Mathematical Tripos: IA Vector Calculus

Contents

0 Introduction

0.1 Schedule

This is a copy from the booklet of schedules. You will observe as we proceed that I have moved various topics around; however, all topics will be covered.

VECTOR CALCULUS (C6) **24 lectures**

This course develops the theory of partial differentiation and the calculus of scalar and vector quantities in two and three dimensions. A sound knowledge of these topics is a ***vital prerequisite*** *for almost all the later courses in applied mathematics and theoretical physics.*

Partial differentiation

Partial derivatives, geometrical interpretation, statement (only) of symmetry of mixed partial derivatives, chain rule. Directional derivatives. The gradient of a real-valued function, its interpretation as normal to level surfaces (examples including the use of cylindrical, spherical *and general orthogonal curvilinear* coordinates). Statement of Taylor series for functions on $\mathbb{R}^n$. Local extrema of real functions, classification using the Hessian matrix. [6]

Curves, line integrals and differentials

Parametrised curves and arc length, tangents and normals to curves in $\mathbb{R}^3$, the radius of curvature. Line integrals, conservative fields. Informal treatment of differentials, exact differentials. [3]

Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$

Surface and volume integrals. Determinants, Jacobians, and change of variables. [4]

Divergence, curl and ∇^2 in Cartesian coordinates, examples; formulae for these operators (statement only) in cylindrical, spherical *and general orthogonal curvilinear* coordinates. Vector derivative identities. The divergence theorem, Green's theorem, Stokes' theorem. Irrotational fields. [6]

Laplace's equation

Laplace's equation in $\mathbb{R}^2$ and $\mathbb{R}^3$. Examples of solutions. Green's (second) theorem. Bounded regions and Dirichlet boundary condition; uniqueness and maximum principle. [3]

Fundamental solutions in $\mathbb{R}^2$ and $\mathbb{R}^3$ with point sources. Poisson's equation and its solution, interpretation in electrostatics and gravity. [2]

0.2 Lectures

- Lectures will start at 10:05 promptly with a summary of the last lecture. Please be on time since it is distracting to have people walking in late.
- I will endeavour to have a 2 minute break in the middle of the lecture for a rest and/or jokes and/or politics and/or paper aeroplanes; students seem to find that the break makes it easier to concentrate throughout the lecture.[1] When I took a vote on this in 1998, one student voted against the break (and for more lecturing time), while $N \gg 1$ students voted in favour of the break.
- If you throw paper aeroplanes please pick them up. I will pick up the first one to stay in the air for 5 seconds.
- I will aim to finish by 10:55, but may continue for upto 2 extra minutes if in the middle of a proof/explanation.
- I will stay around for a few minutes at the front after lectures in order to answer questions.

[1] Having said that, research suggests that within the first 20 minutes I will, at some point, have lost the attention of all of you.

- By all means chat to each other *quietly* if I am unclear, but please do not discuss, say, last night's football results, or who did (or did not) get drunk and/or laid. Such chatting is a distraction.
- I want you to learn. I will do my best to be clear but you *must* read through and understand your notes before the next lecture ... otherwise you will get hopelessly lost. An understanding of your notes will not diffuse into you just because you have carried your notes around for a week ... or put them under your pillow.
- I welcome *constructive* heckling. If I am inaudible, illegible, unclear or just plain wrong then please shout out.
- I aim to avoid the words *trivial*, *easy*, *obvious* and *yes*. Let me know if I fail. I will occasionally use *straightforward* or *similarly to last time*; if it is not, email me (`S.J.Cowley@damtp.cam.ac.uk`) or catch me at the end of the next lecture.
- Sometimes, although I may have confused both you and myself, I will have to plough on as a result of time constraints; however I will clear up any problems at the beginning of the next lecture.
- The course is *applied* rather than *pure* mathematics. Hence do not expect pure mathematical levels of rigour; having said that all the outline/sketch 'proofs' could in principle be tightened up given sufficient time.
- Since the course is *applied* there will be some applied examples (even though it is a *methods* course). If you need motivation remember that the theory of [financial] derivatives is applied mathematics.
- If anyone is colour blind please come and tell me which colour pens you cannot read.
- Finally, I was in your position 25 years ago and nearly gave up the Tripos. If you feel that the course is going over your head, or you are spending more than 12 hours a week on it, come and chat.

0.3 Printed Notes

- Printed notes will be handed out for the course ... so that you can listen to me rather than having to scribble things down. If it is not in the printed notes or on the example sheets it should not be in the exam.
- The notes will only be available in lectures and only *once* for each set of notes.
- I do not keep back-copies (otherwise my office would be an even worse mess) ... from which you may conclude that I will not have copies of last time's notes.
- There will only be approximately as many copies of the notes as there were students at the last lecture. We are going to fell a forest as it is, and I have no desire to be even more environmentally unsound.
- Please do not take copies for your absent friends unless they are *ill*.
- The notes are deliberately *not* available on the WWW; they are an adjunct to lectures and are not meant to be used independently.
- If you do not want to attend lectures then there are a number of excellent textbooks that you can use in place of my notes.
- With one or two exceptions figures/diagrams are deliberately omitted from the notes. I was taught to do this at my teaching course on *How To Lecture* ... the aim being that it might help you to stay awake if you have to write something down from time to time.
- There will be a number of unlectured worked examples in the notes. In the past these were not included because I was worried that students would be unhappy with material in the notes that was not lectured. However, a vote on this in 1999 was was overwhelming in favour of including unlectured worked examples.

- The notes are not stapled together since the photocopier is significantly slower in staple-mode.
- Please email me corrections to the notes and examples sheets (`S.J.Cowley@damtp.cam.ac.uk`).

0.4 Examples Sheets

- There will be four examples sheets. They will be available on the WWW at about the same time as I hand them out (see `http://www.damtp.cam.ac.uk/user/examples/`).
- You should be able to do examples sheets 1/2/3/4 after lectures 7/13/19/24 respectively (actually I hope after lectures 6/12/18/24). Hence I suggest that you do not arrange to have your first supervision before the middle of week 3 of lectures.
- Examples sheet 4 is slightly longer ... I expect most students to tackle this sheet over the Easter vacation.
- There is some repetition on the sheets by design; pianists do scales, athletes do press-ups, mathematicians do algebra/manipulation.
- Your *supervisors* might like to know that I have outline answers for most questions.

0.5 Previous Comments, Comments on Comments and Style of Lectures.

- There was far too much physics in the notes.

  See above.

- I think that students should be told what is the use of what they are learning because in that way they will understand better.

  See above.

- I found that the lecturer's *blitzkrieg* approach to this lecture course meant that I was unable to follow the detailed arguments put across, and hence found it difficult to understand some points which on later examination proved to be relatively simple. ... I thought that there were too many 'sketches' of proofs, which just confused the issue and made it pretty difficult to get an initial grasp of the subject. ... The lecturer would be better teaching us the basics for longer; more time should be added to worked examples and less to some unnecessary further reading. ...

  There were a number of similar comments (and also roughly the same number of balancing comments which are available on request).

  One style of lecturing is to lecture the key points very clearly with the expectation that students will read round the subject for their wider education. Another style is to go faster over the basic stuff, but to include other material

  - *that you ought to see at some point (but you will not unless you read around), and/or*
  - *that motivates a result (e.g. an outline proof) so that the course is more than a series of recipes (since you are mathematicians, indeed* la crème de la crème*, rather than engineers or natural scientists), and/or*
  - *that makes the notes more or less self contained.*

  I favour the second approach, for which it is crucial that you read through and understand your lecture notes before the next lecture. However, I hope to take on board some of the above criticisms and aim to cut out some material.

  Also, some of the comments may have been motivated, at least in part, by a desire that only material necessary to do examination questions is taught. At least in Michaelmas and Lent I have some sympathy with the view that we are here to educate and that examinations are an unfortunate evil which should not determine what is, or is not, taught.

Recall that schedules are minimal for lecturing, and maximal for examining.

- At the time I thought this was a poorly lectured course, but when I came to read the notes thoroughly, I realised that they were coherent and interesting, a good introduction to the subject. ...It turns out that I learnt much more than it seemed at the time.

  The last comment was written after the exams; I fear that often the courses that you have to grapple with at the time are the courses that you learn best.

- I think the bits towards the end: monopoles, solution to Laplaces equation etc, could have been explained better.

  There were lots of similar comments on this part of the course. The lecturer will try to do better.

- Enjoyed it, apart from blatant anti-Trinity under-currents.

  Trinity is big enough to look after itself.

0.6 Books

K.F. Riley, M.P. Hobson and S.J. Bence. *Mathematical Methods for Physicists and Engineers.* Cambridge University Press 1998. *There is an uncanny resemblance between my notes and this book since both used the same source, i.e. previous Cambridge lecture notes. A 'recommended' book; possibly worth buying.*

D.E. Bourne and P.C. Kendall. *Vector Analysis and Cartesian Tensors.* Chapman and Hall 1992. *A 'recommended' book.*

T.M. Apostol. *Calculus.* Wiley 1975. *This is the book I relied on most heavily while writing my notes.*

H. Anton. *Calculus.* Wiley 1995. *Recommended by previous students.*

M.L. Boas. *Mathematical Methods in the Physical Sciences.* Wiley 1983.

E. Kreyszig. *Advanced Engineering Mathematics.* Wiley 1999. *Another of the books I used when writing my notes; however it is not to every student's liking.*

J.E. Marsden and A.J.Tromba. *Vector Calculus.* Freeman 1996.

H.M. Schey. *Div, grad, curl and all that: an informal text on vector calculus.* Norton 1996.

M.R. Spiegel. *Vector Analysis.* McGraw Hill 1974. *Lots of worked examples. I liked this book when I took the course.*

0.7 Revision

- Please reread your *Study Skills in Mathematics* booklet. It's available on the WWW with URL `http://www.maths.cam.ac.uk/undergrad/inductionday/studyskills/text/`.
- Please make sure that you recall the following from IA *Algebra & Geometry*.

0.7.1 Vectors

Like many applied mathematicians we will denote vectors using underline or bold, and be a little carefree with transposes, e.g. in $\mathbb{R}^3$ we will denote a vector $\mathbf{a}$ by

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \tag{0.1a}$$
$$= (a_1, a_2, a_3)\,, \tag{0.1b}$$

where the $\mathbf{e}_j$ $(j = 1, 2, 3)$ define a right-handed orthonormal basis. When using suffix notation we will use a_i to represent $\mathbf{a}$.

Remarks.

1. If $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$ is a *orthonormal* basis for $\mathbb{R}^m$ then the unit vectors $\mathbf{e}_j$ $(j = 1, \ldots, m)$ span $\mathbb{R}^m$ and

$$\mathbf{e}_i.\mathbf{e}_j = \delta_{ij} \quad (i, j = 1, \ldots, m)\,, \tag{0.2}$$

   where δ_{ij} is the *Kronecker delta*.

2. The orthonormal basis $\mathbf{e}_j$ $(j = 1, 2, 3)$ in $\mathbb{R}^3$ is right-handed if

$$\varepsilon_{ijk}\,\mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k \qquad (s.c.)\,, \tag{0.3}$$

where ε_{ijk} is the three-dimensional *alternating tensor* and *s.c.* stands for the *summation convention.*

3. Alternative notation for a right-handed orthonormal basis in $\mathbb{R}^3$ includes

$$\mathbf{e}_1 = \mathbf{e}_x = \mathbf{i}\,, \quad \mathbf{e}_2 = \mathbf{e}_y = \mathbf{j} \quad \text{and} \quad \mathbf{e}_3 = \mathbf{e}_z = \mathbf{k}\,, \tag{0.4}$$

for the unit vectors in the x, y and z directions respectively. Hence from (0.2) and (0.3)

$$\mathbf{i}.\mathbf{i} = \mathbf{j}.\mathbf{j} = \mathbf{k}.\mathbf{k} = 1\,, \quad \mathbf{i}.\mathbf{j} = \mathbf{j}.\mathbf{k} = \mathbf{k}.\mathbf{i} = 0\,, \tag{0.5a}$$
$$\mathbf{i}\times\mathbf{j} = \mathbf{k}\,, \quad \mathbf{j}\times\mathbf{k} = \mathbf{i}\,, \quad \mathbf{k}\times\mathbf{i} = \mathbf{j}\,. \tag{0.5b}$$

4. Pure mathematicians generally do not use underline or bold for vectors, but often are more conscientious about keeping track of the transpose, e.g.

$$a = (a_1, a_2, a_3)^t\,.$$

Addition and multiplication in $\mathbb{R}^3$*.* Vectors add, and are multiplied by a scalar, according to

$$\lambda\mathbf{a} + \mu\mathbf{b} = (\lambda a_1 + \mu b_1, \lambda a_2 + \mu b_2, \lambda a_3 + \mu b_3) \quad (\lambda, \mu \in \mathbb{R})\,. \tag{0.6}$$

The scalar and vector products are defined by

$$\begin{aligned}
\mathbf{a}.\mathbf{b} = \mathbf{b}.\mathbf{a} &= a_1b_1 + a_2b_2 + a_3b_3\,, && (0.7a)\\
\mathbf{a}\wedge\mathbf{b} \equiv \mathbf{a}\times\mathbf{b} &= -\mathbf{b}\times\mathbf{a} && (0.7b)\\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} && (0.7c)\\
&= (a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1)\,, && (0.7d)
\end{aligned}$$

or in terms of suffix notation

$$\mathbf{a}.\mathbf{b} = a_ib_i\,, \quad (\mathbf{a}\times\mathbf{b})_i = \varepsilon_{ijk}a_jb_k \quad (s.c.). \tag{0.8}$$

Alternatively, if $a = |\mathbf{a}|$, $b = |\mathbf{b}|$ and θ is the angle between $\mathbf{a}$ and $\mathbf{b}$, then

$$\mathbf{a}.\mathbf{b} = ab\cos\theta \quad \text{and} \quad \mathbf{a}\times\mathbf{b} = ab\sin\theta\,\widehat{\mathbf{n}}\,, \tag{0.9}$$

where, if $\theta \neq 0$, $\widehat{\mathbf{n}}$ is the unit vector such that $(\mathbf{a}, \mathbf{b}, \widehat{\mathbf{n}})$ are a right-handed basis. $|\mathbf{a}\times\mathbf{b}|$ is the area of the parallelogram having sides $\mathbf{a}$ and $\mathbf{b}$.

Properties of vector products in $\mathbb{R}^3$*.*

$$\mathbf{a}.\mathbf{a} = |a|^2\,, \tag{0.10a}$$
$$\mathbf{a}\times\mathbf{a} = 0\,, \tag{0.10b}$$
$$\mathbf{a}.(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = (\mathbf{a}\times\mathbf{b}).\mathbf{c}\,, \tag{0.10c}$$
$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}.\mathbf{c})\mathbf{b} - (\mathbf{a}.\mathbf{b})\mathbf{c}\,, \tag{0.10d}$$
$$(\mathbf{a}\times\mathbf{b}).(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}.\mathbf{c})(\mathbf{b}.\mathbf{d}) - (\mathbf{a}.\mathbf{d})(\mathbf{b}.\mathbf{c})\,, \tag{0.10e}$$

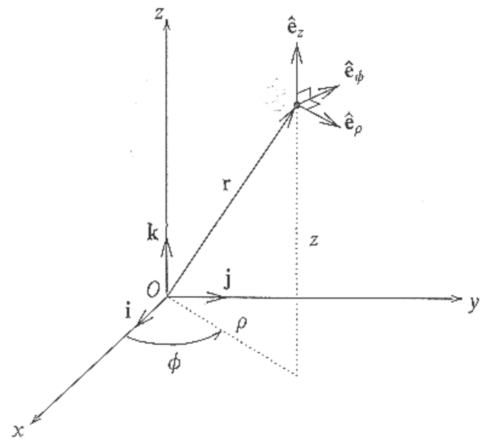
where (0.10d) can be obtained using the identity

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\,. \tag{0.11}$$

Position vectors and increments.

$$\mathbf{r} = \mathbf{x} = (x_1, x_2, x_3) = (x, y, z)\,, \tag{0.12a}$$
$$r = |\mathbf{r}| = |\mathbf{x}| = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}} = (x^2 + y^2 + z^2)^{\frac{1}{2}}\,, \tag{0.12b}$$
$$\mathrm{d}\mathbf{s} = \mathrm{d}\mathbf{r} = \mathrm{d}\mathbf{x} = (\mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3) = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)\,, \tag{0.12c}$$
$$\mathrm{d}s = \mathrm{d}r = (\mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2)^{\frac{1}{2}} = (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)^{\frac{1}{2}}\,. \tag{0.12d}$$

0.7.2 Cylindrical polar co-ordinates (ρ, ϕ, z) in $\mathbb{R}^3$.



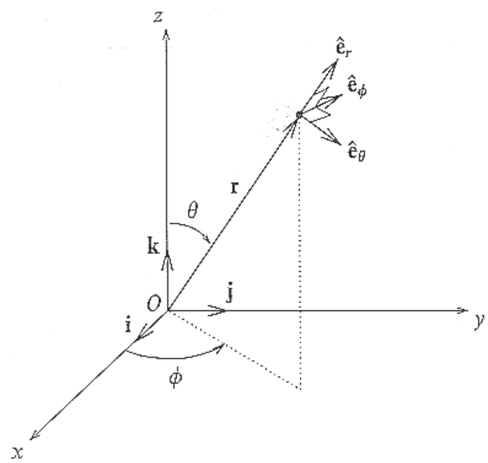
In cylindrical polar co-ordinates the position vector $\mathbf{x}$ is given in terms of a radial distance ρ from an axis $\mathbf{k}$, a polar angle ϕ, and the distance z along the axis:

$$\mathbf{x} = (\rho\cos\phi,\ \rho\sin\phi,\ z)\,, \tag{0.13}$$

where $0 \leqslant \rho < \infty$, $0 \leqslant \phi \leqslant 2\pi$ and $-\infty < z < \infty$.

Remark. Often r and/or θ are used in place of ρ and/or ϕ respectively (but then there is potential confusion with the different definitions of r and θ in spherical polar co-ordinates).

0.7.3 Spherical polar co-ordinates (r, θ, ϕ) in $\mathbb{R}^3$.



In spherical polar co-ordinates the position vector $\mathbf{x}$ is given in terms of a radial distance r from the origin, a 'latitude' angle θ, and a 'longitude' angle ϕ:

$$\mathbf{x} = (r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)\,, \tag{0.14}$$

where $0 \leqslant r < \infty$, $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$.

0.7.4 Determinants.

Suppose that A is a $m \times m$ matrix.

1. If $m = 2$ then the determinant of A is given by

$$|\mathsf{A}| = a_{11}a_{22} - a_{12}a_{21} \qquad \text{where} \qquad \mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

   If $m = 3$ the determinant of A is given by

$$\begin{aligned} |\mathsf{A}| &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ &= \varepsilon_{ijk}a_{1i}a_{2j}a_{3k} \qquad (s.c.) \\ &= \varepsilon_{ijk}a_{i1}a_{j2}a_{k3} \qquad (s.c.)\,, \end{aligned}$$

   where ε_{ijk} is the three-dimensional alternating tensor and

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

You might like to know, although the result will not be used, that for a general $m \times m$ matrix A

$$
\begin{aligned}
|\mathsf{A}| \;=\; \det \mathsf{A} \;=\; \det(A_{ij}) \quad &= \quad \varepsilon_{i_1 i_2 \dots i_m} A_{i_1 1} A_{i_2 2} \dots A_{i_m m} && (s.c.)\\
&= \quad \varepsilon_{j_1 j_2 \dots j_m} A_{1 j_1} A_{2 j_2} \dots A_{m j_m} && (s.c.)\\
&= \quad \sum_{\text{perm } \sigma} \operatorname{sgn}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \dots A_{\sigma(m)m}\\
&= \quad \sum_{\text{perm } \sigma} \operatorname{sgn}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \dots A_{m\sigma(m)}\,.
\end{aligned}
$$

where $\varepsilon_{i_1 i_2 \dots i_m}$ is the m-dimensional alternating tensor.

2. If $|\mathsf{A}| \neq 0$ then
$$|\mathsf{A}^{-1}| = |\mathsf{A}|^{-1}\,.$$

3. If B is another $m \times m$ matrix, then
$$|\mathsf{AB}| = |\mathsf{A}|\;|\mathsf{B}|\,.$$

1 Partial Differentiation

1.1 Scalar Functions of One Variable

You should have met most of the ideas, if not the presentation, of this first subsection at school.

Suppose $f(x)$ is a real function of $x \in \mathbb{R}$.

1.1.1 Limits

ℓ is the limit of $f(x)$ as $x \to a$ if

$$\lim_{x \to a} f(x) = \ell\,. \tag{1.1}$$

Remarks.

- Limits do not necessarily exist. For instance consider $f = 1/x$ as $x \to 0$.
- The left limit, i.e.

  $$\lim_{x \to a-} f(x) = \ell_-\,,$$

  need not equal the right limit, i.e.

  $$\lim_{x \to a+} f(x) = \ell_+\,.$$

  For instance consider $f = \tan^{-1}(1/x)$ as $x \to 0\pm$.

1.1.2 Continuity

$f(x)$ is continuous at a iff

$$\lim_{x \to a} f(x) = f(a)\,. \tag{1.2}$$

Remark.

- The limit may exist at $x = a$ without f being continuous. For instance consider

  $$f(x) = \begin{cases} 0 & x \neq a \\ 1 & x = a \end{cases}\,.$$

1.1.3 Differentiability

$f(x)$ is differentiable at a with derivative $\frac{\mathrm{d}f}{\mathrm{d}x}(a) = f'(a)$ if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a)\,. \tag{1.3}$$

Remarks.

- The derivative measures the rate of change of f with respect to x.
- A differentiable function is continuous (because both the numerator and denominator in (1.3) go to zero in the limit).

1.1.4 Taylor's Theorem

Suppose that f is differentiable $(N+1)$ times; then

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^N}{N!}f^{(N)}(a) + R_N(h)\,, \tag{1.4}$$

where for some $0 < \theta < 1$

$$R_N(h) = \frac{h^{N+1}}{(N+1)!}f^{(N+1)}(a+\theta h)\,. \tag{1.5a}$$

If $f^{(N+1)}$ is continuous then

$$R_N(h) = \frac{h^{N+1}}{N!}\int_0^1 (1-t)^N f^{(N+1)}(a+th)\,\mathrm{d}t\,. \tag{1.5b}$$

Notation.

If $g(x)/x$ is bounded as $x \to 0$ we say that $g(x) = O(x)$,

If $g(x)/x \to 0$ as $x \to 0$ we say that $g(x) = o(x)$.

Remarks.

- If f is $(N+1)$ times differentiable (on a suitable interval) then

$$f(a+h) - f(a) - \sum_{r=1}^{N}\frac{h^r}{r!}f^{(r)}(a) = O(h^{N+1})\,, \tag{1.6a}$$

  while if f is N times differentiable then from the definition of differentiability

$$f(a+h) - f(a) - \sum_{r=1}^{N}\frac{h^r}{r!}f^{(r)}(a) = o(h^N)\,. \tag{1.6b}$$

- If f is differentiable

$$\begin{aligned} f(a+h) &= f(a) + hf'(a) + o(h) \qquad (1.7)\\ &= f(a) + hf'(a) + \ldots\,, \end{aligned}$$

  where the ellipsis $\ldots$ is sloppy (but common) notation to indicate terms smaller than those included.
- Knowledge of the derivative allows us to make a local *linear* approximation of a smooth function.

1.2 Functions of Several Variables

We have so far considered scalar functions of a scalar

$$\text{e.g.}\quad f(x) = x^2 + 1 \quad \text{from } \mathbb{R} \text{ to } \mathbb{R}\,,$$

but it is also possible to have scalar functions of a vector

$$\text{e.g.}\quad f(x,y) = x^2 + y^2 \qquad \text{from } \mathbb{R}^2 \text{ to } \mathbb{R}\,,$$

$$\text{e.g.}\quad f(\mathbf{x}) \equiv f(x_1,\ldots,x_m) = \sum_{j=1}^{m} x_j^2 = x_j x_j \quad (s.c.) \qquad \text{from } \mathbb{R}^m \text{ to } \mathbb{R}\,,$$

and vector functions of a scalar or vector

$$\text{e.g.} \quad \mathbf{f}(x) = \sum_{j=1}^{n} f_j(x)\mathbf{e}_j \equiv (f_1(x), \dots, f_n(x)) \quad \text{from } \mathbb{R} \text{ to } \mathbb{R}^n \ ,$$

$$\text{e.g.} \quad \mathbf{f}(\mathbf{x}) = \sum_{j=1}^{n} f_j(\mathbf{x})\mathbf{e}_j \quad \text{from } \mathbb{R}^m \text{ to } \mathbb{R}^n \ ,$$

where the $\mathbf{e}_j$ $(j = 1, \dots, n)$ are orthonormal basis vectors (see (0.2)).

Remark.

- A function may only be defined on a subset of $\mathbb{R}^m$ called its *domain*. For instance the real function f,

$$f(x, y) = (1 - x^2 - y^2)^{\frac{1}{2}} \,,$$

  is only defined for $x^2 + y^2 \leqslant 1$.

1.2.1 Limits

For a function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$, $\boldsymbol{\ell}$ is the limit of $\mathbf{f}(\mathbf{x})$ as $\mathbf{x} \to \mathbf{a}$ if

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell} \tag{1.8}$$

independently of the way that $\mathbf{x}$ approaches $\mathbf{a}$.

Remark and Example.

- The independence of approach route of $\mathbf{x}$ to $\mathbf{a}$ is crucial (cf. left and right limits being equal for $f : \mathbb{R} \to \mathbb{R}$). To illustrate this, consider the function

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \mathbf{x} \neq 0 \\ 0 & \mathbf{x} = 0 \end{cases} . \tag{1.9}$$

  Suppose that $\mathbf{x} = (\rho\cos\phi, \rho\sin\phi) \to 0$ on lines of constant ϕ. Since

$$f(x, y) = \sin 2\phi \qquad \rho \neq 0 \,,$$

  it follows that $\lim_{\rho\to 0} f(x, y)$ can be anything between -1 and 1. Hence no limit exists as $\mathbf{x} \to 0$.

1.2.2 Continuity

A function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ is continuous at $\mathbf{a}$ iff

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) \,. \tag{1.10}$$

Remark.

- Continuity in $\mathbb{R}^m$ says more than continuity in x_1 for fixed $x_2, x_3, \ldots, x_m$, plus
  continuity in x_2 for fixed $x_1, x_3, \ldots, x_m$, plus … plus
  continuity in x_m for fixed $x_1, x_2, \ldots, x_{m-1}$.

  For instance, in the case of $f(x, y)$ as defined by (1.9) no limit exists as $\mathbf{x} \to 0$, hence f cannot be continuous at $\mathbf{x} = 0$. However $f(x, 0) = 0$ and $f(0, y) = 0$ are continuous functions of x and y respectively.

1.3 Derivatives of Vector Functions of One Variable

Suppose $\mathbf{f}(x)$ is a real vector function from $\mathbb{R}$ to $\mathbb{R}^n$. The derivative of $\mathbf{f}(x)$ at a is defined as

$$\mathbf{f}'(a) = \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x}(a) = \lim_{x\to a} \frac{\mathbf{f}(x) - \mathbf{f}(a)}{x - a}\,, \tag{1.11a}$$

where this limit exists. If we write $\mathbf{f} = f_i\,\mathbf{e}_i$ (*s.c.*), where the $\mathbf{e}_i$ $(i = 1, \ldots, n)$ are *fixed* basis vectors (e.g. $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in $\mathbb{R}^3$), it follows that

$$\begin{aligned}
\mathbf{f}'(a) \;=\; \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x}(a) \;&=\; \lim_{x\to a} \frac{f_i(x)\mathbf{e}_i - f_i(a)\mathbf{e}_i}{x - a} \;=\; \frac{\mathrm{d}f_i}{\mathrm{d}x}(a)\,\mathbf{e}_i & \text{(1.11b)}\\
&=\; \left(\frac{\mathrm{d}f_1}{\mathrm{d}x}(a), \frac{\mathrm{d}f_2}{\mathrm{d}x}(a), \ldots, \frac{\mathrm{d}f_n}{\mathrm{d}x}(a)\right)\,, & \text{(1.11c)}\\
&=\; \left(f_1'(a), f_2'(a), \ldots, f_n'(a)\right). & \text{(1.11d)}
\end{aligned}$$

Remark. If the position a particle in $\mathbb{R}^3$ at time t is given by $\mathbf{x}(t)$, then we define the velocity $\mathbf{u}(t)$ of the particle at time t to be

$$\mathbf{u} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,.$$

1.3.1 Example

Consider the function $g(s) = \mathbf{u}(s).\mathbf{u}(s)$. Then using the *summation convention* (*s.c.*),

$$\frac{\mathrm{d}g}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{u}.\mathbf{u}\right) = \frac{\mathrm{d}}{\mathrm{d}s}(u_i u_i) = 2u_i \frac{\mathrm{d}u_i}{\mathrm{d}s} = 2\mathbf{u}.\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s}\,. \tag{1.12}$$

Suppose further that there is a constraint so that $g = \mathbf{u}.\mathbf{u} = 1$. It follows that $g' = 0$, and hence from
1/00 (1.12) that $\mathbf{u}.\mathbf{u}' = 0$, i.e. that $\mathbf{u}$ and $\mathbf{u}'$ are perpendicular.

1.4 Partial Derivatives

How do we differentiate functions of several variables? The method we will adopt is:

- try the obvious;
- spot something fishy;
- fix it up (but not in this course).

1.4.1 Definition

For a function $f(x,y)$ of two variables define the *partial derivative* with respect to x at (a,b) as

$$\left(\frac{\partial f}{\partial x}\right)_y = \lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}\,. \tag{1.13a}$$

Similarly define the *partial derivative* with respect to y at (a,b) as

$$\left(\frac{\partial f}{\partial y}\right)_x = \lim_{k\to 0}\frac{f(a,b+k)-f(a,b)}{k}\,. \tag{1.13b}$$

More generally for $\mathbf{f}:\mathbb{R}^m\to\mathbb{R}^n$, the j^{th} *partial derivative* of $\mathbf{f}$ at $\mathbf{a}\in\mathbb{R}^m$ is

$$\frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{a}) = \lim_{h_j\to 0}\frac{\mathbf{f}(a_1,\dots,a_{j-1},a_j+h_j,a_{j+1},\dots,a_m)-\mathbf{f}(a_1,\dots,a_{j-1},a_j,a_{j+1},\dots,a_m)}{h_j}\,, \tag{1.14}$$

where this limit exists.

Notation. There are a number of alternative notations for partial derivatives:

$$\begin{aligned}\frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{a}) &= \partial_{x_j}\mathbf{f}(\mathbf{a}) = \partial_j\mathbf{f}(\mathbf{a}) = \mathbf{f}_{x_j}(\mathbf{a}) = \mathbf{f},_j(\mathbf{a}) = D_j\mathbf{f}(\mathbf{a}) &(1.15a)\\ &= \left(\frac{\partial \mathbf{f}}{\partial x_j}\right)_{x_1,\dots,x_{j-1},x_{j+1},\dots,x_m}\!\!(\mathbf{a}) = \left.\frac{\partial \mathbf{f}}{\partial x_j}\right|_{x_1,\dots,x_{j-1},x_{j+1},\dots,x_m}\!\!(\mathbf{a})\,. &(1.15b)\end{aligned}$$

Remark. When doing the question on *Examples Sheet 1* on *Euler's theorem for homogeneous functions*
1/99 you are recommended to use the $\mathbf{f},_j$ (or $\partial_j\mathbf{f}$ or $D_j\mathbf{f}$) notation.

1.4.2 Examples

1. Consider the function $f(x,y)=x^2+yx$. Then

$$\left(\frac{\partial f}{\partial x}\right)_y = 2x+y\,,\qquad \left(\frac{\partial f}{\partial y}\right)_x = x\,.$$

2. Consider the function $f(x,y,z)=r$ where $r=|\mathbf{x}|=(x^2+y^2+z^2)^{\frac12}$. Then

$$\left(\frac{\partial r}{\partial x}\right)_{y,z} = \left(\frac{\partial}{\partial x}\right)_{y,z}(x^2+y^2+z^2)^{\frac12} = \frac{1}{2(x^2+y^2+z^2)^{\frac12}}\;2x = \frac{x}{r}\,, \tag{1.16a}$$

and similarly

$$\left(\frac{\partial r}{\partial y}\right)_{x,z} = \frac{y}{r}\,,\quad \left(\frac{\partial r}{\partial z}\right)_{x,y} = \frac{z}{r}\,. \tag{1.16b}$$

3. Find $\frac{\partial f}{\partial x_k}$ for $f=x_j$, i.e. $\frac{\partial x_j}{\partial x_k}$. First note that for $k=1$

$$\left(\frac{\partial x_1}{\partial x_1}\right)_{x_2,\dots,x_m} = 1\,,\quad \left(\frac{\partial x_2}{\partial x_1}\right)_{x_2,\dots,x_m} = 0\,,\ \text{etc.,}$$

i.e.

$$\left(\frac{\partial x_j}{\partial x_1}\right)_{x_2,\dots,x_m} = \left\{\begin{array}{lll}1 & \text{if} & j=1\\ 0 & \text{if} & j\neq 1\end{array}\right\} = \delta_{j1}\,,$$

where δ_{jk} is the Kronecker delta. Similarly for $k=2$

$$\left(\frac{\partial x_j}{\partial x_2}\right)_{x_1,x_3,\dots,x_m} = \delta_{j2}\,,$$

and for general k

$$\frac{\partial x_j}{\partial x_k} = \delta_{jk}\,. \tag{1.17}$$

We can now use this result to obtain (1.16a) and (1.16b) more cleanly. First we note (invoking the summation convention) that $r^2 = x_j x_j$, and hence that

$$2r\frac{\partial r}{\partial x_k} = \frac{\partial r^2}{\partial x_k} = \frac{\partial}{\partial x_k}(x_j x_j) = 2x_j\frac{\partial x_j}{\partial x_k} = 2x_j\delta_{jk} = 2x_k\,,$$

i.e. on dividing through by $2r$

$$\frac{\partial r}{\partial x_k} = \frac{x_k}{r}\,, \tag{1.18}$$

which, in terms of suffix notation, is the same as (1.16a) and (1.16b).

4. *Unlectured example to demonstrate that the partial derivatives of a function can exist at all points in $\mathbb{R}^3$, including the origin, even though the function is not continuous at the origin.* The partial derivatives of the function defined in (1.9), i.e.

$$f(x,y) = \begin{cases} \dfrac{2xy}{x^2+y^2} & \mathbf{x} \neq 0 \\ 0 & \mathbf{x} = 0 \end{cases}\,, \tag{1.19a}$$

for $\mathbf{x} \neq 0$ are

$$\left(\frac{\partial f}{\partial x}\right)_y = \frac{2y(y^2-x^2)}{(x^2+y^2)^2}\,, \quad \left(\frac{\partial f}{\partial y}\right)_x = \frac{2x(x^2-y^2)}{(x^2+y^2)^2}\,, \tag{1.19b}$$

and at $\mathbf{x} = 0$ are

$$\left(\frac{\partial f}{\partial x}\right)_y = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = 0\,, \tag{1.19c}$$

$$\left(\frac{\partial f}{\partial y}\right)_x = \lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h} = 0\,. \tag{1.19d}$$

Hence $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at all points in $\mathbb{R}^2$. However, in § 1.2.2 we concluded that f as defined by (1.9) is not continuous at 0. This is *fishy*, since for $f : \mathbb{R} \to \mathbb{R}$ differentiability implies continuity.

Remark. The existence of partial derivatives does *not* imply differentiability (although, of course, we have yet to define *differentiability*, so this statement is a little presumptive).

1.4.3 Directional Derivatives

This is the generalisation of [first-order] partial derivatives to directions other than parallel to the co-ordinate vectors $\mathbf{e}_j$.

Definition. Let $\widehat{\mathbf{u}} \in \mathbb{R}^m$ be a unit vector. Then for $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$, the *directional derivative* of $\mathbf{f}$ at $\mathbf{a}$ in direction $\widehat{\mathbf{u}}$ is

$$\mathbf{f}'(\mathbf{a};\widehat{\mathbf{u}}) = \lim_{h\to 0} \frac{\mathbf{f}(\mathbf{a}+h\widehat{\mathbf{u}}) - \mathbf{f}(\mathbf{a})}{h}\,, \tag{1.20}$$

whenever this limit exists.

Remarks.

- It follows from (1.20) and the definition of a partial derivative (1.14) that

$$\frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{a}) = \mathbf{f}'(\mathbf{a};\mathbf{e}_j)\,. \tag{1.21}$$

- The directional derivative is a measure of the rate of change of $\mathbf{f}$ in the direction $\widehat{\mathbf{u}}$.
- The directional derivatives of a function at $\mathbf{a}$ can exist for all directions $\widehat{\mathbf{u}}$, without the function necessarily being continuous at $\mathbf{a}$ (see *Examples Sheet 1* for a function which is not continuous at the origin but for which the directional derivatives exist in all directions at the origin).

1.5 Differentiability

The presence of a margin line indicates material that, in my opinion, does not appear in the schedule, and that, in my opinion, is not examinable.
Remarks

- The key idea behind differentiability is the *local* approximation of a function by a linear function/map.
- Note from definition (1.3) that a function $g(x)$ is differentiable at a if for some $g'(a)$

$$g(a+h) - g(a) = g'(a)h + \epsilon(h)|h|\,,$$

  where $|\epsilon(h)| \to 0$ as $|h| \to 0$.

1.5.1 Scalar Functions of Two Variables

A function $f(x,y)$ is differentiable at (a,b) if

$$f(a+h,b+k) - f(a,b) = T_x h + T_y k + \epsilon(\mathbf{h})|\mathbf{h}|\,, \tag{1.22}$$

where T_x and T_y are independent of $\mathbf{h} = (h,k)$, and $|\epsilon(\mathbf{h})| \to 0$ as $|\mathbf{h}| = (h^2+k^2)^{\frac{1}{2}} \to 0$.

Remarks.

- $T_x h + T_y k$ is a linear function of h and k (i.e. we have our linear approximation).
- Let $z = f(x,y)$ define a surface (cf. $y = g(x)$ defining a curve). Then

$$z = f(a,b) + T_x(x-a) + T_y(y-b)$$

  is the tangent plane touching the surface at the point $(a,b,f(a,b))$ (cf. $y = g(a) + g'(a)(x-a)$ defining the tangent curve at the point $(a,g(a))$).

- Let $|\mathbf{h}| \to 0$ with $k = 0$, then from (1.22) and the definition of the partial derivative (1.13a)
$$T_x(x,y) = \left(\frac{\partial}{\partial x}\right)_y f(x,y)\,. \tag{1.23a}$$
Similarly by letting $|\mathbf{h}| \to 0$ with $h = 0$
$$T_y(x,y) = \left(\frac{\partial}{\partial y}\right)_x f(x,y)\,. \tag{1.23b}$$
- Differentiability implies that the directional derivative exists for all directions $\widehat{\mathbf{u}} = (\widehat{u}_x, \widehat{u}_y)$; in particular
$$f'(\mathbf{x}; \widehat{\mathbf{u}}) = T_x \widehat{u}_x + T_y \widehat{u}_y\,.$$
It follows that differentiability implies that partial derivatives exist for all orientations of the co-ordinate axes.

Unlectured example: show that the function $f(x,y) = xy$ is differentiable. Starting from the definition (1.22) we have that
$$\begin{aligned} f(x+h, y+k) - f(x,y) &= xy + yh + xk + hk - xy \\ &= yh + xk + \frac{hk}{(h^2+k^2)^{\frac{1}{2}}}|\mathbf{h}|\,. \end{aligned}$$
Hence, after writing $h = |\mathbf{h}|\cos\phi$ and $k = |\mathbf{h}|\sin\phi$ where $|\mathbf{h}| = (h^2+k^2)^{\frac{1}{2}}$,
$$\epsilon(\mathbf{h}) = \frac{hk}{(h^2+k^2)^{\frac{1}{2}}} = |\mathbf{h}|\cos\phi\,\sin\phi \to 0 \quad \text{as} \quad |\mathbf{h}| \to 0\,.$$
Therefore f is differentiable, and moreover from definition (1.22) and results (1.23a) and (1.23b)
$$\left(\frac{\partial f}{\partial x}\right)_y = T_x = y\,, \quad \left(\frac{\partial f}{\partial y}\right)_x = T_y = x\,.$$

1.5.2 Functions of Several Variables

A function $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $\mathbf{a} \in \mathbb{R}^m$ if for all $\mathbf{h} \in \mathbb{R}^m$
$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = \mathsf{T}\,\mathbf{h} + \boldsymbol{\epsilon}(\mathbf{h})|\mathbf{h}|\,, \tag{1.24}$$
where T is a $n \times m$ matrix that is independent of $\mathbf{h}$, and
$$|\boldsymbol{\epsilon}(\mathbf{h})| \to 0 \quad \text{as} \quad |\mathbf{h}| \to 0\,.$$

Remarks.

- $\mathsf{T}\,\mathbf{h}$ is a linear function of $\mathbf{h}$.
- From the definition of a partial derivative given in (1.14), it follows from (1.24) for the special case $\mathbf{h} = (0, \dots, 0, h_j, 0, \dots, 0)$ that
$$\frac{\partial f_i}{\partial x_j}(\mathbf{a}) = \lim_{h_j \to 0} \frac{f_i(a_1, \dots, a_j + h_j, \dots, a_m) - f_i(a_1, \dots, a_j, \dots, a_m)}{h} = \mathsf{T}_{ij}\,. \tag{1.25}$$

Definition. The $n \times m$ *Jacobian matrix* T (which is a 'representation' of the derivative) is defined by
$$\mathsf{T} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}. \tag{1.26}$$

Example. For $\mathbf{f}(x,y) = x^2y\mathbf{i} + 2x\mathbf{j} + 2y\mathbf{k}$,

$$\mathsf{T} = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 \\ 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

1.5.3 Properties of Differentiable Functions

We shall just state a couple of these. See Part IB *Analysis* both for more properties (many of which are not unexpected), and for proofs.

Continuity. If a function $\mathbf{f}(\mathbf{x})$ is differentiable at $\mathbf{a}$, then $\mathbf{f}$ is continuous at $\mathbf{a}$.

A sufficient condition for differentiability. If all the partial derivatives of $\mathbf{f}$ exist 'close to' $\mathbf{a}$, and are continuous at $\mathbf{a}$, then $\mathbf{f}$ is differentiable at $\mathbf{a}$; i.e. the continuity of partial derivatives, rather than just their existence, ensures differentiability.

Optional Exercises.

1. Confirm that for f defined by (1.19a), the partial derivatives defined by (1.19b), (1.19c) and (1.19d) are not continuous at the origin. Note that if these partial derivatives were continuous at the origin, then from the above two results, f would be both differentiable and continuous at the origin, which would be in contradiction to the result derived in § 1.2.2.
2. Give an example of a function $f(x,y)$ that is differentiable at $(0,0)$, but for which the partial derivatives are either not continuous at $(0,0)$, or do not exist at other than at $(0,0)$.

1.6 The Chain Rule and Change of Variables

For differentiable $x : \mathbb{R} \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ define the composite function $F : \mathbb{R} \to \mathbb{R}$ by

$$F(t) = f(x(t)).$$

Then the chain rule states that

$$F'(t) \;=\; f'(x(t))\,x'(t) \quad \text{i.e.} \quad \frac{\mathrm{d}F}{\mathrm{d}t} \;=\; \frac{\mathrm{d}f}{\mathrm{d}x}\,\frac{\mathrm{d}x}{\mathrm{d}t}\,. \tag{1.27}$$

2/00 How does this generalise to functions of several variables?

1.6.1 The Chain Rule for $\mathbf{x} : \mathbb{R} \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$

Suppose that $\mathbf{x}$ and f are differentiable, and consider

$$F(t) = f(\mathbf{x}(t)), \tag{1.28}$$

and its derivative

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \lim_{\delta t \to 0} \frac{F(t+\delta t) - F(t)}{\delta t} = \lim_{\delta t \to 0} \frac{f(\mathbf{x}(t+\delta t)) - f(\mathbf{x}(t))}{\delta t} = \lim_{\delta t \to 0} \frac{f(\mathbf{x}(t) + \delta\mathbf{x}(t)) - f(\mathbf{x}(t))}{\delta t},$$

where we have written

$$\mathbf{x}(t+\delta t) = \mathbf{x}(t) + \delta\mathbf{x}(t).$$

Recall also from the definition (1.11a) of a derivative of a vector function of one variable,

$$\lim_{\delta t \to 0} \frac{\delta \mathbf{x}}{\delta t} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}. \tag{1.29}$$

In order to focus ideas first consider $m = 2$ and write $\mathbf{x} = (x, y)$. Then from using (1.29) together with the definitions (1.13a) & (1.13b) of a partial derivative, it follows on assuming continuity of the partial derivatives,[2] that

$$\begin{aligned}
\frac{\mathrm{d}F}{\mathrm{d}t} &= \lim_{\delta t\to 0}\left\{\frac{f(x+\delta x, y+\delta y) - f(x, y+\delta y) + f(x, y+\delta y) - f(x,y)}{\delta t}\right\} \\
&= \lim_{\delta t\to 0}\left\{\frac{f(x+\delta x, y+\delta y) - f(x, y+\delta y)}{\delta x}\;\frac{\delta x}{\delta t}\right\} + \lim_{\delta t\to 0}\left\{\frac{f(x, y+\delta y) - f(x,y)}{\delta y}\;\frac{\delta y}{\delta t}\right\} \\
&= \left(\frac{\partial f}{\partial x}\right)_y \frac{\mathrm{d}x}{\mathrm{d}t} + \left(\frac{\partial f}{\partial y}\right)_x \frac{\mathrm{d}y}{\mathrm{d}t}\,. \qquad (1.30a)
\end{aligned}$$

Similarly (but much more messily) for general m

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t)) = \sum_{j=1}^{m}\left(\frac{\partial f}{\partial x_j}\right)_{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_m}\frac{\mathrm{d}x_j}{\mathrm{d}t}\,, \qquad (1.30b)$$

or, if we invoke the summation convention and do not display the variables that are begin held constant,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{x}(t)) = \frac{\partial f}{\partial x_j}\,\frac{\mathrm{d}x_j}{\mathrm{d}t}\,. \qquad (1.30c)$$

2/99

Example. Suppose that

$$f(x,y) = x^2 + y\,, \quad \mathbf{x}(t) = (x,y) = (t^2, \mathrm{e}^t)\,,$$

and that as above $F(t) = f(\mathbf{x}(t))$. We calculate $\frac{\mathrm{d}F}{\mathrm{d}t}$ in two ways.

1. From the definitions of f and $\mathbf{x}$

$$\left(\frac{\partial f}{\partial x}\right)_y = 2x\,, \quad \left(\frac{\partial f}{\partial y}\right)_x = 1\,, \quad \frac{\mathrm{d}x}{\mathrm{d}t} = 2t\,, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = \mathrm{e}^t\,.$$

   Hence from using the chain rule (1.30a)

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial f}{\partial x}\,\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\,\frac{\mathrm{d}y}{\mathrm{d}t} = 2x\,2t + \mathrm{e}^t = 4t^3 + \mathrm{e}^t\,.$$

2. From direct substitution

$$F = t^4 + \mathrm{e}^t\,, \quad \text{and hence} \quad \frac{\mathrm{d}F}{\mathrm{d}t} = 4t^3 + \mathrm{e}^t\,.$$

Unlectured example: implicit differentiation. Suppose that $f \equiv f(x,y)$ and that $y \equiv y(x)$. Write $x = t$ and define $F(t) = f(\mathbf{x}(t))$ where $\mathbf{x} = (t, y(t))$. Then from (1.30a)

$$\begin{aligned}
\frac{\mathrm{d}F}{\mathrm{d}t} &= \left(\frac{\partial f}{\partial x}\right)_y \frac{\mathrm{d}x}{\mathrm{d}t} + \left(\frac{\partial f}{\partial y}\right)_x \frac{\mathrm{d}y}{\mathrm{d}t} \\
&= \left(\frac{\partial f}{\partial x}\right)_y + \left(\frac{\partial f}{\partial y}\right)_x \frac{\mathrm{d}y}{\mathrm{d}t}\,,
\end{aligned}$$

or since $x = t$

$$\frac{\mathrm{d}}{\mathrm{d}x} f\big(x, y(x)\big) = \left(\frac{\partial f}{\partial x}\right)_y + \left(\frac{\partial f}{\partial y}\right)_x \frac{\mathrm{d}y}{\mathrm{d}x}\,. \qquad (1.31)$$

Suppose further $y(x)$ is defined implicitly by the equation

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

Then $\mathrm{d}f/\mathrm{d}x = 0$, and hence from (1.31)

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\left(\frac{\partial f}{\partial x}\right)_y \Bigg/ \left(\frac{\partial f}{\partial y}\right)_x\,. \qquad (1.32)$$

[2] The continuity assumption is for simplicity; it is not a necessity.

