Mathematics IV (Electrical) MAS243

The point of this course

Vector calculus is the mathematical language of the laws of physics. Electromagnetism is governed by Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \qquad \nabla \times \mathbf{E} = -\dot{\mathbf{B}}$$
$$\nabla \cdot \mathbf{B} = 0 \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \dot{\mathbf{E}}.$$

Incompressible fluid flow is governed by the Navier-Stokes equation:

$$\rho(\dot{\mathbf{v}} + \mathbf{v}.\nabla\mathbf{v}) = -\nabla p + \mu\nabla^2\mathbf{v} + \mathbf{f}.$$

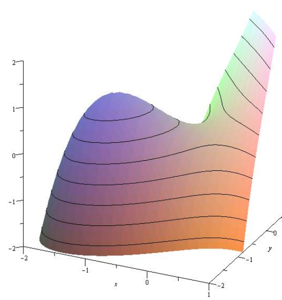
We need to understand the various different manifestations of the differential operator ∇. Moreover, these equations involve quantities like $\mathbf{E}$ and $\mathbf{v}$, the electric field and the fluid velocity at a single point. To calculate the total energy of the electric field in a region, or the total fluid flow through a pipe, we need to integrate (with respect to several variables). We need to understand how integration in several variables relates to differentiation in several variables, generalising the fact that $\int_a^b f'(x)\,dx = f(b) - f(a)$. For this we need Stokes's theorem and the Divergence theorem.

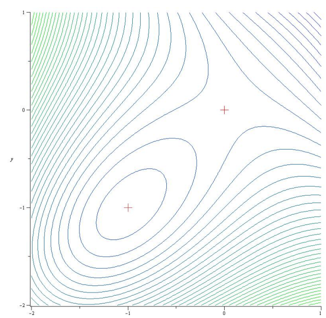
The syllabus

- Maxima and minima of functions of two or more variables.
- Constraints and Lagrange multipliers.
- Integration over two-dimensional regions, starting with rectangles and circles. Change of order, change of variables. Applications.
- Introduction to three-dimensional surfaces.
- Spherical and cylindrical coordinate systems, and integration in terms of such coordinates.
- Revision of vectors. Scalar and vector products, derivatives, vector equations for lines.
- Scalar and vector fields. Examples, including temperature and electric and magnetic fields.
- Div, grad and curl. Geometric and physical significance. Identities.
- Iterated operators: $\nabla.(\nabla \times u) = 0$, $\nabla \times \nabla f = 0$, $\nabla.(\nabla f) = \nabla^2(f)$. Closure and exactness.
- Unit vectors and differential operators in cylindrical and polar coordinates.
- Line integrals of scalars and vectors. Path independence for exact vector fields.
- Stokes's theorem and the Divergence theorem. Examples and applications.

3D Diagrams

There will be many three-dimensional diagrams for this course. It is often helpful to have a version that you can rotate with your mouse to inspect from different angles. Unfortunately I cannot embed such versions in these slides, but they will be available on the course web page.





Reminder on partial derivatives

- A function like $u = x^2y + y^3z$ has three different derivatives, with respect to the three variables involved.

$$u_x = \frac{\partial u}{\partial x} = 2xy \qquad u_y = \frac{\partial u}{\partial y} = x^2 + 3y^2z \qquad u_z = \frac{\partial u}{\partial z} = y^3$$

- These partial derivatives measure the sensitivity of u to small changes in x, y and z. If these variables change by small amounts δx, δy and δz, then the resulting change δu in u is approximately

$$\delta u \simeq u_x\,\delta x + u_y\,\delta y + u_z\,\delta z.$$

- We have the usual product rule and quotient rule:

$$(uv)_x = u_x v + uv_x \qquad (u/v)_x = \frac{u_x v - uv_x}{v^2}$$

and similarly for the partial derivatives with respect to y or z.

- We also have a chain rule: if v is a function of x, y and z, and $u = f(v)$, then

$$u_x = f'(v)v_x \qquad u_y = f'(v)v_y \qquad u_z = f'(v)v_z.$$

Double partial derivatives

Consider again $u = x^2y + y^3z$.

$$\begin{array}{lll} u_x = 2xy & u_y = x^2 + 3y^2z & u_z = y^3 \\ u_{xx} = 2y & u_{xy} = 2x & u_{xz} = 0 \\ u_{yx} = 2x & u_{yy} = 6yz & u_{yz} = 3y^2 \\ u_{zx} = 0 & u_{zy} = 3y^2 & u_{zz} = 0 \end{array}$$

Note that $u_{xy} = u_{yx}$, $u_{xz} = u_{zx}$ and $u_{yz} = u_{zy}$. This is a general principle: if we take partial derivatives with respect to two different variables, then the order does not matter.

We can write all the second-order partial derivatives as a symmetric square matrix, called the *Hessian matrix*:

$$H = \begin{bmatrix} u_{xx} & u_{xy} & u_{xz} \\ u_{yx} & u_{yy} & u_{yz} \\ u_{zx} & u_{zy} & u_{zz} \end{bmatrix} = \begin{bmatrix} 2y & 2x & 0 \\ 2x & 6yz & 3y^2 \\ 0 & 3y^2 & 0 \end{bmatrix}.$$

Partial derivatives example

Take $P = V^2/R$ (power dissipated in a resistor).

$$\begin{aligned} P_V &= 2V/R \\ P_R &= -V^2/R^2 \\ P_{VV} &= 2/R \\ P_{VR} = P_{RV} &= -2V/R^2 \\ P_{RR} &= 2V^2/R^3 \end{aligned}$$

$$H = \begin{bmatrix} P_{VV} & P_{VR} \\ P_{RV} & P_{RR} \end{bmatrix} = \begin{bmatrix} 2/R & -2V/R^2 \\ -2V/R^2 & 2V^2/R^3 \end{bmatrix}$$

Partial derivatives example

Consider the function $u = a + ab^2 + ab^2c^3$.

$$\begin{aligned} u_a &= 1 + b^2 + b^2c^3 \\ u_b &= 2ab + 2abc^3 \\ u_c &= 3ab^2c^2 \\ u_{aa} &= 0 \\ u_{ab} = u_{ba} &= 2b + 2bc^3 \\ u_{ac} = u_{ca} &= 3b^2c^2 \\ u_{bb} &= 2a + 2ac^3 \\ u_{bc} = u_{cb} &= 6abc^2 \\ u_{cc} &= 6ab^2c. \end{aligned}$$

The Hessian is therefore

$$H = \begin{bmatrix} u_{aa} & u_{ab} & u_{ac} \\ u_{ba} & u_{bb} & u_{bc} \\ u_{ca} & u_{cb} & u_{cc} \end{bmatrix} = \begin{bmatrix} 0 & 2b + 2bc^3 & 3b^2c^2 \\ 2b + 2bc^3 & 2a + 2ac^3 & 6abc^2 \\ 3b^2c^2 & 6abc^2 & 6ab^2c \end{bmatrix}.$$

Partial derivatives example

Consider $f(x,y,z) = \ln(ax+by+cz)$ (a,b,c constant).Consider $f(x,y,z) = \ln(v)$, where $v = ax+by+cz$

$$f_x(x,y,z) = \ln'(v)\,\partial v/\partial x = a/v$$
$$f_y(x,y,z) = \ln'(v)\,\partial v/\partial y = b/v$$
$$f_z(x,y,z) = \ln'(v)\,\partial v/\partial z = c/v$$

For the second derivatives, we have

$$f_{xy}(x,y,z) = \frac{\partial}{\partial y} f_x(x,y,z) = \frac{\partial}{\partial y}\left(\frac{a}{v}\right) = \frac{-a}{v^2}\frac{\partial v}{\partial y} = \frac{-ab}{v^2}.$$

Proceeding in the same way, we see that

$$\begin{array}{lll} f_{xx}(x,y,z) = -a^2/v^2 & f_{xy}(x,y,z) = -ab/v^2 & f_{xz}(x,y,z) = -ac/v^2 \\ f_{yx}(x,y,z) = -ab/v^2 & f_{yy}(x,y,z) = -b^2/v^2 & f_{yz}(x,y,z) = -bc/v^2 \\ f_{zx}(x,y,z) = -ac/v^2 & f_{zy}(x,y,z) = -bc/v^2 & f_{zz}(x,y,z) = -c^2/v^2 \end{array}$$

This means that the Hessian matrix is

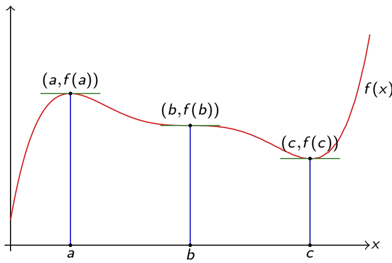
$$H = \frac{-1}{(ax+by+cz)^2}\begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}.$$

Optimisation

- Often we have a function f of one or more variables, and we want to know the maximum and minimum possible values of f.
- Consider a device involving a strong electric field of strength $E(x,y,z)$. To check whether there will be arcing, we need to know the maximum value of E.
- Consider a circuit with resistors R_1, R_2, R_3 and capacitors C_1 and C_2. The quality of the circuit is measured by some function $Q(R_1,R_2,R_3,C_1,C_2)$. To choose the best component values, we want to find where Q takes its maximum value.
- For a function $f(x)$ (with only one variable) just find where $f'(x) = 0$.
- To distinguish between maxima and minima, consider $f''(x)$.
- More detail on the following slides.
- For a function of several variables, look for points where *all* partial derivatives vanish.
- To distinguish between maxima, minima and saddles, consider the Hessian matrix of all second-order partial derivatives.

Optimisation with one variable

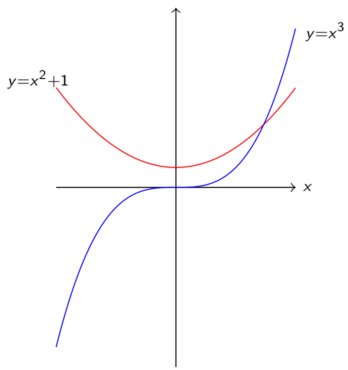
- The *critical points* of $f(x)$ are the values of x where $f'(x) = 0$.



  In the picture, a, b and c are the critical points.
- There is a *local maximum* at a: for x close to a we have $f(x) \leq f(a)$. Local maxima are always critical points.
- There is a *local minimum* at c: for x close to c we have $f(x) \geq f(c)$. Local minima are always critical points.
- There is an *inflection point* at b: a critical point that is neither a local minimum nor a local maximum. These are rare.
- To find the maximum and minimum of $f(x)$, you should start by solving $f'(x) = 0$ to find the critical points.

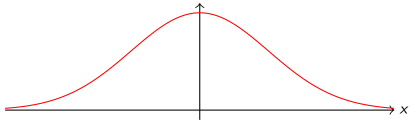
Functions with no maximum or minimum

- It can happen that $f(x)$ does not have a maximum or minimum.
- The function $y = x^2 + 1$ has a minimum value of 1 at $x = 0$, but has no maximum value.
- The function $y = x^3$ has no minimum and no maximum.
- When trying to find maxima and minima, you should remember the possibility that they might not exist.

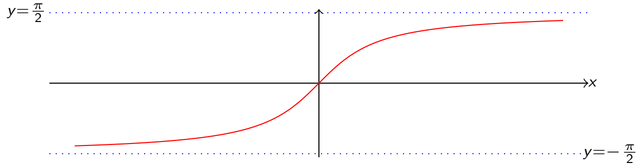


Limiting maxima/minima that are not attained

- Here is the graph of $y = e^{-x^2}$:



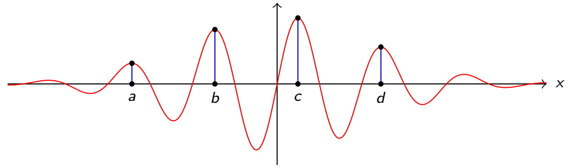
- The maximum value is $y = 1$, attained at $x = 0$.
- The minimum value is $y = 0$, but this is never reached. Instead, the graph approaches $y = 0$ arbitrarily closely for large (positive or negative) x.
- Here is the graph of $y = \arctan(x) = \tan^{-1}(x)$:



- The minimum value is $y = -\pi/2$ and the maximum is $y = \pi/2$, but neither of these limits is ever reached.
- In these cases we cannot find the maximum and minimum values by looking for critical points.

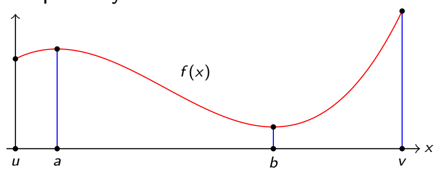
Other complications

- The function $y = e^{-x^2}\sin(10x)$ has local maxima at a, b, c and d. Only the one at $x = c$ is a global maximum.



To find the global maximum, we need to check *all* the critical points.

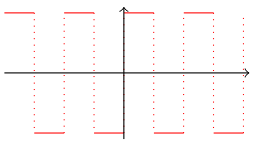
- If we are only interested in $u \leq x \leq v$, we need to check the endpoints $x = u$ and $x = v$ separately.



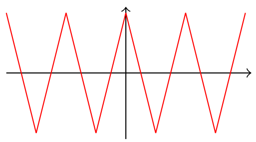
Here the minimum is at $x = b$, which is a critical point. However, the maximum is at the endpoint $x = v$, which is not a critical point. To find the maximum, we need to check the endpoints as well as the critical points.

Bad functions

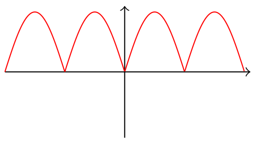
These methods do not work well if the graph of $f(x)$ has jumps or kinks.



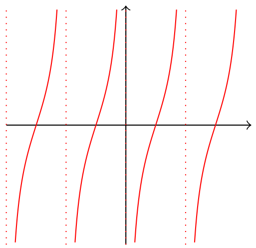
square wave



sawtooth

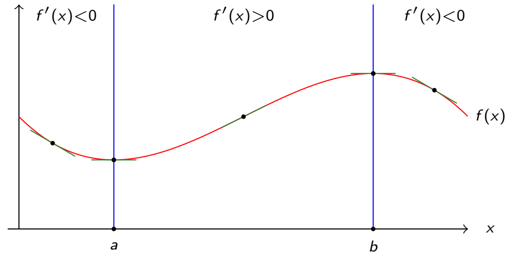


rectified sin wave



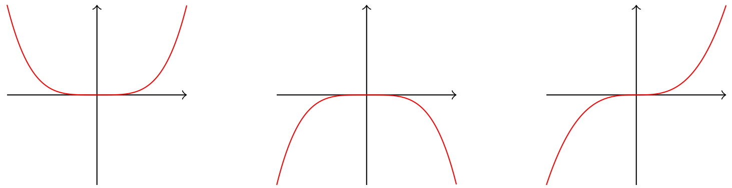
cotangent

Classifying critical points



- The slope $f'(x)$ is negative to the left of a, and zero at a, and positive to the right of a; so $f''(a) > 0$.
- The slope $f'(x)$ is positive to the left of b, and zero at b, and negative to the right of b; so $f''(b) < 0$.
- In general, a critical point with $f'' > 0$ is a local minimum, and a critical point with $f'' < 0$ is a local maximum.
- It can happen that there is a critical point where $f'' = 0$. Such points are rare. They may be local maxima, local minima or neither.

What if $f''(x) = 0$ at a critical point?



$f(x) = x^4$ $\quad$ $f(x) = -x^4$ $\quad$ $f(x) = x^3$

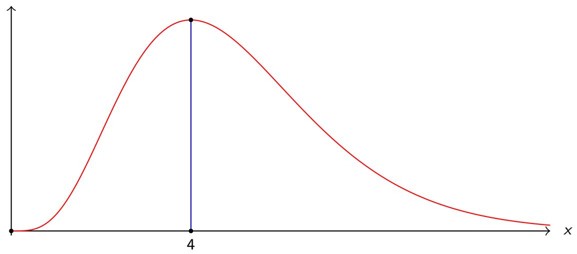
- These three functions each have a critical point at $x = 0$ with $f''(0) = f'(0) = 0$.
- The first has a local minimum at $x = 0$, the second has a local maximum, and the third has an *inflection point* (a critical point that is neither a local maximum nor a local minimum).
- We will not take the time to explain general rules for this situation, as it does not occur very often.

The function $f(x) = x^4e^{-x}$

- Consider $f(x) = x^4e^{-x}$ for $x \geq 0$.
- Note that $f(0) = 0$. When x is large the decay of e^{-x} wipes out the growth of x^4 (exponentials beat polynomials) and so $f(x) \to 0$.
- The first derivative is

$$f'(x) = 4x^3e^{-x} + x^4.(-e^{-x}) = (4x^3 - x^4)e^{-x} = x^3(4-x)e^{-x}.$$

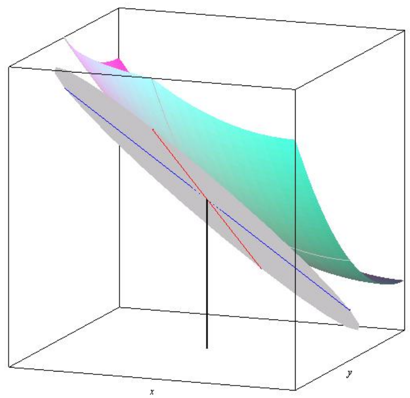
- For $0 < x < 4$ we see that all three factors (x^3, $4 - x$ and e^{-x}) are positive so $f'(x) > 0$ and the function increases from zero.
- For $x > 4$ we have $x^3, e^{-x} > 0$ but $4 - x < 0$ so $f'(x) < 0$ and the function decreases back to zero.
- The minimum is zero, and the maximum is $f(4) = 256e^{-4} \simeq 4.69$.



Critical points in two variables

- For a function $f(x, y)$, a *critical point* is a point where f_x and f_y both vanish.
- Consider $f(x, y) = 2x^2 + 2xy - 6x + y^2 - 4y + 5$, so $f_x(x, y) = 4x + 2y - 6$ and $f_y(x, y) = 2x + 2y - 4$.
- At a critical point we must have (A) $4x + 2y = 6$ and (B) $2x + 2y = 4$ so $(A - B)$ $2x = 2$ so $x = 1$ so $y = 1$. Thus, the only critical point is at $(x, y) = (1, 1)$.
- If we make small changes δx and δy to x and y, the resulting change in f is $\delta f \simeq f_x.\delta x + f_y.\delta y$.
- By taking δx and δy to have the same sign as f_x and f_y, we can make $\delta f > 0$ unless $f_x = f_y = 0$.
- If we are at a local maximum it is impossible to have $\delta f > 0$, so we must have $f_x = f_y = 0$, so we have a critical point.
- Similarly, local minima are always critical points.

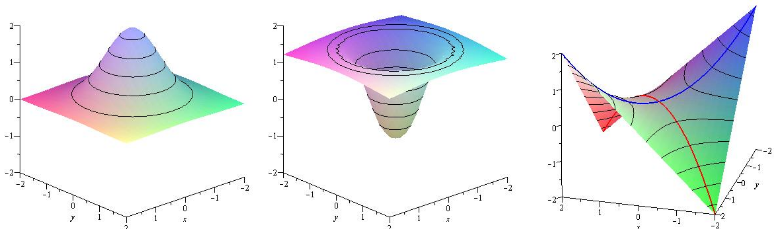
Geometry of partial derivatives



Here is a surface $z = f(x, y)$ with a marked point. The grey disk is the tangent plane to the surface at the point $(a, b, f(a, b))$. This will be horizontal if (a, b) is a critical point (which is clearly not the case for this example.)

Types of critical points

There are three main types of critical point for a function of two variables: local maxima, local minima and saddle points.



If we walk along the blue curve, the saddle looks like a local minimum. If we walk along the red curve, it looks like a local maximum instead. Saddle points are common, not like inflection points. If we have found a critical point and we want to know whether it is a local maximum, a local minimum or a saddle, we need to look at the Hessian matrix.

Classifying critical points

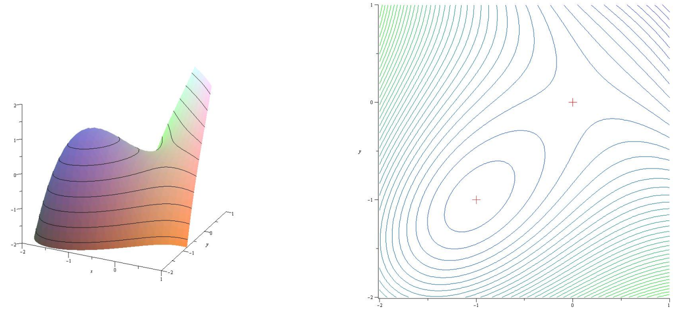
- Suppose $f(x,y)$ has a critical point at (a,b) (so $f_x(a,b) = f_y(a,b) = 0$). Consider the Hessian matrix
$$H = \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix} = \begin{bmatrix} p & q \\ q & r \end{bmatrix} \text{ say.}$$
- The *eigenvalues* of H are the numbers
$$e_1 = (p+r-\sqrt{(p-r)^2+4q^2})/2 \qquad e_2 = (p+r+\sqrt{(p-r)^2+4q^2})/2.$$
- If $e_1, e_2 < 0$ then there is a local maximum at (a,b).
- If $e_1 < 0 < e_2$ then there is a saddle point at (a,b).
- If $0 < e_1, e_2$ then there is a local minimum at (a,b).
- In the rare case where $e_1 = 0$ or $e_2 = 0$, the situation is more complicated (as with inflection points).
- Alternatively, put $A_1 = f_{xx}(a,b) = p$ and $A_2 = \det(H) = pr - q^2$.
- If $A_2 > 0$ and $A_1 < 0$ there is a local maximum.
- If $A_2 > 0$ and $A_1 > 0$ there is a local minimum.
- If $A_2 < 0$ there is a saddle point.

The function $f(x,y) = x^3 + 3xy + y^3$

- Take $f(x,y) = x^3 + 3xy + y^3$. The derivatives are
$$f_x(x,y) = 3x^2 + 3y \qquad f_y(x,y) = 3x + 3y^2$$
$$f_{xx}(x,y) = 6x \qquad f_{xy}(x,y) = 3 \qquad f_{yy}(x,y) = 6y.$$
- At a critical point we must have $3x^2 + 3y = 0$ and $3x + 3y^2 = 0$, so $y = -x^2$ and $x = -y^2$. This gives $x = -y^2 = -(-x^2)^2 = -x^4$, so $x^4 + x = 0$, so $x(x^3+1) = 0$, so $x = 0$ or $x = -1$. We also have $y = -x^2$, so if $x = 0$ we have $y = 0$, and if $x = -1$ we have $y = -1$. This means that the critical points are $(0,0)$ and $(-1,-1)$.
- $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = \begin{bmatrix} 6x & 3 \\ 3 & 6y \end{bmatrix}$
$A_1 = 6x$; $A_2 = 6x \times 6y - 3 \times 3 = 9(4xy - 1)$
- At $(0,0)$: $A_2 = -9 < 0$ so we have a saddle point.
- At $(-1,-1)$: $A_2 = 27 > 0$, $A_1 = -6 < 0$ so we have a local maximum.

The function $f(x,y) = x^3 + 3xy + y^3$

$f = x^3 + 3xy + y^3$; saddle point at $(0,0)$; local maximum at $(-1,-1)$.



$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$; $A_1 = f_{xx}$; $A_2 = \det(H) = f_{xx}f_{yy} - f_{xy}^2$

$A_2 < 0$: saddle; $A_1, A_2 > 0$: local minimum; $A_1 < 0$, $A_2 > 0$: local maximum.

For f as above: $A_1 = 6x$, $A_2 = 9(4xy - 1)$.
At $(0,0)$: $A_1 = 0$, $A_2 = -9 < 0$, saddle point.
At $(-1,1)$: $A_1 = -6 < 0$, $A_2 = 27 > 0$, local maximum.

The function $\sin(x)\sin(y)$

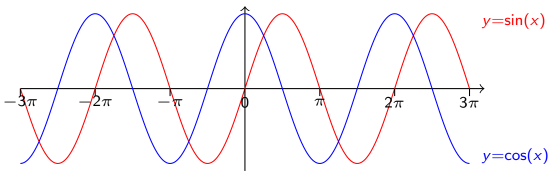
- Put $f(x,y) = \sin(x)\sin(y)$. The derivatives are

$$f_x(x,y) = \cos(x)\sin(y) \qquad\qquad f_y(x,y) = \sin(x)\cos(y)$$
$$f_{xx}(x,y) = -\sin(x)\sin(y) \quad f_{xy}(x,y) = \cos(x)\cos(y) \quad f_{yy}(x,y) = -\sin(x)\sin(y).$$

- At a critical point, we must have $\cos(x)\sin(y) = 0$ and $\sin(x)\cos(y) = 0$. Thus one of the following holds:

$$(p): \quad \cos(x) = \sin(x) = 0 \qquad\qquad (q): \quad \cos(x) = \cos(y) = 0$$
$$(r): \quad \sin(y) = \sin(x) = 0 \qquad\qquad (s): \quad \sin(y) = \cos(y) = 0.$$

but (p) and (s) cannot happen, because $\sin^2 + \cos^2 = 1$.
- $\sin(t) = 0$ for $t = n\pi$; $\cos(t) = 0$ for $t = (n + \frac{1}{2})\pi$.

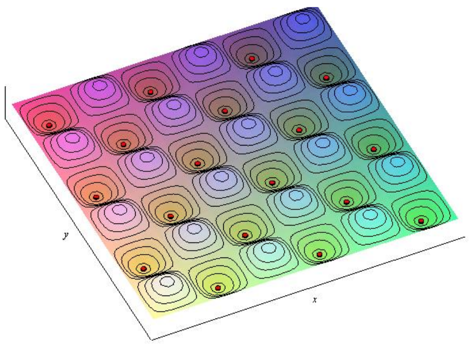


The function $\sin(x)\sin(y)$

$f(x,y) = \sin(x)\sin(y)$; $f_x(x,y) = \cos(x)\sin(y)$; $f_y(x,y) = \sin(x)\cos(y)$.
At critical point, either (r): $\sin(x) = \sin(y) = 0$ or (q): $\cos(x) = \cos(y) = 0$.
$\sin(t) = 0$ for $t = n\pi$; $\cos(t) = 0$ for $t = (n + \frac{1}{2})\pi$.

- In case (r): $x = n\pi$ and $y = m\pi$ for some integers n, m.
- In case (q): $x = (n + \frac{1}{2})\pi$ and $y = (m + \frac{1}{2})\pi$ for some integers n, m.
- $H = \begin{bmatrix} -\sin(x)\sin(y) & \cos(x)\cos(y) \\ \cos(x)\cos(y) & -\sin(x)\sin(y) \end{bmatrix}$
- $A_1 = -\sin(x)\sin(y)$; $A_2 = \sin(x)^2\sin(y)^2 - \cos(x)^2\cos(y)^2$.
- At $(n\pi, m\pi)$: $A_2 = -1 < 0$, so we have a saddle point.
- Note that $\cos((k + \frac{1}{2})\pi) = 0$, $\sin((k + \frac{1}{2})\pi) = (-1)^k$.
- At $((n + \frac{1}{2})\pi, (m + \frac{1}{2})\pi)$: $A_1 = (-1)^{n+m+1}$, $A_2 = 1 > 0$.
- If $n + m$ is even: $A_1 = -1 < 0$ and $A_2 = 1 > 0$ so we have a local maximum.
- If $n + m$ is odd: $A_1 = 1 > 0$ and $A_2 = 1 > 0$ so we have a local minimum.

The function $\sin(x)\sin(y)$



- There are saddle points at $(n\pi, m\pi)$ for all integers n and m.
- There is a local maximum at $((n + \frac{1}{2})\pi, (m + \frac{1}{2})\pi)$ whenever $n + m$ is even.
- There is a local minimum at $((n + \frac{1}{2})\pi, (m + \frac{1}{2})\pi)$ whenever $n + m$ is odd.

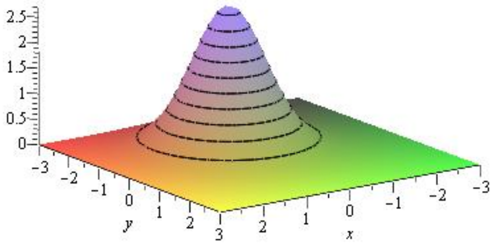
The function $e^{-x^2-y^2-2y}$

- Take $f(x,y) = e^{-x^2-y^2-2y}$.
- $f_x = -2xe^{-x^2-y^2-2y} = -2xf$
- $f_y = (-2y - 2)e^{-x^2-y^2-2y} = (-2y - 2)f$
- For a critical point, we must have $-2xf = 0$ and $(-2y - 2)f = 0$. Here f is never zero, so $-2x = -2y - 2 = 0$ so $(x,y) = (0,-1)$.
- $f_{xx} = (-2xf)_x = -2f - 2xf_x = -2f + 4x^2 f = (4x^2 - 2)f$
- $f_{xy} = (-2xf)_y = -2xf_y = -2x(-2y - 2)f = (4xy + 4x)f$
- $f_{yy} = ((-2y - 2)f)_y = -2f + (-2y - 2)f_y = -2f + (2y + 2)^2 f = (4y^2 + 8y + 2)f$
- At the critical point $(x,y) = (0,-1)$ we have $f = e^{-0-1+2} = e$ so

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = \begin{bmatrix} 4x^2 - 2 & 4xy + 4x \\ 4xy + 4x & 4y^2 + 8y + 2 \end{bmatrix} f = \begin{bmatrix} -2e & 0 \\ 0 & -2e \end{bmatrix}$$

- Now $A_1 = -2e < 0$ and $A_2 = (-2e)^2 - 0^2 = 4e^2 > 0$ so we have a local maximum at $(0,-1)$.

The function $e^{-x^2-y^2-2y}$



There is a local maximum at $(0,-1)$ and no other critical points.

Complications

- As in the one-variable case, there are complications in the relationship between critical points and maxima/minima.
- The function may not have a maximum or minimum.
- It may have a maximum or minimum that is approached arbitrarily closely, but never actually reached.
- A function defined on a restricted region may have a maximum or minimum on the boundary of that region, and this may not be a critical point.
- Differential methods may not work well for functions that are not sufficiently smooth.
- These phenomena can be important, but we will not discuss them further here.

More variables

- Not very much changes if there are more than two variables.
- Critical points are points where all partial derivatives vanish.
- Local maxima and minima are always critical points.
- There may be other critical points, which are saddles of various kinds.
- If there are n variables, then the Hessian is an $n \times n$ symmetric matrix.
- If all the eigenvalues are positive, we have a local minimum. If all are negative, we have a local maximum. If some are negative and some are positive, we have a saddle.
- Let A_k be the determinant of the top left $k \times k$ block in the Hessian. If the variables are w, x, y, z we have

$$A_1 = f_{ww} \quad A_2 = \det\begin{bmatrix} f_{ww} & f_{wx} \\ f_{xw} & f_{xx} \end{bmatrix} \quad A_3 = \det\begin{bmatrix} f_{ww} & f_{wx} & f_{wy} \\ f_{xw} & f_{xx} & f_{xy} \\ f_{yw} & f_{yx} & f_{yy} \end{bmatrix} \quad A_4 = \det\begin{bmatrix} f_{ww} & f_{wx} & f_{wy} & f_{wz} \\ f_{xw} & f_{xx} & f_{xy} & f_{xz} \\ f_{yw} & f_{yx} & f_{yy} & f_{yz} \\ f_{zw} & f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

- If all A_k are positive we have a local minimum. If $A_1 < 0$, $A_2 > 0$, $A_3 < 0$, $A_4 > 0$ etc then we have a local maximum. Otherwise, provided that the last A is nonzero, we have a saddle.

