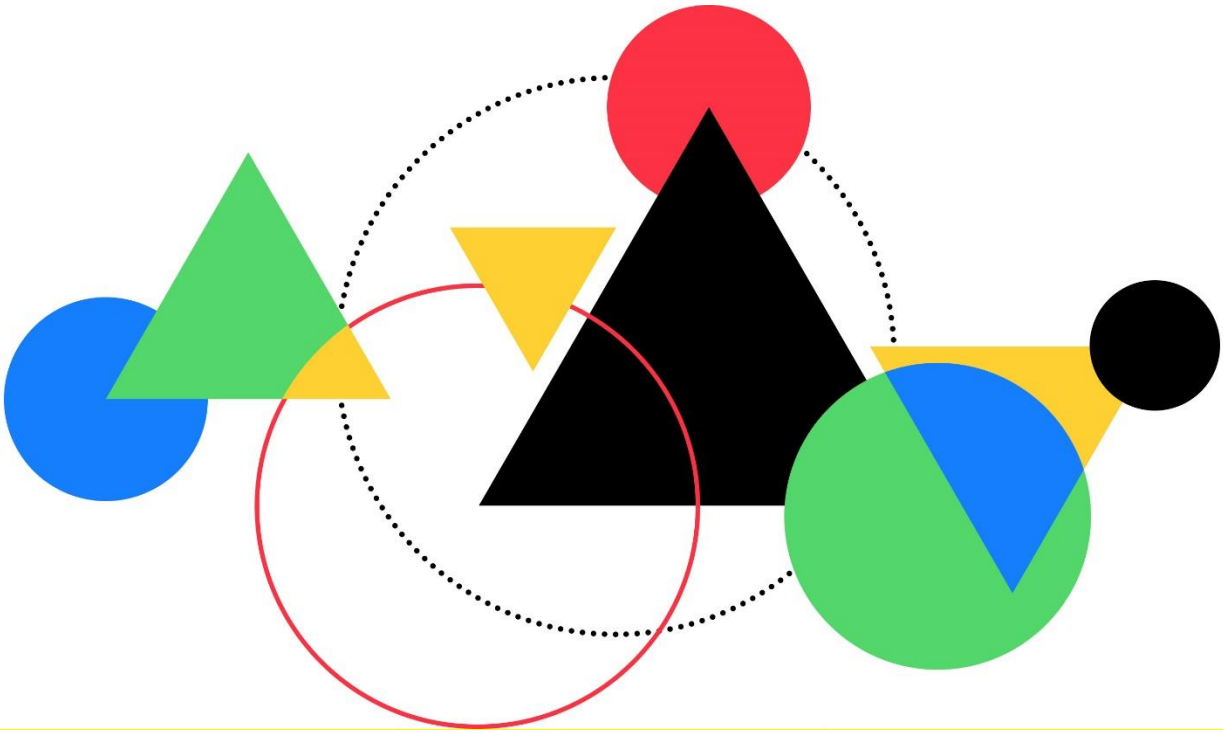
CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Introduction, Vectors refresher

1.i

COURSE SYNOPSIS, RECOMMENDED BOOKS

Course syllabus (on which exams are based):

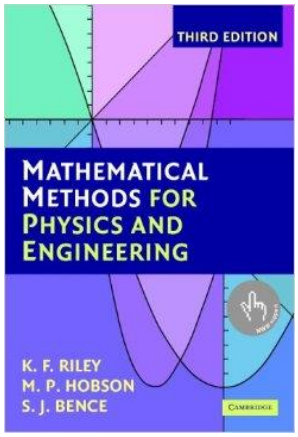
Double integrals and their evaluation by repeated integration in Cartesian, plane polar and other specified coordinate systems. Jacobians. Line, surface and volume integrals, evaluation by change of variables (Cartesian, plane polar, spherical polar coordinates and cylindrical coordinates only unless the transformation to be used is specified). Integrals around closed curves and exact differentials. Scalar and vector fields. The operations of grad, div and curl and understanding and use of identities involving these. The statements of the theorems of Gauss and Stokes with simple applications. Conservative fields.

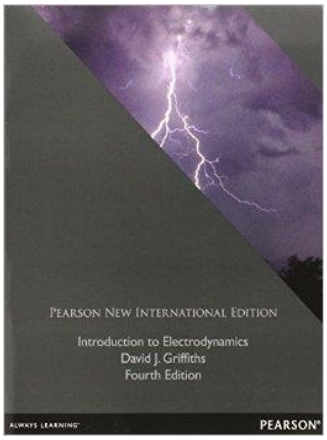
Recommended Books:

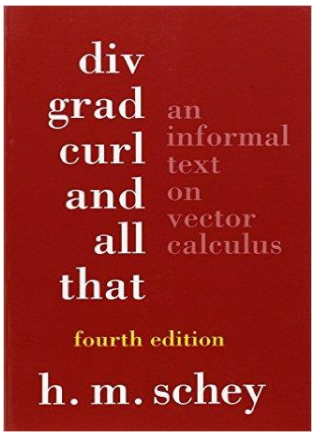
Mathematical Methods for Physics and Engineering (Riley, Hobson and Bence)
This book is lazily referred to as "Riley" throughout these notes (sorry, Drs H and B)
You will all have this book, and it covers all of the maths of this course. However it is rather terse at times and you will benefit from looking at one or both of these:

Introduction to Electrodynamics (Griffiths)
You will buy this next year if you haven't already, and the chapter on vector calculus is very clear

Div grad curl and all that (Schey)
A nice discussion of the subject, although topics are ordered differently to most courses
NB: the latest version of this book uses the opposite convention to polar coordinates to this course (and indeed most of physics), but older versions can often be found in libraries







A vectors refresher

The following is a very brief refresher on the properties of vectors and how to manipulate them.

Vectors

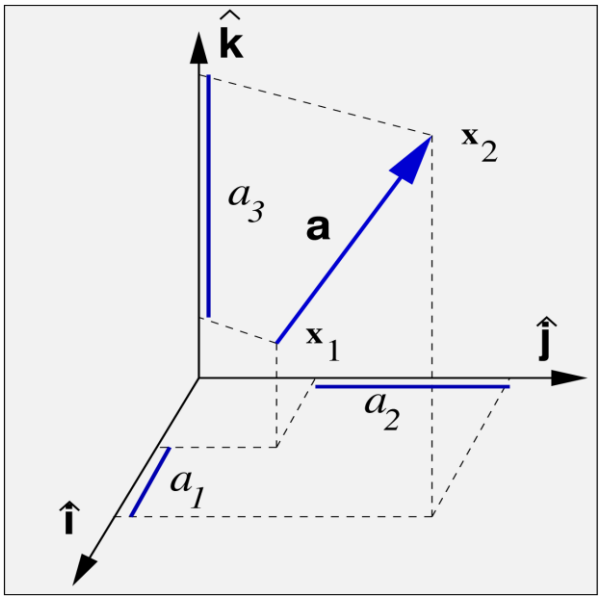
A vector has direction and magnitude and is written in these notes in bold e.g. **F** or underlined. In written material I will use underlining, you may also use an over-arrow (just try to be consistent).

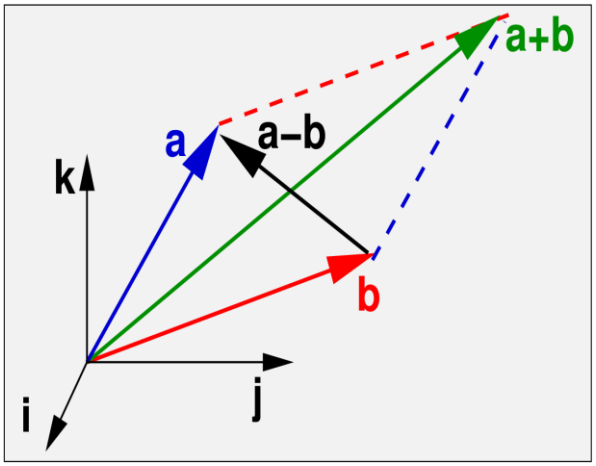
A unit vector has magnitude of one (unity), and is often represented with a hat e.g. $\hat{t}$

The (right-handed) Cartesian coordinate system we use has three orthogonal unit vectors.

Vectors can be described as

F = F_X **i** + F_Y **j** + F_Z **k** (in the diagram below left, $F_X=a_1$, $F_y=a_2$, $F_z=a_3$)





The unit vectors along the x,y,z axes in Cartesian coordinates are $\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}$.

Vectors can be added geometrically, by placing them end-to-end to see the resultant vector (diagram above right), or they can be added algebraically (add the vector components):

F + **G** = (F_X + G_X) **i** + (F_Y + G_Y) **j** + (F_Z + G_Z) **k**

Example 1.1: Write down an expression for the force on Q at **R** due to N charges q_i at $\mathbf{r}_i$ (i=1,2,3.. N)

Answer: The force between two charges is given by:

$$F = K\frac{Qq_1}{r^2}\hat{r}$$

and the vector **from** q_i to Q is **($\mathbf{R}-\mathbf{r}_i$)**, so the sum of the forces

$$F = \sum K\frac{Qq_i}{|R-r_i|^3}(R-r_i)$$

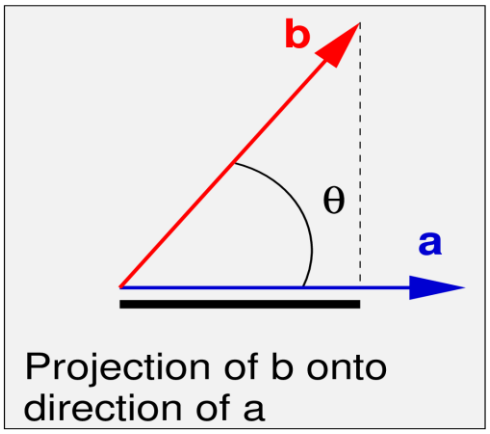
(note that this question makes use of *superposition*)

The scalar ('dot') product of two vectors is a scalar (hence the name), of magnitude

$|\mathbf{a}|\,|\mathbf{b}|\cos\theta$

It is the projection of the vector b in the direction of **a** and vice versa

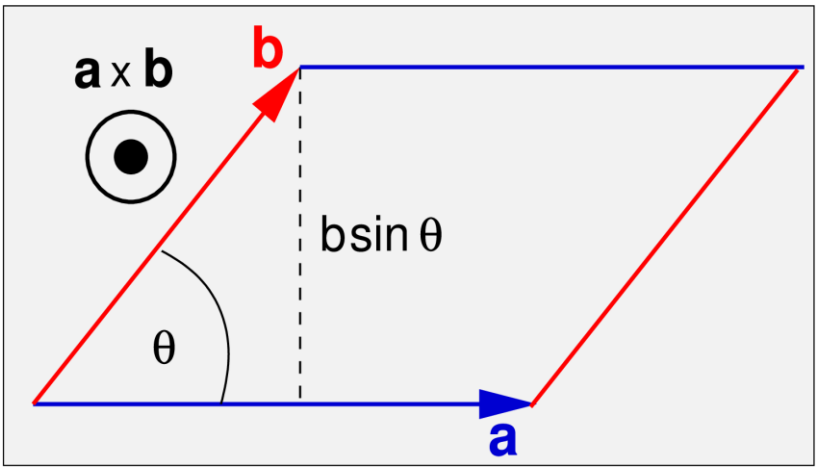
In algebraic form $F \cdot G = F_X G_X + F_Y G_Y + F_Z G_Z$



The vector ('cross') product of two vectors is a **vector** of magnitude

$|\mathbf{a}|\,|\mathbf{b}|\sin\theta$

with direction given by the right-hand screw rule. It is the area of the parallelogram of sides given by the vectors **a** and **b**:





The circle with the point in it in the diagram above represents a vector coming out of the page, a vector going into the page is represented by a circle with a cross in it.

Algebraically, the cross product is given by:

$\mathbf{F} \times \mathbf{G} = (F_y G_z - F_z G_y)\,\mathbf{i} - (F_x G_z - F_z G_z)\,\mathbf{j} + (F_x Gy - F_y G_x)\,\mathbf{k}$

which is most easily remembered as the determinant of a matrix:

$$a \times b = det \begin{pmatrix} i & j & k \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{pmatrix}$$

Useful things to remember for these two products:

Dot product

The dot product of orthogonal vectors is zero (this is used surprisingly often)

It *commutes*: $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$

It *distributes*: $\mathbf{a}\cdot(\mathbf{b} + \mathbf{c}) = \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{c}$

It multiplies by a simple scalar: $(\lambda\mathbf{a})\cdot\mathbf{b} = \lambda(\mathbf{a}\cdot\mathbf{b}) = \mathbf{a}\cdot(\lambda\mathbf{b})$

Cross product

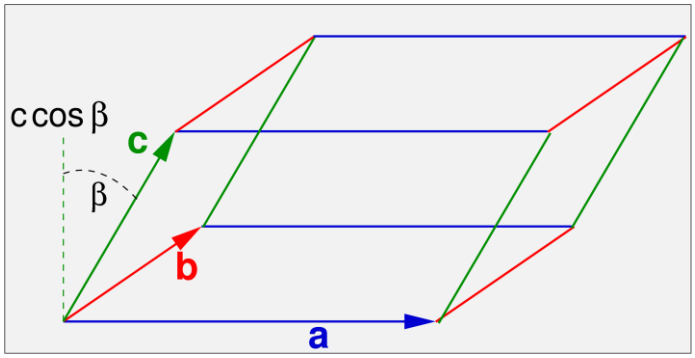
Cross product of parallel vectors is zero

It *anti-commutes*: $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$

It does not *associate*: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$

Scalar triple product

The scalar triple product $|(a \times b).c|$ gives the volume of the parallelepiped who sides are the vectors **a**, **b**, **c**.



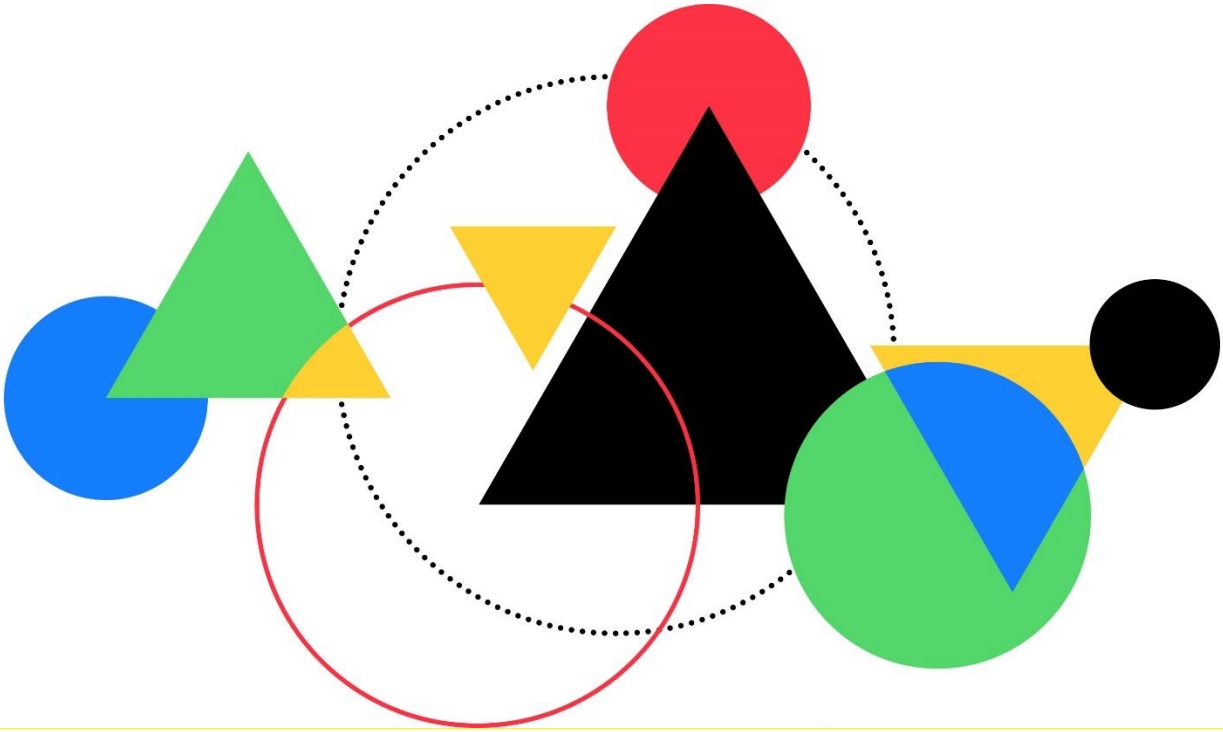
The scalar triple product is equal to the determinant of a 3x3 matrix:

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = det\begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix}$$

(Note that I could equally have written the components of the vector as $a_1 a_2 a_3$)

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Planes, Scalar and Vector fields

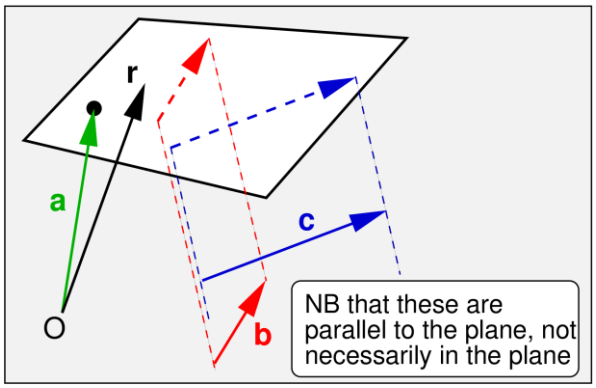
1.ii

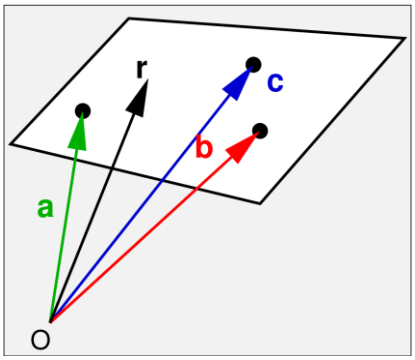
Planes and simple coordinate rotations

In the following, the vector **r** describes a position on the plane.

A plane can be defined in three ways:

(i) If **b** and **c** are parallel to the plane, and **a** is in the plane:
r = **a** + λ**b** + μ**c**





(ii) If three points, **a**, **b** and **c** lie in the plane:

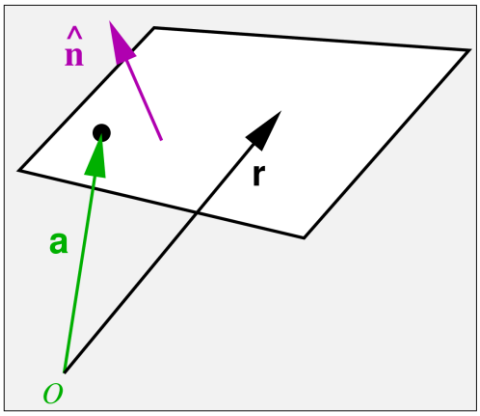
r = **a** + λ(**b**-**a**) + μ(**c**-**a**)

(iii) If the unit normal to the plane is **n** and **a** is in the plane:

r·**n** = **a**·**n** = D

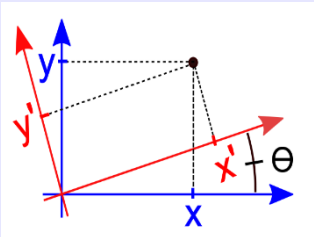
(where D is the perpendicular distance from the origin to the plane; this arises because **n**· (**r**-**a**)=0)

We will later introduce the *gradient*, which will allow you to calculate tangent planes via this approach.



Example 1.2: Coordinate rotation. Show that the matrix transformation of vector components under the illustrated rotation is:

$$\begin{bmatrix} F_X' \\ F_Y' \\ F_Z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_X \\ F_Y \\ F_Z \end{bmatrix}$$



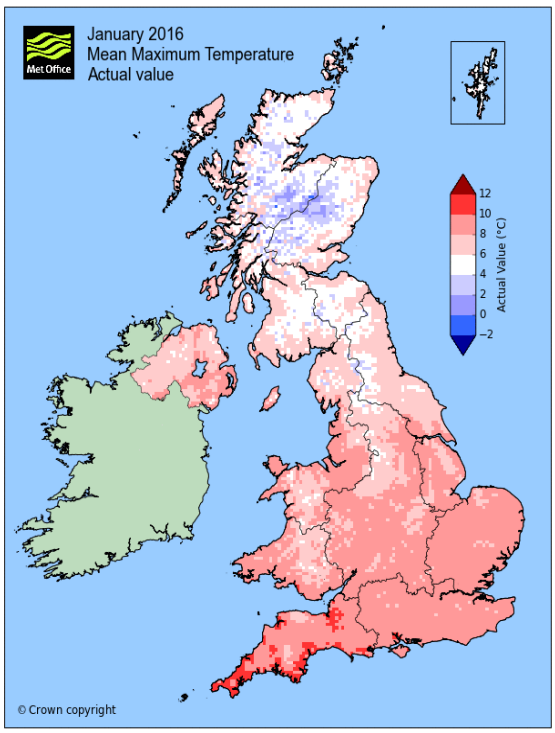
Answer: In the unrotated frame: $x = r\cos\phi$, $y = r\sin\phi$ and in the rotated frame

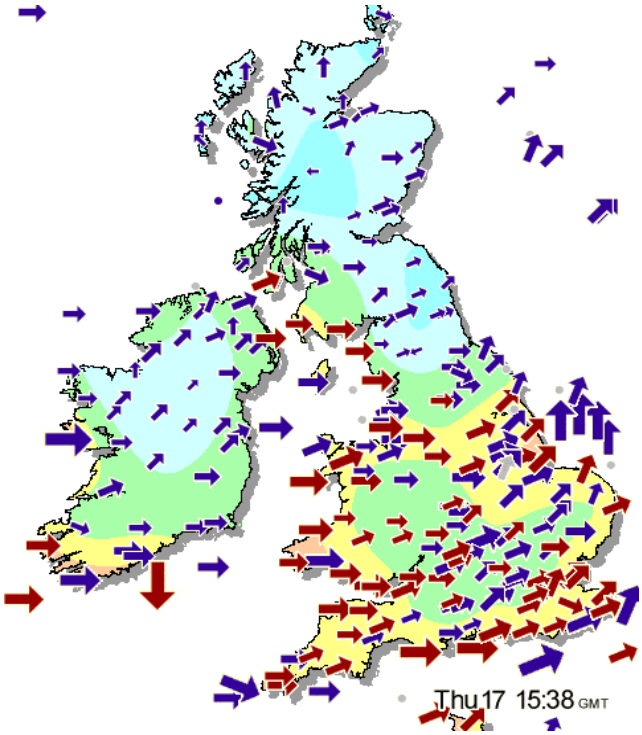
$x' = r\cos(\phi - \theta) = r(\cos\phi\cos\theta + \sin\phi\sin\theta) = x\cos\theta + y\sin\theta$
$y' = r\sin(\phi - \theta) = r(\sin\phi\cos\theta - \cos\phi\sin\theta) = y\cos\theta - x\sin\theta$

The matrix follows , with the minor added complication of a supposed z-axis which is unchanged (e.g. this could be the base plane of something in cylindrical coordinates)

Scalar vs Vector fields

Scalar fields are fields of scalars, vector fields are field of vectors...

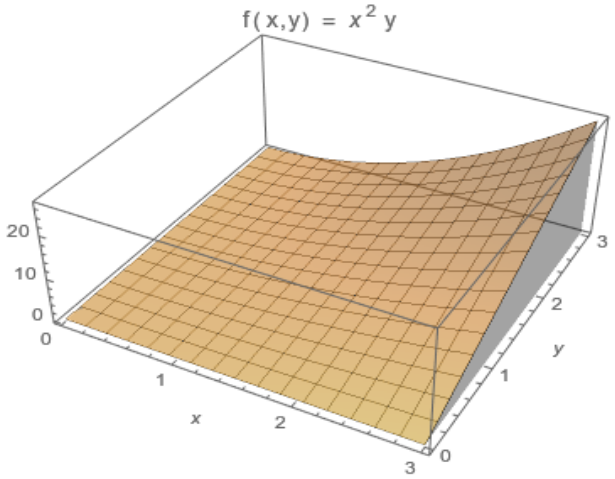




Example scalar field:

$F = x^2y$

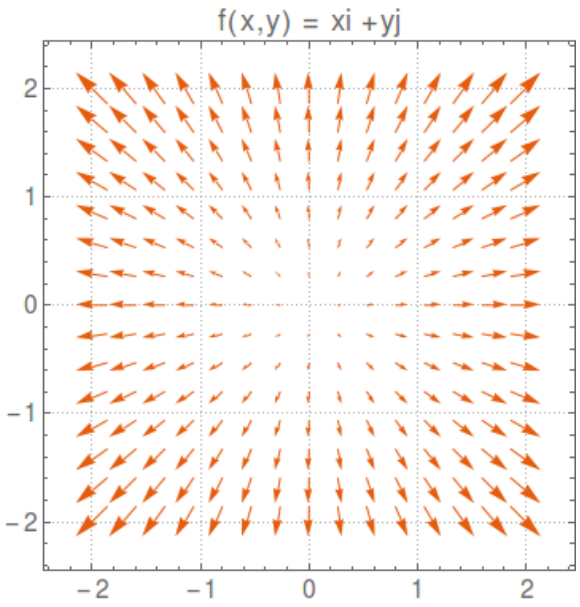
(every x,y pair has a given scalar value associated with it)

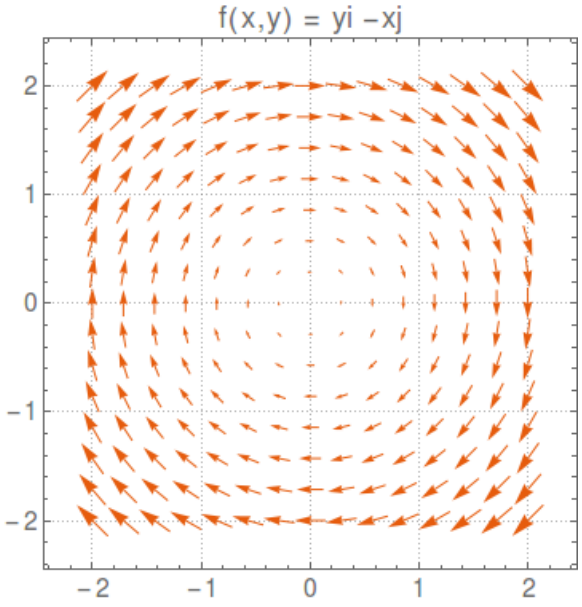


Example vector fields:

$$\boldsymbol{F} = x\boldsymbol{i} + y\boldsymbol{j}$$
$$\boldsymbol{F} = y\boldsymbol{i} - x\boldsymbol{j}$$

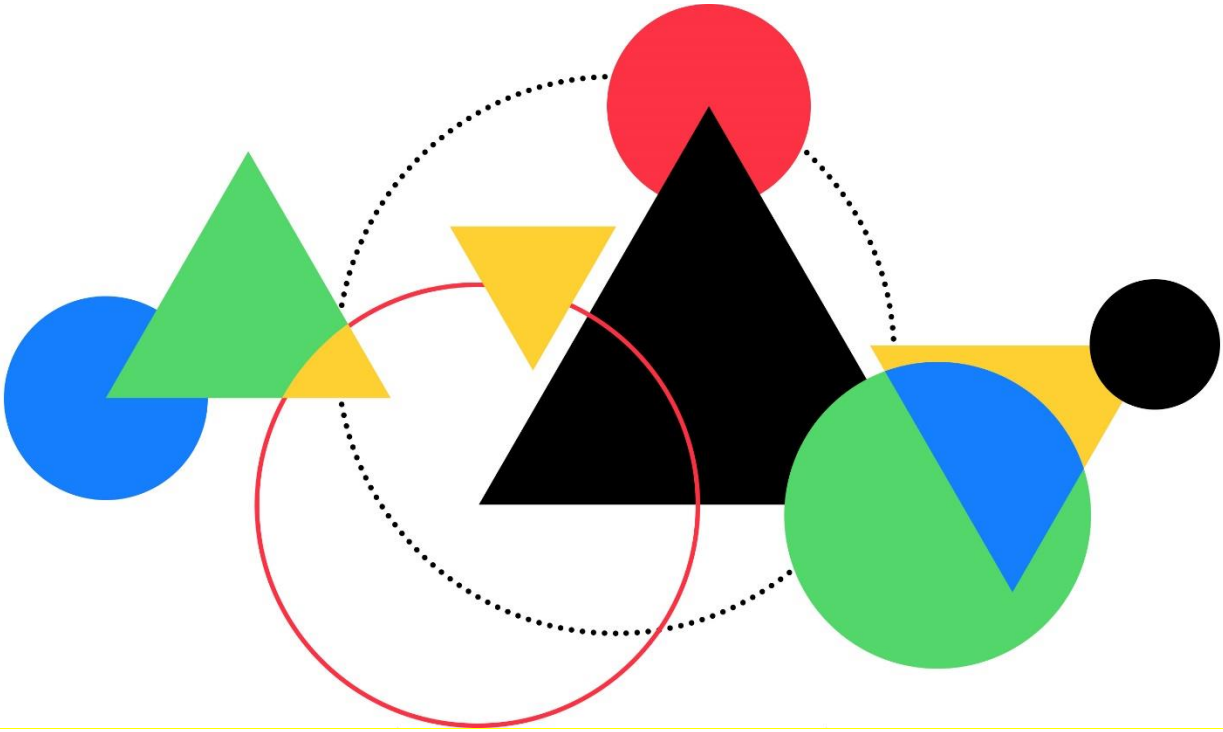
Try and sketch these then look on the next page to check you got it right!


f(x,y) = xi +yj
2
1
0
−1
−2
−2
−1
0
1
2


f(x,y) = yi −xj
2
1
0
−1
−2
−2
−1
0
1
2

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Integrals in more than one variable

2.i

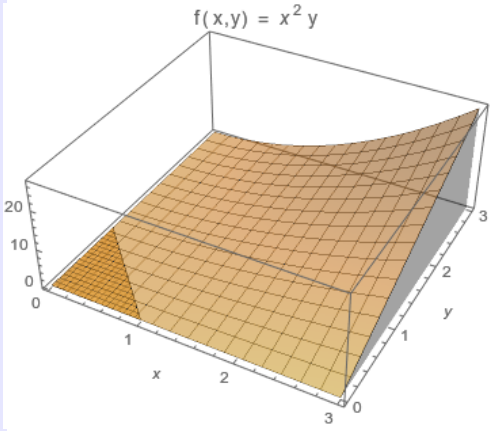
Integrals in more than one variable

When you have been previously integrating a single variable with respect to x, you can consider this to be the area under a 2D graph. Integrating with respect to two variables, typically x and y, can be considered as being the volume under a 3D graph. For a function of (x,y) we may consider integrating between some limits in x, and then for each dx considering y to varying between limits which may vary as a function of x. A good introduction to this is presented in Riley [ch 6]. Care needs to be taken in identifying the exact *region* (rather than just limits). This is best learnt by example, so here are a few:

Example 2.1: Evaluate the double integral

$$\iint_R x^2\, y dx dy$$

where R is the triangular region bounded by the lines x=0, y=0 and x+y=1. This region is illustrated as part of this scalar field in the diagram to the right. [this question is done in Riley section 6.1]



Answer: in the x-y plane, the region is the triangle connecting (0,1) to (1,0) – *sketching the regions is usually an essential first step with multiple integrals.* So if we integrate x between 0 and 1, then for each dx, y will be varying between 0 and (1-x) in each vertical strip.

$$\iint_R x^2\, y dx dy = \int_{x=0}^{1} \int_{y=0}^{1-x} x^2\, y dy dx$$

which can also be written as $\iint_R x^2\, y dx dy = \int_0^1 x^2\, dx \int_0^{1-x} y\, dy$

since x is just a constant when you're integrating x (this separation often makes things easier).

$$\int_0^1 x^2\, dx \int_0^{1-x} y\, dy = \int_0^1 x^2\, dx \left[\frac{y^2}{2}\right]_{y=0}^{y=1-x} = \int_0^1 x^2 \frac{(1-x)^2}{2} dx = \frac{1}{2}\int_0^1 x^2 + x^4 - 2x^3\, dx = \frac{1}{60}$$

This can also be done by integrating y from 0 to 1, and hence x from 0 to 1-y in each horizontal strip, and gives the same answer, as it should.

Answer: Since we are just finding the volume, we can just write this as a 2D integral in (x,y) where the height $f(x) = z = c - x\left(\frac{c}{a}\right) - y\left(\frac{c}{b}\right)$. If we take the x axis to go from 0 to a, then the y axis on each (vertical) strip goes from 0 to (b – bx/a), so the integral becomes:

$$\int_0^a c\, dx \int_0^{b-\left(\frac{bx}{a}\right)} \left(1 - \frac{x}{a} - \frac{y}{b}\right) dy = \int_0^a c\, dx \left[y - \left(\frac{xy}{a}\right) - \left(\frac{y^2}{2b}\right)\right]_0^{b-\left(\frac{bx}{a}\right)}$$

$$= c\int_0^a dx \left[b - \frac{bx}{a} - \frac{bx}{a} + \frac{bx^2}{a^2} - \frac{b}{2} + \frac{bx}{a} - \frac{x^2}{2a^2}\right]$$

$$c\int_0^a dx\left(\frac{bx^2}{2a^2}-\frac{bx}{a}+\frac{b}{2}\right)=\frac{abc}{6}$$

This integral can also be performed by doing a proper (x,y,z) integral of unity (because it is the volume), which gives us more flexibility to put in whatever term we might want to integrate.

This is quite straightforward, since we already have our expression for the z-term as a function of (x,y), so the integral is:

$$\int_0^a dx\int_0^{b-\left(\frac{bx}{a}\right)}dy\int_0^{c\left(1-\frac{y}{b}-\frac{x}{a}\right)}dz$$

which gives the same answer, but now allows us to put some actual function in to integrate (see next example).

Question 2.3: Find the mass of the tetrahedron bounded by the three coordinate surfaces and the plane x/a + y/b + z/c = 1, 1, if its density is given by ρ (x , y , z) = ρ_0 (1 + x/a).

