Vectors

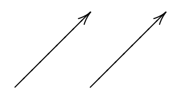
Our very first topic is unusual in that we will start with a brief written presentation. More typically we will begin each topic with a videotaped lecture by Professor Auroux and follow that with a brief written presentation.

As we pointed out in the introduction, vectors will be used throughout the course. The basic concepts are straightforward, but you will have to master some new terminology. Another important point we made earlier is that we can view vectors in two different ways: geometrically and algebraically. We will start with the geometric view and introduce terminology along the way.

Geometric view

A vector is defined as having a magnitude and a direction. We represent it by an arrow in the plane or in space. The length of the arrow is the vector's magnitude and the direction of the arrow is the vector's direction.

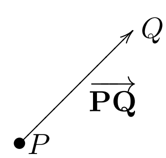
In this way, two arrows with the same magnitude and direction represent the same vector.



(same vector)

We will refer to the start of the arrow as the *tail* and the end as the *tip* or *head.*

The vector between two points will be denoted $\overrightarrow{\mathbf{PQ}}$.

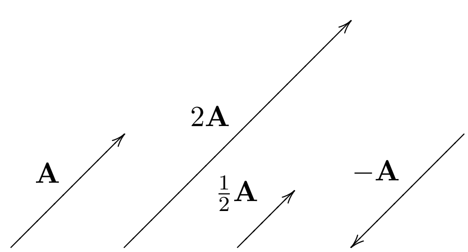


We call P the initial point and Q the terminal point of $\overrightarrow{\mathbf{PQ}}$.

The *magnitude* of the vector $\mathbf{A}$ will be denoted $|\mathbf{A}|$. Magnitude will also be called *length* or *norm*

Scaling, adding and subtracting vectors

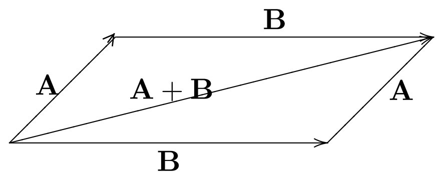
Scaling a vector means changing its length by a scale factor. For example,



(scaling a vector)

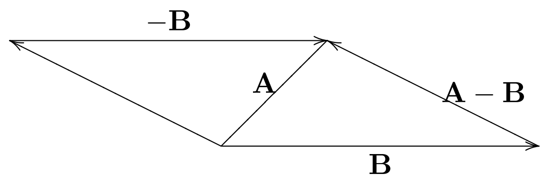
Because we use numbers to scale a vector we will often refer to real numbers as *scalars*

You add vectors by placing them head to tail. As the figure shows, this can be done in either order



It is often useful to think of vectors as *displacements*. In this way, $\mathbf{A}+\mathbf{B}$ can be thought of as the displacement $\mathbf{A}$ followed by the displacement $\mathbf{B}$.

You subtract vectors either by placing the tail to tail or by adding $\mathbf{A}+(-\mathbf{B})$.



Thought of as displacements $\mathbf{A}-\mathbf{B}$ is the displacement from the end of $\mathbf{B}$ to the end of $\mathbf{A}$.

Algebraic view

As is conventional, we label the origin O. In the plane $O=(0,0)$ and in space $O=(0,0,0)$.

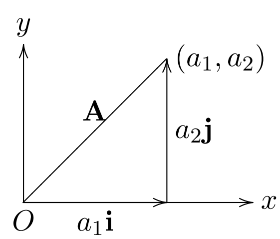
In the xy-plane if we place the tail of $\mathbf{A}$ at the origin, its head will be at the point with coordinates, say, (a_1, a_2). In this way, the coordinates of the head determine the vector $\mathbf{A}$. When we draw $\mathbf{A}$ from the origin we will refer to it as an *origin vector*.

Using the coordinates we write

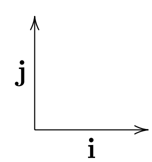
$$\mathbf{A} = \langle a_1,\, a_2\rangle.$$

Addition, subtraction and scaling using coordinates is discussed below.

Graphically:



The vectors $\mathbf{i}$ and $\mathbf{j}$ used in the figure above have coordinates $\mathbf{i}=\langle 1,\,0\rangle$, $\mathbf{j}=\langle 0,\,1\rangle$. We use them so often that they get their own symbols.



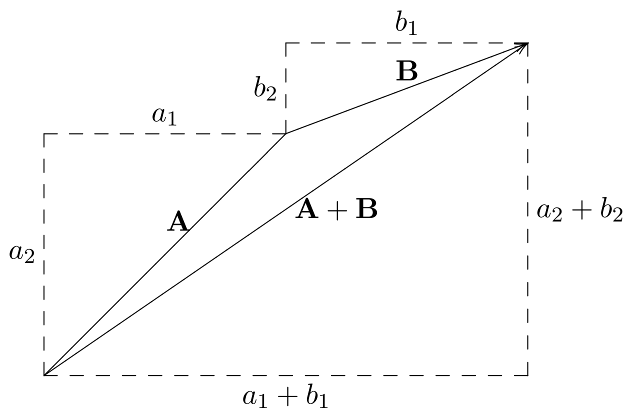
Notation and terminology

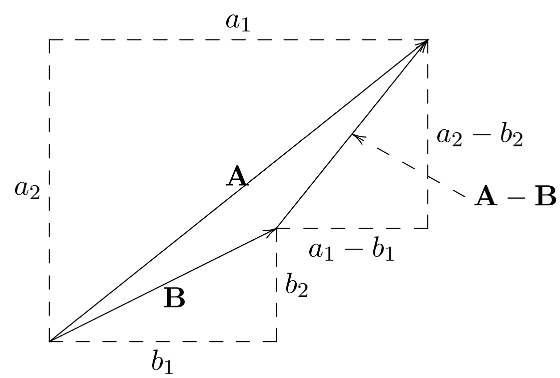
1. (a_1, a_2) indicates a point in the plane.
2. $\langle a_1,\, a_2\rangle = a_1\mathbf{i} + a_2\mathbf{j}$. This is equal to the vector drawn from the origin to the point (a_1, a_2).
3. For $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j}$, a_1 and a_2 are called the $\mathbf{i}$ and $\mathbf{j}$ *components* of $\mathbf{A}$. (Note that they are scalars.)
5. $\overrightarrow{\mathbf{P}} = \overrightarrow{\mathbf{OP}}$ is the vector from the origin to P.

6. On the blackboard vectors will usually have an arrow above the letter. In print we will often drop the arrow and just use the bold face to indicate a vector, i.e. $\mathbf{P} \equiv \overrightarrow{\mathbf{P}}$.
7. A real number is a *scalar*, you can use it to scale a vector.

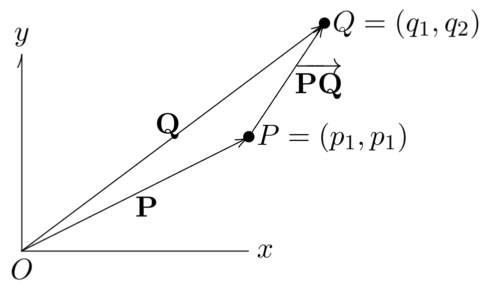
Vector algebra using coordinates
For the vectors $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j}$ and $\mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j}$ we have the following algebraic rules. The figures below connect these rules to the geometric viewpoint.

Magnitude: $|\mathbf{A}| = \sqrt{a_1^2 + a_2^2}$ (this is just the Pythagorean theorem)
Addition: $\mathbf{A}+\mathbf{B} = (a_1+b_1)\mathbf{i} + (a_2+b_2)\mathbf{j}$, that is, $\langle a_1, a_2\rangle + \langle b_1, b_2\rangle = \langle a_1+b_1, a_2+b_2\rangle$
Subtraction: $\mathbf{A}-\mathbf{B} = (a_1-b_1)\mathbf{i} + (a_2-b_2)\mathbf{j}$, that is, $\langle a_1, a_2\rangle - \langle b_1, b_2\rangle = \langle a_1-b_1, a_2-b_2\rangle$



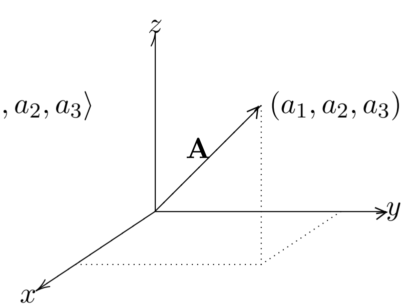


For two points P and Q the vector $\overrightarrow{\mathbf{PQ}} = \overrightarrow{\mathbf{Q}} - \overrightarrow{\mathbf{P}}$ i.e., $\overrightarrow{\mathbf{PQ}}$ is the *displacement* from P to Q.



Vectors in three dimensions
We represent a three dimensional vector as an arrow in space. Using coordinates we need three numbers to represent a vector.

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



Geometrically nothing changes for vectors in three dimensions. They are scaled and added exactly as above.

Algebraically the origin vector $\mathbf{A} = \langle a_1, a_2, a_3\rangle$ starts at the origin and extends to the point (a_1, a_2, a_3). We have the special vectors $\mathbf{i} = \langle 1, 0, 0\rangle$, $\mathbf{j} = \langle 0, 1, 0\rangle$, $\mathbf{k} = \langle 0, 0, 1\rangle$. Using them

$$\langle a_1, a_2, a_3\rangle = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}.$$

Then, for $\mathbf{A} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3\rangle$ we have

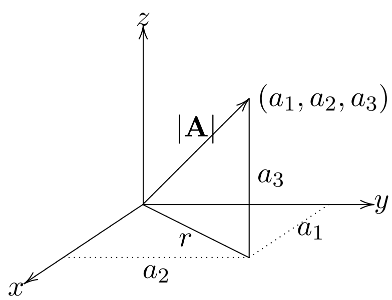
$$\langle a_1, a_2, a_3\rangle + \langle b_1, b_2, b_3\rangle = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3\rangle.$$

exactly as in the two dimensional case.

Magnitude in three dimensions also follows from the Pythagorean theorem.

$$|a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}| = |\langle a_1, a_2, a_3\rangle| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

You can see this in the figure below, where $r = \sqrt{a_1^2 + a_2^2}$ and $|\mathbf{A}| = \sqrt{r^2 + a_3^2} = \sqrt{a_1^2 + a_2^2 + a_3^2}$.



Unit vectors

A unit vector is any vector with unit length. When we want to indicate that a vector is a unit vector we put a hat (circumflex) above it, e.g., $\widehat{\mathbf{u}}$.

The special vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors.

Since vectors can be scaled, any vector can be rescaled to be a unit vector.

Example: Find a unit vector that is parallel to $\langle 3, 4\rangle$.

<u>Answer:</u> Since $|\langle 3, 4\rangle| = 5$ the vector $\frac{1}{5}\langle 3, 4\rangle = \left\langle \frac{3}{5}, \frac{4}{5} \right\rangle$ has unit length and is parallel to $\langle 3, 4\rangle$.

Dot Product

The dot product is one way of combining ("multiplying") two vectors. The output is a scalar (a number). It is called the dot product because the symbol used is a dot. Because the dot product results in a scalar it, is also called the scalar product.

As with most things in 18.02, we have a geometric and algebraic view of dot product.

Algebraic definition (for 2D vectors):

If $\mathbf{A} = \langle a_1, a_2 \rangle$ and $\mathbf{B} = \langle b_1, b_2 \rangle$ then

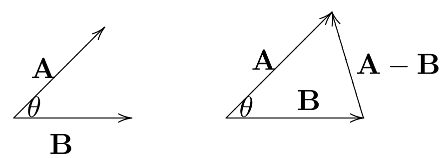
$$\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2.$$

Example: $\langle 6, 5 \rangle \cdot \langle 1, 2 \rangle = 6 \cdot 1 + 5 \cdot 2 = 16.$

Geometric view:

The figure below shows $\mathbf{A}$, $\mathbf{B}$ with the angle θ between them. We get

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos\theta$$



Showing the two views (algebraic and geometric) are the same requires the law of cosines

$$|\mathbf{A} - \mathbf{B}|^2 = |\mathbf{A}|^2 + |\mathbf{B}|^2 - 2|\mathbf{A}||\mathbf{B}| \cos\theta$$

$\Rightarrow\ (a_1^2 + a_2^2) + (b_1^2 + b_2^2) - ((a_1 - b_1)^2 + (a_2 - b_2)^2) = 2|\mathbf{A}||\mathbf{B}| \cos\theta$

$\Rightarrow\ a_1 b_1 + a_2 b_2 = |\mathbf{A}||\mathbf{B}| \cos\theta.$

Since $\langle a_1, a_2 \rangle \cdot \langle b_1, b_2 \rangle = a_1 b_1 + a_2 b_2$, we have shown $\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos\theta$.

From the algebraic definition of dot product we easily get the the following algebraic law

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}.$$

Example: Find the dot product of $\mathbf{A}$ and $\mathbf{B}$.

i) $|\mathbf{A}| = 2$, $|\mathbf{B}| = 5$, $\theta = \pi/4$.

Answer: (draw the picture yourself) $\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos\theta = 10\sqrt{2}/2 = 5\sqrt{2}$.

ii) $\mathbf{A} = \mathbf{i} + 2\mathbf{j}$, $\mathbf{B} = 3\mathbf{i} + 4\mathbf{j}$.

Answer: $\mathbf{A} \cdot \mathbf{B} = 1 \cdot 3 + 2 \cdot 4 = 11.$

Three dimensional vectors

The dot product works the same in 3D as in 2D. If $\mathbf{A} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3 \rangle$ then

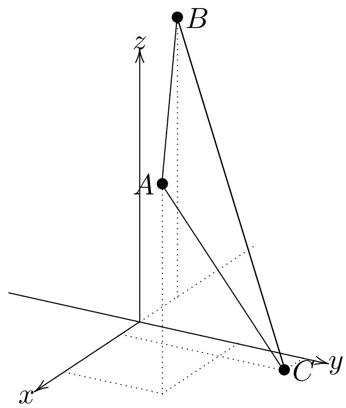
$$\mathbf{A} \cdot \mathbf{B} = a_1 \cdot b_1 + a_2 \cdot b_2 + a_3 \cdot b_3.$$

The geometric view is identical and the same proof shows

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos\theta$$

Example:
Show $A = (4, 3, 6)$, $B = (-2, 0, 8)$, $C = (1, 5, 0)$
are the vertices of a right triangle.

Answer: Two legs of the triangle are $\overrightarrow{\mathbf{AC}} = \langle -3,\, 2,\, -6 \rangle$ and $\overrightarrow{\mathbf{AB}} = \langle -6,\, -3,\, 2 \rangle \;\; \Rightarrow$ $\overrightarrow{\mathbf{AC}} \cdot \overrightarrow{\mathbf{AB}} = 18 - 6 - 12 = 0$. The geometric view of dot product implies the angle between the legs is $\pi/2$ (i.e $\cos\theta = 0$).



Definition of the term orthogonal and the test for orthogonality
When two vectors are perpendicular to each other we say they are *orthogonal.*

As seen in the example, since $\cos(\pi/2) = 0$, the dot product gives a test for orthogonality between vectors:

$$\mathbf{A} \perp \mathbf{B} \;\Leftrightarrow\; \mathbf{A} \cdot \mathbf{B} = 0.$$

Dot product and length
Both the algebraic and geometric formulas for dot product show it is intimately connected to length. In fact, they show for a vector $\mathbf{A}$

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2.$$

Let's show this using both views.
Algebraically: suppose $\mathbf{A} = \langle a_1, a_2, a_3 \rangle$ then

$$\mathbf{A} \cdot \mathbf{A} = \langle a_1, a_2, a_3 \rangle \cdot \langle a_1, a_2, a_3 \rangle = a_1^2 + a_2^2 + a_3^2 = |\mathbf{A}|^2.$$

Geometrically: the angle θ between $\mathbf{A}$ and itself is 0. Therefore,

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}||\mathbf{A}|\cos\theta = |\mathbf{A}||\mathbf{A}| = |\mathbf{A}|^2.$$

As promised both views give the formula.

Uses of dot product

1. Find the angle between $\mathbf{i}+\mathbf{j}+2\mathbf{k}$ and $2\mathbf{i}-\mathbf{j}+\mathbf{k}$.

<u>Answer:</u> We call the angle θ and use both ways of computing the dot product.
Algebraically we have

$$(\mathbf{i}+\mathbf{j}+2\mathbf{k})\cdot(2\mathbf{i}-\mathbf{j}+\mathbf{k}) = 2-1+2 = 3.$$

Geometrically

$$(\mathbf{i}+\mathbf{j}+2\mathbf{k})\cdot(2\mathbf{i}-\mathbf{j}+\mathbf{k}) = |\mathbf{i}+\mathbf{j}+2\mathbf{k}|\cdot|2\mathbf{i}-\mathbf{j}+\mathbf{k}|\cos\theta = \sqrt{6}\,\sqrt{6}\cos\theta.$$

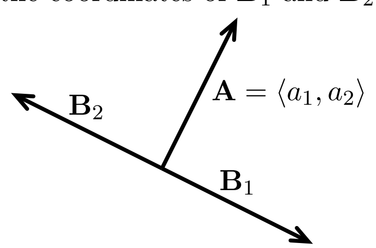
Combining these two we have

$$6\cos\theta = 3 \;\Rightarrow\; \cos\theta = \frac{3}{6} = \frac{1}{2} \;\Rightarrow\; \theta = \frac{\pi}{3}.$$

2. a) Are $\langle 1,3\rangle$ and $\langle -2,2\rangle$ orthogonal?

b) For what value of a are the vectors $\langle 1,a\rangle$ and $\langle 2,3\rangle$ at right angles?

c) In the figure the vectors $\mathbf{A}$ and $\mathbf{B}_1$ are orthogonal as are $\mathbf{A}$ and $\mathbf{B}_2$. If all the vectors are the same length what are the coordinates of $\mathbf{B}_1$ and $\mathbf{B}_2$?



<u>Answer:</u> a) Vectors are orthogonal if their dot product is 0. So, taking the dot product

$$\langle 1,3\rangle\cdot\langle -2,2\rangle = -2+6 = 4 \neq 0.$$

Thus the vectors are not orthogonal.

b) Setting the dot product to 0 and solving for a we get

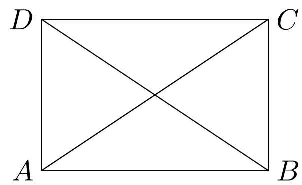
$$\langle 1,a\rangle\cdot\langle 2,3\rangle = 2+3a = 0 \;\Rightarrow\; a = -2/3.$$

c) $\mathbf{B}_1$ is $\mathbf{A}$ rotated 90° clockwise. We will show that $\mathbf{B}_1 = \langle a_2, -a_1\rangle$. It is easy to check that $|\langle a_2,-a_1\rangle| = |\mathbf{A}|$ and $\langle a_2,-a_1\rangle\cdot\mathbf{A} = 0$. The figure above shows that putting the negative sign on the a_1 means $\langle a_2,-a_1\rangle$ is turned clockwise from $\mathbf{A}$. Thus, $\langle a_2,-a_1\rangle = \mathbf{B}_1$.

$\mathbf{B}_2$ is $\mathbf{A}$ rotated 90° counterclockwise. Similarly to $\mathbf{B}_1$, we find $\mathbf{B}_2 = \langle -a_2, a_1\rangle$.

3. Using vectors and dot product show the diagonals of a parallelogram have equal lengths if and only if it's a rectangle

Answer:



We will make use of two properties of the dot product

1. $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2$.

2. $\mathbf{v} \cdot \mathbf{w} = 0 \;\Leftrightarrow\; \mathbf{v} \perp \mathbf{w}$.

Referring to the figure, we will also need to use the fact that $ABCD$ is a parallelogram. That is, $\overrightarrow{\mathbf{AB}} = \overrightarrow{\mathbf{DC}}$.

We have $\overrightarrow{\mathbf{AC}} = \overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}}$ and $\overrightarrow{\mathbf{BD}} = \overrightarrow{\mathbf{BC}} + \overrightarrow{\mathbf{CD}} = \overrightarrow{\mathbf{BC}} - \overrightarrow{\mathbf{AB}}$.

Taking dot products:

$$|\overrightarrow{\mathbf{AC}}|^2 = \overrightarrow{\mathbf{AC}} \cdot \overrightarrow{\mathbf{AC}} = (\overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}}) \cdot (\overrightarrow{\mathbf{AB}} + \overrightarrow{\mathbf{BC}}) = |\overrightarrow{\mathbf{AB}}|^2 + 2\overrightarrow{\mathbf{AB}} \cdot \overrightarrow{\mathbf{BC}} + |\overrightarrow{\mathbf{BC}}|^2.$$

and

$$|\overrightarrow{\mathbf{BD}}|^2 = \overrightarrow{\mathbf{BD}} \cdot \overrightarrow{\mathbf{BD}} + (\overrightarrow{\mathbf{BC}} - \overrightarrow{\mathbf{AB}}) \cdot (\overrightarrow{\mathbf{BC}} - \overrightarrow{\mathbf{AB}}) = |\overrightarrow{\mathbf{BC}}|^2 - 2\overrightarrow{\mathbf{BC}} \cdot \overrightarrow{\mathbf{AB}} + |\overrightarrow{\mathbf{AB}}|^2$$

Comparing the two equations above we see

$$|\overrightarrow{\mathbf{AC}}|^2 = |\overrightarrow{\mathbf{BD}}|^2 \;\Leftrightarrow\; 4\overrightarrow{\mathbf{AB}} \cdot \overrightarrow{\mathbf{BC}} = 0.$$

This shows the diagonals have the same length if and only if $\overrightarrow{\mathbf{AB}} \perp \overrightarrow{\mathbf{BC}}$. That is, if and only if the sides of the parallelogram are orthogonal to each other. QED

Components and Projection

If $\mathbf{A}$ is any vector and $\widehat{\mathbf{u}}$ is a unit vector then the *component* of $\mathbf{A}$ in the direction of $\widehat{\mathbf{u}}$ is

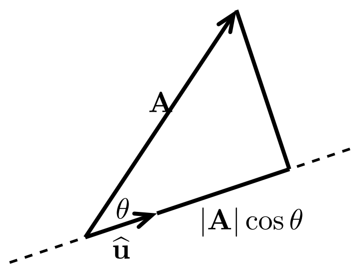
$$\mathbf{A} \cdot \widehat{\mathbf{u}}.$$

(Note: the component is a scalar.)

If θ is the angle between $\mathbf{A}$ and $\widehat{\mathbf{u}}$ then since $|\widehat{\mathbf{u}}| = 1$

$$A \cdot \widehat{\mathbf{u}} = |\mathbf{A}||\widehat{\mathbf{u}}|\cos\theta = |\mathbf{A}|\cos\theta.$$

The figure shows that geometrically this is the length of the leg of the right triangle with hypotenuse $\mathbf{A}$ and one leg parallel to $\widehat{\mathbf{u}}$.



We also call the leg parallel to $\widehat{\mathbf{u}}$ the *orthogonal projection* of $\mathbf{A}$ on $\widehat{\mathbf{u}}$.

For a non-unit vector: the component of $\mathbf{A}$ in the direction of $\mathbf{B}$ is simply the component of $\mathbf{A}$ in the direction of $\widehat{\mathbf{u}} = \frac{\mathbf{B}}{|\mathbf{B}|}$. ($\widehat{\mathbf{u}}$ is the unit vector in the same direction as $\mathbf{B}$.)

Example: Find the component of $\mathbf{A}$ in the direction of $\mathbf{B}$.

i) $|\mathbf{A}| = 2$, $|\mathbf{B}| = 5$, $\theta = \pi/4$.

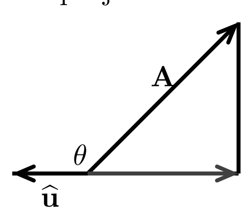
<u>Answer:</u> Referring to the figure above: the component is $|\mathbf{A}|\cos\theta = 2\cos(\pi/4) = \sqrt{2}$. Note, the length of $\mathbf{B}$ given is irrelevant, since we only care about the unit vector parallel to $\mathbf{B}$.

ii) $\mathbf{A} = \mathbf{i} + 2\mathbf{j}$, $\mathbf{B} = 3\mathbf{i} + 4\mathbf{j}$.

<u>Answer:</u> Unit vector in direction of $\mathbf{B}$ is $\frac{\mathbf{B}}{|\mathbf{B}|} = \frac{3}{5}\mathbf{i} + \frac{4}{5}\mathbf{j} \;\Rightarrow\;$ component is $\mathbf{A} \cdot \mathbf{B}/|\mathbf{B}| = 3/5 + 8/5 = 11/5$.

iii) Find the component of $\mathbf{A} = \langle 2, 2\rangle$ in the direction of $\widehat{\mathbf{u}} = \langle -1, 0\rangle$

<u>Answer:</u> The vector $\widehat{\mathbf{u}}$ is a unit vector, so the component is $\mathbf{A} \cdot \widehat{\mathbf{u}} = \langle 2, 2\rangle \cdot \langle -1, 0\rangle = -2$. The negative component is okay, it says the projection of $\mathbf{A}$ and $\widehat{\mathbf{u}}$ point in opposite directions.



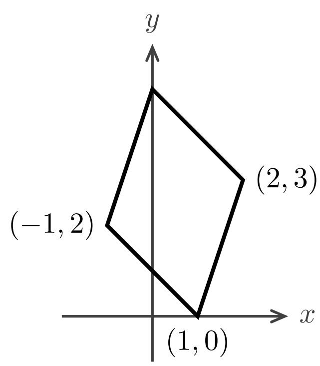
We emphasize one more time that the component of a vector is a *scalar*.

Areas and Determinants

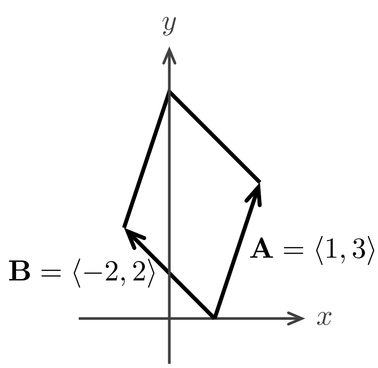
1. Compute $\begin{vmatrix} 6 & 5 \\ 1 & 2 \end{vmatrix}$.

<u>**Answer:**</u> $\begin{vmatrix} 6 & 5 \\ 1 & 2 \end{vmatrix} = 6 \cdot 2 - 5 \cdot 1 = 7.$

2. Compute the area of the parallelogram shown.



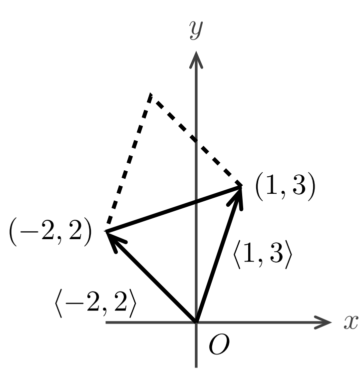
<u>**Answer:**</u> The area is given by the determinant of the vectors determining the parallelogram.



$$\text{Area} = |\det(\mathbf{A}, \mathbf{B})| = \left|\det \begin{pmatrix} 1 & 3 \\ -2 & 2 \end{pmatrix}\right| = 2 + 6 = 8.$$

3. Find the area of the triangle with vertices $(0, 0)$, $(-2, 2)$ and $(1, 3)$.

<u>**Answer:**</u> The triangle is half a parallelogram. So the area is $\dfrac{1}{2}\left|\det \begin{pmatrix} 1 & 3 \\ -2 & 2 \end{pmatrix}\right| = 2.$



Determinants 1.

Given a square array A of numbers, we associate with it a number called the **determinant** of A, and written either $\det(A)$, or $|A|$. For 2×2

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{1}$$

Do not memorize this as a formula — learn instead the pattern which gives the terms. The 2×2 case is easy: the product of the elements on one diagonal (the "main diagonal"), minus the product of the elements on the other (the "antidiagonal").

Below we will see how to compute 3×3 determinants $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$. First, try the following 2×2 example on your own, then check your work against the solution.

Example 1.1 Evaluate $\begin{vmatrix} 1 & -2 \\ -1 & 3 \end{vmatrix}$ using (1).

Solution. Using the same order as in (1), we get $12 + (-8) + 1 - 6 - 8 - (-2) = -7$.

Important facts about $|A|$:

D-1. $|A|$ is multiplied by -1 if we interchange two rows or two columns.

D-2. $|A| = 0$ if one row or column is all zero, or if two rows or two columns are the same.

D-3. $|A|$ is multiplied by c, if every element of some row or column is multiplied by c.

D-4. The value of $|A|$ is unchanged if we add to one row (or column) a constant multiple of another row (resp. column).

All of these facts are easy to check for 2×2 determinants from the formula (1); from this, their truth also for 3×3 determinants will follow from the Laplace expansion.

Though the letters $a, b, c, \ldots$ can be used for very small determinants, they can't for larger ones; it's important early on to get used to the standard notation for the entries of determinants. This is what the common software packages and the literature use. The determinants of order two and three would be written respectively

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

In general, the **ij-entry**, written a_{ij}, is the number in the i-th row and j-th column.

Its **ij-minor**, written $|A_{ij}|$, is the determinant that's left after deleting from $|A|$ the row and column containing a_{ij}.

Its **ij-cofactor**, written here A_{ij}, is given as a formula by $A_{ij} = (-1)^{i+j}|A_{ij}|$. For a 3×3 determinant, it is easier to think of it this way: we put $+$ or $-$ in front of the ij-minor, according to whether $+$ or $-$ occurs in the ij-position in the checkerboard pattern

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix} . \tag{2}$$

Example 1.2 $\quad |A| = \begin{vmatrix} 1 & 0 & 3 \\ 1 & 2 & -1 \\ 2 & 1 & -1 \end{vmatrix} . \qquad$ Find $|A_{12}|,\ A_{12},\ |A_{22}|,\ A_{22}$.

Solution. $|A_{12}| = \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} = 1, \quad A_{12} = -1. \qquad |A_{22}| = \begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} = -7, \quad A_{22} = -7.$

Laplace expansion by cofactors

This is another way to evaluate a determinant; we give the rule for a 3×3. It generalizes easily to an $n \times n$ determinant.

Select any row (or column) of the determinant. Multiply each entry a_{ij} in that row (or column) by its cofactor A_{ij}, and add the three resulting numbers; you get the value of the determinant.

As practice with notation, here is the formula for the Laplace expansion of a third order (i.e., a 3×3) determinant using the cofactors of the first row:

$$a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \ = \ |A| \tag{3}$$

and the formula using the cofactors of the j-th column:

$$a_{1j}A_{1j} + a_{2j}A_{2j} + a_{3j}A_{3j} \ = \ |A| \tag{4}$$

Example 1.3 Evaluate the determinant in Example 1.2 using the Laplace expansions by the first row and by the second column, and check by also using (1).

Solution. The Laplace expansion by the first row is

$$\begin{vmatrix} 1 & 0 & 3 \\ 1 & 2 & -1 \\ 2 & 1 & -1 \end{vmatrix} \ = \ 1 \cdot \begin{vmatrix} 2 & -1 \\ 1 & -1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = 1 \cdot (-1) - 0 \cdot 1 + 3 \cdot (-3) = -10.$$

The Laplace expansion by the second column would be

$$\begin{vmatrix} 1 & 0 & 3 \\ 1 & 2 & -1 \\ 2 & 1 & -1 \end{vmatrix} \ = \ -0 \cdot \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 3 \\ 1 & -1 \end{vmatrix} = 0 + 2 \cdot (-7) - 1 \cdot (-4) = -10.$$

Checking by (1), we have $|A| \ = \ -2 + 0 + 3 - 12 - 0 - (-1) \ = \ -10.$

Example 1.4 Show the Laplace expansion by the first row gives the following formula (which you may have seen before).

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + chd - gec - hfa - ibd$$

Solution. We have

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \cdot \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \cdot \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \cdot \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - fh) - b(di - fg) + c(dh - eg), \end{aligned}$$

whose six terms agree with the six terms on the right of the formula above.

(A similar argument can be made for the Laplace expansion by any row or column.)

For $n \times n$ determinants, the **minor** $|A_{ij}|$ of the entry a_{ij} is defined to be the determinant obtained by deleting the i-th row and j-th column; the **cofactor** A_{ij} is the minor, prefixed by a $+$ or $-$ sign according to the natural generalization of the checkerboard pattern (2). Then the Laplace expansion by the i-th row would be

$$|A| \;=\; a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}\ .$$

This is an inductive calculation — it expresses the determinant of order n in terms of determinants of order $n-1$. Thus, since we can calculate determinants of order 3, it allows us to calculate determinants of order 4; then determinants of order 5, and so on. If we take for definiteness $i = 1$, then the above Laplace expansion formula can be used as the basis of an inductive definition of the $n \times n$ determinant.

Example 1.5 Evaluate $\begin{vmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 1 & 4 \\ -1 & 4 & 1 & 0 \\ 0 & 4 & 2 & -1 \end{vmatrix}$ by its Laplace expansion by the first row.

Solution. $1 \cdot \begin{vmatrix} -1 & 1 & 4 \\ 4 & 1 & 0 \\ 4 & 2 & -1 \end{vmatrix} \;-\; 0 \cdot A_{12} \;+\; 2 \cdot \begin{vmatrix} 2 & -1 & 4 \\ -1 & 4 & 0 \\ 0 & 4 & -1 \end{vmatrix} \;-\; 3 \cdot \begin{vmatrix} 2 & -1 & 1 \\ -1 & 4 & 1 \\ 0 & 4 & 2 \end{vmatrix}$

$$= \; 1 \cdot 21 \;+\; 2 \cdot (-23) \;-\; 3 \cdot 2 \;=\; -31.$$

Cross Product

The cross product is another way of multiplying two vectors. (The name comes from the symbol used to indicate the product.) Because the result of this multiplication is *another vector* it is also called the *vector product*.

As usual, there is an algebraic and a geometric way to describe the cross product. We'll define it algebraically and then move to the geometric description.

Determinant definition for cross product

For the vectors $\mathbf{A} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3\rangle$ we define the cross product by the following formula

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

The bottom three equations above are easily seen to be equivalent and should be taken as the definition of the cross product. The top line is technically flawed because we are not really allowed to use vectors as entries in a determinant. Nonetheless is is an excellent way to remember how to compute the cross product.

Example: $\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 3 & 0 \\ 3 & -2 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 3 & -2 \end{vmatrix}\mathbf{k} = -13\,\mathbf{k}$

Example: Compute $\mathbf{i}\times\mathbf{j}$.

<u>**Answer:**</u> $\mathbf{i} = \langle 1, 0, 0\rangle$ and $\mathbf{j} = \langle 0, 1, 0\rangle$ therefore

$$\mathbf{i}\times\mathbf{j} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \mathbf{i}(0) - \mathbf{j}(0) + \mathbf{k}(1) = \mathbf{k}.$$

Algebraic facts: (these follow easily from properties of determinant).

1. $\mathbf{A}\times\mathbf{A} = \mathbf{0}$
2. Anti-commutivity: $\mathbf{A}\times\mathbf{B} = -\mathbf{B}\times\mathbf{A}$
3. Distributive law: $\mathbf{A}\times(\mathbf{B}+\mathbf{C}) = \mathbf{A}\times\mathbf{B} + \mathbf{A}\times\mathbf{C}$
4. Non-associativity: $(\mathbf{A}\times\mathbf{B})\times\mathbf{C} \neq \mathbf{A}\times(\mathbf{B}\times\mathbf{C})$ (example in a moment).

For the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ we have

$$\mathbf{i}\times\mathbf{j} = \mathbf{k},\quad \mathbf{j}\times\mathbf{k} = \mathbf{i},\quad \mathbf{k}\times\mathbf{i} = \mathbf{j}.$$

Example: (non-associativity) $(\mathbf{i}\times\mathbf{j})\times\mathbf{j} = -\mathbf{i}$ but $\mathbf{i}\times(\mathbf{j}\times\mathbf{j}) = 0$.

Example: It is possible to compute a cross product using the algebraic facts and the known products of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. For example,

$$(2\mathbf{i}+3\mathbf{j})\times(3\mathbf{i}-2\mathbf{j}) = (6\mathbf{i}\times\mathbf{i}) - (4\mathbf{i}\times\mathbf{j}) + (9\mathbf{j}\times\mathbf{i}) - (6\mathbf{j}\times\mathbf{j}) = -13\mathbf{k}.$$

The first equation follows from the distributive law. In the second, we used $\mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = 0$ (algebraic fact 1), $\mathbf{i}\times\mathbf{j} = \mathbf{k}$ (computed above) and $\mathbf{j}\times\mathbf{i} = -\mathbf{k}$ (anti-commutivity).

Geometric description

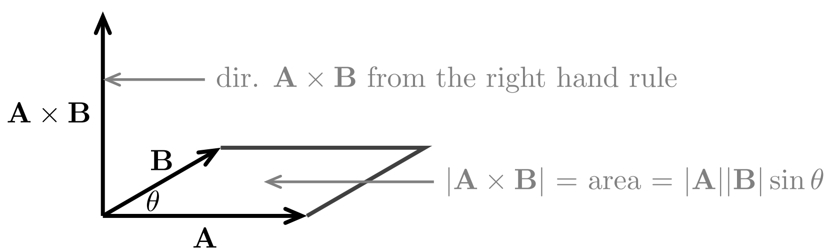
To describe the cross product geometrically we need to describe its magnitude and direction. This is done in the following theorem.

Theorem: The magnitude of $\mathbf{A}\times\mathbf{B}$ is

$$\begin{aligned}|\mathbf{A}\times\mathbf{B}| &= |\mathbf{A}||\mathbf{B}|\sin\theta, \text{ where } \theta \text{ is the angle between them}\\ &= \text{area of the parallelogram spanned by } \mathbf{A} \text{ and } \mathbf{B}.\end{aligned}$$

The direction of $\mathbf{A}\times\mathbf{B}$ is determined as follows.

$\mathbf{A}\times\mathbf{B}$ is perpendicular to the plane of $\mathbf{A}$ and $\mathbf{B}$. In the figure below there are two directions perpendicular to the plane –up and down. The choice is made by the *right hand rule*. This rule says to take your right hand and point your fingers in the direction of $\mathbf{A}$ so that they curl towards $\mathbf{B}$; then your thumb points in the direction of $\mathbf{A}\times\mathbf{B}$.

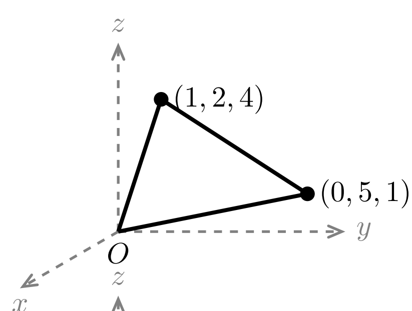


We will not go through the proof of this theorem. It makes use of the Lagrange identity

$$|\mathbf{A}\times\mathbf{B}|^2 = |\mathbf{A}|^2|\mathbf{B}|^2 - (\mathbf{A}\cdot\mathbf{B})^2.$$

This identity is easily show by expanding both sides using components.

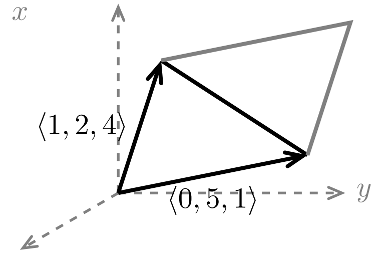
Example: Find the area of the triangle shown.



<u>**Answer:**</u>

The area of the triangle is half the area of the parallelogram (see figure).

So, area triangle $= \dfrac{1}{2}|\langle 1,2,4\rangle \times \langle 0,5,1\rangle|$.



$$\langle 1,2,4\rangle \times \langle 0,5,1\rangle = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 4 \\ 0 & 5 & 1 \end{vmatrix} = \mathbf{i}(-18) - \mathbf{j} + 5\mathbf{k}.$$

Area triangle $= \dfrac{1}{2}\sqrt{18^2+1^2+5^2} = \dfrac{1}{2}\sqrt{350}.$

DON'T FORGET THE GEOMETRY -it will be used to solve problems.

Equation of a Plane

1. Later we will return to the topic of planes in more detail. Here we will content ourself with one example.

Find the equation of the plane containing the three points $P_1 = (1,3,1)$, $\quad P_2 = (1,2,2)$, $P_3 = (2,3,3)$.

<u>**Answer:**</u>

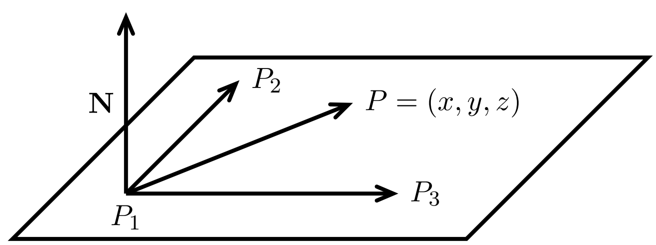
The vectors $\overrightarrow{\mathbf{P_1P_2}}$ and $\overrightarrow{\mathbf{P_1P_3}}$ are in the plane, so

$$\mathbf{N} = \overrightarrow{\mathbf{P_1P_2}} \times \overrightarrow{\mathbf{P_1P_3}} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & -1 & 1 \\ 1 & 0 & 2 \end{vmatrix} = \mathbf{i}(-2) - \mathbf{j}(-1) + \mathbf{k}(1) = \langle -2, 1, 1 \rangle.$$

is orthogonal to the plane.

Now for any point $P = (x,y,z)$ in the plane, the vector $\overrightarrow{\mathbf{P_1P}}$ is also in the plane and is therefore orthogonal to $\mathbf{N}$. Expressing this with the dot product we get

$$\begin{aligned} & \mathbf{N} \cdot \overrightarrow{\mathbf{P_1P}} = 0 \\ \Leftrightarrow \quad & \langle -2, 1, 1 \rangle \cdot \langle x-1, y-3, z-1 \rangle = 0 \\ \Leftrightarrow \quad & -2(x-1) + (y-3) + (z-1) = 0 \\ \Leftrightarrow \quad & -2x + y + z = 2. \end{aligned}$$



The equation of the plane is $-2x + y + z = 2$. You should check that the three points P_1, P_2, P_3 do, in fact, satisfy this equation.

The standard terminology for the vector $\mathbf{N}$ is to call it a *normal* to the plane.

Matrices 1. Matrix Algebra

Matrix algebra.

Previously we calculated the determinants of square arrays of numbers. Such arrays are important in mathematics and its applications; they are called *matrices*. In general, they need not be square, only rectangular.

A rectangular array of numbers having m rows and n columns is called an $m \times n$ **matrix**. The number in the i-th row and j-th column (where $1 \leq i \leq m,\ 1 \leq j \leq n$) is called the **ij-entry**, and denoted a_{ij}; the matrix itself is denoted by A, or sometimes by (a_{ij}).

Two matrices of the same size are *equal* if corresponding entries are equal.

Two special kinds of matrices are the **row-vectors**: the $1 \times n$ matrices $(a_1, a_2, \dots, a_n)$; and the **column vectors**: the $m \times 1$ matrices consisting of a column of m numbers.

> From now on, row-vectors or column-vectors will be indicated by boldface small letters; when writing them by hand, put an arrow over the symbol.

Matrix operations

There are four basic operations which produce new matrices from old.

1. Scalar multiplication: Multiply each entry by c : $cA = (ca_{ij})$

2. Matrix addition: Add the corresponding entries: $A + B = (a_{ij} + b_{ij})$; the two matrices must have the same number of rows and the same number of columns.

3. Transposition: The *transpose* of the $m \times n$ matrix A is the $n \times m$ matrix obtained by making the rows of A the columns of the new matrix. Common notations for the transpose are A^T and A'; using the first we can write its definition as $A^T = (a_{ji})$.

If the matrix A is square, you can think of A^T as the matrix obtained by flipping A over around its main diagonal.

Example 1.1 Let $A = \begin{pmatrix} 2 & -3 \\ 0 & 1 \\ -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 5 \\ -2 & 3 \\ -1 & 0 \end{pmatrix}$. Find $A+B,\ A^T,\ 2A - 3B$.

Solution. $A + B = \begin{pmatrix} 3 & 2 \\ -2 & 4 \\ -2 & 2 \end{pmatrix}$; $A^T = \begin{pmatrix} 2 & 0 & -1 \\ -3 & 1 & 2 \end{pmatrix}$;

$$2A + (-3B) = \begin{pmatrix} 4 & -6 \\ 0 & 2 \\ -2 & 4 \end{pmatrix} + \begin{pmatrix} -3 & -15 \\ 6 & -9 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -21 \\ 6 & -7 \\ 1 & 4 \end{pmatrix}.$$

4. Matrix multiplication This is the most important operation. Schematically, we have

$$\begin{array}{ccccc} A & \cdot & B & = & C \\ m \times n & & n \times p & & m \times p \end{array}$$

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

The essential points are:

1. For the multiplication to be defined, A must have as many *columns* as B has *rows*;

2. The ij-th entry of the product matrix C is the dot product of the i-th row of A with the j-th column of B.

Example 1.2 $\begin{pmatrix} 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -2+4-2 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix};$

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \begin{pmatrix} 4 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 10 \\ -4 & -5 \end{pmatrix}; \qquad \begin{pmatrix} 2 & 0 & 1 \\ 1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -2 & -6 \\ 0 & 2 & 2 \end{pmatrix}$$

The two most important types of multiplication, for multivariable calculus and differential equations, are:

1. AB, where A and B are two *square* matrices of the same size — these can always be multiplied;

2. $A\mathbf{b}$, where A is a square $n \times n$ matrix, and $\mathbf{b}$ is a column n-vector.

Laws and properties of matrix multiplication

M-1. $A(B+C) = AB + AC, \qquad (A+B)C = AC + BC$ *distributive laws*

M-2. $(AB)C = A(BC); \qquad (cA)B = c(AB).$ *associative laws*

In both cases, the matrices must have compatible dimensions.

M-3. Let $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$ then $AI = A$ and $IA = A$ for any 3×3 matrix.

I is called the **identity** matrix of order 3. There is an analogously defined square identity matrix I_n of any order n, obeying the same multiplication laws.

M-4. In general, for two square $n \times n$ matrices A and B, $AB \neq BA$: *matrix multiplication is not commutative.* (There are a few important exceptions, but they are very special — for example, the equality $AI = IA$ where I is the identity matrix.)

M-5. For two square $n \times n$ matrices A and B, we have the *determinant law*:

$$|AB| = |A||B|, \qquad \text{also written} \qquad \det(AB) = \det(A)\det(B)$$

For 2×2 matrices, this can be verified by direct calculation, but this naive method is unsuitable for larger matrices; it's better to use some theory. We will simply assume it in these notes; we will also assume the other results above (of which only the associative law **M-2** offers any difficulty in the proof).

M-6. A useful fact is this: matrix multiplication can be used to pick out a row or column of a given matrix: you multiply by a simple row or column vector to do this. Two examples should give the idea:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} \qquad \text{the second column}$$

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \qquad \text{the first row}$$

Meaning of matrix multiplication

In these examples we will explore the effect of matrix multiplication on the xy-plane.

Example 1: The matrix $A = \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}$ transforms the unit square into a parallelogram as follows.

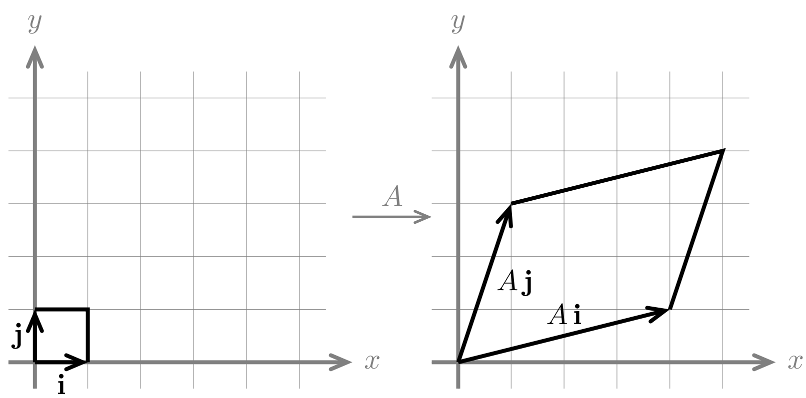
The unit square has sides $\mathbf{i}$ and $\mathbf{j}$. In order multiply a matrix times a vector we write them as column vectors. For example, $\mathbf{i} = \langle 1, 0 \rangle$, $\mathbf{j} = \langle 0, 1 \rangle$ and $\mathbf{v} = \langle a_1, a_2 \rangle$ are written

$$\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

The matrix multiplication then becomes

$$A\,\mathbf{i} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}; \quad A\,\mathbf{j} = A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

We think of the all the points in the square as the endpoints of origin vectors. If we multiply A by all of these vectors we get the following picture.



The square is mapped to the parallelogram. We know that the area of the parallelogram is $|A| = 11$. (Think about the 2×2 determinant you would use to compute the area of the parallelogram.)

18.02SC Multivariable Calculus
Fall 2010

Matrices 2. Solving Square Systems of Linear Equations; Inverse Matrices

Solving square systems of linear equations; inverse matrices.

Linear algebra is essentially about solving systems of linear equations, an important application of mathematics to real-world problems in engineering, business, and science, especially the social sciences. Here we will just stick to the most important case, where the system is *square*, i.e., there are as many variables as there are equations. In low dimensions such systems look as follows (we give a 2×2 system and a 3×3 system):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \tag{7}$$

In these systems, the a_{ij} and b_i are given, and we want to solve for the x_i.

As a simple mathematical example, consider the linear change of coordinates given by the equations

$$\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3 \\ x_2 &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3 \\ x_3 &= a_{31}y_1 + a_{32}y_2 + a_{33}y_3 \end{aligned}$$

If we know the y-coordinates of a point, then these equations tell us its x-coordinates immediately. But if instead we are given the x-coordinates, to find the y-coordinates we must solve a system of equations like (7) above, with the y_i as the unknowns.

Using matrix multiplication, we can abbreviate the system on the right in (7) by

$$A\mathbf{x} = \mathbf{b}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \tag{8}$$

where A is the square matrix of coefficients $(\, a_{ij}\,)$. (The 2×2 system and the $n \times n$ system would be written analogously; all of them are abbreviated by the same equation $A\mathbf{x} = \mathbf{b}$, notice.)

You have had experience with solving small systems like (7) by *elimination*: multiplying the equations by constants and subtracting them from each other, the purpose being to eliminate all the variables but one. When elimination is done systematically, it is an efficient method. Here however we want to talk about another method more compatible with hand-held calculators and MatLab, and which leads more rapidly to certain key ideas and results in linear algebra.

Inverse matrices.

Referring to the system (8), suppose we can find a square matrix M, the same size as A, such that

$$MA = I \qquad \text{(the identity matrix).} \tag{9}$$

We can then solve (8) by matrix multiplication, using the successive steps,

$$\begin{aligned} A\,\mathbf{x} &= \mathbf{b} \\ M(A\,\mathbf{x}) &= M\,\mathbf{b} \\ \mathbf{x} &= M\,\mathbf{b}; \end{aligned} \tag{10}$$

where the step $M(A\mathbf{x}) = \mathbf{x}$ is justified by

$$\begin{aligned} M(A\,\mathbf{x}) &= (MA)\mathbf{x}, && \text{by associative law;} \\ &= I\,\mathbf{x}, && \text{by (9);} \\ &= \mathbf{x}, && \text{because } I \text{ is the identity matrix .} \end{aligned}$$

Moreover, the solution is unique, since (10) gives an explicit formula for it.

The same procedure solves the problem of determining the inverse to the linear change of coordinates $\mathbf{x} = A\mathbf{y}$, as the next example illustrates.

Example 2.1 Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ and $M = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}$. Verify that M satisfies (9) above, and use it to solve the first system below for x_i and the second for the y_i in terms of the x_i:

$$\begin{aligned} x_1 + 2x_2 &= -1 \qquad\qquad & x_1 &= \ \ y_1 + 2y_2 \\ 2x_1 + 3x_2 &= 4 & x_2 &= 2y_1 + 3y_2 \end{aligned}$$

Solution. We have $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, by matrix multiplication. To solve the first system, we have by (10), $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 11 \\ -6 \end{pmatrix}$, so the solution is $x_1 = 11, x_2 = -6$. By reasoning similar to that used above in going from $A\mathbf{x} = \mathbf{b}$ to $\mathbf{x} = M\mathbf{b}$, the solution to $\mathbf{x} = A\mathbf{y}$ is $\mathbf{y} = M\mathbf{x}$, so that we get

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

as the expression for the y_i in terms of the x_i.

Our problem now is: how do we get the matrix M? In practice, you mostly press a key on the calculator, or type a Matlab command. But we need to be able to work abstractly with the matrix — i.e., with symbols, not just numbers, and for this some theoretical ideas are important. The first is that M doesn't always exist.

$$M \ \text{ exists} \quad \Leftrightarrow \quad |A| \neq 0.$$

The implication $\Rightarrow$ follows immediately from the law **M-5** in section M.1 ($\det(AB) = \det(A)\det(B)$) , since

$$MA = I \quad \Rightarrow \quad |M||A| = |I| = 1 \quad \Rightarrow \quad |A| \neq 0.$$

The implication in the other direction requires more; for the low-dimensional cases, we will produce a formula for M. Let's go to the formal definition first, and give M its proper name, A^{-1}:

Definition. Let A be an $n \times n$ matrix, with $|A| \neq 0$. Then the **inverse** of A is an $n \times n$ matrix, written A^{-1}, such that

$$A^{-1}A = I_n, \qquad AA^{-1} = I_n \tag{11}$$

(It is actually enough to verify either equation; the other follows automatically — see the exercises.)

Using the above notation, our previous reasoning (9) - (10) shows that

$$|A| \neq 0 \quad \Rightarrow \quad \text{the unique solution of } \; A\,\mathbf{x} = \mathbf{b} \; \text{ is } \; \mathbf{x} = A^{-1}\mathbf{b}; \tag{12}$$

$$|A| \neq 0 \quad \Rightarrow \quad \text{the solution of } \; \mathbf{x} = A\,\mathbf{y} \; \text{ for the } y_i \text{ is } \; \mathbf{y} = A^{-1}\mathbf{x}. \tag{12}$$

Calculating the inverse of a 3×3 matrix

Let A be the matrix. The formulas for its **inverse** A^{-1} and for an auxiliary matrix $\operatorname{adj} A$ called the **adjoint** of A (or in some books the **adjugate** of A) are

$$A^{-1} \;=\; \frac{1}{|A|}\operatorname{adj} A \;=\; \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}^T . \tag{13}$$

In the formula, A_{ij} is the cofactor of the element a_{ij} in the matrix, i.e., its minor with its sign changed by the checkerboard rule (see section 1 on determinants).

Formula (13) shows that the steps in calculating the inverse matrix are:

1. Calculate the matrix of minors.
2. Change the signs of the entries according to the checkerboard rule.
3. Transpose the resulting matrix; this gives $\operatorname{adj} A$.
4. Divide every entry by $|A|$.

(If inconvenient, for example if it would produce a matrix having fractions for every entry, you can just leave the $1/|A|$ factor outside, as in the formula. Note that step 4 can only be taken if $|A| \neq 0$, so if you haven't checked this before, you'll be reminded of it now.)

The notation A_{ij} for a cofactor makes it look like a matrix, rather than a signed determinant; this isn't good, but we can live with it.

Example 2.2 Find the inverse to $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

Solution. We calculate that $|A| = 2$. Then the steps are (T means transpose):

$$\underset{\text{matrix } A}{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}} \to \underset{\text{cofactor matrix}}{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}} \underset{\text{T}}{\to} \underset{\text{adj } A}{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ -1 & 0 & 1 \end{pmatrix}} \to \underset{\text{inverse of } A}{\begin{pmatrix} \frac12 & 0 & \frac12 \\ \frac12 & 1 & -\frac12 \\ -\frac12 & 0 & \frac12 \end{pmatrix}}$$

To get practice in matrix multiplication, check that $A \cdot A^{-1} = I$, or to avoid the fractions, check that $A \cdot \text{adj}\,(A) = 2I$.

The same procedure works for calculating the inverse of a 2×2 matrix A. We do it for a general matrix, since it will save you time in differential equations if you can learn the resulting formula.

$$\underset{\text{matrix } A}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \to \underset{\text{cofactors}}{\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}} \underset{\text{T}}{\to} \underset{\text{adj } A}{\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}} \to \underset{\text{inverse of } A}{\frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

Example 2.3 Find the inverses to: a) $\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix}$

Solution. a) Use the formula: $|A| = 2$, so $A^{-1} = \dfrac12 \begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac32 & \frac12 \end{pmatrix}$.

b) Follow the previous scheme:

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix} \to \begin{pmatrix} -5 & -3 & 7 \\ 2 & 0 & -1 \\ 4 & 3 & -5 \end{pmatrix} \to \begin{pmatrix} -5 & 2 & 4 \\ -3 & 0 & 3 \\ 7 & -1 & -5 \end{pmatrix} \to \frac13 \begin{pmatrix} -5 & 2 & 4 \\ -3 & 0 & 3 \\ 7 & -1 & -5 \end{pmatrix} = A^{-1}.$$

Both solutions should be checked by multiplying the answer by the respective A.

Proof of formula (13) for the inverse matrix.

We want to show $A \cdot A^{-1} = I$, or equivalently, $A \cdot \text{adj}\ A = |A|I$; when this last is written out using (13) (remembering to transpose the matrix on the right there), it becomes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} |A| & 0 & 0 \\ 0 & |A| & 0 \\ 0 & 0 & |A| \end{pmatrix}. \tag{14}$$

To prove (14), it will be enough to look at two typical entries in the matrix on the right — say the first two in the top row. According to the rule for multiplying the two matrices on the left, what we have to show is that

$$a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} = |A|; \tag{15}$$

$$a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0 \tag{16}$$

These two equations are both evaluating determinants by Laplace expansions: the first equation (15) evaluates the determinant on the left below by the cofactors of the first row; the second equation (16) evaluates the determinant on the right below by the cofactors of the second row (notice that the cofactors of the second row don't care what's actually in the second row, since to calculate them you only need to know the other two rows).

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad\qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

The two equations (15) and (16) now follow, since the determinant on the left is just $|A|$, while the determinant on the right is 0, since two of its rows are the same. □

The procedure we have given for calculating an inverse works for $n \times n$ matrices, but gets to be too cumbersome if $n > 3$, and other methods are used. The calculation of A^{-1} for reasonable-sized A is a standard package in computer algebra programs and MatLab. Unfortunately, social scientists often want the inverses of very large matrices, and for this special techniques have had to be devised, which produce approximate but acceptable results.

Equations of planes

We have touched on equations of planes previously. Here we will fill in some of the details.

Planes in point-normal form

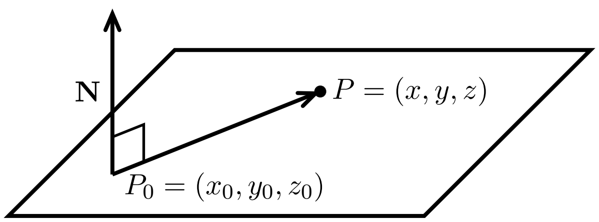
The basic data which determines a plane is a point P_0 in the plane and a vector $\mathbf{N}$ orthogonal to the plane. We call $\mathbf{N}$ a *normal* to the plane and we will sometimes say $\mathbf{N}$ is *normal* to the plane, instead of orthogonal.

Now, suppose we want the equation of a plane and we have a point $P_0 = (x_0, y_0, z_0)$ in the plane and a vector $\overrightarrow{\mathbf{N}} = \langle a, b, c\rangle$ normal to the plane.