An example with three variables

- Take $f(x,y,z) = 8(x^2+y^2+z^2) - (z+1)^3$
- We have $f_x = 16x$ and $f_y = 16y$ and

$$f_z = 16z - 3(z+1)^2 = 16z - 3z^2 - 6z - 3 = -3 + 10z - 3z^2 = (z-3)(1-3z).$$

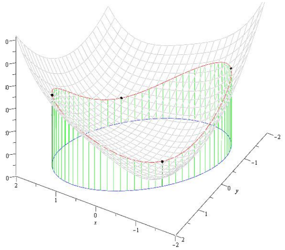
- We have critical points where $x = y = 0$ and $(z-3)(1-3z) = 0$ so $z = 3$ or $z = 1/3$.
- The Hessian matrix is

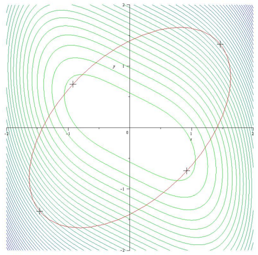
$$H = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} = \begin{bmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 10-6z \end{bmatrix}$$

- $A_1 = 16$ and $A_2 = 256$ and $A_3 = 256(10-6z)$.
- At $(0,0,1/3)$ we get $A_3 = 2048$ so A_1, A_2 and A_3 are all positive, so we have a local minimum.
- At $(0,0,3)$ we have $A_3 = -2048 \neq 0$ and the signs do not alternate so we have some kind of saddle.

Constrained optimisation

- So far we have tried to find the maximum value of a function $f(x, y)$, where both x and y can vary freely.
- This is like looking for the highest point in a certain area of land.
- What if we want to find the highest point on the road instead?
- The road will be given by some equation which we can put in the form $g(x, y) = 0$. For example, $g(x, y) = x^2 + y^2 - 4$ corresponds to a circular road, and $g(x, y) = x + y - 6$ corresponds to an infinite straight road.
- We want to maximise $f(x, y)$ subject to the constraint $g(x, y) = 0$.
- The maximum and minimum occur at points where the road is tangent to the contours.





Constrained optimisation - applications

- Suppose we want to build a 5kW motor that is as light as possible. We have come up with a design with parameters a, b and c that we can adjust. The weight is $W(a, b, c)$ and the power (in kW) is $P(a, b, c)$. We want to minimise $W(a, b, c)$ subject to the constraint $P(a, b, c) - 5 = 0$.
- More generally, whenever we design a device, there will be some requirements that are not negotiable; these will be expressed by constraint equations. There will be other functions that measure the effectiveness of the device. We want to maximise these, but we have to do so subject to the constraints.

The Lagrange multiplier method

- To maximise or minimise $f(x, y)$ subject to $g(x, y) = 0$, we find the (unconstrained) critical points of the function $L(\lambda, x, y) = f(x, y) - \lambda g(x, y)$.
- For example, suppose we want to minimise $f(x, y) = x^2 + y^2$ subject to $3x + 4y = 5$. Then $g(x, y) = 3x + 4y - 5$ so

$$L(\lambda, x, y) = x^2 + y^2 - \lambda(3x + 4y - 5).$$

- For a critical point, the derivatives must vanish:

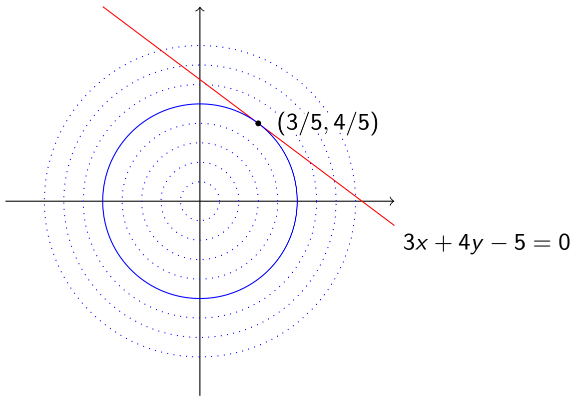
$$\begin{aligned} L_\lambda &= -3x - 4y + 5 = 0 && \text{(A)} \\ L_x &= 2x - 3\lambda = 0 && \text{(B)} \\ L_y &= 2y - 4\lambda = 0. && \text{(C)} \end{aligned}$$

- Here (B) and (C) give $x = 3\lambda/2$ and $y = 2\lambda$. Substituting these values in (A) gives $-9\lambda/2 - 8\lambda + 5 = 0$, which simplifies to $\lambda = 2/5$. This in turn gives $x = 3\lambda/2 = 3/5$ and $y = 2\lambda = 4/5$. Thus, the only critical point is $(x, y) = (3/5, 4/5)$. The value of f here is $(3/5)^2 + (4/5)^2 = 1$.

Geometric interpretation

$f(x, y) = x^2 + y^2 =$ squared distance from (x, y) to $(0, 0)$
$g(x, y) = 3x + 4y - 5 = 0$; minimum value of f is 1 at $(3/5, 4/5)$.

Geometrically, we have found the closest point to the origin on the line $3x + 4y = 5$.

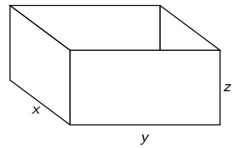


Why does the method (usually) work?

- Suppose that (λ, x, y) is a critical point for $L(\lambda, x, y) = f(x, y) - \lambda g(x, y)$.
- We have $L_\lambda(\lambda, x, y) = -g(x, y)$, and this must be zero as we are at a critical point of L. This means that we are on the constraint curve.
- We also have $L_x = L_y = 0$, which means that $f_x = \lambda g_x$ and $f_y = \lambda g_y$ (at this point).
- Now suppose we move a little way along the constraint curve, by $(\delta x, \delta y)$ say.
- The change in g is $\delta x.g_x + \delta y.g_y$.
- As we are staying on the constraint curve, g is still zero, so we must have $\delta x.g_x + \delta y.g_y = 0$.
- This means that $\delta x.\lambda g_x + \delta y.\lambda g_y = 0$, so $\delta x.f_x + \delta y.f_y = 0$, so $\delta f = 0$.
- We see from this that (x, y) is a critical point for the constrained problem.
- Geometrically, the vector $\mathbf{u} = \begin{bmatrix} g_x \\ g_y \end{bmatrix}$ is normal to the constraint curve, and $\mathbf{v} = \begin{bmatrix} f_x \\ f_y \end{bmatrix}$ is normal to the contour of f. The equations $f_x = \lambda g_x$ and $f_y = \lambda g_y$ say that $\mathbf{v}$ is a multiple of $\mathbf{u}$, so the constraint curve is running parallel to the contour.

A constrained optimization example

Consider a metal tank, open at the top. The volume is $V = xyz$, and the area is $S = xy + 2xz + 2yz$. We want the volume to be $4m^3$, and we want to minimise S, to use as little metal as possible.



- We are minimising S subject to $V - 4 = 0$, so $L = xy + 2xz + 2yz - \lambda(xyz - 4)$.
- Equations are

$$
\begin{aligned}
L_\lambda &= 4 - xyz = 0 & xyz &= 4 & \text{(A)} \\
L_x &= y + 2z - \lambda yz = 0 & z^{-1} + 2y^{-1} &= \lambda & \text{(B)} \\
L_y &= x + 2z - \lambda xz = 0 & z^{-1} + 2x^{-1} &= \lambda & \text{(C)} \\
L_z &= 2x + 2y - \lambda xy = 0 & 2y^{-1} + 2x^{-1} &= \lambda. & \text{(D)}
\end{aligned}
$$

- Subtract (B) and (C) to get $x^{-1} = y^{-1}$ so $x = y$. Substitute in (D) to get $4x^{-1} = 4y^{-1} = \lambda$, so $x = y = 4/\lambda$. Substitute in (C) to get $z^{-1} + \lambda/2 = \lambda$, so $z = 2/\lambda$. Substitute in (A) to get $32 = 4\lambda^3$, so $\lambda = 2$, so $(x, y, z) = (2, 2, 1)$.
- For these values, we have $S = 12$. Thus, the minimum possible area of metal sheet that we need is $12m^2$.

A constrained optimisation example

- Problem: maximise $f(x, y) = x + y$ subject to $x^2/a + y^2/b = 1$ (for some constants $a, b > 0$).
- Take $L = x + y - \lambda(x^2/a + y^2/b - 1)$. For a critical point:

$$
\begin{aligned}
L_\lambda &= 1 - x^2/a - y^2/b = 0 & x^2/a + y^2/b &= 1 & \text{(A)} \\
L_x &= 1 - 2x\lambda/a = 0 & x &= a/(2\lambda) & \text{(B)} \\
L_y &= 1 - 2y\lambda/b = 0 & y &= b/(2\lambda). & \text{(C)}
\end{aligned}
$$

- Substitute (B) and (C) in (A) to get $(a + b)/(4\lambda^2) = 1$, so $\lambda = \pm\sqrt{a + b}/2$. As $x = a/(2\lambda)$ and $y = b/(2\lambda)$ this gives

$$(x, y) = \pm\left(\frac{a}{\sqrt{a + b}}, \frac{b}{\sqrt{a + b}}\right).$$

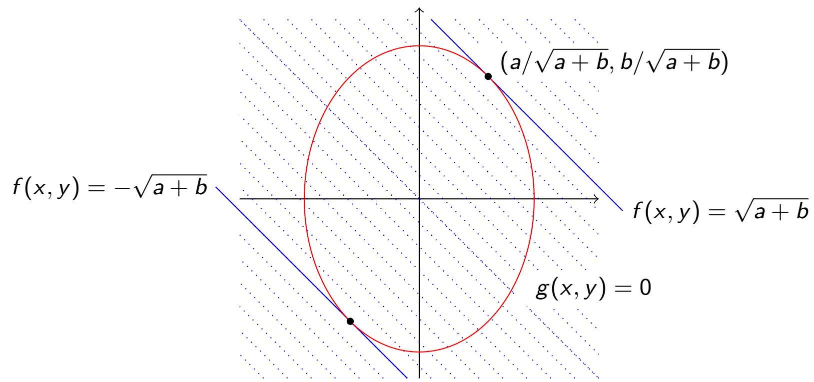
- For these points we have

$$f(x, y) = x + y = \pm(a + b)/\sqrt{a + b} = \pm\sqrt{a + b}.$$

This means that the maximum possible value of f (subject to the constraint) is $\sqrt{a + b}$, and the minimum is $-\sqrt{a + b}$.

A constrained optimisation example

Problem: maximise $f(x, y) = x + y$ subject to $x^2/a + y^2/b = 1$
Maximum and minimum values are $\pm\sqrt{a + b}$, at the points $\pm(a, b)/\sqrt{a + b}$.



Several constraints

- There is a similar method for problems with several constraints.
- For example: maximise z subject to $x^2 + y^2 + z^2 = 9$ and $x + 2y + 4z = 3$.
- Method: find unconstrained critical points of

$$L = z - \lambda(x^2 + y^2 + z^2 - 9) - \mu(x + 2y + 4z - 3)$$

- Equations:

$$\begin{aligned} L_\lambda &= 9 - x^2 - y^2 - z^2 = 0 && \text{(A)} \\ L_\mu &= 3 - x - 2y - 4z = 0 && \text{(B)} \\ L_x &= -2x\lambda - \mu = 0 && \text{(C)} \\ L_y &= -2y\lambda - 2\mu = 0 && \text{(D)} \\ L_z &= -2z\lambda - 4\mu = 0 && \text{(E)} \end{aligned}$$

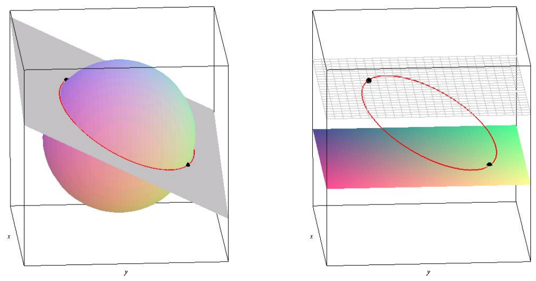
- These can be solved: use (B) to eliminate x and (C) to eliminate μ, then it works out that (D) rearranges to give $y = (6 - 8z)/5$. Substituting these into (A) gives $-21z^2/5 + 24z/5 + 36/5 = 0$, which gives $z = 2$ or $z = -6/7$. After a few more steps, we see that the solutions are

$$\begin{aligned} (\lambda, \mu, x, y, z) &= (1/12, 1/6, -1, -2, 2) \\ (\lambda, \mu, x, y, z) &= (-1/12, 3/14, 9/7, 18/7, -6/7). \end{aligned}$$

- Thus, the minimum value of z is $-6/7$ at $(9/7, 18/7, -6/7)$, and the maximum is 2 at $(-1, 2, 2)$.

Several constraints

We wanted to maximise z subject to $x^2 + y^2 + z^2 = 9$ and $x + 2y + 4z = 3$.



The equation $x^2 + y^2 + z^2 = 9$ defines a sphere (shown in colour) and $x + 2y + 4z = 3$ defines a plane (shown in grey). The two constraints together give the intersection of the sphere and the plane, which is the red curve. The optimisation problem is to find the highest and lowest points on that curve. The right-hand picture shows the planes $z = -6/7$ and $z = 2$, which we saw were the minimum and maximum values.

Integrals over plane regions

Let D be a region in the plane, and let $f(x, y)$ be a function defined for points (x, y) in D. We define the integral $\iint_D f(x, y)dA$ as follows. First, we divide the region D into a large number of small regions $D_1, \dots, D_N$. As each region D_i is small, the value of f will not change much as we move around D_i, so it makes approximate sense to talk about the value of f on D_i as a single number. The integral is approximately defined by

$$\iint_D f(x, y)\, dA = \sum_{i=1}^{N} (\text{ value of } f \text{ on } D_i\ ) \times (\text{ area of } D_i).$$

To get the exact value, we divide D into a larger and larger number of smaller and smaller pieces, and then pass to the limit.

Applications

Some applications of this kind of integration are as follows.

(a) Suppose that the region D is a charged plate, and that the charge density at a point (x, y) is $q(x, y)$; then the total charge is $Q = \iint_D q(x, y)\, dA$.

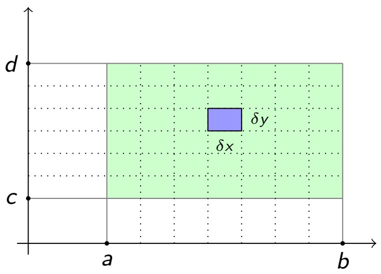
(b) Suppose that the region D represents a structure of constant density ρ and vertical thickness $f(x, y)$ attached to an axle passing vertically through the origin. Then the mass of the structure is $\iint_D \rho f(x, y)\, dA$, whereas the moment of inertia (which measures the difficulty of turning the axle) is $\iint_D \rho f(x, y)(x^2 + y^2)dA$.

(c) Suppose that the region D represents a large solar cell, with the brightness of light arriving at (x, y) being given by the function $f(x, y)$. Then the total incident power on the cell will be (a constant times) $\iint_D f(x, y)\, dA$.

(d) The total area of a region D is just $\iint_D 1\, dA$.

Rectangular regions

In the simplest case, the region D is a rectangle aligned with the axes, given by $a \leq x \leq b$ and $c \leq y \leq d$ say. In this case we can just divide the horizontal interval $[a, b]$ into small intervals of length δx, and divide the vertical interval $[c, d]$ into small intervals of length δy. This divides D into small rectangles of area $\delta A = \delta x.\delta y$.



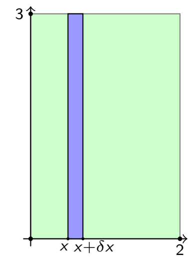
Using this kind of subdivision, we see that the area integral is just obtained by integrating with respect to both variables x and y:

$$\iint_D f(x,y)\, dA = \int_{x=a}^{b} \left(\int_{y=c}^{d} f(x,y)\, dy \right) dx.$$

Rectangular example

D = rectangle where $0 \leq x \leq 2$ and $0 \leq y \leq 3$.

$$\iint_D x^3 + y^2\, dA = \int_{x=0}^{2} \left(\int_{y=0}^{3} x^3 + y^2\, dy \right) dx$$



In the inner integral, we treat x as a constant and y as a variable. This gives

$$\int_{y=0}^{3} x^3 + y^2\, dy = \left[x^3 y + y^3/3 \right]_{y=0}^{3} = 3x^3 + 9.$$

Meaning: if we take a thin strip running horizontally from x to $x + \delta x$, and vertically all the way from 0 to 3, then the sum of the corresponding contributions is approximately $(3x^3 + 9)\delta x$ (and the approximation becomes exact in the limit as $\delta x \to 0$). Outer integral: add up the contributions from all such vertical strips.

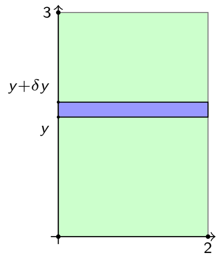
$$\int_{x=0}^{2} 3x^3 + 9\, dx = \left[3x^4/4 + 9x \right]_{x=0}^{2} = (12 + 18) - (0) = 30.$$

The conclusion is that $\iint_D x^3 + y^2\, dA = 30$.

Rectangular example — horizontal strips

D = rectangle where $0 \leq x \leq 2$ and $0 \leq y \leq 3$.

$$\iint_D x^3 + y^2\, dA = \int_{y=0}^{3} \left(\int_{x=0}^{2} x^3 + y^2\, dx \right) dy$$



In the inner integral, we treat y as a constant and x as a variable. This gives

$$\int_{x=0}^{2} x^3 + y^2\, dx = \left[x^4/4 + xy^2 \right]_{x=0}^{2} = 4 + 2y^2.$$

Meaning: if we take a thin strip running vertically from y to $y + \delta y$, and horizontally all the way from 0 to 2, then the sum of the corresponding contributions is approximately $(4 + 2y^2)\delta y$ (and the approximation becomes exact in the limit as $\delta y \to 0$). Outer integral: add up the contributions from all such horizontal strips.

$$\int_{y=0}^{3} 4 + 2y^2\, dy = \left[4y + 2y^3/3 \right]_{y=0}^{3} = (12 + 2 \times 27/3) - (0) = 30.$$

The conclusion is again that $\iint_D x^3 + y^2\, dA = 30$.

Square region example

Let E be the square where $0 \leq x \leq \pi$ and $-\pi/2 \leq y \leq \pi/2$.

$$\iint_E \sin(x)\cos(y)\, dA = \int_{x=0}^{\pi} \left(\int_{y=-\pi/2}^{\pi/2} \sin(x)\cos(y)\, dy \right) dx.$$

In the inner integral, we treat x as a constant and y as a variable. This gives

$$\int_{y=-\pi/2}^{\pi/2} \sin(x)\cos(y)\, dy = \sin(x) \left[\sin(y) \right]_{y=-\pi/2}^{\pi/2} = \sin(x)(1 - (-1)) = 2\sin(x).$$

Again, this means that the contribution coming from a vertical strip of width δx is approximately $2\sin(x)\delta x$. We can now perform the outer integral to add up the contributions from all such vertical strips:

$$\int_{x=0}^{\pi} 2\sin(x)\, dx = 2\left[-\cos(x) \right]_0^{\pi} = 2(1 - (-1)) = 4.$$

The conclusion is that $\iint_E \sin(x)\cos(y)\, dA = 4$.

Answer is constant

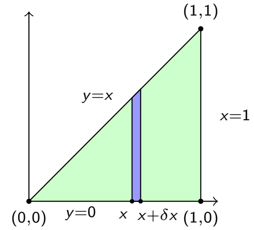
$$\iint_D x^3 + y^2 \, dA = 30$$
$$\iint_E \sin(x)\cos(y)\, dA = 4.$$

Note that in the last two examples, the final answer is just a number, not a function of x and y. We have integrated over all relevant values of x and y, so there is no remaining dependence on x and y. Some common mistakes lead to an answer that does depend on x and/or y; if you get such an answer, you need to look for the mistake.

Triangular example

D = triangle with vertices $(0,0)$, $(1,0)$ and $(1,1)$.

$$\iint_D e^{2x-2y}\, dA = \int_{x=0}^{1} \left(\int_{y=0}^{x} e^{2x-2y}\, dy \right) dx$$



Limits in the inner integral are the range of y values for a particular x. In this integral, we treat x as a constant and y as a variable. This gives

$$\int_{y=0}^{x} e^{2x-2y}\, dy = \left[e^{2x-2y}/(-2) \right]_{y=0}^{x} = (e^0 - e^{2x})/(-2)$$
$$= \tfrac{1}{2}(e^{2x} - 1).$$

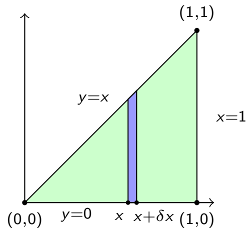
Limits for the outer integral are the full range of x values anywhere in the region, which means $0 \le x \le 1$ in this example.

$$\iint_D e^{2x-2y}\, dA = \int_{x=0}^{1} \tfrac{1}{2}(e^{2x} - 1)\, dx = \left[\tfrac{1}{2}(\tfrac{1}{2}e^{2x} - x) \right]_{x=0}^{1} = \tfrac{1}{2}(\tfrac{1}{2}e^2 - 1) - \tfrac{1}{2}(\tfrac{1}{2} - 0)$$
$$= (e^2 - 3)/4.$$

Triangular example

D = triangle with vertices $(0,0)$, $(1,0)$ and $(1,1)$.

$$\iint_D e^{2x-2y}\, dA = \int_{x=0}^{1} \left(\int_{y=0}^{x} e^{2x-2y}\, dy \right) dx$$



Limits in the inner integral are the range of y values for a particular x. In this integral, we treat x as a constant and y as a variable. This gives

$$\int_{y=0}^{x} e^{2x-2y}\, dy = \left[e^{2x-2y}/(-2) \right]_{y=0}^{x} = (e^0 - e^{2x})/(-2)$$
$$= \tfrac{1}{2}(e^{2x} - 1).$$

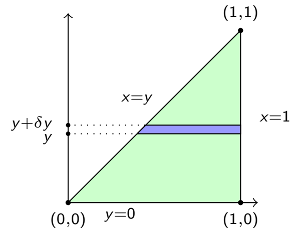
Limits for the outer integral are the full range of x values anywhere in the region, which means $0 \le x \le 1$ in this example.

$$\iint_D e^{2x-2y}\, dA = \int_{x=0}^{1} \tfrac{1}{2}(e^{2x} - 1)\, dx = \left[\tfrac{1}{2}(\tfrac{1}{2}e^{2x} - x) \right]_{x=0}^{1} = \tfrac{1}{2}(\tfrac{1}{2}e^2 - 1) - \tfrac{1}{2}(\tfrac{1}{2} - 0)$$
$$= (e^2 - 3)/4.$$

Triangular example — horizontal strips

D = triangle with vertices $(0,0)$, $(1,0)$ and $(1,1)$.

$$\iint_D e^{2x-2y}\, dA = \int_{y=0}^{1} \left(\int_{x=y}^{1} e^{2x-2y}\, dx \right) dy$$



Limits in the inner integral are the range of x values for a particular y. In this integral, we treat y as a constant and x as a variable. This gives

$$\int_{x=y}^{1} e^{2x-2y}\, dx = \left[\tfrac{1}{2} e^{2x-2y} \right]_{x=y}^{1} = \tfrac{1}{2}(e^{2-2y} - 1)$$

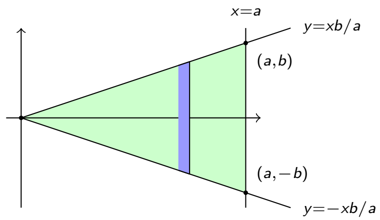
Limits for the outer integral are the full range of y values anywhere in the region, which means $0 \le y \le 1$ in this example.

$$\iint_D e^{2x-2y}\, dA = \int_{y=0}^{1} \tfrac{1}{2}(e^{2-2y} - 1)\, dy = \left[\tfrac{1}{2}(-\tfrac{1}{2}e^{2-2y} - y) \right]_{y=0}^{1}$$
$$= \tfrac{1}{2}(-\tfrac{1}{2} - 1) - \tfrac{1}{2}(-\tfrac{1}{2}e^2 - 0) = (e^2 - 3)/4.$$

Moment of inertia

For the moment of inertia, we need $\iint_D x^2 + y^2\, dA$.

If we fix x with $0 \leq x \leq a$, then y will run from $-xb/a$ to $+xb/a$.



$$\iint_D x^2{+}y^2\, dA = \int_{x=0}^{a} \int_{y=-xb/a}^{xb/a} x^2{+}y^2\, dy\, dx.$$

For the inner integral we have

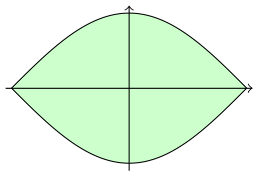
$$\int_{y=-xb/a}^{xb/a} x^2{+}y^2\, dy = \left[x^2 y + \tfrac{1}{3} y^3\right]_{y=-xb/a}^{xb/a} = \frac{2x^3 b}{a} + \frac{2x^3 b^3}{3a^3} = \left(\frac{2b}{a} + \frac{2b^3}{3a^3}\right) x^3.$$

Using this we get

$$\iint_D x^2 + y^2\, dA = \left(\frac{2b}{a} + \frac{2b^3}{3a^3}\right) \int_{x=0}^{a} x^3\, dx = \left(\frac{2b}{a} + \frac{2b^3}{3a^3}\right) \frac{a^4}{4} = \tfrac{1}{2} a^3 b + \tfrac{1}{6} a b^3.$$

Area of a curved region

Let D be the region where $-\pi/2 \leq x \leq \pi/2$ and $-\cos(x) \leq y \leq \cos(x)$.

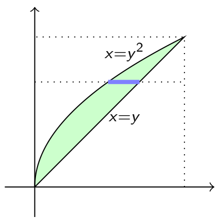


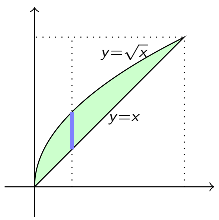
We will find the area of D, or in other words the integral $\iint_D 1\, dA$. Using vertical strips we have

$$\begin{aligned}
\iint_D 1\, dA &= \int_{x=-\pi/2}^{\pi/2} \int_{y=-\cos(x)}^{\cos(x)} 1\, dy\, dx = \int_{x=-\pi/2}^{\pi/2} [y]_{-\cos(x)}^{\cos(x)}\, dx \\
&= \int_{x=-\pi/2}^{\pi/2} 2\cos(x)\, dx = [2\sin(x)]_{x=-\pi/2}^{\pi/2} \\
&= 2 - (-2) = 4.
\end{aligned}$$

Reversing the order of integration

Consider $I = \int_{y=0}^{1} \int_{x=y^2}^{y} x^{-1} y e^x\, dx\, dy$. Inner integral not possible.





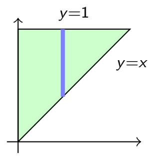
Rewrite in the opposite order:

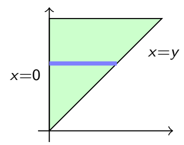
$$\begin{aligned}
I &= \int_{x=0}^{1} \int_{y=x}^{\sqrt{x}} x^{-1} y e^x\, dy\, dx = \int_{x=0}^{1} \left[\tfrac{1}{2} x^{-1} y^2 e^x\right]_{y=x}^{\sqrt{x}}\, dx \\
&= \frac{1}{2} \int_{x=0}^{1} (x^{-1} (\sqrt{x})^2 e^x - x^{-1} x^2 e^x)\, dx = \frac{1}{2} \int_{x=0}^{1} (e^x - x e^x)\, dx \\
&= \frac{1}{2} \Big[(2-x) e^x\Big]_{x=0}^{1} = (e-2)/2.
\end{aligned}$$

($u = x$ and $dv/dx = e^x$; $du/dx = 1$ and $v = e^x$;
$\int x\, e^x\, dx = x e^x - \int e^x\, dx = x e^x - e^x$.)

Reversing the order of integration

Consider $I = \int_{x=0}^{1} \int_{y=x}^{1} \frac{xy}{\sqrt{1+y^4}}\, dy\, dx$.





Rewrite in the opposite order:

$$I = \int_{y=0}^{1} \int_{x=0}^{y} \frac{xy}{\sqrt{1+y^4}}\, dx\, dy = \int_{y=0}^{1} \left[\frac{x^2 y}{2\sqrt{1+y^4}}\right]_{x=0}^{y} dy = \int_{y=0}^{1} \frac{y^3}{2\sqrt{1+y^4}}\, dy.$$

We now substitute $u = 1 + y^4$, so $du/dy = 4y^3$, so $y^3\, dy = du/4$ and $\sqrt{1+y^4} = u^{1/2}$. The limits $y = 0$ and $y = 1$ correspond to $u = 1$ and $u = 2$. This gives

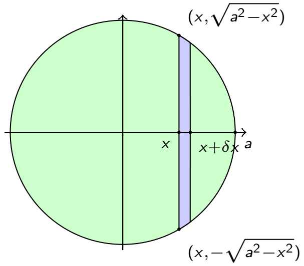
$$\begin{aligned}
I &= \int_{u=1}^{2} \frac{du/4}{2u^{1/2}} = \frac{1}{8} \int_{u=1}^{2} u^{-1/2} du = \frac{1}{8} \left[2u^{1/2}\right]_{u=1}^{2} = (2\sqrt{2} - 2)/8 \\
&= (\sqrt{2} - 1)/4 \simeq 0.1036.
\end{aligned}$$

Integral over a disk

$$I = \iint_D x^2\, dA = \int_{x=-a}^{a} \int_{y=-\sqrt{a^2-x^2}}^{+\sqrt{a^2-x^2}} x^2\, dy\, dx.$$

In the inner integral x^2 counts as a constant:

$$\int_{y=-\sqrt{a^2-x^2}}^{+\sqrt{a^2-x^2}} x^2\, dy = \left[x^2 y\right]_{y=-\sqrt{a^2-x^2}}^{+\sqrt{a^2-x^2}} = 2x^2\sqrt{a^2-x^2}.$$



Now $I = \int_{x=-a}^{a} 2x^2\sqrt{a^2-x^2}\, dx$. Substitute $x = a\sin(\theta)$, so $-a \leq x \leq a$ becomes $-\pi/2 \leq \theta \leq \pi/2$. Also $dx/d\theta = a\cos(\theta)$, so $dx = a\cos(\theta)d\theta$; and $\sqrt{a^2-x^2} = a\sqrt{1-\sin^2(\theta)} = a\cos(\theta)$. This gives

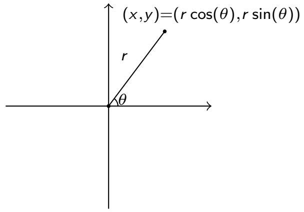
$$I = \int_{\theta=-\pi/2}^{\pi/2} 2a^2\sin^2(\theta).a\cos(\theta).a\cos(\theta)\, d\theta = 2a^4\int_{-\pi/2}^{\pi/2} \sin^2(\theta)\cos^2(\theta)\, d\theta.$$

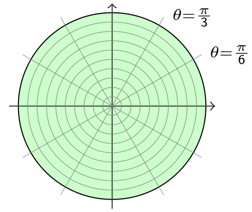
Here $\sin(\theta)\cos(\theta) = \frac{1}{2}\sin(2\theta)$; $\sin^2(\theta)\cos^2(\theta) = \frac{1}{4}\sin^2(2\theta) = (1-\cos(4\theta))/8$. This gives

$$I = \frac{a^4}{4}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} 1-\cos(4\theta)\, d\theta = \frac{a^4}{4}\left[\theta - \tfrac{1}{4}\sin(4\theta)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{a^4}{4}\left(\frac{\pi}{2} - \frac{-\pi}{2}\right) = \frac{\pi a^4}{4}.$$

Polar coordinates

We describe points using the distance r from the origin and the angle θ anticlockwise from the x-axis.





