1. Vectors and the dot product

A vector $\vec{v}$ in $\mathbb{R}^3$ is an arrow. It has a direction and a length (aka the magnitude), but the position is not important. Given a coordinate axis, where the x-axis points out of the board, a little towards the left, the y-axis points to the right and the z-axis points upwards, there are three standard vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$, which have unit length and point in the direction of the x-axis, the y-axis and z-axis. Any vector in $\mathbb{R}^3$ may be written uniquely as a combination of these three vectors. For example, the vector $\vec{v} = 3\hat{\imath} - 2\hat{\jmath} + 4\hat{k}$ represents the vector obtained by moving 3 units along the x-axis, two units backwards along the y-axis and four units upwards.

If we imagine moving the vector so it's tail is at the origin then the endpoint P determines the vector. The point $P = (x, y, z)$ determines the vector $\vec{P} = \langle x, y, z\rangle$ starting at the origin and ending at the point P. Obviously,

$$\langle x, y, z\rangle = x\hat{\imath} + y\hat{\jmath} + z\hat{k} \qquad \text{so that} \qquad \langle 3, -2, 4\rangle = 3\hat{\imath} - 2\hat{\jmath} + 4\hat{k}.$$

One advantage of this algebraic approach is that we can write down vectors in $\mathbb{R}^4$, for example, $\langle 2, 1, -3, 5\rangle$, $\langle \pi, \sin 2, -3, e^3\rangle$.

Question 1.1. *What is the direction of the zero vector which starts and ends at the origin?*

We will adopt the convention that the zero vector points in every direction. In coordinates the zero vector in $\mathbb{R}^3$ is given by $\langle 0, 0, 0\rangle$.

The length of the vector $\vec{v} = \langle a, b, c\rangle$ is the scalar

$$|\vec{v}| = (a^2 + b^2 + c^2)^{1/2}.$$

This is what you get if you apply Pythagoras' Theorem, twice.

One can add vectors in $\mathbb{R}^3$. If you want to add $\vec{u}$ and $\vec{v}$, move the starting point of $\vec{v}$ to the endpoint of $\vec{u}$; the sum is the arrow you get by first going along $\vec{u}$ and then along $\vec{v}$. To subtract two vectors is even easier. The vector $\vec{v} - \vec{u}$ is the vector starting at the endpoint of $\vec{u}$ and ending at the endpoint of $\vec{v}$.

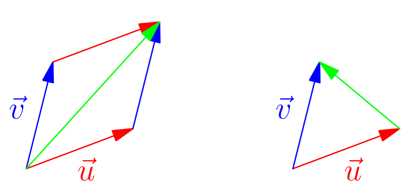


Figure 1. Addition and subtraction of vectors

Algebraically, it is easy to add vectors, add them component by component:

$$\begin{aligned}(3\hat{\imath} - 2\hat{\jmath} + 5\hat{k}) + (-4\hat{\imath} + 4\hat{\jmath} + 3\hat{k}) &= (3-4)\hat{\imath} + (-2+4)\hat{\jmath} + (5+3)\hat{k} \\ &= -\hat{\imath} + 2\hat{\jmath} + 8\hat{k}.\end{aligned}$$

More compactly,

$$\langle 3, -2, 5\rangle + \langle -4, 4, 3\rangle = \langle -1, 2, 8\rangle.$$

Note that it doesn't make sense to add a vector in $\mathbb{R}^2$ and a vector in $\mathbb{R}^3$. You can see this either algebraically or geometrically.

One can also multiply a scalar λ by a vector $\vec{v}$. $\lambda\vec{v}$ is the vector which is λ times as long as $\vec{v}$. If $\lambda > 0$, $\lambda\vec{v}$ has the same direction as $\vec{v}$ and if $\lambda < 0$, then $\lambda\vec{v}$ has the opposite direction. Either way, we will say that $\lambda\vec{v}$ is parallel to $\vec{v}$.

Algebraically, it is again easy to multiply a scalar by a vector,

$$\lambda\langle a, b, c\rangle = \langle \lambda a, \lambda b, \lambda c\rangle \qquad \text{so that} \qquad -3\langle 1, 2, -3\rangle = \langle -3, -6, 9\rangle.$$

The direction is what is left after you remove the length,

$$\hat{u} = \frac{\vec{v}}{|\vec{v}|}.$$

Note that $\hat{u}$ is a unit vector; it's length is one. Vectors will always have arrows on top of them, unit vectors hats.

We can always write a vector as a product of its length times its direction,

$$\vec{v} = |\vec{v}|\left(\frac{\vec{v}}{|\vec{v}|}\right).$$

For example,

$$\langle 1, -2, 2\rangle = 3\langle 1/3, -2/3, 2/3\rangle \qquad \text{and} \qquad \langle 3, 4\rangle = 5\langle 3/5, 4/5\rangle.$$

Question 1.2. *Let M be the midpoint of the line segment AB. Find the vector $\vec{M}$ in terms of the vectors $\vec{A}$ and $\vec{B}$.*

To get to M, from A, one has to go half way from A to B. The vector from A to B is $\overrightarrow{AB} = \vec{B} - \vec{A}$. Halfway means

$$\frac{1}{2}(\vec{B} - \vec{A}),$$

and so this is the vector from A to M. Therefore

$$\vec{M} = \vec{A} + \overrightarrow{AM} = \vec{A} + \frac{1}{2}(\vec{B} - \vec{A}) = \frac{1}{2}(\vec{A} + \vec{B}).$$

Question 1.3. *Show that the diagonals of a parallelogram bisect each other.*

Let's give names to the usual suspects. Let's call the vertices of the parallelogram A, B, C and D. Let X and Y be the midpoints of the diagonals. It is enough to show that $X = Y$ (naming the midpoints of the diagonals is the sneakiest part of the solution to this problem). What do we know? Well, since we have a parallelogram,

$$\begin{aligned} \overrightarrow{AB} &= \overrightarrow{CD} \\ \vec{B} - \vec{A} &= \vec{D} - \vec{C} \\ \vec{B} + \vec{C} &= \vec{A} + \vec{D}. \end{aligned}$$

(assuming we have labelled the vertices appropriately). We have

$$\vec{X} = \frac{1}{2}(\vec{A} + \vec{D}) \qquad \text{and} \qquad \vec{Y} = \frac{1}{2}(\vec{B} + \vec{C}).$$

So

$$\begin{aligned} \vec{Y} &= \frac{1}{2}(\vec{B} + \vec{C}) \\ &= \frac{1}{2}(\vec{A} + \vec{D}) \\ &= \vec{X}. \end{aligned}$$

Since the vectors $\vec{X}$ and $\vec{Y}$ both start at the origin we must have $X = Y$, which is what we want.

Question 1.4. *How do we multiply two vectors?*

Actually there are two answers to this question. The first answer is to take the dot product. If the vectors are

$$\vec{v}_1 = \langle a_1, b_1, c_1 \rangle \qquad \text{and} \qquad \vec{v}_2 = \langle a_2, b_2, c_2 \rangle,$$

then the dot product is the scalar

$$\vec{v}_1 \cdot \vec{v}_2 = a_1 a_2 + b_1 b_2 + c_1 c_2.$$

For example,

$$\langle 1, -2, 4 \rangle \cdot \langle 3, 1, -2 \rangle = 1 \cdot 3 + 1 \cdot -2 + 4 \cdot -2 = 3 - 2 - 8 = -7.$$

Note that

$$\vec{v} \cdot \vec{v} = |v|^2.$$

The usual rules of algebra apply to the dot product:

(1) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.
(2) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$.
(3) $(\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$.

Theorem 1.5 (Geometric interpretation of the dot product)**.** *If θ is the angle between the two vectors $\vec{u}$ and $\vec{v}$, then*

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos \theta.$$

Proof. If either $\vec{u}$ or $\vec{v}$ is the zero vector, then both sides are zero, and we certainly have equality (and we can take θ to be any angle we please, which is consistent with our convention that the zero vector points in every direction). So we may assume that $\vec{u}$ and $\vec{v}$ are both non-zero. If $\vec{u}$ and $\vec{v}$ are parallel, then $\theta = 0$ or π and it is straightforward to check that both sides are equal.

Otherwise, let $\vec{w} = \vec{v} - \vec{u}$, the third side of the triangle with sides given by $\vec{u}$ and $\vec{v}$. Then the square of the length of the third side is

$$\begin{aligned} |\vec{w}|^2 &= \vec{w} \cdot \vec{w} \\ &= (\vec{v} - \vec{u}) \cdot (\vec{v} - \vec{u}) \\ &= |\vec{v}|^2 + |\vec{u}|^2 - 2\vec{u} \cdot \vec{v}. \end{aligned}$$

Compare this with the formula given by the cosine rule. If the lengths of the three sides are are u, v and w, the cosine rule says,

$$w^2 = u^2 + v^2 - 2uv \cos \theta.$$

Now $u = |\vec{u}|$, $v = |\vec{v}|$ and $w = |\vec{w}|$, so putting these two formulae side by side, we see

$$\begin{aligned} w^2 &= u^2 + v^2 - 2\vec{u} \cdot \vec{v} \\ w^2 &= u^2 + v^2 - 2uv \cos \theta, \end{aligned}$$

so that subtracting we get

$$0 = 2(\vec{u} \cdot \vec{v} - uv \cos \theta),$$

whence the result. □

The virtue of (1.5) is that we can use it to find the angle between two vectors.

Question 1.6. *Consider the triangle in space with vertices $A = (1, 0, 0)$, $B = (1, 1, -1)$ and $C = (-1, 1, 0)$.*

What is the angle at A?

Let $\vec{u} = \overrightarrow{AB} = \langle 0, 1, -1 \rangle$ and $\vec{v} = \overrightarrow{AC} = \langle -2, 1, 0 \rangle$. We want the angle θ between $\vec{u}$ and $\vec{v}$. Well,

$$\begin{aligned}
\cos\theta &= \frac{\vec{u} \cdot \vec{v}}{|\vec{u}||\vec{v}|} \\
&= \frac{\langle 0, 1, -1 \rangle \cdot \langle -2, 1, 0 \rangle}{|\langle 0, 1, -1 \rangle||\langle -2, 1, 0 \rangle|} \\
&= \frac{1}{\sqrt{2} \cdot \sqrt{5}} \\
&= \frac{1}{\sqrt{10}}.
\end{aligned}$$

In this case

$$\theta = \cos^{-1}(\frac{1}{\sqrt{10}}) \approx 1.25$$

radians, which in degrees is about 71.57.

What can we say about the sign of

$$\frac{\vec{u} \cdot \vec{v}}{|\vec{u}||\vec{v}|}?$$

If this is positive we have an angle less than $\pi/2$ and if this is negative an angle greater than $\pi/2$. It is zero if and only if the angle is $\pi/2$ and the vectors are orthogonal.

Question 1.7. *Fix a point Q and a vector $\vec{n}$. What is the set of points $P = (x, y, z)$ such that the vector $\overrightarrow{PQ}$ is orthogonal to a vector $\vec{n}$?*

This is a plane.

Question 1.8. *What is the set of points where $2x - y + 3z = 0$?*

One way to answer this question is to guess using an analogy. If we were to drop a variable, we'd get $2x - y = 0$, which represents a line through the origin in $\mathbb{R}^2$. A reasonable guess is that this represents a plane through the origin.

Suppose we put $\vec{n} = \langle 2, -1, 3 \rangle$. Let $P = (x, y, z)$ so that $\vec{P} = \langle x, y, z \rangle$. Then

$$\vec{P} \cdot \vec{n} = \langle x, y, z \rangle \cdot \langle 2, -1, 3 \rangle = 2x - y + 3z.$$

The condition that this is zero, represents the condition that the vector $\vec{P}$ is orthogonal to $\vec{n}$. So this represents the plane through the origin orthogonal to the vector $\langle 2, -1, 3 \rangle$.

Let $\vec{F}$ represent a force. Notice that this makes sense; forces have a direction and a magnitude.

Question 1.9. *What is the component of the force $\vec{F}$ in the direction $\hat{u}$ (this is a direction, so $\hat{u}$ is a unit vector)?*

This is a scalar, a number. If one draws a triangle, with hypotenuse given by $\vec{F}$ and one side parallel to $\hat{u}$ and θ is the angle between $\vec{F}$ and $\hat{u}$ we want the length of the adjacent side. By the usual rules for trigonometry this is the length of the hypotenuse times the cosine of the angle θ, that is $|\vec{F}|\cos\theta$. But $\hat{u}$ has length one, so that $|\hat{u}| = 1$. So the component of $\vec{F}$ in the direction $\hat{u}$ is the dot product $\vec{F} \cdot \hat{u}$.

Even if $\vec{F}$ is not a force, one can always take the component of $\vec{F}$ in the direction of $\vec{u}$.

2. Cross products

Let's suppose you want to calculate the area of a polygon in the plane. Nothing easier, break the polygon into triangles and calculate the area of each triangle.

Question 2.1. *What is the area of a triangle, with two sides determined by the vectors $\vec{u}$ and $\vec{v}$?*

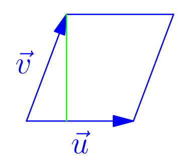


Figure 1. Area of parallelogram

The area of a triangle is 1/2 base $\times$ height, which is half of the area of the corresponding parallelogram, where $|\vec{u}|$ is the base and the length of the green line is the height, $|\vec{v}|\sin\theta$. The area of the parallelogram is $|\vec{u}||\vec{v}|\sin\theta$. How to get our hands on $\sin\theta$? One could use the identity,

$$\cos^2\theta + \sin^2\theta = 1,$$

to get a formula for $\sin\theta$ in terms of $\cos\theta$ and use the dot product, but it is pretty clear that it is going to give an ugly formula.

We prefer cosines to sines, since cosines turn up in dot products. If we have the complementary angle

$$\phi = \frac{\pi}{2} - \theta,$$

then we could use the fact that

$$\sin\theta = \cos(\frac{\pi}{2} - \theta) = \cos\phi.$$

In essence we want to rotate the vector $\vec{u}$ through $\pi/2$ radians counterclockwise, Suppose the vector we get this way is $\vec{u}'$.

Note that the angle between $\vec{u}'$ and $\vec{v}$ is ϕ, which is what we want. On the other hand, $\vec{u}$ and $\vec{u}'$ have the same length. It follows that the area of the parallelogram is

$$|\vec{u}||\vec{v}|\sin\theta = |\vec{u}'||\vec{v}|\cos\phi = \vec{u}'\cdot\vec{v}.$$

Question 2.2. *If $\vec{u} = \langle a_1, a_2\rangle$, then what is the vector $\vec{u}'$?*

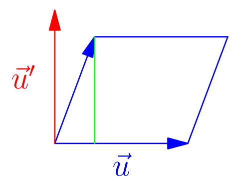


FIGURE 2. Rotated vector

Well by trial and error $\vec{u}' = \langle -a_2, a_1\rangle$ is at right angles to $\vec{u}$ (the dot product is zero). The vector $\hat{\imath} = \langle 1, 0\rangle$ gets sent to $\hat{\jmath} = \langle 0, 1\rangle$, which is the right orientation (counterclockwise versus clockwise).

So the answer is $\vec{u}' = \langle -a_2, a_1\rangle$.

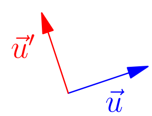


FIGURE 3. $\vec{u}$ and $\vec{u}'$

If $\vec{v} = \langle b_1, b_2\rangle$, then putting all of this together, the formula for the area of the parallelogram is simply

$$a_1b_2 - a_2b_1.$$

Let

$$A = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$$

be a 2×2 matrix. The determinant of A is

$$\det A = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

So the determinant of the matrix A is the $\pm$ the area of the parallelogram determined by the first and second row, $\vec{u} = \langle a_1, a_2\rangle$ and $\vec{v} = \langle b_1, b_2\rangle$. The sign depends on whether or not $\vec{u}$ comes before or after $\vec{v}$ (clockwise versus anticlockwise).

Now suppose we are given three vectors in $\mathbb{R}^3$,

$$\vec{u} = \langle a_1, a_2, a_3\rangle \qquad \vec{v} = \langle b_1, b_2, b_3\rangle \qquad \text{and} \qquad \vec{w} = \langle c_1, c_2, c_3\rangle.$$

Put them into a 3×3 matrix, whose rows are the three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$,

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}.$$

The determinant of A is

$$\det A = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

It is $\pm$ the volume of the parallelepiped determined by the three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$. The sign of the determinant is determined by whether or not the three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ form a right handed set.

Rule 2.3 (Right hand rule)**.** *The three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ form a right handed set if when you point your right hand in the direction of $\vec{u}$, curl your fingers in the direction of $\vec{v}$ then your thumb points in the direction of $\vec{w}$.*

The cross product of two vectors $\vec{v}$ and $\vec{w}$ is the vector given by the formula

$$\vec{v} \times \vec{w} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \hat{\imath} \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - \hat{\jmath} \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + \hat{k} \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

Geometrically, $\vec{v} \times \vec{w}$ is the vector whose length is the area of the parallelogram with sides $\vec{v}$ and $\vec{w}$ and whose direction is orthogonal to the plane spanned by $\vec{v}$ and $\vec{w}$, such that $\vec{v}$, $\vec{w}$ and $\vec{v} \times \vec{w}$ form a right handed set.

Question 2.4. *What is $\hat{\imath} \times \hat{\jmath}$?*

Here are the algebraic rules to manipulate the cross product:

(1) $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$.
(2) $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$.
(3) $(\lambda\vec{u}) \times \vec{v} = \lambda(\vec{u} \times \vec{v})$.

Note that $\vec{w} \times \vec{v} = -\vec{v} \times \vec{w}$. In particular $\vec{v} \times \vec{v} = \vec{0}$. One of the most useful features of the cross product is that the cross product $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$.

Question 2.5. *What is the equation of the plane containing $P_1 = (1, 1, 1)$, $P_2 = (1, 2, 3)$ and $P_3 = (-1, -2, 1)$?*

I will give you two methods to answer this question, the first using determinants the second using the cross product.

Two vectors in the plane are

$$\vec{v} = \overrightarrow{P_1P_2} = \langle 0, 1, 2 \rangle \qquad \text{and} \qquad \vec{w} = \overrightarrow{P_1P_3} = \langle -2, -3, 0 \rangle.$$

Let $P = (x, y, z)$ be a general point of space. P belongs to the plane if and only if the vector

$$\overrightarrow{P_1P} = \langle x - 1, y - 1, z - 1 \rangle,$$

lies in the plane. But this is the case if and only if the volume of the parallelepiped spanned by $\vec{v}$, $\vec{w}$ and $\langle x-1, y-1, z-1\rangle$ is zero, that is when

$$\begin{vmatrix} x-1 & y-1 & z-1 \\ 0 & 1 & 2 \\ -2 & -3 & 0 \end{vmatrix} = 0.$$

If we expand the determinant, we get

$$(x-1)\begin{vmatrix} 1 & 2 \\ -3 & 0 \end{vmatrix} - (y-1)\begin{vmatrix} 0 & 2 \\ -2 & 0 \end{vmatrix} + (z-1)\begin{vmatrix} 0 & 1 \\ -2 & -3 \end{vmatrix} = 6(x-1) - 4(y-1) + 2(z-1).$$

So

$$3(x-1) - 2(y-1) + (z-1) = 0,$$

and expanding we get

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

Here is the second method. The plane is specified by fixing one point P_1 in the plane and requiring that every vector in the plane with tail P_1 is orthogonal to a fixed vector $\vec{n}$, a normal vector to the plane.

$\vec{n}$ is orthogonal to $\vec{v}$ and $\vec{w}$, so we could take $\vec{n}$ to be the cross product.

$$\vec{v}\times\vec{w} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 0 & 1 & 2 \\ -2 & -3 & 0 \end{vmatrix} = \hat{\imath}\begin{vmatrix} 1 & 2 \\ -3 & 0 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} 0 & 2 \\ -2 & 0 \end{vmatrix} + \hat{k}\begin{vmatrix} 0 & 1 \\ -2 & -3 \end{vmatrix} = 6\hat{\imath} - 4\hat{\jmath} + 2\hat{k}.$$

Therefore $\vec{n} = 3\hat{\imath} - 2\hat{\jmath} + \hat{k}$ is a normal vector to the plane. P lies in the plane if and only if $\overrightarrow{P_1P}$ is orthogonal to $\vec{n}$, if and only if

$$\langle x-1, y-1, z-1\rangle \cdot \langle 3, -2, 1\rangle = 0.$$

Expanding gives the same equation as before.

There is one more product, which is sometimes useful. Given three vectors, $\vec{u}$, $\vec{v}$ and $\vec{w}$, the scalar triple product is the scalar

$$\vec{u}\cdot(\vec{v}\times\vec{w}).$$

It is the signed volume of the parallelepiped determined by $\vec{u}$, $\vec{v}$ and $\vec{w}$ (the sign is positive if $\vec{u}$, $\vec{v}$ and $\vec{w}$ form a right handed set and negative if they form a left handed set).

3. Matrices

Often if one starts with a coordinate system (x_1, x_2, x_3), sometimes it is better to work in a coordinate system (y_1, y_2, y_3) related to the old coordinate system in a simple way:

$$\begin{aligned} 2x_1 - x_2 + x_3 &= y_1 \\ -3x_1 + x_2 + 4x_3 &= y_2 \\ 2x_1 - x_2 + x_3 &= y_3. \end{aligned}$$

Matrices are simply a way to encode this transformation in a compact form

$$\begin{pmatrix} 2 & -1 & 1 \\ -3 & 1 & 4 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

In even more compact notation, $A\vec{x} = \vec{y}$, where A is a 3×3 matrix, $\vec{x}$ is a column vector, a 3×1 matrix (3 rows, 1 column) and $\vec{y}$ has the same shape. To get the entries of the product $A\vec{x}$ take the dot product of a row from A and a column from $\vec{x}$.

More generally, if we want to multiply two matrices A and B, take the dot product of the rows of A and the columns of B:

$$\begin{pmatrix} 2 & 3 & 1 & 2 \\ -1 & -1 & 3 & 4 \\ 0 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 3 \\ -3 & -2 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \\ * & x \end{pmatrix}$$

Question 3.1. *What is x?*

It is the entry obtained by taking the dot product of the 3rd row of A and the 2nd column of B:

$$x = \langle 0, -1, 1, 1 \rangle \cdot \langle 1, 3, -2, 0 \rangle = 0 - 3 - 2 + 0 = -5.$$

For the product AB to make sense, A must have the same number of columns as B has rows.

Question 3.2. *Is $AB = BA$ in general?*

No, for four different reasons.

Sometimes the product make sense one way but not the other way. For example if A is 4×2 and B is 2×3 the product AB is a 4×3 matrix but the product BA does not make sense (3 does not match 4).

Sometimes the product makes sense both ways but the shape is different. For example if A is 3×1 and B is 1×3, AB is a 3×3 (consisting of the nine dot products obtained by multiplying an entry of A with an entry of B). But BA has shape 1×1, one dot product in $\mathbb{R}^3$.

If A and B are both square the product makes sense both ways and has the same shape, but it is still not the same. For example, take

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Then the first entry of AB is 4 but the first entry of BA is 1.

Finally, it is clear that matrix multiplication does not commute if one thinks about transformations. E.g. if A corresponds to reflection in the y-axis and B to rotation through $\pi/4$, then AB represents rotation through $\pi/4$ and reflection in the y-axis and BA represents reflection in the y-axis followed by rotation through $\pi/4$.

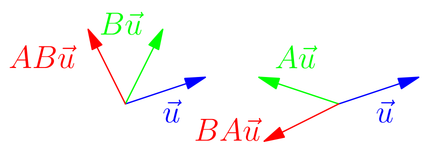


Figure 1. AB vs BA

There is one very special square matrix, called the identity matrix. For example,

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

As a transformation, I_3 does not do anything. In fact,

$$I_3 B = B \qquad \text{and} \qquad AI_3 = A,$$

whenever these products make sense.

Question 3.3. *What does the matrix*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

do as a transformation?

Well it replaces the vector $\langle a_1, a_2 \rangle$ by the vector $\langle a_2, a_1 \rangle$. So $A\hat{\imath} = \hat{\jmath}$ and $A\hat{\jmath} = \hat{\imath}$. A represents reflection in the line $y = x$. So $A^2 = I_2$.

Probably the most important property of the determinant of a matrix is the following

Theorem 3.4. *Let A and B be square matrices of the same size. Then*

$$\det(AB) = \det A \det B.$$

Some transformations are reversible. If A represents a reversible transformation, the matrix corresponding to the inverse transformation is called the inverse matrix A^{-1}. We have

$$A\vec{x} = \vec{y} \qquad \text{and} \qquad \vec{x} = A^{-1}\vec{y},$$

and

$$A^{-1}A = I \qquad \text{and} \qquad AA^{-1} = I.$$

In words the inverse matrix undoes the effect of A.

There is a very useful characterisation of which matrices are invertible (have inverses):

Theorem 3.5. *A is invertible if and only if* $\det A \neq 0$.

Here is a recipe for calculating the inverse of a matrix. This recipe is perfect for 2×2 matrices, (barely) acceptable for 3×3 matrices and simply diabolical for anything larger.

First 2×2. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{then} \qquad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In general, one adopts the following procedure, which we illustrate with the following 3×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 12 \\ 5 & 6 & 0 \end{pmatrix}.$$

Step 1: Form the matrix of minors. In the (i, j) entry, put the determinant you get by erasing the ith row and jth column of A:

$$\begin{pmatrix} -72 & -60 & -15 \\ -18 & -15 & -4 \\ 15 & 12 & 3 \end{pmatrix}.$$

For example the entry in the second row, third column is obtained by taking the matrix A and deleting the second row and third column to get the matrix

$$\begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix}.$$

Now just take the determinant of this 2×2 matrix

$$\begin{vmatrix} 1 & 2 \\ 5 & 6 \end{vmatrix} = 6 - 10 = -4.$$

Step 2: Now flip the signs of the matrix of minors according to the following pattern:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

The result is the matrix of cofactors:

$$\begin{pmatrix} -72 & 60 & -15 \\ 18 & -15 & 4 \\ 15 & -12 & 3 \end{pmatrix}.$$

Step 3: Now take the transpose (flip the matrix about its main diagonal) to get the adjoint matrix:

$$\mathrm{Adj}(A) = \begin{pmatrix} -72 & 18 & 15 \\ 60 & -15 & -12 \\ -15 & 4 & 3 \end{pmatrix}.$$

Step 4: Divide by the determinant to get the inverse matrix. In our case the determinant is 3. So we divide by 3,

$$A^{-1} = \frac{1}{3}\begin{pmatrix} -72 & 18 & 15 \\ 60 & -15 & -12 \\ -15 & 4 & 3 \end{pmatrix} = \begin{pmatrix} -24 & 6 & 5 \\ 20 & -5 & -4 \\ -5 & 4/3 & 1 \end{pmatrix}.$$

Finally, let's check that this is the right answer. We should have

$$A^{-1}A = \begin{pmatrix} -24 & 6 & -5 \\ 20 & -5 & 4 \\ -5 & 4/3 & -1 \end{pmatrix}\begin{pmatrix} -72 & -60 & -15 \\ -18 & -15 & -4 \\ 15 & 12 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

Let's pick an entry at random. Let's calculate the entry in the 2nd row, 3rd column. We should get 0. In fact we get the dot product of the 2nd row of the first matrix and the 3rd column of the second matrix:

$$\langle 20, -5, 4\rangle \cdot \langle -15, -4, 3\rangle = 0,$$

which is indeed correct.

4. Square systems of linear equations

We have already seen that equations of the form

$$ax + by + cz = d,$$

represent planes in $\mathbb{R}^3$.

Question 4.1. *What is the plane through the origin normal to the vector $\vec{n} = \langle 1, 4, 8\rangle$?*

If P is a point in $\mathbb{R}^3$ then P lies in the plane if and only if $\vec{P}$ is orthogonal to $\vec{n}$. So the equation of the plane is given by

$$\vec{P} \cdot \vec{n} = 0.$$

If $P = (x, y, z)$, then $\vec{P} = \langle x, y, z, \rangle$ and

$$\vec{P} \cdot \vec{n} = \langle x, y, z\rangle \cdot \vec{n} = \langle x, y, z\rangle \cdot \langle 1, 4, 8\rangle = x + 4y + 8z.$$

So the equation of the plane is

$$x + 4y + 8z = 0.$$

Note that we can recover the vector $\vec{n}$ orthogonal to the plane from the coefficients of x, y and z.

Question 4.2. *What is the plane through the point $P_0 = (-2, 1, 6)$ normal to the vector $\vec{n} = \langle 1, 4, 8\rangle$?*

Now note that $P = (x, y, z)$ is in the plane if and only if the vector $\overrightarrow{P_0P} = \langle x+2, y-1, z-6\rangle$ is in the plane if and only if $\overrightarrow{P_0P}$ is orthogonal to $\vec{n}$ if and only if $\overrightarrow{P_0}P. \cdot \vec{n} = 0$ if and only if

$$(x + 2) + 4(y - 1) + 8(z - 6) = 0 \qquad \text{so that} \qquad x + 4y + 8z = 50.$$

Once again we can recover $\vec{n}$ from the coefficients of x, y and z.

To determine the equation of a plane, the crucial piece of data is therefore the vector $\vec{n}$. For example, if we are given two vectors living in the plane, then the cross product of these vectors gives us $\vec{n}$.

Question 4.3. *What is the relation between the vector $\vec{v} = \langle -1, 1, 1\rangle$ and the plane $2x - y + 3z = 7$?*

Well, the vector $\vec{n} = \langle 2, -1, 3\rangle$ is orthogonal to the plane. Visibly $\vec{v}$ is not a multiple of $\vec{n}$, so $\vec{v}$ is not orthogonal to the plane. It is parallel to the plane, since

$$\vec{n} \cdot \vec{v} = \langle 2, -1, 3\rangle \cdot \langle -1, 1, 1\rangle = -2 - 1 + 3 = 0.$$

Suppose we are given a 3×3 system of equations,

$$\begin{aligned} a_1x + a_2y + a_3z &= a \\ b_1x + b_2y + b_3z &= b \\ c_1x + c_2y + c_3z &= c. \end{aligned}$$

Compactly,

$$A\vec{x} = \vec{b},$$

where

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

is the coefficient matrix, and

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{and} \qquad \vec{b} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

are column vectors. Each equation represents a plane. Most of the time two planes intersect in a line and then the third plane intersects the line in a single point. Three equations, three unknowns, with any luck there is a single solution. In fact if A is invertible, the unique solution is

$$\vec{x} = A^{-1}\vec{b}.$$

What can go wrong? Let suppose the three planes are $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$. Here is one possible exception. Let's suppose that the first two planes intersect in a line

$$\mathcal{P}_1 \cap \mathcal{P}_2.$$

But suppose that the third plane is parallel to this line. There are two possibilities; the line is contained in the plane (in which case there are infinitely many solutions) or misses the plane entirely (in which case there are no solutions).

Question 4.4. *How many possible configurations are there?*

Here, I am interested in qualitatively different geometric configurations of all three planes.

Let's list the possible configurations according to the solution set.

If the solution is a single point, there is only one configuration. Similarly if the solution is a whole plane, there is only one configuration, all three planes are the same plane. Suppose the solution is a line. There are two possibilities. All three planes contain this line but are different. Or, two planes are the same and the third plane is different.