Let $P = (x, y, z)$ be an arbitrary point in the plane. Then the vector $\overrightarrow{\mathbf{P_0P}}$ is in the plane and therefore orthogonal to $\mathbf{N}$. This means

$$\begin{aligned} & \mathbf{N} \cdot \overrightarrow{\mathbf{P_0P}} = 0 \\ \Leftrightarrow \quad & \langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0 \\ \Leftrightarrow \quad & a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \end{aligned}$$

We call this last equation the point-normal form for the plane.



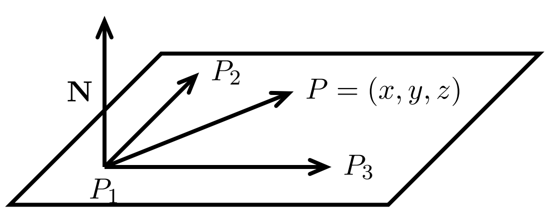
Example 1: Find the plane through the point (1,4,9) with normal $\langle 2, 3, 4\rangle$.

<u>Answer:</u> Point-normal form of the plane is $2(x-1)+3(y-4)+4(z-9) = 0$. We can also write this as $2x + 3y + 4z = 50$.

Example 2: Find the plane containing the points $P_1 = (1, 2, 3)$, $P_2 = (0, 0, 3)$, $P_3 = (2, 5, 5)$.

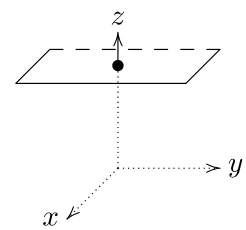
<u>Answer:</u> The goal is to find the basic data, i.e. a point in the plane and a normal to the plane. The point is easy, we already have three of them. To get the normal we note (see figure below) that $\overrightarrow{\mathbf{P_1P_2}}$ and $\overrightarrow{\mathbf{P_1P_3}}$ are vectors in the plane, so their cross product is orthogonal (normal) to the plane. That is,

$$\mathbf{N} = \overrightarrow{\mathbf{P_1P_2}} \times \overrightarrow{\mathbf{P_1P_3}} = \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & -2 & 0 \\ 1 & 3 & 2 \end{pmatrix} = -4\mathbf{i} - \mathbf{j}(-2) + \mathbf{k}(-1) = \langle -4, 2, -1\rangle.$$



Using point-normal form (with point P_1) the equation of the plane is

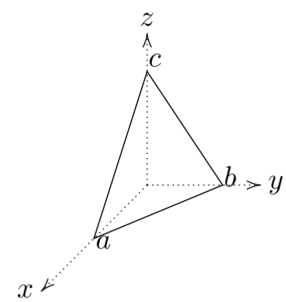
$$-4(x-1)+2(y-2)-(z-3)=0, \text{ or equivalently } \quad -4x+2y-z=-3.$$



Example 3: Find the plane with normal $\mathbf{N}=\widehat{\mathbf{k}}$ containing the point (0,0,3)
Eq. of plane: $\langle 0,0,1\rangle \cdot \langle x,y,z-3\rangle = 0 \Leftrightarrow z=3$.

Example 4: Find the plane with x, y and z intercepts a, b and c.
__**Answer:**__ We could find this using the method example 1. Instead, we'll use a shortcut that works when all the intercepts are known. In this case, the intercepts are

$$(a,0,0), \quad (0,b,0), \quad (0,0,c)$$



and we simply write the plane as

$$x/a+y/b+z/c=1.$$

You can easily check that each of the given points is on the plane.
For completeness we'll do this using the more general method of example 1.
The 3 points give us 2 vectors in the plane, $\langle -a,b,0\rangle$ and $\langle -a,0,c\rangle$.
$\Rightarrow$ $\mathbf{N}=\langle -a,b,0\rangle \times \langle -a,0,c\rangle = \langle bc,ac,ab\rangle$.
Point-normal form: $bc(x-a)+ac(y-0)+ab(z-0)=0$
$\Leftrightarrow$ $bc\,x+ac\,y+ab\,z=abc \Leftrightarrow x/a+y/b+z/c=1$.

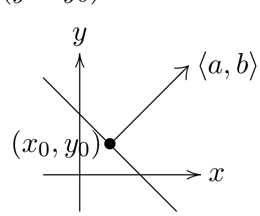
Lines in the plane

While we're at it, let's look at two ways to write the equation of a line in the xy-plane.

Slope-intercept form: Given the slope m and the y-intercept b the equation of a line can be written $y=mx+b$.

Point-normal form:
We can also use point-normal form to find the equation of a line.

Given a point (x_0,y_0) on the line and a vector $\langle a,b\rangle$ normal to the line the equation of the line can be written $a(x-x_0)+b(y-y_0)=0$.



Distances to planes and lines

In this note we will look at distances to planes and lines. Our approach is geometric. Very broadly, we will draw a sketch and use vector techniques.

Please note is that our sketches are not oriented, drawn to scale or drawn in perspective. Rather they are a simple 'cartoon' which shows the important features of the problem.

1. *Distance: point to plane:*
Ingredients: i) A point P, ii) A plane with normal $\overrightarrow{\mathbf{N}}$ and containing a point Q.

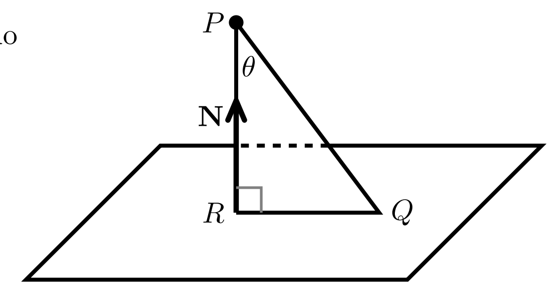
The distance from P to the plane is $d = |\overrightarrow{\mathbf{PQ}}|\cos\theta = \left|\overrightarrow{\mathbf{PQ}} \cdot \dfrac{\overrightarrow{\mathbf{N}}}{|\mathbf{N}|}\right|$.

We will explain this formula by way of the following example.

Example 1: Let $P = (1, 3, 2)$. Find the distance from P to the plane $x + 2y = 3$.

<u>**Answer:**</u> First we gather our ingredients.
$Q = (3, 0, 0)$ is a point on the plane (it is easy to find such a point).
$\mathbf{N}$ = normal to plane = $\mathbf{i} + 2\mathbf{j}$.
R = point on plane closest to P (this is point unknown and we do not need to find it to find the distance). The figure shows that

$$\text{distance} = |PR| = \left|\overrightarrow{\mathbf{PQ}}\right|\cos\theta = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right|.$$



Computing $\overrightarrow{\mathbf{PQ}} = 2\mathbf{i} - 3\mathbf{j} - 2\mathbf{k}$ gives

$$\text{distance} = \left|\overrightarrow{\mathbf{PQ}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}\right| = \left|\langle 2, -3, -2\rangle \cdot \frac{\langle 1, 2, 0\rangle}{\sqrt{5}}\right| = \frac{4}{\sqrt{5}}.$$

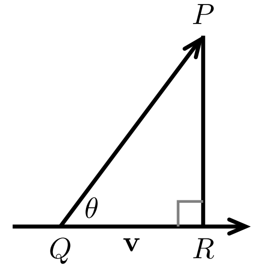
2. *Distance: point to line:*
Ingredients: i) A point P, ii) A line with direction vector $\mathbf{v}$ and containing a point Q.

The distance from P to the line is $d = |\mathbf{QP}|\sin\theta = \left|\overrightarrow{\mathbf{QP}} \times \dfrac{\mathbf{v}}{|\mathbf{v}|}\right|$.

We will explain this formula by way of the following example.

Example 2: Let $P = (1, 3, 2)$, find the distance from the point P to the line through $(1, 0, 0)$ and $(1, 2, 0)$.

<u>**Answer:**</u> First we gather our ingredients.
$Q = (1, 0, 0)$ (this is easy to find).
$\mathbf{v} = \langle 1,\ 2,\ 0\rangle - \langle 1,\ 0,\ 0\rangle = 2\mathbf{j}$ is parallel to the line.
R = point on line closest to P (this is point is unknown).
Using the relation $|\mathbf{A} \times \mathbf{B}| = |\mathbf{A}||\mathbf{B}|\sin\theta$, the figure shows that
$\text{distance} = |PR| = \left|\overrightarrow{\mathbf{PQ}}\right|\sin\theta = \left|\overrightarrow{\mathbf{QP}} \times \dfrac{\mathbf{v}}{|\mathbf{v}|}\right|$.



Computing: $\overrightarrow{\mathbf{PQ}} = 3\mathbf{j} + 2\mathbf{k}$, which implies $\left|\overrightarrow{\mathbf{QP}} \times \frac{\mathbf{v}}{|\mathbf{v}|}\right| = |(3\mathbf{j} + 2\mathbf{k}) \times \mathbf{j}| = |-2\mathbf{i}| = 2$.

3. *Distance between parallel planes:*

The trick here is to reduce it to the distance from a point to a plane.

Example 3: Find the distance between the planes $x + 2y - z = 4$ and $x + 2y - z = 3$.

Both planes have normal $\mathbf{N} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$ so they are parallel.

Take any point on the first plane, say, $P = (4, 0, 0)$.

Distance between planes = distance from P to second plane.

Choose $Q = (1, 0, 0)$ = point on second plane

$\Rightarrow \ d = |\overrightarrow{\mathbf{QP}} \cdot \frac{\mathbf{N}}{|\mathbf{N}|}| = |3\mathbf{i} \cdot (\mathbf{i} + 2\mathbf{j} - \mathbf{k})|/\sqrt{6} = \sqrt{6}/2.$

4. *Distance between skew lines:*

We place the lines in parallel planes and find the distance between the planes as in the previous example

As usual it's easy to find a point on each line. Thus, to find the parallel planes we only need to find the normal.

$$\mathbf{N} = \mathbf{v_1} \times \mathbf{v_2},$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the direction vectors of the lines.

18.02SC Multivariable Calculus
Fall 2010

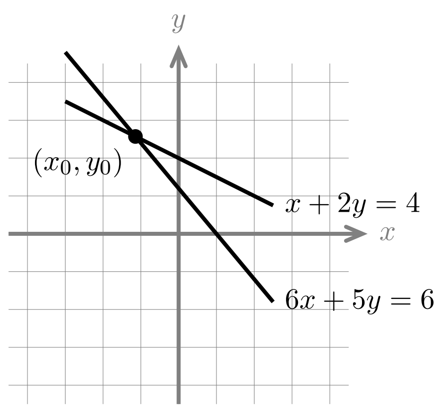
Geometry of linear systems of equations

Very often in math, science and engineering we need to solve a linear system of equations. A simple example of such a system is given by

$$\begin{array}{rcrcl} 6x & + & 5y & = & 6 \\ x & + & 2y & = & 4. \end{array}$$

You have probably already learned algebraic techniques to solve such a system. Later we will also learn to solve such a system using matrix algebra. For now we will focus on the geometric view of this system.

Solving the system means finding a pair (x_0, y_0) which satisfies both equations. Geometrically each of the equations represents a line. That is, each pair (x, y) satisfying the equation is a point on the line. Thus, the solution (x_0, y_0) is the point where the two lines intersect.

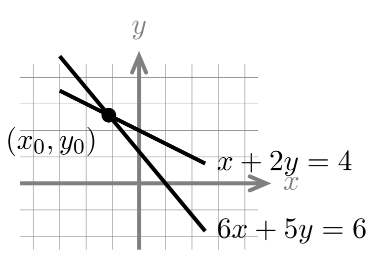


From the graph we can approximate the solution (the exact solution is $(-8/7, 18/7)$), but our interest here is in how many solutions there can be.

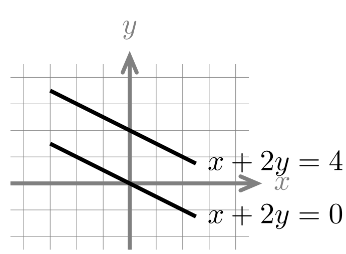
The geometric picture makes this obvious. Here are the three possibilities.

1. The two lines intersect in a point, so there is one solution.
2. The two lines are parallel (and not the same), so there are no solutions.
3. The two lines are the same, so there are an infinite number of solutions.

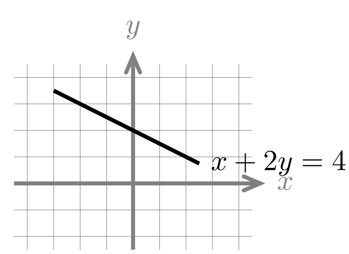
Here are example systems and graphs.



$6x + 5y = 6$
$x + 2y = 4$
(intersecting lines: 1 solution)



$x + 2y = 4$
$x + 2y = 0$
(parallel lines: no solutions)



$x + 2y = 4$
$x + 2y = 4$
(the same line: ∞ solutions)

3 × 3 systems

For 3×3 systems there are more possibilities. For example, consider the system

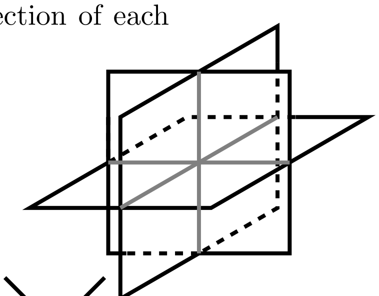
$$\begin{array}{rcrcrcr} 6x & + & 5y & + & 3z & = & 1 \\ x & + & 2y & + & z & = & 4 \\ 2x & - & 2y & - & 2z & = & 8. \end{array}$$

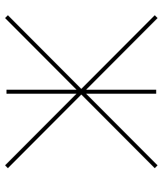
Each equation is the equation of a plane, so, geometrically, solving the system means finding the intersection of three planes, i.e., the point or points which lie on all three planes.

Usually, three planes intersect in a point. You can visualize this by first imagining two of the planes intersecting in a line and then the line intersecting the third plane in a point. Altogether there are four possibilities.

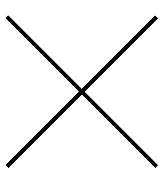
1. Intersect in a point (1 solution to system).
2. Intersect in a line (∞ solutions).
   a) Three different planes, the third plane contains the line of intersection of the first two.
   b) Two planes are the same, the third plane intersects them in a line.
3. Intersect in a plane (∞ solutions)
   a) All three planes are the same.
4. The planes don't all intersect at any point (0 solutions).
   a) The planes are different, but all parallel.
   b) Two planes are parallel, the third crosses them.
   c) The planes are different and none are parallel. but the lines of intersection of each pair are parallel.
   d) Two planes are the same and parallel to the third.

To visualize this we could draw three dimensional figures, for example the figure at the right shows three planes intersecting in a point. Instead, we will visualize the other cases by drawing lines on the page and imagining the planes as extending vertically out of the page.

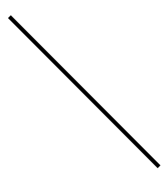




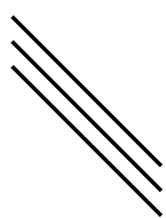
Case (2a)



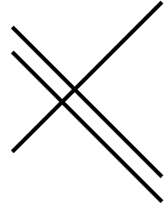
Case (2b)



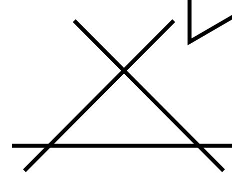
Case (3a)



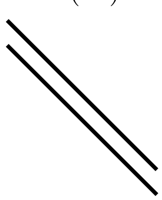
Case (4a)



Case (4b)



Case (4c)



Case (4d)

Matrices 3. Homogeneous and Inhomogeneous Systems

Theorems about homogeneous and inhomogeneous systems.

On the basis of our work so far, we can formulate a few general results about square systems of linear equations. They are the theorems most frequently referred to in the applications.

Definition. The linear system $A\mathbf{x} = \mathbf{b}$ is called **homogeneous** if $\mathbf{b} = \mathbf{0}$; otherwise, it is called **inhomogeneous**.

Theorem 1. *Let A be an $n \times n$ matrix.*

(20) $\qquad |A| \neq 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{b}$ *has the unique solution,* $\mathbf{x} = A^{-1}\mathbf{b}$.

(21) $\qquad |A| \neq 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{0}$ *has only the trivial solution,* $\mathbf{x} = \mathbf{0}$.

Notice that (21) is the special case of (20) where $\mathbf{b} = \mathbf{0}$. Often it is stated and used in the contrapositive form:

(21′) $\qquad A\mathbf{x} = \mathbf{0}$ has a non-zero solution $\quad \Rightarrow \quad |A| = 0$.

(The contrapositive of a statement $P \Rightarrow Q$ is not-$Q \Rightarrow$ not-P; the two statements say the same thing.)

Theorem 2. *Let A be an $n \times n$ matrix.*

(22) $\qquad |A| = 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{0}$ *has non-trivial (i.e., non-zero) solutions.*

(23) $\qquad |A| = 0 \quad \Rightarrow \quad A\,\mathbf{x} = \mathbf{b}$ *usually has no solutions, but has solutions for some* $\mathbf{b}$.

In (23), we call the system **consistent** if it has solutions, **inconsistent** otherwise.

This probably seems like a maze of similar-sounding and confusing theorems. Let's get another perspective on these ideas by seeing how they apply separately to homogeneous and inhomogeneous systems.

Homogeneous systems: $A\mathbf{x} = \mathbf{0}$ *has non-trivial solutions* $\Leftrightarrow |A| = 0$.

Inhomogeneous systems: $A\mathbf{x} = \mathbf{b}$ *has the unique solution* $\mathbf{x} = A^{-1}\mathbf{b}$, *if* $|A| \neq 0$. *If* $|A| = 0$, *then* $A\mathbf{x} = \mathbf{b}$ *usually has no solutions, but does have solutions for some* $\mathbf{b}$.

The statements (20) and (21) are proved, since we have a formula for the solution, and it is easy to see by multiplying $A\,\mathbf{x} = \mathbf{b}$ by A^{-1} that if $\mathbf{x}$ is a solution, it must be of the form $\mathbf{x} = A^{-1}\mathbf{b}$.

We prove (22) just for the 3×3 case, by interpreting it geometrically. We will give a partial argument for (23), based on both algebra and geometry.

Proof of (22).

We represent the three rows of A by the row vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and we let $\mathbf{x} = (x, y, z)$; think of all these as origin vectors, i.e., place their tails at the origin. Then, considering the homogeneous system first,

(24) $\qquad A\,\mathbf{x} = \mathbf{0} \qquad$ is the same as the system $\qquad \mathbf{a} \cdot \mathbf{x} = 0, \quad \mathbf{b} \cdot \mathbf{x} = 0, \quad \mathbf{c} \cdot \mathbf{x} = 0.$

In other words, we are looking for a row vector $\mathbf{x}$ which is orthogonal to three given vectors, namely the three rows of A. By Section 1, we have

$$|A| = \mathbf{a}\cdot\mathbf{b}\times\mathbf{c} \;=\; \text{volume of parallelepiped spanned by } \mathbf{a},\mathbf{b},\mathbf{c}.$$

It follows that if $|A| = 0$, the parallelepiped has zero volume, and therefore the origin vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ lie in a plane. Any non-zero vector $\mathbf{x}$ which is orthogonal to this plane will then be orthogonal to $\mathbf{a},\mathbf{b},\mathbf{c}$, and therefore will be a solution to the system (24). This proves (22): if $|A| = 0$, then $A\,\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

Partial proof of (23). We write the system as $A\,\mathbf{x} = \mathbf{d}$, where $\mathbf{d}$ is the column vector $\mathbf{d} = (d_1, d_2, d_3)^T$.

Writing this out as we did in (24), it becomes the system

(25) $$\mathbf{a}\cdot\mathbf{x} = d_1, \quad \mathbf{b}\cdot\mathbf{x} = d_2, \quad \mathbf{c}\cdot\mathbf{x} = d_3 \ .$$

If $|A| = 0$, the three origin vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ lie in a plane, which means we can write one of them, say $\mathbf{c}$, as a linear combination of $\mathbf{a}$ and $\mathbf{b}$:

(26) $$\mathbf{c} = r\mathbf{a} + s\mathbf{b}, \qquad r.s \text{ real numbers.}$$

Then if $\mathbf{x}$ is any vector, it follows that

(27) $$\mathbf{c}\cdot\mathbf{x} = r(\mathbf{a}\cdot\mathbf{x}) + s(\mathbf{b}\cdot\mathbf{x}) \ .$$

Now if $\mathbf{x}$ is also a solution to (25), we see from (25) and (27) that

(28) $$d_3 = rd_1 + sd_2;$$

this shows that unless the components of $\mathbf{d}$ satisfy the relation (28), there cannot be a solution to (25); thus in general there are no solutions.

If however, $\mathbf{d}$ does satisfy the relation (28), then the last equation in (25) is a consequence of the first two and can be discarded, and we get a system of two equations in three unknowns, which will in general have a non-zero solution, unless they represent two planes which are parallel.

Singular matrices; computational difficulties.

Because so much depends on whether $|A|$ is zero or not, this property is given a name. We say the square matrix A is **singular** if $|A| = 0$, and **nonsingular** or **invertible** if $|A| \neq 0$.

> Indeed, we know that A^{-1} exists if and only if $|A| \neq 0$, which explains the term "invertible"; the use of "singular" will be familiar to Sherlock Holmes fans: it is the 19th century version of "peculiar" or the late 20th century word "special".

Even if A is nonsingular, the solution of $A\mathbf{x} = \mathbf{b}$ is likely to run into trouble if $|A| \approx 0$, or as one says, A is *almost-singular*. Namely, in the formula for A^{-1} the $|A|$ occurs in the denominator, so that unless there is some sort of compensation for this in the numerator, the solutions are likely to be very sensitive to small changes in the coefficients of A, i.e., to the coefficients of the equations. Systems (of any kind) whose solutions behave this way are said to be **ill-conditioned**; it is difficult to solve such systems numerically and special methods must be used.

To see the difficulty geometrically, think of a 2×2 system $A\mathbf{x} = \mathbf{b}$ as representing a pair of lines; the solution is the point in which they intersect. If $|A| \approx 0$, but its entries are not small, then its two rows must be vectors which are almost parallel (since they span a parallelogram of small area). The two lines are therefore almost parallel; their intersection point exists, but its position is highly sensitive to the exact positions of the two lines, i.e., to the values of the coefficients of the system of equations.

Solutions to linear systems

1. Consider the system of equations

$$\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 1 \\ 4x & + & 5y & + & 6z & = & 2 \\ 7x & + & 8y & + & cz & = & 3. \end{array}$$

a) Write the system in matrix form.

b) For which values of c is there exactly one solution?

c) For which values of c are there either 0 or infinitely many solutions?

d) Take the corresponding homogeneous system

$$\begin{array}{rcrcrcl} x & + & 2y & + & 3z & = & 0 \\ 4x & + & 5y & + & 6z & = & 0 \\ 7x & + & 8y & + & cz & = & 0. \end{array}$$

For the value(s) of c found in part (c) give *all* the solutions.

<u>Answer:</u> a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

b) There is exactly one solution when the coefficient matrix has an inverse (i.e., is *invertible*). This happens when the determinant is not zero.

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & c \end{vmatrix} = 1(5c - 48) - 2(4c - 42) + 3(32 - 35) = -3c + 27 = 0 \;\Leftrightarrow\; c = 9.$$

There is exactly one solution as long as $c \neq 9$.

c) This is just the complement of part (b): there are zero or infinitely many solutions when $c = 9$.

d) Setting $c = 9$ our coefficient matrix is $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Thinking of matrix multiplication as a series of dot products between rows of the left matrix and column(s) of the right one we see that in solving

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

we are looking for vectors $\langle x, y, z\rangle$ that are orthogonal to each of the rows of A. Since $\det(A) = 0$, the rows are all in a plane and we can find orthogonal vectors by taking a cross product of (say) the first two rows.

$$\langle 1, 2, 3\rangle \times \langle 4, 5, 6\rangle = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} = \langle -3, 6, -3\rangle.$$

Since scaling will preserve orthogonality, all the solutions are scalar multiples, i.e., all the solutions are of the form $(x, y, z) = (-3a, 6a, -3a)$. We can make this a little nicer by removing the common factor of three,

$$(x, y, z) = (-a, 2a, -a) = a(-1, 2, -1).$$

Parametric equations of lines

General parametric equations

In this part of the unit we are going to look at parametric curves. This is simply the idea that a point moving in space traces out a path over time. Thus there are four variables to consider, the position of the point (x, y, z) and an independent variable t, which we can think of as time. (If the point is moving in plane there are only three variables, the position of the point (x, y) and the time t.)

Since the position of the point depends on t we write

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

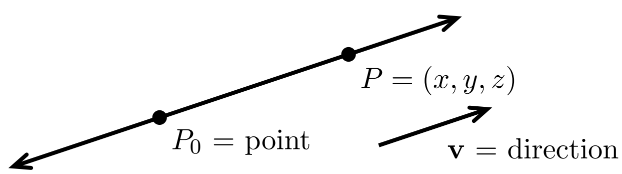
to indicate that x, y and z are functions of t. We call t the parameter and the equations for x, y and z are called *parametric equations*.

In physical examples the parameter often represents time. We will see other cases where the parameter has a different interpretation, or even no interpretation.

Parametric equations of lines

Later we will look at general curves. Right now, let's suppose our point moves on a line.

The basic data we need in order to specify a line are a point on the line and a vector parallel to the line. That is, we need a point and a direction.



Example 1: Write parametric equations for a line through the point $P_0 = (1, 2, 3)$ and parallel to the vector $\mathbf{v} = \langle 1, 3, 5 \rangle$.

<u>Answer:</u> If $P = (x, y, z)$ is on the line then the vector

$$\overrightarrow{\mathbf{P_0P}} = \langle x - 1, y - 2, z - 3 \rangle$$

is parallel to $\langle 1, 3, 5 \rangle$. That is, $\overrightarrow{\mathbf{P_0P}}$ is a scalar multiple of $\langle 1, 3, 5 \rangle$. We call the scale t and write:

$$\begin{aligned} & \langle x, y, z \rangle = \langle x - 1, y - 2, z - 3 \rangle = t\langle 1, 3, 5 \rangle \\ \Leftrightarrow \quad & x - 1 = t, \; y - 2 = 3t, \quad z - 3 = 5t \\ \Leftrightarrow \quad & x = 1 + t, \; y = 2 + 3t, \quad z = 3 + 5t. \end{aligned}$$

Example 2: In example 1, if our direction vector was $\langle 2, 6, 10 \rangle = 2\mathbf{v}$ we would get the same line with a different parametrization. That is, the moving point's trajectory would follow the same path as the trajectory in example 1 , but would arrive at each point on the line at a different time.

Example 3: In general, the line through $P_0 = (x_0, y_0, z_0)$ in the direction of (i.e., parallel to) $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$ has parametrization

$$\begin{aligned} & \langle x, y, z \rangle = \langle x_0 + tv_1, y_0 + tv_2, z_0 + tv_3 \rangle \\ \Leftrightarrow \quad & x = x_0 + tv_1, \quad y = y_0 + tv_2 \quad z = z_0 + tv_3. \end{aligned}$$

Example 4: Find the line through the point $P_0 = (1,2,3)$ and $P_1 = (2,5,8)$.

Answer: We use the data given to find the basic data (a point and direction vector) for the line.

We're given a point, $P_0 = (1,2,3)$. The direction vector $\mathbf{v} = \overrightarrow{\mathbf{P_0P_1}} = \langle 1,3,5\rangle$. So, we get

$$\begin{aligned} & \langle x,y,z\rangle = \overrightarrow{\mathbf{OP_0}} + t\mathbf{v} = \langle 1+t, 2+3t, 3+5t\rangle \\ \Leftrightarrow \quad & x = 1+t, \quad y = 2+3t, \quad z = 3+5t. \end{aligned}$$

Intersection of a line and a plane

1. Consider the plane $\mathcal{P} = 2x + y - 4z = 4$.

a) Find all points of intersection of $\mathcal{P}$ with the line

$$x = t, \quad y = 2 + 3t, \quad z = t.$$

b) Find all points of intersection of $\mathcal{P}$ with the line

$$x = 1 + t, \quad y = 4 + 2t, \quad z = t.$$

c) Find all points of intersection of $\mathcal{P}$ with the line

$$x = t, \quad y = 4 + 2t, \quad z = t.$$

Answer: a) To find the intersection we substitute the formulas for x, y and z into the equation for $\mathcal{P}$ and solve for t.

$$2(t) + (2 + 3t) - 4(t) = 4 \;\Leftrightarrow\; t = 2.$$

Now use $t = 2$ to find the point of intersection: $(x, y, z) = (2, 8, 2)$.

b) Substituting gives

$$2(1 + t) + (4 + 2t) - 4(t) = 4 \;\Leftrightarrow\; 6 = 4. \;\Leftrightarrow\; \text{no values of } t \text{ satisfy this equation.}$$

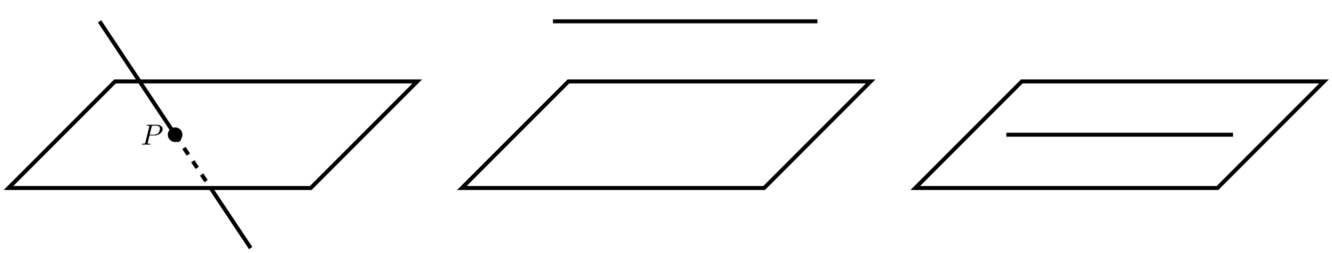
There are no points of intersection.

c) Substituting gives

$$2(t) + (4 + 2t) - 4(t) = 4 \;\Leftrightarrow\; 4 = 4. \;\Leftrightarrow\; \text{all values of } t \text{ satisfy this equation.}$$

The line is contained in the plane, i.e., all points of the line are in its intersection with the plane.

Here are cartoon sketches of each part of this problem.



(a) line intersects the plane in a point

(b) line is parallel to the plane

(c) line is in the plane

Parametric Curves

General parametric equations
We have seen parametric equations for lines. Now we will look at parametric equations of more general trajectories. Repeating what was said earlier, a parametric curve is simply the idea that a point moving in the space traces out a path.

We can use a parameter to describe this motion. Quite often we will use t as the parameter and think of it as time. Since the position of the point depends on t we write

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

to indicate that x, y and z are functions of t. We call t the parameter and the equations for x, y and z are called *parametric equations*.

It is not always necessary to think of the parameter as representing time. We will see cases where it is more convenient to express the position as a function of some other variable.

The position vector
In order to use vector techniques we define the *position vector*

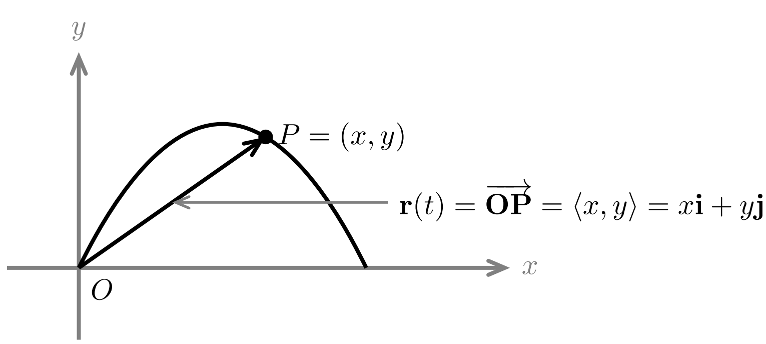
$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} = \langle x(t), y(t), z(t)\rangle.$$

This is just the vector from the origin to the moving point. As the point moves so does the position vector –see the figure with example 1.

Example 1: Thomas Pynchon fires a rocket from the origin. Its initial x-velocity is $v_{0,x}$ and its initial y-velocity is $v_{0,y}$.
You've probably seen this, but in any case, physics tells us that the parametric equations for its parabolic trajectory are

$$x(t) = v_{0,x}t, \quad y(t) = -\frac{1}{2}gt^2 + v_{0,y}t.$$

At time t the rocket is at point $P = (x(t), y(t))$. The position vector can be written in many different ways: $\mathbf{r}(t) = \overrightarrow{\mathbf{OP}} = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} = \langle x, y\rangle$.

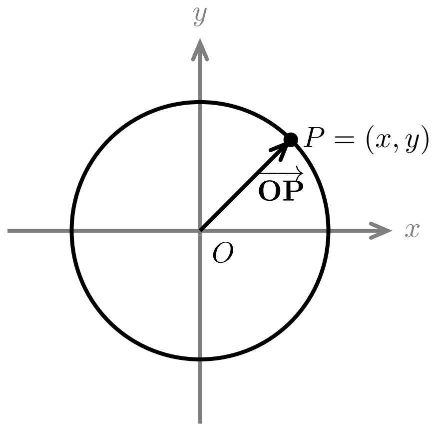


Next we will give a series of examples of parametrized curves. The most important are circles and lines. The last one is the *cycloid*. It is an important example which combines lines and circles.

Circles and ellipses
Consider the parametric curve in the plane

$$x(t) = a\cos t, \quad y(t) = a\sin t.$$

Easily we get the relation $x^2 + y^2 = a^2\cos^2 t + a^2\sin^2 t = a^2$. Therefore the trajectory is on a circle of radius a centered at O.



We will call $x(t) = a\cos t$, $y(t) = a\sin t$ the *parametric form* of the curve and $x^2 + y^2 = a^2$ the *symmetric form*.

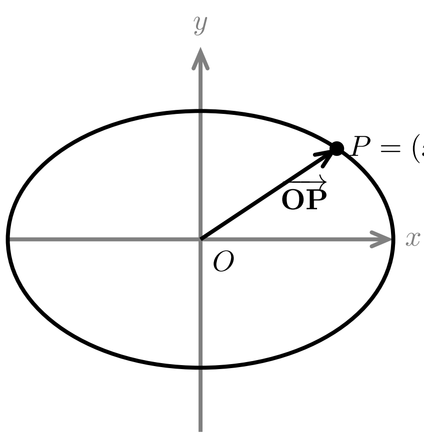
Note, a different parametrization, say

$$x(t) = a\cos(3t), \quad y(t) = a\sin(3t)$$

results in the same path, i.e. the circle $x^2 + y^2 = a^2$, but the two trajectories differ by how fast they travel around the circle.

The circle is easily changed to an ellipse by

$$\begin{aligned} &\text{parametric form:} \quad x(t) = a\cos t, \quad y(t) = b\cos t \\ &\text{symmetric form:} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \end{aligned}$$



Lines

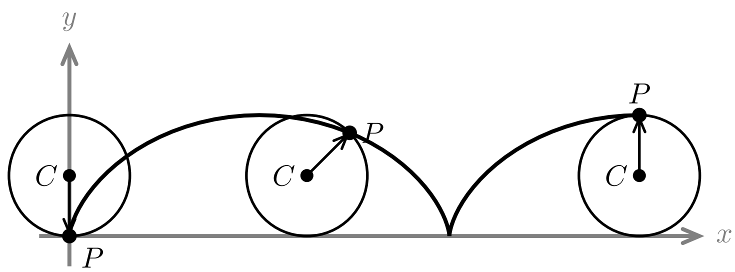
We review parametric equations of lines by writing the the equation of a general line in the plane. We know we can parametrize the line through (x_0, y_0) parallel to $\langle b_1, b_2\rangle$ by

$$x(t) = x_0 + tb_1, \quad y(t) = y_0 + tb_2 \;\Leftrightarrow\; \mathbf{r}(t) = \langle x, y\rangle = \langle x_0 + tb_1, y_0 + tb_2\rangle = \langle x_0, y_0\rangle + t\langle b_1, b_2\rangle.$$

The cycloid

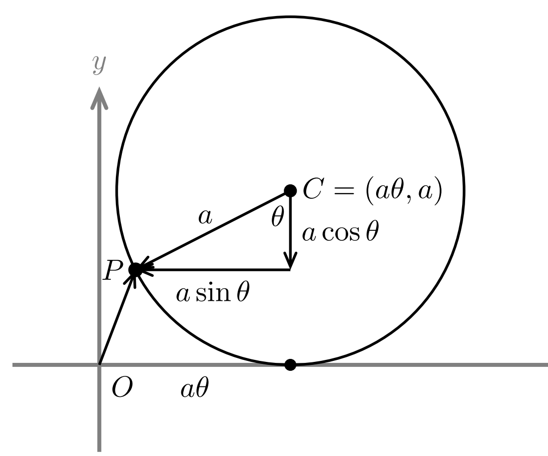
The cycloid has a long and storied history and comes up surprisingly often in physical problems. For us it is a curve that has no simple symmetric form, so we will only work with it in its parametric form.

The cycloid is the trajectory of a point on a circle that is rolling without slipping along the x-axis. To be specific, we'll follow the point P that starts at the origin.



The natural parameter to use is the angle θ that the wheel has turned. We'll use vector methods to find the position vector for P as a function of θ.

Our strategy is to break the motion up into translation of the center and rotation about the center. The figure shows the wheel after it has turned through a small θ. We see the position vector



$$\overrightarrow{\mathbf{OP}} = \overrightarrow{\mathbf{OC}} + \overrightarrow{\mathbf{CP}}.$$

We'll compute each piece separately.

After turning θ radians the wheel has rolled a distance $a\theta$, so the center of the circle is at $(a\theta, a)$, i.e.,

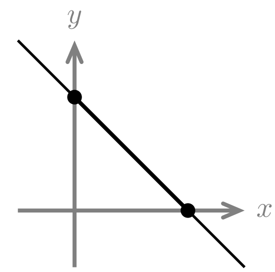
$$\overrightarrow{\mathbf{OC}} = \langle a\theta, a\rangle.$$

The figure also shows that

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

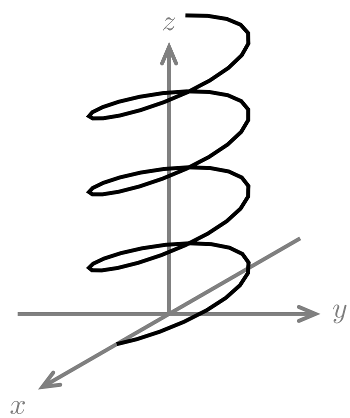
Putting the pieces together we get parametric equations for the cycloid

$$\overrightarrow{\mathbf{OP}} = \langle a\theta - a\sin\theta, a - a\cos\theta\rangle$$
$$\Leftrightarrow \quad x(\theta) = a\theta - a\sin\theta, \quad y(\theta) = a - a\cos\theta.$$



Example 2: (Where the symmetric form loses information.)
Find the symmetric form for $x = 3\cos^2 t,\ y = 3\sin^2 t$.
Easily we get: $x + y = 3$, with x, y non-negative.
The symmetric form shows a line, but the parametric trajectory only traces out a part of the line. In fact, it goes back an forth over the part of the line in the first quadrant.

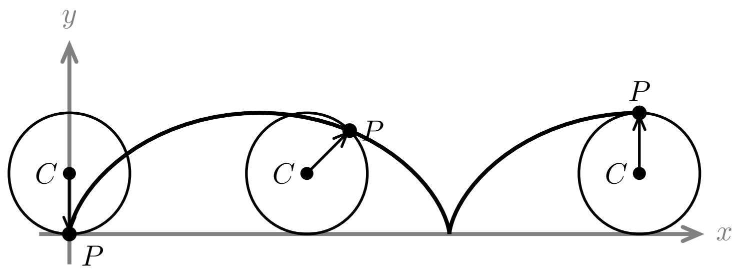
Example 3: The curve $\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + at\,\mathbf{k}$ is a helix winding around the z-axis.



Cusp on the cycloid

The graph of the cycloid has point where the graph touches the x-axis. These points are usually called *cusps*.

What you saw in the previous video was an analysis of the behavior of the trajectory near the cusps. We will go through that analysis again and discuss what's happening physically on a rolling wheel.



In order to simplify the way our equations look, let's take the radius of the wheel to be $a = 1$. Then the parametric equations for the cycloid are

$$x(\theta) = \theta - \sin\theta, \quad y(\theta) = 1 - \cos\theta.$$

Taking derivatives we get $\dfrac{dx}{d\theta} = 1 - \cos\theta$ and $\dfrac{dy}{d\theta} = \sin\theta$.
Thus the slope of the curve is

$$\frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta} = \frac{\sin\theta}{1 - \cos\theta}.$$

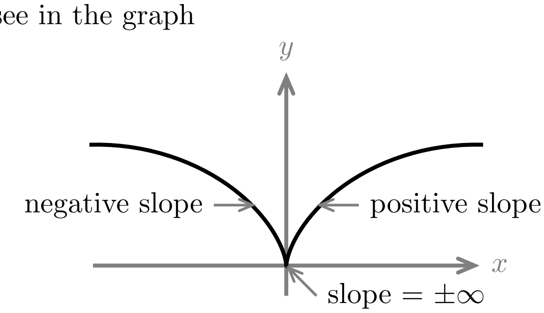
As $\theta \to 0$ this is of indeterminant form $0/0$. Using L'Hospital's rule we get

$$\lim_{\theta\to 0} \frac{\sin\theta}{1-\cos\theta} = \lim_{\theta\to 0} \frac{\cos\theta}{\sin\theta}$$

Since $\cos\theta$ goes to 1 and $\sin\theta$ goes to 0 this limit does not exist.

But looking at it more carefully we see that as $\theta \to 0^-$ the limit goes to $-\infty$ and as $\theta \to 0^+$ it goes to $+\infty$. That is, right at the cusp the slope of the curve is $-\infty$ to the left and $+\infty$ to the right.

This mirrors what we see in the graph



Later, when we learn about velocity we'll see that, at $\theta = 0$, $\dfrac{dx}{d\theta} = 0$ and $\dfrac{dy}{d\theta} = 0$ means the velocity is 0. At the cusp, the point changes abruptly from moving down to moving up. Physically this can only happen if the velocity is 0 at the changeover point.

Velocity and acceleration

Now we will see one of the benefits of using the position vector. Let's assume we have a moving point with position vector

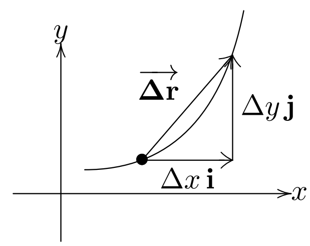
$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}.$$

(We assume the point moves in the plane. The extension to a point moving in space is trivial.)

Velocity

Over a short time Δt the position changes by $\Delta\mathbf{r}$. The average velocity over this time is simply

$$\frac{\Delta\mathbf{r}}{\Delta t}, \text{ i.e., displacement/time.}$$



The figure shows $\Delta\mathbf{r} = \Delta x\,\mathbf{i} + \Delta y\,\mathbf{j}$. Dividing by Δt we get

$$\text{average velocity } = \frac{\Delta\mathbf{r}}{\Delta t} = \frac{\Delta x}{\Delta t}\,\mathbf{i} + \frac{\Delta y}{\Delta t}\,\mathbf{j}$$

Now, as we usually do in calculus, we let $\Delta t \to 0$. The average velocity becomes the (instantaneous) velocity and the ratios in the formula above become derivatives. For completeness we write the velocity vector in a number of different forms

$$\text{velocity } = \frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\,\mathbf{i} + \frac{dy}{dt}\,\mathbf{j} = x'\mathbf{i} + y'\mathbf{j} = \langle x', y'\rangle.$$

Tangent vector: (same thing as velocity)

In the picture above, we see that as Δt shrinks to 0 the vector $\frac{\Delta\mathbf{r}}{\Delta t}$ becomes tangent to the curve. When the parameter is time we can rightfully refer to $\mathbf{r}'(t)$ as the *velocity.* In general, we will abuse the language and refer to the derivative of position with respect to any parameter as velocity. If we are thinking geometrically or want to be precise, we will call the derivative by its geometric name: the *tangent vector.*

As always, we encourage you to remember the geometric view of the velocity vector. Knowing it is tangent to the curve will be important as we develop the subject and solve problems.

Acceleration

There is no reason to stop taking derivatives after one. Since acceleration is change in velocity per unit time, we get

$$\text{acceleration } = \mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = x''(t)\,\mathbf{i} + y''(t)\,\mathbf{j} = \langle x'', y''\rangle.$$

Example: A rocket follows a trajectory

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} = v_{0,x}\,t\,\mathbf{i} + (-\frac{g}{2}t^2 + v_{0,y}t)\mathbf{j}.$$

Find its velocity and acceleration vectors.

<u>**Answer:**</u>

$$\text{velocity } = \mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = v_{0,x}\mathbf{i} + (-gt + v_{0,y})\mathbf{j},$$

$$\text{acceleration } = \mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = -g\mathbf{j}.$$

Example: Find the velocity and acceleration vectors for the cycloid

$$x = \theta - \sin\theta, \quad y = 1 - \cos\theta.$$

<u>**Answer:**</u> As noted, this a slight abuse of language,

$$\text{velocity } = \text{tangent vector } = \mathbf{v}(\theta) = \frac{d\mathbf{r}}{d\theta} = \langle x'(\theta), y'(\theta)\rangle = \langle 1 - \cos\theta, \sin\theta\rangle.$$

$$\text{acceleration } = \mathbf{a}(\theta) = \frac{d\mathbf{v}}{d\theta} = \langle \sin\theta, \cos\theta\rangle.$$

Example: In the cycloid above, suppose the wheel rolls at 3 revolutions per second. Write the parametric equations in terms of time, and compute the velocity.

<u>**Answer:**</u> Since 3 rev/second = 6π radians/sec, we have $\theta = 6\pi t$. Therefore,

$$x(t) = 6\pi t - \sin(6\pi t), \quad y(t) = 1 - \cos(6\pi t).$$

$$\mathbf{v}(t) = \langle 6\pi - 6\pi\cos(6\pi t), 6\pi\sin(6\pi t)\rangle.$$

Velocity, speed and arc length

Speed

Velocity, being a vector, has a magnitude and a direction. The direction is tangent to the curve traced out by $\mathbf{r}(t)$. The magnitude of its velocity is the speed.

$$\text{speed } = |\mathbf{v}| = \left|\frac{d\mathbf{r}}{dt}\right|.$$

Speed is in units of distance per unit time. It reflects how fast our moving point is moving.

Example: A point goes one time around a circle of radius 1 unit in 3 seconds. What is its average velocity and average speed.

Answer: The distance the point traveled equals the circumference of the circle, 2π. Its net displacement is $\mathbf{0}$, since it ends where it started. Thus, its average speed = distance/time = $2\pi/3$ and its average velocity = displacement/time = $\mathbf{0}$.

If you look carefully, we've used a boldface $\mathbf{0}$ because velocity is a vector.

Our usual symbol for distance traveled is s. For a point moving along a curve the distance traveled is the length of the curve. Because of this we also refer to s as *arc length.*

Notation and nomenclature summary:

Since we will use a variety of notations, we'll collect them here. The unit tangent vector will be explained below. As you should expect, we will also be able to view everything from a geometric perspective.

$\mathbf{r}(t) =$ position.

In the plane $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} = \langle x, y\rangle$

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

$\dfrac{d\mathbf{r}}{dt} = \mathbf{v}(t) =$ velocity = tangent vector.

In the plane $\mathbf{v} = x'(t)\mathbf{i} + y'(t)\mathbf{j} = \langle x', y'\rangle$

In space $\mathbf{v} = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k} = \langle x', y', z'\rangle$.

$\mathbf{T} = \dfrac{\mathbf{v}}{|\mathbf{v}|} =$ unit tangent vector.

$s =$ arclength, speed $= \dfrac{ds}{dt} = |\mathbf{v}|$.

In the plane $\dfrac{ds}{dt} = \sqrt{(x')^2 + (y')^2}$.

In space $\dfrac{ds}{dt} = \sqrt{(x')^2 + (y')^2 + (z')^2}$.

$\mathbf{v} = \dfrac{ds}{dt}\mathbf{T}, \quad \mathbf{T} = \dfrac{\mathbf{v}}{ds/dt}$

$\mathbf{a}(t) = \dfrac{d\mathbf{v}}{dt} = \dfrac{d^2\mathbf{r}}{dt^2} =$ acceleration.

In the plane $\mathbf{a}(t) = x''(t)\mathbf{i} + y''(t)\mathbf{j} = \langle x'', y''\rangle$

In space $\mathbf{a} = x''(t)\mathbf{i} + y''(t)\mathbf{j} + z''(t)\mathbf{k} = \langle x'', y'', z''\rangle$.

Unit tangent vector

As its name implies, the *unit tangent vector* is a unit vector in the same direction as the tangent vector. We usually denote it $\mathbf{T}$. We compute it by dividing the tangent vector by its length. Here are several ways of writing this.

$$\mathbf{T} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{d\mathbf{r}/dt}{ds/dt} = \frac{\mathbf{v}}{ds/dt}.$$

Multiply $\mathbf{T}$ by ds/dt gives the formula

$$\mathbf{v} = \frac{ds}{dt}\mathbf{T},$$

which expresses velocity as a magnitute, ds/dt and a direction $\mathbf{T}$.

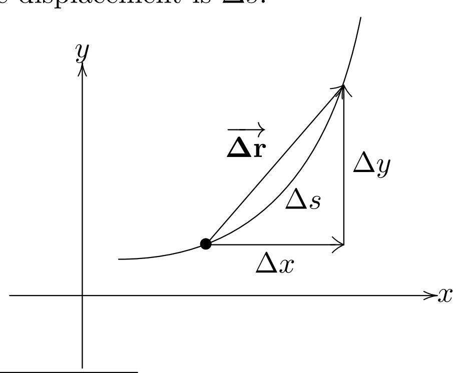
Geometric considerations

Here we'll offer a mathematical justification for our statement that

$$\text{speed } = \frac{ds}{dt} = |\mathbf{v}|.$$

We'll work in two dimensions. The extension to 3D is straightforward.

The figure below shows a curve, and a small displacement $\Delta\mathbf{r}$. The length along the curve from the start to end of the displacement is Δs.



We see $\Delta s \approx |\Delta\mathbf{r}| = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. Dividing by Δt gives

$$\frac{\Delta s}{\Delta t} \approx \left|\frac{\Delta \mathbf{r}}{\Delta t}\right| = \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2}$$

Taking the limit as $\Delta t \to 0$ gives

$$\frac{ds}{dt} = \left|\frac{d\mathbf{r}}{dt}\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}.$$

Kepler's Second Law

By studying the Danish astronomer Tycho Brahe's data about the motion of the planets, Kepler formulated three empirical laws; two of them can be stated as follows:

Second Law A planet moves in a plane, and the radius vector (from the sun to the planet) sweeps out equal areas in equal times.

First Law The planet's orbit in that plane is an ellipse, with the sun at one focus.

From these laws, Newton deduced that the force keeping the planets in their orbits had magnitude $1/d^2$, where d was the distance to the sun; moreover, it was directed toward the sun, or as was said, *central*, since the sun was placed at the origin.

Using a little vector analysis (without coordinates), this section is devoted to showing that *the Second Law is equivalent to the force being central.*

It is harder to show that an elliptical orbit implies the magnitude of the force is of the form K/d^2, and vice-versa; this uses vector analysis in polar coordinates and requires the solution of non-linear differential equations.

1. Differentiation of products of vectors

Let $\mathbf{r}(t)$ and $\mathbf{s}(t)$ be two differentiable vector functions in 2- or 3-space. Then

$$\text{(1)} \qquad \frac{d}{dt}(\mathbf{r}\cdot\mathbf{s}) \;=\; \frac{d\mathbf{r}}{dt}\cdot\mathbf{s} \;+\; \mathbf{r}\cdot\frac{d\mathbf{s}}{dt}; \qquad \frac{d}{dt}(\mathbf{r}\times\mathbf{s}) \;=\; \frac{d\mathbf{r}}{dt}\times\mathbf{s} \;+\; \mathbf{r}\times\frac{d\mathbf{s}}{dt}.$$

These rules are just like the product rule for differentiation. Be careful in the second rule to get the multiplication order correct on the right, since $\mathbf{a}\times\mathbf{b} \neq \mathbf{b}\times\mathbf{a}$ in general. The two rules can be proved by writing everything out in terms of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ components and differentiating. They can also be proved directly from the definition of derivative, without resorting to components, as follows:

Let t increase by Δt. Then $\mathbf{r}$ increases by $\Delta\mathbf{r}$, and $\mathbf{s}$ by $\Delta\mathbf{s}$, and the corresponding change in $\mathbf{r}\cdot\mathbf{s}$ is given by

$$\Delta(\mathbf{r}\cdot\mathbf{s}) \;=\; (\mathbf{r}+\Delta\mathbf{r})\cdot(\mathbf{s}+\Delta\mathbf{s}) \;-\; \mathbf{r}\cdot\mathbf{s}\;,$$

so if we expand the right side out and divide all terms by Δt, we get

$$\frac{\Delta(\mathbf{r}\cdot\mathbf{s})}{\Delta t} \;=\; \frac{\Delta\mathbf{r}}{\Delta t}\cdot\mathbf{s} \;+\; \mathbf{r}\cdot\frac{\Delta\mathbf{s}}{\Delta t} \;+\; \frac{\Delta\mathbf{r}}{\Delta t}\cdot\Delta\mathbf{s}\;.$$

Now let $\Delta t \to 0$; then $\Delta\mathbf{s} \to 0$ since $\mathbf{s}(t)$ is continuous, and we get the first equation in (1). The second equation in (1) is proved the same way, replacing $\cdot$ by $\times$ everywhere.

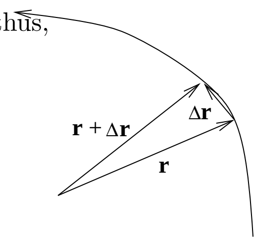
2. Kepler's second law and the central force. To show that the force being central (i.e., directed toward the sun) is equivalent to Kepler's second law, we need to translate that law into calculus. "Sweeps out equal areas in equal times" means:

the radius vector sweeps out area at a constant rate .

The first thing therefore is to obtain a mathematical expression for this rate. Referring to the picture, we see that as the time increases from t to $t+\Delta t$, the corresponding change in the area A is given approximately by

$$\Delta A \;\approx\; \text{area of the triangle} \;=\; \frac{1}{2}\,|\mathbf{r}\times\Delta\mathbf{r}|\ ,$$

since the triangle has half the area of the parallelogram formed by $\mathbf{r}$ and $\Delta\mathbf{r}$; thus,



$$2\,\frac{\Delta A}{\Delta t} \;\approx\; \left|\mathbf{r}\times\frac{\Delta\mathbf{r}}{\Delta t}\right|,$$

and as $\Delta t \to 0$, this becomes

$$(2) \qquad 2\,\frac{dA}{dt} \;=\; \left|\mathbf{r}\times\frac{d\mathbf{r}}{dt}\right| \;=\; |\mathbf{r}\times\mathbf{v}|. \qquad \text{where } \mathbf{v}=\frac{d\mathbf{r}}{dt}\ .$$

Using (2), we can interpret Kepler's second law mathematically. Since the area is swept out at a constant rate, dA/dt is constant, so according to (2),

$$(3) \qquad |\mathbf{r}\times\mathbf{v}| \quad \textit{is a constant.}$$

Moreover, since Kepler's law says $\mathbf{r}$ lies in a plane, the velocity vector $\mathbf{v}$ also lies in the same plane, and therefore

$$(4) \qquad \mathbf{r}\times\mathbf{v} \quad \textit{has constant direction (perpendicular to the plane of motion).}$$

Since the direction and magnitude of $\mathbf{r}\times\mathbf{v}$ are both constant,

$$(5) \qquad \mathbf{r}\times\mathbf{v} \;=\; \mathbf{K},\ \text{a constant vector,}$$

and from this we see that

$$(6) \qquad \frac{d}{dt}\,(\mathbf{r}\times\mathbf{v}) \;=\; \mathbf{0}\ .$$

But according to the rule (1) for differentiating a vector product,

$$(7) \qquad \begin{aligned} \frac{d}{dt}\,(\mathbf{r}\times\mathbf{v}) \;&=\; \mathbf{v}\times\mathbf{v}+\mathbf{r}\times\mathbf{a}, &\qquad& \text{where } \mathbf{a}=\frac{d\mathbf{v}}{dt},\\ &=\; \mathbf{r}\times\mathbf{a}, && \text{since } \mathbf{s}\times\mathbf{s}=\mathbf{0} \text{ for any vector } \mathbf{s}. \end{aligned}$$

Now (6) and (7) together imply

$$(8) \qquad \mathbf{r}\times\mathbf{a} \;=\; \mathbf{0},$$

which shows that *the acceleration vector* $\mathbf{a}$ *is parallel to* $\mathbf{r}$, but in the opposite direction, since the planets do go around the sun, not shoot off to infinity.

Thus $\mathbf{a}$ is directed toward the center (i.e., the sun), and since $\mathbf{F}=m\mathbf{a}$, the force $\mathbf{F}$ is also directed toward the sun. (Note that "center" does not mean the center of the elliptical orbits, but the mathematical origin, i.e., the tail of the radius vector $\mathbf{r}$, which we are taking to be the sun's position.)

The reasoning is reversible, so for motion under any type of central force, the path of motion will lie in a plane and area will be swept out by the radius vector at a constant rate.

Functions of two variables

Examples: Functions of several variables
$f(x, y) = x^2 + y^2 \;\Rightarrow\; f(1, 2) = 5$ etc.
$f(x, y) = xy^2 e^{x+y}$
$f(x, y, z) = xy \log z$
Ideal gas law: $P = kT/V$.

Dependent and independent variables
In $z = f(x, y)$ we say x, y are independent variables and z is a dependent variable. This indicates that x and y are free to take any values and then z depends on these values. For now it will be clear which are which, later we'll have to take more care.

Graphs
For the function $y = f(x)$: there is one independent variable and one dependent variable, which means we need 2 dimensions for its graph.

Graphing technique:

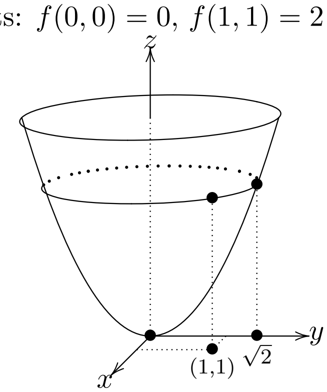
go to x then compute $y = f(x)$ then go up to height y.

For $z = f(x, y)$ we have two independent and one dependent variable, so we need 3 dimensions to graph the function. The technique is the same as before.

Example: Consider $z = f(x, y) = x^2 + y^2$.
To make the graph:

go to (x, y) then compute $z = f(x, y)$ then go up to height z.

We show the plot of three points: $f(0, 0) = 0$, $f(1, 1) = 2$ and $f(0, \sqrt{2}) = 2$.



The figure above shows more than just the graph of three points. Here are the steps we used to draw the graph. Remember, this is just a sketch, it should suggest the shape of the graph and some of its features.