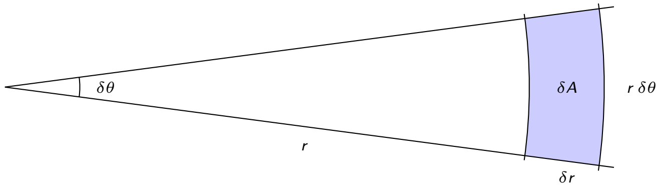
Polar coordinates are related to ordinary (cartesian) coordinates by the formulae

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$
$$r = \sqrt{x^2+y^2} \qquad \theta = \arctan(y/x).$$

(Care is needed to choose the right value of $\arctan(y/x)$.) In the diagram on the right above, we have divided a disk into small pieces using lines of constant θ and circles of constant r.To use this kind of subdivision for integration, we need to know the area of the small pieces.

The polar area element

Consider a piece of angular width $\delta\theta$, where the radius runs from r to $r+\delta r$.



Provided that $\delta\theta$ is small this will be approximately rectangular. If we measure angles in radians (as we always will) then the length of the curved side will be $r\,\delta\theta$, and the straight side has length δr, so the area is approximately $\delta A = r\,\delta r\,\delta\theta$. In the limit this becomes $dA = r\,dr\,d\theta$, so we have the following prescription: if D is a region that is conveniently described in polar coordinates, then

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

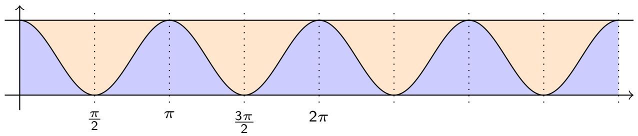
where the limits need to be filled in in accordance with the geometry of the region.

Disk integral of x^2, again

Consider again $\iint_D x^2\, dA$, where D is a disk of radius a around the origin. Here the appropriate limits are just $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq a$. The integral is

$$\begin{aligned}
\iint_D x^2\, dA &= \int_{\theta=0}^{2\pi}\int_{r=0}^{a} r^2\cos^2(\theta)\, r\, dr\, d\theta \\
&= \int_{\theta=0}^{2\pi}\cos^2(\theta)\int_{r=0}^{a} r^3\, dr\, d\theta \\
&= \frac{a^4}{4}\int_{\theta=0}^{2\pi}\cos^2(\theta)\, d\theta = \frac{a^4}{4}\int_{\theta=0}^{2\pi}\frac{1+\cos(2\theta)}{2}\, d\theta \\
&= \frac{a^4}{4}\left[\frac{1}{2}\theta + \frac{1}{4}\sin(2\theta)\right]_0^{2\pi} = \frac{\pi a^4}{4}
\end{aligned}$$

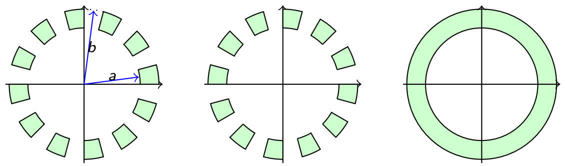
as before.
The following picture shows why $\int_0^{2\pi}\cos^2(\theta)\, d\theta = \pi$:



Each region has the same area, namely $\pi/4$.

A slotted rotor

Moment of inertia $I = \iint_D (x^2+y^2)dA$.



We first use a simplifying trick. Let D' be the region in the middle picture, and put $I' = \iint_{D'}(x^2+y^2)dA$. As D' is just obtained by turning D slightly, the moment of inertia will be the same, so $I' = I$. On the other hand, $2I = I + I'$ is just the integral over the simpler region D'' shown on the right. We thus have $I = \frac{1}{2}\iint_{D''}(x^2+y^2)dA$. For D'' the limits are just $0 \leq \theta \leq 2\pi$ and $a \leq r \leq b$. The integrand is

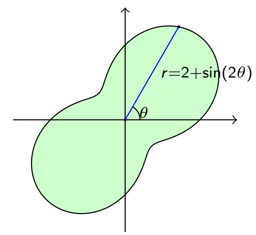
$$x^2 + y^2 = (r\cos(\theta))^2 + (r\sin(\theta))^2 = r^2,$$

and the area element is $dA = r\,dr\,d\theta$. We thus have

$$\begin{aligned} I &= \tfrac{1}{2}\int_{\theta=0}^{2\pi}\int_{r=a}^{b} r^3\,dr\,d\theta = \tfrac{1}{2}\int_{\theta=0}^{2\pi}\frac{b^4-a^4}{4}d\theta \\ &= \frac{1}{2}\,\frac{b^4-a^4}{4}\,2\pi = \pi(b^4-a^4)/4. \end{aligned}$$

Area of a curved region

The picture shows the region D given in polar coordinates by $0 \leq r \leq 2+\sin(2\theta)$. We would like to find the area of D, or in other words $A = \iint_D 1\,dA$. Here $dA = r\,dr\,d\theta$ as usual, and the relevant limits are $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 2+\sin(\theta)$



$$\begin{aligned} A &= \int_{\theta=0}^{2\pi}\int_{r=0}^{2+\sin(2\theta)} r\,dr\,d\theta = \int_{\theta=0}^{2\pi}\left[\frac{r^2}{2}\right]_{r=0}^{2+\sin(2\theta)} d\theta \\ &= \frac{1}{2}\int_{\theta=0}^{2\pi}(2+\sin(2\theta))^2\,d\theta = \frac{1}{2}\int_{\theta=0}^{2\pi} 4 + 4\sin(2\theta) + \sin^2(2\theta)\,d\theta \\ &= \frac{1}{2}\int_{\theta=0}^{2\pi} 4 + 4\sin(2\theta) + \tfrac{1}{2} - \tfrac{1}{2}\cos(2\theta)\,d\theta. \end{aligned}$$

The integral of $\sin(k\theta)$ or $\cos(k\theta)$ over a whole number of complete cycles is zero. Thus, only the terms 4 and $\frac{1}{2}$ contribute to the integral, and we have

$$A = \tfrac{1}{2}(2\pi.(4+\tfrac{1}{2})) = 9\pi/2.$$

The Gaussian integral

It is an important fact (for the theory of the normal distribution in statistics, the analysis of heat flow, the pricing of financial derivatives, and other applications) that $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$. We will explain one way to calculate this. Put $I = \int_{-\infty}^{\infty} e^{-x^2}\,dx$. It obviously does not matter what we call the variable, so we also have $I = \int_{-\infty}^{\infty} e^{-y^2}\,dy$. We can now multiply these two expressions together to get

$$I^2 = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} e^{-x^2-y^2}\,dx\,dy = \iint_{\text{whole plane}} e^{-x^2-y^2}\,dA.$$

We can rewrite this using polar coordinates, noting that $x^2+y^2 = r^2$ and $dA = r\,dr\,d\theta$. We get

$$I^2 = \int_{r=0}^{\infty}\int_{\theta=0}^{2\pi} r\,e^{-r^2}\,d\theta\,dr = 2\pi\int_{r=0}^{\infty} r\,e^{-r^2}\,dr.$$

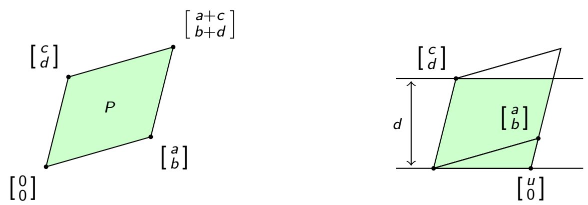
We now substitute $u = r^2$, so u also runs from 0 to ∞ and $du = 2r\,dr$. The integral becomes

$$I^2 = 2\pi\int_{u=0}^{\infty} e^{-u}.\tfrac{1}{2}du = \pi\Big[-e^{-u}\Big]_{u=0}^{\infty} = \pi((-0)-(-1)) = \pi$$

so $I = \sqrt{\pi}$ as claimed.

Area of a parallelogram

The area of the parallelogram P is $|ad-bc| = \left|\det\begin{bmatrix} a & c \\ b & d\end{bmatrix}\right|$.



Indeed, P consists of the top triangle (shown in yellow) together with the middle region (shown in green). The top triangle has the same area as the bottom one, so we may as well consider the parallelogram P' consisting of the bottom triangle together with the middle region. This parallelogram has a base of length u and a perpendicular height of d, so the area is ud. Note that $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]$ is reached from $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$ by moving in the opposite direction to the vector $\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$, so $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] - t\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$ for some t. By comparing the y-coordinates we see that $t = b/d$, and by looking at the x-coordinates we deduce that $u = a - bc/d$, so the area is $ud = ad - bc$. This works when $\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$ is anticlockwise from $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$. If $\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$ is clockwise from $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$ it works out instead that $ad - bc < 0$ and the area is $-(ad-bc)$. In all cases: the area is $|ad-bc|$.

Change of variables in a double integral

Suppose x and y can be expressed in terms of some other variables u and v. The *Jacobian matrix*:

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}.$$

For small changes δu and δv to u and v, resulting changes in x and y are

$$\delta x \simeq x_u\,\delta u + x_v\,\delta v \qquad\qquad \delta y \simeq y_u\,\delta u + y_v\,\delta v.$$

These equations can be combined as a single matrix equation:

$$\begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \begin{bmatrix} \delta u \\ \delta v \end{bmatrix} = \frac{\partial(x,y)}{\partial(u,v)} \begin{bmatrix} \delta u \\ \delta v \end{bmatrix}.$$

Now let the change in u vary between 0 and δu, and let the change in v vary between 0 and δv. The resulting changes in $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ then cover a small parallelogram spanned by $\begin{bmatrix} x_u \\ y_u \end{bmatrix} \delta u$ and $\begin{bmatrix} x_v \\ y_v \end{bmatrix} \delta v$, and the area of this parallelogram is $|x_u y_v - x_v y_u|\delta u\,\delta v$, or in other words $\left|\det\left(\frac{\partial(x,y)}{\partial(u,v)}\right)\right| \delta u\,\delta v$.

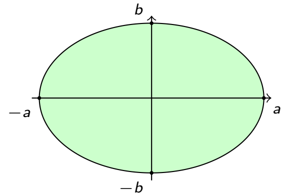
Using this: $\qquad dA = \left|\det\left(\dfrac{\partial(x,y)}{\partial(u,v)}\right)\right| du\,dv.$

This is a key ingredient for double integrals by substitution.

Area of an ellipse

We will find the area of an ellipse E with equation $x^2/a^2 + y^2/b^2 \leq 1$ (for some $a, b > 0$). For this it is best to use a kind of distorted polar coordinates:

$$x = ar\cos(\theta) \qquad\qquad y = br\sin(\theta).$$



Then $x^2/a^2 + y^2/b^2 = r^2\cos^2(\theta) + r^2\sin^2(\theta) = r^2$, so $x^2/a^2 + y^2/b^2 \leq 1$ becomes $0 \leq r \leq 1$. Partial derivatives:

$$x_r = a\cos(\theta) \qquad x_\theta = -ar\sin(\theta) \qquad y_r = b\sin(\theta) \qquad y_\theta = br\cos(\theta),$$

$$J = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{bmatrix} a\cos(\theta) & -ar\sin(\theta) \\ b\sin(\theta) & br\cos(\theta) \end{bmatrix}.$$

This means that the absolute value of the determinant is

$$|\det(J)| = |abr\cos^2(\theta) - (-abr\sin^2(\theta))| = |abr| = abr,$$

so $dA = abr\,dr\,d\theta$. We therefore have

$$\text{area} = \iint_E 1\,dA = \int_{\theta=0}^{2\pi}\int_{r=0}^{1} abr\,dr\,d\theta = ab\int_{\theta=0}^{2\pi}\left[\frac{r^2}{2}\right]_{r=0}^{1} d\theta = ab\int_{\theta=0}^{2\pi}\frac{1}{2}d\theta = \pi ab.$$

Integration over three-dimensional regions

Suppose we have a solid region E in 3-dimensional space, and a function $f(x,y,z)$. We can define the volume integral of f (written $\iiint_E f(x,y,z)\,dV$) in a very similar way to the area integrals that we discussed before: we divide E into a large number of small regions $E_1, \dots, E_N$, and then

$$\iiint_E f(x,y,z)\,dV \simeq \sum_{i=1}^{N} (\text{ value of } f \text{ on } E_i\ ) \times (\text{ volume of } E_i),$$

with the approximation becoming exact in the limit where the size of the subregions tends to zero. As in the plane case, such integrals can usually be evaluated by integrating over three different variables with suitable limits depending on the geometry of the region E.

(a) Total energy of a magnetic field: integrate (field strength)^2.

(b) Moment of inertia of a rotor: integrate (distance from the axis)^2.

(c) Total mass of a star: integrate the density.

(d) Centre of mass $(\overline{x}, \overline{y}, \overline{z})$ of an object: $\overline{x} = \left(\iiint_E x\,dV\right) / \left(\iiint_E 1\,dV\right)$ and similarly for $\overline{y}$ and $\overline{z}$.

Microwave oven

Let E be the inside of a microwave of length a, width b and height c (integers). For the total energy of the microwaves:

$$I = \iiint_E (\sin(k\pi x)\sin(m\pi y)\sin(n\pi z))^2\,dV,$$

where k, n and m are also integers. This just reduces to

$$I = \int_{x=0}^{a}\int_{y=0}^{b}\int_{z=0}^{c} \sin^2(k\pi x)\sin^2(m\pi y)\sin^2(n\pi z)\,dz\,dy\,dx$$

For the innermost integral, we have $\sin^2(n\pi z) = (1 - \cos(2n\pi z))/2$, so

$$\int_{z=0}^{c} \sin^2(n\pi z)\,dz = \left[\frac{z}{2} - \frac{\sin(2n\pi z)}{4n\pi}\right]_{z=0}^{c}$$

As n and c are integers, the sin() term is zero at both endpoints, and we just get $\int_{z=0}^{c}\sin^2(n\pi z)\,dz = c/2$. The terms $\sin^2(x)$ and $\sin^2(y)$ are just carried along as constants, so we get

$$I = \int_{x=0}^{a}\int_{y=0}^{b} \sin^2(k\pi x)\sin^2(m\pi y)\frac{c}{2}\,dy\,dx.$$

We can integrate over y and then over x in the same way, giving

$$I = \int_{x=0}^{a} \sin^2(k\pi x)\frac{b}{2}\frac{c}{2}\,dx. = \frac{a}{2}.\frac{b}{2}.\frac{c}{2} = \frac{abc}{8}.$$

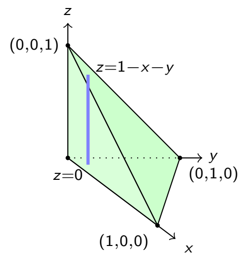
Moment of inertia of a cube

Let E be the cube given by $-1 \leq x, y, z \leq 1$. The moment of inertia about the z-axis is

$$
\begin{aligned}
\iiint_E (x^2+y^2)dV &= \int_{x=-1}^{1}\int_{y=-1}^{1}\int_{z=-1}^{1}(x^2+y^2)dz\,dy\,dx \\
&= \int_{x=-1}^{1}\int_{y=-1}^{1}\left[x^2z+y^2z\right]_{z=-1}^{1}dy\,dx \\
&= \int_{x=-1}^{1}\int_{y=-1}^{1}2(x^2+y^2)dy\,dx \\
&= \int_{x=-1}^{1}\left[2x^2y+2y^3/3\right]_{y=-1}^{1}dx \\
&= \int_{x=-1}^{1}4x^2+4/3\,dx \\
&= \left[4x^3/3+4x/3\right]_{x=-1}^{1} = 8/3+8/3=16/3.
\end{aligned}
$$

Integral over a tetrahedron

Let E be the tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$.



The shadow in the (x,y)-plane is the triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$, which means that x varies from 0 to 1, and y varies from 0 to $1-x$. Each of the points $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ satisfies $x+y+z=1$, which means that the equation of the top face is $x+y+z=1$, or in other words $z=1-x-y$. The equation of the bottom face is $z=0$, so overall z varies from 0 to $1-x-y$. Thus, for any function $f(x,y,z)$ we have

$$
\iiint_E f(x,y,z)\,dV = \int_{x=0}^{1}\int_{y=0}^{1-x}\int_{z=0}^{1-x-y}f(x,y,z)\,dz\,dy\,dx.
$$

Volume of a tetrahedron

For a tetrahedron E with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$:

$$
\iiint_E f(x,y,z)\,dV = \int_{x=0}^{1}\int_{y=0}^{1-x}\int_{z=0}^{1-x-y}f(x,y,z)\,dz\,dy\,dx.
$$

For the volume of E: take $f(x,y,z)=1$. The innermost integral is then

$$
\int_{z=0}^{1-x-y}1\,dz = 1-x-y.
$$

Thus, the integral with respect to y is

$$
\begin{aligned}
\int_{y=0}^{1-x}(1-x-y)\,dy &= \left[(1-x)y-y^2/2\right]_{y=0}^{1-x} = ((1-x)(1-x)-(1-x)^2/2)-0 \\
&= 1/2-x+x^2/2.
\end{aligned}
$$

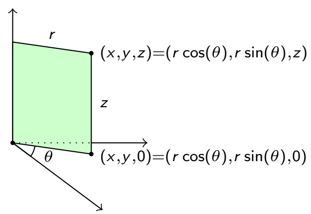
Finally, the outermost integral (with respect to x) is

$$
\int_{x=0}^{1}(1/2-x+x^2/2)\,dx = \left[x/2-x^2/2+x^3/6\right]_{x=0}^{1} = 1/2-1/2+1/6=1/6.
$$

We conclude that the volume of the tetrahedron is $1/6$.

Cylindrical polar coordinates

When using cylindrical polar coordinates we describe points in terms of the distance r from the z-axis, the angle θ anticlockwise from the (x,z)-plane, and the height z above the (x,y)-plane.



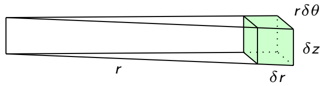
Just as in the two-dimensional case, r and θ are related to x and y by the equations

$$
\begin{aligned}
x &= r\cos(\theta) & y &= r\sin(\theta) \\
r &= \sqrt{x^2+y^2} & \theta &= \arctan(y/x).
\end{aligned}
$$

Applications: rotating machines, fibre-optic cables, dish-shaped antennas.

Cylindrical polar volume element

If we allow r, θ and z to vary by small amounts δr, $\delta\theta$ and δz, then the corresponding region is approximately a right-angled box with sides of length δr, δz and $r\delta\theta$. The volume is thus $\delta V \simeq r\delta r\,\delta\theta\,\delta z$.



This means that for a function f on a 3-dimensional region E, we have

$$\iiint_E f(x,y,z)\,dV = \int_{z=\cdots}^{\cdots}\int_{\theta=\cdots}^{\cdots}\int_{r=\cdots}^{\cdots} f(r\cos(\theta), r\sin(\theta), z)\,{\color{red}r}\,dr\,d\theta\,dz,$$

where the limits must be determined using the geometry of the region.

Jacobian method

$$\iiint_E f(x,y,z)\,dV = \int_{z=\cdots}^{\cdots}\int_{\theta=\cdots}^{\cdots}\int_{r=\cdots}^{\cdots} f(r\cos(\theta), r\sin(\theta), z)\,{\color{red}r}\,dr\,d\theta\,dz,$$

The formula for dV can also be obtained using a three-dimensional of the Jacobian matrix. We have

$$J = \frac{\partial(x,y,z)}{\partial(r,\theta,z)} = \begin{bmatrix} x_r & x_\theta & x_z \\ y_r & y_\theta & y_z \\ z_r & z_\theta & z_z \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -r\sin(\theta) & 0 \\ \sin(\theta) & r\cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{aligned}\det(J) &= \cos(\theta)\ \det\begin{bmatrix} r\cos(\theta) & 0 \\ 0 & 1\end{bmatrix} - (-r\sin(\theta))\ \det\begin{bmatrix}\sin(\theta) & 0 \\ 0 & 1\end{bmatrix} + 0\ \det\begin{bmatrix}\sin(\theta) & r\cos(\theta) \\ 0 & 0\end{bmatrix} \\ &= r\cos^2(\theta) + r\sin^2(\theta) = r.\end{aligned}$$

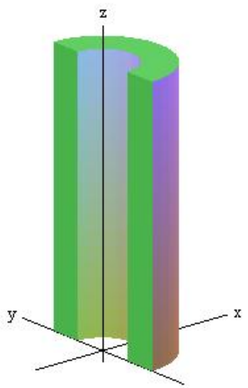
(Tidier approach: expand along the bottom row.)
Now $|\det(J)| = |r| = r$ as $r \geq 0$. We conclude that

$$dV = dx\,dy\,dz = \left|\det\left(\frac{\partial(x,y,z)}{\partial(r,\theta,z)}\right)\right| dr\,d\theta\,dz = |\det(J)|\,dr\,d\theta\,dz = r\,dr\,d\theta\,dz,$$

just as we saw before by a more geometric argument.

Centre of mass of a half-pipe

Consider a region E as shown on the right.
(Inner radius 1, outer radius 2, height 8.)
$0 \leq z \leq 8; \qquad -\pi/2 \leq \theta \leq \pi/2; \qquad 1 \leq r \leq 2.$
Centre of mass $(\overline{x}, 0, 4)$, where $\overline{x} = X/V$,
$X = \iiint_E x\,dV$ and $V = \iiint_E 1\,dV$.



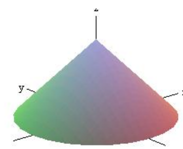
$$\begin{aligned} V &= \int_{z=0}^{8}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_{r=1}^{2} r\,dr\,d\theta\,dz \\ &= \int_{z=0}^{8}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{3}{2}\,d\theta\,dz = \int_{z=0}^{8}\frac{3\pi}{2}\,dz = 12\pi \\ X &= \int_{z=0}^{8}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_{r=1}^{2} r\cos(\theta).r\,dr\,d\theta\,dz \\ &= \int_{z=0}^{8}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}}\left[\frac{1}{3}r^3\cos(\theta)\right]_{r=1}^{2} d\theta\,dz \end{aligned}$$

$$= \frac{7}{3}\int_{z=0}^{8}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos(\theta)\,d\theta\,dz = \frac{7}{3}\int_{z=0}^{8}\Big[\sin(\theta)\Big]_{-\pi/2}^{\pi/2} dz = \frac{14}{3}\int_{z=0}^{8} 1\,dz = \frac{112}{3}$$

$$\overline{x} = \frac{112}{3\times 12\pi} = \frac{28}{9\pi} \simeq 0.99.$$

Centre of mass of a cone

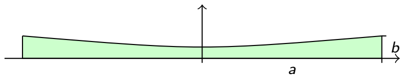
Consider a cone E as on the right (base radius=height=1).
Radius at height z is $1-z$.
$0 \leq z \leq 1; \qquad 0 \leq \theta \leq 2\pi; \qquad 0 \leq r \leq 1-z.$
Centre of mass $(0,0,\overline{z})$, where $\overline{z} = Z/V$,
$Z = \iiint_E z\,dV$ and $V = \iiint_E 1\,dV$.



$$\begin{aligned} V &= \int_{z=0}^{1}\int_{\theta=0}^{2\pi}\int_{r=0}^{1-z} r\,dr\,d\theta\,dz = \int_{z=0}^{1}\int_{\theta=0}^{2\pi}\frac{(1-z)^2}{2}\,d\theta\,dz \\ &= \int_{z=0}^{1}\pi(1-z)^2\,dz = \pi\int_{z=0}^{1} 1-2z+z^2\,dz = \pi\left[z - z^2 + \tfrac{1}{3}z^3\right]_{z=0}^{1} = \pi/3 \\ Z &= \int_{z=0}^{1}\int_{\theta=0}^{2\pi}\int_{r=0}^{1-z} zr\,dr\,d\theta\,dz = \int_{z=0}^{1}\int_{\theta=0}^{2\pi} z\frac{(1-z)^2}{2}\,d\theta\,dz \\ &= \int_{z=0}^{1}\pi z(1-z)^2\,dz = \pi\int_{z=0}^{1} z - 2z^2 + z^3\,dz \\ &= \pi\left[\tfrac{1}{2}z^2 - \tfrac{2}{3}z^3 + \tfrac{1}{4}z^4\right]_{z=0}^{1} = \pi/12 \\ \overline{z} &= Z/V = \frac{\pi}{12}\Big/\frac{\pi}{3} = \frac{\pi}{12}\Big/\frac{4\pi}{12} = \frac{1}{4}. \end{aligned}$$

Mass of a parabolic mirror

Telescope mirrors always have a parabolic cross-section. We could make such a mirror by starting with a large flat cylinder of radius a and thickness b, and grinding the top until it fits the surface $z = b(r^2 + a^2)/(2a^2)$.

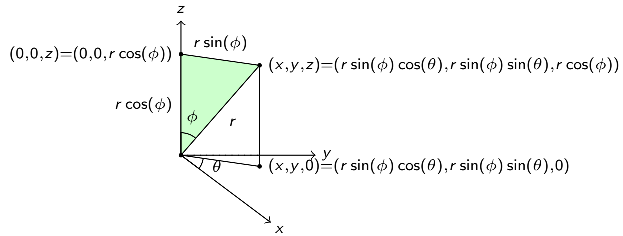


Write E for the region filled by the remaining material. It is easiest to integrate over E using vertical strips. Inner integral over z, outer integrals over r and θ. If the density of the material is ρ, the total mass of the mirror will be

$$\begin{aligned} M &= \iiint_E \rho\, dV = \int_{r=0}^{a} \int_{\theta=0}^{2\pi} \int_{z=0}^{\frac{b(r^2+a^2)}{2a^2}} \rho r\, dz\, d\theta\, dr \\ &= \int_{r=0}^{a} \int_{\theta=0}^{2\pi} \frac{\rho r b(r^2+a^2)}{2a^2}\, d\theta\, dr = \frac{b\rho}{2a^2} \int_{r=0}^{a} \int_{\theta=0}^{2\pi} r^3 + a^2 r d\theta\, dr \\ &= \frac{b\rho\pi}{a^2} \int_{r=0}^{a} r^3 + a^2 r\, dr = \frac{b\rho\pi}{a^2} \left[\frac{r^4}{4} + \frac{a^2 r^2}{2} \right]_{r=0}^{a} \\ &= \frac{b\rho\pi}{a^2} \left(\frac{a^4}{4} + \frac{a^4}{2} \right) = \frac{3a^2 b\rho\pi}{4}. \end{aligned}$$

Spherical polar coordinates

In spherical polar coordinates we describe a point (x, y, z) by giving the distance r from the origin, the angle θ anticlockwise from the xz plane, and the angle ϕ from the z-axis.



The variables r, θ and ϕ are related to x and y by the equations

$$\begin{aligned} &x = r\sin(\phi)\cos(\theta) && y = r\sin(\phi)\sin(\theta) && z = r\cos(\phi) \\ &r = \sqrt{x^2+y^2+z^2} && \theta = \arctan(y/x) && \phi = \arctan(\sqrt{x^2+y^2}/z). \end{aligned}$$

Note that ϕ ranges from 0 (on the positive z-axis) to π (on the negative z-axis), whereas θ ranges from 0 to 2π (or equivalently, from $-\pi$ to π). It is also useful to observe that $\sqrt{x^2+y^2} = r\sin(\phi)$.

Spherical polar volume element

For these coordinates it is easiest to find the area element using the Jacobian. We have $x = r\sin(\phi)\cos(\theta)$, $y = r\sin(\phi)\sin(\theta)$, $z = r\cos(\phi)$ so

$$J = \frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{bmatrix} x_r & x_\theta & x_\phi \\ y_r & y_\theta & y_\phi \\ z_r & z_\theta & z_\phi \end{bmatrix} = \begin{bmatrix} \sin(\phi)\cos(\theta) & -r\sin(\phi)\sin(\theta) & r\cos(\phi)\cos(\theta) \\ \sin(\phi)\sin(\theta) & r\sin(\phi)\cos(\theta) & r\cos(\phi)\sin(\theta) \\ \cos(\phi) & 0 & -r\sin(\phi) \end{bmatrix}.$$

We will expand the determinant along the bottom row. This gives

$$\det(J) = \cos(\phi)\det(A) - 0\det(B) + (-r\sin(\phi))\det(C),$$

where

$$A = \begin{bmatrix} -r\sin(\phi)\sin(\theta) & r\cos(\phi)\cos(\theta) \\ r\sin(\phi)\cos(\theta) & r\cos(\phi)\sin(\theta) \end{bmatrix} \quad B = \begin{bmatrix} \sin(\phi)\cos(\theta) & r\cos(\phi)\cos(\theta) \\ \sin(\phi)\sin(\theta) & r\cos(\phi)\sin(\theta) \end{bmatrix} \quad C = \begin{bmatrix} \sin(\phi)\cos(\theta) & -r\sin(\phi)\sin(\theta) \\ \sin(\phi)\sin(\theta) & r\sin(\phi)\cos(\theta) \end{bmatrix}.$$

$$\det(A) = -r^2\sin(\phi)\cos(\phi)\sin^2(\theta) - r^2\sin(\phi)\cos(\phi)\cos^2(\theta) = -r^2\sin(\phi)\cos(\phi)$$

$$\det(C) = r\sin^2(\phi)\cos^2(\theta) - (-r\sin^2(\phi)\sin^2(\theta)) = r\sin^2(\phi)$$

$$\begin{aligned} \det(J) &= \cos(\phi)\det(A) - 0\det(B) + (-r\sin(\phi))\det(C) \\ &= -r^2\sin(\phi)\cos^2(\phi) - r^2\sin(\phi)\sin^2(\phi) = -r^2\sin(\phi). \end{aligned}$$

As $0 \leq \phi \leq \pi$ we have $\sin(\phi) \geq 0$ so $|-r^2\sin(\phi)| = r^2\sin(\phi)$. We conclude that

$$dV = |\det(J)|dr\, d\theta\, d\phi = r^2\sin(\phi)\, dr\, d\theta\, d\phi.$$

Spherical polar volume element

$$x = r\sin(\phi)\cos(\theta) \qquad y = r\sin(\phi)\sin(\theta) \qquad z = r\cos(\phi)$$

$$dV = |\det(J)|dr\, d\theta\, d\phi = r^2\sin(\phi)\, dr\, d\theta\, d\phi.$$

This means that for a function f on a 3-dimensional region E, we have

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

$$\int_{\phi=\cdots}^{\cdots} \int_{\theta=\cdots}^{\cdots} \int_{r=\cdots}^{\cdots} f(r\cos(\theta)\sin(\phi), r\sin(\theta)\sin(\phi), r\cos(\phi))\, r^2\sin(\phi)\, dr\, d\theta\, d\phi,$$

where the limits must be determined using the geometry of the region.