Finally how many possibilities when there are no solutions? Well we could have three parallel planes. Or two planes are the same plane and the third plane is parallel to this plane. Or the three planes could intersect in three parallel lines (the tobelerone solution). Or two planes are parallel and the third plane intersects the planes in two parallel lines. Four possibilities for no solutions.

In total this makes $1 + 1 + 2 + 4 = 8$ possible configurations.

All seven of the possible degenerate configurations correspond to a matrix which is not invertible, that is, a matrix whose determinant is zero.

It is interesting to distinguish a special case from the general case.

We say that we have a homogeneous system if $\vec{b} = \vec{0}$, so that we are trying to solve the system

$$A\vec{x} = \vec{0}.$$

In this case $\vec{x} = \vec{0}$ is always a solution. If $\det A \neq 0$, then this solution is unique. Otherwise we get a line or a plane through the origin.

Theorem 4.5. *The homogeneous system of equations*

$$A\vec{x} = \vec{0},$$

always has at least one solution, $\vec{x} = 0$.

$\vec{x} = \vec{0}$ is the unique solution if $\det A \neq 0$. Otherwise there are infinitely many solutions.

Now suppose we consider the general case,

$$A\vec{x} = \vec{b}.$$

If $\vec{b} \neq \vec{0}$, the system is called an inhomogeneous system. Here we shift the three planes from the origin. There might be zero, one or infinitely many solutions. If $\det A \neq 0$, then $\vec{x} = A^{-1}\vec{b}$ is the unique solution. Otherwise there might be no solutions or infinitely many solutions.

Theorem 4.6. *The inhomogeneous system of equations*

$$A\vec{x} = \vec{b},$$

might have zero, one or infinitely many solutions.

If $\det A \neq 0$ then $\vec{x} = A^{-1}\vec{b}$ is the unique solution. If $\det A = 0$ then either there are no solutions or infinitely many solutions.

To see why (4.5) and (4.6) are true, at least in dimension three, we have to go through the eight configurations and see what happens in each case. Note that there are only four cases to consider for (4.5), since the four cases where the planes have no common points doesn't occur.

Now the rows of the matrix A are the normal vectors to the three planes. The determinant is the (signed) volume of the parallelepiped spanned by the three normals.

If there is one solution the three planes all point in different directions, we have a parallelepiped and its volume is non-zero.

If there is a plane of solutions, the three planes are the same and the three normal directions are parallel. The determinant is zero, either algebraically because one row is a multiple, or geometrically, since the parallelepiped is degenerate and the volume is zero.

If there is a line of solutions the normal directions are orthogonal to any vector in the line, that is, the normal directions all live in a plane. But then the volume of the parallelepiped is zero.

This proves (4.5).

Finally, suppose that there are no solutions. If two planes are parallel the corresponding normal directions are parallel, so the parallelepiped is degenerate, it lives in a plane and so the parallelepiped has zero volume. This only leaves one, the case where the three planes meet in three parallel lines. In this case the three normal directions live in the plane orthogonal to the direction of the three lines. So the parallelepiped is again degenerate and the volume is zero.

To summarise. The parallelepiped is non-degenerate in only one case, so that the determinant is non-zero in only one case, when the three planes meet in one point. This is the content of (4.5) and (4.6).

5. Parametric curves

We have already seen that one way to represent lines in $\mathbb{R}^3$ is to think of them as being the intersection of two planes. Another approach is to parametrise the line.

Pick two points $Q_0 = (1, -2, 4)$ and $Q_1 = (3, -1, 3)$ and consider the line which contains both points. Imagine a particle traveling along the line at constant speed, which is at Q_0 at time $t = 0$ and at Q_1 at time $t = 1$. In general the position vector of the particle at time t is

$$\vec{Q}(t) = \vec{Q}_0 + t\overrightarrow{Q_0Q_1} = \langle 1, -2, 4\rangle + t\langle 2, 1, -1\rangle = \langle 1 + 2t, -2 + t, 4 - t\rangle.$$

In other words, if $\vec{Q}(t) = \langle x(t), y(t), z(t)\rangle$, then

$$\begin{aligned} x(t) &= 1 + 2t \\ y(t) &= -2 + t \\ z(t) &= 4 - t. \end{aligned}$$

Note that the velocity velocity $\vec{v}$ of the particle is $\overrightarrow{Q_0Q_1} = \langle 2, 1, -1\rangle$. Indeed it is traveling with constant velocity and this is how far the particle moves in unit time. Note that $\vec{v}$ is parallel to the line (or points in the direction of the line).

Question 5.1. *What are the positions of Q_0 and Q_1 relative to the plane $2x - y - z = 3$?*

Well, plug in the coordinates of both points into the equation of the plane. The first point gives $2 + 2 - 4 = 0 < 3$ and the second point gives $6 + 1 - 3 = 4 > 3$. Note that every point is contained in a plane parallel to the plane $2x - y - z = 3$ (think of a stack of pancakes, an infinite stack of pancakes). Q_0 is contained in the plane $2x - y - z = 0$ and Q_1 is contained in the plane $2x - y - z = 4$. So the points are opposite sides of the plane.

It follows that the particle is on the plane at some time t between 0 and 1, so that the line meets the plane.

To find the point of intersection of the plane with the line, plug in $\vec{Q}(t)$ into the equation of the plane and solve for t,

$$3 = 2(1 + 2t) - (-2 + t) - (4 - t) = 4t \qquad \text{and so} \qquad t = \frac{3}{4},$$

which is indeed between 0 and 1. The point is

$$(\frac{5}{2}, -\frac{5}{4}, \frac{13}{4}) = \frac{1}{4}(10, -5, 13).$$

Suppose we tried the same trick with a line parallel to this plane. What would happen? Well, if the line misses the plane, we couldn't solve for t. So we would get an equation of the form

$$a = 3,$$

where a is a constant, not equal to 3. If the line is contained in the plane, then we would get the equation

$$3 = 3,$$

which is valid for any t.

Suppose we are given a line as the intersection of two planes,

$$2x - y + z = 3 \qquad \text{and} \qquad x + 3y - z = 1.$$

How can we find a parametric form of the line? There are two methods.

One is to find two points on this line. Pick another plane and intersect with these two planes. It is convenient to pick the plane $x = 0$. The two equations above reduce to

$$\begin{aligned} -y + z &= 3 \\ 3y - z &= 1. \end{aligned}$$

Adding we get $2y = 4$, so that $y = 2$. This gives $z = 5$. So one point is $Q_0 = (0, 2, 5)$. Now let's pick the plane $x = 1$. The two equations above reduce to

$$\begin{aligned} -y + z &= 1 \\ 3y - z &= 0. \end{aligned}$$

Adding we get $2y = 1$, so that $y = 1/2$. This gives $z = 3/2$. So the other point is $Q_1 = (1, 1/2, 3/2)$.

The line is given parametrically as

$$\vec{Q}(t) = \vec{Q}_0 + t\overrightarrow{Q_0Q_1} = \langle 0, 2, 5\rangle + t\langle 1, -3/2, -7/2\rangle = \langle t, 2-3t/2, 5-7t/2\rangle.$$

Another method is to use the cross product to find the direction of the line. A normal vector to the first plane is $\vec{n}_1 = \langle 2, -1, 1\rangle$ and a normal vector to the second plane is $\vec{n}_2 = \langle 1, 3, -1\rangle$. The line lies in both planes so its direction is orthogonal to both planes. In other words the line is parallel to the cross product:

$$\vec{v} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 2 & -1 & 1 \\ 1 & 3 & -1 \end{vmatrix} = \hat{\imath}\begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} + \hat{k}\begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} = -2\hat{\imath} + 3\hat{\jmath} + 7\hat{k}.$$

Together with the point Q_0 this gives us another way to parametrise the lines

$$\vec{P}(t) = \vec{Q}_0 + t\vec{v} = \langle 0, 2, 5\rangle + t\langle -2, 3, 7\rangle = \langle -2, 2 + 3t, 5 + 7t\rangle.$$

Notice that this is the same line with a different parametrisation.

Question 5.2. *How are the lines*

$\vec{P}(t) = \langle 1+2t, -2+t, 2+5t \rangle$ *and* $\vec{Q}(t) = \langle -2+t, -6+3t, -4+t \rangle$

related?

These two lines intersect. So they are neither skew nor parallel. The key point is to use two different parameters. We want to know if we can find s and t such that

$$\langle 1 + 2s, -2 + s, 2 + 5s \rangle = \langle -2 + t, -6 + 3t, -4 + t \rangle.$$

This gives us three simultaneous linear equations for s and t,

$$\begin{aligned} 1 + 2s &= -2 + t \\ -2 + s &= -6 + 3t \\ 2 + 5s &= -4 + t. \end{aligned}$$

With a little bit of work, one can check that $s = -1$ and $t = 1$ is a solution. So the lines intersect.

Note that we can parametrise a lot more curves than just lines. Consider the example of a cycloid. Here we have a wheel rolling along the ground, and we keep track of a point on the rim of the wheel. What sort of curve does this point trace out?

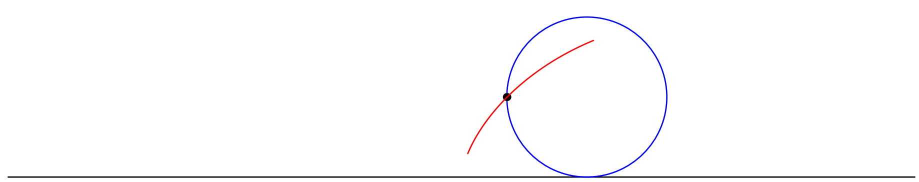


FIGURE 1. A rolling stone

Let's suppose that the wheel has radius a. We will parametrise the motion using the angle θ which the wheel has turned since the start. Then the centre of the wheel has moved a distance of $a\theta$. Let's suppose that the point on the rim starts at $(0, 0)$, so that the centre of the wheel starts at $(0, a)$.

Call P the point on the rim, A the point of contact of the wheel with the floor and B the centre of the wheel.

Then

$$\vec{P} = \vec{A} + \overrightarrow{AB} + \overrightarrow{BP}.$$

Now

$$\vec{A} = \langle a\theta, 0 \rangle \qquad \text{and} \qquad \overrightarrow{AB} = \langle 0, a \rangle,$$

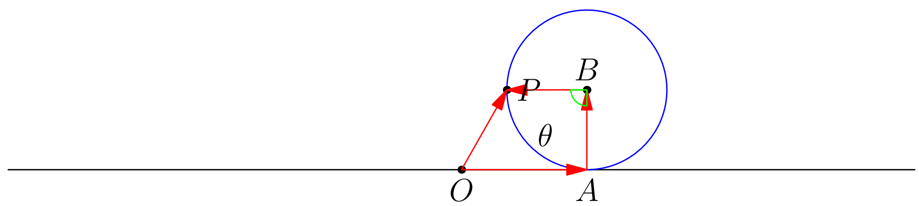


Figure 2. Labels

since the centre of the wheel is always directly above the point of contact with the floor. Now the length of $\overrightarrow{BP}$ is a and the angle θ is the angle from the $-y$-axis.

$$\overrightarrow{BP} = \langle -a\sin\theta, -a\cos\theta\rangle.$$

Putting all of this together,

$$\vec{P} = \langle a(\theta - \sin\theta), a(1-\cos\theta)\rangle.$$

Question 5.3. *What is happening when the marked point is touching the floor?*

Use Taylor series approximation. To simplify the computation, let's take $a = 1$. For t close to zero,

$$f(t) = f(0) + f'(0)t + f''(0)t^2/2 + \dots.$$

This gives

$$\sin\theta \approx \theta - \theta^3/6 \qquad \text{and} \qquad \cos\theta \approx 1 - \theta^2/2.$$

So

$$x(\theta) \approx \theta^3/6 \qquad \text{and} \qquad y(\theta) \approx \theta^2/2.$$

So

$$\frac{y(\theta)}{x(\theta)} \approx \frac{3}{\theta},$$

which as $\theta \to 0$ tends to ∞. So we have a vertical tangent.

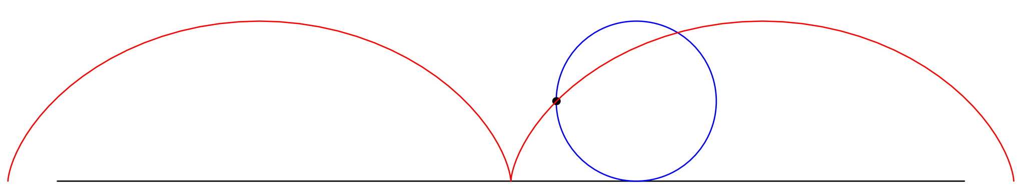


Figure 3. Cycloid

6. Velocity and Acceleration

A particle moving in space sweeps out a curve. The position vector

$$\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k} = \langle x(t), y(t), z(t)\rangle,$$

is naturally a function of time t. For example, the cycloid

$$\vec{r}(t) = \langle t - \sin t, 1 - \cos t\rangle.$$

The velocity vector is simply the derivative of the position vector with respect to time

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \langle \frac{dx}{dt}, \frac{dy}{dt}\rangle.$$

Notice that velocity is a vector; it has a magnitude and a direction. In the case of the cycloid,

$$\vec{v}(t) = \langle 1 - \cos t, \sin t\rangle.$$

At time $t = 0$, the rotation and the motion of the wheel cancel

$$\vec{v}(0) = \langle 0, 0\rangle,$$

and at time $t = \pi$

$$\vec{v}(\pi) = \langle 2, 0\rangle,$$

they combine.

The magnitude of the velocity is the speed $|\vec{v}|$. For the cycloid, the speed is

$$|\langle 1 - \cos t, \sin t\rangle| = ((1 - \cos t)^2 + \sin^2 t)^{1/2} = \sqrt{2}(1 - \cos t)^{1/2}.$$

At $t = 0$ the speed is zero and at $t = \pi$ the speed is 2.

The acceleration vector is simply the derivative of the velocity vector with respect to time,

$$\vec{a} = \frac{d\vec{v}}{dt}.$$

For the cycloid the acceleration vector is

$$\vec{a} = \langle \sin t, \cos t\rangle.$$

Question 6.1. *What is the speed?*

It is

$$|\frac{d\vec{r}}{dt}|.$$

It is not

$$\frac{d|\vec{r}|}{dt}.$$

For example, imagine going around a circle. Then $|\vec{r}|$ is constant, so that the second expression is zero. But you can speed up and slow down even if you are going around a circle.

Let s be the distance travelled along the path. Then

$$\frac{ds}{dt} = |\vec{v}|,$$

is the speed. So we can recover the distance travelled by integrating the speed. However it is not always so easy to do this. For the cycloid, one full revolution is

$$s = \int_0^{2\pi} \sqrt{2 - 2\cos t}\, \mathrm{d}t.$$

If one uses the trigonometric identity,

$$1 - \cos t = 2\sin^2(t/2),$$

one see that the integrand is

$$2\sin(t/2),$$

at least when $\sin(t/2) \geq 0$. So the integral is

$$4\int_0^{\pi/2} 2\sin(t/2)\, \mathrm{d}t = 4\left[-4\cos(t/2)\right]_0^{\pi/2} = 4 - 2^{3/2}.$$

Most of the time it is not possible to calculate the integral. s is called the arclength parameter.

The unit tangent vector is the direction of the velocity,

$$\hat{T} = \frac{\vec{v}}{|\vec{v}|}.$$

In fact

$$\frac{d\vec{r}}{dt} = \frac{d\vec{r}}{ds}\frac{ds}{dt} = \hat{T}\frac{ds}{dt} = |\vec{v}|\hat{T}.$$

Based on some astronomical observations of Tycho Brahe, Johannes Kepler formulated three laws of planetary motion. The second law states that the area swept out by the vector from the sun to the planet sweeps out equal area in equal time. Newton derived all of Kepler's laws from calculus and his universal theory of gravitation.

It is relatively easy to derive Kepler's second law using vector calculus.

The area swept out in time Δt is approximately

$$\frac{1}{2}|\vec{r} \times \Delta\vec{r}| \approx \frac{1}{2}|\vec{r} \times \vec{v}|\Delta t.$$

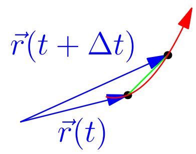


FIGURE 1. Area swept out in time Δt

So Kepler's second law may be restated as saying

$$\frac{dA}{dt} = \frac{1}{2}|\vec{r} \times \vec{v}|,$$

is constant. Now $\vec{r} \times \vec{v}$ is perpendicular to the plane of motion, so the direction of the cross product $\vec{r} \times \vec{v}$ is constant. Putting all of this together, Kepler's second law says that the cross product

$$\vec{r} \times \vec{v}$$

is a constant vector.

The usual Leibniz rule applies to both the dot and the cross product.

$$\frac{d(\vec{a} \cdot \vec{b})}{dt} = \vec{a} \cdot \frac{d\vec{b}}{dt} + \frac{d\vec{a}}{dt} \cdot \vec{b},$$

and

$$\frac{d(\vec{a} \times \vec{b})}{dt} = \vec{a} \times \frac{d\vec{b}}{dt} + \frac{d\vec{a}}{dt} \times \vec{b}.$$

For dot products, we don't need to be careful about the order but the order is important for the cross product.

$$\begin{aligned} \vec{0} &= \frac{d(\vec{r} \times \vec{v})}{dt} \\ &= \vec{r} \times \frac{d\vec{v}}{dt} + \frac{d\vec{r}}{dt} \times \vec{v}. \\ &= \vec{r} \times \vec{a} + \vec{v} \times \vec{v}. \\ &= \vec{r} \times \vec{a}. \end{aligned}$$

It follows that $r \times \vec{v}$ is constant if and only if the acceleration vector is parallel to $\vec{r}$. Note that Newton's second law has a vector form

$$\vec{F} = m\vec{a}.$$

So the acceleration $\vec{a}$ is parallel to $\vec{r}$ if and only if the force $\vec{F}$ is parallel to $\vec{r}$. In other words, Kepler's second law is equivalent to the statement that the force is directed to the sun.

7. Functions of more than one variable

Most functions in nature depend on more than one variable. Pressure of a fixed amount of gas depends on the temperature and the volume; increase the temperature and pressure goes up; increase the volume and the pressure goes down.

To understand a function of one variable, $f(x)$, look at its graph, $y = f(x)$. This is a curve in the plane.

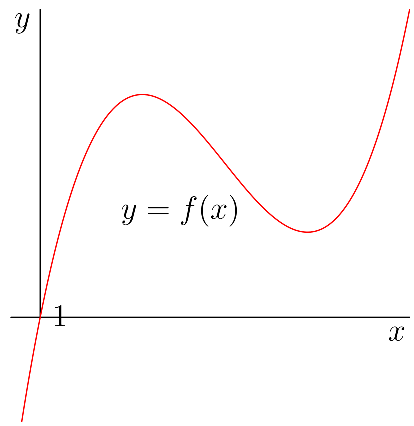


Figure 1. Graph of a function of one variable

To understand a function of two variables, $f(x, y)$, look at its graph $z = f(x, y)$. This is a surface in $\mathbb{R}^3$.

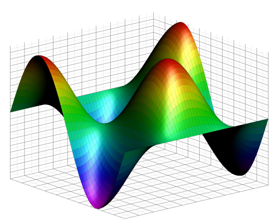


Figure 2. Graph of a function of two variables

Let's do a couple of examples. $f(x, y) = -x$. The graph is $z = -x$. What does this surface look like in $\mathbb{R}^3$? Well, $x + z = 0$ is the equation of a plane. Normal vector $\vec{n} = \langle 1, 0, 1\rangle$ and it passes through the origin.

One way to get a picture is to slice by coordinate planes. If we slice by $y = 0$, we get $z = -x$, a line of slope -1 in the xz-plane. In fact if we slice by any coordinate plane $y = a$, a a constant, we get the same line $z = -x$. If we slice by $x = 0$, we get $z = 0$, a horizontal line in the

yz-plane. If we slice by $x = 1$, we get $z = -1$, a different horizontal line.

How about $f(x, y) = 1 - x^2 - y^2$? If we slice by $y = 0$, we get $z = 1 - x^2$, an upside down parabola. If we slice by $y = 1$, we get $z = -x^2$, another upside down parabola. Similarly if we slice by $y = a$, we get the parabola, $z = -x^2 - a^2$. By symmetry in x and y, we get the same picture if we slice by $x = a$.

How about if we fix z? Then $x^2 + y^2 = 1 - z$. So we only get a non-empty slice, if we take $z \leq 1$. If $z = 0$, we get the circle $x^2 + y^2 = 1$. If we increase z, we get circles of smaller radii. If we decrease z they get bigger.

In fact the graph is a paraboloid.

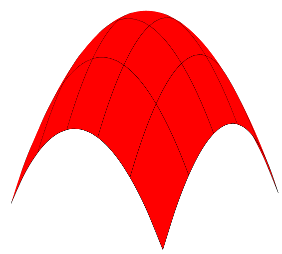


FIGURE 3. Paraboloid

One way to get a picture of the graph is to look at the contour lines. These are lines in the xy-plane of constant height. Formally, they are the solutions to the equation

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

where c is fixed. The contour lines of $f(x, y) = 1 - x^2 - y^2$ are concentric circles centred at the origin:

What does

$$z = \sqrt{x^2 + y^2},$$

look like? Well the contour lines are circles, so it looks like a paraboloid. But if we cut by coordinate planes, we get a different picture. If we take the plane $y = 0$, we get $z = \sqrt{x^2}$, or what comes to the same thing $z = |x|$. The graph of this look like a V. In fact $z = \sqrt{x^2 + y^2}$ is the graph of a cone.

It is not hard to see that $z = x^2 + y^2$ is another paraboloid. It is the same story as $z = 1 - x^2 - y^2$. The contour lines are the circles $x^2 + y^2 = c$. Cutting by coordinate hyperplanes, we get parabolas, but this time the right way up, so that the graph of $z = x^2 + y^2$ is a paraboloid the other way up to $z = 1 - x^2 - y^2$.

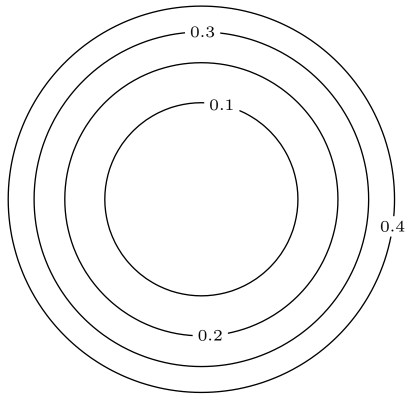


FIGURE 4. Contour lines of paraboloid

What does

$$z = y^2 - x^2,$$

look like? Well the contour lines are hyperbolae:

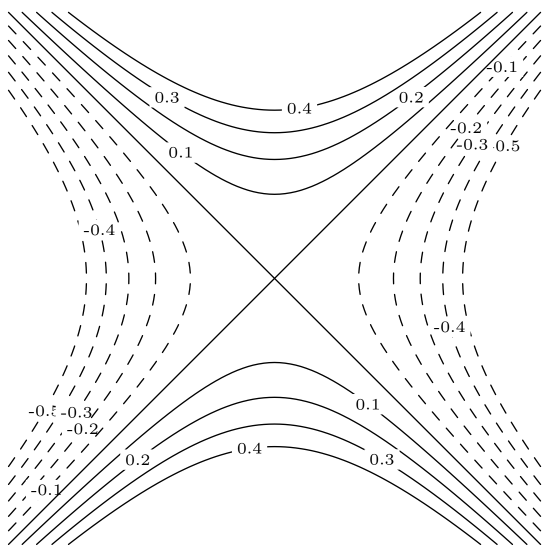


FIGURE 5. Contour lines for $y^2 - x^2$

How about if we take cross sections? Fix $x = a$, we get parabolas $z = y^2 - a^2$. Fix $y = a$, we get upside down parabolas $z = a^2 - x^2$.

The graph of this function is called a saddle point:

One way to understand a function of one variable is to differentiate. The derivative is the slope of the tangent line.

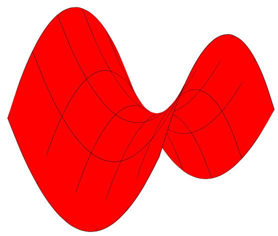


FIGURE 6. Saddle point

If we have a function of two variables, there are two obvious derivatives. We could fix y and vary x, to get a partial derivative

$$f_x(x_0, y_0) = \left.\frac{\partial f}{\partial x}\right|_{x=x_0, y=y_0} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

Similarly, we can fix x and vary y.

$$f_y(x_0, y_0) = \left.\frac{\partial f}{\partial y}\right|_{x=x_0, y=y_0} = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}.$$

f_x is the slope of the tangent line if you cut by the plane $y = y_0$; f_y is the slope of the tangent line to if you cut by the plane $x = x_0$.

It is straightforward to calculate partial derivatives. Let $f(x, y) = x^2 y - \sin(x + y^2)$.

$$f_x = 2xy - \cos(x + y^2) \qquad \text{and} \qquad f_y = x^2 - 2y\cos(x + y^2).$$

$$\frac{\partial(\ln(x \cos y))}{\partial x} = \cos y \frac{1}{x \cos y} = \frac{1}{x},$$

and

$$\frac{\partial(\ln(x \cos y))}{\partial y} = -x \sin y \frac{1}{x \cos y} = -\tan y.$$

We can use partial derivatives to estimate the change in f, if we change x and y by a small amount.

$$\Delta f \approx f_x \Delta x + f_y \Delta y.$$

In fact, we can calculate the tangent plane at a point (x_0, y_0, z_0), where $z_0 = f(x_0, y_0)$. One way to calculate the tangent plane is to use the approximation formula,

$$(\dagger) \qquad z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

In fact the approximation formula works by approximating Δf by using linear approximation. The tangent plane is the best linear approximation to the function f.

The tangent plane is the plane which should contain the tangent line to any curve in the graph. You can get two curves easily, either by fixing y and varying x or by fixing x and varying y. These are the curves you get by cutting by either the plane $y = y_0$ or the plane $x = x_0$. The tangent line to the first curve is

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

and the tangent line to the second curve is

$$z - z_0 = f_y(x_0, y_0)(y - y_0).$$

Visibly (†) contains both tangent lines.

8. Review

Two ways to multiply vectors $\vec{v}$ and $\vec{w}$.

The dot product $\vec{v} \cdot \vec{w}$ takes two vectors and spits out a scalar, a number. Most important identity:

$$\vec{v} \cdot \vec{w} = |\vec{v}||\vec{w}| \cos \theta,$$

where θ is the angle between $\vec{v}$ and $\vec{w}$. Use this to compute θ.

Most important property:

$\vec{v}$ and $\vec{w}$ are orthogonal if and only if $\vec{v} \cdot \vec{w} = 0$.

Question 8.1. *What is the cosine of the angle between the vectors*

$$\vec{v} = \langle -1, 2, 2 \rangle \qquad \textit{and} \qquad \vec{w} = \langle 1, -4, 8 \rangle?$$

$$\cos \theta = \frac{\langle -1, 2, 2 \rangle \cdot \langle 1, -4, 8 \rangle}{|\langle -1, 2, 2 \rangle||\langle 1, -4, 8 \rangle|} = \frac{-1 - 8 + 16}{\sqrt{1 + 4 + 4}\sqrt{1 + 16 + 64}} = \frac{7}{27}.$$

The cross product $\vec{v} \times \vec{w}$ takes two vectors in $\mathbb{R}^3$ and spits out another vector in $\mathbb{R}^3$. Algebraically defined by determinants:

$$\vec{v} \times \vec{w} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Geometrically determined by:

magnitude of $\vec{v} \times \vec{w}$ is the area of the parallelogram given by $\vec{v}$ and $\vec{w}$, that is, $|\vec{v}||\vec{w}| \sin \theta$.

direction is determined by the following two properties:

(i) orthogonal to both $\vec{v}$ and $\vec{w}$,

(ii) the vectors $\vec{v}$, $\vec{w}$ and $\vec{v} \times \vec{w}$ form a right handed set.

Two important properties

$$\vec{v} \times \vec{w} = -\vec{w} \times \vec{v} \qquad \text{so that} \qquad \vec{v} \times \vec{v} = \vec{0}.$$

One can see the first property one of two ways. If you swap two rows of a determinant, the sign changes (the determinant is the signed volume of a parallelepiped). On the other hand as $\vec{v}$, $\vec{w}$ and $\vec{v} \times w$ are a right handed set, $\vec{w}$, $\vec{v}$ and $-\vec{v} \times \vec{w}$ are a right handed set.

Question 8.2. *What is the area of the triangle with sides*

$$\vec{v} = \langle -1, -2, 2 \rangle \qquad \textit{and} \qquad \langle 1, -2, 3 \rangle?$$

We want half the magnitude of the cross product. The cross product is

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ -1 & -2 & 2 \\ 1 & -2 & 3 \end{vmatrix} = \hat{\imath}\begin{vmatrix} -2 & 2 \\ -2 & 3 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} -1 & 2 \\ 1 & 3 \end{vmatrix} + \hat{k}\begin{vmatrix} -1 & -2 \\ 1 & -2 \end{vmatrix} = -2\hat{\imath} + 5\hat{\jmath} + 4\hat{k}.$$

Half the magnitude is

$$\frac{1}{2}(4+25+16)^{1/2} = \frac{1}{2}\sqrt{45} = \frac{3}{2}\sqrt{5}.$$

Planes in $\mathbb{R}^3$ are given by a single linear equation:

$$ax + by + cz = d.$$

Geometrically a plane is a determined by a point P_0 on the plane and a normal direction, $\vec{n}$. The point $P = (x, y, z)$ lies in the plane if and only if the vector $\overrightarrow{P_0P}$ is parallel to the plane, that is, if and only if

$$\overrightarrow{P_0P} \cdot \vec{n} = 0.$$

One can rewrite this equation as

$$\vec{P} \cdot \vec{n} = \vec{P}_0 \cdot \vec{n}.$$

If $\vec{n} = \langle 6, -2, -3\rangle$, and $P_0 = (4, -1, 3)$ then

$$\langle x-4, y+1, z-3\rangle \cdot \langle 6, -2, -3\rangle = 0,$$

that is

$$6(x-4) - 2(y+1) - 3(z-3) = 0,$$

that is

$$6x - 2y - 3z = 31.$$

Note that one can read off a vector orthogonal to the plane from the equation immediately. The plane $ax + by + cz = d$ is orthogonal to $\vec{n} = \langle a, b, c\rangle$. d is a measure of how far the plane is from the origin; if $d = 0$ the plane passes through the origin. If d is not zero the plane has been translated in the direction of $\vec{n}$ (for example, consider horizontal planes, given by $z = 0$, $z = 1$, $z = 2$, $z = -1$, etc. They are planes translated up and down, that is, in the direction of $\hat{k}$.

How can one represent a line? One possibility is as the intersection of two planes. Each plane is determined by a single equation, so a line may be given to you as the set of solutions to two equations. For example, the solutions of the two equations

$$\begin{aligned} 2x - y + z &= 3 \\ 3x + y + z &= 1, \end{aligned}$$

represents a line.

Can one manipulate these two equations to get a single equation?

NO!

This is important (because if you try to eliminate one equation, it is guaranteed you made a mistake and that you were wasting your time). There are lots of ways to see that this is not possible.

