_1 + C_2 + C_3$ and

C_1 is the segment of the x-axis from (0,0,0) to (1,0,0);
C_2 is the segment of the line $x = 1,\ z = 0$ from (1,0,0) to (1,1,0);
C_3 is the segment of the line $x = 1,\ y = 1$ from (1,1,0) to (1,1,1).

3. Let $\mathbf{F} = -y\mathbf{I} + x\mathbf{J}$. Let C be the boundary of the unit square with vertices at (0,0) and (1,1). Find the circulation and flux of $\mathbf{F}$ around C traversed counterclockwise.

18.4 Green's Theorem in the Plane

Suppose that D is a region in the plane whose boundary is a curve, which we will always consider to be directed so that D always lies to the left of its boundary. We use the notation ∂D to represent the boundary of D so directed. To put it another way: for $\mathbf{T}$ and $\mathbf{N}$ the unit tangent and normal to C as defined in the preceding section, $\mathbf{N}$ is to the right of $\mathbf{T}$, so points out of D. For this reason $\mathbf{N}$ is called the exterior normal.The boundary of a domain is a closed curve (or several closed curves). From the discussion in the preceding section, we know that if $\mathbf{F}$ is a gradient field defined on D, then $\oint_{\partial D} \mathbf{F} \cdot d\mathbf{X} = 0$ and curl $\mathbf{F} = 0$. The connection between these two statements is much deeper and is embodied in Green's theorem which relates the line integral on ∂D with the double integral of curl $\mathbf{F}$ on the domain D. First we state the theorem in differential form.

Proposition 18.6 (Green's Theorem). Let D be a region, whose boundary ∂D is oriented so that D lies to the left of ∂D. Suppose that $Pdx + Qdy$ is a differential defined on the region D. Then

$$\oint_{\partial D} Pdx + Qdy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA \ . \tag{18.25}$$

Example 18.15. Let's redo example 18.14 using Green's theorem, where E represents the region bounded by the ellipse:

$$\oint_C \mathbf{F} \cdot d\mathbf{X} = \oint_C -ydx + xdy = \int\int_E (1+1)dxdy = 4\pi \ ,$$

since the area of E is 2π.

Example 18.16. Given the differential $x^2dx - xdy$, and D be the rectangle $1 \le x \le 3,\ 1 \le y \le 4$, we have

$$\oint_{\partial D} x^2dx - xdy = \int\int_D (-1+2x)dxdy = \int_1^3 \int_1^4 (-1+2x)dydx = 18 \ .$$

We now restate Green's theorem in two ways in vector form.

Proposition 18.7 (Stokes' Theorem in the Plane). Let D be a region with boundary ∂D. Let $\mathbf{F}$ be a vector field defined on D. Then

$$\oint_{\partial D} \mathbf{F} \cdot d\mathbf{X} = \oint_{\partial D} \mathbf{F} \cdot \mathbf{T} ds = \int\int_D \text{curl } \mathbf{F} \cdot \mathbf{K} dA \ . \tag{18.26}$$

This follows directly from (18.25), for if we write $\mathbf{F} = P\mathbf{I} + Q\mathbf{J}$ in component form, we have $\mathbf{F} \cdot d\mathbf{X} = Pdx + Qdy$ and curl $\mathbf{F} \cdot \mathbf{K} = \partial Q/\partial x - \partial P/\partial y$. In terms of fluid flows, this theorem state that the circlation of the fluid around the curve C can be obtained by integrating the curl over the region bounded by C. If we think of C as the boundary of a small disc around a point, this explains the definition of curl: its value is approximately the rate at which the fluid "curls" around the point.