1.6.2 The Chain Rule for $\mathbf{x} : \mathbb{R}^\ell \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$

Suppose $\mathbf{x}$ and f are differentiable and consider

$$F(u_1, \dots, u_\ell) = f(\mathbf{x}(u_1, \dots u_\ell)) \,,$$

where the vector $\mathbf{u} = (u_1, \dots u_\ell)$ is the independent variable. By keeping $u_2, \dots, u_\ell$ constant and differentiating with respect to u_1, we have from (1.30b) that

$$\left(\frac{\partial F}{\partial u_1}\right)_{u_2,\dots,u_\ell} = \sum_{j=1}^{m} \left(\frac{\partial f}{\partial x_j}\right)_{x_1,\dots,x_{j-1},x_{j+1},\dots,x_m} \left(\frac{\partial x_j}{\partial u_1}\right)_{u_2,\dots,u_\ell} ,$$

where we write $\frac{\partial F}{\partial u_1}$ rather than $\frac{\mathrm{d}F}{\mathrm{d}u_1}$ since $F(\mathbf{u})$ is a function of two or more variables. Similarly

$$\frac{\partial F}{\partial u_k} = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u_k} \,, \tag{1.33a}$$

where we have not displayed the variables that are being held constant — this is a ***dangerous*** notational 'simplification', but an often used one. Alternatively if we invoke the *s.c.*

$$\frac{\partial F}{\partial u_k} = \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u_k} \,. \tag{1.33b}$$

Example. Consider the transformation (actually, as discussed later, a rotation of the co-ordinate axes through an angle θ)

$$\mathbf{x} \equiv (x, y) = (u\cos\theta - v\sin\theta,\ u\sin\theta + v\cos\theta) \,, \tag{1.34}$$

where we have written (x, y) for (x_1, x_2) and (u, v) for (u_1, u_2). Then for $F(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}))$, it follows from (1.33a) or (1.33b), that

$$\left(\frac{\partial F}{\partial u}\right)_v = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v = \cos\theta \left(\frac{\partial f}{\partial x}\right)_y + \sin\theta \left(\frac{\partial f}{\partial y}\right)_x , \tag{1.35a}$$

$$\left(\frac{\partial F}{\partial v}\right)_u = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u = -\sin\theta \left(\frac{\partial f}{\partial x}\right)_y + \cos\theta \left(\frac{\partial f}{\partial y}\right)_x . \tag{1.35b}$$

Sloppy Notation. Sometimes, indeed often, $F(u, v) = f(x(u, v), y(u, v))$ is written as $f(u, v)$. Admittedly this can be confusing at first sight, but context is all and otherwise it is possible to run out of alphabet! The results (1.35a) & (1.35b) then become

$$\left(\frac{\partial f}{\partial u}\right)_v = \cos\theta \left(\frac{\partial f}{\partial x}\right)_y + \sin\theta \left(\frac{\partial f}{\partial y}\right)_x ,$$

$$\left(\frac{\partial f}{\partial v}\right)_u = -\sin\theta \left(\frac{\partial f}{\partial x}\right)_y + \cos\theta \left(\frac{\partial f}{\partial y}\right)_x ,$$

or for more general $f \equiv f(x, y)$ and $\mathbf{x} \equiv \mathbf{x}(u, v)$

$$\left(\frac{\partial f}{\partial u}\right)_v = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial u}\right)_v + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial u}\right)_v ,$$

$$\left(\frac{\partial f}{\partial v}\right)_u = \left(\frac{\partial f}{\partial x}\right)_y \left(\frac{\partial x}{\partial v}\right)_u + \left(\frac{\partial f}{\partial y}\right)_x \left(\frac{\partial y}{\partial v}\right)_u .$$

Similarly for $f \equiv f(\mathbf{x})$ and $\mathbf{x} \equiv \mathbf{x}(\mathbf{u})$, (1.33b) becomes

$$\left(\frac{\partial f}{\partial u_k}\right)_{u_1,\dots,u_{k-1},u_{k+1},\dots,u_\ell} = \left(\frac{\partial f}{\partial x_j}\right)_{x_1,\dots,x_{j-1},x_{j+1},\dots,x_m} \left(\frac{\partial x_j}{\partial u_k}\right)_{u_1,\dots,u_{k-1},u_{k+1},\dots,u_\ell} ,$$

or just

$$\frac{\partial f}{\partial u_k} = \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial u_k} \,. \tag{1.33c}$$

1.6.3 The Chain Rule for $\mathbf{x} : \mathbb{R}^\ell \to \mathbb{R}^m$ and $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$

For differentiable $\mathbf{x}(\mathbf{u})$ and $\mathbf{f}(\mathbf{x})$, let

$$\mathbf{F}(\mathbf{u}) = \mathbf{f}(\mathbf{x}(\mathbf{u}))\,.$$

Then by applying (1.33b) componentwise and using the *s.c.*, we have that

$$\frac{\partial F_i}{\partial u_k} = \frac{\partial f_i}{\partial x_j}\frac{\partial x_j}{\partial u_k}\,, \tag{1.36a}$$

or in more sloppy notation

$$\frac{\partial f_i}{\partial u_k} = \frac{\partial f_i}{\partial x_j}\frac{\partial x_j}{\partial u_k}\,, \tag{1.36b}$$

1.6.4 The Matrix Form of the Chain Rule

Denote the Jacobian matrices of	$\mathbf{F}(\mathbf{u})$,	$\mathbf{f}(\mathbf{x})$	and	$\mathbf{x}(\mathbf{u})$	
by	S,	T	and	U	respectively.
These matrices are of size	$n \times \ell$,	$n \times m$	and	$m \times \ell$	respectively.

The matrix T is given by (1.26), while S and U are given by

$$\mathsf{S} = \begin{pmatrix} \frac{\partial F_1}{\partial u_1} & \frac{\partial F_1}{\partial u_2} & \dots & \frac{\partial F_1}{\partial u_\ell} \\ \frac{\partial F_2}{\partial u_1} & \frac{\partial F_2}{\partial u_2} & \dots & \frac{\partial F_2}{\partial u_\ell} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_n}{\partial u_1} & \frac{\partial F_n}{\partial u_2} & \dots & \frac{\partial F_n}{\partial u_\ell} \end{pmatrix}, \qquad \mathsf{U} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \dots & \frac{\partial x_1}{\partial u_\ell} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \dots & \frac{\partial x_2}{\partial u_\ell} \\ \dots & \dots & \dots & \dots \\ \frac{\partial x_m}{\partial u_1} & \frac{\partial x_m}{\partial u_2} & \dots & \frac{\partial x_m}{\partial u_\ell} \end{pmatrix}. \tag{1.37}$$

Using these definitions, (1.36a) can be rewritten (using the *s.c.*) as

$$\mathsf{S} = \mathsf{T}\mathsf{U} \quad \text{or, using the } s.c.\text{, as} \quad \mathsf{S}_{ik} = \mathsf{T}_{ij}\mathsf{U}_{jk}\,. \tag{1.38}$$

This is the matrix form of the chain rule.

Remarks.

- From the property of determinants
$$|\mathsf{S}| = |\mathsf{T}||\mathsf{U}|\,. \tag{1.39}$$
- If $n = m$ and T is non-singular then
$$\mathsf{U} = \mathsf{T}^{-1}\mathsf{S} \qquad \text{and} \qquad |\mathsf{U}| = |\mathsf{S}|/|\mathsf{T}|\,. \tag{1.40a}$$
Similarly if $m = \ell$ and U is non-singular then
$$\mathsf{T} = \mathsf{S}\mathsf{U}^{-1} \qquad \text{and} \qquad |\mathsf{T}| = |\mathsf{S}|/|\mathsf{U}|\,. \tag{1.40b}$$

1.6.5 Change of Variables

To fix ideas, initially work with a scalar function/field f.

Digression. There are many important physical examples of scalar fields, e.g. the gravitational potential, the electric potential, density, temperature, salinity, the ozone mixing ratio (i.e. the concentration of ozone), the partial pressure of oxygen in blood (P_{O_2}).

For instance, suppose that $f(x, y, z)$ describes the ozone mixing ratio. Rather than using Cartesian co-ordinates (x, y, z) centred, say, at the centre of the earth, it is more natural to use spherical polar co-ordinates (r, θ, ϕ), or a modified version of spherical polar co-ordinates that measures distances above the surface of the earth. We need to be able to transform f and its partial derivatives from one co-ordinate system to another.

For simplicity we will consider a two dimensional example, in which instead of using Cartesian co-ordinates (x, y), we wish to use (u, v), say because of the physical properties of the system. For instance:

Rotation — also see (1.34)

$$x = u\cos\theta - v\sin\theta$$
$$y = u\sin\theta + v\cos\theta$$

Plane polar co-ordinates

$$x = u\cos v$$
$$y = u\sin v$$

For U defined by (1.37), i.e. $U_{ij} = \frac{\partial x_i}{\partial u_j}$,

Rotation

$$\mathsf{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Plane polar co-ordinates

$$\mathsf{U} = \begin{pmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{pmatrix}$$

while from (1.37) and (1.26) respectively

$$\mathsf{S} = \left(\frac{\partial F}{\partial u}, \frac{\partial F}{\partial v}\right) \quad \text{and} \quad \mathsf{T} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right).$$

Hence from (1.38), or directly from (1.33a) or (1.33b)

Rotation — see also (1.35a) & (1.35b)

$$\begin{aligned} \frac{\partial F}{\partial u} &= \cos\theta\,\frac{\partial f}{\partial x} + \sin\theta\,\frac{\partial f}{\partial y} \\ \frac{\partial F}{\partial v} &= -\sin\theta\,\frac{\partial f}{\partial x} + \cos\theta\,\frac{\partial f}{\partial y} \end{aligned}$$

Plane polar co-ordinates

$$\begin{aligned} \frac{\partial F}{\partial u} &= \cos v\,\frac{\partial f}{\partial x} + \sin v\,\frac{\partial f}{\partial y} \\ \frac{\partial F}{\partial v} &= -u\sin v\,\frac{\partial f}{\partial x} + u\cos v\,\frac{\partial f}{\partial y} \end{aligned}$$

Remarks.

- For rotations it is more conventional to use (x', y') or (X, Y) rather than (u, v).
- For plane polar co-ordinates it is more conventional to use (ρ, ϕ) or (r, θ) rather than (u, v).

In the case of (ρ, ϕ) we have after identifying F with f that

$$\begin{aligned}
\rho\left(\frac{\partial f}{\partial \rho}\right)_\phi &= \rho\cos\phi\left(\frac{\partial f}{\partial x}\right)_y + \rho\sin\phi\left(\frac{\partial f}{\partial y}\right)_x \\
&= x\left(\frac{\partial f}{\partial x}\right)_y + y\left(\frac{\partial f}{\partial y}\right)_x, && (1.41a)\\
\left(\frac{\partial f}{\partial \phi}\right)_\rho &= -\rho\sin\phi\left(\frac{\partial f}{\partial x}\right)_y + \rho\cos\phi\left(\frac{\partial f}{\partial y}\right)_x \\
&= -y\left(\frac{\partial f}{\partial x}\right)_y + x\left(\frac{\partial f}{\partial y}\right)_x. && (1.41b)
\end{aligned}$$

From solving (1.41a) and (1.41b) for f_x and f_y, or equivalently from calculating U^{-1} and using (1.40b),

$$\begin{aligned}
\left(\frac{\partial f}{\partial x}\right)_y &= \cos\phi\left(\frac{\partial f}{\partial \rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial f}{\partial \phi}\right)_\rho, && (1.42a)\\
\left(\frac{\partial f}{\partial y}\right)_x &= \sin\phi\left(\frac{\partial f}{\partial \rho}\right)_\phi + \frac{\cos\phi}{\rho}\left(\frac{\partial f}{\partial \phi}\right)_\rho. && (1.42b)
\end{aligned}$$

A Common Error. An alternative method of calculating (1.42a) and (1.42b) would be to use the chain rule in the form

$$\begin{aligned}
\left(\frac{\partial f}{\partial x}\right)_y &= \left(\frac{\partial f}{\partial \rho}\right)_\phi\left(\frac{\partial \rho}{\partial x}\right)_y + \left(\frac{\partial f}{\partial \phi}\right)_\rho\left(\frac{\partial \phi}{\partial x}\right)_y, \\
\left(\frac{\partial f}{\partial y}\right)_x &= \left(\frac{\partial f}{\partial \rho}\right)_\phi\left(\frac{\partial \rho}{\partial y}\right)_x + \left(\frac{\partial f}{\partial \phi}\right)_\rho\left(\frac{\partial \phi}{\partial y}\right)_x.
\end{aligned}$$

It is then *crucial* to observe that $\frac{\partial\rho}{\partial x}$ and $\frac{\partial\phi}{\partial x}$ are to be evaluated at constant y, and $\frac{\partial\rho}{\partial y}$ and $\frac{\partial\phi}{\partial y}$ are to be evaluated at constant x. After recalling that $\rho = (x^2+y^2)^{\frac{1}{2}}$ and $\phi = \tan^{-1}(y/x)$, it is straightforward to show that

$$\left(\frac{\partial \rho}{\partial x}\right)_y = \frac{x}{(x^2+y^2)^{\frac{1}{2}}} = \cos\phi, \quad \left(\frac{\partial \phi}{\partial x}\right)_y = -\frac{y}{x^2+y^2} = -\frac{\sin\phi}{\rho}, \quad \text{etc.}$$

A common ***error*** is to try to calculate $\left(\frac{\partial\rho}{\partial x}\right)_y$ as follows:

$$x = \rho\cos\phi \quad\Rightarrow\quad \rho = \frac{x}{\cos\phi} \quad\Rightarrow\quad \frac{\partial\rho}{\partial x} = \frac{1}{\cos\phi}.$$

The final expression is $\left(\frac{\partial\rho}{\partial x}\right)_\phi$ ***not*** the required $\left(\frac{\partial\rho}{\partial x}\right)_y$.

When evaluating a partial derivative remember to hold the right variables constant.

1.7 The Gradient of a Scalar Field

Let $f : \mathbb{R}^m \to \mathbb{R}$ be differentiable at $\mathbf{a} \in \mathbb{R}^m$, and let $(\mathbf{e}_1, \ldots, \mathbf{e}_m)$ be an orthonormal basis of $\mathbb{R}^m$.

Definition. The *gradient* of f at $\mathbf{a}$ is

$$\begin{aligned}
\text{grad} f = \boldsymbol{\nabla} f &= \sum_{j=1}^{m} \mathbf{e}_j \frac{\partial f}{\partial x_j}(\mathbf{a}) && (1.43a)\\
&= \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \frac{\partial f}{\partial x_2}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_m}(\mathbf{a})\right). && (1.43b)
\end{aligned}$$

3/00

Remark. For Cartesian co-ordinates and $f \equiv f(x,y,z)$, the gradient of f can be written

$$\begin{aligned}\boldsymbol{\nabla} f &= \mathbf{i}\,\frac{\partial f}{\partial x} + \mathbf{j}\,\frac{\partial f}{\partial y} + \mathbf{k}\,\frac{\partial f}{\partial z} && (1.44\text{a})\\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right). && (1.44\text{b})\end{aligned}$$

Example. Suppose that

$$f = \frac{1}{r} = \frac{1}{|\mathbf{x}|} = \frac{1}{(x^2+y^2+z^2)^{\frac{1}{2}}} \quad \text{for} \quad \mathbf{x} \neq 0\,; \tag{1.45}$$

note and remember that the function is singular, indeed not defined, at $\mathbf{x} = 0$. Then

$$\frac{\partial f}{\partial x} = -\frac{x}{(x^2+y^2+z^2)^{\frac{3}{2}}} = -\frac{x}{r^3}\,,$$

and so

$$\boldsymbol{\nabla} f = \left(-\frac{x}{r^3}, -\frac{y}{r^3}, -\frac{z}{r^3}\right) = -\frac{\mathbf{x}}{r^3}\,. \tag{1.46a}$$

Alternatively, using example (1.18),

$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{1}{r}\right) = -\frac{1}{r^2}\frac{\partial r}{\partial x_k} = -\frac{1}{r^2}\frac{x_k}{r} = -\frac{x_k}{r^3}\,. \tag{1.46b}$$

Properties of the Gradient. We shall just state a couple of these (they follow in a straightforward manner from, say, use of suffix notation). For differentiable functions $f : \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}$

$$\begin{aligned}\boldsymbol{\nabla}(\lambda f + \mu g) &= \lambda \boldsymbol{\nabla} f + \mu \boldsymbol{\nabla} g \quad \text{for} \quad \lambda, \mu \in \mathbb{R}\,,\\ \boldsymbol{\nabla}(fg) &= g \boldsymbol{\nabla} f + f \boldsymbol{\nabla} g\,.\end{aligned}$$

3/99

1.7.1 The Gradient is a Vector

Revision: Orthonormal Basis Vectors.

Suppose that for $\mathbb{R}^m$ $\mathbf{e}_1, \ldots, \mathbf{e}_m$ are orthonormal basis vectors with Cartesian co-ordinates $\mathbf{x}$,
and that $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ are orthonormal basis vectors with Cartesian co-ordinates $\mathbf{x}'$.

From *Algebra & Geometry*, if $\mathbf{v} \in \mathbb{R}^m$ then (using the *s.c.*)

$$\mathbf{v} = v_j \mathbf{e}_j = (\mathbf{v}.\mathbf{e}_j)\mathbf{e}_j\,, \tag{1.47a}$$

since

$$\mathbf{v}.\mathbf{e}_j = (v_k \mathbf{e}_k).\mathbf{e}_j = v_k\,(\mathbf{e}_k.\mathbf{e}_j) = v_k \delta_{kj} = v_j\,. \tag{1.47b}$$

In particular from (1.47a) for the vector $\mathbf{v} = \mathbf{e}'_k$, it follows that

$$\mathbf{e}'_k = (\mathbf{e}'_k.\mathbf{e}_j)\,\mathbf{e}_j\,. \tag{1.48}$$

Also, in terms of the $\mathbf{e}'_k$ basis

$$\mathbf{v} = v'_k \mathbf{e}'_k \quad \text{with} \quad v'_k = \mathbf{v}.\mathbf{e}'_k = (v_j \mathbf{e}_j).\mathbf{e}'_k = (\mathbf{e}_j.\mathbf{e}'_k) v_j\,. \tag{1.49}$$

Hence for the vector $\mathbf{x}$

$$x'_k = (\mathbf{e}_j.\mathbf{e}'_k) x_j\,, \tag{1.50a}$$

and so using (1.17)

$$\frac{\partial x'_k}{\partial x_i} = (\mathbf{e}_j.\mathbf{e}'_k)\frac{\partial x_j}{\partial x_i} = (\mathbf{e}_j.\mathbf{e}'_k)\delta_{ji} = \mathbf{e}_i.\mathbf{e}'_k\,. \tag{1.50b}$$

The Result. We wish to show that $\boldsymbol{\nabla} f$ is independent of its basis vectors, i.e. that

$$\boldsymbol{\nabla} f = \mathbf{e}_j \frac{\partial f}{\partial x_j} \quad \text{and} \quad \boldsymbol{\nabla} f = \mathbf{e}'_k \frac{\partial f}{\partial x'_k} \qquad (s.c.)$$

are the same vector. To this end we note that

$$\begin{aligned}
\mathbf{e}_j \frac{\partial f}{\partial x_j} &= \mathbf{e}_j \frac{\partial x'_k}{\partial x_j}\frac{\partial f}{\partial x'_k} && \text{using the chain rule (1.33b)} \\
&= \mathbf{e}_j\,(\mathbf{e}_j.\mathbf{e}'_k)\,\frac{\partial f}{\partial x'_k} && \text{using (1.50b)} \\
&= \mathbf{e}'_k \frac{\partial f}{\partial x'_k} && \text{using (1.48)}\,.
\end{aligned}$$

□

Remark. Similarly (*s.c.*)

$$\mathbf{e}_j \frac{\partial}{\partial x_j} = \mathbf{e}_j \frac{\partial x'_k}{\partial x_j}\frac{\partial}{\partial x'_k} = \mathbf{e}_j\,(\mathbf{e}_j.\mathbf{e}'_k)\,\frac{\partial}{\partial x'_k} = \mathbf{e}'_k \frac{\partial}{\partial x'_k}\,.$$

Hence the differential operator

$$\boldsymbol{\nabla} = \mathbf{e}_j \frac{\partial}{\partial x_j} \qquad (s.c.) \tag{1.51}$$

is a vector (note the order of $\mathbf{e}_j$ and $\frac{\partial}{\partial x_j}$). This operator is referred to as 'del' or 'nabla' (after a harp-like ancient Assyrian/Hebrew musical instrument of similar shape).

1.7.2 Directional Derivatives Revisited

As shown below the directional derivative of a differentiable scalar function $f : \mathbb{R}^m \to \mathbb{R}$ at $\mathbf{a}$ in the direction $\widehat{\mathbf{u}}$ is the component of the gradient vector at $\mathbf{a}$ in the direction $\widehat{\mathbf{u}}$, i.e.

$$f'(\mathbf{a}; \widehat{\mathbf{u}}) = \widehat{\mathbf{u}}.\boldsymbol{\nabla} f\,. \tag{1.52}$$

Thus the directional derivative has a maximum when $\widehat{\mathbf{u}}$ is parallel to $\boldsymbol{\nabla} f$, i.e. $\boldsymbol{\nabla} f$ is in the direction of the maximum rate of change of f.

To derive (1.52) start from the definition of a directional derivative (1.20), and use the definition of differentiability (1.24):[3]

$$f'(\mathbf{a};\widehat{\mathbf{u}}) \quad = \quad \lim_{h\to 0}\frac{f(\mathbf{a}+h\widehat{\mathbf{u}})-f(\mathbf{a})}{h} \;=\; \mathsf{T}\widehat{\mathbf{u}}\,.$$

For a scalar function f

$$\mathsf{T} = \left(\begin{array}{cccc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_m}\end{array}\right),$$

and hence (with a slightly cavalier attitude to column and row vectors)

$$\mathsf{T}\widehat{\mathbf{u}} = \frac{\partial f}{\partial x_i}\,\widehat{u}_i = \widehat{\mathbf{u}}.\boldsymbol{\nabla} f\,.$$

□

1.8 Partial Derivatives of Higher Order

The partial derivatives of the first-order partial derivatives of a function are called the second-order partial derivatives; e.g. for a function $f:\mathbb{R}^m\to\mathbb{R}$, the j^{th} partial derivative of $\frac{\partial f}{\partial x_i}$ is written

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j\partial x_i} = \left(f_{x_i}\right)_{x_j} = f_{x_i x_j}\,. \tag{1.53}$$

There is a straightforward generalisation to n^{th} order partial derivatives.

Notation. If $i\neq j$ then a second-order partial derivative is called a mixed second-order partial derivative.

Example. For the scalar function $f(x,y)=x\mathrm{e}^y$,

$$\begin{aligned} f_x &= \mathrm{e}^y\,, & f_{xy} &= \mathrm{e}^y\,, & f_{xx} &= 0\,,\\ f_y &= x\mathrm{e}^y\,, & f_{yx} &= \mathrm{e}^y\,, & f_{yy} &= x\mathrm{e}^y\,. \end{aligned}$$

Note that $f_{xy}=f_{yx}$, i.e. that the mixed second-order partial derivatives are equal.

Unlectured example showing that mixed partial derivatives need ***not*** *be equal.* For the scalar function

$$f(x,y) = \begin{cases} xy\left(\dfrac{x^2-y^2}{x^2+y^2}\right) & (x,y)\neq 0\\[2ex] \qquad 0 & (x,y)=0\end{cases}\quad,$$

[3] The proof of this result without using the definition of differentiability (1.24) is messier. For $m=3$ one can proceed along the lines:

$$\begin{aligned} f'(\mathbf{a};\widehat{\mathbf{u}}) &= \lim_{h\to 0}\frac{f(\mathbf{a}+h(\widehat{u}_1\mathbf{e}_1+\widehat{u}_2\mathbf{e}_2+\widehat{u}_3\mathbf{e}_3))-f(\mathbf{a})}{h}\\ &= \lim_{h\to 0}\Bigg\{\quad \widehat{u}_1\;\frac{f(\mathbf{a}+h(\widehat{u}_1\mathbf{e}_1+\widehat{u}_2\mathbf{e}_2+\widehat{u}_3\mathbf{e}_3))-f(\mathbf{a}+h(\widehat{u}_2\mathbf{e}_2+\widehat{u}_3\mathbf{e}_3))}{h\widehat{u}_1}\\ &\qquad\quad +\widehat{u}_2\;\frac{f(\mathbf{a}+h(\widehat{u}_2\mathbf{e}_2+\widehat{u}_3\mathbf{e}_3))-f(\mathbf{a}+h\widehat{u}_3\mathbf{e}_3)}{h\widehat{u}_2}\\ &\qquad\quad +\widehat{u}_3\;\frac{f(\mathbf{a}+h\widehat{u}_3\mathbf{e}_3)-f(\mathbf{a})}{h\widehat{u}_3}\Bigg\}\\ &= \widehat{u}_1\frac{\partial f}{\partial x}+\widehat{u}_2\frac{\partial f}{\partial y}+\widehat{u}_3\frac{\partial f}{\partial z}\,. \end{aligned}$$

we have from the definitions (1.13a) and (1.13b) of a partial derivative, that

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,y) &= \frac{(3x^2-y^2)y}{x^2+y^2} - \frac{2x^2y(x^2-y^2)}{(x^2+y^2)^2} = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} \qquad (x,y)\neq 0\,,\\
\frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0\,,\\
\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) &= \lim_{k\to 0}\frac{\frac{\partial f}{\partial x}(0,k)-\frac{\partial f}{\partial x}(0,0)}{k} = \lim_{k\to 0}\frac{\frac{k(0+0-k^4)}{(0+k^2)^2}-0}{k} = -1\,.
\end{aligned} \tag{1.54a}$$

Similarly

$$\begin{aligned}
\frac{\partial f}{\partial y}(x,y) &= -\frac{x(y^4+4y^2x^2-x^4)}{(y^2+x^2)^2} \qquad (x,y)\neq 0\,,\\
\frac{\partial f}{\partial y}(0,0) &= \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{0-0}{k} = 0\,,\\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) &= \lim_{h\to 0}\frac{\frac{\partial f}{\partial y}(h,0)-\frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h\to 0}\frac{\frac{-h(0+0-h^4)}{(0+h^2)^2}-0}{h} = +1\,.
\end{aligned} \tag{1.54b}$$

A comparison of (1.54a) and (1.54b) illustrates that mixed partial derivatives need **not** be equal.

1.8.1 A Sufficient Condition for the Equality of Mixed Partial Derivatives

For a function $f:\mathbb{R}^m\to\mathbb{R}$ suppose that the partial derivatives $\frac{\partial^2 f}{\partial x_j\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i\partial x_j}$ exist at all points sufficiently close to $\mathbf{a}$ and that they are continuous at $\mathbf{a}$, then

$$\frac{\partial^2 f}{\partial x_j\partial x_i}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_i\partial x_j}(\mathbf{a}).$$

Remarks.

- See Part IB *Analysis* for a proof.
- The condition can be relaxed to one of the mixed partial derivatives being continuous.
- Henceforth we shall assume, unless stated otherwise, that mixed partial derivatives are equal.

1.8.2 Change of Variables: Example

Under a change of variables, it is often necessary to transform second order, or higher order, partial derivatives. We illustrate the procedure with an example, namely the transformation from Cartesian co-ordinates (x,y), to plane polar co-ordinates (ρ,ϕ), where

$$x=\rho\cos\phi \quad\text{and}\quad y=\rho\sin\phi\,.$$

In particular suppose for a function f we wish to calculate $\left(\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}\right)$ in terms of the partial derivatives in plane polar co-ordinates. First, we recall from (1.42a) that

$$\left(\frac{\partial f}{\partial x}\right)_y = \cos\phi\left(\frac{\partial f}{\partial \rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial f}{\partial \phi}\right)_\rho\,,$$

or

$$\left(\frac{\partial}{\partial x}\right)_y = \cos\phi\left(\frac{\partial}{\partial\rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial}{\partial\phi}\right)_\rho . \tag{1.55}$$

Hence

$$\begin{aligned}
\left(\frac{\partial^2 f}{\partial x^2}\right)_y = \left(\frac{\partial}{\partial x}\right)_y\left(\frac{\partial f}{\partial x}\right)_y &= \left(\cos\phi\left(\frac{\partial}{\partial\rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial}{\partial\phi}\right)_\rho\right)\left(\frac{\partial f}{\partial x}\right)_y \\
&= \cos\phi\left(\frac{\partial}{\partial\rho}\right)_\phi\left(\cos\phi\left(\frac{\partial f}{\partial\rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial f}{\partial\phi}\right)_\rho\right) \\
&\quad - \frac{\sin\phi}{\rho}\left(\frac{\partial}{\partial\phi}\right)_\rho\left(\cos\phi\left(\frac{\partial f}{\partial\rho}\right)_\phi - \frac{\sin\phi}{\rho}\left(\frac{\partial f}{\partial\phi}\right)_\rho\right) \\
&= \cos^2\phi\left(\frac{\partial^2 f}{\partial\rho^2}\right)_\phi + \frac{\sin\phi\,\cos\phi}{\rho^2}\left(\frac{\partial f}{\partial\phi}\right)_\rho \\
&\quad - \frac{\sin\phi\,\cos\phi}{\rho}\left(\frac{\partial}{\partial\rho}\right)_\phi\left(\frac{\partial f}{\partial\phi}\right)_\rho + \frac{\sin^2\phi}{\rho}\left(\frac{\partial f}{\partial\rho}\right)_\phi \\
&\quad - \frac{\sin\phi\,\cos\phi}{\rho}\left(\frac{\partial}{\partial\phi}\right)_\rho\left(\frac{\partial f}{\partial\rho}\right)_\phi + \frac{\sin\phi\,\cos\phi}{\rho^2}\left(\frac{\partial f}{\partial\phi}\right)_\rho \\
&\quad + \frac{\sin^2\phi}{\rho^2}\left(\frac{\partial^2 f}{\partial\phi^2}\right)_\rho .
\end{aligned}$$

On assuming the equality of mixed derivatives we obtain

$$\frac{\partial^2 f}{\partial x^2} = \cos^2\phi\,\frac{\partial^2 f}{\partial\rho^2} + \frac{\sin^2\phi}{\rho}\,\frac{\partial f}{\partial\rho} - \frac{2\sin\phi\,\cos\phi}{\rho}\,\frac{\partial^2 f}{\partial\rho\partial\phi} + \frac{2\sin\phi\,\cos\phi}{\rho^2}\,\frac{\partial f}{\partial\phi} + \frac{\sin^2\phi}{\rho^2}\,\frac{\partial^2 f}{\partial\phi^2} . \tag{1.56a}$$

We can calculate $\frac{\partial^2 f}{\partial y^2}$ by observing that $x = \rho\sin\left(\phi + \frac{\pi}{2}\right)$ and $y = -\rho\cos\left(\phi + \frac{\pi}{2}\right)$. Hence from (1.56a), after applying the transformations $x \to -y$, $y \to x$, $\phi \to \left(\phi + \frac{\pi}{2}\right)$, we have that

$$\frac{\partial^2 f}{\partial y^2} = \sin^2\phi\,\frac{\partial^2 f}{\partial\rho^2} + \frac{\cos^2\phi}{\rho}\,\frac{\partial f}{\partial\rho} + \frac{2\sin\phi\,\cos\phi}{\rho}\,\frac{\partial^2 f}{\partial\rho\partial\phi} - \frac{2\sin\phi\,\cos\phi}{\rho^2}\,\frac{\partial f}{\partial\phi} + \frac{\cos^2\phi}{\rho^2}\,\frac{\partial^2 f}{\partial\phi^2} . \tag{1.56b}$$

and hence that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial\rho^2} + \frac{1}{\rho}\,\frac{\partial f}{\partial\rho} + \frac{1}{\rho^2}\,\frac{\partial^2 f}{\partial\phi^2} . \tag{1.57}$$

1.9 Taylor's Theorem

Recall from (1.4) that if $f : \mathbb{R} \to \mathbb{R}$ is differentiable $(N+1)$ times then Taylor's Theorem states that

$$f(a+h) = f(a) + \sum_{r=1}^{N} \frac{h^r}{r!}\,\frac{\mathrm{d}^r f}{\mathrm{d}x^r}(a) + R_N(h) ,$$

where for some $0 < \theta < 1$

$$R_N(h) = \frac{h^{N+1}}{(N+1)!}\,\frac{\mathrm{d}^{N+1} f}{\mathrm{d}x^{N+1}}(a+\theta h) .$$

There is a generalisation of this result to functions of several variables. Namely, if for a function $f : \mathbb{R}^m \to \mathbb{R}$ the partial derivatives of orders up to and including the $(N+1)^{\text{th}}$ are continuous at $\mathbf{a}$, then there is some $0 < \theta < 1$ such that

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \sum_{r=1}^{N} \frac{1}{r!}(\mathbf{h}.\boldsymbol{\nabla})^r f(\mathbf{a}) + R_N(\mathbf{h}) , \tag{1.58a}$$

where

$$R_N(\mathbf{h}) \quad = \quad \frac{1}{(N+1)!}(\mathbf{h}.\boldsymbol{\nabla})^{N+1} f(\mathbf{a}+\theta\mathbf{h})\,, \tag{1.58b}$$

and

$$\mathbf{h}.\boldsymbol{\nabla} \quad = \quad \sum_{j=1}^{m} h_j \frac{\partial}{\partial x_j} \;=\; h_j \frac{\partial}{\partial x_j} \qquad (s.c.)\,. \tag{1.58c}$$

If for a function $f : \mathbb{R}^m \to \mathbb{R}$ the partial derivatives of orders up to and including the N^{th} are continuous at $\mathbf{a}$, then there is the weaker result

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \sum_{r=1}^{N} \frac{1}{r!}(\mathbf{h}.\boldsymbol{\nabla})^r f(\mathbf{a}) + o(|\mathbf{h}|^N)\,. \tag{1.58d}$$

Remarks.

- See Part IB *Analysis* for a proof.
- Since f is sufficiently differentiable, the partial derivatives commute. This allows some simplification in the calculation of
$$(\mathbf{h}.\boldsymbol{\nabla})^r f \quad \text{for} \quad r = 1, \ldots N+1\,.$$
For instance, for $m = 2$ the binomial theorem gives
$$\begin{aligned}(\mathbf{h}.\boldsymbol{\nabla})^r f \quad &= \quad \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^r f \\ &= \quad \sum_{s=0}^{r} \frac{r!\, h_1^s h_2^{r-s}}{(r-s)!s!}\, \frac{\partial^r f}{\partial x_1^s \partial x_2^{r-s}}\,,\end{aligned} \tag{1.59a}$$
while the summation convention yields
$$\begin{aligned}(\mathbf{h}.\boldsymbol{\nabla})^r \quad &= \quad \left(h_{j_1} \frac{\partial}{\partial x_{j_1}}\right)\left(h_{j_2} \frac{\partial}{\partial x_{j_2}}\right) \ldots \left(h_{j_r} \frac{\partial}{\partial x_{j_r}}\right) \\ &= \quad h_{j_1} h_{j_2} \ldots h_{j_r} \frac{\partial}{\partial x_{j_1}} \frac{\partial}{\partial x_{j_2}} \ldots \frac{\partial}{\partial x_{j_r}}\,.\end{aligned} \tag{1.59b}$$
- A form of Taylor's Theorem holds for $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ by applying the above result componentwise:
$$f_i(\mathbf{a}+\mathbf{h}) = f_i(\mathbf{a}) + \sum_{r=1}^{N} \frac{1}{r!}(\mathbf{h}.\boldsymbol{\nabla})^r f_i(\mathbf{a}) + \ldots\,. \tag{1.60}$$

4/99
4/00

Example. Suppose that $f : \mathbb{R}^3 \to \mathbb{R}$ is given by

$$f(\mathbf{x}) = \frac{1}{r} \quad \text{for} \quad r \neq 0\,,$$

where $r = |\mathbf{x}|$. Then, from using example (1.46b),

$$(\mathbf{h}.\boldsymbol{\nabla}) f(\mathbf{x}) \quad = \quad h_j \frac{\partial}{\partial x_j}\left(\frac{1}{r}\right) = -\frac{h_j x_j}{r^3} \tag{1.61a}$$

$$\begin{aligned}(\mathbf{h}.\boldsymbol{\nabla})^2 f(\mathbf{x}) \quad &= \quad (\mathbf{h}.\boldsymbol{\nabla})((\mathbf{h}.\boldsymbol{\nabla})f) \\ &= \quad h_k \frac{\partial}{\partial x_k}\left(-\frac{h_j x_j}{r^3}\right) \\ &= \quad -h_j h_k x_j . \left(-\frac{3}{r^4}\right) . \frac{x_k}{r} - h_j h_k \left(\frac{\delta_{jk}}{r^3}\right) \\ &= \quad \frac{h_j h_k}{r^5}\left(3 x_j x_k - r^2 \delta_{jk}\right)\,.\end{aligned} \tag{1.61b}$$

Hence for $|\mathbf{h}| \ll |\mathbf{x}|$,

$$\begin{aligned}\frac{1}{|\mathbf{x}+\mathbf{h}|} = f(\mathbf{x}+\mathbf{h}) &= f(\mathbf{x}) + (\mathbf{h}.\boldsymbol{\nabla})f(\mathbf{x}) + \tfrac{1}{2}(\mathbf{h}.\boldsymbol{\nabla})^2 f(\mathbf{x}) + \dots \\ &= \frac{1}{r} - \frac{h_j x_j}{r^3} + \frac{h_j h_k}{2r^5}\left(3x_j x_k - r^2\delta_{jk}\right) + \dots\,. \end{aligned} \tag{1.62}$$

We will make use of this result when we come to discuss monopoles, dipoles, quadrupoles, etc. at the end of the course.

2 Curves and Surfaces

2.1 Curves

A curve may be represented parametrically by

$$\mathbf{x} = \boldsymbol{\psi}(t)\,, \tag{2.1}$$

where $\mathbf{x}$ is a point on the curve, and t is a parameter that varies along the curve.

Remarks.

- Pure mathematicians define a 'curve' (or 'path' if the curve is 'smooth' enough) to be a mapping, while the image of the mapping is a 'set of points of the curve'. Applied mathematicians tend to identify the set of points with the curve.
- The parametrisation of a given curve is not unique.
- We will concentrate on $\mathbb{R}^3$, but there are straightforward extensions to $\mathbb{R}^m$.
- One of our aims is to identify what characterises a curve.

Examples.

1. $$\mathbf{x} = \boldsymbol{\psi}(t) = (at, bt^2, 0)\,.$$
   This is the parabola $a^2y = bx^2$ in the xy-plane.

2. $$\mathbf{x} = \boldsymbol{\psi}(t) = (\mathrm{e}^t\cos t, \mathrm{e}^t \sin t, 0)\,.$$
   This is a spiral.

3. $$\mathbf{x} = \boldsymbol{\psi}(t) = (\cos t, \sin t, t)\,.$$
   This is a helix on the cylinder of radius one, i.e.
   the cylinder $\rho^2 = x^2 + y^2 = 1$.

4. Suppose

$$\mathbf{x} \;=\; (\rho(t)\,\cos\phi(t),\; \rho(t)\,\sin\phi(t),\; z(t))\,,$$

where

$$\rho \;=\; \gamma + \cos\alpha t\,, \quad \phi = \beta t\,, \quad z = \sin\alpha t\,,$$

and α, β and $\gamma > 1$ are constants. This curve lies on the torus $(\rho-\gamma)^2 + z^2 = 1$.

Unlectured paragraph discussing when the curve is closed. For the curve defined by $0 \leqslant t \leqslant t_c$, to be closed it is necessary that the curve

(a) returns to $x = \gamma + 1$ at t_c, which implies that

$$(\gamma + \cos\alpha t_c)\cos\beta t_c = \gamma + 1\,,$$

i.e. that $\alpha t_c = p\pi$ and $\beta t_c = q\pi$, where p and q are integers;

(b) returns to $y = 0$ at t_c, which is ensured if $\beta t_c = q\pi$ since $\sin q\pi = 0$;

(c) returns to $z = 0$ at t_c, which is ensured if $\alpha t_c = p\pi$ since $\sin p\pi = 0$.

It follows from the above three results that for the curve to be closed it is necessary that

$$\frac{\alpha}{\beta} = \frac{p}{q}\,,$$

i.e. that α/β is rational.

Reading Exercise. What happens if α/β is irrational?

Definitions.

- A *simple* curve is one that does not intersect or touch itself.

simple intersects touches

- A curve is *smooth* if the derivative $\boldsymbol{\psi}'(t)$ exists. It is *piecewise smooth* if the curve can be divided into a finite number of smooth curves.

smooth piecewise smooth

2.1.1 The Length of a Simple Curve

Let $\mathcal{C}$ be a simple curve defined by

$$\mathbf{x} = \boldsymbol{\psi}(t) \quad \text{for} \quad t_a \leqslant t \leqslant t_b.$$

Form a *dissection* D by splitting $[t_a, t_b]$ into N intervals

$$t_a = t_0 < t_1 < \ldots < t_N = t_b\,.$$

The points $\{\boldsymbol{\psi}(t_0), \boldsymbol{\psi}(t_1), \ldots \boldsymbol{\psi}(t_N)\}$ can be viewed as the vertices of an inscribed polygonal path. The length of this path is

$$\ell_D = \sum_{j=1}^{N} |\boldsymbol{\psi}(t_j) - \boldsymbol{\psi}(t_{j-1})|,. \tag{2.2}$$

Define

$$\ell(\mathcal{C}) = \sup_D \ell_D\,.$$

Then if $\ell(\mathcal{C}) < \infty$, $\mathcal{C}$ is said to be *rectifiable* and $\ell(\mathcal{C})$ is the length of the curve.

Remarks.

1. If $\mathcal{C}$ is not simple, but can be split into a finite number of simple curves $\mathcal{C}_j$ of length $\ell(\mathcal{C}_j)$, then

$$\ell(\mathcal{C}) = \sum_j \ell(\mathcal{C}_j)\,.$$

2. If $\boldsymbol{\psi}$ is continuously differentiable (i.e. if the vector derivative $\boldsymbol{\psi}'(t) = (\psi_1'(t), \psi_2'(t), \psi_3'(t))$ is continuous), then

$$\ell = \int_{t_a}^{t_b} |\boldsymbol{\psi}'(t)|\,\mathrm{d}t\,. \tag{2.3}$$

This result can be derived from (2.2) using the definition of a Riemann integral and the idea of uniform continuity. See also the comments after equation (2.5).

2.1.2 The Arc Length

Definition. The *arc length [function]* of a smooth curve $\mathcal{C}$ parameterised by $\boldsymbol{\psi}(t)$ is defined to be

$$s(t) = \int_{t_a}^{t} |\boldsymbol{\psi}'(\tau)|\,\mathrm{d}\tau\,. \tag{2.4}$$

Remarks

1. At one end of the curve the arc length is zero, while at the other end the arc length equals the length of the curve, i.e. from (2.3)

$$s(t_a) = 0\,, \quad \text{and} \quad s(t_b) = \ell\,.$$

2. The arc length is an increasing function since

$$\frac{\mathrm{d}s}{\mathrm{d}t} = |\boldsymbol{\psi}'(t)| \geqslant 0\,.$$

3. Using (2.1), i.e. $\mathbf{x} = \boldsymbol{\psi}(t)$, the infinitesimal arc length can be re-expressed (admittedly slightly cavalierly) as

$$\mathrm{d}s = |\boldsymbol{\psi}'(t)|\,\mathrm{d}t = \left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2\right]^{1/2} \mathrm{d}t = \left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right)^{\frac{1}{2}}\,. \tag{2.5}$$

This is as expected on geometrical grounds. Further, on geometrical grounds we would expect that

$$\begin{aligned} \ell &= \int (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2)^{1/2} \\ &= \int_{t_a}^{t_b} |\boldsymbol{\psi}'(t)|\,\mathrm{d}t \qquad \text{from using (2.5)}\,. \end{aligned}$$

This is reassuringly consistent with (2.3).

4. If $\frac{\mathrm{d}s}{\mathrm{d}t} = |\boldsymbol{\psi}'(t)| > 0$ then (by the inverse function theorem ... see this term's *Analysis* course) $s(t)$ is invertible, and there is an inverse function such that

$$t = \lambda(s)\,.$$

We can then *parametrise the curve in terms of* s:

$$\mathbf{x} = \boldsymbol{\psi}(t) = \boldsymbol{\psi}(\lambda(s)) = \mathbf{r}(s)\,. \tag{2.6}$$

Further, since from (2.4)

$$\frac{\mathrm{d}\lambda}{\mathrm{d}s} = \frac{\mathrm{d}t}{\mathrm{d}s} = \frac{1}{\mathrm{d}s/\mathrm{d}t} = \frac{1}{|\boldsymbol{\psi}'(t)|} = \frac{1}{|\boldsymbol{\psi}'(\lambda(s))|}\,,$$

it follows from using the chain rule that

$$\mathbf{r}'(s) = \boldsymbol{\psi}'(\lambda(s))\,\frac{\mathrm{d}\lambda}{\mathrm{d}s} = \frac{\boldsymbol{\psi}'(\lambda(s))}{|\boldsymbol{\psi}'(\lambda(s))|}\,,$$

and hence that (as required for consistency)

$$|\mathbf{r}'(s)| = 1\,. \tag{2.7}$$

2.1.3 The Tangent

Definition. Let L be the straight line through $\boldsymbol{\psi}(t)$ and $\boldsymbol{\psi}(t+h)$. The *tangent* at $\boldsymbol{\psi}(t)$ is the limiting position of L as $h \to 0$.

Remarks

- L has the direction of the vector
$$\tfrac{1}{h}(\boldsymbol{\psi}(t+h) - \boldsymbol{\psi}(t)).$$
Hence if
$$\boldsymbol{\psi}'(t) = \lim_{h\to 0} \frac{(\boldsymbol{\psi}(t+h) - \boldsymbol{\psi}(t))}{h}$$
is non-zero, $\boldsymbol{\psi}'(t)$ has the direction of the tangent at $\boldsymbol{\psi}(t)$.
- For a smooth curve $\mathcal{C}$ parameterised using arc length, it follows from (2.7) that $\mathbf{r}'(s) \neq 0$; and moreover that
$$\mathbf{r}' \equiv \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \tag{2.8}$$
is the *unit tangent vector*.

2.1.4 The Osculating Plane

If $\boldsymbol{\psi}''(t) \neq 0$, then three neighbouring points, say t_1, t_2 and t_3, on a curve determine a plane. The limiting plane as t_2, t_3 tend to t_1 is called the *osculating plane* at the point t_1.

The Normal of the Osculating Plane. A normal to the plane defined by t_1, t_2 and t_3 is given by

$$\mathbf{n} = (\boldsymbol{\psi}(t_2) - \boldsymbol{\psi}(t_1)) \times (\boldsymbol{\psi}(t_3) - \boldsymbol{\psi}(t_1))\,.$$

Let $t_2 = t_1 + \delta t_2$ and $t_3 = t_1 + \delta t_3$, then using Taylor's theorem (1.60) for vector functions, or alternatively (1.4) applied componentwise,

$$\begin{aligned} \boldsymbol{\psi}(t_2) - \boldsymbol{\psi}(t_1) &= \boldsymbol{\psi}(t_1 + \delta t_2) - \boldsymbol{\psi}(t_1) \\ &= \delta t_2\, \boldsymbol{\psi}'(t_1) + \tfrac{1}{2}\,\delta t_2^2\, \boldsymbol{\psi}''(t_1) + \ldots\,, \end{aligned}$$

and so

$$\begin{aligned} \mathbf{n} &= (\delta t_2\, \boldsymbol{\psi}'(t_1) + \tfrac{1}{2}\,\delta t_2^2\, \boldsymbol{\psi}''(t_1) + \ldots) \times (\delta t_3\, \boldsymbol{\psi}'(t_1) + \tfrac{1}{2}\,\delta t_3^2\, \boldsymbol{\psi}''(t_1) + \ldots) \\ &= \tfrac{1}{2}\,\delta t_2 \delta t_3 (\delta t_3 - \delta t_2)(\boldsymbol{\psi}'(t_1) \times \boldsymbol{\psi}''(t_1)) + \ldots\,. \end{aligned}$$

5/99 Hence the unit normal of the osculating plane at t_1 is given, up to a sign, by

$$\widehat{\mathbf{n}} = \frac{\boldsymbol{\psi}'(t_1) \times \boldsymbol{\psi}''(t_1)}{|\boldsymbol{\psi}'(t_1) \times \boldsymbol{\psi}''(t_1)|}, \tag{2.9}$$

5/00 and the plane itself has the equation

$$(\mathbf{x} - \boldsymbol{\psi}(t_1)).\widehat{\mathbf{n}} = 0 .$$

Remark. We should check under what conditions the normal is well defined, i.e. we should determine under what conditions $\boldsymbol{\psi}'(t) \times \boldsymbol{\psi}''(t) \neq 0$.

We have already seen that there is an advantage in parameterising the curve using arc length, since we then know from (2.7) that $\mathbf{r}'(s) \neq 0$. Moreover, if we define

$$\mathbf{u}(s) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}(s) ,$$

then because $\mathbf{u}.\mathbf{u} = 1$ from (2.7), it follows from using our earlier result (1.12) for the derivative of a scalar product, that

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}(\mathbf{u}.\mathbf{u}) = \frac{\mathrm{d}}{\mathrm{d}s}(u_j u_j) = 2u_j \frac{\mathrm{d}u_j}{\mathrm{d}s} = 2\mathbf{u}.\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} , \tag{2.10a}$$

and that $\mathbf{u}$ is perpendicular to

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s}(s) = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}(s) . \tag{2.10b}$$

Hence, except when $\mathbf{r}''(s) = 0$, we conclude that $\mathbf{r}'(s) \times \mathbf{r}''(s) \neq 0$ and that the normal of the osculating plane is well defined.

2.1.5 The Serret-Frenet Equations

The first equation is

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} = \kappa\mathbf{p} , \tag{2.11}$$

where κ is defined by

$$\kappa(s) = |\mathbf{r}''(s)| \geqslant 0 ,$$

and $\mathbf{p}$ is a unit vector.

Definitions.

Curvature. $\kappa(s)$ is referred to as the *curvature*, while $\rho(s) = 1/\kappa(s)$ is referred to as the *radius of curvature*.

Principle Normal. If $\kappa > 0$, i.e. if $\mathbf{r}''(s) \neq 0$, then $\mathbf{p}$ is defined to be the *unit principle normal*.

Bi-normal. Let

$$\mathbf{b} = \mathbf{u} \times \mathbf{p} = \frac{\mathbf{r}'(s) \times \mathbf{r}''(s)}{\kappa} . \tag{2.12}$$

Then $\mathbf{b}$ is defined to be the *unit bi-normal vector*. From (2.9) it follows that $\mathbf{b}$ is, up to a sign, the unit normal of the osculating plane.

Moving Trihedral. Since $\mathbf{u}$ is perpendicular to $\mathbf{p}$ from (2.10a) and (2.11), it follows from (2.12) that $\mathbf{u}$, $\mathbf{p}$ and $\mathbf{b}$ form a right handed co-ordinate system — the *moving trihedral*.

Remark. Knowledge of the the curvature is necessary when calculating certain physical forces such as surface tension or bending moments.

Figure of tangent, osculating plane, bi-normal, etc. from Kreyszig.

Torsion. Since $\mathbf{b}.\mathbf{b} = 1$ it follows as above that

$$\mathbf{b} \cdot \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} = 0\,. \tag{2.13}$$

Further, since $\mathbf{u}$, $\mathbf{p}$ and $\mathbf{b}$ are mutually orthogonal, and in particular $\mathbf{u}.\mathbf{b} = 0$ and $\mathbf{p}.\mathbf{b} = 0$, it follows using definition (2.11) of $\mathbf{p}$ that

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}(\mathbf{u} \cdot \mathbf{b}) = \mathbf{u} \cdot \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} + \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} \cdot \mathbf{b} = \mathbf{u} \cdot \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} + \kappa\mathbf{p} \cdot \mathbf{b} = \mathbf{u} \cdot \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s}\,. \tag{2.14}$$

Thus $\frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s}$ is perpendicular to both $\mathbf{b}$ and $\mathbf{u}$, i.e. $\frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s}$ is parallel to $\mathbf{p}$. We write

$$\frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} = -\tau\mathbf{p}\,, \tag{2.15}$$

where τ is defined to be the *torsion*; this is the second Serret-Frenet equation.

The Third Equation. Further, from (2.12)

$$\mathbf{b} \times \mathbf{u} = (\mathbf{u} \times \mathbf{p}) \times \mathbf{u} = (\mathbf{u} \cdot \mathbf{u})\,\mathbf{p} - (\mathbf{p} \cdot \mathbf{u})\,\mathbf{u} = \mathbf{p}\,.$$

Hence

$$\begin{aligned} \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} &= \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} \times \mathbf{u} + \mathbf{b} \times \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} \\ &= -\tau\,\mathbf{p} \times \mathbf{u} + \kappa\,\mathbf{b} \times \mathbf{p}\,. \end{aligned}$$

i.e. from using (2.12) again

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = \tau\,\mathbf{b} - \kappa\,\mathbf{u}\,. \tag{2.16}$$

This is the third Serret-Frenet equation.