(1) We already decided that one equation represents a plane.

(2) Let's look at a concrete example. Suppose we start with the x-axis. Parametrically this is given as $\vec{r}(t) = t\hat{\imath} = \langle t, 0, 0\rangle$. How does one describe this by equations? Well $y = 0$ and $z = 0$ are two obvious equations. Clearly one cannot do better than this; no single linear equation will force both the component of $\hat{\jmath}$ and $\hat{k}$ to be zero.

(3) $\mathbb{R}^3$ is three dimensional. There are three degrees of freedom. Up-down, left-right, front-back. A plane has two degrees of freedom and a line one.

One equation imposes one condition, we lose one degree of freedom. So there are two degrees of freedom left. For example, the equation $y = 0$ means we can no longer go left-right, one constraint. We can still go up-down and front-back, so we still have two degrees of freedom. $y = 0$ represents a plane.

If we have two equations, each equation imposes one condition, so a pair of equations imposes two conditions. This leaves one degree of freedom. For example, $y = 0$ and $z = 0$ impose two conditions; you cannot move left-right and you cannot go up-down. This leaves one degree of freedom, front-back. The pair of equations $y = 0$ and $z = 0$ represents a line.

Question 8.3. *What is the equation of the plane containing the point $P_0 = (3, -4, 1)$ and the line given as the intersection of the two planes*

$$\begin{aligned} 2x - y + z &= 3 \\ 3x + y + z &= 1? \end{aligned}$$

We need to find the normal direction $\vec{n}$ of the plane. For this we need two vectors $\vec{v}$ and $\vec{w}$ parallel to the plane. For this we need two points P_1 and P_2 in the plane.

Obviously we want to choose two points P_1 and P_2 belonging to the line. Intersect the line with a plane to get a point. Take $x = 0$. Put this into the two equations we get

$$\begin{aligned} -y + z &= 3 \\ y + z &= 1. \end{aligned}$$

$z = 2$ and $y = -1$. So $P_1 = (0, -1, 2)$ is a point on the line.

Or we could take $x = 2$.

$$\begin{aligned} -y + z &= -1 \\ y + z &= -5. \end{aligned}$$

In this case $2z = -6$, $z = -3$ and so $y = -2$. So $P_2 = (2, -2, -3)$ is a point on the plane.

The vectors

$$\vec{v} = \overrightarrow{P_0P_1} = \langle -3, 3, 1\rangle \qquad \text{and} \qquad \vec{w} = \overrightarrow{P_0P_2} = \langle -1, 2, -4\rangle$$

are parallel to the plane. The cross-product is orthogonal to the plane:

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ -3 & 3 & 1 \\ -1 & 2 & -4 \end{vmatrix} = \hat{\imath}\begin{vmatrix} 3 & 1 \\ 2 & -4 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} -3 & 1 \\ -1 & -4 \end{vmatrix} + \hat{k}\begin{vmatrix} -3 & 3 \\ -1 & 2 \end{vmatrix} = -14\hat{\imath} - 13\hat{\jmath} - 3\hat{k}.$$

So the equation of the plane is

$$\langle x - 3, y + 4, z - 1\rangle \cdot \langle -14, -13, -3\rangle = 0,$$

so that

$$-14(x - 3) - 13(y + 4) - 3(z - 1) = 0.$$

Rearranging, we get

$$14x + 13y + 3z = -7.$$

There is another way to represent lines, we can parametrise a line. If Q_0 and Q_1 are two points in $\mathbb{R}^3$, then

$$\vec{r}(t) = \vec{Q}_0 + t\overrightarrow{Q_0Q_1}.$$

When $t = 0$, $\vec{r}(0) = Q_0$ and when $t = 1$, $\vec{r}(1) = Q_1$. Given a value for t, we get a point of the line. If we put $\overrightarrow{Q_0Q_1} = \vec{v}$, then we rewrite this parametrisation as

$$\vec{r}(t) = \vec{Q}_0 + t\overrightarrow{Q_0Q_1} = \vec{Q}_0 + t\vec{v}.$$

Here $\vec{v} = \overrightarrow{Q_0Q_1}$ is the velocity vector of the particle (at time $t = 0$, it is at Q_0 and time $t = 1$ at Q_1, or one could just differentiate).

If $Q_0 = (1, 2, 3)$ and $Q_1 = (2, -5, 2)$, then a parametrisation of the line through Q_0 and Q_1 is

$$\vec{r}(t) = \langle 1, 2, 3\rangle + t\langle 1, -7, -1\rangle = \langle 1 + t, 2 - 7t, 3 - t\rangle.$$

Question 8.4. *What is the shortest distance between the two lines*

$$\vec{r}_1(t) = \langle 6+2t, -1+t, 8+2t\rangle \qquad \textit{and} \qquad \vec{r}_2(t) = \langle 5-2t, -3+2t, 1+t\rangle?$$

We first check that these two lines are not parallel. The lines are parallel to

$$\vec{v} = \langle 2, 1, 2\rangle \qquad \text{and} \qquad \vec{w} = \langle -2, 2, 1\rangle.$$

$\vec{v}$ and $\vec{w}$ are not parallel ($\vec{w}$ is not a multiple of $\vec{v}$) so the lines are not parallel.

There are at least three different ways to solve this problem. All of them rely on the following basic observation. Suppose that P_1 and P_2 are the two closest points on either line. Then $\overrightarrow{P_0P_1}$ is orthogonal to both $\vec{v}$ and $\vec{w}$.

Now we describe the three methods. First the basic principle.

(1) Pick two random points R_1 and R_2 on the lines. The length of $\overrightarrow{P_1P_2}$ is nothing more than the (absolute value of the) component of $\overrightarrow{R_1R_2}$ in the direction of $\overrightarrow{P_1P_2}$.
(2) There are two parallel planes, containing either line. To find the distance between two parallel planes is relatively easy.
(3) $\overrightarrow{P_1P_2}$ is orthogonal to both $\vec{v}$ and $\vec{w}$. This gives two equations for the position of P_1 and P_2 and using this we can find P_1 and P_2.

Now to the execution.

Method #1: If we set $t = 0$ then we get two random points,

$$R_1 = (6, -1, 8) \qquad \text{and} \qquad R_2 = (5, -3, 1).$$

As the vector $\overrightarrow{P_1P_2}$ is orthogonal to $\vec{v}$ and $\vec{w}$, it is parallel to the cross product:

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{vmatrix} = \hat{\imath}\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} 2 & 2 \\ -2 & 1 \end{vmatrix} + \hat{k}\begin{vmatrix} 2 & 1 \\ -2 & 2 \end{vmatrix} = -3\hat{\imath} - 6\hat{\jmath} + 6\hat{k}.$$

So $\langle 1, 2, -2\rangle$ is othogonal to both $\vec{v}$ and $\vec{w}$. Dividing by the length, we get a unit vector orthogonal to both $\vec{v}$ and $\vec{w}$,

$$\hat{u} = \frac{1}{3}\langle 1, 2, -2\rangle.$$

The component of

$$\overrightarrow{R_1R_2} = \langle -1, -2, -7\rangle,$$

in the direction of $\hat{u}$, is

$$|\overrightarrow{R_1R_2}|\cos\theta.$$

But

$$\overrightarrow{R_1R_2}\cdot\hat{u} = |\overrightarrow{R_1R_2}||\hat{u}|\cos\theta = |\overrightarrow{R_1R_2}|\cos\theta.$$

So we just need to take the dot product:

$$\langle -1, -2, -7\rangle \cdot \frac{1}{3}\langle 1, 2, -2\rangle = \frac{1}{3}(-1 - 4 + 14) = 3.$$

This distance is 3.

Method #2: If $\mathcal{P}_1$ contains the first line and is parallel to the second line, it must be parallel to $\vec{v}$ and $\vec{w}$. So it must be orthogonal to $\vec{n} = \langle 1, 2, -2\rangle$, the cross product. The first plane has normal vector $\vec{n}$ and passes through $R_1 = (6, -1, 8)$. Hence

$$0 = \langle x - 6, y + 1, z - 8\rangle \cdot \langle 1, 2, -2\rangle = (x - 6) + 2(y + 6) - 2(z - 8).$$

Rearranging, we get

$$x + 2y - 2z = -12.$$

Similarly the second plane contains $R_2 = (5, -3, 1)$ and has normal vector $\vec{n}$,

$$0 = \langle x - 5, y - 3, z - 1\rangle \cdot \langle 1, 2, -2\rangle = (x - 5) + 2(y - 3) - 2(z - 1).$$

Rearranging, we get

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

Pick any point on the first plane. $P = (0, 0, 6)$ lies on the first plane. The line through this point parallel to $\vec{n}$ meets the second plane at a point Q whose distance from P is the distance between the two planes (whence the two lines).

The line through P parallel to $\vec{n}$ is given by

$$\vec{r}(t) = \langle 0, 0, 6\rangle + t\langle 1, 2, -2\rangle = \langle t, 2t, 6 - 2t\rangle.$$

This is on the second plane when

$$t + 4t - 12 + 4t = -3 \qquad \text{so that} \qquad t = 1.$$

The point $Q = (1, 2, 4)$. $\overrightarrow{PQ} = \langle 1, 2, -2\rangle$, which has length 3.

Method #3: We first parametrise the first line with a different parameter s.

$$\overrightarrow{P_1P_2} = \vec{r}_2(t) - \vec{r}_1(s) = \langle -1, -2, -7\rangle - s\langle 2, 1, 2\rangle + t\langle -2, 2, 1\rangle.$$

$\overrightarrow{P_1P_2}$ is orthogonal to $\vec{v}$ and $\vec{w}$ if and only if

$$\overrightarrow{P_1P_2} \cdot \vec{v} = 0 \qquad \text{and} \qquad \overrightarrow{P_1P_2} \cdot \vec{w} = 0.$$

This gives us two equations for s and t,

$$-9s = 18$$
$$9t = 9$$

Hence $s = -2$, $t = 1$. The vector

$$\overrightarrow{P_1P_2} = \langle -1, -2, -7 \rangle + 2\langle 2, 1, 2 \rangle + \langle -2, 2, 1 \rangle = \langle 1, 2, -2 \rangle.$$

This has length 3.

9. Max-min problems

Let $z = f(x, y)$ be a differentiatiable function. If f has a local minimum or maximum at (x_0, y_0) then $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0)$ are both zero. In this case the tangent plane is horizontal, since the tangent plane is determined by the two tangent lines to the coordinate cross-sections, and these are both horizontal. Equivalently, by the approximation formula, $\Delta f \approx 0$ (or even better, $|\Delta f|$ is much smaller than either $|\Delta x$ or $|\Delta y|$).

We say that (x_0, y_0) is a critical point of $f(x, y)$ if

$$f_x(x_0, y_0) = 0 \qquad \text{and} \qquad f_y(x_0, y_0) = 0.$$

Critical points come in three different flavours: a local minimum, a local maximum or a saddle point.

Question 9.1. *Find the critical points of*

$$f(x, y) = 5x^2 + 2xy + 10y^2 + 18x - 16y + 30.$$

The partials are

$$f_x(x, y) = 10x + 2y + 18 \qquad \text{and} \qquad f_y(x, y) = 2x + 20y - 16.$$

If we set those equal to zero, we get the critical points. We get two simultaneous linear equations,

$$\begin{aligned} 5x + y &= -9 \\ x + 10y &= 8. \end{aligned}$$

Hence $49y = 49$, $y = 1$ and so $x = -2$. $(-2, 1)$ is the only critical point. Note that

$$f(x, y) = (x + 3y - 1)^2 + (2x - y + 5)^2 + 4,$$

and so we have a minimum. The minimum value is 4. Note that we have a paraboloid (shifted and rescaled).

Suppose that we want to maximise and minimise a function of two variables. If we have a function of one variable on an interval we find the critical points, check to see if they are maxima or minima (possibly using the 2nd derivative test) and then compare these with the values at the endpoints.

For a function of two variables we will do the same thing but note that checking to see what happens at the endpoints now means checking to see what happens on the boundary of the region.

Question 9.2. *What is the maximum and minimum value of*

$$f(x, y) = 5x^2 + 2xy + 10y^2 + 18x - 16y + 30.$$

over the region $-3 \leq x \leq 3$ *and* $-2 \leq y \leq 2$*?*

The region is a rectangle with sides $x = 3$, $-2 \leq y \leq 2$; $x = -33$, $-2 \leq y \leq 2$; $-3 \leq x \leq 3$, $y = 2$; $-3 \leq x \leq 3$, $y = -2$:

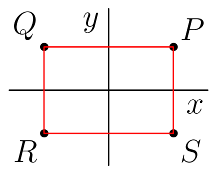


FIGURE 1. Boundary of region in xy-plane

Since there is only one critical point, a local minimum inside the boundary, we know the minimum is 4 (think about why this is true!), but the maximum must occur on the boundary. When $y = 2$ we get

$$g(x) = f(x, 2) = 5x^2 + 4x + 40 + 18x - 32 + 30 = 5x^2 + 22x - 38,$$

a function of one variable. $g(x)$ is a parabola, so has a local minimum. The maximum of $g(x)$ is at one of the endpoints, $x = 3$ or $x = -3$, which correspond to $P = (3, 2)$ and $Q = (-3, 2)$ for $f(x, y)$. A similar analysis along the other three edges tells us the maximum must occur at one of P, Q, R and S. Plugging in values we see that the maximum is at $R = (3, -2)$ and the maximum value is 189.

Suppose we have a bunch of data, a collection of pairs (x_i, y_i), $1 \leq i \leq n$. We try to fit a curve to this data. Quite often we expect the data to lie along a line. We posit that the data fits the equation,

$$y = ax + b.$$

Here a and b are unknowns and we try to find the best choice of a and b. The deviations are

$$y_i - (ax_i + b).$$

We choose a and b to minimise the sum of the squares of the deviations

$$D = \sum_i (y_i - (ax_i + b))^2.$$

We find the critical points of D,

$$\frac{\partial D}{\partial a} = 0 \qquad \text{and} \qquad \frac{\partial D}{\partial b} = 0.$$

This gives two equations

$$\sum_i 2(y_i - (ax_i + b))(-x_i) = 0 \qquad \text{and} \qquad \sum_i 2(y_i - (ax_i + b))(-1) = 0.$$

Collecting together terms in these expressions we get

$$(\sum_i x_i^2)a + (\sum x_i)b = \sum_i x_i y_i$$
$$(\sum_i x_i)a + nb = \sum_i y_i.$$

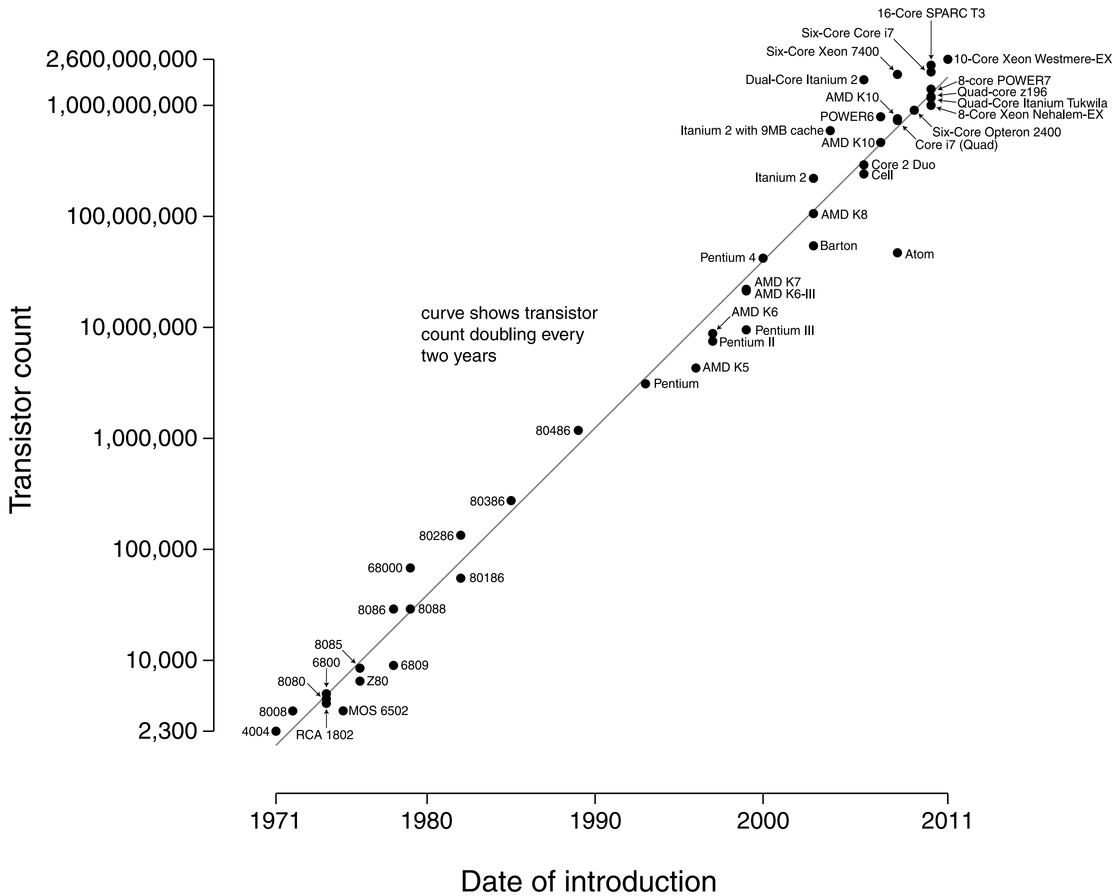
The same method works when we try to fit data to other functions. For example, imagine trying to fit data to exponential functions $y = ce^{ax}$. In this case take logs of both sides to get

$$\ln y = \ln(ce^{ax}) = \ln c + ax.$$

Put $b = \ln c$ and we are down to the method of least squares, with the data $(x_i, \ln y_i)$.

For example, Moore's law predicts that the number of transistors on an integrated circuit doubles roughly every two years. The following picture (source: wikipedia) plots the year against a log of the number of the transitors. Note the almost uncanny fit.

Microprocessor Transistor Counts 1971-2011 & Moore's Law



10. Second derivative test

Let's turn to the problem of determining the nature of the critical points. Recall that there are three possibilities; either we have a local maximum, a local minimum or a saddle point.

Let's start with the key case, a quadratic polynomial.

$$f(x, y) = ax^2 + bxy + cy^2.$$

The basic trick is to complete the square. For example,

$$f(x, y) = x^2 - 2xy + 3y^2 = (x - y)^2 + 2y^2.$$

If we take the partials, we get

$$2ax + by \qquad \text{and} \qquad bx + 2cy.$$

Setting these equal to zero, we get

$$\begin{aligned} 2ax + by &= 0 \\ bx + 2cy &= 0. \end{aligned}$$

A homogeneous pair of linear equations. We can rewrite this as a matrix equation:

$$\begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Provided

$$\begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}$$

is an invertible matrix (that is, the determinant $4ac - b^2 \neq 0$), this has the unique solution

$$(x, y) = (0, 0).$$

So we just want to know what sort of critical point we have at the origin. Let's suppose that $a \neq 0$, then

$$\begin{aligned} z &= a\left(x^2 + \frac{b}{a}xy\right) + cy^2 \\ &= a\left(x + \frac{b}{2a}y\right)^2 + \left(c - \frac{b^2}{4a}\right)y^2 \\ &= \frac{1}{4a}\left(4a^2\left(x + \frac{b}{2a}y\right)^2 + \left(4ac - b^2\right)y^2\right). \end{aligned}$$

To get from the first line to the second line we completed the square. The advantage of the third line is that we know that $4a^2$ is always positive (we are assuming that $a \neq 0$).

There are three cases.

Case 1: $a > 0$, $4ac - b^2 > 0$. $f(x, y)$ is a sum of two squares and $(0, 0)$ is a local minimum.

Case 2: $a < 0$, $4ac - b^2 > 0$. $-f(x, y)$ is a sum of two squares and $(0, 0)$ is a local maximum.

Case 3: $4ac - b^2 < 0$. $f(x, y)$ is the difference of two squares and $f(x, y)$ is a saddle point.

The case $4ac - b^2 = 0$ is a degenerate case (the second derivative test fails).

For the second derivative test, one looks at the second derivatives of f. There are four second derivatives,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx} \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{xy} \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{yx}.$$

Example 10.1. *Let $f(x, y) = x^3 y + xy$.*

We have

$$f_x = 3x^2 y + y \qquad \text{and} \qquad f_y = x^3 + x$$

$$f_{xx} = 6xy \qquad f_{yy} = 0$$

$$f_{xy} = 3x^2 + 1 \qquad f_{yx} = 3x^2 + 1.$$

In general, $f_{xy} = f_{yx}$ (except for some pathological functions). Suppose that (x_0, y_0) is a critical point of $f(x, y)$. Let $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$ and $C = f_{yy}(x_0, y_0)$.

If $AC - B^2 > 0$ then we have a local maximum or minimum; a minimum if $A > 0$ and a maximum if $A < 0$.

If $AC - B^2 < 0$ then we have a saddle point.

If $AC - B^2 = 0$ then the second derivative test is inconclusive (a nice way to say that we know absolutely nothing and we were wasting our time needlessly computing 2nd derivatives).

Let's see what happens if

$$f(x, y) = ax^2 + bxy + cy^2.$$

Then

$$A = f_{xx}(0, 0) = 2a \qquad B = f_{xy}(0, 0) = b \qquad \text{and} \qquad C = f_{yy}(0, 0) = 2c.$$

So

$$AC - B^2 = 4ac - b^2,$$

and the test works.

In general, we have

$$\Delta f \approx f_x(p)(x-x_0)+f_y(y-y_0)+\frac{1}{2}f_{xx}(p)(x-x_0)^2+f_{xy}(p)(x-x_0)(y-y_0)+\frac{1}{2}f_{yy}(p)(y-y_0)^2,$$

where the partials are all computed at $p = (x_0, y_0)$. At a critical point we have

$$\Delta f \approx \frac{A}{2}(x - x_0)^2 + B(x - x_0)(y - y_0) + \frac{C}{2}(y - y_0)^2.$$

Now if we want to find global maximum or minimum, we also have to worry about what happens on the boundary or as we approach infinity.

Example 10.2. *What are the maxima and minima of*

$$f(x, y) = x + y + \frac{1}{xy}?$$

First find critical points

$$f_x = 1 - \frac{1}{x^2 y} \qquad \text{and} \qquad f_y = 1 - \frac{1}{xy^2}.$$

Setting these equal to zero, we get

$$x^2 y = 1 \qquad \text{and} \qquad xy^2 = 1.$$

Dividing we get $x = y$. Hence $x^3 = 1$ and so $x = y = 1$. $(1, 1)$ is the only critical point.

We now apply the second derivative test.

$$f_{xx} = \frac{2}{x^3 y} \qquad f_{xy} = \frac{1}{x^2 y^2} \qquad \text{and} \qquad f_{yy} = \frac{2}{xy^3}.$$

So

$$A = f_{xx}(1, 1) = 2 \qquad B = f_{xy}(1, 1) = 1 \qquad \text{and} \qquad C = f_{yy}(1, 1) = 2.$$

$AC - B^2 = 4 - 1 = 3 > 0$, so we have either a local maximum or a local minimum. $A = 2 > 0$ so we have a local minimum. But what happens at the boundary and as we go to infinity? As $x \to 0$ $f(x, y) \to \infty$. Similarly if $y \to 0$, $x \to \infty$ or $y \to \infty$.

So $f(x, y)$ has a global minimum at $(x, y) = (1, 1)$ and there is no maximum.

11. Differentials and the chain rule

Let $w = f(x, y, z)$ be a function of three variables. Introduce a new object, called the total differential.

$$\mathrm{d}f = f_x \,\mathrm{d}x + f_y \,\mathrm{d}y + f_z \,\mathrm{d}z.$$

Formally behaves similarly to how Δf behaves,

$$\Delta f \approx f_x \,\Delta x + f_y \,\Delta y + f_z \,\Delta z.$$

However it is a new object (it is not the same as a small change in f as the book would claim), with its own rules of manipulation. For us, the main use of the total differential will be to understand the chain rule.

Suppose that x, y and z are functions of one variable t. Then $w = f(x, y, z)$ becomes a function of t. Divide the equation above to get the derivative of f,

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}.$$

This is an instance of the chain rule.

Example 11.1. *Let $f(x, y, z) = xyz + z^2$. Suppose that $x = t^2$, $y = 3/t$ and $z = \sin t$.*

Then

$$f_x = yz \qquad f_y = xz \qquad \text{and} \qquad f_z = 2z,$$

so that

$$\frac{dw}{dt} = 2yzt - \frac{3xz}{t^2} + (xy + 2z)\cos t = 3\sin t + (3t + 2\sin t)\cos t.$$

On the other hand, if we substitute for x, y and z, we get

$$w = 3t \sin t + \sin^2 t,$$

and we can calculate directly,

$$\frac{dw}{dt} = 3\sin t + 3t\cos t + 2\sin t\cos t.$$

There are two ways to see that the chain rule is correct.

$$\mathrm{d}x = x'(t)\,\mathrm{d}t \qquad \mathrm{d}y = y'(t)\,\mathrm{d}t \qquad \text{and} \qquad \mathrm{d}z = z'(t)\,\mathrm{d}t.$$

Substituting we get

$$\begin{aligned} \mathrm{d}w &= f_x \,\mathrm{d}x + f_y \,\mathrm{d}y + f_z \,\mathrm{d}z \\ &= f_x x'(t)\,\mathrm{d}t + f_y y'(t)\,\mathrm{d}t + f_z z'(t)\,\mathrm{d}t, \end{aligned}$$

and dividing by $\mathrm{d}t$ gives us the chain rule.

More rigorously, start with the approximation formula,

$$\Delta w \approx f_x \,\Delta x + f_y \,\Delta y + f_z \,\Delta z,$$

divide both sides by Δt and take the limit as $\Delta t \to 0$.

One can use the chain rule to justify some of the well-known formulae for differentiation.

Let $f(u,v) = uv$. Suppose that $u = u(t)$ and $v = v(t)$ are both functions of t. Then

$$\frac{d(uv)}{dt} = f_u \frac{du}{dt} + f_v \frac{dv}{dt} = vu' + uv',$$

which is the product rule. Similarly if $f = u/v$, then

$$\frac{d(u/v)}{dt} = f_u \frac{du}{dt} + f_v \frac{dv}{dt} = \frac{1}{v} u' - \frac{u}{v^2} v' = \frac{u'v - v'u}{v^2},$$

which is the quotient rule.

Now suppose that $w = f(x,y)$ and $x = x(u,v)$ and $y = y(u,v)$. Then

$$\begin{aligned}
\mathrm{d}w &= f_x \,\mathrm{d}x + f_y \,\mathrm{d}y \\
&= f_x(x_u \,\mathrm{d}u + x_v \,\mathrm{d}v) + f_y(y_u \,\mathrm{d}u + y_v \,\mathrm{d}v) \\
&= (f_x x_u + f_y y_u)\,\mathrm{d}u + (f_x x_v + f_y y_v)\,\mathrm{d}v. \\
&= f_u \,\mathrm{d}u + f_v \,\mathrm{d}v.
\end{aligned}$$

If we write this out in long form, we have

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \qquad \text{and} \qquad \frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}.$$

Example 11.2. *Suppose that $w = f(x,y)$ and we change from Cartesian to polar coordinates,*

$$\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta.
\end{aligned}$$

We have

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \cos\theta & \qquad \frac{\partial x}{\partial \theta} &= -r\sin\theta \\
\frac{\partial y}{\partial r} &= \sin\theta & \qquad \frac{\partial x}{\partial \theta} &= r\cos\theta.
\end{aligned}$$

So

$$\begin{aligned}
f_r &= \cos\theta f_x + \sin\theta f_y \\
f_\theta &= -r\sin\theta f_x + r\cos\theta f_y.
\end{aligned}$$

12. The Gradient and directional derivatives

We have

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}.$$

We can rewrite this as

$$\nabla f \cdot \vec{v}(t),$$

where

$$\nabla f = f_x \hat{\imath} + f_y \hat{\jmath} + f_z \hat{k} \qquad \text{and} \qquad \vec{v} = \frac{d\vec{r}}{dt} = x'(t)\hat{\imath} + y'(t)\hat{\jmath} + z'(t)\hat{k}/$$

∇f is called the gradient of f. For a given point, we get a vector (so that ∇f is a vector valued function). Perhaps one of the most important properties of the gradient is:

Theorem 12.1. *∇f is orthogonal to the level surface $w = c$.*

Example 12.2. *Let $f(x, y, z) = ax + by + cz$.*

The level surface $w = d$ is the plane

$$ax + by + cz = d.$$

The gradient is

$$\nabla f = \langle a, b, c \rangle,$$

which is indeed a normal vector to the plane $ax + by + cz = d$.

Example 12.3. *Let $f(x, y) = x^2 + y^2$.*

The level curve $w = c$ is a circle,

$$x^2 + y^2 = c,$$

centred at the origin of radius $\sqrt{c}$. The gradient is

$$\nabla f = \langle 2x, 2y \rangle,$$

which is a radial vector, orthogonal to the circle.

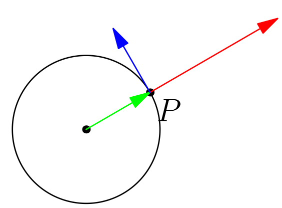


Figure 1. 3 vectors: green position, red gradient, blue velocity

Example 12.4. *Let $f(x, y) = y^2 - x^2$.*

The level curve is a hyperbola,

$$y^2 - x^2 = c,$$

with asymptotes $y = x$ and $y = -x$. The gradient is

$$\nabla f = \langle -2x, 2y \rangle.$$

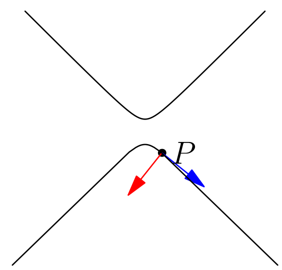


FIGURE 2. Red gradient, blue tangent vector

Proof of (**??**). Pick a curve $\vec{r}(t)$ contained in the level surface $w = c$. The velocity vector $\vec{v} = \vec{r}'(t)$ is contained in the tangent plane. By the chain rule,

$$0 = \frac{dw}{dt} = \nabla f \cdot \vec{v} = 0,$$

so that ∇f is perpendicular to every vector parallel to the tangent plane. □

We can use this to calculate the tangent plane. For example, consider

$$2x^2 - y^2 - z^2 = 6.$$

Let's calculate the tangent plane to this surface at the point $(x_0, y_0, z_0) = (2, 1, 1)$. We have

$$\nabla f = \langle 4x, -2y, -2z \rangle.$$

At $(x_0, y_0, z_0) = (2, 1, 1)$, the gradient is $\langle 8, -2, -2 \rangle$, so that $\vec{n} = \langle 4, -1, -1 \rangle$ is a normal vector to the tangent plane. It follows that the equation of the tangent plane is

$$0 = \langle x - 2, y - 1, z - 1 \rangle \cdot \langle 4, -1, -1 \rangle \qquad \text{so that} \qquad 4x - y - z = 6,$$

is the equation of the tangent plane.

