MATH20901 Multivariable Calculus

Richard Porter
University of Bristol
2018

Course Information

- Prerequistes: Calculus 1 (and Linear Algebra and Geometry, Analysis 1)
- The course develops multivariable calculus from Calculus 1. The main focus of the course is on developing differential vector calculus, tools for changing coordinate systems and major theorems of integral calculus for functions of more than one variable.

  This unit is central to many branches of pure and applied mathematics. For example, in applied mathematics vector calculus is an integral part of describing field theories that model physical processes and dealing with the equations that arise.

  It is used in 2nd year **Applied Partial Differential Equations** and in year 3 **Fluid Dynamics**, **Quantum Mechanics**, **Mathematical Methods**, and **Modern Mathematical Biology**.
- Lecturer: Dr. Richard Porter, Room SM2.7
- Web:

  `https://people.maths.bris.ac.uk/~marp/mvcalc`

  Notes may contain extra sections for interest or additional information. Problem sheets, solutions, homework feedback forms, problems class sheets, past exam papers, video tutorials.
- Email: `richard.porter@bris.ac.uk`
- Books: Lots of books on multivariable/vector calculus. Jerrold E. Marsden & Anthony J. Tromba, "Vector Calculus", ed. 5 , W. H. Freeman and Company, 2003
- Maths Café: TBA
- Office Hours: Tuesday 9-10am.
- Homework set weekly from 5 problems sheets.
- Timetabled problems classes/exercise classes: unseen problems/some from the problem sheets/and as many as possible from past exam papers.
- Exam: Jan 90 mins. 2 compulsory questions. No calculators.

1 A review of differential calculus for functions of more than one variable

Revision and extension of results from Calculus 1.

1.1 General maps from $\mathbb{R}^m$ to $\mathbb{R}^n$

Let $\mathbf{x} \in \mathbb{R}^m = (x_1, x_2, \ldots, x_m)$[1].

Often in 2D write $\mathbf{x} \equiv (x, y)$ or in 3D $\mathbf{x} \equiv (x, y, z)$.

Defn: A **scalar map** or **scalar function**, f, say is defined by $f : \mathbb{R}^m \to \mathbb{R}$ s.t. $\mathbf{x} \to f(\mathbf{x})$. We write it as $f(\mathbf{x})$.

Defn: A general map, or **vector function**, say $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ s.t. $\mathbf{x} \to \mathbf{F}(\mathbf{x})$ is defined as

$$\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_n(\mathbf{x})).$$

and the components are scalar maps denoted by $F_i : \mathbb{R}^m \to \mathbb{R}$ $(i = 1, \ldots, n)$.

Defn: A map $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ is **linear** if $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, and $\lambda, \mu \in \mathbb{R}$, $\mathbf{F}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathbf{F}(\mathbf{x}) + \mu\mathbf{F}(\mathbf{y})$.

Proposition: A map $\mathbf{F}$ is linear iff $\exists$ a matrix $A \in \mathbb{R}^{n\times m}$ s.t. $\mathbf{F} = A\mathbf{x}$.

E.g. 1.1: $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^2$ s.t. $(x_1, x_2, x_3) \to (x_3 - x_1, x_2 + x_1)$. Then

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A\mathbf{x}$$

is a linear map, since $A(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda A\mathbf{x} + \mu A\mathbf{y}$.

E.g. 1.2: $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ s.t. $(x_1, x_2) \to (x_2 x_1, \mathrm{e}^{x_2})$. (Not a linear map since, for example, $\mathbf{F}(2\mathbf{x}) \neq 2\mathbf{F}(\mathbf{x})$)

1.2 The derivative of a map

Defn: The derivative of the map $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ is the $n \times m$ matrix $\mathbf{F}'(\mathbf{x})$ such that the i, jth element is

$$\{\mathbf{F}'(\mathbf{x})\}_{ij} = \frac{\partial F_i}{\partial x_j}.$$

[1] Although written on the page as a row vector, in computations, vectors are actually arranged as column vectors unless indicated by a T for transpose.