Answer: This question has the same geometric region as the previous one, but we now need to integrate a density term which has a dependence on x. The integral is clearly given by:

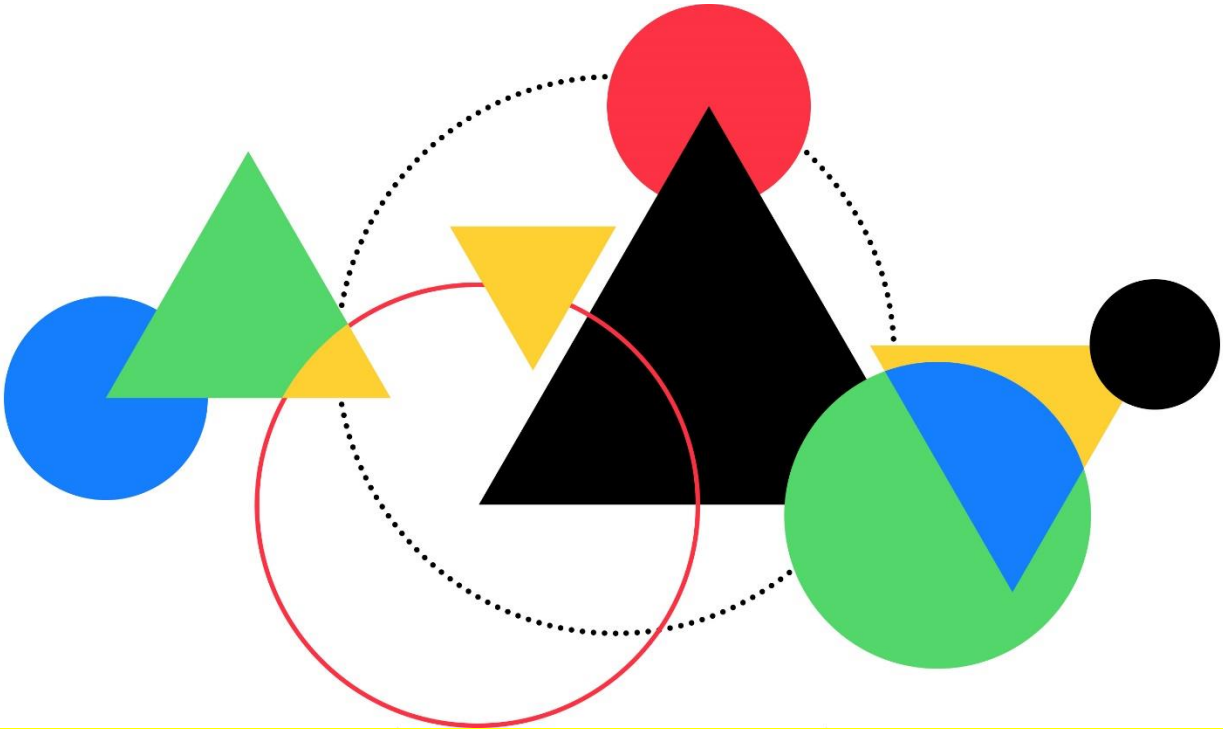
$$\int_0^a \rho_0\left(1+\frac{x}{a}\right)dx\int_0^{b-\left(\frac{bx}{a}\right)}dy\int_0^{c\left(1-\frac{y}{b}-\frac{x}{a}\right)}dz$$

where we've taken the mass function outside of the integrations with respect to y and z since it is independent of them. Then proceed precisely as before, which means that we get to:

$$Mass=c\rho_0\int_0^a\left(1+\frac{x}{a}\right)\left(\frac{bx^2}{2a^2}-\frac{bx}{a}+\frac{b}{2}\right)dx=\frac{5}{24}abc\rho_0$$

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020

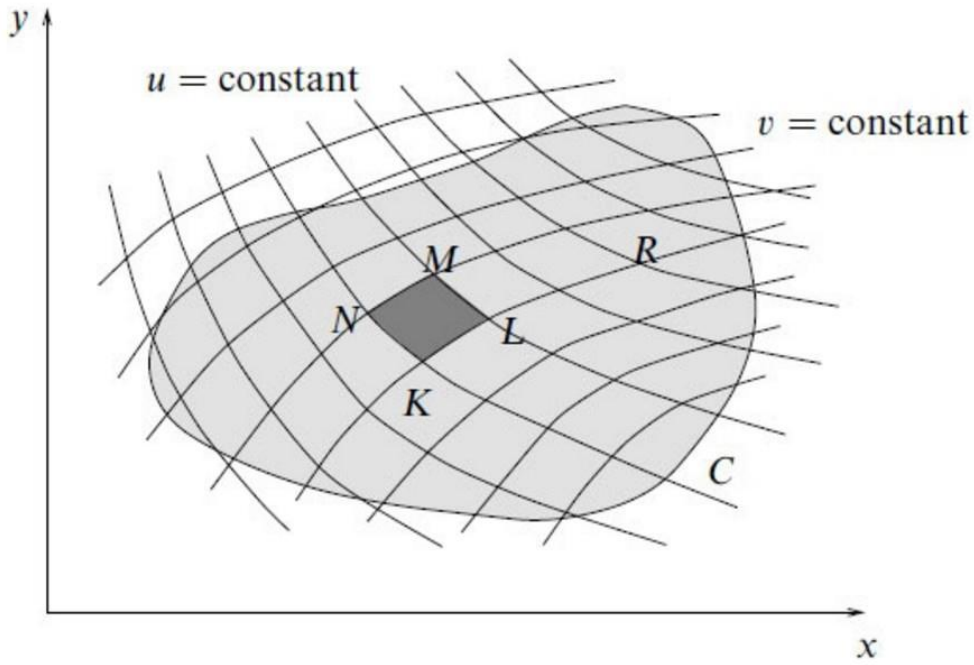


Changing coordinate systems: The Jacobian

2.ii

Changing variables: the Jacobian

It is often the case that you wish to change variables to simplify or otherwise address a problem.



This often occurs whenwhen the relative sizes of areas (or volumes) in two coordinate systems need to be considered when evaluating the integral. Consider the figure below, and the relative sizes of the areas (dx dy) and (du dv). This approach is from Riley (section 6.4).

In the x-y coordinates, the line KL (at constant v) corresponds to

$$KL = \frac{\partial x}{\partial u} du\, i + \frac{\partial y}{\partial u} du\, j$$

and the line KN (at constant u) corresponds to

$$KN = \frac{\partial x}{\partial v} dv\, i + \frac{\partial y}{\partial v} dv\, j$$

The cross product of these two vectors can be evaluated as:

$$KL \times KN = det \begin{pmatrix} i & j & k \\ \frac{\partial x}{\partial u} du & \frac{\partial y}{\partial u} du & 0 \\ \frac{\partial x}{\partial v} dv & \frac{\partial y}{\partial v} dv & 0 \end{pmatrix}$$

$$= k\left(\frac{\partial x}{\partial u} du \frac{\partial y}{\partial v} dv - \frac{\partial y}{\partial u} du \frac{\partial x}{\partial v} dv\right) = k\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right) du dv$$

(note that, since the two vectors were in the x-y plane the cross product only has a z-direction component, as expected). The magnitude of this vector is the area of the parallelogram KLMN.

We can define the *Jacobian* of x-y with respect to u-v as

$$J = \frac{\partial(x,y)}{\partial(u,v)} \equiv \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}$$

which can also be written as the determinant of a matrix:

$$J = det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

(NB 'Jacobian' is used interchangeably sometimes to mean the matrix, and sometimes to mean its determinant. Sorry.) So in converting from x-y to u-v coordinates we replace

$dxdy$ with $|J|dudv$

and convert the limits region from one coordinate system to the other (straightforwardly done in most cases from the provided equations).

Question: Evaluate the Gaussian integral $\int_{-\infty}^{\infty} e^{-x^2}\,dx$

Answer 2.4: This question makes use of integrals in two dimensions and also the Jacobian to transform coordinate systems. The solution is well-known and in most textbooks. Setting the integral equal to I, and squaring it:

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2}\,dx \int_{-\infty}^{\infty} e^{-y^2}\,dy = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dy = \iint_R e^{-(x^2+y^2)}\,dydx$$

We can now switch to plane polar coordinates, in which system $x^2+y^2=r^2$. In doing so we need to evaluate the determinant of the Jacobian which scales the coordinate transform. The 2D Jacobian is given by:

$$\begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{bmatrix}$$

and the determinant of this matrix is written as $\frac{\partial(x,y)}{\partial(r,\theta)}$.

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r$$

Switching coordinate systems using the Jacobian is straightforwardly done as follows:

$$\int f\,(x,y)dxdy = \int f\,(u,v)\frac{\partial(x,y)}{\partial(r,\theta)}dudv$$

– but <u>don't forget to change the limits/region too!</u>

So returning to the Gaussian integral, changing the variables and inserting the Jacobian determinant, we get:

$$I^2 = \int_0^{2\pi} d\,\phi \int_0^{\infty} r\,e^{-r^2}dr = 2\pi\left[\frac{-e^{-r^2}}{2}\right]_0^{\infty} = 2\pi\left(\frac{1}{2}\right) = \pi \text{and so } I = \sqrt{\pi}$$

Appendix C of Blundell & Blundell (your 2nd year Thermodynamics textbook) shows nicely how this result can be worked on to give the answer for different powers of x in front of the integral, useful for working out the means of particle speed distribution etc.

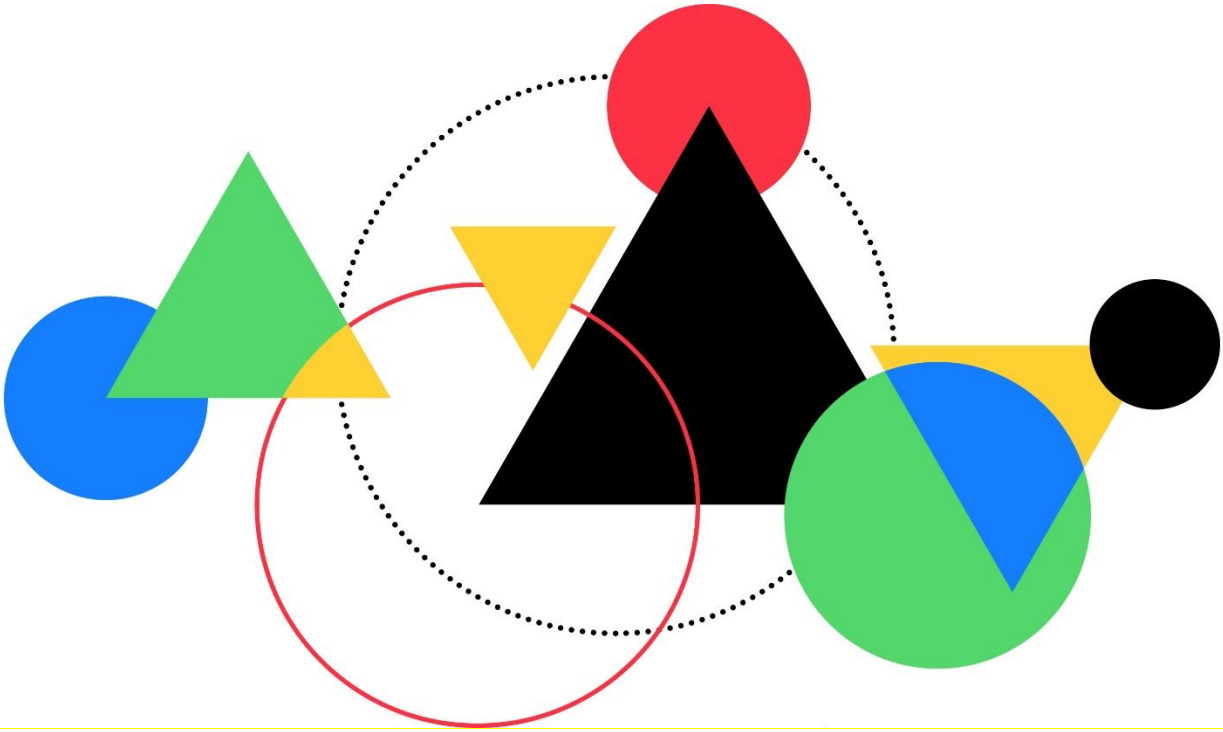
The Jacobian for a three-dimensional transformation is given by:

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

With some thought, and comparison to the derivation of the 2D Jacobian as well as the formula for the scalar triple product, we can see that the determinant of this matrix corresponds to a 'volume'.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Gradient

3.i

The gradient

In this section we introduce for the first time the 'Del' or 'nabla' symbol, which should be viewed as the following vector operator:

$$\boldsymbol{\nabla} = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z} \right)$$

It is really useful to memorise it just like this (which is why I've made it big and centre-stage), as a vector operator. I think the treatment given in Griffiths is the clearest and most to-the-point (at least for physicists). It is the key operator in calculating the gradient, divergence and curl.

The *gradient* is a vector and is the result when you apply this vector operator to a scalar field. It gives you both the magnitude (slope) and direction of greatest change of the scalar field.

So, if ϕ is a scalar field (I will most commonly use U, ϕ, ψ) then its gradient is given, in (x, y, z) Cartesian coordinates, by:

$$\boldsymbol{\nabla}\phi = \left(\boldsymbol{i}\frac{\partial \phi}{\partial x} + \boldsymbol{j}\frac{\partial \phi}{\partial y} + \boldsymbol{k}\frac{\partial \phi}{\partial z} \right)$$

Exploring this a bit further, suppose we have a variable U(x,y,z) i.e. a scalar field, and we want to calculate 'how fast does U vary when we vary x,y,z ?' . The total change in U is given by:

$$dU = dx\left(\frac{\partial U}{\partial x}\right) + dy\left(\frac{\partial U}{\partial y}\right) + dz\left(\frac{\partial U}{\partial z}\right)$$

which is equivalent to:

$$dU = \left(\boldsymbol{i}\frac{\partial U}{\partial x} + \boldsymbol{j}\frac{\partial U}{\partial y} + \boldsymbol{k}\frac{\partial U}{\partial z} \right).(\boldsymbol{i}\, dx + \boldsymbol{j}\, dy + \boldsymbol{k}\, dz)$$

which itself is equivalent to:

$$dU = (\boldsymbol{\nabla} U).(\boldsymbol{dl})$$

where d**l** is the vector representing the small movement by dx, dy, dz in the scalar field. This is an example of a *directional derivative,* which we will shortly define formally.

Since

$$dU = (\boldsymbol{\nabla} U).(\boldsymbol{dl}) = |\nabla U||dl|\cos\theta$$

this demonstrates that the greatest change in U occurs in the direction of the gradient, and has slope equal to the magnitude of the gradient. It also demonstrates that in directions perpendicular to the

gradient there is no change → the gradient is perpendicular to equipotentials (useful in both two and three dimensions in this course).

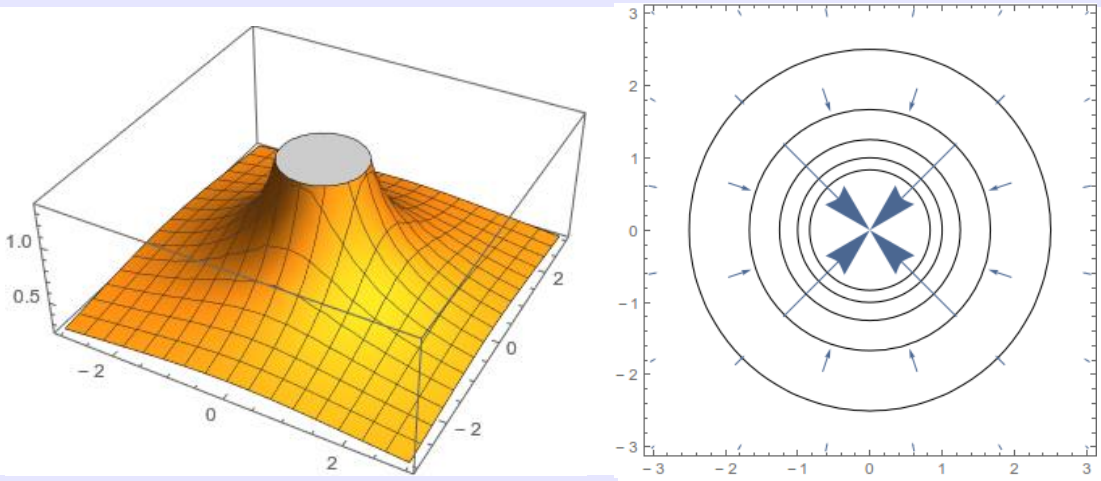
Example 4.1 The electric potential from a point charge has the following form:

$\phi = \frac{C}{(x^2+y^2)^{1/2}}$– calculate the electric field

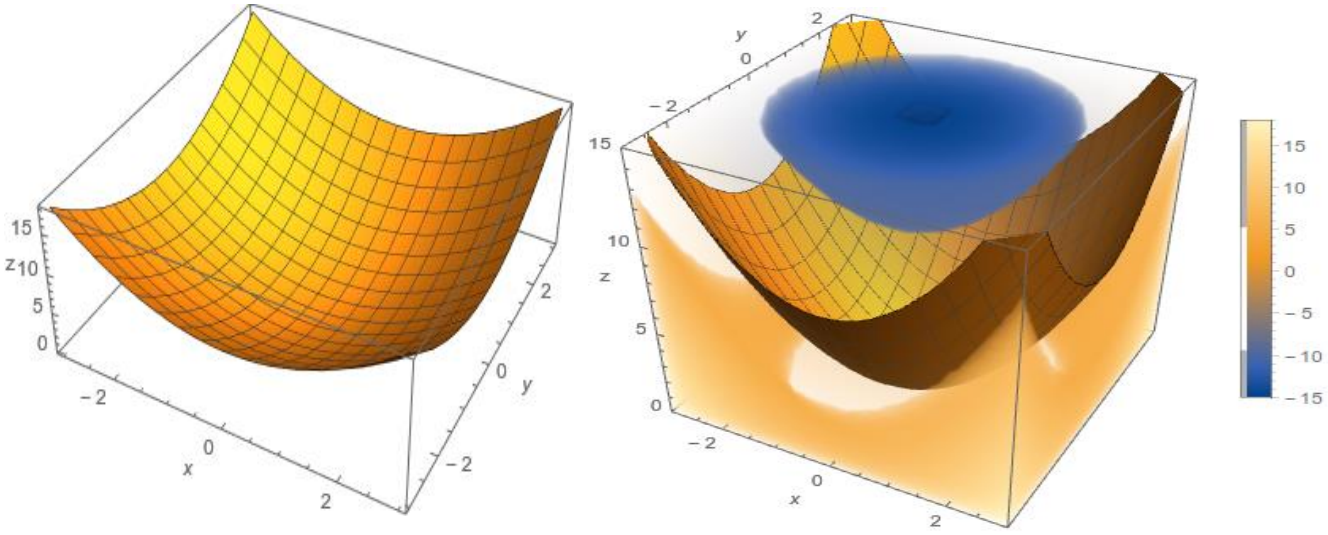
Answer:

$$E = -\boldsymbol{\nabla}\phi = \frac{C(x+y)}{(x^2+y^2)^{3/2}} = \frac{Cr}{|r|^3}$$

We have shown here explicitly that the electric field (-gradient) is always in the radial direction, perpendicular to the (circular) lines of electric equipotential. The figures below plot the form of the potential and the derived gradient (recall, **E** is in the opposite direction to the gradient).



A 3D scalar field can be considered to be a set of nested surfaces, to which the gradient is the normal. Consider the function $f(x,y,z) = x^2 + y^2 - z$,where is the surface $f(x,y,z) = 0$? We can visualise this by plotting $z = x^2 + y^2$ (figure below left). This can be seen as a surface embedded in the 3D scalar field that is $f(x,y,z)$ (figure below right).



By taking the gradient and evaluating it at a given point, the unit normal to a surface can be easily obtained, as in the following example.

Example 4.2 Find the unit normal to the surface $x^2 + y^2 - z = 0$ (illustrated above) at the point (1,1,2)

Answer: This surface is an equipotential of the 3D scalar field $f(x, y, z) = x^2 + y^2 - z$

in the case that $f(x, y, z) = 0$

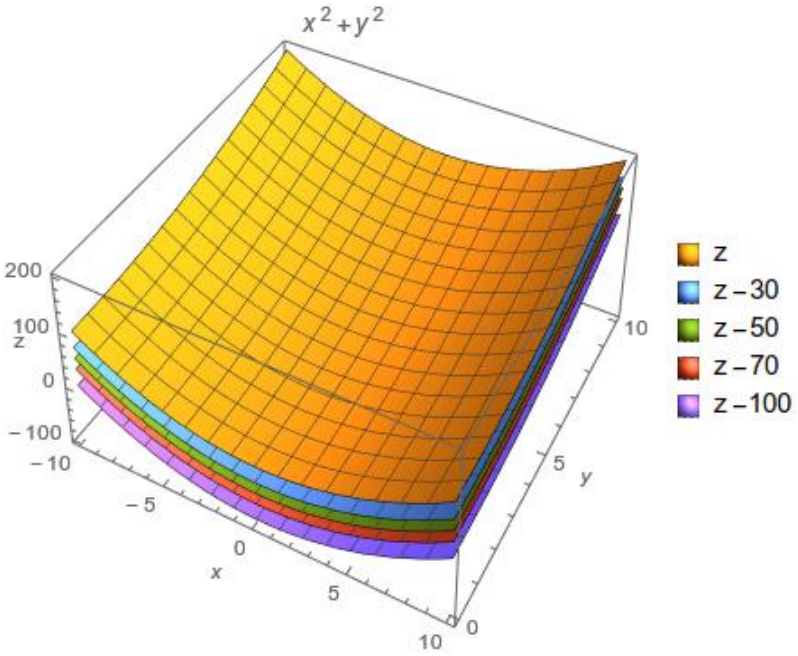
We can take the gradient of this field and divide it by its magnitude to get a unit vector perpendicular to the surface:

$\boldsymbol{\nabla} f(x, y, z) = 2x\boldsymbol{i} + 2y\boldsymbol{j} - \boldsymbol{k}$ and $|\nabla f| = \sqrt{(2x)^2 + (2y)^2 + 1}$

At the point (1,1,2), $\nabla f(x, y, z) = 2\boldsymbol{i} + 2\boldsymbol{j} - \boldsymbol{k}$ and $|\boldsymbol{\nabla} f| = 3$ so the unit normal

$\boldsymbol{n} = \frac{2\boldsymbol{i}+2\boldsymbol{j}-\boldsymbol{k}}{3}$ (note that -**n** is also a unit normal).

Note also that the fact of $f(x, y, z) = 0$ does not turn out to be relevant to the result. This is because this is just one of a set of nested surfaces which obey $f(x, y, z) = constant$ (see figure below). This is only true in this case because the normal has no dependence on z.



The figure left shows the nested set of surfaces which correspond to

$$x^2 + y^2 - z = constant$$

Directional derivatives

The directional derivative is the rate of change of the scalar field in some given direction, say, **s**.

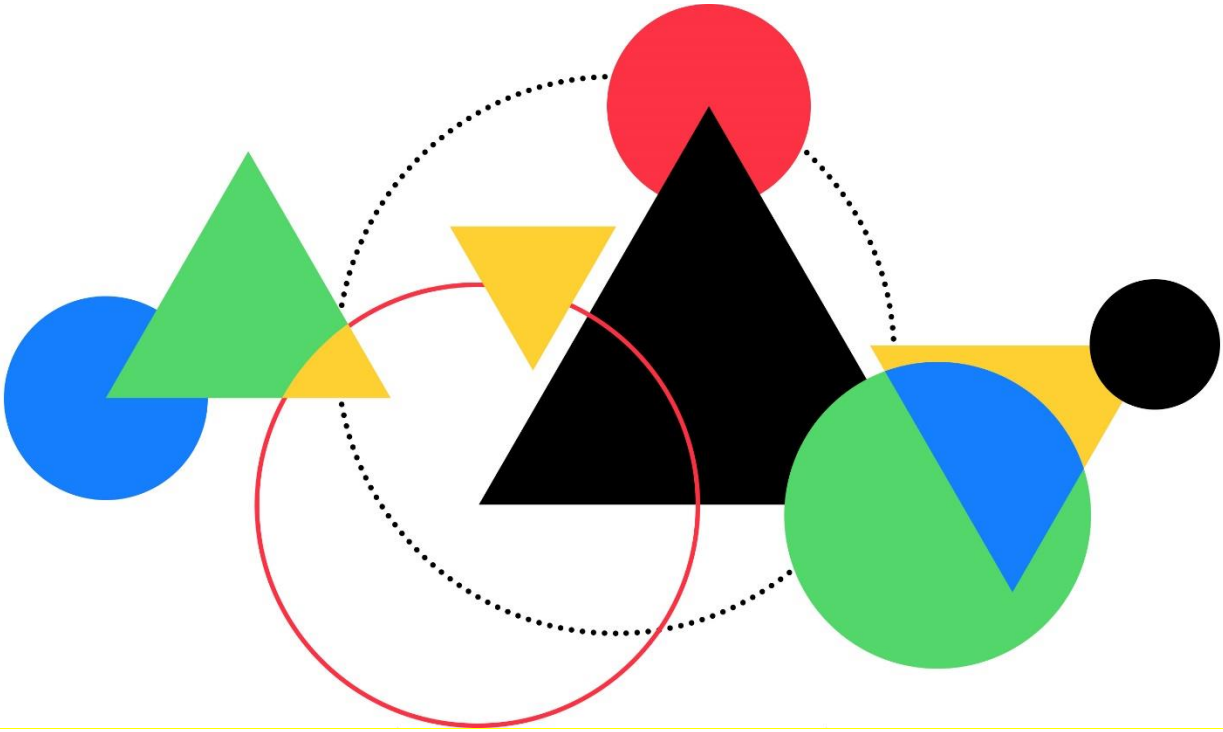
This is simply the dot product of the gradient at that point with the unit vector in the direction **s.**

$$\nabla_s \phi = \nabla \phi . \hat{\boldsymbol{s}}$$

Note the use of the directional derivative in the proof that the gradient points in the direction of greatest rate of change.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



A brief introduction to statistical distributions

3.ii

An introduction to statistical distributions

Continuous distributions, PDF and CDF

The statistical distributions we consider in this small addition to CP4 (since 2019) are of the form where a variable X may be drawn from a population described by a statistical distribution. This *continuous random variable* X (discrete distributions also exist) has a range in the form of an interval or a union of non-overlapping intervals on the real line (possibly the whole real line). The variable x is an instance of this distribution. Note that we can only evaluate the probability of x taking some value in a given range; formally the probability of X taking some precise value x is zero: P(X=x) = 0.

The probability of x taking some value in the range x to x+dx is given by the integral of the *Probability Density Function* (PDF) over that range. The *Cumulative Distribution Function* (CDF) as a function of x is the integral of the PDF from zero to the value of x. The following formal statement relates the CDF and PDF:

Consider a continuous random variable X with an absolutely continuous CDF, $F_X(x)$.

The function $f_X(x)$ *defined by* $f_X(x) = \frac{d(F_X(x))}{dx}$ *(if* $F_X(x)$ *is differentiable at x) is called the probability density function (PDF) of X.*

In other words, the PDF is the derivative of the CDF with respect to x and the CDF is the integral of the PDF between zero and x.

Example: your friend tells you that she will stop by your house sometime after or equal to 1 p.m. and before 2 p.m., but she cannot give you any more information as her schedule is quite hectic. Your friend is very dependable, so you are sure that she will stop by your house, but other than that we have no information about the arrival time. Thus, we assume that the arrival time is completely random in the 1 p.m. and 2 p.m. interval.

1. What is the sample space S?

The sample space is simply S = [1,2)

(note the use of "[" to indicate equal to or greater than and ")" to indicate less than)

2. Sketch the PDF and CDF

The PDF is a uniform top hat function between 1 and 2pm, and zero outside. The CDF is the integral of this PDF to x, and as such is a slope from 0 at 1pm to unity at 2pm. The height of the PDF is determined such that this total integrated area is 1 (since a probability of 1 means that the event *must* happen in that interval).

3. What is the probability of P(1.5)? Why?

P(1.5) = 0, as does P(X=x) for any specific value, since for a continuous probability function, the probability has to be integrated over some interval.

4. What is the probability of T ∈ [1,1.5)?

This is equal to the value of the CDF at X = 1.5, which we can see intuitively will be 0.5. In other words there is a 50% chance that your friend will arrive in the first half of the agreed interval (since the distribution is uniform between 1 – 2pm).

5. For $1 \leq a \leq b \leq 2$, what is $P(a \leq T \leq b) = P([a,b])$?

$P(a,b) = b - a$ for $1 \leq a \leq b \leq 2$. More generally, for a uniform distribution between limits x_1 and x_2, the PDF is a top-hat function of height $(x_2-x_1)^{-1}$ since the total integrated area of the PDF across the sample space = 1. The probability to be in the range (a,b) then corresponds to $P(a,b)=(b-a)/(x_2-x_1)$, which in this case reduces to b – a since $(x_2-x_1 = 1)$.

Expectation value and Variance

For a given PDF $f_X(x)$ the expectation value of x (i.e. the mean) is given by

$$E[X] = \int x\, f_X(x) dx$$

and more generally for any function of x is given by

$$E[g(X)] = \int g\,(x) f_X(x) dx$$

Example: For $f_X(x) = x + \frac{1}{2}$ for $0 \leq x \leq 1$ and zero otherwise, find $E[X^n]$

Solution:

$$E(X^n) = \int x^n \left(x + \frac{1}{2}\right) = \frac{(3n+4)}{2(n+2)(n+1)}$$

The Variance of a distribution is given by

$$Var[X] = EX^2 - (EX)^2 = \int x^2 f_X(x) - \mu_X^2 dx$$

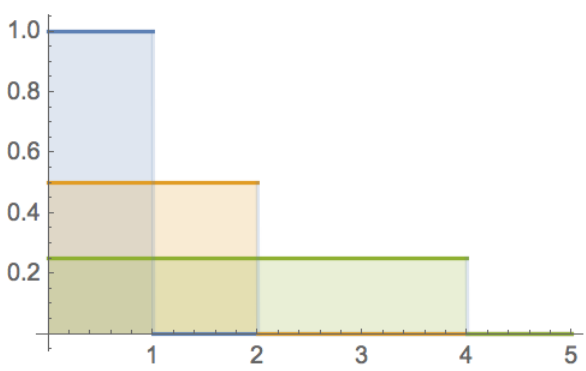
Where μ_X is the mean value (i.e. the expectation value).

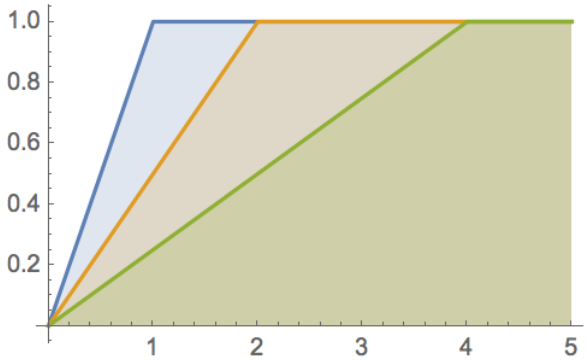
Some common statistical distributions

Uniform distribution (in all figures PDF left, CDF right)

$$PDF\ f(a < x < b) = constant, zero\ otherwise$$

Figures are differing ranges and corresponding constants such that integrated probability = 1

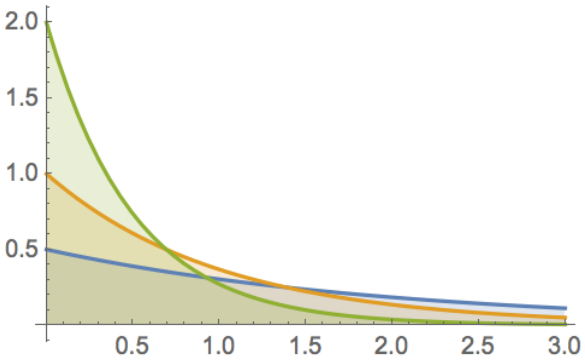


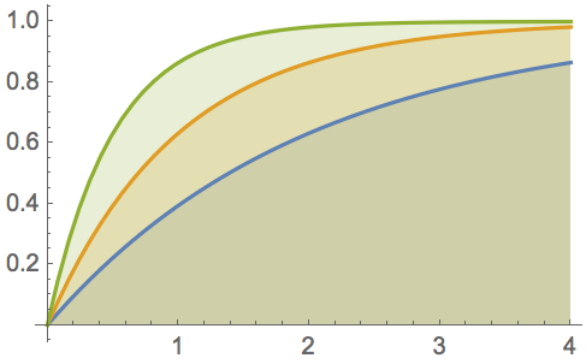


Exponential distribution (PDF left, CDF right)

$$PDF f(x) = \lambda \exp[-\lambda x]$$

Figures are differing values of λ

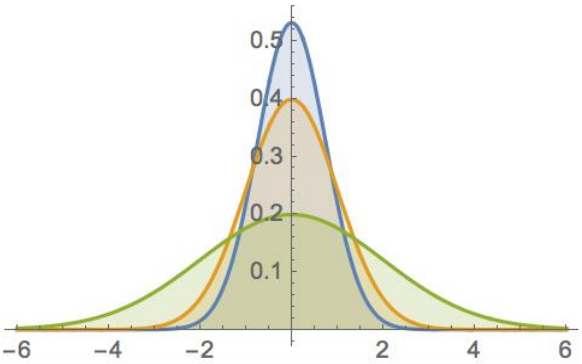


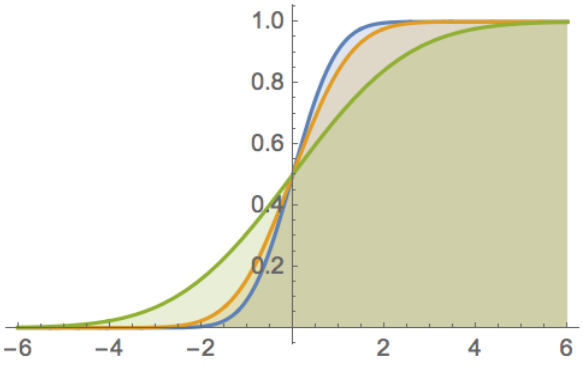


Normal (Gaussian) distribution (PDF left, CDF right)

$PDF f(x) = \frac{1}{(\sigma\sqrt{2\pi})} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$ where μ, σ^2 are the mean and variance respectively.

Figures are Gaussians with zero mean but differing variances.