In this example, there are other ways to figure out an equation for the tangent plane. We could write z as a function of x and y,

$$z = \sqrt{2x^2 - y^2 - 6},$$

and find an equation for the tangent plane in the standard way. Beware that this is not always possible.

Suppose that we are at a point (x_0, y_0) in the plane and we move in a direction $\hat{u} = \langle a, b\rangle$. We can define the directional derivative in the direction $\hat{u}$. Consider the line

$$\vec{r}(s) = \langle x_0, y_0\rangle + s\langle a, b\rangle.$$

The velocity vector is $\hat{u}$, which has unit length, so that the speed is one. In other words, $\vec{r}(s)$ is parametrised by arclength.

$$\left.\frac{dw}{ds}\right|_{\hat{u}} = \lim_{s\to 0} \frac{f(x_0 + sa, y_0 + sb) - f(x_0, y_0)}{\Delta s}.$$

If $\hat{u} = \hat{\imath}$, then the directional derivative is f_x and if $\hat{u} = \hat{\jmath}$ then the directional derivative is f_y. In general, if we slice the graph $w = f(x, y)$ by vertical planes, the directional derivative is the slope of the resulting curve.

$$\left.\frac{dw}{ds}\right|_{\hat{u}} = \nabla f \cdot \frac{d\vec{r}}{ds} = \nabla f \cdot \hat{u}.$$

Question 12.5. *Fix a vector $\vec{v} = \langle c, d\rangle$ in the plane. Which unit vector $\hat{u}$*

(1) maximises $\vec{v} \cdot \hat{u}$?
(2) minimises $\vec{v} \cdot \hat{u}$?
(3) When is $\vec{v} \cdot \hat{u} = 0$?

We know

$$\vec{v} \cdot \hat{u} = |\vec{v}||\hat{u}| \cos\theta = |\vec{v}| \cos\theta.$$

$|\vec{v}|$ is fixed as $\vec{v}$ is fixed. So we want to

(1) maximise $\cos\theta$,
(2) minimise $\cos\theta$
(3) and we want to know when $\cos\theta = 0$.

This happens when

(1) $\theta = 0$, in which case $\cos\theta = 1$,
(2) $\theta = \pi$, in which case $\cos\theta = -1$,
(3) and $\theta = \pi/2$, in which case $\cos\theta = 0$.

Geometrically the three cases correspond to:

(1) $\hat{u}$ points in the same direction as $\vec{v}$,
(2) $\hat{u}$ points in the opposite direction, and
(3) $\hat{u}$ is orthogonal to $\vec{v}$.

Now consider $v = \nabla f$. The directional derivative is

$$\left.\frac{dw}{ds}\right|_{\hat{u}} = \nabla f \cdot \hat{u}.$$

This is maximised when $\hat{u}$ points in the direction of ∇f. In other words, ∇f points in the direction of maximal increase, $-\nabla f$ points in the direction of maximal decrease and it is orthogonal to the level curves. The magnitude $|\nabla f|$ of the gradient is the directional derivative in the direction of ∇f, it is the largest possible rate of change.

In terms of someone climbing a mountain: ∇f points in the direction you need to go straight up the mountain, with magnitude the slope. $-\nabla f$ points straight down and ∇f is orthogonal to the level curve, which is the direction which takes you around the mountain.

13. Lagrange multipliers

If we want to maximise a function over a region, we also need to maximise the function over the boundary of the region, which is often given to us as a level surface. We are asked to solve something like

$$\text{maximise} \quad f(x,y,z) \qquad \text{subject to} \qquad g(x,y,z) = c.$$

Typical problem:

Example 13.1. *Find the closest point to the origin lying on the hyperbola $xy = 4$.*

If we put this in the form above we want to

$$\text{minimise} \quad x^2 + y^2 \qquad \text{subject to} \qquad xy = 4.$$

One obvious way to solve any such problem is to eliminate a variable.

$$y = \frac{4}{x}.$$

So we want to minimise

$$x^2 + \frac{16}{x^2}.$$

It turns out that in even quite simple examples, this can get very messy.

Now we know that one of the closest points is $(2,2)$, and at this point the tangent directions to the curve $xy = 4$ and to the curve $x^2 + y^2 = 8$ (the level curve of the function we want to maximise) are parallel.

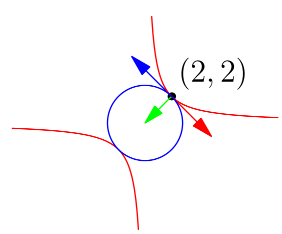


Figure 1. Closest point: tangent vectors are parallel

If the tangent vectors are not parallel, we can do better, we can move along the curve $xy = 4$ to a closer point. Here is what happens at $(4,1)$:

In other words, at the closest point, the normal directions to the level curves are parallel, so that ∇f and ∇g are parallel, that is, there is a scalar λ, called a multiplier, such that

$$\nabla f = \lambda \nabla g.$$

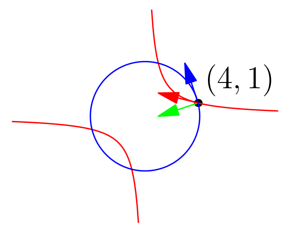


FIGURE 2. Tangent vectors are not parallel: we can do better

This suggests a strategy to solve optimisation problems. Introduce a new variable λ. We want x, y and λ such that

$$\begin{aligned} f_x &= \lambda g_x \\ f_y &= \lambda g_y \\ g &= c. \end{aligned}$$

In our case this means

$$\begin{aligned} 2x &= \lambda y \\ 2y &= \lambda x \\ xy &= 4. \end{aligned}$$

Rearranging, this gives

$$\begin{aligned} 2x - \lambda y &= 0 \\ -\lambda x + 2y &= 0. \end{aligned}$$

A homogeneous system of linear equations. Either $(x, y) = (0, 0)$, which is not possible since $xy = 4$. Or the determinant is zero,

$$0 = \begin{vmatrix} 2 & -\lambda \\ -\lambda & 4 \end{vmatrix} = 4 - \lambda^2.$$

So $\lambda = \pm 2$. If $\lambda = -2$ then x and y have different sign, impossible. So $\lambda = 2$ is the only solution, in which case $x = y$. But then $x^2 = 4$ and $x = y = 2$ (or -2).

Why does this work? Well, if we are at a maximum of f subject to the constraint $g = c$, then if we move in any direction $\hat{u}$ in the surface $g = c$, we must have

$$0 = \left.\frac{df}{ds}\right|_{\hat{u}} = \nabla f \cdot \hat{u}.$$

It follows that ∇f is orthogonal to the level surface $g = c$. As ∇g also has this property, ∇f and ∇g are parallel, so that there is a multiplier.

Warning: The second derivative test won't work if one uses Lagrange multipliers. You just have to look at the critical points and compare them to see which is the maximum and which is the minimum.

Example 13.2. *What are the dimensions of a box with largest volume if the total surface area is* 64*?*

Let x, y and z be the sides of the box. The surface area of the box is

$$2(xy + xz + yz).$$

So we want to

$$\text{maximise} \quad xyz \qquad \text{subject to} \qquad xy + xz + yz = 32.$$

Let's first try to solve this problem without Lagrange multipliers. Use the constraint to express z as a function of x and y,

$$z = \frac{32 - xy}{x + y}.$$

Plug this back into the volume,

$$\text{maximise} \quad \frac{xy(32 - xy)}{x + y}.$$

The next step would be to find the partials with respect to x and y. A mess!

Instead, let's use Lagrange multipliers. Introduce λ. We want to solve $\nabla f = \lambda \nabla g$, so that

$$\begin{aligned} yz &= \lambda(y + z) \\ xz &= \lambda(x + z) \\ yz &= \lambda(x + y) \\ xy + xz + yz &= 32. \end{aligned}$$

Multiply the first equation by x, the second by y and the third by z.

$$\begin{aligned} xyz &= \lambda x(y + z) \\ xyz &= \lambda y(x + z) \\ xyz &= \lambda z(x + y) \\ xy + xz + yz &= 32. \end{aligned}$$

So

$$\lambda x(y + z) = \lambda y(x + z).$$

Suppose that $\lambda = 0$. Then $yz = 0$ so that one of y or z is zero. Similarly one of x and y is zero and one of x and z is zero. So two out of three of x, y and z would be zero, impossible.

If $\lambda \neq 0$, we get

$$xy + xz = yx + yz.$$

So

$$xz = yz.$$

Again, $z \neq 0$, so $x = y$. By symmetry, $x = y = z$. Hence

$$3x^2 = 32.$$

Hence

$$x = \frac{\sqrt{32}}{\sqrt{3}}.$$

Let's check that a cube is a maximum (and not a minimum or a saddle point). We cannot use the 2nd derivative test. We just have to think about it. What happens if we increase x? Since the surface area is constant, y and z must decrease to zero and the volume goes to zero. Similarly as we increase either y or z. Since we approach zero as we go to the boundary, the cube has to correspond to a maximum.

14. Non-independent variables

Consider an ideal gas, which has a pressure P, a volume V and a temperature T. These variables are not independent, they are related by

$$PV = kT \qquad \text{where} \qquad k = nR.$$

Here n and R are constants, which depend on the amount and type of gas.

Many functions depend on all three variables, and finding out how the function varies as one changes the variables is complicated. Let's consider an easy example.

Question 14.1. *Suppose we have a right-angled triangle with sides x, y and hypotenuse z. Then*

$$x^2 + y^2 = z^2,$$

so x, y and z are not independent. Let $P = x + y + z$ the length of the perimeter. What is

$$\frac{\partial P}{\partial x}?$$

There are (at least) four possible answers.

(1) 1, since y and z are constant.

(2) Rewrite as
$$x + y + \sqrt{x^2 + y^2},$$
and take the partial with respect to x (fixing y).

(3) Rewrite as
$$x + \sqrt{z^2 - x^2} + z,$$
and take the partial with respect to x (fixing z).

(4) The question is ambiguous.

The answer cannot be one, since if we vary x, one of at least y or z has to vary. We have to choose to hold either y or z constant.

If we fix y and increase x, then we get wider triangles of the same height. If we fix z and increase x, then our triangle is like a ladder sliding down the wall. So the answer is (4), the question is ambiguous.

To make the question unambiguous, we have to specify what variables we are holding constant. New notation:

$$\left(\frac{\partial P}{\partial x}\right)_y.$$

This means, vary x whilst holding y constant. Implicitly this means we treat z as a function of x and y.

There are three methods to compute this derivative. Perhaps the best is just to use to the total differential:

$$\mathrm{d}P = P_x\,\mathrm{d}x + P_y\,\mathrm{d}y + P_z\,\mathrm{d}z.$$

For us, this means

$$\mathrm{d}P = \mathrm{d}x + \mathrm{d}y + \mathrm{d}z.$$

Now let's differentiate the relation:

$$2x\,\mathrm{d}x + 2y\,\mathrm{d}y = 2z\,\mathrm{d}z.$$

Solving for $\mathrm{d}z$,

$$\mathrm{d}z = \frac{x}{z}\,\mathrm{d}x + \frac{y}{z}\,\mathrm{d}y.$$

Substitute this back into the expression for dP,

$$\begin{aligned}
\mathrm{d}P &= \mathrm{d}x + \mathrm{d}y + \mathrm{d}z \\
&= \mathrm{d}x + \mathrm{d}y + \left(\frac{x}{z}\,\mathrm{d}x + \frac{y}{z}\,\mathrm{d}y\right) \\
&= (1 + \frac{x}{z})\,\mathrm{d}x + (1 + \frac{y}{z})\,\mathrm{d}y.
\end{aligned}$$

So

$$\left(\frac{\partial P}{\partial x}\right)_y = 1 + \frac{x}{z}.$$

Note that we also computed

$$\left(\frac{\partial P}{\partial y}\right)_x = 1 + \frac{y}{z},$$

at the same time.

There are two other methods to compute the derivative. We could eliminate z

$$x + y + \sqrt{x^2 + y^2},$$

and differentiate with respect to x, fixing y,

$$1 + \frac{x}{\sqrt{x^2 + y^2}}.$$

Finally, we could use the chain rule.

$$P = x + y + z.$$

Differentiate both sides with respect to x, fixing y,

$$\left(\frac{\partial P}{\partial x}\right)_y = 1 + \left(\frac{\partial z}{\partial x}\right)_y.$$

So we need to calculate the last term. We know

$$x^2 + y^2 = z^2.$$

Differentiate both sides with respect to x, fixing y,

$$2x = 2z \left(\frac{\partial z}{\partial x} \right)_y .$$

Solving, we get

$$\left(\frac{\partial z}{\partial x} \right)_y = \frac{x}{z}.$$

Putting all of this together, we get

$$\left(\frac{\partial P}{\partial x} \right)_y = 1 + \frac{x}{z}.$$

Example 14.2. *Suppose we have a triangle, sides a, b and angle θ between the sides. The area of the triangle is*

$$A = \frac{1}{2} ab \sin \theta.$$

Suppose also that the triangle is a right-angled triangle, with b the hypotenuse, so that

$$a = b \cos \theta.$$

What is

$$\frac{\partial A}{\partial \theta}?$$

Once again this question is ambiguous. If we forget that the triangle is right-angled, we get

$$\frac{\partial A}{\partial \theta} = \frac{1}{2} ab \cos \theta.$$

But now suppose we vary θ keeping a right-angled triangle. There are two quantities we could compute

$$\left(\frac{\partial A}{\partial \theta} \right)_a .$$

We keep a right-angled triangle and fix one side. Or we could compute

$$\left(\frac{\partial A}{\partial \theta} \right)_b .$$

We keep a right-angled triangle and fix the hypotenuse.

Just to practice, let's compute the first expression. We first use the method of differentials.

$$\begin{aligned} \mathrm{d}A &= A_a \,\mathrm{d}a + A_b \,\mathrm{d}b + A_\theta \,\mathrm{d}\theta \\ &= \frac{1}{2} b \sin \theta \,\mathrm{d}a + \frac{1}{2} a \sin \theta \,\mathrm{d}b + \frac{1}{2} ab \cos \theta \,\mathrm{d}\theta. \end{aligned}$$

If we differentiate the relation, we get

$$\mathrm{d}a = \cos\theta\,\mathrm{d}b - b\sin\theta\,\mathrm{d}\theta.$$

Now we want to view A as a function of θ and a. So we want to get rid of $\mathrm{d}b$,

$$\mathrm{d}b = \sec\theta\,\mathrm{d}a + b\tan\theta\,\mathrm{d}\theta.$$

So

$$\begin{aligned}
\mathrm{d}A &= \frac{1}{2}b\sin\theta\,\mathrm{d}a + \frac{1}{2}a\sin\theta\,\mathrm{d}b + \frac{1}{2}ab\cos\theta\,\mathrm{d}\theta.\\
&= \frac{1}{2}b\sin\theta\,\mathrm{d}a + \frac{1}{2}a\sin\theta\,(\sec\theta\,\mathrm{d}a + b\tan\theta\,\mathrm{d}\theta) + \frac{1}{2}ab\cos\theta\,\mathrm{d}\theta.\\
&= \frac{1}{2}(a\tan\theta + b\sin\theta)\,\mathrm{d}a + \frac{1}{2}ab(\cos\theta + \sin\theta\tan\theta)\,\mathrm{d}\theta.
\end{aligned}$$

Hence

$$\left(\frac{\partial A}{\partial\theta}\right)_a = \frac{1}{2}ab(\cos\theta + \sin\theta\tan\theta).$$

Next, let's try substitution.

$$b = a\sec\theta.$$

So,

$$A = \frac{1}{2}a^2\tan\theta.$$

Hence

$$\left(\frac{\partial A}{\partial\theta}\right)_a = \frac{1}{2}a^2\sec^2\theta.$$

Let's check the two answers are compatible.

$$\frac{1}{2}ab(\cos\theta + \sin\theta\tan\theta) = \frac{1}{2}a^2(1+\tan^2\theta) = \frac{1}{2}a^2\sec^2\theta.$$

Finally, let's use the chain rule.

$$\left(\frac{\partial A}{\partial\theta}\right)_a = \frac{1}{2}a\left(\frac{\partial b}{\partial\theta}\right)_a\sin\theta + \frac{1}{2}ab\cos\theta.$$

Here we applied the product rule to the product $b\sin\theta$, treating $a/2$ as a constant. So we need to compute the derivative on the RHS. We differentiate the relation,

$$0 = \left(\frac{\partial b}{\partial\theta}\right)_a\cos\theta - b\sin\theta,$$

where again we applied the product rule. It follows that

$$\left(\frac{\partial b}{\partial\theta}\right)_a = b\tan\theta.$$

Putting all of this together, we get

$$\left(\frac{\partial A}{\partial \theta}\right)_a = \frac{1}{2}ab\tan\theta\sin\theta + \frac{1}{2}ab\cos\theta.$$

15. Partial differential equations; double integrals

15.1. **Partial differential equations.** Recall that many functions of one variable are characterised by a(n ordinary) differential equation.

$$\frac{dy}{dx} = ky.$$

The solutions to this ODE are

$$y(x) = ae^{kx},$$

where a is a constant.

A partial differential equation is an equation satisfied by a function of several variables; a preposterously large number of problems in nature are described by PDEs.

Heat equation: The heat in a sheet of metal is described by $w(x, y, t)$ where

$$\frac{\partial w}{\partial t} = k\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right).$$

The term on the right has special significance:

Laplace equation: Let $u(x, y)$ be a function in the plane.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

This is a solution to the heat equation which is static in time.

Wave equation: Let $w(x, t)$ represent the displacement of a guitar string, as a function of time:

$$\frac{\partial^2 w}{\partial t^2} = c^2 \frac{\partial^2 w}{\partial x^2}.$$

For example, consider $w(x, t) = \sin(\omega t + kx)$, where ω and k are constants.

Question 15.1. *For which values of ω and k is w a solution of the wave equation?*

We plug in the formula for w and see what we get:

$$\frac{\partial w}{\partial t} = \omega \cos(\omega t + kx),$$

so that

$$\frac{\partial^2 w}{\partial t^2} = -\omega^2 \sin(\omega t + kx).$$

On the other hand,

$$\frac{\partial w}{\partial x} = k \cos(\omega t + kx),$$

so that

$$\frac{\partial^2 w}{\partial x^2} = -k^2 \sin(\omega t + kx).$$

So, if $\omega^2 = c^2 k^2$, we have a solution to the wave equation.

15.2. **Double integrals.** Suppose we have a region R in the plane and a function f defined on R. The double integral is the volume between the graph of f and the region R in the xy-plane:

$$\iint_R f(x, y)\, \mathrm{d}A.$$

As usual this is the signed volume. To compute this integral, imagine cutting the region R into small pieces R_i. Pick a point (x_i, y_i) belonging to each piece. Then the volume is approximately the sum

$$\sum_i f(x_i, y_i)\Delta A_i,$$

where ΔA_i is the area of the region R_i. Taking the limit, as the area of each piece goes to zero, we get the volume.

How do we compute the double integral? First we imagine dividing the region into small rectangles. The area of each rectangle is

$$\Delta A_i = \Delta x_i \Delta y_i.$$

Summing first over y and then x, we can compute the area by first integrating over y and then x.

Example 15.2. *Compute the volume of $f(x, y) = 2x + 3y - 1$ over the rectangle $1 \leq x \leq 3$, $1 \leq y \leq 2$.*

$$\iint_R (2x + 3y - 1)\, \mathrm{d}A = \int_{x=1}^{x=3} \int_{y=1}^{y=2} (2x + 3y - 1)\, \mathrm{d}y\mathrm{d}x.$$

First, we compute the inner integral

$$\int_{y=1}^{y=2} (2x+3y-1)\, \mathrm{d}y = \Big[2xy + 3y^2/2 - y \Big]_1^2 = (4x+6-2) - (2x+3/2-1) = 2x + 7/2.$$

Now compute the outer integral

$$\int_{x=1}^{x=3} 2x + 9/2\, \mathrm{d}x = \Big[x^2 + 7x/2 \Big]_1^3 = (9 + 21/2) - (1 + 7/2) = 15.$$

Note that we can switch the order of integration. If we first sum over x and then y, then we first integrate over x and then over y. Of course we get the same answer:

$$\iint_R (2x+3y-1)\,\mathrm{d}A = \int_{y=1}^{y=2}\int_{x=1}^{x=3} (2x+3y-1)\,\mathrm{d}x\mathrm{d}y.$$

Compute the inner integral

$$\int_{x=1}^{x=3} (2x+3y-1)\,\mathrm{d}x = \Big[x^2+3yx-x\Big]_1^3 = (9+9y-3)-(1+3y-1) = 6+6y.$$

Now compute the outer integral

$$\int_{y=1}^{y=2} 6+6y\,\mathrm{d}y = \Big[6y+3y^2\Big]_1^2 = 12+12-6-3 = 15.$$

What happens when we want to integrate over a region which is not a rectangle?

Example 15.3. *Compute the volume of $f(x,y) = x+3y$ over the circle $x^2+y^2 \le 1$.*

The outer limits for x are relatively straightforward, the largest value of x is 1 and the smallest -1. But the limits for y are dependent on x. Imagine fixing a value for x; we get a vertical line. The lower limit for y represents where this line first meets the region. The upper limit for y represents where this line leaves the region.

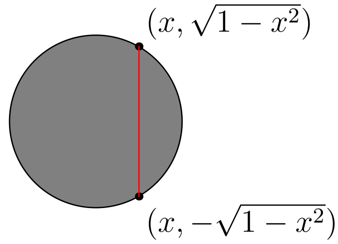


FIGURE 1. Limits for y

$$\iint_R (x+3y)\,\mathrm{d}A = \int_{x=-1}^{x=1}\int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} (x+3y)\,\mathrm{d}y\mathrm{d}x.$$

Inner integral:

$$\int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} (2x+3y)\,\mathrm{d}y = \Big[xy+3y^2/2\Big]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} = 2x\sqrt{1-x^2}.$$

Outer integral:

$$\int_{x=-1}^{x=1} 2x\sqrt{1-x^2}\,\mathrm{d}x = \Big[-2/3(1-x^2)^{3/2}\Big]_{-1}^{1} = 0.$$

Notice that in retrospect we could have predicted that the integral is zero. The function $f(x,y)$ is a sum of x and $3y$. x is odd with respect to x and y is odd with respect to y, so both pieces integrate to zero.

Example 15.4. *Compute*

$$\int_0^1 \int_x^1 e^{y^2}\,\mathrm{d}y\mathrm{d}x.$$

Note that we cannot compute the inner integral,

$$\int_x^1 e^{y^2}\,\mathrm{d}y.$$

since there is no way to integrate e^{y^2}. Instead, let's switch the order of integration. To do this, we first have to determine the region we are integrating over. x ranges from 0 to 1. For a given value of x, y ranges from x to 1. So the region of integration is the triangle $x \geq 0$, $y \leq 1$ and $y \geq x$.

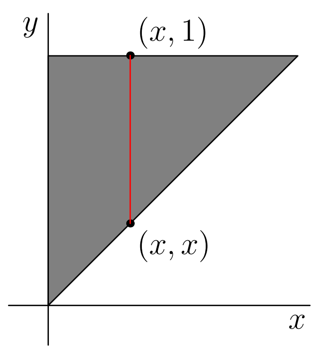


FIGURE 2. Region of integration

So we have

$$\int_0^1 \int_x^1 e^{y^2}\,\mathrm{d}y\mathrm{d}x = \int_0^1 \int_0^y e^{y^2}\,\mathrm{d}x\mathrm{d}y.$$

The inner integral is

$$\int_0^y e^{y^2}\,\mathrm{d}x = \Big[xe^{y^2}\Big]_0^y = ye^{y^2}.$$

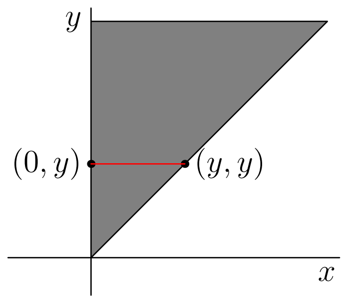


FIGURE 3. Limits of integral

The outer integral is

$$\int_0^1 ye^{y^2}\,\mathrm{d}y = \left[1/2e^{y^2}\right]_0^1 = 1/2(e-1).$$

16. Polar coordinates and applications

Let's suppose that either the integrand or the region of integration comes out simpler in polar coordinates ($x = r\cos\theta$ and $y = r\sin\theta$). Let suppose we have a small change in r and θ. The small change Δr in r gives us two concentric circles and the small change $\Delta\theta$ in θ gives us an angular wedge.

If the changes are small, we almost get a rectangle with sides Δr and $r\Delta\theta$,

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

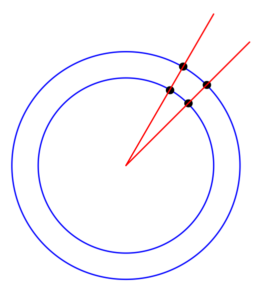


Figure 1. Small changes in r and θ

Taking the limit as Δr and $\Delta\theta$ go to zero, we get

$$\mathrm{d}A = r\mathrm{d}r\mathrm{d}\theta.$$

Example 16.1. *Compute the volume of $f(x, y) = x+3y$ over the circle $x^2 + y^2 \leq 1$.*

$$\iint_R (x + 3y)\,\mathrm{d}A = \int_0^{2\pi}\int_0^1 r^2(\cos\theta + 3\sin\theta)\,\mathrm{d}r\mathrm{d}\theta.$$

The inner integral is

$$\int_0^1 r^2(\cos\theta + 3\sin\theta)\,\mathrm{d}r = \left[\frac{r^3}{3}(\cos\theta + 3\sin\theta)\right]_0^1 = \frac{1}{3}\cos\theta + \sin\theta.$$

The outer integal is

$$\int_0^{2\pi} \frac{1}{3}\cos\theta + \sin\theta\,\mathrm{d}\theta = \left[\frac{1}{3}\sin\theta - \cos\theta\right]_0^{2\pi} = 0.$$

Example 16.2. *What is*

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x?$$

Let

$$I = \int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x.$$

Then

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\,\mathrm{d}y\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2}\,\mathrm{d}x\mathrm{d}y \\ &= \iint_{\mathbb{R}^2} e^{-x^2-y^2}\,\mathrm{d}A \\ &= \int_0^{2\pi}\int_0^{\infty} re^{-r^2}\,\mathrm{d}r\mathrm{d}\theta. \end{aligned}$$

The inner integral is

$$\int_0^{\infty} re^{-r^2}\,\mathrm{d}r = \left[-\frac{1}{2}e^{-r^2}\right]_0^{\infty} = \frac{1}{2}.$$

The outer integral is

$$\int_0^{2\pi} \frac{1}{2}\,\mathrm{d}\theta = \left[\frac{\theta}{2}\right]_0^{2\pi} = \pi.$$

So $I = \sqrt{\pi}$.

We can use double integrals to compute some interesting quantities. If we want to compute the area of a region, then just integrate 1,

$$\iint_R 1\,\mathrm{d}A.$$

The point is that the volume under the graph of a function of constant height is the area of the base times the height.

If we have a material whose mass density,

$$\delta(x,y) = \lim \frac{\Delta m}{\Delta A},$$

is a function of the position, then the total mass is

$$M = \iint_R \delta\,\mathrm{d}A.$$

Recall that the average of a function $f(x,y)$ over the region R is

$$\bar{f} = \frac{1}{A}\iint_R f\,\mathrm{d}A,$$

where A is the area of R. The centre of mass or centroid $(\bar{x}, \bar{y})$ of a plate with density δ is given by

$$\bar{x} = \frac{1}{M}\iint_R x\delta \,\mathrm{d}A \qquad \text{and} \qquad \bar{y} = \frac{1}{M}\iint_R y\delta \,\mathrm{d}A$$

The moment of inertia is a measure of how hard it is to rotate an object (in just the same way that the mass measures how hard it is to move an object). Suppose we have a mass m rotating in a circle of radius r with angular speed

$$\omega = \frac{d\theta}{dt}.$$

The velocity is $v = r\omega$. So the kinetic engergy is

$$\frac{1}{2}mv^2 = \frac{1}{2}mr^2\omega^2.$$

The moment of inertia is

$$I_0 = mr^2.$$

For a solid with density δ, the moment of inertia about the origin is

$$I_0 = \iint_R r^2\delta \,\mathrm{d}A.$$

The moment of inertia about the x-axis is

$$I_0 = \iint_R y^2\delta \,\mathrm{d}A.$$

Here y^2 is the square of the distance to the x-axis. In fact, one can think of the moment of inertia about the origin as the moment of inertia about the z-axis in $\mathbb{R}^3$. For a point in the plane, the square of the distance to the z-axis is r^2.

Example 16.3. *What is the moment of inertia of a disc of radius a about its centre?*

Here we assume that the density is one. Put the disc in the plane centred at the origin. Let R be the circle $x^2 + y^2 \leq a^2$. The moment of inertia is

$$\iint_R r^2 \,\mathrm{d}A = \int_0^{2\pi}\int_0^a r^3 \,\mathrm{d}r\mathrm{d}\theta.$$

The inner integral is

$$\int_0^a r^3 \,\mathrm{d}r = \left[\frac{r^4}{4}\right]_0^a = \frac{a^4}{4}.$$

The outer integral is

$$\int_0^{2\pi} \frac{a^4}{4}\,\mathrm{d}\theta = \left[\frac{a^4}{4}\theta\right]_0^{2\pi} = \frac{\pi a^4}{2}.$$

Now what happens if we try to compute the moment of inertia about a point of the circumference? Put the circle so its centre is at $(a, 0)$, so that the origin is the point on the circumference. Clearly we would like to use polar coordinates again.

The equation of the circle in Cartesian coordinates is

$$(x-a)^2 + y^2 = a^2.$$

Expanding we get

$$x^2 + y^2 = 2ax.$$

So

$$r^2 = 2ar\cos\theta,$$

that is

$$r = 2a\cos\theta.$$

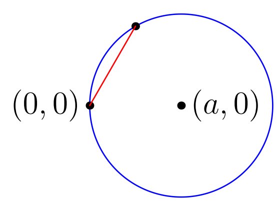


FIGURE 2. Limits of integration

The moment of inertia about the origin is therefore

$$I_0 = \iint_R r^2\,\mathrm{d}A = \int_0^{\pi}\int_0^{2a\cos\theta} r^3\,\mathrm{d}r\mathrm{d}\theta.$$

The inner integral is

$$\int_0^{2a\cos\theta} r^3\,\mathrm{d}r = \left[\frac{1}{4}r^4\right]_0^{2a\cos\theta} = 4a^4\cos^4\theta.$$

The outer integral is

$$\int_0^{\pi} 4a^4\cos^4\theta\,\mathrm{d}\theta = \frac{3}{2}\pi a^4.$$

17. Review

Linear approximation:

$$\Delta f \approx f_x \Delta x + f_y \Delta y.$$

Tangent plane: to $z = f(x, y)$ at (x_0, y_0, z_0)

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

Let $w = f(x, y, z)$. Chain rule:

$$\mathrm{d}w = f_x \, \mathrm{d}x + f_y \, \mathrm{d}y + f_z \, \mathrm{d}z.$$

So

$$\frac{dw}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}$$

We can encode this efficiently using the gradient:

$$\nabla f = \langle f_x, f_y, f_z \rangle = f_x \hat{\imath} + f_y \hat{\jmath} + f_z \hat{k},$$

Then

$$\frac{dw}{dt} = \nabla f \cdot \vec{v}(t).$$

The most important property of the gradient is that it is normal to the level curves, or to the level surfaces.