Summary. To summarise the three Serret-Frenet equations are

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} = \kappa\mathbf{p}\,, \tag{2.17a}$$

$$\frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} = -\tau\mathbf{p}\,, \tag{2.17b}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = \tau\mathbf{b} - \kappa\mathbf{u}\,. \tag{2.17c}$$

Remarks.

- Since $\mathbf{p} = \mathbf{b} \times \mathbf{u}$ we can write the above vector system of equations as

  $$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} = \kappa(\mathbf{b} \times \mathbf{u})\,, \qquad \frac{\mathrm{d}\mathbf{b}}{\mathrm{d}s} = \tau(\mathbf{u} \times \mathbf{b})\,.$$

  This system can be viewed as six scalar equations in six scalar unknowns.[4] Hence if $\kappa(s)$, $\tau(s)$, $\mathbf{u}(0)$ and $\mathbf{b}(0)$ are known, it is possible to solve for $\mathbf{u}(s)$ and $\mathbf{b}(s)$, and subsequently for $\mathbf{r}(s)$.

- $\kappa(s)$ and $\tau(s)$ contain all the information about the geometry of a curve up to translation and orientation.

[4] Alternatively we can reduce the system to four equations in four unknowns if we use the fact that $|\mathbf{u}| = |\mathbf{b}| = 1$.

2.2 Surfaces

A curve is the locus of a point with 1 degree of freedom.
A surface is the locus of a point with 2 degrees of freedom.

2.2.1 Representations

We will concentrate on $\mathbb{R}^3$.

Explicit Representation:

$$z = f(x, y)\,, \tag{2.18a}$$

e.g. a hemisphere of radius a

$$z = (a^2 - x^2 - y^2)^{1/2}\,.$$

Implicit Representation:

$$F(\mathbf{x}) = F(x, y, z) = 0\,, \tag{2.18b}$$

e.g.

- (2.18a) rewritten: $$z - f(x, y) = 0\,;$$
- a sphere of radius a: $$x^2 + y^2 + z^2 - a^2 = 0\,.$$

Parametric Representation:

$$\mathbf{x} \equiv \mathbf{x}(\mathbf{u}) = (x(u, v), y(u, v), z(u, v))\,, \tag{2.18c}$$

e.g. a sphere of radius a

$$\mathbf{x} = (a \sin u \cos v,\ a \sin u \sin v, a \cos u)\,.$$

Remark.

- Implicit and parametric representations are not unique.

2.2.2 Normals

Parametric Representation. Suppose a surface has the parametric representation

$$\mathbf{x} \equiv \mathbf{x}(u, v)\,.$$

Then the functions $\mathbf{x}(u, v_0)$ of u and $\mathbf{x}(u_0, v)$ of v, for fixed v_0 and u_0 respectively, describe curves in the surface.

Further, if we assume that both $\mathbf{x}(u, v_0)$ and $\mathbf{x}(u_0, v)$ are differentiable, then from the definition of a tangent in § 2.1.3, it follows that

a tangent at (u_0, v_0) to the line $v = v_0$ is $\frac{\partial \mathbf{x}}{\partial u}(u_0, v_0)$,

a tangent at (u_0, v_0) to the line $u = u_0$ is $\frac{\partial \mathbf{x}}{\partial v}(u_0, v_0)$.

Hence a *unit normal* to the surface at (u_0, v_0) is given by

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}\,. \tag{2.19}$$

Implicit Representation. For an implicit representation, let $\mathbf{x} = \boldsymbol{\psi}(t)$ describe a curve in the surface. Then, if the surface is described by $F(\mathbf{x}) = 0$,

$$F(\boldsymbol{\psi}(t)) = 0\,.$$

From the chain rule (1.30c) we conclude that

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}F(\boldsymbol{\psi}(t)) = \frac{\partial}{\partial x_j}F(\boldsymbol{\psi}(t))\,\frac{\mathrm{d}\psi_j}{\mathrm{d}t} = \boldsymbol{\nabla}F(\boldsymbol{\psi}(t))\cdot\frac{\mathrm{d}\boldsymbol{\psi}}{\mathrm{d}t}\,.$$

Hence if $\boldsymbol{\nabla}F \neq 0$, $\boldsymbol{\nabla}F$ is perpendicular to the tangent of the curve $\boldsymbol{\psi}(t)$. But $\boldsymbol{\psi}(t)$ can be any curve on the surface, hence $\boldsymbol{\nabla}F$ must be normal to the surface, i.e. up to a sign

$$\mathbf{n} = \frac{\boldsymbol{\nabla}F}{|\boldsymbol{\nabla}F|}\,. \tag{2.20}$$

Remarks.

- Since $\boldsymbol{\nabla}F$ is in the direction of the maximum rate of change of F, the normal to the surface is in the direction of the maximum rate of change of F.
- If $\mathbf{n}$ is a normal, then so is $(-\mathbf{n})$.
- For surfaces it is conventional to denote the *unit* normal by $\mathbf{n}$, rather than by, say, $\widehat{\mathbf{n}}$.
- The plane normal to $\mathbf{n}$ touching the surface is known as the *tangent plane.*

6/99

Examples.

1. Consider a surface of revolution defined parametrically by

$$\mathbf{x}(\phi, t) = (R(t)\cos\phi,\ R(t)\sin\phi, t)\ . \tag{2.21}$$

In cylindrical polar co-ordinates (ρ, ϕ, z),
the surface is given by

$$\begin{aligned} z &= t\,, \\ \rho = (x^2+y^2)^{\frac{1}{2}} &= R(z)\,. \end{aligned}$$

Further, from (2.21)

$$\begin{aligned} \mathbf{x}_\phi(\phi, t) &= (-R\sin\phi,\ R\cos\phi,\ 0)\,, \\ \mathbf{x}_t(\phi, t) &= (R'\cos\phi,\ R'\sin\phi,\ 1)\,, \end{aligned}$$

and hence from (2.19), after identifying ϕ and t with u and v respectively, a normal to the surface is given by

$$\begin{aligned}\mathbf{n} &= \frac{\mathbf{x}_\phi \times \mathbf{x}_t}{|\mathbf{x}_\phi \times \mathbf{x}_t|} = \frac{(R\cos\phi, R\sin\phi, -RR')}{|R(1+R'^2)^{\frac{1}{2}}|} \\ &= \frac{(\cos\phi, \sin\phi, -R')}{(1+R'^2)^{\frac{1}{2}}}\,. \end{aligned} \tag{2.22}$$

Remember that $-\mathbf{n}$ is also a normal.

6/00

2. Consider the surface with the implicit representation

$$F(\mathbf{x}) = \mathbf{a}\cdot\mathbf{x} - c = 0\,, \tag{2.23}$$

where $\mathbf{a} \in \mathbb{R}^3$ and $c \in \mathbb{R}$. Since

$$(\boldsymbol{\nabla} F)_i = \frac{\partial}{\partial x_i}(a_j x_j - c) = a_j \delta_{ij} = a_i\,, \quad \text{i.e.} \quad \boldsymbol{\nabla} F = \mathbf{a}\,,$$

the unit normal is given by

$$\mathbf{n} = \frac{\mathbf{a}}{|\mathbf{a}|}\,.$$

This is as expected since (2.23) is the equation of a plane with normal $\mathbf{a}$.

3. A surface is given by the implicit representation

$$F(x,y,z) = x^2 - y^2 - z^2 - c^2 = 0$$

where c is a real constant. Since

$$x^2 - \sigma^2 = c^2 \quad \text{where} \quad \sigma^2 = y^2 + z^2\,,$$

this is a surface of revolution about the x-axis. Further, since

$$\boldsymbol{\nabla} F = 2(x, -y, -z)\,,$$

the unit normal (up to a sign) is given by

$$\mathbf{n} = \tfrac{1}{r}(x, -y, -z)\,.$$

Unlectured remark. $\boldsymbol{\nabla} F = 0$ at $\mathbf{x} = 0$; hence if $c = 0$ so that the surface is given by

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

the normal at $(0,0,0)$ is not defined (but then this is the apex of a cone).

2.2.3 Definitions

Smooth Surfaces. A *smooth* surface is a surface with a unique (upto a sign) normal whose direction depends continuously on the points of the surface. A *piecewise smooth* surface is one that can be divided into finitely many portions each of which is smooth.

smooth | piecewise smooth

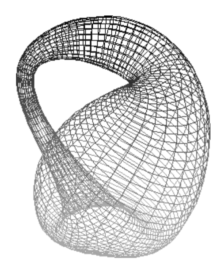
Remark.

- An alternative definition of a smooth surface for a given parametric representation, e.g. $\mathbf{x} \equiv \mathbf{x}(u, v)$ is that $\mathbf{x}_u$ and $\mathbf{x}_v$ are continuous and that $|\mathbf{x}_u \times \mathbf{x}_v| \neq 0$. However, with this definition the same surface can fail to be smooth at different points, depending on the parameterisation.

Orientable Surfaces. Let $\mathcal{S}$ be a smooth surface. Choose a unit normal $\mathbf{n}$ of $\mathcal{S}$ at a point P, i.e. fix on one of the two possible normals at point P. The direction of $\mathbf{n}$ is then said to be the *positive normal direction.*

The surface $\mathcal{S}$ is said to be *orientable* if the positive normal direction once given at an arbitrary point P can be continued in a unique and continuous way to the entire surface.

sphere | mobius strip | klein bottle



orientable | non-orientable | non-orientable

Open/Closed Surfaces. A surface can be either

closed, if 'any' straight line intersects the surface an even number of times (where tangents count twice or not at all), e.g. a sphere,

or *open,* if the surface is bounded by a curve, e.g. a mobius strip or a hemisphere.

A closed surface bounds a *convex* region if 'any' straight line intersects the surface a maximum of twice (where tangents count twice or not at all), e.g. a sphere.

The Orientation of a Bounding Curve. If an open orientable surface is bounded by a simple closed curve $\mathcal{C}$, then an orientation of $\mathcal{C}$ can be identified by the 'right-hand-rule'.

A piecewise smooth surface is orientable if the orientations of common boundaries can be arranged to be opposite.

2.3 Critical Points or Stationary Points

In this subsection we extend the ideas of maxima and minima for functions of one variable, i.e. $f : \mathbb{R} \to \mathbb{R}$, to functions of several variables, i.e. $f : \mathbb{R}^m \to \mathbb{R}$.

2.3.1 Definitions

Local Maximum or Minimum. A function $f : \mathbb{R}^m \to \mathbb{R}$ defined on a domain $\mathcal{D}$ is said to have a *local maximum* or *minimum* at $\mathbf{a} \in \mathcal{D}$ if for all $\mathbf{x} \in \mathcal{D}$ that are sufficiently close[5] to $\mathbf{a}$

$$f(\mathbf{x}) < f(\mathbf{a}) \quad \text{or} \quad f(\mathbf{x}) > f(\mathbf{a}) \tag{2.24}$$

respectively if $\mathbf{x} \neq \mathbf{a}$.

Weak Local Maximum or Minimum. If (2.24) is relaxed to

$$f(\mathbf{x}) \leqslant f(\mathbf{a}) \quad \text{or} \quad f(\mathbf{x}) \geqslant f(\mathbf{a}) \,,$$

then f is said to have a *weak local maximum* or *weak local minimum* respectively.

Extremum. An *extremum* is equivalent to a maximum or minimum, while a *weak extremum* is equivalent to a weak maximum or minimum.

[5] 'Sufficiently close' means that there exists $\delta > 0$ such that (2.24) holds for all $\mathbf{x} \in \mathcal{D}$ such that $\mathbf{x} \neq \mathbf{a}$ and $|\mathbf{x} - \mathbf{a}| < \delta$.

Remarks

- The above are *local* definitions. That a *local* maximum or minimum has been found does not imply that a *global* maximum or minimum has been found.

- For a function defined on a closed subset of $\mathbb{R}^m$, remember that the global maximum or minimum of the function may occur on a boundary.

- For a function defined on an open subset of $\mathbb{R}^m$ it may be more appropriate to seek local suprema or infima (*if* they exist) rather than local maxima or minima.

Critical Point or Stationary Point. If a function $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable sufficiently close to $\mathbf{a}$, then $\mathbf{a}$ is a *critical point*, or *stationary point*, if

$$\frac{\partial f}{\partial x_j}(\mathbf{a}) = 0 \quad \text{for} \quad j = 1, \ldots, m\,, \quad \text{i.e. if} \quad \boldsymbol{\nabla} f(\mathbf{a}) = 0\,. \tag{2.25}$$

Property of a Weak Extremum. If $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable, and f has a weak extremum at a point $\mathbf{a}$ in the *interior* of its domain of definition, then $\mathbf{a}$ is a critical point of f, i.e. $\boldsymbol{\nabla} f(\mathbf{a}) = 0$.

To see this suppose, without loss of generality, that f has a weak local minimum. Then for $|h|$ sufficiently small,

$$f(\mathbf{a} + h\mathbf{e}_j) \geqslant f(\mathbf{a})\,.$$

Hence

$$\lim_{h \to 0+} \frac{f(\mathbf{a} + h\mathbf{e}_j) - f(\mathbf{a})}{h} \quad \geqslant \quad 0\,,$$

and

$$\lim_{h \to 0-} \frac{f(\mathbf{a} + h\mathbf{e}_j) - f(\mathbf{a})}{h} \quad \leqslant \quad 0\,.$$

But f is differentiable, which implies that the above two limits must both be equal to $\frac{\partial f}{\partial x_j}(\mathbf{a})$ (*inter alia* see (1.24)); hence $\frac{\partial f}{\partial x_j}(\mathbf{a}) = 0$.

Remark. A weak extremum at a point $\mathbf{a}$ need not be a critical point if $\mathbf{a}$ lies on the boundary of the domain of f.

7/99 Saddle Point. A critical point that is not a weak extremum is called a *saddle point*.

Level Set. The set $L(c)$ defined by $L(c) = \{\mathbf{x} | f(\mathbf{x}) = c\}$ is called a *level set* of f. In $\mathbb{R}^2$ it is called a *level curve* or *contour*. In $\mathbb{R}^3$ it is called a *level surface*.

2.3.2 Functions of Two Variables

For a function of two variables, $f : \mathbb{R}^2 \to \mathbb{R}$, the function can be visualised by thinking of it as the surface with the explicit representation

$$z = f(x, y)\,;$$

z is constant on the level curves of f. Further, from the equivalent implicit representation

$$F(x, y, z) = z - f(x, y) = 0\,.$$

Figure of $z = 2 - x^2 - y^2$, $z = x^2 + y^2$, $z = xy$ from Apostol.

Figure of $z = x^3 = 3xy^2$, $z = x^2y^2$, $z = 1 - x^2$ from Apostol.

we note that

$$\boldsymbol{\nabla} F = (-f_x, -f_y, 1)\,.$$

Hence from (2.20) and the following remarks, the tangent plane is 'horizontal' at a critical point of f
7/00 (e.g. the tangent plane is 'horizontal' at the top of a 'hill').

Examples.

1. For the function

$$f = 2 - x^2 - y^2\,, \quad \text{we have that} \quad f_x = -2x\,, \quad f_y = -2y\,. \tag{2.26}$$

   Hence there is a critical point at $x = y = 0$. Further, since

$$f(x,y) - f(0,0) = f(x,y) - 2 = -x^2 - y^2 \leqslant 0\,,$$

   the critical point is a maximum.

2. For the function

$$f = x^3 - 3xy^2\,, \tag{2.27}$$

   we have that

$$f_x = 3x^2 - 3y^2\,, \quad f_y = -6xy\,.$$

   Hence, as in the previous example, there is a critical point at $x = y = 0$. However in this case we note that

$$f(x,0) - f(0,0) = x^3 \left\{ \begin{array}{ll} > 0 & \text{if } x > 0 \\ < 0 & \text{if } x < 0 \end{array} \right. ,$$

   and thus the critical point is a saddle point.

2.3.3 Maximum, Minimum or Saddle?

How do we determine if an interior critical point is a maximum, a minimum or a saddle point? First a definition.

Definition. Suppose for a function $f : \mathbb{R}^m \to \mathbb{R}$ that all the second-order partial derivatives exist, then define the *Hessian matrix* $\mathsf{H}(\mathbf{a})$ to have components

$$H_{ij}(\mathbf{a}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})\,. \tag{2.28}$$

Suppose now that a function $f : \mathbb{R}^m \to \mathbb{R}$ is continuously differentiable up to and including second order, and that it has a critical point at $\mathbf{a}$. We examine the variation in f near $\mathbf{a}$ in order to determine the nature of the critical point.

From Taylor's Theorem (1.58d) with $N = 2$, we have with the help of definition (2.25) of a critical point, i.e. $\frac{\partial f}{\partial x_j}(\mathbf{a}) = 0$, that near $\mathbf{a}$

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) &= f(\mathbf{a}) + 0 + \tfrac{1}{2}\left(\sum_{i=1}^m h_i \frac{\partial}{\partial x_i}\right)\left(\sum_{j=i}^m h_j \frac{\partial}{\partial x_j}\right) f(\mathbf{a}) + o(|\mathbf{h}|^2) \\ &= f(\mathbf{a}) + \tfrac{1}{2}\sum_{i,j=1}^m h_i \left(\frac{\partial^2}{\partial x_i \partial x_j}\right) f(\mathbf{a})\; h_j + o(|\mathbf{h}|^2)\,. \end{aligned} \tag{2.29}$$

Hence from using (2.28)

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \tfrac{1}{2}\sum_{i,j=1}^m h_i\, H_{ij}(\mathbf{a})\, h_j + o(|\mathbf{h}|^2) = \tfrac{1}{2}\mathbf{h}^t \mathsf{H}\mathbf{h} + o(|\mathbf{h}|^2)\,, \tag{2.30}$$

where expressions of the general type $\mathbf{h}^t\mathsf{H}\mathbf{h}$ are known as *quadratic forms*. We note that

- for small $|\mathbf{h}|$, we expect the sign of $(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}))$ to be same as that of the quadratic form;
- since f has been assumed to have continuous second order partial derivatives, the mixed second order partial derivatives commute, and hence, in this case, the Hessian matrix H is symmetric.

We can exploit the symmetry of the Hessian matrix in order to simplify the quadratic form. From *Algebra & Geometry* we know that (or at least in principle we know that):

1. the eigenvalues of H are real, say $\lambda_1, \ldots, \lambda_m$, although not necessarily distinct;
2. it is possible to find an orthonormal set of eigenvectors, say $\mathbf{e}^{(j)}$, which span $\mathbb{R}^m$ (even if the λ_j are not distinct), where

$$\mathsf{H}\mathbf{e}^{(j)} = \lambda_j \mathbf{e}^{(j)} \quad \text{for} \quad j = 1, \ldots, m \quad (\text{no } s.c.)\,. \tag{2.31}$$

Thus, because the $\mathbf{e}^{(j)}$ span $\mathbb{R}^m$ and so are a basis, we can write

$$\mathbf{h} = \sum_{j=1}^{m} h_j \mathbf{e}^{(j)}\,,$$

where, from (1.47b) and the fact that the $\mathbf{e}^{(j)}$ are orthonormal (i.e. $\mathbf{e}^{(i)}\cdot \mathbf{e}^{(j)} = \delta_{ij}$),

$$h_i = \mathbf{h}\cdot \mathbf{e}^{(i)}\,.$$

Hence from (2.31)

$$\mathsf{H}\mathbf{h} = \sum_{j=1}^{m} h_j \lambda_j \mathbf{e}^{(j)}\,,$$

and

$$\begin{aligned} \mathbf{h}^t \mathsf{H}\mathbf{h} &= \sum_{i=1}^{m} h_i \mathbf{e}^{(i)}\cdot \sum_{j=1}^{m} h_j \lambda_j \mathbf{e}^{(j)} \\ &= \sum_{i,j} h_i h_j \lambda_j \delta_{ij} && \text{since } \mathbf{e}^{(i)}\cdot \mathbf{e}^{(j)} = \delta_{ij} \\ &= \sum_{j=1}^{m} \lambda_j h_j^2\,. \end{aligned}$$

It follows from (2.30) that close to a critical point

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \tfrac{1}{2}\sum_{j=1}^{m} \lambda_j h_j^2 + o(|\mathbf{h}|^2)\,. \tag{2.32}$$

The Result.

1. If all the eigenvalues of $\mathsf{H}(\mathbf{a})$ are strictly positive, f has a local minimum at $\mathbf{a}$.
2. If all the eigenvalues of $\mathsf{H}(\mathbf{a})$ are strictly negative, f has a local maximum at $\mathbf{a}$.
3. If $\mathsf{H}(\mathbf{a})$ has at least one strictly positive eigenvalue and at least one strictly negative eigenvalue, then f has a saddle point at $\mathbf{a}$.
4. Otherwise, it is not possible to determine the nature of the critical point from the eigenvalues of $\mathsf{H}(\mathbf{a})$.

Outline Proof.

1. Suppose that all eigenvalues are positive, then
$$\lambda_{\min} = \min_j \lambda_j > 0\,.$$
Hence
$$\sum_{j=1}^{m} \lambda_j h_j^2 \geqslant \sum_{j=1}^{m} \lambda_{\min} h_j^2 = \lambda_{\min}|\mathbf{h}|^2\,,$$
and thus from (2.32)
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \geqslant \tfrac{1}{2}\lambda_{\min}|\mathbf{h}|^2 + o(|\mathbf{h}|^2)\,.$$
It follows for $|\mathbf{h}|$ sufficiently small, i.e. for $\mathbf{a}+\mathbf{h}$ sufficiently close to $\mathbf{a}$, that
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) > 0\,,$$
i.e. that $\mathbf{a}$ is a local minimum.

2. Apply the above argument to the function $-f$, since the transformation $f \to -f$ changes the sign of the eigenvalues.

3. Without loss of generality suppose that $\lambda_1 > 0$ and that $\lambda_2 < 0$. For $\mathbf{h} = (h_1, 0, \dots, 0)$ it follows from (2.32) that if h_1 is sufficiently small
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \tfrac{1}{2}\lambda_1 h_1^2 + o(h_1^2) > 0\,.$$
Similarly if $\mathbf{h} = (0, h_2, 0, \dots, 0)$, then if h_2 is sufficiently small
$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \tfrac{1}{2}\lambda_2 h_2^2 + o(h_2^2) < 0\,.$$
Hence $\mathbf{a}$ is not an extremum of f, which means that it's a saddle point.

Examples.

1. For the earlier example (2.26), i.e. $f = 2 - x^2 - y^2$, we have seen that there is a critical point at $(0,0)$. Further
$$\mathsf{H}(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$
Hence the eigenvalues of the Hessian are -2 and -2, i.e. both are strictly negative, which implies from above that the critical point is a maximum.

2. For the earlier example (2.27), i.e. $f = x^3 - 3xy^2$, we have seen that there is again a critical point at $(0,0)$. Further, since $f_x = 3x^2 - 3y^2$ and $f_y = -6xy$,
$$\mathsf{H}(x,y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix},$$
and
$$\mathsf{H}(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$
Hence both eigenvalues of the Hessian are 0; therefore we can deduce nothing from the eigenvalues of the Hessian (actually, as we saw before, the critical point is a saddle).

Unlectured remark. For the case when $f : \mathbb{R}^2 \to \mathbb{R}$, the calculation of the eigenvalues of the Hessian matrix is overkill since there is a simpler 'recipe' (see a question on *Examples Sheet 2*). In particular if $\Delta = \det \mathsf{H}(\mathbf{a})$ then

> if $\Delta > 0$ and $f_{xx}(\mathbf{a}) > 0$, f has a local minimum at $\mathbf{a}$; if $\Delta > 0$ and $f_{xx}(\mathbf{a}) < 0$, f has a local maximum at $\mathbf{a}$; if $\Delta < 0$, then f has a saddle point at $\mathbf{a}$; if $\Delta = 0$, the test is inconclusive.

3 Line Integrals and Exact Differentials

3.1 The Riemann Integral

First some revision of integration via a clarification of definitions (see IA *Analysis* for a formal treatment).

Dissection. A *dissection*, *partition* or *subdivision* D of the interval $[t_a, t_b]$ is a finite set of points $t_0, \ldots, t_N$ such that

$$t_a = t_0 < t_1 < \ldots < t_N = t_b\,.$$

Modulus. Define the *modulus*, *gauge* or *norm* of a dissection D, written $|D|$, to be the length of the longest subinterval $(t_j - t_{j-1})$ of D, i.e.

$$|D| = \max_{1\leqslant j\leqslant N} |t_j - t_{j-1}|\,. \tag{3.1a}$$

Riemann Sum. A *Riemann sum*, $\sigma(D, \boldsymbol{\zeta})$, for a bounded function $f : \mathbb{R} \to \mathbb{R}$ is any sum

$$\sigma(D, \boldsymbol{\zeta}) = \sum_{j=1}^{N} f(\zeta_j)(t_j - t_{j-1}) \quad \text{where} \quad \zeta_j \in [t_{j-1}, t_j]\,. \tag{3.1b}$$

Note that if the subintervals all have the same length so that $t_j = (a + jh)$ and $t_j - t_{j-1} = h$ where $h = (t_N - t_0)/N$, and if we take $\zeta_j = t_j$, then

$$\sigma(D, \boldsymbol{\zeta}) = \sum_{j=1}^{N} f(a + jh)\, h\,. \tag{3.1c}$$

Integrability. A bounded function $f : \mathbb{R} \to \mathbb{R}$ is *integrable* if there exists $I \in \mathbb{R}$ such that

$$\lim_{|D|\to 0} \sigma(D, \boldsymbol{\zeta}) = I\,, \tag{3.1d}$$

where the limit of the Riemann sum must exist independent of the dissection (subject to the condition that $|D| \to 0$) and independent of the choice of $\boldsymbol{\zeta}$ for a given dissection D.[6] Note that this is more restrictive than saying that the sum (3.1c) converges as $h \to 0$.

Definite Integral. For an integrable function f the *Riemann definite integral* of f over the interval $[t_a, t_b]$ is defined to be the limiting value of the Riemann sum, i.e.

$$\int_{t_a}^{t_b} f(t)\, \mathrm{d}t = I\,. \tag{3.1e}$$

An integral should be thought of as the limiting value of a sum, ***NOT*** *as the area under a curve.*

8/00

First Fundamental Theorem of Calculus. This states that the derivative of the integral of f is f, i.e. if f is suitably 'nice' (e.g. f is continuous) then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{t_a}^{x} f(t)\, \mathrm{d}t\right) = f(x)\,. \tag{3.1f}$$

[6] More precisely, f is integrable if given $\varepsilon > 0$, there exists $I \in \mathbb{R}$ and $\delta > 0$ such that whatever the choice of $\boldsymbol{\zeta}$ for a given dissection D

$$|\sigma(D, \boldsymbol{\zeta}) - I| < \varepsilon \quad \text{when} \quad |D| < \delta\,.$$

Second Fundamental Theorem of Calculus. This essentially states that the integral of the derivative of f is f, e.g. if f is differentiable then

$$\int_{t_a}^{t_b} \frac{\mathrm{d}f}{\mathrm{d}t}\,\mathrm{d}t = f(t_b) - f(t_a)\,. \tag{3.1g}$$

8/99

Substitution or Change of Variables. If $f(t)$ and $\frac{\mathrm{d}\lambda}{\mathrm{d}s}(s)$ are continuous functions of t and s respectively, then by setting $t = \lambda(s)$ and by using the chain rule

$$\int f(t)\,\mathrm{d}t = \int f(\lambda(s))\frac{\mathrm{d}\lambda}{\mathrm{d}s}\,\mathrm{d}s\,. \tag{3.1h}$$

Finally the statement above stands repeating

An integral should be thought of as the limiting value of a sum, ***NOT*** *as the area under a curve.*

Indeed, even repeating again.

Life will be simpler if you stop thinking, if you ever did, of an integral as the area under a curve.

Of course this is not to say that integrals are not a handy way of calculating the areas under curves.

3.2 Line Integrals

Let $\mathcal{C}$ be a smooth curve in $\mathbb{R}^3$ parameterised by arc length s:

$$\mathbf{x} = \mathbf{r}(s)\ . \tag{3.2}$$

3.2.1 Line Integrals of Scalar Functions

We could define a line integral by dissecting $\mathcal{C}$ and then taking a limit of a Riemann sum. However, this is unnecessary.

Definition. Suppose

$$f(\mathbf{x})\,:\,\mathbb{R}^3 \to \mathbb{R}\,.$$

Then the *line integral* of the scalar function f along the curve $\mathcal{C}$ is defined to be

$$\int_{\mathcal{C}} f(\mathbf{x})\,\mathrm{d}s = \int_{s_a}^{s_b} f(\mathbf{r}(s))\,\mathrm{d}s\,, \tag{3.3}$$

where the right hand side is a Riemann integral.

Remarks.

- Suppose instead that the curve is parameterised by other than the arc length, e.g. by t where $\mathbf{x} = \boldsymbol{\psi}(t)$ for $t_a \leqslant t \leqslant t_b$. From the definition of arc length (2.4) we have that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = |\boldsymbol{\psi}'(t)|\,, \tag{3.4}$$

and thus from the rules governing the change of variable in a Riemann integral

$$\int_{\mathcal{C}} f(\mathbf{x})\,\mathrm{d}s = \int_{s_a}^{s_b} f(\mathbf{r}(s))\,\mathrm{d}s = \int_{t_a}^{t_b} f(\boldsymbol{\psi}(t))\,|\boldsymbol{\psi}'(t)|\,\mathrm{d}t\,. \tag{3.5a}$$

Alternatively if we employ the more concise notation $\mathbf{x} \equiv \mathbf{x}(\mathbf{t})$, then from (2.5)

$$\int_{\mathcal{C}} f(\mathbf{x})\,\mathrm{d}s = \int_{t_a}^{t_b} f(\mathbf{x}(t)) \left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2\right]^{\frac{1}{2}} \mathrm{d}t\,. \tag{3.5b}$$

- Recall again that the definition of a line integral is best thought of a the limiting value of a Riemann sum. However, if you insist on thinking of integrals as areas, suppose that $\mathbf{x} = \boldsymbol{\psi}(t) = (\psi_1(t), \psi_2(t), 0)$ specifies the base position of a curvy wall, and $f(\mathbf{x})$ is the height of the wall at $\mathbf{x}$; then (3.5a) is the area of the wall.

Properties. For functions f and g, and constants λ and μ:

(i) $$\int_{-\mathcal{C}} f\,\mathrm{d}s = -\int_{\mathcal{C}} f\,\mathrm{d}s\,,$$

(ii) $$\int_{\mathcal{C}} (\lambda f + \mu g)\,\mathrm{d}s = \lambda \int_{\mathcal{C}} f\,\mathrm{d}s + \mu \int_{\mathcal{C}} g\,\mathrm{d}s\,,$$

(iii) $$\int_{\mathcal{C}} f\,\mathrm{d}s = \int_{\mathcal{C}_1} f\,\mathrm{d}s + \int_{\mathcal{C}_2} f\,\mathrm{d}s\,.$$

Example. Suppose that $f(x, y)$ is a scalar function defined by

$$f = 2xy\,.$$

We evaluate the line integral $\int f \mathrm{d}s$ for two different paths.

1. Consider first the path

$$\mathcal{C} \quad : \quad \mathbf{x} = (a\cos t, a\sin t), \qquad 0 \leqslant t \leqslant \pi/2\,,$$

where $a > 0$. Then

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = (-a\sin t, a\cos t)\,,$$

and hence from the definition of arc length (2.4)

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \left|\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right| = a\,.$$

Thus from (3.5a), or equivalently (3.5b),

$$\int_{\mathcal{C}} f\,\mathrm{d}s = \int_0^{\frac{\pi}{2}} 2a^2 \cos t \sin t\; a\,\mathrm{d}t = a^3\,.$$

2. Next consider the path $\widehat{\mathcal{C}} = \widehat{\mathcal{C}}_1 + \widehat{\mathcal{C}}_2$, where

$$\begin{aligned} \widehat{\mathcal{C}}_1 \quad &: \quad \mathbf{x} = (x, 0)\,, \qquad a \geqslant x \geqslant 0\,, \\ \widehat{\mathcal{C}}_2 \quad &: \quad \mathbf{x} = (0, y)\,, \qquad 0 \leqslant y \leqslant a\,. \end{aligned}$$

Then $f = 0$ on both $\widehat{\mathcal{C}}_1$ and $\widehat{\mathcal{C}}_2$, and hence

$$\int_{\widehat{\mathcal{C}}} f\,\mathrm{d}s = 0\,.$$

Remark.

- In general, if different paths $\mathcal{C}$ and $\widehat{\mathcal{C}}$ join the same end points, then the line integrals along these different paths are not equal, i.e. in general
$$\int_{\mathcal{C}} f\,\mathrm{d}s \neq \int_{\widehat{\mathcal{C}}} f\,\mathrm{d}s\,.$$
For instance if $f = 1$, then
$$\begin{aligned}\int_{\mathcal{C}} \mathrm{d}s &= \text{length of } \mathcal{C}\,,\\ \int_{\widehat{\mathcal{C}}} \mathrm{d}s &= \text{length of } \widehat{\mathcal{C}}\,,\end{aligned}$$
which for most choices of two distinct paths $\mathcal{C}$ and $\widehat{\mathcal{C}}$ are not equal.

3.2.2 Line Integrals of Vector Functions

For a vector function $\mathbf{F}(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}^3$, there are at least two types of line integral that can be defined along a path $\mathcal{C}$.

1. The first is a vector, and can be evaluated component by component using definition (3.3) of a line integral of a scalar function:
$$\mathbf{I} = \int_{\mathcal{C}} \mathbf{F}\,\mathrm{d}s = \int_{\mathcal{C}} \sum_{j=1}^{3} (F_j \mathbf{e}_j)\ \mathrm{d}s = \sum_{j=1}^{3} \mathbf{e}_j \int_{\mathcal{C}} F_j\,\mathrm{d}s = \sum_{j=1}^{3} \mathbf{e}_j \int_{s_a}^{s_b} F_j(\mathbf{r}(s))\,\mathrm{d}s\,. \tag{3.6}$$

2. The second integral is a scalar and is defined for a parameterisation t by
$$I = \int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \int_{t_a}^{t_b} \mathbf{F}(\mathbf{x}(t)).\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\mathrm{d}t\,, \tag{3.7a}$$
where
$$\mathrm{d}\mathbf{x} = \sum_{i=1}^{3} \mathbf{e}_i\,\mathrm{d}x_i = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)\,. \tag{3.7b}$$
Hence
$$\int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \int_{\mathcal{C}} (F_1\mathrm{d}x + F_2\mathrm{d}y + F_3\mathrm{d}z) = \int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{s}\,, \tag{3.8}$$
where in the final expression of (3.8) we have used the alternative notation $\mathrm{d}\mathbf{s} = \mathrm{d}\mathbf{x}$; note that then $\mathrm{d}s = |\mathrm{d}\mathbf{s}| = |\mathrm{d}\mathbf{x}|$.

Examples.

1. *Work.* Suppose that we identify $\mathbf{F}$ with the force acting on a particle transversing $\mathcal{C}$. The work done in moving the particle an infinitesimal distance $\mathrm{d}\mathbf{x}$ is known to be $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ (e.g. see IA *Dynamics*); this is the generalisation to 3D of the 1D formula
$$\texttt{work done = force} \times \texttt{distance}\,.$$
Hence the total work done in transversing the curve $\mathcal{C}$ is
$$W = \int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x}\,. \tag{3.9}$$

We can parameterise the curve by the time at which the particle is at a point. Then from (3.7a) and the use of Newton's second law of motion (i.e. $\mathbf{F} = m\ddot{\mathbf{x}}$, assuming no gravity, etc.)

$$\begin{aligned} W &= \int_{t_a}^{t_b} \mathbf{F}.\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\mathrm{d}t \\ &= \int_{t_a}^{t_b} \left(m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2}\right)\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\mathrm{d}t \\ &= \int_{t_a}^{t_b} \frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}m\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right)^2\right)\mathrm{d}t \\ &= \left[\tfrac{1}{2}m\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right)^2\right]_{t_a}^{t_b} \\ &= \text{gain in kinetic energy.} \end{aligned}$$

2. Suppose that $\mathbf{F}$ is a vector function defined by

$$\mathbf{F} = (y, x, 0)\,. \tag{3.10}$$

We evaluate the line integral $\int \mathbf{F}.\mathrm{d}\mathbf{x}$ for the family of paths $\mathcal{C}$ parameterised by $x = t$, $y = t^a$ where $a > 0$ and $0 \leqslant t \leqslant 1$. Then

$$\int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \int_0^1 \left(y\frac{\mathrm{d}x}{\mathrm{d}t} + x\frac{\mathrm{d}y}{\mathrm{d}t}\right)\mathrm{d}t = \int_0^1 (t^a + t\,a\,t^{a-1})\,\mathrm{d}t = \int_0^1 (a+1)t^a\,\mathrm{d}t = 1\,. \tag{3.11}$$

We note that the answer is independent of a, i.e. the same result is obtained for a number of different paths; this is ***not*** a co-incidence (see below). However, in general, if different paths join the same end points, then the line integrals of vector functions along these different paths are not equal.

Unlectured check. As a confirmation of the above result consider the $a \to \infty$ limiting path, i.e. the piecewise smooth curve $\widehat{\mathcal{C}} = \widehat{\mathcal{C}}_1 + \widehat{\mathcal{C}}_2$ defined by

$$\begin{aligned} \widehat{\mathcal{C}}_1 &: \quad x = t_1,\ y = 0 \quad \text{for} \quad 0 \leqslant t_1 \leqslant 1\,, \\ \widehat{\mathcal{C}}_2 &: \quad x = 1,\ y = t_2 \quad \text{for} \quad 0 \leqslant t_2 \leqslant 1\,. \end{aligned}$$

On $\widehat{\mathcal{C}}_1$

$$\mathbf{F}.\,\mathrm{d}\mathbf{x} = \mathbf{F}.\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t_1}\mathrm{d}t_1 = (0.1 + t_1.0)\,\mathrm{d}t_1 = 0\,,$$

while on $\widehat{\mathcal{C}}_2$

$$\mathbf{F}.\,\mathrm{d}\mathbf{x} = \mathbf{F}.\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t_2}\mathrm{d}t_2 = (t_2.0 + 1.1)\,\mathrm{d}t_2 = \mathrm{d}t_2\,.$$

Hence

$$\int_{\widehat{\mathcal{C}}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \int_0^1 \mathrm{d}t_2 = 1\,. \tag{3.12}$$

3. For a scalar differentiable function f consider

$$\begin{aligned} \int_{\mathcal{C}} \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x} &= \int_{t_a}^{t_b} \boldsymbol{\nabla} f(\mathbf{x}(t)).\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\mathrm{d}t \\ &\qquad\qquad \text{from definition (3.7a)} \\ &= \int_{t_a}^{t_b} \frac{\partial f}{\partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}t}\,\mathrm{d}t \\ &= \int_{t_a}^{t_b} \frac{\mathrm{d}}{\mathrm{d}t}\left(f(\mathbf{x}(t))\right)\,\mathrm{d}t \\ &\qquad\qquad \text{from the chain rule (1.30c)} \end{aligned}$$

$$= \quad f(\mathbf{x}_b) - f(\mathbf{x}_a)\,, \tag{3.13}$$

where the final line follows from the second fundamental theorem of calculus (3.1g). Hence, again, the value of the line integral is independent of the path.

Result (3.13) might be viewed as a [line] integral result for gradients, which is equivalent to the second fundamental theorem of calculus (3.1g) for derivatives (i.e. the line integral 'undoes' the gradient).

Exercise. Find the connection between the preceding two examples.

3.3 Differentials

Definition. Suppose the scalar function $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable at $\mathbf{a}$, then the *differential* $\mathrm{d}f$ of $f(x_1, \ldots, x_m)$ at $\mathbf{a}$ is defined to be

$$\mathrm{d}f(\mathbf{a}) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\mathbf{a})\,\mathrm{d}x_j\,. \tag{3.14a}$$

Alternatively, using the summation convention

$$\mathrm{d}f = \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j\,. \tag{3.14b}$$

Hence for $m = 1$, i.e. $f \equiv f(x)$

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x\,; \tag{3.15a}$$

while for $m = 2$, i.e. $f \equiv f(x, y)$

$$\mathrm{d}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y\,; \tag{3.15b}$$

etc.

Remarks.

- Definition (3.14a) is consistent for $f = x_i$, since from (3.14a), and using the *s.c.* and (1.17),

$$\mathrm{d}f = \frac{\partial x_i}{\partial x_j}\,\mathrm{d}x_j = \delta_{ij}\,\mathrm{d}x_j = \mathrm{d}x_i\,.$$

- Definition (3.14a) is sometimes referred to as the *chain rule for differentials.*

Unlectured example: rederivation of the chain rule (1.33b). Suppose that $f(\mathbf{x})$ and $\mathbf{x}(\mathbf{u})$ are differentiable. If $f(\mathbf{x}) = F(\mathbf{u})$, $\mathbf{u} = (u, v)$ and $\mathbf{x} = (x(u, v), y(u, v))$, then from (3.15b) we have for the change in f due to changes in x and y

$$\mathrm{d}f = f_x\,\mathrm{d}x + f_y\,\mathrm{d}y\,. \tag{3.16a}$$

After a notation change we also have from (3.15b) for changes in x and y due to changes in u and v,

$$\mathrm{d}x = x_u\,\mathrm{d}u + x_v\,\mathrm{d}v\,, \qquad \mathrm{d}y = y_u\,\mathrm{d}u + y_v\,\mathrm{d}v\,. \tag{3.16b}$$

Hence from substituting (3.16b) into (3.16a)

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

But it also follows from (3.15b) that the change in $f(x(u, v), y(u, v)) = F(u, v)$ due to changes in u and v is given by

$$\mathrm{d}f = \mathrm{d}F = F_u\,\mathrm{d}u + F_v\,\mathrm{d}v\,.$$

Hence from comparing the above two results and using the fact that $\mathrm{d}u$ and $\mathrm{d}v$ can be varied independently,

$$F_u = f_x\, x_u + f_y\, y_u \quad \text{and} \quad F_v = f_x\, x_v + f_y\, y_v\,.$$

Similarly if $f : \mathbb{R}^m \to \mathbb{R}$ and $\mathbf{x} : \mathbb{R}^\ell \to \mathbb{R}^m$

$$\mathrm{d}f = \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j\,, \quad \mathrm{d}x_j = \frac{\partial x_j}{\partial u_k}\,\mathrm{d}u_k \quad \text{and} \quad \mathrm{d}F = \frac{\partial F}{\partial u_k}\,\mathrm{d}u_k\,.$$

Hence

$$\frac{\partial F}{\partial u_k}\,\mathrm{d}u_k = \frac{\partial f}{\partial x_j}\,\frac{\partial x_j}{\partial u_k}\,\mathrm{d}u_k\,, \tag{3.17a}$$

or, since the $\mathrm{d}u_k$ can be varied independently,

$$\frac{\partial F}{\partial u_k} = \frac{\partial f}{\partial x_j}\,\frac{\partial x_j}{\partial u_k}\,. \tag{3.17b}$$

This is the same as (1.33b) after identifying $F(u, v)$ with $f(u, v)$.

9/00

3.3.1 Interpretations

Local Linear Approximation. Suppose that f is differentiable, and also suppose that we interpret $\mathrm{d}\mathbf{x}$ as the increment in $\mathbf{x}$ from $\mathbf{a}$, i.e. suppose we let

$$\mathrm{d}\mathbf{x} = (\mathbf{x} - \mathbf{a})\,.$$

Then it follows from the definition of a differentiable function, i.e. (1.24), that $\mathrm{d}f$ can be interpreted as the linear local *approximation* to the increment in f corresponding to an increment $\mathrm{d}\mathbf{x}$ in $\mathbf{x}$; i.e. for $m = 1$

$$f(a + \mathrm{d}x) \approx f(a) + \mathrm{d}f(a) = f(a) + \frac{\mathrm{d}f}{\mathrm{d}x}(a)\,\mathrm{d}x\,.$$

9/99

Relationship to Gradient. For a given fixed orthonormal basis, $\mathbf{e}_j$ $(j = 1, \ldots, m)$, the gradient of a scalar function f is defined in (1.43b) as

$$\boldsymbol{\nabla} f = \sum_{j=1}^{m} \mathbf{e}_j\,\frac{\partial f}{\partial x_j}\,.$$

Hence, using the generalisation of (3.7b) to $\mathbb{R}^m$

$$\mathrm{d}\mathbf{x} = \sum_{i=1}^{m} \mathbf{e}_i\,\mathrm{d}x_i\,,$$

and the property of an orthonormal basis that $\mathbf{e}_i.\,\mathbf{e}_j = \delta_{ij}$, it follows that

$$\boldsymbol{\nabla} f.\mathrm{d}\mathbf{x} = \sum_{i,j=1}^{m} \mathbf{e}_j\,\frac{\partial f}{\partial x_j}.\,\mathbf{e}_i\,\mathrm{d}x_i = \sum_{j}^{m} \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j = \mathrm{d}f\,. \tag{3.18}$$

Tangent Plane. If we again interpret $\mathrm{d}\mathbf{x}$ as the increment in $\mathbf{x}$ from $\mathbf{a}$, i.e. $\mathrm{d}\mathbf{x} = (\mathbf{x}-\mathbf{a})$, then the equation

$$\mathrm{d}f(\mathbf{a}) = 0 \tag{3.19a}$$

can be rewritten using (3.18) as

$$\mathrm{d}f(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a}).\mathrm{d}\mathbf{x} = \boldsymbol{\nabla} f(\mathbf{a}).(\mathbf{x} - \mathbf{a}) = 0\,. \tag{3.19b}$$

This is the equation of the plane with normal $\boldsymbol{\nabla} f(\mathbf{a})$ passing through $\mathbf{x} = \mathbf{a}$. Moreover from equation (2.20) for the normal to the surface $F(\mathbf{x}) = 0$, we conclude that (3.19a) is the equation of the tangent plane at $\mathbf{a}$ to the surface

$$F(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{a}) = 0\ .$$

Unlectured Example. For $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ suppose that

$$f(x, y) = y - g(x)\,.$$

Then the tangent 'plane'/line at $\mathbf{x} = \mathbf{a} = (a, b)$ to the 'surface'/curve $f(\mathbf{x}) = f(\mathbf{a})$ is given from (3.19a) by

$$\begin{aligned} 0 \;=\; \mathrm{d}f(\mathbf{a}) \;&=\; -\frac{\mathrm{d}g}{\mathrm{d}x}(a)\,\mathrm{d}x + \mathrm{d}y \\ &=\; -\frac{\mathrm{d}g}{\mathrm{d}x}(a)\,(x-a) + (y-b)\,. \end{aligned}$$

In other words the tangent to the curve

$$y \;=\; b + g(x) - g(a)$$

is

$$y \;=\; b + (x-a)\frac{\mathrm{d}g}{\mathrm{d}x}(a)\,.$$

3.3.2 Properties

For $f, g : \mathbb{R}^m \to \mathbb{R}$ and $\lambda, \mu \in \mathbb{R}$, it follows in a straightforward manner from the definition (3.14a) of a differential that

$$\mathrm{d}(\lambda f + \mu g) \;=\; \lambda \mathrm{d}f + \mu \mathrm{d}g\,, \tag{3.20a}$$

$$\mathrm{d}(fg) \;=\; f\,\mathrm{d}g + g\,\mathrm{d}f\,. \tag{3.20b}$$

Unlectured Example. From (3.20b) with $f = g = r$ it follows that $\mathrm{d}r^2 = 2r\mathrm{d}r$. Hence if

$$r^2 = x^2 + y^2 + z^2\,, \tag{3.21a}$$

then, using in addition (3.20a), it follows that

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

i.e.

$$\mathrm{d}r = \frac{x}{r}\,\mathrm{d}x + \frac{y}{r}\,\mathrm{d}y + \frac{z}{r}\,\mathrm{d}z\,. \tag{3.21b}$$

3.3.3 Exact Differentials

Definitions.

Differential Form. For vector functions $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$, the forms

$$\omega(\mathbf{x}) = \mathbf{F}(\mathbf{x}).\mathrm{d}\mathbf{x} = \sum_{j=1}^{m} F_j(\mathbf{x})\,\mathrm{d}x_j\,, \tag{3.22}$$

that appear in line integrals such as (3.7a) are called *differential [one-] forms*. For instance if $m = 3$:

$$\omega(\mathbf{x}) = F_1(x, y, z)\,\mathrm{d}x + F_2(x, y, z)\,\mathrm{d}y + F_3(x, y, z)\,\mathrm{d}z\,.$$

Exact Differential (2D). Suppose $P(x, y)$ and $Q(x, y)$ are defined in some domain $\mathcal{R}$ of $\mathbb{R}^2$, and that they have continuous first order partial derivatives. Then if for all $(x, y) \in \mathcal{R}$ there exists a *single valued* function $f : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\omega(x, y) = P(x, y)\,\mathrm{d}x + Q(x, y)\,\mathrm{d}y = \mathrm{d}f(x, y)\,, \tag{3.23}$$

then ω is an *exact differential [form]* on $\mathcal{R}$.

Simply Connected. A *simply connected domain* in $\mathbb{R}^2$ is one where any closed curve in the domain can be shrunk to a point in the domain without leaving the domain.

In $\mathbb{R}^3$ the definition is more complicated. First we define *path connected*: a domain $\mathcal{R}$ is said to be *path-connected* if for any two points $\mathbf{a}$ and $\mathbf{b}$ in $\mathcal{R}$ there exists a continuous function $\boldsymbol{\psi}$ from the unit interval $[0,1]$ to $\mathcal{R}$ defining a curve with $\boldsymbol{\psi}(0)=\mathbf{a}$ and $\boldsymbol{\psi}(1)=\mathbf{b}$. A domain is *simply connected* if it is path-connected and every path between two points can be continuously transformed into every other.

simply connected　　　　not simply connected

A Necessary Condition for ω to be an Exact Differential (2D).

If ω is an exact differential, i.e. if there exists a single valued function f such that $\omega = \mathrm{d}f$, then from (3.23) and the definition of a differential (3.15b)

$$P\,\mathrm{d}x + Q\,\mathrm{d}y = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y \ .$$

Since this condition must hold for arbitrary $\mathrm{d}\mathbf{x}$, it follows from considering, say, $\mathrm{d}\mathbf{x} = (\mathrm{d}x, 0)$ and $\mathrm{d}\mathbf{x} = (0, \mathrm{d}y)$, that

$$P = \frac{\partial f}{\partial x} \quad \text{and} \quad Q = \frac{\partial f}{\partial y}$$

respectively. Moreover,

$$\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial Q}{\partial x}\,,$$

since P and Q are assumed to have continuous first order partial derivatives (and hence from § 1.8.1 the mixed second order partial derivatives of f must be equal). We conclude that a necessary condition for ω to be exact is that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}\ . \tag{3.24}$$

A Sufficient Condition for ω to be an Exact Differential (2D).

If the domain $\mathcal{R}$ is simply connected then (3.24) is a sufficient condition for ω to be an exact differential.