Volume and moment of a sphere

The volume of a sphere E of radius a is

$$V = \iiint_E 1\,dV = \int_{\phi=0}^{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{a} r^2\sin(\phi)dr\,d\theta\,d\phi$$
$$= \int_{\phi=0}^{\pi}\int_{\theta=0}^{2\pi}\frac{a^3}{3}\sin(\phi)\,d\theta\,d\phi = \int_{\phi=0}^{\pi}\frac{2\pi a^3}{3}\sin(\phi)\,d\phi$$
$$= \frac{2\pi a^3}{3}\Big[-\cos(\phi)\Big]_{\phi=0}^{\pi} = \frac{2\pi a^3}{3}(1-(-1)) = \frac{4\pi a^3}{3}.$$

Now suppose that the sphere has density ρ. The distance of a point from the z-axis is $r\sin(\phi)$, so the moment of inertia around that axis is

$$I = \iiint_E \rho.(r\sin(\phi))^2\,dV = \int_{\phi=0}^{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{a}\rho r^2\sin^2(\phi)r^2\sin(\phi)dr\,d\theta\,d\phi$$
$$= \int_{\phi=0}^{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{a}\rho r^4\sin^3(\phi)dr\,d\theta\,d\phi.$$

Here the three different variables do not interact in any interesting way so we can rewrite the integral as

$$I = \left(\int_{\phi=0}^{\pi}\sin(\phi)^3\,d\phi\right)\left(\int_{\theta=0}^{2\pi}1\,d\theta\right)\left(\int_{r=0}^{a}r^4\,dr\right)\rho.$$

Volume and moment of a sphere

$$I = \left(\int_{\phi=0}^{\pi}\sin(\phi)^3\,d\phi\right)\left(\int_{\theta=0}^{2\pi}1\,d\theta\right)\left(\int_{r=0}^{a}r^4\,dr\right)\rho.$$

Two of these integrals are easy: we have $\int_{\theta=0}^{2\pi}1\,d\theta = 2\pi$ and $\int_{r=0}^{a}r^4\,dr = a^5/5$. For the integral with respect to ϕ, we recall that $\sin(\phi) = (e^{j\phi}-e^{-j\phi})/(2j)$. We can cube this to get

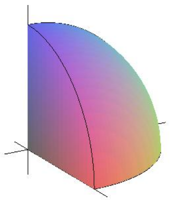
$$\sin^3(\phi) = \frac{1}{8j^3}(e^{3j\phi} - 3e^{2j\phi}e^{-j\phi} + 3e^{j\phi}e^{-2j\phi} - e^{-3j\phi})$$
$$= \frac{-1}{8j}(e^{3j\phi} - 3e^{j\phi} + 3e^{-j\phi} - e^{-3j\phi})$$
$$= \frac{3}{4}\left(\frac{e^{j\phi}-e^{-j\phi}}{2j}\right) - \frac{1}{4}\left(\frac{e^{3j\phi}-e^{-3j\phi}}{2j}\right) = \frac{3}{4}\sin(\phi) - \frac{1}{4}\sin(3\phi).$$

$$\int_{\phi=0}^{\pi}\sin^3(\phi)\,d\phi = \Big[-\frac{3}{4}\cos(\phi) + \frac{1}{12}\cos(3\phi)\Big]_{\phi=0}^{\pi}$$
$$= (3/4 - 1/12) - (-3/4 + 1/12) = 4/3.$$

Combining this with the r and θ integrals gives $I = \frac{4}{3}.2\pi.\frac{a^5}{5}.\rho = \frac{8\pi a^5\rho}{15}$.

Mass centre of an octant

Let E be the part of a sphere of radius 1 where $x \geq 0$, $y \geq 0$ and $z \geq 0$.



The centre of mass of E (assuming constant density) is $(\overline{x},\overline{y},\overline{z})$, where $\overline{x} = (\iiint_E x\,dV)/(\iiint_E 1\,dV)$ and so on. It is clear by symmetry that $\overline{x}$, $\overline{y}$ and $\overline{z}$ are all the same, so we will just calculate $\overline{z} = Z/V$. Here $V = \frac{1}{8}(\text{vol of sphere}) = \frac{1}{8}\frac{4}{3}\pi = \frac{\pi}{6}$. Next $Z = \iiint_E z\,dV$. The restriction $z \geq 0$ means that $0 \leq \phi \leq \pi/2$, and the restrictions $x, y \geq 0$ mean that $0 \leq \theta \leq \pi/2$. Recall also that $z = r\cos(\phi)$ and $dV = r^2\sin(\phi)\,dr\,d\theta\,d\phi$, so

$$z\,dV = r^3\sin(\phi)\cos(\phi)\,dr\,d\theta\,d\phi = \tfrac{1}{2}r^3\sin(2\phi)\,dr\,d\theta\,d\phi.$$

This gives

$$Z = \int_{\phi=0}^{\frac{\pi}{2}}\int_{\theta=0}^{\frac{\pi}{2}}\int_{r=0}^{1}\tfrac{1}{2}r^3\sin(2\phi)\,dr\,d\theta\,d\phi$$

Mass centre of an octant

$$Z = \int_{\phi=0}^{\frac{\pi}{2}}\int_{\theta=0}^{\frac{\pi}{2}}\int_{r=0}^{1}\tfrac{1}{2}r^3\sin(2\phi)\,dr\,d\theta\,d\phi$$
$$= \frac{1}{2}\left(\int_{\phi=0}^{\frac{\pi}{2}}\sin(2\phi)\,d\phi\right)\left(\int_{\theta=0}^{\frac{\pi}{2}}1\,d\theta\right)\left(\int_{r=0}^{1}r^3\,dr\right)$$
$$= \frac{1}{2}\Big[-\frac{\cos(2\phi)}{2}\Big]_{\phi=0}^{\frac{\pi}{2}}.\frac{\pi}{2}.\frac{1}{4}$$
$$= \frac{1}{2}.\frac{2}{2}.\frac{\pi}{2}.\frac{1}{4} = \frac{\pi}{16}$$

so

$$\overline{z} = Z/V = \frac{\pi}{16}/\frac{\pi}{6} = \frac{6}{16} = \frac{3}{8}.$$

We conclude that the centre of mass is $(\frac{3}{8},\frac{3}{8},\frac{3}{8})$.

Basic definitions

Recall that a *vector* is a quantity with both magnitude and direction. Examples include:

(a) The velocity and acceleration of a particle are vectors.

(b) The separation between two particles is a vector.

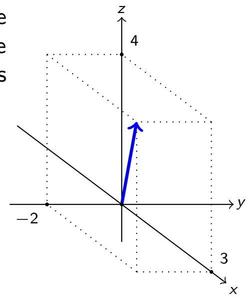
(c) If we have chosen a point to count as the origin, then the displacement of a particle from that origin is also a vector.

(d) The electric field at a point is a vector, and the magnetic field is another vector.

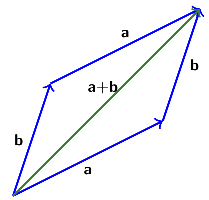
By contrast, a *scalar* is a quantity that has a magnitude, but not a direction. For example, the pressure, temperature and electric potential at a point are scalars.
When answering questions in vector algebra or vector calculus, you should always ask yourself whether your answer should be a scalar or a vector, and make sure that what you have written has the right type. This simple check will detect a substantial fraction of incorrect answers.

Numerical vectors

Normally we will fix a coordinate system, and use it to represent vectors as triples of numbers. For example, the triple $(3, -2, 4)$ represents the vector that goes 3 steps along the x-axis, 2 steps backwards parallel to the y-axis, and 4 steps parallel to the z-axis.



We can add vectors in an obvious way, for example $(3, -2, 4) + (1, 1, 1) = (4, -1, 5)$. Geometrically, this corresponds to joining the vectors together nose to tail.



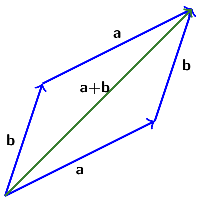
Similarly, we can multiply a vector by a scalar to get a new vector, for example $3(3, -2, 4) = (9, -6, 12)$. The new vector has the same direction as the old one (if the scalar is positive) or the opposite direction (if the scalar is negative).

Length of vectors

The length of a vector $\mathbf{a} = (x, y, z)$ is given by

$$|\mathbf{a}| = \sqrt{x^2 + y^2 + z^2}.$$

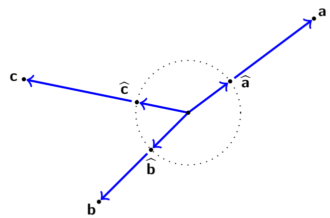
It is a useful fact that we always have $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$; this is called the *triangle inequality*. To see why it is true, consider the parallelogram below.



The distance from the origin to $\mathbf{a} + \mathbf{b}$ in a straight line is $|\mathbf{a} + \mathbf{b}|$, whereas the distance via $\mathbf{a}$ is $|\mathbf{a}| + |\mathbf{b}|$. The inequality just says that it is shorter to go in a straight line.

Unit vectors

A *unit vector* is a vector of length one. We write $\widehat{\mathbf{a}}$ for the unit vector in the same direction as $\mathbf{a}$.



This is given by

$$\widehat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} = \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right).$$

For example, if $\mathbf{a} = (1, -2, 2)$ then

$$|\mathbf{a}| = \sqrt{1^2 + (-2)^2 + 2^2} = \sqrt{1 + 4 + 4} = 3$$
$$\widehat{\mathbf{a}} = \frac{\mathbf{a}}{3} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3} \right).$$

Note that $|\mathbf{a}|$ is a scalar, and $\widehat{\mathbf{a}}$ is a vector.

Vectors along the coordinate axes

The unit vectors along the three coordinate axes are denoted by **i**, **j** and **k**:

$$\begin{aligned}\mathbf{i} &= (1,0,0)\\ \mathbf{j} &= (0,1,0)\\ \mathbf{k} &= (0,0,1).\end{aligned}$$

Note that

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (x,0,0) + (0,y,0) + (0,0,z) = (x,y,z).$$

For example, the vector $(10,0,-20)$ can also be expressed as $10\mathbf{i} - 20\mathbf{k}$.

Dot products

The dot product of vectors $\mathbf{a} = (x,y,z)$ and $\mathbf{b} = (u,v,w)$ is given by

$$\mathbf{a}.\mathbf{b} = (x,y,z).(u,v,w) = xu + yv + zw.$$

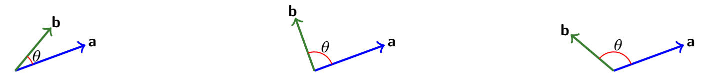
Note that this is a scalar, and that $\mathbf{a}.\mathbf{b}$ is the same as $\mathbf{b}.\mathbf{a}$. For example, we have

$$(1,2,3).(10,100,1000) = 10 + 200 + 3000 = 3210.$$

Note also that $\mathbf{a}.\mathbf{a} = x^2 + y^2 + z^2 = |\mathbf{a}|^2$.
For the unit vectors **i**, **j** and **k** we have

$$\mathbf{i}.\mathbf{i} = 1 \quad \mathbf{i}.\mathbf{j} = 0 \quad \mathbf{i}.\mathbf{k} = 0 \quad \mathbf{j}.\mathbf{i} = 0 \quad \mathbf{j}.\mathbf{j} = 1 \quad \mathbf{j}.\mathbf{k} = 0 \quad \mathbf{k}.\mathbf{i} = 0 \quad \mathbf{k}.\mathbf{j} = 0 \quad \mathbf{k}.\mathbf{k} = 1.$$

Geometrically: $\mathbf{a}.\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\theta)$, where θ is the angle between **a** and **b**. In particular, as $-1 \leq \cos(\theta) \leq 1$ this means that $-|\mathbf{a}||\mathbf{b}| \leq \mathbf{a}.\mathbf{b} \leq |\mathbf{a}||\mathbf{b}|$, or equivalently $|\mathbf{a}.\mathbf{b}| \leq |\mathbf{a}||\mathbf{b}|$. This is called the *Cauchy-Schwartz inequality*. We also see that $\mathbf{a}.\mathbf{b}$ is zero when $\theta = \pi/2$, which means that **a** and **b** are perpendicular to each other.



$\theta < \frac{\pi}{2}; \quad \mathbf{a}.\mathbf{b} > 0 \qquad \theta = \frac{\pi}{2}; \quad \mathbf{a}.\mathbf{b} = 0 \qquad \theta > \frac{\pi}{2}; \quad \mathbf{a}.\mathbf{b} < 0$

Angle example

Consider the vectors $\mathbf{a} = (3,0,4)$ and $\mathbf{b} = (2,-1,2)$. We will find the angle θ between **a** and **b**. The inner products are

$$\begin{aligned}|\mathbf{a}|^2 &= \mathbf{a}.\mathbf{a} = 3^2 + 0^2 + 4^2 = 25\\ |\mathbf{b}|^2 &= \mathbf{b}.\mathbf{b} = 2^2 + (-1)^2 + 2^2 = 9\\ |\mathbf{a}||\mathbf{b}|\cos(\theta) &= \mathbf{a}.\mathbf{b} = 3 \times 2 + 0 \times (-1) + 4 \times 2 = 14.\end{aligned}$$

From this we see that $|\mathbf{a}| = \sqrt{25} = 5$ and $|\mathbf{b}| = \sqrt{9} = 3$, so

$$\cos(\theta) = \frac{\mathbf{a}.\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{14}{5 \times 3} = \frac{14}{15} \simeq 0.933.$$

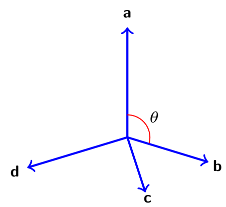
This means that $\theta = \arccos(0.933)$, which is 0.367 radians or 21.04 degrees.

Methane

The hydrogen atoms in a molecule of methane lie at the following positions:

$$\mathbf{a} = (0,0,1) \qquad \mathbf{b} = \left(\frac{2\sqrt{2}}{3}, 0, -\frac{1}{3}\right)$$

$$\mathbf{c} = \left(-\frac{\sqrt{2}}{3}, \frac{\sqrt{6}}{3}, -\frac{1}{3}\right) \qquad \mathbf{d} = \left(-\frac{\sqrt{2}}{3}, -\frac{\sqrt{6}}{3}, -\frac{1}{3}\right).$$



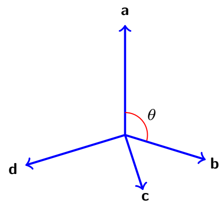
It is clear that **a** is a unit vector. We also have

$$\begin{aligned}|\mathbf{b}|^2 &= \left(\frac{2\sqrt{2}}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{4 \times 2}{9} + \frac{1}{9} = 1\\ |\mathbf{c}|^2 &= \left(\frac{\sqrt{2}}{3}\right)^2 + \left(\frac{\sqrt{6}}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{2}{9} + \frac{6}{9} + \frac{1}{9} = 1\end{aligned}$$

so **b** and **c** are unit vectors, and **d** is also a unit vector by the same calculation as for **c**. It is also clear that $\mathbf{a}.\mathbf{b} = \mathbf{a}.\mathbf{c} = \mathbf{a}.\mathbf{d} = -1/3$. In fact, we also have $\mathbf{b}.\mathbf{c} = \mathbf{b}.\mathbf{d} = \mathbf{c}.\mathbf{d} = -1/3$.

Methane

$$\mathbf{a} = (0,0,1) \qquad \mathbf{b} = \left(\frac{2\sqrt{2}}{3}, 0, -\frac{1}{3}\right)$$
$$\mathbf{c} = \left(-\frac{\sqrt{2}}{3}, \frac{\sqrt{6}}{3}, -\frac{1}{3}\right) \qquad \mathbf{d} = \left(-\frac{\sqrt{2}}{3}, -\frac{\sqrt{6}}{3}, -\frac{1}{3}\right).$$



$$\mathbf{b}.\mathbf{c} = \frac{2\sqrt{2}}{3}.\left(-\frac{\sqrt{2}}{3}\right) + 0.\frac{\sqrt{6}}{3} + \left(-\frac{1}{3}\right)\left(-\frac{1}{3}\right) = \frac{-4}{9} + \frac{1}{9} = -\frac{1}{3}$$
$$\mathbf{b}.\mathbf{d} = \frac{2\sqrt{2}}{3}.\left(-\frac{\sqrt{2}}{3}\right) + 0.\left(-\frac{\sqrt{6}}{3}\right) + \left(-\frac{1}{3}\right)\left(-\frac{1}{3}\right) = \frac{-4}{9} + \frac{1}{9} = -\frac{1}{3}$$
$$\mathbf{c}.\mathbf{d} = \left(-\frac{\sqrt{2}}{3}\right)\left(-\frac{\sqrt{2}}{3}\right) + \left(\frac{\sqrt{6}}{3}\right)\left(-\frac{\sqrt{6}}{3}\right) + \left(-\frac{1}{3}\right)\left(-\frac{1}{3}\right) = \frac{2}{9} - \frac{6}{9} + \frac{1}{9} = -\frac{1}{3}.$$

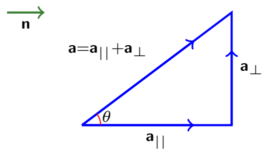
If θ is the angle between $\mathbf{a}$ and $\mathbf{b}$, then we have

$$\cos(\theta) = \frac{\mathbf{a}.\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{-1/3}{1 \times 1} = -\frac{1}{3},$$

so θ is $\arccos(-1/3)$, which is 1.911 radians or 109.5 degrees. By the same calculation, the angle between any two of the atoms is 109.5 degrees.

Parallel and perpendicular components

Now suppose we have a vector $\mathbf{a}$ and a unit vector $\mathbf{n}$. We can write $\mathbf{a}$ as $\mathbf{a}_{||} + \mathbf{a}_{\perp}$, where $\mathbf{a}_{||}$ is the part parallel to $\mathbf{n}$, and $\mathbf{a}_{\perp}$ is the part perpendicular to $\mathbf{n}$.



In the picture, θ is the angle between $\mathbf{a}$ and $\mathbf{a}_{||}$, which is the same as the angle between $\mathbf{a}$ and $\mathbf{n}$. From this (and the fact that $|\mathbf{n}| = 1$) it follows that

$$\mathbf{a}.\mathbf{n} = |\mathbf{a}||\mathbf{n}|\cos(\theta) = |\mathbf{a}|\cos(\theta) = |\mathbf{a}_{||}|.$$

(Equation of scalars, valid for $\theta \leq \frac{\pi}{2}$; for all θ we have $|\mathbf{a}.\mathbf{n}| = |\mathbf{a}_{||}|$.)

$$\mathbf{a}_{||} = (\mathbf{a}.\mathbf{n})\mathbf{n}$$
$$\mathbf{a}_{\perp} = \mathbf{a} - (\mathbf{a}.\mathbf{n})\mathbf{n}.$$

Parallel and perpendicular example

$$\mathbf{a}_{||} = (\mathbf{a}.\mathbf{n})\mathbf{n} \qquad \mathbf{a}_{\perp} = \mathbf{a} - (\mathbf{a}.\mathbf{n})\mathbf{n}.$$

Consider the vector $\mathbf{a} = (3,6,9)$ and the unit vector $\mathbf{n} = (2/3, 2/3, -1/3)$. We have

$$\mathbf{a}.\mathbf{n} = 3.\frac{2}{3} + 6.\frac{2}{3} + 9.\frac{-1}{3} = 2 + 4 - 3 = 3$$
$$\mathbf{a}_{||} = (\mathbf{a}.\mathbf{n})\mathbf{n} = 3\mathbf{n} = (2,2,-1)$$
$$\mathbf{a}_{\perp} = \mathbf{a} - \mathbf{a}_{||} = (3,6,9) - (2,2,-1) = (1,4,10).$$

The cross product

We next recall the cross product operation. For vectors $\mathbf{a} = (x,y,z)$ and $\mathbf{b} = (u,v,w)$, we define

$$\mathbf{a} \times \mathbf{b} = (x,y,z) \times (u,v,w) = (yw - zv, zu - xw, xv - yu).$$

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ u & v & w \end{bmatrix} = \det \begin{bmatrix} y & z \\ v & w \end{bmatrix} \mathbf{i} - \det \begin{bmatrix} x & z \\ u & w \end{bmatrix} \mathbf{j} + \det \begin{bmatrix} x & y \\ u & v \end{bmatrix} \mathbf{k}.$$

Note that $\mathbf{a} \times \mathbf{b}$ is a vector, in contrast to $\mathbf{a}.\mathbf{b}$, which is a scalar.

Example: Consider the vectors $\mathbf{a} = (1,2,3)$ and $\mathbf{b} = (3,2,1)$. We have

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \det \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \mathbf{i} - \det \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \mathbf{j} + \det \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \mathbf{k} =$$
$$-4\mathbf{i} - (-8)\mathbf{j} + (-4)\mathbf{k} = (-4, 8, -4).$$

Example: For the standard unit vectors you can check that

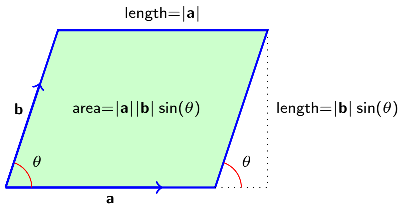
$$\begin{array}{lll} \mathbf{i} \times \mathbf{i} = 0 & \mathbf{i} \times \mathbf{j} = \mathbf{k} & \mathbf{i} \times \mathbf{k} = -\mathbf{j} \\ \mathbf{j} \times \mathbf{i} = -\mathbf{k} & \mathbf{j} \times \mathbf{j} = 0 & \mathbf{j} \times \mathbf{k} = \mathbf{i} \\ \mathbf{k} \times \mathbf{i} = \mathbf{j} & \mathbf{k} \times \mathbf{j} = -\mathbf{i} & \mathbf{k} \times \mathbf{k} = 0. \end{array}$$

Cross product geometry

Geometrically, it can be shown that $\mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, and that

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin(\theta) = \text{ area of the parallelogram spanned by } \mathbf{a} \text{ and } \mathbf{b},$$

where θ is again the angle between $\mathbf{a}$ and $\mathbf{b}$.



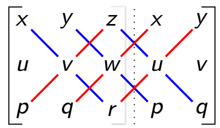
In particular, we see that $\mathbf{a} \times \mathbf{b}$ is zero when $\sin(\theta) = 0$, which means that $\theta = 0$ or $\theta = \pi$, so $\mathbf{a}$ and $\mathbf{b}$ have the same direction or opposite directions. Algebraically, we have the following identities:

$$\mathbf{a} \times \mathbf{a} = \mathbf{0} \qquad \mathbf{b} \times \mathbf{a} = -\mathbf{a} \times \mathbf{b}$$
$$\mathbf{a}.(\mathbf{a} \times \mathbf{b}) = 0 \qquad \mathbf{b}.(\mathbf{a} \times \mathbf{b}) = 0.$$

The scalar triple product

Suppose we have vectors $\mathbf{a} = (x, y, z)$, $\mathbf{b} = (u, v, w)$ and $\mathbf{c} = (p, q, r)$. We can take the cross product $\mathbf{b} \times \mathbf{c}$, which is a vector, and then take the dot product of that vector with $\mathbf{a}$ to get a scalar $\mathbf{a}.(\mathbf{b} \times \mathbf{c})$, which is called the *scalar triple product* of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. Using the determinant formula for $\mathbf{b} \times \mathbf{c}$ we find that $\mathbf{a}.(\mathbf{b} \times \mathbf{c})$ is also a determinant:

$$\mathbf{a}.(\mathbf{b} \times \mathbf{c}) = \det \begin{bmatrix} x & y & z \\ u & v & w \\ p & q & r \end{bmatrix}.$$



A convenient trick for expanding such determinants is as follows. We first expand the matrix by repeating the first two columns at the end, then draw sloping lines as shown. For each of the blue lines sloping down and to the right, we have a term with a plus sign. For example, the first blue line joins x, v and r, giving a term $+xvr$. Each of the red lines sloping down and to the left gives a term with a minus sign. Altogether, the determinant is

$$\mathbf{a}.(\mathbf{b} \times \mathbf{c}) = \det \begin{bmatrix} x & y & z \\ u & v & w \\ p & q & r \end{bmatrix} = xvr + ywp + zuq - zvp - xwq - yur.$$

The scalar triple product

$$\mathbf{a} = (x, y, z) \qquad \mathbf{b} = (u, v, w) \qquad \mathbf{c} = (p, q, r)$$

$$\mathbf{a}.(\mathbf{b} \times \mathbf{c}) = \det \begin{bmatrix} x & y & z \\ u & v & w \\ p & q & r \end{bmatrix}.$$

There are a number of slight variants of the scalar triple product, but they all turn out to be the same, at least up to a plus or minus sign. Specifically, we have

$$\mathbf{a}.(\mathbf{b} \times \mathbf{c}) = \mathbf{b}.(\mathbf{c} \times \mathbf{a}) = \mathbf{c}.(\mathbf{a} \times \mathbf{b}) = -\mathbf{a}.(\mathbf{c} \times \mathbf{b}) = -\mathbf{b}.(\mathbf{a} \times \mathbf{c}) = -\mathbf{c}.(\mathbf{b} \times \mathbf{a}).$$

We also have $\mathbf{a}.(\mathbf{b} \times \mathbf{c}) = (\mathbf{b} \times \mathbf{c}).\mathbf{a}$ and so on, just because $\mathbf{u}.\mathbf{v} = \mathbf{v}.\mathbf{u}$ for any vectors $\mathbf{u}$ and $\mathbf{v}$.

Vector triple products

We can take the cross product of the vector $\mathbf{a}$ with the vector $\mathbf{b} \times \mathbf{c}$ to get another vector $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. **Warning:** this is not the same as $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$. However, both of these iterated cross products, and various variants, can be described in terms of dot products as follows:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}.\mathbf{c})\mathbf{b} - (\mathbf{a}.\mathbf{b})\mathbf{c}$$
$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a}.\mathbf{c})\mathbf{b} - (\mathbf{b}.\mathbf{c})\mathbf{a}.$$

The following observations may help you remember the rules:

(a) The vector outside the brackets on the left occurs in both the dot products on the right.

(b) Each of the vectors inside the brackets on the left occurs in one of the dot products on the right.

(c) The dot product of the first vector with the last vector occurs with a plus sign. The other dot product occurs with a minus sign.

Vector fields and scalar fields

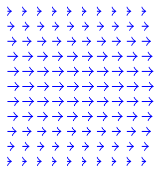
In many applications, we do not consider individual vectors or scalars, but functions that give a vector or scalar at every point. Such functions are called *vector fields* or *scalar fields*. For example:

(a) Suppose we want to model the flow of air around an aeroplane. The velocity of the air flow at any given point is a vector. These vectors will be different at different points, so they are functions of position (and also of time). Thus, the air velocity is a vector field. Similarly, the pressure and temperature are scalar quantities that depend on position, or in other words, they are scalar fields.

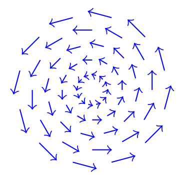
(b) The magnetic field inside an electrical machine is a vector that depends on position, or in other words a vector field. The electric potential is a scalar field.

Although we will mainly be concerned with scalar and vector fields in three-dimensional space, we will sometimes use two-dimensional examples because they are easier to visualise.

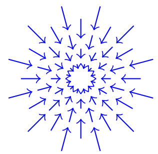
Example vector fields in two dimensions



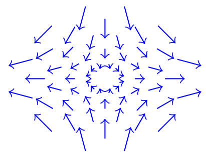
$\mathbf{u} = (1 - y^2, 0)$



$\mathbf{u} = (-y/3, x/3)$



$\mathbf{u} = (-x/3, -y/3)$



$\mathbf{u} = (x/3, -y/3)$

The gradient of a scalar field

If f is a scalar field, then we define $\nabla(f) = (f_x, f_y, f_z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$.

(A vector field, the *gradient* of f, sometimes written $\mathrm{grad}(f)$ rather than $\nabla(f)$.)

(a) For the function $f = x^3 + y^4 + z^5$, we have $\nabla(f) = (3x^2, 4y^3, 5z^4)$.

(b) For the function $f = \sin(x)\sin(y)\sin(z)$ we have

$$\nabla(f) = (\cos(x)\sin(y)\sin(z),\ \sin(x)\cos(y)\sin(z),\ \sin(x)\sin(y)\cos(z)).$$

(c) For the function $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$ we have

$$r_x = \tfrac{1}{2}(x^2 + y^2 + z^2)^{-\frac{1}{2}}.2x = \frac{x}{(x^2 + y^2 + z^2)^{\frac{1}{2}}} = x/r,$$

and similarly $r_y = y/r$ and $r_z = z/r$. This means that

$$\nabla(r) = (x/r, y/r, z/r).$$

More generally, for any n we have

$$(r^n)_x = nr^{n-1}r_x = nr^{n-1}x/r = nr^{n-2}x.$$

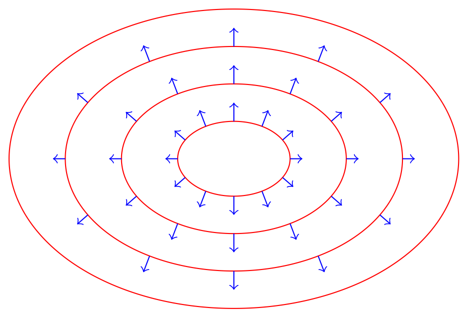
The other two derivatives work in the same way, so

$$\nabla(r^n) = nr^{n-2}\,(x, y, z).$$

Geometry of the gradient

Fact: The vector $\nabla(f)$ points in the direction of maximum increase of f. It is perpendicular to the surfaces where f is constant.

The picture below illustrates the two-dimensional version of this fact in the case where $f = \sqrt{x^2/9 + y^2/4}$.



The four red ovals are given by $f = 1$, $f = 2$, $f = 3$ and $f = 4$. The blue arrows show the vector field $\nabla(f)$, which is perpendicular to the red ovals as expected.

Geometry of the gradient

Fact: The vector $\nabla(f)$ points in the direction of maximum increase of f. It is perpendicular to the surfaces where f is constant.

To see why the above fact is true, remember that if we make small changes δx, δy and δz to x, y and z, then the resulting change in f is approximately given by

$$\delta f = f_x\,\delta x + f_y\,\delta y + f_z\,\delta z.$$

If we write $\mathbf{r}$ for the vector (x, y, z), this becomes

$$\delta f = \nabla(f).\delta\mathbf{r} = |\nabla(f)||\delta\mathbf{r}|\cos(\theta),$$

where θ is the angle between $\delta\mathbf{r}$ and $\nabla(f)$. If we move along a surface where f is constant, then δf will be zero so we must have $\cos(\theta) = 0$, so $\theta = \pm\pi/2$, so $\delta\mathbf{r}$ is perpendicular to $\nabla(f)$. This means that $\nabla(f)$ is perpendicular to the surfaces of constant f, as we stated before. On the other hand, to make δf as large as possible (for a fixed step size $|\delta\mathbf{r}|$) we need to maximise $\cos(\theta)$, which means taking $\theta = 0$, so that $\delta\mathbf{r}$ is in the same direction as $\nabla(f)$. In other words, $\nabla(f)$ points in the direction of maximum increase of f.

Applications of the gradient

(a) We write $\mathbf{E}$ for the electric field (which is a vector field) and ϕ for the electric potential (which is a scalar field). These are related by the equation $\mathbf{E} = \nabla(\phi)$. (All this is valid only when there are no significant time-varying magnetic fields.)