Equally interesting is the rate at which fluid passes through the boundary, given by $\int_C \mathbf{F} \cdot \mathbf{N} ds$. Using the expression $\mathbf{N}ds = dy\mathbf{I} - dx\mathbf{J}$, and $\mathbf{F} = P\mathbf{I} + Q\mathbf{J}$, we have

Proposition 18.8 (Gauss' Divergence Theorem in the Plane). Let D be a region with boundary ∂D. Let $\mathbf{F}$ be a vector field defined on D. Then

$$\oint_{\partial D} \mathbf{F} \cdot \mathbf{N} ds = \oint_{\partial D} (-Q dx + P dy) = \int\int_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dA = \int\int_D \text{div } \mathbf{F} dA \ .$$

This is interpreted as saying (in terms of fluid flow) the rate of change of the amount of fluid inside the region D is equal to the flux of the fluid through the boundary.

Example 18.17. Let D be the disc of radius 1 centered at the point (0,1), and let C be its boundary oriented counter clockwise. Suppose $\mathbf{V} = -y\mathbf{I}$ is the velocity field of a flow in the upper half plane. Calculate the circulation along C and the flux through C.

First of all, we see that the the fluid is moving from right to left along the lines $y = \text{const}$ at speed proportional to the distance to the x-axis. Since fluid enters the disc from the right along any such line at the same speed as it leaves the disc, we should expect the flux to be zero. On the other hand, the fluid is moving to the left faster on the upper part of the circle (which is oriented to the left) than on the lower part of the circle, so we should expect a positive circulation. According to Stokes' theorem, the circulation is

$$\oint_C \mathbf{V} \cdot \mathbf{T} ds = \int\int_D \text{curl } \mathbf{V} \cdot \mathbf{K} dA \ .$$

Now, since curl $\mathbf{V} = \mathbf{K}$, this becomes simply

$$\oint_C \mathbf{V} \cdot \mathbf{T} ds = \int\int_D dA = \pi \ ,$$

the area of D. According to the Divergence theorem, the flux out of D is

$$\oint_C \mathbf{F} \cdot \mathbf{N} ds = \int\int_D \text{div } \mathbf{F} dA = 0 \ ,$$

since the divergence of $\mathbf{V}$ is zero.

As a verification of these theorems, we also compute the line integrals. For that we use this parametrization of C: $\mathbf{X}(t) = \cos t\mathbf{I} + (1 + \sin t)\mathbf{J}$. Then $d\mathbf{X} = (-\sin t\mathbf{I} + \cos t\mathbf{J})dt$, and since $\mathbf{V} = -y\mathbf{I} = -(1 + \sin t)\mathbf{I}$ along C, we have

$$\oint_C \mathbf{V} \cdot d\mathbf{X} = \int_0^{2\pi} (1 + \sin t)(\sin t) dt = \int_0^{2\pi} \frac{1}{2} dt = \pi$$

Now, to calculate the flux through C out of D, we have $\mathbf{N}ds = \cos t\mathbf{I} + \sin t\mathbf{J}$, and

$$\oint_C \mathbf{V} \cdot \mathbf{N} ds = \int_0^{2\pi} -(1 + \sin t)(\cos t) dt = 0 \ .$$

A simple application of Green's theorem leads to a way of calculating area by line integrals.

Proposition 18.8. Let D be a region in the plane. Then the area of D is given by any of these line integrals over its boundary, ∂D:

$$Area(D) \ = \oint_{\partial D} xdy = -\oint_{\partial D} ydx = \frac{1}{2}\oint_{\partial D} -ydx + xdy \ ,$$

for in each of these cases the form $\partial Q/\partial y - \partial P/\partial x = 1$.

Example 18.18. Find the area of the region R bounded by the curves $y = x^2$ and $y = 1$.

We do this using Green's theorem. The boundary of R is in two pieces: $C_1 : y = 1$, with x going from 1 to -1, and $C_2 : y = x^2, -1 \leq x \leq 1$. Since $dy = 0$ on C_1, we have

$$Area = \oint_{\partial R} xdy = \int_{C_2} xdy = \int_{-1}^{1} x(2xdx) = \frac{4}{3} \ .$$

