Prepared for submission to JHEP

Vector Calculus IA

B. C. Allanach[a] **J.M. Evans**[a]

[a]*Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

E-mail: B.C.Allanach@damtp.cam.ac.uk

Abstract: These are lecture notes for the Cambridge mathematics tripos Part IA Vector Calculus course. There are four examples sheets for this course. These notes are pretty much complete.

Books See the schedules for a list, but particularly:

- "Mathematical Methods for Physics and Engineering", CUP 2002 by Riley, Hobson and Bence £28.
- "Vector Analysis and Cartesian Tensors", Bourne and Kendall 1999 by Nelson Thomas £30.

Future lecturers: please feel free to use/modify this resource. If you do though, please append your name to the current author list.

Contents

---

1 Derivatives and Coordinates

1.1 Differentiation Using Vector Notation

1.1.1 Vector function of a scalar

A vector function $\underline{F}(u)$ is 'differentiable' at u if

$$\delta \underline{F} = \underline{F}(u+\delta u) - \underline{F}(u) = \underline{F}'(u)\delta u + o(\delta u) \text{ as } \delta u \to 0,$$

and the <u>derivative</u> is the vector

$$\underline{F}'(u) = \frac{d\underline{F}}{du} = \lim_{\delta u \to 0} \frac{1}{\delta u}\left[\underline{F}(u+\delta u) - \underline{F}(u)\right].$$

Limits of vectors are defined using the norm(length) so $\underline{v} \to c$ iff $|\underline{v} - \underline{c}| \to 0$ and $\underline{a}(h) = \underline{b}(h) + \underline{n}o(h)$ iff $|\underline{a}(h) - \underline{b}(h)| = o(h)$, for $\underline{n}$ some unit vector.

<u>Leibniz identities</u> hold for appropriate products of scalar functions $f(u)$ and vectors $\underline{F}(u), \underline{G}(u)$:

$$(f\underline{F})' = f'\underline{F} + f\underline{F}', \qquad (\underline{F}\cdot\underline{G})' = \underline{F}'\cdot\underline{G} + \underline{F}\cdot\underline{G}',$$
$$(\underline{F}\wedge\underline{G})' = \underline{F}'\wedge\underline{G} + \underline{F}\wedge\underline{G}'.$$

Vectors can be differentiated component by component:

$$\underline{F}(u) = F_i(u)\underline{e}_i \Rightarrow \underline{F}'(u) = F_i'(u)\underline{e}_i$$

provided the basis $\{\underline{e}_i\}$ is independent of u (note the implicit summation convention over repeated indices).

The definition of the derivative can be abbreviated using differential notation $d\underline{F} = \underline{F}'(u)\ du$ in which $o-$terms are suppressed, compared to the relation above with small but finite changes.

Example: A point particle of mass m has position $\underline{r}(t)$ (a function of time t), velocity[a] $\dot{\underline{r}}(t)$, acceleration $\ddot{\underline{r}}(t)$ and momentum $\underline{p} = m\dot{\underline{r}}(t)$. If $\underline{F}(\underline{r})$ is the force on a particle, then $\dot{\underline{p}} = m\ddot{\underline{r}} = \underline{F}(\underline{r})$ is Newton's Second Law. We define the angular momentum about the origin $\underline{L} = \underline{r}\wedge\underline{p} = m\underline{r}\wedge\dot{\underline{r}} \Rightarrow \dot{\underline{L}} = \cancel{m\dot{\underline{r}}\wedge\dot{\underline{r}}} + m\underline{r}\wedge\ddot{\underline{r}}$, using Leibniz. Hence $\dot{\underline{L}} = \underline{r}\wedge\underline{F}$, or the torque of $\underline{F}$ about the origin.

[a] Note that derivatives with respect to t are conventionally denoted by dots instead of dashes.

1.1.2 Scalar function of position; gradient and directional derivatives

A scalar function $f(\underline{r})$ is differentiable at $\underline{r}$ if

$$\delta f(\underline{r}+\delta\underline{r}) - f(\underline{r}) = (\underline{\nabla} f)\cdot\delta\underline{r} + o(|\delta\underline{r}|) \text{ as } |\delta\underline{r}|\to 0.$$

$\underline{\nabla} f$ is a vector, the gradient of f at $\underline{r}$. The definition says that δf depends linearly on $\delta\underline{r}$, up to smaller $o-$terms.

Taking $\delta\underline{r} = h\underline{n}$ with $\underline{n}$ a unit vector,

$$\begin{aligned}\delta f = f(\underline{r}+h\underline{n}) - f(\underline{r}) &= (\underline{\nabla} f)\cdot(h\underline{n}) + o(h)\\ \Rightarrow \underline{n}\cdot\underline{\nabla f} &= \lim_{h\to 0}\frac{1}{h}\left[f(\underline{r}+h\underline{n}) - f(\underline{r})\right],\end{aligned}$$

the directional derivative of f along $\underline{n}$. Thus, the gradient contains information about how f changes as we move away from $\underline{r}$ to first order in the displacement.

Let $\underline{r} = x_i\underline{e}_i$ with $\{\underline{e}_i\}$ orthonormal. Setting $\underline{n} = \underline{e}_i$ for fixed i gives

$$\underline{e}_i\cdot\underline{\nabla} f = \lim_{h\to 0}\frac{1}{h}\left[f(\underline{r}+h\underline{e}_i) - f(\underline{r})\right] = \frac{\partial f}{\partial x_i}.$$

Hence the gradient in Cartesian coordinates is

$$\underline{\nabla} f = \frac{\partial f}{\partial x_i}\underline{e}_i. \tag{1.1}$$

With this choice of basis and coordinates, the definition of differentiability becomes

$$\delta f = \frac{\partial f}{\partial x_i}\delta x_i + o(\delta x)$$

as $\delta x = \sqrt{\delta x_i \delta x_i}\to 0$. In differential notation, we suppress $o-$terms:

$$df = \underline{\nabla} f\cdot d\underline{r} = \frac{\partial f}{\partial x_i}dx_i.$$

Example: $f(x,y,z) = x + e^{xy}\sin z$ at $(x,y,z) = (0,1,0)$.

$$\underline{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \left(1 + ye^{xy}\sin z,\ xe^{xy}\sin z,\ e^{xy}\cos z\right),$$

thus $\underline{\nabla} f = (1,\ 0,\ 1)$ at $(x,y,z) = (0,1,0)$. The rate of change of f along direction $\underline{n}$ is given by the directional derivative $\underline{n}\cdot\underline{\nabla} f$. So f increases most rapidly for $\underline{n} = +\frac{1}{\sqrt{2}}(1,0,1)$ and decreases most rapidly for $\underline{n} = -\frac{1}{\sqrt{2}}(1,0,1)$ at a rate $\underline{n}\cdot\underline{\nabla} f = \pm\sqrt{2}$. To first order, there is *no* change in f if $\underline{n}\perp(1,0,1)$.

1.1.3 The chain rule: a particular case

Consider a composition of differentiable functions $f(\underline{r}(u))$. A change δu produces a change $\delta\underline{r} = \underline{r}'\delta u + o(\delta u)$ and

$$\delta f = \underline{\nabla} f \cdot \delta\underline{r} + o(|\delta\underline{r}|) = \underline{\nabla} f \cdot \underline{r}'(u)\delta u + o(\delta u).$$

This shows that f is differentiable as a function of u and

$$\frac{df}{du} = \underline{\nabla} f \cdot \frac{d\underline{r}}{du},$$

a <u>chain rule</u>. In Cartesian coordinates (setting $\underline{r} = x_i\underline{e}_i$), $\frac{df}{du} = \frac{\partial f}{\partial x_i}\frac{dx_i}{du}$.

1.2 Differentiation Using Coordinate Notation

1.2.1 Differentiable functions $\Re^n \to \Re^m$

The functions $\underline{F}(u)$ and $f(\underline{r})$ discussed in sections 1.1.1,1.1.2 are maps $\Re \to \Re^3$ and $\Re^3 \to \Re$ while a vector field $\underline{F}(\underline{r})$ is a function $\Re^3 \to \Re^3$.

In general, $f : \Re^n_{\{x_i\}} \to \Re^m_{\{y_r\}}$ is equivalent to a set of functions $y_r = F_r(x_1, \ldots, x_n) = F_r(x_i)$ where $r = 1, \ldots, m$ and $i = 1, \ldots, n$. The function f is <u>differentiable</u> if $\delta y_r = M(f)_{ri}\delta x_i + o(\delta x)$ as $\delta x = \sqrt{\delta x_i \delta x_i} \to 0$, with <u>derivative</u> $M(f)_{ri} = \frac{\partial y_r}{\partial x_i} = \frac{\partial F_r}{\partial x_j}$, a $m \times n$ matrix of partial derivatives. For example, for $\Re^3 \to \Re$, we get a 1×3 matrix, or vector, gradient i.e. $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$.

A convenient abbreviation of the definition: *replace small changes by differentials and drop the $o-$terms, which are understood.*

$$dy_r = M(f)_{ri}dx_i = \frac{\partial y_r}{\partial x_i}dx_i.$$

A function is <u>smooth</u> if it can be differentiated any number of times, i.e. if all partial derivatives exist, for example $\frac{\partial^2 F_r}{\partial x_i \partial x_j}$, $\frac{\partial^3 F_r}{\partial x_i \partial x_j \partial x_k}$ etc, and these are totally symmetric in $i, j, k, \ldots$ For example, $\frac{\partial^2 F_r}{\partial x_i \partial x_j} = \frac{\partial^2 F_r}{\partial x_j \partial x_i}$. The functions that we will consider here will be smooth except for where things obviously go wrong, eg $f(x) = 1/x$ is smooth except at $x = 0$.

1.2.2 The chain rule - general version

Consider the functions

$$\begin{array}{rccccc} & & g & & f & \\ & \Re^p & \to & \Re^n & \to & \Re^m \\ \text{coordinates} & \{u_a\} & & \{x_i\} & & \{y_r\} \end{array},$$

with the ranges of the labels a, i, r understood. If f and g are smooth, then so is $f \circ g$, with derivative

$$M(f \circ g)_{ra} = M(f)_{ri}M(g)_{ia},$$

or matrix multiplication

$$\underset{m\times p}{\frac{\partial y_r}{\partial u_a}} = \underset{m\times n}{\frac{\partial y_r}{\partial x_i}}\ \underset{n\times p}{\frac{\partial x_i}{\partial u_a}} \text{ matrix type}$$

Thus, if we compare the functions, we also compose the derivatives (matrices) as linear maps.

$$\left.\begin{aligned} dy_r &= M(f)_{ri}dx_i \\ dx_i &= M(g)_{ia}du_a \end{aligned}\right\} \text{ combine and compare with } dy_r = M(f\circ g)_{ra}du_a.$$

In operator form, $\frac{\partial}{\partial u_a} = \frac{\partial x_i}{\partial u_a}\frac{\partial}{\partial x_i}$, which holds when acting on any function which depends on u_a through x_i.

For example, $y = f(x_i(u))$ $\quad \Re \to \Re^3 \to \Re$

$$\underset{1\times 1}{\frac{dy}{du}} = \underset{1\times 3}{\frac{\partial y}{\partial x_i}}\ \underset{3\times 1}{\frac{dx_i}{du}}.$$

Compare with section 1.1.3.

1.2.3 Inverse functions

With the notation of the last section, take $m = n = p$ and let f, g be inverse functions, both smooth, with $y_r = u_r$:

$$\begin{array}{ccc} & g & \\ & \rightarrow & \\ \Re^n & & \Re^n \\ & \leftarrow & \\ & f & \\ \{u_a\} & & \{x_i\} \end{array}$$

Both $f\circ g$ and $g\circ f$ are identity functions on $\Re^n$. $M(f\circ g)_{ia}$ and $M(g\circ f)$ are $n\times n$ identity matrices. $M(f)_{ai}$ and $M(g)_{ia}$ are inverse matrices of each other. Equivalently, $\frac{\partial u_b}{\partial u_a} = \frac{\partial u_b}{\partial x_i}\frac{\partial x_i}{\partial u_a} = \delta_{ab}$; $\frac{\partial x_j}{\partial x_i} = \frac{\partial x_j}{\partial u_a}\frac{\partial u_a}{\partial x_i} = \delta_{ij}$.
For $n = 1$, we get the familiar result $\frac{du}{dx} = \frac{1}{dx/du}$.
For $n > 1$, we must invert *matrices* to relate $\frac{\partial x_i}{\partial u_a}$ and $\frac{\partial u_a}{\partial x_i}$.

For example: for $n = 2$, we write $u_1 = \rho$, $u_2 = \varphi$ and let $x_1 = \rho\cos\varphi, x_2 = \rho\sin\varphi$.

$$M(g) = \begin{pmatrix} \frac{\partial x_1}{\partial \rho} & \frac{\partial x_1}{\partial \varphi} \\ \frac{\partial x_2}{\partial \rho} & \frac{\partial x_2}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}.$$

The relations can be inverted i.e. $\rho = \sqrt{x_1^2 + x_2^2}$ and $\varphi = \tan^{-1}(x_2/x_1)$ (except that φ is not defined at $\rho = 0$ and it is defined up to a multiple of 2π for $\rho \neq 0$). Then, we can compute directly

$$M(f) = \begin{pmatrix} \frac{\partial \rho}{\partial x_1} & \frac{\partial \rho}{\partial x_2} \\ \frac{\partial \varphi}{\partial x_1} & \frac{\partial \varphi}{\partial x_2} \end{pmatrix} = M(g)^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{\rho}\sin\varphi & \frac{1}{\rho}\cos\varphi \end{pmatrix} \qquad (\rho \neq 0).$$

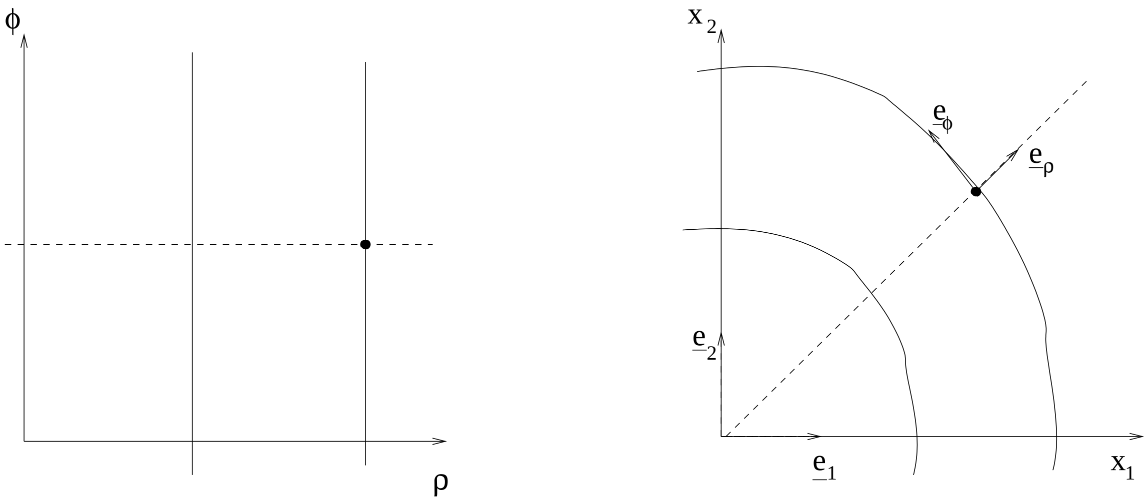


Figure 1. Plane polar coordinates. $\rho > 0$ and $0 \leq \varphi < 2\pi$ for uniqueness.

In general, for invertible smooth maps, as above, $\det M(g) = \det\left(\frac{\partial x_i}{\partial u_a}\right)$, $\det M(f) = \det\left(\frac{\partial u_a}{\partial x_i}\right)$ are called the Jacobians. They are non-zero, with

$$\det\left(\frac{\partial x_i}{\partial u_a}\right)\det\left(\frac{\partial u_a}{\partial x_i}\right) = 1, \text{ (no summation convention).}$$

1.3 Coordinate Systems

Now, we apply the results of the last section to changes of coordinates on Euclidean space. We can always choose to use Cartesian coordinates, but other choices are often more useful. In general, the coordinates $\{u_a\}$ on Euclidean space (or some subset of it) can be invertible functions of some Cartesian coordinates $\{x_i\}$ such that $u_a(\{x_i\})$ and $x_i(\{u_a\})$ are smooth.

In 2 dimensions, we define plane polar coordinates ρ, φ by

$$x_1 = \rho\cos\varphi, \qquad x_2 = \rho\sin\varphi,$$

as in Section 1.2.3, see Fig. 1. ρ and φ are coordinates (or parameters) labelling points in the plane, but they are not components of a position vector (whereas $\underline{r} = x_1\underline{e}_1 + x_2\underline{e}_2$ *is*). Nevertheless, we can associate basis vectors with these coordinates:

$$\underline{e}_\rho = \hat{\rho} = \cos\varphi\underline{e}_1 + \sin\varphi\underline{e}_2, \qquad \underline{e}_\varphi = \hat{\varphi} = -\sin\varphi\underline{e}_1 + \cos\varphi\underline{e}_2$$

are unit vectors in the directions of increasing ρ and φ, respectively, with the other held fixed. These vectors vary with position and are ill-defined at the origin.

In 3 dimensions, we define the coordinates as in Table 1.

cylindrical polars	spherical polars
ρ, φ, z $x_1 = \rho\cos\varphi$ $x_2 = \rho\sin\varphi$ $x_3 = z$	r, θ, φ $x_1 = r\sin\theta\cos\varphi$ $x_2 = r\sin\theta\sin\varphi,\ r = \lvert\underline{r}\rvert$ $x_3 = r\cos\theta$
Basis vectors	
$\underline{e}_\rho = \cos\varphi\underline{e}_1 + \sin\varphi\underline{e}_2$ $\underline{e}_\varphi = -\sin\varphi\underline{e}_1 + \cos\varphi\underline{e}_2$ $\underline{e}_z = \underline{e}_3$	$\underline{e}_r = \cos\theta\underline{e}_z + \sin\theta\underline{e}_\rho$ $\underline{e}_\varphi = -\sin\varphi\underline{e}_1 + \cos\varphi\underline{e}_2$ $\underline{e}_\theta = -\sin\theta\underline{e}_z + \cos\theta\underline{e}_\rho$.

Table 1. 3-dimensional cylindrical and polar coordinates.

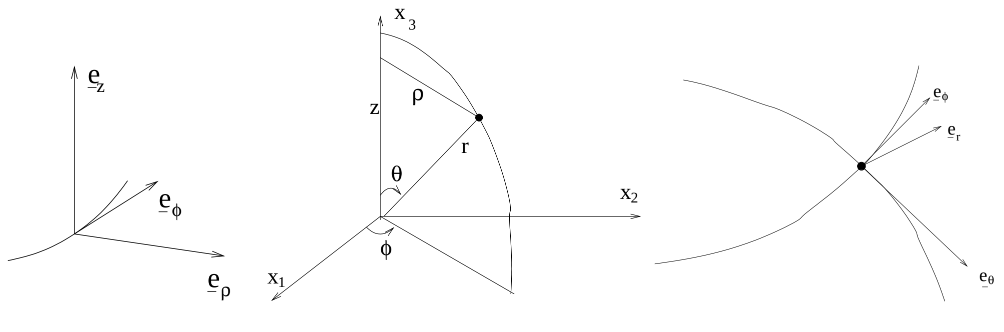


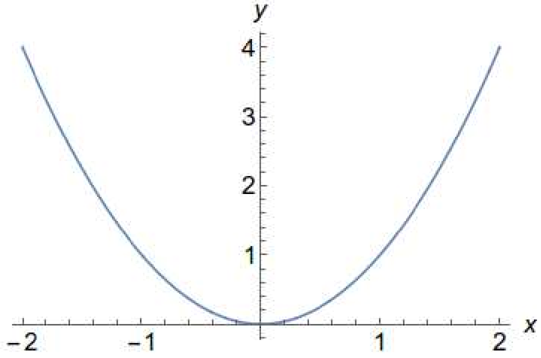
Figure 2. Cylindrical polar coordinates and spherical polar coordinates. $r = |\underline{r}|$, $0 \le \varphi < 2\pi$ and $0 \le \theta \le \pi$.

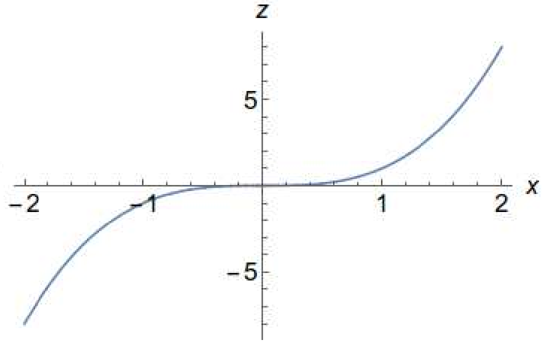
2 Curves and Line Integrals

2.1 Parameterised Curves, Tangents and Arc Length

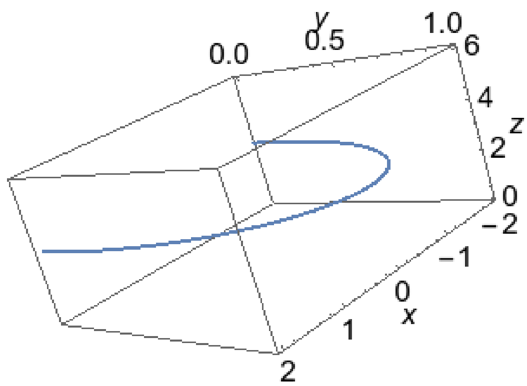
A parameterised curve C is given by a smooth function $\underline{r}(u)$ with u belong to some interval $u \subseteq \Re$.

Examples:
1. $\underline{r}(u) = u\hat{\underline{i}} + u^2\hat{\underline{j}} + u^3\hat{\underline{k}}$ for $-\infty < u < \infty$, a curve without end-points.

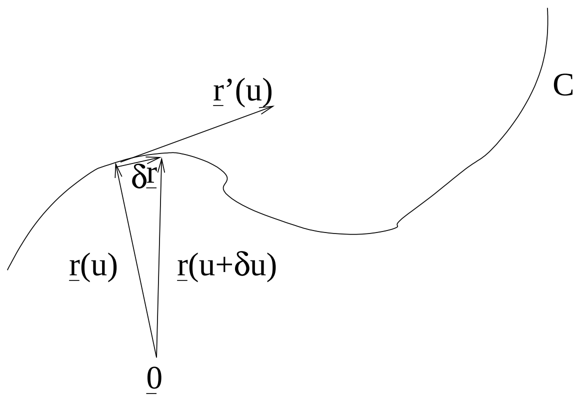




2. $\underline{r}(u) = 2\cos u\hat{\underline{i}} + \sin u\hat{\underline{j}} + 3\hat{\underline{k}}$ where $0 \le u \le \pi$, a curve with end-points.



This is a parameterisation of $x^2/4 + y^2 = 1,\ y \ge 0,\ z = 3$.



The derivative $\underline{r}'(u)$ is a vector tangent to the curve at each point, provided it is not zero. The parameterisation is called regular if $\underline{r}'(u) \neq 0$. We will assume this holds except perhaps at isolated points, so that the curve can always be divided into segments on which the parameterisation is regular.

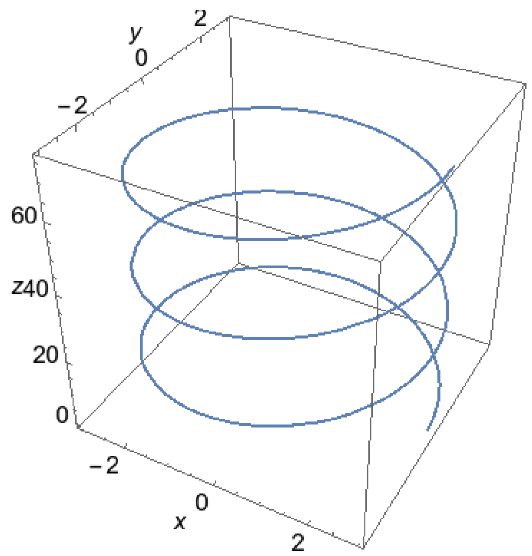
The arc length s measured along C, from some fixed point satisfies

$$\delta s = |\delta \underline{r}| + o(|\delta \underline{r}|) = |\underline{r}'(u)\delta u| + o(\delta u) \tag{2.1}$$

$$\Rightarrow \frac{ds}{du} = \pm\left|\frac{d\underline{r}}{du}\right| = \pm|\underline{r}'(u)|. \tag{2.2}$$

The sign fixes the direction of measuring s, for increasing or decreasing u.

Example: A helix $\underline{r}(u) = (3\cos u,\ 3\sin u,\ 4u) \Rightarrow \underline{r}'(u) = (-3\sin u,\ 3\cos u,\ 4)$.



$\frac{ds}{du} = |\underline{r}'(u)| = \sqrt{3^2+4^2} = 5$. Thus the arc length, measured from $\underline{r}(0) = 3\hat{\underline{i}}$ to $\underline{r}(u)$ $(u > 0)$ is $s(u) = 5u$. Hence, the distance from $\underline{r}(0) = 3\hat{\underline{i}}$ to $\underline{r}(2\pi) = 3\hat{\underline{i}} + 8\pi\hat{\underline{k}}$ is $s(2\pi) = 10\pi$.

The parameterisation of a curve can be changed: take $u \leftrightarrow \tilde{u}$ invertible and smooth (each is a smooth function of the other). The chain rule implies that

$$\frac{d\underline{r}}{d\tilde{u}} = \frac{d\underline{r}}{du}\Big/\frac{d\tilde{u}}{du}$$

so the tangent vector changes size, but not direction. Choosing $\tilde{u} = s$, the arc length, gives a tangent vector $\underline{t} = \frac{d\underline{r}}{ds}$ of unit length along the direction of increasing s.

$ds = \pm|\underline{r}'(u)|du$ is a scalar line element on C.

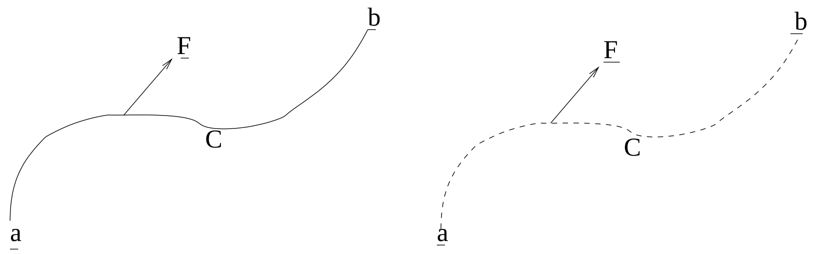
2.2 Line Integrals of Vector Fields

2.2.1 Definitions and Examples

Let C be a smooth curve $\underline{r}(u)$ with end points $\underline{r}(\alpha) = \underline{a}$ and $\underline{r}(\beta) = \underline{b}$, so that u runs over an interval with ends α and β. A choice of orientation is a direction $\underline{a} \to \underline{b}$ along C. The line integral of a smooth vector field $\underline{F}(\underline{r})$ along C with this orientation is defined by

$$\int_C \underline{F}(\underline{r}) \cdot d\underline{r} = \int_\alpha^\beta \underline{F}(\underline{r}(u)) \cdot \underline{r}'(u)du. \tag{2.3}$$

This can be regarded as the limit of a sum over small displacements corresponding to dividing the interval for u into small segments.



Each term is

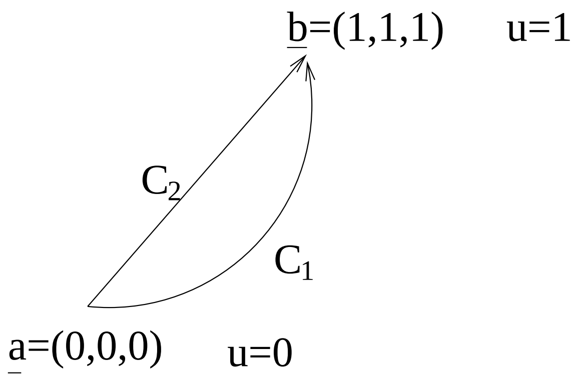
$$\underline{F}(\underline{r}) \cdot \delta \underline{r} = \underline{F}(\underline{r}(u)) \cdot \frac{d\underline{r}}{du} + o(\delta u).$$

The left hand side does not depend upon the parameterisation, and so neither does Eq. 2.3. This can also be checked by the chain rule:

$$\underline{F}(\underline{r}) \cdot \frac{d\underline{r}}{du} du = \underline{F}(\underline{r}) \cdot \frac{d\underline{r}}{d\tilde{u}} d\tilde{u}.$$

Using components, with $\underline{r} = x_i \underline{e}_i$ and $\underline{F} = F_i \underline{e}_i$, the line integral is $\int_C F_i dx_i = \int_\alpha^\beta F_i \frac{dx_i}{du}\ du$.
$d\underline{r} = \underline{r}'(u) du$ is the <u>line element</u> on C.

Examples: $\underline{F}(\underline{r}) = (xe^y,\ z^2,\ xy)$, so $\int_C \underline{F} \cdot d\underline{r} = \int_C xe^y\ dx + z^2\ dy + xy\ dz$.