Joint (2D) probability distributions

We may also consider a 2D joint probability distribution,

$$P\big((x,y) \in A\big) = f_{XY}(x,y)dxdy$$

which is normalised in an identical way to the 1D probability distributions

$$\int\int f_{XY}(x,y)dxdy = 1$$

Example: For $f_{XY}(x,y) = x + cy^2$ for ($0 \le x \le 1$ and $0 \le y \le 1$) and zero otherwise, find (a) the constant c and (b) P ($0 \le X \le 1/2$ and $0 \le Y \le 1/2$)

Solution

(a) NormaIising, using $\int\int f_{XY}(x,y)dxdy = 1$ (which is the same as above but now we know the sampled space is restricted to the range 0 – 1 for each variable), we get c = 3/2.

(b) This is a straightforward 2D integral, making use of the value of c we found in (a)

$$\int\int \left(x + \frac{3}{2}y^2\right)dxdy = \frac{3}{32}$$

We may also *marginalize* our multi-dimensional PDF in order to evaluate it simply in terms of a single variable:

$f_X(x) = \int f_{XY}(x,y)dy$ for all x, and $f_Y(y) = \int f_{XY}(x,y)dx$ for all y.

Example: Find the marginal PDFs $f_X(x)$ and $f_Y(y)$ for the 2D joint PDF in the problem above.

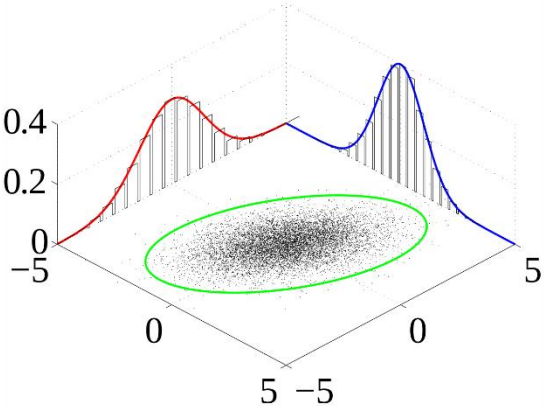
Solution

$$f_X(x) = \int \left(x + \frac{3}{2}y^2\right)dy = x + \frac{1}{2}$$

for $0 \le x \le 1$ and zero otherwise

$$f_Y(y) = \int \left(x + \frac{3}{2}y^2\right)dx = \frac{3}{2}y^2 + \frac{1}{2}$$

for $0 \leq y \leq 1$ and zero otherwise

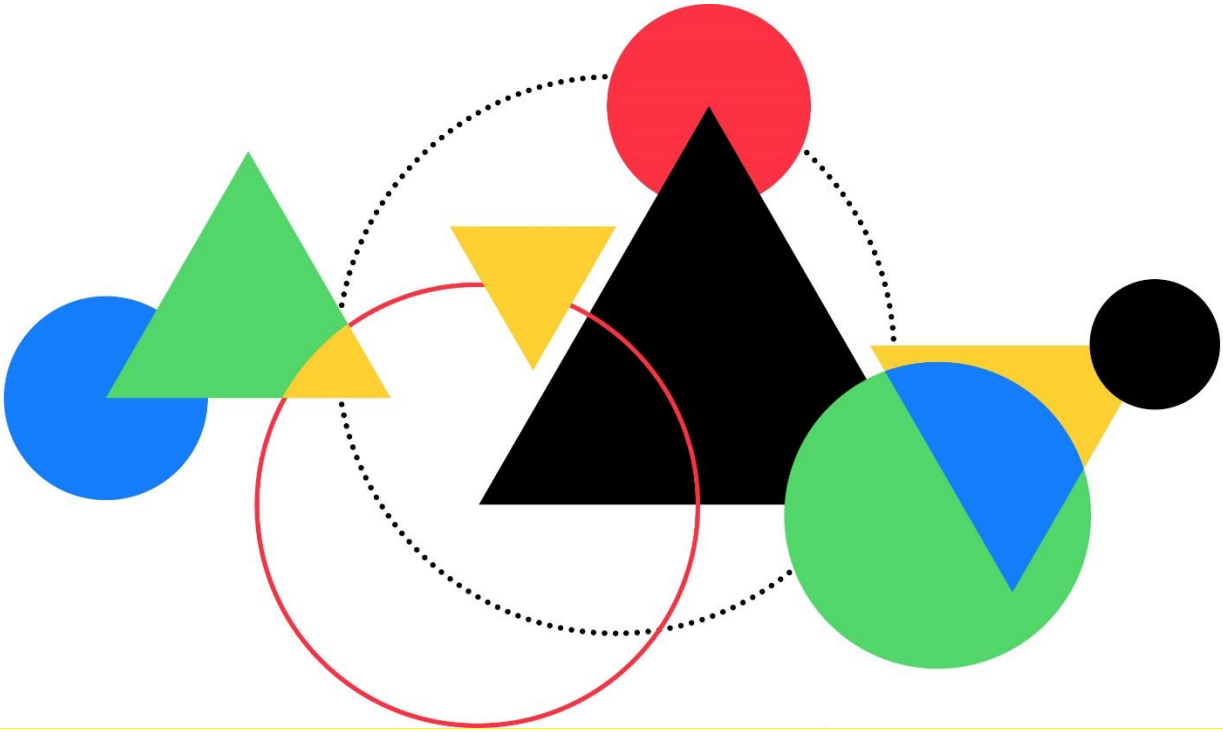


An example of a joint probability distribution measured from data (the points), and some contour of a 2D distribution (likely Gaussian) used to fit them.

The associated marginal probability distributions in x and y are indicated by the red and blue plots on the sides.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Calculating the gradient in other coordinate systems, scale factors

4.i

The gradient in other coordinate systems

We have seen the form for the Gradient in the x,y,z cartesian coordinate system, how about in other coordinate systems? We can do this individually for each transformation, but it is worth taking the time to understand the general transformation.

Consider first the transformation from one coordinate system to another, e.g. from $(x, y, z)$to (x', y', z'). Our scalar field is ψ. We can use the chain rule as follows:

$$\frac{\partial \psi}{\partial x'} = \frac{\partial \psi}{\partial x}\frac{\partial x}{\partial x'} + \frac{\partial \psi}{\partial z}\frac{\partial z}{\partial x'} + \frac{\partial \psi}{\partial y}\frac{\partial y}{\partial x'}$$

$$\frac{\partial \phi}{\partial y'} = \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial y'} + \frac{\partial \phi}{\partial z}\frac{\partial z}{\partial y'} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial y'}$$

$$\frac{\partial \psi}{\partial z'} = \frac{\partial \psi}{\partial x}\frac{\partial x}{\partial z'} + \frac{\partial \psi}{\partial z}\frac{\partial z}{\partial z'} + \frac{\partial \psi}{\partial y}\frac{\partial y}{\partial z'}$$

Which relations we see can be summarised by the following matrix relation:

$$\begin{pmatrix} \frac{\partial \psi}{\partial x'} \\ \frac{\partial \psi}{\partial y'} \\ \frac{\partial \psi}{\partial z'} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial x'} & \frac{\partial y}{\partial x'} & \frac{\partial z}{\partial x'} \\ \frac{\partial x}{\partial y'} & \frac{\partial y}{\partial y'} & \frac{\partial z}{\partial y'} \\ \frac{\partial x}{\partial z'} & \frac{\partial y}{\partial z'} & \frac{\partial z}{\partial z'} \end{pmatrix} \begin{pmatrix} \frac{\partial \psi}{\partial x} \\ \frac{\partial \psi}{\partial y} \\ \frac{\partial \psi}{\partial z} \end{pmatrix}$$

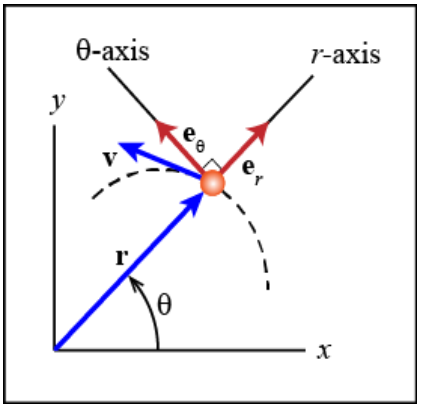
(Note is that the 3 x 3 matrix on the right hand side is the transpose of the Jacobian $\frac{\partial(x,y,z)}{\partial(x',y',z')}$)

The second thing to note is that there's a lot of repetition here, which could be simplified with some index notation (**N.B.** index notation not introduced cleanly and separately in most text books. One which does give it a dedicated chapter, however, is *Vector Calculus* by Paul C. Matthews). Index notation is not formally part of this however and I will not continue with it here.

First, let's show how we can use this approach to calculate the gradient in a new coordinate system.

Example 4.3 Calculate the gradient in plane polar coordinates

Answer: The switch from $(x, y)$to $(r, \theta)$and the associated unit vectors are shown in the diagram to the right.



Our scalar field is again ψ. We know that

$$x = r\cos\theta$$

$$y = r\sin\theta$$

and can also see that the unit vectors for plane polars are

$$\hat{\boldsymbol{r}} = \boldsymbol{i}\cos\theta + \mathbf{j}\sin\theta$$

$$\hat{\boldsymbol{\theta}} = -\boldsymbol{i}\sin\theta + \mathbf{j}\cos\theta$$

Using the matrix approach outlined above, we see that

$$\begin{pmatrix} \frac{\partial\psi}{\partial r} \\ \frac{\partial\psi}{\partial\theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial\theta} & \frac{\partial y}{\partial\theta} \end{pmatrix} \begin{pmatrix} \frac{\partial\psi}{\partial x} \\ \frac{\partial\psi}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} \frac{\partial\psi}{\partial x} \\ \frac{\partial\psi}{\partial y} \end{pmatrix}$$

We actually want to replace $\frac{\partial\phi}{\partial x}$ and $\frac{\partial\psi}{\partial y}$ in our expression for the gradient, so we use the inverse of the transformation matrix to get these:

[remember in matrix algebra, if AB=P then $A=PB^{-1}$, see e.g. Riley 8.10]

$$\begin{pmatrix} \frac{\partial\psi}{\partial x} \\ \frac{\partial\psi}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \frac{-1}{r}\sin\theta \\ \sin\theta & \frac{1}{r}\cos\theta \end{pmatrix} \begin{pmatrix} \frac{\partial\psi}{\partial r} \\ \frac{\partial\psi}{\partial\theta} \end{pmatrix}$$

(note that we already knew the determinant of the original transformation matrix, namely r, as it is the Jacobian for transforming from (x, y) to (r, θ) coordinates which we used in example 2.4.)

So now we state the expression for the gradient in two dimensions, and insert our substitutions.

$$\boldsymbol{\nabla}\phi = \left(\boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y}\right) = \left(\cos\theta\frac{\partial\phi}{\partial r} - \frac{1}{r}\sin\phi\frac{\partial\phi}{\partial\theta}\right)\boldsymbol{i} + \left(\sin\theta\frac{\partial\phi}{\partial r} + \frac{1}{r}\cos\theta\frac{\partial\phi}{\partial\theta}\right)\boldsymbol{j}$$

Rearranging... and recalling our earlier definitions of $\hat{\boldsymbol{r}}$ and $\widehat{\boldsymbol{\theta}}$ we get:

$$\boldsymbol{\nabla}\phi = \frac{\partial\phi}{\partial r}(\boldsymbol{i}\cos\theta + \boldsymbol{j}\sin\theta) + \frac{\partial\phi}{\partial\theta}\left(\frac{1}{r}\right)(-\boldsymbol{i}\sin\theta + \boldsymbol{j}\cos\theta)$$

$$\nabla\phi = \hat{\boldsymbol{r}}\frac{\partial\phi}{\partial r} + \widehat{\boldsymbol{\theta}}\left(\frac{1}{r}\right)\frac{\partial\phi}{\partial\theta}$$

which is the expression for the gradient in plane polar coordinates

Scale factors

There is another way to think about calculating the terms of the gradient in a new coordinate system, which we will approach more intuitively (and less rigorously) here:

<u>The component of **∇**f in the direction of any variable will be the partial derivative of f with respect to that variable, divided by the scale factor for that variable.</u>

Let's see what this means in the case of the gradient in plane polars which was derived above. We know that in plane polar coordinates if the position vector changes by dr in the radial direction then that point has moved by dr (in the direction $\hat{\boldsymbol{r}}$). However, if angle changes by $d\theta$ then the distance moved is $rd\theta$ (in the direction $\widehat{\boldsymbol{\theta}}$ at an infinitesimal scale), i.e. a change in angle results in a movement which is the change in the variable multiplied by the length r. This is known as the scale

factor. The inverse of this scale factor is the $\frac{1}{r}$ before the $\frac{\partial\psi}{\partial\theta}$ term in our derived expression for the gradient in plane polar coordinates.

The gradient in cylindrical and spherical polar coordinates

We have provided the tools (in fact more than one set) to allow the derivation of the expressions for the gradient in any (3D) coordinate system. Below we give the expressions in cylindrical and spherical polar coordinates, and encourage the students to confirm them via both the matrix and scale factor approaches (more of which soon).

The gradient in cylindrical coordinates is given by:

$$\boldsymbol{\nabla}\psi = \hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\left(\frac{1}{r}\right)\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{z}}\frac{\partial\psi}{\partial z}$$

The gradient in spherical polar coordinates is given by:

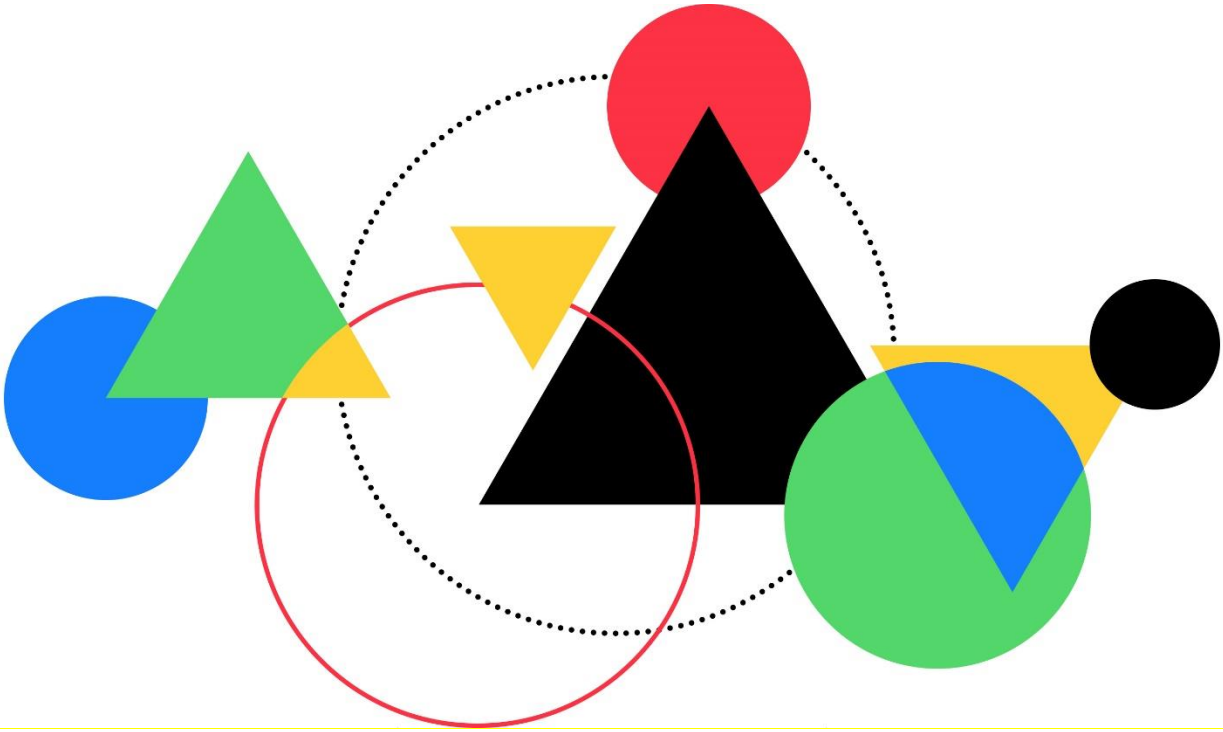
$$\boldsymbol{\nabla}\psi = \hat{\boldsymbol{r}}\ \frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\left(\frac{1}{r}\right)\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{\varphi}}\left(\frac{1}{r\sin\theta}\right)\frac{\partial\psi}{\partial\phi}$$

These formulae are given in the exam data and formula sheet, and do not need to be learned.

Important: while in plane polars and cylindrical, there is only one angle and hence no ambiguity, in spherical polars it is important to be absolutely clear to which angle you are referring, polar or azimuthal. In the above expression, the term with $\frac{1}{r}$ before it is the polar angle.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Curvilinear coordinate systems: cartesian, cylindrical and spherical polar

4.ii

Curvilinear coordinate systems

As already noted, there are three 3D coordinate systems used in this course (and in general in physics): cartesian, cylindrical and spherical polar. All are examples of *curvilinear* coordinate systems. For each system we need to be familiar with the line, surface and volume elements and well as the unit vectors.

The line element is the distance in coordinate space associated with infinitesimal change in each of the coordinates. The *surface elements* are the three different surface sizes associated with dV, the *volume element*, which is the volume enclosed by said infinitesimal change in coordinates.

Cartesian coordinates

Line element:

$$ds^2 = dx^2 + dy^2 + dz^2$$

Surface elements:

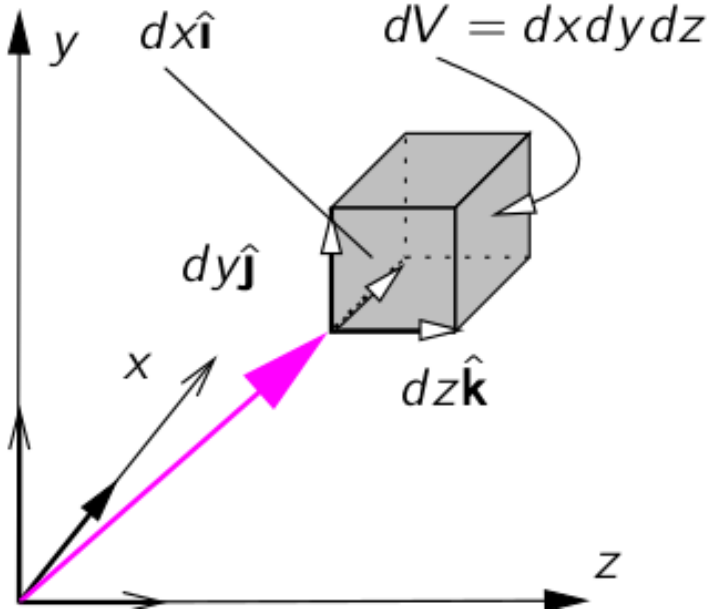
$$(dxdy, dzdx, dydz)$$

Volume element:

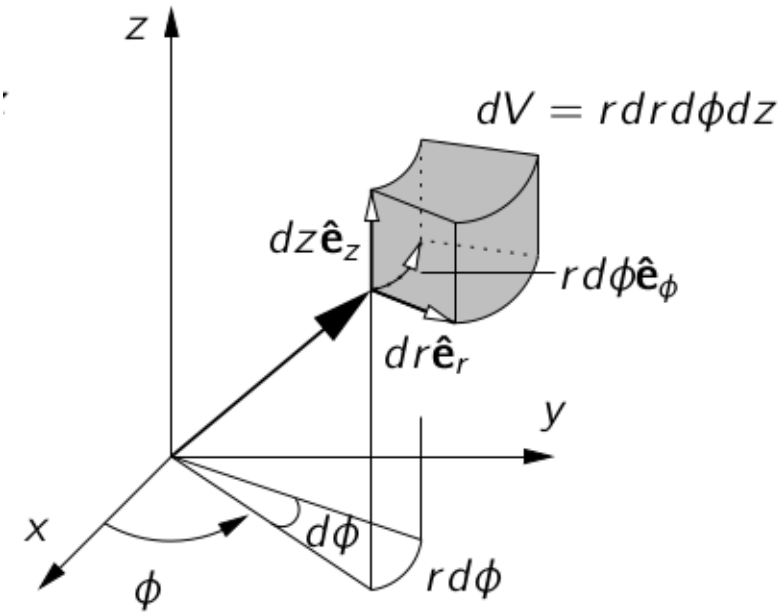
$$dV = dxdydz$$

Scale factors:

$$(1,1,1)$$



Cylindrical coordinates



Line element:

$$ds^2 = dr^2 + (rd\phi)^2 + dz^2$$

Surface elements:

$$(drdz, rd\phi dz, rd\phi dr)$$

Volume element:

$$dV = rdrd\phi dz$$

Scale factors:

$$(1, r, 1) \rightarrow \nabla\psi = \frac{\partial\psi}{\partial r}r + \left(\frac{1}{r}\right)\frac{\partial\psi}{\partial\phi}\phi + \frac{\partial\psi}{\partial z}\hat{z}$$

Notice that the **general form** of the line element is

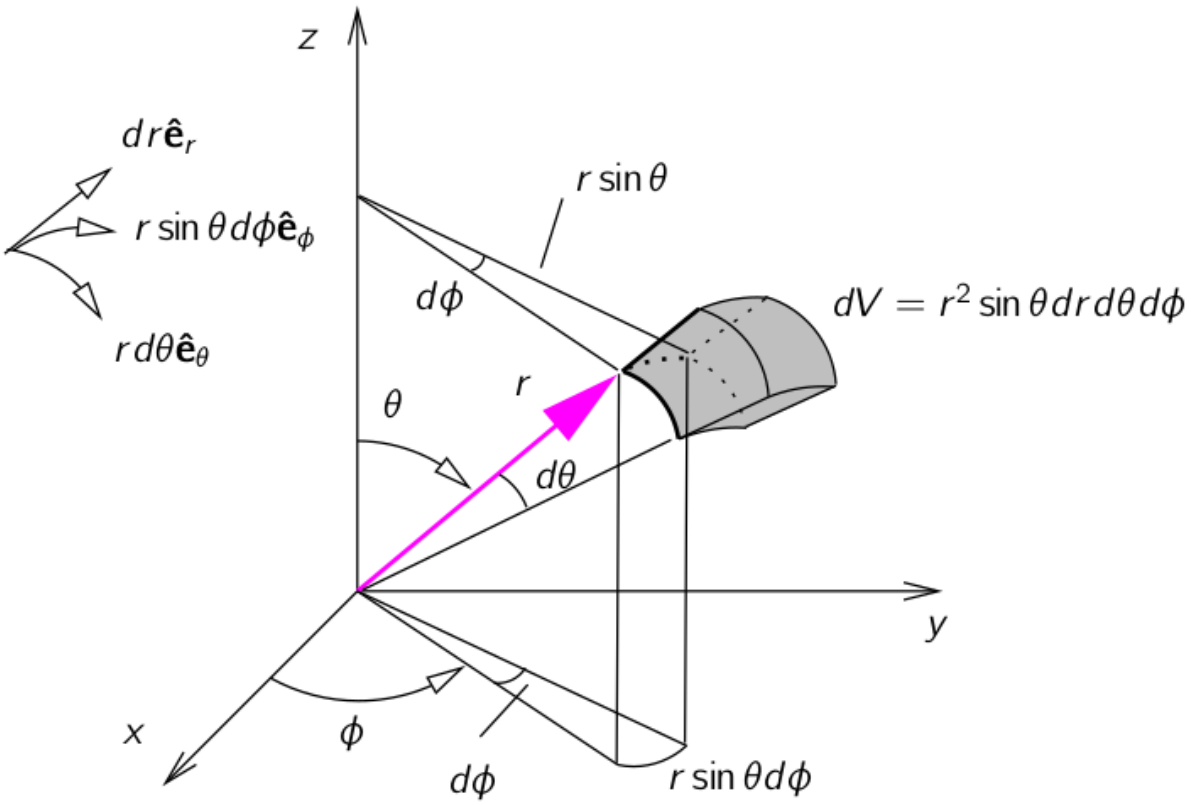
$ds^2 = f^2x_1^2 + g^2x_2^2 + h^2x_3^2$ where (f, g, h) are the scale factors.

Similarly the general form of the volume element is

$$dV = fghx_1x_2x_3$$

and, as previously noted, the form of the gradient is

$$\boldsymbol{\nabla}\psi = \left(\frac{1}{f}\right)\frac{\partial\psi}{\partial x_1}\hat{\boldsymbol{x}}_1 + \left(\frac{1}{g}\right)\frac{\partial\psi}{\partial x_2}\hat{\boldsymbol{x}}_2 + \left(\frac{1}{h}\right)\frac{\partial\psi}{\partial x_3}\hat{\boldsymbol{x}}_3$$



Spherical polar coordinates

Line element:

$$ds^2 = dr^2 + (rd\theta)^2 + (r\sin\theta d\phi)^2$$

Surface elements:

$$(rdrd\theta, r\sin\theta d\phi dr, r^2\sin\theta d\phi d\theta)$$

Volume element:

$$dV = r^2\sin\theta dr d\theta d\phi$$

Scale factors:

$$(1, r, r\sin\theta) \rightarrow \boldsymbol{\nabla}\psi = \hat{\mathbf{r}}\frac{\partial\psi}{\partial r} + \hat{\boldsymbol{\theta}}\left(\frac{1}{r}\right)\frac{\partial\psi}{\partial\theta} + \hat{\boldsymbol{\phi}}\left(\frac{1}{r\sin\theta}\right)\frac{\partial\psi}{\partial\phi}$$

Volume elements using the Jacobian

Here we have constructed the volume elements geometrically, but of course they can be calculated directly using the Jacobian, which allows us to convert from one coordinate system to another (one can think of the determinant of the 3 x 3 Jacobian as its "volume").

From earlier in the course we know that in converting between coordinate systems:

$dx\ dy\ dz = J\ du\ dv\ dw$ where $J = \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}$

Since $dxdydz = Jdudvdw$is the volume element we're looking for, we just need to calculate the Jacobian.

Exercise 6.1 Calculate directly the volume element in spherical polar coordinates using the Jacobian.

Answer First we need to know how to write x,y,z in terms of the spherical polar coordinates. Referring to the figure above, we see that:

$x = r\sin\theta\cos\phi,\ y = r\sin\theta\sin\phi, z = r\cos\theta$

The Jacobian is therefore:

$$J = \frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

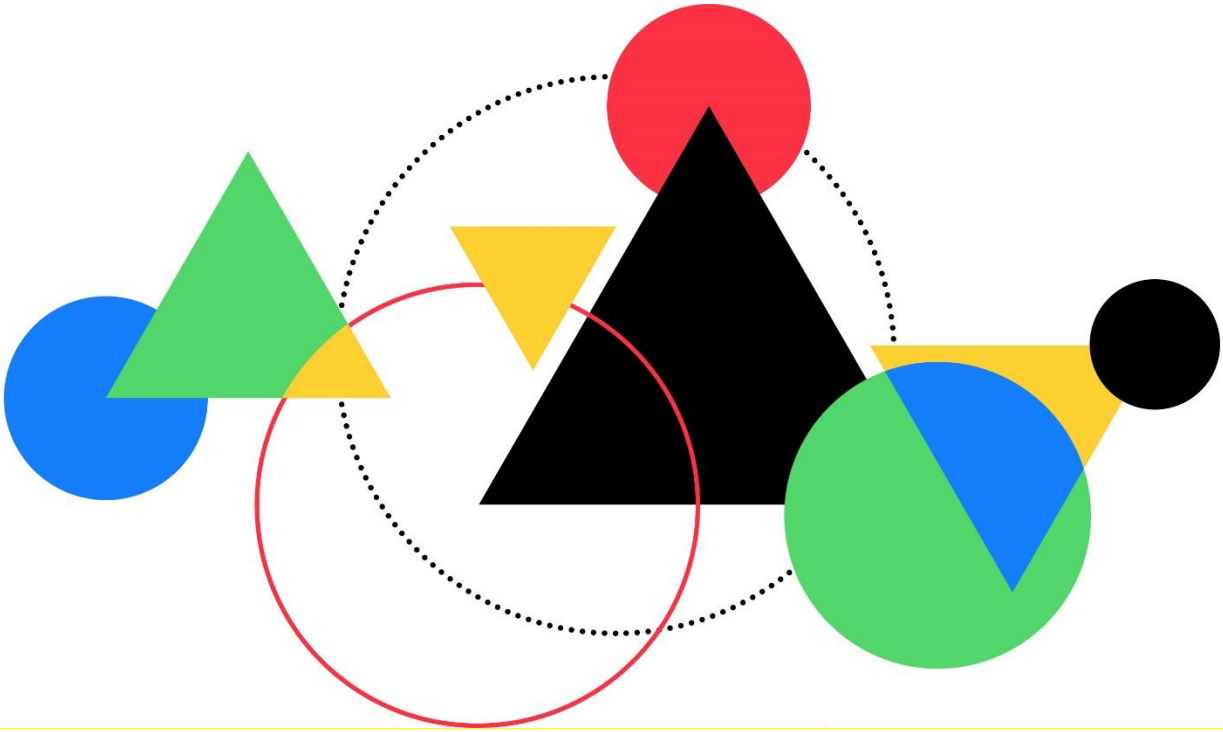
Algebra, as they say, ensues... (I recommend everyone work through this otherwise you'll never quite believe it)... and we get

$$J = \frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = r^2\sin\theta$$

and hence $dV = r^2\sin\theta dr d\theta d\phi$ as already demonstrated geometrically.

CP4: Vector Calculus and Multiple Integrals

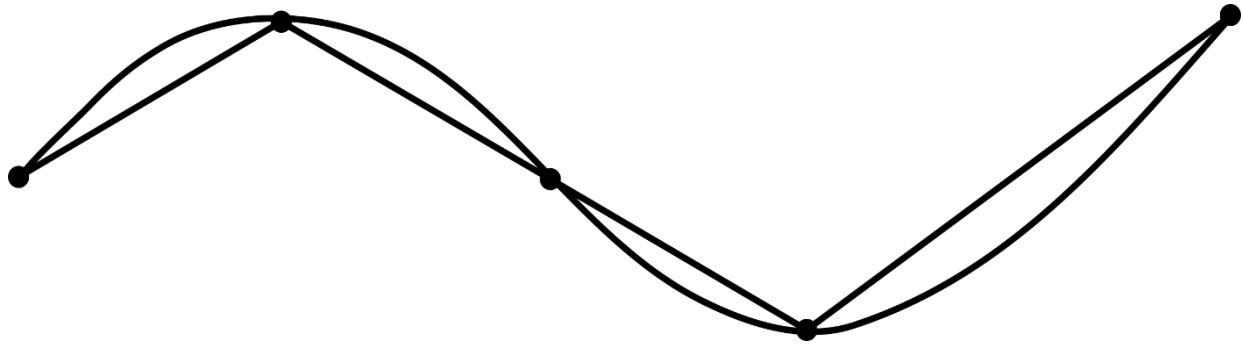
Rob Fender, HT 2020



Arc length, Differentiating vectors

5.i

Arc length



The expression for arc length in non-parametrised and parametrised curves, respectively, is given by:

$$L = \int ds = \int \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\, dx = \int \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\, dt$$

where dS is the *arc length* , equivalent to the *line element* when we were discussing coordinate systems.

For example, using the expression for the line element for plane polar coordinates,

$$ds^2 = dr^2 + (rd\theta)^2$$

we can see that the path length in plane polar coordinates for a position vector varying in length as a function of polar angle is given by

$$L = \int ds = \int \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\, d\theta$$

Differentiating vectors

Consider a function **r**(t) which describes the position of something as a function of time

$\frac{d}{dt}r(t)$= speed
$\frac{d^2}{dt^2}r(t)$= acceleration
$\frac{d^3}{dt^3}r(t)$= jerk

We differentiate a vector just like we differentiate scalar functions:

$$\frac{d}{dp}\mathbf{a}(p)=\lim_{dp\to 0}\frac{\mathbf{a}(p+dp)-\mathbf{a}(p)}{dp}$$

If we write **a** in terms of components relative to a fixed coordinate system (e.g. **i, j, k**) then:

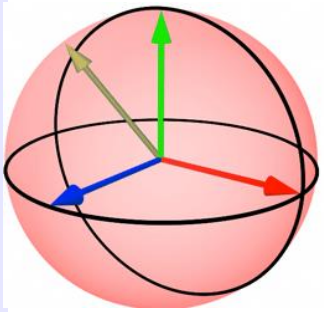
$$a(p) = a_1(p)\hat{\imath} + a_2(p)\hat{\jmath} + a_3(p)\hat{k}$$

and so

$$\frac{d}{dp}a(p) = \frac{da_1(p)}{dp}\hat{\imath} + \frac{da_2(p)}{dp}\hat{\jmath} + \frac{da_3(p)}{dp}\hat{k}$$

So: **to differentiate a vector function with respect to a fixed coordinate system, differentiate each component separately**

Example 3.1: A 3D vector **a** has constant magnitude, but is varying over time. What can you say about the direction of d**a**/dt?



Answer: Think of a radius vector tracing out the surface of a sphere – intuitively you can see that its only motion can be perpendicular to itself, as it can't get any longer or shorter.

We can demonstrate this mathematically as follows:

$\frac{d}{dt}(a\cdot a) = a\cdot\frac{da}{dt} + \frac{da}{dt}\cdot a = 2a\cdot\frac{da}{dt}$ and we know that (**a**·**a**) = a² and so $\frac{d}{dt}(a\cdot a) = 0$ which implies $2a\cdot\frac{da}{dt} = 0$ and so the rate of change of **a** must be perpendicular to **a**.

Other good news: The product rule works on dot and cross products; the chain rule also works.

As shown above the rate of change of a vector of constant magnitude is perpendicular to the vector (obvious since the length is fixed)

Parametrized curves:
If a vector **r**(p) traces some space curve, then

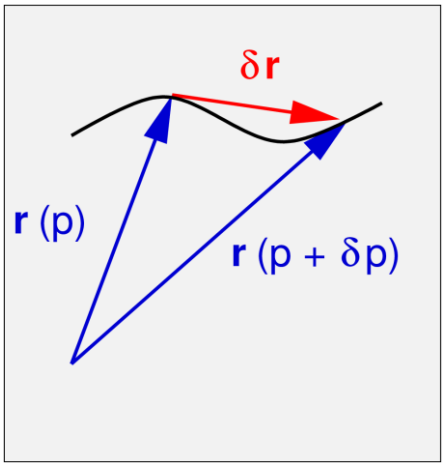
$\frac{dr(p)}{dp}$ is a tangent to the curve

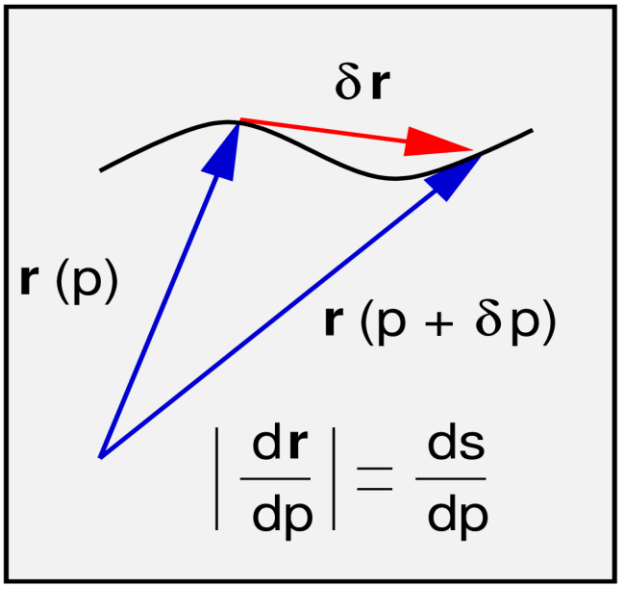
The parameter 'p' can be anything (e.g. time, height up a spiral) but there is a special case.