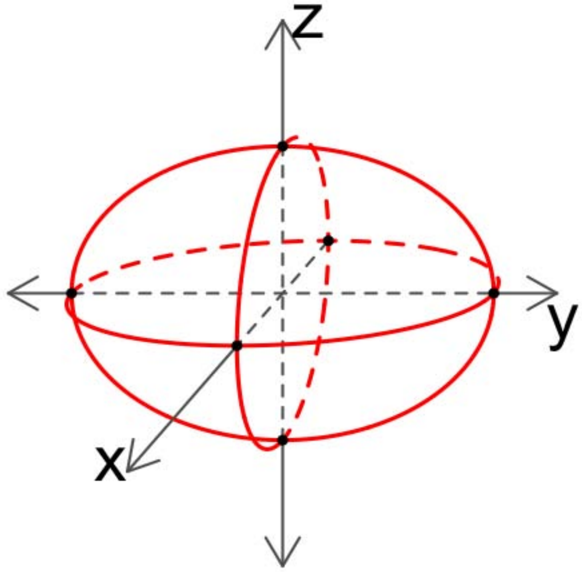
1. First we draw the axes. The z-axis points up, the y-axis is to the right and the x-axis comes out of the page, so it is drawn at the angle shown. This gives a perspective with the eye somewhere in the first octant.

2. The yz-traces are those curves found by setting $x =$ a constant. We start with the trace when $x = 0$. This is an upward pointing parabola in the yz-plane.

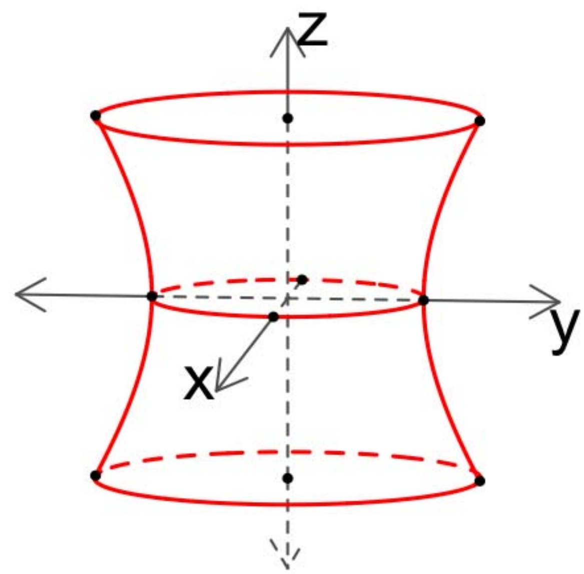
3. Next we sketch the trace with $z = 3$. This is a circle of radius $\sqrt{3}$ at height $z = 3$. Note, the traces where $z =$ constant are generally called *level curves*.

This is enough for this graph. Other graphs take other traces. You should expect to do a certain amount of trial and error before your figure looks right.

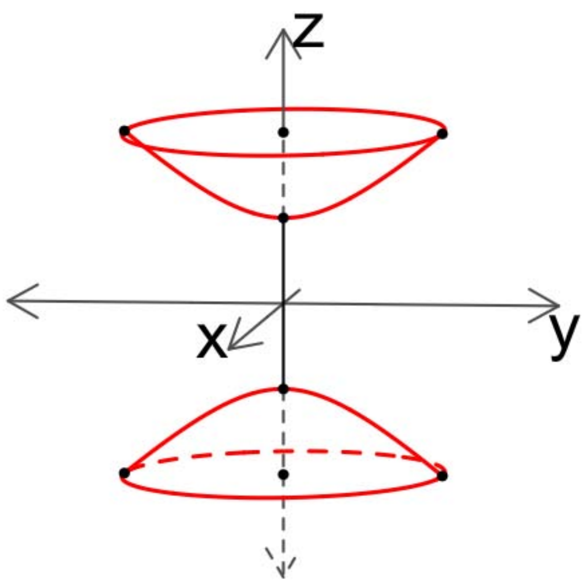
Gallery of graphs



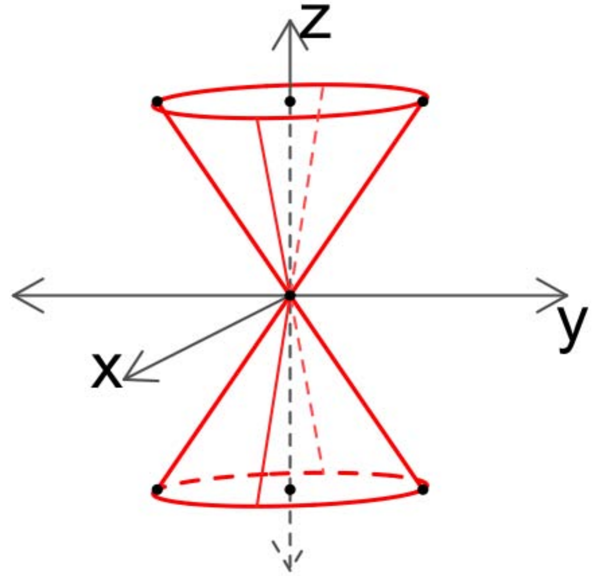
Ellipsoid: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$



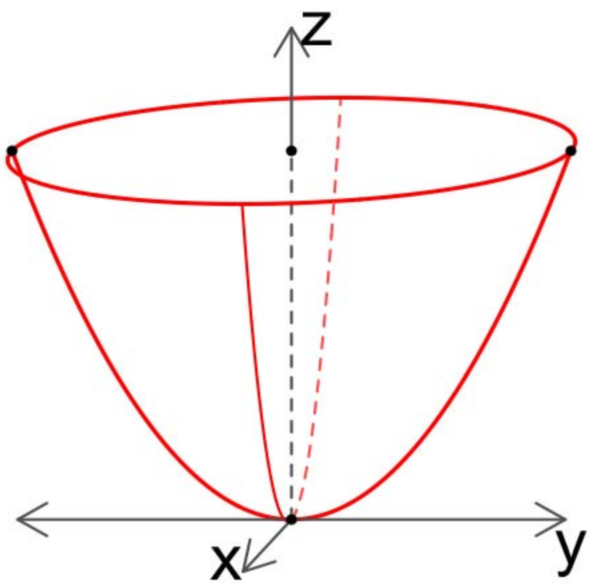
Hyperboloid of one sheet:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$



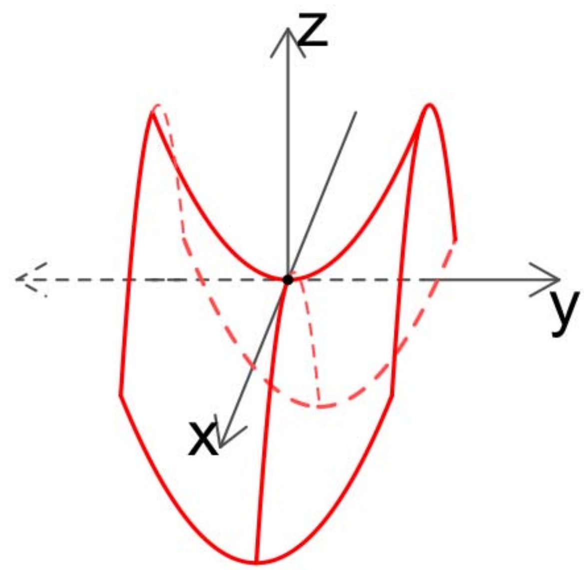
Hyperboloid of two sheets:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$



Elliptic cone: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = \dfrac{z^2}{c^2}$



Elliptic paraboloid: $z = ax^2 + by^2$



Hyperbolic paraboloid: $z = by^2 - ax^2$

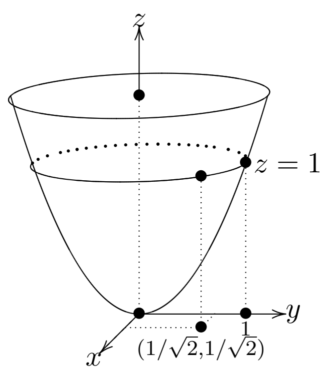
Level Curves and Contour Plots

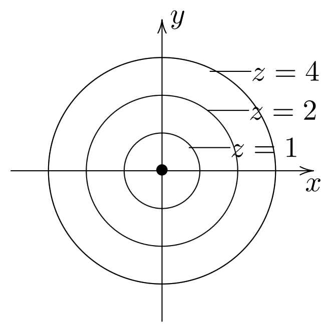
Level curves and *contour plots* are another way of visualizing functions of two variables. If you have seen a topographic map then you have seen a contour plot.

Example: To illustrate this we first draw the graph of $z = x^2 + y^2$. On this graph we draw *contours*, which are curves at a fixed height z = constant.

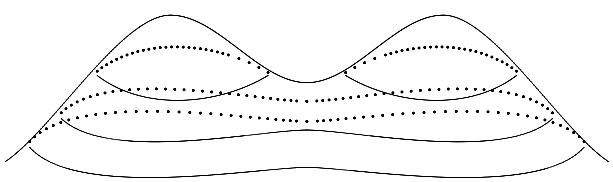
For example the curve at height $z = 1$ is the circle $x^2 + y^2 = 1$. On the graph we have to draw this at the correct height. Another way to show this is to draw the curves in the xy-plane and label them with their z-value. We call these curves *level curves* and the entire plot is called a *contour plot*.

For this example they are shown in the plot on the right. Notice that the 3D graph is simply the level curves 'pulled out' each to its correct height.

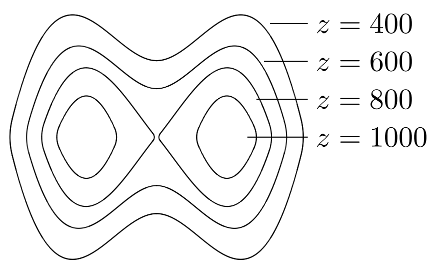




Here is another plot of a 'mountain pass'. Notice that in the contour plot the mountain pass is represented by a level curve that crosses itself. Moving up or down from the cross level curves heights decrease and moving right or left in the other they increase.



Mountain pass

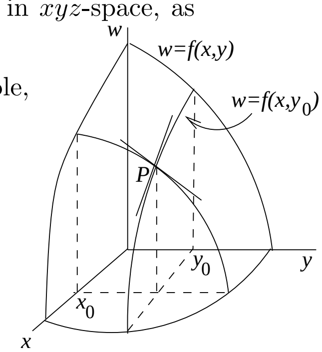


Level curves

Partial derivatives

Partial derivatives

Let $w = f(x, y)$ be a function of two variables. Its graph is a surface in xyz-space, as pictured.



Fix a value $y = y_0$ and just let x vary. You get a function of *one* variable,

(1) $$w = f(x, y_0), \qquad \text{the } \textbf{partial function} \text{ for } y = y_0.$$

Its graph is a curve in the vertical plane $y = y_0$, whose slope at the point P where $x = x_0$ is given by the derivative

(2) $$\frac{d}{dx} f(x, y_0)\bigg|_{x_0}, \quad \text{or} \quad \frac{\partial f}{\partial x}\bigg|_{(x_0,y_0)}.$$

We call (2) the **partial derivative** of f with respect to x at the point (x_0, y_0); the right side of (2) is the standard notation for it. The partial derivative is just the ordinary derivative of the partial function — it is calculated by holding one variable fixed and differentiating with respect to the other variable. Other notations for this partial derivative are

$$f_x(x_0, y_0), \qquad \frac{\partial w}{\partial x}\bigg|_{(x_0,y_0)}, \qquad \left(\frac{\partial f}{\partial x}\right)_0, \qquad \left(\frac{\partial w}{\partial x}\right)_0;$$

the first is convenient for including the specific point; the second is common in science and engineering, where you are just dealing with relations between variables and don't mention the function explicitly; the third and fourth indicate the point by just using a single subscript.

Analogously, fixing $x = x_0$ and letting y vary, we get the partial function $w = f(x_0, y)$, whose graph lies in the vertical plane $x = x_0$, and whose slope at P is the *partial derivative of f with respect to y*; the notations are

$$\frac{\partial f}{\partial y}\bigg|_{(x_0,y_0)}, \qquad f_y(x_0, y_0), \qquad \frac{\partial w}{\partial y}\bigg|_{(x_0,y_0)}, \qquad \left(\frac{\partial f}{\partial y}\right)_0, \qquad \left(\frac{\partial w}{\partial y}\right)_0.$$

The partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ depend on (x_0, y_0) and are therefore functions of x and y.

Written as $\partial w/\partial x$, the partial derivative gives the rate of change of w with respect to x alone, at the point (x_0, y_0): it tells how fast w is increasing as x increases, when y is held constant.

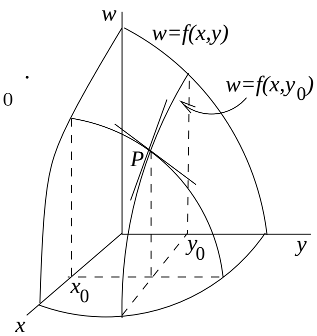
For a function of three or more variables, $w = f(x, y, z, \dots)$, we cannot draw graphs any more, but the idea behind partial differentiation remains the same: to define the partial derivative with respect to x, for instance, hold all the other variables constant and take the ordinary derivative with respect to x; the notations are the same as above:

$$\frac{d}{dx} f(x, y_0, z_0, \dots) = f_x(x_0, y_0, z_0, \dots), \qquad \left(\frac{\partial f}{\partial x}\right)_0, \qquad \left(\frac{\partial w}{\partial x}\right)_0.$$

The Tangent Approximation

1. The tangent plane.

For a function of one variable, $w = f(x)$, the tangent line to its graph at a point (x_0, w_0) is the line passing through (x_0, w_0) and having slope $\left(\frac{dw}{dx}\right)_0$.



For a function of two variables, $w = f(x, y)$, the natural analogue is the *tangent plane* to the graph, at a point (x_0, y_0, w_0). What's the equation of this tangent plane? Referring to the picture at right (this figure was also used when we introduced partial derivatives), we see that the tangent plane

(i) must pass through (x_0, y_0, w_0), where $w_0 = f(x_0, y_0)$;

(ii) must contain the tangent lines to the graphs of the two partial functions — this will hold if the plane has the same slopes in the $\mathbf{i}$ and $\mathbf{j}$ directions as the surface does.

Using these two conditions, it is easy to find the equation of the tangent plane. The general equation of a plane through (x_0, y_0, w_0) is

$$A(x - x_0) + B(y - y_0) + C(w - w_0) = 0 .$$

Assume the plane is not vertical; then $C \neq 0$, so we can divide through by C and solve for $w - w_0$, getting

$$(3) \qquad w - w_0 = a(x - x_0) + b(y - y_0), \qquad a = A/C,\ b = B/C.$$

The plane passes through (x_0, y_0, w_0); what values of the coefficients a and b will make it also tangent to the graph there? We have

$$\begin{aligned} a &= \text{slope of plane (3) in the } \mathbf{i}\text{-direction} && \text{(by putting } y = y_0 \text{ in (3));} \\ &= \text{slope of graph in the } \mathbf{i}\text{-direction,} && \text{(by (ii) above)} \\ &= \left(\frac{\partial w}{\partial x}\right)_0; && \text{(by the definition of partial derivative);} \quad \text{similarly,} \\ b &= \left(\frac{\partial w}{\partial y}\right)_0 . \end{aligned}$$

Therefore the equation of the **tangent plane** to $w = f(x, y)$ at (x_0, y_0) is

$$(4) \qquad w - w_0 = \left(\frac{\partial w}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial w}{\partial y}\right)_0 (y - y_0)$$

2. The approximation formula.

The most important use for the tangent plane is to give an approximation that is the basic formula in the study of functions of several variables — almost everything follows in one way or another from it.

The intuitive idea is that if we stay near (x_0, y_0, w_0), the graph of the tangent plane (4) will be a good approximation to the graph of the function $w = f(x, y)$. Therefore if the point (x, y) is close to (x_0, y_0),

$$(5) \qquad \begin{array}{ccc} f(x,y) & \approx & w_0 + \left(\dfrac{\partial w}{\partial x}\right)_0 (x - x_0) + \left(\dfrac{\partial w}{\partial y}\right)_0 (y - y_0) \\ \text{height of graph} & \approx & \text{height of tangent plane} \end{array}$$

The function on the right side of (5) whose graph is the tangent plane is often called the **linearization** of $f(x, y)$ at (x_0, y_0): it is the linear function which gives the best approximation to $f(x, y)$ for values of (x, y) close to (x_0, y_0).

An equivalent form of the approximation (5) is obtained by using Δ notation; if we put

$$\Delta x = x - x_0, \qquad \Delta y = y - y_0, \qquad \Delta w = w - w_0,$$

then (5) becomes

$$(6) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y, \qquad \text{if} \quad \Delta x \approx 0,\ \Delta y \approx 0\ .$$

This formula gives the approximate change in w when we make a small change in x and y. We will use it often.

The analogous approximation formula for a function $w = f(x, y, z)$ of three variables would be

$$(7) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y + \left(\frac{\partial w}{\partial z}\right)_0 \Delta z, \qquad \text{if} \quad \Delta x, \Delta y \Delta z \approx 0\ .$$

Unfortunately, for functions of three or more variables, we can't use a geometric argument for the approximation formula (7); for this reason, it's best to recast the argument for (6) in a form which doesn't use tangent planes and geometry, and therefore can be generalized to several variables. This is done at the end of this Chapter TA; for now let's just assume the truth of (7) and its higher-dimensional analogues.

Here are two typical examples of the use of the approximation formula. Other examples are in the Exercises. In the rest of your study of partial differentiation, you will see how the approximation formula is used to derive the important theorems and formulas.

Example 1. Give a reasonable square, centered at $(1, 1)$, over which the value of $w = x^3 y^4$ will not vary by more than $\pm$ .1 .

Solution. We use (6). We calculate for the two partial derivatives

$$w_x = 3x^2 y^4 \qquad\qquad w_y = 4x^3 y^3$$

and therefore, evaluating the partials at $(1, 1)$ and using (6), we get

$$\Delta w \approx 3\Delta x + 4\Delta y.$$

Thus if $|\Delta x| \leq .01$ and $\Delta y| \leq .01$, we should have

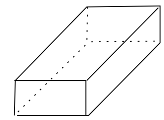
$$|\Delta w| \leq 3|\Delta x| + 4|\Delta y| \leq .07,$$

which is within the bounds. So the answer is the square with center at (1,1) given by

$$|x - 1| \leq .01, \qquad |y - 1| \leq .01\ .$$

Example 2. The sides a, b, c of a rectangular box have lengths measured to be respectively 1, 2, and 3. To which of these measurements is the volume V most sensitive?

Solution. $V = abc$, and therefore by the approximation formula (7),



$$\begin{aligned} \Delta V &\approx bc\,\Delta a + ac\,\Delta b + ab\,\Delta c \\ &\approx 6\,\Delta a + 3\,\Delta b + 2\,\Delta c, \qquad \text{at } (1,2,3); \end{aligned}$$

thus it is most sensitive to small changes in side a, since Δa occurs with the largest coefficient. (That is, if one at a time the measurement of each side were changed by say .01, it is the change in a which would produce the biggest change in V, namely .06 .)

The result may seem paradoxical — the value of V is most sensitive to the length of the *shortest* side — but it's actually intuitive, as you can see by thinking about how the box looks.

Sensitivity Principle *The numerical value of $w = f(x, y, \dots)$, calculated at some point $(x_0, y_0, \dots)$, will be most sensitive to small changes in that variable for which the corresponding partial derivative $w_x, w_y, \dots$ has the largest absolute value at the point.*

The Tangent approximation

4. Critique of the approximation formula.

First of all, the approximation formula for functions of two or three variables

$$(6) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y, \qquad \text{if} \quad \Delta x \approx 0,\ \Delta y \approx 0\ .$$

$$(7) \qquad \Delta w \approx \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y + \left(\frac{\partial w}{\partial z}\right)_0 \Delta z, \qquad \text{if} \quad \Delta x, \Delta y \Delta z \approx 0\ .$$

is not a precise mathematical statement, since the symbol $\approx$ does not specify exactly how close the quantitites on either side of the formula are to each other. To fix this up, one would have to specify the error in the approximation. (This can be done, but it is not often used.)

A more fundamental objection is that our discussion of approximations was based on the assumption that the tangent plane is a good approximation to the surface at (x_0, y_0, w_0). Is this really so?

Look at it this way. The tangent plane was determined as the plane which has the same slope as the surface in the $\mathbf{i}$ and $\mathbf{j}$ directions. This means the approximation (6) will be good if you move away from (x_0, y_0) in the $\mathbf{i}$ direction (by taking $\Delta y = 0$), or in the $\mathbf{j}$ direction (putting $\Delta x = 0$). But does the tangent plane have the same slope as the surface in all the other directions as well?

Intuitively, we should expect that this will be so if the graph of $f(x, y)$ is a "smooth" surface at (x_0, y_0) — it doesn't have any sharp points, folds, or look peculiar. Here is the mathematical hypothesis which guarantees this.

Smoothness hypothesis. We say $f(x, y)$ is **smooth** at (x_0, y_0) if

(8) $\qquad$ *f_x and f_y are continuous in some rectangle centered at (x_0, y_0).*

If (8) holds, the approximation formula (6) will be valid.

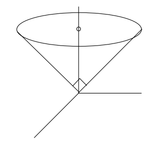
Though pathological examples can be constructed, in general the normal way a function fails to be smooth (and in turn (6) fails to hold) is that one or both partial derivatives fail to exist at (x_0, y_0). This means of course that you won't even be able to write the formula (6), unless you're sleepy. Here is a simple example.

Example 3. Where is $w = \sqrt{x^2 + y^2}$ smooth? Discuss.

Solution. Calculating formally, we get

$$\frac{\partial w}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}\ , \qquad \frac{\partial w}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}.$$

These are continuous at all points except $(0, 0)$, where they are undefined. So the function is smooth except at the origin; the approximation formula (6) should be valid everywhere except at the origin.



Indeed, investigating the graph of this function, since $w = \sqrt{x^2+y^2}$ says that

$$\text{height of graph over } (x,y) \ = \ \text{distance of } (x,y) \text{ from } w\text{-axis},$$

the graph is a right circular cone, with vertex at $(0,0)$, axis along the w-axis, and vertex angle a right angle. Geometrically the graph has a sharp point at the origin, so there should be no tangent plane there, and no valid approximation formula (6) — there is no linear function which approximates a cone at its vertex.

A non-geometrical argument for the approximation formula

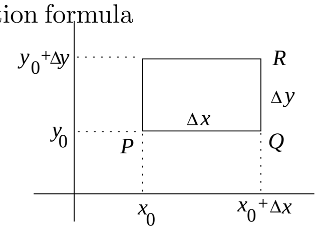
We promised earlier a non-geometrical approach to the approximation formula (6) that would generalize to higher-dimensions, in particular to the 3-variable formula (7). This approach will also show why the hypothesis (8) of *smoothness* is needed. The argument is still imprecise, since it uses the symbol $\approx$, but it can be refined to a proof (which you will find in your book, though it's not easy reading).

It uses the one-variable approximation formula for a differentiable function $w = f(u)$:

$$\Delta w \approx f'(u_0)\Delta u, \qquad \text{if } \Delta u \approx 0 \ . \tag{9}$$

We wish to justify — without using reasoning based on 3-space — the approximation formula

$$\Delta w \ \approx \ \left(\frac{\partial w}{\partial x}\right)_0 \Delta x + \left(\frac{\partial w}{\partial y}\right)_0 \Delta y, \qquad\qquad \text{if} \quad \Delta x \approx 0,\ \Delta y \approx 0\ . \tag{6}$$



We are trying to calculate the change in w as we go from P to R in the picture, where $P = (x_0, y_0)$, $R = (x_0 + \Delta x, y_0 + \Delta y)$. This change can be thought of as taking place in two steps:

$$\Delta w \ = \ \Delta w_1 \ + \ \Delta w_2, \tag{10}$$

the first being the change in w as you move from P to Q, the second the change as you move from Q to R. Using the one-variable approximation formula (9) :

$$\Delta w_1 \ \approx \ \frac{d}{dx} f(x, y_0)\bigg|_{x_0} \cdot \Delta x \ = \ f_x(x_0, y_0)\,\Delta x; \tag{11}$$

similarly,

$$\begin{aligned} \Delta w_2 \ &\approx \ \frac{d}{dy} f(x_0+\Delta x, y)\bigg|_{y_0} \cdot \Delta y \ = \ f_y(x_0+\Delta x, y_0)\,\Delta y \\ &\approx f_y(x_0, y_0)\,\Delta y, \end{aligned} \tag{12}$$

if we assume that f_y is continuous (i.e., f is smooth), since the difference between the two terms on the right in the last two lines will then be like $\epsilon\,\Delta y$, which is negligible compared with either term itself. Substituting the two approximate values (11) and (12) into (10) gives us the approximation formula (6). $\square$

To make this a proof, the error terms in the approximations have to be analyzed, or more simply, one has to replace the $\approx$ symbol by equalities based on the Mean-Value Theorem of one-variable calculus.

This argument readily generalizes to the higher-dimensional approximation formulas, such as (7); again the essential hypothesis would be smoothness: the three partial derivatives w_x, w_y, w_z should be continuous in a neighborhood of the point (x_0, y_0, z_0).

Critical Points

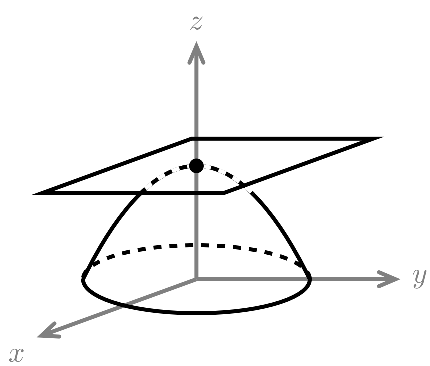
Critical points:
A standard question in calculus, with applications to many fields, is to find the points where a function reaches its relative maxima and minima.

Just as in single variable calculus we will look for maxima and minima (collectively called *extrema*) at points (x_0, y_0) where the first derivatives are 0. Accordingly we define a *critical point* as any point (x_0, y_0) where

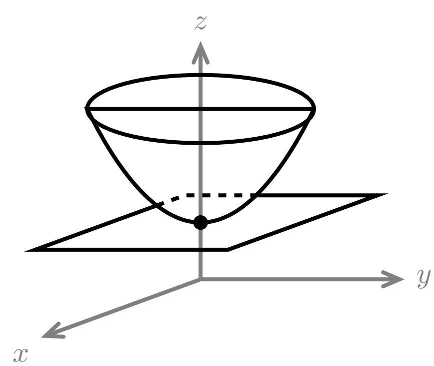
$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \text{ and } \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Often we will abbreviate this as $f_x = 0$ and $f_y = 0$.

Our first job is to verify that relative maxima and minima occur at critical points. The figures below illustrates that they occur at places where the tangent plane is horizontal.



Max. with horizontal tang. plane



Min. with horizontal tang. plane

Since horizontal planes are of the form $z =$ constant. and the equation of the tangent plane at (x_0, y_0, z_0) is

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

we see it is horizontal when

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

Thus, extrema occur at critical points. But, just as in single variable calculus, not all critical points are extrema.

Example: Find the critical points of $z = x^2 + y^2 + .5$.

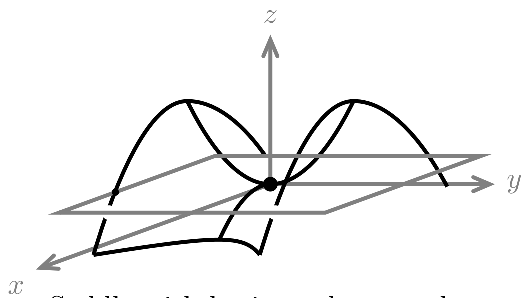
<u>**Answer:**</u> $\dfrac{\partial z}{\partial x} = 2x$ and $\dfrac{\partial z}{\partial y} = 2y$. Clearly the only point where both derivatives are 0 is $(0, 0)$. Thus, there is a single critical point at $(0, 0)$. The figure shows it is clearly the point where z reaches a minimum value. (See the figure above on the right.)

Example: Find the critical points of $z = 1 - x^2 - y^2$.

<u>**Answer:**</u> $\dfrac{\partial z}{\partial x} = -2x$ and $\dfrac{\partial z}{\partial y} = -2y$. Clearly the only point where both derivatives are 0 is $(0, 0)$. Thus, there is a single critical point at $(0, 0)$. The figure shows it is clearly the point where z reaches a maximum value. (See the figure above on the left.)

Example: Find the critical points of $z = -x^2 + y^2$.

Answer: $\dfrac{\partial z}{\partial x} = -2x$ and $\dfrac{\partial z}{\partial y} = 2y$. Clearly the only point where both derivatives are 0 is $(0, 0)$. Thus, there is a single critical point at $(0, 0)$. The figure shows it is neither a minimum or a maximum.



Saddle with horizontal tang. plane

Example: Making a box with minimum material.

A box is made of cardboard with double thick sides, a triple thick bottom, single thick front and back and no top. It's volume = 3.

What dimensions use the least amount of cardboard?

Answer: The box shown has dimensions x, y, and z.

The area of one side $= yz$. There are two double thick sides $\Rightarrow$ cardboard used $= 4yz$.

The area of the front (and back) $= xz$. It is single thick $\Rightarrow$ cardboard used $= 2xz$.

The area of the bottom $= xy$. It is triple thick $\Rightarrow$ cardboard used $= 3xy$.

Thus, the total cardboard used is

$$w = 4yz + 2xz + 3xy.$$

The volume $= 3 = xyz \Rightarrow z = \frac{3}{xy}$. Substituting this in the formula for w gives

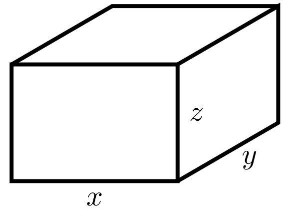
$$w = \frac{12}{x} + \frac{6}{y} + 3xy.$$

We find the critical points of w.

$$w_x = -\frac{12}{x^2} + 3y = 0, \quad w_y = -\frac{6}{y^2} + 3x = 0.$$

The first equation implies $y = \dfrac{4}{x^2}$. Substituting this in the second equation gives $-\dfrac{6}{16}x^4 + 3x = 0$. Thus, $x = 0$ or 2. We reject 0 since then y is undefined. Using $x = 2$ we find $y = 1$. Thus, there there is one critical point at (2,1). and at this point we have $z = 3/2$.

This point gives the box with minimum cardboard used because physically we know it must have a minimum somewhere. Later we will learn to check this with the second derivative test.

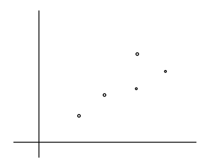


Least Squares Interpolation

1. The least-squares line.

Suppose you have a large number n of experimentally determined points, through which you want to pass a curve. There is a formula (the Lagrange interpolation formula) producing a polynomial curve of degree $n-1$ which goes through the points exactly. But normally one wants to find a simple curve, like a line, parabola, or exponential, which goes approximately through the points, rather than a high-degree polynomial which goes exactly through them. The reason is that the location of the points is to some extent determined by experimental error, so one wants a smooth-looking curve which averages out these errors, not a wiggly polynomial which takes them seriously.

In this section, we consider the most common case — finding a line which goes approximately through a set of data points.



Suppose the data points are

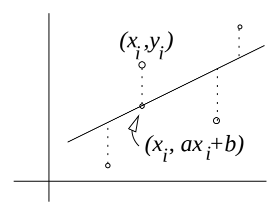
$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

and we want to find the line

$$y = ax + b \tag{1}$$

which "best" passes through them. Assuming our errors in measurement are distributed randomly according to the usual bell-shaped curve (the so-called "Gaussian distribution"), it can be shown that the right choice of a and b is the one for which the sum D of the squares of the deviations

$$D = \sum_{i=1}^{n} \big(y_i - (ax_i + b)\big)^2 \tag{2}$$



is a *minimum*. In the formula (2), the quantities in parentheses (shown by dotted lines in the picture) are the **deviations** between the observed values y_i and the ones $ax_i + b$ that would be predicted using the line (1).

The deviations are squared for theoretical reasons connected with the assumed Gaussian error distribution; note however that the effect is to ensure that we sum only positive quantities; this is important, since we do not want deviations of opposite sign to cancel each other out. It also weights more heavily the larger deviations, keeping experimenters honest, since they tend to ignore large deviations ("I had a headache that day").

This prescription for finding the line (1) is called the **method of least squares**, and the resulting line (1) is called the **least-squares** line or the **regression** line.

To calculate the values of a and b which make D a minimum, we see where the two partial derivatives are zero:

$$\begin{aligned} \frac{\partial D}{\partial a} &= \sum_{i=1}^{n} 2(y_i - ax_i - b)(-x_i) = 0 \\ \frac{\partial D}{\partial b} &= \sum_{i=1}^{n} 2(y_i - ax_i - b)(-1) = 0\,. \end{aligned} \tag{3}$$

These give us a pair of *linear* equations for determining a and b, as we see by collecting terms and cancelling the 2's:

$$\begin{aligned} \left(\sum x_i^2\right) a \;+\; \left(\sum x_i\right) b \;&=\; \sum x_i y_i \\ \left(\sum x_i\right) a \;+\; n\, b \;&=\; \sum y_i\ . \end{aligned} \tag{4}$$

(Notice that it saves a lot of work to differentiate (2) using the chain rule, rather than first expanding out the squares.)

The equations (4) are usually divided by n to make them more expressive:

$$\begin{aligned} \bar{s}\, a \;+\; \bar{x}\, b \;&=\; \frac{1}{n}\sum x_i y_i \\ \bar{x}\, a \;+\; b \;&=\; \bar{y}\ , \end{aligned} \tag{5}$$

where $\bar{x}$ and $\bar{y}$ are the average of the x_i and y_i, and $\bar{s} = \sum x_i^2/n$ is the average of the squares.

From this point on use linear algebra to determine a and b. It is a good exercise to see that the equations are always solvable unless all the x_i are the same (in which case the best line is vertical and can't be written in the form (1)).

In practice, least-squares lines are found by pressing a calculator button, or giving a MatLab command. Examples of calculating a least-squares line are in the exercises accompanying the course. Do them from scratch, starting from (2), since the purpose here is to get practice with max-min problems in several variables; don't plug into the equations (5). Remember to differentiate (2) using the chain rule; don't expand out the squares, which leads to messy algebra and highly probable error.

2. Fitting curves by least squares.

If the experimental points seem to follow a curve rather than a line, it might make more sense to try to fit a second-degree polynomial

$$y = a_0 + a_1 x + a_2 x^2 \tag{6}$$

to them. If there are only three points, we can do this exactly (by the Lagrange interpolation formula). For more points, however, we once again seek the values of a_0, a_1, a_2 for which the sum of the squares of the deviations

$$D \;=\; \sum_1^n \big(y_i - (a_0 + a_1 x_i + a_2 x_i^2)\big)^2 \tag{7}$$

is a minimum. Now there are three unknowns, a_0, a_1, a_2. Calculating (remember to use the chain rule!) the three partial derivatives $\partial D/\partial a_i,\ i = 0, 1, 2$, and setting them equal to zero leads to a square system of three linear equations; the a_i are the three unknowns, and the coefficients depend on the data points (x_i, y_i). They can be solved by finding the inverse matrix, elimination, or using a calculator or MatLab.

If the points seem to lie more and more along a line as $x \to \infty$, but lie on one side of the line for low values of x, it might be reasonable to try a function which has similar behavior, like

$$y \;=\; a_0 + a_1 x + a_2\, \frac{1}{x} \tag{8}$$

and again minimize the sum of the squares of the deviations, as in (7). In general, this method of least squares applies to a trial expression of the form

$$y = a_0 f_0(x) + a_1 f_1(x) + \ldots + a_r f_r(x), \tag{9}$$

where the $f_i(x)$ are given functions (usually simple ones like $1, x, x^2, 1/x, e^{kx}$, etc. Such an expression (9) is called a **linear combination** of the functions $f_i(x)$. The method produces a square inhomogeneous system of linear equations in the unknowns $a_0, \ldots, a_r$ which can be solved by finding the inverse matrix to the system, or by elimination.

The method also applies to finding a linear function

$$z = a_1 + a_2 x + a_3 y \tag{10}$$

to fit a set of data points

$$(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)\ . \tag{11}$$

where there are two independent variables x and y and a dependent variable z (this is the quantity being experimentally measured, for different values of (x, y)). This time after differentiation we get a 3×3 system of linear equations for determining a_1, a_2, a_3 .

The essential point in all this is that the unknown coefficients a_i should occur *linearly* in the trial function. Try fitting a function like ce^{kx} to data points by using least squares, and you'll see the difficulty right away. (Since this is an important problem — fitting an exponential to data points — one of the Exercises explains how to adapt the method to this type of problem.)

Least squares interpolation

1. Use the method of least squares to fit a line to the three data points

$$(0,0), \quad (1,2), \quad (2,1).$$

<u>Answer:</u> We are looking for the line $y = ax + b$ that best models the data. The deviation of a data point (x_i, y_i) from the model is

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

By best we mean the line that minimizes the sum of the squares of the deviation. That is we want to minimize

$$\begin{aligned} D &= (0 - (a \cdot 0 + b))^2 + (2 - (a \cdot 1 + b))^2 + (1 - (a \cdot 2 + b))^2 \\ &= b^2 + (2 - a - b)^2 + (1 - 2a - b)^2. \end{aligned}$$

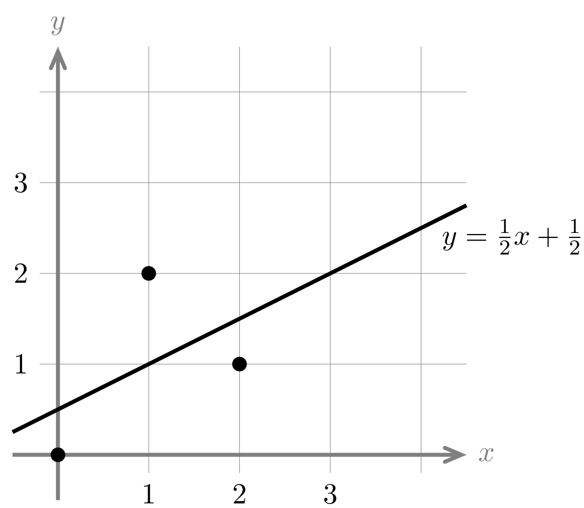
(Remember, the variables whose values are to be found are a and b.) We do not expand out the squares, rather we take the derivatives first. Setting the derivatives equal to 0 gives

$$\begin{aligned} \tfrac{\partial D}{\partial a} &= -2(2 - a - b) - 4(1 - 2a - b) = 0 \;\Rightarrow\; 10a + 6b = 8 \;\Rightarrow\; 5a + 3b = 4 \\ \tfrac{\partial D}{\partial b} &= 2b - 2(2 - a - b) - 2(1 - 2a - b) = 0 \;\Rightarrow\; 6a + 6b = 6 \;\Rightarrow\; 3a + 3b = 3. \end{aligned}$$

This linear system of two equations in two unknowns is easy to solve. We get

$$a = \frac{1}{2}, \qquad b = \frac{1}{2}.$$

Here is a plot of the problem.



Second Derivative Test

1. The Second Derivative Test

We begin by recalling the situation for twice differentiable functions $f(x)$ of one variable. To find their local (or "relative") maxima and minima, we

1. find the critical points, i.e., the solutions of $f'(x) = 0$;
2. apply the second derivative test to each critical point x_0:

$$f''(x_0) > 0 \;\Rightarrow\; x_0 \text{ is a local minimum point;}$$
$$f''(x_0) < 0 \;\Rightarrow\; x_0 \text{ is a local maximum point.}$$

The idea behind it is: at x_0 the slope $f'(x_0) = 0$; if $f''(x_0) > 0$, then $f'(x)$ is strictly increasing for x near x_0, so that the slope is negative to the left of x_0 and positive to the right, which shows that x_0 is a minimum point. The reasoning for the maximum point is similar.

If $f''(x_0) = 0$, the test fails and one has to investigate further, by taking more derivatives, or getting more information about the graph. Besides being a maximum or minimum, such a point could also be a horizontal point of inflection.

The analogous test for maxima and minima of functions of two variables $f(x, y)$ is a little more complicated, since there are several equations to satisfy, several derivatives to be taken into account, and another important geometric possibility for a critical point, namely a **saddle point**. This is a local minimax point; around such a point the graph of $f(x, y)$ looks like the central part of a saddle, or the region around the highest point of a mountain pass. In the neighborhood of a saddle point, the graph of the function lies both above and below its horizontal tangent plane at the point.

The second-derivative test for maxima, minima, and saddle points has two steps.

1. Find the critical points by solving the simultaneous equations $\begin{cases} f_x(x, y) = 0, \\ f_y(x, y) = 0. \end{cases}$

Since a critical point (x_0, y_0) is a solution to both equations, both partial derivatives are zero there, so that the tangent plane to the graph of $f(x, y)$ is horizontal.

2. To test such a point to see if it is a local maximum or minimum point, we calculate the three second derivatives at the point (we use subscript 0 to denote evaluation at (x_0, y_0), so for example $(f)_0 = f(x_0, y_0)$), and denote the values by A, B, and C:

$$A = (f_{xx})_0, \qquad B = (f_{xy})_0 = (f_{yx})_0, \qquad C = (f_{yy})_0, \tag{1}$$

(we are assuming the derivatives exist and are continuous).

Second-derivative test. *Let (x_0, y_0) be a critical point of $f(x, y)$, and A, B, and C be as in (1). Then*

$$AC - B^2 > 0, \quad A > 0 \text{ or } C > 0 \;\Rightarrow\; (x_0, y_0) \text{ is a minimum point;}$$
$$AC - B^2 > 0, \quad A < 0 \text{ or } C < 0 \;\Rightarrow\; (x_0, y_0) \text{ is a maximum point;}$$
$$AC - B^2 < 0 \;\Rightarrow\; (x_0, y_0) \text{ is a saddle point.}$$

If $AC - B^2 = 0$, the test fails and more investigation is needed.

Note that if $AC - B^2 > 0$, then $AC > 0$, so that A and C must have the same sign.

Example 1. Find the critical points of $w = 12x^2 + y^3 - 12xy$ and determine their type.

Solution. We calculate the partial derivatives easily:

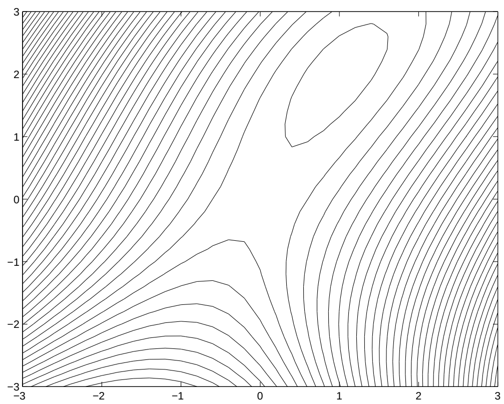
$$\begin{array}{ll} w_x = 24x - 12y & A = w_{xx} = 24 \\ w_y = 3y^2 - 12x & B = w_{xy} = -12 \\ & C = w_{yy} = 6y \end{array} \tag{2}$$

To find the critical points we solve simultaneously the equations $w_x = 0$ and $w_y = 0$; we get

$$\begin{array}{c} w_x = 0 \\ w_y = 0 \end{array} \Rightarrow \begin{array}{c} y = 2x \\ y^2 = 4x \end{array} \Rightarrow 4x^2 = 4x \Rightarrow x = 0, 1 \Rightarrow \begin{array}{c} (x, y) = (0, 0) \\ (x, y) = (1, 2) \end{array}.$$

Thus there are two critical points: $(0, 0)$ and $(1, 2)$. To determine their type, we use the second derivative test: we have $AC - B^2 = 144y - 144$, so that

at $(0, 0)$, we have $AC - B^2 = -144$, so it is a saddle point;
at $(1, 2)$, we have $AC - B^2 = 144$ and $A > 0$, so it is a a minimum point.



A plot of the level curves is given at the right, which confirms the above. Note that the behavior of the level curves near the origin can be determined by using the approximation $w \approx 12x^2 - 12xy$; this shows the level curves near $(0, 0)$ look like those of the function $x(x - y)$: the family of hyperbolas $x(x - y) = c$, with asymptotes given by the degenerate hyperbola $x(x - y) = 0$, i.e., the pair of lines $x = 0$ (the y-axis) and $x - y = 0$ (the diagonal line $y = x$).

2. Justification for the Second-derivative Test.

The test involves the quantity $AC - B^2$. In general, whenever we see the expressions $B^2 - 4AC$ or $B^2 - AC$ or their negatives, it means the quadratic formula is involved, in one of its two forms (the second is often used to get rid of the excess two's):

$$Ax^2 + Bx + C = 0 \quad \Rightarrow \quad x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \tag{3}$$

$$Ax^2 + 2Bx + C = 0 \quad \Rightarrow \quad x = \frac{-B \pm \sqrt{B^2 - AC}}{A} \tag{4}$$

This is what is happening here. We want to know whether, near a critical point P_0, the graph of our function $w = f(x, y)$ always stays on one side of its horizontal tangent plane (P_0 is then a maximum or minimum point), or whether it lies partly above and partly below the tangent plane (P_0 is then a saddle point). As we will see, this is determined by how the graph of a quadratic function $f(x)$ lies with respect to the x-axis. Here is the basic lemma.

Lemma. For the quadratic function $Ax^2 + 2Bx + C$,

$$AC - B^2 > 0, \quad A > 0 \text{ or } C > 0 \quad \Rightarrow \quad Ax^2 + 2Bx + C > 0 \quad \text{for all } x; \tag{5}$$

$$AC - B^2 > 0, \quad A < 0 \text{ or } C < 0 \quad \Rightarrow \quad Ax^2 + 2Bx + C < 0 \quad \text{for all } x; \tag{6}$$

$$AC - B^2 < 0 \quad \Rightarrow \quad \begin{cases} Ax^2 + 2Bx + C > 0, & \text{for some } x; \\ Ax^2 + 2Bx + C < 0, & \text{for some } x. \end{cases} \tag{7}$$

Proof of the Lemma. To prove (5), we note that the quadratic formula in the form (4) shows that the zeros of $Ax^2 + 2Bx + C$ are imaginary, i.e., it has no real zeros. Therefore its graph must lie entirely on one side of the x-axis; which side can be determined from either A or C, since

$$A > 0 \;\Rightarrow\; \lim_{x\to\infty} Ax^2 + 2Bx + C = \infty; \qquad C > 0 \;\Rightarrow\; Ax^2 + 2Bx + C > 0 \text{ when } x = 0.$$

If $A < 0$ or $C < 0$, the reasoning is analogous and proves (6).

If on the other hand $AC - B^2 < 0$, formula (4) shows the quadratic function has two real roots, so that its parabolic graph crosses the x-axis twice, and hence lies partly above and partly below it. This proves (7). □

Proof of the Second-derivative Test in a special case.

The simplest function is a linear function, $w = w_0 + ax + by$, but it does not in general have maximum or minimum points and its second derivatives are all zero. The simplest functions to have interesting critical points are the quadratic functions, which we write in the form (the 2's will be explained momentarily):

$$w = w_0 + ax + by + \frac{1}{2}(Ax^2 + 2Bxy + Cy^2). \tag{8}$$

Such a function has in general a unique critical point, which we will assume is $(0,0)$; this gives the function a special form, which we can determine by evaluating its partial derivatives at $(0,0)$:

$$\begin{array}{ll} & w_{xx} = A \\ (w_x)_0 = a & \\ & w_{xy} = B \\ (w_y)_0 = b & \\ & w_{yy} = C \end{array} \tag{9}$$

(The neat look of the above explains the $\frac{1}{2}$ and $2B$ in (8).) Since $(0,0)$ is a critical point, (9) shows that $a = 0$ and $b = 0$, so our quadratic function has the form

$$w - w_0 = \frac{1}{2}(Ax^2 + 2Bxy + Cy^2). \tag{10}$$

We moved w_0 to the left side since the tangent plane at $(0,0)$ is the horizontal plane $w = w_0$, and we are interested in whether the graph of the quadratic function lies above or below this tangent plane, i.e., whether $w - w_0 > 0$ or $w - w_0 < 0$ at points other than the origin.

If $(x,y) \neq (0,0)$, then either $x \neq 0$ or $y \neq 0$; say $y \neq 0$. Then we write (10) as

$$w - w_0 \;=\; \frac{y^2}{2}\left[A\left(\frac{x}{y}\right)^2 + 2B\left(\frac{x}{y}\right) + C\right] \tag{11}$$

We know that $y^2 > 0$ if $y \neq 0$; applying our previous lemma to the factor on the right of (11), (or if $y = 0$, switching the roles of x and y in (11) and applying the lemma), we get

$$\begin{array}{rcl}
AC - B^2 > 0, \quad A > 0 \text{ or } C > 0 & \Rightarrow & w - w_0 > 0 \quad \text{for all } (x,y) \neq (0,0); \\
 & \Rightarrow & (0,0) \text{ is a minimum point;} \\
AC - B^2 > 0, \quad A < 0 \text{ or } C < 0 & \Rightarrow & w - w_0 < 0 \quad \text{for all } (x,y) \neq (0,0); \\
 & \Rightarrow & (0,0) \text{ is a maximum point;} \\
AC - B^2 < 0 & \Rightarrow & \begin{cases} w - w_0 > 0, & \text{for some } (x,y); \\ w - w_0 < 0, & \text{for some } (x,y); \end{cases} \\
 & \Rightarrow & (0,0) \text{ is a saddle point.}
\end{array}$$

Argument for the Second-derivative Test for a general function.

This part won't be rigorous, only suggestive, but it will give the right idea.

We consider a general function $w = f(x, y)$, and assume it has a critical point at (x_0, y_0), and continuous second derivatives in the neighborhood of the critical point. Then by a generalization of Taylor's formula to functions of several variables, the function has a best quadratic approximation at the critical point. To simplify the notation, we will move the critical point to the origin by making the change of variables

$$u = x - x_0, \qquad v = y - y_0.$$

Then the best quadratic approximation is (if the x, y on the left and u, v on the right is upsetting, just imagine u and v replaced everywhere by $x - x_0$ and $y - y_0$):

$$w \;=\; f(x,y) \;\approx\; w_0 + \frac{1}{2}\big(Au^2 + 2Buv + Cv^2\big); \tag{13}$$

here the coefficients A, B, C are given as in (1) by the second partial derivatives with respect to u and v at $(0, 0)$, or what is the same (according to the chain rule—see the footnote below), by the second partial derivatives with respect to x and y at (x_0, y_0).

> (Intuitively, one can see the coefficients have these values by differentiating both sides of (13) and pretending the approximation is an equality. There are no linear terms in u and v on the right since $(0, 0)$ is a critical point.)

Since the quadratic function on the right of (13) is the best approximation to $w = f(x, y)$ for (x, y) close to (x_0, y_0), it is reasonable to suppose that their graphs are essentially the same near (x_0, y_0), so that if the quadratic function has a maximum, minimum or saddle point there, so will $f(x, y)$. Thus our results for the special case of a quadratic function having the origin as critical point carry over to the general function $f(x, y)$ at a critical point (x_0, y_0), if we interpret $A,\ B,\ C$ as the second partial derivatives at (x_0, y_0).

This is what the second derivative test says. $\square$

Footnote: Using $u = x - x_0$ and $v = y - y_0$, we can apply the chain rule for partial derivatives, which tells us that for all x, y and the corresponding u, v, we have

$$w_x = w_u \frac{\partial u}{\partial x} + w_v \frac{\partial v}{\partial x} = w_u, \ \text{ since } u_x = 1 \text{ and } v_x = 0,$$

and similarly, $w_y = w_v$. Therefore at the corresponding points,

$$(w_x)_{(x_0,y_0)} = (w_u)_{(0,0)}, \qquad\qquad (w_y)_{(x_0,y_0)} = (w_v)_{(0,0)},$$

and differentiating once more and using the same reasoning,

$$(w_{xx})_{(x_0,y_0)} = (w_{uu})_{(0,0)}, \qquad (w_{xy})_{(x_0,y_0)} = (w_{uv})_{(0,0)}, \qquad (w_{yy})_{(x_0,y_0)} = (w_{vv})_{(0,0)}.$$

Second derivative test

1. Find and classify all the critical points of

$$f(x,y) = x^6 + y^3 + 6x - 12y + 7.$$

<u>**Answer:**</u> Taking the first partials and setting them to 0:

$$\frac{\partial z}{\partial x} = 6x^5 + 6 = 0 \quad \text{and} \quad \frac{\partial z}{\partial y} = 3y^2 - 12 = 0.$$

The first equation implies $x = -1$ and the second implies $y = \pm 2$. Thus, the critical points are $(-1, 2)$ and $(-1, -2)$.

Taking second partials:

$$\frac{\partial^2 z}{\partial x^2} = 30x^4, \quad \frac{\partial^2 z}{\partial xy} = 0, \quad \frac{\partial^2 z}{\partial y^2} = 6y.$$

We analyze each critical point in turn.
At (-1,-2): $A = z_{xx}(-1,-2) = 30, \quad B = z_{xy}(-1,-2) = 0, \quad C = z_{yy}(-1,-2) = -12.$
Therefore $AC - B^2 = -360 < 0$, which implies the critical point is a saddle.

At (-1,2): $A = z_{xx}(-1,2) = 30, \quad B = z_{xy}(-1,2) = 0, \quad C = z_{yy}(-1,2) = 12.$
Therefore $AC - B^2 = 360 > 0$ and $A > 0$, which implies the critical point is a minimum.

Chain Rule and Total Differentials

1. Find the total differential of $w = x^3yz + xy + z + 3$ at $(1, 2, 3)$.

Answer: The total differential at the point (x_0, y_0, z_0) is

$$dw = w_x(x_0, y_0, z_0)\,dx + w_y(x_0, y_0, z_0)\,dy + w_z(x_0, y_0, z_0)\,dz.$$

In our case,

$$w_x = 3x^2yz + y, \quad w_y = x^3z + x, \quad w_z = x^3y + 1.$$

Substituting in the point $(1, 2, 3)$ we get: $w_x(1, 2, 3) = 20,\ w_y(1, 2, 3) = 4,\ w_z(1, 2, 3) = 3$. Thus,

$$dw = 20\,dx + 4\,dy + 3\,dz.$$

2. Suppose $w = x^3yz + xy + z + 3$ and

$$x = 3\cos t, \quad y = 3\sin t, \quad z = 2t.$$

Compute $\dfrac{dw}{dt}$ and evaluate it at $t = \pi/2$.

Answer: We do not substitute for x, y, z before differentiating, so we can practice the chain rule.

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= (3x^2yz + y)(-3\sin t) + (x^3z + x)(3\cos t) + (x^3y + 1)(2).\end{aligned}$$

At $t = \pi/2$ we have $x = 0,\ y = 3,\ z = \pi, \qquad \sin\pi/2 = 1, \cos\pi/2 = 0.$
Thus,

$$\left.\frac{dw}{dt}\right|_{\pi/2} = 3(-3) + 3(0) + (1)2 = -7.$$

3. Show how the tangent approximation formula leads to the chain rule that was used in the previous problem.

Answer: The approximation formula is

$$\Delta w \approx \left.\frac{\partial f}{\partial x}\right|_o \Delta x + \left.\frac{\partial f}{\partial y}\right|_o \Delta y + \left.\frac{\partial f}{\partial z}\right|_o \Delta z.$$

If x, y, z are functions of time then dividing the approximation formula by Δt gives

$$\frac{\Delta w}{\Delta t} \approx \left.\frac{\partial f}{\partial x}\right|_o \frac{\Delta x}{\Delta t} + \left.\frac{\partial f}{\partial y}\right|_o \frac{\Delta y}{\Delta t} + \left.\frac{\partial f}{\partial z}\right|_o \frac{\Delta z}{\Delta t}.$$

In the limit as $\Delta t \to 0$ we get the chain rule.

Note: we use the regular 'd' for the derivative $\frac{dw}{dt}$ because in the chain of computations

$$t \to x, y, z \to w$$

the dependent variable w is ultimately a function of exactly one independent variable t. Thus, the derivative with respect to t is not a partial derivative.

Chain rule

Now we will formulate the chain rule when there is more than one independent variable.

We suppose w is a function of x, y and that x, y are functions of u, v. That is,

$$w = f(x, y) \text{ and } x = x(u, v),\ y = y(u, v).$$

The use of the term chain comes because to compute w we need to do a chain of computations

$$(u, v) \to (x, y) \to w.$$

We will say w is a *dependent* variable, u and v are *independent* variables and x and y are *intermediate* variables.

Since w is a function of x and y it has partial derivatives $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$.

Since, ultimately, w is a function of u and v we can also compute the partial derivatives $\dfrac{\partial w}{\partial u}$ and $\dfrac{\partial w}{\partial v}$. The chain rule relates these derivatives by the following formulas.

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v}.
\end{aligned}$$

Example: Given $w = x^2y + y^2 + x$, $x = u^2v$, $y = uv^2$ find $\dfrac{\partial w}{\partial u}$.

<u>**Answer:**</u> First we compute

$$\frac{\partial w}{\partial x} = 2xy + 1,\ \frac{\partial w}{\partial y} = x^2 + 2y,\ \frac{\partial x}{\partial u} = 2uv,\ \frac{\partial y}{\partial u} = v^2,\ \frac{\partial x}{\partial v} = u^2,\ \frac{\partial y}{\partial v} = 2uv.$$

The chain rule then implies

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
&= (2xy + 1)2uv + (x^2 + 2y)v^2 \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} \\
&= (2xy + 1)u^2 + (x^2 + 2y)2uv.
\end{aligned}$$

Often, it is okay to leave the variables mixed together. If, for example, you wanted to compute $\frac{\partial w}{\partial u}$ when $(u, v) = (1, 2)$ all you have to do is compute x and y and use these values, along with u, v, in the formula for $\dfrac{\partial w}{\partial u}$.

$$x = 2,\ y = 4 \ \Rightarrow\ \frac{\partial w}{\partial u} = (5)(4) + (12)(4) = 68.$$

If you actually need the derivatives expressed in just the variables u and v then you would have to substitute for x, y and z.

Proof of the chain rule:
Just as before our argument starts with the tangent approximation at the point (x_0, y_0).

$$\Delta w \approx \left.\frac{\partial w}{\partial x}\right|_o \Delta x + \left.\frac{\partial w}{\partial y}\right|_o \Delta y.$$

Now hold v constant and divide by Δu to get

$$\frac{\Delta w}{\Delta u} \approx \left.\frac{\partial w}{\partial x}\right|_o \frac{\Delta x}{\Delta u} + \left.\frac{\partial w}{\partial y}\right|_o \frac{\Delta y}{\Delta u}.$$

Finally, letting $\Delta u \to 0$ gives the chain rule for $\dfrac{\partial w}{\partial u}$.

Ambiguous notation

Often you have to figure out the dependent and independent variables from context.

Thermodynamics is a big player here. It has, for example, the variables P, T, V, U, S. and *any* two can be taken to be independent and the others are functions of those two.

We will do more with this topic in the future.

Chain rule with more variables

1. Let $w = xyz$, $\quad x = u^2 v$, $y = uv^2$, $z = u^2 + v^2$.

a) Use the chain rule to find $\dfrac{\partial w}{\partial u}$.

b) Find the total differential dw in terms of du and dv.

c) Find $\dfrac{\partial w}{\partial u}$ at the point $(u, v) = (1, 2)$.