Unlectured Outline Proof. *This outline proof is for those cannot wait for an idea of why this result is true. If you are prepared to wait then, once Stokes' theorem has been discussed in § 6.3.1, we will obtain a more general result in § 6.3.6, of which this is a special case.*

In what follows we will assume for simplicity that $\mathcal{R} = \mathbb{R}^2$. We can then be certain that all the line integrals of P and Q below remain within the domains of P and Q.

If ω is to be an exact differential then from (3.15b) and (3.23) we require that

$$f_x(x,y) = P(x,y)\,.$$

This suggests that we consider the function

$$f(x,y) = \int_{x_0}^{x} P(x,y)\,\mathrm{d}x + Y(y)\,,$$

where x_0 is a constant and Y is to be determined. By differentiating with respect to y and using condition (3.24) it follows that for this function

$$\begin{aligned} f_y(x,y) &= \int_{x_0}^{x} P_y(x,y)\,\mathrm{d}x + Y'(y) \\ &= \int_{x_0}^{x} Q_x(x,y)\,\mathrm{d}x + Y'(y) \\ &= Q_x(x,y) - Q_x(x_0,y) + Y'(y)\,. \end{aligned}$$

Now if ω is to be an exact differential then we also require that $f_y = Q_x$. This is achieved if

$$Y = \int_{y_0}^{y} Q_x(x_0,y)\,\mathrm{d}y\,,$$

where y_0 is a constant. Hence, for the function f defined by

$$\begin{aligned} f(x,y) &= \int_{x_0}^{x} P(x,y)\,\mathrm{d}x + \int_{y_0}^{y} Q_x(x_0,y)\,\mathrm{d}y \\ &= \int_{\mathcal{C}_1+\mathcal{C}_2} \mathbf{F}.\mathrm{d}\mathbf{x}\,, \end{aligned} \tag{3.25}$$

where $\mathbf{F} = (P,Q)$, we conclude that

$$\mathrm{d}f = f_x\mathrm{d}x + f_y\mathrm{d}y = P\mathrm{d}x + Q\mathrm{d}y = \omega\,.$$

Thus, by the construction of an appropriate function (which, in the terminology introduced in § 3.4.1 below, is referred to as a *scalar potential*), we have shown that if condition (3.24) is satisfied, then ω is an exact differential.

Exact Differentials (3D).

In an analogous way to above, the differential form

$$\omega = F_1\,\mathrm{d}x + F_2\,\mathrm{d}y + F_3\,\mathrm{d}z \tag{3.26a}$$

can be shown to be exact in a simply connected domain $\mathcal{R}$ of $\mathbb{R}^3$ if and only if

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}\,, \qquad \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}\,, \qquad \frac{\partial F_3}{\partial x} = \frac{\partial F_1}{\partial z}\,. \tag{3.26b}$$

We shall see in § 5.3.1 and equation (5.9b) that condition (3.26b) is equivalent to $\boldsymbol{\nabla}\times\mathbf{F} = 0$.

Unlectured Proof of Necessity. If ω is exact then for some function f

$$F_1\,\mathrm{d}x + F_2\,\mathrm{d}y + F_3\,\mathrm{d}z = f_x\,\mathrm{d}x + f_y\,\mathrm{d}y + f_z\,\mathrm{d}z\,.$$

Hence from independently varying $\mathrm{d}x$, $\mathrm{d}y$ and $\mathrm{d}z$,

$$F_1 = f_x\,, \quad F_2 = f_y \quad \text{and} \quad F_3 = f_z\,.$$

Condition (3.26b) then follows from the requirements that

$$f_{xy} = f_{yx}\,, \quad f_{yz} = f_{zy} \quad \text{and} \quad f_{zx} = f_{xz}\,.$$

Worked Exercise.

Question. Is the differential form

$$\omega = (\sin y)\,\mathrm{d}x + (2y + x\cos y)\,\mathrm{d}y \tag{3.27a}$$

exact on $\mathbb{R}^2$? If so, find all functions f such that $\omega = \mathrm{d}f$.

Answer. In terms of the notation of (3.23)

$$\begin{array}{rclcrcl} P &=& \sin y\,, & & P_y &=& \cos y\,, \\ Q &=& 2y + x\cos y\,, & & Q_x &=& \cos y\,. \end{array}$$

Thus since $P_y = Q_x$, it follows from the sufficiency condition on page 49 that ω is exact. Further, from (3.15b) and (3.23)

$$\frac{\partial f}{\partial x} = P = \sin y\,,$$

which implies that

$$f = x\sin y + Y(y)\,, \tag{3.27b}$$

where Y is a yet to be determined function. Similarly

$$\frac{\partial f}{\partial y} = Q = 2y + x\cos y\,,$$

and hence

$$f = y^2 + x\sin y + X(x)\,. \tag{3.27c}$$

Combining (3.27b) and (3.27c) it follows that

$$Y(y) - y^2 = X(x) = k\,,$$

where k is a constant, and hence that

$$f = y^2 + x\sin y + k\,. \tag{3.27d}$$

3.3.4 Solution of Differential Equations

The differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{P(x,y)}{Q(x,y)} \tag{3.28}$$

can alternatively be written as

$$\omega = P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y = 0\,.$$

The case when ω is an exact differential. If ω is an exact differential and equal to $\mathrm{d}f$, then since $\mathrm{d}f = 0$ the solution to the differential equation (3.28) is

$$f(x,y) = \kappa\,,$$

where κ is a constant.

Unlectured Example. Suppose that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\sin y}{2y + x\cos y}\,, \tag{3.29}$$

then ω is given by (3.27a) and f by (3.27d). Hence solutions to (3.29) are given by

$$y^2 + x\sin y = \kappa \qquad \text{i.e.} \qquad x = \frac{\kappa - y^2}{\sin y}\,.$$

The case when ω is not an exact differential. If the [initial] rewrite of (3.28) does not yield an exact differential, then the aim is to find an *integrating factor*, $\mu(x,y)$, such that

$$\mu\omega = \mu(x,y)P(x,y)\,\mathrm{d}x + \mu(x,y)Q(x,y)\,\mathrm{d}y$$

is exact. It follows from the necessary and sufficient condition (3.24) that we then require

$$\frac{\partial}{\partial y}(\mu P) = \frac{\partial}{\partial x}(\mu Q)\,,$$

i.e. that

$$\mu\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) + P\frac{\partial \mu}{\partial y} - Q\frac{\partial \mu}{\partial x} = 0\,.$$

This is a *partial differential equation* for μ, which it is not necessarily straightforward to solve. However if either

$$\frac{1}{Q}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = X(x) \qquad \text{or} \qquad \frac{1}{P}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = Y(y)\,,$$

where X and Y are functions only of x and y respectively, then an integrating factor exists that is only a function of x or y respectively, i.e. either

$$\mu \equiv \mu(x) \quad \text{and} \quad \frac{1}{\mu}\frac{\mathrm{d}\mu}{\mathrm{d}x} = X\,, \qquad \text{or} \qquad \mu \equiv \mu(y) \quad \text{and} \quad \frac{1}{\mu}\frac{\mathrm{d}\mu}{\mathrm{d}y} = -Y\,,$$

respectively.

Unlectured Example. Note that while this example illustrates integrating factors, the underlying differential equation is actually best solved using separation of variables.

Suppose that

$$\omega = y\,\mathrm{d}x - x\,\mathrm{d}y\,, \quad \text{i.e.} \quad P = y\,, \quad Q = -x\,.$$

Then

$$\frac{1}{Q}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = -\frac{2}{x} \qquad \text{and} \qquad \frac{1}{P}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = \frac{2}{y}\,.$$

Hence in this case we can choose either $\mu \equiv \mu(x)$ or $\mu \equiv \mu(y)$; specifically either

$$\frac{1}{\mu}\frac{\mathrm{d}\mu}{\mathrm{d}x} = -\frac{2}{x} \quad \Rightarrow \quad \mu = \frac{1}{x^2} \qquad \text{or} \qquad \frac{1}{\mu}\frac{\mathrm{d}\mu}{\mathrm{d}y} = -\frac{2}{y} \quad \Rightarrow \quad \mu = \frac{1}{y^2}\,.$$

As a check note that we then have either

$$\frac{y\,\mathrm{d}x}{x^2} - \frac{\mathrm{d}y}{x} = -\mathrm{d}\left(\frac{y}{x}\right) \qquad \text{or} \qquad \frac{\mathrm{d}x}{y} - \frac{x\,\mathrm{d}y}{y^2} = \mathrm{d}\left(\frac{x}{y}\right)\,.$$

3.4 Line Integrals of Exact Differentials

Suppose that

$$\omega = P\,\mathrm{d}x + Q\,\mathrm{d}y$$

is an exact differential. Then from definition (3.23) there exists a *single valued* function $f(\mathbf{x})$ such that

$$\mathrm{d}f = P\,\mathrm{d}x + Q\,\mathrm{d}y\,.$$

Hence using (i) the relationship (3.18) between the differential of a scalar function and the gradient of the function, and (ii) example (3.13), it follows that

$$\begin{aligned}\int_{\mathcal{C}}(P\,\mathrm{d}x + Q\,\mathrm{d}y) &= \int_{\mathcal{C}}\mathrm{d}f \\ &= \int_{\mathbf{x}_a}^{\mathbf{x}_b}\boldsymbol{\nabla}f.\mathrm{d}\mathbf{x} \\ &= f(\mathbf{x}_b) - f(\mathbf{x}_a)\,. \qquad (3.30)\end{aligned}$$

Thus in $\mathbb{R}^2$ the line integral of an exact differential depends only on the end points and is independent of the path. The same result holds in $\mathbb{R}^m$.

Example. Reconsider example (3.10), for which

$$\mathbf{F}.\,\mathrm{d}\mathbf{x} = y\,\mathrm{d}x + x\,\mathrm{d}y = \mathrm{d}(xy)\,.$$

Hence $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is an exact differential (with $f = xy$), and thus from (3.30)

$$\int\limits_{(0,0)\to(1,1)} (y\,\mathrm{d}x + x\,\mathrm{d}y) = \Big[xy\Big]_{(0,0)}^{(1,1)} = 1\ ,$$

independent of path chosen for the integral.

Remarks.

- Similarly in 3D, if $\mathbf{F}.\,\mathrm{d}\mathbf{x} = \mathrm{d}f$ then $\int_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x}$ depends only on the end points.
- The integral of an exact differential around a closed curve in a simply connected region must vanish because

$$\begin{aligned} \oint_{\mathcal{C}} \mathrm{d}f &= \int_{\mathcal{C}_1} \mathrm{d}f + \int_{\mathcal{C}_2} \mathrm{d}f \\ &= \int_{\mathcal{C}_1} \mathrm{d}f - \int_{-\mathcal{C}_2} \mathrm{d}f \\ &= (f(\mathbf{x}_b) - f(\mathbf{x}_a)) - (f(\mathbf{x}_b) - f(\mathbf{x}_a)) \\ &= 0\,. \end{aligned} \tag{3.31}$$

10/00

3.4.1 Conservative Fields

Definitions.

Conservative Field. If the line integral

$$\int_{\mathbf{x}_a}^{\mathbf{x}_b} \mathbf{F}.\,\mathrm{d}\mathbf{x}$$

is independent of the path taken from $\mathbf{x}_a$ to $\mathbf{x}_b$ in a simply connected domain, then the vector function $\mathbf{F}(\mathbf{x})$ if said to be a *conservative field*.

Scalar Potential. The *scalar potential* f of a conservative field $\mathbf{F}(\mathbf{x})$ is defined to be

$$f(\mathbf{x}) = \int_{\mathbf{x}_0}^{\mathbf{x}} \mathbf{F}.\,\mathrm{d}\mathbf{x}\,, \tag{3.32a}$$

where $\mathbf{x}_0$ is some fixed point. The scalar potential defined by (3.32a) is guaranteed to be a single valued function because, from the definition of a conservative field, the value of the integral is independent of the path.

Property. For the scalar potential defined by (3.32a)

$$\boldsymbol{\nabla} f = \mathbf{F}\,. \tag{3.32b}$$

Verification. From (3.32a) and the definition of a partial derivative (1.14)

$$\begin{aligned}\frac{\partial f}{\partial x}(\mathbf{a}) &= \lim_{h\to 0}\frac{f(\mathbf{a}+h\mathbf{e}_1)-f(\mathbf{a})}{h}\\ &= \lim_{h\to 0}\frac{1}{h}\int_{\mathbf{a}}^{\mathbf{a}+h\mathbf{e}_1}\mathbf{F}(\mathbf{x}).\,\mathrm{d}\mathbf{x}\,.\end{aligned}$$

Since $\mathbf{F}$ is a conservative field, we may choose the most convenient path for evaluating the integral (see the definition of a conservative field). In this case we take a straight line defined parametrically by $\mathbf{x}(t)=\mathbf{a}+ht\mathbf{e}_1$ for $0\leqslant t\leqslant 1$. Then, since $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=h\mathbf{e}_1$, it follows from (3.7a) that

$$\begin{aligned}\frac{\partial f}{\partial x}(\mathbf{a}) &= \lim_{h\to 0}\frac{1}{h}\int_{\mathbf{a}}^{\mathbf{a}+h\mathbf{e}_1}\mathbf{F}(\mathbf{x}).\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\mathrm{d}t\\ &= \lim_{h\to 0}\frac{1}{h}\int_0^1 F_1(a_1+ht,a_2,a_3)\,h\,\mathrm{d}t\\ &= \lim_{h\to 0}\int_0^1\big(F_1(a_1,a_2,a_3)+o(1)\big)\,\mathrm{d}t\\ &\qquad\qquad\qquad\text{from using Taylor's theorem (1.58a)}\\ &= F_1(\mathbf{a})\,.\end{aligned}$$

Similarly for the other components we deduce that

$$\frac{\partial f}{\partial y}=F_2\quad\text{and}\quad\frac{\partial f}{\partial z}=F_3\ .$$

Putting these results together it follows that

$$\boldsymbol{\nabla} f=\mathbf{F}\,. \qquad \square$$

Property. From (3.32b) and the relationship (3.18) between the differential and the gradient of a function,

$$\mathbf{F}.\,\mathrm{d}\mathbf{x}=\boldsymbol{\nabla} f.\mathrm{d}\mathbf{x}=\mathrm{d}f\,. \tag{3.32c}$$

Thus it follows that the differential form $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ of a conservative field $\mathbf{F}$ is exact.

Remarks on Work. From (3.9) the work done on a particle transversing a curve $\mathcal{C}$ is

$$W=\int_{\mathcal{C}}\mathbf{F}.\,\mathrm{d}\mathbf{x}\ . \tag{3.33}$$

- When a force $\mathbf{F}$ is a conservative field we refer to it as a conservative force.
- Friction is not a conservative force because the work done in transversing a closed loop is non-zero (because friction opposes motion), and hence $\int_{\mathbf{x}_a}^{\mathbf{x}_b}\mathbf{F}.\,\mathrm{d}\mathbf{x}$ cannot be independent of the path taken.
- If $\mathbf{F}=(0,0,-g)$ is the gravitational force at the surface of the earth, it follows from the necessary and sufficient conditions for a differential form to be exact, i.e. (3.26b), that $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is exact. Hence from the 3D extension of (3.30), i.e. the result that the line integral of an exact differential is independent of path, it follows that that the work done in moving from $\mathbf{x}_a$ to $\mathbf{x}_b$ against gravity is independent of path. The gravitational force at the earth's surface is thus a conservative force.

3.5 Non Simply Connected Domains

Suppose that the vector function $\mathbf{F}$ is defined by

$$\mathbf{F} = \left(-\frac{y}{x^2+y^2}\,,\ \frac{x}{x^2+y^2}\,,\ 0 \right) \quad \text{for} \quad \mathbf{x} \neq 0\,.$$

Then $\mathbf{F}\,.\,\mathrm{d}\mathbf{x}$ is exact in any simply connected domain *excluding the origin* since

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

Further

$$f = \tan^{-1}\left(\frac{y}{x}\right) \quad \text{satisfies} \quad \mathrm{d}f = \boldsymbol{\nabla} f\,.\,\mathrm{d}\mathbf{x} = \mathbf{F}\,.\,\mathrm{d}\mathbf{x}\,.$$

The question then arises: what if we drop the simply connectedness?

For instance, suppose we consider the domain $\mathcal{R} = \mathbb{R}^2/(0,0)$ (i.e. the real plane minus the origin ... this is *not* a simply connected domain since a curve encircling the origin cannot be shrunk to a point in the domain), and the line integral round the path $\mathcal{C} : \mathbf{x} = (a\cos\phi, a\sin\phi)$ where $0 \leqslant \phi \leqslant 2\pi$. Then

$$\begin{aligned}
\oint_{\mathcal{C}} \mathbf{F}\,.\,\mathrm{d}\mathbf{x} &= \int_0^{2\pi} \mathbf{F}\,.\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\phi}\,\mathrm{d}\phi \\
&= \int_0^{2\pi} \left(\left(\frac{-\sin\phi}{a} \right).\,(-a\sin\phi) + \frac{\cos\phi}{a}.\,a\cos\phi \right) \mathrm{d}\phi \\
&= 2\pi\ .
\end{aligned}$$

We conclude that the simply connectedness is crucial for the integral to depend only on the end points.

Optional Exercise. Show that for any curve $\mathcal{C}$ that encircles the origin exactly once

$$\oint_{\mathcal{C}} \mathbf{F}\,.\,\mathrm{d}\mathbf{x} = 2\pi\,.$$

4 Multiple Integrals

Having studied single integrals over a straight line (pre-university), and single integrals over a curve in three-dimensional space (§ 3.2), we will now look at double integrals over an area (§ 4.1), triple integrals over a volume (§ 4.2), and double integrals over a surface in three-dimensional space (§ 4.4).

There are a couple of extra books that are helpful for these subjects:

- *Calculus* by Boyce & DiPrima (*lots of nice pictures*);
- *Differential & Integral Calculus* by Courant (*old fashioned, but precise*).

4.1 Double Integrals (Area Integrals)

Suppose that we wish to integrate $f : \mathbb{R}^2 \to \mathbb{R}$ over a region $\mathcal{R}$ of the xy-plane, where $\mathcal{R}$ is closed and bounded, where

- by bounded we mean that there exists a circumscribing circle/rectangle.
- by closed we mean that the boundary is part of region.

Subdivide $\mathcal{R}$ by sectionally smooth curves into N subregions $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_N$ with areas $\delta S_1, \dots, \delta S_N$. For a subregion $\mathcal{S}_j$ define its greatest diameter $|\mathcal{S}_j|$, to be the greatest distance between two points in $\mathcal{S}_j$. In each $\mathcal{S}_j$ choose an arbitrary point (ξ_j, η_j).

Definition. By analogy with the Riemann integral (3.1d), define the double integral

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S = \lim_{\substack{N\to\infty \\ |\mathcal{S}_j|\to 0}} \sum_j f(\xi_j,\eta_j)\,\delta S_j\,, \tag{4.1}$$

if the limit exists as the greatest of the diameters of the subregions tends to zero, and if the limit is both

- independent of the subdivision of $\mathcal{R}$, and
- independent of the choice of the (ξ_j, η_j).

Properties. Double integrals have the expected properties, e.g. for functions f and g, and real constants λ and μ:

$$\iint_{\mathcal{R}} (\lambda f + \mu g)\,\mathrm{d}S = \lambda \iint_{\mathcal{R}} f\,\mathrm{d}S + \mu \iint_{\mathcal{R}} g\,\mathrm{d}S\,.$$

4.1.1 Evaluation by Successive Integration

One way to evaluate double integrals is first to keep one variable fixed and to evaluate the *inner integral*, and then to evaluate the remaining *outer integral*.

For simplicity assume the region $\mathcal{R}$ is convex[a] and is described by

$$a \leqslant x \leqslant b\,, \quad y_1(x) \leqslant y \leqslant y_2(x)\,. \tag{4.2a}$$

[a] A bounded and closed region $\mathcal{R}$ is convex if it contains the line segment joining any two points in it.

Then

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S = \int_{x=a}^{b} \underbrace{\left[\int_{y=y_1(x)}^{y_2(x)} f(x,y)\,\mathrm{d}y\right]}_{\text{only a function of } x} \mathrm{d}x\,. \tag{4.2b}$$

Suppose instead that the region is equivalently specified as

$$c \leqslant y \leqslant d\,, \quad x_1(y) \leqslant x \leqslant x_2(y)\,. \tag{4.2c}$$

Then

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S = \int_{y=c}^{d} \underbrace{\left[\int_{x=x_1(y)}^{x_2(y)} f(x,y)\,\mathrm{d}x\right]}_{\text{only a function of } y} \mathrm{d}y\,. \tag{4.2d}$$

Remarks.

- For nice functions (e.g. f continuous), (4.2b) and (4.2d) give the same result. We will assume that this is true for this course; however it is not always so. Consider

  $$f = \begin{cases} +1 & y < x < y+1 \\ -1 & y+1 < x < y+2 \\ 0 & \text{otherwise} \end{cases}$$

  Then

  $$\int_{y=0}^{\infty} \left[\int_{x=0}^{\infty} f\mathrm{d}x\right] \mathrm{d}y = 0\,,$$

  while using the formula for the area of a triangle

  $$\int_{x=0}^{\infty} \left[\int_{y=0}^{\infty} f\mathrm{d}y\right] \mathrm{d}x = 3(\tfrac{1}{2}) - \tfrac{1}{2} = 1\,.$$

  Comment. Ok, fair cop, the region is unbounded. However, see *Examples Sheet 2* for an integral in a bounded region where care is needed in changing the order of integration.

- If $\mathcal{R}$ is not convex rotate the co-ordinate axes so that lines of either constant x or constant y cut boundaries no more than twice. If that is not possible then split $\mathcal{R}$ into regions where the axes can be rotated so that lines of either constant x or constant y cut boundaries no more than twice.

- **NEVER EVER** write a double integral as

  $$\int_{y=y_1(x)}^{y_2(x)} \left[\int_{x=a}^{b} f(x,y)\,\mathrm{d}x\right] \mathrm{d}y\,.$$

  This is **MEANINGLESS**.

- An alternative notation is to write the d$\bullet$ immediately following the integration sign it refers to, viz.

  $$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S = \int_a^b \mathrm{d}x \int_{y_1(x)}^{y_2(x)} \mathrm{d}y\ f(x,y)\,.$$

Examples.

1. For the region $\mathcal{R}$ specified by (4.2a), i.e. $a \leqslant x \leqslant b$ and $y_1(x) \leqslant y \leqslant y_2(x)$

$$\begin{aligned}\iint_{\mathcal{R}} \frac{\partial P}{\partial y}\,\mathrm{d}y\,\mathrm{d}x &= \int_{x=a}^{b}\left[\int_{y=y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\mathrm{d}y\right]\mathrm{d}x \\ &= \int_{x=a}^{b}\big[P(x,y_2(x)) - P(x,y_1(x))\big]\mathrm{d}x\,. \end{aligned} \tag{4.3}$$

   While this is not a complete evaluation, the double integral has been reduced to a standard Riemann integral. We will use this result in the proof of Green's theorem (see (6.11)).

2. If $f(x,y) = 1$, then

$$\iint_{\mathcal{R}} \mathrm{d}S = \text{ area of } \mathcal{R}\,.$$

   Special Case. For $\mathcal{R}$ given by $a \leqslant x \leqslant b$ and $0 \leqslant y \leqslant y(x)$,

$$\begin{aligned}\iint_{\mathcal{R}} \mathrm{d}S &= \int_a^b\left[\int_0^{y(x)}\mathrm{d}y\right]\mathrm{d}x \\ &= \int_a^b y(x)\,\mathrm{d}x\,.\end{aligned}$$

   Some care is needed with signs if $y(x)$ is negative (since areas are by convention positive).

3. Suppose that $f(x,y) = y$, and that the region is $\mathcal{R}$ is specified by

$$1 \leqslant x \leqslant 2 \quad \text{and} \quad 1/x \leqslant y \leqslant e^x\,.$$

   Do the y-integration first; then

$$\begin{aligned}\int_{x=1}^{2}\mathrm{d}x\int_{y=1/x}^{e^x}\mathrm{d}y\;y &= \int_{x=1}^{2}\mathrm{d}x\,\left[\tfrac{1}{2}y^2\right]_{1/x}^{e^x} \\ &= \tfrac{1}{2}\int_1^2 \mathrm{d}x(e^{2x} - 1/x^2) \\ &= \tfrac{1}{4}(e^4 - e^2 - 1)\,.\end{aligned}$$

   Exercise. Do the x-integration first.

   Moral. Think before deciding on the order of integration.

11/99
11/00 4. The volume V beneath the surface $z = f(x,y) > 0$ and above a region $\mathcal{R}$ in the xy-plane is (cf. example 2)

$$V = \iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S\,.$$

4.2 Triple Integrals (Volume Integrals)

Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ and $\mathcal{V}$ is a domain in $\mathbb{R}^3$.
Subdivide $\mathcal{V}$ into N sub-domains $\mathcal{V}_1, \ldots, \mathcal{V}_N$,
with volumes $\delta\tau_1, \ldots, \delta\tau_N$, and let $\boldsymbol{\xi}_j \in \mathcal{V}_j$.

Definition. Define the triple integral

$$\iiint_{\mathcal{V}} f(\mathbf{x})\,\mathrm{d}\tau = \lim_{N\to\infty} \sum_j f(\boldsymbol{\xi}_j)\,\delta\tau_j \tag{4.4}$$

if the limit exists as the maximum dimension of the largest $\mathcal{V}_j$ tends to zero, and the limit is independent both of the subdivision of $\mathcal{V}$ and the choice of the $\boldsymbol{\xi}_j$.

4.2.1 Evaluation

As for double integrals one can often evaluate triple integrals by repeated integration, e.g. for a convex region

$$\iiint_{\mathcal{V}} f(\mathbf{x})\mathrm{d}\tau = \int_{x=a}^{b} \mathrm{d}x \int_{y=y_1(x)}^{y_2(x)} \mathrm{d}y \int_{z=z_1(x,y)}^{z_2(x,y)} \mathrm{d}z\ f(x,y,z)\,. \tag{4.5}$$

For 'nice' functions the order of integration is unimportant (although it is, of course, important to ensure that the integration ranges are correct).

Example. Let $\mathcal{V}$ be the tetrahedron bounded by the four planes

$$x = 0,\ y = 0,\ z = 0 \quad \text{and} \quad x + y + z = a\,.$$

The vertices are at O $(0,0,0)$
P $(a,0,0)$
Q $(0,a,0)$
R $(0,0,a)$

Suppose that we wish to evaluate

$$J_x = \iiint x^2\mathrm{d}\tau\,,\ J_y = \iiint y^2\mathrm{d}\tau\,,\ J_z = \iiint z^2\mathrm{d}\tau\,.$$

First we note that by symmetry $J_x = J_y = J_z$. We will concentrate on J_x.

If we are to evaluate the J_x by repeated integration, which integral should we do first?

1. The z integral? Looks good because the integrand has no explicit z dependence.
2. The y integral? Looks good because the integrand has no explicit y dependence.
3. The x integral? Looks bad because the integrand has explicit x dependence.

We will do 3!

$$\begin{aligned}
J_x &= \int_0^a \mathrm{d}z \int_0^{a-z} \mathrm{d}y \int_0^{a-z-y} \mathrm{d}x \; x^2 \\
&= \int_0^a \mathrm{d}z \int_0^{a-z} \mathrm{d}y \; \tfrac{1}{3}(a-z-y)^3 \\
&= \int_0^a \mathrm{d}z \; \tfrac{1}{12}(a-z)^4 \\
&= \tfrac{1}{60}a^5 \,.
\end{aligned}$$

Remark. We can combine four of these shapes to form a pyramid, $\widehat{\mathcal{V}}$. The moment of inertia of the pyramid about Oz is

$$I_z = \iiint_{\widehat{\mathcal{V}}} (x^2+y^2)\; \mathrm{d}\tau = \frac{2a^5}{15}\,.$$

Optional Exercise.

1. Calculate the moments of inertia about Ox and Oy for the pyramid.
2. Find the centre of mass of the pyramid.
3. Find the moments of inertia about centre of mass using the *Theorem of Parallel Axes*.

4.3 Jacobians and Change of Variables

The evaluation of some multiple integrals can sometimes be simplified by a change of variables, e.g. $(x,y) \to (u,v)$.

As well as accounting for the change in the region of integration as a result of the change of variables, we also need to account for any change in the elementary area or volume, cf. change of variable for ordinary integrals where the transformation $x = g(u)$ is accompanied by $\mathrm{d}x = g'(u)\,\mathrm{d}u$ so that

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(u))\; g'(u)\,\mathrm{d}u\,.$$

4.3.1 Jacobians

Start in 2D and suppose that

$$\mathbf{x}(\mathbf{u}) : \mathbb{R}^2 \to \mathbb{R}^2$$

is a transformation representing a change of variables. Then in the case of a function of $\mathbf{x}$,

$$f(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}\,,$$

when viewed as a function of $\mathbf{u}$,

$$f(\mathbf{u}) : \mathbb{R}^2 \to \mathbb{R}\,,$$

the partial derivatives are related by the the chain rule (1.33a); in matrix form this states that

$$\begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

Definition. The determinant of the Jacobian matrix

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

is called the *Jacobian*, or *Jacobian determinant*, of the transformation $\mathbf{x}(\mathbf{u})$, and is written

$$J = J(\mathbf{x},\mathbf{u}) = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \det\left(\frac{\partial x_i}{\partial u_j}\right) = \det(U_{ij}) = \det \mathsf{U}\,. \tag{4.6}$$

Remarks.

- There is an important distinction between the Jacobian matrix and the Jacobian (or Jacobian determinant).
- For a three dimensional transformation $\mathbf{x}(\mathbf{u}) : \mathbb{R}^3 \to \mathbb{R}^3$

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

Co-ordinate Transformations.

For a change of variables given by the transformation $\mathbf{x}(\mathbf{u})$ to be 'useful', there needs to be a 1-1 correspondence between points in the $\mathbf{u}$ co-ordinate system and those in the $\mathbf{x}$ co-ordinate system.

In order to appreciate when this is so, start from the definition of a differential (3.14a) applied componentwise to $\mathbf{x} : \mathbb{R}^m \to \mathbb{R}^m$:

$$\mathrm{d}x_i = \sum_{\alpha=1}^{m} \frac{\partial x_i}{\partial u_\alpha}\mathrm{d}u_\alpha = \sum_{\alpha=1}^{m} U_{i\alpha}\mathrm{d}u_\alpha\,, \quad \text{i.e.} \quad \mathrm{d}\mathbf{x} = \mathsf{U}\,\mathrm{d}\mathbf{u}\,. \tag{4.7}$$

Short Observation. If there is a 1-1 correspondence between points in the $\mathbf{u}$ co-ordinate system and those in the $\mathbf{x}$ co-ordinate system, then as well as expressing $\mathrm{d}\mathbf{x}$ in terms of $\mathrm{d}\mathbf{u}$ it should be possible to express $\mathrm{d}\mathbf{u}$ in terms of $\mathrm{d}\mathbf{x}$. If so, then from (4.7) it follows that

$$\mathrm{d}\mathbf{u} = \mathsf{U}^{-1}\mathrm{d}\mathbf{x}\,.$$

This suggests that there will be a 1-1 correspondence between points in the two co-ordinate systems if $\det(\mathsf{U}) \neq 0$.

More Lengthy Comment. If $\mathrm{d}\mathbf{x}$ and $\mathrm{d}\mathbf{u}$ are interpreted as increments in $\mathbf{x}$ and $\mathbf{u}$ in their respective co-ordinate systems, equation (4.7) provides a means for calculating the local change in $\mathbf{x}$ corresponding to a given local change in $\mathbf{u}$. Denote by $\mathrm{d}\mathbf{x}^{(\alpha)}$ for $\alpha = 1, \ldots, m$ the increment in $\mathbf{x}$ corresponding to an increment

$$\mathrm{d}\mathbf{u}^{(\alpha)} = (0, \ldots, \mathrm{d}u_\alpha, \ldots, 0) \neq 0\,,$$

where the $\mathrm{d}\mathbf{u}^{(\alpha)}$ for $\alpha = 1, \ldots, m$ can be viewed as a local basis in the $\mathbf{u}$ co-ordinate system. Then from (4.7)

$$\mathrm{d}x_i^{(\alpha)} = \sum_{\alpha=1}^{m} U_{i\alpha} \mathrm{d}u_i^{(\alpha)} = U_{i\alpha}\mathrm{d}u_\alpha \quad (\text{no } s.c.).$$

The condition for the transformation $\mathbf{x}(\mathbf{u})$ to be 'useful' is that the $\mathrm{d}\mathbf{x}^{(\alpha)}$ for $\alpha = 1, \ldots, m$ are a local basis in the $\mathbf{x}$ co-ordinate system, i.e. that they are independent, or equivalently that the only solution to

$$\sum_{\alpha=1}^{m} \mathrm{d}\mathbf{x}^{(\alpha)}\lambda_\alpha = 0\,, \quad \text{or} \quad \sum_{\alpha=1}^{m} \mathrm{d}x_i^{(\alpha)}\lambda_\alpha = 0 \quad \text{for } i = 1, \ldots, m\,,$$

is $\lambda_\alpha = 0$ for $\alpha = 1, \ldots, m$. From the results for the solution of linear systems in *Algebra & Geometry* it is thus necessary that

$$\det(\mathrm{d}x_i^{(\alpha)}) = \det(U_{i\alpha}\mathrm{d}u_\alpha) = \det(U_{i\alpha})\mathrm{d}u_1 \ldots \mathrm{d}u_m \neq 0 \quad (\text{no } s.c.),$$

i.e. that $\det(\mathsf{U}) \neq 0$. Moreover, if $\det \mathsf{U} \neq 0$ then U^{-1} exists, and it then follows from (4.7) that the inverse transformation,

$$\mathrm{d}\mathbf{u} = \mathsf{U}^{-1}\mathrm{d}\mathbf{x}\,,$$

is well defined.

It is in fact possible to *prove* that if $\mathbf{x}(\mathbf{u})$ is continuously differentiable and $\det(\mathsf{U}) \neq 0$, then the transformation $\mathbf{x}(\mathbf{u})$ is non-singular and there exists a 1-1 correspondence between points in the $\mathbf{u}$ plane and those in the $\mathbf{x}$ plane. This is the generalisation to more than one variable of the Inverse Function Theorem (see *Analysis I*).

4.3.2 Chain Rule for Jacobians

Start in 2D, and suppose that $\mathbf{x} = (x, y)$ is a function of $\mathbf{u} = (u, v)$, and $\mathbf{u} = (u, v)$ is a function of $\mathbf{s} = (s, t)$, i.e. $\mathbf{x} \equiv \mathbf{x}(\mathbf{u})$ and $\mathbf{u} \equiv \mathbf{u}(\mathbf{s})$. Write

$$\begin{array}{lll} x_1 = x, & u_1 = u & s_1 = s \\ x_2 = y, & u_2 = v & s_2 = t \end{array}\,,$$

and let $\mathsf{X_s}$, $\mathsf{X_u}$ and $\mathsf{U_s}$ deonte the Jacobian matrices of $\mathbf{x} \equiv \mathbf{x}(\mathbf{s})$, $\mathbf{x} \equiv \mathbf{x}(\mathbf{u})$ and $\mathbf{u} \equiv \mathbf{u}(\mathbf{s})$ respectively. Then from the chain rule (1.36a) and (1.38) for two vector functions ($s.c.$)

$$\frac{\partial x_i}{\partial s_k} = \frac{\partial x_i}{\partial u_j}\frac{\partial u_j}{\partial s_k} \qquad \text{i.e.} \qquad \mathsf{X_s} = \mathsf{X_u}\,\mathsf{U_s}\,.$$

Hence from (1.39) and from the definition of the Jacobian (4.6)

$$\frac{\partial(x,y)}{\partial(s,t)} = |\mathsf{X_s}| = |\mathsf{X_u}\,\mathsf{U_s}| = |\mathsf{X_u}|\,|\mathsf{U_s}| = \frac{\partial(x,y)}{\partial(u,v)}\,\frac{\partial(u,v)}{\partial(s,t)}\,,$$

i.e.

$$\frac{\partial(x,y)}{\partial(s,t)} = \frac{\partial(x,y)}{\partial(u,v)}\,\frac{\partial(u,v)}{\partial(s,t)}\,. \tag{4.8a}$$

Remark.

- The extension to three and higher dimensions is straightforward:

$$\frac{\partial(x,y,z)}{\partial(r,s,t)} = \frac{\partial(x,y,z)}{\partial(u,v,w)}\,\frac{\partial(u,v,w)}{\partial(r,s,t)}\,. \tag{4.8b}$$

Inverse. Suppose we identify (s,t) with (x,y) in (4.8a), then

$$\frac{\partial(x,y)}{\partial(u,v)}\,\frac{\partial(u,v)}{\partial(x,y)} = \frac{\partial(x,y)}{\partial(x,y)} = \begin{vmatrix} \left(\frac{\partial x}{\partial x}\right)_y & \left(\frac{\partial x}{\partial y}\right)_x \\ \left(\frac{\partial y}{\partial x}\right)_y & \left(\frac{\partial y}{\partial y}\right)_x \end{vmatrix} = 1\,.$$

Hence

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{\dfrac{\partial(u,v)}{\partial(x,y)}}\,. \tag{4.9}$$

Remark.

- Calculate the easier of $\frac{\partial(x,y)}{\partial(u,v)}$ and $\frac{\partial(u,v)}{\partial(x,y)}$, and use (4.9) if necessary.

12/99

Example: Plane Polar Co-ordinates.

In plane polar co-ordinates $x = \rho\cos\phi$ and $y = \rho\sin\phi$; so

$$\frac{\partial x}{\partial\rho} = \cos\phi\,, \qquad \frac{\partial x}{\partial\phi} = -\rho\sin\phi\,,$$
$$\frac{\partial y}{\partial\rho} = \sin\phi\,, \qquad \frac{\partial y}{\partial\phi} = \rho\cos\phi\,.$$

Hence

$$\frac{\partial(x,y)}{\partial(\rho,\phi)} = \begin{vmatrix} \frac{\partial x}{\partial\rho} & \frac{\partial x}{\partial\phi} \\ \frac{\partial y}{\partial\rho} & \frac{\partial y}{\partial\phi} \end{vmatrix} = \rho\cos^2\phi + \rho\sin^2\phi = \rho\,, \tag{4.10a}$$

and from (4.9)

$$\frac{\partial(\rho,\phi)}{\partial(x,y)} = \frac{1}{\rho}\,. \tag{4.10b}$$

Exercise. Calculate $\frac{\partial(\rho,\phi)}{\partial(x,y)}$ using $\rho = (x^2+y^2)^{\frac{1}{2}}$ and $\phi = \arctan(y/x)$, and remembering that

$$\left(\frac{\partial x}{\partial\rho}\right)_y \neq \frac{1}{\left(\frac{\partial\rho}{\partial x}\right)_\phi}\,.$$

12/00

4.3.3 Change of Variable in Multiple Integrals

Start in 2D. Consider

$$I = \iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S\,,$$

and suppose, as above, that $\mathbf{x} \equiv \mathbf{x}(\mathbf{u})$. The question is then how do we express the integral in terms of the new co-ordinates u and v (in terms of which we hope the integral is easier to evaluate)?

Three Tasks.

1. Find $F(u, v) \equiv f(x(u, v),\, y(u, v))$.
2. Identify the region of integration $\widehat{\mathcal{R}}$ in the uv-plane.

3. Find how the elementary areas $\mathrm{d}S_{xy}$ and $\mathrm{d}S_{uv}$ map between the different co-ordinate planes.

Three Answers.

1. Usually straightforward.
2. To find the transformed region $\widehat{\mathcal{R}}$, consider each component of boundary in turn.

   For instance, suppose that one boundary in the xy-plane is given by $g_1(x, y) = 0$. Then in the uv-plane the boundary is given by

   $$G_1(u, v) = g_1(x(u, v), y(u, v)) = 0\,.$$

   Warning.

   - Care is needed in the case of transformation with singular points, i.e. points where $J = 0$. Such a transformation may not be 1-1 because one region may map to two or more regions, e.g. if there is a $\sqrt{\ }$ fix on one sign!

3. In order to determine how elementary areas transform, consider a rectangular subdivision of the integration region $\widehat{\mathcal{R}}$ in the uv-plane based on curves defined by $u =$ constant and $v =$ constant. In the xy-plane consider a corresponding subdivision also based on curves defined by $u =$ constant and $v =$ constant, i.e. curves defined by $\mathbf{x} = \mathbf{V}(v)$ and $\mathbf{x} = \mathbf{U}(u)$ where

   $$\mathbf{V}(v) = \mathbf{x}(u_0, v)\,, \quad \mathbf{U}(u) = \mathbf{x}(u, v_0)\,,$$

for fixed u_0 and v_0.

In particular consider a rectangle with area $\delta u \delta v$ in the uv-plane and its transformation in the xy-plane:

If $|\delta u| \ll 1$ and $|\delta v| \ll 1$ then by Taylor's theorem

$$\mathbf{x}(u_0+\delta u, v_0) - \mathbf{x}(u_0, v_0) = \frac{\partial \mathbf{x}}{\partial u}(u_0, v_0)\,\delta u + o(\delta u)\,, \tag{4.11a}$$

$$\mathbf{x}(u_0, v_0+\delta v) - \mathbf{x}(u_0, v_0) = \frac{\partial \mathbf{x}}{\partial v}(u_0, v_0)\,\delta v + o(\delta v)\,. \tag{4.11b}$$

Hence as $\delta u,\ \delta v \to 0$, the transformed area approaches that of a parallelogram with sides

$$\frac{\partial \mathbf{x}}{\partial u}\,\delta u \quad \text{and} \quad \frac{\partial \mathbf{x}}{\partial v}\,\delta v\,,$$

and area

$$\begin{aligned}
\left|\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}\right| \delta u\, \delta v &= \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right| \delta u\, \delta v \\
&= \operatorname{abs}\left(\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \right) \delta u\, \delta v \\
&= \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \delta u\, \delta v\,.
\end{aligned}$$

We infer that infinitesimal areas in the xy and uv planes are related by

$$\mathrm{d}S_{xy} = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \mathrm{d}S_{uv} \qquad \left(\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \neq 0 \right)\,. \tag{4.12}$$

Remarks.

- The absolute value of the Jacobian determinant is appropriate, because areas are positive by convention.

- The Jacobian determinant *is* positive if the 'handedness' of the co-ordinate system is preserved. In particular, if a point in the uv-plane traces out a curve in a counterclockwise direction, then the image point in the xy-plane traces out a curve in the counterclockwise/clockwise direction according as the Jacobian determinant is positive/negative.

Putting the above three components together, we conclude that

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}S_{xy} = \iint_{\widehat{\mathcal{R}}} F(u,v)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|\mathrm{d}S_{uv}\,. \tag{4.13a}$$

Three Dimensions. Similarly in 3D, except we need the volume of the parallelepiped with generators.

$$\frac{\partial\mathbf{x}}{\partial u}\,\delta u\,,\quad \frac{\partial\mathbf{x}}{\partial v}\,\delta v\,,\quad \frac{\partial\mathbf{x}}{\partial w}\,\delta w\,.$$

Exercise. Show that this volume is given by

$$\left|\left(\frac{\partial\mathbf{x}}{\partial u}\times\frac{\partial\mathbf{x}}{\partial v}\right)\cdot\frac{\partial\mathbf{x}}{\partial w}\right|\delta u\,\delta v\,\delta w = \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right|\delta u\,\delta v\,\delta w\,.$$

Hence

$$\iiint_{\mathcal{V}} f(\mathbf{x})\,\mathrm{d}\tau_{\mathbf{x}} = \iiint_{\widehat{\mathcal{V}}} F(\mathbf{u})\;\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right|\mathrm{d}\tau_{\mathbf{u}}\,, \tag{4.13b}$$

where $F(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}))$, and it has been assumed that [small] volumes are positive.

Examples.

1. Evaluate

$$I = \iint_{\mathcal{R}} \frac{1}{x^2}\,\mathrm{d}S\,, \tag{4.14a}$$

where the region $\mathcal{R}$ is bounded by the planes

$$y=0\,,\ y=x\,,\ x+y=1 \quad\text{and}\quad x+y=2\,. \tag{4.14b}$$

Since the boundaries can be expressed as either $y = kx$ or $x + y = c$, this suggests trying the new variables

$$u = x+y \quad\text{and}\quad v = y/x\,. \tag{4.15}$$

Then

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -\dfrac{y}{x^2} & \dfrac{1}{x} \end{vmatrix} = \frac{x+y}{x^2}\,,$$

and from the result for the inverse of Jacobians (4.9)

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{x^2}{x+y}\,.$$

Thus using (4.13a) and (4.15) it follows that (4.14a) can be transformed to

$$I = \iint_{\widehat{\mathcal{R}}} \frac{1}{x^2}\,\frac{x^2}{x+y}\,\mathrm{d}u\,\mathrm{d}v = \iint_{\widehat{\mathcal{R}}} \frac{1}{u}\,\mathrm{d}u\,\mathrm{d}v\,.$$

From (4.14b) and (4.15) the region of integration in the uv-plane is

$$1 \leqslant u \leqslant 2 \quad\text{and}\quad 0 \leqslant v \leqslant 1\,.$$

Hence

$$I = \int_1^2 \mathrm{d}u \int_0^1 \mathrm{d}v\;\frac{1}{u} = \ln 2\,.$$

Comment. If instead we had chosen $u = y/x$ and $v = x+y$ in (4.15) then the Jacobian would have changed sign since the handedness of the system would not have been preserved. As a result a little extra care concerning signs would have been required.

2. *Plane Polar Co-ordinates:* $(\rho,\ \phi)$.

Suppose

$$f(x,y) \equiv F(\rho,\phi)$$

then using (4.13a) and (4.10a)

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}x\mathrm{d}y = \iint_{\widehat{\mathcal{R}}} F(\rho,\phi)\,\rho\,\mathrm{d}\rho\mathrm{d}\phi\,. \qquad (4.16)$$

Example. Evaluate

$$I^2 = \int_0^\infty \mathrm{d}x \int_0^\infty \mathrm{d}y\ \mathrm{e}^{-x^2-y^2}\,. \qquad (4.17\mathrm{a})$$

A transformation to plane polar co-ordinates using (4.16) yields

$$I^2 = \int_0^\infty \mathrm{d}\rho \int_0^{\frac{\pi}{2}} \mathrm{d}\phi\ \rho\mathrm{e}^{-\rho^2} = \frac{\pi}{2}\Big[-\tfrac{1}{2}\mathrm{e}^{-\rho^2}\Big]_0^\infty = \frac{\pi}{4}\,. \qquad (4.17\mathrm{b})$$

Further we note that

$$I^2 = \int_0^\infty \mathrm{d}x\ \mathrm{e}^{-x^2}\ \int_0^\infty \mathrm{d}y\ \mathrm{e}^{-y^2} = \left(\int_0^\infty \mathrm{d}t\ \mathrm{e}^{-t^2}\right)^2,$$

and thus

$$I = \int_0^\infty \mathrm{d}t\ \mathrm{e}^{-t^2} = \tfrac{1}{2}\pi^{\frac{1}{2}}\,. \qquad (4.18)$$

3. *Cylindrical Polar Co-ordinates:* $(\rho,\ \phi,\ z)$.

There is potential for notational confusion here. Many use $(r,\ \theta,\ z)$ for cylindrical polar co-ordinates, while others use $(\rho,\ \phi,\ z)$ — the advantage of the latter is that a notational clash with spherical polar co-ordinates over the meanings of r and θ is then avoided (see below). The relationship to Cartesian co-ordinates is as follows:

	(ρ,ϕ,z)	(r,θ,z)
radial co-ordinate	$\rho^2 = x^2+y^2 \geqslant 0$	$r^2 = x^2+y^2 \geqslant 0$
polar angle	$0 \leqslant \phi = \tan^{-1}\frac{y}{x} < 2\pi$	$0 \leqslant \theta = \tan^{-1}\frac{y}{x} < 2\pi$
axial co-ordinate	z	z

From result (4.10a) for plane polar co-ordinates (since the axial co-ordinate z does not depend on ρ or ϕ), or from a re-derivation using $x = \rho\cos\phi$ and $y = \rho\sin\phi$, the Jacobian for the transformation between Cartesian co-ordinates and cylindrical polar co-ordinates is given by

$$\frac{\partial(x,y,z)}{\partial(\rho,\phi,z)} = \rho\,, \qquad (4.19\mathrm{a})$$

and so the infinitesimal volume transforms according to

$$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \ \to\ \rho\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}z\,. \qquad (4.19\mathrm{b})$$

Thus the integral of

$$f(x,y,z) \equiv F(\rho,\phi,z)$$

over a volume $\mathcal{V}$ becomes in cylindrical polar co-ordinates

$$\iiint_{\mathcal{V}} f(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z = \iiint_{\widehat{\mathcal{V}}} F(\rho,\phi,z)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}z\,.$$

Example. Suppose that $f=1$ and $\mathcal{V}$ is the cylindrical region $0 \leqslant z \leqslant h$ and $x^2+y^2 \leqslant a^2$, then

$$\iiint_{\mathcal{V}} \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_0^h \mathrm{d}z \int_0^{2\pi} \mathrm{d}\phi \int_0^a \mathrm{d}\rho\ \rho = \pi a^2 h\,.$$

13/99

4. *Spherical Polar Co-ordinates:* $(r,\,\theta,\,\phi)$

	(r,θ,ϕ)
radial co-ordinate	$r^2 = x^2+y^2+z^2 \geqslant 0$
polar/latitudinal angle	$0 \leqslant \theta \leqslant \pi$
azimuthal/longitudinal angle	$0 \leqslant \phi < 2\pi$

In terms of Cartesian co-ordinates

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

Hence the Jacobian for the transformation between Cartesian co-ordinates and spherical polar co-ordinates is given by

$$\begin{aligned} \frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} &= \begin{vmatrix} s_\theta c_\phi & r c_\theta c_\phi & -r s_\theta s_\phi \\ s_\theta s_\phi & r c_\theta s_\phi & r s_\theta c_\phi \\ c_\theta & -r s_\theta & 0 \end{vmatrix} \\ &= r^2 \sin\theta\,, \end{aligned} \tag{4.20a}$$

and so the infinitesimal volume transforms according to

$$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \ \to\ r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\,. \tag{4.20b}$$

Thus the integral of

$$f(x,y,z) \equiv F(r,\theta,\phi)$$

13/00

over a volume $\mathcal{V}$ becomes in spherical polar co-ordinates

$$\iiint_{\mathcal{V}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iiint_{\widehat{\mathcal{V}}} F(r,\theta,\phi)\,r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi\,.$$

Example. Suppose that $f=1$, and that $\mathcal{V}$ is the interior of the sphere of radius a. Then

$$\begin{aligned} \iiint_{\mathcal{V}} f\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_0^a \mathrm{d}r \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\, r^2\sin\theta \\ &= \tfrac{1}{3}a^3\,.\,2\,.\,2\pi \ = \ \tfrac{4}{3}\pi a^3\,. \end{aligned}$$

4.4 Surface Integrals

So far we have evaluated integrals over areas in 2D, and volumes in 3D. We consider now integrals over a general surface in 3D.