Example 18.19. We can verify that the area of an ellipse with major radius a and minor radius b is πab by Green's theorem and this parametrization of the boundar of the ellipse:

$$x = a\cos t \ , \quad y = b\sin t \ , \quad 0 \leq t \leq 2\pi \ .$$

Then

$$Area = \frac{1}{2}\oint_{\partial E} -ydx + xdy = \frac{1}{2}\int_0^{2\pi} (-b\sin t)(-a\sin t)dt + (a\cos t)(b\cos t)dt =$$

$$\frac{1}{2}\int_0^{2\pi} ab(\sin^2 t + \cos^2 t)dt = \pi ab \ .$$

Problems 18.4

1. Let C be the boundary of the triangle with vertices at (0,0), (3,0), (4,5), oriented in the counterclockwise sense. Find $\int_C 3ydx + 6xdy$.

2. Find $\int_C xy^2dx + x^2ydy$ where C is the line from (2,3) to (5,1).

3. Let C be the boundary of the square $0 \leq x \leq \pi, \ 0 \leq x \leq \pi$, traversed in the counterclockwise sense. Find

$$\int_C \sin(x+y)dx + \cos(x+y)dy \ .$$

4. Let C be the boundary of the triangle with vertices (0,0), (1,0), (1,2), traversed in the counterclockwise sense. Find $\int_C x^2dx + xydy$.

5. The *cycloid* is the curve given parametrically in the plane by

$$x = t - \sin t \ , \quad y = 1 - \cos t \ , \quad t \geq 0 \ .$$

Find the area under one arch of the cycloid.

6. A function of two variables is *harmonic* if

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \; .$$

Let R be a regular region in the plane with boundary C, $\mathbf{N}$ the exterior normal to C. Show that

$$\oint_C \nabla f \cdot \mathbf{N} ds = 0$$

for a harmonic function f.

7. Let $\mathbf{F}(\mathbf{X}) = e^{xy}(\mathbf{I} + \mathbf{J})$, and let R be the rectangle $0 \leq x \leq 2, \; -1 \leq y \leq 1$. For C the boundary of R traversed counterclockwise, Calculate

$$\text{a)} \quad \oint_C \mathbf{F} \cdot \mathbf{T} ds \qquad \text{b)} \quad \oint_C \mathbf{F} \cdot \mathbf{N} ds \; ,$$

where $\mathbf{T}$ and $\mathbf{N}$ are the tangent and normal to the curve C.

8. Let $\mathbf{F} = x\mathbf{I} + xy^2\mathbf{J}$. Let C be the circle $x^2 + y^2 = 1$ traversed in the counterclockwise sense. Find

$$\int_C \mathbf{F} \cdot \mathbf{T} ds \; , \qquad \int_C \mathbf{F} \cdot \mathbf{N} ds$$

where s represents arclength along the circle, $\mathbf{T}$ is the unit tangent vector, and $\mathbf{N}$ is the unit external normal vector.

9. Let $\mathbf{F} = y\mathbf{I} + 2x\mathbf{J}$. Let C be the curve $y = x^2$ from $x = 0$ to $x = 2$. Find the flux of $\mathbf{F}$ across C from left to right, that is, for $\mathbf{N}$ the unit normal to the right along C, find

$$\int_C \mathbf{F} \cdot \mathbf{N} ds \; .$$

10. Let $\mathbf{F} = x^3\mathbf{I} + y^3\mathbf{J}$. Let C be the circle $x^2 + y^2 = 9$ traversed in the counterclockwise sense. Find

$$\int_C \mathbf{F} \cdot \mathbf{N} ds \; .$$

where s represents arclength along the circle, and $\mathbf{N}$ is the unit external normal vector.

18.5. Stokes' and Gauss' theorems in three dimensions

When we move from two to three variables, the two interpretations of Green's theorem become two quite different theorems. Stokes' theorem relates integration on a surface with an integral on its bounding curve, and Gauss' theorem relates integration over a region with an integral on its bounding surface. We shall state these theorems and illustrate their use through examples, but shall not attempt to give proofs.