(b) Similarly, there is a gravitational potential function ψ, and the gravitational force field is proportional to $\nabla(\psi)$.

(c) The net force on a particle of air involves $\nabla(p)$, where p is the pressure.

Electric field of a point charge

If we have a single charge at the origin, then the resulting electric potential function is $\phi = Ar^{-1}$ for some constant A, where $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$ as usual. Note that

$$(r^{-1})_x = -\tfrac{1}{2}(x^2 + y^2 + z^2)^{-\frac{3}{2}}.2x = -x/r^3$$

and similarly $(r^{-1})_y = -y/r^3$ and $(r^{-1})_z = -z/r^3$. This gives the electric field:

$$\mathbf{E} = \nabla(\phi) = -Ar^{-3}(x, y, z) = -A\mathbf{r}/r^3.$$

Electric field of a line charge

Suppose instead that we have a whole line of charges distributed along the z-axis. It works out that the corresponding electric potential function is $\phi = -\frac{1}{2}A\ln(x^2 + y^2)$ for some constant A. This is independent of z, so $\phi_z = 0$. On the other hand, we have

$$\phi_x = \frac{-\frac{1}{2}A}{x^2 + y^2}.2x = -\frac{Ax}{x^2 + y^2}.$$

By a similar calculation we have $\phi_y = -Ay/(x^2 + y^2)$, so

$$\mathbf{E} = \nabla(\phi) = \left(-\frac{Ax}{x^2 + y^2}, -\frac{Ay}{x^2 + y^2}, 0\right).$$

Uniform electric field

Suppose we have an electric potential of the form $\phi = ax + by + cz$, where a, b and c are constant. The corresponding electric field is

$$\mathbf{E} = \nabla(\phi) = (a, b, c).$$

In other words, we have a uniform electric field everywhere. If we put $\mathbf{u} = (a, b, c)$ we can write the above in vector notation as $\phi = \mathbf{u}.\mathbf{r}$ and $\nabla(\phi) = \nabla(\mathbf{u}.\mathbf{r}) = \mathbf{u}$.

grad(θ)

Consider the function

$$\theta(x, y, z) = \text{ angle between the } x\text{-axis and } (x, y, 0) \; = \arctan(y/x)$$

(as used in polar coordinates).
It is a standard fact that $\arctan'(t) = 1/(1 + t^2)$. Using this, we get

$$\begin{aligned}
\theta_x &= \arctan'\left(\frac{y}{x}\right)\frac{\partial}{\partial x}\left(\frac{y}{x}\right) = \frac{1}{1 + (y/x)^2}\frac{-y}{x^2} = \frac{-y}{x^2 + y^2} \\
\theta_y &= \arctan'\left(\frac{y}{x}\right)\frac{\partial}{\partial y}\left(\frac{y}{x}\right) = \frac{1}{1 + (y/x)^2}\frac{1}{x} = \frac{x}{x^2 + y^2} \\
\theta_z &= 0,
\end{aligned}$$

so

$$\nabla(\theta) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0\right).$$

div and curl

Now suppose we have a vector field $\mathbf{u} = (f, g, h)$, so f, g and h are all functions of x, y and z. We can think of ∇ as itself being a strange kind of vector, in which the entries are differential operators:

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right).$$

This means we can make sense of the dot product $\nabla.\mathbf{u}$ and the cross product $\nabla \times \mathbf{u}$ as follows:

$$\begin{aligned}
\nabla.\mathbf{u} &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right).(f, g, h) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} = f_x + g_y + h_z \\
\nabla \times \mathbf{u} &= \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f & g & h \end{bmatrix} = (h_y - g_z, f_z - h_x, g_x - f_y).
\end{aligned}$$

Note that $\nabla.\mathbf{u}$ is a scalar field, and $\nabla \times \mathbf{u}$ is a vector field. The scalar field $\nabla.\mathbf{u}$ is called the *divergence* of $\mathbf{u}$, and is sometimes written as $\mathrm{div}(\mathbf{u})$. The vector field $\nabla \times \mathbf{u}$ is called the *curl* of $\mathbf{u}$, and is sometimes written $\mathrm{curl}(\mathbf{u})$.

Examples of div and curl

(a) For the vector field $\mathbf{u} = (x^2 + y^2, y^2 + z^2, z^2 + x^2)$ we have

$$\nabla.\mathbf{u} = \frac{\partial}{\partial x}(x^2 + y^2) + \frac{\partial}{\partial y}(y^2 + z^2) + \frac{\partial}{\partial z}(z^2 + x^2) = 2x + 2y + 2z$$

$$\nabla \times \mathbf{u} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 + y^2 & y^2 + z^2 & z^2 + x^2 \end{bmatrix} = (-2z, -2x, -2y).$$

(b) For the vector field $\mathbf{u} = (\sin(x), \sin(x), \sin(x))$ we have

$$\nabla.\mathbf{u} = \frac{\partial}{\partial x}\sin(x) + \frac{\partial}{\partial y}\sin(x) + \frac{\partial}{\partial z}\sin(x) = \cos(x) + 0 + 0 = \cos(x)$$

$$\nabla \times \mathbf{u} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \sin(x) & \sin(x) & \sin(x) \end{bmatrix} = (0, -\cos(x), \cos(x)).$$

(c) For the vector field $\mathbf{u} = (-y, x, z)$ we have

$$\nabla.\mathbf{u} = \frac{\partial}{\partial x}(-y) + \frac{\partial}{\partial y}(x) + \frac{\partial}{\partial z}(z) = 0 + 0 + 1 = 1$$

$$\nabla \times \mathbf{u} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y & x & z \end{bmatrix} = (0, 0, 2).$$

grad, div and curl in two dimensions

(a) For a scalar field f in two dimensions, $\mathrm{grad}(f) = \nabla(f) = (f_x, f_y)$ (a vector field).

(b) For a vector field $\mathbf{u} = (p, q)$ in two dimensions, $\mathrm{div}(\mathbf{u}) = \nabla.\mathbf{u} = p_x + q_y$ (a scalar field).

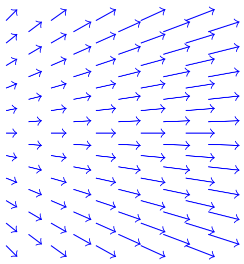
(c) For a vector field $\mathbf{u} = (p, q)$ in two dimensions,

$$\mathrm{curl}(\mathbf{u}) = \det \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ p & q \end{bmatrix} = q_x - p_y$$

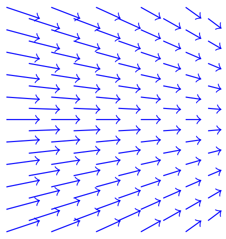
(a *scalar* field, not a vector field as in three dimensions).

Geometric interpretation of div($\mathbf{u}$)

It works out that the divergence $\mathrm{div}(\mathbf{u}) = \nabla.\mathbf{u}$ is positive when the vectors $\mathbf{u}$ are spreading out, and negative when they are coming together.



diverging: $\nabla.\mathbf{u} > 0$

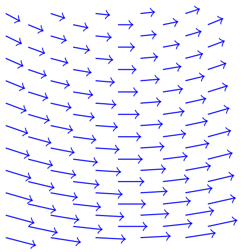


converging: $\nabla.\mathbf{u} < 0$

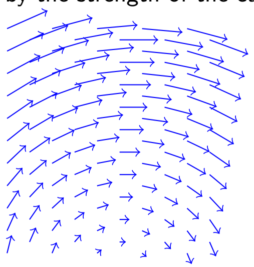
For the velocity field of an incompressible fluid we will have $\nabla.\mathbf{u} = 0$.

Geometric interpretation of curl($\mathbf{u}$)

In two dimensions, it works out that $\mathrm{curl}(\mathbf{u}) > 0$ in regions where the field is curling anticlockwise, and $\mathrm{curl}(\mathbf{u}) < 0$ in regions where it is curling clockwise, and the absolute value of $\mathrm{curl}(\mathbf{u})$ is determined by the strength of the curling.



$\mathrm{curl}(\mathbf{u}) > 0$, smaller



$\mathrm{curl}(\mathbf{u}) < 0$, larger

In three dimensions, the field $\mathbf{u}$ can curl around any axis. In this context, $\mathrm{curl}(\mathbf{u})$ is also a vector field, and it will point along the axis of the curling.

Maxwell's equations

These involve:

- The electric field $\mathbf{E}$, which is a vector field.
- The magnetic field $\mathbf{B}$, which is another vector field.
- The current density $\mathbf{J}$, which is also a vector field.
- The charge density ρ, which is a scalar field.
- Two constants: $\epsilon_0 \simeq 8.854 \times 10^{-12} F/m^2$ and $\mu_0 \simeq 1.257 \times 10^{-6} Hm^{-1}$.

The quantities $\mathbf{E}$, $\mathbf{B}$, $\mathbf{J}$ and ρ may also depend on time; we write $\dot{\mathbf{E}}$ for $\partial\mathbf{E}/\partial t$ and so on. The various fields are related by the following equations:

$$\nabla.\mathbf{E} = \rho/\epsilon_0 \qquad \nabla \times \mathbf{E} = -\dot{\mathbf{B}}$$

$$\nabla.\mathbf{B} = 0 \qquad \nabla \times \mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\dot{\mathbf{E}}.$$

This means that:

- The electric field diverges in regions where there is positive charge, and converges in regions where there is negative charge.
- The magnetic field never diverges or converges.
- Changing magnetic fields cause the electric field to curl.
- Currents cause the magnetic field to curl. Changing electric fields also cause the magnetic field to curl, but the effect is usually much weaker, because ϵ_0 is small.

Plane wave solution to Maxwell's equations

One class of solutions to Maxwell's equations is as follows. Put $c = 1/\sqrt{\mu_0\epsilon_0} \simeq 3 \times 10^8 ms^{-1}$ (which turns out to be the speed of light), and let α be any constant. We can take $\mathbf{J} = 0$ and $\rho = 0$ and

$$\mathbf{E} = (0, \sin(\alpha(x - ct)), 0) \qquad \mathbf{B} = (0, 0, \sin(\alpha(x - ct))/c).$$

We find that

$$\nabla.\mathbf{E} = \frac{\partial}{\partial y}\sin(\alpha(x - ct)) = 0 = \rho/\epsilon_0 \qquad \dot{\mathbf{E}} = (0, -\alpha c\cos(\alpha(x - ct)), 0)$$

$$\nabla.\mathbf{B} = \frac{\partial}{\partial z}\sin(\alpha(x - ct))/c = 0 \qquad \dot{\mathbf{B}} = (0, 0, -\alpha\cos(\alpha(x - ct))$$

$$\nabla \times \mathbf{E} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & \sin(\alpha(x - ct)) & 0 \end{bmatrix} = (0, 0, \alpha\cos(\alpha(x - ct))) = -\dot{\mathbf{B}}$$

$$\nabla \times \mathbf{B} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & \sin(\alpha(x - ct))/c \end{bmatrix} = (0, -\alpha\cos(\alpha(x - ct))/c, 0) = \dot{\mathbf{E}}/c^2 = \mu_0\epsilon_0\dot{\mathbf{E}}.$$

This shows that we do indeed have a solution to the equations. It represents an electromagnetic wave of wavelength $1/\alpha$ moving at speed c in the x-direction.

Stationary charged particle

Another solution to Maxwell's equations has $\mathbf{E} = (-xr^{-3}, -yr^{-3}, -zr^{-3})$ with all other fields ($\mathbf{B}$, $\mathbf{J}$ and ρ) being zero. It is clear that $\dot{\mathbf{E}} = 0$ and $\dot{\mathbf{B}} = 0$, so the only equations that we need to check are that $\nabla.\mathbf{E} = 0$ and $\nabla \times \mathbf{E} = 0$. For this we recall that $r_x = x/r$, so $(r^{-3})_x = -3r^{-4}r_x = -3xr^{-5}$. In the same way, we have $(r^{-3})_y = -3yr^{-5}$ and $(r^{-3})_z = -3zr^{-5}$. Using this we find that

$$\begin{array}{lll} (-xr^{-3})_x = 3x^2r^{-5} - r^{-3} & (-xr^{-3})_y = 3xyr^{-5} & (-xr^{-3})_z = 3xzr^{-5} \\ (-yr^{-3})_x = 3xyr^{-5} & (-yr^{-3})_y = 3y^2r^{-5} - r^{-3} & (-yr^{-3})_z = 3yzr^{-5} \\ (-zr^{-3})_x = 3xzr^{-5} & (-zr^{-3})_y = 3yzr^{-5} & (-zr^{-3})_z = 3z^2r^{-5} - r^{-3}. \end{array}$$

$$\begin{aligned} \nabla.\mathbf{E} &= (-xr^{-3})_x + (-yr^{-3})_y + (-zr^{-3})_z \\ &= 3x^2r^{-5} - r^{-3} + 3y^2r^{-5} - r^{-3} + 3z^2r^{-5} - r^{-3} \\ &= 3(x^2 + y^2 + z^2)r^{-5} - 3r^{-3} = 3r^2r^{-5} - 3r^{-3} = 0. \end{aligned}$$

$$\begin{aligned} \nabla \times \mathbf{E} &= \Big((-zr^{-3})_y - (-yr^{-3})_z, (-xr^{-3})_z - (-zr^{-3})_x, (-yr^{-3})_x - (-xr^{-3})_y\Big) \\ &= \Big(3yzr^{-3} - 3yzr^{-3}, 3xzr^{-3} - 3xzr^{-3}, 3xyr^{-3} - 3xyr^{-3}\Big) = (0, 0, 0). \end{aligned}$$

This shows that we have a solution to the equations, as claimed. This one represents the electric field of a single stationary particle at the origin, with no magnetic field.

Identities involving div, grad and curl

Let $\mathbf{u}$ and $\mathbf{v}$ be vector fields, let f be a scalar field, and let p be a function of one variable. Then:

$$\begin{array}{rr} \nabla(f + g) = \nabla(f) + \nabla(g) & \nabla(fg) = f\,\nabla(g) + g\,\nabla(f) \\ \nabla.(\mathbf{u} + \mathbf{v}) = \nabla.\mathbf{u} + \nabla.\mathbf{v} & \nabla.(f\mathbf{u}) = f\nabla.\mathbf{u} + \nabla(f).\mathbf{u} \\ \nabla \times (\mathbf{u} + \mathbf{v}) = \nabla \times \mathbf{u} + \nabla \times \mathbf{v} & \nabla \times (f\mathbf{u}) = f\nabla \times \mathbf{u} + \nabla(f) \times \mathbf{u} \\ \nabla(p(f)) = p'(f)\,\nabla(f) & \nabla.(\mathbf{u} \times \mathbf{v}) = \mathbf{v}.(\nabla \times \mathbf{u}) - \mathbf{u}.(\nabla \times \mathbf{v}). \end{array}$$

Example: we will check $\nabla.(\mathbf{u} \times \mathbf{v}) = \mathbf{v}.(\nabla \times \mathbf{u}) - \mathbf{u}.(\nabla \times \mathbf{v})$.
Suppose that $\mathbf{u} = (f, g, h)$ and $\mathbf{v} = (p, q, r)$. Then

$$\mathbf{u} \times \mathbf{v} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ f & g & h \\ p & q & r \end{bmatrix} = (gr - hq, hp - fr, fq - gp)$$

$$\begin{aligned} &\nabla.(\mathbf{u} \times \mathbf{v}) \\ &= (gr - hq)_x + (hp - fr)_y + (fq - gp)_z \\ &= (g_xr + gr_x - h_xq - hq_x) + (h_yp + hp_y - f_yr - fr_y) + (f_zq + fq_z - g_zp - gp_z) \\ &= p(h_y - g_z) + q(f_z - h_x) + r(g_x - f_y) + f(q_z - r_y) + g(r_x - p_z) + h(p_y - q_x) \\ &= (p, q, r).(h_y - g_z, f_z - h_x, g_x - f_y) - (f, g, h).(r_y - q_z, p_z - r_x, q_x - p_y) \\ &= \mathbf{v}.(\nabla \times \mathbf{u}) - \mathbf{u}.(\nabla \times \mathbf{v}). \end{aligned}$$

Second-order operators

There are several different ways to combine the div, grad and curl operators:

$$\begin{array}{l} \text{scalar field} \xrightarrow{\text{grad}} \text{vector field} \xrightarrow{\text{div}} \text{scalar field} \\ \text{scalar field} \xrightarrow{\text{grad}} \text{vector field} \xrightarrow{\text{curl}} \text{vector field} \\ \text{vector field} \xrightarrow{\text{div}} \text{scalar field} \xrightarrow{\text{grad}} \text{vector field} \\ \text{vector field} \xrightarrow{\text{curl}} \text{vector field} \xrightarrow{\text{div}} \text{scalar field} \\ \text{vector field} \xrightarrow{\text{curl}} \text{vector field} \xrightarrow{\text{curl}} \text{vector field}\ . \end{array}$$

(No other combinations make sense.
For example, we cannot define curl(div($\mathbf{u}$)), because div($\mathbf{u}$) is a scalar field, and we can only take the curl of a vector field.)
It is important that two of above combinations are automatically zero.

Fact

(a) *For any scalar field f we have* curl(grad(f)) $= \nabla \times (\nabla(f)) = 0$.

(b) *For any vector field $\mathbf{u}$ we have* div(curl($\mathbf{u}$)) $= \nabla.(\nabla \times \mathbf{u}) = 0$.

Second-order operators

(a) For any scalar field f we have $\mathrm{curl}(\mathrm{grad}(f)) = \nabla \times (\nabla(f)) = 0$.
(b) For any vector field $\mathbf{u}$ we have $\mathrm{div}(\mathrm{curl}(\mathbf{u})) = \nabla.(\nabla \times \mathbf{u}) = 0$.

These can be checked directly. For a scalar field f, we have $\nabla(f) = (f_x, f_y, f_z)$. After remembering that $f_{xy} = f_{yx}$ and so on, we find that

$$\nabla \times (\nabla(f)) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_x & f_y & f_z \end{bmatrix} = (f_{zy} - f_{yz}, f_{xz} - f_{zx}, f_{yx} - f_{xy}) = (0, 0, 0).$$

Now consider instead a vector field $\mathbf{u} = (p, q, r)$. We have

$$\nabla \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ p & q & r \end{bmatrix} = (r_y - q_z, p_z - r_x, q_x - p_y),$$

so

$$\begin{aligned} \nabla.(\nabla \times \mathbf{u}) &= (r_y - q_z)_x + (p_z - r_x)_y + (q_x - p_y)_z \\ &= r_{yx} - q_{zx} + p_{zy} - r_{xy} + q_{xz} - p_{yz} \\ &= p_{zy} - p_{yz} + q_{xz} - q_{zx} + r_{yx} - r_{xy} = 0. \end{aligned}$$

Second-order operators

There are three more possible combinations.

(a) For a scalar field f we have $\mathrm{div}(\mathrm{grad}(f)) = \nabla.(\nabla(f)) = f_{xx} + f_{yy} + f_{zz}$. This is usually written as $\nabla^2(f)$, and called the *Laplacian* of f. Note that the Laplacian of a scalar field is a scalar field. We can also define the Laplacian of a vector field by the rule

$$\nabla^2(p, q, r) = (\nabla^2(p), \nabla^2(q), \nabla^2(r)) = (p_{xx}+p_{yy}+p_{zz}, q_{xx}+q_{yy}+q_{zz}, r_{xx}+r_{yy}+r_{zz}).$$

Note that the Laplacian of a vector field is again a vector field.

(b) For a vector field $\mathbf{u} = (p, q, r)$ we have

$$\begin{aligned} \mathrm{grad}(\mathrm{div}(\mathbf{u})) &= \nabla(\nabla.\mathbf{u}) = \nabla(p_x + q_y + r_z) \\ &= (p_{xx} + q_{yx} + r_{zx}, p_{xy} + q_{yy} + r_{zy}, p_{xz} + q_{yz} + r_{zz}). \end{aligned}$$

(c) The last remaining combination can be expressed in terms of (a) and (b), by the equation

$$\mathrm{curl}(\mathrm{curl}(\mathbf{u})) = \nabla \times (\nabla \times \mathbf{u}) = \nabla(\nabla.\mathbf{u}) - \nabla^2(\mathbf{u}).$$

It is straightforward but somewhat lengthy to check this; we will not give the details.

Second-order operators in two dimensions

In two dimensions, the situation is similar but simpler:

(a) For any scalar field f we have

$$\mathrm{div}(\mathrm{grad}(f)) = \nabla.(\nabla(f)) = f_{xx} + f_{yy},$$

which is again called the Laplacian and denoted by $\nabla^2(f)$.

(b) We also have

$$\mathrm{curl}(\mathrm{grad}(f)) = \mathrm{curl}(f_x, f_y) = f_{yx} - f_{xy} = 0.$$

Incompressible and irrotational fields

We will say that a vector field $\mathbf{u}$ is *incompressible* (or *solenoidal*) if $\mathrm{div}(\mathbf{u}) = 0$, and that it is *irrotational* (or *conservative*) if $\mathrm{curl}(\mathbf{u}) = 0$.

Example

(a) For any scalar field f (in two or three dimensions) we have a vector field $\nabla(f) = \mathrm{grad}(f)$. The rule $\mathrm{curl}(\mathrm{grad}(f)) = 0$ tells us that $\mathrm{grad}(f)$ is irrotational.

(b) For any vector field $\mathbf{v}$ in three dimensions we have another vector field $\mathrm{curl}(\mathbf{v})$. The rule $\mathrm{div}(\mathrm{curl}(\mathbf{v})) = \nabla.(\nabla \times \mathbf{v}) = 0$ tells us that $\mathrm{curl}(\mathbf{v})$ is incompressible.

Example

The two-dimensional vector field $\mathbf{u} = (x^2 - y^2 + 2xy, x^2 - y^2 - 2xy)$ has

$$\mathrm{div}(\mathbf{u}) = \frac{\partial}{\partial x}(x^2 - y^2 + 2xy) + \frac{\partial}{\partial y}(x^2 - y^2 - 2xy) = (2x + 2y) + (-2y - 2x) = 0$$

$$\mathrm{curl}(\mathbf{u}) = \frac{\partial}{\partial x}(x^2 - y^2 - 2xy) - \frac{\partial}{\partial y}(x^2 - y^2 + 2xy) = (2x - 2y) - (-2y + 2x) = 0,$$

so it is both incompressible and irrotational.

Examples of incompressible and irrotational fields

Incompressible: $\nabla.\mathbf{u} = 0$; irrotational/conservative: $\nabla \times \mathbf{u} = 0$.

Consider a vector field of the form
$\mathbf{u} = (ax + by + cz,\ dx + ey + fz,\ gx + hy + iz)$
(where $a, b, \ldots, i$ are constants). We have

$$\nabla.\mathbf{u} = (ax + by + cz)_x + (dx + ey + fz)_y + (gx + hy + iz)_z = a + e + i$$

$$\nabla \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ ax + by + cz & dx + ey + fz & gx + hy + iz \end{bmatrix} = (h - f, c - g, d - b).$$

Thus, $\mathbf{u}$ is incompressible when $a + e + i = 0$, and it is irrotational when $h = f$, $g = c$ and $d = b$. In the irrotational case, we can rewrite the equation for $\mathbf{u}$ as

$$\mathbf{u} = (ax + by + cz,\quad bx + ey + fz,\quad cx + fy + iz).$$

If we put $p = \frac{1}{2}(ax^2 + ey^2 + iz^2) + bxy + cxz + fyz$, we find that

$$p_x = ax + by + cz \qquad p_y = bx + ey + fz \qquad p_z = cx + fy + iz,$$

so

$$\nabla(p) = (ax + by + cz,\quad bx + ey + fz,\quad cx + fy + iz) = \mathbf{u}.$$

Potential functions

If $\mathbf{u}$ is an irrotational vector field, a *potential function* for $\mathbf{u}$ is a scalar field p such that $\nabla(p) = \mathbf{u}$. (Because $\mathrm{curl}(\mathrm{grad}(p)) = 0$, only irrotational fields can have a potential.) Potential functions always exist (but they may be multi-valued), and it is often useful to find them.

Example: Consider the vector field $\mathbf{u} = (y + z, z + x, x + y)$. This has

$$\nabla \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y+z & z+x & x+y \end{bmatrix} = (1 - 1, 1 - 1, 1 - 1) = 0,$$

so it is irrotational. It therefore makes sense to look for a potential function, or in other words a function $p(x, y, z)$ with $(p_x, p_y, p_z) = (y + z, z + x, x + y)$. As we want $p_x = y + z$, we must have

$$p = \int y + z\, dx = xy + xz + q(y, z).$$

We thus have $q_x = 0$, and the equation $p_y = z + x$ becomes $x + q_y = z + x$, or in other words $q_y = z$. Integrating this, we get

$$q = \int z\, dy = yz + r(z).$$

We now have $p = xy + xz + q = xy + xz + yz + r$, so the equation $p_z = x + y$ becomes $x + y + r_z = x + y$, so $r_z = 0$. As r can only depend on z and we have $r_z = 0$ we see that r is a genuine constant. We can choose it to be zero, and we find that the function $p = xy + xz + yz$ is a potential function for $\mathbf{u}$.

A vector field with no potential function

Consider the vector field $\mathbf{u} = (0, 0, x^2)$. This has

$$\nabla \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & x^2 \end{bmatrix} = (0, -2x, 0) \neq 0,$$

so it is not irrotational, so it cannot have a potential function. We will nonetheless try to find one, and see what goes wrong. A potential function p would have to have $(p_x, p_y, p_z) = (0, 0, x^2)$. As $p_x = p_y = 0$, we see that p can only depend on z. That means that the derivative p_z also depends only on z, so we cannot have $p_z = x^2$. Thus, there is no potential function.

Another potential example

Consider again the two-dimensional vector field

$$\mathbf{u} = (x^2 - y^2 + 2xy,\quad x^2 - y^2 - 2xy).$$

We saw earlier that this is irrotational, so it has a potential function p, satisfying $p_x = x^2 - y^2 + 2xy$ and $p_y = x^2 - y^2 - 2xy$. Integrating the first of these gives

$$p = \int x^2 - y^2 + 2xy\, dx = \tfrac{1}{3}x^3 - xy^2 + x^2y + q,$$

where q depends only on y. This gives $p_y = -2xy + x^2 + q_y$, but p_y is supposed to be equal to $x^2 - y^2 - 2xy$, so we must have $q_y = -y^2$, which gives $q = -\frac{1}{3}y^3$ (plus a constant, which we may take to be zero). Altogether this gives

$$p = \tfrac{1}{3}x^3 - xy^2 + x^2y - \tfrac{1}{3}y^3.$$

Another potential example

Consider the vector field $\mathbf{u} = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right)$. We have

$$\nabla \times \mathbf{u} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \end{bmatrix} = \left(0, 0, \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) + \frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right)\right)$$

The relevant partial derivatives are

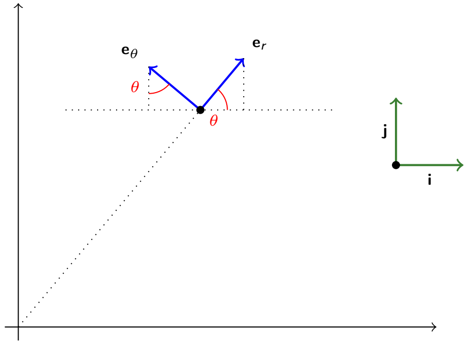
$$\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{1.(x^2+y^2) - x.2x}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

$$\frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right) = \frac{1.(x^2+y^2) - y.2y}{(x^2+y^2)^2} = \frac{x^2-y^2}{(x^2+y^2)^2},$$

and when we add these together we get zero. This means that $\nabla \times \mathbf{u} = 0$, so $\mathbf{u}$ is irrotational. It therefore makes sense to look for a potential function p, which must satisfy $\nabla(p) = (p_x, p_y, p_z) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right)$.
Looking back to an earlier example, we see that the required function is $p = \theta = \arctan(y/x)$. This is most naturally thought of as a multivalued function: for example, the value at $(-1, 0, 0)$ could be any odd multiple of π. This is bound up with the fact that $\mathbf{u}$ is not well-defined on the z-axis (where the formula $x/(x^2+y^2)$ involves division by zero). There is much more that could be said about this kind of phenomenon (with applications to magnetic fields around superconductors, for example) but we will not explore that here.

Two dimensions

At any point in the plane, we can define vectors $\mathbf{r}_r$ and $\mathbf{e}_\theta$ as shown:



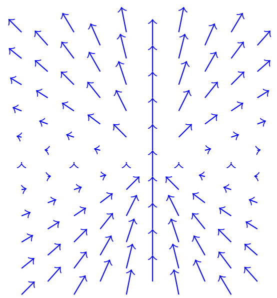
In situations with circular symmetry, it is often more natural to describe vector fields in terms of $\mathbf{e}_r$ and $\mathbf{e}_\theta$ rather than $\mathbf{i}$ and $\mathbf{j}$. One can translate between the two descriptions as follows:

$$\mathbf{e}_r = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j} \qquad \mathbf{e}_\theta = -\sin(\theta)\mathbf{i} + \cos(\theta)\mathbf{j}$$

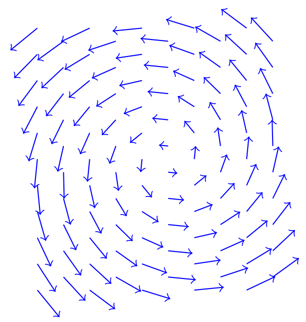
$$\mathbf{i} = \cos(\theta)\mathbf{e}_r - \sin(\theta)\mathbf{e}_\theta \qquad \mathbf{j} = \sin(\theta)\mathbf{e}_r + \cos(\theta)\mathbf{e}_\theta.$$

Examples

Here are two examples of vector fields described in terms of $\mathbf{e}_r$ and $\mathbf{e}_\theta$:



$\mathbf{u} = \sin(\theta)\mathbf{e}_r$



$\mathbf{u} = \sqrt{r}(\mathbf{e}_\theta + \mathbf{e}_r/10)$

Div, grad and curl in polar coordinates

We will need to express the operators grad, div and curl in terms of polar coordinates.

(a) For any two-dimensional scalar field f (expressed as a function of r and θ) we have

$$\nabla(f) = \mathrm{grad}(f) = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta.$$

(b) For any 2-dimensional vector field $\mathbf{u} = m\,\mathbf{e}_r + p\,\mathbf{e}_\theta$ (where m and p are expressed as functions of r and θ) we have

$$\mathrm{div}(\mathbf{u}) = r^{-1}m + m_r + r^{-1}p_\theta = r^{-1}\left((rm)_r + p_\theta\right)$$

$$\mathrm{curl}(\mathbf{u}) = r^{-1}p + p_r - r^{-1}m_\theta = r^{-1}\left((rp)_r - m_\theta\right) = \frac{1}{r}\det\begin{bmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} \\ m & rp \end{bmatrix}.$$

Note that the product rule gives $(rm)_r = m + r\,m_r$ and $(rp)_r = p + r\,p_r$.

(c) For any two-dimensional scalar field f we have

$$\nabla^2(f) = r^{-1}f_r + f_{rr} + r^{-2}f_{\theta\theta} = r^{-1}(rf_r)_r + r^{-2}f_{\theta\theta}$$

Note: in the exam, if you need these formulae, they will be provided.