We define $C_1 : \underline{r}(u) = (u, u^2, u^3) \Rightarrow \underline{r}'(u) = (1, 2u, 3u^2)$. On C_1, $\underline{F}(\underline{r}) = (ue^{u^2},\ u^6,\ u^3)$. Hence

$$\int_{C_1} \underline{F} \cdot d\underline{r} = \int_0^1 \underline{F} \cdot \underline{r}'(u)\ du = \int_0^1 ue^{u^2} + 2u^7 + 3u^5\ du = e/2 + 1/4.$$

For $C_2 : \underline{r}(t) = (t, t, t) \Rightarrow \underline{r}'(t) = (1, 1, 1)$ (we are now using t as the independent variable). On C_2, $\underline{F}(\underline{r}(t)) = (te^t, t^2, t^2)$. Hence

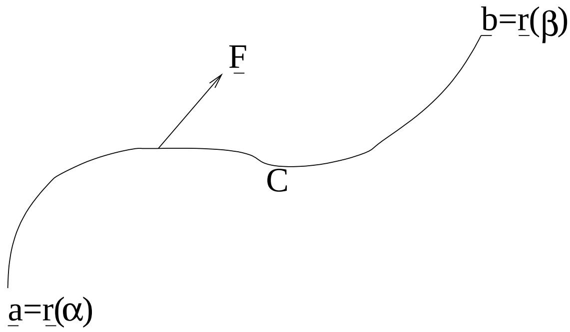
$$\int_{C_2} \underline{F} \cdot d\underline{r} = \int_0^1 (te^t + 2t^2)\ dt = [te^t]_0^1 - \int_0^1 e^t\ dt + 2/3 = 5/3.$$

A closed curve has $\underline{b} = \underline{a}$. The line integral is sometimes then written $\oint_C \underline{F} \cdot d\underline{r}$ and is called the circulation of $\underline{F}$ around C.

2.2.2 Comments

1. $\int_C \underline{F} \cdot d\underline{r}$ depends on C in general and not just upon the end points $\underline{a}$ and $\underline{b}$.

2. $\int_C \underline{F} \cdot d\underline{r} = \int_\alpha^\beta \underline{F} \cdot \underline{r}'(u)\ du$ for orientation $\underline{a} \to \underline{b}$. This holds whether $\alpha < \beta$ or $\alpha > \beta$.



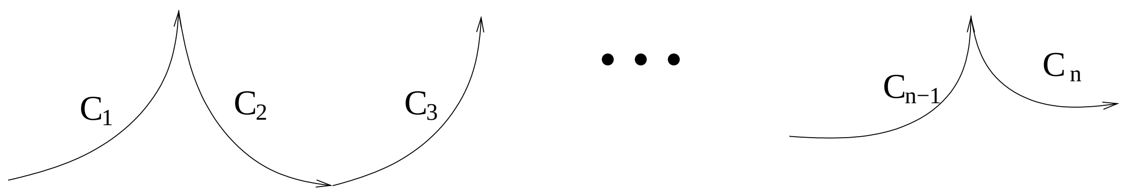
3. If C is the curve with orientation $\underline{a} \to \underline{b}$, we write $-C$ for a curve with orientation $\underline{b} \to \underline{a}$, and $\int_{-C} \underline{F} \cdot d\underline{r} = \int_\beta^\alpha \underline{F} \cdot \underline{r}'(u)\ du = -\int_C \underline{F} \cdot d\underline{r}$.

4. The orientation of a curve is also equivalent to a choice of unit tangent vector. Measuring the arc length s along the direction given by the orientation $d\underline{r} = \underline{t}\ ds$ defines a unit tangent vector $\underline{t}$ and $\int_C \underline{F} \cdot d\underline{r} = \int_C \underline{F} \cdot \underline{t}\ ds$. The integral can be regarded as the limit of a sum of terms $\underline{F} \cdot \delta\underline{r} = \underline{F} \cdot \underline{t}\delta s$. By convention, $\int_C f(s)\ ds$ is always evaluated in the sense of increasing s, so that $\int_C 1\ ds$ is the length of C and is greater than 0.

2.3 Sums of Curves and Integrals

A piecewise smooth curve C consists of a number of segments; we write $C = C_1 + C_2 + \cdots + C_n$, where each C_i is a smooth curve with a regular parameterisation. The end-points of successfive segments coincide, and there is a compatible choice of orientations:

C:

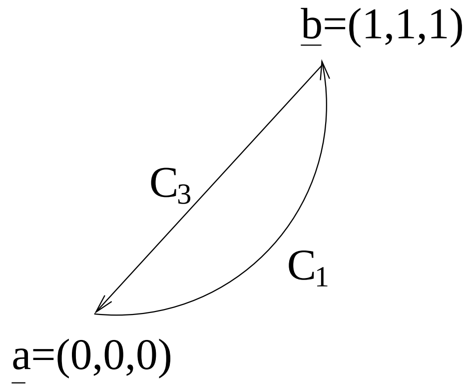


The line integral over a piecewise smooth curve C is

$$\int_C \underline{F} \cdot d\underline{r} = \int_{C_1} \underline{F} \cdot d\underline{r} + \int_{C_2} \underline{F} \cdot d\underline{r} + \ldots + \int_{C_n} \underline{F} \cdot d\underline{r}.$$

The parameterisations on each C_i can be chosen independently.

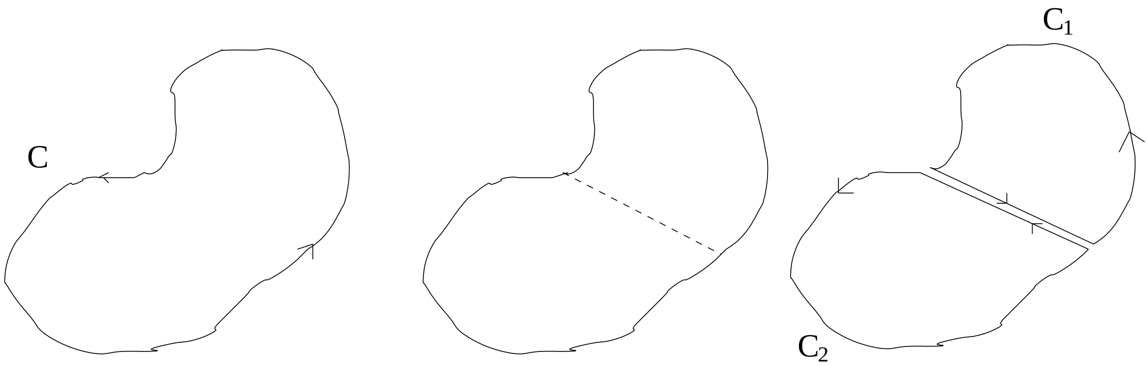
Example: (see Section 2.2.1)



$\int_C \underline{F}\cdot d\underline{r} = \int_C xe^y\ dx + z^2\ dy + xy\ dz$, where $C = C_1 + C_3$ is a piecewise smooth closed curve as shown with $C_3 = -C_2$. From our previous results,

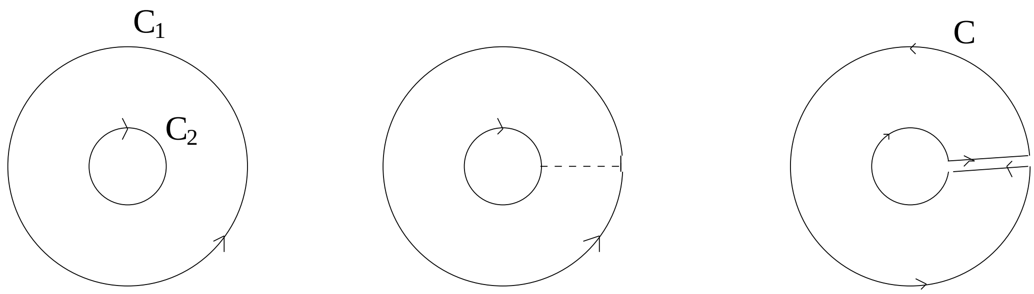
$$\oint_C \underline{F}\cdot d\underline{r} = \int_{C_1} \underline{F}\cdot d\underline{r} + \int_{C_3} \underline{F}\cdot d\underline{r} = \int_{C_1} \underline{F}\cdot d\underline{r} - \int_{C_2} \underline{F}\cdot d\underline{r}$$
$$= (e/2 + 1/4) - 5/3 = e/2 - 17/12.$$

It is convenient to extend sums to any sets of piecewise smooth curves, even if they are not connected. This allows for some useful constructions, for example (these rely on segments with the opposite orientations cancelling):



$$\oint_C \underline{F}\cdot d\underline{r} = \oint_{C_1} \underline{F}\cdot d\underline{r} + \oint_{C_2} \underline{F}\cdot d\underline{r}.$$

or



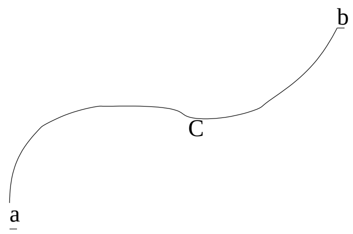
2.4 Gradients and Exact Differentials

2.4.1 Line Integrals and Gradients

If $\underline{F} = \underline{\nabla} f$ for some scalar field $f(\underline{r})$ then

$$\int_C \underline{F} \cdot d\underline{r} = f(\underline{b}) - f(\underline{a}), \tag{2.4}$$

for $\underline{\text{any}}$ curve C from $\underline{a}$ to $\underline{b}$.



Proof:

$$\int_C \underline{F} \cdot d\underline{r} = \int_C \underline{\nabla} f \cdot d\underline{r} = \int_\alpha^\beta \underline{\nabla} f \cdot \frac{d\underline{r}}{du}\ du$$

for any parameterisation of C with $\underline{a} = \underline{r}(\alpha)$ and $\underline{b} = \underline{r}(\beta)$.

$$= \int_\alpha^\beta \frac{d}{du}\left(f(\underline{r}(u))\right)\ du$$

by the chain rule, see section 1.1.3,

$$= [f(\underline{r}(u))]_\alpha^\beta = f(\underline{b}) - f(\underline{a}),$$

as required.

Note that

- When $\underline{F}$ is a gradient, the line integral depends only on the end points, not on the curve joining them.
- When C is a closed curve, $\oint_C \underline{F} \cdot d\underline{r} = 0$ if $\underline{F}$ is a gradient.
- If $\underline{F} = \underline{\nabla} f$, $\underline{F}$ is called a conservative vector field. There are a number of alternative definitions which are equivalent (with suitable assumptions) - see later.

2.4.2 Differentials

It is often convenient to work with differentials $\underline{F} \cdot d\underline{r} = F_i \, dx_i$ as objects that can be integrated along curves. Such a differential is called <u>exact</u> if it has the form $df = \underline{\nabla} f \cdot d\underline{r} = \frac{\partial f}{\partial x_i} \, dx_i$. So $\underline{F} = \underline{\nabla} f \Leftrightarrow F_i = \frac{\partial f}{\partial x_i} \Leftrightarrow F_i \, dx_i = df$ is exact. To test if this holds, we can use the necessary condition $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ (since both are equal to $\frac{\partial^2 f}{\partial x_i \partial x_j}$ for 'nice' functions). We will see later that, locally, this is also sufficient. For an exact differential, Eq. 2.4 is

$$\int_C \underline{F} \cdot d\underline{r} = \int_C df = f(\underline{b}) - f(\underline{a}).$$

Differentials can be manipulated using

$$d(\lambda f + \mu g) = \lambda df + \mu dg$$

for constants μ and λ. There is also a Leibniz rule

$$d(fg) = (df)g + f(dg).$$

Using these, it may be possible to find f by inspection.

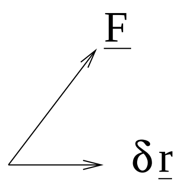
Example:

$$\int_C (3x^2 y \sin z \; dx + x^3 \sin z \; dy + x^3 y \cos z \; dz) = \int_C d(x^3 y \sin z) = \left[x^3 y \sin z \right]_{\underline{a}}^{\underline{b}} = 1$$

if $\underline{a} = (0,0,0)$ and $\underline{b} = (1,1,\pi/2)$, for instance.

2.5 Work and Potential Energy

If $\underline{F}(\underline{r})$ is a force then $\int_C \underline{F} \cdot d\underline{r}$ is the <u>work</u> the force does along the curve C. This is the limit of a sum of terms $\underline{F} \cdot d\underline{r}$, i.e. (the component of $\underline{F}(\underline{r})$ along $\delta\underline{r}$)$\times|\delta\underline{r}|$.



Consider the position of a point particle moving under $\underline{F}(\underline{r})$ according to Newton's second law: $m\ddot{\underline{r}} = \underline{F}(\underline{r})$. The kinetic energy of the particle is

$$K(t) = \frac{1}{2} m \dot{\underline{r}}^2 \Rightarrow \frac{d}{dt} K(t) = m \dot{\underline{r}} \cdot \ddot{\underline{r}} = \underline{F} \cdot \dot{\underline{r}}$$

If the path of a particle is a curve C from $\underline{a} = \underline{r}(\alpha)$ to $\underline{b} = \underline{r}(\beta)$, then

$$K(\beta) - K(\alpha) = \int_\alpha^\beta \frac{d}{dt} K(t) \; dt = \int_\alpha^\beta \underline{F} \cdot \dot{\underline{r}} \; dt = \int_C \underline{F} \cdot d\underline{r}$$

i.e. the change in the kinetic energy is equal to the work done by the force.

For a conservative force, $\underline{F} = -\nabla V$, where $V(\underline{r})$ is the potential energy (setting $f = -V$ in Eq. 2.4), $\int_C \underline{F} \cdot d\underline{r} = V(\underline{a}) - V(\underline{b})$, or the work done is equal to the loss in potential energy. So, in this case,

$$K(\beta) + V(\underline{r}(\beta)) = K(\alpha) + V(\underline{r}(\alpha)),$$

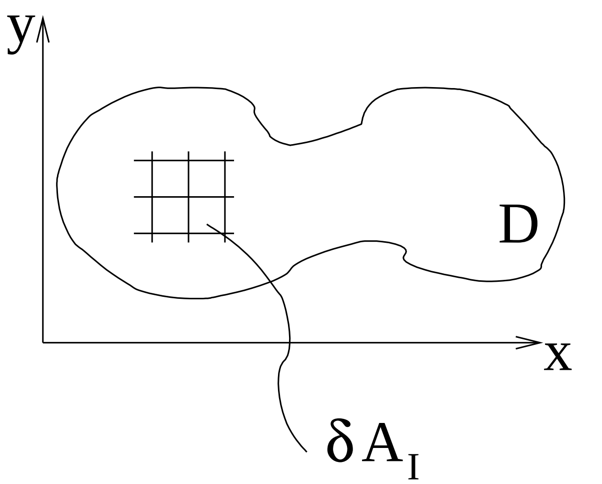
i.e. the total energy $K + V$ is conserved (in other words, it is constant during the motion).

3 Integration in $\Re^2$ and $\Re^3$

3.1 Integrals over subsets of $\Re^2$

3.1.1 Definition as the limit of a sum

Let D be a subset of $\Re^2$. We use the position vector $\underline{r} = (x, y)$ with Cartesian coordinates. Consider approximating D by N small, disjoint subsets of simple shape, e.g. triangles or parallelograms, labelled by I with area δA_I: each small enough to be contained within a disc of diameter l.

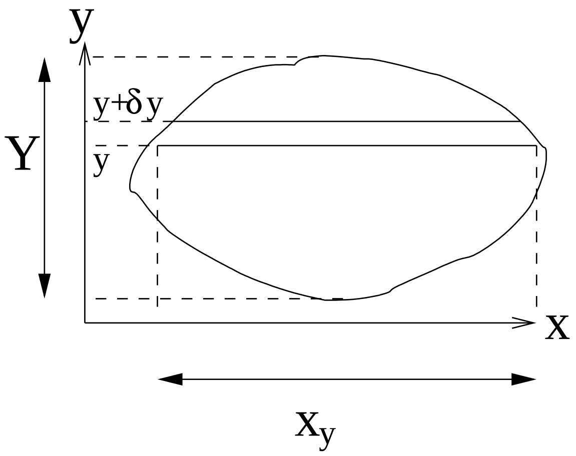


We assume that as $l \to 0$ and $N \to \infty$ the union of all of the small sets $\to D$. For a function or scalar field $f(\underline{r})$, we define $\int_D f(\underline{r}) dA = \lim_{l\to 0} \sum_I f(\underline{r}_I^*)\delta A_I$ where $\underline{r}_I^*$ is some point within each small set. We assume that the limit exists for well-behaved functions f (i.e. smooth) and D (for example the interior of a non-intersecting closed curve) independent of all of the choices we have made before taking the limit $l \to 0$ and $N \to \infty$.

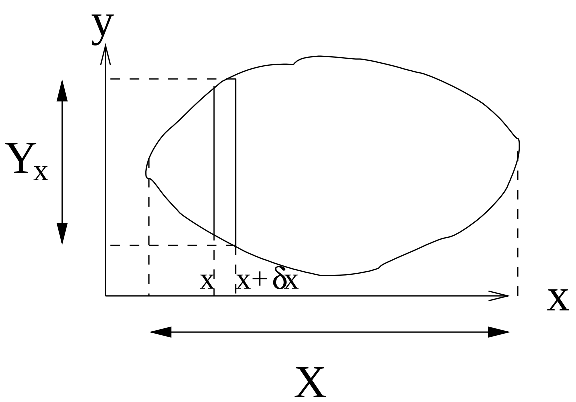
Taking $f = 1$ yields $\int_D dA$ =area of D. Integrals with $f \neq 1$ are also known as area integrals, referring to the region, or 'domain' of integration, D.

3.1.2 Evaluation as multiple integrals

Area integrals can be expressed as successive integrals over the coordinates x and y as follows: choose the small sets in the definition to be rectangles, each of size $\delta A_I = \delta x \delta y$. Summing over subsets in a narrow horizontal strip with y and δy fixed and taking $\delta x \to 0$ gives a contribution $\delta y \int_{x_y} f(x,y)\ dx$ with range $x_y = \{x : (x,y) \in D\}$. Then, summing over all such strips and taking $\delta y \to 0$ gives $\int_D f(x,y)\ dA = \int_Y \left(\int_{x_y} f(x,y)\ dx \right)\ dy$ where Y is the range of all y in D:

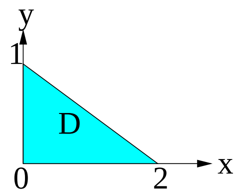


Alternatively, we can first sum over subsets in a narrow vertical strip with x and δx fixed and take $\delta y \to 0$ to get $\delta x \int_{Y_x} f(x,y)\ dy$ with range $Y_x = \{y : (x,y) \in D\}$. Then, we sum over all such strips and take $\delta x \to 0$ to get $\int_D f(x,y)\ dA = \int_X \left(\int_{Y_x} f(x,y) dy \right) dx$, where X is the range of all x in D:



We can summarise all of this by the statement that the area element is $dA = dx\ dy$ in Cartesian coordinates.

Example: $f(x,y) = x^2 y$.



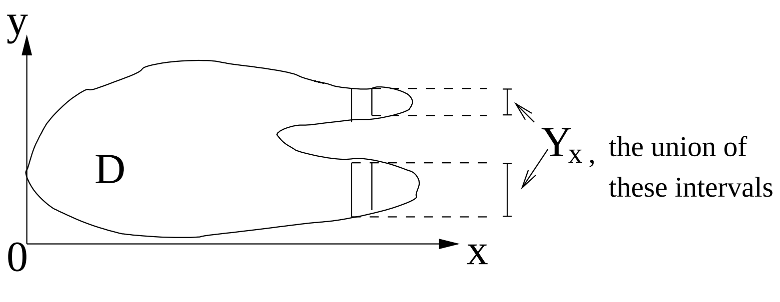
$$\int_D f(x,y)\ dA = \int_0^1 \left(\int_0^{2-2y} x^2 y\ dx \right)\ dy = \int_0^1 y \left[\frac{x^3}{3} \right]_0^{2-2y}\ dy = \frac{8}{3} \int_0^1 y(1-y)^3\ dy$$
$$= \frac{2}{15}.$$

Alternatively,

$$\int_D f(x,y)\ dA = \int_0^2 \left(\int_0^{1-x/2} x^2 y\ dy \right)\ dx = \int_0^2 x^2 \left[\frac{y^2}{2} \right]_0^{1-x/2}\ dx$$
$$= \frac{1}{2} \int_0^2 x^2 (1 - \frac{x}{2})^2\ dx = \frac{2}{15}.$$

3.1.3 Comments

1. We have adopted a notation with the range of a one dimensional integral given by a set, which may consist of disconnected intervals. For example, $X = [a_1, b_1] \cup [a_2, b_2] \Rightarrow \int_X f(x)dx = \int_{a_1}^{b_1} f(x)dx + \int_{a_2}^{b_2} f(x)dx$, and similarly for any number of intervals. The simple pictures we drew to illustrate the arguments in Section 3.1.2 involve ranges which are single intervals, but the arguments hold more generally, for example:

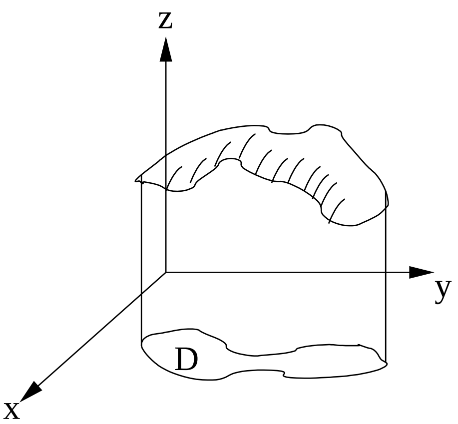


2. The result that we can integrate over x and y in either order is called Fubini's theorem. It holds if f is a continuous function and D is a compact subset of $\Re^2$, and under some more general conditions too. In some cases, f can diverge, or D can become infinite, and $\int_D f\ dA$ still exists, but these must be handled on a case-by-case basis.

   Less well behaved examples do exist where a different answer is obtained on integrating in one order or the other, just like in conditionally convergent series.

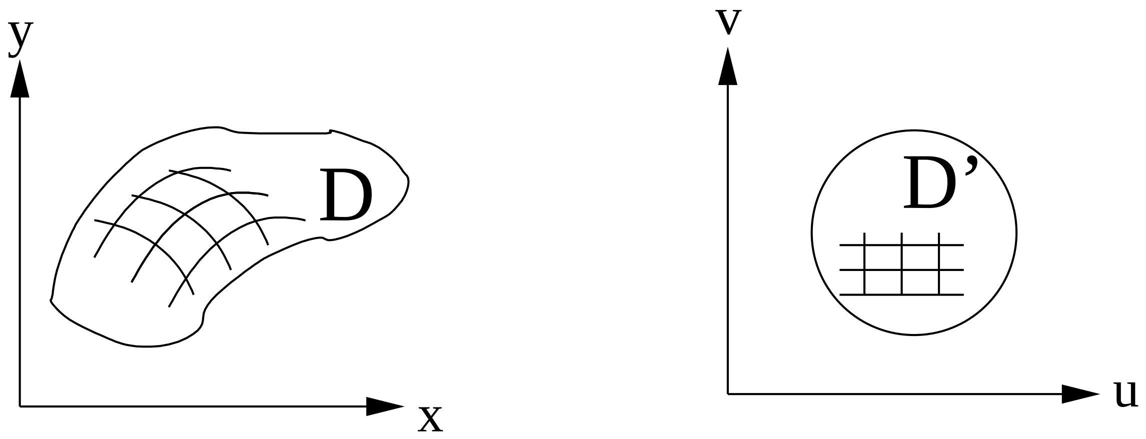
3. In the special case $f(x,y) = g(x)h(y)$ and $D = \{(x,y) :\ a \le x \le b,\ c \le y \le d\}$, $\int_D f(x,y)\ dx\ dy = \int_a^b g(x)dx \int_c^d h(y)dy$.

4. By plotting the graph of $z = f(x,y)$ in three dimensions, we get a surface and $\int_D f(x,y)\ dA$ can be interpreted as the volume beneath it:



3.2 Change of Variables for an Integral in $\Re^2$

Consider a smooth invertible transformation $(x,y) \leftrightarrow (u,v)$ with regions D and D' in one-to-one correspondence:



Then for a function (or scalar field) f, $\int_D f(x,y)\ dx\ dy = \int_{D'} f(x(u,v), y(u,v))|J|\ du\ dv$, where

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

is the <u>Jacobian</u>. In the formula, $|J|$ is the modulus of the Jacobian.

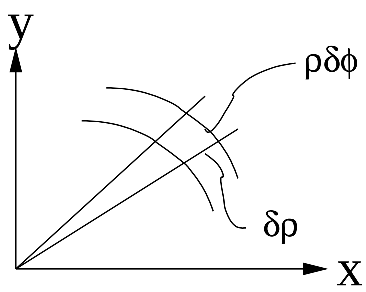
Proof in outline: write $\int_{D'}$ as the limit of the sum of rectangles of area $\delta A' = \delta u \delta v$ in the (u,v) plane. Each rectangle is mapped approximately to a parallelogram of area $A = |J|\delta u \delta v$ in the (x,y) plane, since up to $o-$terms,

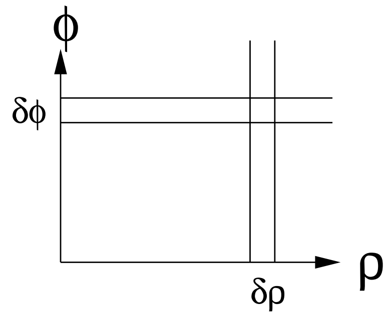
$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix},$$

and $|J|$ is the factor by which the areas are related. Summing over parallelograms in the (x, y) plane gives the integral over D as claimed, with the Jacobian factor giving the correct area for each parallelogram. The result can be summarised as

$$dA = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \; du \; dv.$$

Example: Cartesian coordinates $(x, y) \leftrightarrow$ plane polar coordinates (ρ, φ).

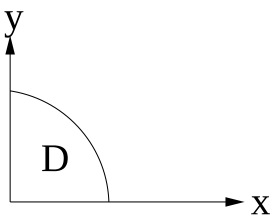


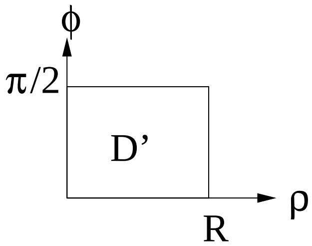


$$\begin{matrix} x = \rho\cos\varphi \\ y = \rho\sin\varphi \end{matrix} \quad \text{and } J = \frac{\partial(x,y)}{\partial(\rho,\varphi)} = \begin{vmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{vmatrix} = \rho \Rightarrow dA = \rho \; d\rho \; d\varphi.$$

In this simple case, we see that small changes $\delta\rho$ and $\delta\varphi$ produce a small area in the xy plane of approximate size $\delta A = \rho \; \delta\rho \; \delta\varphi$, confirming the Jacobian factor. Now we apply this to regions D and D' for $f = e^{-(x^2+y^2)/2} = e^{-\rho^2/2}$:

$$D : x^2 + y^2 \le R^2;\; x, y \ge 0 \qquad D' : 0 \le \rho \le R;\; 0 \le \varphi \le \pi/2.$$





$$\int f \, dA = \int_{D'} f\rho \, d\rho \, d\varphi = \int_{\rho=0}^{R}\int_{\varphi=0}^{\pi/2} e^{-\rho^2/2}\rho \, d\rho \, d\varphi = \left[-e^{-\rho^2/2}\right]_0^R [\varphi]_0^{\frac{\pi}{2}} = \frac{\pi}{2}(1-e^{-R^2/2}).$$

For this function (a Gaussian), we can take $R \to \infty$ and the integral still exists. In this limit, $0 \le x, y < \infty$ so

$$\int_D f \; dA = \int_{x=0}^{\infty}\int_{y=0}^{\infty} e^{-(x^2+y^2)/2} \; dx \; dy = \left(\int_0^{\infty} e^{-x^2/2} \; dx\right)\left(\int_0^{\infty} e^{-y^2/2} \; dy\right) = \frac{\pi}{2}.$$

Hence, we derive the well known result for the Gaussian integral $\int_0^{\infty} e^{-x^2/2} \; dx = \sqrt{\pi/2}$.