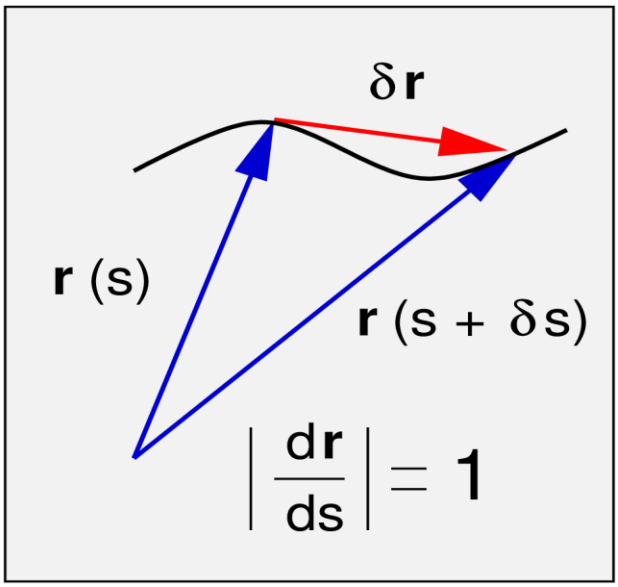
If the parameter is

s = arc length ('metric distance') then

$\frac{dr}{ds}$ is a **unit** tangent to the curve







In the figure above, the left hand figure illustrates that for a parametrised curve varying with some parameter p, the rate of change of the magnitude of the position vector **r** with p is equal to the rate of change of arc length s with p. The right hand figure illustrates the special case for which p=s.

Example 3.2: Consider the following curve (not done in lecture):

$$\boldsymbol{r} = a\cos\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{i} + a\sin\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{j} + \left(\frac{hs}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{k}$$

where a and h are constants and s is path length.

1. Sketch this curve (it is a **helix**)
2. Show that the tangent d**r**/ds to the curve has a constant elevation angle w.r.t the xy -plane, and determine its magnitude.

Answer

$$\frac{d\boldsymbol{r}}{ds} = \frac{-a}{\sqrt{(a^2+h^2)}}\sin\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{i} + \frac{a}{\sqrt{(a^2+h^2)}}\cos\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{j} + \left(\frac{h}{\sqrt{(a^2+h^2)}}\right)k$$

which shows that there is a circle projected onto the x,y plane of extent $\left(\frac{a}{\sqrt{(a^2+h^2)}}\right)$, and with z extent $\left(\frac{h}{\sqrt{(a^2+h^2)}}\right)$ and therefore rising pitch angle $\theta = \tan^{-1}\left(\frac{h}{a}\right)$

Arc-length s is special because ds = | d**r** |, but *whatever the parameter p*:

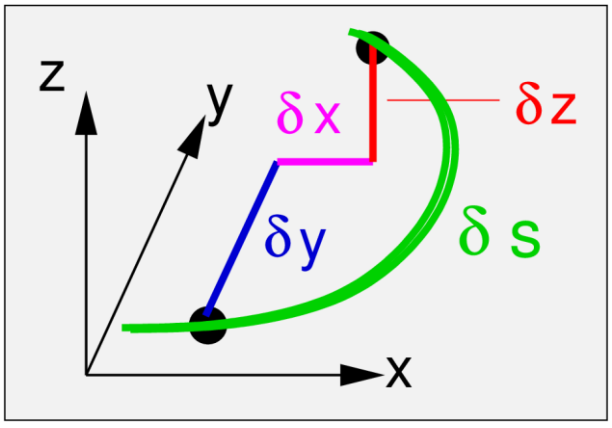
$\int_{p1}^{p2}\left|\frac{dr}{dp}\right| dp$ = accumulated path length

Using Pythagoras' theorem on a short piece of curve, in the limit where ds → 0, we have:

$$ds^2 = dx^2 + dy^2 + dz^2$$

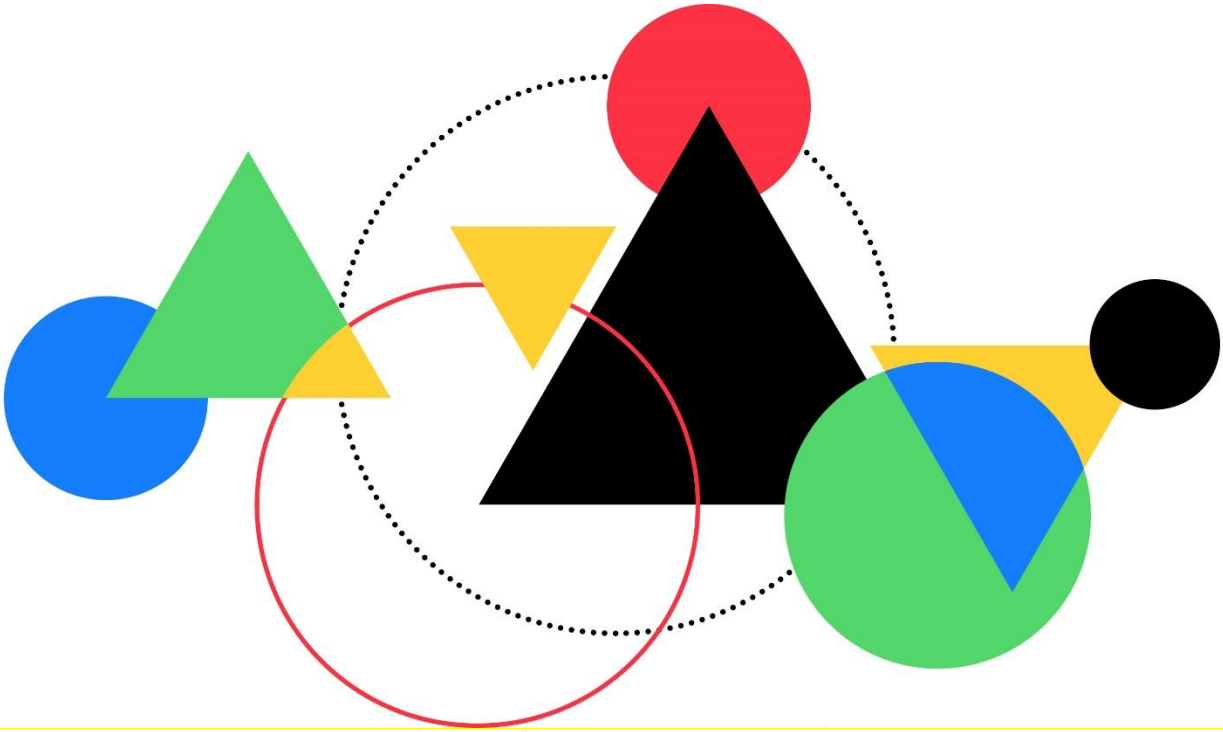
(which is the *line element* in Cartesian coordinates), and so if a curve is parametrised in terms of p:

$$\frac{ds}{dp} = \sqrt{\left[\frac{dx}{dp}\right]^2 + \left[\frac{dy}{dp}\right]^2 + \left[\frac{dz}{dp}\right]^2}$$



CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Frénet-Serret coordinate system

5.ii

The Frénet-Serret coordinate system

Can we create a coordinate system which varies with s, i.e. three mutually perpendicular directions which are intrinsic to the curve and not fixed in an external reference frame?

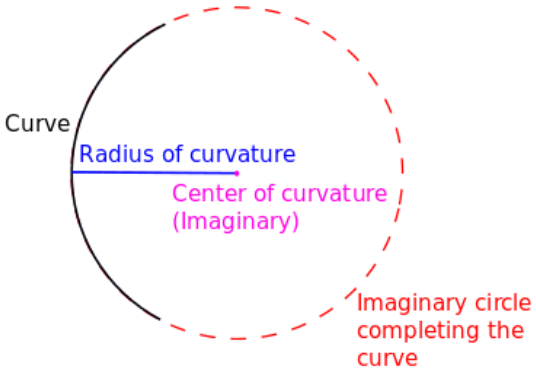
The first direction is obvious:

$\hat{\boldsymbol{t}} = \frac{d\boldsymbol{r}(s)}{ds}$= unit tangent

We already established that the derivative of vector of fixed length (in this case unity) is perpendicular to the vector, so a second direction is:

$\hat{\boldsymbol{n}} = \frac{1}{\kappa}\frac{d\hat{\boldsymbol{t}}(s)}{ds}$= the 'principal normal', where κ is the curvature.

The curvature is the reciprocal of the *radius of curvature*, which is the radius of a circle which would fit that portion of the curve (see figure right)



The plane defined by $(\hat{\boldsymbol{t}}, \hat{\boldsymbol{n}})$ is called the *osculating* plane.

The third coordinate is the cross product of the other two:

$\hat{\boldsymbol{b}} = \hat{\boldsymbol{t}} \times \hat{\boldsymbol{n}}$= the 'binormal', which acts out of the osculating plane.

This leads to a right-handed set of orthogonal unit vectors $(\hat{\boldsymbol{t}}, \hat{\boldsymbol{n}}, \hat{\boldsymbol{b}})$ such that:

$$\hat{\boldsymbol{b}} = \hat{\boldsymbol{t}} \times \hat{\boldsymbol{n}}, \quad \hat{\boldsymbol{t}} = \hat{\boldsymbol{n}} \times \hat{\boldsymbol{b}}, \quad \hat{\boldsymbol{n}} = \hat{\boldsymbol{b}} \times \hat{\boldsymbol{t}}$$

Example 3.3: Differentiate further to derive the full set of Frenet-Serret equations

Answer: Consider first the derivative of **n** with respect to s, $\frac{d\hat{n}}{ds}$,this must be perpendicular to **n** (as we showed above, the derivative with respect to s for a unit of fixed [unit] length) and so must lie in the (**t,b**) plane. So we can define in terms of some constants and the unit vectors **t** and **b**:

$$\frac{d\hat{n}}{ds} = \mu\hat{t} + \tau\hat{b}$$

Differentiating **b** with respect to s, we get

$$\frac{d\hat{b}}{ds} = \frac{d}{ds}(\hat{t} \times \hat{n}) = \frac{d\hat{t}}{ds} \times \hat{n} + \hat{t} \times \frac{d\hat{n}}{ds}$$

Since d**t**/ds is in the direction of **n**, the first term is zero. Substituting in our expression for d**n**/ds, we arrive at

$$\frac{d\hat{b}}{ds} = \hat{t} \times \ (\mu\hat{t} + \tau\hat{b}) \ = \ -\tau(\hat{\boldsymbol{b}} \times \hat{\boldsymbol{t}}) \ = \tau\hat{\boldsymbol{b}} \ - \ \boldsymbol{\kappa}\hat{\boldsymbol{t}}$$

Finally we may differentiate **n** with respect to s and substitute in previously identified terms as above, to get

$$\frac{d\hat{\boldsymbol{n}}}{ds} = \tau\hat{\boldsymbol{b}} \ - \ \boldsymbol{\kappa}\hat{\boldsymbol{t}}$$

$$\begin{pmatrix} d\boldsymbol{t}/ds \\ d\boldsymbol{n}/ds \\ d\boldsymbol{b}/ds \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{t} \\ \boldsymbol{n} \\ \boldsymbol{b} \end{pmatrix}$$

Example 3.4: Find the curvature and torsion for the helix

$$\boldsymbol{r} = a\cos\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{i} + a\sin\left(\frac{s}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{j} + \left(\frac{hs}{\sqrt{(a^2+h^2)}}\right)\boldsymbol{k}$$

Answer: Start by making the simplification that $\beta = \sqrt{(a^2+h^2)}$ so equation neatens down to:

$$\boldsymbol{r} = a\cos\left(\frac{s}{\beta}\right)\boldsymbol{i} + a\sin\left(\frac{s}{\beta}\right)\boldsymbol{j} + \left(\frac{hs}{\beta}\right)\boldsymbol{k}$$

Find unit normal first

$$\frac{d\boldsymbol{r}}{ds} = \hat{\boldsymbol{t}} = -\left(\frac{a}{\beta}\right)\sin\left(\frac{s}{\beta}\right)\boldsymbol{i} + \left(\frac{a}{\beta}\right)\cos\left(\frac{s}{\beta}\right)\boldsymbol{j} + \left(\frac{h}{\beta}\right)\boldsymbol{k}$$

$$\kappa\hat{\boldsymbol{n}} = \frac{dt}{ds} = -\left(\frac{a}{\beta^2}\right)\cos\left(\frac{s}{\beta}\right)\boldsymbol{i} - \left(\frac{a}{\beta^2}\right)\sin\left(\frac{s}{\beta}\right)\boldsymbol{j}$$

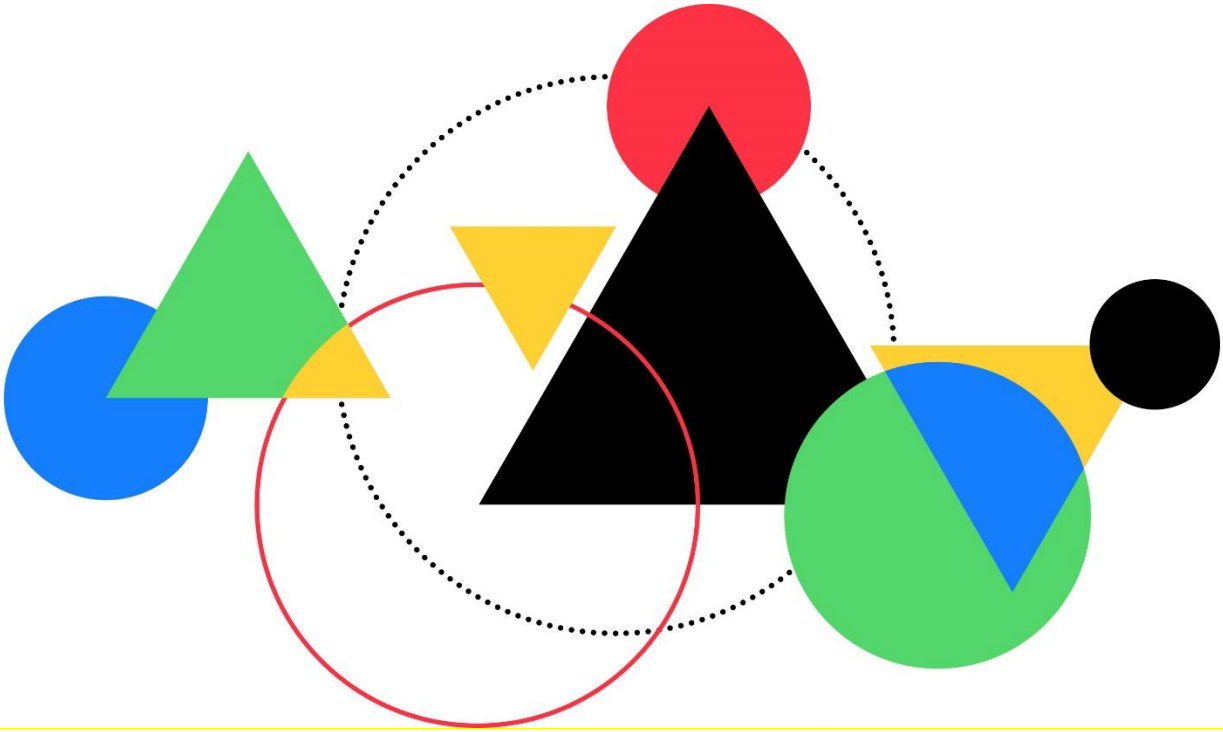
Since the curvature is a positive number and is the magnitude of **n**, we conclude

$\kappa = \left(\frac{a}{\beta^2}\right)$and $\hat{\boldsymbol{n}} = -\cos\left(\frac{s}{\beta}\right)\boldsymbol{i} - \sin\left(\frac{s}{\beta}\right)\boldsymbol{j}$

Which gives us that the curvature is constant. Next calculate **b**, its rate of change with s, and from this and the expression above we find that the torsion $\tau = (h/\beta^2)$.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Line Integrals

6.i

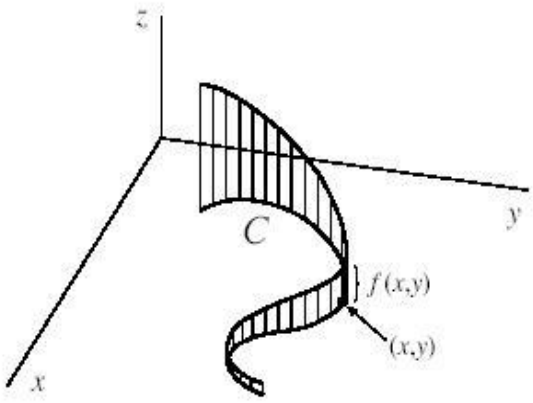
Line integrals

We can integrate things along a path C (described by a line). These 'things' can be

– Scalar fields

– Vector fields (dot product)

A *parameter* such as t is often used to describe progression along the line.



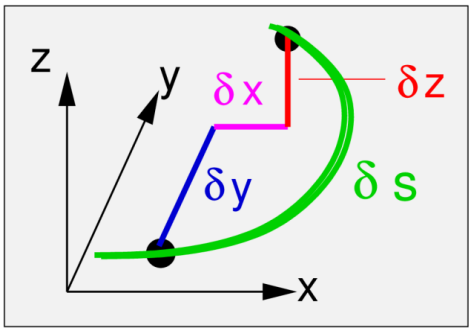
Parametrized lines and scalar line integrals

Functions describing lines are often parametrized by some variable, e.g. t, which does not have a one-to-one correspondence with line length, s. Problems are often posed in the form:

Evaluate $I = \int_C \psi\,(x,y,z)ds$, where the subscript C below the integral tells you that we are integrating along some curve, which will be specified in the problem, and $\psi(x,y,z)$ is some scalar function of position. What we need to do here is to parametrize Ψ, calculate the limits t1, t2 of the parameter which correspond to the curve C, and substitute

$$\frac{ds}{dt} = \sqrt{\left[\frac{dx}{dt}\right]^2 + \left[\frac{dy}{dt}\right]^2 + \left[\frac{dz}{dt}\right]^2}$$

which is the relation between path length and a change in the parameter.



This leads us to:

$$I = \int_C ds\,\psi(x,y,z) = \int_{t_i}^{t_f} dt \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\,\psi(t)$$

Example 6.1 Evaluate $\int_c xy^4\,dS$ where C is the right half of the circle $x^2+y^2=16$

Answer: This question is particularly straightforward because it is only in 2D. First lets parametrize this:

$x=4\cos t, y=4\sin t$. For the curve C as described, $-\pi/2 \le t \le \pi/2$.

Now the derivatives:

$\frac{dx}{dt}=-4\sin t, \frac{dy}{dt}=4\cos t \rightarrow \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}=4$

and so $\int_c xy^4\,dS=\int_{-\pi/2}^{+\pi/2}(4\cos t)\,(4\sin t)^4(4)dt=4096\int_{-\pi/2}^{+\pi/2}\cos t\sin^4 t dt=\frac{8192}{5}$

Line integrals are sometimes given in terms of cartesian coordinates rather than parametrizations:

Example 6.2 Evaluate $\int_c \sin(\pi y)dy+yx^2dx$ where C is the line segment from (0,2) to (1,4)

This line segment is clearly a straight line of intercept 2 and gradient 2. So we see that the parametrization is:

$r(t)=ti+(2t+2)j$ for $0\le t\le 1$

So in the above parametization, we identify $x=t$ and $y=2t+2$

Please do stop and make sure you see how $r(t)$ with these parameters describes the line segment C.

So now we can separate the integral:

$\int_C \sin(\pi y)dy+yx^2dx=\int_C \sin(\pi y)dy+\int_C y\,x^2dx$

Because we want to replace the dx and dy with dt, we need to differentiate them:

$\frac{dy}{dt}=2, \frac{dx}{dt}=1 \rightarrow$ dy = 2 dt and dx = dt

$$\int_C \sin(\pi y)dy+\int_C y\,x^2dx=\int_0^1 \sin\big(\pi(2t+2)\big)(2)dt+\int_0^1 t^2\,(2t+2)(1)dt$$

Taking each integral in turn, we see that

$$2\int_0^1 \sin(2\pi t+2\pi)dt=2\left[\frac{-\cos(2\pi t+2\pi)}{2\pi}\right]_0^1=0$$

and

$$\int_0^1 2\,t^3 + 2t^2 dt = \left[\frac{t^4}{2} + \frac{2t^3}{3}\right]_0^1 = \frac{7}{6}$$

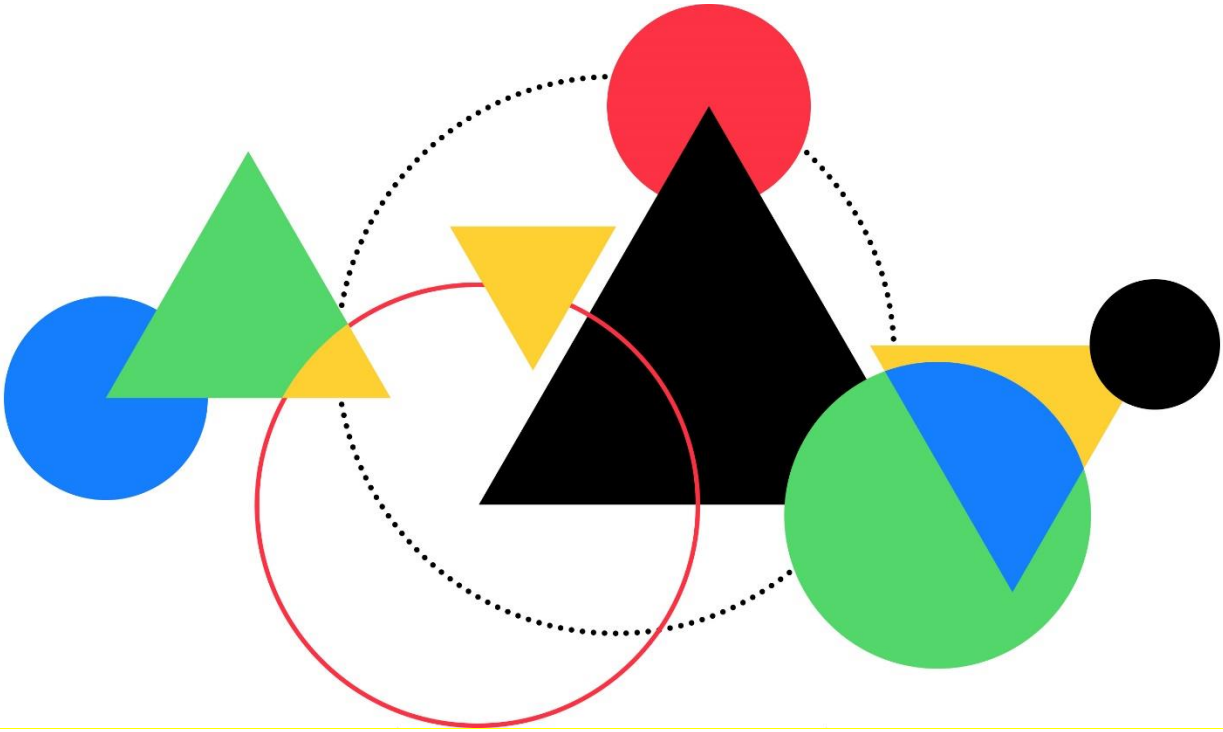
so the total value of the scalar line integral is $7/6$.

Note that for scalar line integrals such as these, the direction in which the integral is made along the curve C does not matter. However, be aware that the parameter (e.g. t) needs to be set so that it is *increasing*, since your integral is with respect to *arc length* (if you do not do this, you are integrating in a negative direction, and will get the negative of your original integral; this is the reason $\int_b^a f(x)dx = -\int_a^b f(x)dx$. even though the area under the graph is the same). Considering the counter-clockwise semi-circle described previously, to go the other way we would have to define a new coordinate system such that t was still increasing from $-\pi/2$ to $+\pi/2$ while going the other way around the loop, and so now we would have $y = -\sin t$ (x would remain the same). Test it out to make sure it makes sense to you.

For *vector* line integrals, which we look at next, the direction does matter.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020

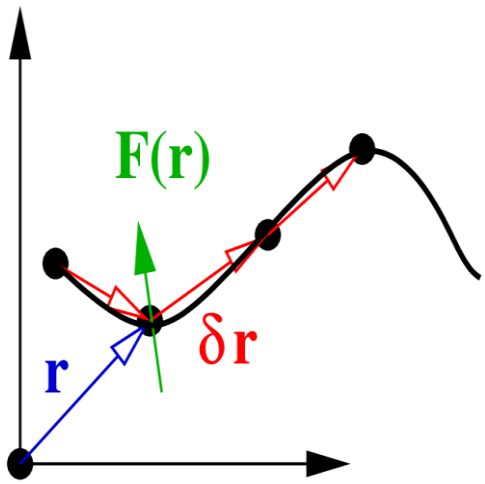


Line Integrals

6.ii

Vector line integrals

$$W = \int_A^B F \,.\, \boldsymbol{dr}$$



This kind of integral is very common in physics.

In some fields the line integral is independent of path, in others it depends on the path.

Example 6.3 A force ***F*** acts on a body as it moves between (0,0) and (1,1). The force has the form:

$$F = x^2 y\boldsymbol{i} + xy^2\boldsymbol{j}$$

Referring to the figure, find the work done when the path is
1. along the line y=x
2. along the curve $y=x^n$
3. along the x-axis to (1,0) and then up to (1,1)

Answer: First note that in planar cartesian coordinates $dr = \boldsymbol{i}dx + \boldsymbol{j}dy$ and then the work done is

$$W = \int_L F.dr = \int_L (x^2 y\hat{\imath} + xy^2\hat{\jmath}) \,.\, (idx + jdy) = \int_L (x^2 ydx + xy^2 dy)$$

PATH 1: For path y=x we note that dy=dx and so we can convert all y references to x:

$$W = \int_{(0,0)}^{(1,1)} (x^2 ydx + xy^2 dy) = \int_{(x=0)}^{(x=1)} (x^2 xdx + xx^2 dx) = \int_{(x=0)}^{(x=1)} 2\,x^3 dx = \frac{1}{2}$$

PATH 2: For the path $y=x^n$, $dy=n\,x^{n-1}\,dx$ and again convert all the y and dy references:

$$W = \int_{(0,0)}^{(1,1)} (x^2 x^n dx + xx^{2n} n x^{n-1} dx) = \int_{(x=0)}^{(x=1)} (x^{n+2} dx + n x^{3n} dx) = \frac{1}{n+3} + \frac{n}{3n+1}$$

(which we see gives the correct answer for path 1 if we set n=1)

PATH 3: This path we split into two parts. Along the x-axis part y=dy=0 so there is no work done at all in moving along the x-axis in this field. Moving vertically from (0,1) to (1,1), dx=0 so the first term of the work is again zero, but x is *not* zero, so there is some work done (but note that work along the y-axis would also have been zero)

$$W_A = \int_0^1 (1)\, y^2 dy = \frac{1}{3}$$

Clearly then – in this field at least – the work done depends upon the path taken

But this is not always the case...

Example 6.4 Repeat the following calculations but this time in terms of a different field:

$$F = xy^2 i + x^2 y j$$

Answer: You'll already be anticipating the answer, and can work it out thoroughly to test, but the point can be illustrated simply by re-evaluating path 2 in the new field.

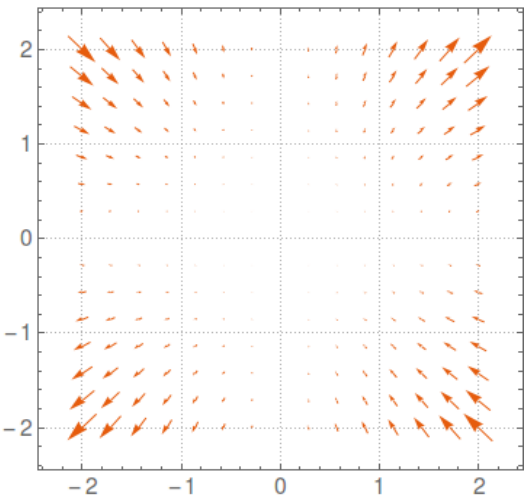
Again, for the path $y=x^n$, $dy=n\,x^{n-1}\,dx$, and again convert all the y and dy references:

$$W = \int_{(x=0)}^{(x=1)} (x^{2n+1}dx + nx^{2n+1}dx) = \int_{(x=0)}^{(x=1)} (n+1)\,x^{2n+1}dx = \frac{1}{2}$$

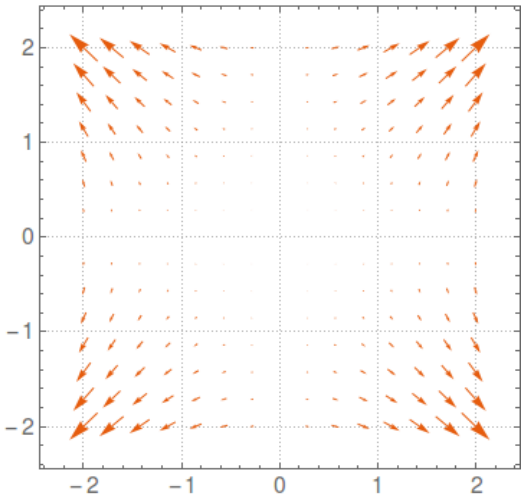
So now whatever the index of the relation between x and y, the integrated work is a constant → it is now independent of path

Here is what those fields look like:

$F = x^2 y\hat{\imath} + xy^2\hat{\jmath}$ $\quad F = xy^2\hat{\imath} + x^2 y\hat{\jmath}$



Non-conservative



Conservative

The field $F = xy^2\hat{\imath} + x^2 y\hat{\jmath}$, for which the work done is path independent, is an example of a special type of field called a conservative field.

If a field is conservative, then the integral between two points is independent of the path taken. An obvious example is a gravitational field: you know that the difference in gravitational potential energy between two positions depends only on the height (distance from the gravitating centre of mass) of the object.

If a field is conservative then clearly:

$$\oint_C F \,.\, dr = 0$$

where C is a closed path.

Conservative fields can always be written as the gradient of a scalar potential. As we have seen earlier, the electric field is the gradient of the electric potential, and so clearly the electric field is a conservative field (as is, for the same reasons, the gravitational field).

Considering this in more detail; since the integral in a conservative field is independent of path taken then:

$$\int_A^B F \,.\, dr = \psi(B) - \psi(A) = \Delta\psi$$

where Ψ is some scalar field. Since A and B are arbitrary, we can consider the case where A is infinitesimally close to B, in which case:

$$F.\, dr = \psi(B) - \psi(A) = \Delta\psi$$

But we also know from the directional derivative that:

$$\Delta\psi = \nabla\psi.\, dr$$

and hence

$$F = \nabla\psi$$

If $F = \nabla\psi$then:

$$\frac{\partial F_x}{\partial y} = \frac{\partial}{\partial y}\frac{\partial \psi}{\partial x} = \frac{\partial^2 \psi}{\partial x \partial y} = \frac{\partial}{\partial x}\frac{\partial \psi}{\partial y} = \frac{\partial F_y}{\partial x}$$

(because the second derivatives commute)

Example 6.5 Demonstrate analytically the conclusions from example 6.3, namely that

$F_1 = x^2 y\hat{\imath} + xy^2\hat{\jmath}$ is non-conservative and

$F_2 = xy^2\hat{\imath} + x^2 y\hat{\jmath}$ is conservative

Answer:

For field F_1:$F_x = x^2 y$and $F_y = xy^2$, so $\frac{\partial F_x}{\partial y} = x^2$and $\frac{\partial F_y}{\partial x} = y^2$.

Since t$\frac{\partial F_x}{\partial y} \neq \frac{\partial F_y}{\partial x}$the field is not conservative.

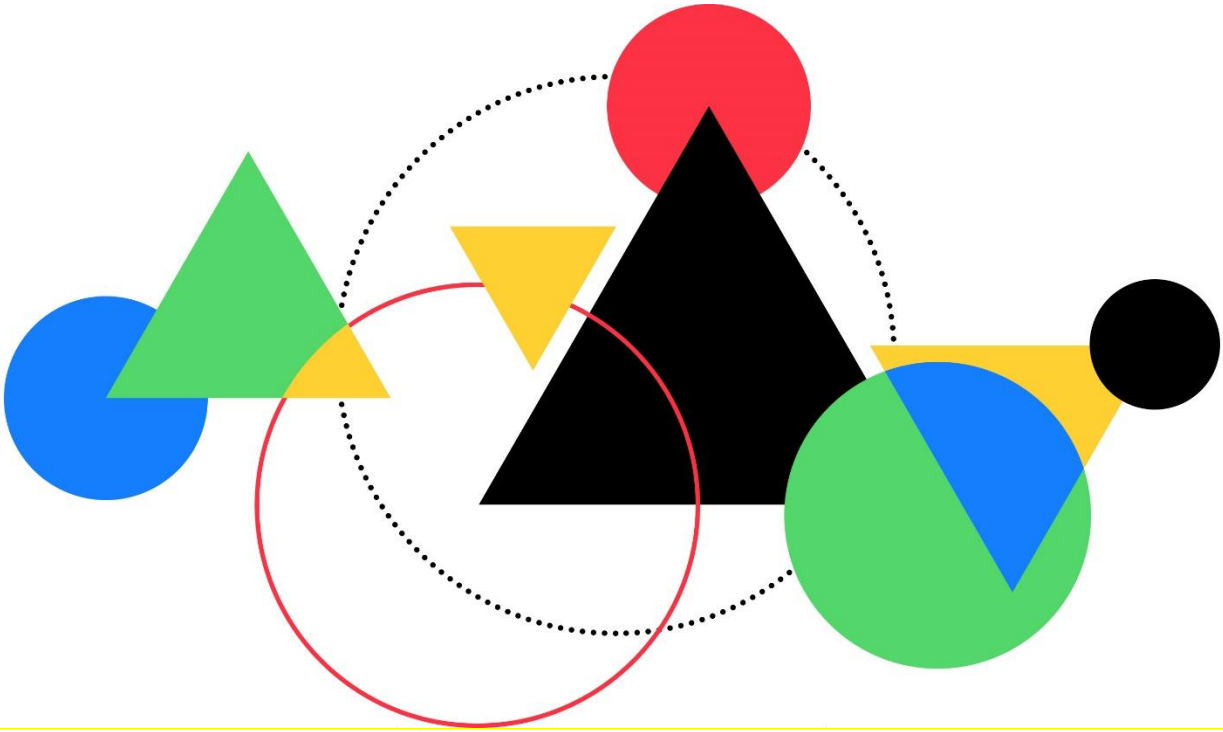
For field F_1:$F_x = xy^2$and $F_y = x^2 y$, so $\frac{\partial F_x}{\partial y} = 2xy$ and $\frac{\partial F_y}{\partial x} = 2xy$

Since t$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}$he field is conservative. Furthermore, we can directly calculate the underlying scalar field, since $\psi = \int F_x\, dx = \int F_y\, dy = \frac{x^2 y^2}{2} + C$where C is some constant.