Grad in polar coordinates

For any two-dimensional scalar field f (as a function of r and θ) we have

$$\nabla(f) = \operatorname{grad}(f) = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta.$$

Justification: Consider the field $\mathbf{u} = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta$; we show that this is the same as $\operatorname{grad}(f)$. Two-variable chain rule: suppose we make a small change δr to r. This causes a change $\delta x \simeq x_r\,\delta r$ to x, which in turn causes a change $\simeq f_x\,\delta x \simeq f_x\,x_r\,\delta r$ to f. At the same time, our change in r also causes a change $\delta y \simeq y_r\,\delta r$ to x, which causes a change $\simeq f_y\,\delta y = f_y\,y_r\,\delta r$ to f. Altogether, the change in f is $\delta f \simeq (f_x x_r + f_y y_r)\delta r$. By passing to the limit $\delta r \to 0$, we get $f_r = f_x x_r + f_y y_r$. Similarly, $f_\theta = f_x x_\theta + f_y y_\theta$. Moreover, we can differentiate the formulae

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

to get

$$x_r = \cos(\theta) \qquad y_r = \sin(\theta)$$
$$x_\theta = -r\sin(\theta) \qquad y_\theta = r\cos(\theta), \text{ so}$$

$$f_r = f_x x_r + f_y y_r = \cos(\theta) f_x + \sin(\theta) f_y$$
$$f_\theta = f_x x_\theta + f_y y_\theta = -r\sin(\theta) f_x + r\cos(\theta) f_y$$
$$\begin{aligned}\mathbf{u} &= f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta = f_x\cos(\theta)\mathbf{e}_r + f_y\sin(\theta)\mathbf{e}_r - f_x\sin(\theta)\mathbf{e}_\theta + f_y\cos(\theta)\mathbf{e}_\theta \\ &= f_x\,(\cos(\theta)\mathbf{e}_r - \sin(\theta)\mathbf{e}_\theta) + f_y\,(\sin(\theta)\mathbf{e}_r + \cos(\theta)\mathbf{e}_\theta) = f_x\mathbf{i} + f_y\mathbf{j} = \operatorname{grad}(f).\end{aligned}$$

Examples of polar div, grad and curl

Example: Consider $f = r^n$. Clearly $f_r = nr^{n-1}$ and $f_\theta = 0$, so

$$\operatorname{grad}(f) = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta = nr^{n-1}\mathbf{e}_r.$$

Note also that $\mathbf{r} = (x, y) = (r\cos(\theta), r\sin(\theta)) = r\,\mathbf{e}_r$, so $\mathbf{e}_r = \mathbf{r}/r$, so we can rewrite as $\operatorname{grad}(r^n) = nr^{n-2}\mathbf{r}$. (Obtained earlier using rectangular coordinates.)

Example: Consider $f = \theta$. Clearly $f_r = 0$ and $f_\theta = 1$, so

$$\operatorname{grad}(f) = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta = r^{-1}\mathbf{e}_\theta = r^{-2}(-r\sin(\theta), r\cos(\theta)) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right).$$

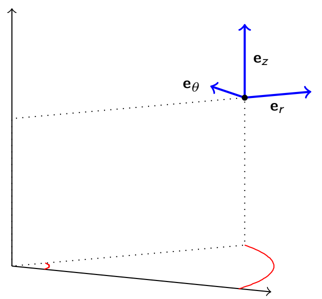
(Obtained earlier using rectangular coordinates.)

Example: Consider $\mathbf{u} = \sqrt{r}(\mathbf{e}_\theta + \mathbf{e}_r/10)$ from the plot above. This is $\mathbf{u} = p\mathbf{e}_r + q\mathbf{e}_\theta$ where $p = r^{\frac{1}{2}}/10$ and $q = r^{\frac{1}{2}}$, so $p_\theta = q_\theta = 0$ and $p_r = r^{-\frac{1}{2}}/20$ and $q_r = r^{-\frac{1}{2}}/2$. It follows that

$$\operatorname{div}(\mathbf{u}) = r^{-1}p + p_r + r^{-1}q_\theta = r^{-1}r^{\frac{1}{2}}/10 + r^{-\frac{1}{2}}/20 + 0 = 3r^{-\frac{1}{2}}/20$$
$$\operatorname{curl}(\mathbf{u}) = r^{-1}q + q_r - r^{-1}p_\theta = r^{-1}r^{-\frac{1}{2}} + r^{-\frac{1}{2}}/2 - 0 = 3r^{-\frac{1}{2}}/2.$$

Cylindrical polar coordinates

In cylindrical polar coordinates we use unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_z$ as shown below:



Thus, $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are the same as for two-dimensional polar coordinates, and $\mathbf{e}_z$ is just the vertical unit vector $\mathbf{k}$. The equations are:

$$\mathbf{e}_r = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j} \qquad \mathbf{e}_\theta = -\sin(\theta)\mathbf{i} + \cos(\theta)\mathbf{j} \qquad \mathbf{e}_z = \mathbf{k}$$
$$\mathbf{i} = \cos(\theta)\mathbf{e}_r - \sin(\theta)\mathbf{e}_\theta \qquad \mathbf{j} = \sin(\theta)\mathbf{e}_r + \cos(\theta)\mathbf{e}_\theta \qquad \mathbf{k} = \mathbf{e}_z.$$

Div, grad and curl in cylindrical polar coordinates

The rules for div, grad and curl are as follows:

(a) For any three-dimensional scalar field f (expressed as a function of r, θ and z) we have

$$\nabla(f) = \operatorname{grad}(f) = f_r\,\mathbf{e}_r + r^{-1} f_\theta\,\mathbf{e}_\theta + f_z\mathbf{e}_z.$$

(b) For any three-dimensional vector field $\mathbf{u} = m\,\mathbf{e}_r + p\,\mathbf{e}_\theta + q\,e_z$ (where m, p and q are expressed as functions of r, θ and z) we have

$$\operatorname{div}(\mathbf{u}) = r^{-1}m + m_r + r^{-1}p_\theta + q_z = r^{-1}(rm)_r + r^{-1}p_\theta + q_z$$
$$\operatorname{curl}(\mathbf{u}) = \frac{1}{r}\det\begin{bmatrix}\mathbf{e}_r & r\mathbf{e}_\theta & \mathbf{e}_z \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ m & rp & q\end{bmatrix}.$$

(c) For any three-dimensional scalar field f we have

$$\nabla^2(f) = r^{-1}f_r + f_{rr} + r^{-2}f_{\theta\theta} + f_{zz} = r^{-1}(rf_r)_r + r^{-2}f_{\theta\theta} + f_{zz}.$$

Example of curl in cylindrical polar coordinates

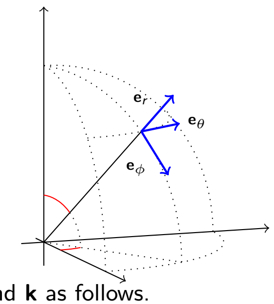
Consider the vector field $\mathbf{u}$ given in cylindrical polar coordinates by $\mathbf{u} = r(\mathbf{e}_\theta + \mathbf{e}_z)$. This is $\mathbf{u} = m\mathbf{e}_r + p\mathbf{e}_\theta + q\mathbf{e}_z$, where $m = 0$ and $p = q = r$, so

$$
\begin{aligned}
&\operatorname{curl}(\mathbf{u}) \\
&= \frac{1}{r}\det\begin{bmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & \mathbf{e}_z \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial z} \\ 0 & r^2 & r \end{bmatrix} \\
&= \frac{1}{r}\left(\left(\frac{\partial}{\partial\theta}(r) - \frac{\partial}{\partial z}(r^2)\right)\mathbf{e}_r - \left(\frac{\partial}{\partial r}(r) - \frac{\partial}{\partial z}(0)\right) r\mathbf{e}_\theta + \left(\frac{\partial}{\partial r}(r^2) - \frac{\partial}{\partial\theta}(0)\right)\mathbf{e}_z\right) \\
&= \frac{1}{r}\left(-r\mathbf{e}_\theta + 2r\mathbf{e}_z\right) = 2\mathbf{e}_z - \mathbf{e}_\theta.
\end{aligned}
$$

Spherical polar coordinates

In spherical polar coordinates we use unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ as on the right:

Note that $\mathbf{e}_\theta$ has the same meaning as it did in the cylindrical case, but $\mathbf{e}_r$ has changed. It used to be the unit vector pointing horizontally away from the z-axis, but now it points directly away from the origin.



The vectors $\mathbf{e}_r$, $\mathbf{e}_\phi$ and $\mathbf{e}_\theta$ are related to $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as follows.

$$
\begin{aligned}
\mathbf{e}_r &= \sin(\phi)\cos(\theta)\mathbf{i} + \sin(\phi)\sin(\theta)\mathbf{j} + \cos(\phi)\mathbf{k} \\
\mathbf{e}_\phi &= \cos(\phi)\cos(\theta)\mathbf{i} + \cos(\phi)\sin(\theta)\mathbf{j} - \sin(\phi)\mathbf{k} \\
\mathbf{e}_\theta &= -\sin(\theta)\mathbf{i} + \cos(\theta)\mathbf{j} \\
\mathbf{i} &= \sin(\phi)\cos(\theta)\mathbf{e}_r + \cos(\phi)\cos(\theta)\mathbf{e}_\phi - \sin(\theta)\mathbf{e}_\theta \\
\mathbf{j} &= \sin(\phi)\sin(\theta)\mathbf{e}_r + \cos(\phi)\sin(\theta)\mathbf{e}_\phi + \cos(\theta)\mathbf{e}_\theta \\
\mathbf{k} &= \cos(\phi)\mathbf{e}_r - \sin(\phi)\mathbf{e}_\phi.
\end{aligned}
$$

Div, grad and curl in spherical polar coordinates

The rules for div, grad and curl in spherical polar coordinates are as follows.

(a) For any three-dimensional scalar field f (expressed as a function of r, ϕ and θ) we have

$$\nabla(f) = \operatorname{grad}(f) = f_r\,\mathbf{e}_r + r^{-1}f_\phi\,\mathbf{e}_\phi + (r\,\sin(\phi))^{-1}f_\theta\mathbf{e}_\theta.$$

(b) For any three-dimensional vector field $\mathbf{u} = m\,\mathbf{e}_r + p\,\mathbf{e}_\phi + q\,e_\theta$ (where m, p and q are expressed as functions of r, ϕ and θ) we have

$$\operatorname{div}(\mathbf{u}) = r^{-2}(r^2m)_r + (r\sin(\phi))^{-1}(\sin(\phi)p)_\phi + (r\sin(\phi))^{-1}q_\theta$$

$$\operatorname{curl}(\mathbf{u}) = \frac{1}{r^2\sin(\phi)}\det\begin{bmatrix} \mathbf{e}_r & r\mathbf{e}_\phi & r\sin(\phi)\mathbf{e}_\theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\phi} & \frac{\partial}{\partial\theta} \\ m & rp & r\sin(\phi)q \end{bmatrix}.$$

(c) For any three-dimensional scalar field f we have

$$\nabla^2(f) = r^{-2}(r^2f_r)_r + (r^2\sin(\phi))^{-1}(\sin(\phi)f_\phi)_\phi + (r^2\sin^2(\phi))^{-1}f_{\theta\theta}.$$

Example of div, grad and curl in spherical polar coordinates

Potential of a point charge at the origin is $V = A/r$, (A constant, $r = \sqrt{x^2+y^2+z^2}$). The electric field is $\mathbf{E} = \operatorname{grad}(V)$. No magnetism or other charges, so Maxwell says $\operatorname{div}(\mathbf{E}) = 0$ and $\operatorname{curl}(\mathbf{E}) = 0$. We will check this. First, we have $V_r = -A/r^2$ and $V_\phi = V_\theta = 0$, so the rule

$$\operatorname{grad}(V) = V_r\,\mathbf{e}_r + r^{-1}V_\phi\,\mathbf{e}_\phi + (r\,\sin(\phi))^{-1}V_\theta\mathbf{e}_\theta$$

just gives $\mathbf{E} = \operatorname{grad}(V) = -Ar^{-2}\mathbf{e}_r$. In other words, we have $\mathbf{E} = m\mathbf{e}_r + p\mathbf{e}_\phi + q\mathbf{e}_\theta$ with $m = -Ar^{-2}$ and $p = q = 0$. The general rule for the divergence is

$$\operatorname{div}(\mathbf{E}) = r^{-2}(r^2m)_r + (r\sin(\phi))^{-1}(\sin(\phi)p)_\phi + (r\sin(\phi))^{-1}q_\theta.$$

As $p = q = 0$, the second and third terms are zero. In the first term, we have $r^2m = -A$, which is constant, so $(r^2m)_r = 0$ as well. This means that $\operatorname{div}(\mathbf{E}) = 0$ as expected. Finally, $\operatorname{curl}(\mathbf{E})$ is

$$\frac{1}{r^2\sin(\phi)}\det\begin{bmatrix} \mathbf{e}_r & r\mathbf{e}_\phi & r\sin(\phi)\mathbf{e}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\phi} & \frac{\partial}{\partial\theta} \\ m & rp & r\sin(\phi)q \end{bmatrix} = \frac{1}{r^2\sin(\phi)}\det\begin{bmatrix} \mathbf{e}_r & r\mathbf{e}_\phi & r\sin(\phi)\mathbf{e}_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\phi} & \frac{\partial}{\partial\theta} \\ -Ar^{-2} & 0 & 0 \end{bmatrix}.$$

As $\frac{\partial}{\partial\phi}(-Ar^{-2}) = \frac{\partial}{\partial\theta}(-Ar^{-2}) = 0$, all terms vanish so $\operatorname{curl}(\mathbf{E}) = 0$ as well.

Curves

Often we need to deal with curves in three-dimensional space. For example:

(a) A wire in an electrical machine is a curve. To calculate the magnetic field created by a current in the wire, or the force exerted on the wire by an externally applied magnetic field, we need equations for the curve.

(b) The path of a moving particle over time defines a curve. If the particle is charged then it will feel a force from any electric or magnetic fields; to understand the effect of this, we need various equations relating the position, velocity, force and acceleration to the fields.

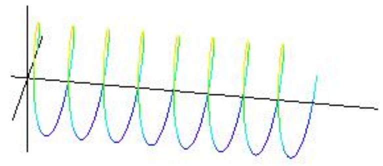
We can describe a curve by giving the x, y and z coordinates (or equivalently, the position vector $\mathbf{r} = (x, y, z)$) in terms of another parameter t. (In the case of a moving particle we often take t to be time, but that is not compulsory.)

Helix

The equation

$$\mathbf{r} = (x, y, z) = (at,\ b\cos(t),\ b\sin(t))$$

describes a helix winding around the x-axis.

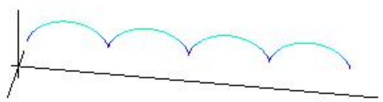


This is the path followed by an electron moving in a uniform magnetic field. It could also describe a wire wound round a cylinder.

Cycloid

Suppose that a car with axles of length a and wheels of radius b drives at constant speed c along the x-axis. A pebble stuck in the front left tyre will move along the curve with equation

$$\begin{aligned}\mathbf{r} &= (ct, a/2, b) - b(\sin(ct/b), 0, \cos(ct/b)) \\ &= (ct - b\sin(ct/b), a/2, b - b\cos(ct/b)).\end{aligned}$$



The first term $(ct, a/2, b)$ reflects the overall motion of the car, and the second term comes from the rotation of the wheel.

Projectile

A thrown ball will follow a parabolic path like

$$\mathbf{r} = (at, bt, ct - dt^2)$$

for some constants $a, \ldots, d$.

Integration along curves

To integrate along a curve C, we divide C into many small pieces, each running from some position $\mathbf{r}$ to a nearby position $\mathbf{r} + \delta\mathbf{r}$. Each such piece will give a contribution to the integral, and we add up the contributions to get an approximation to the required value. For the exact value, we pass to the limit where the length of the small pieces tends to zero.

(a) The length of the curve is approximately the sum of the lengths $|\delta\mathbf{r}|$ over all the small pieces. The exact length is denoted by $\int_C |d\mathbf{r}|$.

(b) The vector from the beginning to the end of the curve is the sum of the vectors $\delta\mathbf{r}$ over all the small pieces. In the limit we denote this by $\int_C d\mathbf{r}$.

(c) If a particle moves along a curve C through a force field $\mathbf{F}$, then the work done against the force is $-\int_C \mathbf{F}.d\mathbf{r}$.

For integrals of type (b) and (c), it makes a difference which direction we follow when traversing the curve: the answer we get when traversing the curve backwards will be the negative of the answer we get when traversing the curve forwards.

Calculating integrals along curves

In practice, we calculate these integrals as follows. We parametrise the curve as $\mathbf{r} = (x(t), y(t), z(t))$ for some range of values of t (say $a \leq t \leq b$), and we write $\dot{x} = dx/dt$ and so on. We then have

$$\begin{aligned} d\mathbf{r} &= \frac{d\mathbf{r}}{dt}dt = \dot{\mathbf{r}}dt = (\dot{x}\,dt,\ \dot{y}\,dt,\ \dot{z}\,dt) \\ |d\mathbf{r}| &= \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\,dt, \end{aligned}$$

so

$$\begin{aligned} \text{length}(C) &= \int_C |d\mathbf{r}| = \int_{t=a}^{b} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\,dt \\ \text{work} &= \int_C \mathbf{F}.d\mathbf{r} = \int_{t=a}^{b} \mathbf{F}.\dot{\mathbf{r}}\,dt \end{aligned}$$

and so on.

Example of integration along a curve

Let C be the curve given by

$$\mathbf{r} = (x, y, z) = (6t, 3\sqrt{2}t^2, 2t^3)$$

for $0 \leq t \leq 1$. We will calculate the length of this curve. We have

$$\begin{aligned} d\mathbf{r} &= (6, 6\sqrt{2}t, 6t^2)\,dt \\ |d\mathbf{r}| &= \sqrt{36 + 72t^2 + 36t^4}\,dt = 6\sqrt{1 + 2t^2 + t^4}\,dt = 6(1 + t^2)\,dt, \end{aligned}$$

so

$$\text{length} = \int_C |d\mathbf{r}| = \int_{t=0}^{1} 6(1 + t^2)\,dt = \Big[6t + 2t^3\Big]_{t=0}^{1} = 8.$$

Example of integration along a curve

Consider a particle moving along a path $\mathbf{r} = (x, y, z) = (t, 0, t/2)$ (for $0 \leq t \leq 1$) against a force field $\mathbf{F} = (y^2 + z^2 - 1, 0, 0)$.
(This could reasonably model the wind force in a wind tunnel of radius one centred on the x-axis.)
Note that

$$\begin{aligned} d\mathbf{r} &= (1, 0, 1/2)\,dt \\ \mathbf{F} &= (y^2 + z^2 - 1, 0, 0) = (0^2 + (t/2)^2 - 1, 0, 0) = (t^2/4 - 1, 0, 0) \\ \mathbf{F}.d\mathbf{r} &= (t^2/4 - 1)\,dt. \end{aligned}$$

The work done against the force is therefore

$$\text{work} = -\int_C \mathbf{F}.d\mathbf{r} = \int_{t=0}^{1} (1 - t^2/4)\,dt = \Big[t - t^3/12\Big]_{t=0}^{1} = \frac{11}{12}.$$

The Fundamental Theorem of Calculus

If f is a function of one variable, it is basic that $\int_{x=a}^{b} f'(x)\,dx = f(b) - f(a)$. (This is the *Fundamental Theorem of Calculus.*) Similarly:

Fact: For any curve C from $\mathbf{a}$ to $\mathbf{b}$, and any scalar field p, we have

$$\int_C \nabla(p).d\mathbf{r} = p(\mathbf{b}) - p(\mathbf{a}).$$

The reason is simple: the change in p along a short piece of the curve is approximately

$$\delta p \simeq p_x\delta x + p_y\delta y + p_z\delta z = \nabla(p).\delta\mathbf{r}.$$

If we add up these small changes, we get the overall change in p from $\mathbf{a}$ to $\mathbf{b}$. This gives:

Method: Suppose we have a curve C from $\mathbf{a}$ to $\mathbf{b}$, and we want to calculate the integral $I = \int_C \mathbf{F}.d\mathbf{r}$ for some vector field $\mathbf{F}$. Suppose that $\mathbf{F}$ is conservative (ie $\text{curl}(\mathbf{F}) = 0$). We can then find a potential function p with $\nabla(p) = \mathbf{F}$, and it will follow that $\int_C \mathbf{F}.d\mathbf{r} = p(\mathbf{b}) - p(\mathbf{a})$.
Note that in this method, *we do not need to know anything about C except where it starts and ends*. This often makes calculations much easier.

Example calculation using a potential

If C goes from $\mathbf{a}$ to $\mathbf{b}$ and $\mathbf{F} = \nabla(p)$ then

$$\int_C \mathbf{F}.d\mathbf{r} = p(\mathbf{b}) - p(\mathbf{a})$$

Let C be given by $(x,y,z) = (1-2t^2, 1, 2t^3)$ for $0 \le t \le 1$, and consider the vector field $\mathbf{F} = (-y/(x^2+y^2), x/(x^2+y^2), 0)$. It would be very unpleasant to calculate $\int_C \mathbf{F}.d\mathbf{r}$ directly. However, we saw earlier that $\mathbf{F}$ is conservative, with the polar coordinate function θ as a potential, so $\mathbf{F} = \nabla(\theta)$, so $\int_C \mathbf{F}.d\mathbf{r}$ is just the change in θ from the start of C to the end of C. At the start of C we have $t = 0$ so $(x,y,z) = (1,1,0)$, so $\theta = \pi/4$. At the end we have $t = 1$, so $(x,y,z) = (-1,1,2)$, which lies above the point $(-1,1,0)$ in the xy-plane; this means that $\theta = 3\pi/4$. This means that

$$\int_C \mathbf{F}.d\mathbf{r} = \frac{3\pi}{4} - \frac{\pi}{4} = \frac{\pi}{2}.$$

We have cheated a little bit here (although our answer is in fact correct) by ignoring the multi-valued nature of θ. This becomes important if we need to deal with curves that wind several times around the z-axis. However, we will not explore this further at the moment.

Using a simpler path

If we have trouble finding a potential function, it may be better to use the following approach:

Method: Suppose we have a curve C from $\mathbf{a}$ to $\mathbf{b}$, and we want to calculate the integral $I = \int_C \mathbf{F}.d\mathbf{r}$ for some vector field $\mathbf{F}$. Suppose that $\mathbf{F}$ is conservative. We can then find a different curve C' from $\mathbf{a}$ to $\mathbf{b}$ for which the calculation is easier, and then I will be equal to $\int_{C'} \mathbf{F}.d\mathbf{r}$.

The reason why this method works is that both $\int_C \mathbf{F}.d\mathbf{r}$ and $\int_{C'} \mathbf{F}.d\mathbf{r}$ are equal to $p(\mathbf{b}) - p(\mathbf{a})$, where p is the potential function. For this to be valid, we need to know that p exists (so we must check that $\mathbf{F}$ is conservative) but we do not actually need to find p.

Example of using a simpler path

Let C be the helical path given by $\mathbf{r} = (t, \cos(10\pi t), \sin(10\pi t))$ for $0 \le t \le 1$, which runs from $\mathbf{a} = (0,1,0)$ to $\mathbf{b} = (1,1,0)$. Let $\mathbf{F}$ be the vector field (yz, xz, xy). We would like to calculate $\int_C \mathbf{F}.d\mathbf{r}$. We first check whether $\mathbf{F}$ is conservative, by finding the curl:

$$\nabla \times \mathbf{F} = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz & xz & xy \end{bmatrix} = (x-x, y-y, z-z) = 0.$$

As $\mathbf{F}$ is conservative, we can replace C by a simpler path without changing the integral. In particular, we can use the straight line L given by $\mathbf{r} = (x,y,z) = (t,1,0)$. On L we have $x = t$ and $y = 1$ and $z = 0$ so $\mathbf{F} = (yz, xz, xy) = (0,0,t)$ and $d\mathbf{r} = (1,0,0)dt$ so $\mathbf{F}.d\mathbf{r} = 0$, so we conclude that $\int_C \mathbf{F}.d\mathbf{r} = \int_L \mathbf{F}.d\mathbf{r} = 0$.

A non-conservative example

L : straight line $\mathbf{r} = (t, 1, 0)$ from $\mathbf{a} = (0, 1, 0)$ at $t = 0$ to $\mathbf{b} = (1, 1, 0)$ at $t = 1$
C : helical path $\mathbf{r} = (t, \cos(10\pi t), \sin(10\pi t))$ (same limits)

Now consider the vector field $\mathbf{G} = (0, -z, y)$. This one has

$$\nabla \times \mathbf{G} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & -z & y \end{bmatrix} = (1 - (-1), 0 - 0, 0 - 0) = (2, 0, 0) \neq 0.$$

As $\mathbf{G}$ is not conservative, the integrals $\int_C \mathbf{G}.d\mathbf{r}$ and $\int_L \mathbf{G}.d\mathbf{r}$ need not be the same. On C we have

$$\begin{aligned} \mathbf{r} &= (x, y, z) = (t, \cos(10\pi t), \sin(10\pi t)) \\ d\mathbf{r} &= (\dot{x}, \dot{y}, \dot{z})\, dt = (1, -10\pi \sin(10\pi t), 10\pi \cos(10\pi t))\, dt \\ \mathbf{G} &= (0, -z, y) = (0, -\sin(10\pi t), \cos(10\pi t)) \\ \mathbf{G}.d\mathbf{r} &= 10\pi(\sin^2(10\pi t) + \cos^2(10\pi t))\, dt = 10\pi\, dt \\ \int_C \mathbf{G}.d\mathbf{r} &= \int_{t=0}^{1} 10\pi\, dt = 10\pi. \end{aligned}$$

A non-conservative example

L : straight line $\mathbf{r} = (t, 1, 0)$ from $\mathbf{a} = (0, 1, 0)$ at $t = 0$ to $\mathbf{b} = (1, 1, 0)$ at $t = 1$
C : helical path $\mathbf{r} = (t, \cos(10\pi t), \sin(10\pi t))$ (same limits)
$\mathbf{G} = (0, -z, y)$ $\int_C \mathbf{G}.d\mathbf{r} = 10\pi$

On L we have

$$\begin{aligned} \mathbf{r} &= (x, y, z) = (t, 1, 0) \\ d\mathbf{r} &= (\dot{x}, \dot{y}, \dot{z})\, dt = (1, 0, 0)\, dt \\ \mathbf{G} &= (0, -z, y) = (0, 0, 1) \\ \mathbf{G}.d\mathbf{r} &= 0 \\ \int_L \mathbf{G}.d\mathbf{r} &= 0. \end{aligned}$$

Thus, the integrals over C and L are different, as expected.

Finding a potential

Method: Let $\mathbf{F}$ be an conservative vector field. We can then define a potential function p for $\mathbf{F}$ by the rule

$$p(a, b, c) = \text{ the integral } \int_C \mathbf{F}.d\mathbf{r}, \text{ for any curve } C \text{ from } (0, 0, 0) \text{ to } (a, b, c)\ .$$

The answer will not depend on the choice of curve, so we can choose whichever curve makes the integral easiest. A straight line is often good, but sometimes a broken line (from $(0, 0, 0)$ to $(a, 0, 0)$ to $(a, b, 0)$ to (a, b, c), for example) is better.

Note again that this is only valid for conservative fields. Fields that are not conservative do not have a potential function.

Example of finding a potential

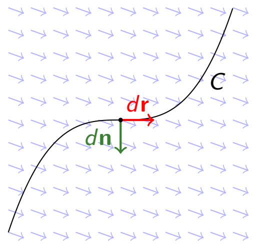
We checked previously that the vector field $\mathbf{F} = (yz, xz, xy)$ has $\text{curl}(\mathbf{F}) = 0$, so $\mathbf{F}$ is conservative. It follows that there is a potential function p with $\text{grad}(p) = \mathbf{F}$. To find $p(a, b, c)$, we evaluate $\int_L \mathbf{F}.d\mathbf{r}$, where L is the straight line from $(0, 0, 0)$ to (a, b, c). This can be parametrised by $\mathbf{r} = (x, y, z) = (ta, tb, tc)$ for $0 \leq t \leq 1$, which gives

$$\begin{aligned} d\mathbf{r} &= (a, b, c)dt \\ \mathbf{F} &= ((tb)(tc), (ta)(tc), (ta)(tb)) = (t^2bc, t^2ac, t^2ab) \\ \mathbf{F}.d\mathbf{r} &= 3t^2abc\, dt \\ p(a, b, c) &= \int_L \mathbf{F}.d\mathbf{r} = \int_{t=0}^{1} 3t^2abc\, dt = \Big[t^3abc\Big]_{t=0}^{1} = abc. \end{aligned}$$

It is convenient to write this calculation in terms of a, b and c, to avoid confusion between the end of the path (where $(x, y, z) = (a, b, c)$) and the points along the path (where $(x, y, z) = (ta, tb, tc)$). However, we can restate the final answer as $p(x, y, z) = xyz$, which is more convenient for later use.

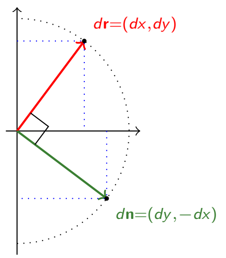
Flux across a curve

The picture shows a vector field $\mathbf{F}$ and a curve C, with the vector $d\mathbf{r}$ pointing along the curve, and another vector $d\mathbf{n}$ of the same length perpendicular to $d\mathbf{r}$.
The integral $\int_C \mathbf{F}.d\mathbf{r}$ measures the extent to which $\mathbf{F}$ points along the curve. For some purposes, however, we want to measure the flow of $\mathbf{F}$ *across* the curve, in which case we want to evaluate $\int_C \mathbf{F}.d\mathbf{n}$ rather than $\int_C \mathbf{F}.d\mathbf{r}$.