3.3 Generalisation to $\Re^3$

3.3.1 Definitions

Let a volume $V \in \Re^3$ and a position vector $\underline{r} = (x, y, z)$. We approximate V by N small disjoint subsets of simple shape (for example, cuboids) labelled by I with volume δV_I, each contained within a solid sphere of diameter l. We assume that as $l \to 0$ and $N \to \infty$, the union of the small subsets tends to V. Then, the integral of a function (or scalar field) $f(\underline{r})$ over V is defined

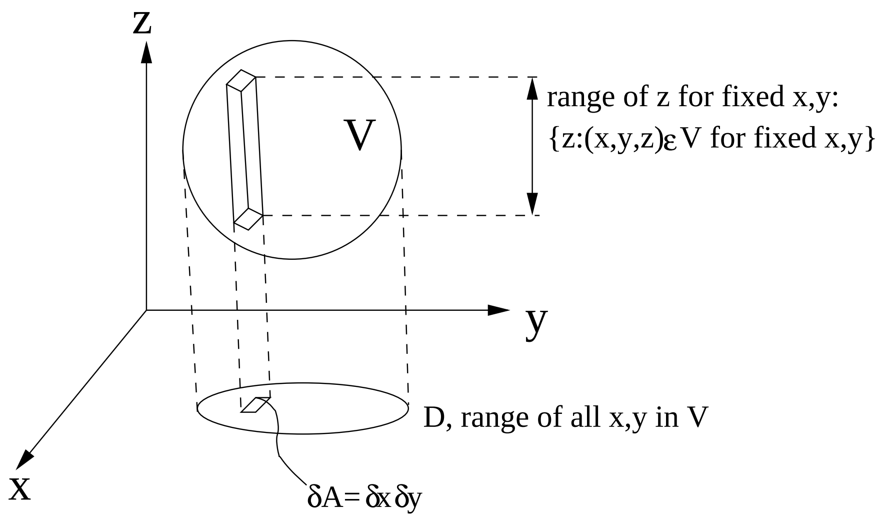
$$\int_V f(\underline{r})\ dV = \lim_{l\to 0} \sum_I f(\underline{r}_I^*)\delta V_I,$$

where $\underline{r}_I^*$ is any chosen point in each small subset.

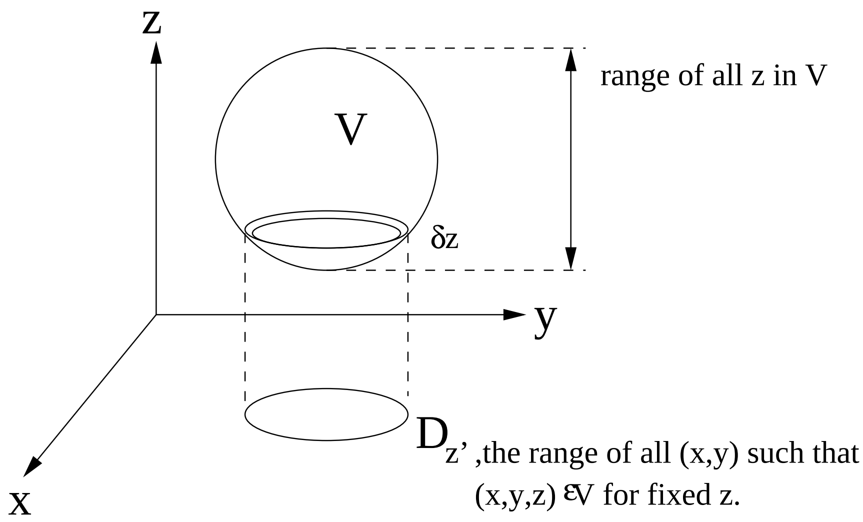
The definition can be reduced to successive integration over x, y, z by taking cuboids of volume $\delta V_I = \delta x\ \delta y\ \delta z$ and taking $\delta x \to 0$, $\delta y \to 0$, $\delta z \to 0$ in some order, for example

$$\int_V f(\underline{r})\ dV = \int_D \left(\int_{z_{xy}} f(x, y, z)\ dz \right) dx\ dy,$$

where we have taken $\delta z \to 0$ first, then $\delta x \to 0$ and $\delta y \to 0$ to get an area integral over D:



Alternatively, $\int_V f(\underline{r})\, dV = \int_z (\int_{D_z} f(x,y,z)\, dx\, dy) dz$, i.e. take δx and δy to zero first to get an area integral over slice D_z, then sum over slices and take $\delta z \to 0$:



Taking $f = 1$ gives $\int_V dV$ =volume of V. $f \neq 1$ are also known as volume integrals. Such integrals arise with $f(\underline{r})$ being the density for some quantity, for example mass, charge or probability, often denoted ρ. Then, $\rho(\underline{r})$ is the amount of the quantity in a small volume δV at $\underline{r}$. $\int_V \rho(\underline{r})\, dV$ is the total amount of the quantity in V. To summarise: the volume element in Cartesian coordinates $dV = dx\, dy\, dz$.

3.3.2 Change of Variables in $\Re^3$

Consider an invertible transformation $(x, y, z) \leftrightarrow (u, v, w)$, which is smooth both ways, with regions V and V' in $1 : 1$ correspondence. Then, for a function f, $\int_V f\, dx\, dy\, dz = \int_{V'} f|J|\, du\, dv\, dw$ where

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

is the Jacobian.

The justification (an informal proof) is similar to the 2-dimensional case: given a small cuboid with sides $\delta u, \delta v, \delta w$ respectively, we get an approximate small parallelapiped of volume $|J|\delta u \delta v \delta w$ in (x, y, z) space, and we can use these as small sets in the definition of the integral, hence the result.

Important special cases are cylindrical polar or spherical polar coordinates, as in Section 1.3. For cylindrical polar coordinates

$$x = \rho \cos\varphi,\ y = \rho \sin\varphi,\ z, \tag{3.1}$$

which means

$$\frac{\partial(x,y,z)}{\partial(\rho,\varphi,z)} = \rho \Rightarrow dV = \rho d\rho\, d\varphi\, dz.$$

On the other hand, in spherical polar coordinates,

$$x = r\sin\theta\cos\varphi,\ \ y = r\sin\theta\sin\varphi,\ \ z = r\cos\theta. \tag{3.2}$$

i.e.

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = r^2\sin\theta \Rightarrow dV = r^2\sin\theta dr\ d\theta\ d\varphi.$$

In summary, for a general coordinate system, the volume element is:

$$dV = \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right|\ du\ dv\ dw.$$

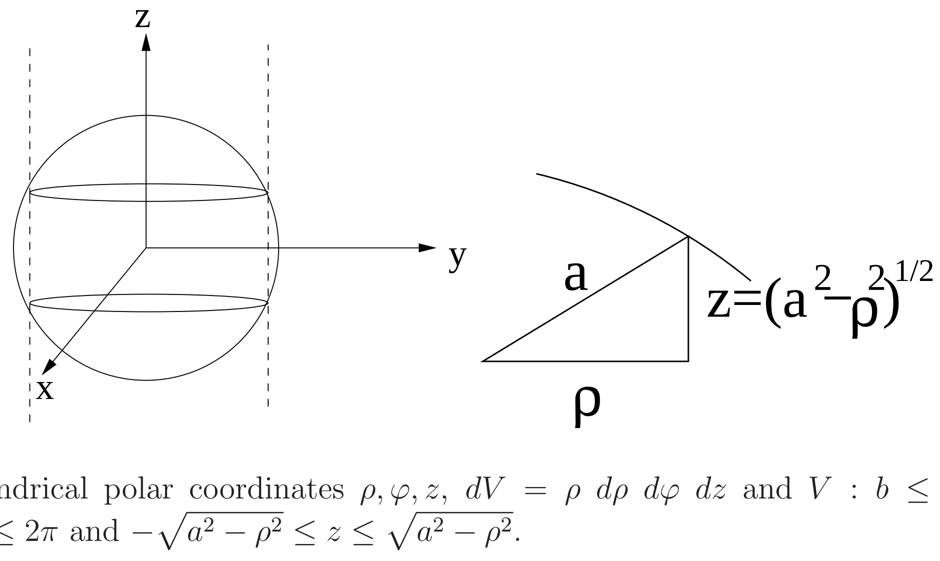
3.3.3 Examples

Example: $f(\underline{r})$ is spherically symmetric and V is a sphere of radius a.

$$\begin{aligned}\int_V f\ dV &= \int_{r=0}^{a}\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} f(r)r^2\sin\theta\ dr\ d\theta\ d\varphi = \int_0^a dr\int_0^\pi d\theta\int_0^{2\pi} d\varphi\ r^2 f(r)\sin\theta\\ &= \int_0^a dr\ r^2 f(r)[-\cos\theta]_0^\pi[\varphi]_0^{2\pi} = 4\pi\int_0^a f(r)r^2\ dr,\end{aligned}$$

a useful general result. We can understand it as the sum of spherical shells of thickness δr and volume $4\pi r^2\delta r$, taking $\delta r \to 0$. Note that taking $f = 1$ gives the volume of a sphere, $4\pi a^3/3$.

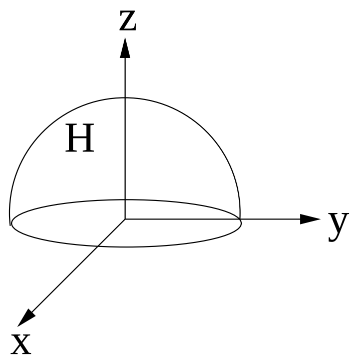
Example: Consider the volume within a sphere of radius a with a cylinder of radius b removed (where $b < a$). Thus $V : x^2 + y^2 + z^2 \leq a^2$ and $x^2 + y^2 \ \geq b^2$.



In cylindrical polar coordinates ρ, φ, z, $dV = \rho\ d\rho\ d\varphi\ dz$ and $V : b \leq \rho \leq a$, $0 \leq \varphi \leq 2\pi$ and $-\sqrt{a^2-\rho^2} \leq z \leq \sqrt{a^2-\rho^2}$.

$$\int_V dV = \int_{\rho=a}^{b}\int_{\varphi=0}^{2\pi}\int_{z=-\sqrt{a^2-\rho^2}}^{z=\sqrt{a^2-\rho^2}} \rho\ d\rho\ d\varphi\ dz = 2\pi\int_b^a d\rho\ 2\rho\sqrt{a^2-\rho^2} = \frac{4\pi}{3}(a^2-b^2)^{3/2}$$

Example: Consider the density of electric charge $\rho(\underline{r}) = \rho_0 z/a$ in a hemisphere H of radius a, with $z \geq 0$ and ρ_0 a constant. What is the total charge of H?



In spherical polar coordinates, $r \leq a$, $0 \leq \varphi \leq 2\pi$ and $0 \leq \theta \leq \pi/2$. Also, $dV = r^2 \sin\theta \; dr \; d\theta \; d\varphi$.
The total charge in H is

$$Q = \int_H \rho \; dV = \int_{r=0}^{a} \int_{\theta=0}^{\pi/2} \int_{\varphi=0}^{2\pi} \frac{\rho_0}{a} r \cos\theta r^2 \sin\theta \; dr \; d\theta \; d\varphi$$

$$= \frac{\rho_0}{a} \int_0^a r^3 \; dr \int_0^{\pi/2} \sin\theta\cos\theta \; d\theta \int_0^{2\pi} d\varphi = \frac{\rho_0}{a}\left[\frac{r^4}{4}\right]_0^a \left[\frac{1}{2}\sin^2\theta\right]_0^{\pi/2} [\varphi]_0^{2\pi} = \frac{1}{4}\pi a^3 \rho_0.$$

Checking the dimensions of our answer, we have a volume times a charge density (which is correct).

3.4 Further Generalisations and Comments

3.4.1 Integration in $\Re^n$

The definitions of integrals in $\Re^2$ and $\Re^3$ extend easily to $\Re^n$ with $n-$dimensional Euclidean length and 'volume'. The integrals are evaluated by successive integration over the Cartesian coordinates $x_1, x_2, \ldots, x_n : dV = dx_1 \; dx_2 \cdots dx_n$. For a change of variables $\{x_i\} \leftrightarrow \{u_a\}$ with the regions $D \leftrightarrow D'$ in 1:1 correspondence,

$$\int_D f(x_1, x_2, \ldots, x_n) \; dx_1 \; dx_2 \ldots dx_n = \int_{D'} f(\{x_i(\{u_a\})\})|J| \; du_1 \; du_2 \ldots du_n$$

with $J = \frac{\partial(x_1,\ldots,x_n)}{\partial(u_1,\ldots,u_n)} = \det\left(\frac{\partial x_i}{u_a}\right)$, the Jacobian. As before, $|J|$ enters as a scale factor which gives the correct $n-$dimensional Euclidean volume for small changes in the coordinates u_a relative to the product $\delta u_1 \ldots \delta u_n$.

3.4.2 Change of variables for the case $n = 1$

For a single variable $n = 1$, the Jacobian is just $J = dx/du$ and the general formula reads

$$\int_D f(x)dx = \int_{D'} f(x(u)) \left| \frac{dx}{du} \right| du.$$

The modulus of dx/du is correct because of our natural convention about integrating over subsets D and D'.

For example, suppose D is the interval $a \leq x \leq b$ so $\int_D f(x)\ dx = \int_a^b f(x)\ dx$. Let $u = \alpha, \beta$ correspond to $x = a, b$, respectively. Since $x \leftrightarrow u$ is invertible and smooth in both directions, $\frac{dx}{du} \neq 0$. Either $\alpha < \beta$ and $\frac{dx}{du} > 0$, in which case $\int_{D'} f|\frac{dx}{du}|\ du = \int_\alpha^\beta f\frac{dx}{du}\ du$ or $\alpha > \beta$ and $\frac{dx}{du} < 0$, so $\int_{D'} f|\frac{dx}{du}|\ du = \int_\beta^\alpha f(-\frac{dx}{du})\ du$. Either way, the formula yields $\int_a^b f(x)\ dx = \int_\alpha^\beta f(x(u))\frac{dx}{du}\ du$, a familiar result.

3.4.3 Vector valued integrals

Because limits apply to vectors as well as scalars, we can define for instance $\int_V \underline{F}(\underline{r})dV$ in a similar way to $\int_V f(\underline{r})dV$ as the limit of a sum over contributions from small volumes. With Cartesian components $\underline{F}(\underline{r}) = F_i(\underline{r})\underline{e}_i$, we have $\int_V \underline{F}(\underline{r})dV = \left(\int_V F_i(\underline{r})dV\right)\underline{e}_i$, i.e. we can integrate component by component.

Take the case of $\rho(\underline{r})$ being a mass density for a body occupying volume V. Then,

$$M = \int_V \rho(\underline{r})dV$$

is the total mass and

$$\underline{R} = \frac{1}{M}\int_V \underline{r}\rho(\underline{r})dV$$

is the <u>centre of mass</u>. This generalises definitions for a set of point particles at positions $\underline{r}_\alpha$ (α labels the particle, not the component here):

$$M = \sum_\alpha m_\alpha, \qquad \underline{R} = \frac{1}{M}\sum_\alpha m_a\underline{r}_\alpha.$$

Example: Consider a solid hemisphere H with $r \leq a$ and $z \geq 0$, with uniform density ρ. The total mass $M = \int_H \rho \, dV = 2\pi a^3 \rho/3$. If we put $\underline{R} = (X, Y, Z)$:

$$X = \frac{1}{M} \int_H x\rho \, dV = \frac{\rho}{M} \int_{r=0}^{a} \int_{\theta=0}^{\pi/2} \int_{\varphi=0}^{2\pi} xr^2 \sin\theta \, dr \, d\theta \, d\varphi$$
$$= \frac{\rho}{M} \int_0^a r^3 \, dr \int_0^{\pi/2} \sin^2\theta \, d\theta \cancel{\int_0^{2\pi} \cos\varphi \, d\varphi} = 0.$$

$Y = 0$ in a similar way.

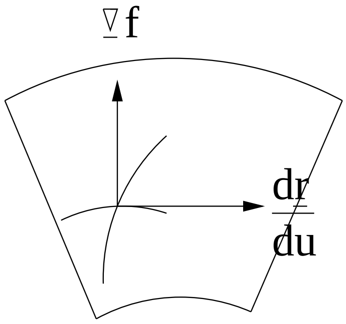
$$Z = \frac{\rho}{M} \int_0^a r^3 \, dr \int_0^{\pi/2} \sin\theta \cos\theta d\theta \int_0^{2\pi} d\varphi = \frac{\rho}{M} \frac{a^4}{4} \frac{1}{2} 2\pi = \frac{3a}{8},$$

i.e. the centre of mass is at $\underline{R} = (0, 0, 3a/8)$.

4 Surfaces and Surface Integrals

4.1 Surfaces and Normals

For an appropriate smooth function f on $\Re^3$ and a suitable constant c, the equation $f(\underline{r}) = c$ defines a smooth surface S. Consider any curve $\underline{r}(u)$ in S.

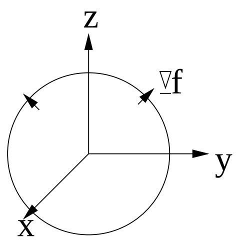


By the chain rule, $\frac{d}{du} f(\underline{r}(u)) = \underline{\nabla} f \cdot \frac{d\underline{r}}{du}$. This is zero, since the curve is in S and f does not change along it. At any $\underline{r}$ on S with $\underline{\nabla} f \neq 0$, there exist two linearly independent tangent directions $\perp$ to $\underline{\nabla} f$. Hence

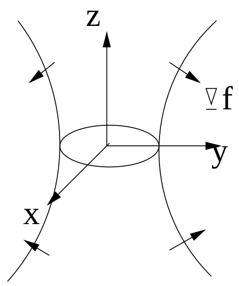
1. In the neighbourhood of this point, the equation indeed describes a surface (it has two independent tangent directions).

2. $\underline{\nabla} f$ is <u>normal</u> to the surface at this point.

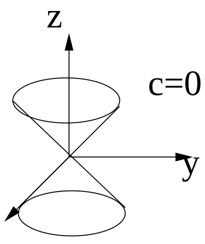
Examples: $f(\underline{r}) = x^2 + y^2 + z^2 = c$. For $c > 0$, this describes a sphere of radius $\sqrt{c}$. $\underline{\nabla} f = 2(x, y, z) = 2\underline{r}$.



Taking $f(\underline{r}) = x^2 + y^2 - z^2 = c$ results in a hyperboloid for $c > 0$. $\underline{\nabla} f = 2(x, y, -z)$.

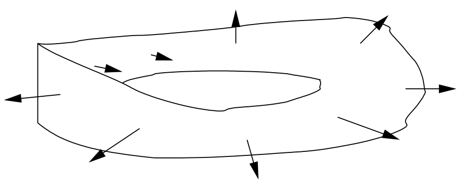


For $c = 0$, the surface degenerates to a single point $\underline{0}$ in the first case, and to a double cone in the second (it is singular at the apex $\underline{0}$, where $\underline{\nabla} f = \underline{0}$.

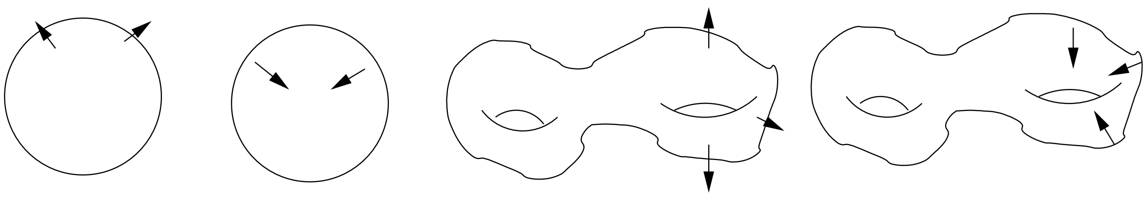


A surface S can be defined to have a boundary ∂S consisting of a piecewise smooth closed curve. Defining S as above but with $z \geq 0$, ∂S is the circle $x^2 + y^2 = c$ and $z = 0$ in the first two cases. A surface is bounded if it can be contained within some solid sphere, and unbounded otherwise. A bounded surface with no boundary is called closed.

At each point on a surface, there is a unit normal $\underline{n}$ (a normal of unit length) that is unique up to a sign. The surface is called orientable if there is a consistent choice of unit normal $\underline{n}$ which varies smoothly. We shall deal exclusively with orientable surfaces, which rules out examples like the Möbius band:



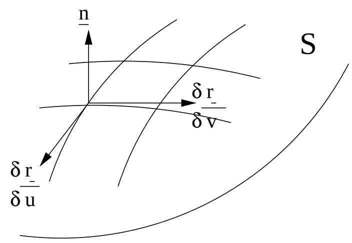
For an orientable surface, the unit normal is then determined everywhere by a choice of sign. The choce of normal is called the orientation of the surface. For simple examples, we can specify the orientation by referring to the outward or inward normal:



4.2 Parameterised Surfaces and Area

A surface S can also be defined by a position vector $\underline{r}(u,v)$ depending smoothly on the parameters u and v. S is swept out as u and v vary in some region D.

$$\delta\underline{r} = \frac{\partial \underline{r}}{\partial u}\delta u + \frac{\partial \underline{r}}{\partial v}\delta v + o\text{-terms}.$$



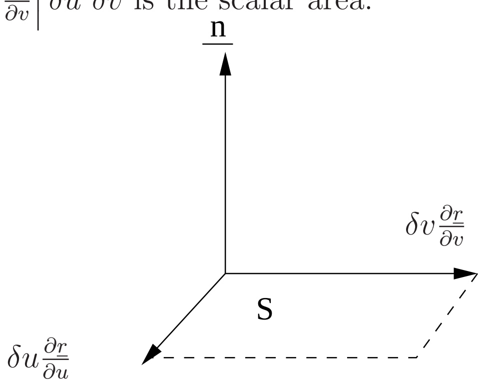
$\frac{\partial \underline{r}}{\partial u}$ and $\frac{\partial \underline{r}}{\partial v}$ are tangent vectors to curves on S with v and u constant, respectively. To define a surface, we need two linearly independent directions at each point, so we require a regular parameterisation:

$$\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} \neq 0$$

and this defines a unit normal $\underline{n}$ on S. The small changes $\delta u, \delta v$ produce a small parallelogram on S, up to $o-$terms with vector area

$$\delta \underline{S} = \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} = \delta S\ \underline{n},$$

where the $\delta S = \left|\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right| \delta u\ \delta v$ is the scalar area.



Note that with these conventions, the choice of the unit normal is determined by the order of the parameters u and v. By summing and taking limits, the area of S is

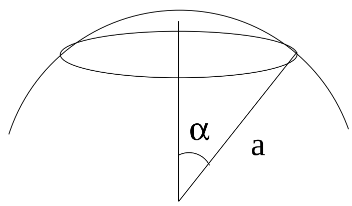
$$\int_S dS = \int_D \left|\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right| du\ dv,$$

where $dS = \left|\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right| du\ dv$ is the <u>scalar area element</u>.

Example: S being part of a sphere of radius a, parametrised by polar angles θ, φ (see section 1.3).

$$\underline{r}(\theta, \varphi) = a\left(\cos\varphi \sin\theta \underline{i} + \sin\theta \sin\varphi \underline{j} + \cos\theta \underline{k}\right) = a\underline{e}_r.$$

$\frac{\partial \underline{r}}{\partial \theta} = a\underline{e}_\theta$, $\underline{e}_\theta = \cos\theta(\cos\varphi \hat{\underline{i}} + \sin\varphi \hat{\underline{j}}) - \sin\theta \hat{\underline{k}}$ and $\frac{\partial \underline{r}}{\partial \varphi} = a \sin\theta \underline{e}_\varphi$, where $\underline{e}_\varphi = -\sin\varphi \hat{\underline{i}} + \cos\varphi \hat{\underline{j}}$. Thus $\frac{\partial \underline{r}}{\partial \theta} \wedge \frac{\partial \underline{r}}{\partial \varphi} = a^2 \sin\theta \underline{e}_r$, where $\underline{e}_r = \underline{n}$, the outward normal. The scalar area element is $dS = a^2 \sin\theta\ d\theta\ d\varphi$. Suppose that S is the region $0 \le \theta \le \alpha$, $0 \le \varphi \le 2\pi$.



Then the area is

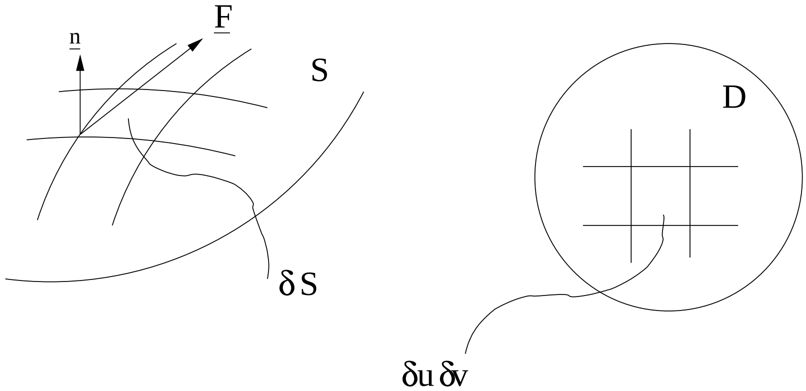
$$\int_0^{2\pi} d\varphi \int_0^{\alpha} d\theta a^2 \sin\theta = 2\pi a^2(1 - \cos\alpha).$$

Checking this result, $\alpha = \pi/2$ results in a hemisphere of area $2\pi a^2$ and $\alpha = \pi$ results in a sphere of area $4\pi a^2$ (both of which are known results).

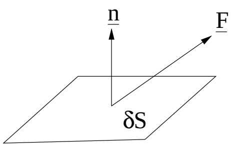
4.3 Surface Integrals of Vector Fields

Let S be a smooth surface defined by $\underline{r}(u, v)$ with S being the appropriate region in (u, v) parameter space. Again, the order of the parameters specifies an orientation for S equivalent to a choice of unit normal $\underline{n}$. The <u>surface integral</u> of a vector field $\underline{F}(\underline{r})$ over S with this orientation is defined by

$$\begin{aligned}\int_S \underline{F}(\underline{r}) \cdot d\underline{S} &= \int_S \underline{F}(\underline{r}) \cdot \underline{n}\ dS \\ &= \int_D \underline{F}(\underline{r}(u, v)) \cdot \left(\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right) du\ dv\end{aligned}$$



By writing the integral over D as the limit of a sum over small rectangles $\delta u \delta v$, the integral over S becomes the limit of a sum of contributions



$$
\begin{aligned}
\underline{F}(\underline{r}) \cdot \delta \underline{S} &= \underline{F}(\underline{r}) \cdot \underline{n}\ \delta S \\
&= \underline{F}(\underline{r(u,v)}) \cdot \left(\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} \right)\ \delta u\ \delta v + o\text{-terms.}
\end{aligned}
$$

From the geometrical nature of the left hand side, the integral $\int_S \underline{F} \cdot d\underline{S}$ is independent of the parameterisation, for a given orientation. It is called the <u>flux</u> of $\underline{F}$ through S.

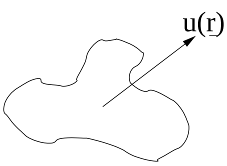
$$
d\underline{S} = \underline{n} dS = \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\ du\ dv
$$

is the <u>vector area element</u>.

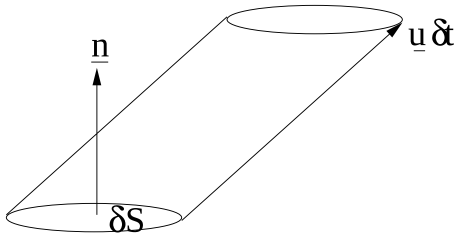
Note that changing the orientation of S is equivalent to changing the sign of the unit normal $\underline{n}$, which is equivalent to changing the order of u and v in the definition of S, which is also equivalent to changing the sign of any flux integral.

Example: For a sphere of radius a we found from $\underline{r}(\theta, \varphi)$ in section 4.2 that $\frac{\partial \underline{r}}{\partial \theta} = a \underline{e}_\theta$, $\frac{\partial \underline{r}}{\partial \varphi} = a \sin\theta \underline{e}_\varphi$. The vector area element is then $d\underline{S} = \underline{n} dS = \frac{\partial \underline{r}}{\partial \theta} \wedge \frac{\partial \underline{r}}{\partial \varphi}\ d\theta\ d\varphi = a^2 \sin\theta \underline{e}_r\ d\theta\ d\varphi$, where $\underline{n} = \underline{e}_r = \underline{r}/a$, the outward normal.

The fluid flux through a surface provides a physical example. The <u>velocity field</u> $\underline{u}(\underline{r})$ of a fluid gives the motion of any small volume of fluid at $\underline{r}$.

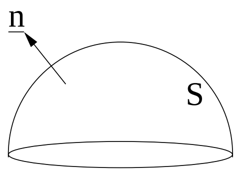


We assume that it depends smoothly upon $\underline{r}$ (note that it will also depend upon t in general). For any small area δS on the surface S, the volume of fluid crossing it in time δt, in the sense given by normal $\underline{n}$, is the volume of the cylinder shown:



$\underline{u}\delta t \cdot \underline{n}\delta S = \delta t\underline{u} \cdot \underline{\delta S}$, so the volume crossing the whole surface S in time δt is $\delta t \int_S \underline{u} \cdot \underline{dS} \Rightarrow$ the flux of $\underline{u}$, $\int_S \underline{u} \cdot \underline{dS}$, is the <u>rate</u> of volume crossing S (i.e. the amount per unit time).

Example: $\underline{u} = (-x, 0, z)$ in Cartesian coordinates, and S is the section of a sphere of radius a with $0 \leq \varphi \leq 2\pi$ and $0 \leq \theta \leq \alpha$, as in the example in section 4.2.