For scalar functions of single variables, the derivative $f'(x_0)$ is defined to be precisely the function such that the line formed by

$$f(x_0) + (x - x_0)f'(x_0)$$

is tangent at $x = x_0$ to the curve $f(x)$.

For vector functions of multiple variables, $\mathbf{F}'(\mathbf{x}_0)$ is defined to be precisely the matrix such that

$$\mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \tag{1}$$

defines the tangent plane at $\mathbf{x} = \mathbf{x}_0$ to the hypersurface formed by $\mathbf{F}(\mathbf{x})$.

E.g. 1.3: If $\mathbf{F} = A\mathbf{x}$ (a linear map) then $\mathbf{F}'(\mathbf{x}) = A$.

Proof: We have

$$F_i = \sum_{k=1}^{m} A_{ik} x_k$$

where A_{ij} is the i, jth element of A and so

$$\{\mathbf{F}'(\mathbf{x})\}_{ij} = \frac{\partial F_i}{\partial x_j} = \sum_{k=1}^{m} A_{ik} \frac{\partial x_k}{\partial x_j} = A_{ij}$$

since $\partial x_k / \partial x_j = 1$ if $j = k$ else zero.

E.g. 1.4: With $\mathbf{F}(\mathbf{x}) = (x_2 x_1, \mathrm{e}^{x_2})$ we have

$$\mathbf{F}'(\mathbf{x}) = \begin{pmatrix} x_2 & x_1 \\ 0 & \mathrm{e}^{x_2} \end{pmatrix}.$$

Defn: The matrix $\mathbf{F}'(\mathbf{x})$ with elements $\partial F_i / \partial x_j$ is called the **Jacobian matrix**.

1.3 The gradient of a function

Defn: The **gradient** of a scalar function $f : \mathbb{R}^m \to \mathbb{R}$ (i.e. $f(\mathbf{x})$) is denoted

$$\boldsymbol{\nabla} f \equiv (\partial f / \partial x_1, \partial f / \partial x_2, \ldots, \partial f / \partial x_m).$$

Note: The rows of the Jacobian matrix are formed by gradients of the components of $\mathbf{F}$, viz

$$\mathbf{F}'(\mathbf{x}) = \begin{pmatrix} (\boldsymbol{\nabla} F_1)^T \\ (\boldsymbol{\nabla} F_2)^T \\ \vdots \\ (\boldsymbol{\nabla} F_n)^T \end{pmatrix}.$$

(More on this later.)

1.4 The directional derivative

Defn: The **directional derivative** of $\mathbf{F}$ at $\mathbf{x}_0$ along $\mathbf{v}$ (such that $|\mathbf{v}| = 1$) is a vector in $\mathbb{R}^n$ given by

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{x}_0) = \left(\frac{dF_1(\mathbf{x}_0 + t\mathbf{v})}{dt}, \ldots, \frac{dF_n(\mathbf{x}_0 + t\mathbf{v})}{dt} \right)_{t=0} \equiv \left. \frac{d\mathbf{F}(\mathbf{x}_0 + t\mathbf{v})}{dt} \right|_{t=0}.$$

It measures the *rate of change of* $\mathbf{F}$ *in the direction of* $\mathbf{v}$ and it is formulated in terms of ordinary 1D derivatives.

Note: Can be shown that

$$D_{\mathbf{v}}\mathbf{F}(\mathbf{x}_0) = \mathbf{F}'(\mathbf{x}_0)\mathbf{v}.$$

Proof: (informal)

$$\left. \frac{d\mathbf{F}(\mathbf{x}_0 + t\mathbf{v})}{dt} \right|_{t=0} = \lim_{t \to 0} \left(\frac{\mathbf{F}(\mathbf{x}_0 + t\mathbf{v}) - \mathbf{F}(\mathbf{x}_0)}{t} \right) = \lim_{t \to 0} \left(\frac{\mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x}_0 + t\mathbf{v} - \mathbf{x}_0) - \mathbf{F}(\mathbf{x}_0)}{t} \right)$$

which gives the result. In above, we replace $\mathbf{F}$ by equation of tangent plane (1) which coincides in limit $t \to 0$.