Answer: a) The chain rule says

$$\begin{aligned}\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial u} \\ &= (yz)(2uv) + (xz)(v^2) + (xy)(2u).\end{aligned}$$

b) Using the formulas given we get

$$dw = yz\,dx + xz\,dy + xy\,dz$$

and

$$dx = 2uv\,du + u^2\,dv, \quad dy = v^2\,du + 2uv\,dv, \quad dz = 2u\,du + 2v\,dv.$$

Substituting for dx, dy, dz in the equation for dw gives

$$\begin{aligned}dw &= (yz)(2uv\,du + u^2\,dv) + (xz)(v^2\,du + 2uv\,dv) + (xy)(2u\,du + 2v\,dv). \\ &= (2yzuv + xzv^2 + 2xyu)\,du + (yzu^2 + 2xzuv + 2xyv)\,dv.\end{aligned}$$

Therefore

$$\frac{\partial w}{\partial u} = 2yzuv + xzv^2 + 2xyu \text{ and } \frac{\partial w}{\partial v} = yzu^2 + 2xzuv + 2xyv.$$

c) We do the chain of computations to compute the partial.

$$(u, v) = (1, 2) \Rightarrow (x, y, z) = (2, 4, 5) \Rightarrow \frac{\partial w}{\partial u} = (20)(4) + (10)(4) + (8)(2) = 136.$$

Gradient: definition and properties

Definition of the gradient

If $w = f(x, y)$, then $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$ are the rates of change of w in the $\mathbf{i}$ and $\mathbf{j}$ directions.

It will be quite useful to put these two derivatives together in a vector called the *gradient* of w.

$$\operatorname{grad} w = \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \right\rangle.$$

We will also use the symbol $\boldsymbol{\nabla} w$ to denote the gradient. (You read this as 'gradient of w' or 'grad w'.)

Of course, if we specify a point $P_0 = (x_0, y_0)$, we can evaluate the gradient at that point. We will use several notations for this

$$\operatorname{grad} w(x_0, y_0) = \boldsymbol{\nabla}\ w|_{P_0} = \boldsymbol{\nabla}\ w|_o = \left\langle \left.\frac{\partial w}{\partial x}\right|_o, \left.\frac{\partial w}{\partial y}\right|_o \right\rangle.$$

Note well the following: (as we look more deeply into properties of the gradient these can be points of confusion).

1. The gradient takes a scalar function $f(x, y)$ and produces a vector $\boldsymbol{\nabla} f$.
2. The vector $\boldsymbol{\nabla} f(x, y)$ lies in the plane.

For functions $w = f(x, y, z)$ we have the gradient

$$\operatorname{grad} w = \boldsymbol{\nabla} w = \left\langle \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z} \right\rangle.$$

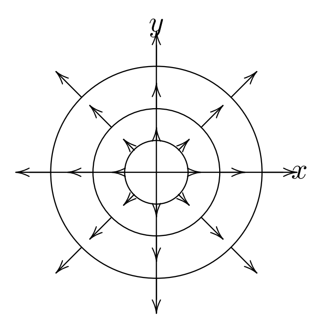
That is, the gradient takes a scalar function of three variables and produces a three dimensional vector.

The gradient has many geometric properties. In the next session we will prove that for $w = f(x, y)$ the gradient is perpendicular to the level curves $f(x, y) = c$. We can show this by direct computation in the following example.

Example 1: Compute the gradient of $w = (x^2 + y^2)/3$ and show that the gradient at $(x_0, y_0) = (1, 2)$ is perpendicular to the level curve through that point.

<u>**Answer:**</u> The gradient is easily computed

$$\boldsymbol{\nabla} w = \langle 2x/3, 2y/3 \rangle = \frac{2}{3}\langle x, y \rangle.$$



$z = (x^2 + y^2)/3$

At $(1, 2)$ we get $\boldsymbol{\nabla} w(1, 2) = \frac{2}{3}\langle 1, 2 \rangle$. The level curve through $(1, 2)$ is

$$(x^2 + y^2)/3 = 5/3,$$

which is identical to $x^2 + y^2 = 5$. That is, it is a circle of radius $\sqrt{5}$ centered at the origin. Since the gradient at (1,2) is a multiple of $\langle 1, 2 \rangle$, it points radially outward and hence is perpendicular to the circle. Below is a figure showing the gradient field and the level curves.

Example 2: Consider the graph of $y = \mathrm{e}^x$. Find a vector perpendicular to the tangent to $y = \mathrm{e}^x$ at the point $(1, \mathrm{e})$.

Old method: Find the slope take the negative reciprocal and make the vector.

New method: This graph is the level curve of $w = y - \mathrm{e}^x = 0$.

$\nabla w = \langle -\mathrm{e}^x, 1 \rangle \;\Rightarrow\;$ (at $x = 1$) $\nabla w(1, e) = \langle -\mathrm{e}, 1 \rangle$ is perpendicular to the tangent vector to the graph, $\mathbf{v} = \langle 1, \mathrm{e} \rangle$.

Higher dimensions

Similarly, for $w = f(x, y, z)$ we get level surfaces $f(x, y, z) = c$. The gradient is perpendicular to the level surfaces.

Example 3: Find the tangent plane to the surface $x^2 + 2y^2 + 3z^2 = 6$ at the point $P = (1, 1, 1)$.

<u>Answer:</u> Introduce a new variable

$$w = x^2 + 2y^2 + 3z^2.$$

Our surface is the level surface $w = 6$. Saying the gradient is perpendicular to the surface means exactly the same thing as saying it is normal to the tangent plane. Computing

$$\nabla w = \langle 2x, 4y, 6z \rangle \;\Rightarrow\; \nabla w|_P = \langle 2, 4, 6 \rangle.$$

Using point normal form we get the equation of the tangent plane is

$$2(x-1) + 4(y-1) + 6(z-1) = 0, \qquad \text{or} \qquad 2x + 4y + 6z = 12.$$

18.02SC Multivariable Calculus
Fall 2010

Gradient: proof that it is perpendicular to level curves and surfaces

Let $w = f(x, y, z)$ be a function of 3 variables. We will show that at any point $P = (x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ (so $f(x_0, y_0, z_0) = c$) the gradient $\nabla f|_P$ is perpendicular to the surface.

By this we mean it is perpendicular to the tangent to any curve that lies on the surface and goes through P. (See figure.)

This follows easily from the chain rule: Let

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

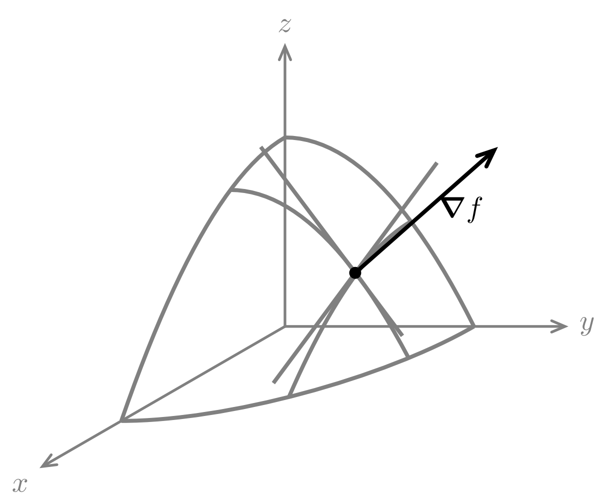
be a curve on the level surface with $\mathbf{r}(t_0) = \langle x_0, y_0, z_0 \rangle$. We let $g(t) = f(x(t), y(t), z(t))$. Since the curve is on the level surface we have $g(t) = f(x(t), y(t), z(t)) = c$. Differentiating this equation with respect to t gives

$$\frac{dg}{dt} = \left.\frac{\partial f}{\partial x}\right|_P \left.\frac{dx}{dt}\right|_{t_0} + \left.\frac{\partial f}{\partial y}\right|_P \left.\frac{dy}{dt}\right|_{t_0} + \left.\frac{\partial f}{\partial z}\right|_P \left.\frac{dz}{dt}\right|_{t_0} = 0.$$

In vector form this is

$$\begin{aligned} &\left\langle \left.\frac{\partial f}{\partial x}\right|_P, \left.\frac{\partial f}{\partial y}\right|_P, \left.\frac{\partial f}{\partial z}\right|_P \right\rangle \cdot \left\langle \left.\frac{dx}{dt}\right|_{t_0}, \left.\frac{dy}{dt}\right|_{t_0}, \left.\frac{dz}{dt}\right|_{t_0} \right\rangle = 0 \\ \Leftrightarrow \quad &\nabla f|_P \cdot \mathbf{r}'(t_0) = 0. \end{aligned}$$

Since the dot product is 0, we have shown that the gradient is perpendicular to the tangent to any curve that lies on the level surface, which is exactly what we needed to show.



Tangent Plane to a Level Surface

1. Find the tangent plane to the surface $x^2 + 2y^2 + 3z^2 = 36$ at the point $P = (1, 2, 3)$.

Answer: In order to use gradients we introduce a new variable

$$w = x^2 + 2y^2 + 3z^2.$$

Our surface is then the the level surface $w = 36$. Therefore the normal to surface is

$$\boldsymbol{\nabla} w = \langle 2x, 4y, 6z \rangle.$$

At the point P we have $\boldsymbol{\nabla} w|_P = \langle 2, 8, 18 \rangle$. Using point normal form, the equation of the tangent plane is

$$2(x - 1) + 8(y - 2) + 18(z - 3) = 0, \text{ or equivalently } 2x + 8y + 18z = 72.$$

2. Use gradients and level surfaces to find the normal to the tangent plane of the graph of $z = f(x, y)$ at $P = (x_0, y_0, z_0)$.

Answer: Introduce the new variable

$$w = f(x, y) - z.$$

The graph of $z = f(x, y)$ is just the level surface $w = 0$. We compute the normal to the surface to be

$$\boldsymbol{\nabla} w = \langle f_x, f_y, -1 \rangle.$$

At the the point P the normal is $\langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle$, so the equation of the tangent plane is

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

We can write this in a more compact form as

$$(z - z_0) = \left.\frac{\partial f}{\partial x}\right|_0 (x - x_0) + \left.\frac{\partial f}{\partial y}\right|_0 (y - y_0),$$

which is exactly the formula we saw earlier for the tangent plane to a graph.

Directional Derivatives

Directional derivative

Like all derivatives the *directional derivative* can be thought of as a ratio. Fix a unit vector $\mathbf{u}$ and a point P_0 in the *plane*. The **directional derivative** of w at P_0 in the direction $\mathbf{u}$ is defined as

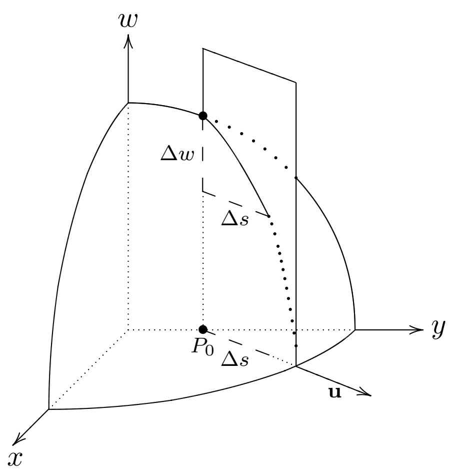
$$\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \lim_{\Delta s \to 0} \frac{\Delta w}{\Delta s}.$$

Here Δw is the change in w caused by a step of length Δs in the direction of $\mathbf{u}$ (all in the xy-plane).

Below we will show that

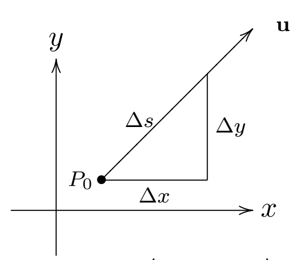
$$\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \boldsymbol{\nabla} w(P_0) \cdot \mathbf{u}. \tag{1}$$

We illustrate this with a figure showing the graph of $w = f(x, y)$. Notice that Δs is measured in the plane and Δw is the change of w on the graph.



Proof of equation 1

The figure below represents the change in position from P_0 resulting from taking a step of size Δs in the $\mathbf{u}$ direction.



Since $(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2$ we have that $\left\langle \dfrac{\Delta x}{\Delta s}, \dfrac{\Delta y}{\Delta s} \right\rangle$. is a unit vector, so

$$\mathbf{u} = \left\langle \frac{\Delta x}{\Delta s}, \frac{\Delta y}{\Delta s} \right\rangle.$$

The tangent plane approximation at P_0 is

$$\Delta w \approx \left.\frac{\partial w}{\partial x}\right|_{P_0} \Delta x + \left.\frac{\partial w}{\partial y}\right|_{P_0} \Delta y$$

Dividing this approximation by Δs gives

$$\frac{\Delta w}{\Delta s} \approx \left.\frac{\partial w}{\partial x}\right|_{P_0} \frac{\Delta x}{\Delta s} + \left.\frac{\partial w}{\partial y}\right|_{P_0} \frac{\Delta y}{\Delta s}.$$

We can rewrite this as a dot product

$$\frac{\Delta w}{\Delta s} \approx \left\langle \left.\frac{\partial w}{\partial x}\right|_{P_0}, \left.\frac{\partial w}{\partial y}\right|_{P_0} \right\rangle \cdot \left\langle \frac{\Delta x}{\Delta s}, \frac{\Delta y}{\Delta s} \right\rangle.$$

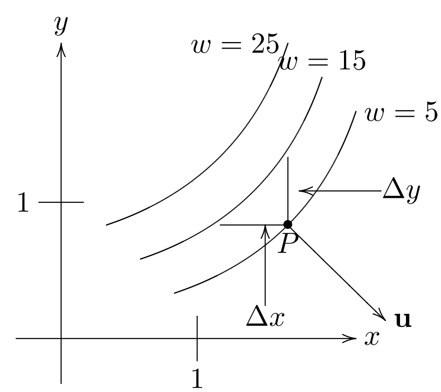
In the dot product the first term is $\nabla w|_{P_0}$ and the second is just $\mathbf{u}$, so,

$$\frac{\Delta w}{\Delta s} \approx \nabla w|_{P_0} \cdot \mathbf{u}.$$

Now taking the limit we get equation (1).

Example: (Algebraic example) Let $w = x^3 + 3y^2$.
Compute $\frac{dw}{ds}$ at $P_0 = (1, 2)$ in the direction of $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j}$.
<u>Answer:</u> We compute all the necessary pieces:
i) $\nabla w = \langle 3x^2, 6y \rangle \Rightarrow \nabla w|_{(1,2)} = \langle 3, 12 \rangle$.
ii) $\mathbf{u}$ must be a unit vector, so $\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \langle \frac{3}{5}, \frac{4}{5} \rangle$.
iii) $\left.\frac{dw}{ds}\right|_{P_0,\mathbf{u}} = \nabla w|_{(1,2)} \cdot \mathbf{u} = \langle 3, 12 \rangle \cdot \langle \frac{3}{5}, \frac{4}{5} \rangle = \boxed{\frac{57}{5}.}$

Example: (Geometric example) Let $\mathbf{u}$ be the direction of $\langle 1, -1 \rangle$.
Using the picture at right estimate $\left.\frac{\partial w}{\partial x}\right|_P$, $\left.\frac{\partial w}{\partial y}\right|_p$, and $\left.\frac{dw}{ds}\right|_{P,\mathbf{u}}$.



By measuring from P to the next in level curve in the x direction we see that $\Delta x \approx -.5$.
$\Rightarrow \boxed{\left.\frac{\partial w}{\partial x}\right|_P \approx \frac{\Delta w}{\Delta x} \approx \frac{10}{-.5} = -20.}$

Similarly, we get $\boxed{\left.\frac{\partial w}{\partial y}\right|_P \approx 20.}$

Measuring in the $\mathbf{u}$ direction we get $\Delta s \approx -.3$
$\Rightarrow \boxed{\left.\frac{dw}{ds}\right|_{P,\mathbf{u}} \approx \frac{\Delta w}{\Delta s} \approx \frac{10}{.3} = -33.3.}$

Direction of maximum change:
The direction that gives the maximum rate of change is in the same direction as ∇w. The proof of this uses equation (1). Let θ be the angle between ∇w and $\mathbf{u}$. Then the geometric form of the dot product says

$$\left.\frac{dw}{ds}\right|_{\mathbf{u}} = \nabla w \cdot \mathbf{u} = |\nabla w||\mathbf{u}| \cos\theta = |\nabla w| \cos\theta.$$

(In the last equation we dropped the $|\mathbf{u}|$ because it equals 1.) Now it is obvious that this is greatest when $\theta = 0$. That is, when $\boldsymbol{\nabla} w$ and $\mathbf{u}$ are in the same direction.

18.02SC Multivariable Calculus
Fall 2010

Lagrange Multipliers

We will give the argument for why Lagrange multipliers work later. Here, we'll look at where and how to use them. Lagrange multipliers are used to solve constrained optimization problems. That is, suppose you have a function, say $f(x, y)$, for which you want to find the maximum or minimum value. But, you are not allowed to consider all (x, y) while you look for this value. Instead, the (x, y) you can consider are constrained to lie on some curve or surface. There are lots of examples of this in science, engineering and economics, for example, optimizing some utility function under budget constraints.

Lagrange multipliers problem:
Minimize (or maximize) $w = f(x, y, z)$ constrained by $g(x, y, z) = c$.

Lagrange multipliers solution:
Local minima (or maxima) must occur at a *critical point*. This is a point where $\nabla f = \lambda \nabla g$, and $g(x, y, z) = c$.

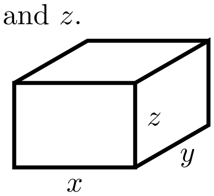
Example: Making a box using a minimum amount of material.

A box is made of cardboard with double thick sides, a triple thick bottom, single thick front and back and no top. It's volume is fixed at 3.

What dimensions use the least amount of cardboard?

Answer: We did this problem once before by solving for z in terms of x and y and substituting for it. That led to an unconstrained optimization problem in x and y. Here we will do it as a constrained problem. It is important to be able to do this because eliminating one variable is not always easy.

The box shown has dimensions x, y, and z.



The area of one side $= yz$. There are two double thick sides $\Rightarrow$ cardboard used $= 4yz$.

The area of the front (and back) $= xz$. It is single thick $\Rightarrow$ cardboard used $= 2xz$.

The area of the bottom $= xy$. It is triple thick $\Rightarrow$ cardboard used $= 3xy$.

Thus, the total cardboard used is

$$w = f(x, y, z) = 4yz + 2xz + 3xy.$$

The fixed volume acts as the constraint. It forces a relation between x, y and z so they can't all be varied independently. The constraint is

$$V = xyz = 3.$$

Our first job is to set up the equations to look for critical points. $\nabla f = \langle 2z + 3y, 4z + 3x, 4y + 2x\rangle$ and $\nabla V = \langle yz, xz, xy\rangle$.

The Lagrange multiplier equations are then

$$\begin{aligned} & \nabla f = \lambda \nabla V, \text{ and } V = 3 \\ \Leftrightarrow \quad & \langle 2z + 3y,\ 4z + 3x,\ 4y + 2x\rangle = \lambda\langle yz,\ xz,\ xy\rangle, \quad xyz = 3 \end{aligned}$$

Next we solve these equations for critical points. We do this by solving for λ in each equation (we call this *solving symmetrically*).

$\frac{2z+3y}{yz} = \lambda \quad \frac{4z+3x}{xz} = \lambda, \quad \frac{4y+2x}{xy} = \lambda, \quad xyz = 3 \quad \Rightarrow \frac{2}{y} + \frac{3}{z} = \frac{4}{x} + \frac{3}{z} = \frac{4}{x} + \frac{2}{y}$

$\Rightarrow \frac{2}{y} = \frac{4}{x} \Rightarrow x = 2y$ and $\frac{3}{z} = \frac{2}{y} \Rightarrow z = \frac{3}{2}y$

Now, $xyz = 3 \Rightarrow 3y^3 = 3 \Rightarrow y = 1$

Answer: $x = 2, \quad y = 1, \quad z = \frac{3}{2}, \quad w = 18.$

Sphere example:

Minimize $w = y$ constrained to $x^2 + y^2 + z^2 = 1$.

Answer: $\nabla f = \langle 0, 1, 0\rangle$, $\nabla g = \langle 2x, 2y, 2z\rangle$

$\nabla f = \lambda \nabla g \Rightarrow \langle 0, 1, 0\rangle = \lambda\langle 2x, 2y, 2z\rangle \Rightarrow x = z = 0.$

Constraint $\Rightarrow y = \pm 1$. (Gives the minimum and maximum respectively).

Example: (**checking the boundary**)

A rectangle in the plane is placed in the first quadrant so that one corner O is at the origin and the two sides adjacent to O are on the axes. The corner P opposite O is on the curve $x + 2y = 1$. Using Lagrange multipliers find for which point P the rectangle has maximum area. Say how you know this point gives the maximum.

Answer: We need some names

$g(x, y) = x + 2y = 1 =$ the constraint and $f(x, y) = xy =$ the area.

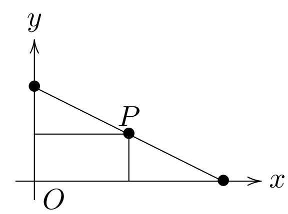
The gradients are: $\nabla g = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}, \quad \nabla f = y\,\hat{\mathbf{i}} + x\,\hat{\mathbf{j}}$.

Lagrange multipliers: $\Rightarrow y = \lambda, \quad x = 2\lambda, \quad x + 2y = 1.$

The first two equations $\Rightarrow x = 2y$;

Combine this with the third equation $\Rightarrow 4y = 1$.

$\Rightarrow y = 1/4, \quad x = 1/2 \Rightarrow P = (1/2, 1/4).$



We know this is a maximum because the maximum occurs either at a critical point or on the boundary. In this case, the boundary points are on the axes at (1,0) and (0,1/2), which gives a rectangle with area = 0.

Example: (**boundary at ∞**)

A rectangle in the plane is placed in the first quadrant so that one corner O is at the origin and the two sides adjacent to O are on the axes. The corner P opposite O is on the curve $xy = 1$. Using Lagrange multipliers find for which point P the rectangle has minimum perimeter. Say how you know this point gives the minimum.

Answer: Let $g(x, y) = xy = 1 =$ the constraint and $f(x, y) = 2x + 2y =$ the perimeter.

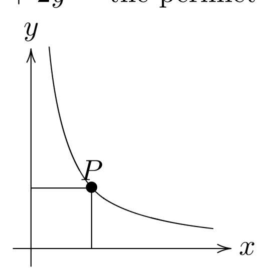
Gradients: $\nabla g = y\,\hat{\mathbf{i}} + x\,\hat{\mathbf{j}}, \quad \nabla f = 2\hat{\mathbf{i}} + 2\hat{\mathbf{j}}$.

Lagrange multipliers: $\Rightarrow$
$$2 = \lambda y$$
$$2 = \lambda x$$
$$xy = 1$$

The first two equations $\Rightarrow x = y$;

Combine this with the third equation $\Rightarrow x^2 = 1$.

$\Rightarrow x = 1, \quad x = 1 \Rightarrow P = (1, 1).$



We know this is a minimum because the minimum occurs either at a critical point or on the boundary. In this case the boundary points are infinitely far out on the axes which gives a rectangle with perimeter = ∞.

Proof of Lagrange Multipliers

Here we will give two arguments, one geometric and one analytic for why Lagrange multipliers work.

Critical points
For the function $w = f(x, y, z)$ constrained by $g(x, y, z) = c$ (c a constant) the critical points are defined as those points, which satisfy the constraint and where ∇f is parallel to ∇g. In equations:

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \quad \textit{and} \quad g(x, y, z) = c.$$

Statement of Lagrange multipliers
For the constrained system local maxima and minima (collectively extrema) occur at the critical points.

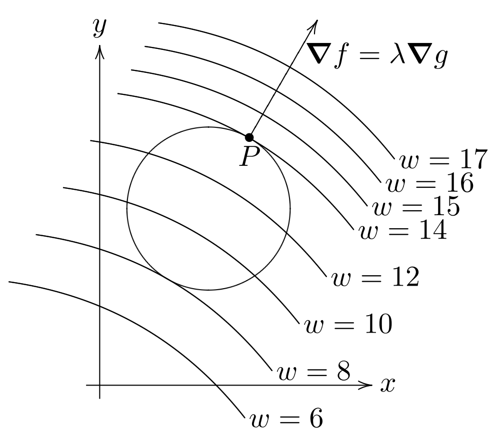
Geometric proof for Lagrange
(We only consider the two dimensional case, $w = f(x, y)$ with constraint $g(x, y) = c$.)

For concreteness, we've drawn the constraint curve, $g(x, y) = c$, as a circle and some level curves for $w = f(x, y) = c$ with explicit (made up) values. Geometrically, we are looking for the point on the circle where w takes its maximum or minimum values.

Now, start at the level curve with $w = 17$, which has no points on the circle. So, clearly, the maximum value of w on the constraint circle is less than 17. Move down the level curves until they first touch the circle when $w = 14$. Call the point where the first touch P. It is clear that P gives a local maximum for w on $g = c$, because if you move away from P in either direction on the circle you'll be on a level curve with a smaller value.

Since the circle is a level curve for g, we know ∇g is perpendicular to it. We also know ∇f is perpendicular to the level curve $w = 14$, since the curves themselves are tangent, these two gradients must be parallel.

Likewise, if you keep moving down the level curves, the last one to touch the circle will give a local minimum and the same argument will apply.



Analytic proof for Lagrange (in three dimensions)

Suppose f has a local maximum at P on the constraint surface.

Let $\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$ be an arbitrary parametrized curve which lies on the constraint surface and has $(x(0), y(0), z(0)) = P$. Finally, let $h(t) = f(x(t), y(t), z(t))$. The setup guarantees that $h(t)$ has a maximum at $t = 0$.

Taking a derivative using the chain rule in vector form gives

$$h'(t) = \nabla f|_{\mathbf{r}(t)} \cdot \mathbf{r}'(t).$$

Since $t = 0$ is a local maximum, we have

$$h'(0) = \nabla f|_P \cdot \mathbf{r}'(0) = 0.$$

Thus, $\nabla f|_P$ is perpendicular to any curve on the constraint surface through P.

This implies $\nabla f|_P$ is perpendicular to the surface. Since $\nabla g|_P$ is also perpendicular to the surface we have proved $\nabla f|_P$ is parallel to $\nabla g|_P$. QED

Non-independent Variables

1. We give a worked example here. A fuller explanation will be given in the next session. Let

$$w = x^3y^2 + x^2y^3 + y$$

and assume x and y satisfy the relation

$$x^2 + y^2 = 1.$$

We consider x to be the independent variable, then, because y depends on x we have w is ultimately a function of the single variable x.

a) Compute $\dfrac{dw}{dx}$ using implicit differentiation.

b) Compute $\dfrac{dw}{dx}$ using total differentials.

<u>**Answer:**</u>

a) Implicit differentiation means remembering that y is a function of x, e.g., $\dfrac{dy^2}{dx} = 2y\dfrac{dy}{dx}$.
Thus,

$$\frac{dw}{dx} = 3x^2y^2 + 2x^3y\frac{dy}{dx} + 2xy^3 + 3x^2y^2\frac{dy}{dx} + \frac{dy}{dx}.$$

Now we differentiate the constraint to find $\dfrac{dy}{dx}$.

$$x^2 + y^2 = 1 \;\Rightarrow\; 2x + 2y\frac{dy}{dx} = 0 \;\Rightarrow\; \frac{dy}{dx} = -\frac{x}{y}.$$

Substituting this in the equation for $\dfrac{dw}{dx}$ gives

$$\frac{dw}{dx} = 3x^2y^2 - 2x^3y\frac{x}{y} + 2xy^3 - 3x^2y^2\frac{x}{y} - \frac{x}{y} = 3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y}.$$

b) Taking total differentials of both w and the constraint equation gives

$$\begin{aligned} dw &= 3x^2y^2\,dx + 2x^3y\,dy + 2xy^3\,dx + 3x^2y^2\,dy + dy \\ &= (3x^2y^2 + 2xy^3)\,dx + (2x^3y + 3x^2y^2 + 1)\,dy \end{aligned}$$

$$2x\,dx + 2y\,dy = 0.$$

We can solve the second equation for dy and substitute in the equation for dw.

$$\begin{aligned} dy &= -\frac{x}{y}\,dx \;\Rightarrow \\ dw &= (3x^2y^2 + 2xy^3)\,dx + (2x^3y + 3x^2y^2 + 1)\left(-\frac{x}{y}\right)\,dx \\ &= (3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y})\,dx \end{aligned}$$

Thus,

$$\frac{dw}{dx} = 3x^2y^2 - 2x^4 + 2xy^3 - 3x^3y - \frac{x}{y}.$$

18.02SC Multivariable Calculus
Fall 2010

Non-independent Variables

1. Partial differentiation with non-independent variables.

Up to now in calculating partial derivatives of functions like $w = f(x, y)$ or $w = f(x, y, z)$, we have assumed the variables x, y (or x, y, z) were independent. However in real-world applications this is frequently not so. Computing partial derivatives then becomes confusing, but it is better to face these complications now while you are still in a calculus course, than wait to be hit with them at the same time that you are struggling to cope with the thermodynamics or economics or whatever else is involved.

For example, in thermodynamics, three variables that are associated with a contained gas are its

$$p = \text{pressure}, \qquad v = \text{volume}, \qquad T = \text{temperature},$$

and you can express other thermodynamic variables like the internal energy U and entropy S in terms of p, v, and T.

However, p, v, and T are not independent variables. If the gas is a so-called "ideal gas", they are related by the equation

$$pv = nRT \qquad (n, R \text{ constants}). \tag{1}$$

To see what complications this produces, let's consider first a purely mathematical example.

Example 1. Let $w = x^2 + y^2 + z^2$, where $z = x^2 + y^2$. Calculate $\dfrac{\partial w}{\partial x}$.

Discussion.

(a) If we think of x and y as the independent variables, then we can calculate $\dfrac{\partial w}{\partial x}$ by two different methods:

(i) using $z = x^2 + y^2$ to get rid of z, we get

$$\begin{aligned} w &= x^2 + y^2 + (x^2 + y^2)^2 \\ &= x^2 + y^2 + x^4 + 2x^2y^2 + y^4; \\ \frac{\partial w}{\partial x} &= 2x + 4x^3 + 4xy^2 \end{aligned}$$

(ii) or by using the chain rule, remembering z is a function of x and y,

$$\begin{aligned} w &= x^2 + y^2 + z^2 \\ \frac{\partial w}{\partial x} &= 2x + 2z\frac{\partial z}{\partial x} = 2x + 2z \cdot 2x \\ &= 2x + 2(x^2 + y^2) \cdot 2x, \end{aligned}$$

so the two methods agree.

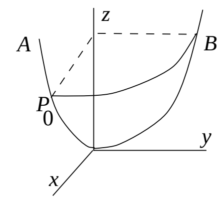
(b) On the other hand, if we think of x and z as the independent variables, using say method (i) above, we get rid of y by using the relation $y^2 = z - x^2$, and get

$$\begin{aligned} w &= x^2 + y^2 + z^2 = x^2 + (z - x^2) + z^2 \\ &= z + z^2; \\ \frac{\partial w}{\partial x} &= 0. \end{aligned}$$

These answers are genuinely different — we cannot convert one into the other by using the relation $z = x^2 + y^2$. Will the right $\partial w/\partial x$ please stand up?

The answer is, that there *is* no one right answer, because the problem was not well-stated. *When the variables are not independent, an expression like $\partial w/\partial x$ has no definite meaning.*

To see why this is so, we interpret the above example geometrically. Saying that x, y, z satisfy the relation $z = x^2 + y^2$ means that the point (x, y, z) lies on the paraboloid surface formed by rotating $z = y^2$ about the z-axis. The function



$$w \;=\; x^2 + y^2 + z^2$$

measures the square of the distance from the origin. To be definite, let's suppose we are at the starting point $P = P_0 : (1, 0, 1)$ indicated, and we want to calculate $\partial w/\partial x$ at this point.

Case (a) If we take x and y to be the independent variables, then to find $\partial w/\partial x$, we *hold y fixed and let x vary.* So P moves in the xz-plane towards A, along the path shown.

As P moves along this path, evidently w, the square of its distance from the origin, is steadily increasing: $\dfrac{\partial w}{\partial x} > 0$ and in fact the calculations for **(a)** on the previous page show that $\dfrac{\partial w}{\partial x} = 6$.

Case (b) If we take x and z to be the independent variables, then to find $\partial w/\partial x$, we *hold z fixed and let x vary.* Now P moves in the plane $z = 1$, along the circular path towards B.

As P moves on this path, the square of its distance from the origin is *not changing*, and therefore $\dfrac{\partial w}{\partial x} = 0$, as we calculated in **(b)** before.

To sum up, the value of $\partial w/\partial x$ depends on which variables we take to be independent, because we are measuring different rates of change, as P moves along different paths.

There is only one way out of our difficulty. When we ask for $\partial w/\partial x$, we must at the same time specify which variables are to be taken as the independent ones. This is done by using the following notation:

Case **(a)**: x, y are the independent variables: $\left(\dfrac{\partial w}{\partial x}\right)_y$

Case **(b)**: x, z are the independent variables: $\left(\dfrac{\partial w}{\partial x}\right)_z$

These are read, "the partial of w with respect to x, with y (resp. z) held constant".

Note how in each case the two lower letters give you the two independent variables. If we had more variables, we would use a similar notation. For instance if

$$w \;=\; f(x, y, z, t), \qquad \text{where } \; xy = zt, \tag{2}$$

then only three of the variables x, y, z, t can be independent; the fourth is then determined by the equation on the right of (2). Thus we would write expressions like

$$\left(\frac{\partial w}{\partial x}\right)_{y,t} \qquad \text{“partial of } w \text{ with respect to } x;\ y \text{ and } t \text{ held constant”;}$$

$$\left(\frac{\partial w}{\partial y}\right)_{x,z} \qquad \text{“partial of } w \text{ with respect to } y;\ x \text{ and } z \text{ held constant”;}$$

in the first, x, y, t are the independent variables; in the second, x, y, z are independent.

2. Differentials vs. Chain Rule

An alternative way of calculating partial derivatives uses total differentials. We illustrate with an example, doing it first with the chain rule, then repeating it using differentials. By definition, the differential of a function of several variables, such as $w = f(x, y, z)$ is

$$dw = f_x dx + f_y dy + f_z dz, \tag{3}$$

where the three partial derivatives $f_x,\ f_y,\ f_z$ are the *formal* partial derivatives, i.e., the derivatives calculated as if x, y, z were independent.

Example 2. Find $\left(\frac{\partial w}{\partial y}\right)_{x,t}$, where $w = x^3y - z^2t$ and $xy = zt$.

Solution 1. Using the chain rule and the two equations in the problem, we have

$$\left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zt\left(\frac{\partial z}{\partial y}\right)_{x,t} = x^3 - 2zt\,\frac{x}{t} = x^3 - 2zx.$$

Solution 2. We take the differentials of both sides of the two equations in the problem:

$$dw = 3x^2y\,dx + x^3dy - 2zt\,dz - z^2dt, \qquad y\,dx + x\,dy = z\,dt + t\,dz. \tag{4}$$

Since the problem indicates that x, y, t are the independent variables, we eliminate dz from the equations in (4) by multiplying the second equation by $2z$, adding it to the first, then grouping the terms, which gives

$$dw = (3x^2y - 2zy)\,dx + (x^3 - 2zx)\,dy + z^2dt$$

Comparing this with (3) — after replacing z by t in (3) — we see that

$$\left(\frac{\partial w}{\partial x}\right)_{y,t} = 3x^2y - 2zy, \qquad \left(\frac{\partial w}{\partial y}\right)_{x,t} = x^3 - 2zx, \qquad \left(\frac{\partial w}{\partial t}\right)_{x,y} = z^2\ .$$

(The actual partial derivatives are the same as the formal partial derivatives w_x, w_y, w_t because x, y, t are independent variables.)

Notice that the differential method here takes a bit more calculation, but gives us three derivatives, not just one; this is fine if you want all three, but a little wasteful if you don't. The main thing to keep in mind for the method is that differentials are treated like vectors, with the $dx,\ dy,\ dz, \ldots$ playing the role of $\mathbf{i}\,,\mathbf{j}\,,\mathbf{k}\,,\ldots.$ That is:

D1. *Differentials can be added, subtracted, and multiplied by scalar functions;*

D2. *If the variables $x, y, \ldots$ are independent, two differentials are equal if and only if their corresponding coefficients are equal:*

$$A\,dx + B\,dy + \ldots = A_1\,dx + B_1\,dy + \ldots \qquad \Leftrightarrow \qquad A = A_1,\ B = B_1, \ldots; \tag{5}$$

D3. *One differential can be substituted into another.*

Remarks.

1. In Example 2, Solution 2, we used the operations in **D1** to do the calculations. We used **D2** in the last step, taking advantage of the fact that the x, y, t were independent.

We could have done the calculations using **D3** instead, by solving the second equation in (4) for dz and substituting it into the first equation. **D3** is a consequence of the chain rule. Illustrations of its use will be given in the next section.

2. The main advantage of calculating with differentials is that one need not take into account whether the variables are dependent or not, or which variables depend on which others; the method does this automatically for you. Examples will illustrate.

3. If the variables are not independent, **D2** is emphatically *not* true; the second equation in (4) gives a counterexample.

Note also that in **D1**, there is no attempt to include a "multiplication" or "division" of differentials to the list of operations. If u and v are functions of several variables, then their "product" $du\,dv$ makes no sense as a differential, nor does their "quotient" du/dv, which despite appearances is not in general related to any derivative, or function, or even defined. (There is no elementary analogue of the dot and cross product of vectors, though in advanced differential geometry courses a certain type of product for differentials is defined and used for multiple integration.)

Example 3. Let $w = x^2 - yz + t^2$, where x, y, z, t satisfy the two equations

$$z^2 = x + y^2 \qquad \text{and} \qquad xy = zt.$$

Using these equations, we can express first z and then t in terms of x and y; this means that w can also be expressed in terms of x and y. Without actually calculating $w(x, y)$ explicitly, find its gradient vector $\nabla w(x, y)$.

Solution. Since we need both partial derivatives $(\partial w/\partial x)_y$ and $(\partial w/\partial y)_x$, it makes sense to use the differential method. Taking the differential of w and of the two equations connecting the variables gives us

$$dw = 2x\,dx - z\,dy - y\,dz + 2t\,dt, \qquad x\,dy + y\,dx = z\,dt + t\,dz, \qquad 2z\,dz = dx + 2y\,dy. \tag{6}$$

We want x and y to be the independent variables; using the operations in **D1**, first eliminate dt by solving for it in the second equation, and substituting for it into the first equation; then eliminate dz by solving for it in the last equation and substituting into the first equation; the result is

$$dw = \left(2x - \frac{y}{2z} + \frac{2ty}{z} - \frac{t^2}{z^2}\right) dx + \left(-z - \frac{y^2}{z} + \frac{2xt}{z} - \frac{2t^2 y}{z^2}\right) dy. \tag{7}$$

Since x and y are independent, comparing the two expressions for dw in (7) and (3) (using x and y), and then using **D2**, shows that the two coefficients in (7) are respectively the two partial derivatives w_x and w_y, i.e., the two components of the gradient ∇w.

Example 4. Suppose the variables x, y, z satisfy an equation $g(x, y, z) = 0$. Assume the point $P : (1, 1, 1)$ lies on the surface $g = 0$ and that $(\nabla g)_P = \langle -1, 1, 2 \rangle$.

Let $f(x, y, z)$ be another function, and assume that $(\nabla f)_P = \langle 1, 2, 1 \rangle$.

Find the gradient of the function $w = f(x, y, z(x, y))$ of the two independent variables x and y, at the point $x = 1,\ y = 1$.

Solution. Using differentials, we have, by (3) and our hypotheses,

$$(dw)_P = dx + 2dy + dz; \qquad (dg)_P = -dx + dy + 2dz = 0, \quad \text{since } dg = 0 \text{ for all } x, y, z;$$

eliminating dz by solving the second equation for it and substituting into the first, or by dividing the second equation by 2 and substracting it from the first, we get

$$(dw)_P = \tfrac{3}{2}dx + \tfrac{3}{2}dy; \qquad\qquad (\nabla w)_P = \tfrac{3}{2}\,\mathbf{i} + \tfrac{3}{2}\,\mathbf{j}\,.$$

Non-independent Variables

3. Abstract partial differentiation; rules relating partial derivatives

Often in applications, the function w is not given explicitly, nor are the equations connecting the variables. Thus you need to be able to work with functions and equations just given abstractly. The previous ideas work perfectly well, as we will illustrate. However, we will need (as in section 2) to distinguish between

formal partial derivatives, written here $f_x,\ f_y, \ldots$ (calculated as if all the variables were independent), and

actual partial derivatives, written $\partial f/\partial x, \ldots$, which take account of any relations between the variables.

Example 5. If $f(x,y,z) = xy^2z^4$, where $z = 2x + 3y$, the three formal derivatives are

$$f_x = y^2z^4, \qquad f_y = 2xyz^4, \qquad f_z = 4xy^2z^3,$$

while three of the many possible actual partial derivatives are (we use the chain rule)

$$\begin{aligned}
\left(\frac{\partial f}{\partial x}\right)_y &= f_x + f_z\left(\frac{\partial z}{\partial x}\right)_y = y^2z^4 + 8xy^2z^3;\\
\left(\frac{\partial f}{\partial y}\right)_x &= f_y + f_z\left(\frac{\partial z}{\partial y}\right)_x = 2xyz^4 + 12xy^2z^3;\\
\left(\frac{\partial f}{\partial z}\right)_x &= f_y\left(\frac{\partial y}{\partial z}\right)_x + f_z = \tfrac{2}{3}xyz^4 + 4xy^2z^3.
\end{aligned}$$

Rules connecting partial derivatives. These rules are widely used in the applications, especially in thermodynamics. Here we will use them as an excuse for further practice with the chain rule and differentials.

With an eye to thermodynamics, we assume a set of variables $t, u, v, w, x, y, z, \ldots$ connected by several equations in such a way that

- any *two* are independent;
- any *three* are connected by an equation.

Thus, one can choose any two of them to be the independent variables, and then each of the other variables can be expressed in terms of these two.

We give each rule in two forms—the second form is the one ordinarily used, while the first is easier to remember. (The first two rules are fairly simple in either form.)

(8a,b) $$\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial x}\right)_z = 1 \qquad \left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{(\partial y/\partial x)_z}$$ **reciprocal rule**

(9a,b) $$\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial t}\right)_z = \left(\frac{\partial x}{\partial t}\right)_z \qquad \left(\frac{\partial x}{\partial y}\right)_z = \frac{(\partial x/\partial t)_z}{(\partial y/\partial t)_z},$$ **chain rule**

(10a,b) $$\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y = -1 \qquad \left(\frac{\partial x}{\partial y}\right)_z = -\frac{(\partial x/\partial z)_y}{(\partial y/\partial z)_x},$$ **cyclic rule**

Note how the successive factors in the cyclic rule are formed: the variables are used in the successive orders $x, y, z;\ \ y, z, x;\ \ z, x, y;$ one says they are permuted cyclically, and this explains the name.

Proof of the rules. The first two rules are simple: since z is being held fixed throughout, each variable becomes a function of just one other variable, and (9) is just the one-variable chain rule. Then (8) is just the special case of (9) where $x = t$.

The cyclic rule is less obvious — on the right side it looks almost like the chain rule, but different variables are being held constant in each of the differentiations, and this changes it entirely. To prove it, we suppose $f(x, y, z) = 0$ is the equation satisfied by x, y, z; taking y and z as the independent variables and differentiating $f(x, y, z) = 0$ with respect to y gives:

$$f_x \left(\frac{\partial x}{\partial y}\right)_z + f_y = 0; \qquad \text{therefore} \quad \left(\frac{\partial x}{\partial y}\right)_z = -\frac{f_y}{f_x} . \tag{11}$$

Permuting the variables in (11) and multiplying the resulting three equations gives (10a):

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -\frac{f_y}{f_x} \cdot -\frac{f_z}{f_y} \cdot -\frac{f_x}{f_z} = -1.$$

Example 6. Suppose $w = w(x, r)$, with $r = r(x, \theta)$. Give an expression for $\left(\frac{\partial w}{\partial r}\right)_\theta$ in terms of formal partial derivatives of w and r.

Solution. Evidently the independent variables are to be r and θ, since these are the ones that occur in the lower part of the partial derivative, with x dependent on them. Since θ is viewed as a constant, the chain rule gives

$$\left(\frac{\partial w}{\partial r}\right)_\theta = w_x \left(\frac{\partial x}{\partial r}\right)_\theta + w_r ;$$
$$\left(\frac{\partial x}{\partial r}\right)_\theta = \frac{1}{(\partial r/\partial x)_\theta},$$

by the reciprocal rule (8). and therefore finally,

$$\left(\frac{\partial w}{\partial r}\right)_\theta = \frac{w_x}{r_x} + w_r .$$

Definition of double integration

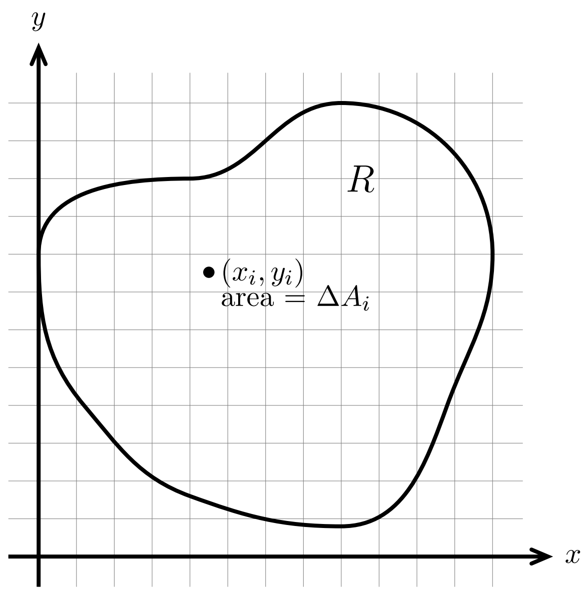
In this note we will work abstractly, defining double integration as a sum, technically a limit of Riemann sums. It is best to learn this first before getting into the details of computing the value of a double integral –we will learn how to do that next.

Definition of double integrals
Suppose we have a region in the plane R and a function $f(x, y)$, Then the double integral

$$\int\int_R f(x, y)\, dA$$

is defined as follows.

Divide the region R into small pieces, numbered from 1 to n. Let ΔA_i be the area of the i^{th} piece and also pick a point (x_i, y_i) in that piece. The figure shows a region R divided into small pieces and shows the i^{th} piece with its area, and choice of a point in the little region.



Now form the sum

$$\sum_{i=1}^{n} f(x_i, y_i)\, \Delta A_i,$$

and then, finally

$$\int\int_R f(x, y)\, dA = \lim_{\Delta A \to 0} \sum_{i=1}^{n} f(x_i, y_i)\, \Delta A_i.$$

Here, the limit is taken by letting the number of pieces go to infinity and the area of each piece go to 0. There are technical requirements that the limit exist and be independent of the specific limiting process. In 18.02 these requirements are always met. (Later you might study fractals and other strange objects which don't satisfy them.)

Interpretations of the double integral

As you saw in single variable calculus, these sums can be used to compute areas, volumes, mass, work, moment of inertia and many other quantities. Again, before focusung on some computational issues we will show you how easy it is to setup a double integral to compute certain quantities.

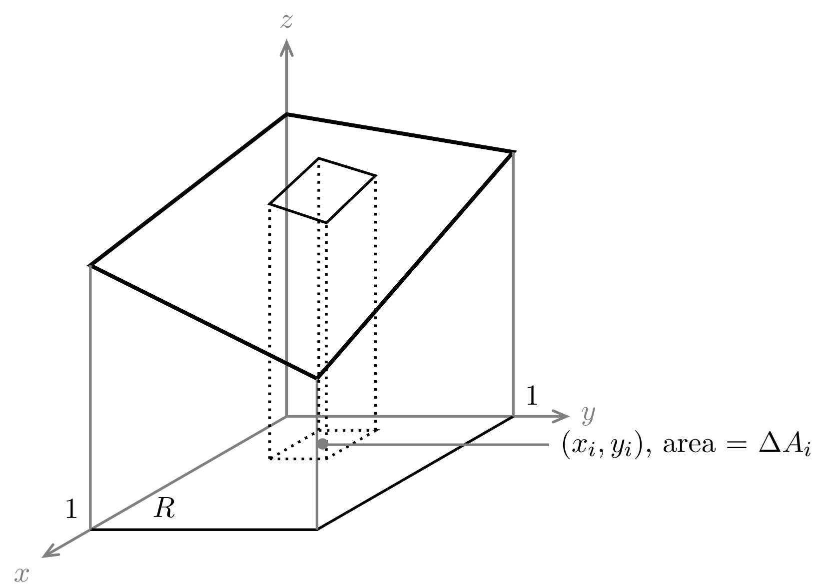
Example 1: Set up a double integral to compute the area of a region R in the plane.

<u>**Answer:**</u> Use the figure above for visualization. The area of R is just the sum of the areas of the pieces. That is,

$$\text{area } = \int\int_R dA.$$

Example 2: Set up a double integral to compute the volume of the solid below the graph of $z = f(x, y) = 2 - .5(x + y)$ and above the unit square in the xy-plane.

<u>**Answer:**</u> The figure below shows the graph of $f(x, y)$ above the unit square in the plane. The unit square is labeled R. We also show a little piece of the R and the solid region above that piece. We are imagining we've divided R into n small pieces and this is the i^{th} one. It contains the point (x_i, y_i) and has area ΔA_i.



The small solid region is almost a box and so its volume, ΔV_i, is roughly its base times its height, i.e.,

$$\Delta V_i \approx \Delta A_i \times f(x_i, y_i).$$

The total volume is the sum of the volumes of all the small pieces, i.e.,

$$\text{volume } = \sum_{i=1}^{n} \Delta V_i \approx \sum_{i=1}^{n} \Delta A_i \times f(x_i, y_i).$$

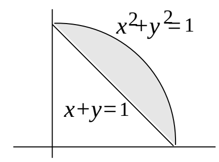
In the limit this becomes an exact integral for volume

$$\text{volume } = \int\int_R f(x, y)\, dA.$$

Limits in Iterated Integrals

For most students, the trickiest part of evaluating multiple integrals by iteration is to put in the limits of integration. Fortunately, a fairly uniform procedure is available which works in any coordinate system. *You must always begin by sketching the region; in what follows we'll assume you've done this.*

1. Double integrals in rectangular coordinates.



Let's illustrate this procedure on the first case that's usually taken up: double integrals in rectangular coordinates. Suppose we want to evaluate over the region R pictured the integral

$$\iint_R f(x,y)\,dy\,dx \;, \qquad\qquad R = \quad \text{region between } x^2+y^2=1 \quad \text{and} \quad x+y=1\;;$$

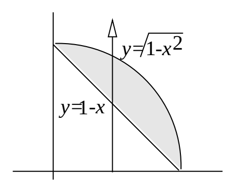
we are integrating first with respect to y. Then to put in the limits,

1. Hold x fixed, and let y increase (since we are integrating with respect to y). As the point (x,y) moves, it traces out a vertical line.

2. Integrate from the y-value where this vertical line enters the region R, to the y-value where it leaves R.

3. Then let x increase, integrating from the lowest x-value for which the vertical line intersects R, to the highest such x-value.

Carrying out this program for the region R pictured, the vertical line enters R where $y = 1-x$, and leaves where $y = \sqrt{1-x^2}$.



The vertical lines which intersect R are those between $x = 0$ and $x = 1$. Thus we get for the limits:

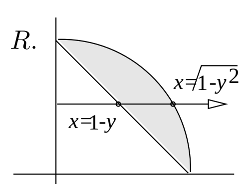
$$\iint_R f(x,y)\,dy\,dx \;=\; \int_0^1 \int_{1-x}^{\sqrt{1-x^2}} f(x,y)\,dy\,dx.$$

To calculate the double integral, integrating in the reverse order $\iint_R f(x,y)\,dx\,dy$,

1. Hold y fixed, let x increase (since we are integrating first with respect to x). This traces out a horizontal line.

2. Integrate from the x-value where the horizontal line enters R to the x-value where it leaves.

3. Choose the y-limits to include all of the horizontal lines which intersect R.



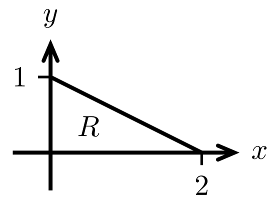
Following this prescription with our integral we get:

$$\iint_R f(x,y)\,dx\,dy \;=\; \int_0^1 \int_{1-y}^{\sqrt{1-y^2}} f(x,y)\,dx\,dy.$$

Limits for double integrals

1. Evaluate $\iint_R x\,dA$, where R is the finite region bounded by the axes and $2y + x = 2$.

<u>Answer:</u>
First we sketch the region.



Next, we find limits of integration. By using vertical stripes we get limits
Inner: y goes from 0 to $1 - x/2$; outer: x goes from 0 to 2.
Thus the integral is

$$\int_0^2 \int_0^{1-x/2} x\,dy\,dx$$

Finally, we compute the inner, then the outer integrals.
Inner: $xy|_0^{1-x/2} = x - \dfrac{x^2}{2}$.
Outer: $\dfrac{x^2}{2} - \dfrac{x^3}{6}\Big|_0^2 = \dfrac{2}{3}$.

Changing the order of integration

1. Evaluate

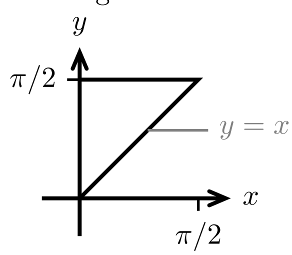
$$I = \int_0^{\pi/2} \int_x^{\pi/2} \frac{\sin y}{y} \, dy \, dx$$

by changing the order of integration.

__**Answer:**__

The given limits are (inner) y from x to $\pi/2$; (outer) x from 0 to $\pi/2$.

We use these to sketch the region of integration.



The given limits have inner variable y. To reverse the order of integration we use horizontal stripes. The limits in this order are

(inner) x from 0 to y; (outer) y from 0 to $\pi/2$.

So the integral becomes

$$I = \int_0^{\pi/2} \int_0^{y} \frac{\sin y}{y} \, dx \, dy$$

We compute the inner, then the outer integrals.

Inner: $\left. \frac{\sin y}{y} x \right|_0^y = \sin y$. Outer: $-\cos y|_0^{\pi/2} = 1$.

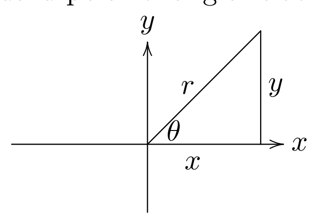
Integration in polar coordinates

Polar Coordinates

Polar coordinates are a different way of describing points in the plane. The polar coordinates (r, θ) are related to the usual *rectangular coordinates* (x, y) by by

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

The figure below shows the standard polar triangle relating x, y, r and θ.



Because cos and sin are periodic, different (r, θ) can represent the same point in the plane. The table below shows this for a few points.

(x, y)	$(1, 0)$	$(0, 1)$	$(2, 0)$	$(1, 1)$	$(-1, 1)$	$(-1, -1)$	$(0, 0)$
(r, θ)	$(1, 0)$	$(1, \pi/2)$	$(2, 0)$	$(\sqrt{2}, \pi/4)$	$(\sqrt{2}, 3\pi/4)$	$(\sqrt{2}, 5\pi/4)$	$(0, \pi/2)$
(r, θ)	$(1, 2\pi)$			$(\sqrt{2}, 9\pi/4)$		$(-\sqrt{2}, \pi/4)$	$(0, -7.2)$
(r, θ)	$(1, 4\pi)$						

In fact, you can add any multiple of 2π to θ and the polar coordinates will still represent the same point.

Because θ is not uniquely specified it's a little trickier going from rectangular to polar coordinates. The equations are easily deduced from the standard polar triangle.

$$r = \sqrt{x^2 + y^2}, \qquad \text{"}\theta = \tan^{-1}(y/x)\text{"}.$$

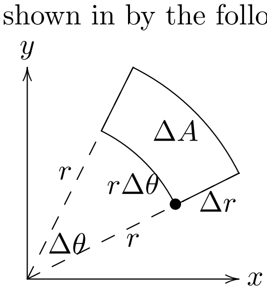
We use quotes around $\tan^{-1}$ to indicate it is not a single valued function.

The area element in polar coordinates

In polar coordinates the area element is given by

$$dA = r\,dr\,d\theta.$$

The geometric justification for this is shown in by the following figure.

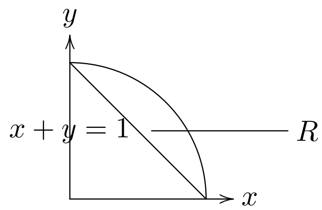


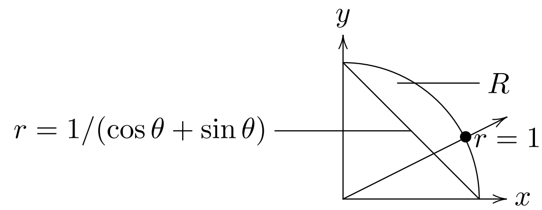
The small curvy rectangle has sides Δr and $r\Delta\theta$, thus its area satisfies $\Delta A \approx (\Delta r)(r\,\Delta\theta)$. As usual, in the limit this becomes $dA = r\,dr\,d\theta$.

Double integrals in polar coordinates

The area element is one piece of a double integral, the other piece is the limits of integration which describe the region being integrated over.

Finding procedure for finding the limits in polar coordinates is the same as for rectangular coordinates. Suppose we want to evaluate $\int\int_R dr\,d\theta$ over the region R shown.





(The integrand, including the r that usually goes with $r\,dr\,d\theta$, is irrelevant here, and therefore omitted.)
As usual, we integrate first with respect to r. Therefore, we

1. Hold θ fixed, and let r increase (since we are integrating with respect to r). As the point moves, it traces out a ray going out from the origin.

2. Integrate from the r-value where the ray enters R to the r-value where it leaves. This gives the limits on r.

3. Integrate from the lowest value of θ for which the corresponding ray intersects R to the highest value of θ.

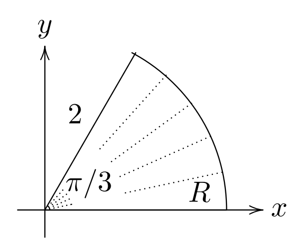
To follow this procedure, we need the equation of the line in polar coordinates. We have

$$x + y = 1 \quad \to \quad r\cos\theta + \mathbf{r}\sin\theta = 1, \quad \text{or} \quad r = \frac{1}{\cos\theta + \sin\theta}.$$

This is the r value where the ray enters the region; it leaves where $r = 1$. The rays which intersect R lie between $\theta = 0$ and $\theta = \pi/2$. Thus the double iterated integral in polar coordinates has the limits

$$\int_0^{\pi/2}\int_{1/(\cos\theta+\sin\theta)}^{1} dr\,d\theta.$$

Example: Find the mass of the region R shown if it has density $\delta(x,y) = xy$ (in units of mass/unit area)
In polar coordinates: $\delta = r^2\cos\theta\sin\theta$.
Limits of integration: (radial lines sweep out R):
inner (fix θ): $0 < r < 2$, outer: $0 < \theta < \pi/3$.



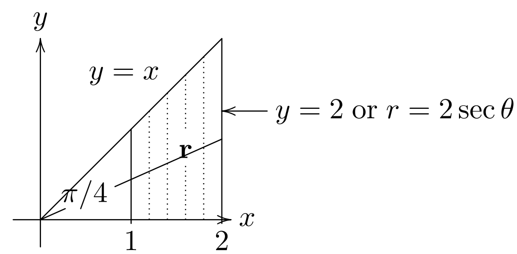
$\Rightarrow$ Mass $M = \int\int_R \delta(x,y)\,dA = \int_{\theta=0}^{\pi/3}\int_{r=0}^{2} r^2\cos\theta\sin\theta\, r\,d\theta\,dr$

Inner: $\int_0^2 r^3\cos\theta\sin\theta\,dr = \frac{r^4}{4}\cos\theta\sin\theta\Big|_0^2 = 4\cos\theta\sin\theta$

Outer: $M = \int_0^{\pi/3} 4\cos\theta\sin\theta\,d\theta = 2\sin^2\theta\big|_0^{\pi/3} = \frac{3}{2}.$

Example: Let $I = \int_1^2 \int_0^x \frac{1}{(x^2+y^2)^{3/2}}\, dy\, dx$. Compute I using polar coordinates.