$\underline{dS} = a^2 \sin\theta \underline{n}\, d\theta\, d\varphi$, where $\underline{n}$ is the outward normal, $\underline{n} = \underline{r}/a = (x, y, z)/a$ in Cartesian coordinates. We have $\underline{n} \cdot \underline{u} = \underline{r} \cdot \underline{u}/a = (-x^2 + z^2)a = a(-\sin^2\theta\cos^2\varphi + \cos^2\theta)$, so

$$\begin{aligned}
\int \underline{u} \cdot \underline{dS} &= \int_0^\alpha d\theta \int_0^{2\pi} d\varphi a^3 \sin\theta \left[(\cos^2\theta - 1)\cos^2\varphi + \cos^2\theta\right] \\
&= \int_0^\alpha d\theta\; a^3 \sin\theta \left[\pi(\cos^2\theta - 1) + 2\pi\cos^2\theta\right], \text{ since } \int_0^{2\pi} \cos^2\varphi\, d\varphi = \pi \\
&= \int_0^\alpha d\theta\; a^3\pi(3\cos^2\theta - 1)\sin\theta = \pi a^3[\cos\theta - \cos^3\theta]_0^\alpha = \pi a^3 \cos\alpha \sin^2\alpha
\end{aligned}$$

is the <u>flux</u> of $\underline{u}$ through S, for a velocity field u.

Any change in parameterisation is done explicitly. Let $\underline{r}(u, v)$ and $\underline{r}(\tilde{u}, \tilde{v})$ be two regular parameterisations for the surface S. By the chain rule

$$\begin{aligned}
\frac{\partial \underline{r}}{\partial u} &= \frac{\partial \underline{r}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial u} + \frac{\partial \underline{r}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial u} \qquad \text{and} \\
\frac{\partial \underline{r}}{\partial v} &= \frac{\partial \underline{r}}{\partial \tilde{u}}\frac{\partial \tilde{u}}{\partial v} + \frac{\partial \underline{r}}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial v}.
\end{aligned} \tag{4.1}$$

These imply that

$$\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} = \frac{\partial(\tilde{u}, \tilde{v})}{\partial(u, v)} \frac{\partial \underline{r}}{\partial \tilde{u}} \wedge \frac{\partial \underline{r}}{\partial \tilde{v}}.$$

The scalar area element is therefore

$$dS = \left|\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right| du\ dv = \left|\frac{\partial \underline{r}}{\partial \tilde{u}} \wedge \frac{\partial \underline{r}}{\partial \tilde{v}}\right| d\tilde{u}\ d\tilde{v}$$

since $\left|\frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}\right| du\ dv = d\tilde{u}\ d\tilde{v}$.

The vector area element is

$$d\underline{S} = \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} du\ dv = \frac{\partial \underline{r}}{\partial \tilde{u}} \wedge \frac{\partial \underline{r}}{\partial \tilde{v}} d\tilde{u}\ d\tilde{v}$$

provided that (u, v) and $(\tilde{u}, \tilde{v})$ correspond to the same orientation on S, since then $\frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)} > 0$.

4.4 Comparing Line, Surface and Volume Integrals

4.4.1 Line and surface integrals and orientations

We have analogous definitions for integrals of scalar functions f or vector functions $\underline{F}$ along a curve C or over a surface S:

$\int_C f\ ds$ $\int_S f\ dS$	$\int_C \underline{F} \cdot d\underline{r} = \int_C \underline{F} \cdot \underline{t}\ ds$ $\int_S \underline{F} \cdot d\underline{S} = \int_S \underline{F} \cdot \underline{n}\ dD$
independent of orientation $f = 1$ gives length or area	depend on orientation from definition of integral or the choice of unit tangent $\underline{t}$ or normal $\underline{n}$.

The integrals can be regarded directly as limits of sums. To evaluate them using other parameterisations, use $ds = \left|\frac{\partial \underline{r}}{\partial u}\right| du$, $d\underline{r} = \frac{d\underline{r}}{du} du$, $dS = \left|\frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\right| du\ dv$ and $d\underline{S} = \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v}\ du\ dv$. By convention, the limits on u and v integrals are in the natural order $\int_\alpha^\beta$ with $\alpha < \beta$ except for the $d\underline{r}$ integral: we also allow $\alpha > \beta$ there if necessary. Other kinds of integrals can be considered, for example $\int_C \underline{F}\ ds$, $\int_S f\ d\underline{S}$. They can be reduced to previous cases, for instance $\underline{a} \cdot \int_C \underline{F}\ ds = \int_C (\underline{a} \cdot \underline{F})\ ds$.

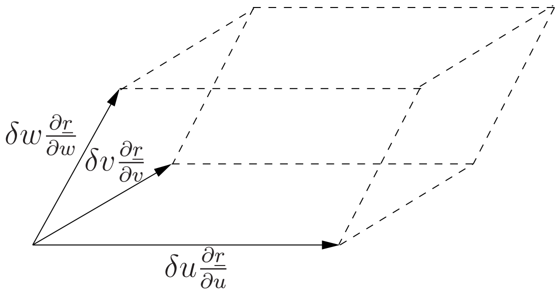
4.4.2 Change of variables in $\Re^2$ and $\Re^3$ revisited

Consider the special case of a surface S which is a subset of the plane: $\underline{r}(u, v) = (x(u, v),\ y(u, v),\ 0) \Rightarrow \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} = (0, 0, J)$ with $J = \frac{\partial(x,y)}{\partial(u,v)}$. From the definition in section 4.2, $\int_S f(\underline{r}) dS = \int_D f(\underline{r}(u, v))|J|\ du\ dv$ and we recover the formula for changing variables in $\Re^2$ found in section 3.2 (with $S \to D, D \to D'$ and $dS \to dA$).

The results for $\Re^3$ can be derived by starting from a smooth parameterisation $\underline{r}(u, v, w)$ so that

$$\delta \underline{r} = \frac{\partial \underline{r}}{\partial u}\delta u + \frac{\partial \underline{r}}{\partial v}\delta v + \frac{\partial \underline{r}}{\partial w}\delta w + o - terms.$$

The small changes $\delta u, \delta v, \delta w$ correspond to a cuboid in parameter space, producing a parallelepiped:



of volume $\delta V = \left| \frac{\partial r}{\partial u}\delta u \cdot \frac{\partial r}{\partial v}\delta v \wedge \frac{\partial r}{\partial w}\delta w \right| = |J|\delta u \ \delta v \ \delta w$. We deduce that the volume element is $dV = |J| du \ dv \ dw$ where $J = \frac{\partial r}{\partial u} \cdot \frac{\partial r}{\partial v} \wedge \frac{\partial r}{\partial w} = \frac{\partial(x,y,z)}{\partial(u,v,w)}$. This is an alternative derivation of the result in section 3.3.

5 Geometry of Curves and Surfaces

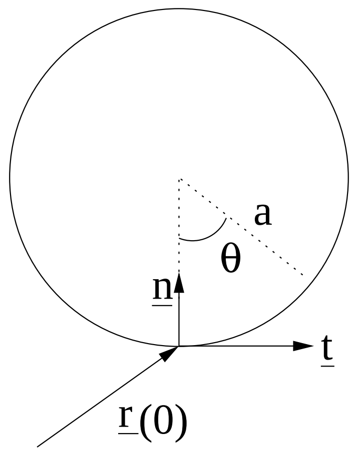
5.1 Curves, Curvature and Normals

Let $\underline{r}(s)$ be a curve parameterised by the arc length s. Since $\underline{t}(s) = \frac{dr}{ds}$ is a unit tangent vector, $\underline{t}^2 = 1 \Rightarrow \underline{t} \cdot \underline{t}' = 0$ specifies a direction normal to the curve if $\underline{t}' \neq 0$. Let $\underline{t}' = \kappa \underline{n}$ where the unit vector $\underline{n}(s)$ is called the principal normal and $\kappa(s)$ is called the curvature (we can make κ positive by choosing an appropriate direction for $\underline{n}$). These quantities can vary from point to point on the curve.

Suppose that the curve can be Taylor expanded around $s = 0$:

$$\begin{aligned} \underline{r}(s) &= \underline{r}(0) + s\underline{r}'(0) + \frac{1}{2}s^2\underline{r}''(0) + O(s^3) \\ &= \underline{r}(0) + st|_{s=0} + \frac{1}{2}s^2\kappa|_{s=0}\underline{n} + O(s^3). \end{aligned} \tag{5.1}$$

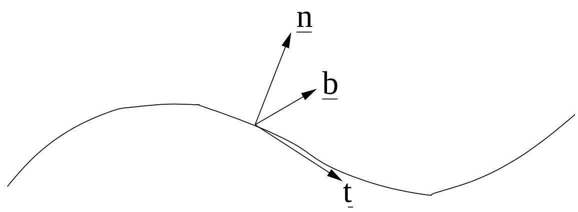
Compare this with the vector equation for a circle passing through $\underline{r}(0)$ with radius a, in the plane defined by $\underline{t}$ and $\underline{n}$, as shown:



By comparison with Eq. 5.1, we see that we can match a general curve to a circle, to second order in arc length, by taking $a = 1/\kappa$, the radius of curvature at the point $s = 0$.

In practice, for a curve $\underline{r}(u)$ given in terms of some parameter u, we can calculate $\underline{t}, s, \underline{n}, \kappa$ by taking derivatives d/du and then using the Chain Rule to convert the expression to the derivatives d/ds afterwards.

Given $\underline{t}(s)$ and $\underline{n}(s)$ we can extend to an orthonormal basis by defining $\underline{b}(s)$, the binormal to be $\underline{b} = \underline{t} \wedge \underline{n}$

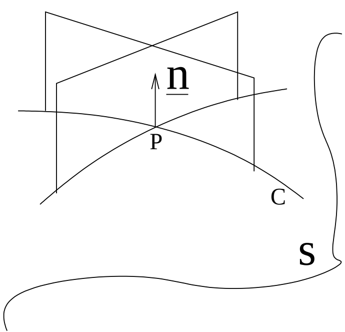


The geometry of the curve is encoded in how this basis changes along it. This can be specified by two scalar functions of arc length: the curvature $\kappa(s)$ and the torsion $\tau(s) = -\underline{b}' \cdot \underline{n}$.

[*Non-examinable and not lectured:*] The torsion determines what the curve looks like to third order in its Taylor expansion, and how the curve lifts out of the $\underline{t}, \underline{n}$ plane. Analogously to curvature, one defines a radius of torsion $\sigma = 1/\tau$. If a curve is in a plane (and the curvature is non-zero), the torsion is zero. A helix has constant torsion (and it is positive for a right-handed helix, and negative for a left-handed one).

5.2 Surfaces and Intrinsic Geometry (non-examinable)

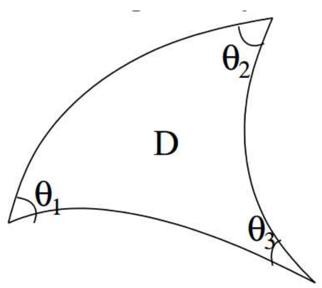
We can study the geometry of surfaces through the curves which lie on them. At a given point P on a surface S, with normal $\underline{n}$, consider the curve defined by the intersection of S with a plane containing $\underline{n}$:



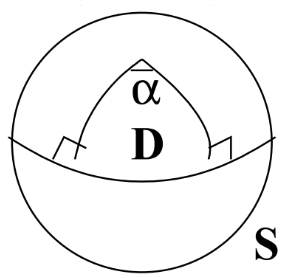
If we move the plane, we also move the curve C produced by the intersection. Let κ be the curvature of C at P, defined above. As the plane is rotated about $\underline{n}$, we find a range $\kappa_{min} \le \kappa \le \kappa_{max}$ where κ_{min} and $\kappa_{\max}$ are the principal curvatures. Then $K = \kappa_{min}\kappa_{max}$ is called the Gaussian curvature.

The *Theorema Egregium* is that K is intrinsic to the surface S, i.e. it can be expressed in terms of lengths, angles etc. which are measured entirely on the surface.

This is the start of intrinsic geometry - embedding a surface in Euclidean space determines length and angles on it, but then we study only geometric aspects of this intrinsic structure. Eventually, we do away with this embedding altogether, in Riemannian geometry and general relativity.



For an example using intrinsic geometry: consider a geodesic triangle D on a surface S. This means that ∂D consists of geodesics: the shortest curves between two points. Let $\theta_1, \theta_2, \theta_3$ be the interior angles (defined using scalar products of tangent vectors). Then $\theta_1+\theta_2+\theta_3 = \pi + \int_D K\ dS$ is the Gauss-Bonnet theorem, generalising the angle sum of a triangle to curved space.



We can check this theorem when S is a sphere of radius a, for which geodesics are great circles. It's easy to see that $\kappa_{min} = \kappa_{max} = 1/a$ so $K = 1/a^2$, a constant. We have a family of geodesic triangles D, as shown, with $\theta_1 = \alpha,\ \theta_2 = \theta_3 = \pi/2$. Since K is constant over S, $\int_D K\ dA = K \times \text{area of D} = \frac{1}{a^2}(a^2\alpha) = \alpha$. Then $\theta_1+\theta_2+\theta_3 = \pi+\alpha$, agreeing with the prediction of the theorem. By contrast, the extrinsic geometry is to do with how surfaces are embedded in the space.

6 Grad, Div and Curl

6.1 Definitions and Notation

We regard the gradient $\underline{\nabla} f$ as obtained from a scalar field $f(\underline{r})$ by applying $\underline{\nabla} = e_i \frac{\partial}{\partial x_i}$ for any Cartesian coordinates x_i and orthonormal basis $\underline{e}_i$ (see section 1.1), or

$$\underline{\nabla} = \underline{i}\frac{\partial}{\partial x} + \underline{j}\frac{\partial}{\partial y} + \underline{k}\frac{\partial}{\partial z} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}).$$

Let us use basis vectors that are orthonormal and right-handed: $\underline{e}_i \cdot \underline{e}_j = \delta_{ij}$ and $\underline{e}_i \wedge \underline{e}_j = \epsilon_{ijk}\underline{e}_k$.

$\underline{\nabla}$ (nabla, or del) is both an operator and a vector. We can therefore apply it to a vector field $\underline{F}(\underline{r}) = F_i(\underline{r})\underline{e}_i$ using either the scalar or vector product. The divergence, or div of $\underline{F}$ is:

$$\underline{\nabla} \cdot \underline{F} = (\underline{e}_i \frac{\partial}{\partial x_i}) \cdot (F_j \underline{e}_j) = \frac{\partial F_i}{\partial x_i}$$

and is a scalar field.

The curl of $\underline{F}$ is defined to be

$$\underline{\nabla} \wedge \underline{F} = (\underline{e}_i \frac{\partial}{\partial x_i}) \wedge (F_j \underline{e}_j) = \epsilon_{ijk} \frac{\partial F_j}{\partial x_i} \underline{e}_k$$

and is a vector field. It can also be written as

$$\begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{vmatrix}.$$

Example: $\underline{F} = xe^z \hat{\underline{i}} + y^2 \sin x \hat{\underline{j}} + xyz \hat{\underline{k}}$.

$$\begin{aligned} \underline{\nabla} \cdot \underline{F} &= \frac{\partial}{\partial x}(xe^z) + \frac{\partial}{\partial y}(y^2 \sin x) + \frac{\partial}{\partial z}(xyz) = e^z + 2y \sin x + xy. \\ \underline{\nabla} \wedge \underline{F} &= \left(\frac{\partial}{\partial y}(xyz) - \frac{\partial}{\partial z}(y^2 \sin x)\right) \hat{\underline{i}} + \left(\frac{\partial}{\partial z}(xe^z) - \frac{\partial}{\partial x}(xyz)\right) \hat{\underline{j}} + \\ &\quad \left(\frac{\partial}{\partial x}(y^2 \sin x) - \frac{\partial}{\partial y}(xe^z)\right) \hat{\underline{k}} \\ &= xz\hat{\underline{i}} + (xe^z - yz)\hat{\underline{j}} + y^2 \cos x \hat{\underline{k}}. \end{aligned}$$

Since $\underline{\nabla}$ is an operator, ordering is important, so $\underline{F} \cdot \underline{\nabla} = F_i \frac{\partial}{\partial x_i}$ is a scalar differential operator and $\underline{F} \wedge \underline{\nabla} = \underline{e}_k \epsilon_{ijk} F_i \frac{\partial}{\partial x_j}$ is a vector differential operator. In the example just above, $\underline{F} \cdot \underline{\nabla} = xe^z \frac{\partial}{\partial z} + y^2 \sin x \frac{\partial}{\partial y} + xyz \frac{\partial}{\partial z}$.

Grad, div and curl are linear differential operators:

$$\begin{aligned} \underline{\nabla}(\lambda f + \mu g) &= \lambda \underline{\nabla} f + \mu \underline{\nabla} g \\ \underline{\nabla} \cdot (\lambda \underline{F} + \mu \underline{G}) &= \lambda \underline{\nabla} \cdot \underline{F} + \mu \underline{\nabla} \cdot \underline{G} \\ \underline{\nabla} \wedge (\lambda \underline{F} + \mu \underline{G}) &= \lambda \underline{\nabla} \wedge \underline{F} + \mu \underline{\nabla} \wedge \underline{G} \end{aligned}$$

for any constants λ and μ. Grad and div can be analogously defined in any dimension n, but curl is specific to $n = 3$ because it uses the vector product.

Examples: Consider $\underline{r} = x_i \underline{e}_i$, $r^2 = x_i x_i \Rightarrow 2r\frac{\partial r}{\partial x_i} = 2x_i$, or $\frac{\partial r}{\partial x_i} = x_i/r$.

1. $\underline{\nabla} r^\alpha = \underline{e}_i \frac{\partial}{\partial x_i}(r^\alpha) = \underline{e}_i \alpha r^{\alpha-1}\frac{\partial r}{\partial x_i} = \underline{e}_i \alpha r^{\alpha-1} x_i/r = \alpha r^{\alpha-2}\underline{r} = \alpha r^{\alpha-1}\underline{\hat{r}}$
2. $\underline{\nabla}\cdot\underline{r} = \frac{\partial x_i}{\partial x_i} = \delta_{ii} = 3.$
3. $\underline{\nabla}\wedge\underline{r} = \epsilon_{ijk}\frac{\partial}{\partial x_i}(x_j)\underline{e}_k = \underline{0}.$

6.2 Leibniz Properties

Let us take scalar fields f and g and vector fields $\underline{F}, \underline{G}$. Their Leibniz properties can be proved using index notation. The properties

$$\begin{aligned}
\underline{\nabla}(fg) &= (\underline{\nabla} f)g + f(\underline{\nabla} g), \\
\underline{\nabla}\cdot(f\underline{F}) &= (\underline{\nabla} f)\cdot\underline{F} + f(\underline{\nabla}\cdot\underline{F}) \text{ and} \\
\underline{\nabla}\wedge(f\underline{F}) &= (\underline{\nabla} f)\wedge\underline{F} + f(\underline{\nabla}\wedge\underline{F})
\end{aligned}$$

are frequently useful. There are also some less commonly used properties too:

$$\begin{aligned}
\underline{\nabla}(\underline{F}\cdot\underline{G}) &= \underline{F}\wedge(\underline{\nabla}\wedge\underline{G}) + \underline{G}\wedge(\underline{\nabla}\wedge\underline{F}) + (\underline{F}\cdot\underline{\nabla})\underline{G} + (\underline{G}\cdot\underline{\nabla})\underline{F} \qquad (6.1) \\
\underline{\nabla}\wedge(\underline{F}\wedge\underline{G}) &= \underline{F}(\underline{\nabla}\cdot\underline{G}) - \underline{G}(\underline{\nabla}\cdot\underline{F}) + (\underline{G}\cdot\underline{\nabla})\underline{F} - (\underline{F}\cdot\underline{\nabla})\underline{G} \\
\underline{\nabla}\cdot(\underline{F}\wedge\underline{G}) &= (\underline{\nabla}\wedge\underline{F})\cdot\underline{G} - \underline{F}\cdot(\underline{\nabla}\wedge\underline{G}).
\end{aligned}$$

Examples:

1. $\underline{\nabla}\cdot(\underline{F}\wedge\underline{G}) = \frac{\partial}{\partial x_i}(\underline{F}\wedge\underline{G})_i = \frac{\partial}{\partial x_i}(\epsilon_{ijk}F_j G_k) = \epsilon_{kij}\frac{\partial F_j}{\partial x_i}G_k - F_j\epsilon_{jik}\frac{\partial G_k}{\partial x_i}.$
2. $[\underline{F}\wedge(\underline{\nabla}\wedge\underline{G})]_i = \epsilon_{ijk}F_j(\underline{\nabla}\wedge G)_i = \epsilon_{kij}\epsilon_{kpq}F_j\frac{\partial G_q}{\partial x_p} = (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})F_j\frac{\partial G_q}{\partial x_p} = F_j\frac{\partial G_j}{\partial x_i} - F_j\frac{\partial G_i}{\partial x_j}$. Hence

$$[\underline{F}\wedge(\underline{\nabla}\wedge\underline{G})]_i = F_j\frac{\partial G_j}{\partial x_i} - (\underline{F}\cdot\underline{\nabla})\underline{G_i} \qquad (6.2)$$

Taking Eq. 6.2 (along with the $\underline{F} \leftrightarrow \underline{G}$ counterpart and substituting into the right hand side of Eq. 6.1 yields its left hand side, i.e. the formula for

$$[\underline{\nabla}(\underline{F}\cdot\underline{G})]_i = \frac{\partial F_j}{\partial x_i}G_j + F_j\frac{\partial G_j}{\partial x_i}.$$

Using the Leibniz properties, we can extend our results to some previous examples:

Examples:

1. $\underline{\nabla}\cdot(r^\alpha\underline{r}) = \underline{\nabla}(r^\alpha)\cdot\underline{r} + r^\alpha\underline{\nabla}\cdot\underline{r} = (\alpha r^{\alpha-2}\underline{r})\cdot\underline{r} + r^\alpha(3) = (\alpha+3)r^\alpha.$
2. $\underline{\nabla}\wedge(r^\alpha\underline{r}) = \underline{\nabla}(r^\alpha)\wedge\underline{r} + r^\alpha\cancel{\underline{\nabla}\wedge\underline{r}} = \alpha r^{\alpha-2}\cancel{\underline{r}\wedge\underline{r}} = \underline{0}.$

6.3 Second Order Derivatives

If $\underline{F} = \underline{\nabla}$, the field is called conservative. If $\underline{\nabla} \wedge \underline{F} = \underline{0}$, a field is called or irrotational. Two different combinations of grad, div and curl vanish identically: for any scalar field f, $\underline{\nabla} \wedge (\underline{\nabla} f) = \underline{0}$ and for any vector field $\underline{A}$, $\underline{\nabla} \cdot (\underline{\nabla} \wedge \underline{A}) = 0$. Both of these results follow from the symmetry of mixed partial derivatives:

$$\epsilon_{kij} \frac{\partial^2 f}{\partial x_i \partial x_j} = 0, \qquad \epsilon_{kij} \frac{\partial^2 A_k}{\partial x_i \partial x_j} = 0.$$

The converse of each result also holds for fields defined on all of $\Re^3$:

$$\underline{\nabla} \wedge \underline{F} = \underline{0} \Leftrightarrow \underline{F} = \underline{\nabla} f$$

for some f, a scalar potential. Also, again for $\underline{H}$ defined on[1] $\Re^3$,

$$\underline{\nabla} \cdot \underline{H} = 0 \Leftrightarrow \underline{H} = \underline{\nabla} \wedge \underline{A},$$

for some vector potential $\underline{A}$. In this case, $\underline{H}$ is called solenoidal.

If $\underline{F}$ or $\underline{H}$ are defined only on subsets of $\Re^3$, then f or $\underline{A}$ may be defined only on smaller subsets (see later).

The Laplacian operator is defined by

$$\nabla^2 = \underline{\nabla} \cdot \underline{\nabla} = \frac{\partial^2}{\partial x_i \partial x_i} \text{ or } (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}).$$

On a scalar field, $\nabla^2 f = \underline{\nabla} \cdot (\underline{\nabla} f)$ whereas on a vector field,

$$\nabla^2 \underline{A} = \underline{\nabla}(\underline{\nabla} \cdot \underline{A}) - \underline{\nabla} \wedge (\underline{\nabla} \wedge \underline{A}).$$

Again, this may be checked by using components.

7 Integral Theorems

The following closely-related results generalise the Fundamental Theorem of calculus: that integration is the inverse of differentiation. To generalise to higher dimensions, we need the right kind of derivative to match the right kind of integral.

7.1 Statements and Examples

7.1.1 Green's theorem (in the plane)

For smooth functions $P(x, y)$ and $Q(x, y)$:

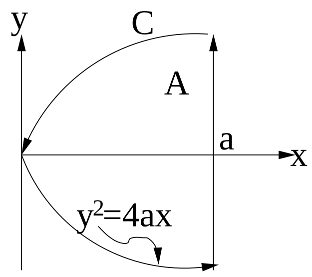
$$\int_A (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\, dA = \int_C (P\, dx + Q\, dy),$$

[1]We shall see in section 8.2 how this can change for a field which is not defined on all of $\Re^3$.

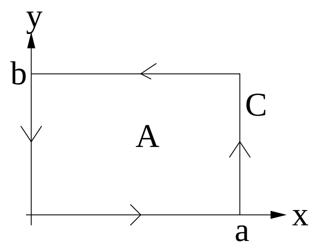
where A is a bounded region in the (x, y) plane with boundary $\partial A = C$, a piecewise smooth, non-intersecting closed curve, traversed anti-clockwise.

Examples:

1. $P = x^2 y$ and $Q = xy^2 \Rightarrow \int_A (y^2 - x^2)\ dA = \int_C x^2 y\ dx + xy^2\ dy$. If C is the parabola $y^2 = 4ax$ and the line $x = a$, both with $2a \geq y \geq -2a$, then each integral above gives $\frac{104}{105} a^4$ (see Examples Sheet I).

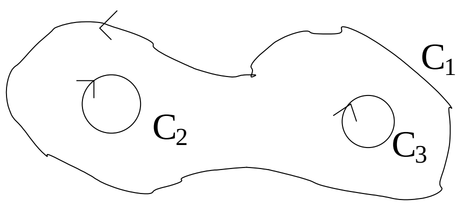


2. If A is a rectangle with $0 \leq x \leq a, 0 \leq y \leq b$, then Green's theorem reduces immediately to the Fundamental Theorem of Calculus for integrals in one dimension:
$\int_A -\frac{\partial P}{\partial y}\ dA = \int_0^a dx \int_0^b dy (-\frac{\partial P}{\partial y}) = \int_0^a dx[-P(x, b) + P(x, 0)] = \int_C P\ dx$, since only the horizontal segments contribute.

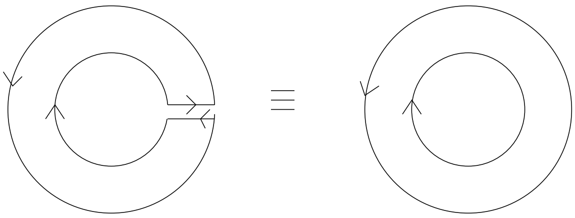


Also, $\int_A \frac{\partial Q}{\partial x}\ dA = \int_0^b dy \int_0^a dx (\frac{\partial Q}{\partial x}) = \int_0^b [Q(a, x) - Q(0, y)] = \int_C Q dy$, since only the vertical segments contribute. Adding these two integrals gives the result in the theorem.

Green's theorem also holds for a bounded region A with the boundary ∂A consisting of a number of disconnected components (each being a piecewise smooth, non-intersecting closed curve) with anti-clockwise orientations on the exterior boundary and clockwise orientations on the interior boundary. For example,



These can be related to the case of a <u>single</u> boundary component using the construction from section 2.3:



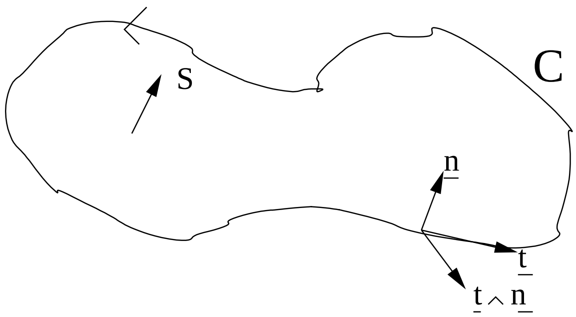
since the horizontal parts *cancel* in the opposite directions (in the limit that they are brought close together).

7.1.2 Stokes' theorem

For a smooth vector field $\underline{F}(\underline{r})$:

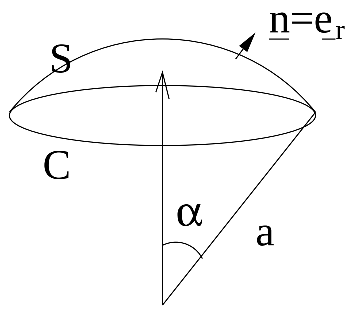
$$\int_S \underline{\nabla} \wedge \underline{F} \cdot d\underline{S} = \int_C \underline{F} \cdot d\underline{r},$$

where S is a bounded smooth surface with boundary $\partial S = C$, a piecewise smooth curve, and where S and C have compatible orientations. In detail: $\underline{n}$ is normal to S ($dS = \underline{n}\ dS$), $\underline{t}$ is tangent to C ($d\underline{r} = \underline{t}\ ds$). $\underline{t}$ and $\underline{n}$ have compatible orientations if $\underline{t} \wedge \underline{n}$ points out of S. Alternatively, moving around C in direction $\underline{t}$ with $\underline{n}$ 'up' $\Rightarrow$ that S is on the <u>left</u>.