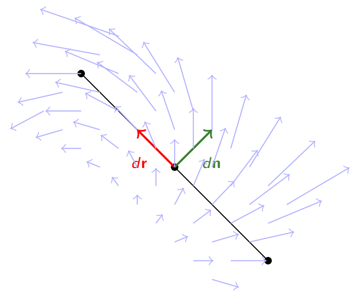


Note that $d\mathbf{r} = (dx, dy) = (\dot{x}, \dot{y})dt$, and $d\mathbf{n}$ is obtained by rotating this a quarter turn clockwise, so $d\mathbf{n} = (dy, -dx) = (\dot{y}, -\dot{x})dt$.
Thus, if $\mathbf{F} = (P, Q)$ we have

$$\int_C \mathbf{F}.d\mathbf{n} = \int_{t=\cdots}^{\cdots} (P, Q).(\dot{y}, -\dot{x})dt$$
$$= \int_{t=\cdots}^{\cdots} \dot{y}P - \dot{x}Q \; dt = \int_C (-Q, P).d\mathbf{r}$$



Example of flux across a curve



Let L be the straight line from $(1, 0)$ to $(0, 1)$, so $\mathbf{r} = (x, y) = (1 - t, t)$ for $0 \leq t \leq 1$, so $d\mathbf{r} = (-1, 1)dt$, so $d\mathbf{n} = (1, 1)dt$. Let $\mathbf{F}$ be the vector field $(x^2 - y^2, 2xy)$, so on L we have

$$\mathbf{F} = ((1-t)^2 - t^2, 2t(1-t)) = (1 - 2t + t^2 - t^2, 2t - 2t^2) = (1 - 2t, 2t - 2t^2),$$

so $\mathbf{F}.d\mathbf{n} = ((1 - 2t) + (2t - 2t^2))dt = (1 - 2t^2)dt$ so

$$\int_C \mathbf{F}.d\mathbf{n} = \int_{t=0}^{1} (1 - 2t^2)\, dt = \left[t - \tfrac{2}{3}t^3\right]_{t=0}^{1} = 1 - \tfrac{2}{3} = \tfrac{1}{3}$$

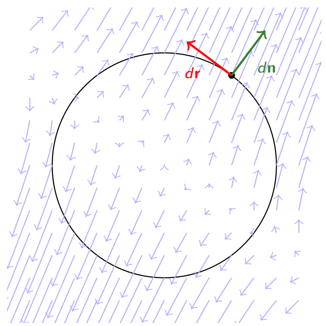
Flow out of a circle

We will calculate the flow of the field

$$\mathbf{F} = (x + 2y, 3x + 4y)$$

out of the unit circle C. We parametrise C as $\mathbf{r} = (x, y) = (\cos(t), \sin(t))$ for $0 \leq t \leq 2\pi$.
This gives

$$d\mathbf{r} = (\dot{x}, \dot{y})dt = (-\sin(t), \cos(t))\, dt$$
$$d\mathbf{n} = (\dot{y}, -\dot{x})dt = (\cos(t), \sin(t))\, dt$$
$$\mathbf{F} = (\cos(t) + 2\sin(t), 3\cos(t) + 4\sin(t))$$
$$\mathbf{F}.d\mathbf{n} = (\cos^2(t) + 5\sin(t)\cos(t) + 4\sin^2(t))dt$$



Now

$$\int_0^{2\pi} \sin(t)\cos(t)\, dt = \tfrac{1}{2}\int_0^{2\pi} \sin(2t)\, dt = 0$$
$$\int_0^{2\pi} \sin^2(t)\, dt = \int_0^{2\pi} \cos^2(t)\, dt = \pi$$

so

$$\int_C \mathbf{F}.d\mathbf{n} = \int_0^{2\pi} (\cos^2(t) + 5\sin(t)\cos(t) + 4\sin^2(t))dt = \pi + 0 + 4\pi = 5\pi.$$

Surfaces

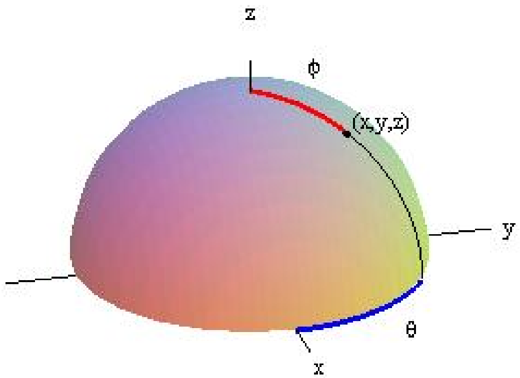
As well as considering curved paths, we also need to consider curved surfaces in three-dimensional space. Such a surface can be parametrised as $\mathbf{r} = (x(s, t), y(s, t), z(s, t))$ for some pair of parameters s and t.

A hemisphere

The upper half of a spherical shell of radius 2 can be described in terms of parameters ϕ and θ by

$$(x, y, z) = (2\sin(\phi)\cos(\theta),\ 2\sin(\phi)\sin(\theta),\ 2\cos(\phi))$$

(for $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi/2$).

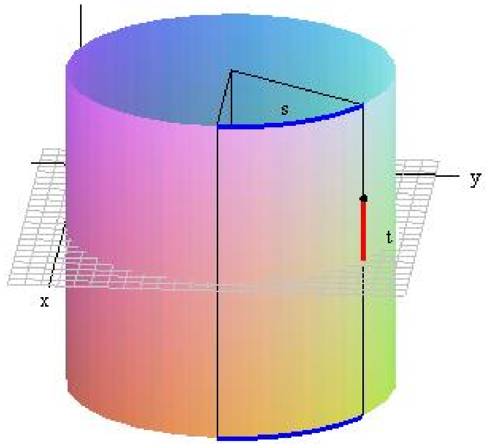


An off-centre cylinder

Let S be a cylindrical surface of radius 1, centred on the line joining $(1, 1, -1)$ to $(1, 1, 1)$. Then S can be described in terms of parameters s and t by

$$(x, y, z) = (1 + \cos(s), 1 + \sin(s), t)$$

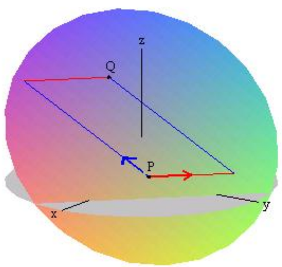
(for $0 \le s \le 2\pi$ and $-1 \le t \le 1$).



(Part of) a plane

Let S be the plane where $x + y + z = 3$. This can be parametrised in many different ways, one of which is

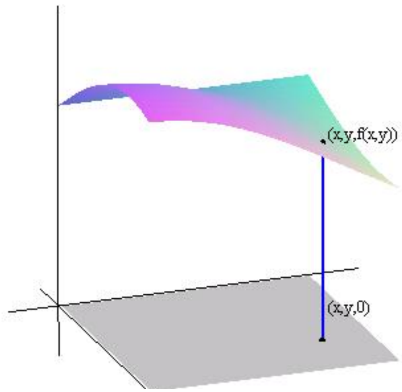
$$(x, y, z) = (1 - s, 1 + s - t, 1 + t) = (1, 1, 1) + s(-1, 1, 0) + t(0, -1, 1).$$



The picture shows the point $P = (1, 1, 1)$, which lies on S. Any other point on S (such as Q) can be reached from P by adding a multiple of the red vector $(-1, 1, 0)$ and a multiple of the blue vector $(0, -1, 1)$.

The graph of $f(x, y)$

For any function $f(x, y)$, the equation $z = f(x, y)$ defines a surface.

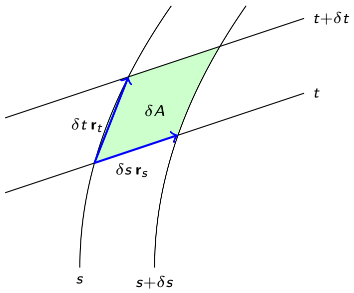


We can use the variables x and y themselves as parameters, and then the full parametrisation is

$$(x, y, z) = (x, y, f(x, y)).$$

Integration over surfaces

To integrate over S, we need a formula for the area of a small piece of S in terms of a parametrisation $\mathbf{r} = (x(s,t), y(s,t), z(s,t))$. If s and t vary by δs and δt, then the corresponding part of the surface will be a small parallelogram spanned by the vectors $\mathbf{r}_s\,\delta s = (x_s \delta s,\ y_s\,\delta s)$ and $\mathbf{r}_t\,\delta t = (x_t \delta t,\ y_t\,\delta t)$.



We write δA for the area of this parallelogram. We also write $\delta\mathbf{A}$ for the vector $(\mathbf{r}_s \times \mathbf{r}_t)\delta s\,\delta t$. This is perpendicular to $\mathbf{r}_s$ and $\mathbf{r}_t$(which means that it is normal to the surface), and $|\delta\mathbf{A}| = \delta A$. In the limit we get $d\mathbf{A} = (\mathbf{r}_s \times \mathbf{r}_t) ds\,dt$ and $dA = |d\mathbf{A}| = |\mathbf{r}_s \times \mathbf{r}_t| ds\,dt$. Also $d\mathbf{A} = \mathbf{n}\,dA$, where $\mathbf{n}$ is the unit normal to S.

Area of a hemisphere

Consider again a hemispherical shell of radius a. We have

$$\begin{aligned}
\mathbf{r} &= (a\sin(\phi)\cos(\theta),\ a\sin(\phi)\sin(\theta),\ a\cos(\phi))\\
\mathbf{r}_\phi &= (a\cos(\phi)\cos(\theta),\ a\cos(\phi)\sin(\theta),\ -a\sin(\phi))\\
\mathbf{r}_\theta &= (-a\sin(\phi)\sin(\theta),\ a\sin(\phi)\cos(\theta),\ 0)\\
\mathbf{r}_\phi \times \mathbf{r}_\theta &= \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k}\\ a\cos(\phi)\cos(\theta) & a\cos(\phi)\sin(\theta) & -a\sin(\phi)\\ -a\sin(\phi)\sin(\theta) & a\sin(\phi)\cos(\theta) & 0\end{bmatrix}\\
&= (a^2\sin^2(\phi)\cos(\theta),\ a^2\sin^2(\phi)\sin(\theta),\ a^2\sin(\phi)\cos(\phi))\\
&= a^2\sin(\phi)\mathbf{e}_r\\
d\mathbf{A} &= a^2\sin(\phi)\mathbf{e}_r\,d\phi\,d\theta\\
dA &= |d\mathbf{A}| = a^2\sin(\phi)d\theta\,d\phi.
\end{aligned}$$

It follows that the area of the surface is

$$\begin{aligned}
A &= \iint_S 1\,dA = \int_{\theta=0}^{2\pi}\int_{\phi=0}^{\frac{\pi}{2}} a^2\sin(\phi)d\theta\,d\phi\\
&= 2a^2\pi\int_{\phi=0}^{\frac{\pi}{2}}\sin(\phi)\,d\phi = 2a^2\pi\Big[-\cos(\phi)\Big]_{\phi=0}^{\frac{\pi}{2}} = 2a^2\pi.
\end{aligned}$$

Area of a cylinder

Consider a cylindrical surface as before. We have

$$\begin{aligned}
\mathbf{r} &= (1+\cos(s), 1+\sin(s), t) \qquad\qquad (0 \le s \le 2\pi,\ -1 \le t \le 1)\\
\mathbf{r}_s &= (-\sin(s), \cos(s), 0)\\
\mathbf{r}_t &= (0,0,1)\\
\mathbf{r}_s \times \mathbf{r}_t &= \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k}\\ -\sin(s) & \cos(s) & 0\\ 0 & 0 & 1\end{bmatrix} = (\cos(s), \sin(s), 0)\\
d\mathbf{A} &= (\cos(s),\sin(s),0)\,ds\,dt\\
|\mathbf{r}_s \times \mathbf{r}_t| &= |(\cos(s),\sin(s),0)| = \sqrt{\cos^2(s)+\sin^2(s)} = 1\\
dA &= |\mathbf{r}_s \times \mathbf{r}_t|\,ds\,dt = ds\,dt.
\end{aligned}$$

It follows that the area of the surface is

$$\iint_S 1\,dA = \int_{s=0}^{2\pi}\int_{t=-1}^{1} 1\,ds\,dt = 2\pi(1-(-1)) = 4\pi.$$

The area element for $z = f(x,y)$

Consider a surface of the form $z = f(x,y)$. We have

$$\begin{aligned}
\mathbf{r} &= (x, y, f(x,y))\\
\mathbf{r}_x &= (1, 0, f_x)\\
\mathbf{r}_y &= (0, 1, f_y)\\
\mathbf{r}_x \times \mathbf{r}_y &= \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k}\\ 1 & 0 & f_x\\ 0 & 1 & f_y\end{bmatrix} = (-f_x, -f_y, 1)\\
d\mathbf{A} &= (\mathbf{r}_x \times \mathbf{r}_y)\,dx\,dy = (-f_x,-f_y,1)\,dx\,dy\\
|\mathbf{r}_x \times \mathbf{r}_y| &= \sqrt{f_x^2+f_y^2+1}\\
dA &= \sqrt{f_x^2+f_y^2+1}\,dx\,dy.
\end{aligned}$$

The area of the surface $z = \cosh(x+y)/\sqrt{2}$

If $z = f(x,y)$ then $d\mathbf{A} = (-f_x, -f_y, 1)dx\,dy$ and $dA = \sqrt{1+f_x^2+f_y^2}\,dx\,dy$.

Now consider the case where $z = f(x,y) = \cosh(x+y)/\sqrt{2}$ for $0 \le x, y \le 1$.

$$f_x = \sinh(x+y)/\sqrt{2} \qquad f_y = \sinh(x+y)/\sqrt{2}$$

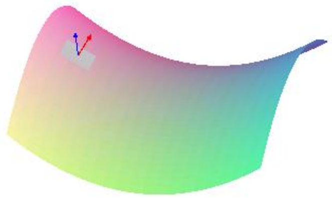
$$\begin{aligned}
\sqrt{1+f_x^2+f_y^2} &= \sqrt{1+\tfrac{1}{2}\sinh^2(x+y)+\tfrac{1}{2}\sinh^2(x+y)} = \sqrt{1+\sinh^2(x+y)} \\
&= \sqrt{\cosh^2(x+y)} = \cosh(x+y) \\
dA &= \sqrt{1+f_x^2+f_y^2}\,dx\,dy = \cosh(x+y)\,dx\,dy.
\end{aligned}$$

It follows that the area of the surface is

$$\begin{aligned}
A &= \iint_S 1\,dA = \int_{x=0}^{1}\int_{y=0}^{1}\cosh(x+y)dy\,dx \\
&= \int_{x=0}^{1}\Big[\sinh(x+y)\Big]_{y=0}^{1}dx = \int_{x=0}^{1}\sinh(x+1)-\sinh(x)\,dx \\
&= \Big[\cosh(x+1)-\cosh(x)\Big]_{x=0}^{1} = (\cosh(2)-\cosh(1))-(\cosh(1)-\cosh(0)) \\
&= \cosh(2)-2\cosh(1)+1 \simeq 1.676
\end{aligned}$$

Flow across a surface

Consider a surface S in a region where there is a vector field $\mathbf{F}$.



We want to calculate the flow of $\mathbf{F}$ across S. At each point on S there is a normal vector $\mathbf{n}$. The flow across S only involves the component of $\mathbf{F}$ in the normal direction, or in other words $\mathbf{F}.\mathbf{n}$. We need to integrate this with respect to area, giving $\iint_S \mathbf{F}.\mathbf{n}dA$. However, $\mathbf{n}dA$ is the same as the vector $d\mathbf{A}$ considered earlier, which can be calculated from a parametrisation by the rule $d\mathbf{A} = (\mathbf{r}_s \times \mathbf{r}_t)ds\,dt$. Note that if $\mathbf{F} = (P, Q, R)$ then the scalar triple product $\mathbf{F}.d\mathbf{A} = \mathbf{F}.(\mathbf{r}_s \times \mathbf{r}_t)ds\,dt$ can be expressed as a determinant:

$$\mathbf{F}.d\mathbf{A} = \det\begin{bmatrix} P & Q & R \\ x_s & y_s & z_s \\ x_t & y_t & z_t \end{bmatrix} ds\,dt.$$

Example

Let S be the surface given by $z = f(x,y) = xy$ for $0 \le x, y \le 1$, and let $\mathbf{F}$ be the vector field $(x+y+z, x+y+z, x+y+z)$.

$$\begin{aligned}
\mathbf{r} &= (x, y, xy) \\
\mathbf{r}_x \times \mathbf{r}_y &= (-f_x, -f_y, 1) = (-y, -x, 1) \\
d\mathbf{A} &= (-y, -x, 1)\,dx\,dy \\
\mathbf{F} &= (x+y+xy, x+y+xy, x+y+xy) \\
\mathbf{F}.d\mathbf{A} &= (-y(x+y+xy)-x(x+y+xy)+(x+y+xy))\,dx\,dy \\
&= \left(x+y-x^2-y^2-xy-xy^2-x^2y\right)\,dx\,dy \\
\iint_S \mathbf{F}.d\mathbf{A} &= \int_{x=0}^{1}\int_{y=0}^{1}\left(x+y-x^2-y^2-xy-xy^2-x^2y\right)\,dy\,dx \\
&= \int_{x=0}^{1}\left(x+\tfrac{1}{2}-x^2-\tfrac{1}{3}-\tfrac{1}{2}x-\tfrac{1}{3}x-\tfrac{1}{2}x^2\right)\,dx \\
&= \int_{x=0}^{1}\left(\tfrac{1}{6}+\tfrac{1}{6}x-\tfrac{3}{2}x^2\right)\,dx \\
&= \tfrac{1}{6}+\tfrac{1}{6}.\tfrac{1}{2}-\tfrac{3}{2}.\tfrac{1}{3} = \tfrac{2}{12}+\tfrac{1}{12}-\tfrac{6}{12} \\
&= -\tfrac{1}{4}.
\end{aligned}$$

Integral theorems — Introduction

Some important facts about electromagnetism are as follows:

(a) For any three-dimensional region, the total electric field crossing the boundary of the region is ϵ_0^{-1} times the total charge in the region.

(b) On the other hand, the magnetic field crossing the boundary always cancels out to give a total of zero.

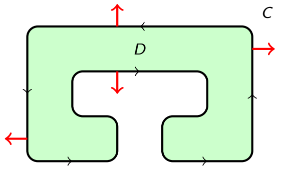
(c) Now suppose we have a surface S in three-dimensional space. Suppose that has a boundary that is a closed curve C (so the surface could be a disk or a hemispherical bowl, but not a complete sphere). Then the circulation of $\mathbf{E}$ around C is minus the rate of change of the total magnetic field passing through S.

(d) Similarly, the circulation of $\mathbf{B}$ around C is μ_0 times the rate of change of the current passing through S (including the "displacement current" $\epsilon_0\dot{\mathbf{E}}$).

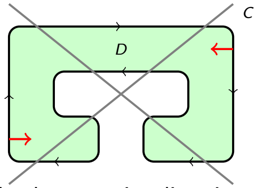
These are not really new physical facts; they are mathematically equivalent to Maxwell's equations. Maxwell's equations told us about the values of scalar and vector fields and their derivatives at every point in space. The above statements are about various kinds of integrals of such scalar and vector fields over curves, surfaces and three-dimensional regions. The main point of this final section of the course is to understand why these integral statements are the same as the earlier differential statements.

The sign convention for closed curves

Let D be a region in the plane. The edge of the region will be a curve, which we call C. For any vector field $\mathbf{u}$, we can consider the integral $\int_C \mathbf{u}.d\mathbf{n}$ measuring the flux of $\mathbf{u}$ across C. This kind of integral depends on the direction in which we traverse the curve. We will always traverse in the direction which keeps the region D on our left. This means that we are basically going anticlockwise, although it may not always seem that way if C has a complicated shape.



In this direction
we keep the region on the left
so $d\mathbf{n}$ points outwards



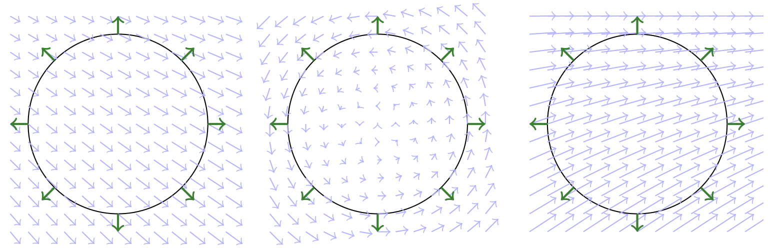
In the opposite direction
we keep the region on the right
so $d\mathbf{n}$ points inwards

The two-dimensional divergence theorem

Let D be a region in the plane whose boundary is a closed curve C. The two-dimensional divergence theorem says that for any vector field $\mathbf{u}$ that is well-behaved everywhere in D, we have

$$\iint_D \mathrm{div}(\mathbf{u})\, dA = \int_C \mathbf{u}.d\mathbf{n}.$$

Here "well-behaved" means that there are no discontinuous jumps (as with a square wave) or kinks (as with a sawtooth). Functions like $1/(x^2+y^2)$ (which blows up to infinity at the origin) are allowed if the origin lies outside D, but disallowed if the origin is inside D.



$\mathrm{div}(\mathbf{u}) > 0$ $\quad$ $\mathrm{div}(\mathbf{u}) = 0$ $\quad$ $\mathrm{div}(\mathbf{u}) < 0$

The two-dimensional divergence theorem

Claim: $\displaystyle\iint_D \mathrm{div}(\mathbf{u})\, dA = \int_C \mathbf{u}.d\mathbf{n}$ $\qquad$ ($C =$ boundary of D, anticlockwise)

Let $\mathbf{u}$ be (p,q). We then have

$$\mathrm{div}(\mathbf{u}) = p_x + q_y$$
$$\mathbf{u}.d\mathbf{n} = (p,q).(dy,-dx) = p\,dy - q\,dx.$$

The claim is thus that

$$\iint_D (p_x + q_y)\, dA = \int_C (p\,dy - q\,dx).$$

It will be enough to show that

$$\iint_D q_y\, dA = -\int_C q\,dx \qquad \text{(A)}$$

$$\iint_D p_x\, dA = \int_C p\,dy \qquad \text{(B)}$$

The two-dimensional divergence theorem

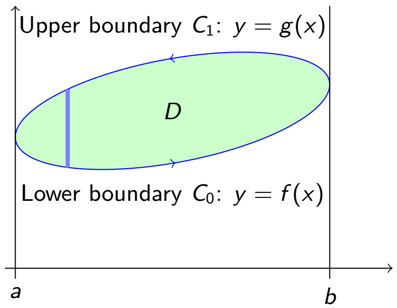
Claim (A): $\displaystyle\iint_D q_y\, dA = -\int_C q\,dx$ $\qquad$ ($C =$ boundary of D, anticlockwise)

$\iint_D q_y\,dA = \int_{x=a}^b \int_{y=f(x)}^{g(x)} q_y(x,y)\,dy\,dx$
$= \int_{x=a}^b \left[q(x,y)\right]_{y=f(x)}^{g(x)}\,dx$
$= \int_{x=a}^b (q(x,g(x)) - q(x,f(x)))\,dx$ (A).

On C_0 we have $y = f(x)$ so
$-\int_{C_0} q\,dx = -\int_{x=a}^b q(x,f(x))\,dx$. (B)

Similarly
$-\int_{C_1} q\,dx = +\int_{x=a}^b q(x,g(x))\,dx$. (C)

(Sign has changed because $\int_a^b$ goes left to right whereas C_1 goes right to left.)
Add (B) and (C) and compare with (A):

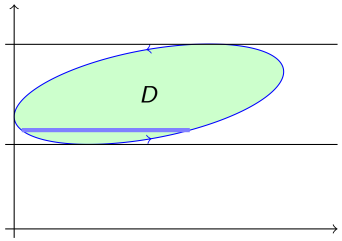


$$\iint_D q_y\,dA \overset{(A)}{=} \int_{x=a}^b q(x,g(x))\,dx - \int_{x=a}^b q(x,f(x))\,dx \overset{(B),(C)}{=} -\int_{C_1} q\,dx - \int_{C_0} q\,dx$$

$= -\int_C q\,dx$ as claimed.

The two-dimensional divergence theorem

We just proved using vertical strips that $\iint_D q_y \, dA = -\int_C q \, dx$



Similarly, with horizontal strips: $\iint_D p_x \, dA = \int_C p \, dy$

Adding these gives

$$\iint_D \mathrm{div}(\mathbf{u}) dA = \iint_D (p_x + q_y) dA = \int_C (p\,dy - q\,dx) = \int_C (p,q).(dy, -dx) = \int_C \mathbf{u}.d\mathbf{n}$$

which is the two-dimensional divergence theorem.

Example

Let D be the disc where $x^2 + y^2 \leq m^2$, so C is a circle of radius m.
Take $\mathbf{u} = (ax + by, cx + dy)$ for some constants a, b, c and d.
Then $\mathrm{div}(\mathbf{u}) = (ax + by)_x + (cx + dy)_y = a + d$, so
$\iint_D \mathrm{div}(\mathbf{u}) dA = (a + d)\mathrm{area}(D) = \pi m^2 (a + d)$.
On the other hand, we can parametrise C by $\mathbf{r} = (x, y) = (m\cos(t), m\sin(t))$,
so $d\mathbf{n} = (\dot{y}, -\dot{x})dt = (m\cos(t), m\sin(t))\,dt$. On C we also have

$$\mathbf{u} = (ax + by, cx + dy) = (am\cos(t) + bm\sin(t), cm\cos(t) + dm\sin(t))$$

so

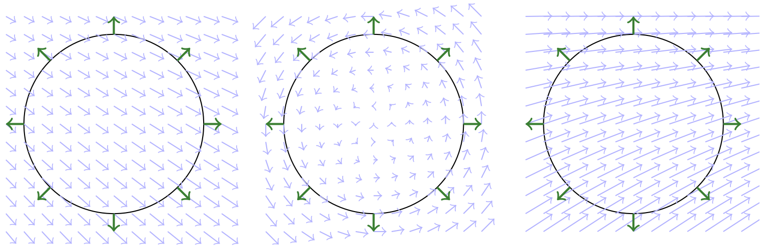
$$\begin{aligned}
\mathbf{u}.d\mathbf{n} &= (am\cos(t) + bm\sin(t))(m\cos(t))dt + \\
&\qquad (cm\cos(t) + dm\sin(t))(m\sin(t))dt \\
&= m^2(a\cos^2(t) + (b + c)\sin(t)\cos(t) + d\sin^2(t))dt \\
&= \frac{m^2}{2}(a + a\cos(2t) + (b + c)\sin(2t) + d - d\cos(2t)) \\
&= \frac{m^2}{2}((a + d) + (a - d)\cos(2t) + (b + c)\sin(2t)) \\
\int_C \mathbf{u}.d\mathbf{n} &= \frac{m^2}{2}\left[(a + d)t + \tfrac{1}{2}(a - d)\sin(2t) - \tfrac{1}{2}(b + c)\cos(2t)\right]_{t=0}^{2\pi} \\
&= \frac{m^2}{2} 2\pi(a + d) = \pi m^2(a + d).
\end{aligned}$$

The two-dimensional divergence theorem

Let D be a region in the plane whose boundary is a closed curve C. The two-dimensional divergence theorem says that for any vector field $\mathbf{u}$ that is well-behaved everywhere in D, we have

$$\iint_D \mathrm{div}(\mathbf{u}) \, dA = \int_C \mathbf{u}.d\mathbf{n}.$$

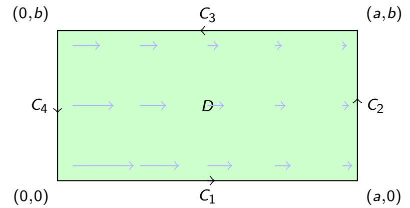
Here “well-behaved” means that there are no discontinuous jumps (as with a square wave) or kinks (as with a sawtooth). Functions like $1/(x^2 + y^2)$ (which blows up to infinity at the origin) are allowed if the origin lies outside D, but disallowed if the origin is inside D.



$\mathrm{div}(\mathbf{u}) > 0$ $\qquad$ $\mathrm{div}(\mathbf{u}) = 0$ $\qquad$ $\mathrm{div}(\mathbf{u}) < 0$

Example

Let D be the rectangle as shown below. The boundary consists of $C_1, \ldots, C_4$.

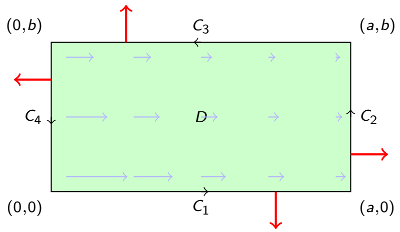


Consider the horizontal vector field $\mathbf{u} = (e^{-x-y}, 0)$. This has
$\mathrm{div}(\mathbf{u}) = \frac{\partial}{\partial x}(e^{-x-y}) = -e^{-x-y} = -e^{-x}e^{-y}$, so

$$\begin{aligned}
\iint_D \mathrm{div}(\mathbf{u}) dA &= -\int_{x=0}^{a} e^{-x}\,dx \int_{y=0}^{b} e^{-y}\,dy = -\Big[-e^{-x}\Big]_{x=0}^{a}\Big[-e^{-y}\Big]_{y=0}^{b} \\
&= -(1 - e^{-a})(1 - e^{-b}) = e^{-a} + e^{-b} - e^{-a-b} - 1.
\end{aligned}$$

Example

$\mathbf{u} = (e^{-x-y}, 0)$



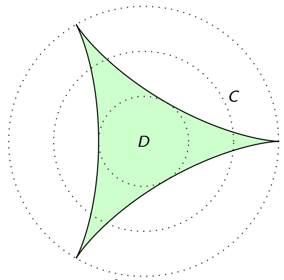
On C_1 and C_3 the normal $d\mathbf{n}$ is vertical but $\mathbf{u}$ is horizontal so $\mathbf{u}.d\mathbf{n} = 0$.
On C_2 we have $d\mathbf{n} = (1,0)dy$ and $x = a$ so $\mathbf{u} = (e^{-a-y}, 0)$ so $\mathbf{u}.d\mathbf{n} = e^{-a-y}dy$
so $\int_{C_2} \mathbf{u}.d\mathbf{n} = \int_{y=0}^{b} e^{-a-y}dy = \left[-e^{-a-y}\right]_{y=0}^{b} = e^{-a} - e^{-a-b}$
We can parametrise C_4 in the right direction by $(x,y) = (0, b-t)$ for $0 \leq t \leq b$. This gives $d\mathbf{n} = (\dot{y}, -\dot{x})dt = (-1,0)dt$ and $\mathbf{u} = (e^{-x-y}, 0) = (e^{t-b}, 0)$ so $\mathbf{u}.d\mathbf{n} = -e^{t-b}$ so
$\int_{C_4} \mathbf{u}.d\mathbf{n} = \int_{t=0}^{b} -e^{t-b}\,dt = \left[-e^{t-b}\right]_{t=0}^{b} = -1 + e^{-b}$. This gives

$$\int_C \mathbf{u}.d\mathbf{n} = (e^{-a} - e^{-a-b}) + (-1 + e^{-b}) = e^{-a} + e^{-b} - e^{-a-b} - 1 = \iint_D \mathrm{div}(\mathbf{u})dA$$

Area of a deltoid

The picture shows the deltoid curve C:

$x = 2\cos(t) + \cos(2t) \qquad y = 2\sin(t) - \sin(2t)$.



It is hard to find the area of D directly. However, we can evaluate it by a trick using the divergence theorem. Consider the vector field $\mathbf{F} = (x, 0)$, so $\mathrm{div}(\mathbf{F}) = \partial x/\partial x + \partial 0/\partial y = 1$, so $\iint_D \mathrm{div}(\mathbf{F})dA = \mathrm{area}(D)$.