Example 17.1. *What is the tangent plane to the ellipsoid*

$$3x^2 + 5y^2 + 3z^2 = 11,$$

at the point $(x_0, y_0, z_0) = (1, 1, 1)$*?*

Well, this is a level surface of the function $f(x, y, z) = 3x^2 + 5y^2 + 3z^2$.

$$\nabla f = \langle 6x, 10y, 6z \rangle.$$

At the point $(1, 1, 1)$, we have

$$\nabla f = \langle 6, 10, 6 \rangle.$$

So $\vec{n} = \langle 3, 5, 3 \rangle$ is a normal vector the tangent plane. So the equation of the tangent plane is

$$\langle x-1, y-1, z-1 \rangle \cdot \langle 3, 5, 3 \rangle = 0 \qquad \text{so that} \qquad 3(x-1)+5(y-1)+3(z-1) = 0.$$

Rearranging, we get $3x + 5y + 3z = 11$.

Directional derivative: Let $w = f(x, y)$ be a function of two variables. Let $\hat{u} = \langle a, b \rangle$ be a direction in the plane. The directional derivative, in the direction of $\hat{u}$,

$$\left. \frac{dw}{ds} \right|_{\hat{u}} = \lim_{s \to 0} \frac{f(x_0 + sa, y_0 + sb) - f(x_0, y_0)}{s}.$$

If $\hat{u} = \hat{\imath}$, we get $f_x(x_0, y_0)$ and if $\hat{u} = \hat{\jmath}$, we get $f_y(x_0, y_0)$.

To compute, use the gradient:

$$\left.\frac{dw}{ds}\right|_{\hat{u}} = \nabla f \cdot \hat{u}.$$

So the gradient points in the direction of maximum increase of w and the magnitude of the gradient is the rate of change in this direction. The direction of maximum decrease of f is given by $-\nabla f$.

Example 17.2. *What is the closest point to $p = (1, -1)$ on the curve $x^3 - x + 2y^2 = 1.9$?*

At $(1, -1)$ we have $f(1, -1) = 2$, so we want $\Delta f = -0.1$. From p we should go in the direction to decrease f the most:

$$\nabla f = \langle 3x^2 - 1, 4y \rangle \qquad \text{so that} \qquad \nabla f_{(1,-1)} = \langle 2, -4 \rangle.$$

We want go in the direction of $-\nabla f_{(1,-1)} = \langle -2, 4 \rangle$. The magnitude is $2\sqrt{5}$, so want to go in the direction

$$\hat{u} = \frac{1}{\sqrt{5}} \langle -1, 2 \rangle.$$

If we go in this direction f decreases by $2\sqrt{5}$. So we want to go a distance of

$$\frac{1}{20\sqrt{5}}.$$

That is we want a displacement of

$$\frac{1}{100} \langle -1, 2 \rangle.$$

So we want the point

$$\langle 1, -1 \rangle + \frac{1}{100} \langle -1, 2 \rangle = \langle 0.99, -0.98 \rangle$$

To find the maximum and the minimum of a function $w = f(x, y)$, first find the critical points, the solutions to $f_x = 0$ and $f_y = 0$. To analyse the type of the critical points (local minimum, local maximum or saddle point), use the 2nd derivative test. Let $A = f_{xx}(x_0, y_0)$, $B = f_{xy}(x_0, y_0)$ and $C = f_{yy}(x_0, y_0)$. If $AC - B^2 > 0$ we have a maximum or minimum. $A > 0$ is a minimum and $A < 0$ is a maximum. If $AC - B^2 < 0$ we have a saddle point.

Next check what happens at the boundary, including infinity.

Example 17.3.

maximise and minimise $\qquad x+y+z \qquad$ *subject to* $\qquad x^2y^3z^5 = 2^2 3^3 5^5,$

where x, y and $z \geq 0$.

Use equation to eliminate x,

$$x = \sqrt{\frac{2^2 3^3 5^5}{y^3 z^5}}.$$

So we want to maximise

$$h(y,z) = \sqrt{\frac{2^2 3^3 5^5}{y^3 z^5}} + y + z.$$

Find the critical points:

$$h_y = -\frac{3}{2}\sqrt{\frac{2^2 3^3 5^5}{y^5 z^5}} + 1 \qquad \text{and} \qquad h_z = -\frac{5}{2}\sqrt{\frac{2^2 3^3 5^5}{y^3 z^7}} + 1.$$

So we want

$$0 = -\frac{3}{2}\sqrt{\frac{2^2 3^3 5^5}{y^5 z^5}} + 1 \qquad \text{and} \qquad 0 = -\frac{5}{2}\sqrt{\frac{2^2 3^3 5^5}{y^3 z^7}} + 1.$$

Rearranging, we get

$$\sqrt{\frac{2^2 3^3 5^5}{y^5 z^5}} = \frac{2}{3} \qquad \text{and} \qquad \sqrt{\frac{2^2 3^3 5^5}{y^3 z^7}} = \frac{2}{5}.$$

Squaring, we get

$$\frac{2^2 3^3 5^5}{y^5 z^5} = \frac{2^2}{3^2} \qquad \text{and} \qquad \frac{2^2 3^3 5^5}{y^5 z^7} = \frac{2^2}{5^2}.$$

Taking the reciprocal

$$\frac{y^5 z^5}{2^2 3^3 5^5} = \frac{3^2}{2^2} \qquad \text{and} \qquad \frac{y^5 z^7}{2^2 3^3 5^5} = \frac{5^2}{2^2}.$$

Simplifying

$$y^5 z^5 = 3^5 5^5 \qquad \text{and} \qquad y^3 z^7 = 3^3 5^7.$$

We guess $y = 3$ and $z = 5$. This works and it is clear the solution is unique. $x = 2$ is the other value. Let's try the 2nd derivative test.

$$h_{yy} = \frac{15}{4}\sqrt{\frac{2^2 3^3 5^5}{y^7 z^5}} \qquad h_{yz} = \frac{15}{4}\sqrt{\frac{2^2 3^3 5^5}{y^5 z^7}} \qquad \text{and} \qquad h_{zz} = \frac{35}{4}\sqrt{\frac{2^2 3^3 5^5}{y^3 z^9}}.$$

We have

$$A = \frac{15}{4}\sqrt{\frac{2^2 3^3 5^5}{3^7 5^5}} \qquad B = \frac{15}{4}\sqrt{\frac{2^2 3^3 5^5}{3^5 5^7}} \qquad \text{and} \qquad C = \frac{35}{4}\sqrt{\frac{2^2 3^3 5^5}{3^3 5^9}}.$$

We have $AC - B^2 > 0$. $A > 0$, so have a local minimum. There are no other critical points, so this is a global minimum. Minimum value is 10.

At the boundary, one of the variables goes to ∞ and the sum goes to ∞. No maximum.

Let's use Lagrange multipliers insead. We add a variable λ and solve

$$\begin{aligned} f_x &= \lambda g_x \\ f_y &= \lambda g_y \\ f_z &= \lambda g_z \\ g &= c. \end{aligned}$$

In our case $f(x, y, z) = x + y + z$ and $g(x, y, z) = x^2y^3z^5$. We get

$$\begin{aligned} 1 &= \lambda 2xy^3z^5 \\ 1 &= \lambda 3x^2y^2z^5 \\ 1 &= \lambda 5x^2y^3z^4 \\ x^2y^3z^5 &= 2^2 3^3 5^5. \end{aligned}$$

Multiply the first three equations by x, y and z:

$$\begin{aligned} x &= \lambda 2x^2y^3z^5 \\ y &= \lambda 3x^2y^3z^5 \\ z &= \lambda 5x^2y^3z^5. \end{aligned}$$

So $3x = 2y$, $5x = 2z$. Multiply constraint by 2^3,

$$3^3x^5z^5 = 2^5 3^3 5^5.$$

Cancelling, we get

$$x^5z^5 = 2^5 5^5.$$

Multiply both sides by 2^5,

$$2^5x^5z^5 = 2^{10}5^5.$$

We get

$$5^5x^{10} = 2^{10}5^5.$$

Hence $x = 2$. Thus $y = 3$ and $z = 5$.

18. Change of variables

Question 18.1. *What is the area of the ellipse*

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1?$$

The area is

$$\begin{aligned}\iint_R 1\,\mathrm{d}A &= \iint_{\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2\leq 1} 1\,\mathrm{d}x\mathrm{d}y \\ &= \iint_{u^2+v^2\leq 1} ab\,\mathrm{d}u\mathrm{d}v \\ &= \pi ab.\end{aligned}$$

Here we changed variable from x and y to $u = x/a$ and $v = y/b$. We have

$$\mathrm{d}u = \frac{\mathrm{d}x}{a} \qquad \text{and} \qquad \mathrm{d}v = \frac{\mathrm{d}y}{b}.$$

It follows that

$$\mathrm{d}u\,\mathrm{d}v = \frac{1}{ab}\mathrm{d}x\,\mathrm{d}y.$$

How about if the change of variables is more complicated?

To warm up, let's consider a linear transformation.

$$\begin{aligned}u &= 2x - y \\ v &= x + y.\end{aligned}$$

In this case, a rectangle in the xy-plane gets mapped to a parallelogram. In terms of matrices,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

It follows that the square given by $\hat{\imath}$ and $\hat{\jmath}$ gets mapped to the parallelogram with sides $2\hat{\imath} + \hat{\jmath} = \langle 2, 1\rangle$ and $-\hat{\imath} + \hat{\jmath} = \langle -1, 1\rangle$. The area of this parallelogram is the absolute value of the determinant:

$$\begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 3.$$

So

$$\mathrm{d}u\,\mathrm{d}v = 3\mathrm{d}x\,\mathrm{d}y.$$

(Since the map is linear, every rectangle gets rescaled by the same factor of 3).

In general, by the approximation formula,

$$\begin{aligned}\Delta u &\approx u_x\Delta x + u_y\Delta y \\ \Delta v &\approx v_x\Delta x + v_y\Delta y.\end{aligned}$$

In terms of matrices

$$\begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} \approx \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}.$$

Then a small rectangle in the xy-plane gets mapped approximately to a parallelogram of area the absolute value of the determinant

$$\begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}.$$

The determinant is called the Jacobian,

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

Taking the limit as Δx and Δy go to zero, we get

$$\mathrm{d}u\,\mathrm{d}v = |J|\mathrm{d}x\,\mathrm{d}y.$$

Note that we take the absolute value, as area is always positive.

Let's see what happens if we go from Cartesian coordinates to polar.

$$x = r\cos\theta$$
$$y = r\sin\theta.$$

The determinant is

$$\begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r,$$

so that

$$\mathrm{d}x\,\mathrm{d}y = r\mathrm{d}r\,\mathrm{d}\theta,$$

as expected.

Question 18.2. *Let R be the square with vertices $(\pm 1, 0)$, $(0, \pm 1)$. What is*

$$\iint_R \frac{\sin^2(x-y)}{x+y+2}\,\mathrm{d}x\mathrm{d}y?$$

Let's change coordinates to $u = x - y$ and $v = x + y$. Note that this has two benefits. The integrand simplifies and the sides of the square are given by u or v constant. The side from $(0, 1)$ to $(1, 0)$ corresponds to $v = 1$. u ranges from -1 to 1. Similarly the four sides are $u = \pm 1$ and $v = \pm 1$. The Jacobian is

$$J = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2.$$

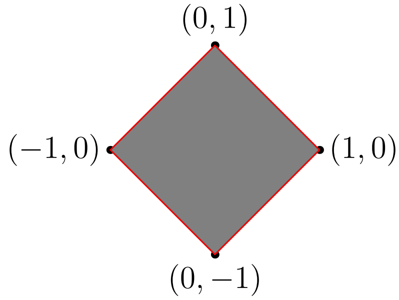


FIGURE 1. Region of integration

So
$$\mathrm{d}u\,\mathrm{d}v = 2\mathrm{d}x\,\mathrm{d}y.$$
$$\iint_R \frac{\sin^2(x-y)}{x+y+2}\,\mathrm{d}x\mathrm{d}y = \int_{-1}^{1}\int_{-1}^{1} \frac{1}{2}\frac{\sin^2 u}{v+2}\,\mathrm{d}u\mathrm{d}v.$$
It is then straightforward to finish off.

One more example. Let's compute
$$\int_0^1\int_0^1 x^2 y\,\mathrm{d}x\mathrm{d}y,$$
by using the change of variable $u = x$ and $v = xy$. The Jacobian is the absolute value of
$$\begin{vmatrix} 1 & 0 \\ y & x \end{vmatrix} = x.$$
Note that x is positive over the square, so no need to take the absolute value.
$$x^2 y\,\mathrm{d}x\mathrm{d}y = x^2 y \frac{1}{x}\,\mathrm{d}u\mathrm{d}v = v\,\mathrm{d}u\mathrm{d}v.$$
Now we figure out the range of integration. First the outer limits. What is the maximum value of v over the square? Well 1, achieved at the point $(1,1)$. And the minimum value is 0, achieved at $(0,0)$. So v ranges from 0 to 1. What about u? Well if we fix a value of v, we get a hyperbola. The maximum value of $u = x$ is always 1. We have $xy = v$. The minimum value is when $x = v$.

So the integral in uv-coordinates is
$$\int_0^1\int_0^1 x^2 y\,\mathrm{d}x\mathrm{d}y = \int_0^1\int_v^1 v\,\mathrm{d}u\mathrm{d}v.$$

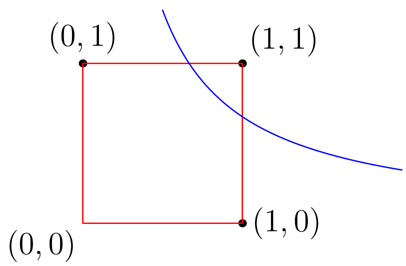


FIGURE 2. Limits for $u = x$

19. Vector fields

Suppose we have a function

$$\vec{F} = M\hat{\imath} + N\hat{\jmath},$$

which assigns to every point of the plane a vector. Here $M = M(x, y)$ and $N(x, y)$ are scalar functions, which are the components of $\vec{F}$. Such a function $\vec{F}$ is called a vector field.

Vector fields appear in many different guises. If you look at the flow of water in a river, every point of the river has a velocity vector. When the wind blows, it blows in different directions and different speeds at every point. Force often forms a vector field. Gravitation always pulls you to the centre of the earth, inversely proportional to the square of the distance to the origin.

It is interesting to draw pictures of vector fields;

(1) $\vec{F} = \hat{\imath} - 3\hat{\jmath}$.
(2) $\vec{F} = x\hat{\imath}$.
(3) $\vec{F} = x\hat{\imath} + y\hat{\jmath}$.
(4) $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$.

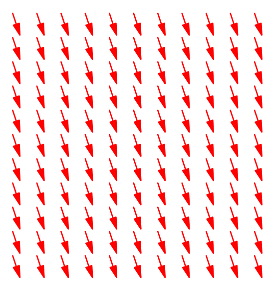


Figure 1. Picture of $\vec{F} = \hat{\imath} - 3\hat{\jmath}$

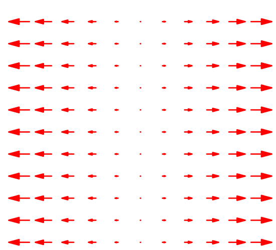


Figure 2. Picture of $\vec{F} = x\hat{\imath}$

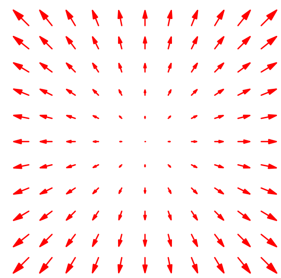


FIGURE 3. Picture of $\vec{F} = x\hat{\imath} + y\hat{\jmath}$

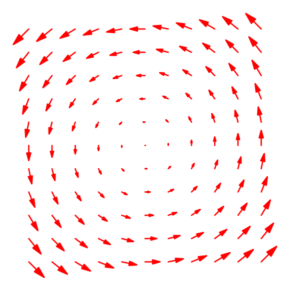


FIGURE 4. Picture of $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$

Recall that the work done is the dot product of the force and the displacement,

$$W \approx \vec{F} \cdot \Delta\vec{r},$$

for a small displacement $\Delta\vec{r}$. If we sum all of the displacements along a trajectory C, we get a Riemann sum and taking the limit as $\Delta\vec{r}$ goes to zero, we get an integral, called a line integral

$$W = \int_C \vec{F} \cdot \mathrm{d}\vec{r} = \lim_{\Delta\vec{r} \to 0} \sum_i \vec{F}_i \cdot \Delta\vec{r}_i.$$

To calculate the line integral, choose a parametrisation $\vec{r}(t)$ of C (you can think of t as the time);

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_{t_1}^{t_2} \left(\vec{F} \cdot \frac{d\vec{r}}{dt} \right) \mathrm{d}t.$$

For example, suppose $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$ and C is given by $x = t$ and $y = t^2$, $0 \leq t \leq 1$. So C is part of the parabola $y = x^2$, starting at $(0, 0)$ and ending at $(1, 1)$.

We calculate everything in terms of t,

$$\vec{F} = \langle -y, x \rangle = \langle -t^2, t \rangle \qquad \text{and} \qquad \frac{d\vec{r}}{dt} = \langle 1, 2t \rangle.$$

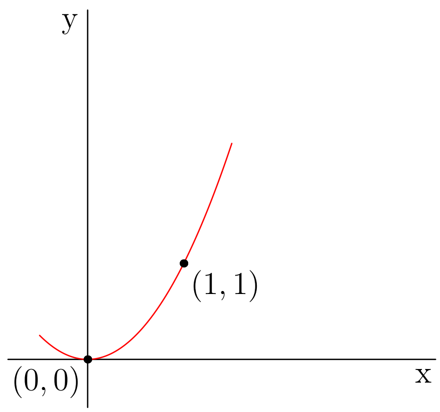


FIGURE 5. The curve C

Hence

$$\int_C \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^1 \vec{F}\cdot\frac{d\vec{r}}{dt}\,\mathrm{d}t = \int_0^1 \langle -t^2, t\rangle\cdot\langle 1, 2t\rangle\,\mathrm{d}t = \int_0^1 t^2\,\mathrm{d}t = \left[\frac{t^3}{3}\right]_0^1 = \frac{1}{3}.$$

Note that we could parametrise C in many different ways. For example, we could choose $x = \sin\theta$, $y = \sin^2\theta$. In this case,

$$\vec{F} = \langle -y, x\rangle = \langle -\sin^2\theta, \sin\theta\rangle \qquad \text{and} \qquad \frac{d\vec{r}}{dt} = \langle \cos\theta, 2\cos\theta\sin\theta\rangle.$$

Hence

$$\int_C \vec{F}\cdot\mathrm{d}\vec{r} = \int_0^{\pi/2} \langle -\sin^2\theta, \sin\theta\rangle\cdot\langle\cos\theta, 2\cos\theta\sin\theta\rangle\,\mathrm{d}\theta = \int_0^{\pi/2}\cos\theta\sin^2\theta\,\mathrm{d}\theta.$$

If we make the substitution $t = \sin\theta$ we get back to the old integral. In practice we always try to use the simplest parametrisation.

There is an alternative and quite pervasive notation for line integrals. We have $\vec{F} = \langle M, N\rangle$ and $\mathrm{d}\vec{r} = \langle \mathrm{d}x, \mathrm{d}y\rangle$. So the line integral is

$$\int_C \vec{F}\cdot\mathrm{d}\vec{r} = \int_C M\mathrm{d}x + N\mathrm{d}y.$$

Note that this notation is a little confusing; it is important to realise we still have a line integral. In the example above we have

$$\int_C -y\,\mathrm{d}x + x\,\mathrm{d}y = \int_0^1 -t^2\,\mathrm{d}t + t\,\mathrm{d}t^2 = \int_0^1 t^2\,\mathrm{d}t^2 = \frac{1}{3}.$$

Here we used the fact that

$$\mathrm{d}y = \mathrm{d}t^2 = 2t\,\mathrm{d}t.$$

Sometimes it is better to use the arclength parametrisation. Recall we have

$$\frac{d\vec{r}}{dt} = \frac{ds}{dt}\hat{T},$$

where s is the arclength parameter and $\hat{T}$ is the unit tangent vector. So

$$\mathrm{d}\vec{r} = \hat{T}\mathrm{d}s.$$

In this case,

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_C \vec{F} \cdot \hat{T}\,\mathrm{d}s.$$

For example, suppose C is a circle of radius a centred at the origin. If $\vec{F} = x\hat{\imath} + y\hat{\jmath}$, then $\vec{F}$ is orthogonal to the unit tangent vector, so that

$$\vec{F} \cdot \hat{T} = 0.$$

In this case,

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_C \vec{F} \cdot \hat{T}\,\mathrm{d}s = 0.$$

Another way to think of this is as follows. If the force is radial then it is perpendicular to the direction in which we move, so the work done is zero.

Now suppose that $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$. Then $\vec{F} \cdot \hat{T} = a$, the magnitude of $\vec{F}$. In this case

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_C a\,\mathrm{d}s = 2\pi a^2.$$

On the other hand, let's choose the parametrisation $x = a\cos\theta$, $y = a\sin\theta$. Then

$$\vec{F} = \langle -a\sin\theta, a\cos\theta \rangle \qquad \text{and} \qquad \frac{d\vec{r}}{dt} = \langle -a\sin\theta, a\cos\theta \rangle.$$

So we get

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_0^{2\pi} a^2 \mathrm{d}\theta = 2\pi a^2.$$

20. Line integrals

Let's look more at line integrals. Let's suppose we want to compute the line integral of $\vec{F} = y\hat{\imath} + x\hat{\jmath}$ around the curve C which is the sector of the unit circle whose angle is $\pi/4$, starting and ending at the origin. We break C into three curves,

$$C = C_1 + C_2 + C_3.$$

The line C_1 from $(0,0)$ to $(1,0)$, the arc C_2 of the unit circle starting at $(1,0)$ and ending at $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and the line from this point back to the origin C_3.

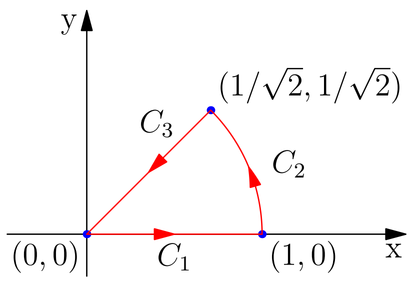


Figure 1. The curve C

We have

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_{C_1} \vec{F} \cdot \mathrm{d}\vec{r} + \int_{C_2} \vec{F} \cdot \mathrm{d}\vec{r} + \int_{C_3} \vec{F} \cdot \mathrm{d}\vec{r}.$$

We parametrise each curve separately.

The curve C_1: For the x-axis, $x(t) = t$, $y(t) = 0$, $0 \leq t \leq 1$. In this case

$$\vec{F} = \langle y, x \rangle = \langle 0, t \rangle \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle 1, 0 \rangle \, \mathrm{d}t.$$

So

$$\int_{C_1} \vec{F} \cdot \mathrm{d}\vec{r} = \int_0^1 \langle 0, t \rangle \cdot \langle 1, 0 \rangle \, \mathrm{d}t = \int_0^1 0 \, \mathrm{d}t = 0.$$

In fact there are two other ways to see that we must get zero. We could take the arclength parametrisation. In this case $\hat{T} = \hat{\imath}$ and $\vec{F} = t\hat{\jmath}$, so that $\vec{F} \cdot \hat{T} = 0$. Or observe that the work done is zero, since the force is orthogonal to the velocity vector.

The curve C_2: For the arc of the circle, $x(t) = \cos t$, $y(t) = \sin t$, $0 \leq t \leq \pi/4$. In this case

$$\vec{F} = \langle y, x \rangle = \langle \sin t, \cos t \rangle \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle -\sin t, \cos t \rangle \, \mathrm{d}t.$$

So

$$\int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^{\pi/4} \langle \sin t, \cos t\rangle \cdot \langle -\sin t, \cos t\rangle \,\mathrm{d}t = \int_0^{\pi/4} \cos(2t)\,\mathrm{d}t = \left[\frac{\sin(2t)}{2}\right]_0^{\pi/4} = \frac{1}{2}.$$

The curve C_3: For the straight line segment starting at $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and ending at the origin, we have $x(t) = t$, $y(t) = t$, $0 \le t \le 1/\sqrt{2}$.

$$\vec{F} = \langle y, x\rangle = \langle t, t\rangle \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle 1, 1\rangle\,\mathrm{d}t.$$

So,

$$\int_{C_3} \vec{F}\cdot \mathrm{d}\vec{r} = \int_{1/\sqrt{2}}^{0} \langle t, t\rangle \cdot \langle 1, 1\rangle \,\mathrm{d}t = \int_{1/\sqrt{2}}^{0} 2t\,\mathrm{d}t = \Big[t^2\Big]_{1/\sqrt{2}}^{0} = -\frac{1}{2}.$$

Note that the limits start at $1/\sqrt{2}$ and end at 0.

Putting all of this together, we get

$$\int_{C} \vec{F}\cdot \mathrm{d}\vec{r} = \int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{C_3} \vec{F}\cdot \mathrm{d}\vec{r} = 0 + 1/2 - 1/2 = 0.$$

We say that $\vec{F}$ is a gradient field if $\vec{F} = \nabla f$, for some scalar function f.

Theorem 20.1 (Fundamental Theorem of Calculus for line integrals)**.** *If $\vec{F} = \nabla f$ is a gradient vector field then*

$$\int_{C} \vec{F}\cdot \mathrm{d}\vec{r} = \int_{C} \nabla f\cdot \mathrm{d}\vec{r} = f(P_1) - f(P_0),$$

where C is a path from P_0 to P_1.

For example, suppose we take $f(x, y) = xy$. Then

$$\nabla f = y\hat{\imath} + x\hat{\jmath} = \vec{F},$$

the vector field above. Using (20.1), we see that

$$\int_{C} \vec{F}\cdot \mathrm{d}\vec{r} = f(0,0) - f(0,0) = 0.$$

On the other hand,

$$\int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} = f(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) - f(1,0) = \frac{1}{2}.$$

In the language of differentials, one can restate (20.1) as

$$\int_{C} f_x\,\mathrm{d}x + f_y\,\mathrm{d}y = \int_{C} \mathrm{d}f = f(P_1) - f(P_0).$$

Proof of (20.1).

$$\begin{aligned}\int_C \nabla f \cdot \mathrm{d}\vec{r} &= \int_{t_0}^{t_1} \left(f_x \frac{dx}{dt} + f_y \frac{dy}{dt} \right) \mathrm{d}t \\ &= \int_{t_0}^{t_1} \frac{d}{dt} \left(f(x(t), y(t)) \right) \mathrm{d}t \\ &= \Big[f(x(t), y(t)) \Big]_{t_0}^{t_1} \\ &= f(P_1) - f(P_0). \end{aligned}$$ □

(20.1) has some very interesting consequences:

Path independence: If C_1 and C_2 are two paths starting and ending at the same point, then

$$\int_{C_1} \nabla f \cdot \mathrm{d}\vec{r} = \int_{C_2} \nabla f \cdot \mathrm{d}\vec{r}.$$

In other words, the line integral

$$\int_C \nabla f \cdot \mathrm{d}\vec{r},$$

depends only on the endpoints, not on the trajectory.

Gradient fields are conservative: If C is a closed loop, then

$$\int_C \nabla f \cdot \mathrm{d}\vec{r} = 0.$$

We already saw that if C is a circle of radius a centred at the circle and $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$, then

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = 2\pi a^2 \neq 0.$$

So the vector field $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$ is not conservative. It follows that $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$ is not the gradient of any scalar field.

If $\vec{F} = \nabla f$ is a gradient field, and $\vec{F}$ is the force, then f has an interesting physical interpretation, it is called the potential. In this case the work done is nothing more than the change in the potential. For example, if $\vec{F}$ is the force due to gravity, f is inversely proportional to the height. If $\vec{F}$ is the electric field, f is the voltage. (Note the annoying fact that mathematicians and physicists use a different sign convention; for physicists $\vec{F} = -\nabla f$).

To summarise, we have four equivalent properties:

(1) $\vec{F}$ is conservative, that is, $\int_C \vec{F} \cdot \mathrm{d}\vec{r} = 0$ for any closed loop.
(2) $\int_C \vec{F} \cdot \mathrm{d}\vec{r}$ is path independent.

(3) $\vec{F} = \nabla f$ is a gradient vector field.

(4) $M\,\mathrm{d}x + N\,\mathrm{d}y$ is an exact differential, equal to $\mathrm{d}f$.

(1) and (2) are equivalent by considering the closed loop $C = C_1 - C_2$. (3) implies (2) by (20.1). We will see (2) implies (3) in the next lecture. (3) and (4) are the same statement, using different notation.

21. Potential Functions

Suppose that $\vec{F} = M\hat{\imath} + N\hat{\jmath} = \nabla f$ is a gradient vector field. Then

$$M_y = f_{xy} = f_{yx} = N_x.$$

So, if $\vec{F}$ is a gradient vector field then $M_y = N_x$.

Theorem 21.1. *Let $\vec{F} = M\hat{\imath} + N\hat{\jmath}$ be a vector field which is defined and differentiable on the whole of $\mathbb{R}^2$.*

Then $\vec{F}$ is a gradient vector field if and only if $M_y = N_x$.

Example 21.2. *Let $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$. Then $M = -y$ and $N = x$. So*

$$M_y = -1 \qquad \textit{and} \qquad N_x = 1,$$

which are not equal. So $\vec{F}$ is not a gradient vector field.

Question 21.3. *For which values of a is $\vec{F} = (4x^2+axy)\hat{\imath}+(3y^2+4x^2)\hat{\jmath}$ a gradient field?*

We have

$$M = 4x^2 + axy \qquad \text{and} \qquad N = 3y^2 + 4x^2.$$

So

$$M_y = ax \qquad \text{and} \qquad N_x = 8x.$$

It follows that $M_y = N_x$ if and only if $a = 8$.

Given that (21.1) is true, it follows that if $M_y = N_x$ then $\vec{F} = \nabla f$, for some scalar function $f(x, y)$. We give two methods to calculate f, when

$$\vec{F} = (4x^2 + 8xy)\hat{\imath} + (3y^2 + 4x^2)\hat{\jmath}.$$

Method 1: We could use the fundamental theorem of calculus for line integrals. Suppose we want to determine the value of $f(x, y)$ at a point (x_1, y_1). Pick a curve C starting at $(0, 0)$ and ending at (x_1, y_1). We have

$$f(x_1, y_1) - f(0, 0) = \int_C \vec{F} \cdot \mathrm{d}\vec{r}.$$

Note $f(0, 0)$ is an integration constant. If f is a potential function then so is $f + c$.

Note that we get to choose C. A sensible choice in this example is to decompose C as the straight line C_1 from $(0, 0)$ to $(x_1, 0)$ and the vertical line from $(x_1, 0)$ to (x_1, y_1), $C = C_1 + C_2$.