Note: If $|\mathbf{v}| \neq 1$ then redefine $\mathbf{v}$ by $\mathbf{v}/|\mathbf{v}|$ where $|\mathbf{v}| = \sqrt{v_1^2 + v_2^2 + \ldots + v_m^2}$.

Note: If $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^m$ with $|\mathbf{v}| = 1$ and $f : \mathbb{R}^m \to \mathbb{R}$ is a scalar function then

$$D_{\mathbf{v}} f = (\boldsymbol{\nabla} f)^T \mathbf{v} \equiv \mathbf{v} \cdot \boldsymbol{\nabla} f. \tag{2}$$

1.5 Operations on maps and their derivatives

1. (Addition of maps) Let $\mathbf{F}, \mathbf{G}$ be maps from $\mathbb{R}^m \to \mathbb{R}^n$. Then if $\mathbf{H} : \mathbb{R}^m \to \mathbb{R}^n$ is the new function defined as

   $$\mathbf{H}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) + \mathbf{G}(\mathbf{x})$$

   it follows

   $$\mathbf{H}'(\mathbf{x}) = \mathbf{F}'(\mathbf{x}) + \mathbf{G}'(\mathbf{x}).$$

   Proof: (simple)

   $$\frac{\partial H_i}{\partial x_j}(\mathbf{x}) = \frac{\partial F_i}{\partial x_j}(\mathbf{x}) + \frac{\partial G_i}{\partial x_j}(\mathbf{x}).$$

2. (Product of maps) Let $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ and $f : \mathbb{R}^m \to \mathbb{R}$, then if $\mathbf{H} : \mathbb{R}^m \to \mathbb{R}^n$ is defined by

   $$\mathbf{H}(\mathbf{x}) = f(\mathbf{x})\mathbf{F}(\mathbf{x})$$

   it follows that $\mathbf{H}'(\mathbf{x})$ is the matrix whose $i, j$the element is defined by

   $$\frac{\partial H_i}{\partial x_j}(\mathbf{x}) = \frac{\partial f}{\partial x_j}(\mathbf{x}) F_i(\mathbf{x}) + f(\mathbf{x}) \frac{\partial F_i}{\partial x_j}(\mathbf{x})$$

   using **product rule** for differentiation. No simple representation for the result using standard linear algebra.

3. (Composition of maps) If $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{G} : \mathbb{R}^n \to \mathbb{R}^p$, then if we define $\mathbf{H} : \mathbb{R}^m \to \mathbb{R}^p$ by

$$\mathbf{H}(\mathbf{x}) = (\mathbf{G} \circ \mathbf{F})(\mathbf{x}) = \mathbf{G}(\mathbf{F}(\mathbf{x}))$$

it follows that

$$\mathbf{H}'(\mathbf{x}) = \mathbf{G}'(\mathbf{F}(\mathbf{x}))\,\mathbf{F}'(\mathbf{x}) \tag{3}$$

where the right-hand side denotes the product of the $p \times n$ matrix $\mathbf{G}'(\mathbf{F}(\mathbf{x}))$ with the $n \times m$ matrix $\mathbf{F}'(\mathbf{x})$.

Proof: From the definition, $H_i = G_i(F_1(x_1, \ldots, x_m), F_2(x_1, \ldots, x_m), \ldots, F_n(x_1, \ldots, x_m))$. So

$$\begin{aligned} \{\mathbf{H}'(\mathbf{x})\}_{ij} &= \frac{\partial H_i}{\partial x_j}(\mathbf{x}) = \frac{\partial G_i}{\partial x_1}\frac{\partial F_i}{\partial x_j} + \frac{\partial G_i}{\partial x_2}\frac{\partial F_2}{\partial x_j} + \ldots + \frac{\partial G_i}{\partial x_n}\frac{\partial F_n}{\partial x_j} \\ &= \sum_{k=1}^{n} \frac{\partial G_i}{\partial x_k}(\mathbf{F}(\mathbf{x}))\frac{\partial F_k}{\partial x_j}(\mathbf{x}) \end{aligned}$$

using the chain rule. This summation can be interpreted as the ith row of $\mathbf{G}'(\mathbf{F}(\mathbf{x}))$ multiplied by the jth column of $\mathbf{F}'(\mathbf{x})$ and this gives the result.