The theorem also holds if ∂S consists of a number of disconnected piecewise smooth closed curves, with the same rule determining the orientation on each component.

Example: Take S to be the section of a sphere of radius a: $0 \le \theta \le \alpha$:

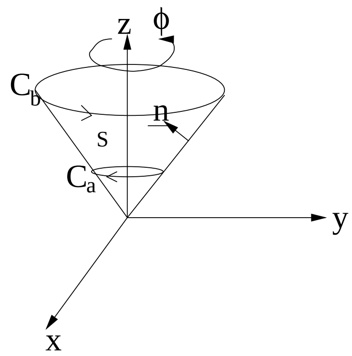


$d\underline{S} = a^2 \sin\theta \underline{e}_r \, d\theta \, d\varphi$, $\underline{F} = (0, xz, 0) \Rightarrow \underline{\nabla} \wedge \underline{F} = (-x, 0, z)$, as in section 4.3 with $u = \underline{\nabla} \wedge \underline{F}$. Thus $\int_S \underline{\nabla} \wedge \underline{F} \cdot d\underline{S} = \pi a^3 \cos\alpha \sin\alpha$. The boundary is $\partial S = C$: $\underline{r}(\varphi) = (a \sin\alpha \cos\varphi, \ a \sin\alpha \sin\varphi, \ a\cos\alpha)$, $0 \le \varphi \le 2\pi$.

$$\int_C \underline{F} \cdot d\underline{r} = \int_0^{2\pi} \underbrace{a \sin\alpha \cos\varphi}_{x} \underbrace{a \cos\alpha}_{z} \underbrace{a \sin\alpha \cos\varphi \; d\varphi}_{dy}$$

$$= a^3 \sin^2\alpha \cos\alpha \int_0^{2\pi} \cos^2\varphi \; d\varphi = \pi a^3 \sin^2\alpha \cos\alpha.$$

Example: Take S to be the section of the cone $z^2 = x^2 + y^2$, $a \le z \le b$ (where $b > a > 0$). In cylindrical polar coordinates $\underline{r}(\rho, \varphi) = (\rho \cos\varphi, \rho \sin\varphi, \rho)$, $a \le \rho \le b$, $0 \le \varphi \le 2\pi$.

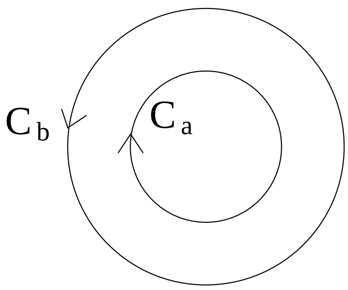


$$\frac{\partial \underline{r}}{\partial \rho} \wedge \frac{\partial \underline{r}}{\partial \varphi} = (\cos\varphi, \sin\varphi, 1) \wedge (-\rho\sin\varphi, \rho\cos\varphi, 0) = (-\rho\cos\varphi, -\rho\sin\varphi, \rho)$$
$$\Rightarrow d\underline{S} = (-\rho\cos\varphi, -\rho\sin\varphi, \rho)\ d\rho\ d\varphi,$$

which gives the normal $\underline{n}$ as shown on the diagram. Taking $\underline{F} = (0, xz, 0) \Rightarrow \underline{\nabla} \wedge \underline{F} = (-x, 0, z)$.

$$\underline{\nabla} \wedge \underline{F} \cdot d\underline{S} = (\rho^2 \cos^2\varphi + \rho^2)\ d\rho\ d\varphi$$
$$\Rightarrow \int_S \underline{\nabla} \wedge \underline{F} \cdot d\underline{S} = \int_a^b d\rho \int_0^{2\pi} d\varphi \rho^2 (3/2 + \cancel{\cos 2\varphi}/2) = \pi(b^3 - a^3).$$

The boundary $\partial S = C_b - C_a$, where the circles have the orientations as shown. For a circle of radius R, C_R : $\underline{r} = (R\cos\varphi, R\sin\varphi, R)$ $0 \le \varphi \le 2\pi \Rightarrow \underline{r}'(\varphi) = (-R\sin\varphi, R\cos\varphi, 0)$. Then $\int_{C_R} \underline{F} \cdot d\underline{r} = \int_0^{2\pi} R^3 \cos^2\varphi\ d\varphi = \pi R^3$, hence $\int_{\partial S} \underline{F} \cdot d\underline{r} = \int_{C_b} \underline{F} \cdot d\underline{r} - \int_{C_a} \underline{F} \cdot d\underline{r} = \pi(b^3 - a^3)$. For the orientations, we look down the $z-$axis and $\underline{n}$ points upwards.



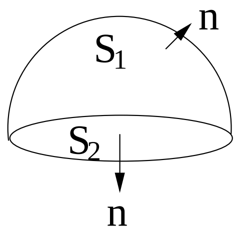
7.1.3 Divergence, or Gauss' theorem

For a smooth vector field $\underline{F}(\underline{r})$,

$$\int_C \underline{\nabla} \cdot \underline{F}\ dV = \int_S \underline{F} \cdot d\underline{S}$$

where V is a bounded volume with boundary $\partial V = S$, a piecewise smooth closed surface with normal $\underline{n}$ pointing outwards.

Example: Take V to be the solid hemisphere $x^2+y^2+z^2 \le a^2,\ z \ge 0$. $\partial V = S = S_1(\text{hemisphere}) + S_2(\text{disc})$.



$\underline{F} = (0,0,z+a) \Rightarrow \underline{\nabla} \cdot \underline{F} = 1$. Then, $\int_V \underline{\nabla} \cdot \underline{F}\ dV = 2/3\pi a^3$, the volume of the hemisphere.
On S_1: $d\underline{S} = \underline{n}\ dS = (x,y,z)/a\ dS$.

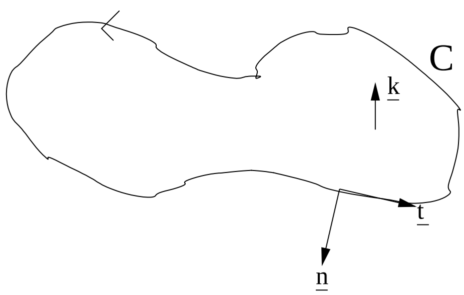
$$\underline{F} \cdot d\underline{S} = z(z+a)/a\ dS = a\cos\theta(\cos\theta+1)a^2\sin\theta\ d\theta\ d\varphi.$$
$$\Rightarrow \int_{S_1} \underline{F} \cdot d\underline{S} = \int_0^{2\pi} d\varphi a^3 \int_0^{\pi/2} d\theta \sin\theta(\cos^2\theta + \cos\theta)$$
$$= 2\pi a^3 \left[-\frac{1}{3}\cos^3\theta - \frac{1}{2}\cos^2\theta\right]_0^{\pi/2} = \frac{5}{3}\pi a^3.$$

On S_2 : $d\underline{S} = \underline{n}\ dS = -(0,0,1)\ dS$. $\underline{F} \cdot d\underline{S} = -a\ dS$ and so $\int_{S_2} \underline{F} \cdot d\underline{S} = -\pi a^3$. Hence $\int_{S_1} \underline{F} \cdot d\underline{S} + \int_{S_2} \underline{F} \cdot d\underline{S} = 5/3\pi a^3 - \pi a^3 = 2/3\pi a^3$, agreeing with the prediction of the theorem.

7.2 Relating and Proving the Integral Theorems

7.2.1 Proving Green's theorem from Stokes' theorem or the 2d divergence theorem

Let A be a region in the (x,y) plane with boundary $C = \partial A$ parameterised by arc length, with the sense shown: $(x(s), y(s), 0)$:



$\underline{t} = (dx/ds, dy/ds, 0)$ is the unit tangent to C and $\underline{n} = (dy/ds, -dx/ds, 0)$ is the unit normal to C out of A. $\underline{n} = \underline{t} \wedge \underline{k}$, where $\underline{k} = (0,0,1)$.

Given any $P(x,y)$ and $Q(x,y)$, consider the vector field $\underline{F} = (P, Q, 0) \Rightarrow \underline{\nabla} \wedge \underline{F} = (0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})$. Then,

$$\int_C \underline{F} \cdot \underline{t}\ ds = \int_C P\ dx + Q\ dy \tag{7.1}$$

and

$$\int_A (\underline{\nabla} \cdot \underline{G}\ dA = \int_A (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dA. \tag{7.2}$$

Thus an application of Stokes' theorem for $\underline{F}$ equates the two left-hand sides of Eqs. 7.1, 7.2 whereas Greens theorem equates the two right-hand sides. Therefore,

$$\boxed{\textbf{Stokes' theorem} \Rightarrow \textbf{Green's theorem}}$$

Now consider $\underline{G} = (Q, -P, 0) \Rightarrow \underline{\nabla} \cdot \underline{G} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$. We have

$$\int_C \underline{G} \cdot \underline{n}\ ds = \int_C P\ dx + Q\ dy \tag{7.3}$$

and

$$\int_A (\underline{\nabla} \wedge \underline{G}) \cdot \underline{k}\ dA = \int_A (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\ dA. \tag{7.4}$$

Again, Green's theorem equates the right hand sides of Eqs. 7.3,7.4 whereas the two-dimensional version of the divergence theorem for a vector field $\underline{G}$ in the (x, y) plane

$$\int_A \underline{\nabla} \cdot \underline{G}\ dA = \int_{C=\partial A} \underline{G} \cdot \underline{n} ds \tag{7.5}$$

relates the left-hand sides:

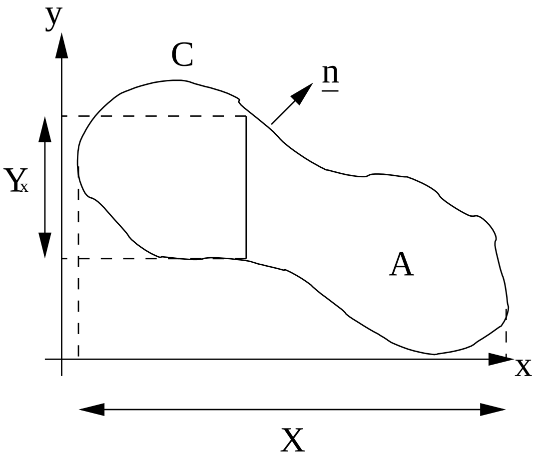
$$\boxed{\textbf{Green's theorem} \Leftrightarrow \textbf{2d divergence theorem}}$$

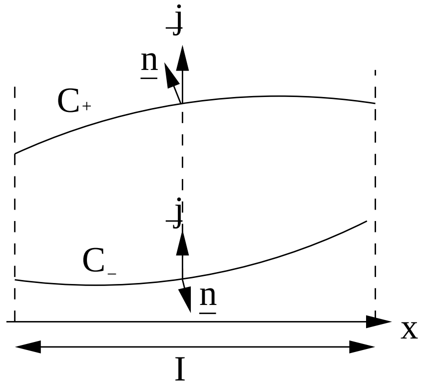
7.2.2 Proving Green's theorem by Proving the 2d Divergence Theorem

We seek to prove the 2d divergence theorem $\int_A \underline{\nabla} \cdot \underline{G}\ dA = \int_{C=\partial A} \underline{G} \cdot \underline{n} ds$, where $\underline{n}$ is the outward normal. It suffices to consider $\underline{G} = G(x,y)\underline{j}$, because a field $\underline{H}$ in the direction $\underline{i}$ can be treated similarly, exchanging x and y, and then we obtain a general vector field from a linear combination of the two. Then,

$$\int_A \underline{\nabla} \cdot \underline{G}\ dA = \int_X \left(\int_{Y_x} \frac{\partial G}{\partial y}\ dy \right)\ dx.$$

We have chosen this order of integration because $\underline{\nabla} \cdot \underline{G} = \frac{\partial G}{\partial y}$. Depending on x, the range Y_x may be a single interval or a union of intervals:

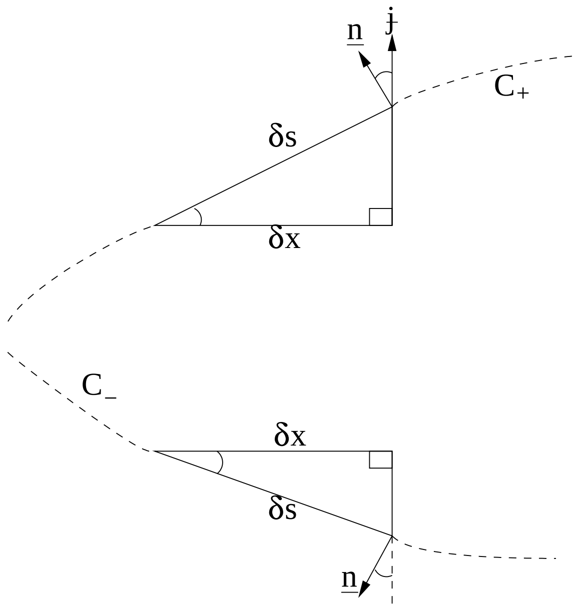




Consider x in some interval I such that Y_x is a single interval $y_+(x) \geq y \geq y_-(x)$, corresponding to the boundary segments $C_\pm$ as we have drawn above.

$$\int_{Y_x} \frac{\partial G}{\partial y}\, dy = \int_{y_-(x)}^{y_+(x)} \frac{\partial G}{\partial y}\, dy = G(x, y_+(x)) - G(x, y_-(x))$$

for any x in I.



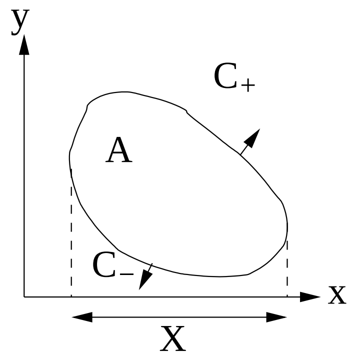
On C_+: $\underline{n} \cdot \underline{j} \geq 0$ and $dx = \underline{n} \cdot \underline{j} ds$

On C_-: $\underline{n} \cdot \underline{j} \leq 0$ and $dx = -\underline{n} \cdot \underline{j} ds$

Hence

$$\begin{aligned} \int_I \left(\frac{\partial G}{\partial y}\, dy \right) dx &= \int_I \left(G(x, y_+(x)) - G(x, y_-(x)) \right) dx \\ &= \int_{C+} \underline{G} \cdot \underline{n}\, ds + \int_{C_-} \underline{G} \cdot \underline{n}\, ds. \end{aligned}$$

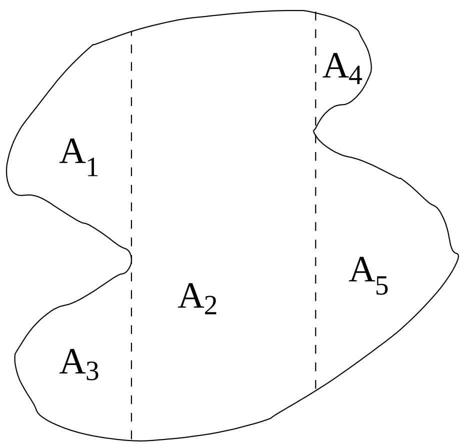
For a region A of 'simple shape'



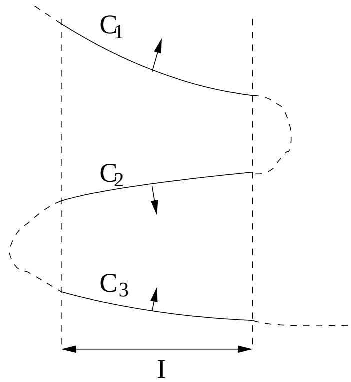
as shown, we take $I = X$ and $\partial A = C_+ + C_-$ and the integrals over $C_\pm$ combine to give $\int_C \underline{G} \cdot \underline{n}\; ds$, proving the result.

For regions A of more general shape, either:

1. Divide A into simpler subsets which do have the property above. The result holds on each subset $A_i \Rightarrow$ it holds on A



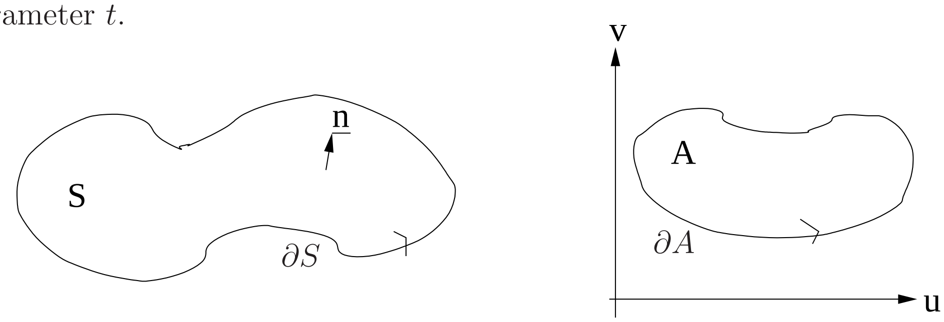
2. Proceed directly, even if Y_x is a sum of intervals and get a sum over <u>all</u> boundary segments corresponding to x in the range I



$$\int_I \left(\int_{Y_x} \frac{\partial G}{\partial y}\, dy \right)\, dx = \sum_i \int_{C_i} \underline{G} \cdot \underline{n}\; ds.$$

7.2.3 Green's theorem ⇒ Stokes' theorem

Consider a parameterised surface given by $\underline{r}(u,v)$, corresponding to a region A in the u,v plane. Suppose the boundaries are given by $\partial A : u(t), v(t)$ and $\partial S : \underline{r}(u(t), v(t))$ for parameter t.



For a smooth vector field $\underline{F}(\underline{r})$, we define

$$F_u = \underline{F} \cdot \frac{\partial \underline{r}}{\partial u}, \qquad F_v = \underline{F} \cdot \frac{\partial \underline{r}}{\partial v}.$$

Then,

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = (\underline{\nabla} \wedge \underline{F}) \cdot (\frac{\partial \underline{r}}{\partial u} - \frac{\partial \underline{r}}{\partial v}). \tag{7.6}$$

To show this, we note that

$$\frac{\partial F_v}{\partial u} = \frac{\partial}{\partial u}\left(F_i \frac{\partial x_i}{\partial v}\right) = \frac{\partial F_i}{\partial x_j}\frac{\partial x_k}{\partial u}\frac{\partial x_i}{\partial v} + F_i \frac{\partial^2 x_i}{\partial u \partial v}.$$

Thus, we obtain that

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = \frac{\partial x_j}{\partial u}\frac{\partial x_i}{\partial v}\left(\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}\right). \tag{7.7}$$

The right-hand side of Eq. 7.6 is

$$(\underline{\nabla} \wedge \underline{F}) \cdot (\frac{\partial \underline{r}}{\partial u} - \frac{\partial \underline{r}}{\partial v}) = \epsilon_{kij}\frac{\partial F_j}{\partial x_i}\epsilon_{kpq}\frac{\partial x_p}{\partial u}\frac{\partial x_q}{\partial v} = \frac{\partial x_j}{\partial u}\frac{\partial x_i}{\partial v}\left(\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}\right), \tag{7.8}$$

by using an $\epsilon\epsilon$ identity. We see that the right-hand sides of Eq. 7.8 and 7.7 are equal, therefore we have shown Eq. 7.6.

We now integrate each side of Eq. 7.6 with respect to $du\, dv$:

$$\int_A \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) dA = \int_s \underline{\nabla} \wedge \underline{F} \cdot d\underline{S}. \tag{7.9}$$

By the Chain Rule,

$$F_u \frac{du}{dt} + F_v \frac{dv}{dt} = \underline{F} \cdot \frac{d\underline{r}}{dt}$$

on the boundary. Integrating with respect to t, we obtain:

$$\int_A F_u du + F_v dv = \int_{\partial S} \underline{F} \cdot d\underline{r}. \tag{7.10}$$

Green's theorem equates the left-hand sides of Eqs. 7.9,7.10 whereas Stokes' equates their right-hand sides.

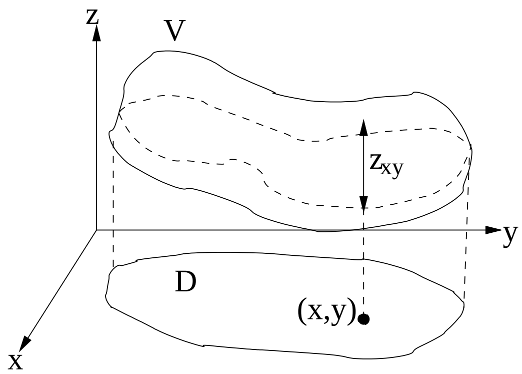
The conclusion is that the left hand sides being equal (i.e. Green's theorem for F_i, F_v on $A \Leftrightarrow$ Stokes' theorem for $\underline{F}$ on S).

7.2.4 Proving the divergence (or Gauss') theorem in 3d: outline

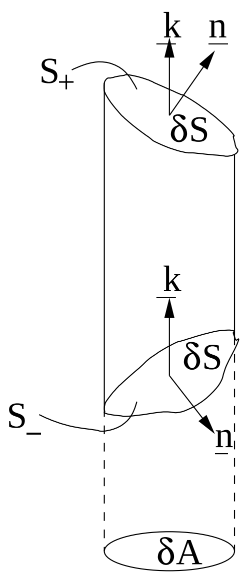
We extend the approach in section 7.2.2 from two dimensions to three. Since the divergence theorem is linear in the vector field, we consider $\underline{F} = F(x,y,z)\underline{k}$ (again, we provide vector fields in other directions and treat them similarly, taking linear combinations to get the general case). With $\underline{\nabla} \cdot \underline{F} = \frac{\partial F}{\partial z}$, we consider the integral

$$\int_V \underline{\nabla} \cdot \underline{F}\, dV = \int_D \left(\int_{Z_{xy}} \frac{\partial F}{\partial z} \right) dA,$$

where Z_{xy} is the range of the z integration for fixed x, y (in general this will be the union of disjoint intervals).



The boundary surface $S = \partial V$ can be divided into surfaces $S_\pm$ on which $\pm \underline{k} \cdot \underline{n} \geq 0$. $S_\pm$ is given by $z = z_\pm(x,y)$ at the ends of some interval in Z_{xy}.



From the diagram, $\delta A = \pm \underline{k} \cdot \underline{n}\ \delta S$ on $S_\pm$ (the signs are given by the fact that $\underline{n}$ is always the <u>outward</u> normal). The integral of $\frac{\partial F}{\partial z}$ then gives

$$F(x,y,z_+(x,y)) - F(x,y,z_-(x,y))$$

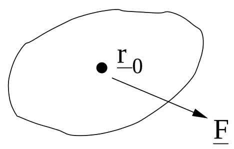
and the integral of this over D can be written $\int_{S_+} F\underline{k} \cdot \underline{n}\, dS + \int_{S_-} F\underline{k} \cdot \underline{n}\, dS$, or in total $\int_{\partial V} \underline{F} \dot{d} \underline{S}$, as required.

8 Some Applications of Integral Theorems

8.1 Integral Expressions of Div and Curl

The divergence theorem for a vector field $\underline{F}$ applied to a small volume V containing a point $\underline{r}_0$ gives

$$\int_{\partial V} \underline{F} \cdot d\underline{S} = \int_V \underline{\nabla} \cdot \underline{F}\, dV \approx (\underline{\nabla} \cdot \underline{F})(\underline{r}_0)V.$$

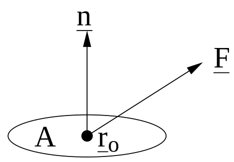


This becomes exact on taking the limit at $\underline{r}_0$:

$$\underline{\nabla} \cdot \underline{F} = \lim_{V \to 0} \tfrac{1}{V} \int_{\partial V} \underline{F} \cdot d\underline{S}.$$

Similarly, Stokes' theorem applied to a small part of a plane with area A and normal $\underline{n}$ gives

$$\int_{\partial A} \underline{F} \cdot d\underline{r} = \int_A (\underline{\nabla} \wedge \underline{F}) \cdot \underline{n}\, dA \approx (\underline{n} \cdot \underline{\nabla} \wedge \underline{F})(\underline{r}_0)A.$$



At $\underline{r}_0$,

$$\underline{n} \cdot \underline{\nabla} \wedge \underline{F} = \lim_{A \to 0} \tfrac{1}{A} \int_{\partial A} \underline{F} \cdot d\underline{r}.$$

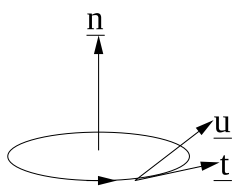
These are manifestly coordinate independent definitions of div and curl.

Examples: The expressions above allow interpretations of $\underline{\nabla} \cdot \underline{u}$ and $\underline{\nabla} \wedge \underline{u}$ when $\underline{u}$ is a velocity field, describing fluid flow.

1. Divergence: we know from section 4.3 that $\int_S \underline{u} \cdot d\underline{S}$ is the rate of fluid crossing δS (i.e. per unit time). Taking V to be the volume occupied by a fixed material quantity of fluid, we therefore have $\dot{V} = \int_{\partial S} \underline{u} \cdot d\underline{S}$ (for a boundary ∂V at some fixed time interval, $\delta t \dot{V}$ is the approximate volume that crosses in time δt, as the volume occupied by the fixed material quantity changes). Then, from above, at $\underline{r}_0$, $\underline{\nabla} \cdot \underline{u} = \lim_{V \to 0} \dot{V}/V$, the local (i.e. at $\underline{r}_0$) relative rate of change of volume occupied by a fixed material quantity of fluid.

   For example, $\underline{u}(\underline{r}) = \alpha \underline{r}$, fluid flows out of the origin. $\underline{\nabla} \cdot u = 3\alpha$, i.e. the volume increases at the same constant relative rate everywhere.

2. Curl: Take the plane area A to be a disc of radius a.



   $\int_{\partial A} \underline{u} \cdot d\underline{r} = \int_{\partial A} \underline{u} \cdot \underline{t}\, ds = (2\pi a) \times$ the average of $\underline{u} \cdot \underline{t}$, the component tangential to the boundary $= 2\pi a^2 \omega$, where we define $\omega = 1/a \times$ (the average of $\underline{u} \cdot \underline{t}$), as the local angular velocity around $\underline{n}$. Then

$$\underline{n} \cdot \underline{\nabla} \wedge \underline{u} = \lim_{a \to 0} \frac{1}{\pi a^2}(2\pi a^2 \omega) = 2\omega$$

   at $\underline{r}_0$, i.e. twice the local rate of rotation. For example, $\underline{u} = \underline{\omega} \wedge \underline{r}$, rigid motion with constant angular velocity $\underline{\omega}$ about $\underline{0}$, and $\underline{\nabla} \wedge (\underline{\omega} \wedge \underline{r}) = 2\underline{\omega}$, twice the rate of rotation.

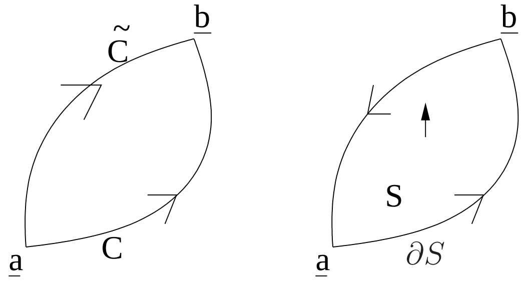
8.2 Conservative and Irrotational Fields, and Scalar Potentials

Consider three ways to define a vector field $\underline{F}$, (A) and (B) are equivalent definitions of a conservative vector field, and (C) defines an irrotational field:

(A) $\underline{F} = \underline{\nabla} f$ for some scalar field $f(\underline{r})$.
(B) $\int_C \underline{F} \cdot d\underline{r}$ is independent of C, for some given end points and orientation.
(C) $\underline{\nabla} \wedge \underline{F} = 0$.

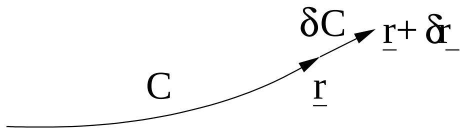
(A), (B) and (C) are all equivalent for $\underline{F}(\underline{r})$ on $\Re^3$. (A)$\Rightarrow$(B) with $\int_C \underline{F} \cdot d\underline{r} = f(\underline{b}) - f(\underline{a})$ was shown in section 2.4. (A)$\Rightarrow$(C) was shown in section 6.3 (via $\underline{\nabla} \wedge \underline{\nabla} f = 0$.)