4.4.1 Surfaces

For a smooth surface $\mathcal{S}$ described by the parametric representation (2.18c), i.e.

$$\mathbf{x} \equiv \mathbf{x}(u, v) \tag{4.21}$$

where u and v are independent variables, we have from (2.19) that a unit normal is given by

$$\mathbf{n} = \frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}\,. \tag{4.22}$$

In calculations of surface integrals we need, in addition to the normal, the size of an 'elementary area' in the surface.

By an analogous argument to that in the previous subsection, § 4.3.3, a rectangle of area $\delta u \delta v$ at (u_0, v_0) in the uv-plane is mapped by the the transformation (4.21) into a region that, for small $\delta u \delta v$, is roughly an almost planar parallelogram in the surface $\mathcal{S}$. As in (4.11a) and (4.11b) the directions of the sides of the parallelogram are given by

$$\begin{aligned} \mathbf{x}(u_0 + \delta u, u_0) - \mathbf{x}(u_0, v_0) &= \frac{\partial \mathbf{x}}{\partial u}(u_0, v_0)\delta u + o(\delta u)\,, && (4.23a) \\ \mathbf{x}(u_0, v_0 + \delta v) - \mathbf{x}(u_0, v_0) &= \frac{\partial \mathbf{x}}{\partial v}(u_0, v_0)\delta v + o(\delta v)\,. && (4.23b) \end{aligned}$$

Hence as δu, $\delta v \to 0$ the elementary area within the surface $\mathcal{S}$ is given by

$$\mathrm{d}S = \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| \mathrm{d}u\,\mathrm{d}v\,, \tag{4.24}$$

where $\mathrm{d}u\,\mathrm{d}v$ is [sloppyish] notation for $\mathrm{d}S_{uv}$.

Combining (4.22) and (4.24) we define the vector elementary area to be

$$\begin{aligned} \mathrm{d}\mathbf{S} &= \mathbf{n}\,\mathrm{d}S && (4.25a) \\ &= \left(\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right) \mathrm{d}u\,\mathrm{d}v\,, && (4.25b) \end{aligned}$$

or possibly the negative of this depending on the choice of the positive normal direction.

Examples.

1. Consider a surface specified 'explicitly' by $$\mathbf{x}(x,y) = (x, y, z(x,y)).$$

   Then
$$\begin{aligned}\frac{\partial \mathbf{x}}{\partial x} &= (1, 0, z_x)\,,\\ \frac{\partial \mathbf{x}}{\partial y} &= (0, 1, z_y)\,,\end{aligned}$$
   so that, after identifying x and y with u and v respectively in (4.25b), we have that
$$\mathrm{d}\mathbf{S} = (-z_x, -z_y, 1)\,\mathrm{d}x\,\mathrm{d}y\,, \tag{4.26a}$$
   and
$$\mathrm{d}S = (1 + z_x^2 + z_y^2)^{\frac{1}{2}}\,\mathrm{d}x\,\mathrm{d}y \;\geqslant \mathrm{d}x\,\mathrm{d}y\,. \tag{4.26b}$$

2. *Spherical Surface.* Consider a spherical surface of radius a represented parametrically by
$$\mathbf{x}(\theta,\phi) = (a\sin\theta\cos\phi\,,\ a\sin\theta\sin\phi\,,\ a\cos\theta)\,.$$
   Then
$$\begin{aligned}\mathbf{x}_\theta &= (\ \ a\cos\theta\cos\phi\,,\ a\cos\theta\sin\phi\,,\ -a\sin\theta)\,,\\ \mathbf{x}_\phi &= (-a\sin\theta\sin\phi\,,\ a\sin\theta\cos\phi\,,\ 0)\,,\end{aligned}$$
   so that after identifying θ and ϕ with u and v respectively in (4.25b), we have that
$$\begin{aligned}\mathrm{d}\mathbf{S} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,a^2\sin\theta\,\mathrm{d}\theta\mathrm{d}\phi\\ &= \widehat{\mathbf{x}}\,a^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\,,\end{aligned} \tag{4.27a}$$
   where $\widehat{\mathbf{x}}$ is the unit vector in the radial direction, and
$$\mathrm{d}S = a^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\,. \tag{4.27b}$$

3. *Surface of Revolution.* Consider the surface specified by (see (2.21) with $t = z$)
$$\mathbf{x}(\phi, z) = (R(z)\cos\phi, R(z)\sin\phi, z)\,. \tag{4.28}$$
   In terms of cylindrical polar co-ordinates, the surface is given by
$$\rho = (x^2+y^2)^{\frac{1}{2}} = R(z)\,, \quad 0 \leqslant \phi < 2\pi\,.$$
   Proceed geometrically. In a plane $\phi =$ constant, the length of a line element on the surface is given by
$$\mathrm{d}s = \left(\mathrm{d}\rho^2 + \mathrm{d}z^2\right)^{\frac{1}{2}} = \left(R_z^2 + 1\right)^{\frac{1}{2}}\mathrm{d}z\,.$$
   Hence
$$\mathrm{d}S = R\,\mathrm{d}\phi\,\left(R_z^2+1\right)^{\frac{1}{2}}\mathrm{d}z = \left(1+R_z^2\right)^{\frac{1}{2}}R\,\mathrm{d}\phi\,\mathrm{d}z\,. \tag{4.29}$$

   Exercise. Obtain same result using (4.25b), and also show that
$$\mathrm{d}\mathbf{S} = (\cos\phi, \sin\phi, -R_z)R\,\mathrm{d}\phi\,\mathrm{d}z\,. \tag{4.30}$$

Answer. From (4.28)

$$\begin{aligned}\mathbf{x}_\phi &= (-R\sin\phi, R\cos\phi, 0)\,,\\ \mathbf{x}_z &= (R_z\cos\phi, R_z\sin\phi, 1)\,.\end{aligned}$$

Hence from (4.25b), and compare with (2.22),

$$\begin{aligned}\mathrm{d}\mathbf{S} &= (R\cos\phi, R\sin\phi, -RR_z)\,\mathrm{d}\phi\,\mathrm{d}z\\ &= (\cos\phi, \sin\phi, -R_z)R\,\mathrm{d}\phi\,\mathrm{d}z\,,\\ \mathrm{d}S &= (1+R_z^2)^{\frac{1}{2}}R\,\mathrm{d}\phi\,\mathrm{d}z\,.\end{aligned}$$

4.4.2 Evaluation of Surface Integrals

The above expressions for $\mathrm{d}S$ and $\mathrm{d}\mathbf{S}$ effectively reduce surface integrals to area integrals. There are at least five types of surface integral encountered in physical applications:

Scalars: $\displaystyle\iint_{\mathcal{S}} f\,\mathrm{d}S\,,\qquad \iint_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S}$

Vectors: $\displaystyle\iint_{\mathcal{S}} f\,\mathrm{d}\mathbf{S}\,,\qquad \iint_{\mathcal{S}} \mathbf{u}\,\mathrm{d}S\,,\qquad \iint_{\mathcal{S}} \mathbf{u}\times\mathrm{d}\mathbf{S}$

Examples.

1. Calculate
$$I = \iint_{\mathcal{S}} xy\,\mathrm{d}S\,,$$
where $\mathcal{S}$ is the portion of the surface $x+y+z=a$ in $x\geqslant 0$, $y\geqslant 0$, $z\geqslant 0$.

   From (4.26b) with $z = a-x-y$,
$$\mathrm{d}S = (1+z_x^2+z_y^2)^{\frac{1}{2}}\,\mathrm{d}x\,\mathrm{d}y = \sqrt{3}\,\mathrm{d}x\,\mathrm{d}y\,.$$
   Hence, if $\mathcal{A}$ is the projection of $\mathcal{S}$ onto the xy-plane, then
$$\begin{aligned}I &= \sqrt{3}\iint_{\mathcal{A}} xy\,\mathrm{d}x\,\mathrm{d}y\\ &= \sqrt{3}\int_0^a \mathrm{d}x\int_0^{a-x}\mathrm{d}y\;xy\\ &= \sqrt{3}\int_0^a \mathrm{d}x\;\tfrac{1}{2}x(a-x)^2\\ &= \frac{a^4}{8\sqrt{3}}\,.\end{aligned}$$

2. Integrals such as
$$\mathbf{I} = \iint_{\mathcal{S}} \mathbf{u}\,\mathrm{d}S$$
can be evaluated component by component.

3. Evaluate
$$\mathbf{I} = \iint_{\mathcal{S}} \mathrm{d}\mathbf{S} \tag{4.31}$$

where $\mathcal{S}$ is a closed surface.

Do this component by component starting with $\mathbf{I}.\,\widehat{\mathbf{z}}$. Then, noting that the scalar product $\widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S}$ gives the area of the element of surface when projected onto the xy-plane, we split $\mathbf{I}.\,\widehat{\mathbf{z}}$ into two:

$$\mathbf{I}.\,\widehat{\mathbf{z}} = \iint_{\widehat{\mathbf{z}}.\mathrm{d}\mathbf{S}>0} \widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S} + \iint_{\widehat{\mathbf{z}}.\mathrm{d}\mathbf{S}<0} \widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S}\,,$$

where the first/second integral is the projection of the 'upper'/'lower' surface onto (x, y).

For a closed surface the two projections must be equal and opposite, hence

$$\mathbf{I}.\,\widehat{\mathbf{z}} = 0\,. \tag{4.32}$$

Similarly, $\mathbf{I}.\,\widehat{\mathbf{x}} = 0$ and $\mathbf{I}.\,\widehat{\mathbf{y}} = 0$. Hence $\mathbf{I} = 0$.

4. Evaluate
$$I = \iint_{\mathcal{S}} \mathbf{x}.\,\mathrm{d}\mathbf{S}$$
where $\mathcal{S}$ is the surface of the cone bounded by $(x^2 + y^2)^{\frac{1}{2}} + z = 1$ and $z = 0$, and the positive normal direction is *out of* the body.

On $z = 0$, including the unit disc $\mathcal{S}_1$ that is the base of the surface,
$$\mathbf{x}.\,\mathbf{n} = -z = 0\,.$$

Hence
$$I = \iint_{\mathcal{S}_1} \mathbf{x}.\,\mathrm{d}\mathbf{S} + \iint_{\mathcal{S}_2} \mathbf{x}.\,\mathrm{d}\mathbf{S} = \iint_{\mathcal{S}_2} \mathbf{x}.\,\mathrm{d}\mathbf{S}\,.$$
On $\mathcal{S}_2$, which is a surface of revolution,
$$\rho = (x^2 + y^2)^{\frac{1}{2}} = R(z) \equiv 1 - z\,. \tag{4.33}$$

Hence from (4.30)
$$\begin{aligned} \mathrm{d}\mathbf{S} &= (\cos\phi, \sin\phi, -R_z) R\,\mathrm{d}\phi\,\mathrm{d}z \\ &= (\cos\phi, \sin\phi, 1)\, R\,\mathrm{d}\phi\,\mathrm{d}z\,, \end{aligned}$$

where we have ensured that the normal direction is *out of* the body (see also (2.22)). Hence on $\mathcal{S}_2$
$$\begin{aligned} \mathbf{x}.\,\mathrm{d}\mathbf{S} &= (R\cos^2\phi + R\sin^2\phi + z)\, R\,\mathrm{d}\phi\,\mathrm{d}z \\ &= (1 - z)\,\mathrm{d}\phi\,\mathrm{d}z\,, \end{aligned}$$

from using (4.33). It follows that
$$\iint_{\mathcal{S}_2} \mathbf{x}.\,\mathrm{d}\mathbf{S} \;=\; \int_0^1 \mathrm{d}z \int_0^{2\pi} \mathrm{d}\phi\;(1 - z) = \pi\,.$$

Exercise. Repeat using (4.26a) and
$$\mathbf{x}(x, y) = \left(x, y, 1 - \left(x^2 + y^2\right)^{\frac{1}{2}}\right)\,.$$

14/00

4.4.3 Flux

Definition. Let $\mathbf{v}(\mathbf{x}, t)$ be a vector field. Then the *flux*, Q, of $\mathbf{v}$ through a surface $\mathcal{S}$ is defined to be

$$Q = \iint_{\mathcal{S}} \mathbf{v} \,.\, \mathrm{d}\mathbf{S} \,. \tag{4.34}$$

14/99

Example. Suppose that $\mathbf{v}(\mathbf{x}, t)$ describes the velocity field of the flow of a fluid, e.g. air through or out of a jet engine, water through a hose-pipe, blood flow through arteries/veins, sweat, tears, etc. Suppose we wish to measure how much fluid is coming out of a surface $\mathcal{S}$ per unit time.

Fluid flowing parallel to a surface does not cross it, hence it is the component of velocity normal to $\mathcal{S}$ that matters, i.e. $\mathbf{v} \,.\, \mathbf{n}$. In time δt, the volume of fluid crossing a small area $\mathrm{d}S$ is

$$(\mathbf{v} \,.\, \mathbf{n}) \, \delta t \, \mathrm{d}S = (\mathbf{v} \,.\, \mathrm{d}\mathbf{S}) \, \delta t \,.$$

Therefore in unit time the volume of fluid crossing $\mathcal{S}$ is

$$Q = \iint_{\mathcal{S}} \mathbf{v} \,.\, \mathrm{d}\mathbf{S} \,.$$

This is referred to as the *volume flux* out of $\mathcal{S}$.

4.5 Unlectured Worked Exercise: Examples Sheet 2, Question 18.

Question. Consider the integral

$$I = \iint_{\mathcal{D}} \frac{ky \sin kr_1 \sin kr_2}{r_1 r_2} \, \mathrm{d}x \, \mathrm{d}y \,,$$

where k is a constant,

$$r_1 = [(x+1)^2 + y^2]^{\frac{1}{2}} , \quad r_2 = [(x-1)^2 + y^2]^{\frac{1}{2}} ,$$

and $\mathcal{D}$ is the half of the ellipse $r_1 + r_2 \leqslant 4$ which is in $y \geqslant 0$. By means of the transformation

$$u = \tfrac{1}{2}(r_1 + r_2) \,, \quad v = \tfrac{1}{2}(r_1 - r_2) \,,$$

or otherwise, show that the integral is equal to $4 \sin^3 k \cos k$.

Answer. First consider the $y = 0$ boundary, on which

$$r_1 = |x+1| \quad \text{and} \quad r_2 = |x-1| \,.$$

Hence on $y = 0$,

$$\begin{array}{llcllcll}
\text{for} & 1 \leqslant x \leqslant 2 & : & r_1 = x+1 \,, & r_2 = x-1 & \& & r_1 - r_2 = 2 \,, & r_1 + r_2 = 2x \,, \\
\text{for} & -1 \leqslant x \leqslant 1 & : & r_1 = x+1 \,, & r_2 = 1-x & \& & r_1 - r_2 = 2x \,, & r_1 + r_2 = 2 \,, \\
\text{for} & -2 \leqslant x \leqslant -1 & : & r_1 = -x-1 \,, & r_2 = 1-x & \& & r_2 - r_1 = -2 \,, & r_1 + r_2 = -2x \,.
\end{array}$$

Thus if, as suggested,

$$u = \tfrac{1}{2}(r_1 + r_2)\,, \quad v = \tfrac{1}{2}(r_1 - r_2)\,,$$

it follows that the boundary can be split into four parts specified by:

$$\begin{array}{rrcl} \text{(i)} & v & = & -1\,, \\ \text{(ii)} & u & = & 1\,, \\ \text{(iii)} & v & = & 1\,, \\ \text{(iv)} & u & = & 2\,. \end{array}$$

Further, since

$$r_1^2 = (x+1)^2 + y^2\,, \quad r_2^2 = (x-1)^2 + y^2\,,$$

it follows that

$$r_1 r_{1x} = x + 1\,, \quad r_1 r_{1y} = y\,, \quad r_2 r_{2x} = x - 1\,, \quad r_2 r_{2y} = y\,,$$

and hence that

$$\frac{\partial\,(r_1, r_2)}{\partial\,(x, y)} = r_{1x} r_{2y} - r_{1y} r_{2x} = \frac{(x+1)\,y}{r_1 r_2} - \frac{(x-1)\,y}{r_1 r_2} = \frac{2y}{r_1 r_2}\,.$$

Also

$$\frac{\partial\,(u, v)}{\partial\,(r_1 r_2)} = u_{r_1} v_{r_2} - u_{r_2} v_{r_1} = \tfrac{1}{2}.\left(-\tfrac{1}{2}\right) - \tfrac{1}{2}.\tfrac{1}{2} = -\tfrac{1}{2}\,,$$

and so

$$\frac{\partial\,(u, v)}{\partial\,(x, y)} = \frac{\partial\,(u, v)}{\partial\,(r_1, r_2)}\,\frac{\partial\,(r_1, r_2)}{\partial\,(x, y)} = -\frac{y}{r_1 r_2}\,.$$

We note that the Jacobian is negative in $\mathcal{D}$ since the handedness is not preserved under the co-ordinate transformation. Further, we also note that for $0 < k \ll 1$

$$I \approx \iint_{\mathcal{D}} k^3 y \,\mathrm{d}x\,\mathrm{d}y > 0\,.$$

Hence we conclude that

$$\begin{aligned} I &= \int_1^2 \mathrm{d}u \int_{-1}^1 \mathrm{d}v \;\left|\frac{\partial\,(x, y)}{\partial\,(u, v)}\right| \frac{ky \sin kr_1 \,\sin kr_2}{r_1 r_2} \\ &= \tfrac{1}{2}k \int_1^2 \mathrm{d}u \int_{-1}^1 \mathrm{d}v \,[\cos k\,(r_1 - r_2) - \cos k\,(r_1 + r_2)] \\ &= \tfrac{1}{2}k \int_1^2 \mathrm{d}u \int_{-1}^1 \mathrm{d}v \,[\cos 2kv - \cos 2ku] \\ &= \tfrac{1}{2}k \left[\left[\frac{1}{2k}\sin 2kv\right]_{-1}^1 - 2\left[\frac{1}{2k}\sin 2ku\right]_1^2\right] \\ &= \sin 2k - \tfrac{1}{2}\sin 4k = \sin 2k\,(1 - \cos 2k) = 4\sin^3 k \cos k\,, \end{aligned}$$

where we have checked that the sign is correct by noting that $I > 0$ when $0 < k \ll 1$.

5 Vector Differential Operators

Recall that there are various product operations that we can perform on vectors:

$$\begin{aligned} \text{scalar multiplication}: &\quad \mathbf{a}\lambda \\ \text{scalar product}: &\quad \mathbf{a}.\mathbf{b} \\ \text{vector product}: &\quad \mathbf{a}\times\mathbf{b} \end{aligned}$$

The vector differential operator 'del' or 'nabla', defined in (1.51),

$$\boldsymbol{\nabla} = \sum_{j=1}^{m} \mathbf{e}_j \frac{\partial}{\partial x_j} = \sum_{j=1}^{m} \mathbf{e}_j \,\partial_{x_j} = \sum_{j=1}^{m} \mathbf{e}_j \,\partial_j \,, \tag{5.1}$$

can be involved in such products. For instance, 'grad' is analogous to scalar multiplication: $\boldsymbol{\nabla} f$. As we shall see 'div' and 'curl' are analogous to the scalar and vector product respectively.

In some ways (e.g. at a technical or 'recipe' level) these new operators are not difficult, but in other ways (e.g. at an interpretive level) they require more thought. In order to do most examination questions expertise at a technical level is what is required.

5.1 Revision of Gradient

For $f:\mathbb{R}^m \to \mathbb{R}$, in (1.43a) we defined the gradient of f at $\mathbf{a}$ to be

$$\operatorname{grad} f = \boldsymbol{\nabla} f = \sum_{j=1}^{m} \mathbf{e}_j \frac{\partial f}{\partial x_j}(\mathbf{a}) \,, \quad \text{i.e.} \quad \left[\operatorname{grad} f\right]_i = \left[\boldsymbol{\nabla} f\right]_i = \frac{\partial f}{\partial x_i} \,. \tag{5.2}$$

For $f:\mathbb{R}^3 \to \mathbb{R}$, we noted in (1.44a) and (1.44b) that the gradient of f can be written

$$\boldsymbol{\nabla} f = \mathbf{i}\frac{\partial f}{\partial x} + \mathbf{j}\frac{\partial f}{\partial y} + \mathbf{k}\frac{\partial f}{\partial z} \tag{5.3a}$$

$$= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \,. \tag{5.3b}$$

5.1.1 Summary of Properties

1. In § 1.7.1 we showed that the gradient is a vector, i.e. that for two orthonormal bases $\mathbf{e}_j$ and $\mathbf{e}'_j$ $(j = 1, \dots, m)$
$$\boldsymbol{\nabla} f = \sum_{j=1}^{m} \mathbf{e}_j \frac{\partial f}{\partial x_j} = \sum_{j=1}^{m} \mathbf{e}'_j \frac{\partial f}{\partial x'_j} \,.$$
2. From (1.52), the directional derivative in the direction $\widehat{\mathbf{u}}$ is
$$f'(\mathbf{a};\widehat{\mathbf{u}}) = \widehat{\mathbf{u}}.\boldsymbol{\nabla} f \,.$$
This has maximum value when $\widehat{\mathbf{u}}$ is parallel to $\boldsymbol{\nabla} f$.
3. From (2.20), for a function $F:\mathbb{R}^3 \to \mathbb{R}$ and a surface defined by $F(\mathbf{x}) = 0$, $\boldsymbol{\nabla} F$ is normal to the surface.
4. From (2.25), at a critical/stationary point
$$\boldsymbol{\nabla} f = 0 \,,$$
and from the property listed after (2.20), $\boldsymbol{\nabla} f = 0$ at an interior maximum or minimum of f.
5. From (3.18), for differentiable $f:\mathbb{R}^3 \to \mathbb{R}$
$$\mathrm{d}f = \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x} = f_x\,\mathrm{d}x + f_y\,\mathrm{d}y + f_z\,\mathrm{d}z \,.$$
6. From (3.13) and (3.18) for a curve $\mathcal{C}$ running from $\mathbf{x}_a$ to $\mathbf{x}_b$
$$\int_{\mathcal{C}} \mathrm{d}f = \int_{\mathcal{C}} \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x} = f(\mathbf{x}_b) - f(\mathbf{x}_a) \,.$$

5.2 Divergence

5.2.1 Divergence in Cartesian Co-ordinates

The divergence is analogous to the scalar product. Let the vector function $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$ be differentiable, then for a *fixed* orthonormal basis $\mathbf{e}_j$ $(j = 1, \ldots, m)$ the divergence of $\mathbf{F}$ is defined to be (see (5.6) for a 'better' definition)

$$\begin{aligned} \operatorname{div}\mathbf{F} = \boldsymbol{\nabla}.\,\mathbf{F} &= \sum_{j=1}^{m} \mathbf{e}_j \frac{\partial}{\partial x_j}.\sum_{k=1}^{m} \mathbf{e}_k\, F_k && (5.4a)\\ &= \sum_{j,k=1}^{m} \mathbf{e}_j.\,\mathbf{e}_k \frac{\partial F_k}{\partial x_j} = \sum_{j,k=1}^{m} \delta_{jk} \frac{\partial F_k}{\partial x_j} \\ &= \sum_{j=1}^{m} \frac{\partial F_j}{\partial x_j} \\ &= \frac{\partial F_j}{\partial x_j} \qquad (s.c.)\,. && (5.4b) \end{aligned}$$

For a vector function $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$:

$$\begin{aligned} \operatorname{div}\mathbf{F} = \boldsymbol{\nabla}.\,\mathbf{F} &= (\mathbf{i}\,\partial_x + \mathbf{j}\,\partial_y + \mathbf{k}\,\partial_z)\,.\,(\mathbf{i}F_1 + \mathbf{j}F_2 + \mathbf{k}F_3) && (5.5a)\\ &= \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\,. && (5.5b) \end{aligned}$$

Remarks.

- While the scalar product is commutative (i.e. $\mathbf{a}.\mathbf{b} = \mathbf{b}.\mathbf{a}$ for two constant vectors), $\boldsymbol{\nabla}.\,\mathbf{F} \neq \mathbf{F}.\,\boldsymbol{\nabla}$.
- Let T be the Jacobian matrix of $\mathbf{F}(\mathbf{x})$, i.e. from (1.26) $T_{ij} = \frac{\partial F_i}{\partial x_j}$. Then from (5.4b) the divergence is seen to be the trace of the Jacobian matrix, i.e. $(s.c.)$

$$\operatorname{div}\mathbf{F} = \boldsymbol{\nabla}.\,\mathbf{F} = \frac{\partial F_j}{\partial x_j} = T_{jj} = \operatorname{tr}(\mathsf{T})\,. \qquad (5.6)$$

- For $\mathbf{F}, \mathbf{G} : \mathbb{R}^m \to \mathbb{R}^m$ and $\lambda, \mu \in \mathbb{R}$ it follows from definition (5.4b) that

$$\boldsymbol{\nabla}.\,(\lambda\mathbf{F} + \mu\mathbf{G}) = \frac{\partial}{\partial x_j}(\lambda F_j + \mu G_j) = \lambda\frac{\partial F_j}{\partial x_j} + \mu\frac{\partial G_j}{\partial x_j} = \lambda\boldsymbol{\nabla}.\,\mathbf{F} + \mu\boldsymbol{\nabla}.\,\mathbf{G}\,.$$

Definition.

If $\boldsymbol{\nabla}.\,\mathbf{F} = 0$ in a region, $\mathbf{F}$ is said to be *solenoidal* in that region.

5.2.2 Examples

1. Let

$$\mathbf{F} = (x^2y,\, y^2z,\, z^2x)\,,$$

then from definition (5.5b)

$$\boldsymbol{\nabla}.\,\mathbf{F} = \frac{\partial(x^2y)}{\partial x} + \frac{\partial(y^2z)}{\partial y} + \frac{\partial(z^2x)}{\partial z} = 2xy + 2yz + 2zx\,.$$

2. In $\mathbb{R}^3$ let $\mathbf{F} = \mathbf{x}$, then from definition (5.4b) and (1.17)

$$\boldsymbol{\nabla} . \mathbf{F} \;=\; \boldsymbol{\nabla} . \mathbf{x} \;=\; \frac{\partial x_j}{\partial x_j} \;=\; \delta_{jj} \;=\; 3\,, \tag{5.7a}$$

i.e.

$$\boldsymbol{\nabla} . \mathbf{x} \;=\; \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} \;=\; 3\,. \tag{5.7b}$$

3. In $\mathbb{R}^3$ let $\mathbf{F} = \frac{\mathbf{x}}{r^3}$, then from definition (5.4b), (1.17) and (1.18)

$$\boldsymbol{\nabla} . \left(\frac{\mathbf{x}}{r^3}\right) = \frac{\partial}{\partial x_j}\left(\frac{x_j}{r^3}\right) = \frac{\delta_{jj}}{r^3} - 3\frac{x_j}{r^4}\,\frac{x_j}{r} = \frac{3}{r^3} - \frac{3r^2}{r^5} = 0\,. \tag{5.7c}$$

5.2.3 Another Basis

Suppose that $\mathbf{e}'_j$ $(j = 1, \ldots, m)$ is another *fixed* orthonormal basis in $\mathbb{R}^3$. Earlier in § 1.7.1 we saw that

$$\boldsymbol{\nabla} = \mathbf{e}_j\,\frac{\partial}{\partial x_j} = \mathbf{e}'_k\,\frac{\partial}{\partial x'_k} \qquad (s.c.)$$

is a vector. Hence from the definition (5.4a) we conclude for a vector field

$$\mathbf{F} = \mathbf{e}_j\,F_j = \mathbf{e}'_k\,F'_k \qquad (s.c.)$$

that

$$\boldsymbol{\nabla} . \mathbf{F} = \frac{\partial F_j}{\partial x_j} = \frac{\partial F'_k}{\partial x'_k} \qquad (s.c.)\,.$$

5.2.4 Physical Interpretation

Let $\mathbf{v}(\mathbf{x})$ be a vector field. Consider the flux, Q, of $\mathbf{v}$ out of the cuboid with a vertex at (x_0, y_0, z_0), and sides of length δx, δy and δz parallel to the co-ordinate axes. From definition (4.34)

$$Q = \iint\limits_{\mathcal{S}} \mathbf{v} . \,\mathrm{d}\mathbf{S}\,.$$

Consider the flux through each of the six faces of the cuboid in turn. In the case of the surfaces at $x_0 + \delta x$ and x_0 the flux is given respectively by

$$Q_{x_0+\delta x} \equiv \iint\limits_{\mathcal{S}_{x_0+\delta x}} \mathbf{v} . \,\mathrm{d}\mathbf{S} \;=\; \int_{y_0}^{y_0+\delta y} \mathrm{d}y \int_{z_0}^{z_0+\delta z} \mathrm{d}z\;\, v_1(x_0 + \delta x, y, z)\,,$$

$$Q_{x_0} \equiv \iint\limits_{\mathcal{S}_{x_0}} \mathbf{v} . \,\mathrm{d}\mathbf{S} \;=\; -\int_{y_0}^{y_0+\delta y} \mathrm{d}y \int_{z_0}^{z_0+\delta z} \mathrm{d}z\;\, v_1(x_0, y, z)\,,$$

where in the second integral the minus sign arises because the normal is in the negative x-direction. Next add the two integrals in order to give the net flux in the x-direction, assume that

$$\delta x, \delta y, \delta z \to 0,$$

and expand the integrand using Taylor's theorem, to obtain

$$\begin{aligned}
Q_{x_0+\delta x} + Q_{x_0} &= \int_{y_0}^{y_0+\delta y} \mathrm{d}y \int_{z_0}^{z_0+\delta z} \mathrm{d}z \left[\frac{\partial v_1}{\partial x}(x_0, y, z)\ \delta x + o(\delta x)\right] \\
&= \int_0^1 d\beta\, \delta y \int_0^1 d\gamma\, \delta z \left[\frac{\partial v_1}{\partial x}(x_0, y_0 + \beta\delta y, z_0 + \gamma\delta z)\, \delta x + o(\delta x)\right] \\
&\qquad \text{where} \quad y = y_0 + \beta\delta y,\ z = z_0 + \gamma\delta z \quad (\text{cf. } \S\, 3.4.1) \\
&= \frac{\partial v_1}{\partial x}(x_0, y_0, z_0)\, \delta x\, \delta y\, \delta z + o(\delta x\, \delta y\, \delta z) \\
&\qquad \text{by a further application of Taylor's theorem.}
\end{aligned}$$

Similarly for $(Q_{y_0+\delta y} + Q_{y_0})$ and $(Q_{z_0+\delta z} + Q_{z_0})$. Let $\mathcal{V} = \delta x\, \delta y\, \delta z$ denote the volume of the cuboid, then

$$\begin{aligned}
Q &= Q_{x_0+\delta x} + Q_{x_0} + Q_{y_0+\delta y} + Q_{y_0} + Q_{z_0+\delta z} + Q_{z_0} \\
&= \left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right) \mathcal{V} + o(\mathcal{V}) \\
&= (\boldsymbol{\nabla} . \mathbf{v})\, \mathcal{V} + o(\mathcal{V}),
\end{aligned}$$

and so

$$\boldsymbol{\nabla} . \mathbf{v} = \lim_{\mathcal{V}\to 0} \frac{1}{\mathcal{V}} \iint_{\mathcal{S}} \mathbf{v} . \mathrm{d}\mathbf{S}. \tag{5.8}$$

Thus $\boldsymbol{\nabla} . \mathbf{v}$ can be interpreted as the net rate of flux outflow at $\mathbf{x}_0$ per unit volume.

Remark. If $\mathbf{v}$ is a an incompressible velocity field then the flux Q is the volume of fluid coming out of $\mathcal{S}$ per unit time. Hence

$$\begin{aligned}
\boldsymbol{\nabla} . \mathbf{v} > 0 &\quad \Rightarrow \text{net flow out} \quad \Rightarrow \text{there exists a source at } \mathbf{x}_0; \\
\boldsymbol{\nabla} . \mathbf{v} < 0 &\quad \Rightarrow \text{net flow in} \quad\ \ \Rightarrow \text{there exists a sink at } \mathbf{x}_0.
\end{aligned}$$

5.3 Curl

5.3.1 Curl in Cartesian Co-ordinates

The curl is analogous to the vector product. We confine attention to $\mathbb{R}^3$, and suppose that the vector function $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is differentiable, then the curl of $\mathbf{F}$ is defined to be

$$\begin{aligned}
\operatorname{curl} \mathbf{F} = \boldsymbol{\nabla} \times \mathbf{F} &= (\mathbf{i}\, \partial_x + \mathbf{j}\, \partial_y + \mathbf{k}\, \partial_z) \times (\mathbf{i}\, F_1 + \mathbf{j}\, F_2 + \mathbf{k}\, F_3) \\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ F_1 & F_2 & F_3 \end{vmatrix} \qquad (5.9\text{a}) \\
&= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right), \qquad (5.9\text{b})
\end{aligned}$$

where the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ define a fixed right-handed orthonormal basis. Suppose instead we denote the *fixed* orthonormal basis by $\mathbf{e}_j$ $(j = 1, 2, 3)$, then if it is right-handed

$$\varepsilon_{ijk}\, \mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k \qquad (s.c.)\,, \tag{5.10}$$

and hence an alternative expression for the curl of $\mathbf{F}$ is

$$\begin{aligned}
\operatorname{curl}\mathbf{F} \;=\; \boldsymbol{\nabla}\times\mathbf{F} \;&=\; \sum_{j=1}^{3}\mathbf{e}_j\,\frac{\partial}{\partial x_j}\times\sum_{k=1}^{3}\mathbf{e}_k\,F_k & (5.11a)\\
&=\; \sum_{j,k=1}^{3}\mathbf{e}_j\times\mathbf{e}_k\,\frac{\partial F_k}{\partial x_j}\\
&=\; \sum_{i,j,k=1}^{3}\varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial F_k}{\partial x_j}\\
&=\; \varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial F_k}{\partial x_j}\qquad (s.c.)\,, & (5.11b)
\end{aligned}$$

i.e.

$$[\operatorname{curl}\mathbf{F}]_i \;=\; [\boldsymbol{\nabla}\times\mathbf{F}]_i \;=\; \varepsilon_{ijk}\,\frac{\partial F_k}{\partial x_j}\qquad (s.c.)\,. \tag{5.11c}$$

Remarks.

- Let T be the Jacobian matrix of $\mathbf{F}(\mathbf{x})$, i.e. from (1.26) $T_{ij} = \frac{\partial F_i}{\partial x_j}$. Then from (5.11b) an alternative expression for the curl is

  $$\operatorname{curl}\mathbf{F} \;=\; \boldsymbol{\nabla}\times\mathbf{F} \;=\; \varepsilon_{ijk}\,\mathbf{e}_i\,T_{kj}\qquad (s.c.)\,.$$

  Hence the components of curl are the components (upto a sign) of the skew-symmetric part of the Jacobian matrix (cf. (5.9b)), i.e. the components of the matrix

  $$\tfrac{1}{2}(\mathsf{T}-\mathsf{T}^t) = \tfrac{1}{2}\begin{pmatrix} 0 & \frac{\partial F_1}{\partial y}-\frac{\partial F_2}{\partial x} & \frac{\partial F_1}{\partial z}-\frac{\partial F_3}{\partial x}\\ \frac{\partial F_2}{\partial x}-\frac{\partial F_1}{\partial y} & 0 & \frac{\partial F_2}{\partial z}-\frac{\partial F_3}{\partial y}\\ \frac{\partial F_3}{\partial x}-\frac{\partial F_1}{\partial z} & \frac{\partial F_3}{\partial y}-\frac{\partial F_2}{\partial z} & 0\end{pmatrix}.$$

  15/99 It follows that if the Jacobian matrix of $\mathbf{F}$ is symmetric, then $\operatorname{curl}\mathbf{F}=0$.

- For $\mathbf{F},\mathbf{G}:\mathbb{R}^3\to\mathbb{R}^3$ and $\lambda,\mu\in\mathbb{R}$ it follows from definition (5.11b) that $(s.c.)$

  $$\operatorname{curl}(\lambda\mathbf{F}+\mu\mathbf{G}) = \varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial}{\partial x_j}(\lambda F_k+\mu G_k) = \lambda\varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial F_k}{\partial x_j} + \mu\varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial G_k}{\partial x_j} = \lambda\operatorname{curl}\mathbf{F}+\mu\operatorname{curl}\mathbf{G}\,.$$

15/00

- The necessary and sufficient condition for $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ to be exact in 3D, i.e. (3.26b), can alternatively be written as

  $$\operatorname{curl}\mathbf{F}=0\,. \tag{5.12}$$

5.3.2 Irrotational Fields

Definition.

If $\operatorname{curl}\mathbf{F}=0$ in a region, the vector field $\mathbf{F}(\mathbf{x})$ is said to be *irrotational* in that region.

Poincaré's Lemma.

If $\operatorname{curl}\mathbf{F}=0$ in a simply connected region, then there exists a scalar potential $f(\mathbf{x})$ such that

$$\mathbf{F}=\boldsymbol{\nabla}f\,. \tag{5.13}$$

A proof of this lemma will be given in § 6.3.6 once we have Stokes' theorem at our disposal.[7]

[7] *For those who desire instant gratification.* First note that if $\operatorname{curl}\mathbf{F}=0$, then from (3.26b) and (5.12) we can conclude from the [unproved] sufficiency condition that $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is exact. Thence from (3.4) it follows that $\int_C \mathbf{F}.\,\mathrm{d}\mathbf{x}$ depends only on the end points, i.e. it follows that $\mathbf{F}$ is a conservative field. Thus from (3.32a) a scalar potential f can be defined such that

$$\mathbf{F}=\boldsymbol{\nabla}f\,.$$

5.3.3 Examples

A Useful Result. First let us prove a useful result. If A is a symmetric matrix so that $A_{jk} = A_{kj}$, then

$$\varepsilon_{ijk} A_{jk} = 0\,. \tag{5.14}$$

To see this note that

$$\begin{aligned} \varepsilon_{ijk}\,A_{jk} &= \varepsilon_{ikj}\,A_{kj} && \text{from swapping the } j \text{ and } k \text{ indices} \\ &= -\varepsilon_{ijk}\,A_{jk} && \text{from } \varepsilon_{ijk} = -\varepsilon_{ikj} \text{ and the symmetry of } \mathsf{A}. \end{aligned}$$

Hence

$$2\varepsilon_{ijk}\,A_{jk} = 0\,.$$

1. Suppose

   $$\mathbf{F} = (x^2y,\, y^2z,\, z^2x)\,,$$

   then from (5.9a)

   $$\boldsymbol{\nabla}\times\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ x^2y & y^2z & z^2x \end{vmatrix} = (-y^2, -z^2, -x^2)\,.$$

2. Suppose $\mathbf{F} = \mathbf{x}$, then from (5.11c) and (5.14)

   $$[\boldsymbol{\nabla}\times\mathbf{F}]_i = \varepsilon_{ijk}\,\frac{\partial}{\partial x_j}\,x_k = \varepsilon_{ijk}\,\delta_{jk} = 0\,. \tag{5.15}$$

3. Suppose $\mathbf{F} = \frac{\mathbf{x}}{r^3}$, then from (1.18), (5.11c) and (5.14)

   $$[\boldsymbol{\nabla}\times\mathbf{F}]_i = \varepsilon_{ijk}\,\frac{\partial}{\partial x_j}\left(\frac{x_k}{r^3}\right) = \varepsilon_{ijk}\,\frac{\delta_{jk}}{r^3} - 3\,\varepsilon_{ijk}\,\frac{x_k}{r^4}\,\frac{x_j}{r} = 0\,. \tag{5.16}$$

5.3.4 Another Basis

Suppose that $\mathbf{e}'_j$ $(j = 1, \ldots, m)$ is another *fixed* orthonormal basis in $\mathbb{R}^3$, and that it is right-handed, i.e.

$$\varepsilon_{ijk}\,\mathbf{e}'_i = \mathbf{e}'_j \times \mathbf{e}'_k\,.$$

Then since

$$\boldsymbol{\nabla} = \mathbf{e}_j\,\frac{\partial}{\partial x_j} = \mathbf{e}'_j\,\frac{\partial}{\partial x'_j} \qquad (s.c.),$$

it follows from the definition (5.11a) of curl and the lines leading to (5.11b), that if $\mathbf{F} = \mathbf{e}_k\,F_k = \mathbf{e}'_k\,F'_k$ is a vector field then

$$\operatorname{curl}\mathbf{F} = \varepsilon_{ijk}\,\mathbf{e}_i\,\frac{\partial F_k}{\partial x_j} = \varepsilon_{ijk}\,\mathbf{e}'_i\,\frac{\partial F'_k}{\partial x'_j}\,.$$

5.3.5 Physical Interpretation

1. Consider a rigid body rotating with angular velocity $\boldsymbol{\omega}$ about an axis through $\mathbf{0}$. Then the velocity at a point $\mathbf{x}$ in the body is given by

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{x} \,.$$

Moreover from (5.11c)

$$\begin{aligned} [\boldsymbol{\nabla} \times \mathbf{v}]_i &= \varepsilon_{ijk}\, \partial_j v_k \\ &= \varepsilon_{ijk}\, \partial_j\, \varepsilon_{k\ell m}\, \omega_\ell x_m \\ &= (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})\, \underbrace{\partial_j(\omega_\ell x_m)}_{\omega_\ell \delta_{jm}} \\ &= \omega_i \delta_{jj} - \delta_{ij}\omega_j \\ &= 2\omega_i \,, \end{aligned}$$

i.e.

$$\boldsymbol{\nabla} \times \mathbf{v} = 2\boldsymbol{\omega} \,.$$

Hence the curl at a point gives twice the angular velocity; for a fluid $\operatorname{curl}\mathbf{v}$ is related to twice the *local* angular velocity.

The curl is a measure of the rotation of a vector field.

2. Let $\mathbf{v}$ be a vector field, and define the *circulation*, Γ, of $\mathbf{v}$ around a closed curve $\mathcal{C}$ to be

$$\Gamma = \oint_{\mathcal{C}} \mathbf{v}.\,\mathrm{d}\mathbf{x} \,. \tag{5.17}$$

Suppose that $\mathcal{C}$ is a rectangular curve in the xy-plane with a vertex at $\mathbf{x}_0$, and with sides of length δx and δy parallel to the x and y co-ordinate axes respectively.

Then for small δx and δy and by repeated application of Taylor's theorem

$$\begin{aligned} \left(\int_{\mathcal{C}_1} + \int_{\mathcal{C}_3}\right) \mathbf{v}.\,\mathrm{d}\mathbf{x} &= \int_{x_0}^{x_0+\delta x} \big[v_1(x, y_0, z_0) - v_1(x, y_0+\delta y, z_0)\big]\,\mathrm{d}x \\ &= \int_0^1 \left[-\frac{\partial v_1}{\partial y}(x_0 + \alpha\,\delta x, y_0, z_0)\,\delta y + o(\delta y)\right]\delta x\,\mathrm{d}\alpha \\ &\qquad \text{where} \quad x = x_0 + \delta x\,\alpha \\ &= -\frac{\partial v_1}{\partial y}(x_0, y_0, z_0)\,\delta x\,\delta y + o(\delta x\,\delta y) \,. \end{aligned} \tag{5.18a}$$

Similarly

$$\left(\int_{\mathcal{C}_2} + \int_{\mathcal{C}_4}\right) \mathbf{v}.\,\mathrm{d}\mathbf{x} = \frac{\partial v_2}{\partial x}(x_0, y_0, z_0)\,\delta x\,\delta y + o(\,\delta x\,\delta y) \,. \tag{5.18b}$$

Let $\mathcal{A} = \delta x\,\delta y$ be the area enclosed by $\mathcal{C}$, then from (5.18a) and (5.18b), and using (5.9b),

$$\lim_{\mathcal{A}\to 0} \frac{1}{\mathcal{A}} \oint_{\mathcal{C}} \mathbf{v}.\,\mathbf{dx} = \widehat{\mathbf{z}}.\operatorname{curl}\mathbf{v} \,. \tag{5.19}$$

$\widehat{\mathbf{z}}.\operatorname{curl}\mathbf{v}$ can thus be interpreted as the circulation about $\widehat{\mathbf{z}}$ at $\mathbf{x}_0$ per unit area. Analogous results hold for the other components of $\operatorname{curl}\mathbf{v}$.

5.4 Vector Differential Identities

There are a large number of these, for instance

$$\begin{aligned}
\boldsymbol{\nabla}.\,(f\,\mathbf{F}) &= f\,\boldsymbol{\nabla}.\,\mathbf{F} + (\mathbf{F}.\,\boldsymbol{\nabla})f\,, &\quad (5.20a)\\
\boldsymbol{\nabla}\times(f\mathbf{F}) &= f\,(\boldsymbol{\nabla}\times\mathbf{F}) + (\boldsymbol{\nabla}f)\times\mathbf{F}\,, &\quad (5.20b)\\
\boldsymbol{\nabla}.\,(\mathbf{F}\times\mathbf{G}) &= \mathbf{G}.\,(\boldsymbol{\nabla}\times\mathbf{F}) - \mathbf{F}.\,(\boldsymbol{\nabla}\times\mathbf{G})\,, &\quad (5.20c)\\
\boldsymbol{\nabla}\times(\mathbf{F}\times\mathbf{G}) &= \mathbf{F}\,(\boldsymbol{\nabla}.\,\mathbf{G}) - \mathbf{G}\,(\boldsymbol{\nabla}.\,\mathbf{F}) + (\mathbf{G}.\,\boldsymbol{\nabla})\mathbf{F} - (\mathbf{F}.\,\boldsymbol{\nabla})\mathbf{G}\,, &\quad (5.20d)\\
\boldsymbol{\nabla}(\mathbf{F}.\,\mathbf{G}) &= (\mathbf{F}.\,\boldsymbol{\nabla})\mathbf{G} + (\mathbf{G}.\,\boldsymbol{\nabla})\mathbf{F} + \mathbf{F}\times(\boldsymbol{\nabla}\times\mathbf{G}) + \mathbf{G}\times(\boldsymbol{\nabla}\times\mathbf{F})\,. &\quad (5.20e)
\end{aligned}$$

Such identities are not difficult to prove/derive using suffix notation, so they do not need to be memorised.

Notation. As far as notation is concerned, for scalar f

$$\mathbf{F}.(\boldsymbol{\nabla}f) = F_j\left(\frac{\partial f}{\partial x_j}\right) = \left(F_j\frac{\partial}{\partial x_j}\right)f = (\mathbf{F}.\boldsymbol{\nabla})f\,.$$

However, the right hand form is preferable since for vector $\mathbf{G}$, the ith component of $(\mathbf{F}.\boldsymbol{\nabla})\mathbf{G}$ is unambiguous being $F_j\frac{\partial G_i}{\partial x_j}$, while that of $\mathbf{F}.(\boldsymbol{\nabla}\mathbf{G})$ is not, i.e. it is not clear whether the ith component of $\mathbf{F}.(\boldsymbol{\nabla}\mathbf{G})$ is

$$F_j\frac{\partial G_i}{\partial x_j} \quad \text{or} \quad F_j\frac{\partial G_j}{\partial x_i}\,.$$

Example Verifications.

(5.20a). From the definition of divergence (5.4b)

$$\begin{aligned}
\boldsymbol{\nabla}.\,(f\mathbf{F}) &= \frac{\partial}{\partial x_i}(fF_i)\\
&= f\frac{\partial}{\partial x_i}F_i + F_i\frac{\partial}{\partial x_i}f\\
&= f\,\boldsymbol{\nabla}.\mathbf{F} + (\mathbf{F}.\boldsymbol{\nabla})f\,.
\end{aligned}$$

(5.20e). Start with part of the right-hand-side:

$$\begin{aligned}
\left[\mathbf{F}\times(\boldsymbol{\nabla}\times\mathbf{G}) + \mathbf{G}\times(\boldsymbol{\nabla}\times\mathbf{F})\right]_i &= \varepsilon_{ijk}\,F_j\;\varepsilon_{k\ell m}\frac{\partial G_m}{\partial x_\ell} + \varepsilon_{ijk}\,G_j\;\varepsilon_{k\ell m}\frac{\partial F_m}{\partial x_\ell}\\
&= (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})\left(F_j\frac{\partial G_m}{\partial x_\ell} + G_j\frac{\partial F_m}{\partial x_\ell}\right)\\
&= F_j\frac{\partial G_j}{\partial x_i} + G_j\frac{\partial F_j}{\partial x_i} - F_j\frac{\partial G_i}{\partial x_j} - G_j\frac{\partial F_i}{\partial x_j}\\
&= \frac{\partial}{\partial x_i}(F_j\,G_j) - \left(F_j\frac{\partial}{\partial x_j}\right)G_i - \left(G_j\frac{\partial}{\partial x_j}\right)F_i\\
&= \left[\boldsymbol{\nabla}(\mathbf{F}.\mathbf{G}) - (\mathbf{F}.\boldsymbol{\nabla})\mathbf{G} - (\mathbf{G}.\boldsymbol{\nabla})\mathbf{F}\right]_i\,.
\end{aligned}$$

Rearrange to complete the verification.

Warnings.

1. Always remember what terms the differential operator is acting on, e.g. is it all terms to the right or just some?

2. Be very very careful when using standard vector identities where you have just replaced a vector with $\boldsymbol{\nabla}$. Sometimes it works, sometimes it does not! For instance for constant vectors $\mathbf{D}$, $\mathbf{F}$ and $\mathbf{G}$

$$\mathbf{F}\,.\,(\mathbf{D}\times\mathbf{G}) \;=\; \mathbf{D}\,.\,(\mathbf{G}\times\mathbf{F}) = -\mathbf{D}\,.\,(\mathbf{F}\times\mathbf{G})\,.$$

However for $\boldsymbol{\nabla}$ and vector functions $\mathbf{F}$ and $\mathbf{G}$

$$\mathbf{F}\,.\,(\boldsymbol{\nabla}\times\mathbf{G}) \neq \boldsymbol{\nabla}\,.\,(\mathbf{G}\times\mathbf{F}) = -\boldsymbol{\nabla}\,.\,(\mathbf{F}\times\mathbf{G})\,,$$

since

$$\begin{aligned}\mathbf{F}\,.\,(\boldsymbol{\nabla}\times\mathbf{G}) &= F_i\,\varepsilon_{ijk}\,\frac{\partial G_k}{\partial x_j}\\ &= \varepsilon_{ijk}\,F_i\,\frac{\partial G_k}{\partial x_j}\,,\end{aligned}$$

while

$$\begin{aligned}\boldsymbol{\nabla}\,.\,(\mathbf{G}\times\mathbf{F}) &= \frac{\partial}{\partial x_j}\,(\varepsilon_{jki}\,G_kF_i)\\ &= \varepsilon_{ijk}\,\frac{\partial}{\partial x_j}(F_iG_k)\,.\end{aligned}$$

3. Think carefully before deciding not to use suffix notation.

5.5 Second Order Vector Differential Operators

There are a number of possible ways of combining the first order differential operators grad, div and curl to form second order differential operators, namely

grad(scalar)	div(vector)	curl(vector)
$\boldsymbol{\nabla}(\boldsymbol{\nabla}\,.\,\mathbf{F})$	$\boldsymbol{\nabla}\,.\,(\boldsymbol{\nabla}f)$	$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}f)$
	$\boldsymbol{\nabla}\,.\,(\boldsymbol{\nabla}\times\mathbf{F})$	$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{F})$

5.5.1 grad (div $\mathbf{F}$)

From the definitions of grad and div, i.e. (5.2) and (5.4b) respectively,

$$\begin{aligned}[\boldsymbol{\nabla}(\boldsymbol{\nabla}\,.\,\mathbf{F})]_i &= \frac{\partial}{\partial x_i}\,\frac{\partial F_j}{\partial x_j}\\ &= \frac{\partial^2 F_j}{\partial x_i\partial x_j}\,,\end{aligned}$$

i.e.

$$[\boldsymbol{\nabla}(\boldsymbol{\nabla}\,.\,\mathbf{F})]_x = \frac{\partial}{\partial x}\left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}\right)\,.$$

This is nothing special!