Surface Integrals

Let $\mathbf{F}$ be the velocity field of a flow in three dimensions, and S a surface in the region of flow. We want to calculate the rate at which fluid is passing through the surface - this is called the **flux** of the flow through S. Take a small rectangle of area ΔS on the surface. In an in interval of time of length Δt, the fluid which passes through the sruface is very nearly that inside the parallelipiped whose base is the rectangle and whose side is the vector $\mathbf{V}\Delta t$. This volume is $\Delta V = (\mathbf{F}\cdot\mathbf{N})\Delta S\Delta t$, so

$$\frac{\Delta V}{\Delta t} = (\mathbf{F}\cdot\mathbf{N})\Delta S \ .$$

Now, if we sum these terms over a grid of rectangles on S, and take the limit as the grid becomes fine we get

Proposition 18.9. Let $\mathbf{F}$ be a vector field defined in a neighborhood of the surface S. Choose a normal $\mathbf{N}$ to S. The **flux** of $\mathbf{F}$ over S in the direction $\mathbf{N}$ is

$$Flux = \int\int_S (\mathbf{F}\cdot\mathbf{N})dS \ . \tag{18.27}$$

In order to calculate this, we assume that the surface S is given parametrically by $\mathbf{X} = \mathbf{X}(u,v)$, for (u,v) in a region R in u,v space. We have

$$\mathbf{N} = \frac{\mathbf{X}_u \times \mathbf{X}_v}{|\mathbf{X}_u \times \mathbf{X}_v|} \ , \quad dS = |\mathbf{X}_u \times \mathbf{X}_v| dudv \ ,$$

so

$$Flux = \int\int_S (\mathbf{F}\cdot\mathbf{N})dS = \int\int_R \mathbf{F}\cdot(\mathbf{X}_u \times \mathbf{X}_v)dudv \ . \tag{18.28}$$

Example 18.20 Let $\mathbf{F} = z^2\mathbf{I} + \mathbf{J} + x^2\mathbf{K}$, and H the upper hemisphere $x^2+y^2+z^2 = 1, \ z \geq 0$. Find the flux of $\mathbf{F}$ through H from the inside of the sphere.

We parametrize H using spherical coordinates:

$$H: \quad \mathbf{X}(\phi,\theta) = \cos\theta\sin\phi\mathbf{I} + \sin\theta\sin\phi\mathbf{J} + \cos\phi\mathbf{K}$$

for $0 \leq \phi \leq \pi/2, \ 0 \leq \theta \leq 2\pi$. Differentiating:

$$\mathbf{X}_\phi = \cos\theta\cos\phi\mathbf{I} + \sin\theta\cos\phi\mathbf{J} - \sin\phi\mathbf{K} \ ,$$

$$\mathbf{X}_\theta = -\sin\theta\sin\phi\mathbf{I} + \cos\theta\sin\phi\mathbf{J} \ .$$

Check that the direction through H from the interior of the sphere is that of $\mathbf{X}_\phi \times \mathbf{X}_\theta$. Thus we must compute

$$\mathbf{F}\cdot(\mathbf{X}_\phi \times \mathbf{X}_\theta) = \det\begin{pmatrix} \cos^2\phi & 1 & \cos^2\theta\sin^2\phi \\ \cos\theta\cos\phi & \sin\theta\cos\phi & -\sin\phi \\ -\sin\theta\sin\phi & \cos\theta\sin\phi & 0 \end{pmatrix}$$

$$= \cos^2\phi\sin^2\phi\cos\theta - \sin^2\sin\theta + \sin^3\phi\cos\phi\cos^2\theta \ .$$

To calculate the integral (18.28), we first integrate with respect to θ. The first two terms integrate to zero, and since $\int_0^{2\pi}\cos^2\theta d\theta = \pi$, we obtain

$$\int\int_S (\mathbf{F}\cdot\mathbf{N})dS = \pi\int_0^{\pi/2}\sin^3\phi\cos\phi d\phi = \frac{\pi}{4} \ .$$

Stokes' theorem

Now, suppose that $\mathbf{F}$ is a vector field defined on a surface S in three dimensions, and S is bounded by a curve, denoted ∂S. As in two dimensions, Stokes' theorem relates the circulation about ∂S with the integral of curl $\mathbf{F}$ on S. For this to work we have to be sure that the direction of integration on ∂S is consistent with the choice of normal to S.