We shall now show that (C)$\Rightarrow$(B) and (B)$\Rightarrow$(A) in $\Re^3$. Given $\underline{F}(\underline{r})$ obeying (C) on $\Re^3$, let C and $\tilde{C}$ be any two curves from $\underline{a} \to \underline{b}$:



If S is any surface with boundary $\partial S = C - \tilde{C}$, $\int_C \underline{F} \cdot d\underline{r} - \int_{\tilde{C}} \underline{F} \cdot d\underline{r} = \int_{\partial S} \underline{F} \cdot d\underline{r} = \int_S \underline{\nabla} \wedge \underline{F} \cdot d\underline{S}$ by Stokes' theorem. Hence, (C)⇒(B).

To show (B)⇒(A), we must define $f(\underline{r})$. Let us fix $\underline{a}$ and define $f(\underline{r}) = \int_C \underline{F}(\underline{r}') \cdot d\underline{r}'$ for any curve C from $\underline{a}$ to $\underline{r}$ ($\underline{r}'$ is a dummy variable). The integral is independent of C if (B) holds, so that $f(\underline{r})$ is well defined.



For a small change from $\underline{r}$ to $\underline{r} + \delta\underline{r}$, there is a small extension of C by δC as shown, and

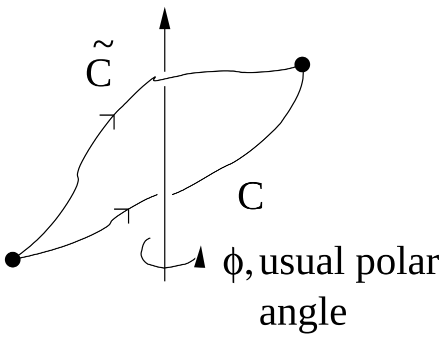
$$\begin{aligned} f(\underline{r} + \delta\underline{r}) &= \int_{C+\delta C} \underline{F} \cdot d\underline{r}' = \int_C \underline{F} \cdot d\underline{r}' + \int_{\delta C} \underline{F} \cdot d\underline{r}' \\ &= f(\underline{r}) + \underline{F}(\underline{r}) \cdot \delta\underline{r} + o(|\delta\underline{r}|). \\ \Rightarrow \delta f &= \underline{F}(\underline{r}) \cdot \delta\underline{r} + o(|\delta\underline{r}|). \end{aligned} \tag{8.1}$$

But we have $\delta f = \underline{\nabla} f \cdot \delta\underline{r} + o(|\delta\underline{r}|)$ from the definition of grad. Since this holds for any $\delta\underline{r}$, we have that $\underline{F} = \underline{\nabla} f$ and hence (B)⇒(A), as claimed.

If $\underline{F}(\underline{r})$ is defined ans satisfies (C) only on a subset D of $\Re^3$, then (B) and (A) will still hold on D if it is <u>simply connected</u>. By definition, this means that any curves C and $\tilde{C}$ with fixed end-points can be smoothly deformed into one another in D. Such a deformation generates a smooth surface S with $\partial S = C - \tilde{C}$, so the argument for (C)⇒(B) holds as above and (B)⇒(A) still follows.

If D is not simply connected, then the construction above produces a multi-valued scalar field $f(\underline{r})$ on D in general. However, we can always choose to restrict to a simply connected subset $D_0 \in D$ such that $f(\underline{r})$ will be single-valued on D_0.

Example: $\underline{F} = (-y/(x^2+y^2), x/(x^2+y^2), 0)$ is well-defined and obeys $\underline{\nabla} \wedge \underline{F} = \underline{0}$ on $D = \Re^3 - \{\text{z-axis}\}$, which is not simply connected.



$\underline{F} = \underline{\nabla} f$ where $f = \tan^{-1}(y/x) = \varphi$ is multi-valued on D, but single-valued on $D_0 = \Re^3 - \{\text{half plane } x \geq 0,\ y = 0\}$, in other words simply connected on D_0.

8.3 Conservation Laws

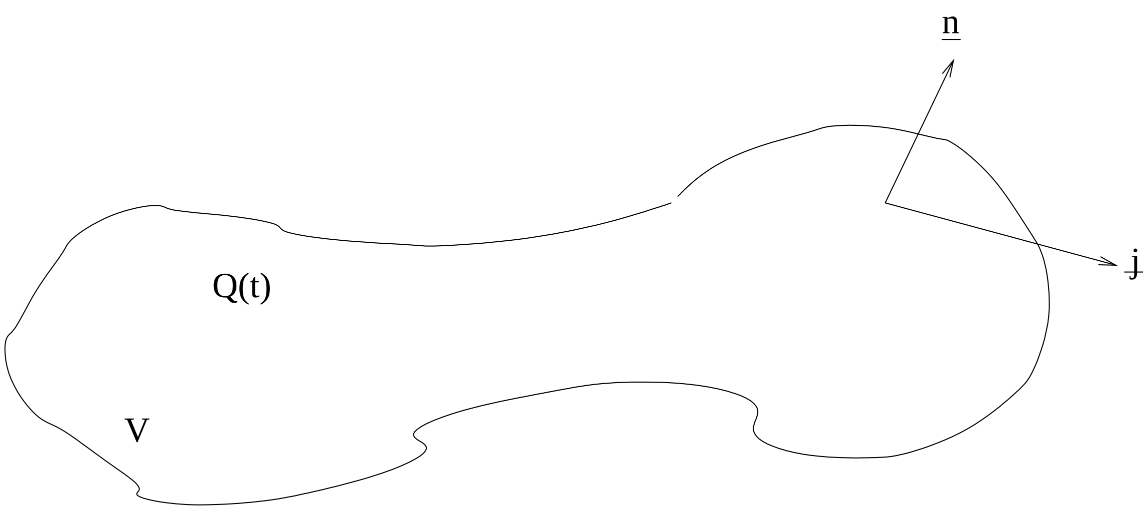
Consider time-dependent scalar and vector fields $\rho(\underline{r}, t)$ and $\underline{j}(\underline{r}, t)$ obeying the conservation equation

$$\frac{\partial \rho}{\partial t} + \underline{\nabla} \cdot \underline{j} = 0.$$

Let V be a fixed $t-$independent volume with boundary $S = \partial V$. Then $Q(t) = \int_V \rho(\underline{r}, t)\ dV$ satisfies

$$\frac{dQ}{dt} = \int_V \frac{\partial \rho}{\partial t}\ dV = -\int_V \underline{\nabla} \cdot \underline{j}\ dV = -\int_S \underline{j} \cdot d\underline{S}$$

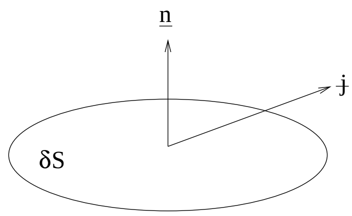
by the divergence theorem at fixed t.



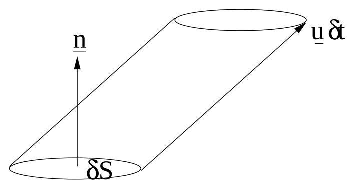
$Q(t)$ is the total amount of some quantity in V; $\rho(\underline{r}, t)$ is the density for this quantity, with $\rho(\underline{r}, t)\delta V$ the amount in a small volume δV; the flux of $\underline{j}$ out of S gives the rate of quantity leaving V, according to the equation above.

Examples:

1. Conservation of electric charge. $\rho(\underline{r},t)$ is the charge density, $\rho(\underline{r},t)\delta V$ is the charge in a small volume δV. $Q(t)$ is the total charge in V. $\underline{j}(\underline{r},t)$ is the electric current density, and $\underline{j}\cdot\delta\underline{S}$ is the charge flowing through $\delta\underline{S}$ per unit time.



2. Conservation of mass for a fluid. $\rho(\underline{r},t)$ is the mass density and $\underline{j}=\rho\underline{u}$ with $\underline{u}(\underline{r},t)$ being a fluid velocity field. By extension of the discussion of fluid volume in section 4.3, $\rho\underline{u}\delta t\cdot\underline{n}\delta S$ is the mass of fluid that crosses δS in time δt, so $\frac{dQ}{dt}=-\int_S\underline{j}\cdot d\underline{S}$ does indeed imply conservation of mass.



The conservation equation in this case is

$$\frac{\partial\rho}{\partial t}+\underline{\nabla}\cdot(\rho\underline{u})=0.$$

For the case in which ρ is constant and uniform (i.e. independent or $\underline{r}$ and t), this becomes $\underline{\nabla}\cdot\underline{u}=0$: the fluid is then called incompressible.

Returning to the general case, take V to be a solid sphere of radius R, $S=\partial V$ is a sphere and $|\underline{r}|=R$. If $|\underline{j}|\to 0$ sufficiently rapidly as $|\underline{r}|=R\to\infty$, then $Q=\int_{\Re^3}\rho(\underline{r},t)\,dV$ is constant $\Rightarrow\dot{Q}=0$ and so, for example, the total mass or total charge is conserved.

9 Othorgonal Curvilinear Coordinates

9.1 Line, Area and Volume Elements

For coordinates u,v,w on $\Re^3$, we have a smooth function $\underline{r}(u,v,w)$ with $d\underline{r}=\frac{\partial\underline{r}}{\partial u}\,du+\frac{\partial\underline{r}}{\partial v}\,dv+\frac{\partial\underline{r}}{\partial w}\,dw$. For a good parameterisation, we require $\frac{\partial\underline{r}}{\partial u}$, $\frac{\partial\underline{r}}{\partial v}$ and $\frac{\partial\underline{r}}{\partial w}$ to be linearly independent

$$\Leftrightarrow\frac{\partial\underline{r}}{\partial u}\cdot\left(\frac{\partial\underline{r}}{\partial v}\wedge\frac{\partial\underline{r}}{\partial w}\right)\neq 0$$

(see the Jacobian condition in section 4.4.2). These vectors are tangent to the curves parameterised by u, v and w when the other two are held fixed.

We say that u, v, w are orthogonal curvilinear coodinates if the tangent vectors above are orthogonal. Then, we can set $\frac{\partial \underline{r}}{\partial u} = h_u \underline{e}_u$, $\frac{\partial \underline{r}}{\partial v} = h_v \underline{e}_v$, $\frac{\partial \underline{r}}{\partial w} = h_w \underline{e}_w$, with $h_u > 0$, $h_v > 0$, $h_w > 0$ and $\underline{e}_u$, $\underline{e}_v$, $\underline{e}_w$ forming an orthonormal and right-handed basis (right-handed meaning $\underline{e}_u \wedge \underline{e}_v = \underline{e}_w$, which may be achieved by ordering the coordinates appropriately). In these coordinates, the line element is

$$d\underline{r} = h_u \underline{e}_u \, du + h_v \underline{e}_v \, dv + h_w \underline{e}_w \, dw$$

and the positive scale factors h_u, h_v, h_w determine the changes in length along each orthogonal direction resulting from changes in u, v and w:

$$|\delta \underline{r}|^2 = h_u^2 (\delta u)^2 + h_v^2 (\delta v)^2 + h_w^2 (\delta w)^2 + \text{o-terms}$$

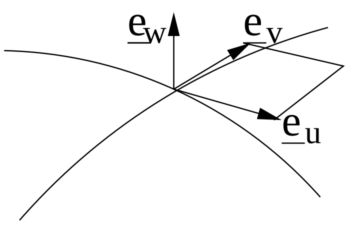
Examples:

1. Cartesian coordinates $\underline{r}(x, y, z) = x\hat{\underline{i}} + y\hat{\underline{j}} + z\hat{\underline{k}}$. $h_x = h_y = h_z = 1$, $\underline{e}_x = \hat{\underline{i}}$, $\underline{e}_y = \hat{\underline{j}}$, $\underline{e}_z = \hat{\underline{k}}$.

2. Cylindrical polar coordinates $\underline{r}(\rho, \varphi, z) = \rho(\cos\varphi \hat{\underline{i}} + \sin\varphi \hat{\underline{j}}) + z\hat{\underline{k}}$, $h_\rho = h_z = 1, h_\varphi = \rho$. $\underline{e}_\rho$, $\underline{e}_\varphi$ and $\underline{e}_z$ are as in section 1.3.

3. Spherical polar coordinates $\underline{r}(r, \theta, \varphi) = r(\cos\varphi \sin\theta \hat{\underline{i}} + \sin\varphi \sin\theta \hat{\underline{j}} + \cos\theta \hat{\underline{k}})$, $h_r = 1, h_\theta = r, h_\varphi = r\sin\theta$. $\underline{e}_r$, $\underline{e}_\theta$, $\underline{e}_\varphi$ are as in section 1.3.

Consider a surface with w constant (say), parameterised by u and v. The vector area element is $d\underline{S} = \frac{\partial \underline{r}}{\partial u} \wedge \frac{\partial \underline{r}}{\partial v} \, du \, dv$ (a general formula) $= h_u \underline{e}_u \wedge h_v \underline{e}_v \, du \, dv = h_u h_v \underline{e}_w \, du \, dv$, for orthogonal coordinates.



$\underline{e}_w$ is normal to the surface and we can see that δS is a small rectangle with sides $h_u \delta u$ and $h_v \delta v$, approximately. The volume element $dV = \frac{\partial \underline{r}}{\partial u} \cdot \left(\frac{\partial \underline{r}}{\partial v} \wedge \frac{\partial \underline{r}}{\partial w} \right) du \, dv \, dw = h_u h_v h_w \, du \, dv \, dw$, which corresponds to a small cuboid with sides $h_u \delta u, h_v \delta v$ and $h_w \delta w$, approximately.

9.2 Grad, Div and Curl

Consider $f(\underline{r}(u, v, w))$ and compare

$$df = \frac{\partial f}{\partial u} du + \frac{\partial f}{\partial v} dv + \frac{\partial f}{\partial w} dw$$

with

$$= (\underline{\nabla} f) \cdot d\underline{r}.$$

(this is the coordinate independent definition of grad, see section 1.1). From the expression for $d\underline{r}$ above, and since the basis vectors are orthonormal,

$$\underline{\nabla} f = \frac{1}{h_u}\underline{e}_u \frac{\partial f}{\partial u} + \frac{1}{h_v}\underline{e}_v \frac{\partial f}{\partial v} + \frac{1}{h_w}\underline{e}_w \frac{\partial f}{\partial w}$$

is the gradient in general orthogonal curvilinear coordinates. To apply this, we need to know the various h factors.

Example: $f = r \sin\theta \cos\varphi$ in spherical polar coordinates. Then,

$$\begin{aligned}\underline{\nabla} f &= \sin\theta\cos\varphi\,\underline{e}_r + \frac{1}{r}(r\cos\theta\cos\varphi)\underline{e}_\theta + \frac{1}{r\sin\theta}(-r\sin\theta\sin\varphi)\underline{e}_\varphi \\ &= \cos\varphi(\sin\theta\,\underline{e}_r + \cos\theta\,\underline{e}_\varphi) - \sin\varphi\,\underline{e}_\varphi.\end{aligned}$$

Now consider the differential operator

$$\underline{\nabla} = \frac{1}{h_u}\underline{e}_u \frac{\partial}{\partial u} + \frac{1}{h_v}\underline{e}_v \frac{\partial}{\partial v} + \frac{1}{h_w}\underline{e}_w \frac{\partial}{\partial w}$$

applied to some vector field $\underline{F} = F_u\underline{e}_u + F_v\underline{e}_v + F_w\underline{e}_w$ using scalar or vector products to obtain div or curl. The results are similar to standard Cartesian formulae but with extra h factors appearing here and there.

$$\begin{aligned}\underline{\nabla} \wedge \underline{F} &= \frac{1}{h_v h_w}\left[\frac{\partial}{\partial v}(h_w F_w) - \frac{\partial}{\partial w}(h_v F_v)\right]\underline{e}_u + \text{(two similar terms)} \\ &= \frac{1}{h_u h_v h_w}\begin{vmatrix} h_u\underline{e}_u & h_v\underline{e}_v & h_w\underline{e}_w \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ h_u F_u & h_v F_v & h_w F_w \end{vmatrix},\end{aligned} \tag{9.1}$$

and

$$\underline{\nabla} \cdot \underline{F} = \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}(h_v h_w F_u) + \text{(two similar terms)}\right]. \tag{9.2}$$

There are a number of ways these can be derived[2]:

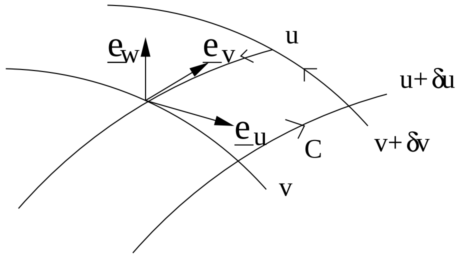
1. Apply $\underline{\nabla}\cdot$ or $\underline{\nabla}\wedge$ and differentiate the basis vectors explicitly. This is most useful when we have explicit formulae for the basis vectors (see example sheet III).

2. Use an algebraic approach: first, apply $\underline{\nabla}\cdot$ or $\underline{\nabla}\wedge$ but calculate the result by first writing $\underline{F}$ in terms of $\underline{\nabla} u$, $\underline{\nabla} v$, $\underline{\nabla} w$, in a suitable way and using $\underline{\nabla}\wedge\underline{\nabla} f = \underline{\nabla}\cdot(\underline{\nabla}\wedge f) = \underline{0}$.

[2] All of which are non-examinable.

3. Use integral expressions for div and curl as in section 8.1. This approach gives good geometric insight.

To follow (3) for curl, consider the area defined by changes δu and δv with w fixed. C is the small circuit shown,



the boundary of an area of approximate size $h_u h_v \delta u \delta v$, with a normal $\underline{e}_w$. We use the standard approximate expressions for integrals along each segment C, and the fact that $h_u \delta u$ and $h_v \delta_v$ are the lengths of the segments, and work to order $\delta u \delta v$ (in order to divide by the area and take the limit) and so we need to keep track of the arguments of h_u, h_v and F_u, F_v: the segments of C in turn are:

$$\int_C \underline{F} \cdot d\underline{r} \approx F_u h_u(u,v)\delta u + F_v h_v(u+\delta u, v)\delta v - F_u h_u(u, v+\delta v)\delta u - F_v h_v(u,v)\delta v$$

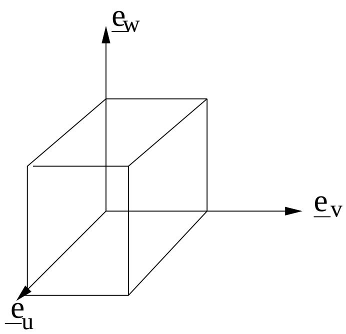
$$\approx \left[\frac{\partial}{\partial u}(h_v F_v) - \frac{\partial}{\partial v}(h_u F_u)\right] \delta u \delta v. \tag{9.3}$$

Dividing by area and taking the limit, we get, from section 8.1

$$\underline{e}_w \cdot \underline{\nabla} \wedge \underline{F} = \frac{1}{h_u h_w}\left[\frac{\partial}{\partial u}(h_v F_v) - \frac{\partial}{\partial v}(h_u F_u)\right],$$

as claimed in Eq. 9.1.

We take a similar approach for div: consider an approximate cuboid volume defined by the changes $\delta u, \delta v, \delta w$ with boundary S.



$$\int_S \underline{F} \cdot d\underline{S} \approx [h_u h_v F_w(u,v,w+\delta w) - h_u h_v F_w(u,v,w)]\, \delta u \delta v + \text{two similar terms}$$

$$\approx \frac{\partial}{\partial w}(h_u h_v F_w)\delta u \delta v \delta w + \text{two similar terms.}$$

Then, dividing by $\delta V = h_u h_v h_w \delta u \delta v \delta w$ and taking the limit gives the formula for $\underline{\nabla} \cdot \underline{F}$ above.

In this course, we need to know how to apply these formulae in particular coordinate systems.

Example: $\underline{A} = \frac{1}{r} \tan \frac{\theta}{2} \underline{e}_\phi$ $(\theta \neq \pi)$ in spherical polars.

$$\underline{\nabla} \wedge \underline{A} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \underline{e}_r & r\underline{e}_\theta & r\sin\theta \underline{e}_\varphi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \varphi} \\ 0 & 0 & r\sin\theta(\frac{1}{r}\tan(\theta/2)) \end{vmatrix} = \frac{1}{r^2 \sin\theta} \underline{e}_r \frac{\partial}{\partial \theta}\left(2\sin^2(\theta/2)\right) = \frac{1}{r^2}\underline{e}_r.$$

10 Gauss' Law and Poisson's Equation

10.1 Laws of Gravitation

Consider a distribution of mass producing a gravitational force $\underline{F}(\underline{r})$ on a point mass m at $\underline{r}$. The total force is a sum of contributions from each part of the distribution, and is proportional to m. Setting $\underline{F}(\underline{r}) = m\underline{g}(\underline{r})$ defines a vector field $\underline{g}(\underline{r})$, the gravitational field, or force per unit mass, or acceleration due to gravity. The gravitational field is conservative:

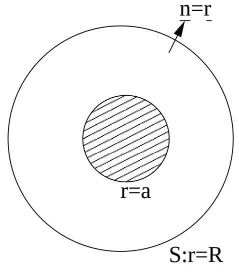
$$\int_C \underline{g} \cdot d\underline{r} = 0$$

for any closed curve C.

$$\int_S \underline{g} \cdot d\underline{S} = -4\pi G M$$

is Gauss' Law in integral form for any closed surface S, which is the boundary of volume V with M being the total mass contained in V. G is Newton's gravitational constant. These equations determine $\underline{g}(\underline{r})$ from the mass distribution. In particular, Newton's law of gravitation follows from Gauss' Law for the flux of $\underline{g}$ as written above, together with an assumption about symmetry.

Consider a total mass M distributed with spherical symmetry about $\underline{0}$, and all of the mass contained within a radius $r = a$. Spherical symmetry implies that $\underline{g}(\underline{r}) = g(r)\hat{\underline{r}}$, with $\hat{\underline{r}} = \underline{r}/|\underline{r}|$ being a radial unit vector and $g(r)$ being some scalar function.



Consider Gauss' Law with S being a sphere of radius $r = R > a$. The outward normal is $\underline{\hat{r}}$.

$$\int_S \underline{g} \cdot d\underline{S} = \int_S g(R)\underline{\hat{r}} \cdot \underline{\hat{r}} dS = \int_S g(R) dS = 4\pi R^2 g(R)$$

by using spherical symmetry. Then, using Gauss' Law,

$$4\pi R^2 g(R) = -4\pi GM \Rightarrow g(R) = -\frac{GM}{R^2}$$

for any $R > a$, or

$$\underline{g}(\underline{r}) = -\frac{GM}{r^2}\underline{\hat{r}} \qquad (r > a).$$

The gravitational force on a mass m at $\underline{r}$ is therefore $\underline{F}(\underline{r}) = -GMm\underline{\hat{r}}/r^2$. In the limit $a \to 0$ we get a point mass M at the origin $\underline{0}$ and we recover Newton's Law of Gravitation for point masses.

For <u>any</u> a, the field and the force for $r > a$ depends only on the total mass M, assuming spherical symmetry (and no mass in $r > a$). This simple behaviour means that we can identify the gravitational field that keeps the moon in orbit (where the earth is approximately point-like) with the one that pulls an apple from a tree (where the earth is certainly not point-like).

The condition $\int_C \underline{g} \cdot d\underline{r} = 0$ for any closed C can be re-written $\int_S \underline{\nabla} \wedge \underline{g} d\underline{S} = 0$ for any S with $\partial S = C$ by Stokes' theorem. Since S is arbitrary, this $\Leftrightarrow \underline{\nabla} \wedge \underline{g} = 0$. In our above example with spherical symmetry, $\underline{\nabla} \wedge \underline{g} = \underline{0}$ automatically.

If the mass distribution is not sufficiently symmetrical, it is difficult to solve for $\underline{g}(\underline{r})$ from Gauss' Law in integral form, but it can be re-cast in differential form. Suppose that the mass M arises from a mass density $\rho(\underline{r})$. Then

$$\int_S \underline{g} \cdot d\underline{S} = -4\pi GM \Rightarrow \int_V \underline{\nabla} \cdot \underline{g} dV = -4\pi G \int_V \rho(\underline{r})\, dV$$

with $S = \partial V$ by Gauss' theorem $\Rightarrow \int_V (\underline{\nabla} \cdot \underline{g} + 4\pi G\rho)\, dV = 0$ for <u>any</u> volume V, hence

$$\Rightarrow \underline{\nabla} \cdot \underline{g} = -4\pi G\rho,$$

<u>Gauss' law in differential form</u>. Furthermore, since $\underline{\nabla} \wedge \underline{g} = 0$, we can introduce the <u>gravitational potential</u> $\varphi(\underline{r})$ with $\underline{g} = -\underline{\nabla}\varphi$ and then Gauss' law becomes

$$\nabla^2 \varphi = 4\pi G\rho,$$

Poisson's equation. In the previous case with spherical symmetry, we can choose $\varphi(r) = -GM/r$ for $r > a$. It obeys the boundary condition $\varphi \to 0$ as $r \to \infty$.

10.2 Laws of Electrostatics

Consider a distribution of electric charges at rest. They produce a force on a charge a at rest at $\underline{r}$ which is proportional to q: $\underline{F} = q\underline{E}(\underline{r})$, where $\underline{E}(\underline{r})$ is the electric field (a vector field), or force per unit charge.

The electric field obeys $\int_C \underline{E} \cdot d\underline{r} = 0$ for any closed curve C, i.e. $\underline{E}(\underline{r})$ is conservative. Also, it obeys Gauss' Law:

$$\int_S \underline{E} \cdot d\underline{S} = \frac{Q}{\epsilon_0}$$

for any closed surface S which is the boundary of a volume V, with Q being the total charge contained within V. ϵ_0 is the permittivity of free space, a constant of nature.

Mathematically speaking, this is identical to the gravitational case in section 10.1, and the laws above can be re-expressed in differential form in the same way: $\underline{\nabla} \wedge \underline{E} = 0$ and

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

(Gauss' Law in differential form), where $\rho(\underline{r})$ is the electric charge density. these can be re-expressed as $\underline{E} = -\underline{\nabla}\varphi$, with $\varphi(\underline{r})$ being the electrostatic potential and $\nabla^2\varphi = -\rho/\epsilon_0$, Poisson's equation.

For example: take a total charge Q distributed with spherical symmetry about $\underline{0}$ and all of the charge being contained within a radius $r = a$. We solve this using Gauss' Law in integral form, just as in the gravitational case. Thus,

$$\underline{E}(\underline{r}) = \frac{Q}{4\pi\epsilon_0}\frac{\hat{\underline{r}}}{r^2}$$

for $r > a$ and

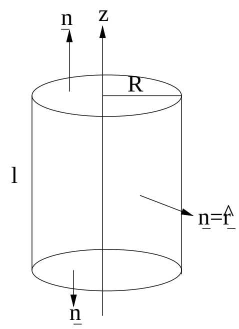
$$\varphi(\underline{r}) = \frac{Q}{4\pi\epsilon_0}\frac{1}{r}.$$

we obtain a point charge at $\underline{0}$ in the limit $a \to 0$. from this result for $\underline{E}$, we recover Coulomb's Law for the force on another point charge q at $\underline{r}$:

$$\underline{F} = q\underline{E} = \frac{qQ}{4\pi\epsilon_0}\frac{\hat{\underline{r}}}{r^2}.$$

Example:

The electric field due to a line of charge. σ is the charge per unit length along the z−axis. We have cylindrical symmetry, but here we shall use r for the radial coordinate in cylindrical polar coordinates. $\underline{E}(\underline{r}) = E(r)\hat{\underline{r}}$ by symmetry. We apply Gauss' law in integral form to a cylinder of radius R and length l as shown.



$$\int_S \underline{E} \cdot d\underline{S} = 2\pi R l E(R) = \underbrace{\sigma l}_{\text{total charge inside cylinder}} /\epsilon_0,$$

where there is no contribution from the end-caps. Thus,

$$\underline{E}(\underline{r}) = \frac{\sigma}{2\pi\epsilon_0}\frac{1}{r}\hat{\underline{r}} \qquad (r > 0).$$

10.3 Poisson's Equation and Laplace's Equation

Consider $\nabla^2\varphi = -\rho$: we aim to solve for $\varphi(\underline{r})$ given $\rho(\underline{r})$, with suitable boundary conditions. We saw this above for $\rho \to -4\pi G\rho$ (gravitation) and $\rho \to \rho/\epsilon_0$ (electrostatics). For $\rho = 0$, we get Laplace's equation $\nabla^2\varphi = 0$, which crops up many times in physics and maths.

One example is irrotational and incompressible fluid flow. The velocity $\underline{u}(\underline{r})$ satisfies $\underline{\nabla} \wedge \underline{u} = \underline{0} \Rightarrow \underline{u} = \underline{\nabla}\varphi$, where φ is the velocity potential and $\underline{\nabla} \cdot \underline{u} = 0 \Rightarrow \nabla^2\varphi = 0$.

It also occurs in two dimensions or greater than three dimensions as well[3] Expressions for ∇^2 are available for non-Cartesian coordinates, but they can be rather complicated.