Note: See the Appendix for a revision of the chain rule and how it applies to multivariable functions.

1.6 Inverse maps

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{G} = \mathbf{F}^{-1}$ be the **inverse** map such that

$$(\mathbf{F}^{-1} \circ \mathbf{F})(\mathbf{x}) = \mathbf{x} \tag{4}$$

Differentiating, applying (3) to (4) and using the fact that $\mathbf{x} = I\mathbf{x}$ where I is the $n \times n$ identity matrix we have

$$(\mathbf{F}^{-1})'(\mathbf{F}(\mathbf{x}))\mathbf{F}'(\mathbf{x}) = I.$$

Thus

$$(\mathbf{F}^{-1})'(\mathbf{F}(\mathbf{x})) = (\mathbf{F}')^{-1}(\mathbf{x}). \tag{5}$$

In other words "*the derivative of the inverse is equal to the inverse of the derivative*".

Note: For scalar maps, we recognise this statement as

$$\frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1} = \frac{1}{\dfrac{dy}{dx}}.$$

and (5) generalises this to functions of more than one variable.

E.g. 1.5: (mapping 2D Cartesian to plane polar coordinates)

Let $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ s.t. $(r, \theta) \to (r\cos\theta, r\sin\theta) \equiv (x, y)$.

This means that

$$\mathbf{F}'(r,\theta) = \begin{pmatrix} \partial x/\partial r & \partial x/\partial \theta \\ \partial y/\partial r & \partial y/\partial \theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Taking inverses

$$(\mathbf{F}'(r,\theta))^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix}.$$

Now consider the inverse map $\mathbf{F}^{-1} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ s.t. $(x,y) \to (\sqrt{x^2+y^2}, \tan^{-1}(y/x)) \equiv (r,\theta)$. Then

$$(\mathbf{F}^{-1})'(x,y) = \begin{pmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial \theta/\partial x & \partial \theta/\partial y \end{pmatrix} = \begin{pmatrix} x/\sqrt{x^2+y^2} & y/\sqrt{x^2+y^2} \\ -y/(x^2+y^2) & x/(x^2+y^2) \end{pmatrix}.$$

Finally,

$$(\mathbf{F}^{-1})'(\mathbf{F}(r,\theta)) = \begin{pmatrix} r\cos\theta/r & r\sin\theta/r \\ -r\sin\theta/r^2 & r\cos\theta/r^2 \end{pmatrix}$$

which is the same as before.

1.7 Solving equations

Question: Given a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, is there always an inverse function $\mathbf{G} \equiv \mathbf{F}^{-1}$, which satisfies

$$(\mathbf{G} \circ \mathbf{F})(\mathbf{x}) = \mathbf{x} \ ?$$

The same question can be stated in terms of a solution to a nonlinear system of equations. Namely, let

$$\mathbf{F}(\mathbf{x}) = \mathbf{y} \tag{6}$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Or, in full,

$$\begin{aligned} F_1(x_1, \ldots, x_n) &= y_1 \\ \vdots \qquad & \quad \vdots \\ F_n(x_1, \ldots, x_n) &= y_n. \end{aligned}$$

Then, given $\mathbf{y}$, is there a $\mathbf{x}$ such that (6) is solved. If so, then $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{y})$.

1.7.1 Inverse function theorem

Let $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$, with $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^n$ such that

$$\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0).$$

If the Jacobian matrix $\mathbf{F}'(\mathbf{x}_0)$ is invertible, then (6) can be solved *uniquely* as

$$\mathbf{x} = \mathbf{F}^{-1}(\mathbf{y}),$$

for $\mathbf{y}$ in the *neighbourhood* of $\mathbf{y}_0$.

Note: A matrix is invertible if and only if its determinant is non-zero. The determinant of the Jacobian matrix $\mathbf{F}'$ is often written as

$$J_{\mathbf{F}}(\mathbf{x}_0) \equiv \left.\frac{\partial(F_1,\ldots,F_n)}{\partial(x_1,\ldots,x_n)}\right|_{\mathbf{x}=\mathbf{x}_0} \tag{7}$$

and called the **Jacobian determinant**.