Proposition 18.10 (Stokes' Theorem). Suppose that $\mathbf{F}$ is a vector field defined on the surface S with the boundary ∂S. Choose the direction of the tangent $\mathbf{T}$ to ∂S and the normal $\mathbf{N}$ to the surface so that the vector $\mathbf{N} \times \mathbf{T}$ points into the surface S. Then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{X} = \int\int_S \text{curl } \mathbf{F} \cdot \mathbf{N} dS \ . \tag{18.29}$$

Example 18.21. Let S be the part of the plane $z = 2x + 3y + z = 12$ which lies in the first quadrant. Let $\mathbf{F} = y\mathbf{I} + z\mathbf{J} + x\mathbf{K}$. Verify Stokes' theorem.

We want to calculate both sides of (18.29) and see that they agree. First, the surface integral. We write the surface parametrically as

$$\mathbf{X}(x, y) = x\mathbf{I} + y\mathbf{J} + (12 - 2x - 3y)\mathbf{K} \ ,$$

for (x, y) in the triangle T with vertices (0,0),(6,0), (0,4). We'll need the partial derivatives

$$\mathbf{X}_x = \mathbf{I} - 2\mathbf{K} \ , \quad \mathbf{X}_y = \mathbf{J} - 3\mathbf{K} \ .$$

Now, we calculate curl $\mathbf{F} = -\mathbf{I} - \mathbf{J} - \mathbf{K}$, so

$$\mathbf{F} \cdot (\mathbf{X}_u \times \mathbf{X}_v) = \det \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -3 \end{pmatrix} = -6 \ .$$

Then, using (18.28)

$$\int\int_S \text{curl } \mathbf{F} \cdot \mathbf{N} dS = -6 \int\int_T dxdy = -72 \ ,$$

since the area of T is 12.

Now, to calculate the boundary integral, we represent the boundary as composed of the three line segments

$$C_1 : 0 \leq x \leq 6, \quad z = 12 - 2x \quad y = 0 \ ,; \quad dz = -2dx \ , dy = 0 \ ,$$

$$C_2 : 0 \leq y \leq 4, \quad x = \frac{12 - 3y}{2} \ , \quad z = 0; \quad dx = -\frac{3}{2}dy, \ dz = 0 \ ,$$

$$C_3 : 0 \leq z \leq 12, \quad y = \frac{12 - z}{3} \ , \quad x = 0; \quad dy = -\frac{dz}{3}, \ dx = 0 \ .$$

Then, recalling that $\mathbf{F} = y\mathbf{I} + z\mathbf{J} + x\mathbf{K}$:

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{X} = \int_0^6 x(-2dx) = -36 \ ,$$

$$\int_{C_2} \mathbf{F}\cdot d\mathbf{X} = \int_0^4 y(-\frac{3}{2}dy) = -12 \ ,$$

$$\int_{C_3} \mathbf{F}\cdot d\mathbf{X} = \int_0^{12} z(-\frac{dz}{3}) = -24 \ .$$

The sum of these is -72, so Stokes' theorem is verified.

Example18.22. Calculate

$$\int_C -ydx + xdy + dz$$

where C is the curve of intersection of the hyperboloid $z = x^2 - y^2$ and the cylinder $x^2 + y^2 = 1$.