5.5.2 div (grad f) — the Laplacian operator

Again from the definitions of grad and div, i.e. (5.2) and (5.4b) respectively,

$$\begin{aligned}\boldsymbol{\nabla}\,.\,(\boldsymbol{\nabla}f) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right) &= \frac{\partial^2 f}{\partial x_j\partial x_j}\\ &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\qquad\text{in 3D}\,.\end{aligned}$$

Definition. The *Laplacian operator* ∇^2 is defined by

$$\nabla^2 = \boldsymbol{\nabla}.\boldsymbol{\nabla} = \frac{\partial^2}{\partial x_j \partial x_j} \qquad (s.c.) \tag{5.21a}$$

$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \qquad \text{in 2D} \tag{5.21b}$$

$$= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad \text{in 3D}\,. \tag{5.21c}$$

Remark. The Laplacian operator ∇^2 is very important in physics. For instance it occurs in

1. Poisson's equation for a potential $\varphi(\mathbf{x})$:

$$\nabla^2\varphi = \rho\,, \tag{5.22a}$$

   where (with a suitable normalisation)

   (a) $\rho(\mathbf{x})$ is charge density in electromagnetism (when (5.22a) relates charge and electric potential),

   (b) $\rho(\mathbf{x})$ is mass density in gravitation (when (5.22a) relates mass and gravitational potential),

   (c) $\rho(\mathbf{x})$ is ....

2. Schrödinger's equation for a non-relativistic quantum mechanical particle of mass m in a potential $V(\mathbf{x})$:

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi = \mathrm{i}\hbar\frac{\partial\psi}{\partial t}\,, \tag{5.22b}$$

   where ψ is the quantum mechanical wave function and $\hbar$ is Planck's constant divided by 2π.

3. Helmholtz's equation

$$\nabla^2 f + \omega^2 f = 0\,, \tag{5.22c}$$

   which governs the propagation of fixed frequency waves (e.g. fixed frequency sound waves). Helmholtz's equation is a 3D generalisation of the simple harmonic resonator

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} + \omega^2 f = 0\,.$$

16/99
16/00
Example. Evaluate in $\mathbb{R}^3$

$$\nabla^2\left(\frac{1}{r}\right) \quad \text{for} \quad r \neq 0\,.$$

From (5.21a) and for $r \neq 0$,

$$\begin{aligned}
\nabla^2\left(\frac{1}{r}\right) &= \frac{\partial^2}{\partial x_i \partial x_i}\left(\frac{1}{r}\right) = \frac{\partial}{\partial x_i}\left(-\frac{1}{r^2}\frac{\partial r}{\partial x_i}\right) = \frac{\partial}{\partial x_i}\left(-\frac{x_i}{r^3}\right) \\
&\qquad \text{using (1.18)} \\
&= \frac{3x_i}{r^4}\frac{\partial r}{\partial x_i} - \frac{1}{r^3}\frac{\partial x_i}{\partial x_i} = \frac{3x_i}{r^4}\frac{x_i}{r} - \frac{1}{r^3}\delta_{ii} \\
&\qquad \text{again using (1.18)} \\
&= \frac{3}{r^3} - \frac{3}{r^3} \\
&= 0\,.
\end{aligned} \tag{5.23}$$

Alternatively obtain the same result by combining (1.46a) and (5.7c), i.e.

$$\boldsymbol{\nabla}\left(\frac{1}{r}\right) = -\frac{\mathbf{x}}{r^3} \qquad \text{and} \qquad \boldsymbol{\nabla}.\left(\frac{\mathbf{x}}{r^3}\right) = 0\,.$$

5.5.3 div (curl $\mathbf{F}$) and the Vector Potential

div (curl $\mathbf{F}$*) = 0.* Suppose that $\mathbf{F}(\mathbf{x})$ is a vector function, for which the the mixed partial derivatives of $\mathbf{F}$ are equal, i.e.

$$\frac{\partial^2 F_k}{\partial x_i \partial x_j} = \frac{\partial^2 F_k}{\partial x_j \partial x_i}\,. \tag{5.24}$$

Then using (5.14) and (5.24), div (curl $\mathbf{F}$) can be evaluated as follows:

$$\begin{aligned}\boldsymbol{\nabla}\,.\,(\boldsymbol{\nabla}\times\mathbf{F}) &= \frac{\partial}{\partial x_i}\left(\varepsilon_{ijk}\,\frac{\partial F_k}{\partial x_j}\right)\\ &= \varepsilon_{ijk}\frac{\partial^2 F_k}{\partial x_i \partial x_j}\\ &= 0\,,\end{aligned} \tag{5.25}$$

i.e. the div of a curl is always zero.

Definition. Conversely suppose that $\boldsymbol{\nabla}\,.\,\mathbf{v} = 0$, i.e. that the vector field $\mathbf{v}(\mathbf{x})$ is solenoidal, then there exists a non-unique *vector potential* $\mathbf{A}(\mathbf{x})$, such that

$$\operatorname{curl}\mathbf{A} = \mathbf{v}\,. \tag{5.26}$$

A Vector Potential. We will not prove that a vector potential always exists when the vector field $\mathbf{v}$ is defined in a general domain. However, it is relatively straightforward to construct a vector potential when the vector field $\mathbf{v} = (u, v, w)$ is defined throughout $\mathbb{R}^3$. In particular, we claim that a possible vector potential is

$$\mathbf{A}(\mathbf{x}) = \left(\int_{z_0}^{z} v(x,y,\zeta)\,\mathrm{d}\zeta,\, \int_{x_0}^{x} w(\xi,y,z_0)\,\mathrm{d}\xi - \int_{z_0}^{z} u(x,y,\zeta)\,\mathrm{d}\zeta, 0\right)\,, \tag{5.27}$$

where $\mathbf{x}_0$ is a constant.

To verify our claim we note from the definition of curl (5.9b) that

$$\begin{aligned}\operatorname{curl}\mathbf{A} &= \left(\frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z}, \frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x}, \frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y}\right)\\ &= \left(u(x,y,z),\, v(x,y,z),\, w(x,y,z_0) - \int_{z_0}^{z} u_x(x,y,\zeta)\,\mathrm{d}\zeta - \int_{z_0}^{z} v_y(x,y,\zeta)\,\mathrm{d}\zeta\right)\\ &\qquad \text{using the first fundamental theorem of calculus (3.1f)}\\ &= \left(u(x,y,z),\, v(x,y,z),\, w(x,y,z_0) + \int_{z_0}^{z} w_z(x,y,\zeta)\,\mathrm{d}\zeta\right)\\ &\qquad \text{using } \boldsymbol{\nabla}\,.\,\mathbf{v} = u_x + v_y + w_z = 0\\ &= (u(x,y,z),\, v(x,y,z),\, w(x,y,z)) = \mathbf{v}\\ &\qquad \text{using the second fundamental theorem of calculus (3.1g).}\end{aligned}$$

Exercise. Write down another vector potential for $\mathbf{v}$.

5.5.4 curl (grad f) and the Vector Potential

For a scalar function f with commutative mixed partial derivatives, it follows from the definitions of grad and curl, i.e. (5.2) and (5.11c) respectively, and lemma (5.14) that

$$[\boldsymbol{\nabla}\times(\boldsymbol{\nabla}f)]_i = \varepsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right) = \varepsilon_{ijk}\frac{\partial^2 f}{\partial x_j \partial x_k} = 0\,, \tag{5.28}$$

i.e. the curl of a grad is always zero.

Non-uniqueness of the Vector Potential. Suppose that

$$\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{v}\,,$$

then for all scalar functions f

$$\boldsymbol{\nabla} \times (\mathbf{A} + \boldsymbol{\nabla} f) = \boldsymbol{\nabla} \times \mathbf{A} = \mathbf{v}\,.$$

Hence the vector potential $\mathbf{A}$ in (5.26) is not unique since for any scalar function f, $(\mathbf{A} + \boldsymbol{\nabla} f)$ is also an acceptable vector potential.

What Vector Potential to Choose? We can fix on a choice of $\mathbf{A}$ by a number of means, for instance by demanding that $\mathbf{A}$ is itself solenoidal, i.e. that

$$\boldsymbol{\nabla}.\,\mathbf{A} = 0\,.$$

For suppose that we have found a vector potential, $\mathbf{A}_1$, such that

$$\mathbf{v} = \boldsymbol{\nabla} \times \mathbf{A}_1 \quad \text{and} \quad \boldsymbol{\nabla}.\,\mathbf{A}_1 \neq 0\,.$$

Further, suppose that φ is a solution to the Poisson equation (see (5.22a) and § 8 below)

$$\nabla^2 \varphi = -\boldsymbol{\nabla}.\,\mathbf{A}_1\,, \tag{5.29}$$

with suitable boundary conditions. Then if we define

$$\mathbf{A}_2 = \mathbf{A}_1 + \boldsymbol{\nabla}\varphi\,,$$

it follows from (5.28), (5.29) and the definition of the Laplacian operator (5.21a), that

$$\boldsymbol{\nabla} \times \mathbf{A}_2 = \mathbf{v} \quad \text{and} \quad \boldsymbol{\nabla}.\,\mathbf{A}_2 = \boldsymbol{\nabla}.\,\mathbf{A}_1 + \nabla^2\varphi = 0\,.$$

5.5.5 curl (curl $\mathbf{F}$) and the Laplacian Acting on a Vector Function

For a vector function $\mathbf{F}$ with equal mixed partial derivatives, it follows from the definition of curl (5.11c) that

$$\begin{aligned}
\left[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{F})\right]_i &= \varepsilon_{ijk}\,\partial_j\,(\varepsilon_{k\ell m}\,\partial_\ell F_m) \\
&= (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})\,\frac{\partial^2 F_m}{\partial x_j \partial x_\ell} \\
&= \frac{\partial^2 F_j}{\partial x_i \partial x_j} - \frac{\partial^2 F_i}{\partial x_j \partial x_j} \\
&= \left[\boldsymbol{\nabla}(\boldsymbol{\nabla}.\,\mathbf{F}) - \nabla^2 \mathbf{F}\right]_i\,,
\end{aligned}$$

i.e. that

$$\operatorname{curl}(\operatorname{curl}\mathbf{F}) = \operatorname{grad}(\operatorname{div}\mathbf{F}) - \nabla^2\mathbf{F}\,.$$

Definition. This identity is normally written as

$$\nabla^2 \mathbf{F} = \operatorname{grad}(\operatorname{div}\mathbf{F}) - \operatorname{curl}(\operatorname{curl}\mathbf{F})\,, \tag{5.30}$$

in which form it is used as the definition of the *Laplacian acting on a vector function.*

5.6 Unlectured Worked Exercise: Yet Another Vector Potential.

Question. Show that if $\mathbf{v}$ is defined throughout $\mathbb{R}^3$, then an alternative vector potential to (5.27) is

$$\mathbf{A} = -\mathbf{x} \times \int_0^1 \lambda\, \mathbf{v}(\lambda\mathbf{x})\,\mathrm{d}\lambda\,. \tag{5.31}$$

Answer. For the vector field $\mathbf{A}$ defined by (5.31)

$$\begin{aligned}
[\operatorname{curl}\mathbf{A}]_i &= \varepsilon_{ijk}\,\frac{\partial A_k}{\partial x_j} \;=\; -\varepsilon_{ijk}\,\frac{\partial}{\partial x_j}\left(\varepsilon_{k\ell m}\;\; x_\ell \int_0^1 \lambda v_m(\lambda\mathbf{x})\,\mathrm{d}\lambda\right)\\
&= (\delta_{im}\delta_{j\ell}-\delta_{i\ell}\delta_{jm})\left(\delta_{j\ell}\int_0^1 \lambda v_m(\lambda\mathbf{x})\,\mathrm{d}\lambda + x_\ell\int_0^1 \lambda\frac{\partial}{\partial x_j}\left(v_m(\lambda\mathbf{x})\right)\,\mathrm{d}\lambda\right)\\
&= 2\int_0^1 \lambda v_i(\lambda\mathbf{x})\,\mathrm{d}\lambda + \int_0^1 \lambda x_j\,\frac{\partial}{\partial x_j}\left(v_i(\lambda\mathbf{x})\right)\,\mathrm{d}\lambda - \int_0^1 \lambda x_i\,\frac{\partial}{\partial x_j}\left(v_j(\lambda\mathbf{x})\right)\,\mathrm{d}\lambda\,.
\end{aligned}\tag{5.32}$$

Let $\boldsymbol{\xi}=\lambda\mathbf{x}$, then

$$\begin{aligned}
\frac{\partial}{\partial x_j}\left(v_j(\lambda\mathbf{x})\right) &= \frac{\partial}{\partial x_j}\left(v_j(\boldsymbol{\xi})\right) \;=\; \frac{\partial v_j}{\partial \xi_k}\,\frac{\partial \xi_k}{\partial x_j} \;=\; \lambda\delta_{jk}\,\frac{\partial v_j}{\partial \xi_k} \;=\; \lambda\frac{\partial v_j}{\partial \xi_j} = 0 \quad \text{since} \quad \boldsymbol{\nabla}.\,\mathbf{v}=0\,,\\
\lambda x_j\,\frac{\partial}{\partial x_j}\left(v_i(\lambda\mathbf{x})\right) &= \xi_j\,\frac{\partial}{\partial x_j}\left(v_i(\boldsymbol{\xi})\right) \;=\; \xi_j\,\frac{\partial v_i}{\partial \xi_k}\,\frac{\partial \xi_k}{\partial x_j} \;=\; \lambda\xi_j\,\delta_{jk}\,\frac{\partial v_i}{\partial \xi_k} \;=\; \lambda\xi_j\,\frac{\partial v_i}{\partial \xi_j}\,,\\
\lambda^2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(v_i(\boldsymbol{\xi})\right) &= \lambda^2\frac{\partial v_i}{\partial \xi_j}\,\frac{\partial \xi_j}{\partial \lambda} \;=\; \lambda\xi_j\,\frac{\partial v_i}{\partial \xi_j}\,.
\end{aligned}$$

On using the above results, (5.32) becomes

$$\begin{aligned}
[\operatorname{curl}\mathbf{A}]_i &= \int_0^1 \left(2\lambda v_i(\lambda\mathbf{x}) + \lambda^2\frac{\mathrm{d}v_i}{\mathrm{d}\lambda}(\lambda\mathbf{x})\right)\mathrm{d}\lambda\\
&= \left[\lambda^2 v_i(\lambda\mathbf{x})\right]_0^1\\
&= v_i(\mathbf{x})\,,
\end{aligned}$$

i.e. for $\mathbf{A}$ defined by (5.31)

$$\operatorname{curl}\mathbf{A} = \mathbf{v}\,. \qquad \square$$

6 The Big Integral Theorems

These results are analogous to the second fundamental theorem of calculus in that they undo a derivative by integrating. They have many important applications in electromagnetism, fluid dynamics, etc.

6.1 Divergence Theorem (Gauss' Theorem)

Divergence Theorem. Let $\mathcal{S}$ be a bounded, piecewise smooth, orientated, non-intersecting surface enclosing a volume $\mathcal{V}$ in $\mathbb{R}^3$, with a normal $\mathbf{n}$ that points outwards from $\mathcal{V}$. Let $\mathbf{u}$ be a vector field with continuous first-order partial derivatives throughout $\mathcal{V}$. Then

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}\tau = \iint_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S}\,. \tag{6.1}$$

The divergence theorem thus states that $\boldsymbol{\nabla}.\,\mathbf{u}$ integrated over a volume $\mathcal{V}$ is equal to the flux of $\mathbf{u}$ across the closed surface $\mathcal{S}$ surrounding the volume.

Outline Proof. Suppose that $\mathcal{S}$ is a surface enclosing a volume $\mathcal{V}$ such that Cartesian axes can be chosen so that any line parallel to any one of the axes meets $\mathcal{S}$ in just one or two points (e.g. a convex surface). We observe that

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}\tau = \iiint_{\mathcal{V}} \left(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}\right)\mathrm{d}\tau\,,$$

comprises of three terms; we initially concentrate on the $\iiint_{\mathcal{V}} \frac{\partial u_3}{\partial z}\mathrm{d}\tau$ term.

Let region $\mathcal{A}$ be the projection of $\mathcal{S}$ onto the xy-plane. As when we first introduced triple integrals, let the lower/upper surfaces, $\mathcal{S}_1/\mathcal{S}_2$ respectively, be parameterised by

$$\begin{aligned} \mathcal{S}_1: \quad \mathbf{x} &= (x, y, z_1(x,y)) \\ \mathcal{S}_2: \quad \mathbf{x} &= (x, y, z_2(x,y))\,. \end{aligned}$$

Then using the second fundamental theorem of calculus (3.1g)

$$\begin{aligned} \iiint_{\mathcal{V}} \frac{\partial u_3}{\partial z}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z &= \iint_{\mathcal{A}} \left[\int_{z=z_1}^{z_2} \frac{\partial u_3}{\partial z}\mathrm{d}z\right]\mathrm{d}x\mathrm{d}y \\ &= \iint_{\mathcal{A}} \left(u_3(x,y,z_2(x,y)) - u_3(x,y,z_1(x,y)\right)\mathrm{d}x\mathrm{d}y\,. \end{aligned} \tag{6.2a}$$

From (4.26a) for $\mathbf{x} = (x, y, z(x, y))$

$$\mathrm{d}\mathbf{S} = (-z_x, -z_y, 1)\,\mathrm{d}x\mathrm{d}y\,,$$

and hence (this is a bit of a dirty trick)

$$\begin{aligned} \widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S} &= \mathrm{d}x\mathrm{d}y \quad \text{on} \quad \mathcal{S}_2\,, \\ \widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S} &= -\mathrm{d}x\mathrm{d}y \quad \text{on} \quad \mathcal{S}_1\,, \end{aligned}$$

where the minus sign in the latter result is because the outward normal points in the opposite direction. Thus (6.2a) can be rewritten as

$$\iiint\limits_{\mathcal{V}} \frac{\partial u_3}{\partial z}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}_2} u_3\,\widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S} \;+\; \iint\limits_{\mathcal{S}_1} u_3\,\widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}} u_3\,\widehat{\mathbf{z}}.\,\mathrm{d}\mathbf{S}\,. \tag{6.2b}$$

Similarly

$$\iiint\limits_{\mathcal{V}} \frac{\partial u_2}{\partial y}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} u_2\,\widehat{\mathbf{y}}.\,\mathrm{d}\mathbf{S}\,, \qquad \iiint\limits_{\mathcal{V}} \frac{\partial u_1}{\partial x}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} u_1\,\widehat{\mathbf{x}}.\,\mathrm{d}\mathbf{S}\,. \tag{6.2c}$$

Adding the above results we obtain

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S}\,.$$

Remarks.

- If the volume is not of the specified type, then divide the volume into parts that are, and add the results for each sub-volume — the 'common' surface integrals cancel.
- The divergence theorem relates a triple integral to a double integral (cf. the second fundamental theorem of calculus which related a single integral to a function).

6.1.1 Examples

1. We confirm the divergence theorem (6.1) for $\mathbf{u} = \mathbf{x}$, and $\mathcal{V}$ a sphere of radius a. First, using (5.7a),

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}\tau = \iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{x}\,\mathrm{d}\tau = 3\iiint\limits_{\mathcal{V}} \mathrm{d}\tau = 3.\,\frac{4}{3}\pi a^3 = 4\pi a^3\,;$$

second, since $\mathbf{n} = \widehat{\mathbf{x}}$ on the surface of a sphere,

$$\iint\limits_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}} \mathbf{x}.\,\mathbf{n}\,\mathrm{d}S = \iint\limits_{\mathcal{S}} r\,\mathrm{d}S = a\iint\limits_{\mathcal{S}} \mathrm{d}S = 4\pi a^3.$$

2. Show that if $\boldsymbol{\nabla}.\,\mathbf{u} = 0$ in $\mathbb{R}^3$, then

$$I = \iint\limits_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S}$$

is the same for every open surface spanning a curve $\mathcal{C}$.

Answer. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two such surfaces, and let $\widehat{\mathcal{S}}$ be any other such surface that does not intersect $\mathcal{S}_1$ and $\mathcal{S}_2$. Let $\mathcal{V}_j$ be the volume enclosed by $\mathcal{S}_j + \widehat{\mathcal{S}}$. Then

$$0 = \iiint_{\mathcal{V}_j} \boldsymbol{\nabla} . \mathbf{u} = \iint_{\mathcal{S}_j} \mathbf{u} . \mathrm{d}\mathbf{S} + \iint_{\widehat{\mathcal{S}}} \mathbf{u} . \mathrm{d}\mathbf{S} \, .$$

Thus

$$\iint_{\mathcal{S}_1} \mathbf{u} . \mathrm{d}\mathbf{S} = - \iint_{\widehat{\mathcal{S}}} \mathbf{u} . \mathrm{d}\mathbf{S} = \iint_{\mathcal{S}_2} \mathbf{u} . \mathrm{d}\mathbf{S} \, .$$

17/99
17/00
3. *Unlectured Example.* Let $\mathbf{u} = (x, y, 0)$ and let $\mathcal{S}$ be the hemispherical surface $(x^2 + y^2 + z^2) = a^2$, $z \geqslant 0$. Evaluate

$$I = \iint_{\mathcal{S}} \mathbf{u} . \mathrm{d}\mathbf{S} \, .$$

Answer. Close the volume by adding $\mathcal{A}$, the disc of radius a in the $z = 0$ plane. Then

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla} . \mathbf{u} \, \mathrm{d}\tau = \iint_{\mathcal{S}} \mathbf{u} . \mathrm{d}\mathbf{S} + \iint_{\mathcal{A}} \mathbf{u} . \mathrm{d}\mathbf{S} \, .$$

On $\mathcal{A}$, the surface element $\mathrm{d}\mathbf{S} = (0, 0, -1)\mathrm{d}x\mathrm{d}y$, and hence $\mathbf{u} . \mathrm{d}\mathbf{S} = 0$. Further, since $\boldsymbol{\nabla} . \mathbf{u} = 2$,

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla} . \mathbf{u} \, \mathrm{d}\tau = 2 . \frac{2}{3} \pi a^3 = \frac{4}{3} \pi a^3 \, .$$

Hence

$$I = \frac{4}{3} \pi a^3 \, .$$

Exercise. Calculate I directly.

6.1.2 Extensions

A boundary with more than one component. Suppose the boundary $\mathcal{S}$ has more than one component, e.g. suppose that $\mathcal{V}$ lies between closed surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Then

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla} . \mathbf{u} \, \mathrm{d}\tau = \iint_{\mathcal{S}_1} \mathbf{u} . \mathrm{d}\mathbf{S} + \iint_{\mathcal{S}_2} \mathbf{u} . \mathrm{d}\mathbf{S} \, , \quad (6.3)$$

where for each surface the normal points away from $\mathcal{V}$.

This result follows from subdividing the volume into surfaces with only one component, and using the fact that the 'common' surface integrals so introduced cancel.

The two dimensional divergence theorem. Suppose that $\mathbf{u} = (u_1(x,y), u_2(x,y), 0)$, and consider a cylinder of volume $\mathcal{V}$, height h, cross section $\mathcal{A}$ with bounding curve $\mathcal{C}$, ends $\mathcal{S}_E$ and side $\mathcal{S}_S$. Then from the z-independence

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}\tau \quad = \quad h \iint_{\mathcal{A}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}S\,.$$

Also at the ends $\mathbf{u}.\,\mathbf{n} = \mathbf{u}.\,\widehat{\mathbf{z}} = 0$, and thus

$$\iint_{\mathcal{S}_E} \mathbf{u}.\,\mathrm{d}\mathbf{S} = 0\,,$$

while on the side

$$\mathrm{d}\mathbf{S} = \mathbf{n}\,\mathrm{d}s\,\mathrm{d}z \quad \text{where} \quad \mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2\,,$$

and hence

$$\iint_{\mathcal{S}_S} \mathbf{u}.\,\mathrm{d}\mathbf{S} = \int_0^h \mathrm{d}z \ \oint_{\mathcal{C}} \mathbf{u}.\,\mathbf{n}\,\mathrm{d}s\,.$$

Combining the above results and using (6.1) yields the two dimensional divergence theorem

$$\iint_{\mathcal{A}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}S = \oint_{\mathcal{C}} \mathbf{u}.\,\mathbf{n}\,\mathrm{d}s\,. \tag{6.4}$$

Unlectured Remark. Now suppose that $\mathbf{u} = (u(x), 0, 0)$ and let the area $\mathcal{A}$ be a rectangle. Then

$$\iint_{\mathcal{A}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}S = \int_c^d \mathrm{d}y \int_a^b \mathrm{d}x\,\frac{\mathrm{d}u}{\mathrm{d}x}\,,$$

and

$$\oint_{\mathcal{C}} \mathbf{u}.\,\mathbf{n}\,\mathrm{d}s = \int_c^d \mathrm{d}y\,u(b) + \int_d^c \mathrm{d}y\,u(a)\,.$$

Hence after factoring $(d-c) \neq 0$,

$$\int_a^b \frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x \quad = \quad (u(b) - u(a))\,,$$

which is just the second fundamental theorem of calculus (3.1g).

The generalisation for a scalar field. For a scalar field $f(\mathbf{x})$ with continuous first-order partial derivatives in $\mathcal{V}$

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla} f\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} f\,\mathrm{d}\mathbf{S}\,. \tag{6.5}$$

To see this let $\mathbf{a}$ be an arbitrary constant vector, and apply the divergence theorem (6.1) to $\mathbf{u} = f\,\mathbf{a}$, to obtain

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}.\,(f\mathbf{a})\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} f\,\mathbf{a}.\,\mathrm{d}\mathbf{S}\,.$$

Now

$$\boldsymbol{\nabla}.\,(f\mathbf{a}) = \frac{\partial}{\partial x_j}\,(f\,a_j) = a_j\,\frac{\partial f}{\partial x_j} = \mathbf{a}.\,(\boldsymbol{\nabla} f)\,,$$

and so because $\mathbf{a}$ is a constant

$$\mathbf{a}.\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla} f\,\mathrm{d}\tau = \mathbf{a}.\iint\limits_{\mathcal{S}} f\,\mathrm{d}\mathbf{S}\,. \tag{6.6}$$

Symbolically we have that $\mathbf{a}.(\mathbf{p}-\mathbf{q})=0$ for arbitrary $\mathbf{a}$. Choose $\mathbf{a}=(\mathbf{p}-\mathbf{q})$ to conclude that $|\mathbf{p}-\mathbf{q}|^2=0$, and thus that

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla} f\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} f\,\mathrm{d}\mathbf{S}\,.$$

Alternatively we may proceed from (6.6) by taking $\mathbf{a}=\mathbf{i}$ to conclude that $\iiint_{\mathcal{V}} \frac{\partial f}{\partial x}\,\mathrm{d}\tau = \iint_{\mathcal{S}} f\,n_1 \mathrm{d}S$, etc.

Exercise. Show by setting $\mathbf{u}=\mathbf{a}\times\mathbf{A}$, where $\mathbf{a}$ is an arbitrary vector and $\mathbf{A}$ is a vector field, that

$$\iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}\times\mathbf{A}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \mathbf{n}\times\mathbf{A}\,\mathrm{d}S\,. \tag{6.7}$$

A further 'generalisation'. Equations (6.5) and (6.7) can be written in component form as

$$\iiint\limits_{\mathcal{V}} \frac{\partial f}{\partial x_j}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} f\,n_j\,\mathrm{d}S \qquad \text{and} \qquad \iiint\limits_{\mathcal{V}} \varepsilon_{ijk}\frac{\partial A_k}{\partial x_j}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \varepsilon_{ijk} n_j A_k\,\mathrm{d}S\,.$$

This suggests the following expression for the general form of the divergence theorem

$$\iiint\limits_{\mathcal{V}} \frac{\partial\,\bullet}{\partial x_j}\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \bullet\,n_j\,\mathrm{d}S\,, \tag{6.8}$$

where $\bullet$ is any function of $\mathbf{x}$, e.g. f, u_j, $\varepsilon_{ijk}A_k$. This is in fact no more than a statement of (6.2b) and (6.2c) in suffix notation with u_j replaced with $\bullet$ in each of the formulae.

An example. For a closed surface $\mathcal{S}$ evaluate (4.31), i.e.

$$\mathbf{I} = \iint\limits_{\mathcal{S}} \mathrm{d}\mathbf{S}\,.$$

From (6.5) with $f=1$

$$\iint\limits_{\mathcal{S}} \mathrm{d}\mathbf{S} = \iiint\limits_{\mathcal{V}} \boldsymbol{\nabla} f\,\mathrm{d}\tau = 0\,,$$

as was done component by component in § 4.4.2, e.g. see (4.32).

6.1.3 Co-ordinate Independent Definition of div

Let a volume $\mathcal{V}$ be enclosed by a surface $\mathcal{S}$, and consider a limit process in which the greatest diameter of $\mathcal{V}$ tends to zero while keeping the point $\mathbf{x}_0$ inside $\mathcal{V}$. If the vector field $\mathbf{u}$ is twice differentiable at $\mathbf{x}_0$, then from Taylor's theorem with $\mathbf{x}=\mathbf{x}_0+\boldsymbol{\delta}\mathbf{x}$,

$$\iiint_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{u}(\mathbf{x})\,\mathrm{d}\tau = \iiint_{\mathcal{V}} \left(\boldsymbol{\nabla}.\,\mathbf{u}(\mathbf{x}_0) + o(|\boldsymbol{\nabla}.\,\mathbf{u}|)\right)\,\mathrm{d}\tau = \boldsymbol{\nabla}.\,\mathbf{u}(\mathbf{x}_0)\,|\mathcal{V}| + o(|\boldsymbol{\nabla}.\,\mathbf{u}|\,|\mathcal{V}|)\,,$$

where $|\mathcal{V}|$ is the volume of $\mathcal{V}$. Thus using the divergence theorem (6.1), a consistent co-ordinate free definition of div is (cf. (5.8) for a cuboid)

$$\boldsymbol{\nabla}.\,\mathbf{u} = \lim_{|\mathcal{V}|\to 0}\frac{1}{|\mathcal{V}|}\iint\limits_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S}\,, \tag{6.9}$$

where $\mathcal{S}$ is any 'nice' small closed surface enclosing a volume $\mathcal{V}$.

6.2 Green's Theorem in the Plane

6.2.1 Green's Theorem in the Plane

Let $\mathcal{C}$ be a simple, piecewise smooth, two-dimensional, closed curve (described in the anti-clockwise sense) bounding a closed region $\mathcal{A}$. Suppose $P(x,y)$ and $Q(x,y)$ have continuous first-order partial derivatives on $\mathcal{A}$. Then

$$\iint\limits_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \oint\limits_{\mathcal{C}} (P\mathrm{d}x + Q\mathrm{d}y)\,. \tag{6.10}$$

Outline Proof. Let $\mathcal{C}$ be a curve such that Cartesian axes can be chosen so that any line parallel to one of the axes meets $\mathcal{C}$ in just one or two points (e.g. suppose $\mathcal{C}$ is convex).

Divide $\mathcal{C}$ into $\mathcal{C}_1$ and $\mathcal{C}_2$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are the lower and upper curves respectively, so that a line of constant x intersects $\mathcal{C}_1$ at a smaller value of y than $\mathcal{C}_2$. Parameterise $\mathcal{C}_j$ by

$$\mathbf{x} = (x, y_j(x), 0)\,.$$

Then

$$\begin{aligned}
\iint\limits_{\mathcal{A}} \frac{\partial P}{\partial y}\mathrm{d}x\mathrm{d}y &= \int_a^b \mathrm{d}x \int_{y_1}^{y_2} \mathrm{d}y\, \frac{\partial P}{\partial y} && (6.11)\\
&= \int_a^b \mathrm{d}x\,[P(x,y_2(x)) - P(x,y_1(x))] \\
&\qquad \text{from (4.3)} \\
&= -\int_b^a \mathrm{d}x\, P(x,y_2) - \int_a^b \mathrm{d}x\, P(x,y_1) \\
&= -\oint\limits_{\mathcal{C}} P(x,y)\,\mathrm{d}x\,.
\end{aligned}$$

Similarly for $\widehat{\mathcal{C}}_j$ parameterised by $\mathbf{x} = (x_j(y), y, 0)$

$$\iint\limits_{\mathcal{A}} \frac{\partial Q}{\partial x}\,\mathrm{d}x\mathrm{d}y = \oint\limits_{\mathcal{C}} Q\mathrm{d}y\,.$$

Subtract the previous result from this to obtain

$$\iint\limits_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \oint\limits_{\mathcal{C}} (P\mathrm{d}x + Q\mathrm{d}y)\,.$$

Remark. This result can be extended to more general curves by surgery.

Unlectured Alternative Outline Proof. From the two-dimensional divergence theorem (6.5)

$$\iint\limits_{\mathcal{A}} \boldsymbol{\nabla}.\,\mathbf{u}\,\mathrm{d}S = \oint\limits_{\mathcal{C}} \mathbf{u}.\,\mathbf{n}\,\mathrm{d}s\,. \tag{6.12}$$

Suppose that the curve $\mathcal{C}$ is parameterised by arclength so that $\mathbf{x} \equiv \mathbf{x}(s)$. Then since $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}$ is tangent to $\mathcal{C}$, the outward normal $\mathbf{n}$ to the curve is given by

$$\mathbf{n} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} \times \widehat{\mathbf{z}} = \left(\frac{\mathrm{d}y}{\mathrm{d}s}, -\frac{\mathrm{d}x}{\mathrm{d}s}, 0 \right)\,.$$

Hence, if we let $\mathbf{u} = (Q, -P, 0)$ so that

$$\boldsymbol{\nabla}.\,\mathbf{u} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\,,$$

then (6.12) becomes

$$\iint_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}S = \oint_{\mathcal{C}} (Q, -P, 0) . (\mathrm{d}y, -\mathrm{d}x, 0) = \oint_{\mathcal{C}} (P \mathrm{d}x + Q \mathrm{d}y) \, .$$

6.2.2 Example.

Express

$$I = \oint_{\mathcal{C}} (x^2 y \, \mathrm{d}x + x y^2 \, \mathrm{d}y) \, ,$$

for some curve $\mathcal{C}$ bounding a surface $\mathcal{A}$, as a surface integral.

Identify

$$P = x^2 y \quad \text{and} \quad Q = x y^2 .$$

Then from Green's theorem (6.10)

$$\begin{aligned} I &= \iint_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x \mathrm{d}y \\ &= \iint_{\mathcal{A}} (y^2 - x^2) \, \mathrm{d}x \mathrm{d}y \, . \end{aligned}$$

Remark. The equivalence of these two expressions for I when $\mathcal{C}$ is one particular curve is a question on *Examples Sheet 2.*

6.2.3 Exact Differentials

Suppose $P\mathrm{d}x + Q\mathrm{d}y$ is exact. From (3.24) we have that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \, .$$

Green's theorem (6.10) then gives

$$\oint_{\mathcal{C}} (P \mathrm{d}x + Q \mathrm{d}y) = \iint_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x \mathrm{d}y = 0 \, .$$

Remark. This is a rederivation of result (3.31) that the line integral of an exact differential around a closed path in a simply connected region is zero.

6.3 Stokes' Theorem

This is a generalisation of Green's theorem to a three dimensional surface.

6.3.1 Stokes' Theorem.

Let $\mathcal{S}$ be a piecewise smooth, open, orientated, non-intersecting surface bounded by a simple, piecewise smooth, closed curve $\mathcal{C}$. Let $\mathbf{u}(\mathbf{x})$ be a vector field with continuous first-order partial derivatives on $\mathcal{S}$. Then

$$\iint_{\mathcal{S}} \boldsymbol{\nabla} \times \mathbf{u} . \, \mathrm{d}\mathbf{S} = \oint_{\mathcal{C}} \mathbf{u} . \, \mathrm{d}\mathbf{x} \,, \tag{6.13}$$

where the line integral is taken in the positive direction of $\mathcal{C}$ (right hand rule).

Stokes' theorem thus states that the flux of $\boldsymbol{\nabla} \times \mathbf{u}$ across an open surface $\mathcal{S}$ is equal to the circulation of $\mathbf{u}$ round the bounding curve $\mathcal{C}$.

Remark. If $\mathcal{S}$ is in the xy-plane, then (6.13) becomes

$$\iint_{\mathcal{S}} \left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right) \mathrm{d}x\mathrm{d}y = \oint_{\mathcal{C}} (u_1 \mathrm{d}x + u_2 \mathrm{d}y) \,,$$

18/99 i.e. Green's theorem (6.10) after writing $\mathbf{u} = (P, Q, 0)$.

18/00

Heuristic Argument. Divide up $\mathcal{S}$ by sectionally smooth curves into N subregions $\mathcal{S}_1, \ldots, \mathcal{S}_N$. For small enough subregions, each $\mathcal{S}_j$ can be approximated by a plane surface $\mathcal{A}_j$, bounded by curves $\mathcal{C}_1, \ldots, \mathcal{C}_N$. Apply Green's theorem in vector form to each individual subregion $\mathcal{A}_j$, then on summing over the subregions

$$\sum_j \iint_{\mathcal{A}_j} (\boldsymbol{\nabla} \times \mathbf{u}) . \, \mathrm{d}\mathbf{S} = \sum_j \oint_{\mathcal{C}_j} \mathbf{u} . \, \mathrm{d}\mathbf{x} \,.$$

Further on letting $N \to \infty$

$$\sum_j \iint_{\mathcal{A}_j} (\boldsymbol{\nabla} \times \mathbf{u}) . \, \mathrm{d}\mathbf{S} \quad \to \quad \iint_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{u}) . \, \mathrm{d}\mathbf{S} \,,$$

while

$$\sum_j \oint_{\mathcal{C}_j} \mathbf{u} . \, \mathrm{d}\mathbf{x} \quad \to \quad \oint_{\mathcal{C}} \mathbf{u} . \, \mathrm{d}\mathbf{x} \,,$$

since the integrals over curves that are 'common' to the plane approximations to two adjacent subregions cancel in the limit. Stokes' theorem follows.

Outline Proof. Let the open surface $\mathcal{S}$ be parameterised by p, q: $\mathbf{x} \equiv \mathbf{x}(p, q)$. Then the local vector elementary area is given by, see (4.25b),

$$\mathrm{d}\mathbf{S} = \left(\frac{\partial \mathbf{x}}{\partial p} \times \frac{\partial \mathbf{x}}{\partial q} \right) \mathrm{d}p \, \mathrm{d}q \,,$$

where the parameters, p and q, are assumed to be ordered so that $\mathrm{d}\mathbf{S}$ is in the chosen positive normal direction.[8]

[8] If this is not the case either interchange p and q, or introduce minus signs appropriately.

Further, from using the definition of curl, i.e. (5.11b), it follows that

$$
\begin{aligned}
\boldsymbol{\nabla}\times\mathbf{u}\,.\,\mathrm{d}\mathbf{S} &= \left(\varepsilon_{ijk}\frac{\partial u_k}{\partial x_j}\right)\left(\varepsilon_{i\ell m}\frac{\partial x_\ell}{\partial p}\,\frac{\partial x_m}{\partial q}\right)\mathrm{d}p\,\mathrm{d}q\\
&= (\delta_{j\ell}\delta_{km}-\delta_{jm}\delta_{k\ell})\,\frac{\partial u_k}{\partial x_j}\,\frac{\partial x_\ell}{\partial p}\,\frac{\partial x_m}{\partial q}\;\mathrm{d}p\,\mathrm{d}q\\
&= \left(\frac{\partial u_k}{\partial x_j}\frac{\partial x_j}{\partial p}\frac{\partial x_k}{\partial q}\;-\;\frac{\partial u_k}{\partial x_j}\frac{\partial x_j}{\partial q}\frac{\partial x_k}{\partial p}\right)\mathrm{d}p\,\mathrm{d}q\\
&= \left(\frac{\partial u_k}{\partial p}\frac{\partial x_k}{\partial q}\;-\;\frac{\partial u_k}{\partial q}\frac{\partial x_k}{\partial p}\right)\mathrm{d}p\,\mathrm{d}q\\
&\qquad \text{using the chain rule (1.36a)}\\
&= \left[\frac{\partial}{\partial p}\left(u_k\frac{\partial x_k}{\partial q}\right)-\frac{\partial}{\partial q}\left(u_k\frac{\partial x_k}{\partial p}\right)\right]\mathrm{d}p\,\mathrm{d}q\\
&\qquad \text{assuming that } \frac{\partial^2 x_k}{\partial p\partial q}=\frac{\partial^2 x_k}{\partial q\partial p}.
\end{aligned}
$$

Hence

$$\iint\limits_{\mathcal{S}}(\boldsymbol{\nabla}\times\mathbf{u})\,.\,\mathrm{d}\mathbf{S}=\iint\limits_{\mathcal{A}}\left[\frac{\partial}{\partial p}\left(u_k\frac{\partial x_k}{\partial q}\right)-\frac{\partial}{\partial q}\left(u_k\frac{\partial x_k}{\partial p}\right)\right]\mathrm{d}p\,\mathrm{d}q\,,$$

where $\mathcal{A}$ is the region in the pq-plane over which $\mathcal{S}$ is parameterised.

Next apply Green's theorem to the right-hand-side by identifying p and q with x and y respectively in (6.10), to obtain

$$\iint\limits_{\mathcal{S}}(\boldsymbol{\nabla}\times\mathbf{u})\,.\,\mathrm{d}\mathbf{S}\;=\;\oint\limits_{\widehat{\mathcal{C}}}\left(u_k\frac{\partial x_k}{\partial p}\mathrm{d}p+u_k\frac{\partial x_k}{\partial q}\mathrm{d}q\right)\;=\;\oint\limits_{\widehat{\mathcal{C}}}u_k\left(\frac{\partial x_k}{\partial p}\mathrm{d}p+\frac{\partial x_k}{\partial q}\mathrm{d}q\right)\,,$$

where $\widehat{\mathcal{C}}$ is the bounding curve for $\mathcal{A}$ and the integral round $\widehat{\mathcal{C}}$ is in an anti-clockwise direction (see that statement of Green's theorem). Moreover, since

$$\mathrm{d}x_k=\frac{\partial x_k}{\partial p}\mathrm{d}p+\frac{\partial x_k}{\partial q}\mathrm{d}q=\,,$$

the right-hand-side can be transformed to a closed line integral around $\mathcal{C}$. Thus

$$\iint\limits_{\mathcal{S}}(\boldsymbol{\nabla}\times\mathbf{u})\,.\,\mathrm{d}\mathbf{S}=\oint\limits_{\mathcal{C}}u_k\mathrm{d}x_k=\oint\limits_{\mathcal{C}}\mathbf{u}\,.\,\mathrm{d}\mathbf{x}\,,$$

since we claim that as $\widehat{\mathcal{C}}$ is traced out in the pq-plane in an anti-clockwise sense, the curve $\mathcal{C}$ is traced out in its oriented direction.[9]

[9] For those of you who are interested, an unlectured discussion of this point is given in § 6.3.2. The rest of you can adopt the position of most books (and possibly even most lecturers) and assume that the statement is 'self-evident'.

6.3.2 For Our Choice of Positive Normal Anti-Clockwise in $\widehat{\mathcal{C}} \Rightarrow$ Oriented in $\mathcal{C}$

Suppose that $\mathbf{p} = \mathbf{P}(s) \equiv (P(s), Q(s))$ parameterises $\widehat{\mathcal{C}}$ in an anti-clockwise sense using arc length. Then from (2.8) the unit tangent, $\widehat{\mathbf{u}}$, to $\widehat{\mathcal{C}}$ is

$$\widehat{\mathbf{u}} = \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}s} = \left(\frac{\mathrm{d}P}{\mathrm{d}s}, \frac{\mathrm{d}Q}{\mathrm{d}s}\right),$$

while the equation of a normal pointing *into* $\mathcal{A}$ from a point $\mathbf{P}(s)$, and parameterised by t, is given by

$$\mathbf{p} = \mathbf{P}(s) + t\widehat{\mathbf{a}}, \quad \text{where} \quad \widehat{\mathbf{a}} = \left(-\frac{\mathrm{d}Q}{\mathrm{d}s}, \frac{\mathrm{d}P}{\mathrm{d}s}\right).$$

Hence a [non-unit] tangent, $\mathbf{u}$, to $\mathcal{C}$, corresponding to $\widehat{\mathcal{C}}$ being traced out in an anti-clockwise sense, is

$$\mathbf{u} = \frac{\mathrm{d}}{\mathrm{d}s}\big(\mathbf{x}(\mathbf{P}(s))\big) = \frac{\partial \mathbf{x}}{\partial p}\frac{\mathrm{d}P}{\mathrm{d}s} + \frac{\partial \mathbf{x}}{\partial q}\frac{\mathrm{d}Q}{\mathrm{d}s},$$

while a [non-unit] normal, $\mathbf{a}$, to $\mathcal{C}$ that is also tangent to the the surface $\mathcal{S}$ is

$$\mathbf{a} = \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{x}(\mathbf{P}(s) + t\widehat{\mathbf{a}})\big) = -\frac{\partial \mathbf{x}}{\partial p}\frac{\mathrm{d}Q}{\mathrm{d}s} + \frac{\partial \mathbf{x}}{\partial q}\frac{\mathrm{d}P}{\mathrm{d}s}.$$

It follows that a [non-unit] normal to $\mathcal{S}$, arranged so that $\mathcal{C}$ is traced out in its oriented direction as $\widehat{\mathcal{C}}$ is traced out in an anti-clockwise sense, is

$$\mathbf{u} \times \mathbf{a} = \left(\frac{\partial \mathbf{x}}{\partial p} \times \frac{\partial \mathbf{x}}{\partial q}\right)\left(\left(\frac{\mathrm{d}P}{\mathrm{d}s}\right)^2 + \left(\frac{\mathrm{d}Q}{\mathrm{d}s}\right)^2\right).$$

We note that this is in the same direction as $\mathbf{n} = \mathrm{d}\mathbf{S}/\mathrm{d}S$. □.

6.3.3 Examples

1. For $\mathbf{u} = (x^2, z^2, -y^2)$ evaluate

   $$I = \iint_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{u}) . \,\mathrm{d}\mathbf{S},$$

   where $\mathcal{S}$ is the surface $z = 4 - x^2 - y^2$, $z \geqslant 0$.

   Since the bounding curve $\mathcal{C}$ is the circle $x^2 + y^2 = 4$ in the plane $z = 0$, it follows from Stokes' theorem (6.13) that

   $$\begin{aligned}\iint_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{u}) . \,\mathrm{d}\mathbf{S} &= \oint_{\mathcal{C}} \mathbf{u} . \,\mathrm{d}\mathbf{x} \\ &= \oint_{\mathcal{C}} (x^2 \mathrm{d}x + z^2 \mathrm{d}y) \\ &\qquad \text{but } z = 0 \text{ on } \mathcal{C} \\ &= \left[\tfrac{1}{3}x^3\right]_2^2 \;=\; 0\,.\end{aligned}$$

   Exercise. Evaluate directly without using Stokes' theorem.

2. *Unlectured Example.* Evaluate $I = \iint_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{u}) . \,\mathrm{d}\mathbf{S}$ where $\mathcal{S}$ is a closed surface.

(a) First divide $\mathcal{S}$ into two parts $\mathcal{S}_1$ and $\mathcal{S}_2$ separated by a curve $\mathcal{C}$ such that

$\mathcal{S}_1$ has oriented boundary curve $\mathcal{C}$
$\mathcal{S}_2$ has oriented boundary curve -$\mathcal{C}$

Then from Stokes' theorem (6.13)

$$\begin{aligned}\iint_{\mathcal{S}} (\boldsymbol{\nabla}\times\mathbf{u}).\,\mathrm{d}\mathbf{S} &= \iint_{\mathcal{S}_1} (\boldsymbol{\nabla}\times\mathbf{u}).\,\mathrm{d}\mathbf{S} + \iint_{\mathcal{S}_2} (\boldsymbol{\nabla}\times\mathbf{u}).\,\mathrm{d}\mathbf{S} \\ &= \oint_{\mathcal{C}} \mathbf{u}.\,\mathrm{d}\mathbf{x} + \oint_{-\mathcal{C}} \mathbf{u}.\,\mathrm{d}\mathbf{x} \\ &= 0\,.\end{aligned}$$

(b) Alternatively from the divergence theorem (6.1)

$$I = \iiint_{\mathcal{V}} \boldsymbol{\nabla}.\,(\boldsymbol{\nabla}\times\mathbf{u})\,\mathrm{d}\tau = 0\,,$$

because div(curl) is always zero from (5.25).

6.3.4 Extensions

1. Suppose $\mathcal{S}$ is bounded by two (or more) curves, say $\mathcal{C}_1$ and $\mathcal{C}_2$. Then

$$\iint_{\mathcal{S}} (\boldsymbol{\nabla}\times\mathbf{u}).\,\mathrm{d}\mathbf{S} = \oint_{\mathcal{C}_1} \mathbf{u}.\,\mathrm{d}\mathbf{x} \;+\; \oint_{\mathcal{C}_2} \mathbf{u}.\,\mathrm{d}\mathbf{x}\,,$$

where the oriented direction of each $\mathcal{C}_j$ is fixed by the right-hand rule (see page 34).

Outline Proof. By surgery/cross-cuts and cancellation.