We shall mainly be concerned with explicit solutions with spherical or cylindrical symmetry in three dimensions, the latter being equivalent to solutions with circular symmetry in two dimensions. Let us use r for the radial coordinate in either spherical or cylindrical polar coordinates, then the solution $\varphi(r)$ has spherical or cylindrical symmetry and $\underline{\nabla}\varphi = \varphi'(r)\hat{\underline{r}}$. Laplace's equation $\nabla^2\varphi = 0$ then becomes an ordinary differential equation for $\varphi(r)$:

1. For spherical symmetry,
$$\varphi'' + \frac{2}{r}\varphi' = \frac{1}{r^2}\left(r^2\varphi'\right)' = 0 \Rightarrow \varphi = \frac{A}{r} + B$$
is the general solution.

2. For cylindrical symmetry,
$$\varphi'' + \frac{1}{r}\varphi' = \frac{1}{r}\left(r\varphi'\right)' = 0 \Rightarrow \varphi = A\log r + B$$

[3]Note that $\nabla^2 = \frac{\partial^2}{\partial x_i \partial x_i}$ for Cartesian coordinates in $\Re^n$.

is the general solution.

Solutions to Poisson's equation can be obtained in a similar way (i.e. by integrating the differential equation directly) or by adding particular integrals to the discussions above.

Example: Consider spherically symmetric solutions of $\nabla^2\varphi = -\rho_0$, a constant. Note that from section 6.2, $\nabla^2(r^\alpha) = \alpha(\alpha+1)r^{\alpha-2}$ so $\alpha = 2$ gives the particular integral $-\frac{1}{6}\rho_0 r^2$. The general solution with spherical symmetry is then

$$\varphi(r) = \underbrace{\frac{A}{r} + B}_{\text{general solution of Laplace's equation}} - \underbrace{\frac{1}{6}\rho_0 r^2}_{\text{particular integral}} .$$

For a unique solution to Poisson's (or Laplace's) equation, we must specify the boundary conditions. If φ is defined in all of $\Re^3$, we often require $\varphi \to 0$ as $|\underline{r}| \to \infty$. If φ is a solution on a bounded volume V, then there are two common kinds of boundary condition on ∂V:

1. Specify φ: a <u>Dirichlet condition</u>
2. Specify[4] $\frac{\partial\varphi}{\partial n} = \underline{n}\cdot\underline{\nabla}\varphi$: a <u>Neumann condition</u>, where $\underline{n}$ is an outward normal on ∂V.

The choice of boundary condition depends on the physical content, for example specifying $\partial\varphi/\partial n$ corresponds to specifying the normal component of $\underline{g}$ or $\underline{E}$. We can also specify different boundary conditions on different boundary components.

[4]Notice the new notation of a derivative with respect to a vector!

Example: We seek a spherical symmetric $\varphi(r)$ with

$$\nabla^2\varphi = \begin{cases} -\rho_0 & r \le a \\ 0 & r > a \end{cases},$$

where ρ_0 is a constant and φ is non-singular at $r = 0$, $\varphi(r \to \infty) \to 0$, and φ and φ' are continuous at $r = a$. From above, the general solution is

$$\varphi = \begin{cases} \cancel{A/r} + B - \frac{1}{6}\rho_0 r^2 & r \le a \\ C/r + \cancel{D} & r > a \end{cases}$$

The physical interpretation is that φ is the gravitational potential inside and outside a planet of constant density ρ_0, radius a and total mass M. We restore gravitational conventions by replacing $\rho_0 \to -4\pi G\rho_0$ and write the results above in terms of $M = 4/3\pi a^3\rho_0$. A and D are fixed to be zero by the boundary conditions at $r = 0, \infty$. The matching at $r = a$ implies

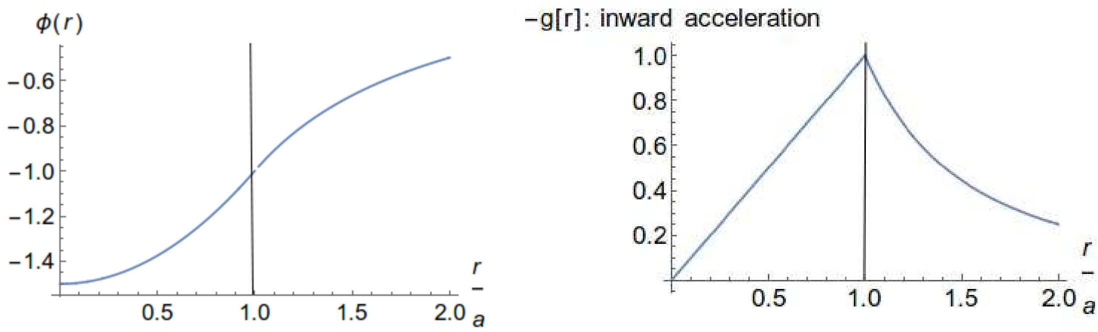
$$\varphi : \; B + \frac{1}{6}4\pi G\rho_0 a^2 = C/a$$

$$\varphi' : \; +\frac{1}{3}4\pi G\rho_0 a = -C/a^2$$

$$\Rightarrow \varphi(r) = \begin{cases} GM/(2a)\,((r/a)^2 - 3) & r \le a \\ -GM/r & r > a \end{cases}$$

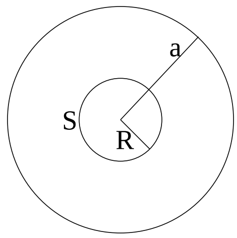
The gravitational field $\underline{g} = -\underline{\nabla}\varphi = g(r)\underline{e}_r$ with

$$g(r) = -\varphi'(r) = \begin{cases} -GMr/a^3 & r \le a \\ -GM/r^2 & r > a \end{cases}.$$



In the left-hand panel, we have fixed $GM/a = 1$ in the vertical axis.

Example: An alternative solution is to apply Gauss' Law for a flux of $\underline{g} = g(r)\underline{e}_r$ out of S, a sphere of radius R. $R \geq a$ has already been covered in section 10.1 (the result is independent of how the mass is distributed, provided there is spherical symmetry).



For $R \leq a$, we get

$$\int_S \underline{g} \cdot d\underline{S} = 4\pi R^2 g(R) = -4\pi G \underbrace{M(R/a)^3}_{\text{mass in } R \leq a} \Rightarrow g(R) = -GMR/a^3,$$

as we found from Poisson's equation.

In general, even if the problem has nothing to do with gravitation or electrostatics, we can solve $\nabla^2\varphi = -\rho$ with ρ and φ exhibiting enough symmetry by considering the flux of $\underline{\nabla}\varphi$ out of a suitable surface $S = \partial V$:

$$\int_S \underline{\nabla}\varphi \cdot d\underline{S} = -\int_V \rho \; dV.$$

This is called the *Gauss Flux Method*.

11 General Results for Laplace's and Poisson's Equations

11.1 Uniqueness Theorems

11.1.1 Statement and proof

Consider $\varphi(\underline{r})$ satisfying Poisson's equation $\nabla^2\varphi = -\rho$ for some $\rho(\underline{r})$ on a bounded volume V with boundary $S = \partial V$ being a closed surface, with outward normal $\underline{n}$. Suppose that φ satisfies either

1. $\varphi(\underline{r}) = f(\underline{r})$ on S (Dirichlet condition, D)

2. $\frac{\partial\varphi(\underline{r})}{\partial n} = \underline{n} \cdot \underline{\nabla}\varphi = g(\underline{r})$ (Neumann condition, N)

where $f(\underline{r})$ or $g(\underline{r})$ are given. Then either

1. $\varphi(\underline{r})$ is unique, or

2. $\varphi(\underline{r})$ is unique up to the replacement $\varphi \to \varphi + c$, where c is a constant.

Proof: Let $\varphi_1(\underline{r})$ and $\varphi_2(\underline{r})$ satisfy Poisson's equation, each obeying the boundary conditions (1) or (2). Then $\psi(\underline{r}) = \varphi_2(\underline{r}) - \varphi_1(\underline{r})$ satisfies $\nabla^2\psi = 0$ on V by linearity and (D) $\psi = 0$ on S or (N) $\partial\psi/\partial n = \underline{n} \cdot \underline{\nabla}\psi = 0$ on S. By the divergence theorem, $\int_V \underline{\nabla} \cdot (\psi\underline{\nabla}\psi)\, dV = \int_S \psi\underline{\nabla}\psi \cdot d\underline{S}$. But

$$\underline{\nabla} \cdot (\psi\underline{\nabla}\psi) = \underline{\nabla}\psi \cdot \underline{\nabla}\psi + \underbrace{\psi\nabla^2\psi}_{\text{0 on V}} = |\underline{\nabla}\psi|^2.$$

So

$$\int_V |\underline{\nabla}\psi|^2\, dV = \int_S \psi\frac{\partial\psi}{\partial n}\, dS = 0$$

by either (N) or (D). Since $|\underline{\nabla}\psi|^2 \geq 0$, the integral can only vanish if $|\underline{\nabla}\psi| = 0 \Rightarrow \underline{\nabla}\psi = \underline{0} \Rightarrow \psi = c$, a constant, on V. So:
(D) $\psi = 0$ on $S \Rightarrow c = 0$ and $\varphi_1 = \varphi_2$ on V, or
(N) $\varphi_2(\underline{r}) = \varphi_1(\underline{r}) + c$ as claimed.

11.1.2 Comments

- The result says nothing about existence. Indeed, for $\nabla^2\varphi = -\rho$ on V with $\frac{\partial\varphi}{\partial n} = g$ on ∂V (Neumann) there can only be a solution if

$$\int_V \nabla^2\varphi\, dV = \int_{\partial V} \frac{\partial\varphi}{\partial n}\, dS \Leftrightarrow \int_V \rho\, dV + \int_{\partial V} g\, dS = 0, \tag{11.1}$$

  but ρ and g are fixed to begin with, so Eq. 11.1 may not necessarily be true.

- The uniqueness result has practical importance: if we can find a solution by <u>any</u> means, then we know it is the <u>only</u> solution. For example: find a solution with symmetry $\varphi(r)$, or more generally, linear combinations of separable solutions[5] $\varphi = X(x)Y(y)$ or $R(r)Y(\theta)$.

- The theorem can be stated and proved similarly for regions in $\Re^{2,3,\cdots}$: it just uses definitions of grad and div, and the divergence theorem.

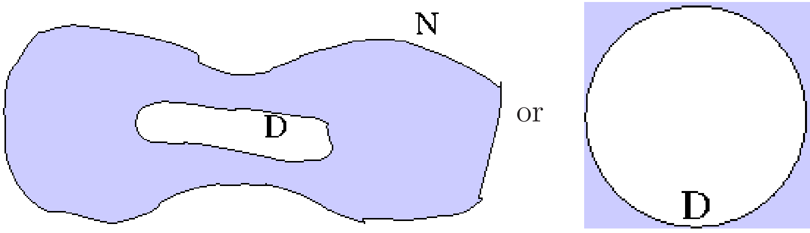
- The results extend to unbounded domains, for example a sphere of radius R, with $|\psi(\underline{r})| = \mathcal{O}(1/R)$ and $|\frac{\partial\psi}{\partial n}(\underline{r}| = \mathcal{O}(1/R^2)$ for $|\underline{r}| = R \Rightarrow |\int_S \psi\frac{\partial\psi}{\partial n}\, dS| = \mathcal{O}(1/R) \to 0$ as $R \to \infty$. This last result follows because

$$\int_S \psi\frac{\partial\psi}{\partial n}\, dS \leq \max_{|\underline{r}|=R}|\psi\frac{\partial\psi}{\partial n}| \times 4\pi R^2 \leq \text{constant} \times (\frac{1}{R}\frac{1}{R^2})4\pi R^2 = \mathcal{O}(\frac{1}{R}).$$

  These conditions on ψ are ensured by taking for example $|\psi(\underline{r})| = \mathcal{O}(1/R)$ and $|\underline{\nabla}\psi(\underline{r})| = \mathcal{O}(1/R^2)$. Therefore, the proof extends to $V = \Re^3$ with these boundary conditions as $R = |\underline{r}| \to \infty$ (S is the "sphere at infinity").

[5] See Part IB Methods.

- Similar results apply to related equations and various kinds of boundary conditions, perhaps with D or N being different parts of the boundary:



- The proof uses a special case of the result

$$\int_S (u\underline{\nabla}v)\cdot d\underline{S} = \int_V \underline{\nabla}u\cdot\underline{\nabla}v\ dV + \int_V u\nabla^2 v\ dV,$$

  sometimes called Green's First Identity. Antisymmetrising in u and v gives

$$\int_S \left(u\underline{\nabla}v - v\underline{\nabla}u\right)\ d\underline{S} = \int_V \left(u\nabla^2 v - v\nabla^2 u\right)\ dV,$$

  sometimes called Green's Second Identity.

11.2 Laplace's Equation and Harmonic Functions

Solutions of Laplace's equation $\nabla^2\varphi = 0$ are also called <u>harmonic</u> functions. They arise in many areas of physics and mathematics and have some very special properties.

11.2.1 The Mean Value Property

Suppose $\varphi(\underline{r})$ is harmonic in a region V containing a solid sphere $V_R:\ |\underline{r}-\underline{a}| \le R$ with boundary $S_R:\ |\underline{r}-\underline{a}| = R$. Then $\varphi(\underline{a}) = \overline{\varphi}(R)$ where

$$\overline{\varphi}(R) = \frac{1}{4\pi R^2}\int_{S_R}\varphi(\underline{r})\ dS,$$

the average of φ over S_R.

Proof: Note that $\overline{\varphi}(R)\to\varphi(a)$ as $R\to 0$. Take spherical polar coordinates (u,θ,χ) centred on $\underline{r}=\underline{a}$. Then, the scalar area element on S_R is $(u=R)$ $dS = R^2\sin\theta\ d\theta\ d\chi \Rightarrow dS/R^2$ is independent of R. Hence

$$\frac{d}{dR}\overline{\varphi}(R) = \frac{1}{4\pi R^2}\int_{S_R}\left.\frac{\partial\varphi}{\partial u}\right|_{u=R}\ dS.$$

But $\frac{\partial\varphi}{\partial u} = \underline{e}_u\cdot\underline{\nabla}\varphi = \underline{n}\cdot\underline{\nabla}\varphi = \frac{\partial\varphi}{\partial n}$ on S_R.

$$\Rightarrow \frac{d\overline{\varphi}(R)}{dR} = \frac{1}{4\pi R^2}\int_{S_R}\underline{\nabla}\varphi\cdot d\underline{S} = \frac{1}{4\pi R^2}\int_{V_R}\nabla^2\varphi\ dV = 0.$$

So $\overline{\varphi}(R)$ is independent of R, and the result follows.

11.2.2 The Maximum (or Minimum) Principle

We say that $\varphi(\underline{r})$ has a local maximum at $\underline{a}$ if, for some $\epsilon > 0$, $\varphi(\underline{r}) < \varphi(\underline{a})$ when $0 < |\underline{r} - \underline{a}| < \epsilon$. The maximum principle states that if a function φ is harmonic on a region V, then φ cannot have a local maximum at an interior point $\underline{a}$ of V.

Proof: If φ has a local maximum at $\underline{a}$ in the interior, we choose ϵ so that $|\underline{r}-\underline{a}| < \epsilon$ lies within V (this is possible, since $\underline{a}$ is an interior point) and then, for any R with $0 < R < \epsilon$, $\varphi(\underline{r}) < \varphi(\underline{a})$ for $|\underline{r} - \underline{a}| = R \Rightarrow \overline{\varphi}(R) < \varphi(\underline{a})$. But this contradicts the mean value property for a harmonic function.

The result is consistent with the usual analysis of stationary points by differentiation. Suppose at $\underline{r} = \underline{a}$, $\underline{\nabla}\varphi = \underline{0}$ and $H_{ij} = \frac{\partial^2\varphi}{\partial x_i \partial x_j}$ has eigenvalues λ_i. φ is harmonic $\Rightarrow \nabla^2\varphi = \frac{\partial^2\varphi}{\partial x_i \partial x_i} = H_{ii} = \sum_i \lambda_i = 0$. (Note that the sum of eigenvalues is equal to the trace of the Hessian matrix H). This implies that there exists at least one eigenvalue of each sign, and hence the stationary point can only be a saddle point (i.e. there is no local maximum or local minimum) unless all $\lambda_i = 0$, in which case this is inconclusive. This standard analysis is consistent with, but not as powerful as, the maximum principle. Everything works analogously for minima instead of maxima.

11.3 Integral Solution of Poisson's Equation

11.3.1 Statement and informal derivation

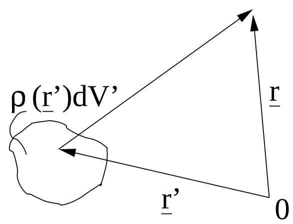
Consider the potential for a point source of strength λ at $\underline{a}$:

$$\psi = \frac{1}{4\pi}\frac{\lambda}{|\underline{r} - \underline{a}|}.$$

$\lambda = -4\pi GM$ for a point mass, or $\lambda = Q/\epsilon_0$ for a point charge, as in section 10. For several sources λ_α at positions $\underline{r}_\alpha$, the potential is a sum of terms

$$\varphi(\underline{r}) = \sum_\alpha \frac{1}{4\pi}\frac{\lambda_\alpha}{|\underline{r} - \underline{r}_\alpha|}.$$

Now consider a general distribution of sources $\rho(\underline{r})$ with $\rho(\underline{r}')dV'$ being the contribution from any small volume at position $\underline{r}'$.



To sum over all such contributions, we replace $\sum_\alpha$ by $\int_{V'} dV'$ in the formula above, which suggests that

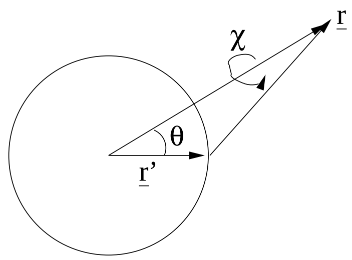
$$\varphi(\underline{r}) = \frac{1}{4\pi}\int_{V'}\frac{\rho(\underline{r}')}{|\underline{r} - \underline{r}'|}dV' \tag{11.2}$$

is the solution to Poisson's equation $\nabla^2\varphi = -\rho$ and is the integral with respect to $\underline{r}'$ taken over the region V' on which $\rho(\underline{r}')$ is non-zero ($\rho = 0$ outside V'). This is indeed <u>the</u> solution with boundary conditions $|\varphi(\underline{r})| = \mathcal{O}(1/|\underline{r}|)$ and $|\underline{\nabla}\varphi(\underline{r})| = \mathcal{O}(1/|\underline{r}|^2)$. For $\rho(\underline{r}')$ non-zero but suitably well-behaved as $|\underline{r}'| \to \infty$ we can also take $V' = \Re^3$, the integral over all space.

Example: Using Eq. 11.2, we can calculate $\varphi(\underline{r})$, the solution to:

$$\nabla^2\varphi = \begin{cases} -\rho_0 & |\underline{r}| \le a \\ \quad 0 & |\underline{r}| > a \end{cases} .$$

We fix $\underline{r}$ and introduce polar coordinates r', θ, χ for $\underline{r}'$ with respect to this fixed direction ($\underline{r}$). Then, $\varphi(\underline{r}) = \frac{1}{4\pi}\int_{V'} \frac{\rho_0}{|\underline{r}-\underline{r}'|}\, dV'$, where $dV' = r'^2 \sin\theta\ d\underline{r}'\ d\theta\ d\chi$, V' : $r' = |\underline{r}'| \le a$ and $|\underline{r} - \underline{r}'| = (r^2 + r'^2 - 2rr'\cos\theta)^{1/2}$ so the integrand is independent of χ.



Therefore,

$$\begin{aligned} \varphi(\underline{r}) &= \frac{1}{4\pi}\int_0^a dr' \int_0^\pi d\theta \int_0^{2\pi} d\chi \frac{\rho_0 r'^2 \sin\theta}{\sqrt{r'^2 + r^2 - 2rr'\cos\theta}} \\ &= \rho_0/2 \int_0^a dr' r'^2 \frac{1}{rr'} \left[\sqrt{r'^2 + r^2 - 2rr'\cos\theta}\right]_{\theta=0}^{\pi} . \end{aligned}$$

$$\left[\sqrt{r'^2 + r^2 - 2rr'\cos\theta}\right]_{\theta=0}^{\pi} = |\underline{r} + \underline{r}'| - |\underline{r} - \underline{r}'| = \begin{cases} 2r' & \text{for } r > r' \\ 2r & \text{for } r < r' \end{cases} ,$$

so

$$\varphi(\underline{r}) = \begin{cases} \rho_0 \int_0^a dr' \frac{r'^2}{r} & = \rho_0 \frac{1}{3} a^3 \frac{1}{r} & (r > a) \\ \rho_0 \left(\int_0^r dr' \frac{r'^2}{r} + \int_r^a dr' r'\right) & = \rho_0 \left(-\frac{1}{6}r^2 + \frac{1}{2}a^2\right) & (r \le a) \end{cases} ,$$

the solution we obtained previously in section 10.3.

11.3.2 Point sources and $\delta-$functions (non-examinable)

We return to $\psi = \lambda/(4\pi|\underline{r} - \underline{a}|)$, a potential for a point source, and note that it obeys $\underline{\nabla}\psi = -\lambda/(4\pi) \times (\underline{r} - \underline{a})/|\underline{r} - \underline{a}|^3$ and $\nabla^2\psi = 0$ for $\underline{r} \ne \underline{a}$. The function ψ is singular when $\underline{r} = \underline{a}$ but the singularity can be given a precise interpretation by considering $\int_S \underline{\nabla}\psi \cdot d\underline{S} = -\lambda$ for any sphere with centre $\underline{a}$, by explicit calculation. Now we compare with $\int_V \nabla^2\psi\ dV$ for V a solid sphere with $\partial V = S$. The divergence theorem holds if we take $\nabla^2\psi = -\lambda\delta(\underline{r}-\underline{a})$ where[6] the three-dimension delta function is a generalised function (or distribution) satisfying $\int_V f(\underline{r})\delta(\underline{r}-\underline{a})\, dV = f(\underline{a})$ for any volume V containing $\underline{a}$. Using this, we can verify the integral solutions of Poisson's equations immediately:

$$\nabla^2\psi(\underline{r}) = \frac{1}{4\pi}\int_{V'} \rho(\underline{r}')\nabla^2\frac{1}{|\underline{r}-\underline{r}'|}dV' = -\int_{V'} \rho(\underline{r}')\delta(\underline{r}-\underline{r}')dV' = -\rho(\underline{r}),$$

as required.

12 Maxwell's Equations

12.1 Laws of Electromagnetism

Static electric charges interact in a way that is mathematically almost identical to gravitational interactions. More generally, moving electric charges interact in a way that can be described by electric and magnetic fields $\underline{E}(\underline{r}, t)$ and $\underline{B}(\underline{r}, t)$.

If $\rho(\underline{r}, t)$ is the charge density and $\underline{j}(\underline{r}, t)$ the current density, then the resulting electric and magnetic fields are determined by Maxwell's equations:

$$\begin{aligned} \underline{\nabla}\cdot\underline{E} &= \tfrac{\rho}{\epsilon_0} \qquad & \underline{\nabla}\wedge\underline{E} + \frac{\partial\underline{B}}{\partial t} &= 0 \\ \underline{\nabla}\cdot\underline{B} &= 0 \qquad & \underline{\nabla}\wedge\underline{B} - \mu_0\epsilon_0\frac{\partial\underline{E}}{\partial t} &= \mu_0\underline{j}, \end{aligned}$$

where ϵ_0 is the permittivity of free space and μ_0 is the permeability of free space (they are both physical constants). Then, the Lorentz Force on a point charge q due to these fields is

$$\underline{F} = q(\underline{E} + \dot{\underline{r}} \wedge \underline{B}).$$

In principle, this gives a complete description of all classical electromagnetic phenomena.

Conservation of electric charge comes from Maxwell's equations:

$$\frac{\partial\rho}{\partial t} + \underline{\nabla}\cdot\underline{j} = 0.$$

Gauss' Law for $\underline{E}$ is unmodified, and now we have a similar law for $\underline{B}$ too (in electrostatics), but with zero 'magnetic charge'. In integral form:

$$\int_S \underline{E}\cdot d\underline{S} = Q/\epsilon_0, \qquad \int_S \underline{B}\cdot d\underline{S} = 0.$$

The remaining Maxwell's equations also have integral forms, for example $\int_C \underline{E}\cdot d\underline{r} = -\frac{d}{dt}\int_S \underline{B}\cdot d\underline{S}$ for any surface S with a boundary of a closed curve C, from Stokes' theorem.

[6] The expression for $\nabla^2\psi$ above makes sense intuitively for a point source concentrated at $\underline{a}$.

12.2 Static Charges and Steady Currents

If $\rho, \underline{j}, \underline{E}$ and $\underline{B}$ are all independent of time, then $\underline{E}$ and $\underline{B}$ are no longer linked, since Maxwell's equations then imply:

Electrostatics	*Magnetostatics*
$\underline{\nabla} \cdot \underline{E} = \rho/\epsilon_0$	$\underline{\nabla} \cdot \underline{B} = 0$
$\underline{\nabla} \wedge \underline{E} = \underline{0}$	$\underline{\nabla} \wedge \underline{B} = \mu_0 \underline{j}$
$\Rightarrow \underline{E} = -\underline{\nabla}\varphi$ and $\nabla^2\varphi = -\rho/\epsilon_0$, the scalar Poisson equation described in section 11.3.	$\Rightarrow \underline{B} = \underline{\nabla} \wedge \underline{A}$, but the vector potential is ambiguous: $\underline{A} \to \underline{A} + \underline{\nabla}\chi$ produces the same $\underline{B}$ field. The scalar field χ can be chosen to make $\underline{\nabla} \cdot \underline{A} = 0 \Rightarrow \nabla^2 \underline{A} = -\mu_0 \underline{j}$, the vector Poisson equation: it has a similar general solution to the scalar case.
ϵ_0 sets the scale of electrostatic effects, for example the Coulomb force.	μ_0 sets the scale of magnetic effects, for example the force between current-carrying wires.

12.3 Electromagnetic Waves

Consider Maxwell's equations with $\rho = 0$ and $\underline{j} = \underline{0}$ (i.e. empty space):

$$\nabla^2 \underline{E} = \underline{\nabla}(\underline{\nabla} \cdot \underline{E}) - \underline{\nabla} \wedge (\underline{\nabla} \wedge \underline{E}) = \underline{\nabla} \wedge \frac{\partial \underline{B}}{\partial t} = \frac{\partial}{\partial t}(\underline{\nabla} \wedge \underline{B}) = \frac{\partial}{\partial t}(\mu_0 \epsilon_0 \frac{\partial \underline{E}}{\partial t}) = \mu_0 \epsilon_0 \frac{\partial^2 \underline{E}}{\partial t^2},$$

or

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\underline{E} = 0, \text{ and similarly we can obtain } \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\underline{B} = 0,$$

where $c^2 = 1/(\epsilon_0 \mu_0)$. These are wave equations in three dimensions describing propagation with speed c. We compare with the one dimensional wave equation for a scalar

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) f = 0, \qquad f = f_\pm(x \mp ct)$$

where f_+ describes a right moving wave and f_- a left moving wave.

So, Maxwell's equations predict electromagnetic waves which move with speed $c = 1/\sqrt{\epsilon_0 \mu_0} \approx 2.998 \times 10^8$ ms^{-1}, i.e. the experimentally measured values of μ_0 and ϵ_0 combine to give the experimental value of the speed of light.

Maxwell famously said "*We can scarcely avoid the inference that light consists of the transverse modulation of the same medium which is the cause of electric and magnetic phenomena*". This was the unification of the theories of electricity, magnetism and light.

13 Tensors and Tensor Fields

13.1 Definitions and Examples

13.1.1 Tensor transformation rule

Consider orthonormal right-handed bases $\{\underline{e}_i\}$ and $\{\underline{e}_i'\}$ in $\Re^3$ with corresponding Cartesian coordinates $\{x_i\}$ and $\{x_i'\}$: $\underline{x} = x_i\underline{e}_i = x_i'\underline{e}_i{}'$. Any such bases are related by a <u>rotation</u>: $\underline{e}_i' = R_{ip}\underline{e}_p$, $x_i' = R_{ip}x_p$ where

1. $R_{ip}R_{jp} = R_{qi}R_{qj} = \delta_{ij}$ or $RR^T = R^TR = I$, in other words R is <u>orthogonal</u>.

2. $\det R = 1$, or R is <u>special</u> orthogonal.

The components of a vector $\underline{v}$ transform in a similar way under this change of basis and coordinates: $\underline{v} = v_i\underline{e}_i = v_i'\underline{e}_i' \Rightarrow v_i' = R_{ip}v_p$.

Tensors are geometric objects which obey a generalised form of this transformation rule. By definition, a <u>tensor</u> T of rank n has components $T_{ij\ldots k}$ (n indices) with respect to each basis $\{\underline{e}_i\}$ or coordinate system $\{x_i\}$ and under a change of basis/coordinates

$$T'_{ij\ldots k} = R_{ip}R_{jq}\ldots R_{kr}T_{pq\ldots r}.$$

This is the <u>tensor transformation rule</u>.