Let $\mathbf{F} = -y\mathbf{I} + x\mathbf{J} + \mathbf{K}$. Then this can be viewed as the integral $\mathbf{F}\cdot d\mathbf{X}$ over the boundary of the piece H of the hyperboloid lying over the disc of radius 1 in the x, y-plane. We calculate that curl $\mathbf{F} = 2\mathbf{K}$, so the integral is, by Stokes' Theorem

$$\int\int_H 2\mathbf{K}\cdot\mathbf{N}dS \ .$$

Now, we can parametrize H by $X(x,y) = x\mathbf{I}+y\mathbf{J}+(x^2-y^2)\mathbf{K}$, with $\mathbf{X}_x = \mathbf{I}+2x\mathbf{K},\ \mathbf{X}_y = \mathbf{J}-2y\mathbf{K}$, so that

$$\int\int_H 2\mathbf{K}\cdot\mathbf{N}dS = \int\int_{x^2+y^2\le 1} 2\mathbf{K}\cdot(\mathbf{I}+2x\mathbf{K})\times\mathbf{J}-2y\mathbf{K})dxdy = \int\int_{x^2+y^2} 2dxdy = 2\pi \ ,$$

since the area of the disc of radius 1 is π.

If we parametrize the curve by $\mathbf{X}(t) = \cos t\mathbf{I} + \sin t\mathbf{J} + (\cos^2 t - \sin^2 t)\mathbf{K},\ 0 \le t \le 2\pi$ and calculate directly, we again get 2π.

Gauss' theorem

Now, suppose that R is a region in three dimensions, and the boundary of R is a surface which we shall denote as ∂R. If we have a fluid in flow, just as in 2 dimensions we expect Gauss' theorem to hold: the calculation of the rate of expansion of the fluid in R, which is the integral of the divergence, is the same as the flux through ∂R.

Proposition 18.11 Gauss' theorem. Let $\mathbf{F}$ be a vector field defined on the region R. We denote the boundary of R as ∂R, and take the normal to be the exterior normal $\mathbf{N}$. Then

(18.30)
$$\int\int_{\partial R} \mathbf{F}\cdot\mathbf{N}dS = \int\int\int_R \text{div } \mathbf{F}dV \ .$$

Example 18.23. Let R be the region inside the cone $z^2 = x^2 + y^2$, bounded by the planes $z = 0$ and $z = 2$. Let $\mathbf{F} = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$. Verify the divergence theorem in this context.

We easily calculate div $\mathbf{F} = 3$, so the right hand side of (28) is 3 times the volume of the cone, so

$$\int\int\int_R \text{div } \mathbf{F}dV = 3(Volume(R) = 3\frac{\pi r^2 h}{3} = 8\pi \ ,$$

since $r = 2,\ h = 2$.

To calculate the boundary integral, we turn to cylindrical coordinates, because of the symmetry around the z-axis. The boundary has two pieces: the disc $D:\ z = 1, r \leq 1$, and the surface of the cone $S:\ z = r \leq 1$. We can see that the integral over S is zero, since the vector field $\mathbf{F}$ is tangent to the cone (it is the tangent vector to the line $z = r, \theta = \theta_0$ which lies on the cone). Thus we need only calculate the boundary integral over D. Since D lies on the plane $z = 2$, its normal is $\mathbf{K}$. Thus since $\mathbf{F} \cdot \mathbf{K} = z = 2$ on the plane $z = 2$,

$$\int\int_{\partial R} \mathbf{F} \cdot \mathbf{N} dS = \int_0^{2\pi} \int_0^2 2r dr d\theta = 4\pi \frac{r^2}{2}\Big|_0^2 = 8\pi \ .$$

One of the main points of the divergence theorem is that informed use of the geometry involved simplifies what could otherwise be a complicated calculation. For example, if we did not observe that $\mathbf{F}$ is orthogonal to the normal to the cone, we'd have to do the calculation. Just to illustrate the methods we do it. First of all, we parametrize the cone using cylindrical coordinates:

$$S \ : \quad \mathbf{X} = r\cos\theta \mathbf{I} + r\sin\theta \mathbf{J} + r\mathbf{K} \ , 0 \leq \theta \leq 2\pi,\ r \leq 2$$

and, differentiating, we find

$$\mathbf{X}_r = \cos\theta \mathbf{I} + \sin\theta \mathbf{J} + \mathbf{K} \ , \quad \mathbf{X}_\theta = -r\sin\theta \mathbf{I} + r\cos\theta \mathbf{J} \ .$$

On the surface, in these coordinates $\mathbf{F} = r\cos\theta \mathbf{I} + r\sin\theta \mathbf{J} + r\mathbf{K}$. Now we calculate $\det(\mathbf{F}, \mathbf{X}_r, \mathbf{X}_\theta) = 0$, or we observe that since $\mathbf{F} = r\mathbf{X}_r$, the determinant must be zero.