2. For a scalar field f with continuous first-order partial derivatives in $\mathcal{S}$

$$\iint_{\mathcal{S}} \boldsymbol{\nabla} f\times\mathrm{d}\mathbf{S} = -\oint_{\mathcal{C}} f\,\mathrm{d}\mathbf{x}\,. \qquad (6.14)$$

To see this let $\mathbf{u} = \mathbf{a}f$ where $\mathbf{a}$ is an arbitrary vector. Then from Stokes' theorem (6.13)

$$\iint_{\mathcal{S}} \boldsymbol{\nabla}\times(\mathbf{a}f).\,\mathrm{d}\mathbf{S} \;=\; \oint_{\mathcal{C}} f\mathbf{a}.\,\mathrm{d}\mathbf{x}\,.$$

But

$$\begin{aligned}\boldsymbol{\nabla}\times(\mathbf{a}f).\,\mathrm{d}\mathbf{S} &= \mathbf{n}.\,\boldsymbol{\nabla}\times(\mathbf{a}f)\,\mathrm{d}S \\ &= n_i\,\varepsilon_{ijk}\,\frac{\partial(a_k f)}{\partial x_j}\,\mathrm{d}S \\ &= -a_k\,\varepsilon_{kji}\,\frac{\partial f}{\partial x_j}\,n_i\,\mathrm{d}S \\ &= -\mathbf{a}.\,(\boldsymbol{\nabla} f\times\mathbf{n})\,\mathrm{d}S\,,\end{aligned}$$

and hence

$$-\mathbf{a}.\iint\limits_{\mathcal{S}} \boldsymbol{\nabla} f \times \mathrm{d}\mathbf{S} = \mathbf{a}.\oint\limits_{\mathcal{C}} f\,\mathrm{d}\mathbf{x}.$$

It follows that since $\mathbf{a}$ is arbitrary (cf. the argument below equation (6.6))

$$\iint\limits_{\mathcal{S}} \boldsymbol{\nabla} f \times \mathrm{d}\mathbf{S} = -\oint\limits_{\mathcal{C}} f\,\mathrm{d}\mathbf{x}.$$

6.3.5 Co-ordinate Independent Definition of curl

Let an open smooth surface $\mathcal{S}$ be bounded by a curve $\mathcal{C}$. Consider a limit process in which the point $\mathbf{x}_0$ remains on $\mathcal{S}$, the greatest diameter of $\mathcal{S}$ tends to zero, and the normals at all points on the surface tend to a specific direction (i.e. the value of $\mathbf{n}$ at $\mathbf{x}_0$).

If vector field $\mathbf{u}$ is twice differentiable at $\mathbf{x}_0$, then from Taylor's theorem with $\mathbf{x} = \mathbf{x}_0 + \boldsymbol{\delta}\mathbf{x}$,

$$\iint\limits_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{u}(\mathbf{x})).\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}} \left(\boldsymbol{\nabla} \times \mathbf{u}(\mathbf{x}_0) + o(|\boldsymbol{\nabla} \times \mathbf{u}|)\right).\,\mathrm{d}\mathbf{S} = \boldsymbol{\nabla} \times \mathbf{u}(\mathbf{x}_0).\mathbf{n}\,|\mathcal{S}| + o(|\boldsymbol{\nabla} \times \mathbf{u}|\,|\mathcal{S}|)\,,$$

where $|\mathcal{S}|$ is the area of $\mathcal{S}$. Thus using Stokes' theorem (6.13), we conclude that a consistent co-ordinate free definition of curl has the form (cf. (5.19) for a rectangle)

$$\mathbf{n}.\,(\boldsymbol{\nabla} \times \mathbf{u}) = \lim_{\mathcal{S}\to 0} \frac{1}{|\mathcal{S}|} \oint_C \mathbf{u}.\,\mathrm{d}\mathbf{x}\,, \tag{6.15}$$

where $\mathcal{S}$ is any 'nice' small open surface with a bounding curve $\mathcal{C}$. All components of $\boldsymbol{\nabla} \times \mathbf{u}$ can be recovered by considering limiting surfaces with normals in different directions.

6.3.6 Irrotational Fields

Suppose that $\mathbf{F}$ has continuous first-order partial derivatives in a simply connected region $\mathcal{V}$. Then the following are equivalent.

(a) $\boldsymbol{\nabla} \times \mathbf{F} = 0$ throughout $\mathcal{V}$, i.e. $\mathbf{F}$ is irrotational throughout $\mathcal{V}$.

(b) $\oint_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = 0$ for all closed curves $\mathcal{C}$ in $\mathcal{V}$.

(c) $\mathbf{F}$ is conservative, i.e. $\int_{\mathbf{x}_a}^{\mathbf{x}_b} \mathbf{F}.\,\mathrm{d}\mathbf{x}$ is independent of path.

(d) There exists a single valued function f called the *scalar potential*, such that $\mathbf{F} = \boldsymbol{\nabla} f$.

(e) $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is an exact differential.

Outline Proof.

(a) $\Rightarrow$ (b). If $\boldsymbol{\nabla} \times \mathbf{F} = 0$, then by Stokes' theorem,

$$\oint\limits_{\mathcal{C}} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \iint\limits_{\mathcal{S}} (\boldsymbol{\nabla} \times \mathbf{F}).\,\mathrm{d}\mathbf{S} = 0\,.$$

(b) $\Rightarrow$ (c). Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be any two paths from $\mathbf{x}_a$ to $\mathbf{x}_b$. Then

$$\int\limits_{\mathcal{C}_1} \mathbf{F}.\,\mathrm{d}\mathbf{x} - \int\limits_{\mathcal{C}_2} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \oint\limits_{\mathcal{C}_1-\mathcal{C}_2} \mathbf{F}.\,\mathrm{d}\mathbf{x} = 0\,.$$

Hence

$$\int\limits_{\mathcal{C}_1} \mathbf{F}.\,\mathrm{d}\mathbf{x} = \int\limits_{\mathcal{C}_2} \mathbf{F}.\,\mathrm{d}\mathbf{x}\,.$$

(c) $\Rightarrow$ (d). This was proved in § 3.4.1. In particular it was shown that if $\mathbf{F}$ is conservative, then for the scalar potential defined by (see (3.32a))

$$f(\mathbf{x}) = \int_{\mathbf{a}}^{\mathbf{x}} \mathbf{F}.\,\mathrm{d}\mathbf{x}\,,$$

it follows that (see (3.32b))

$$\mathbf{F} = \boldsymbol{\nabla} f\,.$$

(d) $\Rightarrow$ (e). This was also proved in § 3.4.1. In particular, if $\mathbf{F} = \boldsymbol{\nabla} f$ then (see (3.32c))

$$\mathbf{F}.\,\mathrm{d}\mathbf{x} = \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x} = \mathrm{d}f\,,$$

i.e. $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is an exact differential.

(e) $\Rightarrow$ (a). We prove this indirectly by showing that '(e) $\Rightarrow$ (d)', and then that '(d) $\Rightarrow$ (a)'.

'(e) $\Rightarrow$ (d)'. If $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is an exact differential, say equal to $\mathrm{d}f$, then from the relationship (3.18) between a differential and a gradient

$$\mathbf{F}.\,\mathrm{d}\mathbf{x} = \mathrm{d}f = \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x}\,.$$

Next we observe that this result can only hold for all $\mathrm{d}\mathbf{x}$ if

$$\mathbf{F} = \boldsymbol{\nabla} f\,.$$

'(d) $\Rightarrow$ (a)'. If $\mathbf{F} = \boldsymbol{\nabla} f$, then from (5.28)

$$\boldsymbol{\nabla} \times \mathbf{F} = \boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) = 0\,.$$

Remarks.

All for one and one for all. We conclude that one result implies all the others.

If $\mathbf{F}$ is irrotational then $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ is exact. The result '(a) $\Rightarrow$ (e)' proves, as promised earlier, that condition (3.26b), i.e. $\boldsymbol{\nabla} \times \mathbf{F} = 0$, is a *sufficient* condition for $\mathbf{F}.\,\mathrm{d}\mathbf{x}$ to be an exact differential in 3D (we proved that it was *necessary* in § 3.3.3). Further, in the special case when $\mathbf{F}$ is a two-dimensional vector field, say $\mathbf{F} = (P, Q, 0)$, we recover equation (3.24), i.e.

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}\,,$$

as a sufficient condition for $P\,\mathrm{d}x + Q\,\mathrm{d}y$ to be exact.

Poincaré's Lemma for irrotational fields. The result '(a) $\Rightarrow$ (d)' proves Poincaré's lemma (see § 5.3.2), i.e. if $\boldsymbol{\nabla} \times \mathbf{F} = 0$ in a simply connected region then there exists a scalar potential f such that

$$\mathbf{F} = \boldsymbol{\nabla} f\,. \tag{6.16}$$

19/00

7 Orthogonal Curvilinear Co-ordinates

So far we have concentrated on Cartesian co-ordinates. What do grad, div and curl look like in cylindrical or spherical polar co-ordinates, etc?

The key difficulty arises because the basis vectors for cylindrical and spherical polar co-ordinates are functions of position (hence the use of the word *curvilinear*).

7.1 Basis Vectors

Denote the orthogonal curvilinear co-ordinates by q_1, q_2, q_3.

	Cartesian Co-ordinates	Cylindrical Polar Co-ordinates	Spherical Polar Co-ordinates
q_1	x	$\rho = (x^2+y^2)^{1/2}$	$r = (x^2+y^2+z^2)^{1/2}$
q_2	y	$\phi = \tan^{-1}(y/x)$	$\theta = \tan^{-1}((x^2+y^2)^{1/2}/z)$
q_3	z	z	$\phi = \tan^{-1}(y/x)$

The position vector $\mathbf{x}$ is a function of $\mathbf{q} = (q_1, q_2, q_3)$, i.e. $\mathbf{x} \equiv \mathbf{x}(\mathbf{q})$. Define

$$h_j(\mathbf{q}) = \left|\frac{\partial \mathbf{x}(\mathbf{q})}{\partial q_j}\right| , \tag{7.1a}$$

and let $\mathbf{e}_j(\mathbf{q})$ be the unit vector in the direction of $\frac{\partial \mathbf{x}}{\partial q_j}$; then (***no*** *s.c.*)

$$\mathbf{e}_\alpha(\mathbf{q}) = \frac{1}{h_\alpha(\mathbf{q})}\frac{\partial \mathbf{x}(\mathbf{q})}{\partial q_\alpha} \qquad \text{or equivalently} \qquad \frac{\partial \mathbf{x}(\mathbf{q})}{\partial q_\alpha} = h_\alpha(\mathbf{q})\mathbf{e}_\alpha(\mathbf{q})\,. \tag{7.1b}$$

For an *orthogonal* curvilinear co-ordinate system we demand that for all $\mathbf{q}$

$$\mathbf{e}_i(\mathbf{q}).\,\mathbf{e}_j(\mathbf{q}) = \delta_{ij}\,. \tag{7.1c}$$

Also by convention we require the system to be right-handed, i.e. (*s.c.*)

$$\varepsilon_{ijk}\,\mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k\,, \quad \text{e.g.} \quad \mathbf{e}_1 = \mathbf{e}_2 \times \mathbf{e}_3\,. \tag{7.1d}$$

19/99

Unlectured Remark. With the right-handed convention (7.1d), the Jacobian [determinant] of the co-ordinate transformation $\mathbf{x} \equiv \mathbf{x}(\mathbf{q})$ is positive (as might be expected), since using (7.1b) and (7.1d) it is given by (*s.c.*)

$$\begin{aligned} J = \det\left(\frac{\partial x_i}{\partial q_j}\right) &= \varepsilon_{ijk}\frac{\partial x_i}{\partial q_1}\frac{\partial x_j}{\partial q_2}\frac{\partial x_k}{\partial q_3} \\ &= \varepsilon_{ijk}\,h_1\,[\mathbf{e}_1]_i\,h_2\,[\mathbf{e}_2]_j\,h_3\,[\mathbf{e}_3]_k \\ &= h_1h_2h_3\,\mathbf{e}_1.\,\mathbf{e}_2\times\mathbf{e}_3 \\ &= h_1h_2h_3\,. \end{aligned} \tag{7.1e}$$

7.1.1 Cartesian Co-ordinates

In this case

$$\mathbf{x} = (x, y, z)\,.$$

Hence

$$\frac{\partial \mathbf{x}}{\partial q_1} = \frac{\partial \mathbf{x}}{\partial x} = \mathbf{i}\,;$$

thus

$$h_1 = \left|\frac{\partial \mathbf{x}}{\partial q_1}\right| = 1\,, \quad \mathbf{e}_1 = \mathbf{i}\,.$$

Similarly

$$h_2 = \left|\frac{\partial \mathbf{x}}{\partial q_2}\right| = 1\,, \quad \mathbf{e}_2 = \mathbf{j}\,, \qquad \text{and} \qquad h_3 = \left|\frac{\partial \mathbf{x}}{\partial q_3}\right| = 1\,, \quad \mathbf{e}_3 = \mathbf{k}\,.$$

Remark.

- $\mathbf{e}_i.\mathbf{e}_j = \delta_{ij}$ and $\varepsilon_{ijk}\,\mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k$, i.e. Cartesian co-ordinates are a right-handed orthogonal co-ordinate system!

7.1.2 Cylindrical Polar Co-ordinates

In this case

$$\mathbf{x} = (\rho\cos\phi,\ \rho\sin\phi, z)\,.$$

Hence

$$\frac{\partial \mathbf{x}}{\partial q_1} = \frac{\partial \mathbf{x}}{\partial \rho} = (\cos\phi, \sin\phi,\ 0)\,,$$
$$\frac{\partial \mathbf{x}}{\partial q_2} = \frac{\partial \mathbf{x}}{\partial \phi} = (-\rho\sin\phi,\ \rho\cos\phi,\ 0)\,,$$
$$\frac{\partial \mathbf{x}}{\partial q_3} = \frac{\partial \mathbf{x}}{\partial z} = (0, 0, 1)\,.$$

It follows that

$$h_1 = \left|\frac{\partial \mathbf{x}}{\partial q_1}\right| = 1\,, \qquad \mathbf{e}_1 = \mathbf{e}_\rho = (\cos\phi, \sin\phi, 0)\,, \tag{7.2a}$$
$$h_2 = \left|\frac{\partial \mathbf{x}}{\partial q_2}\right| = \rho\,, \qquad \mathbf{e}_2 = \mathbf{e}_\phi = (-\sin\phi, \cos\phi, 0)\,, \tag{7.2b}$$
$$h_3 = \left|\frac{\partial \mathbf{x}}{\partial q_3}\right| = 1\,, \qquad \mathbf{e}_3 = \mathbf{e}_z = (0, 0, 1)\,. \tag{7.2c}$$

Remarks.

- $\mathbf{e}_i.\,\mathbf{e}_j = \delta_{ij}$ and $\varepsilon_{ijk}\,\mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k$, i.e. cylindrical polar co-ordinates are a right-handed orthogonal curvilinear co-ordinate system.
- $\mathbf{e}_\rho$ and $\mathbf{e}_\phi$ are functions of position.

7.1.3 Spherical Polar Co-ordinates

In this case

$$\mathbf{x} = (r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)\,.$$

Hence

$$\frac{\partial \mathbf{x}}{\partial q_1} = \frac{\partial \mathbf{x}}{\partial r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,,$$

$$\frac{\partial \mathbf{x}}{\partial q_2} = \frac{\partial \mathbf{x}}{\partial \theta} = (r\cos\theta\cos\phi, r\cos\theta\sin\phi, -r\sin\theta)\,,$$

$$\frac{\partial \mathbf{x}}{\partial q_3} = \frac{\partial \mathbf{x}}{\partial \phi} = (-r\sin\theta\sin\phi, r\sin\theta\cos\phi, 0)\,.$$

It follows that

$$h_1 = 1\,, \qquad \mathbf{e}_1 = \mathbf{e}_r = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,, \tag{7.3a}$$

$$h_2 = r\,, \qquad \mathbf{e}_2 = \mathbf{e}_\theta = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)\,, \tag{7.3b}$$

$$h_3 = r\sin\theta\,, \qquad \mathbf{e}_3 = \mathbf{e}_\phi = (-\sin\phi, \cos\phi, 0)\,. \tag{7.3c}$$

Remarks.

- $\mathbf{e}_i . \mathbf{e}_j = \delta_{ij}$ and $\varepsilon_{ijk}\,\mathbf{e}_i = \mathbf{e}_j \times \mathbf{e}_k$, i.e. spherical polar co-ordinates are a right-handed orthogonal curvilinear co-ordinate system.
- $\mathbf{e}_r$, $\mathbf{e}_\theta$ and $\mathbf{e}_\phi$ are functions of position.

7.2 Incremental Change in Position or Length.

Consider an incremental change in position. Then from (7.1b)

$$\mathrm{d}\mathbf{x} = \sum_j \frac{\partial \mathbf{x}}{\partial q_j}\,\mathrm{d}q_j \;=\; \sum_j h_j\ \mathbf{e}_j\,\mathrm{d}q_j\,. \tag{7.4}$$

Hence from the orthogonality requirement (7.1c), an incremental change in [squared] length is given by

$$\mathrm{d}s^2 = \mathrm{d}\mathbf{x}^2 = \sum_j h_j\ \mathbf{e}_j\,\mathrm{d}q_j \,.\sum_k h_k\ \mathbf{e}_k\,\mathrm{d}q_k = \sum_{j,k} h_j \mathrm{d}q_j\, h_k\,\mathrm{d}q_k\,\delta_{jk} = \sum_j h_j^2\,\mathrm{d}q_j^2\,.$$

Example. For spherical polar co-ordinates from (7.3a), (7.3b) and (7.3c)

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\phi^2\,.$$

7.3 Gradient

We need a co-ordinate independent definition of the gradient for the case when the basis vectors are functions of position.

We recall from the definition of a differential, that for a *fixed* basis system in (3.18) we concluded that

$$\mathrm{d}f = \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x}\,. \tag{7.5a}$$

Definition. For curvilinear orthogonal co-ordinates (for which the basis vectors are in general a function of position), we define $\boldsymbol{\nabla} f$ to be the vector such that for all $\mathrm{d}\mathbf{x}$

$$\mathrm{d}f = \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{x}\,. \tag{7.5b}$$

This is a consistent co-ordinate free definition of grad.

In order to determine the components of $\boldsymbol{\nabla} f$ write

$$\boldsymbol{\nabla} f = \sum_i \mathbf{e}_i \alpha_i \,, \tag{7.6a}$$

then from (7.4) and (7.1c)

$$\mathrm{d}f = \sum_i \mathbf{e}_i \alpha_i . \sum_j h_j \mathbf{e}_j \mathrm{d}q_j = \sum_i \alpha_i h_i \mathrm{d}q_i \,. \tag{7.6b}$$

But from the definition of a differential (3.14a) for $f \equiv f(q_1, q_2, q_3)$,

$$\mathrm{d}f = \sum_i \frac{\partial f}{\partial q_i} \mathrm{d}q_i \,, \tag{7.6c}$$

and hence, since (7.6b) and (7.6c) must hold for all $\mathrm{d}q_i$,

$$\alpha_i = \frac{1}{h_i}\,\frac{\partial f}{\partial q_i} \qquad (\boldsymbol{no}\ \textit{s.c.})\,. \tag{7.6d}$$

So from (7.6a)

$$\boldsymbol{\nabla} f = \sum_i \frac{\mathbf{e}_i}{h_i}\,\frac{\partial f}{\partial q_i} = \left(\frac{1}{h_1}\frac{\partial f}{\partial q_1}, \frac{1}{h_2}\frac{\partial f}{\partial q_2}, \frac{1}{h_3}\frac{\partial f}{\partial q_3}\right) . \tag{7.7}$$

Remark. Each term has dimensions 'f/length'.

7.3.1 Examples

Cylindrical Polar Co-ordinates. In cylindrical polar co-ordinates, the gradient is given from (7.2a), (7.2b) and (7.2c) to be (cf. *Examples Sheet 1* for the special case of two-dimensional plane polar co-ordinates)

$$\boldsymbol{\nabla} = \mathbf{e}_\rho \frac{\partial}{\partial \rho} \;+\; \mathbf{e}_\phi \frac{1}{\rho}\frac{\partial}{\partial \phi} \;+\; \mathbf{e}_z \frac{\partial}{\partial z} \,. \tag{7.8a}$$

Spherical Polar Co-ordinates. In spherical polar co-ordinates the gradient is given from (7.3a), (7.3b) and (7.3c) to be

$$\boldsymbol{\nabla} = \mathbf{e}_r \frac{\partial}{\partial r} \;+\; \mathbf{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} \;+\; \mathbf{e}_\phi \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi} \,. \tag{7.8b}$$

7.4 Divergence

From the co-ordinate independent definition of divergence (6.9), for any 'nice' surface $\mathcal{S}$ enclosing a volume $\mathcal{V}$

$$\boldsymbol{\nabla} . \,\mathbf{u} = \lim_{|\mathcal{V}|\to 0} \frac{1}{|\mathcal{V}|} \iint_{\mathcal{S}} \mathbf{u} . \,\mathrm{d}\mathbf{S} \,.$$

In order to derive an expression for the divergence in orthogonal curvilinear co-ordinates we consider the flux out of a slightly deformed cuboid with one vertex at $\mathbf{x}(\mathbf{p})$, where $\mathbf{p} = (p_1, p_2, p_3)$, and sides of length $h_1(\mathbf{p})\delta q_1$, $h_2(\mathbf{p})\delta q_2$ and $h_3(\mathbf{p})\delta q_3$ (see (7.4)).

On a surface where $q_1 =$constant, it follows from the definition of a vector elementary area (4.25b), and (7.1b), that

$$\begin{aligned} \mathrm{d}\mathbf{S} &= \left(\frac{\partial \mathbf{x}}{\partial q_2} \times \frac{\partial \mathbf{x}}{\partial q_3}\right) \mathrm{d}q_2 \mathrm{d}q_3 \\ &= (h_2\, \mathbf{e}_2 \times h_3\, \mathbf{e}_3) \mathrm{d}q_2 \mathrm{d}q_3 \\ &= \mathbf{e}_1 h_2 h_3 \mathrm{d}q_2 \mathrm{d}q_3 \,. \end{aligned}$$

Hence on the surface $q_1 = p_1 + \delta q_1$

$$\iint\limits_{\mathcal{S}_{p_1+\delta q_1}} \mathbf{u}.\,\mathrm{d}\mathbf{S} \;=\; \int_{p_2}^{p_2+\delta q_2} \mathrm{d}q_2 \int_{p_3}^{p_3+\delta q_3} \mathrm{d}q_3\, u_1(p_1+\delta q_1, q_2, q_3)\, h_2(p_1+\delta q_1, q_2, q_3)\, h_3(p_1+\delta q_1, q_2, q_3)\,.$$

Adding the contributions from the surfaces $q_1 = p_1 + \delta q_1$ and $q_1 = p_1$

$$\begin{aligned}
\iint\limits_{\mathcal{S}_{p_1+\delta q_1}} \mathbf{u}.\,\mathrm{d}\mathbf{S} + \iint\limits_{\mathcal{S}_{p_1}} \mathbf{u}.\,\mathrm{d}\mathbf{S} &= \int_{p_2}^{p_2+\delta q_2} \mathrm{d}q_2 \int_{p_3}^{p_3+\delta q_3} \mathrm{d}q_3 \left[(u_1 h_2 h_3)(p_1+\delta q_1, q_2, q_3) - (u_1 h_2 h_3)(p_1, q_2, q_3)\right] \\
&= \int_0^1 \mathrm{d}\beta \int_0^1 \mathrm{d}\gamma \left[\frac{\partial (u_1 h_2 h_3)}{\partial q_1}(p_1, p_2+\beta\delta q_2, p_3+\gamma\delta q_3)\delta q_1 + o(\delta q_1)\right]\delta q_2\, \delta q_3 \\
&\qquad \text{from Taylor's theorem and writing } q_2 = p_2 + \beta\delta q_2\,,\ q_3 = p_3+\gamma\delta q_3 \\
&= \underbrace{\int_0^1 \mathrm{d}\beta \int_0^1 \mathrm{d}\gamma}_{1}\ \frac{\partial (u_1 h_2 h_3)}{\partial q_1}(p_1, p_2, p_3)\, \delta q_1 \delta q_2 \delta q_3 + o(\delta q_1 \delta q_2 \delta q_3) \\
&\qquad \text{from another application of Taylor's theorem.}
\end{aligned}$$

Similar results can be deduced for the surfaces $q_2 =$ constant and $q_3 =$ constant; hence

$$\iint\limits_{\mathcal{S}} \mathbf{u}.\,\mathrm{d}\mathbf{S} = \left(\frac{\partial (u_1 h_2 h_3)}{\partial q_1} + \frac{\partial (u_2 h_3 h_1)}{\partial q_2} + \frac{\partial (u_3 h_1 h_2)}{\partial q_3}\right)\delta q_1 \delta q_2 \delta q_3 + o(\delta q_1 \delta q_2 \delta q_3)\,.$$

Further from the equation for an incremental change in position (7.4)

$$\mathrm{d}\mathbf{x} = h_1 \mathbf{e}_1 \mathrm{d}q_1 + h_2 \mathbf{e}_2 \mathrm{d}q_2 + h_3 \mathbf{e}_3 \mathrm{d}q_3$$

and thus

$$|\mathcal{V}| = h_1 h_2 h_3\, \delta q_1 \delta q_2 \delta q_3 + o(\delta q_1 \delta q_2 \delta q_3)\,.$$

It follows from the co-ordinate independent definition of divergence (6.9) that

$$\boldsymbol{\nabla}.\,\mathbf{u} = \frac{1}{h_1 h_2 h_3}\left(\frac{\partial (u_1 h_2 h_3)}{\partial q_1} + \frac{\partial (u_2 h_3 h_1)}{\partial q_2} + \frac{\partial (u_3 h_1 h_2)}{\partial q_3}\right). \tag{7.9}$$

20/00 *Remark.* Each term has dimensions 'u/length'.

7.4.1 Examples

Cylindrical Polar Co-ordinates. From (7.2a), (7.2b), (7.2c) and (7.9)

$$\boldsymbol{\nabla}.\,\mathbf{u} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho u_\rho) + \frac{1}{\rho}\frac{\partial u_\phi}{\partial \phi} + \frac{\partial u_z}{\partial z}\,. \tag{7.10a}$$

Spherical Polar Co-ordinates. From (7.3a), (7.3b), (7.3c) and (7.9)

$$\boldsymbol{\nabla}.\,\mathbf{u} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 u_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(\sin\theta\, u_\theta) + \frac{1}{r\sin\theta}\frac{\partial u_\phi}{\partial \phi}\,. \tag{7.10b}$$

7.5 Curl

From the co-ordinate independent definition of curl (6.15), for any 'nice' open surface $\mathcal{S}$ bounded by a curve $\mathcal{C}$

$$\mathbf{n}.\operatorname{curl}\mathbf{u} = \lim_{\mathcal{S}\to 0} \frac{1}{|\mathcal{S}|} \oint_C \mathbf{u}.\,\mathrm{d}\mathbf{x}\,.$$

Without loss of generality take $\mathbf{n} = \mathbf{e}_3$, and consider the circulation round a slightly deformed rectangle $\mathcal{C}$ with a vertex at $\mathbf{x}(\mathbf{p})$, where $\mathbf{p} = (p_1, p_2, p_3)$, and sides of length $h_1(\mathbf{p})\delta q_1$ and $h_2(\mathbf{p})\delta q_2$ (see (7.4)).

On $\mathcal{C}_1$ (where $q_2 = p_2$ and $q_3 = p_3$), $\mathrm{d}\mathbf{x} = h_1\mathbf{e}_1\mathrm{d}q_1$. Hence

$$\int\limits_{\mathcal{C}_1} \mathbf{u}.\,\mathrm{d}\mathbf{x} \quad = \quad \int_{p_1}^{p_1+\delta q_1} \mathrm{d}q_1\, h_1(q_1,p_2,p_3)\,u_1(q_1,p_2,p_3)\,.$$

Adding the contributions from $\mathcal{C}_1$ and $\mathcal{C}_3$:

$$\begin{aligned}
\int\limits_{\mathcal{C}_1} \mathbf{u}.\,\mathrm{d}\mathbf{x} + \int\limits_{\mathcal{C}_3} \mathbf{u}.\,\mathrm{d}\mathbf{x} \quad &= \quad \int_{p_1}^{p_1+\delta q_1} \mathrm{d}q_1\,[(u_1h_1)(q_1,p_2,p_3) - (u_1h_1)(q_1,p_2+\delta q_2,p_3)] \\
&= \quad -\int_0^1 \mathrm{d}\alpha \left[\frac{\partial(u_1h_1)}{\partial q_2}(p_1+\alpha\delta q_1,p_2,p_3)\,\delta q_2 + o(\delta q_2)\right]\delta q_1 \\
&\qquad\qquad \text{from Taylor's theorem and writing } q_1 = p_1 + \alpha\,\delta q_1 \\
&= \quad -\frac{\partial(u_1h_1)}{\partial q_2}(p_1,p_2,p_3)\,\delta q_1\delta q_2 + o(\delta q_1\delta q_2) \\
&\qquad\qquad \text{from another application of Taylor's theorem.}
\end{aligned}$$

Similarly

$$\int\limits_{\mathcal{C}_2} \mathbf{u}.\,\mathrm{d}\mathbf{x} + \int\limits_{\mathcal{C}_4} \mathbf{u}.\,\mathrm{d}\mathbf{x} \quad = \quad \frac{\partial(u_2h_2)}{\partial q_1}\,\delta q_1\delta q_2 + o(\delta q_1\delta q_2)\,.$$

Further

$$|\mathcal{S}| = h_1h_2\,\delta q_1\delta q_2 + o(\delta q_1\delta q_2)\,.$$

20/99 Hence from (6.15)

$$\mathbf{e}_3.\operatorname{curl}\mathbf{u} = \frac{1}{h_1h_2}\left[\frac{\partial}{\partial q_1}(u_2h_2) - \frac{\partial}{\partial q_2}(u_1h_1)\right]\,.$$

All three components of the curl can be written in the concise form

$$\operatorname{curl}\mathbf{u} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\mathbf{e}_1 & h_2\mathbf{e}_2 & h_3\mathbf{e}_3 \\ \partial q_1 & \partial q_2 & \partial q_3 \\ h_1u_1 & h_2u_2 & h_3u_3 \end{vmatrix}\,. \tag{7.11}$$

Remark. Each term has dimensions 'u/length'.

7.5.1 Examples

Cylindrical Polar Co-ordinates. From (7.2a), (7.2b), (7.2c) and (7.11)

$$\operatorname{curl}\mathbf{u} = \frac{1}{\rho}\begin{vmatrix} \mathbf{e}_\rho & \rho\,\mathbf{e}_\phi & \mathbf{e}_z \\ \partial_\rho & \partial_\phi & \partial_z \\ u_\rho & \rho\,u_\phi & u_z \end{vmatrix} \tag{7.12a}$$

$$= \left(\frac{1}{\rho}\frac{\partial u_z}{\partial\phi} - \frac{\partial u_\phi}{\partial z}, \frac{\partial u_\rho}{\partial z} - \frac{\partial u_z}{\partial\rho}, \frac{1}{\rho}\frac{\partial(\rho u_\phi)}{\partial\rho} - \frac{1}{\rho}\frac{\partial u_\rho}{\partial\phi}\right)\,. \tag{7.12b}$$

Spherical Polar Co-ordinates. From (7.3a), (7.3b), (7.3c) and (7.11)

$$\operatorname{curl}\mathbf{u} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\mathbf{e}_\phi \\ \partial_r & \partial_\theta & \partial_\phi \\ u_r & ru_\theta & r\sin\theta u_\phi \end{vmatrix} \tag{7.13a}$$

$$= \left(\frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta\, u_\phi)}{\partial\theta} - \frac{\partial u_\theta}{\partial\phi}\right), \frac{1}{r\sin\theta}\frac{\partial u_r}{\partial\phi} - \frac{1}{r}\frac{\partial(ru_\phi)}{\partial r}, \frac{1}{r}\frac{\partial(ru_\theta)}{\partial r} - \frac{1}{r}\frac{\partial u_r}{\partial\theta}\right). \tag{7.13b}$$

7.6 Laplacian

From (7.7) and (7.9)

$$\nabla^2 f = \boldsymbol{\nabla}.\,(\boldsymbol{\nabla} f) = \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3h_1}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right). \tag{7.14}$$

Cylindrical Polar Co-ordinates. From (7.2a), (7.2b), (7.2c) and (7.14)

$$\nabla^2 f = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial\phi^2} + \frac{\partial^2 f}{\partial z^2}, \tag{7.15a}$$

where the two-dimensional version of this with $\frac{\partial}{\partial z} \equiv 0$ has previously been derived in (1.57).

Spherical Polar Co-ordinates. From (7.3a), (7.3b), (7.3c) and (7.14)

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}. \tag{7.15b}$$

Remark. To find $\nabla^2\mathbf{A}$, for vector $\mathbf{A}$, you *must* use (5.30), i.e.

$$\nabla^2\mathbf{A} = \boldsymbol{\nabla}(\boldsymbol{\nabla}.\,\mathbf{A}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}).$$

7.7 Specific Examples

Evaluate $\boldsymbol{\nabla}.\,\mathbf{x}$, $\boldsymbol{\nabla}\times\mathbf{x}$, and $\nabla^2\left(\frac{1}{r}\right)$ in spherical polar co-ordinates, where

$$\mathbf{x} = (r, 0, 0).$$

From (7.10b)

$$\boldsymbol{\nabla}.\,\mathbf{x} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2.\,r\right) = 3.$$

From (7.13b)

$$\boldsymbol{\nabla}\times\mathbf{x} = \left(0, \frac{1}{r\sin\theta}\frac{\partial r}{\partial\phi}, -\frac{1}{r}\frac{\partial r}{\partial\theta}\right) = (0,0,0).$$

From (7.15b) for $r \neq 0$

$$\begin{aligned}\nabla^2\left(\frac{1}{r}\right) &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\right) \\ &= -\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2.\,\frac{1}{r^2}\right) = 0.\end{aligned} \tag{7.16}$$

$20\frac{1}{2}/00$ ($\frac{1}{2}$ late) All as before: see (5.7a), (5.15) and (5.23) respectively.

8 Poisson's Equation and Laplace's Equation.

In § 5.5 we encountered *Poisson's equation*

$$\nabla^2 f = \rho(\mathbf{x})\,, \tag{8.1a}$$

and *Laplace's equation* (which is a special case of Poisson's equation)

$$\nabla^2 f = 0\,, \tag{8.1b}$$

where the Laplacian operator is given by (see (5.21c), (7.15a) and (7.15b))

$$\nabla^2 = \begin{cases} \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} & \text{2D Cartesians,} \\ \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2} & \text{3D Cartesians,} \\ \dfrac{1}{\rho}\dfrac{\partial}{\partial \rho}\left(\rho\dfrac{\partial}{\partial\rho}\right) + \dfrac{1}{\rho^2}\dfrac{\partial^2}{\partial\phi^2} + \dfrac{\partial^2}{\partial z^2} & \text{Cylindrical Polars,} \\ \dfrac{1}{r^2}\dfrac{\partial}{\partial r}\left(r^2\dfrac{\partial}{\partial r}\right) + \dfrac{1}{r^2\sin\theta}\dfrac{\partial}{\partial\theta}\left(\sin\theta\,\dfrac{\partial}{\partial\theta}\right) + \dfrac{1}{r^2\sin^2\theta}\dfrac{\partial^2}{\partial\phi^2} & \text{Spherical Polars.} \end{cases} \tag{8.2}$$

Equations (8.1a) and (8.1b) are examples of *partial differential equations*.

How do we find solutions to these equations? What are appropriate boundary conditions?

8.1 Linearity of Laplace's Equation

Laplace's equation is *linear*, in that if f_1 and f_2 are solutions to (8.1b), i.e. if

$$\nabla^2 f_1 = 0 \quad \text{and} \quad \nabla^2 f_2 = 0\,,$$

then $(\lambda f_1 + \mu f_2)$, where λ and μ are real constants, is another solution — since from (5.21c)

$$\nabla^2(\lambda f_1 + \mu f_2) = \frac{\partial^2}{\partial x_j \partial x_j}(\lambda f_1 + \mu f_2) = \lambda\frac{\partial^2 f_1}{\partial x_j\partial x_j} + \mu\frac{\partial^2 f_2}{\partial x_j \partial x_j} = \lambda\nabla^2 f_1 + \mu\nabla^2 f_2 = 0\,.$$

Hence if we have a number of solutions of Laplace's equation, we can generate further solutions by adding multiples of the known solutions together.

8.2 Separable Solutions

You have already met the general idea of 'separability' in IA *Differential Equations* where you studied separable equations, i.e. the special differential equations that can be written in the form

$$\underbrace{X(x)\mathrm{d}x}_{\text{function of } x} \quad = \quad \underbrace{Y(y)\mathrm{d}y}_{\text{function of } y} \quad = \quad \text{constant}\,.$$

Sometimes functions can we written in separable form; for instance,

$$f(x,y) = \cos x\, \exp y = X(x)Y(y)\,, \quad \text{where} \quad X = \cos x \text{ and } Y = \exp y\,,$$

is separable in Cartesian co-ordinates, while

$$g(x,y,z) = \frac{1}{(x^2+y^2+z^2)^{\frac{1}{2}}}$$

is not separable in Cartesian co-ordinates, but is separable in spherical polar co-ordinates since

$$g = R(r)\Theta(\theta)\Phi(\phi) \quad \text{where} \quad R = \tfrac{1}{r}\,,\ \Theta = 1 \text{ and } \Phi = 1\,.$$

Solutions to Laplace's equation can be found by seeking solutions that that can be written in separable form, e.g.

$$\begin{aligned}
\text{2D Cartesians:} \quad & f(x,y) &=& \ X(x)Y(y)\,, && \text{(8.3a)}\\
\text{3D Cartesians:} \quad & f(x,y,z) &=& \ X(x)Y(y)Z(z)\,, && \text{(8.3b)}\\
\text{Plane Polars:} \quad & f(\rho,\phi) &=& \ R(\rho)\Phi(\phi)\,, && \text{(8.3c)}\\
\text{Cylindrical Polars:} \quad & f(\rho,\phi,z) &=& \ R(\rho)\Phi(\phi)Z(z)\,, && \text{(8.3d)}\\
\text{Spherical Polars:} \quad & f(r,\theta,\phi) &=& \ R(r)\Theta(\theta)\Phi(\phi)\,. && \text{(8.3e)}
\end{aligned}$$

However, we emphasise that *not* all solutions of Laplace's equation can be written in this form (which will be exploited in IB *Methods*).

8.2.1 Example: Two Dimensional Cartesian Co-ordinates

Seek solutions $f(x, y)$ to Laplace's equation of the form (8.3a). On substituting into (8.1b) we obtain

$$X''(x)Y(y) + X(x)Y''(y) = 0\,,$$

or after rearrangement

$$\underbrace{\frac{X''(x)}{X(x)}}_{\text{function of } x} \;=\; \underbrace{-\frac{Y''(y)}{Y(y)}}_{\text{function of } y} \;=\; C\,, \tag{8.4}$$

where C is a constant (the only function of x that equals a function of y). There are three cases to consider.

$C = 0$. In this case

$$X''(x) = Y''(y) = 0 \quad \Rightarrow \quad X = ax + b \quad \text{and} \quad Y = cy + d\,,$$

where a, b, c and d are constants, i.e.

$$f = (ax + b)(cy + d)\,. \tag{8.5a}$$

$C = k^2 > 0$. In this case

$$X'' - k^2X = 0\,, \quad \text{and} \quad Y'' + k^2Y = 0\,.$$

Hence

$$X = a\mathrm{e}^{kx} + b\mathrm{e}^{-kx}\,, \quad \text{and} \quad Y = c\sin ky + d\cos ky\,,$$

where a, b, c and d are constants, i.e.

$$f = (a\mathrm{e}^{kx} + b\mathrm{e}^{-kx})(c\sin ky + d\cos ky)\,. \tag{8.5b}$$

$C = -k^2 < 0$. Using the $x \leftrightarrow y$ symmetry in (8.4) and (8.5b) it follows that

$$f = (a\mathrm{e}^{ky} + b\mathrm{e}^{-ky})(c\sin kx + d\cos kx)\,. \tag{8.5c}$$

Remarks.

1. Observe from (8.5b) and (8.5c) that solutions to Laplace's equation tend to grow or decay smoothly in one direction while oscillating in the other.
2. Without loss of generality we could also impose, say, $c^2 + d^2 = 1$.

8.3 Isotropic Solutions

The Laplacian operator ∇^2 is *isotropic*, i.e. it has no preferred direction, since the operator is unchanged by the rotation of a Cartesian co-ordinate system. This suggests that there may be solutions to Laplace's equation that also have no preferred direction.

8.3.1 3D: Spherical Symmetry

On this basis in 3D we seek a spherically symmetric solution to Laplace's equation, i.e. we seek a solution independent of θ and ϕ (this is a special case of the separable solution (8.3e) with $\Theta = \Phi = 1$). Write $f = R(r)$, then from (7.15b) and Laplace's equation (8.1b)

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial R}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 R}{\partial\phi^2} = 0\,.$$

Hence we require

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R}{\mathrm{d}r}\right) = 0 \quad &\Rightarrow \quad \frac{\mathrm{d}R}{\mathrm{d}r} = \frac{\beta}{r^2} \\ &\Rightarrow \quad R = \alpha - \frac{\beta}{r}\,, \end{aligned} \tag{8.6}$$

where α and β are constants.

Remarks.

1. There is a 'singularity' at $r = 0$ if $\beta \neq 0$; hence the solution is not valid at $r = 0$ if $\beta \neq 0$.
2. We already knew that $\nabla^2\left(\frac{1}{r}\right) = 0$ for $r \neq 0$ from (5.23).
3. Consider the flux of $\mathbf{u} = \boldsymbol{\nabla} f$ out of a sphere $\mathcal{S}$ of radius r, then using the expression for grad (7.8b)

$$\begin{aligned} \iint_{\mathcal{S}} \boldsymbol{\nabla} f.\,\mathrm{d}\mathbf{S} &= \iint_{\mathcal{S}} \mathbf{n}.\,\boldsymbol{\nabla}\left(\alpha - \frac{\beta}{r}\right)\mathrm{d}S \\ &= \iint_{\mathcal{S}} \frac{\partial}{\partial r}\left(\alpha - \frac{\beta}{r}\right)\mathrm{d}S \\ &= \iint_{\mathcal{S}} \frac{\beta}{r^2}\,\mathrm{d}S = 4\pi\beta\,. \end{aligned} \tag{8.7}$$

The flux is thus independent of r, and there is unit flux if $\beta = \frac{1}{4\pi}$.

Definition. The solution

$$f = -\frac{1}{4\pi r} \tag{8.8}$$

is referred to as the potential of the *unit point source*, or as the *fundamental solution* of Poisson's
21/00 equation in 3D — see also (8.11a).

8.3.2 2D: Radial Symmetry

There is an analogous result in 2D — see *Examples Sheet 4*. In particular radially symmetric solutions to Laplace's equation (8.1b) exist, and are given by (cf. (8.6))

$$f = \alpha + \beta\log\rho\,, \tag{8.9a}$$

where α and β are constants.

Remarks.

1. There is again a 'singularity' at the origin if $\beta \neq 0$, i.e. the solution is not valid at $\rho = 0$ if $\beta \neq 0$.
2. Consider the flux of $\mathbf{u} = \boldsymbol{\nabla} f$ out of a circle $\mathcal{C}$ of radius ρ, then using the expression for grad (7.8a)

$$\begin{aligned}
\int\limits_{\mathcal{C}} \boldsymbol{\nabla} f.\,\mathbf{n}\,\mathrm{d}s &= \int\limits_{\mathcal{C}} \mathbf{n}.\,\boldsymbol{\nabla}\left(\alpha + \beta\log\rho\right)\,\mathrm{d}s \\
&= \int\limits_{\mathcal{C}} \frac{\partial}{\partial\rho}\left(\alpha + \beta\log\rho\right)\,\mathrm{d}s \\
&= \int_0^{2\pi} \frac{\beta}{\rho}\,\rho\,\mathrm{d}\phi = 2\pi\beta\,.
\end{aligned}$$

   The flux is thus independent of ρ, and there is unit flux if $\beta = \frac{1}{2\pi}$.

Definition. The solution

$$f = \frac{1}{2\pi}\log\rho \tag{8.9b}$$

is referred to as the potential of the *unit line source*, or as the *fundamental solution* of Poisson's equation in 2D.

8.4 Monopoles and Dipoles

We begin by noting that if $f(\mathbf{x})$ is a solution to Laplace's equation (8.1b) then further solutions can be generated by differentiating with respect to Cartesian co-ordinates; for suppose that

$$g(\mathbf{x}) = \frac{\partial f}{\partial x_i}\,, \tag{8.10}$$

then from (5.21c)

$$\nabla^2 g = \frac{\partial^2}{\partial x_j \partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial}{\partial x_i}\left(\frac{\partial^2 f}{\partial x_j \partial x_j}\right) = 0\,.$$

Definition. The fundamental solution (8.8), i.e.

$$\varphi_m \equiv -\frac{1}{4\pi r}\,, \quad r \neq 0\,, \tag{8.11a}$$

is sometimes referred to as the *unit monopole.*

Definition. Let $\mathbf{d}$ be a constant vector, then using (1.18) and (8.10), it follows that

$$\varphi_{\mathbf{d}} = -d_i\frac{\partial\varphi_m}{\partial x_i} = -\frac{d_i}{4\pi r^2}\frac{\partial r}{\partial x_i} = -\frac{\mathbf{d}.\,\mathbf{x}}{4\pi r^3}\,, \quad r \neq 0\,, \tag{8.11b}$$

is a solution to Laplace's equation. This is referred to as a *dipole* of strength $\mathbf{d}$.

Remark. A dipole can be viewed as being generated by two point sources that are of large, but opposite,
21/99 strength and very close to each other (see *Examples Sheet 4*).

8.5 Harmonic Functions

Definition. The function $f : \mathbb{R}^m \to \mathbb{R}$ $(m = 2, 3)$ is *harmonic* iff all its second-order partial derivatives exist and are continuous, and

$$\nabla^2 f = 0\,.$$

Remark. In IB *Complex Methods* you will see that the real and imaginary parts of a complex differentiable function are both harmonic (this follows in a straightforward manner from the Cauchy-Riemann equations). Hence, solutions of Laplace's equation can be found by taking the real and imaginary parts of any complex differentiable function.

8.6 Green's Formulae, or Green's Theorem (But A Different One)

8.6.1 Three Dimensions

Suppose $\varphi, \psi, \nabla^2\varphi$ and $\nabla^2\psi$ are defined in a region $\mathcal{V}$ bounded by a piecewise smooth surface $\mathcal{S}$. Let $\frac{\partial\psi}{\partial n}$ denote the component of $\boldsymbol{\nabla}\psi$ in the direction of the local normal of $\mathcal{S}$, then

Green's First Formula

$$\iiint\limits_{\mathcal{V}} (\varphi\nabla^2\psi + \boldsymbol{\nabla}\varphi.\,\boldsymbol{\nabla}\psi)\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \varphi\boldsymbol{\nabla}\psi.\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}} \varphi\frac{\partial\psi}{\partial n}\mathrm{d}S\,; \tag{8.12a}$$

Green's Second Formula

$$\iiint\limits_{\mathcal{V}} (\varphi\nabla^2\psi - \psi\nabla^2\varphi)\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)\mathrm{d}S\,. \tag{8.12b}$$

Remark. Green's First Formula is analogous to integration by parts: $\int(fg'' + f'g')\mathrm{d}x = [fg']$.

Outline Proof.

Green's First Formula. Put $\mathbf{u} = \varphi\boldsymbol{\nabla}\psi$ in the divergence theorem (6.1). Then since

$$\boldsymbol{\nabla}.\,\mathbf{u} = \boldsymbol{\nabla}.\,(\varphi\boldsymbol{\nabla}\psi) = \frac{\partial}{\partial x_j}\left(\varphi\frac{\partial\psi}{\partial x_j}\right) = \varphi\frac{\partial^2\psi}{\partial x_j\partial x_j} \;+\; \frac{\partial\varphi}{\partial x_j}\frac{\partial\psi}{\partial x_j}\,,$$

it follows that

$$\iiint\limits_{\mathcal{V}} (\varphi\nabla^2\psi + \boldsymbol{\nabla}\varphi.\,\boldsymbol{\nabla}\psi)\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \varphi\boldsymbol{\nabla}\psi.\,\mathrm{d}\mathbf{S}\,. \tag{8.13a}$$

Green's Second Formula. Similarly with $\varphi \leftrightarrow \psi$

$$\iiint\limits_{\mathcal{V}} (\psi\nabla^2\varphi + \boldsymbol{\nabla}\psi.\,\boldsymbol{\nabla}\varphi)\,\mathrm{d}\tau = \iint\limits_{\mathcal{S}} \psi\boldsymbol{\nabla}\varphi.\,\mathrm{d}\mathbf{S}\,. \tag{8.13b}$$

Now subtract (8.13b) from (8.13a) to obtain (8.12b). □

8.6.2 Two Dimensions

Suppose φ, ψ, $\nabla^2\varphi$ and $\nabla^2\psi$ are defined on a closed surface $\mathcal{A}$ bounded by a simple closed curve $\mathcal{C}$. Then

Green's First Formula

$$\iint\limits_{\mathcal{A}} (\varphi\nabla^2\psi + \boldsymbol{\nabla}\varphi.\,\boldsymbol{\nabla}\psi)\,\mathrm{d}S = \oint\limits_{\mathcal{C}} \varphi(\boldsymbol{\nabla}\psi.\,\mathbf{n})\mathrm{d}s = \oint\limits_{\mathcal{C}} \varphi\frac{\partial\psi}{\partial n}\,\mathrm{d}s\,. \tag{8.14a}$$

Green's Second Formula

$$\iint\limits_{\mathcal{A}} (\varphi\nabla^2\psi - \psi\nabla^2\varphi)\,\mathrm{d}S = \oint\limits_{\mathcal{C}} \left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)\mathrm{d}s\,. \tag{8.14b}$$

Outline Proof. As above but using the two dimensional divergence theorem (6.4).

8.7 Uniqueness Theorems

You have considered questions of uniqueness before. For instance, in IA *Algebra & Geometry* you saw that if A is a $m \times m$ matrix then the solution to

$$\mathsf{A}\mathbf{x} = \mathbf{b} \quad \text{where} \quad \mathbf{x}, \mathbf{b} \in \mathbb{R}^m \,,$$

is unique iff $\det \mathsf{A} \neq 0$.

How do we know that a solution of Poisson's/Laplace's equations for a given set of boundary conditions is unique? Indeed how can we be sure that the correct number of boundary conditions have been specified? More generally how do we know that a solution to a specified problem exists?

Existence is hard and will not be tackled here — but we can make progress on uniqueness if we already know that a solution exists.

8.7.1 Generic Problems

Suppose that in a volume $\mathcal{V}$ bounded by a surface $\mathcal{S}$

$$\nabla^2 \varphi = \rho \,.$$

One of three common types of boundary condition is normally specified for this Poisson equation.

Dirichlet Condition. In this case φ is specified on $\mathcal{S}$; i.e.

$$\varphi(\mathbf{x}) = g(\mathbf{x}) \quad \text{on} \quad \mathcal{S} \,, \tag{8.15a}$$

where $g(\mathbf{x})$ is a known function.

Neumann Condition. In this case $\frac{\partial \varphi}{\partial n}$ is specified on $\mathcal{S}$; i.e.

$$\frac{\partial \varphi}{\partial n} = \mathbf{n} . \boldsymbol{\nabla} \varphi = h(\mathbf{x}) \quad \text{on} \quad \mathcal{S} \,, \tag{8.15b}$$

where $h(\mathbf{x})$ is a known function.