<u>Answer:</u> Here are the steps we take.
Draw the region.



Find limits in polar coordinates:
Inner (fix θ): $\sec\theta < r < 2\sec\theta$, outer: $0 < \theta < \pi/4$.

$$\Rightarrow \; I = \int_{\theta=0}^{\pi/4} \int_{r=\sec\theta}^{2\sec\theta} \frac{1}{r^3} r\, dr\, d\theta.$$

Compute the integral:

Inner: $\displaystyle\int_{\sec\theta}^{2\sec\theta} \frac{1}{r^2}\, dr = -\frac{1}{r}\Big|_{\sec\theta}^{2\sec\theta} = \frac{1}{2}\cos\theta.$

Outer: $\displaystyle I = \int_0^{\pi/4} \frac{1}{2}\cos\theta\, d\theta = \frac{1}{2}\sin\theta\Big|_0^{\pi/4} = \frac{\sqrt{2}}{4}.$

Example: Find the volume of the region above the xy-plane and below the graph of $z = 1 - x^2 - y^2$.

You should draw a picture of this.

In polar coordinates we have $z = 1 - r^2$ and we want the volume under the graph and above the inside of the unit disk.

$$\Rightarrow \; \text{volume } V = \int_0^{2\pi} \int_0^1 (1 - r^2)\, r dr\, d\theta.$$

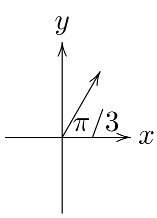
Inner integral: $\displaystyle\int_0^1 (1 - r^2)\, r dr = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$

Outer integral: $\displaystyle V = \int_0^{2\pi} \frac{1}{4}\, d\theta = \frac{\pi}{2}.$

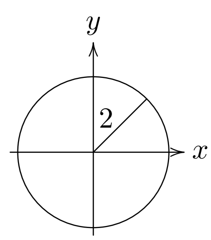
Gallery of polar graphs $(r = f(\theta))$

A point P is on the graph **if any** representation of P satisfies the equation.

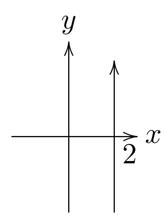
Examples:



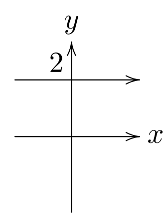
Ray: $\theta = \pi/3$



Circle centered on 0: $r = 2$

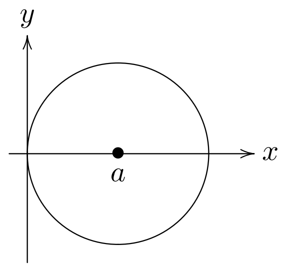


Vertical line $x = 2 \Leftrightarrow r = 2\sec\theta$.

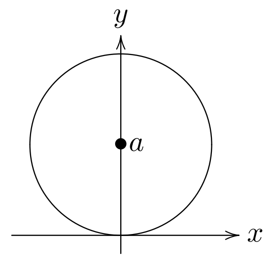


Horizontal line $y = 2 \Leftrightarrow r = 2/\sin\theta$.

Example: Show the graph of $r = 2a\cos\theta$ is a circle of radius a centered at $(a, 0)$.
Some simple algebra gives $r^2 = 2ar\cos\theta = 2ax \quad \Rightarrow \; x^2 + y^2 = 2ax \quad \Rightarrow \; (x-a)^2 + y^2 = a^2$.
This is a circle or radius a centered at $(a, 0)$.
Note: we can determine from the graph that the range of theta is $-\pi/2 \le \theta \le \pi/2$.



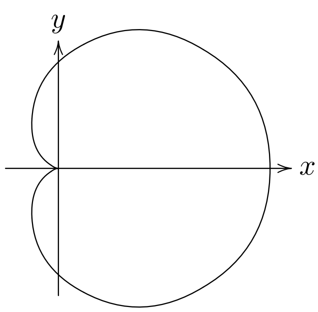
$r = 2a\cos\theta$
$-\pi/2 \le \theta \le \pi/2$.



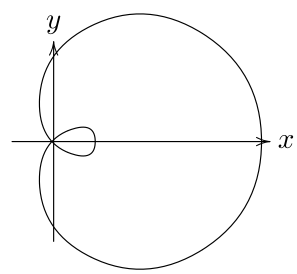
$r = 2a\sin\theta$
$0 \le \theta \le \pi$.

Warning: We can use negative values of r for plotting. You should never use it in integration. In integration it is better to make use of symmetry and only integrate over regions where r is positive.

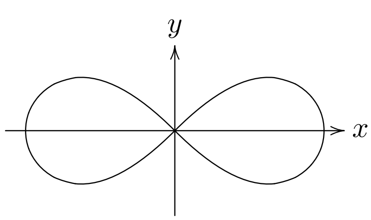
Here are a few more curves.



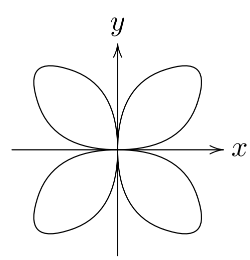
Cardiod: $r = a(1 + \cos\theta)$



Limaçon: $r = a(1 + b\cos\theta)$ $(b > 1)$



Lemniscate: $r^2 = 2a^2\cos 2\theta$

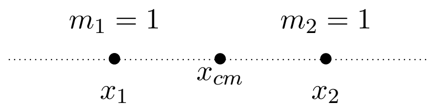


Four leaved rose: $r = a\sin 2\theta$

Mass and average value

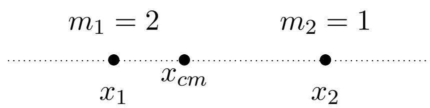
Center of Mass

Example 1: For two equal masses, the center of mass is at the midpoint between them.



$x_{cm} = \frac{x_1+x_2}{2}$.

Example 2: For unequal masses the center of mass is a weighted average of their positions.

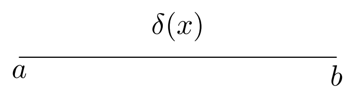


$x_{cm} = \frac{2x_1+x_2}{3}$.

In general, x_{cm} = weighted average of position $= \frac{\sum m_i x_i}{\sum m_i}$.

For a continuous density, $\delta(x)$, on the segment $[a, b]$ (units of density are mass/unit length) the sums become integrals. We will skip running through the logic of this since we are about to show it for two dimensions.

$$M = \int_a^b \delta(x)\,dx, \quad x_{cm} = \frac{1}{M}\int_a^b x\delta(x)\,dx.$$

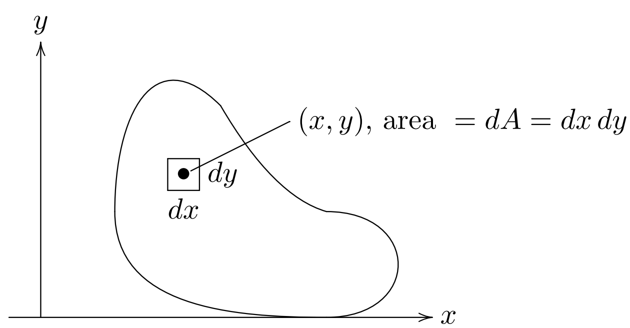


In 2 dimensions we label the center of mass as (x_{cm}, y_{cm}) and we have the following formulas

$$M = \int\int_R \delta(x,y)\,dA, \quad x_{cm} = \frac{1}{M}\int\int_R x\delta(x,y)\,dA, \quad y_{cm} = \frac{1}{M}\int\int_R y\delta(x,y)\,dA.$$

These formulas are easy to justify using our usual method for building integrals.

In this case, we divide our region into little pieces and we sum up the contributions of each piece using an integral. To keep the figure below uncluttered we only show one piece and we don't bother to label it as the i^{th}. In the end we will go directly to the integral, by thinking if it as a sum.



The little piece shown has mass $\delta(x,y)\,dA$ and the total mass is just the sum the pieces. That is, it's just the integral

$$M = \int\int_R \delta\,dA.$$

Likewise the x and y coordinates of the center of mass are just the weighted average of the x and y coordinates of each of the pieces. So, we get the formulas given above.

Example: Suppose the unit square has density $\delta = xy$; Find its mass and center of mass.

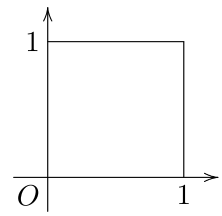
Answer:

$M = \int\int_R \delta\, dA = \int_0^1 \int_0^1 xy\, dx\, dy = \frac{1}{4}.$

$x_{cm} = \frac{1}{M} \int\int x\delta\, dA = \frac{1}{M} \int_0^1 \int_0^1 x^2 y\, dy\, dx.$

This is easily computed as $x_{cm} = \frac{2}{3}$.

By symmetry of the region and the density δ, we also have $y_{cm} = \frac{2}{3}$.



Average Value

We can think of center of mass as the average position of the mass. That is, it's the average of position with respect to mass. We can also take averages of functions with respect to other things. For instance, the *average value* of $f(x, y)$ *with respect to area* on a region R is

$$\frac{1}{\text{area } R} \int\int_R f(x, y)\, dA.$$

In general, if we simply say the average value of a function, it means average value with respect to area (or later, when we do triple integrals it can mean with respect to volume).

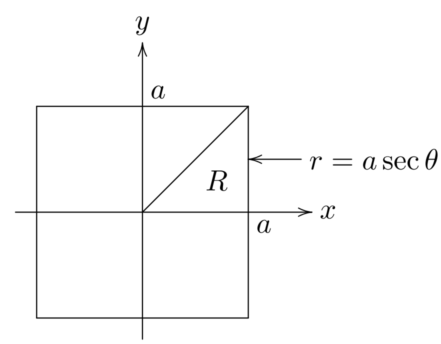
Example: What's the average distance of a point in a unit square from the center?

Answer: We center the square on the origin. Notice each side has length $2a$ and the area of the square is $4a^2$. So,

$$\text{average distance} = \frac{1}{4a^2} \int\int_{\text{square}} \sqrt{x^2 + y^2}\, dx\, dy.$$

By symmetry the integral is 8 times the integral of the triangular region R shown. We actually compute the integral in polar coordinates.

$$\begin{aligned} \text{average} &= \frac{8}{4a^2} \int\int_R r\, r\, dr\, d\theta \\ &= \frac{2}{a^2} \int_0^{\pi/4} \int_0^{a \sec\theta} r^2\, dr\, d\theta \\ &= \frac{2}{a^2} \int_0^{\pi/4} \frac{a^3 \sec^3\theta}{3}\, d\theta \\ &= \frac{a}{3}(\sqrt{2} + \ln(\sqrt{2} + 1)). \end{aligned}$$



This last integral was computed using integration by parts as

$$\int \sec^3\theta\, d\theta = \frac{\int \sec\theta\, d\theta + \sec\theta\tan\theta}{2} = \frac{\ln(\sec\theta + \tan\theta) + \sec\theta\tan\theta}{2}.$$

Note: the center of mass is the average value of x and y with respect to mass.

The *geometric center* has coordinates given by the average value of x and y with respect to area, i.e., the center of mass when $\delta = 1$.

Moment of Inertia

Moment of inertia

We will leave it to your physics class to really explain what moment of inertia means. Very briefly it measures an object's resistance (inertia) to a change in its rotational motion. It is analogous to the way mass measure the resistance to changes in the object's linear motion.

Because it has to do with rotational motion the moment of inertia is always measured about a reference line, which is thought of as the axis of rotation.

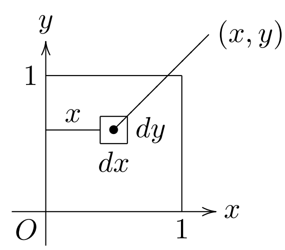
For a point mass, m, the moment of inertia about the line is

$$I = m\,d^2,$$

where d is the distance from the mass to the line. (The letter I is a standard notation for moment of inertia.)

If we have a distributed mass we compute the moment of inertia by summing the contributions of each of its parts. If the mass has a continuous distribution, this sum is, of course, an integral.

Example 1: Suppose the unit square, R, has density $\delta = xy$.
Find its moment of inertia about the y-axis.



Answer: The distance from the small piece of the square (shown in the figure) to the y-axis is x. If the piece has mass dm then its moment of inertia is

$$dI = x^2\,dm = x^2\delta(x, y)\,dA = x^3y\,dx\,dy.$$

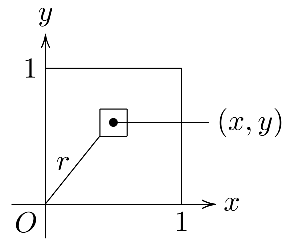
We took a shortcut here: we went straight to the notation for infinitesimal pieces, dI, dA and used equalities. rather than using more formal notation ΔI, ΔA, using approximations and then taking limits.

In the equation above, we used the notation dI to indicate it is just a small bit of moment of inertia. We also used that the mass of a piece is density times area. Now it's a simple matter to sum up all the bits of moment of inertia using an integral

$$I = \int\int_R dI = \int_0^1 \int_0^1 x^3y\,dy\,dx = \frac{1}{8}.$$

(Note: this integral is so easy to compute that we don't give the details.)

Example 2: For the same square as in example 1, find the *polar moment of inertia*.



Answer: The polar moment of inertia of a planar region is the moment of inertia about the origin (the axis of rotation is the z-axis). Finding this is exactly the same as in example 1, except the distance to the axis is now the polar distance r. We get,

$$dI = r^2\, dm = (x^2 + y^2)\delta(x, y)\, dA = (x^3y + xy^3)\, dx\, dy.$$

Summing, using an integral gives

$$I = \int\int_R dI = \int_0^1 \int_0^1 x^3y + xy^3\, dy\, dx = \frac{1}{4}.$$

Changing Variables in Multiple Integrals

1. Changing variables.

Double integrals in x, y coordinates which are taken over circular regions, or have integrands involving the combination $x^2 + y^2$, are often better done in polar coordinates:

$$(1) \qquad \iint_R f(x,y)\,dA = \iint_R g(r,\theta)\,r\,dr d\theta\ .$$

This involves introducing the new variables r and θ, together with the equations relating them to x, y in both the forward and backward directions:

$$(2) \qquad r = \sqrt{x^2+y^2}, \qquad \theta = \tan^{-1}(y/x); \qquad\qquad x = r\cos\theta, \qquad y = r\sin\theta\ .$$

Changing the integral to polar coordinates then requires three steps:

A. Changing the integrand $f(x,y)$ to $g(r,\theta)$, by using (2);

B. Supplying the area element in the r, θ system: $dA = \ r\,dr\,d\theta$;

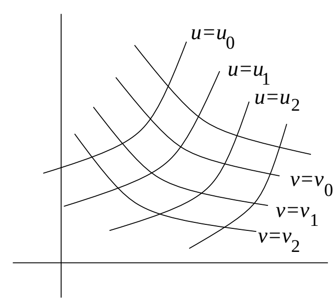
C. Using the region R to determine the limits of integration in the r, θ system.

In the same way, double integrals involving other types of regions or integrands can sometimes be simplified by changing the coordinate system from x, y to one better adapted to the region or integrand. Let's call the new coordinates u and v; then there will be equations introducing the new coordinates, going in both directions:

$$(3) \qquad u = u(x,y), \qquad v = v(x,y); \qquad\qquad x = x(u,v), \qquad y = y(u,v)$$

(often one will only get or use the equations in one of these directions). To change the integral to u, v-coordinates, we then have to carry out the three steps **A**, **B**, **C** above. A first step is to picture the new coordinate system; for this we use the same idea as for polar coordinates, namely, we consider the grid formed by the level curves of the new coordinate functions:

$$(4) \qquad u(x,y) = u_0, \qquad v(x,y) = v_0\ .$$



Once we have this, algebraic and geometric intuition will usually handle steps **A** and **C**, but for **B** we will need a formula: it uses a determinant called the **Jacobian**, whose notation and definition are

$$(5) \qquad \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\ .$$

Using it, the formula for the area element in the u, v-system is

$$(6) \qquad dA = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv\ ,$$

so the change of variable formula is

$$\iint_R f(x,y)\,dx\,dy \;=\; \iint_R g(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|\,du\,dv\;, \tag{7}$$

where $g(u,v)$ is obtained from $f(x,y)$ by substitution, using the equations (3).

We will derive the formula (5) for the new area element in the next section; for now let's check that it works for polar coordinates.

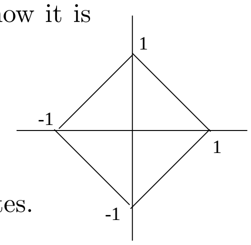
Example 1. Verify (1) using the general formulas (5) and (6).

Solution. Using (2), we calculate:

$$\frac{\partial(x,y)}{\partial(r,\theta)} \;=\; \begin{vmatrix} x_r & x_\theta \\ y_r & y_\theta \end{vmatrix} \;=\; \begin{vmatrix} \cos\theta & -r\,\sin\theta \\ \sin\theta & r\,\cos\theta \end{vmatrix} \;=\; r(\cos^2\theta+\sin^2\theta) \;=\; r\;,$$

so that $dA \;=\; r\,dr\,d\theta$, according to (5) and (6); note that we can omit the absolute value, since by convention, in integration problems we always assume $r \ge 0$, as is implied already by the equations (2).

We now work an example illustrating why the general formula is needed and how it is used; it illustrates step **C** also — putting in the new limits of integration.



Example 2. Evaluate $\displaystyle\iint_R\left(\frac{x-y}{x+y+2}\right)^2 dx\,dy$ over the region R pictured.

Solution. This would be a painful integral to work out in rectangular coordinates. But the region is bounded by the lines

$$x+y=\pm 1, \qquad x-y=\pm 1 \tag{8}$$

and the integrand also contains the combinations $x-y$ and $x+y$. These powerfully suggest that the integral will be simplified by the change of variable (we give it also in the inverse direction, by solving the first pair of equations for x and y):

$$u=x+y, \qquad v=x-y; \qquad\qquad x=\frac{u+v}{2}, \qquad y=\frac{u-v}{2}\;. \tag{9}$$

We will also need the new area element; using (5) and (9) above. we get

$$\frac{\partial(x,y)}{\partial(u,v)} \;=\; \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} \;=\; -\frac{1}{2}\;; \tag{10}$$

note that it is the second pair of equations in (9) that were used, not the ones introducing u and v. Thus the new area element is (this time we do need the absolute value sign in (6))

$$dA \;=\; \frac{1}{2}\,du\,dv\;. \tag{11}$$

We now combine steps **A** and **B** to get the new double integral; substituting into the integrand by using the first pair of equations in (9), we get

$$\iint_R\left(\frac{x-y}{x+y+2}\right)^2 dx\,dy \;=\; \iint_R\left(\frac{v}{u+2}\right)^2\frac{1}{2}\,du\,dv\;. \tag{12}$$

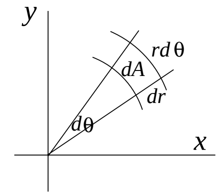
In uv-coordinates, the boundaries (8) of the region are simply $u = \pm 1$, $v = \pm 1$, so the integral (12) becomes

$$\iint_R \left(\frac{v}{u+2}\right)^2 \frac{1}{2}\, du\, dv \;=\; \int_{-1}^{1}\int_{-1}^{1} \left(\frac{v}{u+2}\right)^2 \frac{1}{2}\, du\, dv$$

We have

$$\text{inner integral} \;=\; -\frac{v^2}{2(u+2)}\Bigg]_{u=-1}^{u=1} \;=\; \frac{v^2}{3}\;; \qquad \text{outer integral} \;=\; \frac{v^3}{9}\Bigg]_{-1}^{1} \;=\; \frac{2}{9}\;.$$

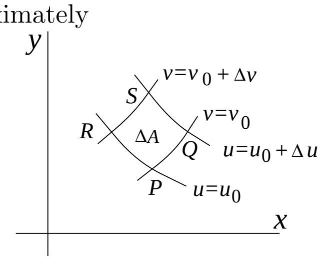
Changing Variables in Multiple Integrals



2. The area element.

In polar coordinates, we found the formula $dA = r\,dr\,d\theta$ for the area element by drawing the grid curves $r = r_0$ and $\theta = \theta_0$ for the r, θ-system, and determining (see the picture) the infinitesimal area of one of the little elements of the grid.

For general u, v-coordinates, we do the same thing. The grid curves (4) divide up the plane into small regions ΔA bounded by these contour curves. If the contour curves are close together, they will be approximately parallel, so that the grid element will be approximately a small parallelogram, and



(13) $$\Delta A \approx \text{area of parallelogram PQRS} = |PQ \times PR|$$

In the uv-system, the points P, Q, R have the coordinates

$$P: (u_0, v_0), \qquad Q: (u_0 + \Delta u, v_0), \qquad R: (u_0, v_0 + \Delta v)\ ;$$

to use the cross-product however in (13), we need PQ and PR in $\mathbf{i}\,\mathbf{j}$- coordinates. Consider PQ first; we have

(14) $$PQ = \Delta x\,\mathbf{i} + \Delta y\,\mathbf{j}\ ,$$

where Δx and Δy are the changes in x and y as you hold $v = v_0$ and change u_0 to $u_0 + \Delta u$. According to the definition of partial derivative,

$$\Delta x \approx \left(\frac{\partial x}{\partial u}\right)_0 \Delta u, \qquad \Delta y \approx \left(\frac{\partial y}{\partial u}\right)_0 \Delta u;$$

so that by (14),

(15) $$PQ \approx \left(\frac{\partial x}{\partial u}\right)_0 \Delta u\ \mathbf{i} + \left(\frac{\partial y}{\partial u}\right)_0 \Delta u\ \mathbf{j}\ .$$

In the same way, since in moving from P to R we hold u fixed and increase v_0 by Δv,

(16) $$PR \approx \left(\frac{\partial x}{\partial v}\right)_0 \Delta v\ \mathbf{i} + \left(\frac{\partial y}{\partial v}\right)_0 \Delta v\ \mathbf{j}\ .$$

We now use (13); since the vectors are in the xy-plane, $PQ \times PR$ has only a $\mathbf{k}$-component, and we calculate from (15) and (16) that

$$\mathbf{k}\text{-component of } PQ \times PR \approx \begin{vmatrix} x_u \Delta u & y_u \Delta u \\ x_v \Delta v & y_v \Delta v \end{vmatrix}_0$$

(17) $$= \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}_0 \Delta u \Delta v\ ,$$

where we have first taken the transpose of the determinant (which doesn't change its value), and then factored the Δu and Δv out of the two columns. Finally, taking the absolute value, we get from (13) and (17), and the definition (5) of Jacobian,

$$\Delta A \approx \left|\frac{\partial(x,y)}{\partial(u,v)}\right|_0 \Delta u \Delta v\ ;$$

passing to the limit as $\Delta u, \Delta v \to 0$ and dropping the subscript 0 (so that P becomes any point in the plane), we get the desired formula for the area element,

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

Changing Variables in Multiple Integrals

3. Examples and comments; putting in limits.

If we write the change of variable formula as

$$\iint_R f(x,y)\,dx\,dy \;=\; \iint_R g(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv\ , \tag{18}$$

where

$$\frac{\partial(x,y)}{\partial(u,v)} \;=\; \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}\ , \qquad g(u,v) = f(x(u,v), y(u,v)), \tag{19}$$

it looks as if the essential equations we need are the inverse equations:

$$x = x(u,v), \qquad y = y(u,v) \tag{20}$$

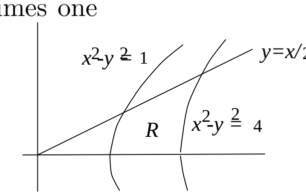
rather than the direct equations we are usually given:

$$u = u(x,y), \qquad v = v(x,y)\ . \tag{21}$$

If it is awkward to get (20) by solving (21) simultaneously for x and y in terms of u and v, sometimes one can avoid having to do this by using the following relation (whose proof is an application of the chain rule, and left for the Exercises):

$$\frac{\partial(x,y)}{\partial(u,v)}\,\frac{\partial(u,v)}{\partial(x,y)} \;=\; 1 \tag{22}$$

The right-hand Jacobian is easy to calculate if you know $u(x,y)$ and $v(x,y)$; then the left-hand one — the one needed in (19) — will be its reciprocal. Unfortunately, it will be in terms of x and y instead of u and v, so (20) still ought to be needed, but sometimes one gets lucky. The next example illustrates.



Example 3. Evaluate $\displaystyle\iint_R \frac{y}{x}\,dx\,dy$, where R is the region pictured, having as boundaries the curves $x^2 - y^2 = 1, \quad x^2 - y^2 = 4, \quad y = 0, \quad y = x/2\ .$

Solution. Since the boundaries of the region are contour curves of $x^2 - y^2$ and y/x, and the integrand is y/x, this suggests making the change of variable

$$u \;=\; x^2 - y^2, \qquad v \;=\; \frac{y}{x}\ . \tag{23}$$

We will try to get through without solving these backwards for x, y in terms of u, v. Since changing the integrand to the u, v variables will give no trouble, the question is whether we can get the Jacobian in terms of u and v easily. It all works out, using (22):

$$\frac{\partial(u,v)}{\partial(x,y)} \;=\; \begin{vmatrix} 2x & -2y \\ -y/x^2 & 1/x \end{vmatrix} \;=\; 2 - 2y^2/x^2 \;=\; 2 - 2v^2; \qquad \text{so} \quad \frac{\partial(x,y)}{\partial(u,v)} \;=\; \frac{1}{2(1-v^2)}\ ,$$

according to (22). We use now (18), put in the limits, and evaluate; note that the answer is positive, as it should be, since the integrand is positive.

$$\begin{aligned}
\iint_R \frac{y}{x}\, dx\, dy &= \iint_R \frac{v}{2(1-v^2)}\, du\, dv \\
&= \int_0^{1/2}\int_1^4 \frac{v}{2(1-v^2)}\, du\, dv \\
&= -\frac{3}{4}\ \ln(1-v^2)\Big]_0^{1/2} \;=\; -\frac{3}{4}\ \ln\ \frac{3}{4}\ .
\end{aligned}$$

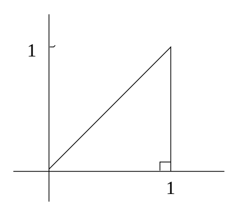
Putting in the limits

In the examples worked out so far, we had no trouble finding the limits of integration, since the region R was bounded by contour curves of u and v, which meant that the limits were constants.

If the region is not bounded by contour curves, maybe you should use a different change of variables, but if this isn't possible, you'll have to figure out the uv-equations of the boundary curves. The two examples below illustrate.

Example 4. Let $u = x+y,\ v = x-y$; change $\displaystyle\int_0^1\int_0^x dy\, dx$ to an iterated integral $du\, dv$.

Solution. Using (19) and (22), we calculate $\dfrac{\partial(x,y)}{\partial(u.v)} = -1/2$, so the Jacobian factor in the area element will be $1/2$.



To put in the new limits, we sketch the region of integration, as shown at the right. The diagonal boundary is the contour curve $v=0$; the horizontal and vertical boundaries are not contour curves — what are their uv-equations? There are two ways to answer this; the first is more widely applicable, but requires a separate calculation for each boundary curve.

Method 1 *Eliminate x and y from the three simultaneous equations $u = u(x,y)$, $v = v(x,y)$, and the xy-equation of the boundary curve.* For the x-axis and $x=1$, this gives

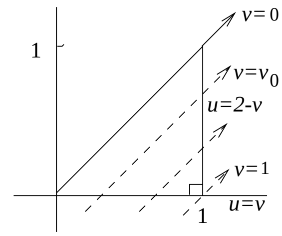
$$\begin{cases} u = x+y \\ v = x-y \\ y = 0 \end{cases} \Rightarrow \quad u = v; \qquad\qquad \begin{cases} u = x+y \\ v = x-y \\ x = 1 \end{cases} \Rightarrow \quad \begin{cases} u = 1+y \\ v = 1-y \end{cases} \Rightarrow \quad u+v = 2.$$

Method 2 *Solve for x and y in terms of u,v; then substitute $x = x(u,v),\ y = y(u,v)$ into the xy-equation of the curve.*

Using this method, we get $x = \frac{1}{2}(u+v),\ \ y = \frac{1}{2}(u-v)$; substituting into the xy-equations:

$$y = 0 \ \Rightarrow\ \frac{1}{2}(u-v) = 0 \ \Rightarrow\ u = v; \qquad\qquad x = 1 \ \Rightarrow\ \frac{1}{2}(u+v) = 1 \ \Rightarrow\ u+v = 2.$$

To supply the limits for the integration order $\iint du\, dv$, we



1. first hold v fixed, let u increase; this gives us the dashed lines shown;
2. integrate with respect to u from the u-value where a dashed line enters R (namely, $u = v$), to the u-value where it leaves (namely, $u = 2 - v$).
3. integrate with respect to v from the lowest v-values for which the dashed lines intersect the region R (namely, $v = 0$), to the highest such v-value (namely, $v = 1$).

Therefore the integral is $\displaystyle\int_0^1 \int_v^{2-v} \frac{1}{2}\, du\, dv$.

(As a check, evaluate it, and confirm that its value is the area of R. Then try setting up the iterated integral in the order $dv\,du$; you'll have to break it into two parts.)

Example 5. Using the change of coordinates $u = x^2 - y^2$, $v = y/x$ of Example 3, supply limits and integrand for $\displaystyle\iint_R \frac{dxdy}{x^2}$, where R is the infinite region in the first quadrant under $y = 1/x$ and to the right of $x^2 - y^2 = 1$.

Solution. We have to change the integrand, supply the Jacobian factor, and put in the right limits.

To change the integrand, we want to express x^2 in terms of u and v; this suggests eliminating y from the u, v equations; we get

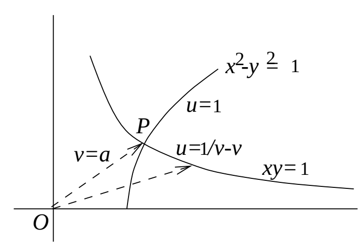
$$u = x^2 - y^2, \quad y = vx \quad \Rightarrow \quad u = x^2 - v^2x^2 \quad \Rightarrow \quad x^2 = \frac{u}{1 - v^2} .$$

From Example 3, we know that the Jacobian factor is $\dfrac{1}{2(1-v^2)}$; since in the region R we have by inspection $0 \le v < 1$, the Jacobian factor is always positive and we don't need the absolute value sign. So by (18) our integral becomes

$$\iint_R \frac{dx\,dxy}{x^2} \;=\; \iint_R \frac{1-v^2}{2u(1-v^2)}\,du\,dv \;=\; \iint_R \frac{du\,dv}{2u}$$

Finally, we have to put in the limits. The x-axis and the left-hand boundary curve $x^2 - y^2 = 1$ are respectively the contour curves $v = 0$ and $u = 1$; our problem is the upper boundary curve $xy = 1$. To change this to $u - v$ coordinates, we follow Method 1:

$$\begin{cases} u = x^2 - y^2 \\ y = vx \\ xy = 1 \end{cases} \quad \Rightarrow \begin{cases} u = x^2 - 1/x^2 \\ v = 1/x^2 \end{cases} \quad \Rightarrow \; u = \frac{1}{v} - v \; .$$



The form of this upper limit suggests that we should integrate first with respect to u. Therefore we hold v fixed, and let u increase; this gives the dashed ray shown in the picture; we integrate from where it enters R at $u = 1$ to where it leaves, at $u = \dfrac{1}{v} - v$.

The rays we use are those intersecting R: they start from the lowest ray, corresponding to $v = 0$, and go to the ray $v = a$, where a is the slope of OP. Thus our integral is

$$\int_0^a \int_1^{1/v - v} \frac{du\,dv}{2u} \; .$$

To complete the work, we should determine a explicitly. This can be done by solving $xy = 1$ and $x^2 - y^2 = 1$ simultaneously to find the coordinates of P. A more elegant approach is to add $y = ax$ (representing the line OP) to the list of equations, and solve all three simultaneously for the slope a. We substitute $y = ax$ into the other two equations, and get

$$\begin{cases} ax^2 = 1 \\ x^2(1-a^2) = 1 \end{cases} \qquad \Rightarrow \qquad a = 1 - a^2 \qquad \Rightarrow \qquad a = \frac{-1+\sqrt{5}}{2} \ ,$$

by the quadratic formula.

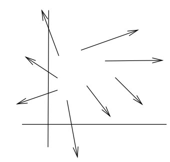
V1. Plane Vector Fields

1. Vector fields in the plane; gradient fields.

We consider a function of the type

$$\mathbf{F}(x,y) \;=\; M(x,y)\,\mathbf{i} + N(x,y)\,\mathbf{j}\ . \tag{1}$$

where M and N are both functions of two variables. To each pair of values (x_0, y_0) for which both M and N are defined, such a function assigns a vector $\mathbf{F}(x_0, y_0)$ in the plane. $\mathbf{F}$ is therefore called a **vector function of two variables**. The set of points (x, y) for which $\mathbf{F}$ is defined is called the *domain* of $\mathbf{F}$.



To visualize the function $\mathbf{F}(x, y)$, at each point (x_0, y_0) in the domain we place the corresponding vector $\mathbf{F}(x_0, y_0)$ so that its tail is at (x_0, y_0). Thus each point of the domain is the tail end of a vector, and what we get is called a **vector field**. This vector field gives a picture of the vector function $\mathbf{F}(x, y)$.

Conversely, given a vector field in a region of the xy-plane, it determines a vector function of the type (1), by expressing each vector of the field in terms of its $\mathbf{i}$ and $\mathbf{j}$ components. Thus there is no real distinction between "vector function" and "vector field". Mindful of the applications to physics, in these notes we will mostly use "vector field". We will use the same symbol $\mathbf{F}$ to denote both the field and the function, saying "the vector field $\mathbf{F}$", rather than "the vector field corresponding to the vector function $\mathbf{F}$".

We say the vector field $\mathbf{F}$ is *continuous* in some region of the plane if both $M(x, y)$ and $N(x, y)$ are continuous functions in that region. The intuitive picture of a continuous vector field is that the vectors associated to points sufficiently near (x_0, y_0) should have direction and magnitude very close to that of $\mathbf{F}(x_0, y_0)$ — in other words, as you move around the field, the vectors should change direction and magnitude smoothly, without sudden jumps in size or direction.

In the same way, we say $\mathbf{F}$ is *differentiable* in some region if M and N are differentiable, that is, if all the partial derivatives

$$\frac{\partial M}{\partial x}\,, \qquad \frac{\partial M}{\partial y}\,, \qquad \frac{\partial N}{\partial x}\,, \qquad \frac{\partial N}{\partial y}$$

exist in the region. We say $\mathbf{F}$ is *continuously differentiable* in the region if all these partial derivatives are themselves continuous there. In general, all the commonly used vector fields are continuously differentiable, except perhaps at isolated points, or along certain curves. But as you will see, these points or curves affect the properties of the field in very important ways.

Where do vector fields arise in science and engineering?

One important way is as **gradient vector fields**. If

$$w = f(x, y) \tag{2}$$

is a differentiable function of two variables, then its *gradient*

$$\nabla w = \frac{\partial w}{\partial x}\,\mathbf{i} + \frac{\partial w}{\partial y}\,\mathbf{j} \tag{3}$$

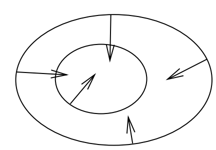
is a vector field, since both partial derivatives are functions of x and y. We recall the geometric interpretation of the gradient:

$$\begin{aligned} \text{dir } \nabla w &= \text{the direction } \mathbf{u} \text{ in which } \left.\frac{dw}{ds}\right|_{\mathbf{u}} \text{ is greatest;} \\ |\nabla w| &= \text{this greatest value of } \left.\frac{dw}{ds}\right|_{\mathbf{u}}, \end{aligned} \tag{4}$$

where $\left.\dfrac{dw}{ds}\right|_{\mathbf{u}} = \nabla w \cdot \mathbf{u}$ is the directional derivative of w in the direction $\mathbf{u}$.

Another important fact about the gradient is that if one draws the contour curves of $f(x, y)$, which by definition are the curves

$$f(x, y) = c, \qquad c \text{ constant},$$

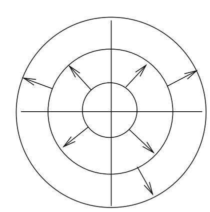


then at every point (x_0, y_0), the gradient vector ∇w at this point is perpendicular to the contour line passing through this point, i.e.,

(5) *the gradient field of f is perpendicular to the contour curves of f .*

Example 1. Let $w = \sqrt{x^2 + y^2} = r$. Using the definition (3) of gradient, we find

$$\nabla w = \frac{x}{r}\,\mathbf{i} + \frac{y}{r}\,\mathbf{j} = \frac{x\,\mathbf{i} + y\,\mathbf{j}}{r} .$$



The domain of ∇w is the xy-plane with $(0, 0)$ deleted, and it is continuously differentiable in this region. Since $|x\,i + y\,j| = r$, we see that $|\nabla w| = 1$. Thus all the vectors of the vector field ∇w are unit vectors, and they point radially outward from the origin. This makes sense by (4), since the definition of w shows that dw/ds should be greatest in the radially outward direction, and have the value 1 in that direction.

Finally, the contour curves for w are circles centered at $(0, 0)$, which are perpendicular to the vectors ∇w everywhere, as (5) predicts.

2. Force and velocity fields.

Continuing our search for ways in which vector fields arise, here are two physical situations which are described mathematically by vector fields. We shall refer to them often in the sequel, using our physical intuition to suggest the sort of mathematical properties that vector fields ought to have.

Force fields.

From physics, we have the two-dimensional electrostatic force fields arising from a distribution of static (i.e., not moving) charges in the plane. At each point (x_0, y_0) of the plane, we put a vector representing the force which would act on a unit positive charge placed at that point.

In the same way, we get vector fields arising from a distribution of masses in the xy-plane, representing the gravitational force acting at each point on a unit mass. There are also the electromagnetic fields arising from moving electric charges and/or a distribution of magnets, representing the magnetic force at each point.

Any of these we shall simply refer to as a **force field**.

Example 2. Express in $\mathbf{i} - \mathbf{j}$ form the electrostatic force field $\mathbf{F}$ in the xy-plane arising from a unit positive charge placed at the origin, given that the force vector at (x, y) is directed radially away from the origin and that it has magnitude c/r^2, c constant.

Solution. Since the vector $x\,\mathbf{i} + y\,\mathbf{j}$ with tail at (x, y) is directed radially outward and has magnitude r, it has the right direction, and we need only change its magnitude to c/r^2. We do this by multiplying it by c/r^3, which gives

$$\mathbf{F} \;=\; \frac{cx}{r^3}\,\mathbf{i} + \frac{cy}{r^3}\,\mathbf{j} \;=\; c\,\frac{x\,\mathbf{i} + y\,\mathbf{j}}{(x^2+y^2)^{3/2}}\;.$$

Flow fields and velocity fields

A second way vector fields arise is as the steady-state *flow fields* and *velocity fields*.

Imagine a fluid flowing in a horizontal shallow tank of uniform depth, and assume that the flow pattern at any point is purely horizontal and not changing with time. We will call this a *two-dimensional steady-state flow* or for short, simply a *flow*. The fluid can either be compressible (like a gas), or incompressible (like water). We also allow for the possibility that at various points, fluid is beiong added to or subtracted from the flow; for instance, someone could be standing over the tank pouring in water at a certain point, or over a certain area. We also allow the density to vary from point to point, as it would for an unevenly heated gas.

With such a flow we can associate two vector fields.

There is the **velocity field** $\mathbf{v}(x, y)$ where the vector $\mathbf{v}(x, y)$ at the point (x, y) represents the velocity vector of the flow at that point — that is, its direction gives the direction of flow, and its magnitude gives the speed of the flow.

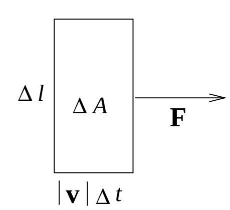
Then there is the **flow field**, defined by

$$\mathbf{F} \;=\; \delta(x, y)\,\mathbf{v}(x, y) \tag{6}$$

where $\delta(x, y)$ gives the density of the fluid at the point (x, y), in terms of mass per unit area. Assuming it is not 0 at a point (x, y), we can interpret $\mathbf{F}(x, y)$ as follows:

$$\text{dir}\,\mathbf{F} \;=\; \text{direction of fluid flow at } (x, y); \tag{7}$$

$$|\mathbf{F}| \;=\; \begin{cases} \text{rate (per unit length per second) of mass transport} \\ \text{across a line perpendicular to the flow direction at } (x, y). \end{cases}$$



Namely, we see that first by (6) and then by the picture,

$$|\mathbf{F}|\,\Delta l\,\Delta t \;=\; \delta\,|\mathbf{v}|\,\Delta t\,\Delta l \;=\; \text{mass in } \Delta A,$$

from which (7) follows by dividing by $\Delta l\,\Delta t$ and letting Δl and $\Delta t \to 0$.

If the density is a constant δ_0, as it would be for an incompressible fluid at a uniform temperature, then the flow field and velocity field are essentially the same, by (6) — the vectors of one are just a constant scalar multiple of the vectors of the other.

Example 3. Describe and interpret $\mathbf{F} = \dfrac{x\,\mathbf{i} + y\,\mathbf{j}}{x^2+y^2}$ as a flow field and a force field.

Solution. As in Example 2, the field $\mathbf{F}$ is defined everywhere except $(0,0)$ and its direction is radially outward; now, however, its magnitude is r/r^2, i.e., $|\mathbf{F}| = 1/r$.

$\mathbf{F}$ is the *flow field* for a source of magnitude 2π at the origin. To see this, look at a circle of radius a centered at the origin. At each point P on this circle, the flow is radially outward and by (7),

$$\begin{aligned}\text{mass transport rate at P} &= \frac{1}{a}, \qquad \text{so that}\\ \text{mass transport rate across circle} &= \frac{1}{a}\cdot 2\pi a \;=\; 2\pi\,.\end{aligned}$$

This shows that in one second, 2π mass flows out through every circle centered at the origin. This is the flow field for a source of magnitude 2π at the origin — for example, one could imagine a narrow pipe placed over the tank, introducting 2π mass units per second at the point $(0,0)$.

We know that $|\mathbf{F}| \;=\; \delta\,|\mathbf{v}| \;=\; 1/r$. Two important cases are:

• if the fluid is incompressible, like water, then its density is constant, so the flow speed must decrease like $1/r$ — the flow outward gets slower the further you are from the origin;

• if it is compressible like a gas, and its flow speed stays constant, then the density must decrease like $1/r$.

We now interpret the same field as a *force field*.

Suppose we think of the z-axis in space as a long straight wire, bearing a uniform positive electrostatic charge. This gives us a vector field in space, representing the electrostatic force field.

Since one part of the wire looks just like any other part, radial symmetry shows first that the vectors in the force field have 0 as their $\mathbf{k}$-component, i.e., they are pointed radially outward from the wire, and second that their magnitude depends only on their distance r from the wire. It can be shown in fact that the resulting force field is $\mathbf{F}$, up to a constant factor.

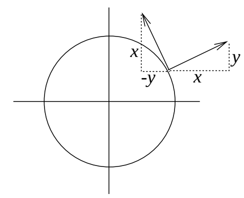
Such a field is called "two-dimensional", even though it is a vector field in space, because z and $\mathbf{k}$ don't enter into its description — once you know how it looks in the xy-plane, you know how it looks all through space.

The important thing to notice is that the magnitude of the force field in the xy-plane decreases like $1/r$, *not* like $1/r^2$, as it would if the charge were all at a point.

In the same way, the gravitational field of a uniform mass distribution along the z-axis would be $-\mathbf{F}$, up to a constant factor, and would be called a "two-dimensional gravitational field". Naturally, we don't have infinite long straight wires, but if you have a long straight wire, and stay away from its ends, or have only a short straight wire, but stay close to it, the force field will look like **F** near the wire.

Example 4. Find the velocity field of a fluid with density 1 in a shallow tank, rotating with constant angular velocity ω counterclockwise around the origin.

Solution. First we find the field direction at each point (x, y).



We know the vector $x\,\mathbf{i} + y\,\mathbf{j}$ is directed radially outward. Therefore a vector perpendicular to it in the counterclockwise direction (see picture) will be $-y\,\mathbf{i} + x\,\mathbf{j}$ (since its scalar product with $x\,\mathbf{i} + y\,\mathbf{j}$ is 0 and the signs are correct).

The preceding vector has magnitude r. If the angular velocity is ω, then the linear velocity is given by

$$|\mathbf{v}| = \omega r,$$

so to get the velocity field, we should multiply the above field by ω :

$$\mathbf{v} = -\omega y\,\mathbf{i} + \omega x\,\mathbf{j} \ .$$

Work and line integrals

Line integrals: (also called *path integrals*)

Ingredients:

Field $\mathbf{F} = M\mathbf{i} + N\mathbf{j} = \langle M, N\rangle$

Curve C: $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} = \langle x, y\rangle \Rightarrow d\mathbf{r} = \langle dx, dy\rangle$.

Line integral:

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \langle M, N\rangle \cdot \langle dx, dy\rangle = \int_C M\,dx + N\,dy.$$

We need to discuss:

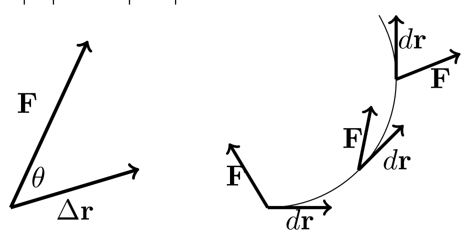
a) What this notation means and how line integrals arise.

b) How to compute them.

c) Their properties and notation.

a) **How line integrals arise.**

The figure on the left shows a force $\mathbf{F}$ being applied over a displacement $\boldsymbol{\Delta}\mathbf{r}$. Work is force times distance, but only the component of the force in the direction of the displacement does any work. So, work $= |\mathbf{F}|\cos\theta\,|\Delta\mathbf{r}| = \mathbf{F}\cdot\Delta\mathbf{r}$.



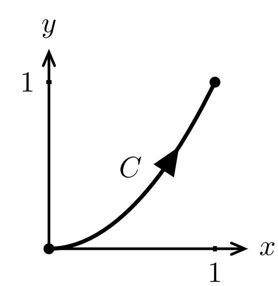
For a variable force applied over a curve the total work is found by 'summing' the infinitesimal pieces. We call this a line integral and denote it

$$\int_C \mathbf{F}\cdot d\mathbf{r}.$$

It is illustrated in the above figure on the right.

b) **Computing line integrals.**

We show this by steps by example.



Example 1: Evaluate $I = \displaystyle\int_C x^2y\,dx + (x-2y)\,dy$

over the part of the parabola $y = x^2$ from $(0,0)$ to $(1,1)$.

First, parametrize the curve:

$$x = t,\ y = t^2, \quad 0 \le t \le 1.$$

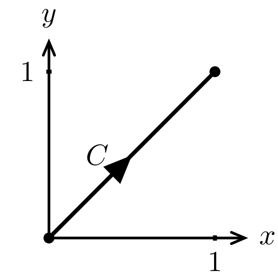
Note, we specified the range of t to get exactly the part of the curve we wanted.

Next, compute the differentials of x and y:

$$dx = dt, \quad dy = 2t\,dt.$$

Finally substitute everything in the integral and compute the standard single variable integral:

$$I = \int_0^1 t^2(t^2)\,dt + (t - 2t^2)2t\,dt = \int_0^1 t^4 + 2t^2 - 4t^3\,dt = -\frac{2}{15}.$$



Example 2: (Line integrals depend on the path.)
Same integral as previous example except C is the straight line from $(0,0)$ to $(1,1)$.
Parametrize curve: $x = t,\quad y = t,\quad 0 \le t \le 1.\ \Rightarrow\ dx = dt,\ dy = dt$

$$\Rightarrow\ I = \int_0^1 t^2 \cdot t\,dt + (t - 2t)\,dt = \int_0^1 t^3 - t\,dt = -\frac{1}{4}.$$

Note: this is a different value from example 1 and illustrates the very important fact that, in general, the line integral depends on the path. Later we will learn how to spot the cases when the line integral will be independent of path.

Example 3: (Line integrals are independent of the parametrization.)
Here we do the same integral as in example 1 except use a different parametrization of C.
Parametrize C: $x = \sin t,\quad y = \sin^2 t,\quad 0 \le t \le \pi/2\ \Rightarrow\ dx = \cos t\,dt,\ dy = 2\sin t\cos t\,dt.$

$$\begin{aligned}\Rightarrow\ I &= \int_0^{\pi/2} \sin^4 t\cos t\,dt + (\sin t - 2\sin^2 t)2\sin t\cos t\,dt \\ &= \int_0^{\pi/2} (\sin^4 t + 2\sin^2 t - 4\sin^3 t)\cos t\,dt = \frac{\sin^5 t}{5} + \frac{2}{3}\sin^3 t - \sin^4 t\Big|_0^{\pi/2} = -\frac{2}{15}.\end{aligned}$$

This is same value as example 1 and illustrates the very important point that the line integral is independent of how the curve is parametrized.

c) **Properties and notation of line integrals**

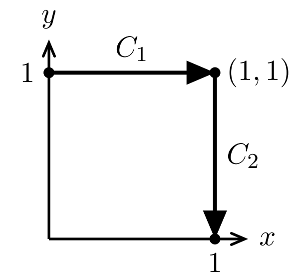
1. They are independent of parametrization.
2. If you reverse direction on curve then the line integral changes sign. That is,

$$\int_{-C} \mathbf{F}\cdot d\mathbf{r} = -\int_C \mathbf{F}\cdot d\mathbf{r}.$$

($-C$ means the same curve traversed in the opposite direction.)

3. If C is closed we use the notation $\oint_C \mathbf{F}\cdot d\mathbf{r} = \oint_C M\,dx + N\,dy.$

Example 4: Evaluate $I = \int_C y\,dx + (x+2y)\,dy$ where C is the curve shown.



Answer: The curve has two pieces so the integral will also

$$I = \int_{C_1} y\,dx + (x+2y)\,dy + \int_{C_2} y\,dx + (x+2y)\,dy.$$

Here we see that we don't always need to introduce a new variable t.

C_1 : $y = 1$, use x as parameter. $0 \le x \le 1 \Rightarrow dx = dx,\ dy = 0$.

Substituting in the integral $\Rightarrow \int_{C_1} y\,dx + (x+2y)\,dy = \int_0^1 dx = 1$.

C_2 : $x = 1$, use y as parameter. y goes from 1 to 0 and $dx = 0$

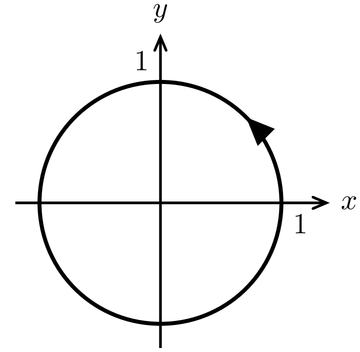
$\Rightarrow \int_{C_2} y\,dx + (x+2y)\,dy = \int_1^0 (1+2y)\,dy = -\int_0^1 1 + 2y\,dy = -2.$

We get $I = 1 - 2 = -1$.

Geometric Approach to Line Integrals

Line integrals are intrinsically geometric, so we should sometimes be able to use geometric reasoning to avoid the tedious calculations used in computing certain line integrals. The geometry can also give us some insight into the situation that calculation sometimes obscures.

We start with a line integral that we compute directly,



Example 1: Evaluate $I = \oint_C -y\,dx + x\,dy$ where C is the unit circle traversed in a counterclockwise (CCW) direction.
Parametrization: $x = \cos t, \quad y = \sin t, \quad 0 \le t \le 2\pi.$
$\Rightarrow \; dx = -\sin t\,dt,\; dy = \cos t\,dt.$
$\Rightarrow \; I = \int_0^{2\pi} -\sin t(-\sin t)\,dt + \cos t(\cos t)\,dt = \int_0^{2\pi} dt = 2\pi.$

The intrinsic formula Recall that we know $\dfrac{d\mathbf{r}}{dt} = \mathbf{T}\dfrac{ds}{dt}$, where $\mathbf{T}$ = unit tangent and s = arclength. Removing the dt gives $d\mathbf{r} = \mathbf{T}\,ds$. We can use this in our formula for line integrals and get a form that we call the *intrinsic formula*

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \mathbf{F}\cdot\mathbf{T}\,ds.$$

Example 2: Redo example 1 using the intrinsic formula: $\displaystyle\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \mathbf{F}\cdot\mathbf{T}\,ds.$
$\mathbf{T}$ = unit tangent $\Rightarrow$ $\mathbf{T} = -y\,\mathbf{i} + x\,\mathbf{j}$ (on the unit circle $x^2 + y^2 = 1$).
$\Rightarrow$ $\mathbf{F}\cdot\mathbf{T} = y^2 + x^2 = 1$ (on the unit circle) $\Rightarrow$ $I = \displaystyle\int_C ds$ = arclength of circle = 2π.
Lesson: it can pay to think geometrically.

Fundamental Theorem for Line Integrals

Gradient fields and potential functions

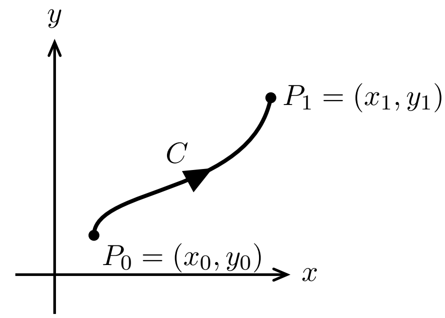
Earlier we learned about the gradient of a scalar valued function

$$\nabla f(x,y) = \langle f_x, f_y \rangle.$$

For example, $\nabla x^3y^4 = \langle 3x^2y^4, 4x^3y^3 \rangle$.

Now that we know about vector fields, we recognize this as a special case. We will call it a *gradient field*. The function f will be called a *potential function* for the field.

For gradient fields we get the following theorem, which you should recognize as being similar to the fundamental theorem of calculus.



Theorem (Fundamental Theorem for line integrals)

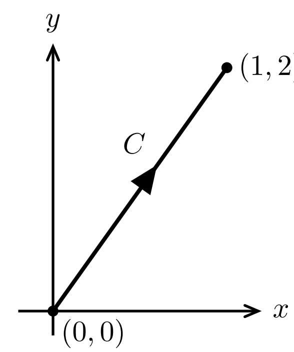
If $\mathbf{F} = \nabla f$ is a gradient field and C is *any* curve with endpoints $P_0 = (x_0, y_0)$ and $P_1 = (x_1, y_1)$ then

$$\int_C \mathbf{F} \cdot d\mathbf{r} = f(x,y)|_{P_0}^{P_1} = f(x_1,y_1) - f(x_0,y_0).$$

That is, for *gradient fields* the line integral is independent of the path taken, i.e., it depends only on the endpoints of C.

Example 1: Let $f(x,y) = xy^3 + x^2 \Rightarrow \mathbf{F} = \nabla f = \langle y^3 + 2x, 3xy^2 \rangle$

Let C be the curve shown and compute $I = \displaystyle\int_C \mathbf{F} \cdot d\mathbf{r}$.



Do this both directly (as in the previous topic) and using the above formula.

Method 1: parametrize C: $x = x$, $y = 2x$, $0 \le x \le 1$.

$$\Rightarrow I = \int_C (y^3 + 2x)\,dx + 3xy^2\,dy = \int_0^1 (8x^3 + 2x)\,dx + 12x^3 2\,dx$$

$$= \int_0^1 32x^3 + 2x\,dx = 9.$$

Method 2: $\displaystyle\int_C \nabla f \cdot d\mathbf{r} = f(1,2) - f(0,0) = 9.$

Proof of the fundamental theorem

$$\begin{aligned}\int_C \nabla f \cdot d\mathbf{r} &= \int_C f_x\,dx + f_y\,dy = \int_{t_0}^{t_1} \left[f_x(x(t),y(t))\,\frac{dx}{dt} + f_y(x(t),y(t))\,\frac{dy}{dt} \right] dt \\ &= \int_{t_0}^{t_1} \frac{d}{dt} f(x(t),y(t))\,dt = f(x(t),y(t))|_{t_0}^{t_1} = f(P_1) - f(P_0) \quad \blacksquare\end{aligned}$$

The third equality above follows from the chain rule.

Significance of the fundamental theorem

For gradient fields $\mathbf{F}$ the work integral $\displaystyle\int_C \mathbf{F} \cdot d\mathbf{r}$ depends only on the endpoints of the path.

We call such a line integral *path independent*.

The special case of this for closed curves C gives:

$$\mathbf{F} = \nabla f \Rightarrow \oint_C \mathbf{F} \cdot d\mathbf{r} = 0 \quad \text{(proof below).}$$

Following physics, where a conservative force does no work around a closed loop, we say $\mathbf{F} = \nabla f$ is a *conservative* field.

Example 1: If $\mathbf{F}$ is the electric field of an electric charge it is conservative.

Example 2: The gravitational field of a mass is conservative.

Differentials: Here we can use differentials to rephrase what we've done before. First recall:

a) $\nabla f = f_x\,\mathbf{i} + f_y\,\mathbf{j} \;\Rightarrow\; \nabla f \cdot d\mathbf{r} = f_x\,dx + f_y\,dy.$

b) $\displaystyle\int_C \nabla f \cdot d\mathbf{r} = f(P_1) - f(p_0).$

Using differentials we have $df = f_x\,dx + f_y\,dy$. (This is the same as $\nabla f \cdot d\mathbf{r}$.) We say $M\,dx + N\,dy$ is an *exact differential* if $M\,dx + N\,dy = df$ for some function f.

As in (b) above we have $\displaystyle\int_C M\,dx + N\,dy = \int_C df = f(P_1) - f(P_0).$

Proof that path independence is equivalent to conservative

We show that

$$\int_C \mathbf{F} \cdot d\mathbf{r} \quad \text{is path independent for any curve } C$$

is equivalent to

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0 \quad \text{for any closed path.}$$

This is not hard, it is really an exercise to demonstrate the logical structure of a proof showing equivalence. We have to show:

i) Path independence $\Rightarrow$ the line integral around any closed path is 0.

ii) The line integral around all closed paths is 0 $\Rightarrow$ path independence.

i) Assume path independence and consider the closed path C shown in figure (i) below. Since the starting point P_0 is the same as the endpoint P_1 we get $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r} = f(P_1) - f(P_0) = 0$ (this proves (i)).

ii) Assume $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed curve. If C_1 and C_2 are both paths between P_0 and P_1 (see fig. 2) then $C_1 - C_2$ is a closed path. So by hypothesis

$$\oint_{C_1 - C_2} \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} - \int_{C_2} \mathbf{F} \cdot d\mathbf{r} = 0 \;\Rightarrow\; \int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

That is the line integral is path independent, which proves (ii).