Example 18.24. Return to example 20, and note that the divergence of that vector field is 0. By applying the divergence theorem, where R is the region bounded by H and the x, y-plane we can replace the integration of example 20 by the easier integration over the planar part of the boundary of H. That surface is the disc $D:\ x^2 + y^2 \leq 1, z = 0$. The normal (pointing outside of the region R) is $-\mathbf{K}$ and on this disc, $\mathbf{F} = \mathbf{J} + x^2\mathbf{K}$. Thus

$$\int\int_D \mathbf{F} \cdot \mathbf{N} dS = -\int\int_D x^2 dA = -\int_0^{2\pi} \int_0^1 r^2 \cos^2\theta r dr d\theta = -\frac{\pi}{4} \ .$$

Now, for this example, the divergence theorem tells us that

$$\int\int_H \mathbf{F} \cdot \mathbf{N} dS + \int\int_D \mathbf{F} \cdot \mathbf{N} dS = 0 \ ,$$

which gives the result $\int\int_H \mathbf{F} \cdot \mathbf{N} dS = \pi/4$.

Problems 18.5

1. A fluid has density 3 and velocity field $\mathbf{V} = 4x\mathbf{I} + 3z\mathbf{J} - z\mathbf{K}$. Find the flux of the fluid out of the ball centered at the origin and of radius 4 through its boundary.

2. Let P be the parabolic cup $z = x^2 + y^2$ lying over the unit disc in the xy-plane. Let $\mathbf{F}(x, y, z) = y\mathbf{I} - x\mathbf{J} + \mathbf{K}$. Calculate

$$\int\int_P \text{curl } \mathbf{F} \cdot \mathbf{N} dS \ .$$

3. Evaluate $\int\int_S \mathbf{F}\cdot\mathbf{N}dS$, where $\mathbf{F}(x,y,z) = x\mathbf{I} + y\mathbf{J} + z\mathbf{K}$ and S is the part of the paraboloid $z = 4 - x^2 - y^2$ which lies above the xy-plane.

4. Evaluate $\int\int_S \sqrt{1+x^2+y^2}dS$ where S is the surface given parametrically by

$$\mathbf{X}(s,t) = s\cos t\mathbf{I} + s\sin t\mathbf{J} + t\mathbf{K}\ , \quad 0 \le s \le 5, 0 \le t \le \pi/2\ .$$

5. Let S be the part of the plane $2x+y+3z = 12$ which lies in the first octant, oriented upward. Let the boundary ∂S of S be oriented so that S is to its left. Given the vector field $\mathbf{F} = 3x\mathbf{I} + \mathbf{J} + y\mathbf{K}$, find $\int_{\partial S} \mathbf{F}\cdot d\mathbf{X}$.

6. Let B^+ be the half-ball $B : x^2+y^2+z^2 \le 1,\ z \ge 0$. Let $\mathbf{F}(x,y,z) = x\mathbf{I} + y\mathbf{J} + \mathbf{K}$. Let H be the hemisphere bounding B^+ above: $H : x^2+y^2+z^2 = 1,\ z \ge 0$. Calculate the flux of $\mathbf{F}$ from B^+ across H.

7. Let $\mathbf{F} = x^2\mathbf{I} + y^2\mathbf{J} + z^2\mathbf{K}$. Calculate the flux of $\mathbf{F}$ out of the sphere S of radius 3.

8. Let P be the piece of the plane $2x + y + 3z = 12$ which lies in the first octant, and let $\mathbf{F} = 3x\mathbf{I} + \mathbf{J} + y\mathbf{K}$. Calculate the flux of $\mathbf{F}$ through P from below.