Mixed Condition. In this case a linear combination of φ and $\frac{\partial \varphi}{\partial n}$ is specified on $\mathcal{S}$; i.e.

$$\alpha(\mathbf{x}) \frac{\partial \varphi}{\partial n} + \beta(\mathbf{x}) \varphi = d(\mathbf{x}) \quad \text{on} \quad \mathcal{S} \,, \tag{8.15c}$$

where $\alpha(\mathbf{x})$, $\beta(\mathbf{x})$ and $d(\mathbf{x})$ are known functions, and $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are not simultaneously zero.

8.7.2 Uniqueness: Three Dimensions

Dirichlet Condition. Suppose φ is a scalar field satisfying $\nabla^2 \varphi = \rho(\mathbf{x})$ in $\mathcal{V}$ and $\varphi = g(\mathbf{x})$ on $\mathcal{S}$. Then φ is unique.

Outline Proof. Suppose that there exist two functions, say φ_1 and φ_2, satisfying

$$\nabla^2 \varphi_j = \rho \;\text{ in }\; \mathcal{V} \,, \quad \text{and} \quad \varphi_j = g \;\text{ on }\; \mathcal{S} \,.$$

Let $\Phi = (\varphi_1 - \varphi_2)$, then

$$\nabla^2 \Phi = 0 \;\text{ in }\; \mathcal{V} \,, \quad \text{and} \quad \Phi = 0 \;\text{ on }\; \mathcal{S} \,. \tag{8.16}$$

From Green's First Formula (8.12a) with $\varphi = \psi = \Phi$,

$$\iiint\limits_{\mathcal{V}} \left(\Phi \nabla^2 \Phi + \boldsymbol{\nabla} \Phi . \boldsymbol{\nabla} \Phi \right) \mathrm{d}\tau = \iint\limits_{\mathcal{S}} \Phi \frac{\partial \Phi}{\partial n} \, \mathrm{d}S \,.$$

Hence from (8.16)

$$\iiint_{\mathcal{V}} |\boldsymbol{\nabla}\Phi|^2 \,\mathrm{d}\tau = 0\,.$$

Suppose $\boldsymbol{\nabla}\Phi \neq 0$ at $\mathbf{x}_0 \in \mathcal{V}$, and assume that the φ_j, and hence Φ, are continuous. Then if

$$|\boldsymbol{\nabla}\Phi(\mathbf{x}_0)| = \varepsilon > 0\,,$$

it follows from the continuity of Φ, that there exists $\delta > 0$ such that

$$|\boldsymbol{\nabla}\Phi(\mathbf{x})| > \tfrac{1}{2}\varepsilon \quad \forall \quad |\mathbf{x}-\mathbf{x}_0| < \delta\,.$$

Hence

$$\iiint_{\mathcal{V}} |\boldsymbol{\nabla}\Phi|^2 \,\mathrm{d}\tau \geqslant \iiint_{|\mathbf{x}-\mathbf{x}_0|<\delta} |\boldsymbol{\nabla}\Phi|^2 \,\mathrm{d}\tau \geqslant \tfrac{4}{3}\pi\delta^3\left(\tfrac{1}{4}\varepsilon^2\right) > 0 \qquad \#\,.$$

It follows that

$$\boldsymbol{\nabla}\Phi = 0 \;\text{ in }\; \mathcal{V} \quad\Rightarrow\quad \Phi = \text{constant} \;\text{ in }\; \mathcal{V}\,.$$

But $\Phi = 0$ on $\mathcal{S}$, hence

$$\Phi = 0 \;\text{ in }\; \mathcal{V} \quad\Rightarrow\quad \varphi_1 = \varphi_2 \;\text{ in }\; \mathcal{V}\,. \qquad \square$$

Neumann Condition. Suppose instead that

$$\nabla^2\varphi = \rho(\mathbf{x}) \;\text{ in }\; \mathcal{V} \quad\text{and}\quad \frac{\partial\varphi}{\partial n} = h(\mathbf{x}) \;\text{ on }\; \mathcal{S}\,.$$

Then φ is unique upto a constant.

Outline Proof. The proof is as above, but with $\frac{\partial\Phi}{\partial n} = 0$ on $\mathcal{S}$, instead of $\Phi = 0$ on $\mathcal{S}$. We again deduce that $\Phi =$ constant in $\mathcal{V}$. However, now there is no means of fixing the constant. Hence

$$\varphi_1 = \varphi_2 + k\,,$$

where k is a constant. $\square$

Mixed Condition. See *Examples Sheet 4*.

8.7.3 Uniqueness: Two Dimensions

Analogous uniqueness results hold in two dimensions, and can be proved using (8.14a) instead of (8.12a). For instance, suppose $\mathcal{S}$ is a closed region of the plane bounded by a simple piecewise smooth closed curve $\mathcal{C}$, and suppose φ is a scalar field satisfying Poisson's equation with a Dirichlet boundary condition, i.e.

$$\nabla^2\varphi = \rho(\mathbf{x}) \quad\text{on}\quad \mathcal{S}\,, \quad\text{and}\quad \varphi = g(\mathbf{x}) \quad\text{on}\quad \mathcal{C},$$

then φ is unique.

8.7.4 Boundary Conditions

The uniqueness results imply that solutions to Poisson's equation, $\nabla^2\varphi = \rho$, are determined by the value of φ or $\frac{\partial\varphi}{\partial n}$ (or a linear combination of φ and $\frac{\partial\varphi}{\partial n}$), on the boundary.

Hence appropriate boundary conditions for Poisson's equation include the specification of φ, or $\frac{\partial\varphi}{\partial n}$, on the *whole* boundary.

8.8 Maxima and Minima of Harmonic Functions

8.8.1 Three Dimensions

Suppose $\varphi(\mathbf{x})$ is harmonic, i.e. $\nabla^2\varphi = 0$, in a volume $\mathcal{V}$ bounded by a piecewise smooth closed surface $\mathcal{S}$. Then φ attains its maximum and minimum values on $\mathcal{S}$.

∇^2 *is a smooth operator — no spikes.*

Outline Proof. Suppose φ has a maximum at an interior point $\mathbf{x}_0$. Surround $\mathbf{x}_0$ by a sphere $\mathcal{S}_\delta$ small enough so that $\frac{\partial\varphi}{\partial n} < 0$ on $\mathcal{S}_\delta$ (always possible if φ differentiable). Also choose $\mathcal{S}_\delta$ so that it is entirely within $\mathcal{V}$. Denote by $\mathcal{V}_\delta$ the volume within the $\mathcal{S}_\delta$.

Let

$$m = \min_{\mathbf{x}\in\mathcal{S}_\delta} \varphi(\mathbf{x}),$$

and define

$$\psi(\mathbf{x}) = \varphi(\mathbf{x}) - m + \varepsilon,$$

where $\varepsilon > 0$. Then

(a) ψ is harmonic in $\mathcal{V}$, since

$$\nabla^2\psi = \nabla^2(\varphi - m + \varepsilon) = 0;$$

(b) $\psi > 0$ on $\mathcal{S}_\delta$, since $\varphi - m \geqslant 0$ on $\mathcal{S}_\delta$;

(c) $\frac{\partial\psi}{\partial n} < 0$ on $\mathcal{S}_\delta$, since by construction $\frac{\partial\psi}{\partial n} = \frac{\partial\varphi}{\partial n} < 0$ on $\mathcal{S}_\delta$.

Set $\varphi = \psi$ in Green's First Formula (8.12a), to obtain

$$\iiint_{\mathcal{V}_\delta} \left(\psi\nabla^2\psi + |\boldsymbol{\nabla}\psi|^2\right) \mathrm{d}\tau = \iint_{\mathcal{S}_\delta} \psi\frac{\partial\psi}{\partial n}\,\mathrm{d}S.$$

Hence using (a)–(c)

$$0 \leqslant \iiint_{\mathcal{V}_\delta} |\boldsymbol{\nabla}\psi|^2\,\mathrm{d}\tau = \iint_{\mathcal{S}_\delta} \psi\frac{\partial\psi}{\partial n}\,\mathrm{d}S < 0. \qquad \#$$

Thus φ cannot have a maximum at an interior point. Similarly for a minimum, or alternatively consider a maximum of $(-\varphi)$. □

Remark. The argument does not apply for $\mathbf{x}_0$ on the surface $\mathcal{S}$, because it is then impossible to surround $\mathbf{x}_0$ by a sphere $\mathcal{S}_\delta$ that lies totally within $\mathcal{V}$.

8.8.2 Two Dimensions

Suppose $u(x,y)$ is harmonic, i.e. $\nabla^2 u = 0$, in a region $\mathcal{A}$ bounded by a simple piecewise smooth closed curve $\mathcal{C}$. Then u attains its maximum and minimum values on $\mathcal{C}$.

Outline Proof. Proceed as in three dimensions.

Example. Find the maximum of

$$u = \sin x\ \cosh y$$

on the unit square $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$.

22/00

Answer. u is a special case of separable solution (8.5c). Thus it is harmonic, and its maximum must be on the boundary. Since both sin and cosh are increasing on $[0,1]$, it follows that

$$\max\ u = \sin 1\ \cosh 1.$$

8.8.3 Unlectured Example: A Mean Value Theorem.

Suppose that the scalar field φ is harmonic in a volume $\mathcal{V}$ bounded by a closed surface $\mathcal{S}$. Then, using the divergence theorem, we note the useful result that

$$\iint_{\mathcal{S}} \frac{\partial\varphi}{\partial n}\,\mathrm{d}S = \iint_{\mathcal{S}} \boldsymbol{\nabla}\varphi.\mathrm{d}\mathbf{S} = \iiint_{\mathcal{V}} \nabla^2\varphi\,\mathrm{d}\tau = 0\,. \tag{8.17}$$

Consider the function $f(r)$ defined to be the mean value of φ on a spherical surface $\mathcal{S}_r$ of radius r and normal $\mathbf{n}$, i.e.

$$\begin{aligned} f(r) &= \frac{1}{4\pi r^2}\iint_{\mathcal{S}_r} \varphi(\mathbf{x})\,\mathrm{d}S \\ &= \frac{1}{4\pi}\int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\;\sin\theta\;\varphi(r,\theta,\phi)\,, \end{aligned} \tag{8.18}$$

where (r,θ,ϕ) are spherical polar co-ordinates. We claim that f is a constant, and that it is equal to the value of φ at the origin.

We can show this result either by using Green's Second Formula (see *Examples Sheet 4*), or directly as follows using (8.17):

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}r} &= \frac{1}{4\pi}\int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\;\sin\theta\;\frac{\partial\varphi}{\partial r} \\ &= \frac{1}{4\pi}\int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\;\sin\theta\;\frac{\partial\varphi}{\partial n} \\ &= \frac{1}{4\pi r^2}\iint_{\mathcal{S}_r} \frac{\partial\varphi}{\partial n}\,\mathrm{d}S \\ &= 0\,. \end{aligned}$$

Further, in the limit $r \to 0$ it follows from (8.18) that

$$f(0) = \varphi(\mathbf{0})\,,$$

and hence

$$\varphi(\mathbf{0}) = \frac{1}{4\pi r^2}\iint_{\mathcal{S}_r} \varphi(\mathbf{x})\,\mathrm{d}S\,.$$

We conclude the mean value result that:

> *if φ is harmonic, the value of φ at a point is equal to the average of the values of φ on any spherical shell centred at that point.*

9 Poisson's Equation: Applications

9.1 Gravity

9.1.1 The Gravitational Potential

The *gravitational acceleration*, or *gravitational field*, $\mathbf{g}$ at $\mathbf{x}$ due to a particle of *mass* m_1 at $\mathbf{x}_1$ is given by Newton's famous inverse square law

$$\mathbf{g} = -Gm_1 \frac{\mathbf{x}-\mathbf{x}_1}{|\mathbf{x}-\mathbf{x}_1|^3} \quad \text{for} \quad \mathbf{x} \neq \mathbf{x}_1\,, \tag{9.1}$$

where G is the *universal gravitational constant* and space is taken to be unbounded (i.e. the whole of $\mathbb{R}^3$).

For the special case $\mathbf{x}_1 = 0$ we have seen in (5.16) that $\operatorname{curl}\mathbf{g} = 0$; the case $\mathbf{x}_1 \neq 0$ follows after a co-ordinate translation, e.g. let $\mathbf{x}' = \mathbf{x} - \mathbf{x}_1$ (see *Examples Sheet 3*). Hence from the results on irrotational fields in § 6.3.6, (a) it follows that gravity as defined by (9.1) is a conservative force, and (b) it follows from Poincaré's lemma (6.16) that there exists a scalar *gravitational potential* φ such that

$$\mathbf{g} = -\boldsymbol{\nabla}\varphi\,. \tag{9.2a}$$

Further, we deduce from the evaluation of the gradient of $1/r$ in (1.46a) that for $\mathbf{x} \neq \mathbf{x}_1$

$$\varphi = -\frac{Gm_1}{|\mathbf{x}-\mathbf{x}_1|}\,. \tag{9.2b}$$

Hence from (9.2a), (9.2b) and (5.23) it follows that

$$\boldsymbol{\nabla}.\,\mathbf{g} = -\nabla^2\varphi = Gm_1\nabla^2\left(\frac{1}{|\mathbf{x}-\mathbf{x}_1|}\right) = 0\,, \qquad \mathbf{x} \neq \mathbf{x}_1\,. \tag{9.3}$$

Remarks.

- The gravitational potential for a point mass (9.2b) is a multiple of the fundamental solution (8.8) for a monopole, i.e. a unit point source.
- The gravitational potential for a point mass (9.2b) satisfies Laplace's equation *except* at the position of the point mass $\mathbf{x} = \mathbf{x}_1$.[10]

22/99

9.1.2 Gauss' Flux Theorem

Point Masses. Let $\mathcal{S}$ be a closed surface, and let $\mathcal{V}$ be the volume interior to $\mathcal{S}$. We start by considering the flux of gravitational field $\mathbf{g}$ out of the surface $\mathcal{S}$ generated by a single particle of mass m_1 at $\mathbf{x}_1$. There are two cases to consider: $\mathbf{x}_1 \notin \mathcal{V}$ and $\mathbf{x}_1 \in \mathcal{V}$ (we will sidestep the case of when the point mass is on the surface).

$\mathbf{x}_1 \notin \mathcal{V}$. Then from the divergence theorem (6.1) and (9.3)

$$\iint\limits_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = \iiint\limits_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{g}\,\mathrm{d}\tau = 0\,. \tag{9.4a}$$

[10] In fact we can write down an equation that is valid for all $\mathbf{x}$, namely

$$\nabla^2\varphi = 4\pi G m_1 \delta(\mathbf{x}-\mathbf{x}_1)\,,$$

where $\delta(\mathbf{x})$ is the three-dimensional Dirac delta function. Hence the gravitational potential for a point mass actually satisfies Poisson's equation.

$\mathbf{x}_1 \in \mathcal{V}$. Before considering a general surface $\mathcal{S}$, suppose that $\mathcal{S} = \mathcal{S}_\delta$ where $\mathcal{S}_\delta$ is a sphere of radius δ centred on $\mathbf{x}_1$. Let $\mathcal{V}_\delta$ be the interior volume to a $\mathcal{S}_\delta$, and note that on $\mathcal{S}_\delta$ the normal $\mathbf{n}$ is parallel to $(\mathbf{x} - \mathbf{x}_1)$. Hence from (9.1), and proceeding as in the calculation leading to (8.7),

$$\begin{aligned} \iint\limits_{\mathcal{S}_\delta} \mathbf{g}.\,\mathrm{d}\mathbf{S} &= -\iint\limits_{\mathcal{S}_\delta} \frac{Gm_1}{|\mathbf{x}-\mathbf{x}_1|^3}\,(\mathbf{x}-\mathbf{x}_1).\,\mathbf{n}\,\mathrm{d}S \\ &= -Gm_1 \iint\limits_{\mathcal{S}_\delta} \frac{1}{\delta^2}\,\mathrm{d}S = -4\pi G m_1\,. \end{aligned} \tag{9.4b}$$

Suppose now that $\mathcal{S}$ is a general surface such that $\mathbf{x}_1 \in \mathcal{V}$. Then define

$$\widehat{\mathcal{V}} = \mathcal{V} - \mathcal{V}_\delta\,, \qquad \widehat{\mathcal{S}} = \mathcal{S} - \mathcal{S}_\delta\,,$$

where δ is chosen to be sufficiently small that $\mathcal{V}_\delta$ lies totally within $\mathcal{V}$ and $-\mathcal{S}_\delta$ means that the normal is directed towards the origin. Since $\mathbf{x}_1 \notin \widehat{\mathcal{V}}$ it follows from (9.4a) that

$$0 = \iiint\limits_{\widehat{\mathcal{V}}} \boldsymbol{\nabla}.\,\mathbf{g}\,\mathrm{d}\tau = \iint\limits_{\widehat{\mathcal{S}}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} - \iint\limits_{\mathcal{S}_\delta} \mathbf{g}.\,\mathrm{d}\mathbf{S}\,.$$

Hence from (9.4b)

$$\iint\limits_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = \iint\limits_{\mathcal{S}_\delta} \mathbf{g}.\,\mathrm{d}\mathbf{S} = -4\pi G m_1\,.$$

We conclude that in both cases that if M is the total mass inside $\mathcal{S}$ then

$$\iint\limits_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = -4\pi G M\,. \tag{9.5a}$$

Suppose now that there is more than one mass, e.g. $m_1, m_2, \dots, m_P$ at points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_P$ respectively. Then, since the gravitational fields of a number of particles can be superimposed, (9.5a) is modified to

$$\iint\limits_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = -4\pi G \sum_{\mathbf{x}_j \in \mathcal{V}} m_j = -4\pi G M\,, \tag{9.5b}$$

where M is again the total mass inside $\mathcal{S}$.

The Continuum Limit. Suppose that $\delta\tau_{\boldsymbol{\xi}}$ is a volume surrounding an arbitrary point $\boldsymbol{\xi}$, and suppose also that $\delta\tau_{\boldsymbol{\xi}}$ contains a mass $\delta m_{\boldsymbol{\xi}}$. The *continuum limit* consists in letting $P \to \infty$ and $m_j \to 0$ $(j = 1, \dots, P)$ in such a way that if we subsequently consider the limit in which $\delta\tau_{\boldsymbol{\xi}}$ contracts to zero volume, then

$$\rho(\boldsymbol{\xi}) = \lim_{\delta\tau_{\boldsymbol{\xi}} \to 0} \frac{\delta m_{\boldsymbol{\xi}}}{\delta\tau_{\boldsymbol{\xi}}}$$

exists for all $\boldsymbol{\xi}$; the quantity $\rho(\boldsymbol{\xi})$ is referred to as the *density of mass* per unit volume.

We wish to determine the total mass, M, in a volume $\mathcal{V}$. By analogy with the definition of a triple integral (4.4), subdivide $\mathcal{V}$ into N sub-domains $\mathcal{V}_1, \dots, \mathcal{V}_N$, and suppose $\boldsymbol{\xi}_i \in \mathcal{V}_i$ $(i = 1, \dots, N)$. Denote the volume of $\mathcal{V}_i$ by $\delta\tau_{\boldsymbol{\xi}_i}$ and the mass contained within it by $\delta m_{\boldsymbol{\xi}_i}$ where

$$\delta m_{\boldsymbol{\xi}_i} = \sum_{\mathbf{x}_j \in \mathcal{V}_i} m_j$$

Then a little arm-waving yields

$$M = \sum_{\mathbf{x}_j \in \mathcal{V}} m_j = \sum_{i=1}^{N} \sum_{\mathbf{x}_j \in \mathcal{V}_i} m_j = \sum_{i=1}^{N} \delta m_{\boldsymbol{\xi}_i} = \lim_{\substack{N\to\infty \\ \delta\tau_{\boldsymbol{\xi}_i}\to 0}} \sum_{i=1}^{N} \rho(\boldsymbol{\xi}_i)\, \delta\tau_{\boldsymbol{\xi}_i} = \iiint_{\mathcal{V}} \rho(\boldsymbol{\xi})\, \mathrm{d}\tau_{\boldsymbol{\xi}} \,.$$

Hence on taking the continuum limit of (9.5b) we have *Gauss' flux theorem*:

$$\iint_{\mathcal{S}} \mathbf{g}.\, \mathrm{d}\mathbf{S} = -4\pi G \iiint_{\mathcal{V}} \rho \, \mathrm{d}\tau \,. \tag{9.6}$$

9.1.3 Poisson's Equation and Solution

Poisson's Equation. Applying the divergence theorem (6.1) to Gauss' flux theorem we have that

$$\iiint_{\mathcal{V}} (\boldsymbol{\nabla}.\, \mathbf{g} + 4\pi G\rho)\, \mathrm{d}\tau = 0\,.$$

But this is true for all volumes $\mathcal{V}$, hence

$$\boldsymbol{\nabla}.\, \mathbf{g} = -4\pi G\rho\,. \tag{9.7a}$$

Further, we know from (9.2a) that $\mathbf{g} = -\boldsymbol{\nabla}\varphi$ for a single point mass. On the assumption that this result continues to hold on taking the continuum limit, e.g. because of the principle of superposition, it follows from (9.7a) that in the continuum limit φ satisfies *Poisson's equation*

$$\nabla^2 \varphi = 4\pi G\rho\,. \tag{9.7b}$$

The Solution in Unbounded Space. We have from (9.2b), and the principle of superposition, that for more than one mass in unbounded space

$$\varphi(\mathbf{x}) = -G \sum_j \frac{m_j}{|\mathbf{x} - \mathbf{x}_j|}\,.$$

We again take the continuum limit in a similar way to above, i.e. from using Taylor's theorem and in particular (1.62),

$$\varphi(\mathbf{x}) = -G \sum_{i=1}^{N} \sum_{\mathbf{x}_j \in \mathcal{V}_i} \frac{m_j}{|\, (\mathbf{x} - \boldsymbol{\xi}_i) + \underbrace{(\boldsymbol{\xi}_i - \mathbf{x}_j)}_{\text{small}}\, |} = -G \lim_{\substack{N\to\infty \\ \delta\tau_{\boldsymbol{\xi}_i}\to 0}} \sum_{i=1}^{N} \left(\frac{\rho(\boldsymbol{\xi}_i)\, \delta\tau_{\boldsymbol{\xi}_i}}{|\, (\mathbf{x} - \boldsymbol{\xi}_i)\, |} + \dots \right),$$

and hence the solution to (9.7b) is

$$\varphi(\mathbf{x}) = -G \iiint_{\mathcal{V}} \frac{\rho(\boldsymbol{\xi})}{|\mathbf{x} - \boldsymbol{\xi}|}\, \mathrm{d}\tau_{\boldsymbol{\xi}}\,, \tag{9.8}$$

where the volume integral extends over all regions where the density is non-zero.

Conditions at a Discontinuity in Density. Suppose that the density ρ is discontinuous (but bounded) across a surface $\mathcal{S}$. How do φ and $\boldsymbol{\nabla}\varphi$ vary across the discontinuity?

Let $\mathbf{x}_0$ be a point on the surface $\mathcal{S}$. Surround $\mathbf{x}_0$ with a surface $\widehat{\mathcal{S}}$ that is a small squashed 'pill-box' of cross-sectional area $\delta A \ll 1$ and thickness $\delta n \ll (\delta A)^{\frac{1}{2}}$. Let the enclosing volume be $\widehat{\mathcal{V}}$, and suppose that the normal of the cross-sectional faces is aligned with the normal of the surface $\mathcal{S}$ at $\mathbf{x}_0$. Then for a very squashed small pill box

$$\begin{aligned} \iint\limits_{\widehat{\mathcal{S}}} \mathbf{g} . \mathrm{d}\mathbf{S} \quad &\approx \quad \big(\mathbf{n} . \mathbf{g}(\mathbf{x}_0+) - \mathbf{n} . \mathbf{g}(\mathbf{x}_0-)\big)\, \delta A \,, \\ \iiint\limits_{\widehat{\mathcal{V}}} \rho \, \mathrm{d}\tau \quad &\leqslant \quad \Big(\max_{\widehat{\mathcal{V}}} \rho\Big)\, \delta n \, \delta A \,, \end{aligned}$$

where $\mathbf{g}(\mathbf{x}_0+)$ and $\mathbf{g}(\mathbf{x}_0-)$ are, respectively, the values of $\mathbf{g}$ on the outside and inside of the surface $\mathcal{S}$ at $\mathbf{x}_0$. On applying Gauss' Flux Theorem (9.6) to $\widehat{\mathcal{S}}$ and factoring δA we have that

$$\big(\mathbf{n} . \mathbf{g}(\mathbf{x}_0+) - \mathbf{n} . \mathbf{g}(\mathbf{x}_0-)\big) = O(\delta n) \,.$$

Thus, on taking the limit $\delta n \to 0$, we conclude that

$$\Big[\mathbf{n} . \mathbf{g}\Big]_{\mathbf{x}_0-}^{\mathbf{x}_0+} = 0 \,, \tag{9.9}$$

i.e. that the normal component of the gravitational field is continuous across a surface of discontinuity in ρ. Moreover, since $\mathbf{g} = -\boldsymbol{\nabla}\varphi$ (see (9.2a)), this condition can also be expressed as

$$\Big[\mathbf{n} . \boldsymbol{\nabla}\varphi\Big]_{\mathbf{x}_0-}^{\mathbf{x}_0+} = \left[\frac{\partial \varphi}{\partial n}\right]_{\mathbf{x}_0-}^{\mathbf{x}_0+} = 0 \,, \tag{9.10}$$

i.e. the normal component of $\boldsymbol{\nabla}\varphi$ is continuous across a surface of discontinuity in ρ.

Further, suppose $\mathcal{C}$ is a straight line between $\mathbf{x}_a$ and $\mathbf{x}_b$, where the two ends are on either side of $\mathcal{S}$. Then using (3.13)

$$\varphi(\mathbf{x}_a) - \varphi(\mathbf{x}_b) = \int_{\mathcal{C}} \boldsymbol{\nabla}\varphi . \mathrm{d}\mathbf{x} \leqslant |\mathbf{x}_b - \mathbf{x}_a| \, \max_{\mathcal{C}} |\boldsymbol{\nabla}\varphi| \,.$$

If we let $\mathbf{x}_a \to \mathbf{x}_0-$ and $\mathbf{x}_b \to \mathbf{x}_0+$ it follows, on the assumption that $|\boldsymbol{\nabla}\varphi|$ is bounded, that

$$\Big[\varphi\Big]_{\mathbf{x}_0-}^{\mathbf{x}_0+} = \varphi(\mathbf{x}_0+) - \varphi(\mathbf{x}_0-) = 0 \,, \tag{9.11}$$

i.e. that φ is continuous across a surface of discontinuity in ρ.

9.1.4 Example

Obtain the gravitational field due to sphere of radius a, uniform density ρ_0, and total mass

$$M = \iiint\limits_{\text{sphere}} \rho_0 \, \mathrm{d}\tau = \tfrac{4}{3}\pi a^3 \rho_0 \,.$$

Without loss of generality we assume that the origin of co-ordinates is at the centre of the sphere.

Using Poisson's Equation. From (9.7b)

$$\nabla^2\varphi = \begin{cases} 4\pi G\rho_0 & \text{for} \quad r < a \\ 0 & \text{for} \quad r > a \end{cases},$$

while from (9.10) and (9.11), φ and $\frac{\partial\varphi}{\partial n}$ will be continuous at the discontinuity in density at $r = a$. Further, from symmetry considerations we anticipate that φ will be a function only of $r = |\mathbf{x}|$, i.e. $\varphi \equiv \varphi(r)$. Then from the expression for the Laplacian in spherical polar co-ordinates (7.15b), for $r < a$ we have that

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\varphi}{\mathrm{d}r}\right) = 4\pi G\rho_0\,,$$

and hence on integrating

$$r^2\frac{\mathrm{d}\varphi}{\mathrm{d}r} = \tfrac{4}{3}\pi G\rho_0 r^3 + \beta_{\mathrm{i}}\,,$$

and

$$\varphi = \tfrac{2}{3}\pi G\rho_0 r^2 + \alpha_{\mathrm{i}} - \frac{\beta_{\mathrm{i}}}{r}\,,$$

where α_{i} and β_{i} are constants. Since there is no point mass at the origin, φ is bounded there, and hence $\beta_{\mathrm{i}} = 0$. Similarly, for $r > a$ we deduce that (see also (8.6))

$$\varphi = \alpha_{\mathrm{o}} - \frac{\beta_{\mathrm{o}}}{r}\,,$$

where α_{o} and β_{o} are constants. If φ and $\frac{\partial\varphi}{\partial n} = \frac{\partial\varphi}{\partial r}$ are to be continuous at $r = a$ it also follows that

$$\tfrac{2}{3}\pi G\rho_0 a^2 + \alpha_{\mathrm{i}} = \alpha_{\mathrm{o}} - \frac{\beta_{\mathrm{o}}}{a}$$

and

$$\tfrac{4}{3}\pi G\rho_0 a = \frac{\beta_{\mathrm{o}}}{a^2}\,,$$

respectively, i.e. that

$$\beta_{\mathrm{o}} = \tfrac{4}{3}\pi G\rho_0 a^3\,, \qquad \alpha_{\mathrm{i}} = \alpha_{\mathrm{o}} - 2\pi G\rho_0 a^2\,.$$

It is usual to normalise the gravitational potential so that $\varphi \to 0$ as $r \to \infty$, in which case $\alpha_{\mathrm{o}} = 0$,

$$\varphi = -\frac{4\pi G\rho_0 a^3}{3r}\,, \qquad \mathbf{g} = -\boldsymbol{\nabla}\varphi = -\frac{4\pi G\rho_0 a^3}{3r^2}\,\widehat{\mathbf{x}} = -\frac{GM}{r^2}\,\widehat{\mathbf{x}} \quad \text{for} \quad r \geqslant a\,,$$

and

$$\varphi = \tfrac{2}{3}\pi G\rho_0(r^2 - 3a^2)\,, \qquad \mathbf{g} = -\boldsymbol{\nabla}\varphi = -\tfrac{4}{3}\pi G\rho_0 r\,\widehat{\mathbf{x}} = -\frac{GMr}{a^3}\,\widehat{\mathbf{x}} \quad \text{for} \quad r \leqslant a\,.$$

Remarks. Outside the sphere, the gravitational acceleration is equal to that of a particle of mass M. Within the sphere, the gravitational acceleration increases linearly with distance from the
23/99 centre.

23/00 *Using Gauss' Flux Theorem.* From symmetry consideration we anticipate that $\mathbf{g}$ will be a function only of $r = |\mathbf{x}|$, and will be parallel to $\widehat{\mathbf{x}}$, i.e. $\mathbf{g} = g(r)\widehat{\mathbf{x}}$. Suppose $\mathcal{S}$ is a spherical surface of radius R. Then

$$\iint_{\mathcal{S}} \mathbf{g}.\,\mathrm{d}\mathbf{S} = \iint_{\mathcal{S}} g(R)\,\widehat{\mathbf{x}}.\,\mathrm{d}\mathbf{S} = g(R)\,4\pi R^2\,.$$

Further,

$$\text{if} \quad R \geqslant a \qquad \text{then} \qquad \iiint_{\mathcal{V}} \rho\,\mathrm{d}\tau = \tfrac{4}{3}\pi a^3\rho_0 = M\,,$$

$$\text{if} \quad R \leqslant a \qquad \text{then} \qquad \iiint_{\mathcal{V}} \rho\,\mathrm{d}\tau = \tfrac{4}{3}\pi R^3\rho_0 = MR^3/a^3\,.$$

Hence from Gauss' flux theorem (9.6), and substituting r for R,

$$g(r) = \begin{cases} -GM/r^2 & r \geqslant a \\ -GMr/a^3 & r \leqslant a \end{cases}. \tag{9.12}$$

9.2 Electrostatics

Similar results hold for the electrostatic force. This is because the *electric field* $\mathbf{E}$ due to a point *charge* q_1 at $\mathbf{x}_1$ is also given by an inverse square law, cf. (9.1),

$$\mathbf{E} = \frac{q_1}{4\pi\varepsilon_0}\,\frac{\mathbf{x}-\mathbf{x}_1}{|\mathbf{x}-\mathbf{x}_1|^3}\,, \tag{9.13}$$

where ε_0 is the *permittivity* of free space. Thus we can read off the equivalent results for electrostatics by means of the transformations

$$\mathbf{g} \to \mathbf{E}\,, \quad m_j \to q_j\,, \quad G \to -\frac{1}{4\pi\varepsilon_0}\,. \tag{9.14}$$

For instance, Gauss' flux theorem (9.6) becomes

$$\iint\limits_{\mathcal{S}} \mathbf{E}.\,\mathrm{d}\mathbf{S} = \frac{1}{\varepsilon_0}\iiint\limits_{\mathcal{V}} \rho\,\mathrm{d}\tau\,, \tag{9.15a}$$

where ρ is the now the *charge density*. From (9.7a), the differential form of this equation is

$$\boldsymbol{\nabla}.\,\mathbf{E} = \frac{\rho}{\varepsilon_0}\,, \tag{9.15b}$$

which is one of Maxwell's equations. Further, if we introduce the *electric potential* φ, where (cf. (9.2a)),

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi\,, \tag{9.15c}$$

then from (9.7b), (9.8) and (9.14)

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0}\,, \tag{9.15d}$$

with solution in unbounded space

$$\varphi(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0}\iiint\limits_{\mathcal{V}} \frac{\rho(\boldsymbol{\xi})}{|\mathbf{x}-\boldsymbol{\xi}|}\,\mathrm{d}\tau_{\boldsymbol{\xi}}\,. \tag{9.15e}$$

9.3 The Solution to Poisson's Equation in $\mathbb{R}^3$

We derived the solution to Poisson's equation in unbounded space, i.e. (9.8) or equivalently (9.15e), as the continuum limit of a sum over point sources. It is also possible to derive it using Green's second formula (8.12b). Specifically we aim to find the solution to

$$\nabla^2\varphi = \rho \text{ in } \mathbb{R}^3\,, \text{ with } \varphi \to 0 \text{ as } |\mathbf{x}| \to \infty\,, \tag{9.16}$$

where for simplicity we will assume that $\rho(\mathbf{x}) = 0$ for $|\mathbf{x}| > R_\rho$ for some real R_ρ.

We begin by supposing that $\mathcal{V}$ is the volume between concentric spheres $\mathcal{S}_\delta$ and $\mathcal{S}_R$ of radii δ and $R > R_\rho$, respectively. Suppose that the spheres are centred on some arbitrary point $\boldsymbol{\xi} \in \mathbb{R}^3$, and let

$$\psi = \frac{1}{|\mathbf{x}-\boldsymbol{\xi}|} = \frac{1}{r}\,, \tag{9.17a}$$

where $r = |\mathbf{x}-\boldsymbol{\xi}|$. Then, since $\boldsymbol{\xi} \notin \mathcal{V}$, it follows from (5.23) that

$$\nabla^2\psi = 0 \text{ in } \mathcal{V}\,. \tag{9.17b}$$

Also, for $\mathbf{x}$ on $\mathcal{S}_\delta$ and for $r = |\mathbf{x}-\boldsymbol{\xi}|$,

$$\psi = \frac{1}{\delta} \quad\text{and}\quad \frac{\partial\psi}{\partial n} = -\frac{\partial\psi}{\partial r} = \frac{1}{\delta^2}\,, \tag{9.18a}$$

where $\mathbf{n}$ is taken to be the normal pointing outward from $\mathcal{V}$. Similarly, for $\mathbf{x}$ on $\mathcal{S}_R$ and for the normal pointing outward from $\mathcal{V}$,

$$\psi = \frac{1}{R} \quad \text{and} \quad \frac{\partial \psi}{\partial n} = \frac{\partial \psi}{\partial r} = -\frac{1}{R^2}\,. \tag{9.18b}$$

Then from Green's second formula (8.12b), (9.16), (9.17a) and (9.17b), we have that

$$-\iiint\limits_{\mathcal{V}} \frac{\rho(\mathbf{x})}{|\mathbf{x}-\boldsymbol{\xi}|}\,\mathrm{d}\tau_{\mathbf{x}} = \iint\limits_{\mathcal{S}_\delta+\mathcal{S}_R} \left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)\mathrm{d}S\,. \tag{9.19}$$

In order to complete the derivation, we need to show that the right-hand side is equal to $4\pi\varphi(\boldsymbol{\xi})$ in the dual limit $\delta \to 0$ and $R \to \infty$. There are four integrals to estimate.

1. $\displaystyle\iint\limits_{\mathcal{S}_\delta} \varphi\frac{\partial\psi}{\partial n}\mathrm{d}S$. *Assume* that φ is continuous (possible to verify *a posteriori*), then for given $\varepsilon > 0$ there exists $\Delta > 0$ such that

$$|\varphi(\mathbf{x}) - \varphi(\boldsymbol{\xi})| < \varepsilon \quad \text{for} \quad |\mathbf{x}-\boldsymbol{\xi}| < \Delta\,.$$

Take $\delta < \Delta$, then using (9.18a)

$$\begin{aligned}
\left|\iint\limits_{\mathcal{S}_\delta} \varphi\frac{\partial\psi}{\partial n}\,\mathrm{d}S - 4\pi\varphi(\boldsymbol{\xi})\right| &= \left|\frac{1}{\delta^2}\iint\limits_{\mathcal{S}_\delta}\varphi(\mathbf{x})\,\mathrm{d}S - \frac{1}{\delta^2}\iint\limits_{\mathcal{S}_\delta}\varphi(\boldsymbol{\xi})\,\mathrm{d}S\right| \\
&\leqslant \frac{1}{\delta^2}\iint\limits_{\mathcal{S}_\delta}|\varphi(\mathbf{x})-\varphi(\boldsymbol{\xi})|\,\mathrm{d}S \\
&\leqslant \frac{1}{\delta^2}\iint\limits_{\mathcal{S}_\delta}\varepsilon\,\mathrm{d}S \\
&\leqslant 4\pi\varepsilon\,.
\end{aligned} \tag{9.20a}$$

2. $\displaystyle\iint\limits_{\mathcal{S}_\delta} \psi\frac{\partial\varphi}{\partial n}\mathrm{d}S$. *Assume* that $\boldsymbol{\nabla}\phi$ is bounded (possible to verify *a posteriori*), then using (9.18a)

$$\begin{aligned}
\left|\iint\limits_{\mathcal{S}_\delta} \psi\frac{\partial\varphi}{\partial n}\,\mathrm{d}S\right| &= \left|\iint\limits_{\mathcal{S}_\delta}\frac{1}{\delta}\,\frac{\partial\varphi}{\partial n}\,\mathrm{d}S\right| \\
&\leqslant \frac{1}{\delta}\iint\limits_{\mathcal{S}_\delta}\max_{\mathcal{S}_\delta}\left|\frac{\partial\varphi}{\partial n}\right|\mathrm{d}S \\
&\leqslant \max_{\mathcal{S}_\delta}\left|\frac{\partial\varphi}{\partial n}\right|\,4\pi\delta\,.
\end{aligned} \tag{9.20b}$$

3. $\displaystyle\iint\limits_{\mathcal{S}_R} \varphi \frac{\partial \psi}{\partial n} \mathrm{d}S$. *Assume* that there exists some $R_m > 0$ such that for $|\mathbf{x} - \boldsymbol{\xi}| > R_m$

$$|\varphi(\mathbf{x})| \leqslant \frac{\kappa}{|\mathbf{x} - \boldsymbol{\xi}|}\,, \qquad |\boldsymbol{\nabla}\varphi(\mathbf{x})| \leqslant \frac{\lambda}{|\mathbf{x} - \boldsymbol{\xi}|^2}\,, \tag{9.20c}$$

where κ and λ are real positive constants (verify *a posteriori*). Then if $R > R_m$ it follows from using (9.18b) that

$$\begin{aligned} \left| \iint\limits_{\mathcal{S}_R} \varphi \frac{\partial \psi}{\partial n}\, \mathrm{d}S \right| &\leqslant \frac{1}{R^2} \iint\limits_{\mathcal{S}_R} |\varphi|\, \mathrm{d}S \\ &\leqslant \frac{1}{R^2} \iint\limits_{\mathcal{S}_R} \frac{\kappa}{R}\, \mathrm{d}S \\ &\leqslant \frac{4\pi\kappa}{R}\,. \end{aligned} \tag{9.20d}$$

4. $\displaystyle\iint\limits_{\mathcal{S}_R} \psi \frac{\partial \varphi}{\partial n} \mathrm{d}S$. Similarly, using (9.18b) and (9.20c),

$$\begin{aligned} \left| \iint\limits_{\mathcal{S}_R} \psi \frac{\partial \varphi}{\partial n}\, \mathrm{d}S \right| &\leqslant \frac{1}{R} \iint\limits_{\mathcal{S}_R} \left| \frac{\partial \varphi}{\partial n} \right| \mathrm{d}S \\ &\leqslant \frac{1}{R} \iint\limits_{\mathcal{S}_R} \frac{\lambda}{R^2}\, \mathrm{d}S \\ &\leqslant \frac{4\pi\lambda}{R}\,. \end{aligned} \tag{9.20e}$$

Now let $\varepsilon \to 0$, $\delta \to 0$ and $R \to \infty$. Then it follows from (9.19), (9.20a), (9.20b), (9.20d) and (9.20e), that

$$\left| 4\pi\varphi(\boldsymbol{\xi}) + \iiint\limits_{\mathcal{V}} \frac{\rho(\mathbf{x})}{|\mathbf{x} - \boldsymbol{\xi}|}\, \mathrm{d}\tau_{\mathbf{x}} \right| \to 0\,. \tag{9.21}$$

Hence the solution in $\mathbb{R}^3$ to

$$\nabla^2\varphi = \rho\,,$$

i.e. the potential generated by a source distribution $\rho(\mathbf{x})$ such as mass or charge, is, after applying the transformation $\mathbf{x} \leftrightarrow \boldsymbol{\xi}$ to (9.21),

$$\varphi(\mathbf{x}) = -\frac{1}{4\pi} \iiint\limits_{\mathcal{V}} \frac{\rho(\boldsymbol{\xi})}{|\mathbf{x} - \boldsymbol{\xi}|}\, \mathrm{d}\tau_{\boldsymbol{\xi}}\,. \tag{9.22}$$

Remark. After accounting for constants and sign changes this is the same result as was obtained earlier by taking the continuum limit — see (9.8) and (9.15e).

9.3.1 More on Monopoles, Dipoles, etc.

We will not explicitly verify from (9.22) that φ is continuous and $\boldsymbol{\nabla}\varphi$ is bounded. However, we will verify, as claimed in (9.20c), that $\varphi = O(|\mathbf{x}|^{-1})$ as $|\mathbf{x}| \to \infty$.

For $|\mathbf{x}| \gg |\boldsymbol{\xi}|$, it follows from Taylor's theorem (1.58a) and (1.62) that (*s.c.*)

$$\begin{aligned}\frac{1}{|\mathbf{x}-\boldsymbol{\xi}|} &= \frac{1}{|\mathbf{x}|} - \xi_i \frac{\partial}{\partial x_j}\left(\frac{1}{|\mathbf{x}|}\right) + \tfrac{1}{2}\xi_i\xi_j \frac{\partial^2}{\partial x_i \partial x_j}\left(\frac{1}{|\mathbf{x}|}\right) + \dots \\ &= \frac{1}{|\mathbf{x}|} + \frac{\xi_j x_j}{|\mathbf{x}|^3} + \frac{\xi_j \xi_k}{2|\mathbf{x}|^5}\left(3x_j x_k - |\mathbf{x}|^2\delta_{jk}\right) + \dots\,, \end{aligned} \tag{9.23}$$

where we have identified $-\boldsymbol{\xi}$ with $\mathbf{h}$ in (1.62).

Hence for $|\mathbf{x}| \gg R_\rho$ we have from (9.22) and (9.23) that

$$\begin{aligned}\varphi(\mathbf{x}) &= -\frac{1}{4\pi}\iiint\limits_{|\boldsymbol{\xi}|\leqslant R_\rho} \rho(\boldsymbol{\xi})\left(\frac{1}{|\mathbf{x}|} + \frac{1}{|\mathbf{x}|^3}\,\xi_i x_i + \frac{1}{2|\mathbf{x}|^5}\,\xi_i\xi_j(3x_i x_j - |\mathbf{x}|^2\delta_{ij}) + \dots\right) \mathrm{d}\tau_{\boldsymbol{\xi}}\,, \\ &= -\frac{Q}{4\pi|\mathbf{x}|} - \frac{x_i d_i}{4\pi|\mathbf{x}|^3} - \frac{(3x_i x_j - |\mathbf{x}|^2\delta_{ij})A_{ij}}{8\pi|\mathbf{x}|^5} + \dots\,, \end{aligned} \tag{9.24a}$$

where

$$Q = \iiint\limits_{|\boldsymbol{\xi}|\leqslant R_\rho} \rho(\boldsymbol{\xi})\,\mathrm{d}\tau_{\boldsymbol{\xi}}\,, \quad \mathbf{d} = \iiint\limits_{|\boldsymbol{\xi}|\leqslant R_\rho} \rho(\boldsymbol{\xi})\,\boldsymbol{\xi}\,\mathrm{d}\tau_{\boldsymbol{\xi}}\,, \tag{9.24b}$$

and

$$A_{ij} = \iiint\limits_{|\boldsymbol{\xi}|\leqslant R_\rho} \rho(\boldsymbol{\xi})\,\xi_i\xi_j\,\mathrm{d}\tau_{\boldsymbol{\xi}}\,, \tag{9.24c}$$

are the so-called *monopole strength*, *dipole strength* and *quadrupole strength* respectively. From a comparison of (9.24a) with (9.20c) we conclude after using (9.23) that for $|\mathbf{x}| \gg R_\rho$ and $|\mathbf{x}| \gg |\boldsymbol{\xi}|$, the *far-field* asymptotic behaviour of φ is as claimed.

By differentiating (9.24a), or by a similar application of Taylor's theorem (better), we conclude that for $|\mathbf{x}| \gg R_\rho$ and $|\mathbf{x}| \gg |\boldsymbol{\xi}|$, the *far-field* asymptotic behaviour of $\boldsymbol{\nabla}\varphi$ is as claimed, i.e. that

$$|\boldsymbol{\nabla}\varphi| = O\left(\frac{1}{|\mathbf{x}|^2}\right) \quad \text{as} \quad |\mathbf{x}| \to \infty\,.$$

Remarks.

- In an order of magnitude sense it follows from the definitions (9.24b) and (9.24c) that

  $$Q = O(\rho R_\rho^3)\,, \quad |\mathbf{d}| = O(\rho R_\rho^4) \quad \text{and} \quad |A_{ij}| = O(\rho R_\rho^5)\,.$$

  Hence at large distances $\mathbf{x}$ from a distributed mass/charge source, we conclude from (9.24a) that the dipole field is less than the monopole field by a factor of $O(R_\rho/|\mathbf{x}|)$, and the quadrupole field is less than the dipole field by a factor of $O(R_\rho/|\mathbf{x}|)$.

- Far from a distributed mass/charge source it follows from (9.24a) that the potential is, to leading order, just that of a point mass/charge of the same strength.

- If the total [mass/]charge of a distributed source is zero then, to leading order, the potential at large distances is that due to a dipole.

- If ρ is the [scaled] mass density of a body, then Q is the total mass of the body, $\mathbf{d}/Q$ is the centre of gravity of the body, and the quadrupole strength is related to the moments of inertia of the body. Hence if you can accurately measure the gravitational potential of the body (e.g. by satellite) you
24/99 can back out the moments of inertia of the body.

24/00

9.4 Heat

What is the equation that governs the flow of heat in a saucepan, an engine block, the earth's core, etc.?

Let $\mathbf{q}(\mathbf{x}, t)$ denote the *flux vector for heat flow*. Then the energy in the form of heat (molecular vibrations) crossing a small surface $\mathrm{d}\mathbf{S}$ in time δt is

$$(\mathbf{q}.\,\mathrm{d}\mathbf{S})\,\delta t\,,$$

and the heat flow out of a closed surface $\mathcal{S}$ in time δt is

$$\left(\iint_{\mathcal{S}} \mathbf{q}.\,\mathrm{d}\mathbf{S}\right)\delta t\,.$$

Let

$E(\mathbf{x}, t)$ denote the *internal energy* per unit mass of the solid,
$Q(\mathbf{x}, t)$ denote any *heat source* per unit time per unit volume of the solid,
$\rho(\mathbf{x}, t)$ denote the *mass density* of the solid (assumed constant here).

The flow of heat in/out of $\mathcal{S}$ must balance the change in internal energy and the heat source. Hence in time δt

$$\left(\iint\limits_{\mathcal{S}} \mathbf{q}.\,\mathrm{d}\mathbf{S}\right)\delta t = -\left(\frac{\mathrm{d}}{\mathrm{d}t}\iiint\limits_{\mathcal{V}} \rho E\,\mathrm{d}\tau\right)\delta t + \left(\iiint\limits_{\mathcal{V}} Q\,\mathrm{d}\tau\right)\delta t\,.$$

For 'slow' changes at constant pressure (1st and 2nd law of thermodynamics)

$$E(\mathbf{x}, t) = c_p\theta(\mathbf{x}, t)\,, \tag{9.25}$$

where θ is the *temperature* and c_p is the *specific heat* (assumed constant here). Hence using the divergence theorem (6.1), and exchanging the order of differentiation and integration,

$$\int\left(\boldsymbol{\nabla}.\,\mathbf{q} + \rho c_p\frac{\partial\theta}{\partial t} - Q\right)\mathrm{d}\tau = 0\,.$$

But this is true for any volume, hence

$$\rho c_p\frac{\partial\theta}{\partial t} = -\boldsymbol{\nabla}.\,\mathbf{q} + Q\,.$$

Experience tells us heat flows from hot to cold. The simplest empirical law relating heat flow to temperature gradient is *Fourier's law* (also known as *Fick's law*)

$$\mathbf{q} = -k\boldsymbol{\nabla}\theta\,, \tag{9.26}$$

where k is the *heat conductivity*. If k is constant then the partial differential equation governing the temperature is

$$\frac{\partial\theta}{\partial t} = \nu\nabla^2\theta + \frac{Q}{\rho c_p} \tag{9.27}$$

where $\nu = k/(\rho c_p)$ is the *diffusivity* (or coefficient of diffusion).

Special Cases.

Diffusion Equation. If there is no heat source then $Q = 0$, and the governing equation becomes the *diffusion equation*

$$\frac{\partial\theta}{\partial t} = \nu\nabla^2\theta\,. \tag{9.28}$$

Poisson's Equation. If the system has reached a steady state (i.e. $\partial_t \equiv 0$), then with $f(\mathbf{x}) = -Q(\mathbf{x})/k$ we recover *Poisson's equation*

$$\nabla^2 \theta = f\,. \tag{9.29}$$

Laplace's Equation. If the system has reached a steady state and there are no heat sources then the temperature is governed by *Laplace's equation*

$$\nabla^2 \theta = 0\,. \tag{9.30}$$

In this case we deduce from § 8.8 that the maximum and minimum temperatures occur on the boundary, i.e. the variations of temperature on the boundary are smoothed out in the interior.