Examples:	T of rank 0, i.e. no indices	$T = T'$, a scalar
	T of rank 1,	$T_i' = R_{ip}T_p$, a vector
	T of rank 2,	$T_{ij}' = R_{ip}R_{jq}T_{pq}$ form a matrix.

13.1.2 Basic examples

1. If $\underline{u}, \underline{v}, \ldots, \underline{w}$ are n vectors, then $T_{ij\ldots k} = u_iv_j\ldots w_k$ defines a tensor of rank n. To show this, we check the transformation rule. For example $n = 2$, $T_{ij} = u_iv_j$.

$$T_{ij}' = u_i'v_j' = (R_{ip}u_p)(R_{jq}v_q) = R_{ip}R_{jq}(u_pv_q) = R_{ip}R_{jq}T_{pq},$$

   i.e the correct rule. Any linear combination of such expressions are also tensors, for example $T_{ij} = u_iv_j + a_ib_j$ for any $\underline{u}, \underline{v}, \underline{a}, \underline{b}$.

2. δ_{ij} and ϵ_{ijk} are tensors of rank 2 and 3 respectively, with the special property that their components are unchanged with respect to any basis/coordinate system: $\delta_{ij}' = R_{ip}R_{jq}\delta_{pq} = R_{ip}R_{jp} = \delta_{ij}$ which is the tensor transformation rule with $\delta_{ij}' = \delta_{ij}$ and $R_{ip}R_{jq}R_{kr}\epsilon_{pqr} = (\det R)\epsilon_{ijk} = \epsilon_{ijk}$ which[7] is the tensor transformation rule with $\epsilon_{ijk}' = \epsilon_{ijk}$.

[7]Note that here, we used a general result from Vectors and Matrices IA, namely $M_{ip}M_{jq}M_{kr}\epsilon_{pqr} = (\det M)\epsilon_{ijk}$.

13.1.3 Rank 2 tensors and matrices

A linear map between vectors $\underline{u}$ and $\underline{v}$ has the general form $v_i = M_{ij}u_j$ or $v'_i = M'_{ij}u'_j$ with respect to different choices of basis/coordinate system. Using the transformation rule for vectors: $v'_i = R_{ip}v_p = R_{ip}(M_{pq}u_q) = M'_{ij}u'_j = M'_{ij}(R_{jq}u_q)$. These expressions are compatible for any $\underline{u}$ iff $R_{ip}M_{pq} = M'_{ij}R_{jq}$ or $M'_{ik} = R_{ip}R_{kq}M_{pq}$; this is the tensor transformation rule for rank 2. In matrix notation $M' = RMR^T$, which is the standard transformation for the matrix of a linear map under an orthogonal change of basis ($R^T = R^{-1}$).

Example: In many substances, an applied electric field $\underline{E}$ induces an electric current density $\underline{j}$ according to the linear relation (Ohm's Law) $j_i = \sigma_{ik}E_k$, where σ_{ik} is the conductivity tensor (it is automatically a tensor, by the argument given above). Substances may conduct differently in different directions (for example if they have a layered structure) and then $\underline{j}$ is not parallel to $\underline{E}$ in general. For an isotropic substance, $\sigma_{ik} = \sigma\delta_{ik} \Rightarrow \underline{j} = \sigma\underline{E}$.

13.2 Tensor Algebra

13.2.1 Addition and scalar multiplication

Tensors T and S of the same rank can be added; $T + S$ is also a tensor of rank n defined by $(T + S)_{ij\ldots k} = T_{ij\ldots k} + S_{ij\ldots k}$ in any coordinate system. To show that this is a tensor, we check the transformation rule, for example for $n = 2$:

$$(T+S)'_{ij} = T'_{ij}+S'_{ij} = R_{ip}R_{jq}T_{pq}+R_{ip}R_{jq}S_{pq} = R_{ip}R_{jq}(T_{pq}+S_{pq}) = R_{ip}R_{jq}(T+S)_{pq}.$$

A tensor T of any rank can be multiplied by a scalar α; αT is a tensor of the same rank, defined by $\alpha T_{ij\ldots k} = \alpha T_{ij\ldots k}$ in any coordinate system (it's easy to check the transformation rule).

13.2.2 Tensor products

If T and S are tensors of rank n and m then the tensor product $T \otimes S$ is a tensor of rank $n + m$ defined by

$$(T \otimes S)_{\underbrace{ij\ldots k}_{n \text{ indices}}\underbrace{pq\ldots r}_{m \text{ indices}}} = T_{ij\ldots k}S_{pq\ldots r}.$$

This is a tensor (check the transformation rule yourself). The tensor product can be defined similarly for any number of tensors, for example for n vectors $\underline{u}, \underline{v}, \ldots \underline{w}$ we define $T = \underline{u} \otimes \underline{v} \cdots \otimes \underline{w}$ by $T_{ij\ldots k} = u_i v_j \cdots w_k$ as in section 13.1.

13.2.3 Contractions

Given a rank n tensor T with components $T_{ijp\ldots q}$, we define a new tensor S of rank $n - 2$ by $S_{p\ldots q} = \delta_{ij}T_{ijp\ldots q} = T_{iip\ldots q}$. This is called contracting on the indices i and

j. Contracting on a different pair of indices results in a different tensor, in general. To check that contraction produces a tensor, we take the example of T_{ij} of rank 2. Contracting T_{ii} of rank 0 (i.e. a scalar), $T'_{ii} = R_{ip}R_{iq}T_{pq} = \delta_{pq}T_{pq} = T_{pp}$. Higher rank tensors involve additional indices that are unaffected by the contraction. Note that with T_{ij} regarded as a 3 by 3 matrix, T_{ii} is the trace. $T'_{ii} = \mathrm{Tr}(T') = \mathrm{Tr}(RTR^T) = \mathrm{Tr}(T) = T_{pp}$, in matrix language.

13.2.4 Symmetric and antisymmetric tensors

A tensor T of rank n obeying $T_{ijp\ldots q} = \pm T_{jip\ldots q}$ is said to be symmetric (+) or anti-symmetric (-) in the indices i and j. (There exists a similar definition for any other index pair). This holds in all coordinate systems if it holds in any one, since

$$T'_{klr\ldots s} = R_{ki}R_{lj}R_{rp}\ldots R_{sq}T_{ijp\ldots q} = \pm R_{ki}R_{lj}R_{rp}\ldots R_{sq}T_{jip\ldots q} = \pm T'_{lkr\ldots s}.$$

A tensor is called totally symmetric/anti-symmetric if it is symmetric/anti-symmetric in every pair of indices. For example $\delta_{ij} = \delta_{ji}$ is totally symmetric and ϵ_{ijk} is totally anti-symmetric. There are totally symmetric tensors of arbitrary rank n, but in $\Re^3$:

1. Any totally antisymmetric tensor of rank 3 has the form $T_{ijk} = \lambda\epsilon_{ijk}$ for some scalar λ.

2. There are no totally antisymmetric tensors of rank $n > 3$ except (trivially) the tensors with all components zero.

13.3 Tensors, Multi-linear Maps and the Quotient Rule

13.3.1 Tensors as multi-linear maps

A tensor T of rank n is equivalent to a multi-linear map from n vectors $\underline{a}, \underline{b}, \ldots \underline{c}$ to $\Re$ defined by

$$T(\underline{a}, \underline{b}, \ldots, \underline{c}) = \overbrace{T_{ij\ldots k}a_i b_j \ldots c_k}^{(1)} = \overbrace{T'_{pq\ldots r}a'_p b'_q \ldots c'_r}^{(2)}.$$

'Multi-linear' means that T is linear in each of the vectors $\underline{a}, \underline{b}, \ldots, \underline{c}$ individually. Note that

1. If the first expression (1) is specified for all vectors $\underline{a}, \underline{b}, \ldots, \underline{c}$ then the components of the tensor are uniquely determined (by taking $\underline{a}, \underline{b}, \ldots, \underline{c}$ to be all possible choices of basis vectors).

2. Using the standard relationship for vector components $v'_p = R_{pi}v_i$ or $v_i = R_{pi}v'_p$, (1) and (2) agree iff $T_{ij\ldots k}R_{pi}a'_p R_{qj}b'_q \ldots R_{rk}c'_r = T'_{pq\ldots r}a'_p b'_q \ldots c'_r$, which holds for any $a'_p b'_q \ldots c'_r$ iff $T'_{pq\ldots r} = R_{pi}R_{qj}\ldots R_{rk}T_{ij\ldots k}$, ie the transformation rule. Thus, there is a one-to-one correspondence between tensors of rank n and multi-linear maps. This gives a way of thinking about tensors' independent of any coordinate system or choice of basis, and the tensor transformation rule emerges naturally.

13.3.2 The quotient rule

If $T_{\underbrace{i\ldots j}_{n}\underbrace{p\ldots q}_{m}}$ is a tensor of rank $n+m$ and $u_{p\ldots q}$ is any tensor of rank m, then

$$v_{i\ldots j} = T_{i\ldots jp\ldots q}u_{p\ldots q}$$

is a tensor of rank n (it is a tensor product of T and u, followed by contractions).

Conversely, suppose that $T_{i\ldots jp\ldots q}$ is an array defined for each coordinate system, and that for any tensor $u_{p\ldots q}$, $v_{i\ldots j}$ as defined above is also a tensor, then $T_{i\ldots jp\ldots q}$ is also a tensor.

This is the quotient rule. We have seen and proved the special case $n=m=1$ in section 13.1.3. The general case can be proved in a similar way, checking the transformation rule directly, but the proliferation of indices makes the argument look more complicated than it really is.

To streamline things, we shall use the ideas in section 13.3.1. It is sufficient to take the special form $u_{p\ldots q} = c_p \ldots d_q$ for any vectors $\underline{c}, \ldots, \underline{d}$; by assumption $v_{i\ldots j} = T_{i\ldots jp\ldots q}c_p \ldots d_q$ is a tensor, but then $v_{i\ldots j}a_i \ldots b_j = T_{i\ldots jp\ldots q}a_i \ldots b_j c_p \ldots d_q$ is a scalar for any vectors $\underline{a}, \ldots, \underline{b}, \underline{c}, \ldots \underline{d}$. This is enough to ensure the tensor transformation rule for the components $T_{i\ldots jp\ldots q}$, as observed in section 13.3.1 above.

13.4 Tensor Calculus

13.4.1 Tensor fields and derivatives

Just as with scalars or vectors, a tensor field is a tensor at each point $T_{ij\ldots k}(\underline{x})$, which we also write as[8] $T_{ij\ldots k}(x_l)$. We assume that fields are smooth so that they can be differentiated any number of times

$$\underbrace{\frac{\partial}{\partial x_p}\cdots\frac{\partial}{\partial x_q}}_{m} T_{\underbrace{ij\ldots k}_{n}},$$

except for where things obviously fail, for example where T is not defined. This is a new tensor of rank $n+m$.

To verify the transformation rule, consider the derivatives: $x'_i = R_{ip}x_p \Rightarrow \frac{\partial x'_i}{\partial x_p} = R_{ip}$, and $x_q = (R^{-1})_{qi}x'_i = R_{iq}x'_i \Rightarrow \frac{\partial x_q}{\partial x'_i} = R_{iq}$. Note that R_{ip} and R_{iq} are constant matrices. Now, by the chain rule

$$\frac{\partial}{\partial x'_i} = \frac{\partial x_q}{\partial x'_i}\frac{\partial}{\partial x_q} = R_{iq}\frac{\partial}{\partial x_q}.$$

This is expected for components of a vector, and as emphasised in section 9.2, $\underline{\nabla} = \underline{e}_i\partial/\partial x_i = \underline{e}'_i\partial/\partial x'_i$ is a vector operator. Using this, it is easy to check the

[8]Meaning that it can depend upon all of the coordinates

transformation rule. The definition of tensor derivatives involves comparing tensors at different points, but otherwise one just uses linear combinations and then limits.

For example: derivatives of a scalar field

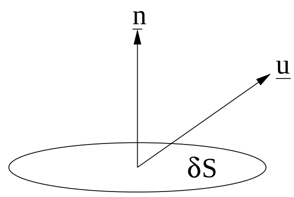
$$\frac{\partial}{\partial x'_i}\frac{\partial}{\partial x'_j}\cdots\frac{\partial}{\partial x'_k}\varphi = \left(R_{ip}\frac{\partial}{\partial x_p}\right)\left(R_{jq}\frac{\partial}{\partial x_q}\right)\cdots\left(R_{kr}\frac{\partial}{\partial x_r}\right)\varphi = R_{ip}R_{jq}\cdots R_{kr}\frac{\partial}{\partial x_p}\frac{\partial}{\partial x_q}\cdots\frac{\partial}{\partial x_r}\varphi$$

which is the required transformation rule (all for Cartesian coordinates related by a constant R). For just one derivative, we have $(\underline{\nabla}\varphi)_i = \frac{\partial\varphi}{\partial x_i}$, a vector.

13.4.2 Integrals and the tensor divergence theorem

Because we can add tensors and take limits (by assumption) the definition of a tensor valued integral is straight forward. For example, $\int_V T_{ij\ldots k}(\underline{x})\ dV$ and the result is a tensor (it is easiest to think of the integral as the limit of a sum).

For a physical example, recall our discussion of the flux of quantities for a fluid with velocity field $\underline{u}(\underline{x})$ through a small surface element (assuming a uniform density ρ)



The flux of volume is $\underline{u}\cdot\underline{n}\delta S = u_j n_j \delta S$, the flux of mass is $\rho\underline{u}\cdot\underline{n}\delta S = \rho u_j n_j \delta S$, whereas the flux of the i^{th} component of momentum is $\rho u_i u_j n_j \delta S = T_{ij} n_j \delta S$, involving $T_{ij} = \rho u_i u_j$. The flux through S is $\int_S T_{ij} n_j dS$.

It is also easy to generalise the divergence theorem from vectors to tensors (and then use it to discuss conservation laws for vector and tensor quantities). Let V be a volume bounded by a smooth surface $S = \partial V$ and $T_{ij\ldots k}$ be a smooth tensor field. Then

$$\int_S T_{ij\ldots kl} n_l\ dS = \int_V \frac{\partial}{\partial x_l}\left(T_{ij\ldots kl}\right)\ dV,$$

where n_l is an outward pointing unit normal.

Proof: apply the usual Divergence Theorem to the vector field $\underline{v}$ defined by $v_l = a_i b_j \cdots c_k T_{ij\ldots kl}$ where $\underline{a}, \underline{b}, \ldots, \underline{c}$ are fixed constant vectors;

$$\begin{aligned}\underline{\nabla}\cdot\underline{v} &= \frac{\partial v_l}{\partial x_l} = a_i b_j \cdots c_k \frac{\partial}{\partial x_l} T_{ij\ldots kl},\\ \underline{n}\cdot\underline{v} &= n_l v_l = a_i b_j \cdots c_k T_{ij\ldots kl} n_l,\end{aligned} \tag{13.1}$$

and we obtain the tensor divergence theorem, but contracted with $a_i b_j \cdots c_l$ on each side. Since $\underline{a}, \underline{b}, \cdots, \underline{c}$ are arbitrary, they can be cancelled off and the result follows.

14 Tensors of Rank 2

14.1 Decomposition of a Second Rank Tensor

Any second rank tensor can be written as a sum of its symmetric and anti-symmetric parts: $T_{ij} = S_{ij} + A_{ij}$ has 9 independent components (IC), where $S_{ij} = \frac{1}{2}(T_{ij} + T_{ji})$ has 6 IC and $A_{ij} = \frac{1}{2}(T_{ij} - T_{ji})$ has 3 IC. In fact, the symmetric part can be reduced further: $S_{ij} = P_{ij} + \frac{1}{3}\delta_{ij}Q$ where $P_{ii} = 0$, or P_{ij} is *traceless*, and $Q = S_{ii}$ is the *trace* of S_{ij}. P_{ij} has 5 IC whereas Q has 1. The antisymmetric part can be re-expressed:

$$A_{ij} = \epsilon_{ijk}B_k \Leftrightarrow B_k = \frac{1}{2}\epsilon_{ijk}A_{ij}$$

from the double-ϵ identities, or

$$(A_{ij}) = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}.$$

To summarise,

$$T_{ij} = \underbrace{P_{ij}}_{\text{symmetric traceless}} + \epsilon_{ijk}\underbrace{B_k}_{\text{vector}} + \frac{1}{3}\delta_{ij}\underbrace{Q}_{\text{scalar}}$$

where $B_k = \frac{1}{2}\epsilon_{ijk}T_{ij}$, and only the antisymmetric parts of T contribute. $Q_k = T_{kk}$ where only symmetric parts of T contribute.

Example: A vector field $F_i(\underline{r})$ derivative $T_{ij} = \frac{\partial F_i}{\partial x_j}$, a tensor field. The decomposition given above has a symmetric traceless piece $P_{ij} = \frac{1}{2}(\frac{\partial F_i}{\partial x_j} + \frac{\partial F_j}{\partial x_i}) - \frac{1}{3}\delta_{ij}\frac{\partial F_k}{\partial x_k}$, an anti-symmetric piece $A_{ij} = \epsilon_{ijk}B_k$ with $B_k = \frac{1}{2}\epsilon_{ijk}\frac{\partial F_i}{\partial x_j} = -\frac{1}{2}(\underline{\nabla} \wedge \underline{F})_k$ and a trace $Q = \frac{\partial F_k}{\partial x_k} = \underline{\nabla} \cdot \underline{F}$. Hence, a complete description of the derivative of a vector field $\underline{F}$ involves a scalar $\underline{\nabla} \cdot \underline{F}$, a vector $\underline{\nabla} \wedge \underline{F}$ and a symmetric traceless tensor P_{ij}.

14.2 The Inertia Tensor

Consider some masses m_α with positions $\underline{r}_\alpha$, all rotating with angular velocity $\underline{\omega}$ about $\underline{0}$, so the velocities are $\underline{v}_\alpha = \underline{\omega} \wedge \underline{r}_\alpha$. The total angular momentum about $\underline{0}$ is

$$\underline{L} = \sum_\alpha \underline{r}_\alpha \wedge m_\alpha \underline{v}_\alpha = \sum_\alpha m_\alpha \underline{r}_\alpha \wedge (\underline{\omega} \wedge \underline{r}_\alpha) = \sum_\alpha m_\alpha \left(|\underline{r}_\alpha|^2 \underline{\omega} - \underline{r}_\alpha(\underline{r}_\alpha \cdot \underline{\omega})\right).$$

In components, $L_i = I_{ij}\omega_j$ where

$$I_{ij} = \sum_\alpha m_\alpha \left(|\underline{r}_\alpha|^2 \delta_{ij} - (\underline{r}_\alpha)_i (\underline{r}_\alpha)_j\right)$$

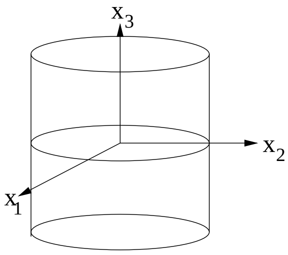
is the <u>inertia tensor</u> about $\underline{0}$. Note that in general $\underline{L}$ is not parallel to $\underline{\omega}$.

For a rigid body occupying volume V with mass density $\rho(\underline{r})$, we replace $\sum_\alpha m_\alpha \to \int_V \rho(\underline{r})\ dV$ above to obtain

$$I_{ij} = \int_V \rho(\underline{r})\,(x_k x_k \delta_{ij} - x_i x_j)\ dV.$$

For example, $I_{33} = \int_V \rho(\underline{r})(x_1^2 + x_2^2)\ dV$ and $I_{12} = -\int_V \rho(\underline{r}) x_1 x_2\ dV$. I_{ij} is a symmetric tensor (by inspection).

Example:



The inertia tensor for a cylinder of radius a and height $2l$, uniform density ρ_0 and mass $M = 2\pi a^2 l \rho_0$. We choose coordinates $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, $dV = r\ dr\ d\theta\ dx_3$, with x_3 being the axis of symmetry.

Then

$$\begin{aligned}
I_{33} &= \int_V \rho_0(x_1^2 + x_2^2)\ dV = \rho_0 \int_0^a dr \int_0^{2\pi} d\theta \int_{-l}^{l} dx_3 r r^2 = \rho_0 \pi l a^4, \\
I_{11} &= \int_V \rho_0(x_2^2 + x_3^2)\ dV = \rho_0 \int_0^a dr \int_0^{2\pi} d\theta \int_{-l}^{l} dx_3 (r^2 \sin^2\theta + x_3^2) r \\
&= \rho_0(\frac{1}{4}a^4 2l\pi + \frac{1}{2}a^2 2\pi \frac{2}{3}l^3) = \rho_0 \pi a^2 l \left(\frac{a^2}{2} + \frac{2}{3}l^2\right).
\end{aligned}$$

We get the same result for I_{22} (which has $\sin^2\theta \to \cos^2\theta$). Also,

$$I_{13} = -\int_V \rho_0 x_1 x_3\ dV = -\rho_0 \int_0^a dr \int_0^{2\pi} d\theta \int_{-l}^{l} dx_3 r^2 \cos\theta x_3 = 0$$

from the θ integration. $I_{12} = I_{13} = 0$, similarly.
In summary: the non-zero components are $I_{33} = Ma^2/2$, $I_{11} = I_{22} = M(\frac{1}{4}a^2 + \frac{1}{3}l^2)$.
Note that for the special case $l = \sqrt{3}a/2$, $I_{ij} = \frac{1}{2}Ma^2\delta_{ij} \Rightarrow \underline{L} = \frac{1}{2}Ma^2\underline{\omega}$ for a rotation about any axis.

14.3 Diagonalisation of a Symmetric Second Rank Tensor

Recall, using matrix notation, $T = (T_{ij})$, $T' = (T'_{ij})$, $R = R_{ij}$, that the tensor transformation rule $T'_{ij} = R_{ip} R_{jq} T_{pq}$ becomes $T' = RTR^T$, where $R^T = R^{-1}$.

If T is symmetric, it can be diagonalised by such an orthogonal transformation. Equivalently, there exists a basis of orthonormal eigenvectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ for T with real eigenvalues $\lambda_1, \lambda_2, \lambda_3$, respectively. The directions defined by $\underline{e}_1, \underline{e}_2, \underline{e}_3$ are the principal axes for T; the tensor is diagonal in Cartesian coordinates along these axes.

This applies to any symmetric tensor. For the special case of the inertia tensor, the eigenvalues are called the principal moments of inertia. As exemplified in section 14.2, we can often guess the correct principal axes for I_{ij} based on the symmetries of the body: with the axes we chose, I_{ij} was found to be diagonal, by direct calculation.

15 Invariant and Isotropic Tensors

15.1 Definitions and Classification Results

A tensor T is invariant under a particular rotation R if

$$T'_{ij\ldots k} = R_{ip}R_{jq}\cdots R_{kr}T_{pq\cdots r} = T_{ij\cdots k},$$

i.e. each component is unchanged. A tensor T which is invariant under every rotation is called isotropic (i.e. the same in all directions). As noted in section 13.1.2, δ_{ij} and ϵ_{ijk} are isotropic tensors. This ensures that the component definitions of the scalar and vector products

$$\underline{a}\cdot\underline{b} = \delta_{ij}a_ib_j, \qquad (\underline{a}\wedge\underline{b})_i = \epsilon_{ijk}a_jb_k$$

are independent of the Cartesian coordinate system. Isotropic tensors in $\Re^3$ can be classified:

1. There are no isotropic tensors of rank 1 (vector) except the zero vector.
2. The most general rank 2 isotropic tensor is $T_{ij} = \alpha\delta_{ij}$ for some scalar α.
3. The most general rank 3 isotropic tensor is $T_{ijk} = \beta\epsilon_{ijk}$ for some scalar β.
4. All isotropic tensors of higher rank are obtained by combining δ_{ij} and ϵ_{ijk} using tensor products, contractions and linear combinations.

Example: The most general isotropic tensor of rank 4 is

$$T_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{jl} + \gamma\delta_{il}\delta_{jk}$$

for some scalar constants α, β, γ. There are no other independent combinations (for example, you might think of $\epsilon_{ijp}\epsilon_{klp}$, but this is equal to $\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$).

15.2 Application to Invariant Integrals

Consider a tensor defined by

$$T_{ij\ldots k} = \int_V f(\underline{x})x_ix_j\cdots x_k dV$$

with $f(\underline{x})$ being a scalar function and V is some volume. Given a rotation R_{ij}, consider an <u>active</u> transformation $\underline{x} = x_i \underline{e}_i$ mapped to $\underline{x}' = x'_i \underline{e}_i$ with $x'_i = R_{ij} x_j$ (NB there is no change of basis vectors here) and V is mapped to V'. Suppose that under this active transformation

1. $f(\underline{x}) = f(\underline{x}')$ and
2. $V' = V$.

Then $T_{ij\ldots k}$ is invariant under R_{ij}.

Proof:

$$R_{ip} R_{jq} \cdots R_{kr} T_{pq\ldots r} = \int_V f(\underline{x}) x'_i x'_j \cdots x'_k dV$$

from the definition of T

$$= \int_V f(\underline{x}') x'_i x'_j \cdots x'_k dV$$

from (1)

$$= \int_{V'} f(\underline{x}') x'_i x'_j \cdots x'_k dV'$$

from (2)

$$= T_{ij\ldots k}$$

(x_i or x'_i are dummy variables in the integral).

NB with the change of variables $x'_i = R_{ip} x_p$, $\frac{\partial x'_i}{\partial x_p} = R_{ip}$ (a constant) and $\det(\frac{\partial x'_i}{\partial x_p}) = \det(R_{ip}) = 1$, i.e. the Jacobian=1 or $dV' = dV$. The result is particularly useful if (1) and (2) hold for <u>any</u> rotation R, in which case $T_{ij\ldots k}$ is isotropic.

Example: $T_{ij} = \int_V x_i x_j \; dV$ with V being a solid sphere with $|\underline{r}| \leq a$. The result above applies with $f = 1$ for any rotation R, so T_{ij} is isotropic $\Rightarrow T_{ij} = \alpha \delta_{ij}$, and we need only determine the scalar α. But taking the trace, $3\alpha = T_{ii} = \int_V x_i x_i \; dV = \int_0^a r^2 4\pi r^2 \; dr = \frac{4}{5}\pi a^5$. Hence $T_{ij} = \frac{4}{15}\pi a^5 \delta_{ij}$.
There is a closely related result for the inertia tensor of a sphere of constant density ρ_0 or mass M:

$$I_{ij} = \int_V \rho_0 \left(x_k x_k \delta_{ij} - x_i x_j \right) \; dV = \rho_0 (3-1) \frac{4}{15} \pi a^5 \delta_{ij} = \frac{2}{5} M a^2 \delta_{ij},$$

which is isotropic.

15.3 A Sketch of a Proof of Classification Results for Rank $n \leq 3$

The results follow straight forwardly by analysing conditions for invariance under specific rotations through π or $\pi/2$ about coordinate axes, then confirming that the most general form allowed is actually isotropic.

1. Suppose that T_i is rank 1 isotropic. Consider

$$(R_{ij}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

a rotation through π about the x_3 axis. Then $T_1 = R_{1p}T_p = R_{11}T_1 = -T_1$ and $T_2 = R_{2p}T_p = R_{22}T_2 = -T_2$. Then $T_1 = T_2 = 0$. Then, requiring isotropy of a rotation through π about the x_1 axis too yields $T_3 = 0$, and therefore $T_i = 0$.

2. Suppose that T_{ij} is rank 2 isotropic. Consider

$$(R_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

a rotation through $\pi/2$ clockwise about the x_3 axis: $T_{13} = R_{1p}R_{3q}T_{pq} = R_{12}R_{33}T_{23} = (+1)(+1)T_{23}$. $T_{23} = R_{2p}R_{3q}T_{pq} = R_{21}R_{33}T_{13} = (-1)(+1)T_{13}$. Hence $T_{13} = T_{23} = 0$ and similarly $T_{31} = T_{32} = 0$. $T_{11} = R_{1p}R_{1q}T_{pq} = R_{12}R_{12}T_{22} = (+1)(+1)T_{22}$. By applying $\pi/2$ rotations about the x_1 and x_2 axes in a similar way, we deduce $T_{11} = T_{22} = T_{33} = \alpha$, say and $T_{ij} \forall i \neq j$. Hence $T_{ij} = \alpha\delta_{ij}$.

3. Suppose that T_{ijk} is rank 3 isotropic. Using the rotation by π as in (1) above, $T_{133} = R_{1p}R_{3q}R_{3r}T_{pqr} = (-1)(+1)(+1)T_{133}$ and $T_{111} = R_{1p}R_{1q}R_{1r}T_{pqr} = (-1)(-1)(-1)T_{111}$, with similar results for π rotations about other axes and with other index positions we write $T_{ijk} = 0$ unless i, j, k are all distinct. Now consider a rotation through $\pi/2$ as in (2) above: $T_{123} = R_{1p}R_{2q}R_{3r}T_{pqr} = R_{12}R_{21}R_{33}T_{213} = (+1)(-1)(+1)T_{213} = -T_{213}$. Along with similar results for other index positions and other axes of rotation, T_{ijk} is totally antisymmetric, hence $T_{ijk} = \beta\epsilon_{ijk}$.

Acknowledgements

We thank Prof. R. Josza for proof reading.