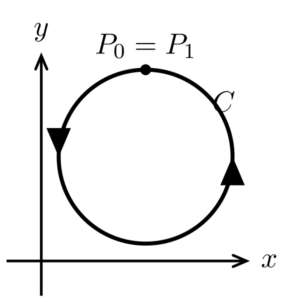


Figure (i)

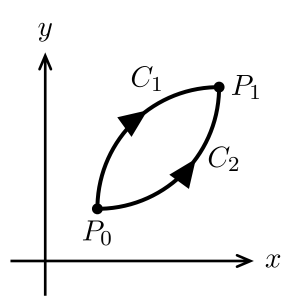


Figure (ii)

18.02SC Multivariable Calculus
Fall 2010

V2.1 Gradient Fields and Exact Differentials

1. Criterion for gradient fields.

Let $\mathbf{F} = M(x,y)\,\mathbf{i} + N(x,y)\,\mathbf{j}$ be a two-dimensional vector field, where M and N are continuous functions. There are three equivalent ways of saying that $\mathbf{F}$ is conservative, i.e., a gradient field:

$$(1) \qquad \mathbf{F} = \nabla f \;\Leftrightarrow\; \int_P^Q \mathbf{F}\cdot d\mathbf{r} \;\text{ is path-independent } \;\Leftrightarrow\; \oint_C \mathbf{F}\cdot d\mathbf{r} = 0 \;\text{ for any closed } C$$

Unfortunately, these equivalent formulations don't give us any effective way of deciding if a given field $\mathbf{F}$ is a conservative field or not. However, if we assume that $\mathbf{F}$ is not just continuous but is even continuously differentiable (meaning: M_x, M_y, N_x, N_y all exist and are continuous), then there is a simple and elegant criterion for deciding whether or not $\mathbf{F}$ is a gradient field in some region.

Criterion. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ be continuously differentiable in a region D. Then, in D,

$$(2) \qquad \mathbf{F} = \nabla f \;\text{ for some } f(x,y) \quad \Rightarrow \quad M_y = N_x\ .$$

Proof. Since $\mathbf{F} = \nabla f$, this means

$$\begin{array}{cccl} M = f_x & \text{and} & N = f_y\ . & \qquad \text{Therefore,} \\ M_y = f_{xy} & \text{and} & N_x = f_{yx}\ . & \end{array}$$

But since these two mixed partial derivatives are continuous (since M_y and N_x are, by hypothesis), a standard 18.02 theorem says they are equal. Thus $M_y = N_x$. $\square$

This theorem may be expressed in a slightly different form, if we define the scalar function called the **two-dimensional curl** of $\mathbf{F}$ by

$$(3) \qquad \text{curl}\,\mathbf{F} \;=\; N_x - M_y\ .$$

Then (2) becomes

$$(2') \qquad \mathbf{F} = \nabla f \quad \Rightarrow \quad \text{curl}\,\mathbf{F} = 0\ .$$

This criterion allows us to test $\mathbf{F}$ to see if it is a gradient field. Naturally, we would also like to know that the converse is true: if curl $\mathbf{F} = 0$, then $\mathbf{F}$ is a gradient field. As we shall see, however, this requires some additional hypotheses on the region D. For now, we will assume D is the whole plane. Then we have

Converse to Criterion. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ be continuously differentiable for all x, y.

$$(4) \qquad M_y = N_x \text{ for all } x,y \quad \Rightarrow \quad \mathbf{F} = \nabla f \;\text{ for some differentiable } f \text{ and all } x,y.$$

The proof of (4) will be postponed until we have more technique. For now we will illustrate the use of the criterion and its converse.

Example 1. For which value(s), if any of the constants a, b will $axy\,\mathbf{i} + (x^2 + by)\,\mathbf{j}$ be a gradient field?

Solution. The partial derivatives are continuous for all x, y and $M_y = ax,\ N_x = 2x$. Thus by (2) and (4), $\mathbf{F} = \nabla f \ \Leftrightarrow\ a = 2;\ b$ is arbitrary.

Example 2. Are the fields $\mathbf{F} \;=\; \dfrac{x\,\mathbf{i} + y\,\mathbf{j}}{x^2+y^2}, \qquad \mathbf{G} \;=\; \dfrac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2+y^2}$ conservative?

Solution. We have (the second line follows from the first by interchanging x and y):

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) &= \frac{-x^2+y^2}{(x^2+y^2)^2}; \qquad & \frac{\partial}{\partial x}\left(\frac{y}{x^2+y^2}\right) &= \frac{2xy}{(x^2+y^2)^2};\\
\frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right) &= \frac{x^2-y^2}{(x^2+y^2)^2}; \qquad & \frac{\partial}{\partial y}\left(\frac{x}{x^2+y^2}\right) &= \frac{2yx}{(x^2+y^2)^2};
\end{aligned}$$

from this, we see immediately that

$$\frac{\partial}{\partial y}\left(\frac{x}{x^2+y^2}\right) = \frac{\partial}{\partial x}\left(\frac{y}{x^2+y^2}\right); \qquad \frac{\partial}{\partial y}\left(-\frac{y}{x^2+y^2}\right) = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right);$$

the two equations in the last line show respectively that $\mathbf{F}$ and $\mathbf{G}$ satisfy the criterion (2). However, neither field is defined at $(0,0)$, so that the converse (4) is not applicable. So the question cannot be decided just on the basis of (2) and (4). In fact, it turns out that $\mathbf{F}$ is a gradient field, since one can check that

$$\mathbf{F} \;=\; \nabla \ln \sqrt{x^2+y^2} \;=\; \nabla \ln r\ .$$

On the other hand, $\mathbf{G}$ is not conservative, since if C is the unit circle $x = \cos t,\ y = \sin t$, we have

$$\oint_C \mathbf{G}\cdot d\mathbf{r} \;=\; \oint_C \frac{-y\,dx + x\,dy}{x^2+y^2} \;=\; \int_0^{2\pi} \frac{\sin^2 t\,dt + \cos^2 t\,dt}{\sin^2 t + \cos^2 t} \;=\; 2\pi \;\neq\; 0.$$

We will return to this example later on.

V2.2-3 Gradient Fields and Exact Differentials

2. Finding the potential function.

In example 2 in the previous reading we saw that

$$\mathbf{F} = \frac{x\,\mathbf{i} + y\,\mathbf{j}}{x^2+y^2} = \nabla \ln \sqrt{x^2+y^2} \;=\; \nabla \ln r.$$

This raises the question of how we found the function $\frac{1}{2}\ln(x^2+y^2)$. More generally, if we know that $\mathbf{F} = \nabla f$ — for example if curl $\mathbf{F} = 0$ in the whole xy-plane — how do we find the function $f(x,y)$? There are two methods; some students prefer one, some the other.

Method 1. Suppose $\mathbf{F} = \nabla f$. By the Fundamental Theorem for Line Integrals,

$$(5) \qquad \int_{(x_0,y_0)}^{(x,y)} \mathbf{F}\cdot d\mathbf{r} \;=\; f(x,y) - f(x_0,y_0)\;.$$

Read from left to right, (5) gives us an easy way of finding the line integral in terms of $f(x,y)$. But read right to left, it gives us a way of finding $f(x,y)$ by using the line integral:

$$(5') \qquad f(x,y) \;=\; \int_{(x_0,y_0)}^{(x,y)} \mathbf{F}\cdot d\mathbf{r} \;+\; c\;.$$

(Here c is an arbitrary constant of integration; as $(5')$ shows, $c = f(x_0,y_0)$.)

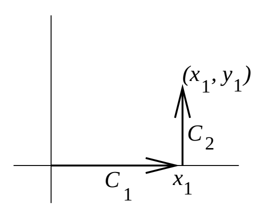
Remark. It is common to refer to $f(x,y)$ as the (mathematical) **potential function**. The potential function used in physics is $-f(x,y)$. The negative sign is used by physicists so that the potential difference will represent work done *against* the field $\mathbf{F}$, rather than work done *by* the field, as the convention is in mathematics.

Example 3. Let $\mathbf{F} = (x+y^2)\,\mathbf{i} + (2xy+3y^2)\,\mathbf{j}$. Verify that $\mathbf{F}$ satisfies the Criterion (2), and use method 1 above to find the potential function $f(x,y)$.

Solution. We verify (2) immediately: $\dfrac{\partial(y^2)}{\partial y} \;=\; 2y \;=\; \dfrac{\partial(2xy)}{\partial x}$.

We use $(5')$. The point (x_0,y_0) can be any convenient starting point; $(0,0)$ is the usual choice, if the integrand is defined there. (We will subscript the variables to avoid confusion with the variables of integration, but you don't have to.) By $(5')$,

$$(6) \qquad f(x_1,y_1) \;=\; \int_{(0,0)}^{(x_1,y_1)} (x+y^2)\,dx + (2xy+3y^2)\,dy\;.$$



Since the integral is path-independent, we can choose any path we like. The usual choice is the one on the right, as it simplifies the computations. (Most of what follows you can do mentally, with a little practice.)

On C_1, we have $y=0$, $dy=0$, so the integral on C_1 becomes $\displaystyle\int_0^{x_1} x\,dx \;=\; \frac{1}{2}x_1^2$.

On C_2, we have $x=x_1$, $dx=0$, so the integral is $\displaystyle\int_0^{y_1}(2x_1y+3y^2)\,dy \;=\; x_1y_1^2 + y_1^3\;.$

Adding the integrals on C_1 and C_2 to get the integral along the entire path, and dropping the subscripts, we get by (6) and (5′)

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

(The constant of integration is added by (5′), since the choice of starting point was arbitrary. You should always confirm the answer by checking that $\nabla f = \mathbf{F}$.) □

Method 2. Once again suppose $\mathbf{F} = \nabla f$, that is $M\,\mathbf{i} + N\,\mathbf{j} = f_x\,\mathbf{i} + f_y\,\mathbf{j}$. It follows that

$$f_x = M \qquad \text{and} \qquad f_y = N\ . \tag{7}$$

These are two equations involving partial derivatives, which we can solve simultaneously by integration. We illustrate using the previous example: $\mathbf{F} = (x+y^2,\ 2xy+3y^2)$.

Solution by Method 2. Using the first equation in (7),

$$\begin{aligned} \frac{\partial f}{\partial x} &= x + y^2. && \text{Hold } y \text{ fixed, integrate with respect to } x: \\ f &= \frac{1}{2}x^2 + y^2x + g(y). && \text{where } g(y) \text{ is an arbitrary function of } y. \end{aligned} \tag{8}$$

To find $g(y)$, we calculate $\dfrac{\partial f}{\partial y}$ two ways:

$$\begin{aligned} \frac{\partial f}{\partial y} &= 2yx + g'(y) && \text{by (8), while} \\ \frac{\partial f}{\partial y} &= 2xy + 3y^2 && \text{from (7), second equation.} \end{aligned}$$

Comparing these two expressions, we see that $g'(y) = 3y^2$, so $g(y) = y^3 + c$. Putting it all together, using (8), we get $f(x,y) = \frac{1}{2}x^2 + y^2x + y^3 + c$, as before. □

In the first method, the answer is written down immediately as a line integral; the rest of the work is in evaluating the integral, which goes quickly, since on a horizontal or vertical path either $dx = 0$ or $dy = 0$.

In the second method, the answer is obtained by an algorithm involving several steps which should be carried out in the right order.

The first method has the advantage of reminding you each time how $f(x,y)$ is defined and what it means, facts of theoretical and practical importance. The second method has the advantage of requiring no knowledge of line integrals, which makes it popular with students; on the other hand, when done in three dimensions, the bookkeeping gets more complicated, whereas in the first method it does not; overall, the first method is faster, provided you are confident enough to do some of the work mentally.

3. Exact differentials.

The formal expressions $M(x,y)\,dx + N(x,y)\,dy$ which have appeared as the integrands in our line integrals are called **differentials**. In some applications, most notably thermodynamics, one usually works directly with the differential $M\,dx + N\,dy$ and its line integral $\int M\,dx + N\,dy$, without considering or using the associated vector field $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$. Therefore it is important to have the results about gradient fields in this section translated into the language of differentials. We do this now.

If $f(x,y)$ is a differentiable function, its **total differential** df (or simply *differential*) is by definition the expression

$$df = f_x dx + f_y dy \ . \tag{9}$$

For example, if $f(x,y) = x^2y^3$, then $d(x^2y^3) = 2xy^3dx + 3x^2y^2dy$.

We call the differential $M\,dx + N\,dy$ **exact**, in a region D where M and N are defined, if it is the total differential of some function $f(x,y)$ in this region, i.e., if in D,

$$M = f_x \qquad \text{and} \qquad N = f_y, \qquad \text{for some } \ f(x,y). \tag{10}$$

From this we see that the relation between differentials and vector fields is

$$\begin{array}{rcl} M\,dx + N\,dy \quad \text{is exact} & \Leftrightarrow & M\,\mathbf{i} + N\,\mathbf{j} \quad \text{is a gradient field} \\ M\,dx + N\,dy \ = \ df & \Leftrightarrow & M\,\mathbf{i} + N\,\mathbf{j} \ = \ \nabla f \ . \end{array}$$

In this language, the criterion (we use the same equation number as in the section where they were first presented)

$$\mathbf{F} = \nabla f \ \text{ for some } f(x,y) \quad \Rightarrow \quad M_y = N_x \ . \tag{2}$$

and its partial converse:

Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ be continuously differentiable for all x, y.

$$M_y = N_x \text{ for all } x, y \quad \Rightarrow \quad \mathbf{F} = \nabla f \ \text{ for some differentiable } f \text{ and all } x, y. \tag{4}$$

become the

Exactness Criterion. *Assume M and N are continuously differentiable in a region D of the plane. Then in this region,*

$$M\,dx + N\,dy \ \ \textit{exact} \quad \Rightarrow \quad M_y = N_x \ ; \tag{11}$$

$$\textit{if } D \textit{ is the whole } xy\textit{-plane}, \quad M_y = N_x \quad \Rightarrow \quad M\,dx + N\,dy \ \ \textit{exact}. \tag{12}$$

If the exactness criterion shows that $M\,dx + N\,dy$ is exact, then the function $f(x,y)$ may be found by either of the two methods previously described.

Two Dimensional Curl

We have learned about the curl for two dimensional vector fields.

By definition, if $\mathbf{F} = \langle M, N \rangle$ then the two dimensional curl of $\mathbf{F}$ is $\operatorname{curl} \mathbf{F} = N_x - M_y$

Example: If $\mathbf{F} = x^3y^2\,\mathbf{i} + x\,\mathbf{j}$ then $M = x^3y^2$ and $N = x$, so $\operatorname{curl} \mathbf{F} = 1 - 2x^3y$.

Notice that $\mathbf{F}(x, y)$ is a vector valued function and its curl is a scalar valued function. It is important that we label this as the two dimensional curl because it is only for vector fields in the plane. Later we will see that the two dimensional curl is really just the $\mathbf{k}$ component of the (vector valued) three dimensional curl.

Green's Theorem

Green's Theorem

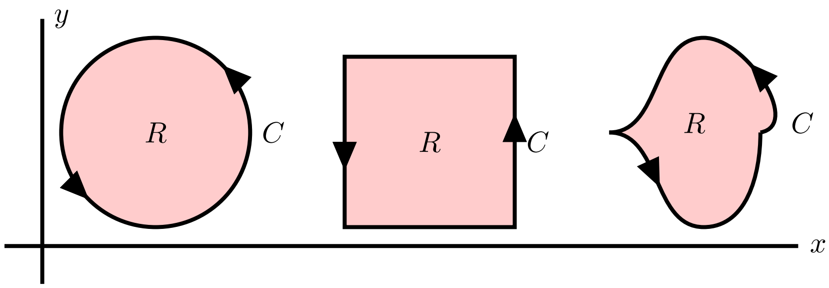
We start with the ingredients for Green's theorem.

(i) C a *simple* closed curve (simple means it never intersects itself)

(ii) R the interior of C.

We also require that C must be *positively oriented*, that is, it must be traversed so its interior is on the left as you move in around the curve. Finally we require that C be *piecewise smooth*. This means it is a smooth curve with, possibly a finite number of corners.

Here are some examples.



Green's Theorem

With the above ingredients for a vector field $\mathbf{F} = \langle M, N \rangle$ we have

$$\oint_C M\,dx + N\,dy = \iint_R N_x - M_y\,dA.$$

We call $N_x - M_y$ the two dimensional curl and denote it $\operatorname{curl}\mathbf{F}$.

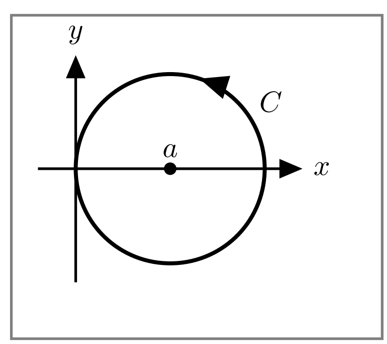
We can write also Green's theorem as

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_R \operatorname{curl}\mathbf{F}\,dA.$$

Example 1: (use the right hand side (RHS) to find the left hand side (LHS))

Use Green's Theorem to compute

$I = \oint_C 3x^2y^2\,dx + 2x^2(1+xy)\,dy$ where C is the circle shown.



By Green's Theorem $I = \iint_R 6x^2y + 4x - 6x^2y\,dA = 4\iint_R x\,dA.$

We could compute this directly, but we know $x_{cm} = \frac{1}{A}\iint_R x\,dA = a$

$\Rightarrow \iint_R x\,dA = \pi a^3 \Rightarrow \boxed{I = 4\pi a^3.}$

Example 2: (Use the LHS to find the RHS.)

Use Green's Theorem to find the area under one arch of the cycloid

$$x = a(\theta - \sin\theta),\ y = a(1-\cos\theta).$$

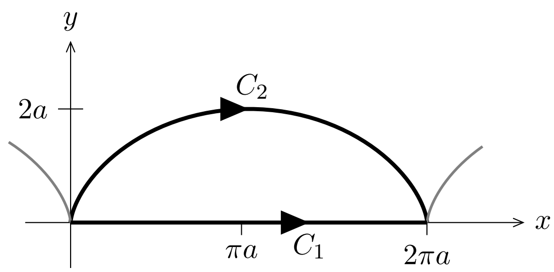
The picture shows the curve $C = C_1 - C_2$ surrounding the area we want to find.(Note the minus sign on C_2.)

By Green's Theorem,

$$\oint_C -y\,dx = \iint_R dA = \text{area}.$$

Thus,

$$\text{area} = \oint_{C_1 - C_2} -y\,dx = \int_{C_1} 0 \cdot dx - \int_{C_2} -y\,dx = \int_0^{2\pi} a^2(1-\cos\theta)^2\,d\theta = 3\pi a^2.$$



Problems: Green's Theorem

Calculate $\oint_C -x^2y\,dx + xy^2\,dy$, where C is the circle of radius 2 centered on the origin.

Answer: Green's theorem tells us that if $\mathbf{F} = \langle M, N \rangle$ and C is a positively oriented simple closed curve, then

$$\oint_C M\,dx + N\,dy = \iint_R N_x - M_y\,dA.$$

We let $M = -x^2y$ and $N = xy^2$ to get:

$$\begin{aligned}
\oint_C -x^2y\,dx + xy^2\,dy &= \iint_R y^2 - (-x^2)\,dA \\
&= \iint_R x^2 + y^2\,dA \\
&= \int_0^{2\pi}\int_0^2 r^2\,r\,dr\,d\theta \\
&= \int_0^{2\pi} \frac{8}{3} d\theta \\
&= \frac{16\pi}{3}.
\end{aligned}$$

This result is 4/3 times the area $\iint_R 1\,dA$ of the circle, and so is a plausible answer.

Green's Theorem and Conservative Fields

We can use Green's theorem to prove the following theorem.

Theorem

Suppose $\mathbf{F} = \langle M, N \rangle$ is a vector field which is defined and with continuous partial derivatives for all (x, y). Then

$$\mathbf{F} \text{ is conservative } \Leftrightarrow \; N_x = M_y \; \text{ or } N_x - M_y = \operatorname{curl} \mathbf{F} = 0.$$

Proof

This is a consequence of Green's theorem. First, suppose $\mathbf{F}$ is conservative, i.e., its work integral is 0 along all simple closed curves. Then Green's theorem says

$$0 = \oint_C \mathbf{F} \cdot d\mathbf{r} = \int\int_R \operatorname{curl} \mathbf{F} \, dA.$$

The only way for the integral of $\operatorname{curl} \mathbf{F}$ to be 0 over all regions R is if $\operatorname{curl} \mathbf{F}$ itself is 0. This implies $N_x = M_y$ as claimed.

For the converse, assume $N_x = M_y$. Then, for any closed curve C surrounding a region R, Green's theorem says,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int\int_R N_x - M_y \, dA = 0.$$

Therefore, the work integral of $\mathbf{F}$ is 0 over any closed curve, which means $\mathbf{F}$ is conservative.

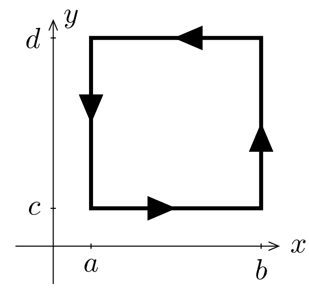
Be careful, the requirement that $\mathbf{F}$ is defined and differentiable everywhere is important. The problem following this note will give an example of a nonconservative field with $\operatorname{curl} \mathbf{F} = 0$.

Later we will learn how to handle fields that aren't defined everywhere.

Green's Theorem: Sketch of Proof

Green's Theorem: $\displaystyle\oint_C M\,dx + N\,dy = \iint_R N_x - M_y\,dA.$

Proof:



i) First we'll work on a rectangle. Later we'll use a lot of rectangles to approximate an arbitrary region.

ii) We'll only do $\displaystyle\oint_C M\,dx$ ($\displaystyle\oint_C N\,dy$ is similar).

By direct calculation the right hand side of Green's Theorem

$$\iint_R -\frac{\partial M}{\partial y}\,dA = \int_a^b\int_c^d -\frac{\partial M}{\partial y}\,dy\,dx.$$

Inner integral: $\quad -M(x,y)|_c^d = -M(x,d) + M(x,c)$

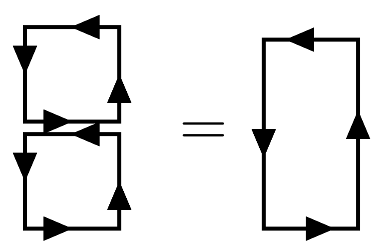
Outer integral: $\quad \displaystyle\iint_R -\frac{\partial M}{\partial y}\,dA = \int_a^b M(x,c) - M(x,d)\,dx.$

For the LHS we have

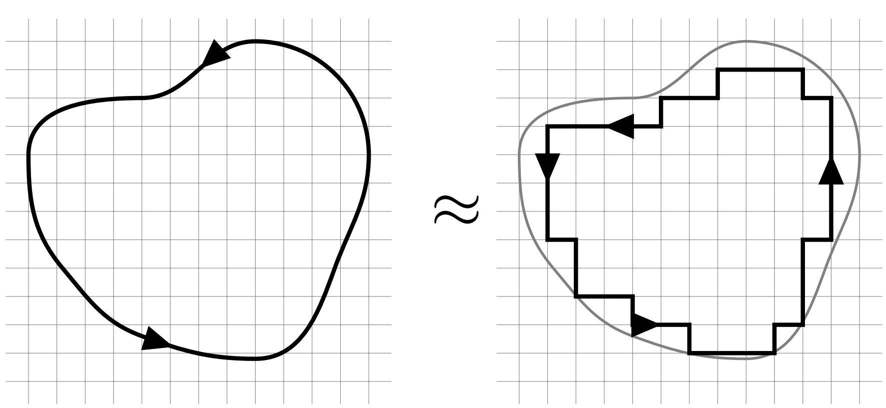
$$\begin{aligned}\oint_C M\,dx &= \int_{bottom} M\,dx + \int_{top} M\,dx \quad (\text{since } dx = 0 \text{ along the sides})\\ &= \int_a^b M(x,c)\,dx + \int_b^a M(x,d)\,dx = \int_a^b M(x,c) - M(x,d)\,dx.\end{aligned}$$

So, for a rectangle, we have proved Green's Theorem by showing the two sides are the same.

In lecture, Professor Auroux divided R into "vertically simple regions". This proof instead approximates R by a collection of rectangles which are especially simple both vertically and horizontally.



For line integrals, when adding two rectangles with a common edge the common edges are traversed in opposite directions so the sum is just the line integral over the outside boundary. Similarly when adding a lot of rectangles: everything cancels except the outside boundary. This extends Green's Theorem on a rectangle to Green's Theorem on a sum of rectangles. Since any region can be approximated as closely as we want by a sum of rectangles, Green's Theorem must hold on arbitrary regions.



Finding Area Using Line Integrals

Use a line integral (and Green's Theorem) to find the area of the unit circle.

Answer: Recall that Green's Theorem tells us $\oint_C M\,dx + N\,dy = \iint_R N_x - M_y\,dA$. To find the area of the unit circle we let $M = 0$ and $N = x$ to get $\iint_R 1\,dA = \oint_C x\,dy$.

We parametrize the circle by $x = \cos\theta$, $y = \sin\theta$, $0 < \theta \le 2\pi$, so $x\,dy = \cos^2\theta\,d\theta$. Then

$$\begin{aligned}
\text{Area} &= \iint_R 1\,dA \\
&= \oint_C x\,dy \\
&= \int_0^{2\pi} \cos^2\theta\,d\theta \\
&= \int_0^{2\pi} \frac{1+\cos 2\theta}{2}\,d\theta \\
&= \frac{1}{2}\left(\theta + \frac{1}{2}\sin 2\theta\right)\Bigg|_0^{2\pi} \\
&= \pi.
\end{aligned}$$

V3. Two-dimensional Flux

In this section and the next we give a different way of looking at Green's theorem which both shows its significance for flow fields and allows us to give an intuitive physical meaning for this rather mysterious equality between integrals.

We have seen that if $\mathbf{F}$ is a force field and C a directed curve, then

$$\text{(1)} \qquad \text{work done by } \mathbf{F} \text{ along } C \;=\; \int_C \mathbf{F}\cdot d\mathbf{r} \;=\; \int_C \mathbf{F}\cdot\mathbf{T}\,ds$$

In words, we are integrating $\mathbf{F}\cdot\mathbf{T}$, the *tangential component* of $\mathbf{F}$, along the curve C. In component notation, if $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$, then the above reads

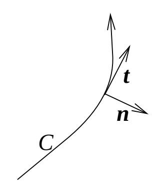
$$\text{(2)} \qquad \text{work} \;=\; \int_C M\,dx + N\,dy \;=\; \int_{t_0}^{t_1}\left(M\frac{dx}{dt} + N\frac{dy}{dt}\right)dt\ .$$

Analogously now, we may integrate $\mathbf{F}\cdot\mathbf{n}$, the *normal component* of $\mathbf{F}$ along C. To describe this, suppose the curve C is parametrized by the arclength s, increasing in the positive direction on C. The position vector for this parametrization and its corresponding tangent vector are given respectively by

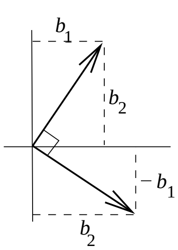
$$\mathbf{r}(s) \;=\; x(s)\,\mathbf{i} + y(s)\,\mathbf{j}\,, \qquad\qquad \mathbf{t}(s) \;=\; \frac{dx}{ds}\,\mathbf{i} + \frac{dy}{ds}\,\mathbf{j}\,;$$

where we have used $\mathbf{t}$ instead of $\mathbf{T}$ since it is a unit vector— its length is 1, as one can see by dividing through by ds on both sides of

$$ds = \sqrt{(dx)^2 + (dy)^2}\ .$$



The unit normal vector $\mathbf{n}$ is the one shown in the picture, obtained by rotating $\mathbf{t}$ clockwise through a right angle.



> Unfortunately, this direction is opposite to the one customarily used in kinematics, where $\mathbf{t}$ and $\mathbf{n}$ form a right-handed coordinate system for motion along C. The choice of $\mathbf{n}$ depends therefore on the context of the problem; the choice we have given is the most natural for applying Green's theorem to flow problems.

The usual formula for rotating a vector clockwise by 90^o (see the figure) shows that

$$\text{(3)} \qquad \mathbf{n}(s) \;=\; \frac{dy}{ds}\,\mathbf{i} - \frac{dx}{ds}\,\mathbf{j}\,.$$

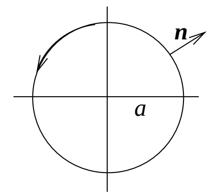
The line integral over C of the normal component $\mathbf{F}\cdot\mathbf{n}$ of the vector field $\mathbf{F}$ is called the **flux of F across** C. In symbols,

$$\text{(4)} \qquad \text{flux of } \mathbf{F} \text{ across } C \;=\; \int_C \mathbf{F}\cdot\mathbf{n}\,ds \;=\; \int_C \left(M\frac{dy}{ds} - N\frac{dx}{ds}\right)ds$$

In the notation of differentials, using (3) we write $\quad \mathbf{n}\,ds = dy\,\mathbf{i} - dx\,\mathbf{j}\,,\quad$ so that

$$\text{(5)} \qquad \text{flux of } F \text{ across } C \;=\; \int_C M\,dy - N\,dx \;=\; \int_C \left(M\frac{dy}{dt} - N\frac{dx}{dt}\right)dt\ ,$$

where $x(t),\ y(t)$ is any parametrization of C. We will need both (4) and (5).



Example 1. Calculate the flux of the field $\mathbf{F} = \dfrac{x\,\mathbf{i} + y\,\mathbf{j}}{x^2+y^2}$ across a circle of radius a and center at the origin, by a) using (4); b) using (5).

Solutions. a) The field is directed radially outward, so that $\mathbf{F}$ and $\mathbf{n}$ have the same direction. (As usual, the circle is directed counterclockwise, which means that $\mathbf{n}$ points outward.) Therefore, at each point of the circle,

$$\mathbf{F}\cdot\mathbf{n} = |\mathbf{F}| = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{a}\ .$$

Therefore, by (4), we get

$$\text{flux} = \oint_C \mathbf{F}\cdot\mathbf{n}\,ds = \oint_C \frac{1}{a}\,ds = 2\pi\ .$$

b) We can also get the same result by straightforward computation using a parametrization of the circle: $x = \cos t,\ y = \sin t$. Using this and (5) above,

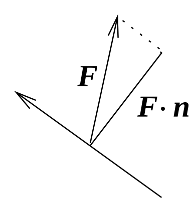
$$\text{flux} = \oint_C \frac{x\,dy - y\,dx}{x^2+y^2} = \int_0^{2\pi} \frac{a^2\cos^2 t + a^2\sin^2 t}{a^2}\,dt = 2\pi\ .$$

The natural physical interpretation for flux calls for thinking of $\mathbf{F}$ as representing a two-dimensional flow field (see section V1). Then the line integral represents the *rate with respect to time at which mass is being transported across* C. (We think of the flow as taking place in a shallow tank of unit depth. The convention about $\mathbf{n}$ makes this mass-transport rate positive if the flow is from left to right as you face in the positive direction on C, and negative in the other case.)

To see this, we follow the same procedure that was used to interpret the tangential integral in a force field as work.

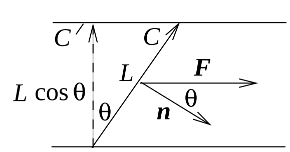
The essential step to see is that if $\mathbf{F}$ is a constant vector field representing a flow, and C is a directed line segment of length L, then

$$\text{(6)} \qquad \text{mass-transport rate across } C = (\mathbf{F}\cdot\mathbf{n})\,L$$



To see this, resolve the flow field into its components parallel to C and perpendicular to C. The component parallel to C contributes nothing to the flow rate *across* C, while the component perpendicular to C is $\mathbf{F}\cdot\mathbf{n}$.

Another way to se (6) is illustrated at the right. Letting C' be as shown, we see by conservation of mass that



$$\begin{aligned}\text{mass-transport rate across } C &= \text{mass-transport rate across } C' \\ &= |\mathbf{F}|(L\cos\theta) \\ &= (\mathbf{F}\cdot\mathbf{n})\,L\ .\end{aligned}$$

Once we have this, we follow the same procedure used to define work as a line integral. We divide up the curve and apply (6) to each of the approximating line segments, the k-th segment being of length approximately Δs_k. Thus

$$\text{mass-transport rate across } k\text{-th line segment} \;\approx\; (\mathbf{F}_k \cdot \mathbf{n}_k)\Delta s_k \; .$$

Adding these up and passing to the limit as the subdivision of the curve gets finer and finer then gives

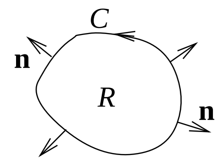
$$\text{mass-transport rate across } C \;=\; \int_C \mathbf{F}\cdot\mathbf{n}\, ds \; .$$

This interpretation shows why we call the line integral the *flux* of $\mathbf{F}$ across C. This terminology however is used even when $\mathbf{F}$ no longer represents a two-dimensional flow field. We speak of the flux of an electromagnetic field, for example.

Referring back to Example 1, the field $\mathbf{F} = \dfrac{x\,\mathbf{i} + y\,\mathbf{j}}{x^2+y^2}$ discussed there represents a flow stemming from a single source of strength 2π at the origin; thus the flux across each circle centered at the origin should also be 2π, regardless of the radius of the circle. This is what we found by actual calculation.

V4.1-2 Green's Theorem in Normal Form

1. Green's theorem for flux.



Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ represent a two-dimensional flow field, and C a simple closed curve, positively oriented, with interior R.

According to the previous section,

$$\text{(1)} \qquad \text{flux of } F \text{ across } C \;=\; \oint_C M\,dy - N\,dx\ .$$

Notice that since the normal vector points outwards, away from R, the flux is positive where the flow is out of R; flow into R counts as negative flux.

We now apply Green's theorem to the line integral in (1); first we write the integral in standard form (dx first, then dy):

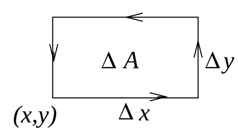
$$\oint_C M\,dy - N\,dx \;=\; \oint_C -N\,dx + M\,dy \;=\; \iint_R \big(M_x - (-N)_y\big)\,dA\ .$$

This gives us **Green's theorem in the normal form**

$$\text{(2)} \qquad \oint_C M\,dy - N\,dx \;=\; \iint_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA\ .$$

Mathematically this is the same theorem as the tangential form of Green's theorem — all we have done is to juggle the symbols M and N around, changing the sign of one of them. What is different is the physical interpretation. The left side represents the flux of $\mathbf{F}$ across the closed curve C. What does the right side represent?

2. The two-dimensional divergence.



Once again, let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$. We give a name to and a notation for the integrand of the double integral on the right of (2):

$$\text{(3)} \qquad \operatorname{div} \mathbf{F} \;=\; \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}, \qquad \text{the } \textbf{divergence} \text{ of } \mathbf{F}\ .$$

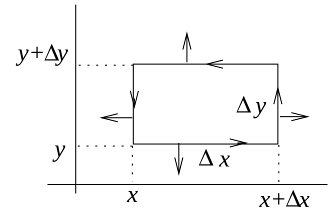
Evidently div $\mathbf{F}$ is a scalar function of two variables. To get at its physical meaning, look at the small rectangle pictured. If $\mathbf{F}$ is continuously differentiable, then div $\mathbf{F}$ is a continuous function, which is therefore approximately constant if the rectangle is small enough. We apply (2) to the rectangle; the double integral is approximated by a product, since the integrand is approximately constant:

$$\text{(4)} \quad \text{flux across sides of rectangle} \;\approx\; \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right)\Delta A\ , \qquad \Delta A \;=\; \text{area of rectangle.}$$

Because of the importance of this approximate relation, we give a more direct derivation of it which doesn't use Green's theorem. The reasoning which follows is widely used in mathematical modeling of physical problems.

Consider the small rectangle shown; we calculate approximately the flux over each side.

$$\begin{aligned}\text{flux across top} &\approx (\mathbf{F}(x,y+\Delta y)\cdot \mathbf{j}\,)\,\Delta x \;=\; N(x,y+\Delta y)\,\Delta x\\ \text{flux across bottom} &\approx (\mathbf{F}(x,y)\cdot -\mathbf{j}\,)\Delta x \;=\; -N(x,y)\Delta x\ ;\end{aligned}$$



adding these up,

$$\begin{matrix}\text{total flux across}\\ \text{top and bottom}\end{matrix} \quad\approx\quad \big(N(x,y+\Delta y)-N(x,y)\big)\Delta x \quad\approx\quad \left(\frac{\partial N}{\partial y}\,\Delta y\right)\Delta x.$$

By similar reasoning applied to the two sides,

$$\begin{matrix}\text{total flux across}\\ \text{left and right sides}\end{matrix} \quad\approx\quad \Big(M(x+\Delta x,y)-M(x,y)\Big)\,\Delta y \quad\approx\quad \left(\frac{\partial M}{\partial x}\,\Delta x\right)\Delta y\ .$$

Adding up the flux over the four sides, we get (4) again:

$$\begin{matrix}\text{total flux over four}\\ \text{sides of the rectangle}\end{matrix} \quad\approx\quad \left(\frac{\partial M}{\partial x}+\frac{\partial N}{\partial y}\right)\Delta x\,\Delta y\ .$$

Continuing our search for a physical meaning for the divergence, if the total flux over the sides of the small rectangle is positive, this means there is a net flow *out* of the rectangle. According to conservation of matter, the only way this can happen is if there is a *source* adding fluid directly to the rectangle. If the flow is taking place in a shallow tank of uniform depth, such a source can be visualized as someone standing over the tank, pouring fluid directly into the rectangle. Similarly, a net flow *into* the rectangle implies there is a *sink* withdrawing fluid from the rectangle. It is best to think of such a sink as a "negative source". The net rate (positive or negative) at which fluid is added directly to the rectangle from above may be called the "source rate" for the rectangle. Thus, since matter is conserved,

$$\text{flux over sides of rectangle} \quad=\quad \text{source rate for the rectangle;}$$

combining this with (4) shows that

$$\text{(5)}\qquad \text{source rate for the rectangle} \quad\approx\quad \left(\frac{\partial M}{\partial x}+\frac{\partial N}{\partial y}\right)\Delta A.$$

We now divide by ΔA and pass to the limit, getting by definition

$$\text{(6)}\qquad \text{the source rate at } (x,y) \quad=\quad \left(\frac{\partial M}{\partial x}+\frac{\partial N}{\partial y}\right) \;=\; \operatorname{div}\mathbf{F}\ .$$

The definition of the double integral as the limit of a sum shows in the usual way now that

$$\text{(7)}\qquad \text{source rate for } R \quad=\quad \iint_R \operatorname{div}\mathbf{F}\,dA\ .$$

These two relations (6) and (7) interpret the divergence physically, for a flow field, and they interpret also Green's theorem in the normal form:

$$\begin{aligned}\text{total flux across } C &\quad=\quad \text{source rate for } R\\ \oint_C M\,dy - N\,dx &\quad=\quad \iint_R \left(\frac{\partial M}{\partial x}+\frac{\partial N}{\partial y}\right) dA\end{aligned}$$

Since Green's theorem is a mathematical theorem, one might think we have "proved" the law of conservation of matter. This is not so, since this law was needed for our interpretation of div $\mathbf{F}$ as the source rate at (x, y).

We give side-by-side the two forms of Green's theorem, first in the vector form, then in the differential form used when calculations are to be done.

Tangential form	**Normal form**
$\displaystyle\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_R \operatorname{curl}\mathbf{F}\,dA$	$\displaystyle\oint_C \mathbf{F}\cdot \mathbf{n}ds = \iint_R \operatorname{div}\mathbf{F}\,dA$
$\displaystyle\oint_C M\,dx + N\,dy = \iint_R \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dA$	$\displaystyle\oint_C M\,dy - N\,dx = \iint_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA$
work by $\mathbf{F}$ around C	flux of $\mathbf{F}$ across C; source rate for R

Verify Green's Theorem in Normal Form

Verify that $\oint_C M\,dy - N\,dx = \iint_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA$ when $\mathbf{F} = x\hat{\mathbf{i}} + x\hat{\mathbf{j}}$ and C is the square with vertices $(0,0)$, (1,0), $(1,1)$ and $(0,1)$.

<u>Answer:</u>

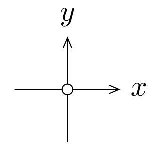
Right hand side: Here $M = N = x$, so $\iint_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dA = \iint_R 1\,dA = 1$.

Left hand side: $\oint_C M\,dy - N\,dx = \oint_C x\,dy - x\,dx$. We evaluate this line integral in four parts.

- $(0,0)$ to $(1,0)$.
$$\int_{x=0}^{x=1} x \cdot 0 - x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}.$$
- $(1,0)$ to $(1,1)$.
$$\int_{y=0}^{y=1} 1\,dy - 1 \cdot 0 = 1.$$
- $(1,1)$ to $(0,1)$.
$$\int_{x=1}^{x=0} x \cdot 0 - x\,dx = -\frac{1}{2}.$$
- $(0,1)$ to $(0,0)$.
$$\int_{y=1}^{y=0} 0\,dy - 0 \cdot 0 = 0.$$

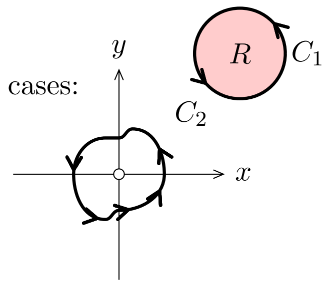
Since the sum of the line integrals along the components of C is 1, $\oint_C x\,dy - x\,dx = 1$. This confirms that the normal form of Green's Theorem is true in this example.

Extended Green's Theorem



Let $\mathbf{F}$ be the "tangential field" $\mathbf{F} = \dfrac{-y\mathbf{i} + x\mathbf{j}}{r^2}$, defined on the punctured plane $D = \mathbf{R}^2 - (0,0)$. It's easy to compute (we've done it before) that $\text{curl}\mathbf{F} = 0$ in D.

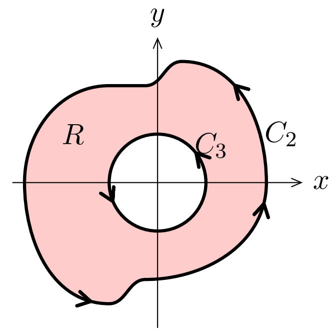
Question: For the tangential field $\mathbf{F}$, what do you think the possible values of $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r}$ could be if C were allowed to be any closed curve?



<u>**Answer:**</u> As we saw in lecture, if C is simple and positively oriented we have two cases:
(i) C_1 not around 0 (ii) C_2 around 0

(i) Green's Theorem $\Rightarrow \displaystyle\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \iint_R \text{curl}\mathbf{F} \cdot \mathbf{k}\, dA = 0.$

(ii) We show that $\displaystyle\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 2\pi.$



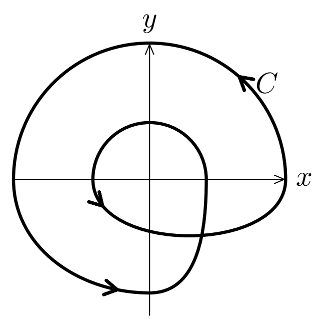
Let C_3 be a small circle of radius a, entirely inside C_2.
By extended Green's Theorem

$$\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} - \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = \iint_R \text{curl}\mathbf{F} \cdot \mathbf{k}\, dA = 0$$

$$\Rightarrow \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_3} \mathbf{F} \cdot d\mathbf{r}.$$

On the circle C_3 we can easily compute the line integral:

$$\mathbf{F} \cdot \mathbf{T} = 1/a \ \Rightarrow \oint_{C_3} \mathbf{F} \cdot \mathbf{T}\, ds = \int_{C_3} \frac{1}{a}\, ds = \frac{2\pi a}{a} = 2\pi. \quad \text{QED}$$



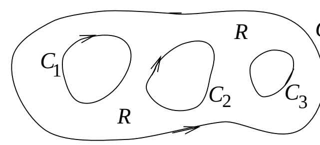
If C is positively oriented but not simple, the figure to the right suggests that we can break C into two curves around the origin at a point where it crosses itself. Repeating this as often as necessary, we find that $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r} = 2\pi n$, where n is the number of times C goes counterclockwise around (0,0).
If C is negatively oriented $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r} = -\oint_{C'} \mathbf{F} \cdot d\mathbf{r}$, where C' is an oppositely oriented copy of C. Hence, our final answer is that $\displaystyle\oint_C \mathbf{F} \cdot d\mathbf{r}$ may equal $2\pi n$ for any integer n.

An interesting aside: n is called the *winding number* of C around 0. n also equals the number of times C crosses the positive x-axis, counting $+1$ from below and -1 from above.

V5. Simply-Connected Regions

1. The Extended Green's Theorem.

In the work on Green's theorem so far, it has been assumed that the region R has as its boundary a single simple closed curve. But this isn't necessary. Suppose the region has a boundary composed of several simple closed curves, like the ones pictured. We suppose these boundary curves $C_1, \ldots, C_m$ all lie within the domain where $\mathbf{F}$ is continuously differentiable. Most importantly, all the curves must be directed so that the normal $\mathbf{n}$ points *away* from R.

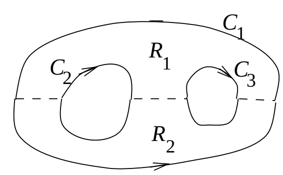


Extended Green's Theorem *With the curve orientations as shown,*

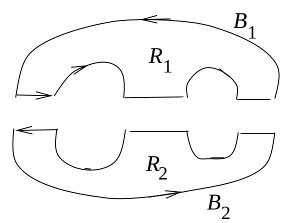
$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \ldots + \int_{C_m} \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_R \text{curl}\, \mathbf{F}\, dA \tag{1}$$

In other words, Green's theorem also applies to regions with several boundary curves, provided that we take the line integral over the complete boundary, with each part of the boundary oriented so the normal $\mathbf{n}$ points outside R.

Proof. We use subdivision; the idea is adequately conveyed by an example. Consider a region with three boundary curves as shown. The three cuts illustrated divide up R into two regions R_1 and R_2, each bounded by a single simple closed curve, and Green's theorem in the usual form can be applied to each piece. Letting B_1 and B_2 be the boundary curves shown, we have therefore



$$\oint_{B_1} \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_{R_1} \text{curl}\, \mathbf{F}\, dA \qquad \oint_{B_2} \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_{R_2} \text{curl}\, \mathbf{F}\, dA \tag{2}$$

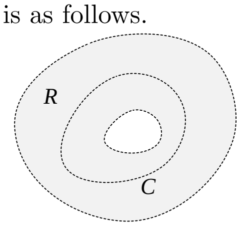


Add these two equations together. The right sides add up to the right side of (1). The left sides add up to the left side of (1) (for $m = 2$), since over each of the three cuts, there are two line integrals taken in opposite directions, which therefore cancel each other out. $\square$

2. Simply-connected and multiply-connected regions.

Though Green's theorem is still valid for a region with "holes" like the ones we just considered, the relation $\text{curl}\, \mathbf{F} = 0 \;\Rightarrow\; \mathbf{F} = \nabla f$ is not. The reason for this is as follows.

We are trying to show that $\quad \text{curl}\, \mathbf{F} = 0 \quad \Rightarrow \quad \oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed curve lying in R. We expect to be able to use Green's theorem. But if the region has a hole, like the one pictured, we cannot apply Green's theorem to the curve C because the interior of C is not entirely contained in R.



To see what a delicate affair this is, consider the earlier Example 2 in Section V2. The field $\mathbf{G}$ there satisfies curl $\mathbf{G} = 0$ everywhere but the origin. The region R is the xy-plane with $(0,0)$ removed. But $\mathbf{G}$ is not a gradient field, because $\oint_C \mathbf{G} \cdot d\mathbf{r} \neq 0$ around a circle C surrounding the origin.

This is clearer if we use Green's theorem in normal form (Section V4). If the flow field satisfies div $\mathbf{F} = 0$ everywhere except at one point, that doesn't guarantee that the flux through every closed curve will be 0. For the spot where div $\mathbf{F}$ is undefined might be a source, through which fluid is being added to the flow.

In order to be able to prove under reasonable hypotheses that curl $\mathbf{F} = 0 \;\Rightarrow\; \mathbf{F} = \nabla f$, we define our troubles away by assuming that R is the sort of region where the difficulties described above cannot occur—i.e., we assume that R has no holes; such regions are called *simply-connected*.

Definition. A two-dimensional region D of the plane consisting of one connected piece is called **simply-connected** if it has this property: whenever a simple closed curve C lies entirely in D, then its interior also lies entirely in D.

As examples: the xy-plane, the right-half plane where $x \geq 0$, and the unit circle with its interior are all simply-connected regions. But the xy-plane minus the origin is not simply-connected, since any circle surrounding the origin lies in D, yet its interior does not.

As indicated, one can think of a simply-connected region as one without "holes". Regions with holes are said to be *multiply-connected*, or *not simply-connected*.

Theorem. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ be continuously differentiable in a simply-connected region D of the xy-plane. Then in D,

(3) $\qquad \text{curl } F = 0 \quad \Rightarrow \quad \mathbf{F} = \nabla f, \qquad \text{for some } f(x,y); \qquad$ in terms of components,

(3′)
$$M_y = N_x \quad \Rightarrow \quad M\,\mathbf{i} + N\,\mathbf{j} = \nabla f, \qquad \text{for some } f(x,y).$$

Proof. Since a field is a gradient field if its line integral around any closed path is 0, it suffices to show

(4) $$\text{curl } \mathbf{F} = 0 \quad \Rightarrow \quad \oint_C \mathbf{F}\cdot d\mathbf{r} = 0 \qquad \text{for every closed curve } C \text{ in } D.$$

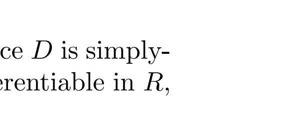
We prove (4) in two steps.

Assume first that C is a simple closed curve; let R be its interior. Then since D is simply-connected, R will lie entirely inside D. Therefore $\mathbf{F}$ will be continuously differentiable in R, and we can use Green's theorem:

$$\oint_C \mathbf{F}\cdot d\mathbf{r} \;=\; \iint_R \text{curl } \mathbf{F}\, dxdy \;=\; 0\ .$$

Next consider the general case, where C is closed but not simple—i.e., it intersects itself. Then C can be broken into smaller simple closed curves for which the above argument will be valid. A formal argument would be awkward to give, but the examples illustrate. In both cases, the path starts and ends at P, and

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \oint_{C_2} \mathbf{F}\cdot d\mathbf{r} + \int_{C_3} \mathbf{F}\cdot d\mathbf{r}\ .$$



In both cases, C_2 is a simple closed path, and also $C_1 + C_3$ is a simple closed path. Since D is simply-connected, the interiors automatically lie in D, so that by the first part of the argument,

$$\int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{C_3} \mathbf{F}\cdot d\mathbf{r} = 0 \qquad \text{and} \qquad \oint_{C_2} \mathbf{F}\cdot d\mathbf{r} = 0$$

Adding these up, we get $\oint_C \mathbf{F}\cdot d\mathbf{r} = 0$. □

The above argument works if C intersects itself a finite number of times. If C intersects itself infinitely often, we would have to resort to approximations to C; we skip this case.

We pause now to summarize compactly the central result, both in the language of vector fields and in the equivalent language of differentials.

Curl Theorem. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j}$ be a continuously differentiable vector field in a simply-connected region D of the xy-plane. Then the following four statements are equivalent — if any one is true for $\mathbf{F}$ in D, so are the other three:

1. $\int_P^Q \mathbf{F}\cdot d\mathbf{r}$ is path-independent — 1.′ $\int_P^Q M\,dx + N\,dy$ is path-independent

for any two points P, Q in D;

2. $\oint_C \mathbf{F}\cdot d\mathbf{r} = 0$, — 2.′ $\oint_C M\,dx + N\,dy = 0$,

for any simple closed curve C lying in D;

3. $\mathbf{F} = \nabla f$ for some f in D — 3.′ $M\,dx + N\,dy = df$ for some f in D

4. curl $\mathbf{F} = 0$ in D — 4.′ $M_y = N_x$ in D.

Remarks. We summarize below what still holds true even if one or more of the hypotheses doesn't hold: D is not simply-connected, or the field $\mathbf{F}$ is not differentiable everywhere in D.

1. Statements 1, 2, and 3 are equivalent even if $\mathbf{F}$ is only continuous; D need not be simply-connected..

2. Statements 1, 2, and 3 each implies 4, if $\mathbf{F}$ is continuously differentiable; D need not be simply-connected. (But 4 implies 1, 2, 3 only if D is simply-connected.)

Example 1. Is $\mathbf{F} = xy\,\mathbf{i} + x^2\,\mathbf{j}$ a gradient field?

Solution. We have curl $\mathbf{F} = x = 0$, so the theorem says it is not.

Example 2. Is $\dfrac{ydx - xdy}{y^2}$ an exact differential? If so, find all possible functions $f(x,y)$ for which it can be written df .

Solution. $M = 1/y$ and $N = -x/y^2$ are continuously differentiable wherever $y = 0$, i.e., in the two half-planes above and below the x-axis. These are both simply-connected. In each of them,

$$M_y = -1/y^2 = N_x$$

Thus in each half-plane the differential is exact, by the theorem, and we can calculate $f(x,y)$ by the standard methods in Sction V2. They give

$$f(x,y) \;=\; \frac{x}{y} + c$$

where c is an arbitrary constant. This constant need not be the same for the two regions, since they do not touch. Thus the most general function is

$$f(x,y) = \begin{cases} x/y \ + c, & y > 0 \\ x/y \ + c', & y < 0 \end{cases}; \qquad c, c' \text{ are arbitrary constants.}$$

Example 3. Let $\mathbf{F} = r^n(x\,\mathbf{i} + y\,\mathbf{j})$, $r = \sqrt{x^2+y^2}$. For which integers n is $\mathbf{F}$ conservative? For each such, find a corresponding $f(x,y)$ such that $\mathbf{F} = \nabla f$.

Solution. By the usual calculation, using the chain rule and the useful polar coordinate relations $r_x = x/r$, $r_y = y/r$, we find that curl $\mathbf{F} = 0$. There are two cases.

Case 1: $n \geq 0$. Then $\mathbf{F}$ is continuously differentiable in the whole xy-plane, which is simply-connected. Thus by the preceding theorem, $\mathbf{F}$ is conservative, and we can calculate $f(x,y)$ as in Section V2.

We use method 1 (line integration). The radial symmetry suggests using the ray C from $(0,0)$ to (x_1,y_1) as the path of integration, with the parametrization

$$x = x_1 t, \quad y = y_1 t, \quad 0 \leq t \leq 1;$$

also, let

$$r_1 = \sqrt{x_1^2 + y_1^2}; \qquad \text{then} \qquad r^n = r_1^n t^n, \quad x\,dx + y\,dy = r_1^2 t\,dt$$

and we get, by method 1 for finding $f(x,y)$,

$$\begin{aligned} f(x_1.y_1) &= \int_C \mathbf{F}\cdot d\mathbf{r} = \int_C r^n(x\,dx + y\,dy) \\ &= \int_0^1 r_1^{n+2} t^{n+1} dt = r_1^{n+2} \frac{t^{n+2}}{n+2}\bigg|_0^1 = \frac{r_1^{n+2}}{n+2}. \end{aligned} \tag{6}$$

so that

$$f(x,y) = \frac{r^{n+2}}{n+2}, \qquad \mathbf{F} = \nabla f, \quad n \geq 0\ . \tag{7}$$

Case 2: $n < 0$. The field $\mathbf{F}$ is not defined at $(0,0)$, so that its domain, the xy-plane with $(0,0)$ removed, is not simply-connected. So even though curl $\mathbf{F} = 0$ in this region, (3) is not immediately applicable.

Nonetheless, if $n = -2$, one can check by differentiation that (7) is still valid.

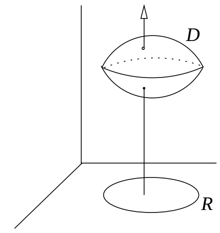
If $n = -2$, guessing, inspection, or method 2 give $f(x,y) = \ln r$.

We conclude that the field in all cases is a gradient field. Note in particular that the two force fields given in section V1, representing respectively (apart from a constant factor) the fields arising from a positive charge at $(0,0)$ and a uniform positive charge along the z-axis, correspond to the respective cases $n = -3$ and $n = -2$, and are both gradient fields:

$$\frac{x\,\mathbf{i} + y\,\mathbf{j}}{r^3} = \nabla\left(-\frac{1}{r}\right) \qquad (n = -3: \text{ positive charge at } (0,0))$$

$$\frac{x\,\mathbf{i} + y\,\mathbf{j}}{r^2} = \nabla(\ln r) \qquad (n = -2: \text{ uniform + charge on } z\text{-axis}).$$

Limits in Iterated Integrals



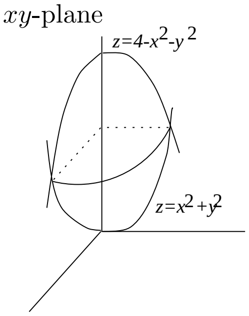
3. Triple integrals in rectangular and cylindrical coordinates.

You do these the same way, basically. To supply limits for $\iiint_D dz\,dy\,dx$ over the region D, we integrate first with respect to z. Therefore we

1. Hold x and y fixed, and let z increase. This gives us a vertical line.

2. Integrate from the z-value where the vertical line enters the region D to the z-value where it leaves D.

3. Supply the remaining limits (in either xy-coordinates or polar coordinates) so that you include all vertical lines which intersect D. This means that you will be integrating the remaining double integral over the region R in the xy-plane which D projects onto.



For example, if D is the region lying between the two paraboloids

$$z = x^2 + y^2 \qquad\qquad z = 4 - x^2 - y^2,$$

we get by following steps 1 and 2,

$$\iiint_D dz\,dy\,dx \;=\; \iint_R \int_{x^2+y^2}^{4-x^2-y^2} dz\,dA$$

where R is the projection of D onto the xy-plane. To finish the job, we have to determine what this projection is. From the picture, what we should determine is the xy-curve over which the two surfaces intersect. We find this curve by eliminating z from the two equations, getting

$$\begin{aligned} x^2 + y^2 &= 4 - x^2 - y^2, \qquad \text{which implies} \\ x^2 + y^2 &= 2. \end{aligned}$$

Thus the xy-curve bounding R is the circle in the xy-plane with center at the origin and radius $\sqrt{2}$.

This makes it natural to finish the integral in polar coordinates. We get

$$\iiint_D dz\,dy\,dx \;=\; \int_0^{2\pi} \int_0^{\sqrt{2}} \int_{x^2+y^2}^{4-x^2-y^2} dz\,r\,dr\,d\theta \;;$$

the limits on z will be replaced by r^2 and $4 - r^2$ when the integration is carried out.

Triple Integrals

1. Find the moment of inertia of the tetrahedron shown about the z-axis. Assume the tetrahedron has density 1.

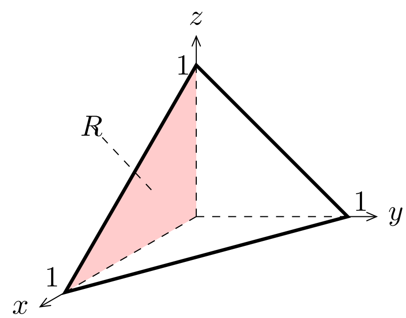


Figure 1: The tetrahedron bounded by $x + y + z = 1$ and the coordinate planes.

Answer: To compute the moment of inertia, we integrate distance squared from the z-axis times mass:

$$\iiint_D (x^2 + y^2) \cdot 1 \, dV.$$

Using cylindrical coordinates about the axis of rotation would give us an "easy" integrand (r) with complicated limits. The integrand $x^2 + y^2$ is not particularly intimidating, so we instead use rectangular coordinates. Integrating first with respect to y or x is preferable; $(x^2 + y^2)(1 - x - y)$ is a somewhat more intimidating integrand.