As we will learn later, the curl of a gradient is always zero. As a result, because a conservative vector field is the gradient of a scalar field, the curl of a conservative field is always zero.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Surfaces

7.i

Surfaces

Before we move on to surface integrals, which are a fundamental part of the course, it is important to ensure we are very clear on surfaces and their geometries.

Describing surfaces: parametric and non-parametric

Surfaces may be described parametrically (e.g. $r(r,\theta)$) or non-parametrically (e.g. $g(x,y)$). Usually all that is needed is some careful thought, and the mathematical tools are straighforward to use. It is important to be able to switch between (x,y) and parametric forms.

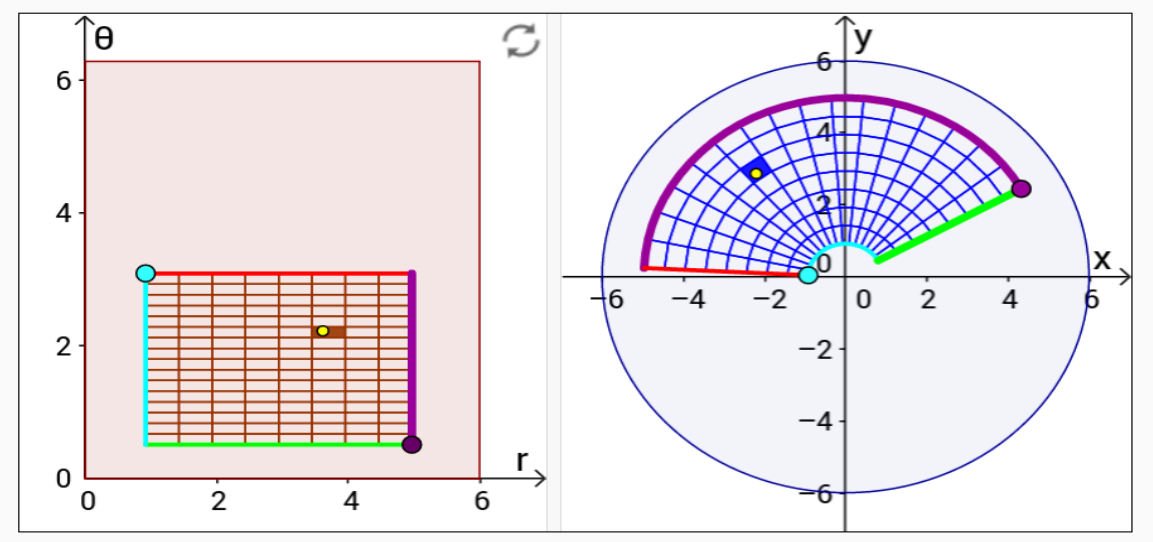
Example 7.1 Write down the parametric form (in spherical polars) of the spherical surface $x^2+y^2+z^2=30$.

Answer: We can see that the radius of the sphere is $\sqrt{30}$. We then need to convert (x,y,z) to spherical polar coordinates:

$x=r\sin\theta\cos\phi$, $y=r\sin\theta\sin\phi$, $z=r\cos\theta$ and so the parametric form is:

$$\boldsymbol{r}(\theta,\phi)=\sqrt{30}\sin\theta\cos\phi\boldsymbol{i}+\sqrt{30}\sin\theta\sin\phi\boldsymbol{j}+\sqrt{30}\cos\theta\boldsymbol{k}$$

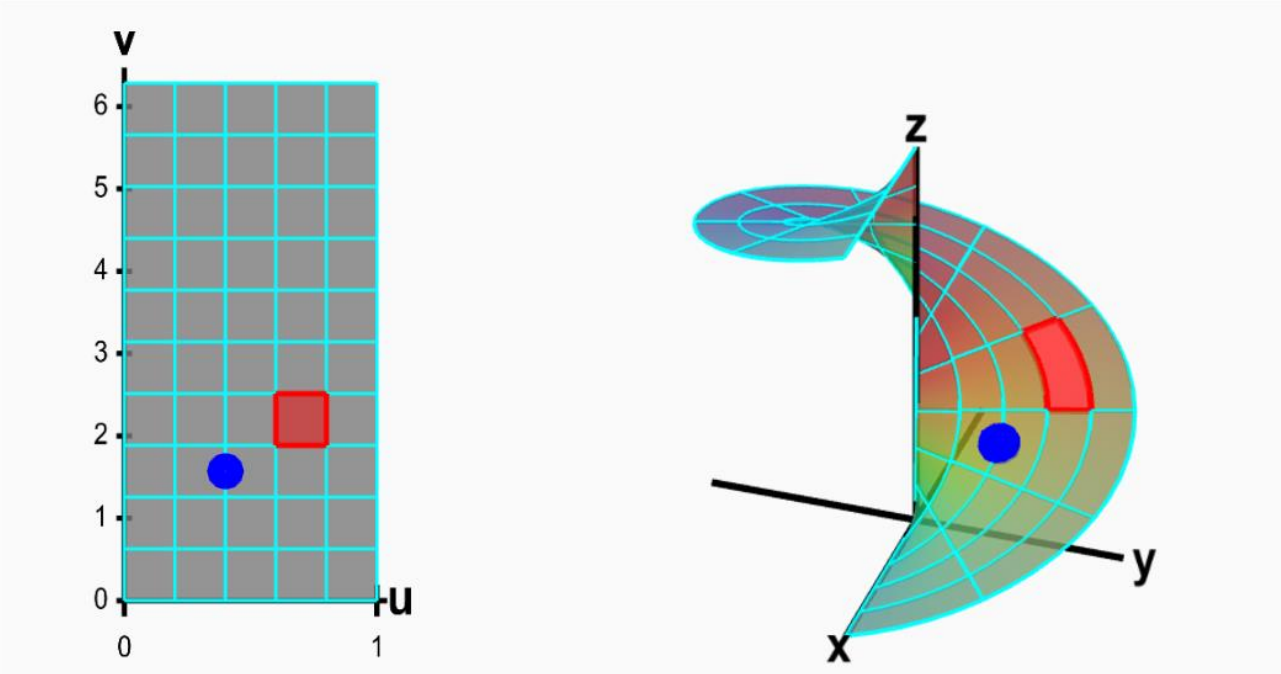
Let's be sure that we are clear about what this mapping means. We'll start by going back to the transformation between plane cartesian and plane polar coordinates. The figure below shows how a circular area in (x,y) coordinates maps onto a rectangular grid in the (r,θ) plane. Note that the coloured boundary lines of the gridded area are mapped exactly between projections. The Jacobian determinant, r, for this transformation shows how the size of the grid squares gets larger in the (x,y) plane as the radius increases.



Now let's consider the mapping which occurs between surfaces which exist in 3D space. Consider the surface

$$\boldsymbol{r}(u,v) = u\cos v\boldsymbol{i} + u\sin v\boldsymbol{j} + v\boldsymbol{k}$$

which describes a helicoid. The plot below shows this surface in the range $0 \leq u \leq 1$and $0 \leq v \leq 2\pi$, plotted in both the $(u,v)$and(x,y)planes. The blue spot and red rectangle are mapped precisely between the two coordinate systems. These figures should help you to visualise what is going on in the two coordinate systems.

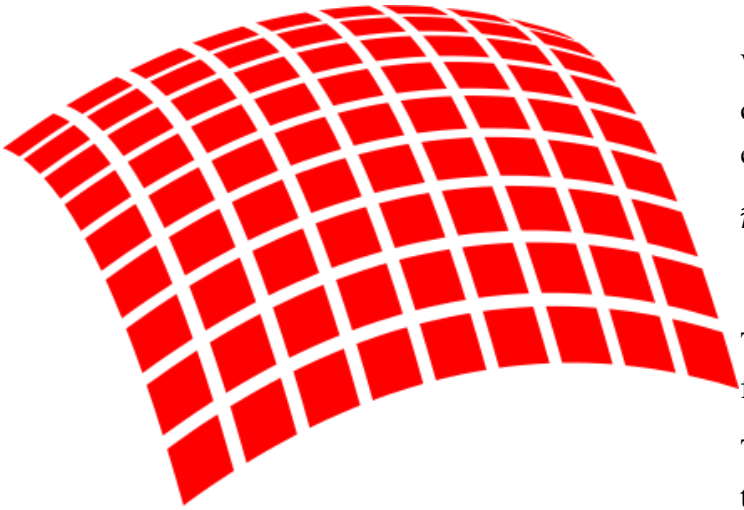


(figures gratefully borrowed from mathinsight.org)

Vector areas

A vector area is defined as an area multiplied by the normal to the surface at that point. When we go on to consider surface integrals of vector functions, we are integrating the dot product of a vector field with the vector areas.

Consider a curved surface in 3D space, such as that shown below.



We may consider that this surface comprises a large number of facets, for each of which we can define a unit normal $\hat{\boldsymbol{n}}$ and an area dS.

The corresponding *vector area* for each facet is $\boldsymbol{dS} = \hat{\boldsymbol{n}} dS$

The vector area of the entire surface is therefore $\sum_i \hat{\boldsymbol{n}}_i \, dS_i$ and in the limit where the facets become infinitesimally small, $S = \int d\,S$

For a curved (or faceted) surface, the vector area is less than the (scalar) area. As an example, consider a sphere, where the vector area is zero. Furthermore, surfaces that share a boundary must have the same vector area even if they have very different shapes and hence (scalar) areas. This is the basis of Stokes' law, which we will return to later.

Example 7.2 Calculate the vector area of a hemisphere

Answer: Again we will use spherical polar coordinates

$x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$

and use the fact that the position vector

$$\boldsymbol{r}(\theta, \phi) = r\sin\phi\cos\theta\boldsymbol{i} + r\sin\phi\sin\theta\boldsymbol{j} + r\cos\theta\boldsymbol{k}$$

will be normal to the surface, and so the unit normal will be

$$\hat{\boldsymbol{n}} = \sin\phi\cos\theta\boldsymbol{i} + \sin\phi\sin\theta\boldsymbol{j} + \cos\theta\boldsymbol{k}$$

If we consider the hemisphere to have its base in the (x,y) plane and to be projected up into the z-direction, we can see that the x- and y-components of the vector area must be zero, as they would cancel out around the circle.

We need only therefore consider the z-component. The projection in the z-direction is given by the dot product of the unit normal component in that direction, with **k**, and so is simply $\cos\theta$and so our integral is

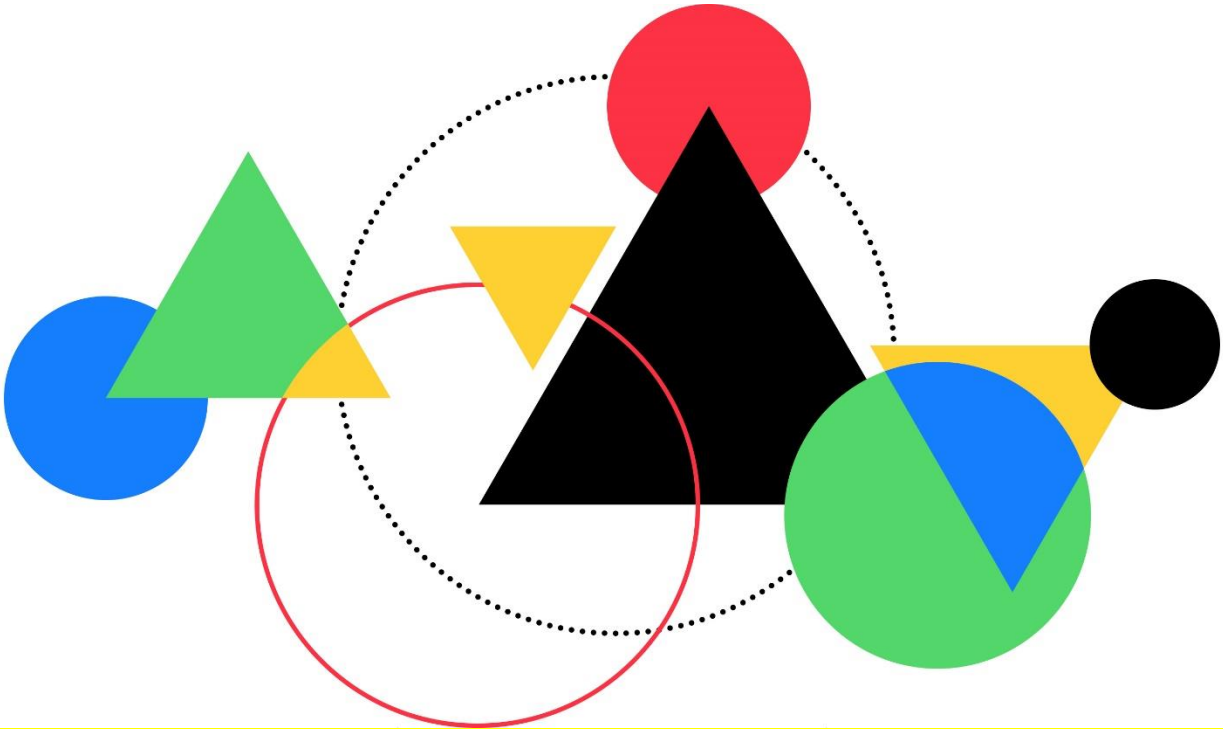
$$\int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} r^2 \sin\theta\cos\theta d\phi d\theta = \pi r^2$$

Thus we see that the vector area of a hemisphere is the area of the circle it projects in the (x,y) plane, direction perpendicular to that plane.

If we extended this analysis to a full sphere, we would see that its vector area is zero.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Surfaces

7.ii

Calculating the normal: planes defined in cartesian coordinates

Being able to determine the unit normal vector $\hat{n}$is central to being able to evaluate surface integrals. There are two ways to approach this, depending on which form (parametric or non-parametric) we are using to describe the surface.

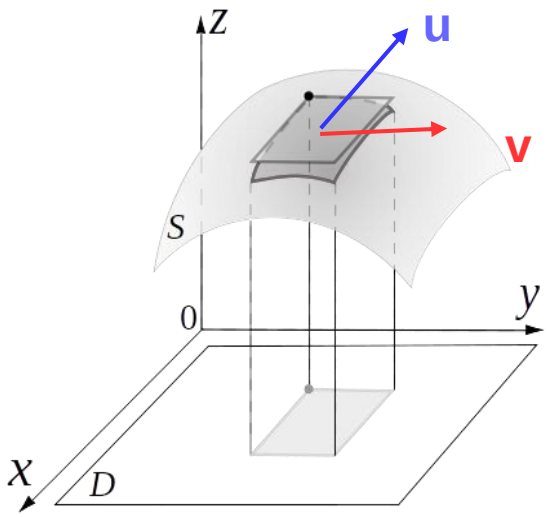
Consider first a surface which is described by $g(x, y)$. We can represent this as a constant surface of a 3D scalar field by simply doing the following:

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

We can now take the gradient of this scalar field in order to obtain the normal vector :

$$\boldsymbol{\nabla} f(x, y, z) = \boldsymbol{k} - \frac{\partial g}{\partial x}\boldsymbol{i} - \frac{\partial g}{\partial y}\boldsymbol{j} = 0$$

and the unit normal is this vector divided by its magnitude:



$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{k} - \frac{\partial g}{\partial x}\boldsymbol{i} - \frac{\partial g}{\partial y}\boldsymbol{j}}{\sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}}$$

This is quite a nasty expression, but as we shall see, the denominator usually cancels out when evaluating surface integrals. We can also derive this expression using an alternative approach.

This time we find two vectors in the tangent plane, and calculate their cross product. The tangent vectors in the x- and y-directions can be written as:

$$\boldsymbol{u} = \boldsymbol{i}u_x + \boldsymbol{k}\left(\frac{\partial g}{\partial x}\right)u_x$$

$$\boldsymbol{v} = \boldsymbol{j}v_y + \boldsymbol{k}\left(\frac{\partial g}{\partial y}\right)v_y$$

The cross product of these two vectors is normal to the tangent plane, and hence normal to the surface.

$$\boldsymbol{n} = \boldsymbol{u} \times \boldsymbol{v} = \left[-\boldsymbol{i}\left(\frac{\partial g}{\partial x}\right) - \boldsymbol{j}\left(\frac{\partial g}{\partial y}\right) + \boldsymbol{k}\right]u_x v_y$$

Normalising, we see again that:

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{k} - \frac{\partial g}{\partial x}\boldsymbol{i} - \frac{\partial g}{\partial y}\boldsymbol{j}}{\sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}}$$

Calculating the normal: planes defined parametrically

What about if the surface is given parametrically, e.g.

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

In this case we will use the cross product of the two vectors in the plane to arrive at the unit normal:

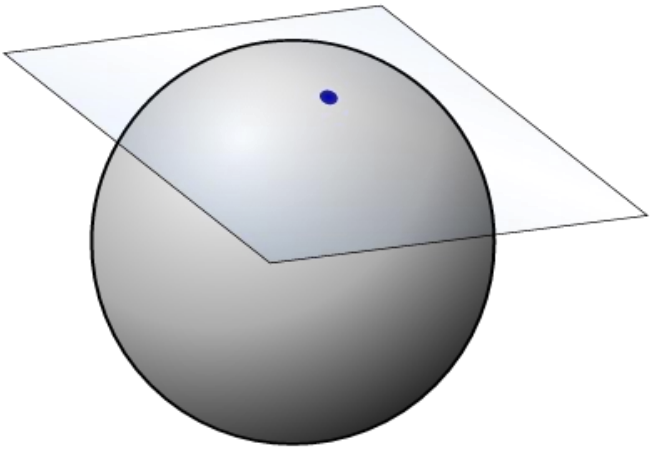
$$\hat{\boldsymbol{n}} = \frac{\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}}{\left|\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right|}$$

Tangent planes

Speaking of tangent planes, we should be clear about them. The normal to a surface at a given point is also normal to the tangent plane at that point. Calculating the tangent plane is straightforward:

1. Calculate a normal vector to the surface using one of the methods outlined above.

2. Use the fact that the dot product of the normal and a vector in the plane must be zero, i.e.



$(r - r_0).n = 0 \rightarrow (x - x_0)n_x + (y - y_0)n_y + (z - z_0)n_z = 0$ where $(x_0, y_0, z_0)$are the points at which the tangent plane is being calculated, and $(n_x, n_y, n_z)$are the components of the calculated normal vector.

Example 7.3 Find the equation of the tangent plane to the parametric surface given by $\boldsymbol{r}(u,v) = u\boldsymbol{i} + 2v^2\,\boldsymbol{j} + (u^2 + v)\boldsymbol{k}$ at the point (2,2,3).

Answer: First take the gradients of the position vector with respect to each of the parameters (u,v):

$\frac{\partial \boldsymbol{r}}{\partial u} = \boldsymbol{i} + 2u\boldsymbol{k}$ and $\frac{\partial r}{\partial v} = 4v\boldsymbol{j} + \boldsymbol{k}$

Now take the cross product to get a vector normal to the plane

$$\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v} = -8uv\boldsymbol{i} - \boldsymbol{j} + 4v\boldsymbol{k}$$

At the point (2,2,3), $\boldsymbol{r}(u,v) = u\boldsymbol{i} + 2v^2\boldsymbol{j} + (u^2 + v)\boldsymbol{k}$ implies that

$u = 2$, and $u^2 + v = 3 \rightarrow v = -1$

(note that equating the **j** term leaves ambiguity about the sign of v, so we skipped to **k**)

So the normal vector is

$$\boldsymbol{n} = 16\boldsymbol{i} - \boldsymbol{j} - 4\boldsymbol{k}$$

How do we get the tangent plane from the normal vector? The dot product of a vector in the plane and a normal vector must be zero.

$$(\boldsymbol{r} - \boldsymbol{r_0}) \cdot \boldsymbol{n} = 0$$

from which we can determine the tangent plane:

$16(x - 2) - (y - 2) - 4(z - 3) = 0 \rightarrow 16x - y - 4z = 18$

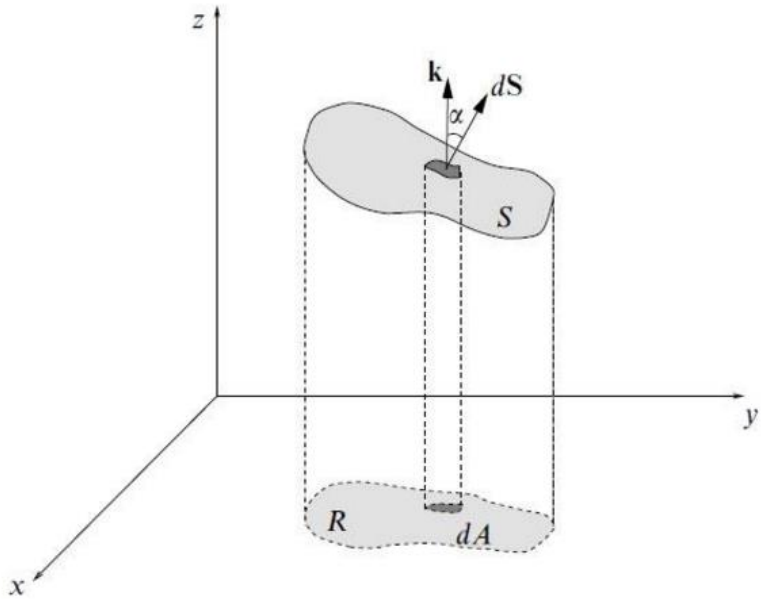
Areas and projection: cartesian surfaces

We often want to evaluate surface integrals which are provided in cartesian coordinates by projecting them onto a plane (e.g. the x,y plane), and so we want to know how a surface area dS might project onto an area dA in this plane. The projected area is $dS\cos\alpha$, i.e.

$$dS = \frac{dA}{\hat{\boldsymbol{n}}.\boldsymbol{k}}$$

Now since we know that for surfaces described in (x,y,z):

$$\hat{\boldsymbol{n}} = \frac{\boldsymbol{k} - \frac{\partial g}{\partial x}\boldsymbol{i} - \frac{\partial g}{\partial y}\boldsymbol{j}}{\sqrt{1+\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2}},$$



we see that

$$\hat{\boldsymbol{n}}.\boldsymbol{k} = \frac{1}{\sqrt{1+\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2}}$$

and therefore

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

In evaluating a surface integral projected in this way, dA would typically become dx dy (depending on the plane on to which the surface is being projected) and the problem evaluated as a standard double integral.

Areas and projection: parametric surfaces

Finally, in a similar way to calculating the Jacobian for the changing of coordinates from (x,y) to (u,v) in Lecture 2, we find that the size of the area element for a parametric surface is:

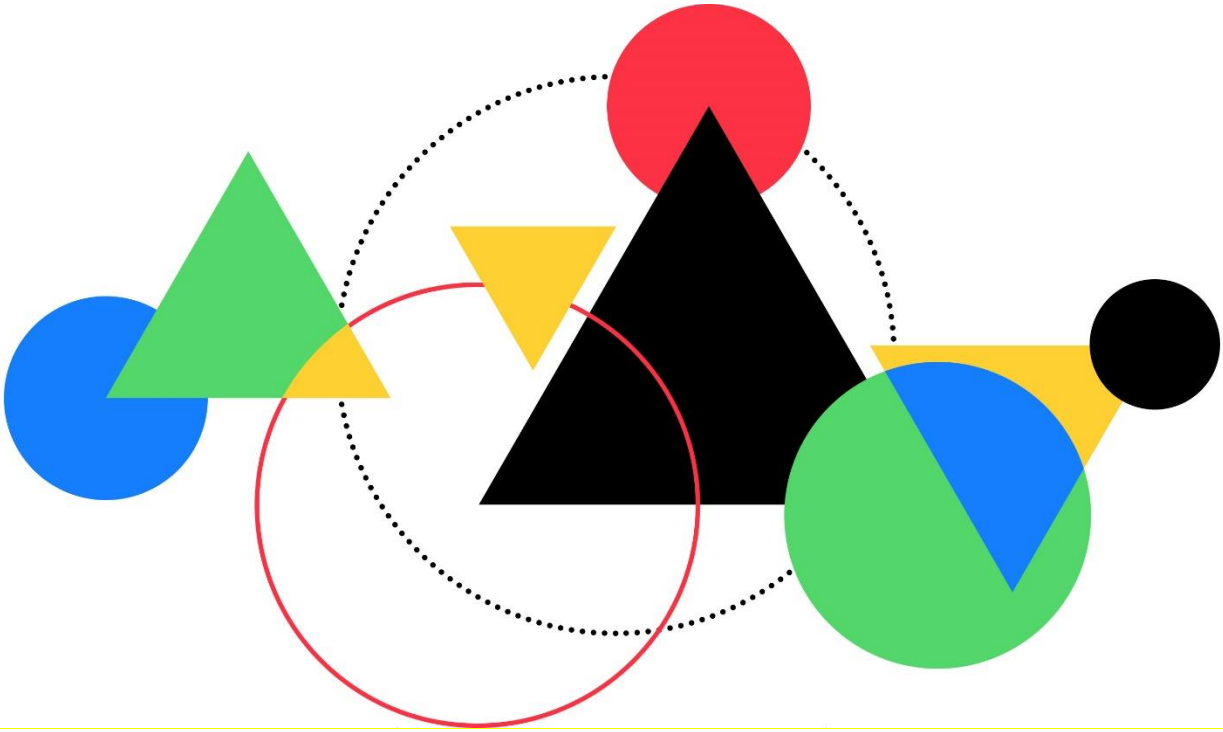
$$dS = \left|\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right| dudv$$

You will notice that this term is going to cancel with the normalisation of the normal vector.

We now have all the tools in hand to deal with surface integrals, which we will tackle next.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Surface integrals (scalar fields)

8.i

Previously, we learned all the tools which are required to be able to evaluate surface integrals. In this lecture we shall work through some examples.

Surface integral of a scalar

There are basic rules for evaluating scalar surface integrals, which results from the tools we developed in the previous lecture.

The surface integral of a scalar over a surface described by z=g(x,y) is given by:

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

where S is the surface described in the question and D is its projection on the (x,y) plane (in this case), so that dA = dx dy. The square root term is the term associated with the projected area as derived in the previous lecture.

The surface integral of a scalar over a surface described parametrically by **r**(u,v) is given by:

$$\iint_S f\,(x,y,z)dS = \iint_D f\left(\boldsymbol{r}(u,v)\right)\left|\frac{\partial \boldsymbol{r}}{\partial u}\times\frac{\partial \boldsymbol{r}}{\partial v}\right|dA$$

where the cross product term represents the area change associated with du, dv. D in this case is the area in (u,v) parameter space which describes the surface and dA = du dv.

Example 8.1 Evaluate

$\iint_S 6\,xydS$

where S is the portion of the plane $x+y+z=1$that lies in the 1st octant and is in front of the (y,z) plane.

Answer: First note that we will be projecting this from the x-axis, and so let's rewrite the plane as $x=g(y,z)=1-y-z$. The region onto which the surface is projected in the (y,z) plane is simply then the triangle bounded by the y- and z-axes and the line from (0,1) to (1,0).

So we can rewrite our surface integral as a projection onto the (y,z) plane thus:

$$\iint_S 6\,xydS = \iint_D 6\,(1-y-z)y\sqrt{1+\left(\frac{\partial g}{\partial y}\right)^2+\left(\frac{\partial g}{\partial z}\right)^2}\,dA$$

where we have substituted for x and dS.

The projection (square root) term comes out to be $\sqrt{3}$and so we get

$$\iint_D 6\,(1-y-z)y\sqrt{1+\left(\frac{\partial g}{\partial y}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2}\,dA = 6\sqrt{3}\int_0^1\int_0^{1-y}(y-y^2-zy)\,dzdy$$

... algebra ensues ...

$$6\sqrt{3}\int_0^1\int_0^{1-y}(y-y^2-zy)\,dzdy=\frac{\sqrt{3}}{4}$$

In the next example we will again evaluate a scalar surface integral, but in this case with a parametrized description of the surface:

Example 8.2 Evaluate

$\iint_S z\,dS$

where S is the upper half of a sphere of radius 2.

We have already reminded ourselves of the parametrization of a sphere, which for this radius is:

$$\boldsymbol{r}(\theta,\phi)=2\sin\theta\cos\phi\boldsymbol{i}+2\sin\theta\sin\phi\boldsymbol{j}+2\cos\theta\boldsymbol{k}$$

and in this case the range of the parameters for the upper half of a sphere is

polar angle:$0\le\theta\le\pi/2$, azimuthal angle: $0\le\phi\le2\pi$

In order to get the normal vector, we need to calculate $\frac{\partial r}{\partial\theta},\frac{\partial \boldsymbol{r}}{\partial\phi}$and the magnitude of their cross product.

... algebra ensues ...

$\left|\frac{\partial \boldsymbol{r}}{\partial\theta}\times\frac{\partial \boldsymbol{r}}{\partial\phi}\right|=4\sin\theta$ which is the projection term.

And now substituting $z=2\cos\theta$ from parametrization, and the area term:

$$\iint_S z\,dS=\iint_D 2\cos\theta(4\sin\theta)dA=\int_0^{\pi/2}\int_0^{2\pi}4\sin(2\theta)d\phi d\theta$$

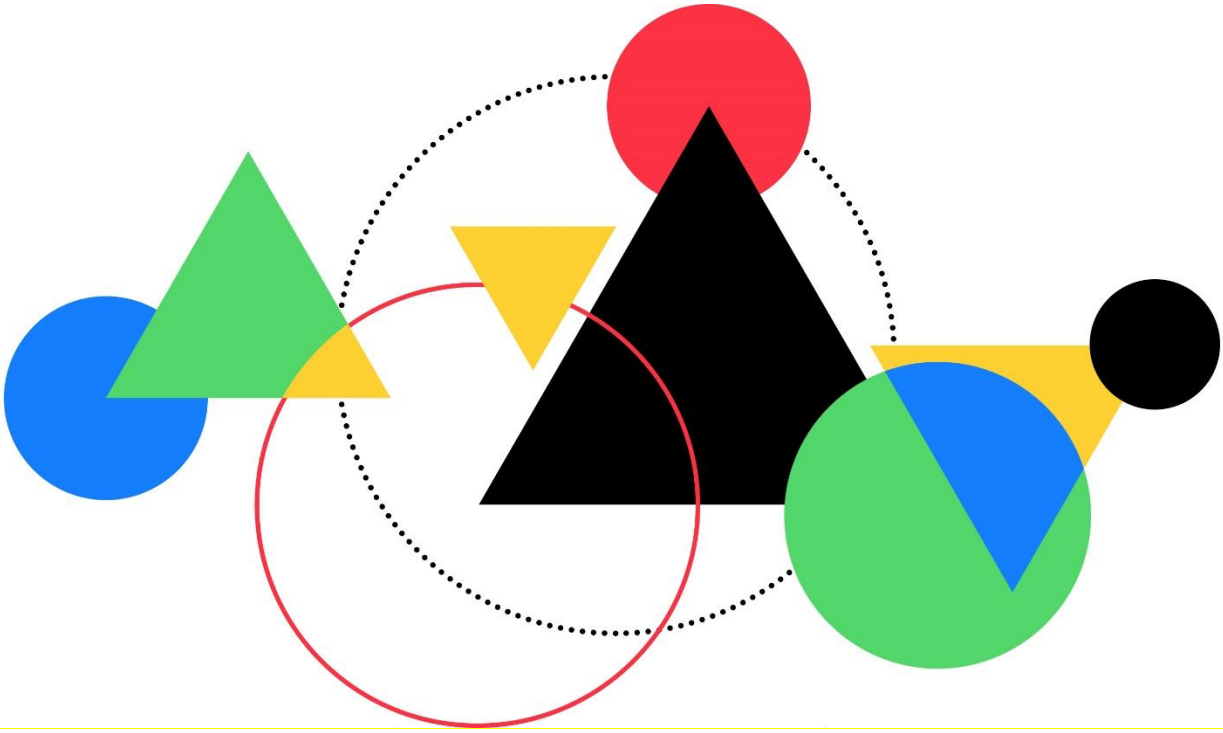
(using the double angle substitution $\sin(2\theta)=2\sin\theta\cos\theta$)

... algebra ensues ... (not very much)

$$\iint_S z\,dS=8\pi$$

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



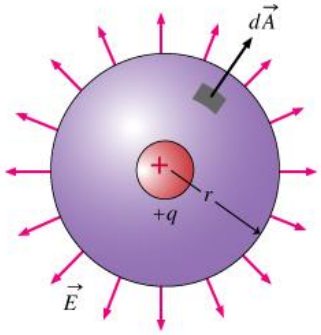
Surface integrals (vector fields)

8.ii

Surface integrals of vector fields

It is often the case in physics that we wish to integrate a vector field over a surface. A notable example of this, which we will return to, is Gauss' law in integral form, which states that:

$$\iint_S \boldsymbol{E}.\hat{\boldsymbol{n}}\, dS = \frac{q}{\epsilon_0}$$



which is telling us that the integral over some surface S of the dot product of the electric field and the normal to the surface, is equal to the charge enclosed divided by the permittivity of free space.

NB: all surfaces have two normal vectors – the convention is that unless otherwise stated, if the surface is closed then the normal vector is pointing outwards. This kind of vector surface integral is said to represent the *flux* through that surface.