We have

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_{C_1} \vec{F} \cdot \mathrm{d}\vec{r} + \int_{C_2} \vec{F} \cdot \mathrm{d}\vec{r}.$$

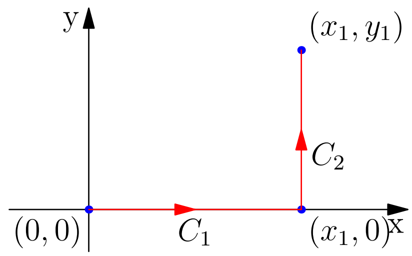


FIGURE 1. The curve C

Let $x(t) = t$, $y(y) = 0$, a parametrisation of C_1. Then

$$\vec{F} = 4t^2\hat{\imath} + 4t^2\hat{\jmath} \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle 1, 0\rangle\, \mathrm{d}t.$$

So

$$\int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^{x_1} \langle 4t^2, 4t^2\rangle \cdot \langle 1, 0\rangle\, \mathrm{dt} = \int_0^{x_1} 4t^2\, \mathrm{dt} = \Big[4t^3/3\Big]_0^{x_1} = 4x_1^3/3.$$

Let $x(t) = x_1$, $y(y) = t$, a parametrisation of C_2. Then

$$\vec{F} = (4x_1^2 + 8x_1 t)\hat{\imath} + (3t^2 + 4x_1^2)\hat{\jmath} \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle 0, 1\rangle\, \mathrm{d}t.$$

So

$$\int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^{y_1} \langle 4x_1^2 + 8x_1 t, 3t^2 + 4x_1^2\rangle \cdot \langle 0, 1\rangle\, \mathrm{dt} = \int_0^{y_1} 3t^2 + 4x_1^2\, \mathrm{dt} = \Big[t^3 + 4x_1^2 t\Big]_0^{y_1} = y_1^3 + 4x_1^2 y_1.$$

So

$$f(x, y) = 4x^3/3 + y^3 + 4x^2 y + c,$$

where c is a constant. Check

$$\nabla f = \langle 4x^2 + 8xy, 3y^2 + 4x^2\rangle = \vec{F},$$

as expected.

Method 2: We want to solve two PDE's

$$f_x = 4x^2 + 8xy \qquad \text{and} \qquad f_y = 3y^2 + 4x^2.$$

Now if we integrate the first equation with respect to x we get

$$f(x, y) = \int f_x(x, y)\, \mathrm{d}x = 4x^3/3 + 4x^2 y + g(y),$$

where $g(y)$ is a function of y. The point here is that for every value of y, we get an integration constant. As we vary y this integration constant can vary. Put differently, if we differentiate $g(y)$ with respect to x then we get zero. So $f(x,y)$ is determined up to $g(y)$. Now plug this value for $f(x,y)$ into the second PDE.

$$4x^2 + \frac{dg}{dy} = 3y^2 + 4x^2.$$

Comparing we have

$$\frac{dg}{dy} = 3y^2.$$

Integrating with respect to y, we get

$$g(y) = y^3 + c,$$

where c is an integration constant. So

$$f(x,y) = 4x^3/3 + 4x^2y + y^3 + c.$$

This is the same solution we got using the other method.

Let's introduce a quantity which measures how far the vector field $\vec{F}$ is from being conservative, the curl of $\vec{F}$,

$$\operatorname{curl}\vec{F} = N_x - M_y.$$

We have $\operatorname{curl}\vec{F} = 0$ if and only if $\vec{F}$ is a gradient field, if and only if $\vec{F}$ is conservative.

The curl of a vector field is a strange beast. If $\vec{F}$ is a velocity vector field, the curl is double the angular velocity of the rotation component of the motion.

Example 21.4. *If $\vec{F} = \langle a, b\rangle$ is a constant vector field, then* $\operatorname{curl}\vec{F}$ *is zero,*

$\vec{F} = \langle x, y\rangle$ *represents expanding motion, which has zero curl.*

$\vec{F} = \langle -y, x\rangle$ *represents rotation around the origin, the curl is* 2*.*

If $\vec{F}$ is a force field then $\operatorname{curl}\vec{F}$ is the torque exerted on a test mass. This measures how much $\vec{F}$ imparts angular momentum. For translation motion, the force divided by the mass is the acceleration, the derivative of the velocity. For rotation, the torque divided by the moment of inertia is the angular acceleration, the derivative of the angular velocity.

22. Greens Theorem

Theorem 22.1 (Green's Theorem)**.** *If C is a positively oriented closed curve enclosing a region R then*

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \iint_R \operatorname{curl} \vec{F} \, \mathrm{d}A.$$

The circle in the centre of the integral sign is simply to emphasize that the line integral is around a closed loop. Here C is oriented so that R is on the left as we go around C.

Green's Theorem, in the language of differentials, comes out as

$$\oint_C M \, \mathrm{d}x + N \, \mathrm{d}y = \iint_R (N_x - M_y) \, \mathrm{d}A.$$

For example, let C be a unit circle centred at $(2, 0)$, oriented counterclockwise and let R be the unit disk, centred at $(2, 0)$.

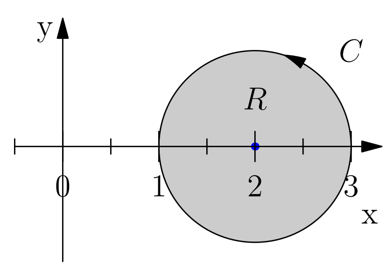


Figure 1. The region R with boundary C

We have

$$\begin{aligned} \oint_C ye^{-x} \, \mathrm{d}x + \left(\frac{1}{2}x^2 - e^{-x}\right) \, \mathrm{d}y &= \oint_C M \, \mathrm{d}x + N \, \mathrm{d}y \\ &= \iint_R (N_x - M_y) \, \mathrm{d}A \\ &= \iint_R (x + e^{-x} - e^{-x}) \, \mathrm{d}A \\ &= \iint_R x \, \mathrm{d}A. \end{aligned}$$

Now one could calculate the last integral by direct calculation of the iterated integral. On the other hand, if we divide the last integral by the area, we get $\bar{x}$, the x-coordinate of the centre of mass. Obviously the centre of mass is at the centre of the circle, so $\bar{x} = 2$. The area is π, so the integral is 2π.

Corollary 22.2. *Let $\vec{F} = M\hat{\imath} + N\hat{\jmath}$ be a vector field which is defined and differentiable on the whole of $\mathbb{R}^2$.*

Then $\vec{F}$ is a gradient vector field if and only if $M_y = N_x$.

Proof. Suppose that $M_y = N_x$. Then $\operatorname{curl} \vec{F} = 0$. By (**??**), we have

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \iint_R \operatorname{curl} \vec{F} \,\mathrm{d}A = \iint_R 0 \,\mathrm{d}A = 0.$$

Hence $\vec{F}$ is conservative. □

Note that this only works if the region R is completely contained in the locus where $\vec{F}$ is defined. In question B5 of last weeks hwk, the integral around the unit circle the vector field $\vec{F}$ is not defined at the origin.

We now describe the proof of (**??**)

Proof of (**??**). First a couple of useful reduction steps. For a start it suffices to prove two separate identities:

$$\oint_C M \,\mathrm{d}x = \iint_R -M_y \,\mathrm{d}A \qquad \text{and} \qquad \oint_C N \,\mathrm{d}y = \iint_R N_x \,\mathrm{d}A.$$

To get the general result, add these two identities.

Secondly, if R is the union of two regions R_1 and R_2 and we know the result for both regions R_1 and R_2 then we know it for R. Indeed,

$$\begin{aligned}
\oint_C \vec{F} \cdot \mathrm{d}\vec{r} &= \oint_{C_1} \vec{F} \cdot \mathrm{d}\vec{r} + \oint_{C_2} \vec{F} \cdot \mathrm{d}\vec{r} \\
&= \iint_{R_1} \operatorname{curl} \vec{F} \,\mathrm{d}A + \iint_{R_2} \operatorname{curl} \vec{F} \,\mathrm{d}A \\
&= \iint_R \operatorname{curl} \vec{F} \,\mathrm{d}A.
\end{aligned}$$

Here C is the boundary of R and C_1, C_2 are the boundaries of R_1 and R_2. The first equality is therefore a little bit more subtle than might first appear; the key thing is that we might get some cancelling.

Using these two reduction steps, we get down to the kernel of the proof. Prove that

$$\oint_C M \,\mathrm{d}x = \iint_R -M_y \,\mathrm{d}A$$

where R is a vertically simple region, that is a region of the form

$$a \leq x \leq b \qquad \text{and} \qquad f_0(x) \leq y \leq f_1(x).$$

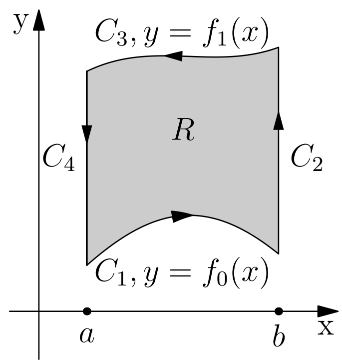


FIGURE 2. Typical vertically simple region

So R is the region between the graph of two functions. Now we calculate both sides. For the LHS break C into four pieces,

$$C = C_1 + C_2 + C_3 + C_4,$$

where C_1 is the lower edge, the graph of $y = f_0(x)$ between a and b, C_2 is the right vertical segment, C_3 is the upper edge, the graph of $y = f_1(x)$ between a and b, and C_4 is the left vertical segment. Now

$$\int_{C_2} M \,\mathrm{d}x = \int_{C_4} M \,\mathrm{d}x = 0,$$

since x is constant on these edges. For the other two edges use the parametrisation $x(t) = t$, $y(t) = f_0(t)$, $a \leq t \leq b$ and $x(t) = t$, $y(t) = f_1(t)$, $a \leq t \leq b$, but with the opposite orientation, so that we get

$$\oint_C M \,\mathrm{d}x = \int_{C_1} M \,\mathrm{d}x + \int_{C_3} M \,\mathrm{d}x = \int_a^b M(t, f_0(t)) \,\mathrm{d}t - \int_a^b M(t, f_1(t)) \,\mathrm{d}t.$$

For the RHS we have

$$-\iint_R M_y \,\mathrm{d}A = -\int_a^b \int_{f_0(x)}^{f_1(x)} M_y \,\mathrm{d}y \,\mathrm{d}x.$$

Now the inner integral is

$$\int_{f_0(x)}^{f_1(x)} M_y \,\mathrm{d}y = -\int_a^b M(x, f_1(x)) - M(x, f_0(x)) \,\mathrm{d}x,$$

and so the outer integral is

$$\int_a^b M(x, f_0(x)) - M(x, f_1(x)) \,\mathrm{d}x,$$

the same as the LHS.

There is a similar calculation with N replacing M, horizontally simple regions replacing vertically simple regions and suitable switching of x and y. $\square$

Example 22.3. *The area of a region R can be evaluated using Green's theorem. For example,*

$$\text{area}(R) = \iint_R 1 \, \mathrm{d}A = \oint_C x \, \mathrm{d}y.$$

One can actually build physical devices that measure area this way. If one has a figure on a piece of paper, a planimeter can be used to find the area. Move the end of the planimeter so it traces out the curve C. At the end one can read off the area.

For a linear planimeter, there is an arm AB. B is constrained to lie in the y-axis and the point A traces out the curve C. Suppose it has coordinates $(0, b)$. Suppose the coordinates of A are (x, y). So $\overrightarrow{AB} = \langle x, y - b\rangle$. It follows that if $\vec{F} = \langle b - y, x\rangle$, then $\vec{F}$ is perpendicular to $\overrightarrow{AB}$. The length of $\vec{F}$ is a constant equal to the length m of the arm. By (**??**)

$$\begin{aligned}
\oint_C M \, \mathrm{d}x + N \, \mathrm{d}y &= \iint_R (N_x - M_y) \, \mathrm{d}A \\
&= \iint_R 1 - \frac{\partial (b - y)}{\partial y} \, \mathrm{d}A \\
&= \iint_R 1 \mathrm{d}A \\
&= \text{area}(R).
\end{aligned}$$

Here we used the fact that

$$\frac{\partial (b - y)}{\partial y} = \frac{\partial \sqrt{m^2 - x^2}}{\partial y} = 0.$$

The flux of a vector field $\vec{F}$ across a curve C is

$$\int_C \vec{F} \cdot \hat{n} \, \mathrm{d}s,$$

where $\hat{n}$ is the unit normal vector to the curve C, obtained from the unit tangent vector $\hat{T}$ by rotating this vector through $\pi/2$ clockwise.

This gives us two line integrals:

We can integrate $\vec{F} \cdot \hat{T}$. In terms of Riemann sums, we add up the contributions from the component of $\vec{F}$ in the direction of $\hat{T}$, that is, along C. This computes the work done.

Or we can integrate $\vec{F} \cdot \hat{n}$. In terms of Riemann sums, we add up the contributions from the component of $\vec{F}$ in the direction of $\hat{n}$, that is, perpendicular to C. This computes the flux.

Suppose that $\vec{F}$ is a velocity vector field. Then the line integral

$$\int_C \vec{F} \cdot \hat{n} \, \mathrm{d}s,$$

represents how much matter crosses C in unit time.

To see this, let's fix ideas and suppose that $\vec{F}$ represents flow of water. Consider a small portion of C. Along this portion, $\vec{F}$ is approximately constant. The amount of water crossing C in unit time is given by a parallelogram with sides $\vec{F}$ and $(\Delta s)\hat{T}$. The area of this parallelogram is

$$(\vec{F} \cdot \hat{n})\Delta s;$$

$\vec{F} \cdot \hat{n}$ is the height of the parallelogram and Δs is the base. Dividing C into small pieces and summing all of these terms, gives a Riemann sum, an approximation to the total amount of water crossing C in unit time. Taking the limit as Δs goes to zero, the line integral

$$\int_C \vec{F} \cdot \hat{n} \, \mathrm{d}s,$$

represents how much water crosses C in unit time.

Note that water flowing left to right across C gets counted positively and water crossing right to left gets counted negatively (from the point of view of a particle travelling along C).

Example 23.1. *Suppose C is a circle of radius a, centre the origin. Let $\vec{F} = x\hat{\imath} + y\hat{\jmath}$. Then $\vec{F}$ points in the same direction as $\hat{n}$. So*

$$\vec{F} \cdot \hat{n} = |\vec{F}| = a.$$

It follows that

$$\oint_C \vec{F} \cdot \hat{n}\,\mathrm{d}s = \oint_C a\,\mathrm{d}s = 2\pi a^2.$$

The flux is $2\pi a^2$.

On the other hand, suppose we start with $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$. *Then*

$$\vec{F} \cdot \hat{n} = 0,$$

since $\vec{F}$ *is perpendicular to* $\hat{n}$. *The flux is zero.*

Note that this makes sense physically. In the first example, water is spewing out of the origin. Lots of it crosses C. In the second example, water is spinning around the origin. None of it crosses C.

Now let's turn to how we would calculate the flux algebraically. We have

$$\mathrm{d}\vec{r} = \hat{T}\,\mathrm{d}s = \langle \mathrm{d}x, \mathrm{d}y \rangle.$$

The vector $\hat{n}$ is obtained from $\hat{T}$ by rotation through $\pi/2$ clockwise. So

$$\hat{n}\,\mathrm{d}s = \langle \mathrm{d}y, -\mathrm{d}x \rangle.$$

So as not to get lost in notation, let's suppose the components of $\vec{F}$ are P and Q,

$$\vec{F} = \langle P, Q \rangle = P\hat{\imath} + Q\hat{\jmath}.$$

We have

$$\int_C \vec{F} \cdot \hat{n}\,\mathrm{d}s = \int_C \langle P, Q \rangle \cdot \langle \mathrm{d}y, -\mathrm{d}x \rangle = \int_C -Q\,\mathrm{d}x + P\,\mathrm{d}y.$$

Theorem 23.2 (Green's Theorem for flux)**.** *If* C *is a positively oriented closed curve enclosing a region* R *and* $\vec{F} = P\hat{\imath} + Q\hat{\jmath}$ *then*

$$\oint_C -Q\,\mathrm{d}x + P\,\mathrm{d}y = \iint_R \operatorname{div} \vec{F}\,\mathrm{d}A, \qquad \textit{where} \qquad \operatorname{div} \vec{F} = P_x + Q_y.$$

Proof. Call $M = -Q$ and $N = P$, so that $\vec{G} = \langle M, N\rangle$ is $\vec{F}$ rotated through $\pi/2$ anticlockwise. Then we have

$$\begin{aligned}
\oint_C -Q\,\mathrm{d}x + P\,\mathrm{d}y &= \oint_C M\,\mathrm{d}x + N\,\mathrm{d}y \\
&= \iint_R \operatorname{curl}\vec{G}\,\mathrm{d}y \\
&= \iint_R N_x - M_y\,\mathrm{d}y \\
&= \iint_R P_x + Q_y\,\mathrm{d}y \\
&= \iint_R \operatorname{div}\vec{F}\,\mathrm{d}A. \qquad \square
\end{aligned}$$

Example 23.3. *Consider the example of a circle C of radius a, centre the origin. Suppose that $\vec{F} = x\hat{\imath} + y\hat{\jmath}$. Then*

$$\operatorname{div}\vec{F} = 1 + 1 = 2.$$

So the RHS of (23.2) *is*

$$\iint_R 2\,\mathrm{d}A = 2\pi a^2.$$

Suppose we move the circle away from the origin. Then computing the LHS becomes quite hard. But the RHS is unchanged.

$\operatorname{div}\vec{F}$ is called the divergence of $\vec{F}$. If $\vec{F}$ is the velocity vector field of water, the divergence measures how much water is being added (or taken away); these are known as sources (or sinks). For the vector field $\vec{F} = x\hat{\imath} + y\hat{\jmath}$, water is being added everywhere (imagine rain falling on the ground and then flowing away from the origin).

For the vector field $\vec{F} = -y\hat{\imath} + x\hat{\jmath}$ the divergence is zero. No water is being added or removed, there are no sources or sinks.

24. Simply connected regions; triple integrals

24.1. **Simply connected regions.** Recall the example of the vector field
$$\vec{F} = \frac{-y\hat{\imath} + x\hat{\jmath}}{x^2 + y^2}.$$
Its curl is zero but this vector field is not conservative. If one looks at the unit circle C centred at the origin, then the line integral of $\vec{F}$ around C is 2π. The point is that Green's theorem does not apply, as $\vec{F}$ is not defined on the whole of the region R enclosed by C; $\vec{F}$ is not defined at the origin.

Definition 24.1. *A region R is simply connected if the interior of every closed curve C is entirely contained in R.*

Informally speaking, the region R does not contain any holes.

Question 24.2. *Which of the following regions are simply connected?*

(1) the whole plane $\mathbb{R}^2$.
(2) the plane minus the origin.
(3) the unit disk.
(4) the plane minus a line.
(5) a comb.

Theorem 24.3. *Suppose the region R is simply connected and $\vec{F}$ is a vector field on R which is differentiable.*

Then $\vec{F}$ is a gradient vector field if and only if $M_y = N_x$.

The proof is the same as when R is the whole of $\mathbb{R}^2$; the same proof works by the very definition of simply connected.

Let's go back to the example above. Let's try to understand what is going on when the region is not simply connected. Start with an annulus, whose boundary is a circle of small radius plus a circle of arbitrary radius. Manufacture a closed curve by going around the big circle counterclockwise, going across a line segment to the small circle, going around the small circle, in the opposite direction and finally going back to the big circle, along the same line segment. Call the whole curve C. This encloses a region R on which $\vec{F}$ is defined everywhere. With this choice of orientation of C, R is always on the left.

The curl is zero, so
$$\oint_C \vec{F}\,\mathrm{d}\vec{r} = \iint_R \operatorname{curl} F\,\mathrm{d}A = 0.$$
But
$$\oint_C \vec{F}\,\mathrm{d}\vec{r} = \int_{C_b} \vec{F}\,\mathrm{d}\vec{r} + \int_L \vec{F}\,\mathrm{d}\vec{r} - \oint_{C_s} \vec{F}\,\mathrm{d}\vec{r} - \int_L \vec{F}\,\mathrm{d}\vec{r},$$

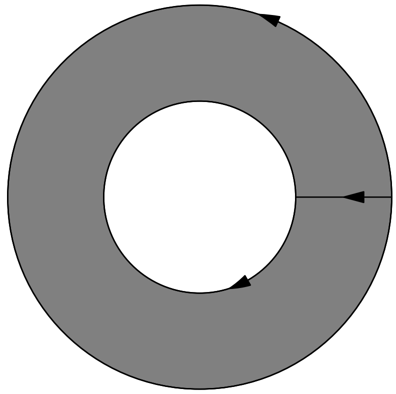


FIGURE 1. The region R with boundary C

where C_b is the big circle, C_s is the small circle, and L is the line segment.

Putting all of this together, we get

$$\int_{C_b} \vec{F}\,\mathrm{d}\vec{r} = \oint_{C_s} \vec{F}\,\mathrm{d}\vec{r}.$$

So the line integral is independent of the size of the circle; in fact any simple closed curve around the origin has integral 2π.

24.2. **Triple integrals.** Suppose we have a function $f(x, y, z)$ of three variables defined on a region R in space $\mathbb{R}^3$. We can dice up R into small regions R_i and form the Riemann sum

$$\sum_i f(x_i, y_i, z_i)\Delta V_i,$$

where ΔV_i is the volume of R_i and (x_i, y_i, z_i) is a random point belonging to R_i. Letting ΔV_i go to zero we get the triple integral

$$\iiint_R f\,\mathrm{d}V.$$

Example 24.4. *What is the volume of the region R between the two paraboloids $z = x^2 + y^2$ and $z = 4 - x^2 - y^2$?*

We want to calculate

$$\mathrm{vol}(R) = \iiint_R 1\,\mathrm{d}V.$$

As usual imagine choosing V_i equal to boxes. First summing over z, then y and then x, the integral reduces to a triple integral

$$\iiint_R 1\,\mathrm{d}V = \int_{x_0}^{x_1}\int_{y_0}^{y_1}\int_{z_0}^{z_1} 1\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x,$$

where we need to determine the limits. The inner limit is the easiest. Given x and y, z ranges between x^2+y^2 and $4-x^2-y^2$. To find the limits for x and y, consider the shadow of R in the xy-plane, the points in the xy-plane which are in the shade, if light comes vertically down.

We want to find those points (x,y) for which the point on the surface $z = 4 - x^2 - y^2$ is above the point on the surface $x^2 + y^2$, that is, we want those points (x,y) such that

$$x^2 + y^2 < 4 - x^2 - y^2.$$

Simplifying, we want

$$x^2 + y^2 < 2,$$

which is a disk of radius $\sqrt{2}$ centred at the origin. So we have

$$\iiint_R 1\,\mathrm{d}V = \int_{-\sqrt{2}}^{\sqrt{2}}\int_{-\sqrt{2-x^2}}^{\sqrt{2-x^2}}\int_{x^2+y^2}^{4-x^2-y^2} 1\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x,$$

We could have found the limits quicker by setting $x^2+y^2 = 4-x^2-y^2$; however this won't always work. Now notice that this would come out much better if we used polar coordinates in the xy-plane.

$$\iiint_R 1\,\mathrm{d}V = \int_0^{2\pi}\int_0^{\sqrt{2}}\int_{r^2}^{4-r^2} r\,\mathrm{d}z\,\mathrm{d}r\,\mathrm{d}\theta,$$

Cylindrical coordinates (r,θ,z) are precisely the coordinates one gets this way. z is the usual height from the xy-plane, and r and θ are the usual polar coordinates of the projection of the point P down to the xy-plane. So r is the distance of P to the z-axis and θ is the angle the shortest line from P to the z-axis makes with the plane $y = 0$, the xz-plane. Of course

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

gets one from cylindrical coordinates to the usual Cartesian coordinates. Note that the volume element

$$\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z.$$

As usual, we get this by considering what happens for a small change in the coordinates.

It is interesting to look at simple equations in cylindrical coordinates. For example, we already know

$$z = c,$$

c a constant, corresponds to a horizontal plane.

$$r = a,$$

a a constant, corresponds to a cylinder of radius a, centred along the z-axis. The equation

$$\theta = b,$$

b a constant, corresponds to a half plane, starting at the z-axis. Even simple equations look quite strange,

$$r = \theta,$$

corresponds to a scroll (infinite at both ends, parallel to the z-axis); if we take a cross-section, we get a spiral. The equation

$$z = ar,$$

represents a cone, vertex at the origin, central axis the z-axis.

Triple integrals have the same sort of applications as double integrals. If R is composed of material, with density δ, the mass of R is the integral,

$$\iiint_R \delta \,\mathrm{d}V.$$

If $f(x, y, z)$ is a function on R, the average value $\bar{f}$ of $f(x, y, z)$ is

$$\bar{f} = \frac{1}{V} \iiint_R f \,\mathrm{d}V.$$

Here V is the volume of R. In particular, the centre of mass, with coordinates $(\bar{x}, \bar{y}, \bar{z})$ is given by

$$\bar{x} = \frac{1}{V} \iiint_R x \,\mathrm{d}V \qquad \bar{y} = \frac{1}{V} \iiint_R y \,\mathrm{d}V \qquad \text{and} \qquad \bar{z} = \frac{1}{V} \iiint_R z \,\mathrm{d}V.$$

Note that in the example above, $\bar{x} = \bar{y} = 0$ by symmetry.

We can also calculate the moment of inertia about an axis;

$$\iiint_R d^2 \delta \,\mathrm{d}V,$$

where δ is the mass density and d^2 is the square of the distance to the axis. For the z-axis, we get

$$I_z = \iiint_R (x^2 + y^2) \delta \,\mathrm{d}V = \iiint_R r^2 \delta \,\mathrm{d}V,$$

which is consistent with the story in the plane. For the x-axis and the y-axis, we get

$$I_x = \iiint_R (y^2 + z^2)\delta \, \mathrm{d}V \qquad \text{and} \qquad I_y = \iiint_R (x^2 + z^2)\delta \, \mathrm{d}V.$$

Note that is consistent with the story in the plane, setting $z = 0$.

Example 24.5. *What is the moment of inertia of the solid cone between $z = ar$ and $z = b$ about the z-axis?*

We have

$$I_z = \iiint_R r^2 \, \mathrm{d}V = \int_0^b \int_0^{2\pi} \int_0^{z/a} r^3 \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}z.$$

The limits for θ are easy, using the symmetry about the z-axis. We want to express r as a function of z. To find the upper and lower limits for r, note that we can squish the cone down to a plane (using the symmetry in θ), say the xz-plane.

In this case, the distance r to the z-axis simplifies to x, that is, $r = x$, when $y = 0$. We get a triangle in the xz-plane. We only take the half of the triangle $x \geq 0$, since take $0 \leq \theta \leq 2\pi$. Equivalently, the solid cone is a solid of revolution. Revolve the triangle in the xz-plane around the z-axis.

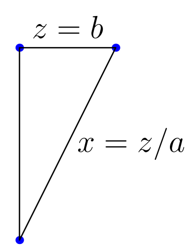


FIGURE 2. Half of squished cone

So $0 \leq r = x \leq z/a$.

Example 24.6. *What is the volume of the region where $z > 1 - y$ and $x^2 + y^2 + z^2 < 1$?*

The bottom surface is the plane $z = 1 - y$ and the top surface is the upper hemisphere $z = \sqrt{1 - x^2 - y^2}$ of the unit sphere. The intersection is a circle. The shadow of R is formed by the image of this circle, which will be an ellipse.

The volume is therefore

$$\iiint_R 1\,\mathrm{d}V = \int_0^1 \int_{-\sqrt{2y(1-y)}}^{\sqrt{2y(1-y)}} \int_{1-y}^{\sqrt{1-x^2-y^2}} 1\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y.$$

The shadow is the region

$$1 - y < \sqrt{1 - x^2 - y^2}.$$

Squaring both sides, we want

$$1 - 2y + y^2 < 1 - x^2 - y^2.$$

That is

$$x^2 < 2y(1 - y),$$

so that

$$-\sqrt{2y(1-y)} < x < \sqrt{2y(1-y)}.$$

Since $x^2 \geq 0$, we must have $2y(1 - y) > 0$. So $0 < y < 1$.

25. Review

Double integrals Integrate function $f(x, y)$ over a region R:

$$\iint_R f \, \mathrm{d}A.$$

Computes the volume of the graph of f lying over R.

Example 25.1. *Evaluate*

$$\int_0^1 \int_0^{x^2} \frac{xe^y}{1-y} \, \mathrm{d}y \, \mathrm{d}x.$$

We cannot caculate this directly.

First we figure out the region of integration. $0 \le x \le 1$. Given x, we have $0 \le y \le x^2$. So we have the region R between $x = 0$ and $x = 1$ under the graph of $y = x^2$. Then we switch the order of integration.

$$\int_0^1 \int_0^{x^2} \frac{xe^y}{1-y} \, \mathrm{d}y \, \mathrm{d}x = \iint_R \frac{xe^y}{1-y} \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 \int_{\sqrt{y}}^1 \frac{xe^y}{1-y} \, \mathrm{d}x \, \mathrm{d}y.$$

The inner integral is

$$\int_{\sqrt{y}}^1 \frac{xe^y}{1-y} \, \mathrm{d}x = \left[\frac{x^2 e^y}{2(1-y)} \right]_{\sqrt{y}}^1 = \frac{e^y(1-y)}{2(1-y)} = \frac{1}{2} e^y.$$

So the outer integral is

$$\int_0^1 \frac{1}{2} e^y \, \mathrm{d}x = \left[\frac{1}{2} e^y \right]_0^1 = \frac{e-1}{2}.$$

We can use the double integral to calculate the mass, centre of mass and moment of inertia:

Example 25.2. *A metal plate is in the shape of a circle of radius* 20*cm. Its density in g/cm*^{2} *at a distance of* r *cm from the centre of the circle is* $10r + 3$.

Find the total mass as an integral.

$$M = \iint_R \delta \, \mathrm{d}A = \int_0^{2\pi} \int_0^{20} (10r + 3) r \, \mathrm{d}r \, \mathrm{d}\theta.$$

Line integrals Integrate a vector field $\vec{F}$ over an oriented curve C.

$$\int_C \vec{F} \cdot \mathrm{d}\vec{r}.$$

Represents the work done.

One can compute directly, by parametrising C. Let $C = C_1 + C_2 + C_3$ be the curve which starts at $(0,0)$ goes along the x-axis to $(1,0)$, goes around the unit circle until $(0,1)$ and comes back to the origin.