To find our limits of integration, we let y go from 0 to the slanted plane $x + y + z = 1$. The x and z coordinates are in R, the *projection* of D to the xz-plane which is bounded by the x and z axes and the line $x + z = 1$.

$$\text{Moment of Inertia } = \int_0^1 \int_0^{1-z} \int_0^{1-x-z} (x^2 + y^2) \, dy \, dx \, dz.$$

Inner: $\left.(x^2 y + \frac{1}{3} y^3)\right|_0^{1-x-z} = x^2 - x^3 - x^2 z + \frac{1}{3}(1 - x - z)^3.$

Middle:

$$\begin{aligned} \int_0^{1-z} x^2(1-z) - x^3 + \frac{1}{3}(1-x-z)^3 \, dx &= \left.\frac{1}{3}x^3(1-z) - \frac{1}{4}x^4 - \frac{1}{12}(1-x-z)^4\right|_0^{1-z} \\ &= \frac{1}{3}(1-z)^4 - \frac{1}{4}(1-z)^4 + \frac{1}{12}(1-z)^4 \\ &= \frac{1}{6}(1-z)^4. \end{aligned}$$

Outer: $\left.\frac{1}{30}(1-z)^5\right|_0^1 = \frac{1}{30}.$

2. Find the mass of a cylinder centered on the z-axis which has height h, radius a and density $\delta = x^2 + y^2$.

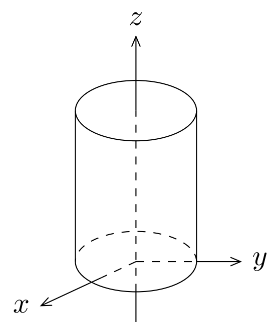


Figure 2: Cylinder.

Answer: To find the mass we integrate the product of density and volume:

$$\text{Mass} = \iiint_D \delta\, dV = \iiint_D r^2\, dV.$$

Naturally, we'll use cylindrical coordinates in this problem. The limits on z run from 0 to h. The x and y coordinates lie in a disk of radius a, so $0 \le r \le a$ and $0 < \theta \le 2\pi$.

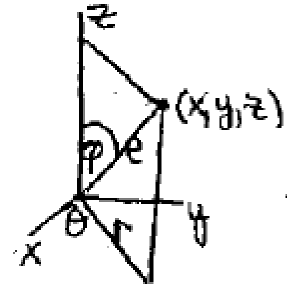
$$\text{Mass} = \iiint_D r^2\, dV = \int_0^{2\pi} \int_0^a \int_0^h r^2\, dz\, r\, dr\, d\theta = \int_0^{2\pi} \int_0^a \int_0^h r^3\, dz\, dr\, d\theta.$$

Inner integral: $r^3 z\big|_0^h = hr^3$.

Middle integral: $\displaystyle\int_0^a hr^3\, dr = \frac{ha^4}{4}$.

Outer integral: $2\pi\dfrac{ha^4}{4} = \dfrac{\pi h a^4}{2}$.

Limits in Spherical Coordinates



Definition of spherical coordinates

$\rho =$ distance to origin, $\rho \geq 0$

$\phi =$ angle to z-axis, $0 \leq \phi \leq \pi$

$\theta =$ usual $\theta =$ angle of projection to xy-plane with x-axis, $0 \leq \theta \leq 2\pi$

Easy trigonometry gives:

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

The equations for x and y are most easily deduced by noticing that for r from polar coordinates we have

$$r = \rho \sin\phi.$$

This implies

$$x = r\cos\theta = \rho\sin\phi\cos\theta, \text{ and } y = r\sin\theta = \rho\sin\phi\sin\theta.$$

Going the other way:

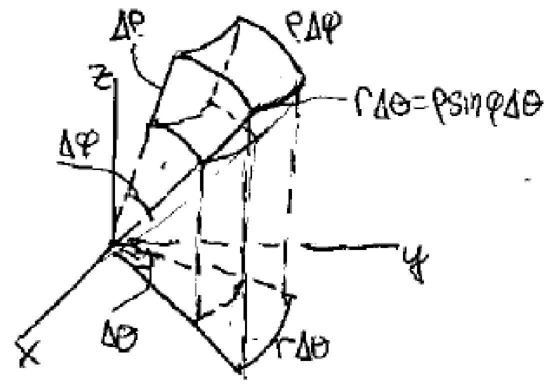
$\rho = \sqrt{z^2 + y^2 + z^2}$ $\quad \phi = \cos^{-1}(z/\rho)$ $\quad \theta = \tan^{-1}(y/x)$.

Example: $(x, y, z) = (1, 0, 0) \Rightarrow \rho = 1,\ \phi = \pi/2,\ \theta = 0$

$(x, y, z) = (0, 1, 0) \Rightarrow \rho = 1,\ \phi = \pi/2,\ \theta = \pi/2$

$(x, y, z) = (0, 0, 1) \Rightarrow \rho = 1,\ \phi = 0,\ \theta$ –can be anything

The volume element in spherical coordinates



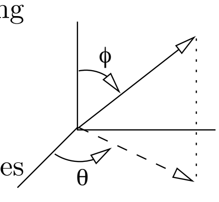
$$dV = \rho^2 \sin\phi\, d\rho\, d\phi\, d\theta.$$

The figure at right shows how we get this. The volume of the curved box is

$$\Delta V \approx \Delta\rho \cdot \rho\Delta\phi \cdot \rho\sin\phi\Delta\theta = \rho^2 \sin\phi\, \Delta\rho\, \Delta\phi\, \Delta\theta.$$

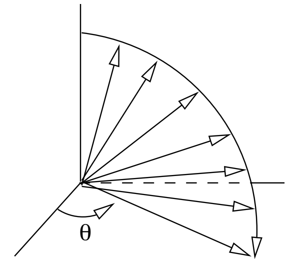
Finding limits in spherical coordinates

We use the same procedure as for rectangular and cylindrical coordinates. To calculate the limits for an iterated integral $\int\int\int_D d\rho\, d\phi\, d\theta$ over a region D in 3-space, we are integrating first with respect to ρ. Therefore we



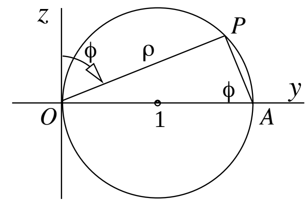
1. Hold ϕ and θ fixed, and let ρ increase. This gives us a ray going out from the origin.

2. Integrate from the ρ-value where the ray enters D to the ρ-value where the ray leaves D. This gives the limits on ρ.



3. Hold θ fixed and let ϕ increase. This gives a family of rays, that form a sort of fan. Integrate over those ϕ-values for which the rays intersect the region D.

4. Finally, supply limits on θ so as to include all of the fans which intersect the region D.



For example, suppose we start with the circle in the yz-plane of radius 1 and center at $(1, 0)$, rotate it about the z-axis, and take D to be that part of the resulting solid lying in the first octant.

First of all, we have to determine the equation of the surface formed by the rotated circle. In the yz-plane, the two coordinates ρ and ϕ are indicated. To see the relation between them when P is on the circle, we see that also angle $OAP = \phi$, since both the angle ϕ and OAP are complements of the same angle, AOP. From the right triangle, this shows the relation is $\rho = 2\sin\phi$.

As the circle is rotated around the z-axis, the relationship stays the same, so $\rho = 2\sin\phi$ is the equation of the whole surface.

To determine the limits of integration, when ϕ and θ are fixed, the corresponding ray enters the region where $\rho = 0$ and leaves where $\rho = 2\sin\phi$.

As ϕ increases, with θ fixed, it is the rays between $\phi = 0$ and $\phi = \pi/2$ that intersect D, since we are only considering the portion of the surface lying in the first octant (and thus above the xy-plane).

Again, since we only want the part in the first octant, we only use θ values from 0 to $\pi/2$. So the iterated integral is

$$\int_0^{\pi/2} \int_0^{\pi/2} \int_0^{2\sin\phi} d\rho\, d\phi\, d\theta.$$

18.02SC Multivariable Calculus
Fall 2010

Changing Variables in Multiple Integrals

4. Changing coordinates in triple integrals

Here the coordinate change will involve three functions

$$u = u(x, y, z), \quad v = v(x, y, z) \quad w = w(x, y, z)$$

but the general principles remain the same. The new coordinates u, v, and w give a three-dimensional grid, made up of the three families of contour surfaces of u, v, and w. Limits are put in by the kind of reasoning we used for double integrals. What we still need is the formula for the new volume element dV.

To get the volume of the little six-sided region ΔV of space bounded by three pairs of these contour surfaces, we note that nearby contour surfaces are approximately parallel, so that ΔV is approximately a parallelepiped, whose volume is (up to sign) the 3×3 determinant whose rows are the vectors forming the three edges of ΔV meeting at a corner. These vectors are calculated as in section 2; after passing to the limit we get

$$dV = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du\, dv\, dw \ , \tag{24}$$

where the key factor is the **Jacobian**

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{vmatrix} . \tag{25}$$

As an example, you can verify that this gives the correct volume element for the change from rectangular to spherical coordinates:

$$x = \rho \sin\phi \cos\theta, \qquad y = \rho \sin\phi \sin\theta, \qquad z = \rho \cos\phi; \tag{26}$$

while this is a good exercise, it will make you realize why most people prefer to derive the volume element in spherical coordinates by geometric reasoning.

Integrals in Spherical Coordinates

1. Find the volume of a sphere of radius a.

Answer: From the problems on limits in spherical coordinates (Session 76), we have limits: inner ρ: 0 to a –radial segments

middle ϕ: 0 to π –fan of rays.

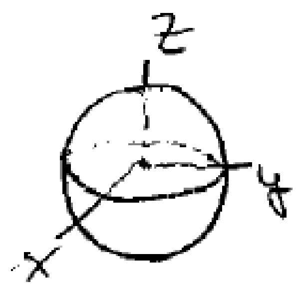
outer θ: 0 to 2π –volume.

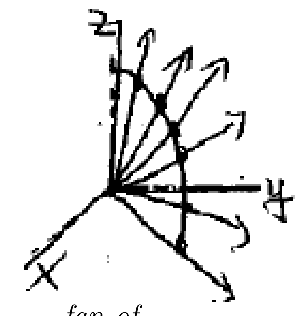
$$V = \iiint_D dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{a} \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$$

Inner: $\left.\dfrac{\rho^3}{3}\sin\phi\right|_0^a = \dfrac{a^3}{3}\sin\phi$

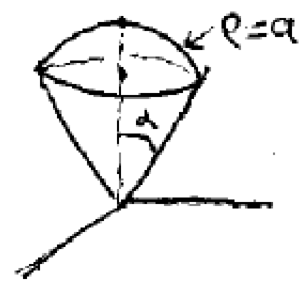
Middle: $\left.-\dfrac{a^3}{3}\cos\phi\right|_0^{\pi} = \dfrac{2}{3}a^3$

Outer: $\frac{4}{3}\pi a^3$ –as it should be.





2. Find the centroid of the region bounded by the sphere $\rho = a$ and the cone $\phi = \alpha$.



Answer: In Session 76 we computed the limits:

ρ: 0 to a, ϕ: 0 to α, θ: 0 to 2π.

By symmetry, $x_{cm} = y_{cm} = 0$.

$$z_{cm} = \frac{1}{V}\iiint_D z \, dV = \frac{1}{V}\int_0^{2\pi}\int_0^{\pi}\int_0^{a} \rho\cos\phi \cdot \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$$

$$= \frac{1}{V}\int_0^{2\pi}\int_0^{\pi}\int_0^{a} \rho^3 \cos\phi \sin\phi \, d\rho \, d\phi \, d\theta.$$

Inner, middle and outer integrals are easy to compute: $z_{cm} = \dfrac{1}{V}\cdot\dfrac{\pi a^4 \sin^2\alpha}{4}$.

$$V = \iiint_D dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{a} \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta = \frac{2}{3}\pi a^3(1 - \cos\alpha).$$

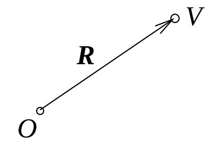
$$\Rightarrow \; z_{cm} = \frac{a^4 \sin^2\alpha \, \pi}{4} \cdot \frac{3}{2\pi a^3(1-\cos\alpha)} = \frac{3a}{8}\cdot\frac{\sin^2\alpha}{1-\cos\alpha} = \frac{3}{8}\,a(1+\cos\alpha).$$

Gravitational Attraction

We use triple integration to calculate the gravitational attraction that a solid body V of mass M exerts on a unit point mass placed at the origin.

If the solid V is also a point mass, then according to Newton's law of gravitation, the force it exerts is given by

$$\mathbf{F} \;=\; \frac{GM}{|\mathbf{R}|^2}\,\mathbf{r}, \tag{1}$$



where $\mathbf{R}$ is the position vector from the origin $\mathbf{0}$ to the point V, and the unit vector $\mathbf{r} = \mathbf{R}/|\mathbf{R}|$ is its direction.

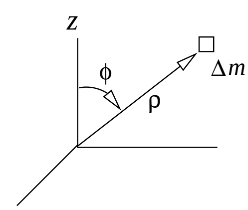
If however the solid body V is not a point mass, we have to use integration. We concentrate on finding just the $\mathbf{k}$ component of the gravitational attraction — all our examples will have the solid body V placed symmetrically so that its pull is all in the $\mathbf{k}$ direction anyway.

To calculate this force, we divide up the solid V into small pieces having volume ΔV and mass Δm. If the density function is $\delta(x, y, z)$, we have for the piece containing the point (x, y, z)

$$\Delta m \;\approx\; \delta(x, y, z)\Delta V, \tag{2}$$

Thinking of this small piece as being essentially a point mass at (x, y, z), the force $\Delta\mathbf{F}$ it exerts on the unit mass at the origin is given by (1), and its $\mathbf{k}$ component ΔF_z is therefore

$$\Delta F_z = \; G\,\frac{\Delta m}{|\mathbf{R}|^2}\,\mathbf{r}\cdot\mathbf{k}\;\;,$$



which in spherical coordinates becomes, using (2), and the picture,

$$\Delta F_z = \; G\,\frac{\cos\phi}{\rho^2}\,\delta\,\Delta V \;\;=\;\; G\,\frac{\delta\,\Delta V}{\rho^2}\;\cos\phi\;.$$

If we sum all the contributions to the force from each of the mass elements Δm and pass to the limit, we get for the $\mathbf{k}$-component of the gravitational force

$$F_z = G\iiint_V \frac{\cos\phi}{\rho^2}\,\delta\,dV\;. \tag{3}$$

If the integral is in spherical coordinates, then $dV = \rho^2\sin\phi\,d\rho\,d\phi\,d\theta$, and the integral becomes

$$F_z = G\iiint_V \delta\cos\phi\sin\phi\,d\rho\,d\phi\,d\theta\;. \tag{4}$$

Example 1. Find the gravitational attraction of the upper half of a solid sphere of radius a centered at the origin, if its density is given by $\delta = \sqrt{x^2+y^2}$.

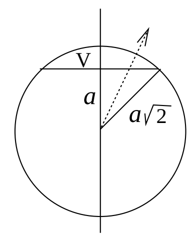
Solution. Since the solid and its density are symmetric about the z-axis, the force will be in the $\mathbf{k}$-direction, and we can use (3) or (4). Since

$$\sqrt{x^2+y^2} = r = \rho \sin\phi\ ,$$

the integral is

$$F_z = G\int_0^{2\pi}\int_0^{\pi/2}\int_0^a \rho\,\sin^2\phi\cos\phi\,d\rho\,d\phi\,d\theta$$

which evaluates easily to $\pi G a^2/3$.



Example 2. Let V be the solid spherical cap obtained by slicing a solid sphere of radius $a\sqrt{2}$ by a plane at a distance a from the center of the sphere. Find the gravitational attraction of V on a unit point mass at the center of the sphere. (Take the density to be 1.)

Solution. To take advantage of the symmetry, place the origin at the center of the sphere, and align the axis of the cap along the z-axis (so the flat side of the cap is parallel to the xy-plane).

We use spherical coordinates; the main problem is determining the limits of integration. If we fix ϕ and θ and let ρ vary, we get a ray which enters V at its flat side

$$z = a, \quad \text{or} \quad \rho\,\cos\phi = a,$$

and leaves V on its spherical side, $\rho = a\sqrt{2}$. The rays which intersect V in this way are those for which $0 \le \phi \le \pi/4$, as one sees from the picture. Thus by (4),

$$F_z \;=\; G\int_0^{2\pi}\int_0^{\pi/4}\int_{a/\cos\phi}^{a\sqrt{2}} \sin\phi\cos\phi\,d\rho\,d\phi\,d\theta,$$

which after integrating with respect to ρ (and θ) becomes

$$\begin{aligned} &= \;2\pi G\int_0^{\pi/4} a\Big(\sqrt{2}-\frac{1}{\cos\phi}\Big)\sin\phi\cos\phi\,d\phi \\ &= \;2\pi G a\Big(\frac{3\sqrt{2}}{4}-1\Big)\ . \end{aligned}$$

Remark. Newton proved that a solid sphere of uniform density and mass M exerts the same force on an external point mass as would a point mass M placed at the center of the sphere. (See Problem 6a in problem section 5C).

This does *not* however generalize to other uniform solids of mass M — it is not true that the gravitational force they exert is the same as that of a point mass M at their center of mass. For if this were so, a unit test mass placed on the axis between two equal point masses M and M' ought to be pulled toward the midposition, whereas actually it will be pulled toward the closer of the two masses.

V8. Vector Fields in Space

Just as in Section V1 we considered vector fields in the plane, so now we consider vector fields in three-space. These are fields given by a vector function of the type

$$(1) \qquad \mathbf{F}(x, y, z) \;=\; M(x, y, z)\,\mathbf{i} + N(x, y, z)\,\mathbf{j} + P(x, y, z)\,\mathbf{k}\ .$$

Such a function assigns the vector $\mathbf{F}(x_0, y_0, z_0)$ to a point (x_0, y_0, z_0) where M, N, and P are all defined. We place the vector so its tail is at (x_0, y_0, z_0), and in this way get the vector field. Such a field in space looks a little like the interior of a haystack.

As before, we say $\mathbf{F}$ is *continuous* in some domain D of 3-space (we will usually use "domain" rather than "region" , when referring to a portion of 3-space) if M, N, and P are continuous in that domain. We say $\mathbf{F}$ is *continuously differentiable* in the domain D if all nine first partial derivatives

$$M_x,\ M_y,\ M_z;\ \ N_x,\ N_y,\ N_z;\ \ P_x,\ P_y,\ P_z$$

exist and are continuous in D.

Again as before, we give two physical interpretations for such a vector field.

The three-dimensional **force fields** of different sorts — gravitational, electrostatic, electromagnetic — all give rise to such a vector field: at the point (x_0, y_0, z_0) we place the vector having the direction and magnitude of the force which the field would exert on a unit test particle placed at the point.

The three-dimensional **flow fields** and **velocity fields** arising from the motion of a fluid in space are the other standard example. We assume the motion is steady-state (i.e., the direction and magnitude of the flow at any point does not change over time). We will call this a *three-dimensional flow.*

As before, we allow sources and sinks — places where fluid is being added to or removed from the flow. Obviously, we can no longer appeal to people standing overhead pouring fluid in at various points (they would have to be aliens in four-space), but we could think of thin pipes inserted into the domain at various points adding or removing fluid.

The velocity field of such a flow is defined just as it was previously: $\mathbf{v}(x, y, z)$ gives the direction and magnitude (speed) of the flow at (x, y, z).

The flow field $\mathbf{F} = \delta\mathbf{v}$, where $\delta(x, y, z)$ is the density) may be similarly interpreted:

$$\begin{aligned} \text{dir}\ \mathbf{F} \;&=\; \text{the direction of flow} \\ (2) \qquad |\mathbf{F}| \;&=\; \text{mass transport rate (per unit area) at } (x, y, z) \text{ in the flow direction;} \end{aligned}$$

that is, $|\mathbf{F}|$ is the rate per unit area at which mass is transported across a small piece of plane perpendicular to the flow at the point (x, y, z)..

The derivation of this interpretation is exactly as in Sections V1 and V3, replacing the small line segment Δl by a small plane area ΔA perpendicular to the flow.

Example 1. Find the three-dimensional electrostatic force field $\mathbf{F}$ arising from a unit positive charge placed at the origin, given that in suitable units $\mathbf{F}$ is directed radially outward from the origin and has magnitude $1/\rho^2$, where ρ is the distance from the origin.

Solution. The vector $x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ with tail at (x, y, z) is directed radially outward and has magnitude ρ. Therefore

$$\mathbf{F} \;=\; \frac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{\rho^3}\;, \qquad \rho = \sqrt{x^2 + y^2 + z^2}$$

Example 2. a) Find the velocity field of a fluid rotating with constant angular velocity ω around the z-axis, in the direction given by the right-hand rule (right-hand fingers curl in direction of flow when thumb points in the $\mathbf{k}$-direction).

b) Find the analogous field if the flow is rotating about the y-axis.

Solution. a) The flow doesn't depend on z — it is really just a two-dimensional problem, whose solution is the same as before (section V1, Example 4):

$$\mathbf{F}(x, y, z) \;=\; \omega(-y\,\mathbf{i} + x\,\mathbf{j}\,)\;.$$

b) If the axis of flow is the y-axis, the flow will have no $\mathbf{j}$-component and will not depend on y. However, by the right-hand rule, the flow in the xz-plane is *clockwise*, when the positive x and z axes are drawn so as to give a right-handed system. Thus

$$\mathbf{F}(x, y, z) \;=\; \omega(z\,\mathbf{i} - x\,\mathbf{k}\,)\;.$$

Example 3. Find the three-dimensional flow field of a gas streaming radially outward with constant velocity from a source at the origin of constant strength.

Solution. This is like the corresponding two-dimensional problem (section V1, Example 3), except that the area of a sphere increases like the *square* of its radius. Therefore, to maintain constant velocity, the density of flow must decrease like $1/\rho^2$ as you go out from the origin; letting δ be the density and c_i be constants, we get

$$\mathbf{F}(x, y, z) \;=\; \delta\mathbf{v} \;=\; \frac{c_1}{\rho^2}\;\frac{c_2(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}\,)}{\rho} \;=\; \frac{c(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}\,)}{\rho^3}\;.$$

Notice that in the three-dimensional case, this field is the same as the one in Example 1 above, with the magnitude falling off like $1/\rho^2$. For the two-dimensional case, the analogue of a point fluid source at the origin is not a point charge at the origin, but a uniform charge along a vertical wire; both give the field whose magnitude falls off like $1/r$.

V9.1 Surface Integrals

Surface integrals are a natural generalization of line integrals: instead of integrating over a curve, we integrate over a surface in 3-space. Such integrals are important in any of the subjects that deal with continuous media (solids, fluids, gases), as well as subjects that deal with force fields, like electromagnetic or gravitational fields.

Though most of our work will be spent seeing how surface integrals can be calculated and what they are used for, we first want to indicate briefly how they are defined. The surface integral of the (continuous) function $f(x, y, z)$ over the surface S is denoted by

$$(1) \qquad \iint_S f(x, y, z)\, dS \ .$$

You can think of dS as the area of an infinitesimal piece of the surface S. To define the integral (1), we subdivide the surface S into small pieces having area ΔS_i, pick a point (x_i, y_i, z_i) in the i-th piece, and form the Riemann sum

$$(2) \qquad \sum f(x_i, y_i, z_i)\Delta S_i \ .$$

w As the subdivision of S gets finer and finer, the corresponding sums (2) approach a limit which does not depend on the choice of the points or how the surface was subdivided. The surface integral (1) is defined to be this limit. (The surface has to be smooth and not infinite in extent, and the subdivisions have to be made reasonably, otherwise the limit may not exist, or it may not be unique.)

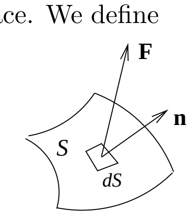
1. The surface integral for flux.

The most important type of surface integral is the one which calculates the flux of a vector field across S. Earlier, we calculated the flux of a plane vector field $\mathbf{F}(x, y)$ across a directed curve in the xy-plane. What we are doing now is the analog of this in space.

We assume that S is *oriented*: this means that S has two sides and one of them has been designated to be the *positive side*. At each point of S there are two unit normal vectors, pointing in opposite directions; the *positively directed* unit normal vector, denoted by $\mathbf{n}$, is the one standing with its base (i.e., tail) on the positive side. If S is a closed surface, like a sphere or cube — that is, a surface with no boundaries, so that it completely encloses a portion of 3-space — then by convention it is oriented so that the outer side is the positive one, i.e., so that $\mathbf{n}$ always points towards the outside of S.

Let $\mathbf{F}(x, y, z)$ be a continuous vector field in space, and S an oriented surface. We define

$$(3) \qquad \textbf{flux of } F \textbf{ through } S \;=\; \iint_S (\mathbf{F}\cdot\mathbf{n})\, dS \;=\; \iint_S \mathbf{F}\cdot d\mathbf{S}\ ;$$



the two integrals are the same, but the second is written using the common and suggestive abbreviation $d\mathbf{S} = \mathbf{n}\, dS$.

If $\mathbf{F}$ represents the velocity field for the flow of an incompressible fluid of density 1, then $\mathbf{F}\cdot\mathbf{n}$ represents the component of the velocity in the positive perpendicular direction to the surface, and $\mathbf{F}\cdot\mathbf{n}\, dS$ represents the flow rate across the little infinitesimal piece of surface having area dS. The integral in (3) adds up these flows across the pieces of surface, so that we may interpret (3) as saying

$$\textit{flux of } F \textit{ through } S \;=\; \textit{net flow rate across } S, \tag{4}$$

where we count flow in the direction of $\mathbf{n}$ as positive, flow in the opposite direction as negative. More generally, if the fluid has varying density, then the right side of (4) is the net mass transport rate of fluid across S (per unit area, per time unit).

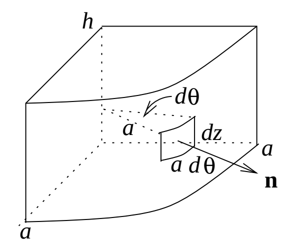
If $\mathbf{F}$ is a force field, then nothing is physically flowing, and one just uses the term "flux" to denote the surface integral, as in (3).

V9.2 Surface Integrals

2. Flux through a cylinder and sphere.

We now show how to calculate the flux integral, beginning with two surfaces where $\mathbf{n}$ and dS are easy to calculate — the cylinder and the sphere.

Example 1. Find the flux of $\mathbf{F} = z\,\mathbf{i} + x\,\mathbf{j} + y\,\mathbf{k}$ outward through the portion of the cylinder $x^2 + y^2 = a^2$ in the first octant and below the plane $z = h$.



Solution. The piece of cylinder is pictured. The word "outward" suggests that we orient the cylinder so that $\mathbf{n}$ points outward, i.e., away from the z-axis. Since by inspection $\mathbf{n}$ is radially outward and horizontal,

$$\mathbf{n} = \frac{x\,\mathbf{i} + y\,\mathbf{j}}{a} . \tag{5}$$

(This is the outward normal to the circle $x^2 + y^2 = a^2$ in the xy-plane; $\mathbf{n}$ has no z-component since it is horizontal. We divide by a to make its length 1.)

To get dS, the infinitesimal element of surface area, we use cylindrical coordinates to parametrize the cylinder:

$$x = a\cos\theta, \qquad y = a\sin\theta \qquad z = z . \tag{6}$$

As the parameters θ and z vary, the whole cylinder is traced out ; the piece we want satisfies $0 \le \theta \le \pi/2,\ 0 \le z \le h$. The natural way to subdivide the cylinder is to use little pieces of curved rectangle like the one shown, bounded by two horizontal circles and two vertical lines on the surface. Its area dS is the product of its height and width:

$$dS = dz \cdot a\,d\theta . \tag{7}$$

Having obtained $\mathbf{n}$ and dS, the rest of the work is routine. We express the integrand of our surface integral (3) in terms of z and θ:

$$\begin{aligned} \mathbf{F}\cdot\mathbf{n}\,dS &= \frac{zx + xy}{a}\cdot a\,dz\,d\theta , && \text{by (5) and (7);} \\ &= (az\cos\theta + a^2\sin\theta\cos\theta)\,dz\,d\theta, && \text{using (6).} \end{aligned}$$

This last step is essential, since the dz and $d\theta$ tell us the surface integral will be calculated in terms of z and θ, and therefore the integrand must use these variables also. We can now calculate the flux through S:

$$\begin{aligned} \int\!\!\int_S \mathbf{F}\cdot\mathbf{n}\,dS &= \int_0^{\pi/2}\int_0^h (az\cos\theta + a^2\sin\theta\cos\theta)\,dz\,d\theta \\ \text{inner integral} &= \frac{ah^2}{2}\cos\theta + a^2h\sin\theta\cos\theta \\ \text{outer integral} &= \left[\frac{ah^2}{2}\sin\theta + a^2h\frac{\sin^2\theta}{2}\right]_0^{\pi/2} = \frac{ah}{2}(a+h) . \end{aligned}$$

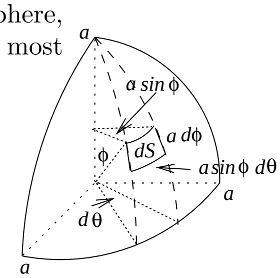
Example 2. Find the flux of $\mathbf{F} = xz\,\mathbf{i} + yz\,\mathbf{j} + z^2\,\mathbf{k}$ outward through that part of the sphere $x^2 + y^2 + z^2 = a^2$ lying in the first octant $(x, y, z, \ge 0)$.

Solution. Once again, we begin by finding $\mathbf{n}$ and dS for the sphere. We take the outside of the sphere as the positive side, so $\mathbf{n}$ points radially outward from the origin; we see by inspection therefore that

$$\mathbf{n} = \frac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{a}\,, \tag{8}$$

where we have divided by a to make $\mathbf{n}$ a unit vector.

To do the integration, we use spherical coordinates ρ, ϕ, θ. On the surface of the sphere, $\rho = a$, so the coordinates are just the two angles ϕ and θ. The area element dS is most easily found using the volume element:



$$dV = \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta = dS \cdot d\rho \quad = \quad \text{area} \cdot \text{thickness}$$

so that dividing by the thickness $d\rho$ and setting $\rho = a$, we get

$$dS = a^2 \sin\phi \, d\phi \, d\theta. \tag{9}$$

Finally since the area element dS is expressed in terms of ϕ and θ, the integration will be done using these variables, which means we need to express x, y, z in terms of ϕ and θ. We use the formulas expressing Cartesian in terms of spherical coordinates (setting $\rho = a$ since (x, y, z) is on the sphere):

$$x = a\sin\phi\cos\theta, \qquad y = a\sin\phi\sin\theta, \qquad z = a\cos\phi\,. \tag{10}$$

We can now calculate the flux integral (3). By (8) and (9), the integrand is

$$\mathbf{F}\cdot\mathbf{n}\,dS \;=\; \frac{1}{a}\,(x^2 z + y^2 z + z^2 z)\cdot a^2 \sin\phi\,d\phi\,d\theta\,.$$

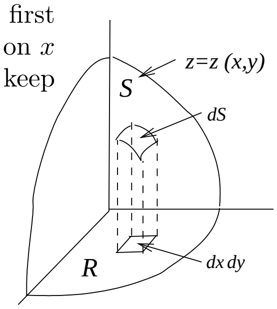
Using (10), and noting that $x^2 + y^2 + z^2 = a^2$, the integral becomes

$$\begin{aligned}\iint_S \mathbf{F}\cdot\mathbf{n}\;dS \;&=\; a^4 \int_0^{\pi/2}\int_0^{\pi/2} \cos\phi\sin\phi\,d\phi\,d\theta \\ &=\; a^4\,\frac{\pi}{2}\,\frac{1}{2}\,\sin^2\phi\,\Big]_0^{\pi/2} \;=\; \frac{\pi a^4}{4}\,.\end{aligned}$$

V9.3-4 Surface Integrals

3. Flux through general surfaces.

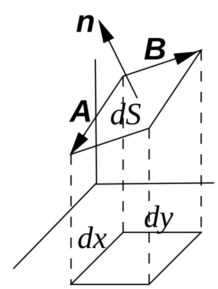
For a general surface, we will use xyz-coordinates. It turns out that here it is simpler to calculate the infinitesimal vector $d\mathbf{S} = \mathbf{n}\,dS$ directly, rather than calculate $\mathbf{n}$ and dS separately and multiply them, as we did in the previous section. Below are the two standard forms for the equation of a surface, and the corresponding expressions for $d\mathbf{S}$. In the first we use z both for the dependent variable and the function which gives its dependence on x and y; you can use $f(x, y)$ for the function if you prefer, but that's one more letter to keep track of.



(11a) $\qquad z = z(x, y), \qquad d\mathbf{S} = (-z_x\,\mathbf{i} - z_y\,\mathbf{j} + \mathbf{k})\,dx\,dy \qquad$ ($\mathbf{n}$ points "up")

(11b) $\qquad F(x, y, z) = c, \qquad d\mathbf{S} = \pm\dfrac{\nabla F}{F_z}\,dx\,dy \qquad$ (choose the right sign);

Derivation of formulas for $d\mathbf{S}$.



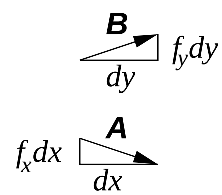
Refer to the pictures at the right. The surface S lies over its projection R, a region in the xy-plane. We divide up R into infinitesimal rectangles having area $dx\,dy$ and sides parallel to the xy-axes — one of these is shown. Over it lies a piece dS of the surface, which is approximately a parallelogram, since its sides are approximately parallel.

The infinitesimal vector $d\mathbf{S} = \mathbf{n}\,dS$ we are looking for has

direction: perpendicular to the surface, in the "up" direction;
magnitude: the area dS of the infinitesimal parallelogram.

This shows our infinitesimal vector is the cross-product

$$d\mathbf{S} = \mathbf{A} \times \mathbf{B}$$



where $\mathbf{A}$ and $\mathbf{B}$ are the two infinitesimal vectors forming adjacent sides of the parallelogram. To calculate these vectors, from the definition of the partial derivative, we have

$\mathbf{A}$ lies over the vector $dx\,\mathbf{i}$ and has slope f_x in the $\mathbf{i}$ direction, so $\mathbf{A} = dx\,\mathbf{i} + f_x\,dx\,\mathbf{k}$;
$\mathbf{B}$ lies over the vector $dy\,\mathbf{j}$ and has slope f_y in the $\mathbf{j}$ direction, so $\mathbf{B} = dy\,\mathbf{j} + f_y\,dy\,\mathbf{k}$.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ dx & 0 & f_x dx \\ 0 & dy & f_y dy \end{vmatrix} = (-f_x\,\mathbf{i} - f_y\,\mathbf{j} + \mathbf{k})\,dx\,dy\ ,$$

which is (11a).

To get (11b) from (11a), , our surface is given by

(12) $$F(x, y, z) = c, \qquad z = z(x, y)$$

where the right-hand equation is the result of solving $F(x, y, z) = c$ for z in terms of the independent variables x and y. We differentiate the left-hand equation in (12) with respect to the independent variables x and y, using the chain rule and remembering that $z = z(x, y)$:

$$F(x, y, z) = c \quad \Rightarrow \quad F_x\frac{\partial x}{\partial x} + F_y\frac{\partial y}{\partial x} + F_z\frac{\partial z}{\partial x} = 0 \quad \Rightarrow \quad F_x + F_z\frac{\partial z}{\partial x} = 0$$

from which we get

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \qquad \text{and similarly,} \qquad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z} \ .$$

Therefore by (11a),

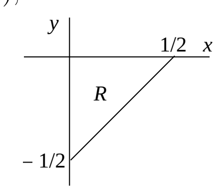
$$d\mathbf{S} \ = \ \left(-\frac{\partial z}{\partial x}\,\mathbf{i} - \frac{\partial z}{\partial y}\,\mathbf{j} + 1\right) dx\, dy \ = \ \left(\frac{F_x}{F_z}\,\mathbf{i} + \frac{F_y}{F_z}\,\mathbf{j} + 1\right) dx\, dy \ = \ \frac{\nabla F}{F_z}\, dx\, dy \ ,$$

which is (11b).

Example 3. The portion of the plane $2x - 2y + z = 1$ lying in the first octant forms a triangle S. Find the flux of $\mathbf{F} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ through S; take the positive side of S as the one where the normal points “up”.

Solution. Writing the plane in the form $z = 1 - 2x + 2y$, we get using (11a),

$$\begin{aligned} d\mathbf{S} \ &= \ (2\,\mathbf{i} \,-\, 2\,\mathbf{j} \,+\, \mathbf{k}\,)\, dx\, dy \ , \qquad \text{so} \\ \iint_S \mathbf{F}\cdot d\mathbf{S} \ &= \ \iint_S (2x - 2y + z)\, dy\, dx \\ &= \ \iint_R \big(2x - 2y + (1 - 2x + 2y)\big)\, dy\, dx \ , \end{aligned}$$



where R is the region in the xy-plane over which S lies. (Note that since the integration is to be in terms of x and y, we had to express z in terms of x and y for this last step.) To see what R is explicitly, the plane intersects the three coordinate axes respectively at $x = 1/2,\ y = -1/2,\ z = 1$. So R is the region pictured; our integral has integrand 1, so its value is the area of R, which is $1/8$.

Remark. When we write $z = f(x, y)$ or $z = z(x, y)$, we are agreeing to parametrize our surface using x and y as parameters. Thus the flux integral will be reduced to a double integral over a region R in the xy-plane, involving only x and y. Therefore you must *get rid of z by using the relation $z = z(x, y)$* after you have calculated the flux integral using (11a). Then determine the region R (the projection of S onto the xy-plane), and supply the limits for the iterated integral over R.

Example 4. Set up a double integral in the xy-plane which gives the flux of the field $\mathbf{F} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ through that portion of the ellipsoid $4x^2 + y^2 + 4z^2 = 4$ lying in the first octant; take $\mathbf{n}$ in the “up” direction.

Solution. Using (11b), we have $d\,\mathbf{S} = \dfrac{\langle 8x, 2y, 8z\rangle}{8z}\, dx\, dy$. Therefore

$$\iint_S \mathbf{F}\cdot d\,\mathbf{S} = \iint_S \frac{8x^2 + 2y^2 + 8z^2}{8z}\, dx\, dy = \iint_S \frac{1}{z}\, dx\, dy = \iint_R \frac{dx\, dy}{\sqrt{1 - x^2 - (y/2)^2}} \ ,$$

where R is the portion of the ellipse $4x^2 + y^2 = 4$ lying the the first quadrant.

The double integral would be most simply evaluated by making the change of variable $u = y/2$, which would convert it to a double integral over a quarter circle in the xu-plane easily evaluated by a change to polar coordinates.

4. General surface integrals.* The surface integral $\iint_S f(x,y,z)\,dS$ that we introduced at the beginning can be used to calculate things other than flux.

a) **Surface area.** We let the function $f(x,y,z)=1$. Then the area of $S = \iint_S dS$.

b) **Mass, moments, charge.** If S is a thin shell of material, of uniform thickness, and with density (in gms/unit area) given by $\delta(x,y,z)$, then

$$\text{mass of } S = \iint_S \delta(x,y,z)\,dS, \tag{13}$$

$$x\text{-component of center of mass} = \overline{x} = \frac{1}{\text{mass } S}\iint_S x\cdot\delta\,dS \tag{14}$$

with the y- and z-components of the center of mass defined similarly. If $\delta(x,y,z)$ represents an electric charge density, then the surface integral (13) will give the total charge on S.

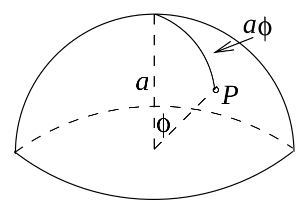
c) **Average value.** The average value of a function $f(x,y,z)$ over the surface S can be calculated by a surface integral:

$$\text{average value of } f \text{ on } S = \frac{1}{\text{area } S}\iint_S f(x,y,z)\,dS\ . \tag{15}$$

Calculating general surface integrals; finding dS.

To evaluate general surface integrals we need to know dS for the surface. For a sphere or cylinder, we can use the methods in section 2 of this chapter.

Example 5. Find the average distance along the earth of the points in the northern hemisphere from the North Pole. (Assume the earth is a sphere of radius a.)



Solution. — We use (15) and spherical coordinates, choosing the coordinates so the North Pole is at $z=a$ on the z-axis. The distance of the point (a,ϕ,θ) from $(a,0,0)$ is $a\phi$, measured along the great circle, i.e., the longitude line — see the picture). We want to find the average of this function over the upper hemisphere S. Integrating, and using (9), we get

$$\iint_S a\phi\,dS = \int_0^{2\pi}\int_0^{\pi/2} a\phi a^2\sin\phi\,d\phi\,d\theta = 2\pi a^3\int_0^{\pi/2}\phi\,\sin\phi\,d\phi = 2\pi a^3\ .$$

(The last integral used integration by parts.) Since the area of $S = 2\pi a^2$, we get using (15) the striking answer: average distance $= a$.

For more general surfaces given in xyz-coordinates, since $d\mathbf{S} = \mathbf{n}\,dS$, the area element dS is the magnitude of $d\mathbf{S}$. Using (11a) and (11b), this tells us

$$z = z(x,y), \qquad dS = \sqrt{z_x^2+z_y^2+1}\;dx\,dy \tag{16a}$$

$$F(x,y,z) = c, \qquad dS = \frac{|\nabla F|}{|F_z|}\;dx\,dy \tag{16b}$$

Example 6. The area of the piece S of $z=xy$ lying over the unit circle R in the xy-plane is calculated by (a) above and (16a) to be:

$$\iint_S dS = \iint_R \sqrt{y^2+x^2+1}\,dx\,dy = \int_0^{2\pi}\int_0^1 \sqrt{r^2+1}\,r\,dr\,d\theta = 2\pi\cdot\frac{1}{3}(r^2+1)^{3/2}\Big]_0^1 = \frac{2\pi}{3}(2\sqrt{2}-1).$$

General Formula for $\mathbf{n}\,dS$

Suppose S is a surface parametrized by x and y and $\mathbf{N}$ is any vector normal to S (not necessarily unit length). Then $\mathbf{n}\,dS = \dfrac{\mathbf{N}}{\mathbf{N}\cdot\mathbf{k}}\,dx\,dy$. Here $\mathbf{n}$ is the upward unit normal

Example: for the sphere $x^2 + y^2 + z^2 = a^2$ with $\mathbf{N} = \langle x, y, z\rangle$, find $\mathbf{n}\,dS$.

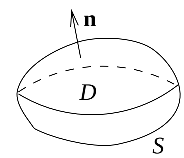
<u>Answer:</u> $\mathbf{n}\,dS = \dfrac{\mathbf{N}}{\mathbf{N}\cdot\mathbf{k}}\,dx\,dy = \langle \frac{x}{z}, \frac{y}{z}, 1\rangle\,dx\,dy$.

(Just like if we wrote $z = \sqrt{a^2 - x^2 - y^2}$, $\mathbf{n}\,dS = \langle -z_x, -z_y, 1\rangle\,dx\,dy$.)

V10.1 The Divergence Theorem

1. Introduction; statement of the theorem.

The divergence theorem is about closed surfaces, so let's start there. By a **closed** surface S we will mean a surface consisting of one connected piece which doesn't intersect itself, and which completely encloses a single finite region D of space called its *interior*. The closed surface S is then said to be the *boundary* of D; we include S in D. A sphere, cube, and torus (an inflated bicycle inner tube) are all examples of closed surfaces. On the other hand, these are not closed surfaces: a plane, a sphere with one point removed, a tin can whose cross-section looks like a figure-8 (it intersects itself), an infinite cylinder.



A closed surface always has two sides, and it has a natural positive direction — the one for which $\mathbf{n}$ points away from the interior, i.e., points toward the outside. We shall always understand that the closed surface has been oriented this way, unless otherwise specified.

We now generalize to 3-space the normal form of Green's theorem (Section V4).

Definition. Let $\mathbf{F}(x, y, z) = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ be a vector field differentiable in some region D. By the **divergence** of $\mathbf{F}$ we mean the scalar function $\operatorname{div} \mathbf{F}$ of three variables defined in D by

$$\text{(1)} \qquad \operatorname{div} \mathbf{F} = \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}\ .$$

The divergence theorem. *Let S be a positively-oriented closed surface with interior D, and let $\mathbf{F}$ be a vector field continuously differentiable in a domain contatining D. Then*

$$\text{(2)} \qquad \iint_S \mathbf{F} \cdot d\mathbf{S} = \iiint_D \operatorname{div} \mathbf{F}\, dV$$

We write dV on the right side, rather than $dx\,dy\,dz$ since the triple integral is often calculated in other coordinate systems, particularly spherical coordinates. The theorem is sometimes called **Gauss' theorem**.

Physically, the divergence theorem is interpreted just like the normal form for Green's theorem. Think of $\mathbf{F}$ as a three-dimensional flow field. Look first at the left side of (2). The surface integral represents the mass transport rate across the closed surface S, with flow out of S considered as positive, flow into S as negative.

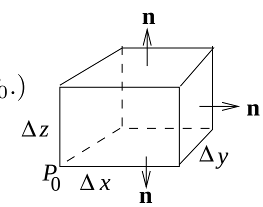
Look now at the right side of (2). In what follows, we will show that the value of div $\mathbf{F}$ at (x, y, z) can be interpreted as the **source rate** at (x, y, z): the rate at which fluid is being added to the flow at this point. (Negative rate means fluid is being removed from the flow.) The integral on the right of (2) thus represents the *source rate for D*. So what the divergence theorem says is:

$$\text{(3)} \qquad \textit{flux across } S = \textit{source rate for } D\ ;$$

i.e., the net flow outward across S is the same as the rate at which fluid is being produced (or added to the flow) inside S.

To complete the argument for (3) we still have to show that

$$\text{div}\,\mathbf{F} \;=\; \textit{source rate at } (x,y,z)\;. \tag{3}$$



To see this, let $P_0 : (x_0, y_0, z_0)$ be a point inside the region D where $\mathbf{F}$ is defined. (To simplify, we denote by $(\text{div}\,\mathbf{F})_0, (\partial M/\partial x)_0$, etc., the value of these functions at P_0.)

Consider a little rectangular box, with edges $\Delta x, \Delta y, \Delta z$ parallel to the coordinate axes, and one corner at P_0. We take $\mathbf{n}$ to be always pointing outwards, as usual; thus on top of the box $\mathbf{n} = \mathbf{k}$, but on the bottom face, $\mathbf{n} = -\mathbf{k}$.

The flux across the top face in the $\mathbf{n}$ direction is approximately

$$\mathbf{F}(x_0, y_0, z_0 + \Delta z)\cdot \mathbf{k}\;\, \Delta x\Delta y \;=\; P(x_0, y_0, z_0 + \Delta z)\;\Delta x\Delta y,$$

while the flux across the bottom face in the $\mathbf{n}$ direction is approximately

$$\mathbf{F}(x_0, y_0, z_0)\cdot -\mathbf{k}\;\, \Delta x\Delta y \;=\; -P(x_0, y_0, z_0)\;\Delta x\Delta y\;.$$

So the net flux across the two faces combined is approximately

$$[P(x_0, y_0, z_0 + \Delta z) - P(x_0, y_0, z_0)]\;\Delta x\Delta y \;=\; \left(\frac{\Delta P}{\Delta z}\right)\Delta x\Delta y\Delta z\;.$$

Since the difference quotient is approximately equal to the partial derivative, we get the first line below; the reasoning for the following two lines is analogous:

$$\begin{aligned}
\textit{net flux across top and bottom} &\approx \left(\frac{\partial P}{\partial z}\right)_0 \Delta x\Delta y\Delta z;\\
\textit{net flux across two side faces} &\approx \left(\frac{\partial N}{\partial y}\right)_0 \Delta x\Delta y\Delta z;\\
\textit{net flux across front and back} &\approx \left(\frac{\partial M}{\partial x}\right)_0 \Delta x\Delta y\Delta z;
\end{aligned}$$

Adding up these three net fluxes, and using (3), we see that

$$\begin{aligned}
\textit{source rate for box} &= \textit{net flux across faces of box}\\
&\approx \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z}\right)_0 \Delta x\Delta y\Delta z\;.
\end{aligned}$$

Using this, we get the interpretation for div $\mathbf{F}$ we are seeking:

$$\textit{source rate at } P_0 \;=\; \lim_{\text{box}\to 0} \frac{\textit{source rate for box}}{\textit{volume of box}} \;=\; (\text{div}\,\mathbf{F})_0\;.$$

Example 1. Verify the theorem when $\mathbf{F} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ and S is the sphere $\rho = a$.

Solution. For the sphere, $\mathbf{n} = \dfrac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{a}$; thus $\mathbf{F}\cdot\mathbf{n} = a$, and $\displaystyle\iint_S \mathbf{F}\cdot d\mathbf{S} = 4\pi a^3$.

On the other side, div $\mathbf{F} = 3$, $\iiint_D 3\, dV = 3 \cdot \frac{4}{3}\pi a^3$; thus the two integrals are equal. $\square$

Example 2. Use the divergence theorem to evaluate the flux of $\mathbf{F} = x^3\,\mathbf{i} + y^3\,\mathbf{j} + z^3\,\mathbf{k}$ across the sphere $\rho = a$.

Solution. Here div $\mathbf{F} = 3(x^2 + y^2 + z^2) = 3\rho^2$. Therefore by (2),

$$\iint_S \mathbf{F} \cdot d\mathbf{S} \;=\; 3\iiint_D \rho^2\, dV \;=\; 3\int_0^a \rho^2 \cdot 4\pi\rho^2\, d\rho \;=\; \frac{12\pi a^5}{5}\;;$$

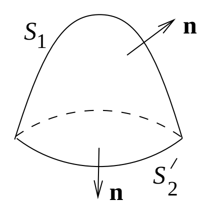
we did the triple integration by dividing up the sphere into thin concentric spheres, having volume $dV = 4\pi\rho^2\, d\rho$.

Example 3. Let S_1 be that portion of the surface of the paraboloid $z = 1 - x^2 - y^2$ lying above the xy-plane, and let S_2 be the part of the xy-plane lying inside the unit circle, directed so the normal $\mathbf{n}$ points upwards. Take $\mathbf{F} = yz\,\mathbf{i} + xz\,\mathbf{j} + xy\,\mathbf{k}$; evaluate the flux of $\mathbf{F}$ across S_1 by using the divergence theorem to relate it to the flux across S_2.

Solution. We see immediately that div $\mathbf{F} = 0$. Therefore, if we let S_2' be the same surface as S_2, but oppositely oriented (so $\mathbf{n}$ points downwards), the surface $S_1 + S_2'$ is a closed surface, with $\mathbf{n}$ pointing outwards everywhere. Hence by the divergence theorem,

$$\iint_{S_1+S_2'} \mathbf{F} \cdot d\mathbf{S} \;=\; 0 \;=\; \iint_{S_1} \mathbf{F} \cdot d\mathbf{S} - \iint_{S_2} \mathbf{F} \cdot d\mathbf{S}$$

Therefore, since we have $\mathbf{n} = \mathbf{k}$ on S_2,



$$\iint_{S_1} \mathbf{F} \cdot d\mathbf{S} \;=\; \iint_{S_2} \mathbf{F} \cdot d\mathbf{S} \;=\; \iint_{S_2} \mathbf{F} \cdot \mathbf{k}\; dS \;=\; \iint_{S_2} xy\, dx\, dy$$
$$=\; 0,$$

by integrating in polar coordinates (or by symmetry).

18.02SC Multivariable Calculus
Fall 2010

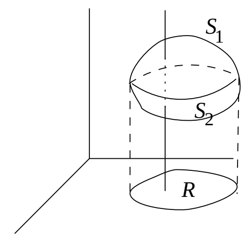
V10.2 The Divergence Theorem

2. Proof of the divergence theorem.

We give an argument assuming first that the vector field $\mathbf{F}$ has only a $\mathbf{k}$-component: $\mathbf{F} = P(x,y,z)\,\mathbf{k}$. The theorem then says

$$\iint_S P\,\mathbf{k}\cdot\mathbf{n}\,dS \;=\; \iiint_D \frac{\partial P}{\partial z}\,dV\;. \tag{4}$$

The closed surface S projects into a region R in the xy-plane. We assume S is vertically simple, i.e., that each vertical line over the interior of R intersects S just twice. (S can have vertical sides, however — a cylinder would be an example.) S is then described by two equations:



(5) $\quad z = g(x,y)$ *(lower surface);* $\quad z = h(x,y)$ *(upper surface)*

The strategy of the proof of (4) will be to reduce each side of (4) to a double integral over R; the two double integrals will then turn out to be the same.

We do this first for the triple integral on the right of (4). Evaluating it by iteration, we get as the first step in the iteration,

$$\begin{aligned}\iiint_D \frac{\partial P}{\partial z}\,dV \;&=\; \iint_R \int_{g(x,y)}^{h(x,y)} \frac{\partial P}{\partial z}\,dz\,dx\,dy\\ &=\; \iint_R \big(P(x,y,h) - P(x,y,g)\big)\,dx\,dy \end{aligned} \tag{6}$$

To calculate the surface integral on the left of (4), we use the formula for the surface area element $d\mathbf{S}$ given in V9, (13):

$$d\mathbf{S} \;=\; \pm(-z_x\,\mathbf{i} - z_y\,\mathbf{j} + k)\,dx\,dy,$$

where we use the $+$ sign if the normal vector to S has a positive k-component, i.e., points generally upwards (as on the upper surface here), and the $-$ sign if it points generally downwards (as it does for the lower surface here).

This gives for the flux of the field $P\,\mathbf{k}$ across the upper surface S_2, on which $z = h(x,y)$,

$$\iint_{S_2} P\,\mathbf{k}\cdot d\mathbf{S} \;=\; \iint_R P(x,y,z)\,dx\,dy \;=\; \iint_R P(x,y,h(x,y))\,dx\,dy\;,$$

while for the flux across the lower surface S_1, where $z = g(x,y)$ and we use the $-$ sign as described above, we get

$$\iint_{S_1} P\,\mathbf{k}\cdot d\mathbf{S} \;=\; \iint_R -P(x,y,z)\,dx\,dy \;=\; \iint_R -P(x,y,g(x,y))\,dx\,dy\;;$$

adding up the two fluxes to get the total flux across S, we have

$$\iint_S P\,\mathbf{k}\cdot d\mathbf{S} \;=\; \iint_R P(x,y,h)\,dx\,dy \;-\; \iint_R P(x,y,g)\,dx\,dy$$

which is the same as the double integral in (6). This proves (4). $\square$

In the same way, if $\mathbf{F} = M(x,y,z)\,\mathbf{i}$ and the surface is simple in the $\mathbf{i}$ direction, we can prove

$$(4') \qquad \iint_S M\,\mathbf{i}\cdot\mathbf{n}\,dS \;=\; \iiint_D \frac{\partial M}{\partial x}\,dV$$

while if $\mathbf{F} = N(x,y,z)\,\mathbf{j}$ and the surface is simple in the $\mathbf{j}$ direction,

$$(4'') \qquad \iint_S N\,\mathbf{j}\cdot\mathbf{n}\,dS \;=\; \iiint_D \frac{\partial N}{\partial y}\,dV\ .$$

Finally, for a general field $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ and a closed surface S which is simple in all three directions, we have only to add up (4), (4′), and (4″). and we get the divergence theorem.

If the domain D is not bounded by a closed surface which is simple in all three directions, it can usually be divided up into smaller domains D_i which are bounded by such surfaces S_i. Adding these up gives the divergence theorem for D and S, since the surface integrals over the new faces introduced by cutting up D each occur twice, with the opposite normal vectors $\mathbf{n}$, so that they cancel out; after addition, one ends up just with the surface integral over the original S.

Partial Differential Equations

An important application of the higher partial derivatives is that they are used in partial differential equations to express some laws of physics which are basic to most science and engineering subjects. In this section, we will give examples of a few such equations. The reason is partly cultural, so you meet these equations early and learn to recognize them, and partly technical: to give you a little more practice with the chain rule and computing higher derivatives.

A **partial differential equation**, PDE for short, is an equation involving some unknown function of several variables and one or more of its partial derivatives. For example,

$$x\,\frac{\partial w}{\partial x} \;-\; y\,\frac{\partial w}{\partial y} \;=\; 0$$

is such an equation. Evidently here the unknown function is a function of two variables

$$w \;=\; f(x,y)\ ;$$

we infer this from the equation, since only x and y occur in it as independent variables. In general a **solution** of a partial differential equation is a differentiable function that satisfies it. In the above example, the functions

$$w \;=\; x^n y^n \qquad \text{any } n$$

all are solutions to the equation. In general, PDE's have many solutions, far too many to find all of them. The problem is always to find the one solution satisfying some extra conditions, usually called either *boundary conditions* or *initial conditions* depending on their nature.

Our first important PDE is the **Laplace equation** in three dimensions:

(1) $$\frac{\partial^2 w}{\partial x^2} \;+\; \frac{\partial^2 w}{\partial y^2} \;+\; \frac{\partial^2 w}{\partial z^2} \;=\; 0\ .$$

Any steady-state temperature distribution in three-space

(2) $$w \;=\; T(x,y,z), \qquad\qquad T = \text{temperature at the point } \ (x,y,z)$$

satisfies Laplace's equation. (Here *steady-state* means that it is unchanging over time, here reflected in the fact that T is not a function of time. For example, imagine a solid object made of some uniform heat-conducting material (say a solid metal ball), and imagine a steady temperature distribution on its surface is maintained somehow (say with some arrangement of wires and thermostats). Then after a while the temperature at each point inside the ball will come to equilibrium — reach a steady state — and the resulting temperature function (2) inside the ball will then satisfy Laplace's equation.

As another example, the *gravitational potential*

$$w \;=\; \phi(x,y,z)$$

resulting from some arrangement of masses in space satisfies Laplace's equation in any region R of space not containing masses. The same is true of the *electrostatic potential* resulting from some collection of electric charges in space: (1) is satisfied in any region which is free of charge. This potential function measures the work done (against the field) carrying a unit test mass (or charge) from a fixed reference point to the point (x, y, z) in the gravitational (or electrostatic) field. Knowing ϕ, the field itself can be recovered as its negative gradient:

$$\mathbf{F} = -\nabla\phi \,.$$

All of this is just to stress the fundamental character of Laplace's equation — we live our lives surrounded by its solutions.

The *two-dimensional* Laplace equation is similar — you just drop the term involving z. The steady-state temperature distribution in a flat metal plate would satisfy the two-dimensional Laplace equation, if the faces of the plate were kept insulated and a steady-state temperature distribution maintained around the edges of the plate.

If in the temperature model we include also heat sources and sinks in the region, unchanging over time, the temperature function satisfies the closely related **Poisson equation**

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} = f(x, y, z), \tag{3}$$

where f is some given function related to the sources and sinks.

Another important PDE is the **wave equation**; given below are the one-dimensional and two-dimensional versions; the three dimensional version would add a similar term in z to the left:

$$\frac{\partial^2 w}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 w}{\partial t^2}; \qquad \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 w}{\partial t^2}. \tag{4}$$

Here $x, y, \ldots$ are the space variables, t is the time, and c is the velocity with which the wave travels — this depends on the medium and the type of wave (light, sound, etc.). A solution, respectively

$$w = w(x, t), \qquad w = w(x, y, t),$$

gives for each moment t_0 of time the shape $w(x, t_0)$, $w(x, y, t_0)$ of the wave.