There are also basic rules for evaluating vector surface integrals, which are very similar to those we developed for evaluating scalar fields, which is not surprising since we are simply replacing the given scalar function by the dot product of the vector function and the unit normal.

The surface integral of a vector field over a surface described by z=g(x,y) is given by:

$$\iint_S \boldsymbol{F} \cdot \hat{\boldsymbol{n}} dS = \iint_D \boldsymbol{F} \cdot \left(\boldsymbol{k} - \frac{\partial g}{\partial x}\boldsymbol{i} - \frac{\partial g}{\partial y}\boldsymbol{j}\right) dA$$

where S is the surface described in the question and D is its projection on the x-axis (in this case), so that dA = dx dy.

The surface integral of a vector field over a surface described parametrically by **r**(u,v) is given by:

$$\iint_S \boldsymbol{F} \cdot \hat{\boldsymbol{n}} dS = \iint_D \boldsymbol{F} \cdot \left(\frac{\partial \boldsymbol{r}}{\partial u} \times \frac{\partial \boldsymbol{r}}{\partial v}\right) dA$$

where the cross product term represents the area change associated with du, dv. D in this case is the area in (u,v) parameter space which describes the surface and dA = du dv.

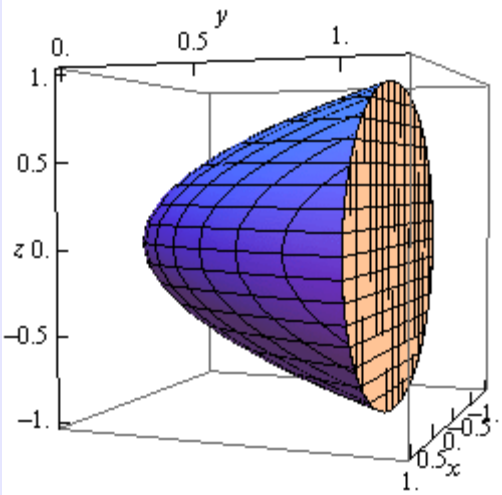
Example 8.3 Evaluate

$\iint_S \boldsymbol{F} \cdot \boldsymbol{dS}$ where $\boldsymbol{F} = y\boldsymbol{j} - z\boldsymbol{k}$

and S is the surface given by the paraboloid $y = x^2 + z^2$, $0 \leq y \leq 1$and the disc $x^2 + z^2 \leq 1$ at $y = 1$. Assume the S has a positive orientation.

Answer:

First, try and sketch the function:



The figure illustrated that the disc is the cap on the bottom of the paraboloid, and so we have a closed surface. This, combined with S having a positive orientation, means that the unit normal vector will point outwards from this closed surface (i.e. normal convention).Let's calculate for the paraboloid first. We have the formula for its surface in the form

$y = g(x, z)$. Since we're going to use the gradient, we want to define a function

$f(x, y, z) = y - g(x, z) = y - x^2 - z^2$, for which we can calculate the gradient:

$\boldsymbol{\nabla} f = -2x\boldsymbol{i} + \boldsymbol{j} - 2z\boldsymbol{k}$. This vector is normal to the paraboloid surface. The unit vector is therefore

$\hat{\boldsymbol{n}} = \frac{\boldsymbol{\nabla} f}{|\boldsymbol{\nabla} f|} = \frac{-2x\boldsymbol{i}+\boldsymbol{j}-2z\boldsymbol{k}}{|\boldsymbol{\nabla} f|}$. We could calculate the magnitude of the normal explicitly (it is 3) but we will find that we do not need to, because it will cancel with the projection term. Note that we should change the sign of $\hat{\boldsymbol{n}}$ to make it point outwards.

So, on to calculate the surface integral for the paraboloid, we can project it on to the (x, z)axis, where it project the disc D.

$$\iint_S F \,.\, ds = \iint_D (y\boldsymbol{j} - z\boldsymbol{k}) \,.\, \frac{(2x\boldsymbol{i} - \boldsymbol{j} + 2z\boldsymbol{k})}{|\boldsymbol{\nabla} f|} |\boldsymbol{\nabla} f| dA$$

where the second $|\boldsymbol{\nabla} f|$is the projection term onto the (x, z) axis, cancelling the normalisation of the gradient for the normal. So we find that the surface integral is just dot product of the vector field with the gradient integrated over projected area.

Evaluating this:

$$(y\boldsymbol{j} - z\boldsymbol{k}).(2x\boldsymbol{i} - \boldsymbol{j} + 2z\boldsymbol{k}) = -y - 2z^2 = -x^2 - z^2 - 2z^2 = -x^2 - 3z^2$$

Now D, the projected region, is the disc of radius 1 centred on the origin, in the (x, z)plane, so let's switch to plane polars to evaluate the integral:

$x = r\cos\theta, z = r\sin\theta, 0 \le r \le 1, 0 \le \theta \le 2\pi$

So we have the following double integral to evaluate:

$$-\int_0^{2\pi}\int_0^1 (r^2\cos^2\theta + 3r^2\sin^2\theta)\, r\, dr d\theta$$

where the r before the dr is the Jacobian for the coordinate conversion.

$$-\int_0^{2\pi}\int_0^1 (\cos^2\theta + 3\sin^2\theta)\, r^3 dr = -\int_0^{2\pi} (\cos^2\theta + 3\sin^2\theta)\, d\theta \int_0^1 r^3\, dr$$

We make the following standard substitutions:

$\cos^2\theta = \frac{1}{2}(\cos 2\theta - 1)$ and $\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta)$

which gives:

$$-\int_0^{2\pi}\left[\frac{1}{2}(\cos 2\theta + 1) + \frac{3}{2}(1 - \cos 2\theta)\right]\frac{1}{4} d\theta = -\pi$$

(where we note that we didn't have to precisely integrate the cos terms since they would become terms involving sin for which the upper and lower limits correspond to zero). So, we have the flux of the vector field through the paraboloidal part of the surface is $-\pi$; now we have to evaluate the flux through the flat cap and add them.

It is clear from the geometry of the problem that the unit normal to the cap is just $\hat{\boldsymbol{n}} = \boldsymbol{j}$ (and in the right orientation, since this is the unit vector which is pointing out of the surface).

$$\iint_S \boldsymbol{F}\cdot\boldsymbol{dS} = \iint_D (y\boldsymbol{j} - z\boldsymbol{k})\cdot(\boldsymbol{j}) = \iint_D y$$

Trivially this is π since it is the area of a circle of radius 1.

Hence the total surface integral is $-\pi + \pi = 0$

Example 8.4 Evaluate

$\iint_S F \,.\, dS$ where $F = xi + yj + z^4k$

where S is the upper half of the sphere $x^2 + y^2 + z^2 = 9$ and the vector field is pointing outwards from the sphere.

Answer:

Use parametric expression for sphere of radius 3:

$$\boldsymbol{r}(\theta, \phi) = 3\sin\theta\cos\phi\boldsymbol{i} + 3\sin\theta\sin\phi\boldsymbol{j} + 3\cos\theta\boldsymbol{k}$$

with the limits $0 \leq \theta \leq \pi/2$ and $0 \leq \phi \leq 2\pi$

We need to determine the normal from the cross product, so first we need

$$\frac{\partial \boldsymbol{r}}{\partial \phi} = -3\sin\theta\sin\phi\boldsymbol{i} + 3\sin\theta\cos\phi\boldsymbol{j}$$

$$\frac{\partial \boldsymbol{r}}{\partial \theta} = 3\cos\theta\cos\phi\boldsymbol{i} + 3\cos\theta\sin\phi\boldsymbol{j} - 3\sin\theta\boldsymbol{k}$$

The cross product (via some … algebra …)

$$\frac{\partial \boldsymbol{r}}{\partial \phi} \times \frac{\partial \boldsymbol{r}}{\partial \theta} = -9\sin^2\theta\cos\phi\boldsymbol{i} - 9\sin^2\theta\sin\phi\boldsymbol{j} - 9\sin\theta\cos\theta\boldsymbol{k}$$

We need to take the negative of this normal vector since this is the *inwards* pointing normal.

$$\boldsymbol{n} = +9\sin^2\theta\cos\phi\boldsymbol{i} + 9\sin^2\theta\sin\phi\boldsymbol{j} + 9\sin\theta\cos\theta\boldsymbol{k}$$

Formally we need the unit vector, but we know that the normalisation is going to cancel with the area term.

We also need **F** in spherical polars, so

$$\boldsymbol{F} = 3\sin\theta\cos\phi\boldsymbol{i} + 3\sin\theta\sin\phi\boldsymbol{j} + (3\cos\theta)^4\boldsymbol{k}$$

So, we can now evaluate the double integral:

$$\iint_{\boldsymbol{S}} \boldsymbol{F} \,.\, \boldsymbol{dS} = \iint_{\boldsymbol{D}} \boldsymbol{F} \,.\, \boldsymbol{n}\, dA$$

which equals:

$$\iint_{\boldsymbol{D}} \boldsymbol{F} \,.\, \boldsymbol{n}\, dA = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi/2} 27\sin^3\theta + 729\sin\theta\cos^5\theta d\theta d\phi$$

… algebra ensues …

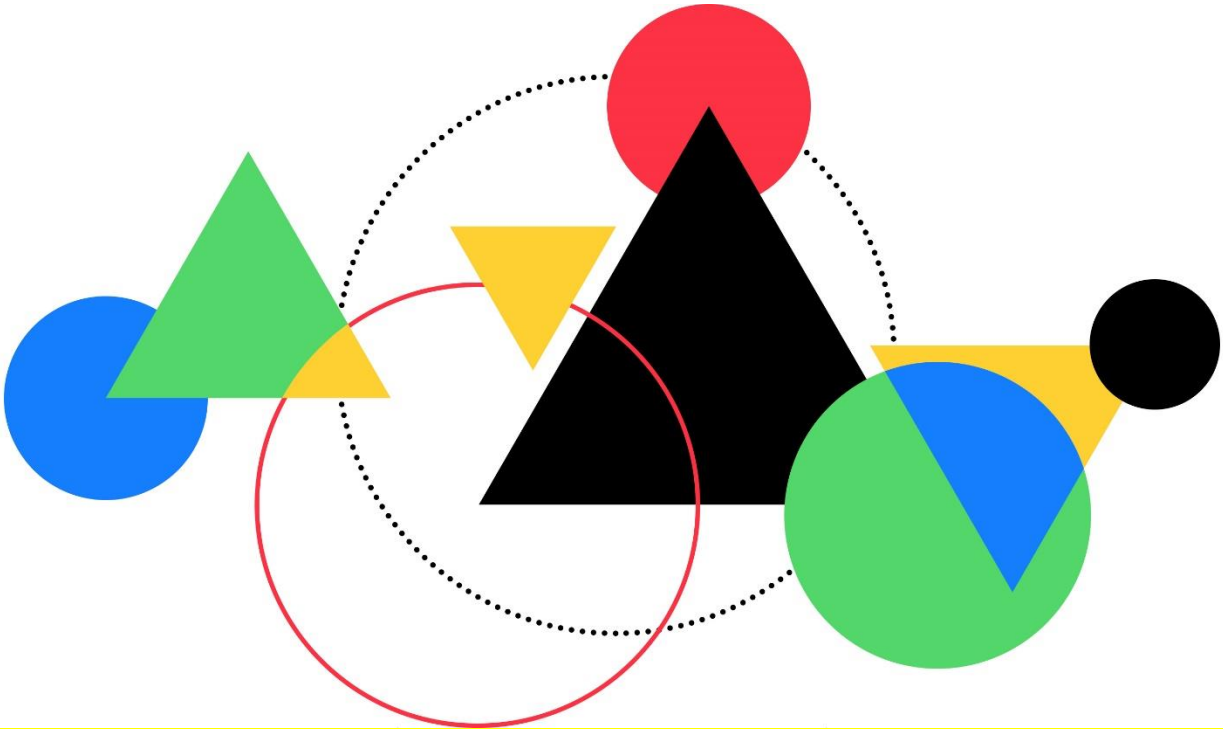
$$\iint_{\boldsymbol{D}} \boldsymbol{F} \,.\, \boldsymbol{n}\, dA = 729\pi$$

What contribution would a flat bottom to the sphere add if the hemisphere was sitting on the plane z=0?

As noted before, the (x,y) components of the surface vector would cancel, and the vector field in the **k** direction at z=0 is zero, so it would add nothing.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Divergence

9.i

The Divergence

We now return to the 'Del' vector operator:

$$\boldsymbol{\nabla} = \left(\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}\right)$$

and instead of applying it to a scalar field to get a vector, we now apply it to a vector field, to get a scalar quantity, the *Divergence*. This can be considered to be the dot product of the vector field with the Del vector operator. In other words:

$$\text{div}\,\boldsymbol{F} = \boldsymbol{\nabla}\cdot\boldsymbol{F} = \left(\frac{\partial F_X}{\partial x} + \frac{\partial F_Y}{\partial y} + \frac{\partial F_Z}{\partial z}\right)$$

which is a *scalar.* What is the meaning of the Divergence? The Divergence is a quantitative measure of how much a field diverges (or converges) at any given point. If we imagine our vector field to be describing the flow of a fluid in some volume centred on a given point, then the Divergence is the net outflow of fluid from the volume, in the limit that the volume goes to zero.

A vector field for which $\nabla.F = 0$ is said to be *solenoidal* (the etymology of this is that solen means pipe → constrained or within a pipe → fixed volume). Sometimes such fields are also said to be *incompressible*.

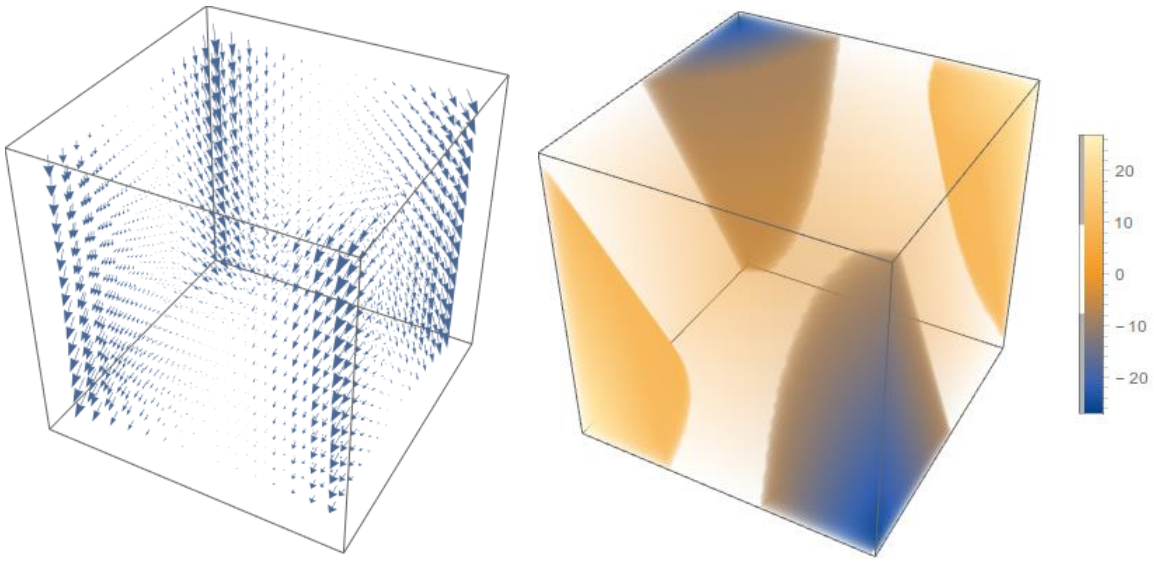
Example 9.1 Evaluate $\boldsymbol{\nabla}\cdot\boldsymbol{F}$ where $\boldsymbol{F} = x^2y\boldsymbol{i} + xyz\boldsymbol{j} - x^2y^2\boldsymbol{k}$

Answer: This is a straightforward calculation:

$$\text{div}\,\boldsymbol{F} = \boldsymbol{\nabla}\cdot\boldsymbol{F} = \left(\frac{\partial(x^2y)}{\partial x} + \frac{\partial(xyz)}{\partial y} + \frac{\partial(-x^2+y^2)}{\partial z}\right) = 2xy + xz$$

We can visualize this vector field and its Divergence in the graphics below.

Note that negative Divergence means that the field is compressing at the point.



Example 9.2 Evaluate $\boldsymbol{\nabla} \cdot \boldsymbol{F}$ where $\boldsymbol{F} = \boldsymbol{r}$ i.e. the position vector field

Answer: If **F** is the position vector, then $\boldsymbol{F} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$

Therefore trivially, the divergence of the position vector = 3

As we have already found, many vector fields (namely conservative vector fields) are themselves already the gradient of a scalar field. So if we start with an initial scalar field, U, then the divergence is

$$\text{div } \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} U) = \nabla^2 U$$

where the ∇^2 operator is known as the Laplacian. We shall return to this later.

So what does the Divergence mean? If we imagine some volume placed in a vector field, then the divergence at a given point is the net flow outwards from that surface, per unit volume. Formally:

$$div\ \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \lim_{V \to 0} \iint_{S(V)} \frac{\boldsymbol{F} \cdot \hat{\boldsymbol{n}}}{|V|} dS$$

where V is the volume, S(V) is the boundary of that volume, and the integral is surface integral of the vector function (which we recognise from earlier in the course).

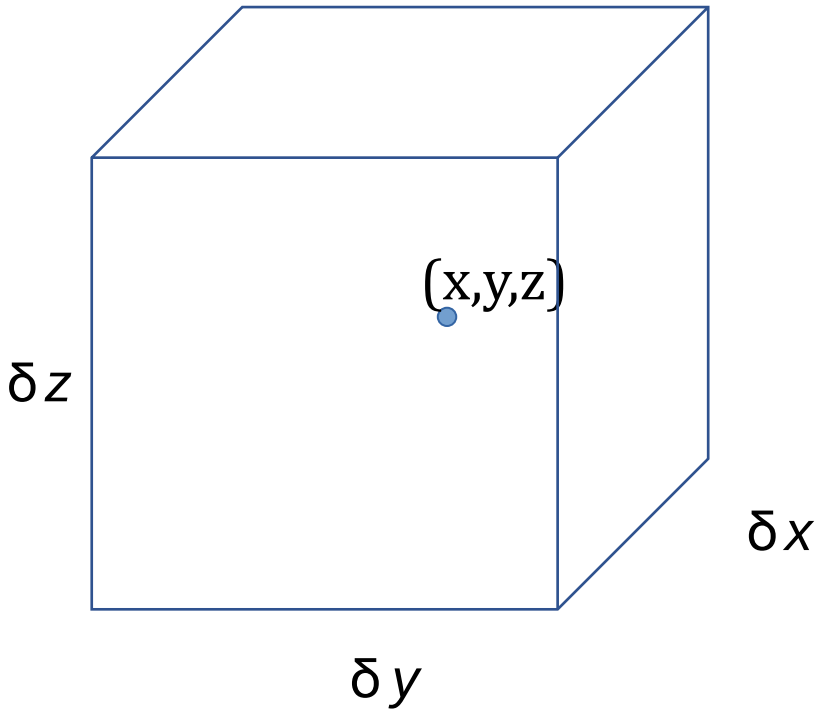
It may help in visualising the Divergence to think of a fluid flow.

Derivation

How can we get from this conceptual definition of the divergence to the formal equations given above, in other words how do we get from.

$div\ \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \lim_{V \to 0} \iint_{S(V)} \frac{\boldsymbol{F} \cdot \hat{\boldsymbol{n}}}{|V|} dS$ to $\text{div}\ \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \left(\frac{\partial F_X}{\partial x} + \frac{\partial F_Y}{\partial y} + \frac{\partial F_Z}{\partial z} \right)$?

We can illustrate how we do this by considering a simple cubic volume, which has its faces oriented along the **i, j, k** directions, has its centre at (x,y,z) and side lengths δx,δy,δz.



Consider first the face which is pointing in the positive **i** (i.e. positive x) direction.

The unit vector normal to the face is simply **i**, so that

$$\boldsymbol{F} \cdot \boldsymbol{i} = F_x$$

where, as usual, F_x is the x-component of the vector field. We can consider this dot product to be acting out of the centre of the face, which has coordinates $(x + \delta x/2, y, z)$

and finally multiplying by the surface area of the face, $\delta y \delta z$, we find that the surface integral from this face is

$$\iint_{+x face} \boldsymbol{F} \cdot \hat{\boldsymbol{n}}\, dS = F_X(x + \delta x/2, y, z)\delta y \delta z$$

We can do the same thing for the opposite face, for which the unit normal is **-i** and find

$$\iint_{-x face} \boldsymbol{F} \cdot \hat{\boldsymbol{n}}\, dS = -F_X(x - \delta x/2, y, z)\delta y \delta z$$

So the total surface integral from these two faces is:

$$\iint_{bothxfaces} \boldsymbol{F}\cdot\hat{\boldsymbol{n}}\,dS = \frac{\left(F_X(x+\delta x/2,y,z)-F_X(x-\delta x/2,y,z)\right)}{\delta x}\delta x\delta y\delta z$$

Recognising that $\delta x\delta y\delta z$ is the volume of the cuboid, we see that in the limit as the side lengths and hence the volume goes to zero, this becomes ($\frac{\partial F_X}{\partial x}\,.volume$).

Of course this works identically for both of the other sides and we arrive at our original equation:

$$\text{div}\,\boldsymbol{F} = \boldsymbol{\nabla}\cdot\boldsymbol{F} = \left(\frac{\partial F_X}{\partial x}+\frac{\partial F_Y}{\partial y}+\frac{\partial F_Z}{\partial z}\right)$$

Divergence in other coordinate systems

The derivation of the divergence in other coordinate systems is not examinable (and indeed these are provided on the information sheet).

In cylindrical coordinates, the divergence is given by:

$$\text{div}\,\boldsymbol{F} = \boldsymbol{\nabla}\cdot\boldsymbol{F} = \left(\frac{1}{r}\frac{\partial(rF_r)}{\partial r}+\frac{1}{r}\frac{\partial F_\phi}{\partial\phi}+\frac{\partial F_Z}{\partial z}\right)$$

In spherical polar coordinates, the divergence is given by:

$$\text{div}\,\boldsymbol{F} = \boldsymbol{\nabla}\cdot\boldsymbol{F} = \left(\frac{1}{r^2}\frac{\partial(r^2F_r)}{\partial r}+\frac{1}{r\sin\theta}\frac{\partial(F_\theta\sin\theta)}{\partial\phi}+\frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi}\right)$$

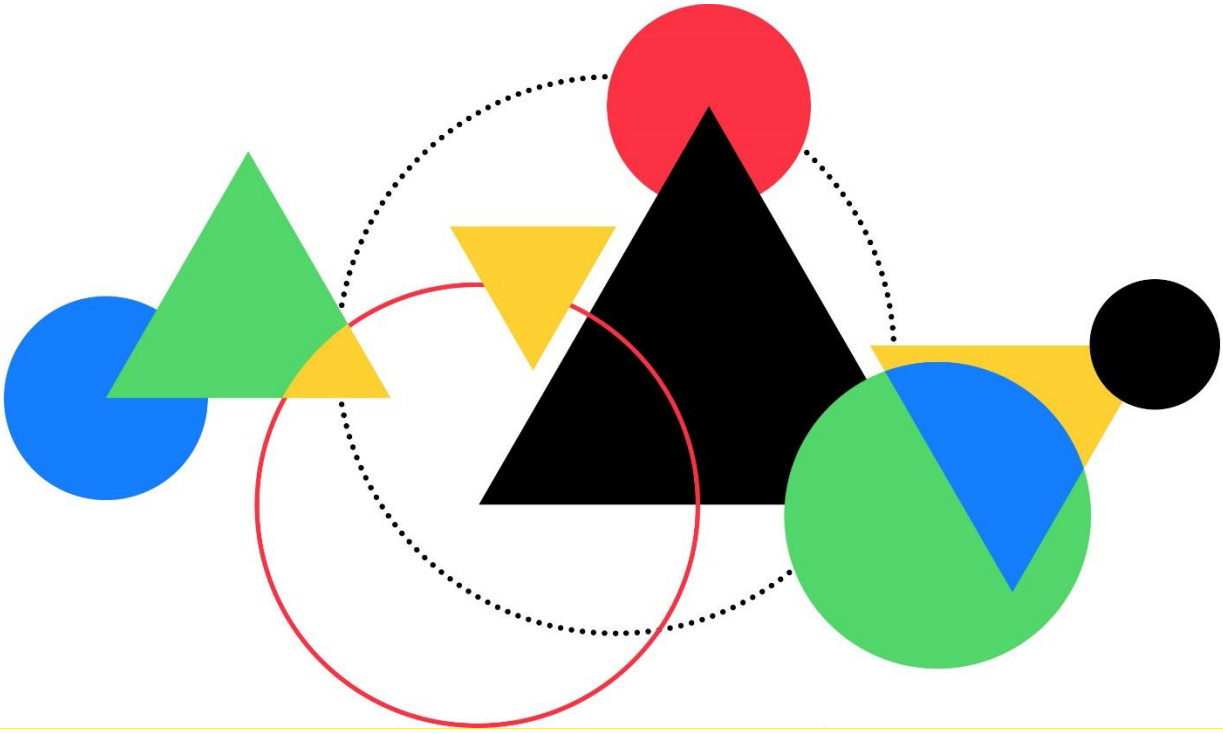
The general form in curvilinear coordinates is given by:

$$\boldsymbol{\nabla}\cdot\boldsymbol{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial(h_2h_3F_1)}{\partial u_1}+\frac{\partial(h_1h_3F_2)}{\partial u_2}+\frac{\partial(h_1h_2F_3)}{\partial u_3}\right]$$

where u are the coordinates and h are the corresponding scale factors. Check yourself that the forms for spherical and cylindrical given above can be arrived at via this form.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Divergence of the electric field, and other things

9.ii

The Divergence of the electric field

Example 9.3 Evaluate the divergence of the electric field

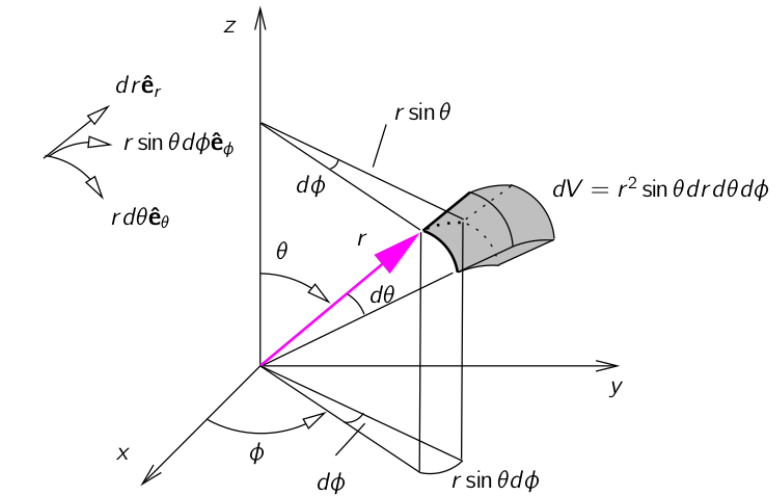
Answer:

The electric field in spherical polar coordinates is given by:

$\boldsymbol{E} = \frac{c\hat{\boldsymbol{r}}}{r^2}$ where c is a bunch of constants for some given configuration.

$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{1}{r^2}\frac{\partial(r^2 F_r)}{\partial r}$ which is … zero.

Wouldn't you have thought the electric field would a diverging field? After all it has a "source" (we will return to this in discussing the Divergence theorem).



Well, appearances can be deceiving – always do the calculation to check! In fact we can see quite simply via geometric reasoning that the divergence of vector field which falls off as $\frac{1}{r^2}$ will be zero. Consider a unit volume in spherical polar coordinates: The surface which is normal to the radial direction has area

$$r^2\sin\theta d\theta d\phi$$

and so we can see that between the inner and outer faces in the radial direction, along some incremental distance in r, the strength of the field will have fallen as r^{-2} just as the area of the face has grown by r^2, hence no net flow in or out of the volume.

As noted above a field like this, for which the divergence is zero, is called *solenoidal.*

This is an early lesson in being cautious about taking things at face value. There are places on the internet where it will tell you that "outflowing" fields, represented by a vector field pointing away from the origin will intuitively have positive divergence, and that "inflowing" vector fields will intuitively have negative divergence.

Example 9.4 What are the conditions for an "outwards-flowing" vector field to actually have negative divergence, and conversely, for an "inflowing field" to actually have positive divergence.

Answer:

An "outward flowing" vector field (like the electric field) will have positive divergence if more flux is flowing out of the outer side than is entering the inner side. This will occur if the flux falls off more slowly than r^{-2}, so for some general function $F_r \propto r^n$ the divergence is positive for $n > -2$. Similarly, for an outflowing field to have negative divergence, $n < -2$. The opposite is true for an inflowing field.

As we know, a conservative vector field is itself the gradient of some scalar potential. Therefore the divergence of such a field is the 2nd derivative of the original scalar function:

If U is scalar field, then the vector field $\boldsymbol{F} = \boldsymbol{\nabla} U$ has the divergence $\boldsymbol{\nabla} \cdot \boldsymbol{F} = \nabla^2 U$.

This term ∇^2 is known as the *Laplacian* and is sometimes also represented by $\boldsymbol{\Delta}$. In rectangular cartesian coordinates, the Laplacian is:

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2}$$

It is straightforward to apply the Laplacian directly to the electric potential to show that the Divergence of the electric field is zero.

Exact differentials

We return briefly to the question of conservative fields. The relation $\boldsymbol{F} = \boldsymbol{\nabla} U$ that holds for conservative fields, implies that those fields are *exact differentials.* In the general case of n dimensions, it must be satisfied that all pairs of the second derivatives must be equal such that if

$$df = \sum_{i=1}^{n} g_i\,(x_1, x_2, \ldots, x_i)dx_i$$

where$(x_1, x_2, \ldots, x_i)$are the variables, then:

$\frac{\partial g_i}{\partial x_j} = \frac{\partial g_j}{\partial x_i}$for all (n(n-1)/2) pairs i,j.

Example 9.5 Establish whether the following differentials are exact or inexact:

(i) $xdy + 3ydx$

(ii)$(y + z)dx + xdy + xdz$

Answer:

For (i), integrating the two components we get

$f = xy + g(y)$and $f = 3xy + h(x)$which cannot be consistent with each other.

For (ii)

$$\frac{\partial(y+z)}{\partial y} = \frac{\partial(x)}{\partial x} = 1$$

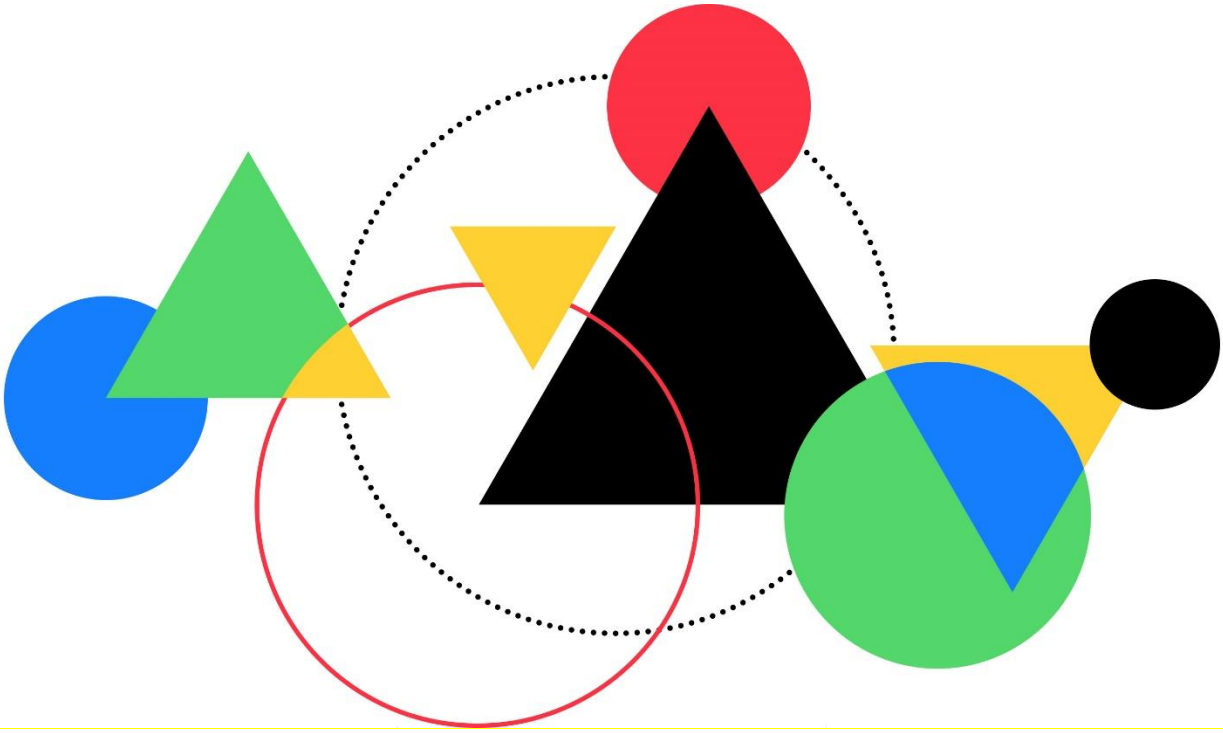
$$\frac{\partial(y+z)}{\partial z} = \frac{\partial(x)}{\partial x} = 1$$

$$\frac{\partial(x)}{\partial z} = \frac{\partial(x)}{\partial y} = 0$$

– all three criteria are satisfied hence this is an exact differential.