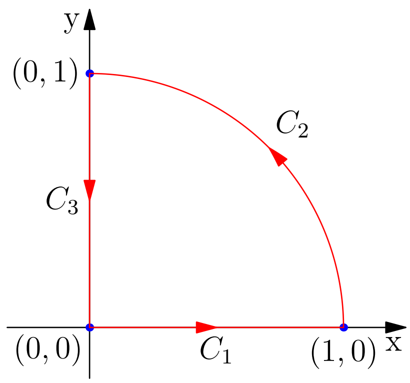


FIGURE 1. The curve C

Let $\vec{F} = -x^3\hat{\imath} + x^2y\hat{\jmath}$.

$$\oint_C \vec{F}\cdot \mathrm{d}\vec{r} = \int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{C_3} \vec{F}\cdot \mathrm{d}\vec{r}.$$

Note that

$$\int_{C_3} \vec{F}\cdot \mathrm{d}\vec{r} = 0,$$

as $\vec{F} = \vec{0}$ along the y-axis. Parametrise C_1 by $x(t) = t$, $y(t) = 0$.

$$\vec{F} = \langle -t^3, 0\rangle \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle 1, 0\rangle\,\mathrm{d}t.$$

So

$$\int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^1 \langle -t^3, 0\rangle \cdot \langle 1, 0\rangle\,\mathrm{d}t = \int_0^1 -t^3\,\mathrm{d}t = \left[-\frac{1}{4}t^4\right]_0^1 = -\frac{1}{4}.$$

Parametrise C_2 by $x(t) = \cos t$, $y(t) = \sin t$.

$$\vec{F} = \langle -\cos^3 t, \cos^2 t \sin t\rangle \qquad \text{and} \qquad \mathrm{d}\vec{r} = \langle -\sin t, \cos t\rangle\,\mathrm{d}t.$$

So

$$\begin{aligned}\int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} &= \int_0^{\pi/2} \langle -\cos^3 t, \cos^2 t \sin t\rangle \cdot \langle -\sin t, \cos t\rangle\,\mathrm{d}t \\ &= \int_0^{\pi/2} 2\cos^3 t \sin t\mathrm{d}t \\ &= \left[-\cos^4 t/2\right]_0^{\pi/2} = 1/2.\end{aligned}$$

In total we get 1/4. We can also use Green's theorem:

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \iint_R \operatorname{curl} \vec{F} \, \mathrm{d}A$$

$$= \int_0^{\pi/2} \int_0^1 r^3 \cos\theta \, \mathrm{d}r\mathrm{d}\theta.$$

The inner integral is

$$\int_0^1 r^3 \cos\theta \, \mathrm{d}r = \left[\frac{1}{4} r^4 \cos\theta\right]_0^1 = \frac{1}{4}\cos\theta.$$

So the outer integral is

$$\int_0^{\pi/2} \frac{1}{4}\cos\theta \, \mathrm{d}\theta = \left[\frac{1}{4}\sin\theta\right]_0^{\pi/2} = \frac{1}{4}.$$

What about the same question, but now let us compute the flux.

$$\oint_C \vec{F} \cdot \hat{n} \, \mathrm{d}s = \int_{C_1} \vec{F} \cdot \hat{n} \, \mathrm{d}s + \int_{C_2} \vec{F} \cdot \hat{n} \, \mathrm{d}s + \int_{C_3} \vec{F} \cdot \hat{n} \, \mathrm{d}s.$$

Once again the flux across C_3 is zero. Along C_1 the normal vector is $-\hat{\jmath}$. So the flux is zero, since $\vec{F}$ is parallel to $\hat{\imath}$ along the x-axis. Along C_2, we have

$$\hat{n} \, \mathrm{d}s = \langle \mathrm{d}y, -\mathrm{d}x \rangle.$$

So

$$\int_{C_1} \vec{F} \cdot \hat{n}\mathrm{d}s = \int_0^{\pi/2} \langle -\cos^3 t, \cos^2 t \sin t \rangle \cdot \langle \cos t, \sin t \rangle \, \mathrm{d}t$$

$$= \int_0^{\pi/2} -\cos^4 t + \cos^2 t \sin^2 t \, \mathrm{d}t$$

$$= \frac{-\pi}{8}.$$

Or we could apply the normal form of Green's theorem:

$$\oint_C \vec{F} \cdot \hat{n} \, \mathrm{d}s = \iint_R \operatorname{div} \vec{F} \, \mathrm{d}A$$

$$= \iint_R -2x^2 \, \mathrm{d}A$$

$$= \int_0^{\pi/2} \int_0^1 -2r^3 \cos^2\theta \, \mathrm{d}r\mathrm{d}\theta.$$

The inner integral is

$$\int_0^1 -2r^3 \cos\theta \, \mathrm{d}r = \left[-\frac{1}{2} r^4 \cos^2\theta\right]_0^1 = -\frac{1}{2}\cos^2\theta.$$

So the outer integral is

$$\int_0^{\pi/2} -\frac{1}{2}\cos^2\theta\,\mathrm{d}\theta = \left[-\frac{t}{4} - \frac{1}{8}\sin(2\theta)\right]_0^{\pi/2} = -\frac{1}{8}\pi.$$

Let

$$\vec{F} = (3x^2 - 2y\sin x\cos x)\hat{\imath} + (a\cos^2 x + 1)\hat{\jmath}.$$

For which values of a is $\vec{F}$ a gradient vector field?

$$M_y = -2\sin x \qquad \text{and} \qquad N_x = -2a\cos x\sin x.$$

These are equal if and only if $a = 1$. For this value of a, what is the integral over the curve C,

$$x(t) = t^2 \qquad \text{and} \qquad y(t) = t^3 - 1,$$

$0 \leq t \leq 1$?

Find a potential function $f(x, y)$. We want

$$f_x = 3x^2 - 2y\sin x\cos x \qquad \text{and} \qquad f_y = \cos^2 x + 1.$$

Integrate the first equation with respect to x,

$$f(x, y) = x^3 - y\cos^2 x + g(y).$$

Use the second equation to determine $g(y)$,

$$-\cos^2 x + \frac{dg}{dy} = \cos^2 x + 1 \qquad \text{so that} \qquad \frac{dg}{dy} = 1.$$

Hence $g(y) = y + c$. So

$$f(x, y) = x^3 - y\cos^2 x + y,$$

will do.

$$\int_C \vec{F}\cdot \mathrm{d}\vec{r} = \int_C \nabla f \cdot \mathrm{d}\vec{r} = f(1, 1) - f(0, 0) = 1.$$

26. Spherical coordinates; applications to gravitation

We have already seen that sometimes it is better to work in cylindrical coordinates. Spherical coordinates (ρ, ϕ, θ) are like cylindrical coordinates, only more so. ρ is the distance to the origin; ϕ is the angle from the z-axis; θ is the same as in cylindrical coordinates.

To get from spherical to cylindrical, use the formulae:

$$\begin{aligned} r &= \rho \sin \phi \\ \theta &= \theta \\ z &= \rho \cos \phi. \end{aligned}$$

As

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

we have

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

On the other hand,

$$\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{r^2 + z^2}.$$

The equation

$$\rho = a,$$

represents the surface of a sphere. On the surface of the sphere, ϕ constant corresponds to *latitude*, although $\phi = 0$ represents the north pole, $\phi = \pi/2$ represents the equator and $\phi = \pi$ represents the south pole. θ constant represents *longitude*.

Question 26.1. *What does the equation*

$$\phi = \pi/4$$

represent?

It represents a cone, through the origin. In cylindrical coordinates we have

$$z = r = \sqrt{x^2 + y^2}.$$

On the other hand, the equation

$$\phi = \pi/2,$$

represents the xy-plane.

We already know the volume element in Cartesian and cylindrical coordinates:

$$\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z.$$

How about in spherical coordinates? We have to calculate the volume of the region when we have a small change in all three coordinates, $\Delta\rho$, $\Delta\theta$ and $\Delta\phi$.

First what happens if we take a sphere of constant radius $\rho = a$? $\Delta\theta$ and $\Delta\phi$ trace out a small region on the surface of the sphere, which is approximately a rectangle. The side corresponding to $\Delta\phi$ is part of the arc of a great circle of radius a. So the length of this side is $a\Delta\phi$. The side corresponding to $\Delta\theta$ is part of the arc of a circle, of radius $r = a\sin\phi$. So the length of this side is $a\sin\phi\Delta\theta$. The area of the region is therefore approximately

$$a^2\sin\phi\Delta\theta\Delta\phi.$$

The volume is then approximately given by

$$\Delta V \approx \rho^2\sin\phi\Delta\theta\Delta\phi\Delta\rho.$$

So

$$\mathrm{d}V = \rho^2\sin\phi\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta.$$

Let's consider again:

Example 26.2. *What is the volume of the region where $z > 1 - y$ and $x^2 + y^2 + z^2 < 1$?*

Note that the closest point on the plane $z = 1 - y$ to the origin is $(1/2, 1/2)$. So the distance of the plane $z = 1 - y$ from the origin is $1/\sqrt{2}$. If we rotate the plane so it is horizontal, we want the volume of the region above the horizontal plane

$$z = \frac{1}{\sqrt{2}},$$

inside the sphere. We can figure this out in cylindrical or spherical coordinates. We carry out the caculation in spherical coordinates for practice.

The plane is given by

$$\rho\cos\phi = z = \frac{1}{\sqrt{2}} \qquad \text{that is} \qquad \rho = \frac{\sec\phi}{\sqrt{2}}.$$

The region is symmetric with respect to θ, so that

$$0 \leq \theta \leq 2\pi.$$

For ϕ we start at the North pole and we go down to $\pi/4$. So the volume is

$$\int_0^{2\pi} \int_0^{\pi/4} \int_{\frac{1}{\sqrt{2}}\sec\phi}^{1} \rho^2 \sin\phi \, \mathrm{d}\rho \, \mathrm{d}\phi \, \mathrm{d}\theta.$$

The force due to gravity on a point mass m at the origin by a body of mass ΔM at (x, y, z) is given by

$$|\vec{F}| = \frac{Gm\Delta M}{\rho^2}.$$

Thus

$$\vec{F} = \frac{Gm\Delta M}{\rho^3} \langle x, y, z \rangle.$$

If we have a body, with mass density δ, then we have to sum together the contributions from each little piece of mass $\Delta M \approx \delta \Delta V$. Thus the force due to gravity on a point mass at the origin is

$$\vec{F} = \iiint_R \frac{Gm\langle x, y, z \rangle}{\rho^3} \delta \, \mathrm{d}V.$$

So the z-component of the force is

$$F_z = \iiint_R \frac{Gmz}{\rho^3} \delta \, \mathrm{d}V.$$

In general, always try to place the point mass at the origin and put the body so that the z-axis is an axis of symmetry (if this is possible). Then

$$\vec{F} = \langle 0, 0, F_z \rangle,$$

and it suffices to compute the z-component. In spherical coordinates, we get

$$\begin{aligned} F_z &= Gm \iiint_R \frac{z}{\rho^3} \delta \, \mathrm{d}V \\ &= Gm \iiint_R \frac{\rho\cos\phi}{\rho^3} \rho^2 \sin\phi \delta \, \mathrm{d}\rho \, \mathrm{d}\phi \, \mathrm{d}\theta \\ &= Gm \iiint_R \delta \cos\phi \sin\phi \, \mathrm{d}\rho \, \mathrm{d}\phi \, \mathrm{d}\theta. \end{aligned}$$

Newton's Theorem To calculate the gravitational attraction of a spherical planet of uniform density, one may treat the sphere as a point mass.

Let's show this is true when the point mass is on the surface of the sphere. Assume the planet has radius a, put the point mass at the origin and make this the south pole of the sphere. Then

$$F_z = Gm \iiint_R \delta \cos\phi \sin\phi \,\mathrm{d}\rho \,\mathrm{d}\phi \,\mathrm{d}\theta$$
$$= Gm \int_0^{2\pi} \int_0^{\pi/2} \int_0^{2a\cos\phi} \delta \cos\phi \sin\phi \,\mathrm{d}\rho \,\mathrm{d}\phi \,\mathrm{d}\theta.$$

The inner integral is

$$\int_0^{2a\cos\phi} \delta \cos\phi \sin\phi \,\mathrm{d}\rho = \Big[\delta \cos\phi \sin\phi \rho\Big]_0^{2a\cos\phi} = 2a\delta \cos^2\phi \sin\phi.$$

The middle integral is

$$\int_0^{\pi/2} 2a\delta \cos^2\phi \sin\phi \mathrm{d}\phi = \Big[-\frac{2}{3}a\delta \cos^3\phi\Big]_0^{\pi/2} = \frac{2}{3}a\delta.$$

The outer integral is

$$\int_0^{2\pi} \frac{2}{3}a\delta \,\mathrm{d}\theta = \Big[\frac{2}{3}a\delta\Big]_0^{2\pi} = \frac{4\pi}{3}a\delta.$$

So the integral is

$$Gm\frac{4\pi}{3}a\delta = \frac{GmM}{a^2},$$

since the mass of the planet is

$$M = \delta\frac{4\pi a^3}{3}.$$

27. Vector fields in space

A vector field in space is given by

$$\vec{F} = P\hat{\imath} + Q\hat{\jmath} + R\hat{k} = \langle P, Q, R\rangle.$$

Here the components, P, Q and R are scalar functions of x, y and z.

$\vec{F}$ could be a force field;

$$\vec{F} = -\frac{c\langle x, y, z\rangle}{\rho^3},$$

is the force due to gravity. There is both an electric $\vec{E}$ and a magnetic field $\vec{B}$. There are velocity fields $\vec{v}$ and gradient vector fields.

In space, we can measure the flux of $\vec{F}$ across a surface S,

$$\iint_S \vec{F} \cdot \hat{n}\, \mathrm{d}S.$$

Here $\hat{n}$ is a unit normal to the surface. There are two choices of $\hat{n}$; we have to choose an orientation, a direction which we decide is positive.

Notation:

$$\mathrm{d}\vec{S} = \hat{n}\, \mathrm{d}S.$$

Suppose that $\vec{F}$ represents the velocity vector field of some fluid. The amount of water that crosses a small piece of surface in unit time is approximately a parallelepiped with area of base ΔS and height $\vec{F} \cdot \hat{n}$,

$$\vec{F} \cdot \hat{n}\Delta S.$$

Suppose

$$\vec{F} = x\hat{\imath} + y\hat{\jmath} + z\hat{k},$$

and S is the surface of a sphere of radius a, centred at the origin. Orient the surface S so that the unit normal points outwards,

$$\hat{n} = \frac{1}{a}\langle x, y, z\rangle.$$

In this case

$$\vec{F} \cdot \hat{n} = \frac{1}{a}(x^2 + y^2 + z^2) = a.$$

Hence

$$\iint_S \vec{F} \cdot \hat{n}\, \mathrm{d}S = \iint_S a\, \mathrm{d}S = 4\pi a^3.$$

Now suppose we work with $\vec{F} = z\hat{k}$. Then

$$\vec{F} \cdot \hat{n} = \frac{z^2}{a}.$$

So the flux is

$$\iint_S \vec{F}\cdot\hat{n}\,\mathrm{d}S = \iint_S \frac{z^2}{a}\,\mathrm{d}S = \int_0^{2\pi}\int_0^{\pi} \frac{a^2\cos^2\phi}{a} a^2\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta.$$

The inner integral is

$$\int_0^{\pi} a^3\cos^2\phi\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta = \left[-\frac{a^3}{3}\cos^3\phi\right]_0^{\pi} = \frac{2a^3}{3}.$$

The outer integral is

$$\int_0^{2\pi}\frac{2a^3}{3}\,\mathrm{d}\theta = \frac{4\pi a^3}{3}.$$

In general, it can be quite hard to parametrise a surface. We will need two parameters to describe the surface and we must express

$$\vec{F}\cdot\hat{n}\,\mathrm{d}S,$$

in terms of them. We must also orient the surface:

Question 27.1. *Can one always orient a surface?*

In fact, somewhat surprisingly, the answer is no. The Möbius band is a surface that cannot be oriented.

To begin with, here are some easy special cases:

(1) If $z = a$ is a horizontal plane then

$$\mathrm{d}\vec{S} = \hat{k}\,\mathrm{d}x\,\mathrm{d}y,$$

(here we choose the upwards orientation).

(2) For the surface of a sphere of radius a centred at the origin then

$$\mathrm{d}\vec{S} = \hat{n}a^2\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta,$$

where

$$\hat{n} = \frac{1}{a}\langle x, y, z\rangle,$$

so that

$$\mathrm{d}\vec{S} = a\sin\phi\langle x, y, z\rangle\,\mathrm{d}\phi\,\mathrm{d}\theta.$$

(3) For a cylinder of radius a centred on the z-axis, use z, θ.

$$\hat{n} = \frac{1}{a}\langle x, y, 0\rangle,$$

which points radially out of the cylinder.

$$\mathrm{d}S = a\,\mathrm{d}z\,\mathrm{d}\theta,$$

so that

$$\mathrm{d}\vec{S} = \langle x, y, 0\rangle\,\mathrm{d}z\,\mathrm{d}\theta.$$

(4) For the graph of a function $f(x, y)$,

$$\mathrm{d}\vec{S} = \langle -f_x, -f_y, 1\rangle \mathrm{d}x\,\mathrm{d}y.$$

28. How to compute the flux

Let's start with the case when S is the graph of a function $z = f(x, y)$ lying over a region R of the plane. If we have a small rectangle with sides Δx and Δy in R then in space we roughly we get a parallelogram with vertices

$$\begin{array}{rl} (x, y, f(x, y)) & \qquad (x + \Delta x, y, f(x + \Delta x, y)) \\ (x, y + \Delta y, f(x, y + \Delta y)) & \qquad (x + \Delta x, y + \Delta y, f(x + \Delta x, y + \Delta y)). \end{array}$$

By linear approximation,

$$f(x+\Delta x, y) \approx f(x, y)+f_x(x, y)\Delta x \qquad \text{and} \qquad f(x, y+\Delta y) \approx f(x, y)+f_y(x, y)\Delta y.$$

and so on. So we have a parallelogram with two sides

$$\vec{v} = \langle \Delta x, 0, f_x(x, y)\Delta x\rangle = \Delta x\langle 1, 0, f_x\rangle \qquad \text{and} \qquad \vec{w} = \langle 0, \Delta y, f_y(x, y)\Delta y\rangle = \Delta y\langle 0, 1, f_y\rangle.$$

The cross product is both a vector normal to the base of the parallelogram and has length the area of the parallelogram. We have

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = \hat{\imath}\begin{vmatrix} 0 & f_x \\ 1 & f_y \end{vmatrix} - \hat{\jmath}\begin{vmatrix} 1 & f_x \\ 0 & f_y \end{vmatrix} + \hat{k}\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -f_x\hat{\imath} - f_y\hat{\jmath} + \hat{k}.$$

It follows that

$$\Delta\vec{S} \approx \vec{v} \times \vec{w} = \Delta x\Delta y\langle -f_x, -f_y, 1\rangle.$$

Taking the limit as Δx and Δy go to zero, we get

$$\mathrm{d}\vec{S} = \langle -f_x, -f_y, 1\rangle\,\mathrm{d}x\,\mathrm{d}y.$$

We can use this to recover

$$\hat{n} = \frac{1}{\sqrt{1 + f_x^2 + f_y^2}}\langle -f_x, -f_y, 1\rangle \qquad \text{and} \qquad \mathrm{d}S = |\mathrm{d}\vec{S}| = \sqrt{1 + f_x^2 + f_y^2}\,\mathrm{d}x\,\mathrm{d}y.$$

In practice, it is usually better not to find the separate pieces.

Question 28.1. *Find the flux of $\vec{F} = z\hat{k}$ across the surface S given by the paraboloid $z = x^2 + y^2$ above the circle R in the xy-plane, given by $x^2 + y^2 \leq 1$, oriented so the normal points upwards (which is into the paraboloid).*

$$\mathrm{d}\vec{S} = \langle -2x, -2y, 1\rangle\,\mathrm{d}x\,\mathrm{d}y.$$

Hence

$$\vec{F} \cdot \mathrm{d}\vec{S} = z\,\mathrm{d}x\,\mathrm{d}y.$$

So the flux is

$$\iint_S \vec{F} \cdot \mathrm{d}\vec{S} = \iint_R z \,\mathrm{d}x\,\mathrm{d}y = \int_0^{2\pi} \int_0^1 r^3 \,\mathrm{d}r\,\mathrm{d}\theta.$$

The inner integral is

$$\int_0^1 r^3 \,\mathrm{d}r = \left[\frac{r^4}{4}\right]_0^1 = \frac{1}{4}.$$

The outer integral is

$$\int_0^{2\pi} \frac{1}{4} \,\mathrm{d}\theta = \frac{\pi}{2}.$$

More generally, suppose S is given as a parametric surface $x(u, v)$, $y(u, v)$ and $z(u, v)$, then we can integrate using u and v, so that $\vec{r} = \vec{r}(u, v)$. Arguing as above,

$$\mathrm{d}\vec{S} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \,\mathrm{d}u\,\mathrm{d}v.$$

Apart from parametrisations, a surface S might be given by a constraint. S might be given implicitly, by an equation $g(x, y, z) = 0$. In this case

$$\vec{N} = \nabla g = \langle g_x, g_y, g_z \rangle,$$

is normal to S and so

$$\hat{n} = \frac{\vec{N}}{|\vec{N}|}$$

is a unit normal.

To get ΔS, consider the projection to the xy-plane (assume that the plane is not vertical; if it is vertical, just project onto the xz-plane or the yz-plane). The key point is to figure out how area changes under projection. I claim

$$\Delta A = \cos\alpha \Delta S,$$

where α is the angle of the surface with the horizontal, that is, the angle between $\vec{N}$ and the vertical $\hat{k}$. The reason for this is the same reason that the projection of a circle, lying in a slanted plane, is an ellipse. Note that every plane contains one horizontal line. To figure out how area changes under projection, one can rotate the plane so that this line is the y-axis. So lengths in the y direction are unchanged. In the x-direction, one gets a right angled triangle. The original length is a hypotenuse and the new length is the adjacent. So lengths in the x-direction scale by $\cos\alpha$. In total the area scales by $\cos\alpha$. But

$$\vec{N} \cdot \hat{k} = |N| \cos\alpha.$$

Putting all of this together,

$$\mathrm{d}\vec{S} = \frac{\vec{N}}{|\vec{N}.\hat{k}|}\,\mathrm{d}x\,\mathrm{d}y.$$

Example 28.2. *Suppose that*

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

so that $g = 0$ defines the graph of f. Then

$$\vec{N} = \nabla g = \langle -f_x, -f_y, 1\rangle \qquad \textit{and} \qquad \vec{N}\cdot\hat{k} = 1,$$

so we get the old formula.

Theorem 28.3 (Divergence Theorem)**.** *Let S be a closed surface bounding a solid D, oriented outwards. Let $\vec{F}$ be a vector field with continuous partial derivatives. Then*

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S} = \iiint_D \operatorname{div}\vec{F}\,\mathrm{d}V \qquad \textit{where} \qquad \operatorname{div}\vec{F} = P_x + Q_y + R_z.$$

This has the same physical interpretation as before. The total amount of material leaving S is equal to the amount of material created (or destroyed) inside the solid D.

Example 28.4. *Let $\vec{F} = z\hat{k}$ and let S be the surface of a sphere of radius a.*

$$\operatorname{div}\vec{F} = 0 + 0 + 1 = 1,$$

and so

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S} = \iiint_D \operatorname{div}\vec{F}\,\mathrm{d}V = \iiint_D 1\,\mathrm{d}V = \frac{4}{3}\pi a^3.$$

It is convenient to introduce some symbolic notation.

$$\nabla = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\rangle$$

is called the del operator.

$$\nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\rangle$$

is the gradient. We have

$$\operatorname{div}\vec{F} = \nabla\cdot\vec{F} = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\rangle \cdot \langle P, Q, R\rangle = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

is the divergence.

29. The divergence theorem

Theorem 29.1 (Divergence Theorem; Gauss, Ostrogradsky)**.** *Let S be a closed surface bounding a solid D, oriented outwards. Let $\vec{F}$ be a vector field with continuous partial derivatives. Then*

$$\iint_S \vec{F} \cdot \mathrm{d}\vec{S} = \iiint_D \nabla \cdot \vec{F}\, \mathrm{d}V.$$

Why is

$$\nabla \cdot \vec{F} = \operatorname{div} \vec{F} = P_x + Q_y + R_z$$

a measure of the amount of material created (or destroyed) at (x, y, z)? Well imagine a small box with one vertex at (x, y, z) and edges Δx, Δy and Δz. The flux through this box is the sum of the flux through the six sides. We can pair off opposite sides. Consider the sides parallel to the xy-plane, that is, orthogonal to the vector $\hat{k}$. Crossing the bottom side is approximately

$$R(x, y, z)\Delta x\Delta y,$$

Crossing the top side is approximately

$$R(x, y, z + \Delta z)\Delta x\Delta y.$$

By linear approximation,

$$R(x, y, z + \Delta z) \approx R(x, y, z) + R_z\Delta z,$$

and so the difference is approximately

$$R_z\Delta x\Delta y\Delta z.$$

By symmetry the contribution from the other two sides is approximately

$$P_x\Delta x\Delta y\Delta z \qquad \text{and} \qquad Q_y\Delta x\Delta y\Delta z.$$

Adding this together, we get

$$(P_x + Q_y + R_z)\Delta x\Delta y\Delta z,$$

which is approximately the amount of water being created or destroyed in the small box. Dividing through by

$$\Delta x\Delta y\Delta z$$

and taking the limit, we get the divergence.

Proof of (29.1). We argue as in the proof of Green's theorem. Firstly, we can prove three separate identities, one for each of P, Q and R. So we just need to prove

$$\iint_S \langle 0, 0, R\rangle \cdot \mathrm{d}\vec{S} = \iiint_D R_z\, \mathrm{d}V.$$

Now divide the region into small pieces, each of which is vertically simple, so that

$$a \leq x \leq b \qquad c \leq y \leq d \qquad \text{and} \qquad f(x,y) \leq z \leq g(x,y),$$

is the region lying over a rectangle in the xy-plane lying between the graph of two functions f and g.

It is enough, because of cancelling, to prove the result for such a region. We calculate both sides explicitly in this case. S has six sides; the four vertical sides and the top and bottom sides. The flux across the four vertical sides is zero, since $\vec{F}$ is moving up and down. The flux across the top side is

$$\begin{aligned}\iint_{S_{\text{top}}} \langle 0,0,R\rangle \cdot \mathrm{d}\vec{S} &= \iint_{S_{\text{top}}} \langle 0,0,R\rangle \cdot \langle -g_x, -g_y, 1\rangle \,\mathrm{d}x\,\mathrm{d}y \\ &= \int_c^d \int_a^b R(x,y,g(x,y))\,\mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

The flux across the bottom side is similar, but it comes with the opposite sign, so the total flux is

$$\int_c^d \int_a^b R(x,y,g(x,y)) - R(x,y,f(x,y))\,\mathrm{d}x\,\mathrm{d}y.$$

For the RHS, we have a triple integral,

$$\iiint_D R_z \,\mathrm{d}V = \int_c^d \int_a^b \int_{f(x,y)}^{g(x,y)} R_z(x,y,z)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y.$$

The inner integral is

$$\int_{f(x,y)}^{g(x,y)} R_z(x,y,z)\,\mathrm{d}z = \Big[R(x,y,z)\Big]_{f(x,y)}^{g(x,y)} = R(x,y,g(x,y)) - R(x,y,f(x,y)),$$

so that both sides are indeed the same. □

What can we say about a radially symmetric vector field $\vec{F}$, such that there is a single unit source at the origin and otherwise the divergence is zero? By radial symmetry, we have

$$\vec{F} = \frac{\langle x,y,z\rangle}{c},$$

for some c, to be determined. Consider the flux across a sphere S of radius a. The flux across S is

$$\iint_S \vec{F} \cdot \mathrm{d}\vec{S}.$$

If we orient S so that the unit normal points outwards, we have

$$\hat{n} = \frac{1}{a}\langle x, y, z\rangle \qquad \text{and} \qquad \vec{F}\cdot\hat{n} = \frac{a}{c}.$$

So the flux is

$$\frac{4\pi a^3}{c}.$$

Since the only source of water is at the origin and we are pumping in water at a rate of one unit there, we must have

$$\frac{4\pi a^3}{c} = 1 \qquad \text{so that} \qquad c = 4\pi a^3.$$

That is

$$\vec{F} = \frac{1}{4\pi}\frac{\langle x, y, z\rangle}{\rho^3}.$$

Typical examples of such force fields are gravity and electric charge.

30. Line integrals

If
$$\vec{F} = P\hat{\imath} + Q\hat{\jmath} + R\hat{k},$$
is a vector field on $\mathbb{R}^3$ and C is a curve in space then we can define the line integral
$$\int_C \vec{F} \cdot \mathrm{d}\vec{r},$$
in the same way as we did in the plane. If we pick a parametrisation of C,
$$\vec{r}(t) = \langle x(t), y(t), z(t)\rangle, \qquad \text{and} \qquad a \leq t \leq b$$
then we can express
$$\mathrm{d}\vec{r} = \langle \mathrm{d}x, \mathrm{d}y, \mathrm{d}z\rangle = \vec{v}(t)\,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle P, Q, R\rangle$$
in terms of t and integrate this. If $\vec{F}$ represents force, the line integral represents the work done moving a particle from the start P_0 to the end P_1.

Example 30.1. *Let C be the parametric curve*
$$\vec{r}(t) = \langle t, t^2, t^3\rangle \qquad 0 \leq t \leq 1 \qquad \textit{and} \qquad \vec{F} = \langle yz, xz, xy\rangle.$$
C is called the twisted cubic.

We express everything in terms of t,
$$\mathrm{d}\vec{r} = \langle 1, 2t, 3t^2\rangle\,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle t^5, t^4, t^3\rangle.$$
So
$$\vec{F} \cdot \mathrm{d}\vec{r} = \langle t^5, t^4, t^3\rangle \cdot \langle 1, 2t, 3t^2\rangle\,\mathrm{d}t = 6t^5\,\mathrm{d}t.$$
The work done is then
$$\int_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_0^1 6t^5\,\mathrm{d}t = \Big[t^6\Big]_0^1 = 1.$$

Now let's suppose we go from $P_0 = (0, 0, 0)$ to $P_1 = (1, 1, 1)$ along a different path. Let's say we go parallel to $\hat{\imath}$, C_1, parallel to $\hat{\jmath}$, C_2 and then parallel to $\hat{k}$, C_3. Let
$$C' = C_1 + C_2 + C_3.$$
So
$$\int_{C'} \vec{F} \cdot \mathrm{d}\vec{r} = \int_{C_1} \vec{F} \cdot \mathrm{d}\vec{r} + \int_{C_2} \vec{F} \cdot \mathrm{d}\vec{r} + \int_{C_3} \vec{F} \cdot \mathrm{d}\vec{r}.$$
Parametrise C_1 in the obvious way
$$\vec{r}(t) = \langle t, 0, 0\rangle \qquad 0 \leq t \leq 1.$$

Then
$$\mathrm{d}\vec{r} = \langle 1, 0, 0\rangle \,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle 0, 0, 0\rangle.$$
So
$$\int_{C_1} \vec{F}\cdot \mathrm{d}\vec{r} = 0.$$
Parametrise C_2
$$\vec{r}(t) = \langle 1, t, 0\rangle \qquad 0 \leq t \leq 1.$$
Then
$$\mathrm{d}\vec{r} = \langle 0, 1, 0\rangle \,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle 0, 0, t\rangle.$$
We have
$$\vec{F}\cdot \mathrm{d}\vec{r} = \langle 0, 0, 1\rangle \cdot \langle 0, 1, 0\rangle \,\mathrm{d}t = 0\,\mathrm{d}t.$$
So
$$\int_{C_2} \vec{F}\cdot \mathrm{d}\vec{r} = 0.$$
Parametrise C_3
$$\vec{r}(t) = \langle 1, 1, t\rangle \qquad 0 \leq t \leq 1.$$
Then
$$\mathrm{d}\vec{r} = \langle 0, 0, 1\rangle \,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle t, t, 1\rangle.$$
We have
$$\vec{F}\cdot \mathrm{d}\vec{r} = \langle t, t, 1\rangle \cdot \langle 0, 0, 1\rangle \,\mathrm{d}t = \mathrm{d}t.$$
So
$$\int_{C_3} \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^1 \mathrm{d}t = 1.$$

The reason why both answers are the same is that $\vec{F}$ is a gradient vector field,
$$\vec{F} = \langle yz, xz, xy\rangle = \nabla(xyz) = \nabla f,$$
where $f(x, y, z) = xyz$. As before we have the fundamental theorem of calculus for line integrals
$$\int_C \nabla f\cdot \mathrm{d}\vec{r} = f(P_1) - f(P_0).$$
As before this means the integral is path independent and $\vec{F}$ is conservative. In our case
$$f(0, 0, 0) = 0 \qquad f(1, 1, 1) = 1, \qquad \text{so that} \qquad f(1, 1, 1) - f(0, 0, 0) = 1.$$

Theorem 30.2. *Let $\vec{F}$ be a vector field on $\mathbb{R}^3$ (or more generally a simply connected region; more about this later).*

$\vec{F}$ is a gradient vector field if and only if
$$P_y = Q_x, \qquad P_z = R_x \qquad \textit{and} \qquad Q_z = R_y.$$

Again, one direction is reasonably clear. If

$$\vec{F} = \nabla f,$$

then

$$P = f_x, \qquad Q = f_y \qquad \text{and} \qquad R = f_z,$$

so that

$$P_y = f_{xy} \qquad \text{and} \qquad Q_x = f_{yx},$$

so that the equality $P_y = Q_x$ is just saying the mixed partials of f are equal.