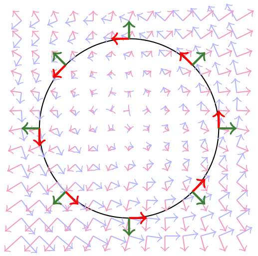
The Divergence Theorem tells us that this is the same as $\int_C \mathbf{F}.d\mathbf{n}$. Here

$$\begin{aligned}
d\mathbf{n} &= (\dot{y}, -\dot{x})\,dt = (2\cos(t) - 2\cos(2t),\ 2\sin(t) + 2\sin(2t))\,dt \\
\mathbf{F} &= (x, 0) = (2\cos(t) + \cos(2t),\ 0) \\
\mathbf{F}.d\mathbf{n} &= (2\cos(t) - 2\cos(2t))(2\cos(t) + \cos(2t)) \\
&= 4\cos^2(t) - 2\cos(t)\cos(2t) - 2\cos^2(2t) \\
&= (2 + 2\cos(2t)) - (\cos(3t) + \cos(t)) - (1 + \cos(4t)) \\
&= 1 - \cos(t) + 2\cos(2t) - \cos(3t) - \cos(4t) \\
\text{area} &= \int_{t=0}^{2\pi} \mathbf{F}.d\mathbf{n} = 2\pi
\end{aligned}$$

Green's theorem

Let D be a region in the plane whose boundary is a closed curve C. Green's theorem says that for any vector field $\mathbf{u}$ that is well-behaved everywhere in D, we have

$$\iint_D \mathrm{curl}(\mathbf{u})\,dA = \int_C \mathbf{u}.d\mathbf{r}.$$



To see this, let $\mathbf{v}$ be the field obtained by turning $\mathbf{u}$ clockwise by $\pi/2$. We can apply the divergence theorem to $\mathbf{v}$ to get $\iint_D \mathrm{div}(\mathbf{v})dA = \int_C \mathbf{v}.d\mathbf{n}$. If $\mathbf{u} = (p,q)$ then $\mathbf{v} = (q, -p)$, so $\mathrm{div}(\mathbf{v}) = q_x - p_y = \mathrm{curl}(\mathbf{u})$ and

$$\mathbf{v}.d\mathbf{n} = (q, -p).(dy, -dx) = p\,dx + q\,dy = (p,q).(dx,dy) = \mathbf{u}.d\mathbf{r}$$

so

$$\iint_D \mathrm{curl}(\mathbf{u})\,dA = \int_C \mathbf{u}.d\mathbf{r}.$$

as claimed

Example of Green's Theorem

Let D be the unit disc, so the boundary curve C is the unit circle. Let $\mathbf{u}$ be the vector field (x^3, x^3). Green's Theorem tells us that $\iint_D \mathrm{curl}(\mathbf{u})\,dA = \int_C \mathbf{u}.d\mathbf{r}$. We will check this by evaluating both sides. First, we have

$$\mathrm{curl}(\mathbf{u}) = \det\begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ x^3 & x^3 \end{bmatrix} = \frac{\partial}{\partial x}(x^3) - \frac{\partial}{\partial y}(x^3) = 3x^2 - 0 = 3x^2.$$

We will evaluate $\iint_D \mathrm{curl}(\mathbf{u})\,dA$ using polar coordinates, so $x = r\cos(\theta)$ and $dA = r\,dr\,d\theta$. This gives

$$\begin{aligned}
\iint_D \mathrm{curl}(\mathbf{u})\,dA &= \int_{\theta=0}^{2\pi}\int_{r=0}^{1} 3(r\cos(\theta))^2\,r\,dr\,d\theta \\
&= 3\left(\int_{\theta=0}^{2\pi}\cos^2(\theta)d\theta\right)\left(\int_{r=0}^{1} r^3\,dr\right) \\
&= 3 \times \pi \times (1/4) = 3\pi/4.
\end{aligned}$$

Example of Green's Theorem

$$D = \text{unit disc}; \qquad \mathbf{u} = (x^3, x^3); \qquad \iint_D \text{curl}(\mathbf{u}) = 3\pi/4$$

Parametrise C as $(x, y) = (\cos(\theta), \sin(\theta))$ (for $0 \le \theta \le 2\pi$)
Then $d\mathbf{r} = (-\sin(\theta), \cos(\theta))\, d\theta$ and $\mathbf{u} = (x^3, x^3) = (\cos^3(\theta), \cos^3(\theta))$ so

$$\mathbf{u}.d\mathbf{r} = (\cos^4(\theta) - \sin(\theta)\cos^3(\theta))\, d\theta$$

Square $\cos^2(\theta) = \frac{1}{2}(1+\cos(2\theta))/2$ to get $\cos^4(\theta) = \frac{1}{4}(1+2\cos(2\theta)+\cos^2(2\theta))$.
Also $\cos^2(2\theta) = (1 + \cos(4\theta))/2$, so $\cos^4(\theta) = (3 + 4\cos(2\theta) + \cos(4\theta))/8$.
On the other hand, we have

$$\begin{aligned} \sin(\theta)\cos^3(\theta) &= (\sin(\theta)\cos(\theta))\ \cos^2(\theta) = \tfrac{1}{2}\sin(2\theta) \times \tfrac{1}{2}(1 + \cos(2\theta)) \\ &= \tfrac{1}{4}(\sin(2\theta) + \sin(2\theta)\cos(2\theta)) = \tfrac{1}{4}\sin(2\theta) + \tfrac{1}{8}\sin(4\theta) \end{aligned}$$

so $\quad \mathbf{u}.d\mathbf{r} = \frac{1}{8}\left(3 + 4\cos(2\theta) + \cos(4\theta) - 2\sin(2\theta) - \sin(4\theta)\right).$

It is standard that $\int_0^{2\pi} \sin(k\theta)d\theta = \int_0^{2\pi} \cos(k\theta)d\theta = 0$ (for $k > 0$) so

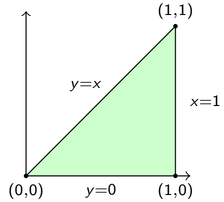
$$\begin{aligned} \int_C \mathbf{u}.d\mathbf{r} &= \tfrac{1}{8}\int_{\theta=0}^{2\pi} 3 + 4\cos(2\theta) + \cos(4\theta) - 2\sin(2\theta) - \sin(4\theta) d\theta \\ &= \tfrac{1}{8}\int_{\theta=0}^{2\pi} 3 d\theta = \tfrac{1}{8} \times 3 \times 2\pi = 3\pi/4. \end{aligned}$$

As expected, this is the same as $\iint_D \text{curl}(\mathbf{u})\, dA$.

Example of Green's Theorem

D : triangle, vertices $(0, 0)$, $(1, 0)$ and $(1, 1)$.
$\mathbf{u}$: vector field $(-y^2, x^2 - xy + y^2)$. This has

$$\text{curl}(\mathbf{u}) = \det \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ -y^2 & x^2 - xy + y^2 \end{bmatrix}$$
$$= (2x - y) - (-2y) = 2x + y$$



so

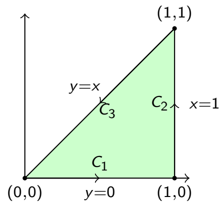
$$\begin{aligned} \iint_D \text{curl}(\mathbf{u})\, dA &= \int_{x=0}^{1}\int_{y=0}^{x} 2x + y\, dy\, dx = \int_{x=0}^{1} \left[2xy + \tfrac{1}{2}y^2\right]_{y=0}^{x} dx \\ &= \int_{x=0}^{1} 2x^2 + \tfrac{1}{2}x^2\, dx = \int_{x=0}^{1} \tfrac{5}{2}x^2\, dx = \left[\tfrac{5}{6}x^3\right]_{x=0}^{1} \\ &= 5/6. \end{aligned}$$

Example of Green's Theorem

D : triangle, vertices $(0, 0)$, $(1, 0)$ and $(1, 1)$.
$\mathbf{u}$: vector field $(-y^2, x^2 - xy + y^2)$.

$$\iint_D \text{curl}(\mathbf{u})\, dA = 5/6$$

The boundary C consists of C_1, C_2 and C_3.



On C_1: $y = 0$; $\mathbf{u} = (-y^2, x^2 - xy + y^2) = (0, x^2)$; $dy = 0$; $d\mathbf{r} = (dx, 0)$;
$\mathbf{u}.d\mathbf{r} = (0, x^2).(dx, 0) = 0$; so $\int_{C_1} \mathbf{u}.d\mathbf{r} = 0$.
On C_2: $x = 1$; $\mathbf{u} = (-y^2, 1 - y + y^2)$; $dx = 0$; $d\mathbf{r} = (0, dy)$;
$\mathbf{u}.d\mathbf{r} = (1 - y + y^2)\, dy$;

$$\int_{C_2} \mathbf{u}.d\mathbf{r} = \int_{y=0}^{1} (1 - y + y^2)\, dy = \left[y - \tfrac{1}{2}y^2 + \tfrac{1}{3}y^3\right]_{y=0}^{1} = 1 - \tfrac{1}{2} + \tfrac{1}{3} = 5/6.$$

On C_3: $y = x$; $\mathbf{u} = (-y^2, y^2)$; $dy = dx$; $d\mathbf{r} = (dx, dx)$;
$\mathbf{u}.d\mathbf{r} = (-y^2, y^2).(dx, dx) = 0$; so $\int_{C_3} \mathbf{u}.d\mathbf{r} = 0$.
Altogether: $\displaystyle\int_C \mathbf{u}.d\mathbf{r} = \int_{C_1} \mathbf{u}.d\mathbf{r} + \int_{C_2} \mathbf{u}.d\mathbf{r} + \int_{C_3} \mathbf{u}.d\mathbf{r} = 0 + 5/6 + 0 = 5/6.$
As expected, this is the same as $\iint_D \text{curl}(\mathbf{u})\, dA$.

The (three-dimensional) Divergence Theorem

Let E be the three-dimensional solid region enclosed by a surface S. Let $\mathbf{u}$ be a vector field that is well-behaved everywhere in E. Then

$$\iiint_E \text{div}(\mathbf{u})\, dV = \iint_S \mathbf{u}.d\mathbf{A}$$

This can be proved by an argument similar to that used for the two-dimensional version. The physical interpretation is also similar: in a steady state, the rate of flow of particles escaping through S must balance the rate of creation of particles in E.

Divergence Theorem Example 1

Let S be the unit sphere, and let E be the solid ball enclosed by S. Consider the vector field $\mathbf{u} = (x, 0, 0)$. This has $\mathrm{div}(\mathbf{u}) = \partial x/\partial x + \partial 0/\partial y + \partial 0/\partial z = 1$, so

$$\iiint_E \mathrm{div}(\mathbf{u}) dV = \iiint_E dV = \text{ volume of } E = 4\pi/3.$$

On S we have

$$\mathbf{r} = (x, y, z) = (\sin(\phi)\cos(\theta),\ \sin(\phi)\sin(\theta),\ \cos(\phi))$$
$$\mathbf{u} = (x, 0, 0) = (\sin(\phi)\cos(\theta),\ 0,\ 0).$$

The unit normal vector is $\mathbf{n} = \mathbf{e}_r = \mathbf{r}$, so $\mathbf{u}.\mathbf{n} = x^2 = \sin^2(\phi)\cos^2(\theta)$. We have also seen before that $dA = \sin(\phi)\, d\phi\, d\theta$, so

$$\begin{aligned}\iint_S \mathbf{u}.d\mathbf{A} &= \iint_S \mathbf{u}.\mathbf{n}\, dA = \int_{\theta=0}^{2\pi}\int_{\phi=0}^{\pi} \sin^3(\phi)\cos^2(\theta)\, d\phi\, d\theta \\ &= \left(\int_{\theta=0}^{2\pi} \cos^2(\theta)\, d\theta\right)\left(\int_{\phi=0}^{\pi} \sin^3(\phi)\, d\phi\right) = \pi \int_{\phi=0}^{\pi} \frac{1}{4}(3\sin(\phi) - \sin(3\phi))d\phi \\ &= \frac{\pi}{4}\left[-3\cos(\phi) + \tfrac{1}{3}\cos(3\phi)\right]_{\phi=0}^{\pi} = \frac{\pi}{4}((3 - \tfrac{1}{3}) - (-3 + \tfrac{1}{3})) = \frac{4\pi}{3}\end{aligned}$$

As expected, this is the same as $\iiint_E \mathrm{div}(\mathbf{u})dV$.

Divergence Theorem Example 2

Let E be the solid vertical cylinder of radius a and height $2b$ centred at the origin, and let S be the surface of E. Consider the vector field $\mathbf{u} = (-y, x, z^3)$. We have

$$\mathrm{div}(\mathbf{u}) = \frac{\partial}{\partial x}(-y) + \frac{\partial}{\partial y}(x) + \frac{\partial}{\partial z}(z^3) = 0 + 0 + 3z^2 = 3z^2.$$

The region E can be described in cylindrical polar coordinates by $0 \leq r \leq a$ and $-b \leq z \leq b$ (with $0 \leq \theta \leq 2\pi$ as usual). Moreover, the volume element in those coordinates is $dV = r\, dr\, d\theta\, dz$. It follows that

$$\begin{aligned}\iiint_E \mathrm{div}(\mathbf{u})\, dV &= \int_{z=-b}^{b}\int_{\theta=0}^{2\pi}\int_{r=0}^{a} 3z^2 r\, dr\, d\theta\, dz \\ &= 2\pi\left(\int_{r=0}^{a} r\, dr\right)\left(\int_{z=-b}^{b} 3z^2\, dz\right) \\ &= 2\pi\left[\tfrac{1}{2}r^2\right]_{r=0}^{a}\left[z^3\right]_{z=-b}^{b} = 2\pi a^2 b^3.\end{aligned}$$

Divergence Theorem Example 2

E : $\quad 0 \leq r \leq a, \quad -b \leq z \leq b, \quad 0 \leq \theta \leq 2\pi.$
$\mathbf{u} = (-y, x, z^3)$. $\iiint_E \mathrm{div}(\mathbf{u})dV = 2\pi a^2 b^3$.

Now consider instead $\iint_S \mathbf{u}.d\mathbf{A} = \iint_S \mathbf{u}.\mathbf{n}\, dA$. Let S_1 be the bottom end of E, where $z = -b$. Let S_2 be the top end, where $z = b$. Let S_3 be the curved outer surface, where $r = a$. On S_1, the outward unit normal is clearly $\mathbf{n} = -\mathbf{k} = (0, 0, -1)$. We also have $z = -b$, so $\mathbf{u} = (-y, x, -b^3)$, so $\mathbf{u}.\mathbf{n} = b^3$. As this is constant, it follows that

$$\iint_{S_1} \mathbf{u}.\mathbf{n}\, dA = \iint_{S_1} b^3\, dA = b^3 \times (\text{ area of } S_1) = \pi a^2 b^3.$$

On S_2 we have $\mathbf{n} = (0, 0, 1)$ and $\mathbf{u} = (-y, x, b^3)$, and it follows easily that $\iint_{S_2} \mathbf{u}.\mathbf{n}\, dA$ is also equal to $\pi a^2 b^3$.
For S_3 it is convenient to work in cylindrical polar coordinates again. The outward unit normal is $\mathbf{n} = \mathbf{e}_r = (\cos(\theta), \sin(\theta), 0)$, and the vector field is

$$\mathbf{u} = (-y, x, z^3) = (-r\sin(\theta), r\cos(\theta), z^3).$$

From this it is clear that $\mathbf{u}.\mathbf{n} = 0$, so $\iint_{S_3} \mathbf{u}.\mathbf{n}\, dA = 0$. Putting this together, we get

$$\iint_S \mathbf{u}.\mathbf{n}\, dA = \pi a^2 b^3 + \pi a^2 b^3 + 0 = 2\pi a^2 b^3,$$

which is the same as $\iiint_E \mathrm{div}(\mathbf{u})\, dV$, as expected.

Divergence Theorem Example 3

Let E be the solid region where $-1 \leq x, y \leq 1$ and $0 \leq z \leq (1 - x^2)(1 - y^2)$. Let S be the surface of E, and let $\mathbf{u}$ be the vector field $(x, y, 0)$. This has $\mathrm{div}(\mathbf{u}) = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(0) = 1 + 1 + 0 = 2$, so

$$\begin{aligned}\iiint_E \mathrm{div}(\mathbf{u})\, dV &= \int_{x=-1}^{1}\int_{y=-1}^{1}\int_{z=0}^{(1-x^2)(1-y^2)} 2\, dz\, dy\, dx \\ &= \int_{x=-1}^{1}\int_{y=-1}^{1} 2(1 - x^2)(1 - y^2)\, dy\, dx \\ &= 2\left(\int_{x=-1}^{1} 1 - x^2\, dx\right)\left(\int_{y=-1}^{1} 1 - y^2\, dy\right) \\ &= 2\left[x - \tfrac{1}{3}x^3\right]_{x=-1}^{1}\left[y - \tfrac{1}{3}y^3\right]_{y=-1}^{1}.\end{aligned}$$

Here

$$\left[x - \tfrac{1}{3}x^3\right]_{x=-1}^{1} = (1 - \tfrac{1}{3}) - (-1 + \tfrac{1}{3}) = \tfrac{2}{3} - (-\tfrac{2}{3}) = \tfrac{4}{3} = \left[y - \tfrac{1}{3}y^3\right]_{y=-1}^{1}$$

$$\iiint_E \mathrm{div}(\mathbf{u})\, dV = 2 \times \tfrac{4}{3} \times \tfrac{4}{3} = 32/9.$$

Divergence Theorem Example 3

$E: \qquad -1 \le x, y \le 1, \qquad 0 \le z \le f(x,y) = (1-x^2)(1-y^2)$
$\mathbf{u} = (x, y, 0) \qquad\qquad \iiint_E \operatorname{div}(\mathbf{u})\, dV = 32/9$

Now consider instead $\iint_S \mathbf{u}.d\mathbf{A} = \iint_S \mathbf{u}.\mathbf{n}\, dA$. Let S_1 be the bottom surface of E (where $z = 0$) and let S_2 be the top surface (where $z = (1-x^2)(1-y^2)$. (Note that there is no side surface, because $(1-x^2)(1-y^2) = 0$ whenever $x = \pm 1$ or $y = \pm 1$.) On S_1 the unit normal vector is $\mathbf{n} = (0, 0, -1)$ but $\mathbf{u} = (x, y, 0)$ so $\mathbf{u}.\mathbf{n} = 0$ so $\iint_{S_1} \mathbf{u}.d\mathbf{A} = 0$. On S_2 we have

$$d\mathbf{A} = (-f_x, -f_y, 1)\, dx\, dy = (2x - 2xy^2, 2y - 2x^2y, 1)\, dx\, dy,$$

$$\mathbf{u}.d\mathbf{A} = (x, y, 0).(2x - 2xy^2, 2y - 2x^2y, 1)\, dx\, dy = (2x^2 + 2y^2 - 4x^2y^2)\, dx\, dy.$$

$$\iint_{S_2} \mathbf{u}.d\mathbf{A} = \int_{y=-1}^{1} \int_{x=-1}^{1} (2x^2 + 2y^2 - 4x^2y^2)\, dx\, dy.$$

$$\int_{x=-1}^{1} (2x^2 + 2y^2 - 4x^2y^2)\, dx = \left[\tfrac{2}{3}x^3 + 2xy^2 - \tfrac{4}{3}x^3y^2\right]_{x=-1}^{1}$$
$$= \left(\tfrac{2}{3} + 2y^2 - \tfrac{4}{3}y^2\right) - \left(-\tfrac{2}{3} - 2y^2 + \tfrac{4}{3}y^2\right) = \tfrac{4}{3}(1 + y^2).$$

Feeding this into the outer integral gives

$$\iint_{S_2} \mathbf{u}.d\mathbf{A} = \tfrac{4}{3} \int_{y=-1}^{1} 1 + y^2\, dy = \tfrac{4}{3}\left[y + \tfrac{1}{3}y^3\right]_{y=-1}^{1} = 32/9 = \iiint_E \operatorname{div}(\mathbf{u})\, dV.$$

Stokes's Theorem

Stokes's Theorem is analogous to Green's Theorem, but it applies to curved surfaces as well as to flat regions in the plane. Suppose we have a surface S whose boundary is a closed curve C, and a well-behaved vector field $\mathbf{u}$. Then

$$\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A} = \pm \int_C \mathbf{u}.d\mathbf{r}.$$

We need a little more discussion to eliminate the ambiguity in the sign. To make sense of the right hand side, we need to specify the direction in which we move around C. The integral in one direction will be the negative of the integral in the opposite direction. Similarly, on the left hand side we have the integral of $\operatorname{curl}(\mathbf{u}).\mathbf{n}\, dA$, where $\mathbf{n}$ is a unit vector normal to the surface. There are two possible directions for $\mathbf{n}$ (each opposite to the other) and there is no natural rule to choose between them. However, the choice of $\mathbf{n}$ can be linked to the choice of direction around the curve as follows: if you walk in the specified direction with your feet on C and your head pointing in the direction of $\mathbf{n}$, then the surface S should be on your left. Provided that we follow this convention, we will have

$$\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A} = \iint_S \operatorname{curl}(\mathbf{u}).\mathbf{n}\, dA = + \int_C \mathbf{u}.d\mathbf{r}.$$

Stokes's Theorem Example 1

Consider the surface S given by $z = x^2 - y^2$ with $x^2 + y^2 \le 1$. We will check Stokes's Theorem for the vector field $(-y, x, 0)$. We parametrise S as

$$\mathbf{r} = (x, y, z) = (r\cos(s),\ r\sin(s),\ r^2\cos^2(s) - r^2\sin^2(s))$$

with $0 \le r \le 1$ and $0 \le s \le 2\pi$. Using $\cos^2(s) - \sin^2(s) = \cos(2s)$:

$$\mathbf{r} = (x, y, z) = (r\cos(s), r\sin(s), r^2\cos(2s)), \text{ which gives}$$

$$\mathbf{r}_r = (\cos(s), \sin(s), 2r\cos(2s))$$
$$\mathbf{r}_s = (-r\sin(s),\ r\cos(s),\ -2r^2\sin(2s))$$

$$\mathbf{r}_r \times \mathbf{r}_s = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos(s) & \sin(s) & 2r\cos(2s)) \\ -r\sin(s) & r\cos(s) & -2r^2\sin(2s) \end{bmatrix}$$

$$= (-2r^2\sin(s)\sin(2s) - 2r^2\cos(s)\cos(2s),\ 2r^2\cos(s)\sin(2s) - 2r^2\sin(s)\cos(2s),\ r\cos^2(s) + r\sin^2(s))$$

Using

$$\sin(a)\sin(b) + \cos(a)\cos(b) = \cos(a-b) = \cos(b-a)$$
$$\sin(a)\cos(b) - \cos(a)\sin(b) = \sin(a-b) = -\sin(b-a)$$

this becomes $\mathbf{r}_r \times \mathbf{r}_s = (-2r^2\cos(s),\ 2r^2\sin(s),\ r)$, so
$d\mathbf{A} = (\mathbf{r}_r \times \mathbf{r}_s)\, dr\, ds = (-2r^2\cos(s),\ 2r^2\sin(s),\ r)\, dr\, ds.$

Stokes's Theorem Example 1

$S: (x, y, z) = (r\cos(s), r\sin(s), r^2\cos(2s)),$
$d\mathbf{A} = (-2r^2\cos(s),\ 2r^2\sin(s),\ r)\, dr\, ds$

Next, we have $\operatorname{curl}(\mathbf{u}) = \det\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y & x & 0 \end{bmatrix} = (0, 0, 2)$, so

$$\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A} = \int_{s=0}^{2\pi} \int_{r=0}^{1} 2r\, dr\, ds = \int_{s=0}^{2\pi} 1\, ds = 2\pi.$$

On the other hand, we can parametrise the boundary curve C (where $r = 1$) as

$$\mathbf{r} = (x, y, z) = (\cos(s), \sin(s), \cos(2s)).$$

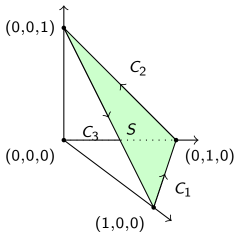
On this curve we have

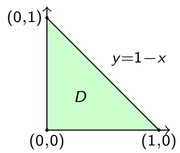
$$\mathbf{u} = (-y,\ x,\ 0) = (-\sin(s),\ \cos(s),\ 0)$$
$$d\mathbf{r} = (-\sin(s), \cos(s), -2\sin(2s))\, ds$$
$$\mathbf{u}.d\mathbf{r} = (\sin^2(s) + \cos^2(s))\, ds = ds$$
$$\int_C \mathbf{u}.d\mathbf{r} = \int_{s=0}^{2\pi} ds = 2\pi.$$

As expected, this is the same as $\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A}$.

Stokes's Theorem Example 2

Let S be the triangular surface shown on the left below, given by $x + y + z = 1$ with $x, y, z \geq 0$. Let $\mathbf{u}$ be the vector field (z, x, y).



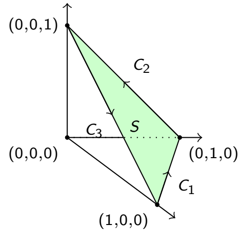


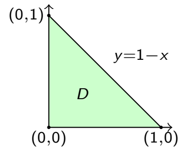
The boundary consists of the edges C_1, C_2 and C_3. We can parametrise C_1 by $\mathbf{r} = (x, y, z) = (1 - t, t, 0)$ for $0 \leq t \leq 1$. This gives $d\mathbf{r} = (-1, 1, 0)dt$. We can also substitute $x = 1 - t$ and $y = t$ and $z = 0$ in the definition $\mathbf{u} = (z, x, y)$ to get $\mathbf{u} = (0, 1 - t, t)$. This gives $\mathbf{u}.d\mathbf{r} = (1 - t)dt$, so

$$\int_{C_1} \mathbf{u}.d\mathbf{r} = \int_{t=0}^{1} (1 - t)\, dt = \left[t - \tfrac{1}{2}t^2 \right]_{t=0}^{1} = 1/2.$$

Stokes's Theorem Example 2

$S : x + y + z = 1$ with $x, y, z \geq 0$; $\quad \mathbf{u} = (z, x, y)$.





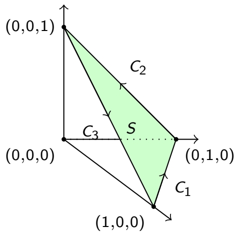
The other edges work in the same way, as in the following table:

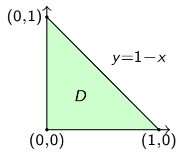
edge	C_1	C_2	C_3
$\mathbf{r}$	$(1-t, t, 0)$	$(0, 1-t, t)$	$(t, 0, 1-t)$
$d\mathbf{r}$	$(-1, 1, 0)dt$	$(0, -1, 1)dt$	$(1, 0, -1)dt$
$\mathbf{u}$	$(0, 1-t, t)$	$(t, 0, 1-t)$	$(1-t, t, 0)$
$\mathbf{u}.d\mathbf{r}$	$(1-t)dt$	$(1-t)dt$	$(1-t)dt$
$\int \mathbf{u}.d\mathbf{r}$	$1/2$	$1/2$	$1/2$

Altogether, we have $\int_C \mathbf{u}.d\mathbf{r} = 3/2$.

Stokes's Theorem Example 2

$S : x + y + z = 1$ with $x, y, z \geq 0$; $\quad \mathbf{u} = (z, x, y)$; $\quad \int_C \mathbf{u}.d\mathbf{r} = 3/2$.





On the other hand, we have $\operatorname{curl}(\mathbf{u}) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z & x & y \end{bmatrix} = (1, 1, 1)$.

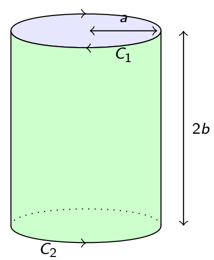
The shadow of S in the xy-plane is the triangle D shown on the right. The surface has the form $z = f(x, y)$, where $f(x, y) = 1 - x - y$ and (x, y) lies in D, so $d\mathbf{A} = (-f_x, -f_y, 1)\, dx\, dy = (1, 1, 1)\, dx\, dy$. This gives

$$\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A} = \int_D (1,1,1).(1,1,1)\, dx\, dy = 3 \int_{x=0}^{1} \int_{y=0}^{1-x} dy\, dx$$

$$= 3 \int_{x=0}^{1} (1 - x)\, dx = 3 \left[x - \tfrac{1}{2}x^2 \right]_{x=0}^{1} = 3/2 = \int_C \mathbf{u}.d\mathbf{r}.$$

Stokes's Theorem Example 3

$S =$ cylinder: $r = a$ with $-b \leq z \leq b$ and $0 \leq \theta \leq 2\pi$. Check Stokes's Theorem for the vector field $\mathbf{u} = (-zy, zx, z^2)$. We parametrise S as



$$\mathbf{r} = (x, y, z) = (a\cos(\theta),\ a\sin(\theta),\ z)$$
$$\mathbf{r}_\theta = (-a\sin(\theta),\ a\cos(\theta),\ 0)$$
$$\mathbf{r}_z = (0, 0, 1)$$
$$\mathbf{r}_\theta \times \mathbf{r}_z = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a\sin(\theta) & a\cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} = (a\cos(\theta),\ a\sin(\theta),\ 0)$$
$$d\mathbf{A} = (\mathbf{r}_\theta \times \mathbf{r}_z)\, d\theta\, dz = a(\cos(\theta), \sin(\theta), 0)\, d\theta\, dz.$$

Note that $d\mathbf{A}$ points outwards, away from the z-axis. Also

$$\operatorname{curl}(\mathbf{u}) = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -zy & zx & z^2 \end{bmatrix} = (0 - x,\ -y - 0,\ z - (-z)) = (-x, -y, 2z).$$

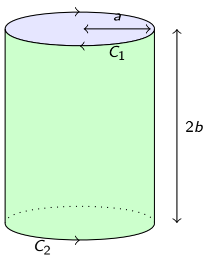
On the surface S this becomes $\operatorname{curl}(\mathbf{u}) = (-a\cos(\theta),\ -a\sin(\theta),\ 2z)$, so

$$\operatorname{curl}(\mathbf{u}).d\mathbf{A} = (-a^2\cos^2(\theta) - a^2\sin^2(\theta))\, d\theta\, dz = -a^2\, d\theta\, dz$$

$$\iint_S \operatorname{curl}(\mathbf{u}).d\mathbf{A} = -a^2 \int_{\theta=0}^{2\pi} \int_{z=-b}^{b} d\theta\, dz = -a^2 \times 2\pi \times 2b = -4\pi a^2 b.$$

Stokes's Theorem Example 3

$S =$ cylinder: $r = a$ with $-b \le z \le b$ and $0 \le \theta \le 2\pi$. $\mathbf{u} = (-zy, zx, z^2)$. $\iint_S \text{curl}(\mathbf{u}).d\mathbf{A} = -4\pi a^2 b$.



Boundary of S: C_1 and C_2. Directions as shown keep S on the left when walking with head in the direction of $d\mathbf{A}$, away from the z-axis. Compatible parametrisations:

$$C_1 : (x, y, z) = (a\cos(t), -a\sin(t), b)$$
$$C_2 : (x, y, z) = (a\cos(t), a\sin(t), -b).$$

On C_1:

$$d\mathbf{r} = (-a\sin(t), -a\cos(t), 0)\,dt$$
$$\mathbf{u} = (-zy, zx, z^2) = (ab\sin(t), ab\cos(t), b^2)$$
$$\mathbf{u}.d\mathbf{r} = -a^2 b\sin^2(t) - a^2 b\cos^2(t) = -a^2 b$$
$$\int_{C_1} \mathbf{u}.d\mathbf{r} = \int_{t=0}^{2\pi} -a^2 b\,dt = -2\pi a^2 b.$$

C_2 is similar: $d\mathbf{r} = (-a\sin(t), a\cos(t), 0)\,dt$, $\mathbf{u} = (ab\sin(t), -ab\cos(t), b^2)$, $\mathbf{u}.d\mathbf{r} = -a^2 b\sin^2(t) - a^2 b\cos^2(t) = -a^2 b$, $\int_{C_2} \mathbf{u}.d\mathbf{r} = -2\pi a^2 b$. Putting these together, we get $\int_C \mathbf{u}.d\mathbf{r} = -4\pi a^2 b$, which is the same as $\iint_S \text{curl}(\mathbf{u}).d\mathbf{A}$, as expected.