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Curl

10.i

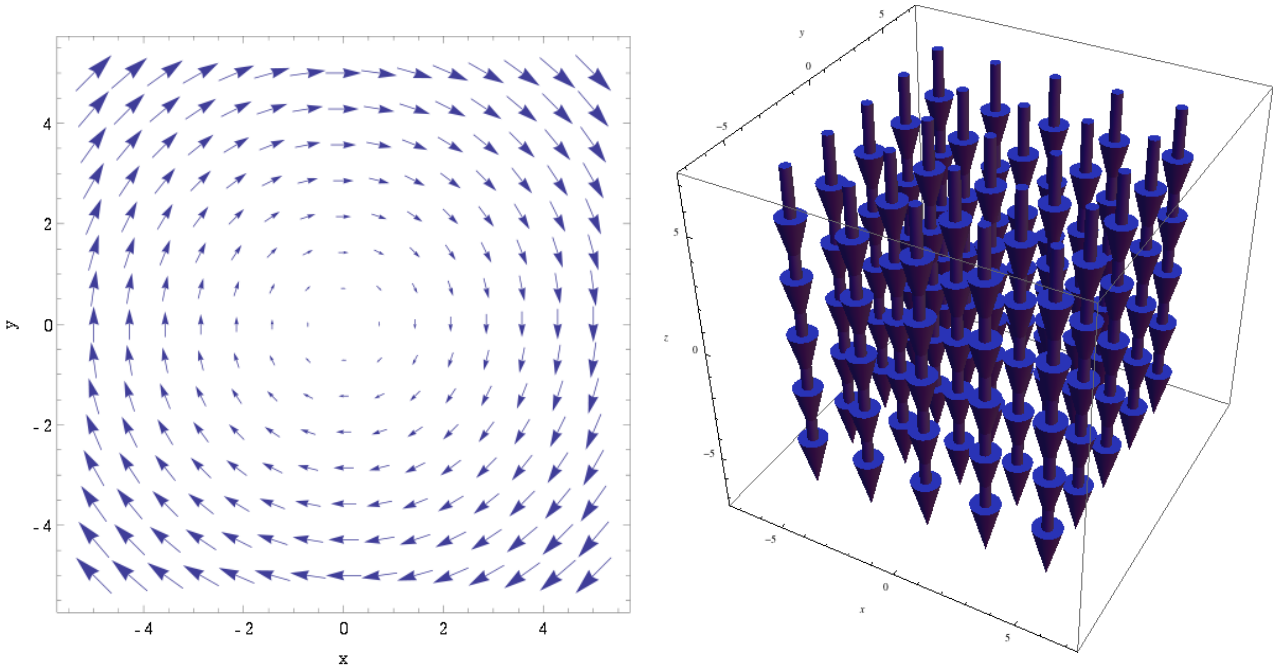
Curl

It is time for the final part of the set of *Div, Grad, Curl all that*, the Curl. The curl measures the microscopic (infinitesimal) rotation of a three-dimensional vector field. It is defined as

$$\text{curl}\,\boldsymbol{F} = \boldsymbol{\nabla} \times \boldsymbol{F} = det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

It is very important to understand that the curl is a measure of the microscopic rather than the macroscopic circulation of the field. Specifically, the curl represents the tendency of "particles" at a point to rotate around the axis that points in the direction of the curl. This is easy to visualise for a 2D field in the (x,y) plane – the curl represents the tendency of the "particles" to rotate about an axis which points in the z-direction.

The vector field $\boldsymbol{F} = y\boldsymbol{i} - x\boldsymbol{j}$ looks like this (left), with its associated curl (right).

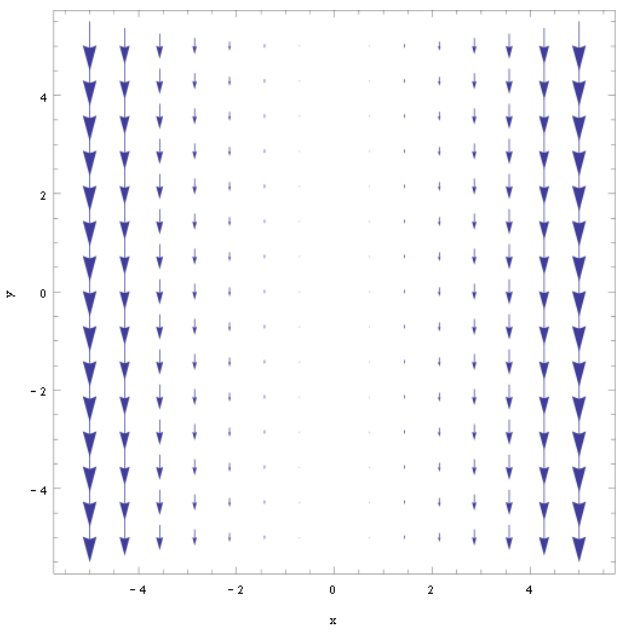


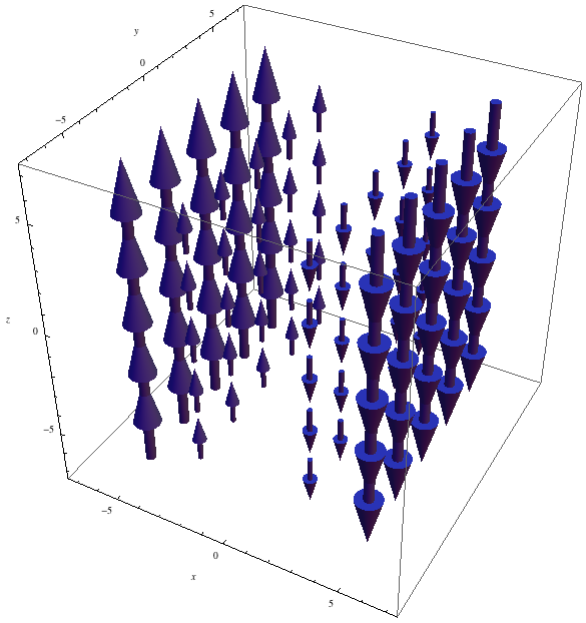
We can easily imagine that a paddle wheel placed anywhere in such a fluid flow would rotate anticlockwise. Directly evaluating the curl gives us:

$$\boldsymbol{\nabla} \times \boldsymbol{F} = -2\boldsymbol{k}$$

i.e. pointing into the page, consistent with the right-hand figure above. The curl has the same magnitude everywhere.

The vector field $\boldsymbol{F} = -x^2\boldsymbol{j}$ looks quite different:





So we can see that in this case a paddle wheel to the right of the y-axis will rotatie clockwise, and one placed to the left of the y-axis will rotate anticlockwise.

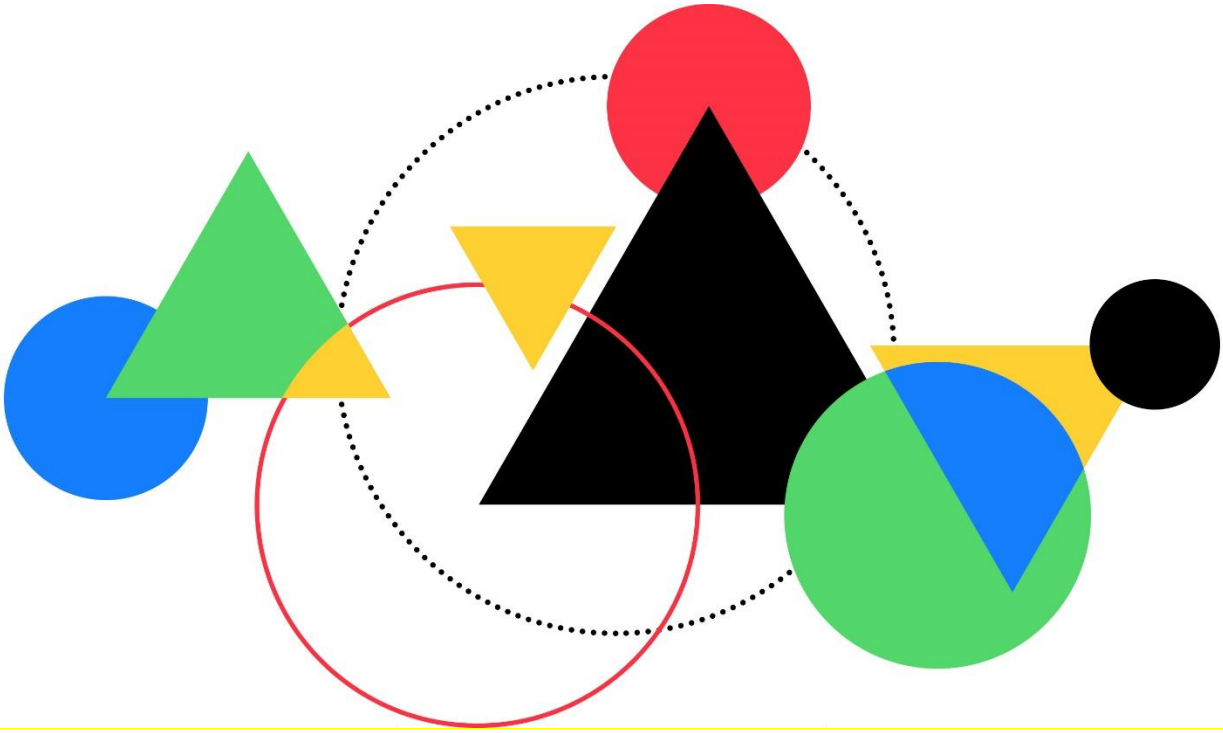
The expression for the curl comes from its formal definition as the infinitesimal area density of the *circulation* of that field.

$$(\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \hat{\boldsymbol{n}} = \lim_{\mathrm{A}\to 0} \int_{C(A)} \frac{\boldsymbol{F}.\boldsymbol{dl}}{|A|} ds$$

The *circulation* of the field along a given axis can be evaluated by taking a closed line integral in the plane orthogonal to that direction, in the limit that the path length C goes to zero. The derivation is partially undertaken in the lecture, and has similarities to that for the Divergence.

The full expression for the curl is therefore the sum of the vectors representing the degree of rotation of the field around the x, y, and z axes. Summed, they indicate the net rotation axis of the field at a given point.

Taking the dot product of the curl with some direction e.g. $\hat{\boldsymbol{n}}$ gives the rotation of the field around that axis.

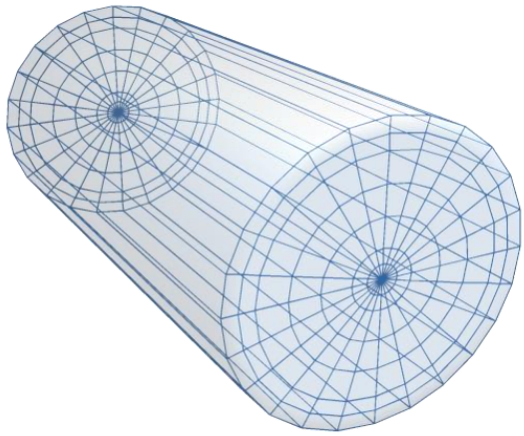


The Curl: examples

10.ii

The relation of curl to the rotation of a rigid body

A solid body is rotating around some axis, which we will align with the z-axis.



The angular velocity vector is then given by

$$\boldsymbol{\omega} = \omega \boldsymbol{k}$$

If we assume that some point in the solid body has a position vector from the axis, i.e. in the x,y-plane, of

$$\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$$

then the velocity of the point is given by

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$$

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r} = det \begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 0 & 0 & \omega \\ x & y & 0 \end{pmatrix} = -y\omega\boldsymbol{i} + x\omega\boldsymbol{j}$$

Now if we take the curl of this velocity field we find:

$$\text{curl } \boldsymbol{v} = \boldsymbol{\nabla} \times \boldsymbol{v} = det \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y\omega & x\omega & 0 \end{bmatrix} = 2\omega\boldsymbol{k}$$

So the curl of the velocity field is a vector which is twice the angular velocity vector of the solid body around its axis of rotation.

The microscopic and macroscopic circulation of the field can be quite different.

Example 10.1 Consider the macroscopic and microscopic circulation of the field given by

$$\boldsymbol{F} = \frac{-y\boldsymbol{i} + x\boldsymbol{j}}{x^2 + y^2}$$

Answer:

Sketching the field shows that a "particle" placed in such a field would circulate anticlockwise around the z-axis (much more strongly closer to the z-axis).

Now let's calculate the curl:

$\boldsymbol{\nabla} \times \boldsymbol{F} = 0$ ... which indicates that the microscopic circulation is zero.

The field described in the above example could be visualised by particle which moves around the field but does not itself rotate.

Example 10.3 Can you consider a field in which the microscopic and macroscopic rotations are in opposite directions?

Answer: Yes. An example is $F = \frac{-y\boldsymbol{i}+x\boldsymbol{j}}{(x^2+y^2)^{3/2}}$

The sense of the macroscopic rotation is the same as the previous example – anticlockwise. However, the field falls off quickly enough that the clockwise push on the inner side is stronger than the anticlockwise push on the outer side, and hence a particle would rotate clockwise while moving anticlockwise.

The curl and conservative fields

1. If some scalar field $U(x, y, z)$ has continuous second order partial derivatives then $\text{curl}(\nabla U) = 0$

2. If **F** is a conservative field then $\text{curl}(\boldsymbol{F}) = 0$ – this is a direct consequence of what it means to be a conservative field, and the above rule

3. As a result, if $\text{curl}(\boldsymbol{F}) = 0$ then **F** is a conservative field (usually, not strictly true).

Example 10.1 Prove that the curl of a conservative field is zero

Answer:

The **i** component of the curl is given by

$$\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}$$

Now we know that $F_z = \frac{\partial U}{\partial z}$ and so $\frac{\partial F_z}{\partial y} = \frac{\partial^2 U}{\partial y \partial z}$, and similarly for $\frac{\partial F_y}{\partial z}$, and so we see that, because of the commutability of the 2nd derivatives, the value of the term is zero. Equally for the other two terms, hence the curl of the conservative field is zero.

Example 10.2 Determine by evaluating the curl if the following vector field is conservative:

$$F = x^2y\boldsymbol{i} + xyz\boldsymbol{j} - x^2y^2\boldsymbol{k}$$

Answer:

Calculate the curl:

$$\text{curl } \boldsymbol{F} = \boldsymbol{i}(-2x^2y - xy) - \boldsymbol{j}(-2xy^2) + \boldsymbol{k}(yz - x^2) \neq 0$$

→ therefore **F** is not a conservative field.

The curl in general curvilinear coordinates can be calculated if the scale factors h are known:

$$\boldsymbol{\nabla} \times \boldsymbol{F} = \frac{1}{h_1h_2h_3} det \begin{bmatrix} h_1\hat{u}_1 & h_2\hat{u}_2 & h_3\hat{u}_3 \\ \dfrac{\partial}{\partial u_1} & \dfrac{\partial}{\partial u_2} & \dfrac{\partial}{\partial u_3} \\ h_1F_1 & h_2F_2 & h_3F_3 \end{bmatrix}$$

Div, Grad, Curl: a summary

Gradient: takes a scalar field and returns a vector whose direction is the rate of greatest change of the scalar field, and whose magnitude is the rate of change in that direction

$$\boldsymbol{grad}\ \phi\ = \boldsymbol{\nabla}\phi = \left(\boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}\right)$$

Divergence: takes a vector field and returns a scalar, which measures the tendency of a field to spread out (diverge). Its value measures the rate of 'flow' of the field out of an infinitesimal volume in the field.

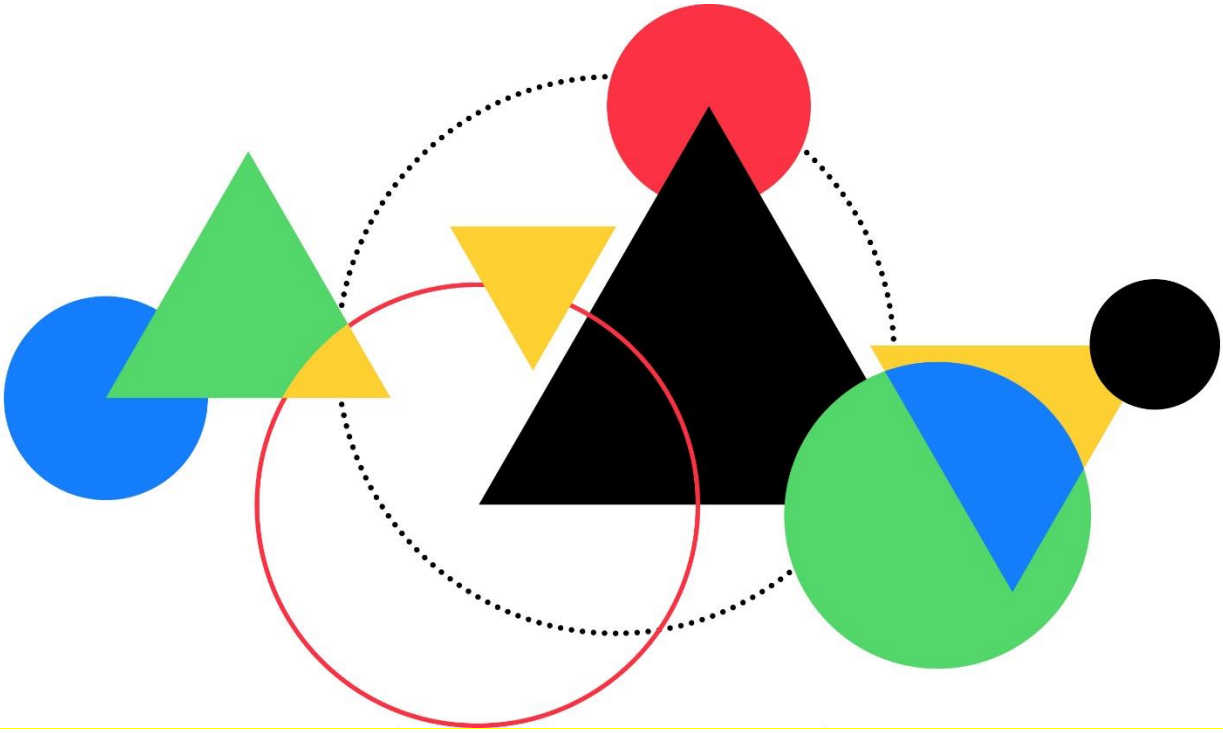
$$\text{div } \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \left(\frac{\partial F_X}{\partial x} + \frac{\partial F_Y}{\partial y} + \frac{\partial F_Z}{\partial z}\right)$$

Curl: takes a vector and returns a vector, which measures the 'microscopic' rotation of a vector field. It is pointed in the direction of the axis around which the field is rotating, and its value measures the rate of that rotation.

$$\text{curl } \boldsymbol{F} = \boldsymbol{\nabla} \times \boldsymbol{F} = det \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



The Divergence Theorem

11.i

The Divergence Theorem

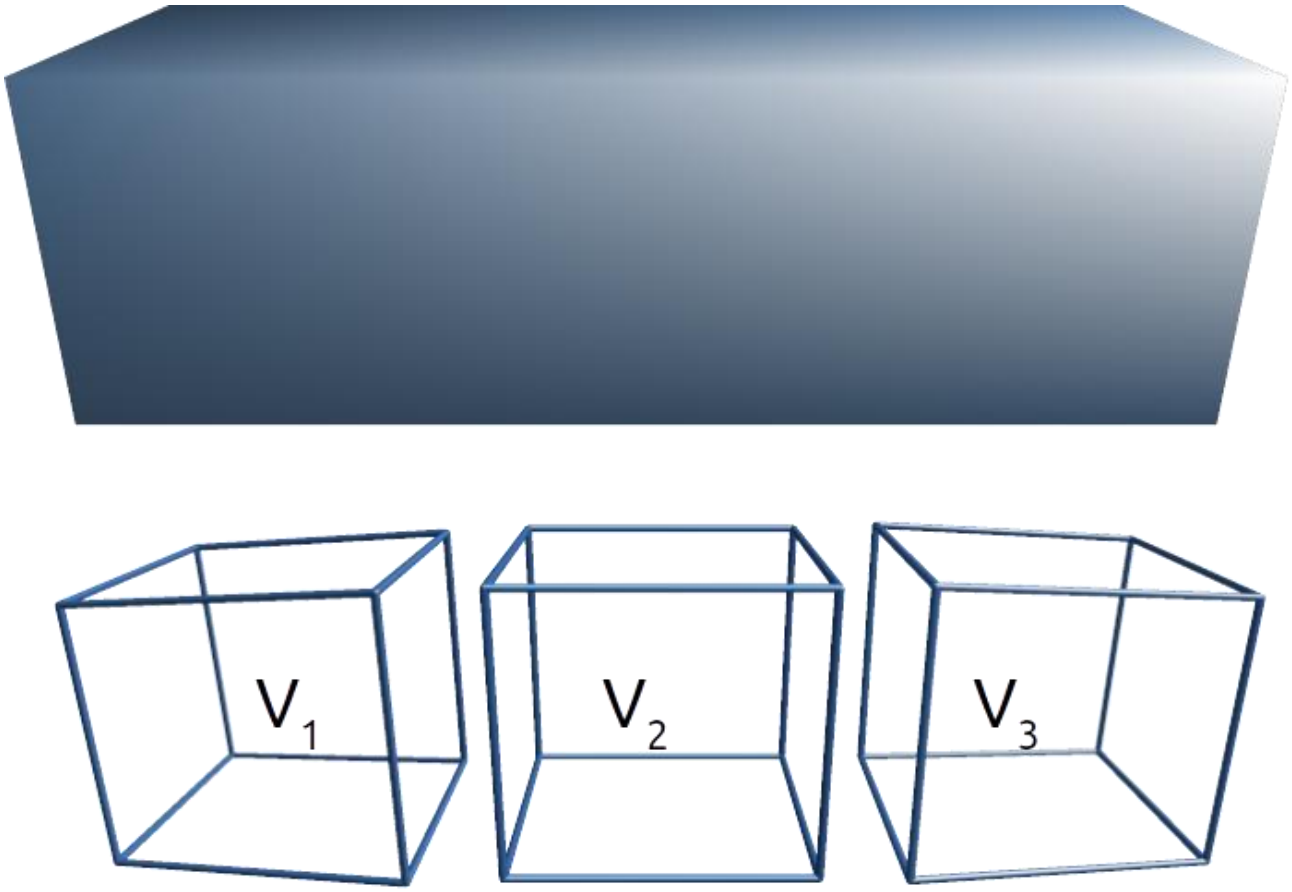
The Divergence theorem relates surface and volume integrals, and feels intuitive. It says that the flux escaping through the boundary of some closed surface is equal to the volume integral of the divergence within that volume.

$$\iint_S \boldsymbol{F}.\hat{\boldsymbol{n}}\, dS = \iiint_V \nabla \cdot \boldsymbol{F} dV$$

You may already see that this make sense. The formal proof is straightforward based on our derivation of the Divergence where we established that:

$$div\, \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \lim_{V \to 0} \iint_{S(V)} \frac{\boldsymbol{F} \cdot \hat{\boldsymbol{n}}}{|V|} dS$$

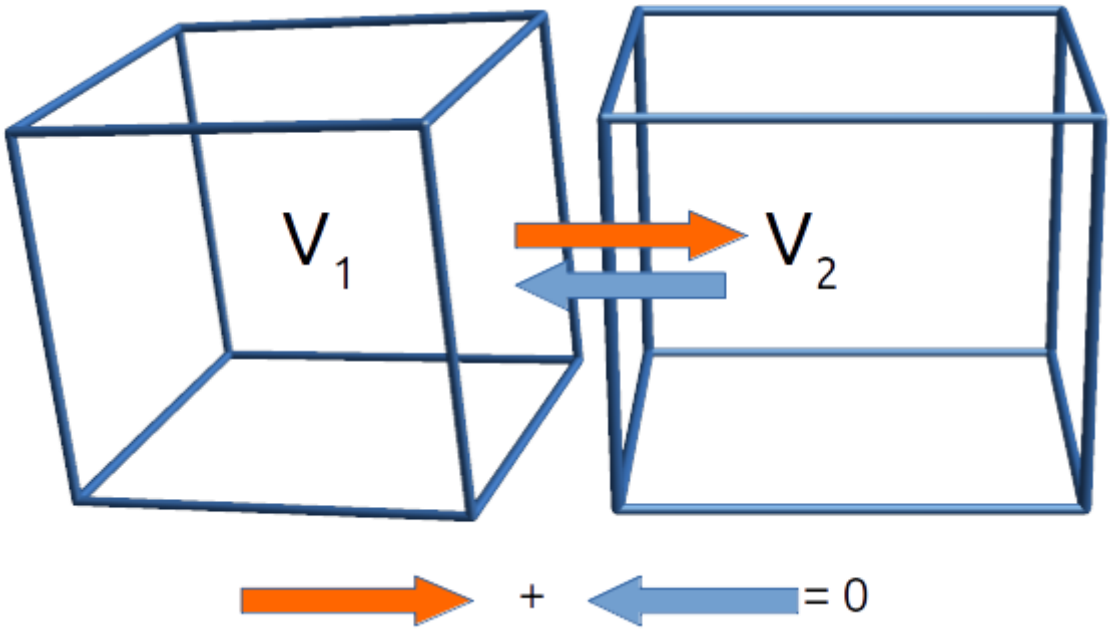
We can imagine that we have some closed surface enclosing a volume which is broken down into small volumes V_i as in the figure below:



We have for each small volume that

$$\boldsymbol{\nabla} \cdot \boldsymbol{F}\, V_i \approx \iint_{S(V_i)} \boldsymbol{F} \cdot \hat{\boldsymbol{n}}\, dS$$

Now if we sum over all the V_i which constitute the total volume, we find that the contributions from touching surfaces cancel out precisely, as they are equal and opposite:



Therefore only the surface terms on the outside of the shape contribute towards the total surface integral and we find that in the limit that each V_i goes to zero we get the Divergence theorem:

$$\iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{F}\, dV = \iint_S \boldsymbol{F} \cdot \hat{n}\, dS$$

More intuition (based on a fluid again): If you imagine some volume, say a box, containing some source of fluid, then it make sense than in the steady state the amount of fluid leaving through a surface surrounding that source, and that source only, would equal the rate at which the source was producing fluid (e.g. if it had a tap in it..).

The divergence theorem is often used to switch problems between surface and volume integrals.

Example 11.1 Use the divergence theorem to evaluate

$\iint_S \boldsymbol{F} \,.\, dS$ where $\boldsymbol{F} = xy\boldsymbol{i} - \frac{1}{2}y^2\boldsymbol{j} + z\boldsymbol{k}$ and the closed surface consists of the three surfaces:

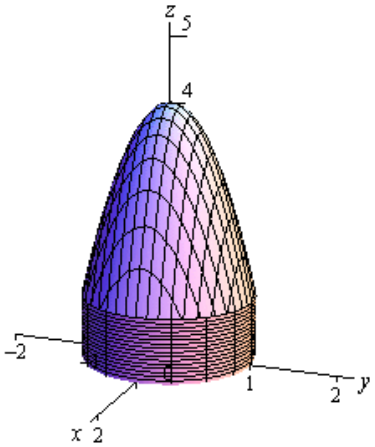
(i) $z = 4 - 3x^2 - 3y^2, 1 \leq z \leq 4$ on the top

(ii) $x^2 + y^2 = 1, 0 \le z \le 1$ on the sides, and

(iii) z=0 on the bottom

Answer:

The surface looks like the figure to the right.



Cylindrical coordinates look like they will be the best for this problem. We can set one range of limits which will describe the entire shape in these coordinates:

$0 \le z \le 4 - 3r^2$ (using the substitution $r^2 = x^2 + y^2$)

$0 \le r \le 1$ and $0 \le \phi \le 2\pi$

Let's evaluate the divergence of the field:: $\nabla . F = y - y + 1$

So the integral is:

$$\iint_S \boldsymbol{F} \cdot \boldsymbol{dS} = \iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{F} = \int_0^{2\pi} d\,\phi \int_0^1 \int_0^{4-3r^2} r\, dzdrd\phi$$

where note that we have added in the r which is the Jacobian term for the switch to cylindrical coordinates.

$$\int_{\phi=0}^{2\pi} d\,\phi \int_{r=0}^{1} \int_{z=0}^{4-3r^2} r\, dzdrd\phi = 2\pi \int_0^1 (4r - 3r^3)\, dr = 2\pi \left[2r^2 - \frac{3}{4}r^4\right]_0^1 = \frac{5}{2}\pi$$

Example 11.2 Show that for any closed surface

$$\iint_S (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{n}\, dS = 0$$

Answer:

We can convert the surface integral to a volume integral using the Divergence theorem:

$$\iint_S (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{n}\, dS = \iiint_V \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{u}) dV$$

The combination *div curl* is always zero, so this integral is zero. <u>You should explicitly calculate this for a cartesian vector field to convince yourself.</u> The proof relies on the fact that the second derivatives commute, just like when we proved that the curl of a conservative field is zero.

This can be used to demonstrate that Stokes' theorem (which we will shortly come to) only works for open surfaces.

Example 11.3 Find the relationship between the surface integral

$\iint_S \boldsymbol{r} \cdot \boldsymbol{n}\, dS$ and the volume V enclosed by the surface, if **r** is the position vector.

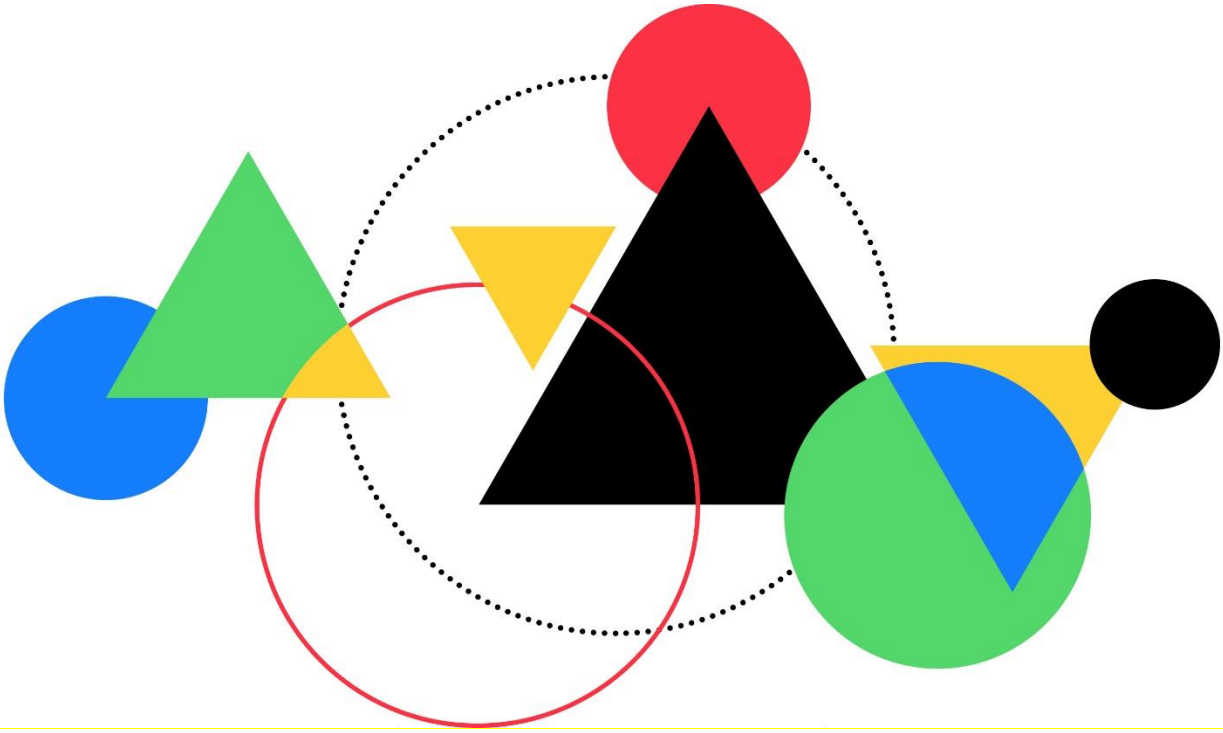
Answer:

Using the Divergence theorem

$$\iint_S \boldsymbol{r} \cdot \boldsymbol{n}\, dS = \iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{r}\, dV$$

Recalling that $\boldsymbol{\nabla} \cdot \boldsymbol{r} = 3$, we see that

$$\iint_S \boldsymbol{r} \cdot \boldsymbol{n}\, dS = \iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{r}\, dV = \iiint_V 3\, dV = 3V$$



The divergence of the electric field and the Dirac delta

11.ii

Gauss' law and the Divergence theorem

Gauss' law in integral form states that the flux of electric field through a surface is proportional to the enclosed charge:

$$\iint_S \boldsymbol{E} \cdot \hat{\boldsymbol{n}}\, dS = \frac{q}{\epsilon_0}$$

From the Divergence theorem, we can rewrite Gauss' law in differential form as:

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0}$$

where ρ is the charge density (charge per unit volume).

However, you will be beginning to worry about an issue here... we have already shown that the divergence of the electric field is zero, which would seem to be a problem when there is some enclosed charge.

The solution to this problem is interesting. Take first some simple vector field **v** which has the form of the electric field, i.e.

$$\boldsymbol{v} = \frac{1}{r^2}\hat{\boldsymbol{r}}$$

If we calculate the surface integral of this field over a sphere of radius R, we see that it is

$$\iint_S \boldsymbol{F} \cdot \hat{\boldsymbol{r}}\, dS = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \left(\frac{1}{r^2}\right) r^2 \sin\theta \, d\theta d\phi = 4\pi$$

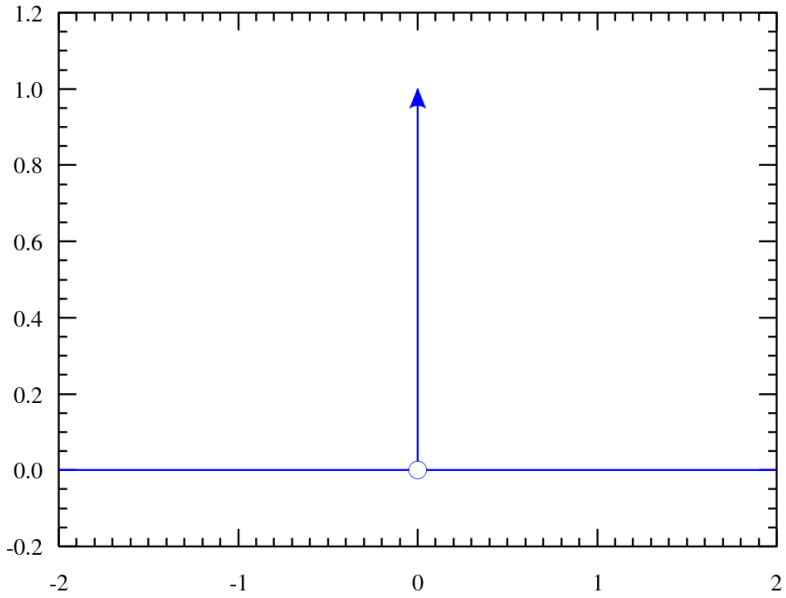
and yet… we know that the volume integral of the divergence is zero. Note also that the result is *independent of R* – the surface integral of such a field on a sphere is 4π regardless of the size of the sphere.