Example 30.3. *For which a and b is*

$$\vec{F} = axy\hat{\imath} + (x^2 + z^3)\hat{\jmath} + (byz^2 - 4z^3)\hat{k}$$

a gradient vector field?

We have

$$P = axy, \qquad Q = x^2 + z^3 \qquad \text{and} \qquad R = byz^2 - 4z^3.$$

$Q_x = 2x$ and $P_y = ax$, so that

$$ax = P_y = Q_x = 2x,$$

and so $a = 2$. $Q_z = 3z^2$ and $R_y = bz^2$ so that

$$bz^2 = R_y = Q_z = 3z^2$$

and so $b = 3$. Finally,

$$P_z = 0 \qquad \text{and} \qquad R_x = 0,$$

so $P_z = R_x$ is clear. Hence

$$\vec{F} = 2xy\hat{\imath} + (x^2 + z^3)\hat{\jmath} + (3yz^2 - 4z^3)\hat{k}$$

is conservative.

Let's look for a potential function $f(x, y, z)$. We want to solve three PDEs

$$\frac{\partial f}{\partial x} = 2xy, \qquad \frac{\partial f}{\partial y} = x^2 + z^3 \qquad \text{and} \qquad \frac{\partial f}{\partial z} = 3yz^2 - 4z^3.$$

We have

$$\frac{\partial f}{\partial x} = 2xy.$$

Integrate both sides with respect to x.

$$f(x, y, z) = \int 2xy \, \mathrm{d}x = x^2y + g(y, z).$$

Note that the constant of integration is actually a function of both y and z. In other words

$$\frac{\partial g(y,z)}{\partial x} = 0.$$

Let's take this answer for f and plug this into the second PDE.

$$\frac{\partial f}{\partial y} = x^2 + z^3 \qquad \text{so that} \qquad x^2 + \frac{\partial g(y,z)}{\partial z} = x^2 + z^3.$$

Cancelling we get

$$\frac{\partial g(y,z)}{\partial y} = z^3.$$

Integrating this with respect to y we get

$$g(y,z) = \int z^3 \,\mathrm{d}y = z^3 y + h(z),$$

where the constant of integration is an abritrary function of z. So now we know

$$f(x,y,z) = x^2 y + z^3 y + h(z).$$

Finally, we plug this back into the third PDE:

$$\frac{\partial f}{\partial z} = 3yz^2 - 4z^3 \qquad \text{so that} \qquad 3yz^2 + \frac{dh}{dz} = 3yz^2 - 4z^3.$$

Cancelling we get

$$\frac{dh}{dz} = -4z^3.$$

Integrating with respect to z, we get

$$h(z) = -z^4 + c.$$

We can take $c = 0$. Putting all of this together

$$f(x,y,z) = x^2 y + z^3 y - z^4.$$

One can define a vector field which measures how far $\vec{F}$ is from being conservative.

$$\begin{aligned}
\operatorname{curl} \vec{F} &= \nabla \times \vec{F} \\
&= \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \\
&= \hat{\imath} \begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ Q & R \end{vmatrix} - \hat{\jmath} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial z} \\ P & R \end{vmatrix} + \hat{k} \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{vmatrix} \\
&= (R_y - Q_z)\hat{\imath} - (R_x - P_z)\hat{\jmath} + (Q_x - P_y)\hat{k}.
\end{aligned}$$

This is the curl of the vector field $\vec{F}$. It measures the angular velocity. For example,

$$\vec{v} = \langle -\omega y, \omega x, 0 \rangle,$$

represents rotation around z-axis with constant angular velocity ω.

$$\operatorname{curl} \vec{v} = \nabla \times \vec{v} = 2\omega \hat{k}.$$

So the magnitude is twice the angular velocity and the direction is the axis of rotation.

31. Stokes Theorem

Stokes' theorem is to Green's theorem, for the work done, as the divergence theorem is to Green's theorem, for the flux. Both are 3D generalisations of 2D theorems.

Theorem 31.1 (Stokes' Theorem)**.** *Let C be any closed curve and let S be any surface bounding C. Let $\vec{F}$ be a vector field on S.*

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \iint_S (\nabla \times \vec{F}) \cdot \hat{n} \,\mathrm{d}S.$$

Note that S is an oriented surface. How do we orient S? We use the orientation on C.

If we drive along C, in the positive direction, with S on the left, then $\hat{n}$ should point upwards (with respect to the driver; that is to say, if you ask the driver to point to the roof of the car, this is the direction we should orient $\hat{n}$).

Put differently, we can use the right hand rule. If our index finger points along C, the middle finger points into S then the thumb points in the direction of $\hat{n}$.

Here are some examples:

(1) If C is the unit circle in the xy-plane, oriented counterclockwise and S is the upper hemisphere of the unit sphere, then $\hat{n}$ points outwards.

(2) If S is the half circular unit cylinder $x^2 + y^2 = 1$, $y \geq 0$, $0 \leq z \leq 1$, and C is the boundary curve, starting at $(1, 0, 0)$, going around to $(-1, 0, 0)$, up to $(-1, 0, 1)$, going around to $(1, 0, 1)$ and down to $(1, 0, 0)$, then $\hat{n}$ points outwards.

(3) If S is the cone with vertex at $(0, 0, 1)$ and base $x^2+y^2 = 1$ in the xy-plane and we are orient the unit circle C counterclockwise, then $\hat{n}$ points outwards.

Suppose that C is a curve in the xy-plane, oriented counterclockwise. C bounds a region S in the xy-plane. If

$$\vec{F} = M\hat{\imath} + N\hat{\jmath},$$

then Green's theorem says

$$\oint_C M \,\mathrm{d}x + N \,\mathrm{d}y = \oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_S \operatorname{curl} \vec{F} \,\mathrm{d}A = \int_S (N_x - M_y) \,\mathrm{d}A.$$

On the other hand, $\hat{n} = \hat{k}$, so that Stokes's theorem says

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_S (\nabla \times \vec{F}) \cdot \hat{n} \,\mathrm{d}A = \int_S (N_x - M_y) \,\mathrm{d}A.$$

So the two theorems are equivalent in this case.

Note one very interesting aspect of Stokes' theorem is that there are very many different surfaces S which bound the same curve C. For example, for the unit circle, the upper hemisphere bounds C, the lower hemisphere bounds C, the cone bounds C, and so on.

Proof of (31.1). We have already seen that if C and S lie in the xy-plane then Stokes' theorem reduces to Green's theorem.

If S and C live in an arbitrary plane, then imagine first translating the plane so that it contains the origin and then rotating it so that it is flat, that is, so it is the xy-plane.

The work done is invariant under both translation and rotation, so that the LHS of Stokes' theorem does not depend on the plane. The same is true for the RHS (this takes a little more justification). So (31.1) holds if S and C lie in any plane.

In the general case, we can subdivide S into small pieces (like a quilted blanket). As usual it is enough to prove (31.1) for each small quilt. If the piece is small enough, S and C live approximately in a plane and we can appeal to Green's theorem. □

Example 31.2. *Let's check that* (31.1) *holds in a special case. Let*

$$\vec{F} = z\hat{\imath} + x\hat{\jmath} + y\hat{k},$$

let C be the unit circle in the xy-plane and S be the paraboloid $z = 1 - x^2 - y^2$, $z > 0$.

We compute the LHS. We parametrise C by

$$x(t) = \cos t, \qquad y(t) = \sin t, \qquad z = 0 \qquad \text{where} \qquad 0 \leq t \leq 2\pi.$$

In this case

$$\mathrm{d}\vec{r} = \langle -\sin t, \cos t, 0\rangle \,\mathrm{d}t \qquad \text{and} \qquad \vec{F} = \langle 0, \cos t, \sin t\rangle.$$

So

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \int_0^{2\pi} \cos^2 t \,\mathrm{d}t = \left[\frac{1}{2}(1 + \frac{1}{2}\sin 2t)\right]_0^{2\pi} = \pi.$$

Now let's compute the RHS.

$$\operatorname{curl} \vec{F} = \nabla \times \vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z & x & y \end{vmatrix} = \hat{\imath} + \hat{\jmath} + \hat{k}.$$

The paraboloid

$$z = 1 - x^2 - y^2 = f(x, y)$$

is the graph of a function. So,

$$\mathrm{d}\vec{S} = \langle -f_x, -f_y, 1\rangle \,\mathrm{d}x\,\mathrm{d}y = \langle 2x, 2y, 1\rangle \,\mathrm{d}x\,\mathrm{d}y.$$

Note that this is the outwards orientation and that C is oriented counterclockwise, so C and S are compatibly oriented. Let R be the unit disk.

$$\iint_S (\nabla \times \vec{F}) \cdot \hat{n}\,\mathrm{d}S = \iint_R (1 + 2x + 2y)\,\mathrm{d}A = \pi.$$

Here we used the fact that x is anti-symmetric about the y-axis, so that

$$\iint_R 2x\,\mathrm{d}A = 0.$$

Similarly

$$\iint_R 2y\,\mathrm{d}A = 0,$$

as y is anti-symmetric about the x-axis. Finally the area of R is π.

32. Topology and beyond

Definition 32.1. *We say that a region R is simply connected if every closed curve C bounds a surface S.*

(1) $\mathbb{R}^3$ is simply connected.
(2) $\mathbb{R}^3$ minus a line is not simply connected.
(3) $\mathbb{R}^3$ minus a point is simply connected.
(4) $\mathbb{R}^3$ minus a circle is not simply connected.
(5) $\mathbb{R}^3$ minus a line segment is simply connected.

This is related to topology, which deals with the classification of geometric objects up to deforming them like pieces of rubber (so you can stretch but not tear). The surface of a sphere is topologically different from the surface of a torus. The sphere is simply connected but the torus is not.

Theorem 32.2. *If R is simply connected region in $\mathbb{R}^3$ then then $\vec{F}$ is conservative if and only if* $\operatorname{curl} \vec{F} = \vec{0}$

Proof. Suppose that $\vec{F}$ is conservative. Then $\vec{F} = \nabla f$. We have already seen that $\operatorname{curl} \vec{F}$ is then zero but it does not hurt to write down the proof again.

$$\vec{F} = \nabla f = \langle f_x, f_y, f_z \rangle.$$

Therefore

$$\operatorname{curl} \vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_x & f_y & f_z \end{vmatrix} = (f_{zy} - f_{yz})\hat{\imath} - (f_{zx} - f_{xz})\hat{\jmath} + (f_{yx} - f_{xy})\hat{k} = \vec{0}.$$

Now suppose $\operatorname{curl} \vec{F} = \vec{0}$. Let C be a closed loop. Pick an orientable surface S which bounds C and orient S compatibly with C.

Then Stokes' theorem says

$$\oint_C \vec{F} \cdot \mathrm{d}\vec{r} = \iint_S (\operatorname{curl} \vec{F}) \cdot \hat{n} \mathrm{d}S = 0. \qquad \square$$

Note that we have to be careful to say that S is orientable. The Möbius strip bounds a closed curve, but it is not orientable.

Suppose that S_1 and S_2 are two surfaces which bound the same curve C. Then Stokes' Theorem says

$$\iint_{S_1} \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S = \oint_C \vec{F} \cdot \mathrm{d}r = \iint_{S_2} \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S.$$

So the integral of the curl of $\vec{F}$ is the same for both surfaces S_1 and S_2.

In fact we can prove this in another way. Imagine constructing a new surface S by joining S_1 and S_2 along their common boundary C. Then S is a closed surface.

Note that one can reverse this process. If you start with a closed surface S and pick a closed curve C in S then you can always cut S along C to obtain two surfaces S_1 and S_2 with a common boundary C. Perhaps the easiest example of all of this is to take a sphere S and the equator C, in which case S_1 and S_2 are the upper and lower hemispheres.

Note that S_1 and S_2 must have opposite orientations to join them together; algebraically $S = S_1 - S_2$. Possibly switching S_1 and S_2 so that the orientation is outwards we can apply the divergence theorem to the closed surface S and the region it bounds D.

$$\iint_S \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S = \iiint_D \operatorname{div}(\operatorname{curl} \vec{F}) \, \mathrm{d}V.$$

Now

$$\operatorname{div}(\operatorname{curl} \vec{F}) = \nabla \cdot (\nabla \times \vec{F}).$$

This suggests that the divergence of a curl is always zero. We check this by direct computation:

$$\nabla \times \vec{F} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = (R_y - Q_z)\hat{\imath} + (P_z - R_x)\hat{\jmath} + (Q_x - P_y)\hat{k}.$$

The divergence of this is

$$R_{yx} - Q_{zx} + P_{zy} - R_{xy} + Q_{xz} - P_{yz} = 0.$$

So

$$\iint_{S_1 - S_2} \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S = 0.$$

But then

$$\iint_{S_1} \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S = \iint_{S_2} \operatorname{curl} \vec{F} \cdot \hat{n} \, \mathrm{d}S.$$

Note that the identity

$$\operatorname{div}(\operatorname{curl} \vec{F}) = 0,$$

is very useful on its own.

Note that one can compose in the opposite direction. If f is a scalar function then ∇f is a vector field, and we can take the curl of this.

$$\operatorname{curl}(\nabla f) = \nabla \times (\nabla f) = \vec{0}.$$

We have already seen that this is the zero vector. This has an interesting physical interpretation. Recall that $\operatorname{curl} \vec{F}$ measures the rotation component of a vector field. So the fact that $\mathrm{curl}(\nabla f) = 0$ says that every force field given by a potential imparts no rotation component. For example gravity imparts no rotation.

33. Review I

Example 33.1. *We have two lines in $\mathbb{R}^3$, one given parametrically by*

$$\vec{r}_1(t) = \langle -4+5t, 1+t, -2-t\rangle,$$

and the other given as the intersection of the two planes:

$$2x - y - z = 6 \qquad \textit{and} \qquad x + y - 2z = 3.$$

What is the shortest distance between these two lines?

There are many different ways to solve this problem but all of them start the same way, by first finding the equation of the second line parametrically.

Note that each equation determines a plane and the intersection of two planes is a line. A line is specified by two points. So we want to find two points on the line.

To get a point on a line intersect with a plane. Let's intersect with the plane $z = 0$. The two equations reduce to

$$\begin{aligned} 2x - y &= 6 \\ x + y &= 3. \end{aligned}$$

This is an inhomogeneous system of linear equation. We can rewrite this as a matrix equation:

$$A\vec{x} = \vec{b} \qquad \text{where} \qquad \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

This has a unique solution if and only if $\det A \neq 0$.

$$\det A = \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 3.$$

As the determinant is not zero, A is invertible:

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

If we are given A^{-1}, then it is easy to solve $A\vec{x} = \vec{b}$:

$$\vec{x} = A^{-1}\vec{b} \qquad \text{that is} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 6 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

So one point on the line is $P = (3, 0, 0)$.

If we take the plane $z = -1$ then

$$\begin{aligned} 2x - y &= 5 \\ x + y &= 1. \end{aligned}$$

Arguing as above,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

So another point on the line is $Q = (2, -1, -1)$. The parametric form of the second line is then

$$\vec{r}_2(t) = \vec{P} + t\overrightarrow{PQ} = \langle 3, 0, 0\rangle + t\langle -1, -1, -1\rangle = \langle 3 - t, -t, -t\rangle.$$

So as to work with pluses rather than minuses, we switch the sign of t to get the same line parametrised with the opposite orientation:

$$\vec{r}_2(t) = \langle 3 + t, t, t\rangle.$$

Now we turn to the problem of finding the two closest points P_1 and P_2 belonging to the two lines. There are many ways to solve this problem. The first three ways use the geometric fact that the vector $\overrightarrow{P_1P_2}$ is orthogonal to the direction of both lines.

Method #1: Find a plane Π containing the second line which is parallel to the first line. Π is orthogonal to the direction of both lines:

$$\vec{u} = \langle 5, 1, -1\rangle \qquad \text{and} \qquad \vec{v} = \langle 1, 1, 1\rangle.$$

The cross product of $\vec{u}$ and $\vec{v}$ is therefore orthogonal to both lines:

$$\begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 5 & 1 & -1 \\ 1 & 1 & 1 \end{vmatrix} = \hat{\imath}\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} - \hat{\jmath}\begin{vmatrix} 5 & -1 \\ 1 & 1 \end{vmatrix} + \hat{k}\begin{vmatrix} 5 & 1 \\ 1 & 1 \end{vmatrix} = 2\hat{\imath} - 6\hat{\jmath} + 4\hat{k}.$$

So

$$\vec{n} = \langle 1, -3, 2\rangle$$

is a normal vector to the plane Π. Π contains $(2, -1, -1)$ (set $t = -1$ just to get a more interesting point than $(3, 0, 0)$). So the equation of the plane is

$$\langle x-2, y+1, z+1\rangle\cdot\langle 1, -3, 2\rangle = 0 \qquad \text{that is} \qquad (x-2)-3(y+1)+2(z+1) = 0,$$

so that rearranging we have $x - 3y + 2z = 3$. Pick any point of the first line. If we set $t = 0$ we get $R = (-4, 1, -2)$.

The line through R parallel to $\vec{n}$ intersects the plane Π at the closest point R' to R. This line is

$$\vec{r}(t) = \langle -4, 1, -2\rangle + t\langle 1, -3, 2\rangle = \langle -4 + t, 1 - 3t, -2 + 2t\rangle.$$

This lies on the plane Π when

$$(t - 4) - 3(1 - 3t) + 2(-2 + 2t) = 3 \qquad \text{so that} \qquad 14t = 14$$

But then $t = 1$. The closest point R' is $(-1, -2, 0)$.

$$\overrightarrow{RR'} = \langle 1, -3, 2\rangle,$$

so the shortest distance is

$$\sqrt{(1+9+4)} = \sqrt{14}.$$

Method #2: Find the two closest points P_1 and P_2 directly. Choose different parametrisations for the first and second line:

$$\vec{r}_1(s) = \langle -4+5s, 1+s, -2-s\rangle \qquad \text{and} \qquad \vec{r}_2(t) = \langle 3+t, t, t\rangle.$$

Then

$$\overrightarrow{P_1P_2} = \langle 7-5s+t, -1-s+t, 2+s+t\rangle.$$

We want this vector to be orthogonal to both $\vec{u}$ and $\vec{v}$:

$$\langle 7-5s+t, -1-s+t, 2+s+t\rangle\cdot\langle 5,1,-1\rangle = 0 \qquad \text{and} \qquad \langle 7-5s+t, -1-s+t, 2+s+t\rangle\cdot\langle 1,1,1\rangle = 0.$$

$$\begin{aligned} -27s + 5t &= -32 \\ -5s + 3t &= -8. \end{aligned}$$

Using guess and check, we see that $s = 1$ and $t = -1$ works. The two closest points are

$$P_1 = (1,2,-3) \qquad \text{and} \qquad P_2 = (2,-1,-1).$$

As before

$$\overrightarrow{P_1P_2} = \langle 1,-3,2\rangle,$$

and the shortest distance if again $\sqrt{14}$.

Method #3: Pick two random points on both lines.

$$R_1 = \langle -4,1,-2\rangle \qquad \text{and} \qquad R_2 = \langle 3,0,0\rangle.$$

Then the distance we want is given by the projection of

$$\overrightarrow{R_1R_2} = \langle 7,-1,2\rangle$$

onto $\vec{n} = \langle 1,-3,2\rangle$. The length of the projection is given by

$$\frac{\langle 7,-1,2\rangle\cdot\langle 1,-3,2\rangle}{|\langle 1,-3,2\rangle} = \frac{7+3+4}{\sqrt{14}} = \sqrt{14}.$$

Method #4: Like method #2, but now use calculus to minimise the distance between P_1 and P_2. Note that if we minimise the distance or the distance squared we get the same points. In practice we minimise the distance squared, since this gives much simpler equations:

$$f(s,t) = (7-5s+t)^2+(-1-s+t)^2+(2+s+t)^2 = 27s^2-10st+3t^2-64s+16t+54.$$

Find the critical points. First find partials:

$$f_s = 54s - 10t - 64 \qquad \text{and} \qquad f_t = -10s + 6t + 16.$$

Set these equal to zero:

$$\begin{aligned} 27s - 5t &= 32 \\ -5s + 3t &= -16 \end{aligned}$$

We already know $s = 1$ and $t = -1$ will work.

Geometrically it is clear that we must have a minimum and there is no maximum. But, again just to practice, let's use the second derivative test to check we have a minimum:

$$f_{ss} = 54 \qquad f_{st} = -10 \qquad \text{and} \qquad f_{tt} = 6.$$

Hence

$$A = 54 \qquad B = -10 \qquad \text{and} \qquad C = 6.$$

$AC - B^2 > 0$. $A > 0$ so we have a minimum.

34. Review II

It is probably helpful to take stock of the various integrals and differentials we have encountered in this course:

Dimension	Standard	Vector
1	$\mathrm{d}t$, $\mathrm{d}s$	$\mathrm{d}\vec{r}$
2	$\mathrm{d}A$	$\mathrm{d}\vec{S}$
3	$\mathrm{d}V$	Not covered

In dimension one the most basic integral is the line integral:

$$\int_C \vec{F}\cdot \mathrm{d}\vec{r} = \int_C M\,\mathrm{d}x + N\,\mathrm{d}y.$$

This integral represents the work done to move a particle along C in a vector field $\vec{F}$. To compute directly, parametrise C. If we use the parameter t, we will get down to a standard one dimensional integral. For example, suppose that

$$\vec{F} = x\hat{\imath} + y\hat{\jmath}$$

and C is the unit circle, oriented counterclockwise. Paremetrise C in the standard way:

$$\vec{r}(t) = \langle \cos t, \sin t\rangle \qquad \text{where} \qquad 0 \leq t \leq 2\pi.$$

Then

$$\mathrm{d}\vec{r} = \langle -\sin t, \cos t\rangle\,\mathrm{d}t \qquad \text{and} \qquad \langle \cos t, \sin t\rangle.$$

Therefore

$$\oint_C \vec{F}\cdot \mathrm{d}\vec{r} = \int_0^{2\pi} 0\mathrm{d}t = 0.$$

One can also use Green's theorem. C bounds the unit disk R:

$$\oint_C \vec{F}\cdot \mathrm{d}\vec{r} = \iint_R \operatorname{curl}\vec{F}\,\mathrm{d}A = \int_0^{2\pi}\int_0^1 (0-0)r\,\mathrm{d}r\,\mathrm{d}\theta = 0,$$

as expected.

A closely related line integral is the flux of $\vec{F}$ across C. We measure the flux from left to right. The flux across C is

$$\int_C \vec{F}\cdot\hat{n}\,\mathrm{d}s.$$

To compute this, use the fact that $\hat{n}$ is the unit tangent vector turned through $\pi/2$ radians clockwise, so

$$\hat{n}\,\mathrm{d}s = \langle \mathrm{d}y, -\mathrm{d}x\rangle.$$

We have

$$\int_C \vec{F}\cdot\hat{n}\,\mathrm{d}s = \int_C M\mathrm{d}y - N\,\mathrm{d}x.$$

In the example above

$$\int_C \vec{F}\cdot\hat{n}\,\mathrm{d}s = \int_0^{2\pi} \cos^2 t + \sin^2 t\,\mathrm{d}t = 2\pi.$$

One can also use Green's theorem in normal form

$$\int_C \vec{F}\cdot\hat{n}\,\mathrm{d}s = \iint_R \operatorname{div}\vec{F}\,\mathrm{d}A = \iint_R 2\,\mathrm{d}A = 2\pi.$$

$\mathrm{d}A$ is the area element in the xy-plane. We have

$$\mathrm{d}A = \mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta.$$

Example 34.1. *What is the area of the ellipse*

$$(2x+y)^2 + (x-y)^2 \leq 5?$$

Use change of variables, $u = 2x + y$ and $v = x - y$. The Jacobian is

$$J = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3.$$

So

$$\mathrm{d}u\,\mathrm{d}v = 3\,\mathrm{d}x\,\mathrm{d}y.$$

So the area of R is

$$\iint_R 1\,\mathrm{d}A = \iint_{(2x+y)^2+(x-y)^2\leq 5} 1\,\mathrm{d}x\,\mathrm{d}y = \iint_{u^2+v^2\leq 5} \frac{1}{3}\,\mathrm{d}u\,\mathrm{d}v = \frac{5}{3}\pi.$$

Example 34.2. *Calculate*

$$\int_0^1 \int_{y^3}^1 \frac{6y^2}{x^2+2}\,\mathrm{d}x\,\mathrm{d}y.$$

We swap the order of integration. The region R of integration is

$$0 \leq y \leq 1 \qquad \text{and} \qquad y^3 \leq x \leq 1.$$

Therefore

$$\int_0^1 \int_{y^3}^1 \frac{6y^2}{x^2+2}\,\mathrm{d}x\,\mathrm{d}y = \iint_R \frac{6y^2}{x^2+2}\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^{x^{1/3}} \frac{6y^2}{x^2+2}\,\mathrm{d}y\,\mathrm{d}x.$$

The inner integral is

$$\int_0^{x^{1/3}} \frac{6y^2}{x^2+2}\,\mathrm{d}y = \left[\frac{2y^3}{x^2+2}\right]_0^{x^{1/3}} = \frac{2x}{2+x^2}.$$

The outer integral is

$$\int_0^1 \frac{2x}{x^2+2}\,\mathrm{d}x = \Big[\ln(x^2+2)\Big]_0^1 = \ln 3/2.$$

In three dimensions, the volume form is

$$\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z = \rho^2\sin\phi\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}\theta.$$

The trickiest thing is to calculate surface integrals in space. The area form on a surface is

$$\mathrm{d}S.$$

It plays the same role as the area form $\mathrm{d}A$ in the plane. More common is

$$\mathrm{d}\vec{S} = \hat{n}\,\mathrm{d}S,$$

which is used to calculate flux out of S:

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S}.$$

Note that we need to choose an orientation of S. There are many ways to calculate the flux. If we parameterise S, $\vec{r}(u,v)$ using two parameters u and v we have

$$\mathrm{d}\vec{S} = \frac{\partial\vec{r}}{\partial u}\times\frac{\partial\vec{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v.$$

If S is given by a single constraint $g(x,y,z)=c$, a constant, then

$$\mathrm{d}\vec{S} = \frac{\vec{N}}{\vec{N}\cdot\hat{k}}\,\mathrm{d}x\,\mathrm{d}y \qquad \text{and} \qquad \mathrm{d}\vec{S} = \frac{\vec{N}}{|\vec{N}\cdot\hat{k}|}\,\mathrm{d}x\,\mathrm{d}y,$$

where $\vec{N} = \nabla g$ and the first form always picks the upwards orientation whilst the second form preserves the orientation. If S is given as the graph of a function $z=f(x,y)$ over a region R in the xy-plane, we have

$$\mathrm{d}\vec{S} = \langle -f_x, -f_y, 1\rangle\,\mathrm{d}x\,\mathrm{d}y.$$

Formulas for spheres centred at the origin and cylinders with central axis the z-axis are simply worth remembering:

$$\mathrm{d}\vec{S} = a\langle x,y,z\rangle\sin\phi\,\mathrm{d}\phi\,\mathrm{d}\theta \qquad \text{and} \qquad \mathrm{d}\vec{S} = \langle x,y,0\rangle\,\mathrm{d}z\,\mathrm{d}\theta.$$

Example 34.3. *Let*

$$\vec{F} = \langle xz^2, yz^2, z^3\rangle.$$

What is the flux out of the cylinder, height 1, radius 1, base in the xy-plane, centred at the origin?

Let's calculate this directly. There are three sides, the two flat ones S_0 and S_1 and the curved one S_2.

For S_2, we have a cylinder, so use second form

$$\mathrm{d}\vec{S} = \langle x, y, 0\rangle \,\mathrm{d}z\,\mathrm{d}\theta.$$

The flux across S_3 is

$$\iint_{S_2} x^2 z^2 + y^2 z^2 \,\mathrm{d}r\,\mathrm{d}\theta = \int_0^{2\pi} \int_0^1 z^2 \,\mathrm{d}z\,\mathrm{d}\theta.$$

The inner integral is

$$\int_0^1 z^2 \,\mathrm{d}z = \left[\frac{z^3}{3}\right]_0^1 = \frac{1}{3}.$$

So the flux across S_2 is $2\pi/3$. $\vec{F}$ is horizontal along S_0, so the flux across S_0 is zero. Across S_1, $\hat{n} = \hat{k}$, so the flux is

$$\iint_{S_1} \vec{F}\cdot \mathrm{d}\vec{S} = \iint_{S_1} 1 \mathrm{d}S = \pi,$$

since the area of S_1 is π.

In total, the flux is

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S} = \iint_{S_0} \vec{F}\cdot\mathrm{d}\vec{S} + \iint_{S_1} \vec{F}\cdot\mathrm{d}\vec{S} + \iint_{S_2} \vec{F}\cdot\mathrm{d}\vec{S} = 0 + \pi + \frac{2\pi}{3} = \frac{5\pi}{3}.$$

Instead we could apply the divergence theorem:

$$\iint_S \vec{F}\cdot\mathrm{d}\vec{S} = \iiint_V \operatorname{div}\vec{F}\,\mathrm{d}V = \iiint_V 5z^2\,\mathrm{d}V.$$

To calculate this integral use cylindrical coordinates

$$\iiint_V 5z^2\,\mathrm{d}V = \int_0^{2\pi}\int_0^1\int_0^1 5z^2 r\,\mathrm{d}z\,\mathrm{d}r\,\mathrm{d}\theta = \frac{5\pi}{3}.$$

Here is a summary of the various fundamental theorems relating integrals in different dimensions:

Dimension	Work done	Flux
0-1	FTC line integrals	
1-2	Green's +Stokes' theorem	Green's theorem (normal form)
2-3	Divergence	Not covered