The third PDE goes by two names, depending on the context: **heat equation** or **diffusion equation**. The one- and two-dimensional versions are respectively

$$\frac{\partial^2 w}{\partial x^2} = \frac{1}{a^2}\frac{\partial w}{\partial t}; \qquad \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \frac{1}{a^2}\frac{\partial w}{\partial t}. \tag{5}$$

It looks a lot like the wave equation (4), but the right-hand side this time involves only the first derivative, which gives it mathematically and physically an entirely different character.

When it is called the (one-dimensional) heat equation, a solution $w(x, t)$ represents a time-varying temperature distribution in say a uniform conducting metal rod, with insulated sides. In the same way, $w(x, y, t)$ would be the time-varying temperature distribution in a flat metal plate with insulated faces. For each moment t_0 in time, $w(x, y, t_0)$ gives the temperature distribution at that moment.

For example, if we assume the distribution is steady-state, i.e., not changing with time, then

$$\frac{\partial w}{\partial t} = 0 \qquad \text{(steady-state condition)}$$

and the two-dimensional heat equation would turn into the two-dimensional Laplace equation (1).

When (5) is referred to as the *diffusion equation*, say in one dimension, then $w(x, t)$ represents the concentration of a dissolved substance diffusing along a uniform tube filled with liquid, or of a gas diffusing down a uniform pipe.

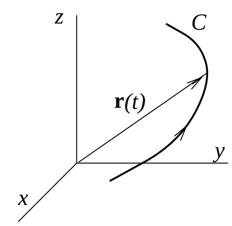
Notice that all of these PDE's are second-order, that is, involve derivatives no higher than the second. There is an important fourth-order PDE in elasticity theory (the bilaplacian equation), but by and large the general rule seems to be either that Nature is content with laws that only require second partial derivatives, or that these are the only laws that humans are intelligent enough to formulate.

V11. Line Integrals in Space

1. Curves in space.

In order to generalize to three-space our earlier work with line integrals in the plane, we begin by recalling the relevant facts about parametrized space curves.

In 3-space, a vector function of one variable is given as

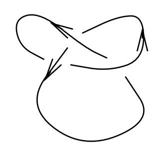


$$\mathbf{r}(t) \;=\; x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}\;. \tag{1}$$

It is called *continuous* or *differentiable* or *continuously differentiable* if respectively $x(t), y(t)$, and $z(t)$ all have the corresponding property. By placing the vector so that its tail is at the origin, its head moves along a curve C as t varies. This curve can be described therefore either by its position vector function (1), or by the three parametric equations

$$x = x(t), \qquad y = y(t), \qquad z = z(t)\;. \tag{2}$$

The curves we will deal with will be finite, connected, and piecewise smooth; this means that they have finite length, they consist of one piece, and they can be subdivided into a finite number of smaller pieces, each of which is given as the position vector of a continuously differentiable function (i.e., one whose derivative is continuous).



In addition, the curves will be *oriented*, or *directed*, meaning that an arrow has been placed on them to indicate which direction is considered to be the positive one. The curve is called *closed* if a point P moving on it always in the positive direction ultimately returns to its starting position, as in the accompanying picture.

The derivative of $\mathbf{r}(t)$ is defined in terms of components by

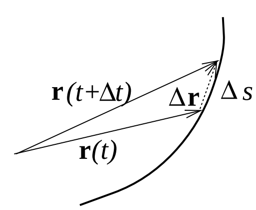
$$\frac{d\mathbf{r}}{dt} \;=\; \frac{dx}{dt}\,\mathbf{i} + \frac{dy}{dt}\,\mathbf{j} + \frac{dz}{dt}\,\mathbf{k}\;. \tag{3}$$

If the parameter t represents time, we can think of $d\mathbf{r}/dt$ as the *velocity vector* $\mathbf{v}$. If we let s denote the arclength along C, measured from some fixed starting point in the positive direction, then in terms of s the magnitude and direction of $\mathbf{v}$ are given by

$$|\mathbf{v}| \;=\; \left|\frac{ds}{dt}\right|, \qquad \text{dir}\;\mathbf{v} \;=\; \begin{cases} \mathbf{t}, & \text{if } ds/dt > 0; \\ -\mathbf{t}, & \text{if } ds/dt < 0\;. \end{cases} \tag{4}$$

Here $\mathbf{t}$ is the unit tangent vector (pointing in the positive direction on C :

$$\mathbf{t} \;=\; \frac{d\mathbf{r}}{ds} \;=\; \frac{d\mathbf{r}/dt}{ds/dt}\;. \tag{5}$$



You can see from the picture that $\mathbf{t}$ is a unit vector, since

$$\left|\frac{d\mathbf{r}}{ds}\right| \;=\; \lim_{\Delta s\to 0}\left|\frac{\Delta\mathbf{r}}{\Delta s}\right| \;=\; 1\;.$$

2. Line integrals in space. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ be a vector field in space, assumed continuous.

We define the *line integral of the tangential component of* $\mathbf{F}$ along an oriented curve C in space in the same way as for the plane. We approximate C by an inscribed sequence of directed line segments $\Delta\mathbf{r}_k$, form the approximating sum, then pass to the limit:

$$\int_C \mathbf{F}\cdot d\mathbf{r} \;=\; \lim_{k\to\infty} \sum_k \mathbf{r}_k \cdot \Delta\mathbf{r}_k \;.$$

The line integral is calculated just like the one in two dimensions:

$$\int_C \mathbf{F}\cdot d\mathbf{r} \;=\; \int_{t_0}^{t_1} \mathbf{F}\cdot \frac{d\mathbf{r}}{dt}\,dt\;, \tag{6}$$

if C is given by the position vector function $\mathbf{r}(t)$, $\quad t_0 \le t \le t_1$. Using x, y, z-components, one would write (6) as

$$\int_C M\,dx + N\,dy + P\,dz \;=\; \int_{t_0}^{t_1} \left(M\frac{dx}{dt} + N\frac{dy}{dt} + P\frac{dz}{dt}\right) dt \tag{6$'$}$$

In particular, if the parameter is the arclength s, then (6) becomes (since $\mathbf{t} = d\mathbf{r}/ds$)

$$\int_C \mathbf{F}\cdot d\mathbf{r} \;=\; \int_{s_0}^{s_1} \mathbf{F}\cdot\mathbf{t}\,ds\;, \tag{7}$$

which shows that the line integral is the integral along C of the tangential component of $\mathbf{F}$. As in two dimensions, this line integral represents the work done by the field $\mathbf{F}$ carrying a unit point mass (or charge) along the curve C.

Example 1. Find the work done by the electrostatic force field $\mathbf{F} = y\,\mathbf{i} + z\,\mathbf{j} + x\,\mathbf{k}$ in carrying a positive unit point charge from $(1, 1, 1)$ to $(2, 4, 8)$ along
a) a line segment b) the twisted cubic curve $\mathbf{r} = t\,\mathbf{i} + t^2\,\mathbf{j} + t^3\,\mathbf{k}$.

Solution. a) The line segment is given parametrically by

$$x - 1 = t, \quad y - 1 = 3t, \quad z - 1 = 7t, \qquad 0 \le t \le 1\;.$$

$$\begin{aligned}\int_C y\,dx + z\,dy + x\,dz \;&=\; \int_0^1 (3t+1)\,dt + (7t+1)\cdot 3\,dt + (t+1)\cdot 7\,dt, \quad \text{using } (6')\\ &=\; \int_0^1 (31t + 11)\,dt \;=\; \frac{31}{2}t^2 + 11t\Big]_0^1 \;=\; \frac{31}{2} + 11 = 26.5\;.\end{aligned}$$

b) Here the curve is given by $x = t,\ y = t^2,\ z = t^3, \quad 1 \le t \le 2$. For this curve, the line integral is

$$\begin{aligned}\int_1^2 t^2 dt + t^3\cdot 2t\,dt + t\cdot 3t^2 dt \;&=\; \int_1^2 (t^2 + 3t^3 + 2t^4)dt\\ &=\; \frac{t^3}{3} + \frac{3t^4}{4} + \frac{2t^5}{5}\Big]_1^2 \;\approx\; 25.18\;.\end{aligned}$$

The different results for the two paths shows that for this field, the line integral between two points depends on the path.

3. Gradient fields and path-independence.

The two-dimensional theory developed for line integrals in the plane generalizes easily to three-space. For the part where no new ideas are involved, we will be brief, just stating the results, and in places sketching the proofs.

Definition. Let $\mathbf{F}$ be a continuous vector field in a region D of space. The line integral $\int_P^Q \mathbf{F} \cdot d\mathbf{r}$ is called **path-independent** if, for any two points P and Q in the region D, the value of $\int_C \mathbf{F} \cdot d\mathbf{r}$ along a directed curve C lying in D and running from P to Q depends only on the two endpoints, and not on C.

An equivalent formulation is (the proof of equivalence is the same as before):

$$(8) \qquad \int_P^Q \mathbf{F} \cdot d\mathbf{r} \text{ is path independent } \Leftrightarrow \oint_C \mathbf{F} \cdot d\mathbf{r} = 0 \text{ for every closed curve } C \text{ in } D$$

Definition Let $f(x, y, z)$ be continuously differentiable in a region D. The vector field

$$(9) \qquad \nabla f \;=\; \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

is called the **gradient field** of f in D. Any field of the form ∇f is called a gradient field.

Theorem. First fundamental theorem of calculus for line integrals. *If $f(x, y, z)$ is continuously differentiable in a region D, then for any two points P_1, P_2 lying in D,*

$$(10) \qquad \int_{P_1}^{P_2} \nabla f \cdot d\mathbf{r} \;=\; f(P_2) - f(P_1),$$

where the integral is taken along any curve C lying in D and running from P_1 to P_2. In particular, the line integral is path-independent.

The proof is exactly the same as before — use the chain rule to reduce it to the first fundamental theorem of calculus for functions of one variable.

There is also an analogue of the second fundamental theorem of calculus, the one where we first integrate, then differentiate.

Theorem. Second fundamental theorem of calculus for line integrals.

Let $\mathbf{F}(x, y, z)$ *be continuous and* $\int_P^Q \mathbf{F} \cdot d\mathbf{r}$ *path-independent in a region D; and define*

$$(11) \qquad \begin{aligned} f(x, y, z) \;&=\; \int_{(x_0,y_0,z_0)}^{(x,y,z)} \mathbf{F} \cdot d\mathbf{r}; \qquad \textit{then} \\ \nabla f \;&=\; \mathbf{F} \qquad \textit{in } D. \end{aligned}$$

Note that since the integral is path-independent, no C need be specified in (11). The theorem is proved in your book for line integrals in the plane. The proof for line integrals in space is analogous.

Just as before, these two theorems produce the three equivalent statements: in D,

$$\text{(12)} \qquad \mathbf{F} = \nabla f \quad \Leftrightarrow \quad \int_P^Q \mathbf{F} \cdot d\mathbf{r} \ \text{ path-independent} \quad \Leftrightarrow \quad \oint_C \mathbf{F} \cdot d\mathbf{r} = 0 \text{ for any closed } C$$

As in the two-dimensional case, if $\mathbf{F}$ is thought of as a force field, then the gradient force fields are called *conservative* fields, since the work done going around any closed path is zero (i.e., energy is conserved). If $\mathbf{F} = \nabla f$, then f is the called the (mathematical) *potential function* for $\mathbf{F}$; the physical potential function is defined to be $-f$.

Example 2. Let $f(x, y, z) = (x + y^2)z$. Calculate $\mathbf{F} = \nabla f$, and find $\int_C \mathbf{F} \cdot d\mathbf{r}$, where C is the helix $x = \cos t, y = \sin t, z = t, 0 \le t \le \pi$.

Solution. By differentiating, $\mathbf{F} = z\,\mathbf{i} + 2yz\,\mathbf{j} + (x + y^2)\,\mathbf{k}$. The curve C runs from $(1, 0, 0)$ to $(-1, 0, \pi)$. Therefore by (10),

$$\int_C \mathbf{F} \cdot d\mathbf{r} \ = \ (x + y^2)z\Big]_{(1,0,0)}^{(-1,0,\pi)} \ = \ -\pi - 0 \ = \ -\pi.$$

No direct calculation of the line integral is needed, notice.

V12. Gradient Fields in Space

1. The criterion for gradient fields. The curl in space.

We seek now to generalize to space our earlier criterion (Section V2) for gradient fields in the plane.

Criterion for a Gradient Field. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ be continuously differentiable. Then

$$(1)\quad \mathbf{F} = \nabla f \quad \text{for some } f(x,y,z) \quad \Rightarrow \quad \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}\,, \quad \frac{\partial M}{\partial z} = \frac{\partial P}{\partial x}\,, \quad \frac{\partial N}{\partial z} = \frac{\partial P}{\partial y}\,.$$

Proof. Since $\mathbf{F} = \nabla f$, when written out this says

$$(2)\qquad M = \frac{\partial f}{\partial x}, \qquad N = \frac{\partial f}{\partial y}. \qquad P = \frac{\partial f}{\partial z}\,; \qquad \text{therefore}$$

$$\frac{\partial M}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial N}{\partial x}\,.$$

The two mixed partial derivatives are equal since they are continuous, by the hypothesis that $\mathbf{F}$ is continuously differentiable.

The other two equalities in (1) are proved similarly. $\square$

Though the criterion looks more complicated to remember and to check than the one in two dimensions, which involves just a single equation, it is not difficult to learn and apply. For theoretical purposes, it can be expressed more elegantly by using the three-dimensional vector curl $\mathbf{F}$.

Definition. Let $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ be differentiable. We define **curl F** by

$$(3)\qquad \text{curl}\,\mathbf{F} = (P_y - N_z)\,\mathbf{i} + (M_z - P_x)\,\mathbf{j} + (N_x - M_y)\,\mathbf{k}$$

$$(3')\qquad = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ M & N & P \end{vmatrix} \qquad \text{(symbolic notation; } \partial_x = \frac{\partial}{\partial x}\text{, etc.)}$$

$$(3'')\qquad = \nabla \times \mathbf{F}, \qquad \text{where } \nabla = \frac{\partial}{\partial x}\,\mathbf{i} + \frac{\partial}{\partial y}\,\mathbf{j} + \frac{\partial}{\partial z}\,\mathbf{k}\,.$$

The equation (3) is the definition. The other two lines give symbolic ways of writing and of remembering the right side of (3). Neither the first nor second row of the determinant contains the sort of thing you are allowed to put into a determinant; however, if you "evaluate" it using the Laplace expansion by the first row, what you get is the right side of (3). Similarly, to evaluate the symbolic cross-product in (3′′), we use the determinant (3′). In doing these, by the "product" of $\dfrac{\partial}{\partial x}$ and M we mean $\dfrac{\partial M}{\partial x}$.

By using the vector field curl $\mathbf{F}$, our criterion (1) becomes

$$(1')\qquad \mathbf{F} = \nabla f \quad \Rightarrow \quad \text{curl}\,\mathbf{F} = \mathbf{0}\,.$$

In dealing with a plane vector field $\mathbf{F} = M(x,y)\,\mathbf{i} + N(x,y)\,\mathbf{j}$, we gave the name curl $\mathbf{F}$ to the *scalar* function $N_x - M_y$, whereas for a vector field $\mathbf{F}$ in space, curl $\mathbf{F}$ is a *vector* function. However, if we think of the two-dimensional field $\mathbf{F}$ as a field in space (i.e., one with zero $\mathbf{k}$-component and not depending on z), then using definition (3) you can compute that

$$\text{curl } \mathbf{F} \;=\; (N_x - M_y)\,\mathbf{k}\ .$$

Thus curl $\mathbf{F}$ has only a $\mathbf{k}$-component, so if we are dealing just with two-dimensional fields, it is natural to give the name curl $\mathbf{F}$ just to this $\mathbf{k}$-component. This is not a universally accepted terminology, however; some call it the "scalar curl", others don't use any name at all for $N_x - M_y$.

Naturally, the question arises as to whether the converse of $(1')$ is true — if curl $\mathbf{F} = \mathbf{0}$, is $\mathbf{F}$ a gradient field? As in two dimensionas, this requires some sort of restriction on the domain, and we will return to this point after we have studied Stokes' theorem. For now we will assume the domain is the whole three-space, in which case it is true:

Theorem. *If* $\mathbf{F}$ *is continuously differentiable for all* x, y, z,

$$(4) \qquad \textit{curl}\,\mathbf{F} \;=\mathbf{0} \quad\Rightarrow\quad \mathbf{F} \;=\; \nabla f, \qquad \textit{for some differentiable } f(x,y,z).$$

We will prove this later. If $\mathbf{F}$ is a gradient field, we can calculate the corresponding (mathematical) potential function $f(x,y,z)$ by the three-dimensional analogue of either of the two methods described before (Section V2). We illustrate with an example.

Example 1. For what value(s), if any, of c will $\mathbf{F} \;=\; y\,\mathbf{i} + (x+cyz)\,\mathbf{j} + (y^2+z^2)\,\mathbf{k}$ be a conservative (i.e., gradient) field? For each such c, find a corresponding potential function $f(x,y,z)$.

Solution. Using (1) and (4), we calculate the relevant partial derivatives:

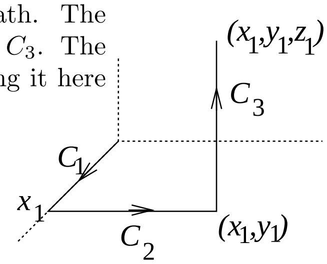
$$M_y = 1,\ N_x = 1; \quad N_z = cy,\ P_y = 2y; \quad M_z = 0,\ P_x = 0$$

Thus all three equations in (1) are satisfied $\Leftrightarrow$ $c = 2$. For this value of c, we now find $f(x,y,z)$ by two methods.

Method 1. We use the second fundamental theorem (Section V11, (11)), taking $(0,0,0)$ as a convenient lower limit for the integral, and using the subscript 1 on the upper limit to avoid confusion with the variables of integration. This gives

$$(5) \qquad f(x_1,y_1,z_1) \;=\; \int_{(0,0,0)}^{(x_1,y_1,z_1)} y\,dx + (x+2yz)\,dy + (y^2+z^2)\,dz$$

Since the integral is path-independent for the choice $c = 2$, we can use any path. The usual choice is the path illustrated, consisting of three line segments C_1, C_2 and C_3. The parametrizations for them are (don't write these out yourself — we are only doing it here this first time to make it clear how the line integral is being calculated):



$$\begin{array}{llll} C_1: & x = x,\ y = 0,\ z = 0; & \text{thus} & dx = dx,\ dy = 0,\ dz = 0; \\ C_2: & x = x_1,\ y = y,\ z = 0; & \text{thus} & dx = 0,\ dy = dy,\ dz = 0; \\ C_3: & x = x_1,\ y = y_1,\ z = z; & \text{thus} & dx = 0,\ dy = 0,\ dz = dz\ . \end{array}$$

Using these, we calculate the line integral (5) over each of the C_i in turn:

$$\begin{aligned}\int_{C_1+C_2+C_3} y\,dx + (x+2yz)\,dy + (y^2+z^2)\,dz &= \int_0^{x_1} 0\cdot dx \;+\; \int_0^{y_1}(x_1+2y\cdot 0)\,dy \;+\; \int_0^{z_1}(y_1^2+z^2)\,dz \\ &= 0 + x_1y_1 + (y_1^2 z_1 + \frac{1}{3}z_1^3)\ .\end{aligned}$$

Dropping subscripts, we have therefore by (5),

$$f(x,y,z) \;=\; xy + y^2 z + \frac{1}{3}z^3 + c, \tag{6}$$

where we have added an arbitrary constant of integration to compensate for our arbitary choice of $(0,0,0)$ as the lower limit of integration — a different choice would have added a constant to the right side of (6).

The work should always be checked; from (6) one sees easily that $\nabla f \;=\; \mathbf{F}$, the field we started with.

Method 2. This requires no line integrals, but the work must be carried out systematically, otherwise you'll get lost in a mess of equations.

We are looking for an $f(x,y,z)$ such that $(f_x, f_y, f_z) \;=\; (y,\; x+2yz,\; y^2+z^2)$. This is equivalent to the three equations

$$f_x \;=\; y, \qquad f_y \;=\; x+2yz, \qquad f_z \;=\; y^2+z^2\ . \tag{7}$$

From the first equation, integrating with respect to x (holding y and z fixed), we get

$$\begin{array}{llcll}
(8) & f(x,y,z) &=& xy + g(y,z), & g \text{ is an arbitrary function} \\
& \dfrac{\partial f}{\partial y} &=& x + \dfrac{\partial g}{\partial y}, & \text{from (8)} \\
& &=& x + 2yz & \text{from (7), second equation; comparing,} \\
& \dfrac{\partial g}{\partial y} &=& 2yz. & \text{Integrating with respect to } y, \\
& g(y,z) &=& y^2 z + h(z), & h \text{ is an arbitrary function; thus} \\
(9) & f(x,y,z) &=& xy + y^2 z + h(z), & \text{from the preceding and (8)} \\
& \dfrac{\partial f}{\partial z} &=& y^2 + h'(z) & \\
& &=& y^2 + z^2, & \text{from (7), third equation; comparing,} \\
& h'(z) &=& z^2, & \\
& h(z) &=& \dfrac{1}{3}z^3 + c; & \text{finally, by (9)} \\
& f(x,y,z) &=& xy + y^2 z + \dfrac{1}{3}z^3 + c & \text{as in Method 1.}
\end{array}$$

2. Exact differentials

Just as we did in the two-dimensional case, we translate the previous ideas into the language of differentials.

The formal expression

$$M(x, y, z)\,dx + N(x, y, z)\,dy + P(x, y, z)\,dz \tag{10}$$

which appears as the integrand in our line integrals is called a **differential**. If $f(x, y, z)$ is a differentiable function, then its **total differential** (or just *differential*) is defined to be

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz \tag{11}$$

The differential (10) is said to be **exact**, in some domain D where M, N and P are defined, if it is the total differential of some differentiable function $f(x, y, z)$ in this domain, that is, if there exists an $f(x, y, z)$ in D such that

$$M = \frac{\partial f}{\partial x}, \quad N = \frac{\partial f}{\partial y}, \quad P = \frac{\partial f}{\partial z}\ . \tag{12}$$

Criterion for exact differentials. *Let D be a domain in which M, N, P are continuously differentiable. Then in D,*

$$M\,dx + N\,dy + P\,dz \quad \text{is exact} \quad \Rightarrow \quad P_y = N_z, \quad M_z = P_x, \quad N_x = M_y\ ; \tag{13}$$

if D is all of 3-space, then the converse is true:

$$P_y = N_z, \quad M_z = P_x, \quad N_x = M_y \quad \Rightarrow \quad M\,dx + N\,dy + P\,dz \quad \text{is exact.} \tag{14}$$

If the test in this criterion shows that the differential (10) is exact, the function $f(x, y, z)$ may be found be either method 1 or method 2. The converse (14) is true under weaker hypotheses about D, which we will come back to after we have taken up Stokes' Theorem.

V15.1 Del Operator

1. Symbolic notation: the del operator

To have a compact notation, wide use is made of the symbolic operator "del" (some call it "nabla"):

$$\nabla \;=\; \frac{\partial}{\partial x}\,\mathbf{i} + \frac{\partial}{\partial y}\,\mathbf{j} + \frac{\partial}{\partial z}\,\mathbf{k} \tag{1}$$

Recall that the "product" of $\dfrac{\partial}{\partial x}$ and the function $M(x,y,z)$ is understood to be $\dfrac{\partial M}{\partial x}$. Then we have

$$\text{grad } f \;=\; \nabla f \;=\; \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} + \frac{\partial f}{\partial z}\,\mathbf{k} \tag{2}$$

The divergence is sort of a symbolic scalar product: if $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$,

$$\text{div } \mathbf{F} \;=\; \nabla\cdot\mathbf{F} \;=\; \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} + \frac{\partial P}{\partial z} \tag{3}$$

while the curl, as we have noted, as a symbolic cross-product:

$$\text{curl } \mathbf{F} \;=\; \nabla\times\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ M & N & P \end{vmatrix}\;.$$

Notice how this notation reminds you that $\nabla\cdot\mathbf{F}$ is a scalar function, while $\nabla\times\mathbf{F}$ is a vector function.

We may also speak of the Laplace operator (also called the "Laplacian"), defined by

$$\text{lap } f \;=\; \nabla^2 f \;=\; (\nabla\cdot\nabla)\,f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\;. \tag{5}$$

Thus, Laplace's equation may be written: $\nabla^2 f = 0$. (This is for example the equation satisfied by the potential function for an electrostatic field, in any region of space where there are no charges; or for a gravitational field, in a region of space where there are no masses.)

In this notation, the divergence theorem and Stokes' theorem are respectively

$$\iint_S \mathbf{F}\cdot d\mathbf{S} \;=\; \iiint_D \nabla\cdot\mathbf{F}\,dV \qquad\qquad \oint_C \mathbf{F}\cdot d\mathbf{r} \;=\; \iint_S \nabla\times\mathbf{F}\cdot d\mathbf{S} \tag{6}$$

Two important relations involving the symbolic operator are:

$$\text{curl (grad} f) \;=\; \mathbf{0} \qquad\qquad \text{div curl } \mathbf{F} \;=\; 0 \tag{7}$$

$$\nabla\times\nabla f \;=\; \mathbf{0} \qquad\qquad \nabla\cdot\nabla\times\mathbf{F} \;=\; 0 \tag{7$'$}$$

The first we have proved (it was part of the criterion for gradient fields); the second is an easy exercise. Note however how the symbolic notation suggests the answer, since we know that for any vector $\mathbf{A}$, we have

$$\mathbf{A}\times\mathbf{A} \;=\; \mathbf{0}, \qquad\qquad \mathbf{A}\cdot\mathbf{A}\times\mathbf{F} \;=\; 0\;,$$

and $(7')$ says this is true for the symbolic vector ∇ as well.

V4.3 Physical meaning of curl

3. An interpretation for curl F.

We will start by looking at the two dimensional curl in the xy-plane. Our interpretation will be that the curl at a point represents twice the angular velocity of a small paddle wheel at that point. At the very end we will indicate how to extend this interpretation to 3 dimensions.

The function curl $\mathbf{F}$ can be thought of as measuring the tendency of $\mathbf{F}$ to produce rotation. Interpreting $\mathbf{F}$ either as a force field or a velocity field, $\mathbf{F}$ will make a suitable test object placed at a point P_0 spin about a vertical axis (i.e., one in the $\mathbf{k}$-direction), and the angular velocity of the spin will be proportional to (curl $\mathbf{F})_0$.

To see this for the velocity field $\mathbf{v}$ of a flowing liquid, place a paddle wheel of radius a so its center is at (x_0, y_0), and its axis is vertical. We ask how rapidly the flow spins the wheel.

If the wheel had only one blade, the velocity of the blade would be $\mathbf{F} \cdot \mathbf{t}$, the component of the flow velocity vector $\mathbf{F}$ perpendicular to the blade, i.e., tangent to the circle of radius a traced out by the blade.

paddlewheel   top view

Since $\mathbf{F} \cdot \mathbf{t}$ is not constant along this circle, if the wheel had only one blade it would spin around at an uneven rate. But if the wheel has many blades, this unevenness will be averaged out, and it will spin around at approximately the *average value* of the tangential velocity $\mathbf{F} \cdot \mathbf{t}$ over the circle. Like the average value of any function defined along a curve, this average tangential velocity can be found by integrating $\mathbf{F} \cdot \mathbf{t}$ over the circle, and dividing by the length of the circle. Thus,

$$\begin{aligned} \text{speed of blade} &= \frac{1}{2\pi a} \oint_C \mathbf{F} \cdot \mathbf{t}\, ds \quad = \frac{1}{2\pi a} \oint_C \mathbf{F} \cdot d\mathbf{r} \\ &= \frac{1}{2\pi a} \iint_R (\text{curl } \mathbf{F})_0 \, dx\, dy, \qquad \text{by Green's theorem,} \\ &\approx \frac{1}{2\pi a} (\text{curl } \mathbf{F})_0 \pi a^2 \ , \end{aligned} \tag{8}$$

where (curl $\mathbf{F})_0$ is the value of the function curl $\mathbf{F}$ at (x_0, y_0). The justification for the last approximation is that if the circle formed by the paddlewheel is small, then curl $\mathbf{F}$ has approximately the value (curl $\mathbf{F})_0$ over the interior R of the circle, so that multiplying this constant value by the area πa^2 of R should give approximately the value of the double integral.

From (8) we get for the tangential speed of the paddlewheel:

$$\text{tangential speed} \quad \approx \quad \frac{a}{2} (\text{curl } \mathbf{F})_0 \ . \tag{9}$$

We can get rid of the a by using the angular velocity ω_0 of the paddlewheel; since the tangential speed is $a\omega_0$, (9) becomes

$$\omega_0 \quad \approx \quad \frac{1}{2} (\text{curl } \mathbf{F})_0 \ . \tag{10}$$

As the radius of the paddlewheel gets smaller, the approximation becomes more exact, and passing to the limit as $a \to 0$, we conclude that, for a two-dimensional velocity field $\mathbf{F}$,

$$\text{(11)} \qquad \boxed{\text{curl } \mathbf{F} = \textit{twice the angular velocity of an infinitesimal paddlewheel at } (x, y)\ .}$$

The curl thus measures the "vorticity" of the fluid flow — its tendency to produce rotation.

A consideration of curl $\mathbf{F}$ for a force field would be similar, interpreting $\mathbf{F}$ as exerting a torque on a spinnable object — a little dumbbell with two unit masses for a gravitational field, or with two unit positive charges for an electrostatic force field.

Example 1. Calculate and interpret curl $\mathbf{F}$ for (a) $x\,\mathbf{i} + y\,\mathbf{j}$ (b) $\omega(-y\,\mathbf{i} + x\,\mathbf{j}\,)$

Solution. (a) curl $\mathbf{F} = 0$; this makes sense since the field is radially outward and radially symmetric, there is no favored angular direction in which the paddlewheel could spin.

(b) curl $\mathbf{F} = 2\omega$ at every point. Since this field represents a fluid rotating about the origin with constant angular velocity ω (see section V1), it is at least clear that curl $\mathbf{F}$ should be 2ω at the origin; it's not so clear that it should have this same value everywhere, but it is true.

Extension to Three Dimensions. To extend this interpretation to three dimensions note that any component of the flow of $\mathbf{F}$ in the $\mathbf{k}$ direction will not have any effect on a paddle wheel in the xy-plane. In fact, for any plane with normal $\mathbf{n}$ the component of $\mathbf{F}$ in the direction of $\mathbf{n}$ has no effect on a paddle wheel in the plane. This leads to the following interpretation of the three dimensional curl:

For any plane with unit normal $\mathbf{n}$, $(\text{curl } \mathbf{F}) \cdot \mathbf{n}$ is two times the angular velocity of a small paddle wheel in the plane.

We could force through a proof along the lines of the 2D proof above. Once we learn Stokes Theorem we can make a much simpler argument.

Testing for a Conservative Field

Let $\mathbf{F} = (3x^2y + az)\mathbf{i} + x^3\mathbf{j} + (3x + 3z^2)\mathbf{k}$.

1. For what value or values of a is $\mathbf{F}$ conservative?

<u>Answer:</u> We know $\mathbf{F}$ is conservative if it's continuously differentiable for all x, y, z and curl$F = 0$. We easily verify that $\mathbf{F}$ is continuously differentiable as required.

$$\text{curl}F = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ (3x^2y + az) & x^3 & (3x + 3z^2) \end{vmatrix} = 0\mathbf{i} - (3 - a)\mathbf{j} + (3x^2 - 3x^2)\mathbf{k} = (a - 3)\mathbf{j}.$$

If $a = 3$, curl$\mathbf{F} = 0$ and so $\mathbf{F}$ must be conservative.

Answer: $a = 3$.

2. Assuming a has the value(s) found in (1), find a potential function f for which $\mathbf{F} = \nabla f$.

<u>Answer:</u> As usual, there are two ways to find such a potential function. For variety, we'll use the second method.

Assume that $\mathbf{F} = \nabla f$. Then $f_x = 3x^2y + 3z$, so we have $f = x^3y + 3xz + g(y, z)$ for some function g.

Combine this with the fact that $f_y = x^3$ to get $x^3 + g_y = x^3$ so $g(y, z) = h(z)$ is constant with respect to y.

Finally, $f_z = 3x + h'(z) = 3x + 3z^2$ implies $h(z) = g(y, z) = z^3 + C$.

We conclude that $f(x, y, z) = x^3y + 3xz + z^3 + C$.

We can now calculate $f_x = 3x^2y + 3z$, $f_y = x^3$ and $f_z = 3x + 3z^2$ to check that $\mathbf{F} = \nabla f$.

V13.1-2 Stokes' Theorem

1. Introduction; statement of the theorem.

The normal form of Green's theorem generalizes in 3-space to the divergence theorem. What is the generalization to space of the tangential form of Green's theorem? It says

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_R \text{curl } \mathbf{F}\, dA \tag{1}$$

where C is a simple closed curve enclosing the plane region R.

Since the left side represents work done going around a closed curve in the plane, its natural generalization to space would be the integral $\oint \mathbf{F} \cdot d\mathbf{r}$ representing work done going around a closed curve in 3-space.

In trying to generalize the right-hand side of (1), the space curve C can only be the boundary of some piece of surface S — which of course will no longer be a piece of a plane. So it is natural to look for a generalization of the form

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_S (\text{something derived from } \mathbf{F})\, dS$$

The surface integral on the right should have these properties:

a) If curl $\mathbf{F} = 0$ in 3-space, then the surface integral should be 0; (for $\mathbf{F}$ is then a gradient field, by V12, (4), so the line integral is 0, by V11, (12)).

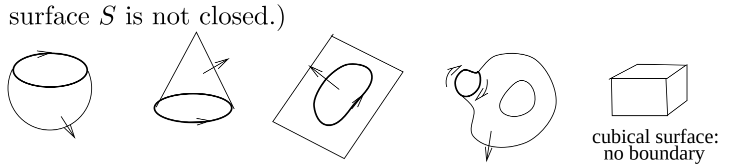
b) If C is in the xy-plane with S as its interior, and the field $\mathbf{F}$ does not depend on z and has only a $\mathbf{k}$-component, the right-hand side should be $\displaystyle\iint_S \text{curl } \mathbf{F}\, dS$.

These things suggest that the theorem we are looking for in space is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} \qquad\qquad \textbf{Stokes' theorem} \tag{2}$$

For the hypotheses, first of all C should be a closed curve, since it is the boundary of S, and it should be oriented, since we have to calculate a line integral over it.

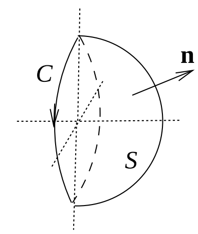
S is an oriented surface, since we have to calculate the flux of curl F through it. This means that S is two-sided, and one of the sides designated as positive; then the unit normal $\mathbf{n}$ is the one whose base is on the positive side. (There is no "standard" choice for positive side, since the surface S is not closed.)



It is important that C and S be *compatibly* oriented. By this we mean that the right-hand rule applies: when you walk in the positive direction on C, keeping S to your left, then your head should point in the direction of $\mathbf{n}$. The pictures give some examples.

The field $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ should have continuous first partial derivatives, so that we will be able to integrate curl F. For the same reason, the piece of surface S should be piecewise smooth and should be finite— i.e., not go off to infinity in any direction, and have finite area.

2. Examples and discussion.



Example 1. Verify the equality in Stokes' theorem when S is the half of the unit sphere centered at the origin on which $y \geq 0$, oriented so $\mathbf{n}$ makes an acute angle with the positive y-axis; take $\mathbf{F} = y\,\mathbf{i} + 2x\,\mathbf{j} + x\,\mathbf{k}$.

Solution. The picture illustrates C and S. Notice how C must be directed to make its orientation compatible with that of S.

We turn to the line integral first. C is a circle in the xz-plane, traced out *clockwise* in the plane. We select a parametrization and calculate:

$$x = \cos t, \qquad y = 0, \qquad z = -\sin t, \qquad 0 \leq t \leq 2\pi\ .$$

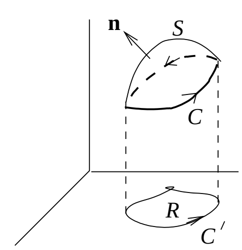
$$\oint_C y\,dx + 2x\,dy + x\,dz \;=\; \oint_C x\,dz \;=\; \int_0^{2\pi} -\cos^2 t\,dt \;=\; \left[-\frac{t}{2} - \frac{\sin 2t}{4}\right]_0^{2\pi} \;=\; -\pi\ .$$

For the surface S, we see by inspection that $\mathbf{n} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$; this is a unit vector since $x^2 + y^2 + z^2 = 1$ on S. We calculate

$$\text{curl}\,\mathbf{F} \;=\; \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ y & 2x & x \end{vmatrix} \;=\; -\mathbf{j} + \mathbf{k}\,; \qquad (\text{curl}\,\mathbf{F}) \cdot \mathbf{n} = -y + z$$

Integrating in spherical coordinates, we have $y = \sin\phi\sin\theta$, $z = \cos\phi$, $dS = \sin\phi\,d\phi\,d\theta$, since $\rho = 1$ on S; therefore

$$\begin{aligned} \iint_S \text{curl}\,\mathbf{F} \cdot d\mathbf{S} &= \iint_S (-y + z)\,dS \\ &= \int_0^{\pi}\int_0^{\pi} (-\sin\phi\sin\theta + \cos\phi)\sin\phi\,d\phi\,d\theta; \\ \text{inner integral} &= \sin\theta\left(\frac{\phi}{2} - \frac{\sin 2\phi}{4}\right) + \frac{1}{2}\sin^2\phi\,\Big]_0^{\pi} \;=\; \frac{\pi}{2}\sin\theta \\ \text{outer integral} &= -\frac{\pi}{2}\cos\theta\,\Big]_0^{\pi} \;=\; -\pi\ , \qquad \text{which checks.} \end{aligned}$$



Example 2. Suppose $\mathbf{F} = x^2\,\mathbf{i} + x\,\mathbf{j} + z^2\,\mathbf{k}$ and S is given as the graph of some function $z = g(x, y)$, oriented so $\mathbf{n}$ points upwards.

Show that $\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\;$ *area of* R, where C is the boundary of S, compatibly oriented, and R is the projection of S onto the xy-plane.

Solution. We have $\text{curl}\,\mathbf{F} \;=\; \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ x^2 & x & z^2 \end{vmatrix} \;=\; \mathbf{k}$. By Stokes' theorem, (cf. V9, (12))

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_S \mathbf{k} \cdot \mathbf{n}\,dS \;=\; \iint_R \mathbf{n} \cdot \mathbf{k}\,\frac{dA}{|\mathbf{n} \cdot \mathbf{k}|},$$

since $\mathbf{n}\cdot\mathbf{k} > 0$, $|\mathbf{n}\cdot\mathbf{k}| = \mathbf{n}\cdot\mathbf{k}$; therefore

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_R dA = \text{area of } R\ .$$

The relation of Stokes' theorem to Green's theorem.

Suppose $\mathbf{F}$ is a vector field in space, having the form $\mathbf{F} = M(x,y)\,\mathbf{i} + N(x,y)\,\mathbf{j}$, and C is a simple closed curve in the xy-plane, oriented positively (so the interior is on your left as you walk upright in the positive direction). Let S be its interior, compatibly oriented — this means that the unit normal $\mathbf{n}$ to S is the vector $\mathbf{k}$, and $dS = dA$.

Then we get by the usual determinant method curl $\mathbf{F} = (N_x - M_y)\,\mathbf{k}$; since $\mathbf{n} = \mathbf{k}$, Stokes theorem becomes

$$\oint \mathbf{F}\cdot d\mathbf{r} = \iint_S \text{curl } \mathbf{F}\cdot\mathbf{n}\, dS = \iint_R (N_x - M_y)\, dA\ ,$$

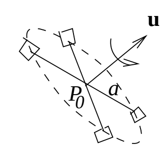
which is Green's theorem in the plane.

The same is true for other choices of the two variables; the most interesting one is $F = M(x,z)\,\mathbf{i} + P(x,z)\,\mathbf{k}$, where C is a simple closed curve in the xz-plane. If careful attention is paid to the choice of normal vector and the orientations, once again Stokes' theorem becomes just Green's theorem for the xz-plane. (See the Exercises.)

Interpretation of curl F.

Suppose now that $\mathbf{F}$ represents the velocity vector field for a three-dimensional fluid flow. Drawing on the interpretation we gave for the two-dimensional curl in Section V4, we can give the analog for 3-space.

The essential step is to interpret the $\mathbf{u}$-component of $(\text{curl } \mathbf{F})_0$ at a point P_0, , where $\mathbf{u}$ is a given unit vector, placed so its tail is at P_0.



Put a little paddlewheel of radius a in the flow so that its center is at P_0 and its axis points in the direction $\mathbf{u}$. Then by applying Stokes' theorem to a little circle C of radius a and center at P_0, lying in the plane through P_0 and having normal direction $\mathbf{u}$, we get just as in Section V4 (p. 4) that

$$\text{angular velocity of the paddlewheel} = \frac{1}{2\pi a^2}\oint_C \mathbf{F}\cdot d\mathbf{r}\ ;$$
$$= \frac{1}{2\pi a^2}\iint_S \text{curl } \mathbf{F}\cdot\mathbf{u}\, dS,$$

by Stokes' theorem, S being the circular disc having C as boundary;

$$\approx \frac{1}{2\pi a^2}(\text{curl } \mathbf{F})_0\cdot\mathbf{u}\ (\pi a^2),$$

since curl $\mathbf{F}\cdot\mathbf{u}$ is approximately constant on S if a is small, and S has area πa^2; passing to the limit as $a \to 0$, the approximation becomes an equality:

$$\text{angular velocity of the paddlewheel} = \frac{1}{2}\,(\text{curl } \mathbf{F})\cdot\mathbf{u}\ .$$

The preceding interprets $(\text{curl } \mathbf{F})_0\cdot\mathbf{u}$ for us. Since it has its maximum value when $\mathbf{u}$ has the direction of $(\text{curl } \mathbf{F})_0$, we conclude

$$\text{direction of } (\text{curl } \mathbf{F})_0 = \text{axial direction in which wheel spins fastest}$$
$$\text{magnitude of } (\text{curl } \mathbf{F})_0 = \text{twice this maximum angular velocity.}$$

V13.3 Stokes' Theorem

3. Proof of Stokes' Theorem.

We will prove Stokes' theorem for a vector field of the form $P(x, y, z)\,\mathbf{k}$. That is, we will show, with the usual notations,

$$(3) \qquad \oint_C P(x, y, z)\, dz \;=\; \iint_S \text{curl}\,(P\,\mathbf{k}) \cdot \mathbf{n}\, dS\ .$$

We assume S is given as the graph of $z = f(x, y)$ over a region R of the xy-plane; we let C be the boundary of S, and C' the boundary of R. We take $\mathbf{n}$ on S to be pointing generally upwards, so that $|\mathbf{n} \cdot \mathbf{k}| = \mathbf{n} \cdot \mathbf{k}$.

To prove (3), we turn the left side into a line integral around C', and the right side into a double integral over R, both in the xy-plane. Then we show that these two integrals are equal by Green's theorem.

To calculate the line integrals around C and C', we parametrize these curves. Let

$$C': \ x = x(t),\ y = y(t), \qquad\qquad t_0 \le t \le t_1$$

be a parametrization of the curve C' in the xy-plane; then

$$C: \ x = x(t),\ y = y(t),\ z = f(x(t), y(t)), \quad t_0 \le t \le t_1$$

gives a corresponding parametrization of the space curve C lying over it, since C lies on the surface $z = f(x, y)$.

Attacking the line integral first, we claim that

$$(4) \qquad \oint_C P(x, y, z)\, dz \;=\; \oint_{C'} P(x, y, f(x, y))(f_x dx + f_y dy)\ .$$

This looks reasonable purely formally, since we get the right side by substituting into the left side the expressions for z and dz in terms of x and y: $z = f(x, y)$, $\quad dz = f_x dx + f_y dy$. To justify it more carefully, we use the parametrizations given above for C and C' to calculate the line integrals.

$$\begin{aligned}
\oint_C P(x, y, z)\, dz \;&=\; \int_{t_0}^{t_1} \big(P(x(t), y(t), z(t)\big)\, \frac{dz}{dt}\, dt \\
&=\; \int_{t_0}^{t_1} \big(P(x(t), y(t), z(t)\big)\, (f_x \frac{dx}{dt} + f_y \frac{dy}{dt})\, dt, \ \text{by the chain rule} \\
&=\; \oint_{C'} P(x, y, f(x, y))(f_x dx + f_y dy), \quad \text{the right side of (4).}
\end{aligned}$$

We now calculate the surface integral on the right side of (3), using x and y as the variables. In the calculation, we must distinguish carefully between such expressions as $P_1(x, y, f)$ and $\frac{\partial}{\partial x} P(x, y, f)$. The first of these means: calculate the partial derivative with respect to the first variable x, treating x, y, z as independent; then substitute $f(x, y)$ for z. The second means: calculate the partial with respect to x, after making the substitution $z = f(x,y)$; the answer is

$$\frac{\partial}{\partial x} P(x,y,f) \;=\; P_1(x,y,f) + P_3(x,y,f)\, f_x\ .$$

(We use P_1 rather than P_x since the latter would be ambiguous — when you use numerical subscripts, everyone understands that the variables are being treated as independent.)

With this out of the way, the calculation of the surface integral is routine, using the standard procedure of an integral over a surface having the form $z = f(x,y)$ given in Section V9. We get

$$d\mathbf{S} \;=\; (-f_x\,\mathbf{i} - f_y\,\mathbf{j} + \mathbf{k})\,dx\,dy, \qquad \text{by V9, (13);}$$

$$\text{curl }(P(x,y,z)\ \mathbf{k}) \;=\; \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ 0 & 0 & P \end{vmatrix} \;=\; P_2(x,y,z)\,\mathbf{i} - P_1(x,y,z)\,\mathbf{j}$$

$$\iint_S \text{curl }(P(x,y,z)\ \mathbf{k}) \cdot d\mathbf{S} \;=\; \iint_S \big(-P_2(x,y,z)f_x + P_1(x,y,z)f_y\big)\,dx\,dy$$

$$(5) \qquad\qquad =\; \iint_R \big(-P_2(x,y,f)f_x + P_1(x,y,f)f_y\big)\,dx\,dy$$

We have now turned the line integral into an integral around C' and the surface integral into a double integral over R. As the final step, we show that the right sides of (4) and (5) are equal by using Green's theorem:

$$\oint_{C'} U\,dx + V\,dy \;=\; \iint_R (V_x - U_y)\,dx\,dy\ .$$

(We have to state it using U and V rather than M and N, or P and Q, since in three-space we have been using these letters for the components of the general three-dimensional field $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$.) To substitute into the two sides of Green's theorem, we need four functions:

$$\begin{aligned} V &= P(x,y,f(x,y))f_y, &\quad \text{so} \quad V_x &= (P_1 + P_3 f_x)f_y + P(x,y,f)f_{yx} \\ U &= P(x,y,f(x,y))f_x, &\quad \text{so} \quad U_y &= (P_2 + P_3 f_y)f_x + P(x,y,f)f_{xy} \end{aligned}$$

Therefore, since $f_{xy} = f_{yx}$, four terms cancel, and the right side of Green's theorem becomes

$$V_x - U_y \;=\; P_1(x,y,f)f_y - P_2(x,y,f)f_x,$$

which is precisely the integrand on the right side of (5). This completes the proof of Stokes' theorem when $\mathbf{F} = P(x,y,z)\,\mathbf{k}$.

In the same way, if $\mathbf{F} = M(x,y,z)\,\mathbf{i}$ and the surface is $x = g(y,z)$, we can reduce Stokes' theorem to Green's theorem in the yz-plane.

If $\mathbf{F} = N(x,y,z)\,\mathbf{j}$ and $y = h(x,z)$ is the surface, we can reduce Stokes' theorem to Green's theorem in the xz-plane.

Since a general field $\mathbf{F} = M\,\mathbf{i} + N\,\mathbf{j} + P\,\mathbf{k}$ can be viewed as a sum of three fields, each of a special type for which Stokes' theorem is proved, we can add up the three Stokes' theorem equations of the form (3) to get Stokes' theorem for a general vector field.

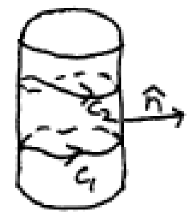
A difficulty arises if the surface cannot be projected in a 1-1 way onto each of three coordinate planes in turn, so as to express it in the three forms needed above:

$$z = f(x,y), \qquad x = g(y,z), \qquad y = h(x,z)\ .$$

In this case, it can usually be divided up into smaller pieces which can be so expressed (if some of these are parallel to one of the coordinate planes, small modifications must be made in the argument). Stokes' theorem can then be applied to each piece of surface, then the separate equalities can be added up to get Stokes' theorem for the whole surface (in the addition, line integrals over the cut-lines cancel out, since they occur twice for each cut, in opposite directions). This completes the argument, *manus undulans*, for Stokes' theorem.

Extended Stokes' Theorem

Let $\mathbf{F} = \langle 2xz + 2y, 2yz + 2yx, x^2 + y^2 + z^2 \rangle$. Take C_1 and C_2 two curves going around the circular cylinder of radius a as shown. Show $\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_2} \mathbf{F} \cdot d\mathbf{r}$.



<u>Answer:</u> We easily compute $\text{curl}\mathbf{F} = (2y - 2x)\mathbf{k} \Rightarrow \text{curl}\mathbf{F} \cdot \mathbf{n} = 0$, where $\mathbf{n}$ is the normal to the cylinder. Let S be the part of the cylinder between C_1 and C_2 then Stokes' theorem says

$\oint_{C_1 - C_2} \mathbf{F} \cdot d\mathbf{r} = \iint_S \text{curl}\mathbf{F} \cdot \mathbf{n}\, dS = 0. \Rightarrow \oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_2} \mathbf{F} \cdot d\mathbf{r}$. QED

V14. Some Topological Questions

We consider once again the criterion for a gradient field. We know that

$$\mathbf{F} = \nabla f \quad \Rightarrow \quad \text{curl } \mathbf{F} = \mathbf{0} \,, \tag{1}$$

and inquire about the converse. It is natural to try to prove that

$$\text{curl } \mathbf{F} = \mathbf{0} \quad \Rightarrow \quad \mathbf{F} = \nabla f \tag{2}$$

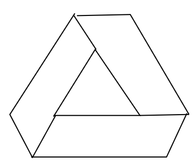
by using Stokes' theorem: if curl $\mathbf{F} = \mathbf{0}$, then for any closed curve C in space,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} \;=\; \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} \;=\; 0. \tag{3}$$

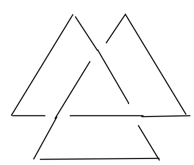
The difficulty is that we are given C, but not S. So we have to ask:

Question. Let D be a region of space in which $\mathbf{F}$ is continuously differentiable. Given a closed curve lying in D; is it the boundary of some two-sided surface S lying inside D?

We explain the "two-sided" condition. Some surfaces are only one-sided: if you start painting them, you can use only one color, if you don't allow abrupt color changes. An example is S below, formed giving three half-twists to a long strip of paper before joining the ends together.



Surface S



Boundary C (trefoil knot)

S has only one side. This means that it cannot be oriented: there is no continuous choice for the normal vector $\mathbf{n}$ over this surface. (If you start with a given $\mathbf{n}$ and make it vary continuously, when you return to the same spot after having gone all the way around, you will end up with $-\mathbf{n}$, the oppositely pointing vector.) For such surfaces, it makes no sense to speak of "the flux through S", because there is no consistent way of deciding on the positive direction for flow through the surface. Stokes' theorem does not apply to such surfaces.

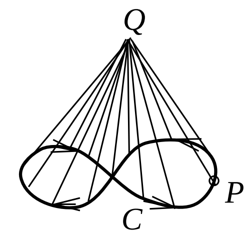
To see what practical difficulties this causes, even when the domain is all of 3-space, consider the curve C in the above picture. It's called the trefoil knot. We know it is the boundary of the one-sided surface S, but this is no good for equation (3), which requires that we find a *two-sided* surface which has C for boundary.

> There are such surfaces; try to find one. It should be smooth and not cross itself. If successful, consider yourself a brown-belt topologist.

The preceding gives some ideas about the difficulties involved in finding a two-sided surface whose boundary is a closed curve C when the curve is knotted, i.e., cannot be continuously deformed into a circle without crossing itself at some point during the deformation. It is by no means clear that such a two-sided surface exists in general.

There are two ways out of the dilemma.

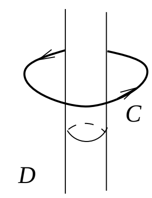
1. If we allow the surface to cross itself, and allow it to be not smooth along some lines, we can easily find such a two-sided surface whose boundary is a given closed curve C. The procedure is simple. Pick some fixed point Q not on the curve C, and join it to a point P on the curve (see the figure). Then as P moves around C, the line segment QP traces out a surface whose boundary is C. It will not be smooth at Q, and it will cross itself along a certain number of lines, but it's easy to see that this is a two-sided surface.



The point now is that Stokes' theorem can still be applied to such a surface: just use subdivision. Divide up the surface into skinny "triangles", each having one vertex at Q, and include among the edges of these triangles the lines where the surface crosses itself. Apply Stokes' theorem to each triangle, and add up the resulting equations.

2. Though the above is good enough for our purposes, it's an amazing fact that for any C there is always a smooth two-sided surface which doesn't cross itself, and whose boundary is C. (This was first proved around 1930 by van Kampen.)

The above at least answers our question affirmatively when D is all of 3-space. Suppose however that it isn't. If for instance D is the exterior of the cylinder $x^2 + y^2 = 1$, then it is intuitively clear that a circle C around the outside of this cylinder isn't the boundary of any finite surface lying entirely inside D.



A class of domains for which it is true however are the *simply-connected* ones.

Definition. A domain D in 3-space is **simply-connected** if each closed curve in it can be shrunk to a point without ever getting outside of D during the shrinking.

For example, 3-space itself is simply-connected, as is the interior or the exterior of a sphere. However the interior of a torus (a bagel, for instance) is not simply-connected, since any circle in it going around the hole cannot be shrunk to a point while staying inside the torus.

If D is simply-connected, then any closed curve C is the boundary of a two-sided surface (which may cross itself) lying entirely inside D. We can't prove this here, but it gives us the tool we need to establish the converse to the criterion for gradient fields in 3-space.

Theorem. *Let D be a simply-connected region in 3-space, and suppose that the vector field $\mathbf{F}$ is continuously differentiable in D. Then in D,*

$$\text{curl } \mathbf{F} = \mathbf{0} \quad \Rightarrow \quad \mathbf{F} = \nabla f\ . \tag{5}$$

Proof. According to the two fundamental theorems of calculus for line integrals (section V11.3), it is enough to prove that $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed curve C lying in D.

Since D is simply-connected, given such a curve C, we can find a two-sided surface S lying entirely in D and having C as its boundary. Applying Stokes' theorem,

$$\oint_C \mathbf{F} \cdot d\mathbf{R} \;=\; \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} \;=\; 0,$$

which shows that $\mathbf{F}$ is conservative, and hence that $\mathbf{F}$ is a gradient field. $\square$

Summarizing, we can say that *if D is simply-connected, the following statements are equivalent—if one is true, so are the other two:*

$$(6) \qquad \mathbf{F} = \nabla f \quad \Leftrightarrow \quad \text{curl}\, \mathbf{F} = \mathbf{0} \quad \Leftrightarrow \quad \int_P^Q \mathbf{F} \cdot d\mathbf{r} \ \ \textit{is path independent.}$$

Concluding remarks about Stokes' theorem.

Just as problems of sources and sinks lead one to consider Green's theorem in the plane for regions which are not simply-connected, it is important to consider such domains in connection with Stokes' theorem.

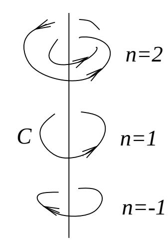
For example, if we put a closed loop of wire in space, the exterior of this loop — the region consisting of 3-space with the wire removed — is not simply-connected. If the wire carries current, the resulting electromagnetic force field $\mathbf{F}$ will satisfy curl $\mathbf{F} = \mathbf{0}$, but $\mathbf{F}$ will not be conservative. In particular, the value of $\oint \mathbf{F} \cdot d\mathbf{r}$ around a closed path which links with the loop will *not* in general be zero, (which explains why you can get power from a wire carrying current, even though the curl of its electromagnetic field is zero).

As an example, consider the vector field in 3-space

$$\mathbf{F} \;=\; \frac{-y\,\mathbf{i} + x\,\mathbf{j}}{r^2}, \qquad\qquad r = \sqrt{x^2 + y^2}\ .$$

The domain of definition is xyz-space, with the z-axis removed (since the z-axis is where $r = 0$). Just as before, (Section V2,p.2), we can calculate that $\oint_C \mathbf{F} \cdot d\mathbf{r} = 2\pi$ if C is a suitably directed circle lying in a plane $z = z_0$ and centered on the z-axis.

> **Exercise.** By using Stokes' theorem for a surface with more than one boundary curve, show informally that for the field above, $\oint_C \mathbf{F} \cdot d\mathbf{r} = 2\pi$ for any closed curve C going once around the z-axis, oriented so when that when the right thumb points in the direction $\mathbf{k}$, the fingers curl in the positive direction on C. Then show that if C goes around n times, the value of the integral will be $2n\pi$.



Suppose now the wire is a closed curve that is *knotted*, i.e., it cannot be continuously deformed to a circle, without crossing itself at some point in the deformation. Let D be the exterior of the wire loop, and consider the value of $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for a vector field $\mathbf{F}$ defined in D and having curl $\mathbf{F} = \mathbf{0}$. If one closed curve C_1 can be deformed into another closed curve C_2 without leaving D (i.e., without crossing the wire), then by using Stokes' theorem for surfaces with two boundary curves, we conclude

$$(6) \qquad \oint_{C_1} \mathbf{F} \cdot d\mathbf{r} \;=\; \oint_{C_2} \mathbf{F} \cdot d\mathbf{r}\ .$$

More generally, two closed curves C_1 and C_2 are called *homologous*, written $C_1 \sim C_2$, if C_1 and C_2' (this means C_2 with its direction reversed) form the complete boundary of some surface lying entirely in D. Then by an extended form of Stokes' theorem, (6) will be true whenever $C_1 \sim C_2$. Thus the problem of determining the possible values for the line integral is reduced to the purely topological problem of finding a set of closed curves in D, no two of which are homologous, but such that every other closed curve is homologous to one of the curves in the set. For any particular knot in 3-space, such a set can be determined by an algorithm, but if one asks for general results relating the appearance of the knot to the number of such basic curves that will be needed, one gets into unsolved problems of topology.

In another vein, the theorems of Green, Stokes, and Gauss (as the divergence theorem is often called) all relate an integral over the interior of some closed curve or surface with an integral over its boundary. There is a much more general result — the generalized Stokes' theorem — which connects an integral over an n-dimensional hypersurface with an integral taken over its $n-1$-dimensional boundary. Green's and Stokes' theorems are the case $n = 2$ of this result, while the divergence theorem is closely related to the case $n = 3$ in 3-space. Just as the theorems we have studied are the key to an understanding of geometry and analysis in the plane and space, so this theorem is central to an understanding of n-dimensional space, and more general sorts of spaces called "n-dimensional manifolds".