This would appear to break the divergence theorem. The solution lies in the fact that the vector function is not defined at all at the origin. In fact the divergence theorem *is* correct: such a field has a divergence which is zero everywhere except at the origin, and yet the volume integral of the divergence over any volume which contain the origin is 4π.

This leads us to the **dirac delta function** δ(x):

This is a function which has the value zero everywhere except at the origin, where it is an infinitely narrow spike of infinite height and integrated area 1.

The most important characteristic of the dirac delta function is that it 'picks' out the value of a function a zero.



As a result,

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = f(0)\int_{-\infty}^{+\infty} \delta(x)dx = f(0)$$

we may also define δ(x-a) which has the same properties but now with the 'delta function' at x=a rather than at the origin. Hence:

$$\int_{-\infty}^{+\infty} f(x)\delta(x-a)dx = f(a)$$

In three dimensions, the Dirac delta function is defined as $\delta^3(\mathbf{r})=\delta(x)\delta(y)\delta(z)$, which is again zero everywhere except the origin.

What does this have to do with the divergence theorem and the electric field? We can now use the Dirac delta to formally solve the problem.

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\boldsymbol{r}}}{r^2}\right) = 4\pi\delta^3(\boldsymbol{r})$$

In the specific case of the electric field, $E = \frac{Q\hat{r}}{4\pi\epsilon_0 r^2}$, the divergence is now given by:

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{Q}{\epsilon_0}\delta^3(\boldsymbol{r})$$

and so

$\int_V \nabla \cdot \boldsymbol{E} = \int_V \frac{Q}{\epsilon_0}\delta^3(\boldsymbol{r})$ which is equivalent to $\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0}$.

Finally, you may wish to consider that the Kronecker delta is the discrete form of the Dirac delta.

Physical application of the divergence theorem – continuity equation

Consider a compressible fluid with position-dependent and time-varying density and velocity fields given by $\rho(\boldsymbol{r}, t)$ and $v(\boldsymbol{r}, t)$.

For an arbitrary volume V, with surface area S, in the fluid, conservation of mass tells us that

$$\frac{dM}{dt} = -\iint_S \rho \boldsymbol{v} \cdot \boldsymbol{dS}$$

i.e. the rate of change of enclosed mass is equal to the rate at which it is flowing out via the surface.

Now the mass M equals

$M = \iiint_V \rho \, dV$ and so

$$\frac{d}{dt}\iiint_V \rho \, dV + \iint_S \rho \boldsymbol{v} \cdot \boldsymbol{dS} = 0$$

We can take the time derivative inside the left-hand term, and replace the right-hand term using the Divergence theorem:

$$\iiint_V \frac{\partial \rho}{\partial t} dV + \iiint_V \boldsymbol{\nabla} \cdot \rho \boldsymbol{v} \, dV = 0$$

Combining the terms which are now under exactly the same volume integral, we see that

$\iiint_V \left(\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \rho \boldsymbol{v}\right) dV = 0$ and hence $\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \rho \boldsymbol{v} = 0$, which is the *continuity equation* which is used in many places in physics.

We may expand the divergence term:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla}\rho + \rho \, \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0$$

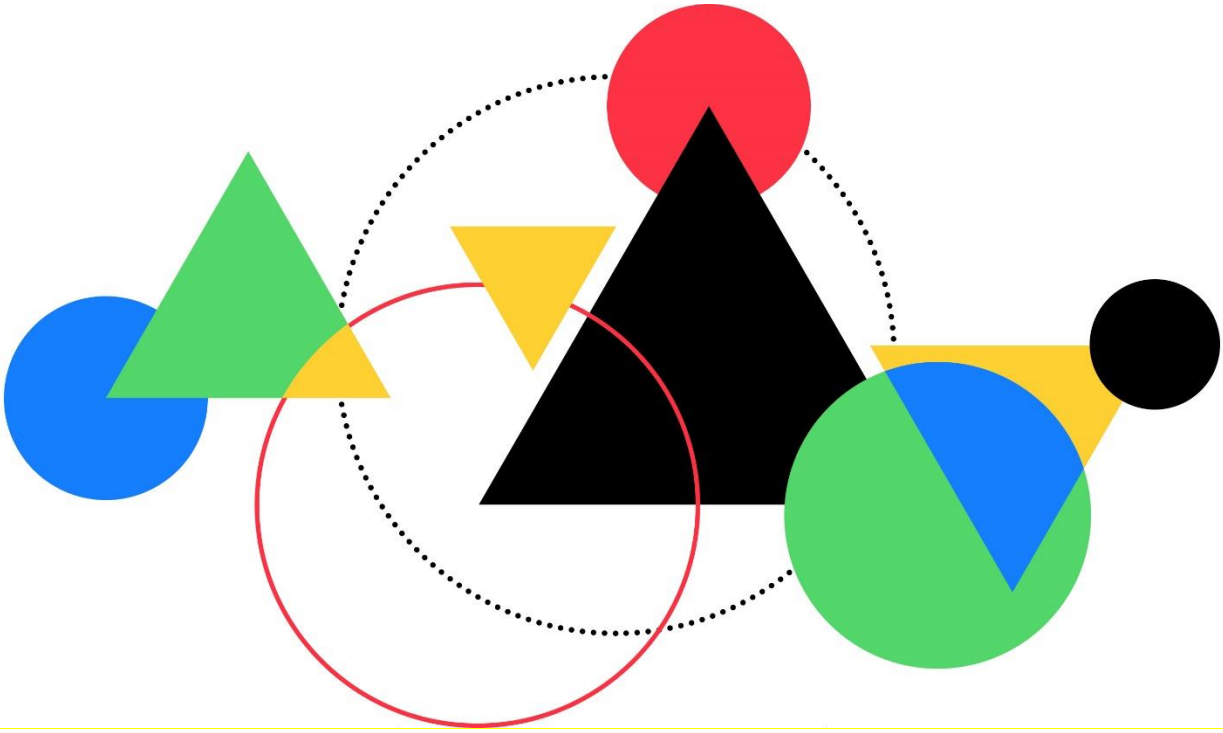
and if we assert that the fluid is constant and uniform then the equation simplifies to

$$\boldsymbol{\nabla} \cdot \boldsymbol{v} = 0$$

– and a fluid obeying this is said to be *incompressible.*

CP4: Vector Calculus and Multiple Integrals

Rob Fender, HT 2020



Stokes' theorem

12.i

Stokes' Theorem

The 2nd integral theorem we are going to deal with is Stokes' Theorem. Stokes' theorem could well be referred to as the 'curl theorem' and relates the surface integral of the curl of a field to the line integral around its boundary.

It states

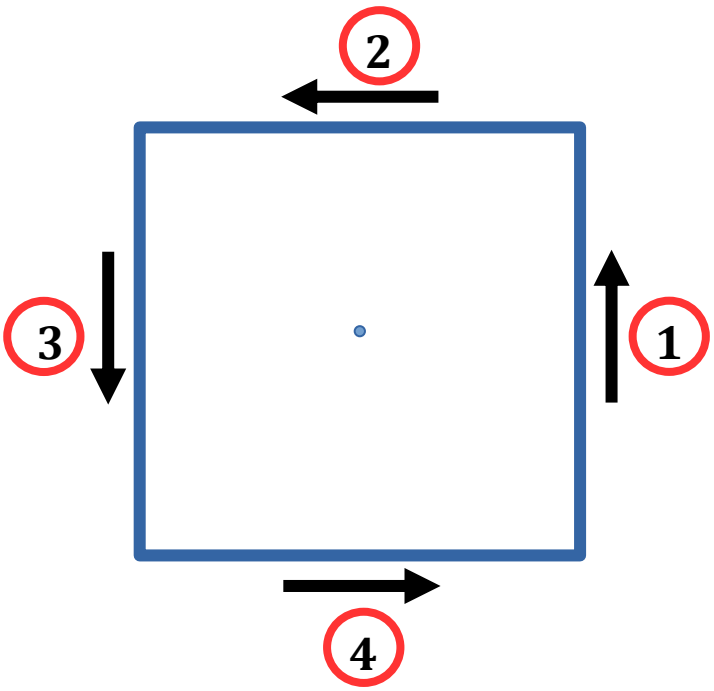
$$\iint_S (\nabla \times F) \,.\, dS = \oint_C F \,.\, dr$$

– the vector surface integral over some open surface S of the curl of the vector field **F** is the line integral of the vector field around the boundary C of S.

How does this come about? We can revisit our definition of the curl, which we used in its derivation earlier in the lectures:

$$(\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot d\hat{\boldsymbol{n}} = \lim_{A \to 0} \left(\frac{1}{A} \oint_C \mathbf{F} \cdot \boldsymbol{dr} \right)$$

To understand this form of the curl, we can consider a square in the (x,y) plane, in a vector field **F**, with its centre at (x,y) and side lengths of δx,δy.



The line integral around this square is the sum of that around the four sides. The curl would be in the **k** direction (so that this can be considered one component of the 3D curl).

1. Along path 1, the line integral is $F_y(x + \delta x, y)\delta y$, and along path 3 it is -$-F_y(x - \delta x, y)\delta y$. The sum of the paths 1 and 3 is therefore

$$\left[F_y(x+\delta x, y) - F_y(x-\delta x, y)\right]\delta y = \frac{\partial F_y}{\partial x} dA$$

using the same sense of substitutions we used in the derivation of the divergence.

2. Similarly the sum of paths 2 and 4 is

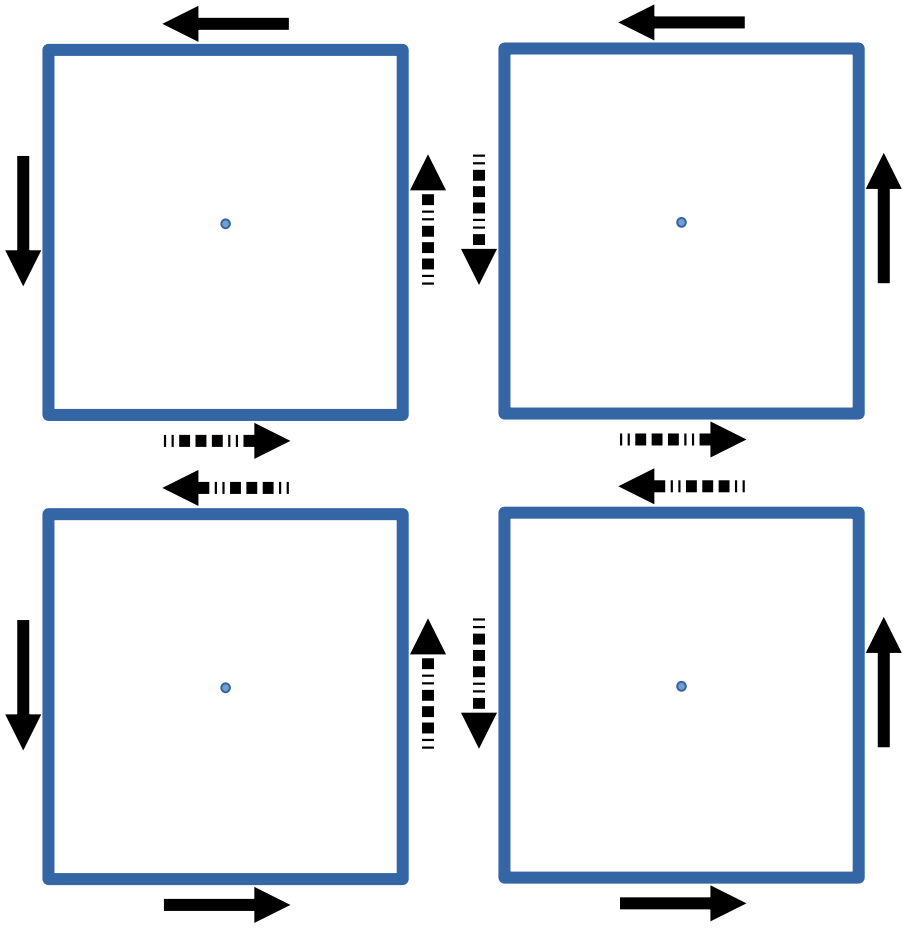
$$-\left[F_x(x, y+\delta y) + F_x(x, y-\delta y)\right]\delta x = \frac{-\partial F_x}{\partial t} dA$$

We therefore see that the total line integral is equal to the component of the curl in the **k** direction multiplied by the area of the square.

This is of course only strictly correct in the limit that the area becomes infinitesimally small, and hence:

$$(\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot d\hat{\boldsymbol{n}} = \lim_{A \to 0} \left(\frac{1}{A} \oint_C \mathbf{F} \cdot \boldsymbol{dr} \right)$$

If we now consider a surface composed of a large number of such squares, we can see that all of the 'internal' components of the line integral will cancel (much as the surface integrals of facing sides cancelled in our derivation of the divergence theorem). This is illustrated in the figure below: the dashed arrows are exactly cancelled out and the only contributor to the line integral is the outside.



And so we arrive at Stokes' theorem where we find that the surface integral of the normal component of the curl is equal to the line integral around the perimeter of the shape.

$$\iint_S (\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \boldsymbol{dS} = \oint_C \boldsymbol{F} \cdot \boldsymbol{dr}$$

In two dimensions, Stokes' theorem reduces to *Green's theorem:*

$$\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = \iint_S \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) dA$$

which is more commonly written as

$$\oint_C L\,dx + M dy = \iint_S \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} \right) dx dy$$

Example 12.1 Use Green's theorem to evaluate

$\oint_C y^3\,dx - x^3 dy$ where C is a circle of radius 2, with the correct orientation, centred at the origin.

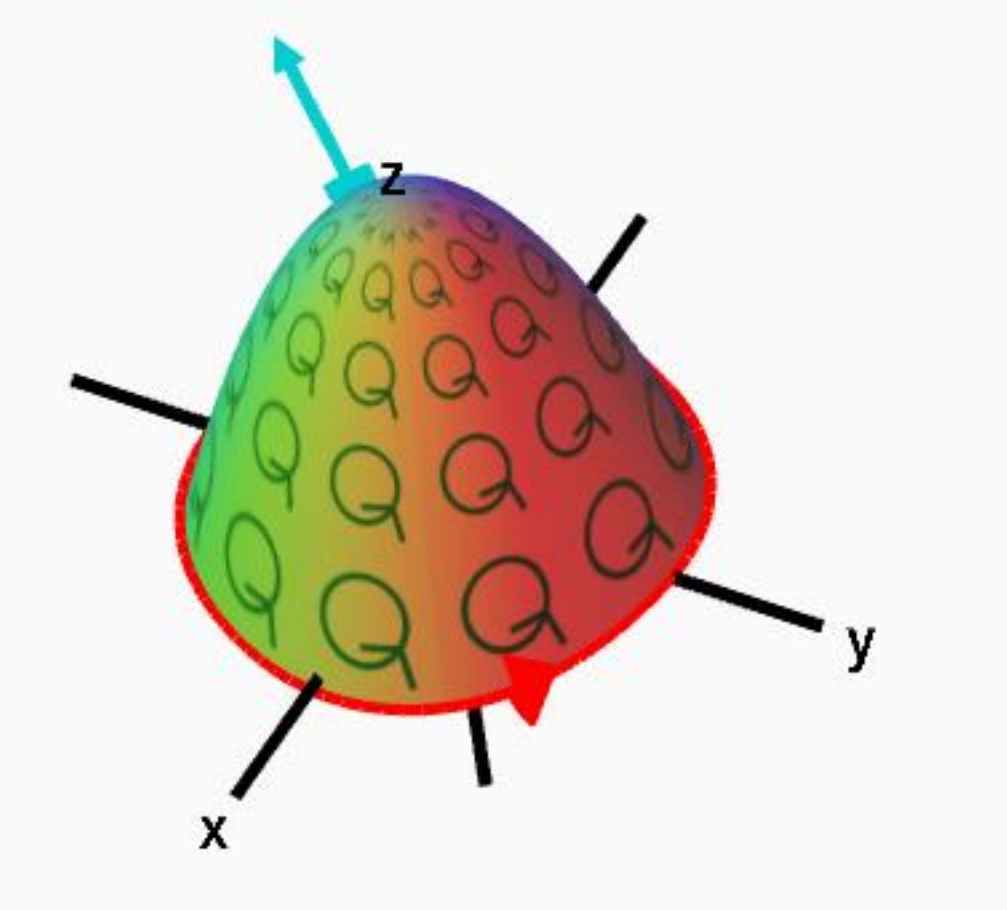
Answer:

Identifying $L = y^3$ and $M = x^3$, and switching to plane polar coordinates, we get

$$\oint_C y^3\,dx - x^3 dy = \int_{\phi=0}^{2\pi} \int_{r=0}^{2} -3r^3 dr = -24\pi$$

(by now you will be used to inserting the Jacobian "r" term when converting from cartesian to plane polar).

Stokes' theorem, though, does not just count for areas in a plane – it works for surfaces in 3D space and works for *any* surface (the 'capping surface') which is bounded by the same perimeter.

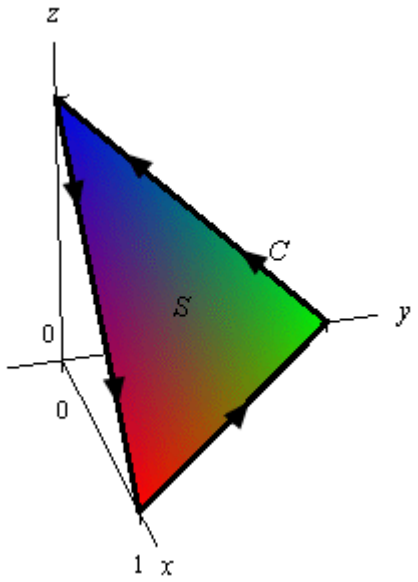


Note that there is a sign convention which is important for Stokes' theorem, which can be remembered thus: if you are walking around the loop which defines the line integral, and your head is in the direction of the positive normal vectors, as you walk around the loop the surface should be on your left. Or you can just remember what the diagram above looks like.

Example 12.2 Use Stokes' theorem to evaluate

$\int_C \boldsymbol{F} \cdot \boldsymbol{dr}$ where $\boldsymbol{F} = z^2\boldsymbol{i} + y^2\boldsymbol{j} + x\boldsymbol{k}$, and C is the triangle with vertices (1,0,0), (0,1,0) and (0,0,1) with counter-clockwise rotation.

Answer: We will need to define a surface over which to evaluate the surface integral (since we are going to use Stokes' theorem). The figure on the right shows this plane, which we can see helpfully projects in a straightforward way on to the (x,y) plane.



The equation of this plane is

$$x + y + z = 1$$

Stokes' theorem states

$\iint_S (\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \boldsymbol{dS} = \oint_C \boldsymbol{F} \cdot \boldsymbol{dr}$ and so we are going to need to evaluate the curl of the vector field **F** and then project it onto the (x,y) plane (since we are tackling it in a non-parametrized way).

$$\text{curl } F = \boldsymbol{\nabla} \times \boldsymbol{F} = det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z^2 & y^2 & x \end{bmatrix} = (2z-1)\boldsymbol{j}$$

$$\iint_S (\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \boldsymbol{dS} = \iint_S (2z-1)\,\boldsymbol{j} \cdot \boldsymbol{dS} = \iint_D (2z-1)\,\boldsymbol{j} \cdot \frac{\boldsymbol{\nabla} f}{|\boldsymbol{\nabla} f|} |\boldsymbol{\nabla} f| dA$$

The scalar function for which we are calculating the gradient is simply

$f(x,y,z) = x + y + z - 1$. In case you are concerned by this leap, if we follow the steps we learned earlier for surface integrals, we can see that this is consistent with saying $g(x,y) = z = 1 - x - y$, and $f(x,y,z) = z - g(x,y)$. So, we find that

$\boldsymbol{\nabla} f = \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}$

(which we note points upwards). The projection onto the (x,y) plane is defined by

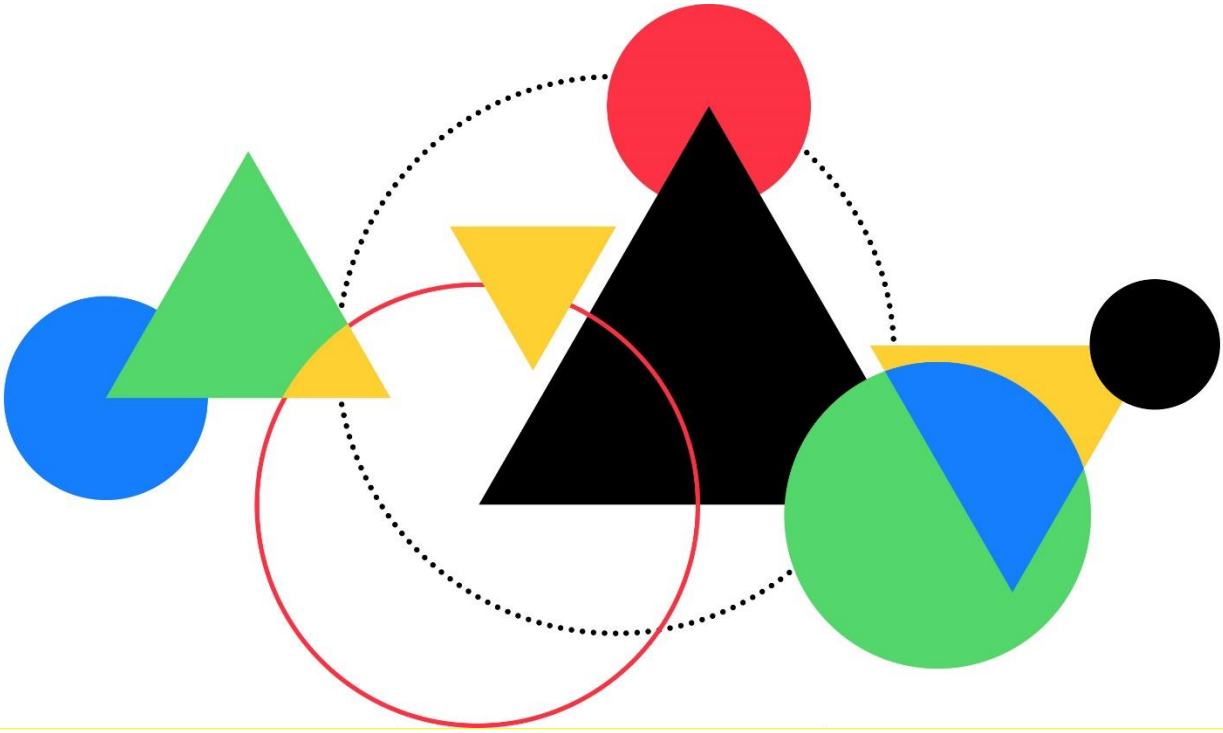
$0 \le x \le 1$ and $0 \le y \le 1 - x$

Our integral is now ready for evaluation:

$$\iint_D (2z-1)\,\boldsymbol{j} \cdot \frac{\boldsymbol{\nabla} f}{|\boldsymbol{\nabla} f|} |\boldsymbol{\nabla} f| dA = \int_{x=0}^{1} \int_{y=0}^{1-x} (2z-1)\,\boldsymbol{j} \cdot (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) dx dy$$

making the substitution z=1-x-y, we get:

[algebra]… =-1/6

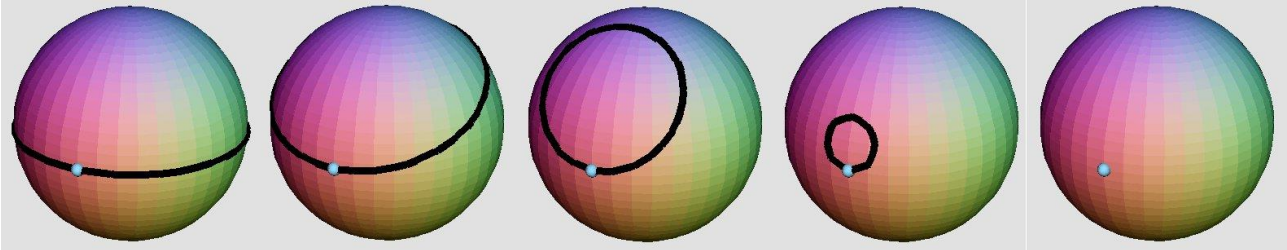


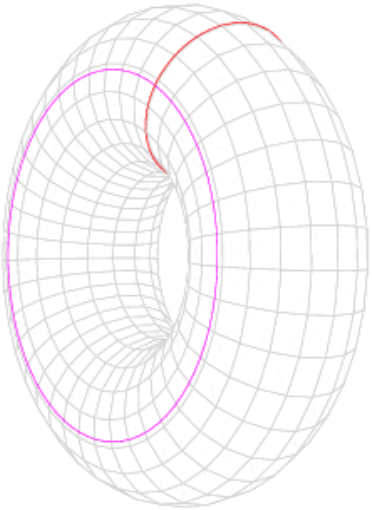
The End

12.ii

Simply- and not-simply connected surfaces

Note that Stokes' theorem only works for *simply-connected spaces*. These are surfaces in which a loop could be (slowly) contracted to zero length without ever leaving that space. The surface or interior of a sphere are examples of simply-connected spaces.





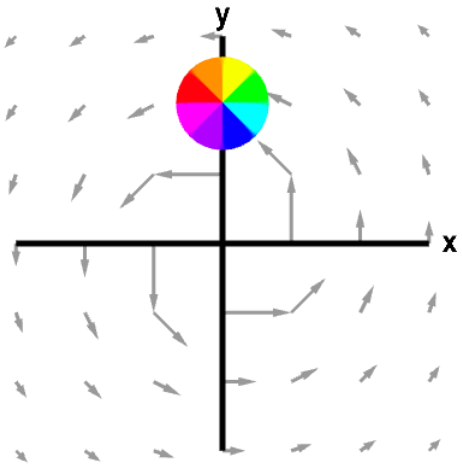
Consider however the surface or interior of a torus, these are not simply connected. Neither the red nor magenta lines could be shrunk to infinitesimally small length without leaving the surface. Equally, a loop inside the torus could not shrink beyond the point at which it reached the central hole.

Take a moment to consider how this is different to the case of the sphere.

Curl zero field with large-scale circulation

In lecture 10 we considered the following field:

$$F = \frac{-y\boldsymbol{i} + x\boldsymbol{j}}{x^2 + y^2}$$



which has curl zero and yet has large-scale circulation. Some questions arise – how, looking at the sketch, does the coloured ball not rotate (even though we have satisfied ourselves that *formally* the curl is zero), and – more fundamentally – does this violate Stokes' law?

Firstly – why do we state that the ball would not rotate when clearly the arrow (representing the fluid flow) are larger on the inside face than the outside, which should result in rotation. The reason here is that the figure is misleading – the ball is meant simply to represent the curl. In reality a surface element would have an outer face which was large than the inside face by a factor $(r_{out}/r_{\text{in}})^2$ which precisely compensates for the reduction in the strength of

the vector field with radial distance. This is directly analogous to the reason that the divergence of the electric field is zero.

But what about the large-scale circulation? The line integral around a closed circular loop centred on the origin would clearly not equal zero, and yet the curl is zero as we have already calculated. This is an example of a field with curl zero which is *not conservative.* We can calculate directly the integral around the line by noting that it is equivalent to:

$$F = \frac{\hat{\boldsymbol{\phi}}}{r}$$

(check you understand why), and we can evaluate the integral around a loop centred on the origin:

$$\oint_{\phi=0} \frac{1}{r} r \, d\phi = 2\pi$$

independent of the radius at which the loop is made (this should be reminiscent of the surface integral of the electric field through a sphere of radius R, which was $4\pi, independent\ of\ R$)

This seems to directly contradict Stokes' law. In fact, just like the case where the divergence theorem was in trouble (previous lecture), we see that this is because the field is not defined at the origin. In fact, this field is not defined *at all* on the z-axis. This in turn means that such a loop around the origin would not be in a simply-connected space (but in fact something like a torus), hence saving Stokes' theorem.

Loose ends

Combinations of div, grad and curl. Vector calculus and electromagnetism. Some further notes on index notation.

Second order and combined terms

Let's consider combinations of div, grad and curl.

Four combinations can be discarded as they do not make sense because they have the wrong form.

Firstly, recalling that the gradient only works on scalar fields, then we can see that

1. grad (grad U)

2. grad (curl **F**)

do not exist. Nor do

3. curl (div **F**)

4. div (div **F**)

because curl and div need to operate on a vector, but div returns a scalar.

We have already shown that the curl of a conservative field is zero, hence:

5. curl (grad U) = 0

and also that

6. div (curl **F**) = 0

There are three, more complex, combinations of terms:

7. div (grad U) = $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} U = \boldsymbol{\nabla}^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2}$

which is the *Laplacian.*

Sometimes we will speak of the Laplacian of a vector field. This is simply a **vector** field whose x-component is the Laplacian of F_x etc:

$$\boldsymbol{\nabla}^2 \boldsymbol{F} = (\nabla^2 F_x)\boldsymbol{i} + \left(\nabla^2 F_y\right)\boldsymbol{j} + (\nabla^2 F_z)\boldsymbol{k}$$

8. grad (div **F**) = $\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{F})$

$$= \left(\frac{\partial^2 F_x}{\partial x^2} + \frac{\partial^2 F_y}{\partial x \partial y} + \frac{\partial^2 F_z}{\partial x \partial z}\right)\boldsymbol{i} + \left(\frac{\partial^2 F_x}{\partial x \partial y} + \frac{\partial^2 F_y}{\partial y^2} + \frac{\partial^2 F_z}{\partial y \partial z}\right)\boldsymbol{j} + \left(\frac{\partial^2 F_x}{\partial x \partial z} + \frac{\partial^2 F_y}{\partial y \partial z} + \frac{\partial^2 F_z}{\partial z^2}\right)\boldsymbol{k}$$

this is rarely encountered, and has no special name of its own. It is **not** the same as the Laplacian of a vector.

9. curl (curl **F**) = $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{F}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{F}) - \boldsymbol{\nabla}^2 \boldsymbol{F}$

Maxwell's equations

The classical equations of electromagnetism, known as Maxwell's equations, are written in terms of vector calculus expressions:

$$\boldsymbol{\nabla}.\boldsymbol{E} = \frac{\rho}{\epsilon_0}$$

$$\boldsymbol{\nabla}.\boldsymbol{B} = 0$$

$$\boldsymbol{\nabla} \times \boldsymbol{E} = \frac{-\partial B}{\partial t}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_o j + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}$$

where ρ is the charge density, and **j** is the current density.

Consider the case of a vacuum with no enclosed charge, the equations then become

$\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0$ (i)

$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ (ii)

$\boldsymbol{\nabla} \times \boldsymbol{E} = \frac{-\partial B}{\partial t}$ (iii)

$\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_0 \epsilon_0 \frac{\partial E}{\partial t}$ (iv)

If we take the curl of the curl of the electric field, we get:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{E}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{E}) - \boldsymbol{\nabla}^2 E$$

We know from (i) that div(E)=0 and, combining with (iii), we get:

$$\boldsymbol{\nabla}^2 \boldsymbol{E} = \boldsymbol{\nabla} \times \frac{\partial B}{\partial t}$$

$$\boldsymbol{\nabla}^2 E = \frac{\partial}{\partial t}\left(\mu_0 \epsilon_0 \frac{\partial E}{\partial t}\right)$$

which leads to

$$\frac{\partial^2 E}{\partial t^2} = c^2 \boldsymbol{\nabla}^2 E$$

which is the equation for a wave propagating at speed c, where

$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$ is the speed of light.

Notes on index notation in vector calculus

In the following we will relabel the (x,y,z) axes as (x_1, x_2, x_3).

Gradient

The gradient is

$$\nabla U = \left(\boldsymbol{i}\frac{\partial U}{\partial x} + \boldsymbol{j}\frac{\partial U}{\partial y} + \boldsymbol{k}\frac{\partial U}{\partial z}\right)$$

So we see that component i of grad(U) is equal to the partial derivative of U with respect to x_i

$$|\boldsymbol{\nabla} U|_i = \frac{\partial U}{\partial x_i}$$

thus in fact we can see that the vector operator can itself be written in index notation:

$$|\boldsymbol{\nabla}|_i = \frac{\partial}{\partial x_i}$$

Divergence

The divergence

$$\text{div}\, \boldsymbol{F} = \boldsymbol{\nabla} \cdot \boldsymbol{F} = \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\right)$$

may be written in index notation as

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \frac{\partial F_j}{\partial x_j}$$

note the summation convention – indicated by the repeated j indices.

Curl

The curl is given by:

$$\text{curl}\,\boldsymbol{F} = \boldsymbol{\nabla} \times \boldsymbol{F} = det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

Consider the first component of the curl:

$$|\boldsymbol{\nabla} \times \boldsymbol{F}|_1 = \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} = \epsilon_{1jk}\frac{\partial F_k}{\partial x_j}$$

(recall our definition of the cross product in terms of index notation)

We can then see that the general form is:

$$|\boldsymbol{\nabla} \times \boldsymbol{F}|_i = \epsilon_{ijk}\frac{\partial F_k}{\partial x_j}$$

Example 13.1 Evaluate curl (curl **F**) using index notation

Answer:

$$[\nabla \times (\nabla \times F)]_i = \epsilon_{ijk}\frac{\partial}{\partial x_i}\epsilon_{klm}\frac{\partial F_m}{\partial x_l} = \epsilon_{ijk}\epsilon_{klm}\frac{\partial^2 F_m}{\partial x_j \partial x_l}$$

Using the result

$$\epsilon_{ijk}\epsilon_{klm} = \left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)$$

We can see that

$$\epsilon_{ijk}\epsilon_{klm}\frac{\partial^2 F_m}{\partial x_j \partial x_l} = \left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)\frac{\partial^2 F_m}{\partial x_i \partial x_l} = \frac{\partial^2 F_j}{\partial x_j \partial x_i} - \frac{\partial^2 F_i}{\partial x_j \partial x_j}$$

(note that this works because the first term only has non-zero terms when m=j and l=i, and the second when m=i and l=j)

This corresponds to the earlier stated result:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{F}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{F}) - \boldsymbol{\nabla}^2 \boldsymbol{F}$$

Consider this for a few moments and it should be clear.

This is the end of the course.

Acknowledgements: I have made use extensively of material provided by Steve Roberts, whom I would like to thank for his generosity. I have also ‘borrowed’ material from the lecture notes of Prof R Golestanian, the late Prof Steve Rawlings and the recommended course books. Many thanks to Tony Middleton for proof-reading the first half of these notes.

Comments, corrections and constructive criticisms welcome at **rob.fender@physics.ox.ac.uk**