Proof: (informal, but instructive)

For $\mathbf{x}$ close to $\mathbf{x}_0$, we can use (1) to locally approximate $\mathbf{F}(\mathbf{x})$ so that

$$\mathbf{y} \approx \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)$$

which means

$$\mathbf{x} \approx \mathbf{x}_0 + (\mathbf{F}'(\mathbf{x}_0))^{-1}(\mathbf{y}-\mathbf{y}_0)$$

since $\mathbf{y}_0 = \mathbf{F}(\mathbf{x}_0)$ but relies on the existence of the inverse of the Jacobian. I.e. given $\mathbf{y}$, $\mathbf{x}$ can be determined.

Note: The theorem tells us nothing about what happens if the inverse does not exist.

E.g. 1.6: Consider the system of equations

$$\frac{x^2+y^2}{x} = u, \quad \sin x + \cos y = v.$$

Q: Given (u,v), we want to solve for (x,y). Near which points is this guaranteed to define a unique function ?

A: We define $\mathbf{F} : \mathbb{R}^2\backslash\{0\} \to \mathbb{R}^2$ s.t.

$$\mathbf{y} \equiv \mathbf{F}(\mathbf{x}) = \left(\frac{x^2+y^2}{x}, \sin x + \cos y\right)$$

(so that $\mathbf{y} \equiv (u,v)$ and $\mathbf{x} = (x,y)$.)

The Jacobian determinant is

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} (x^2-y^2)/x^2 & 2y/x \\ \cos x & -\sin y \end{vmatrix} = \frac{y^2-x^2}{x^2}\sin y - \frac{2y}{x}\cos x.$$

E.g. (i) near $\mathbf{x}_0 = (1,1)$ (where $\mathbf{y} = (2, \sin(1)+\cos(1))$) we can solve for $\mathbf{x}$ in a neighborhood of $\mathbf{x}_0$; E.g. (ii) near $\mathbf{x}_0 = (\pi/2, \pi/2)$ (where $\mathbf{y} = (\pi, 1)$) where $J_{\mathbf{F}} = 0$ we cannot say anything about existence or uniqueness of solutions.

1.7.2 Implicit function theorem

Similar to above. Consider an equation for $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$ in the form

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = 0 \tag{8}$$

where $\mathbf{F} : \mathbb{R}^{m+n} \to \mathbb{R}^n$.

Note: If $\mathbf{F}$ is linear in $\mathbf{y}$ then (8) can be written in the form $\mathbf{y} = \mathbf{G}(\mathbf{x})$ for some $\mathbf{G}$ and the inverse function theorem applies. We suppose that this is not the case.

Suppose that (8) is satisfied by the pair $\mathbf{x}_0$, $\mathbf{y}_0$ (i.e. $\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = 0$.) Then we can express solutions of this as $\mathbf{y} = \mathbf{y}(\mathbf{x})$ for $\mathbf{y} : \mathbb{R}^m \to \mathbb{R}^n$ in the neighbourhood of $\mathbf{y}_0$ provided the Jacobian determinant

$$\left.\frac{\partial(F_1, \ldots, F_n)}{\partial(y_1, \ldots, y_n)}\right|_{\mathbf{x}=\mathbf{x}_0, \mathbf{y}=\mathbf{y}_0} \tag{9}$$

is non-zero.

Proof: (informal but instructive)

The ith component of (8) is

$$F_i(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0$$

and we suppose that

$$y_1 = y_1(x_1, \ldots, x_m), \qquad y_2 = y_2(x_1, \ldots, x_m), \quad \ldots \quad y_n = y_n(x_1, \ldots, x_m).$$

Taking the partial derivative of F_i w.r.t x_k gives (by the chain rule)

$$\frac{\partial F_i}{\partial x_j} + \frac{\partial F_i}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \ldots \frac{\partial F_i}{\partial y_n}\frac{\partial y_n}{\partial x_j} = 0$$

and this can be interpreted as the matrix equation

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix} + \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial y_1} & \cdots & \frac{\partial F_n}{\partial y_n} \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_m} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

and therefore

$$\begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_m} \end{pmatrix} = - \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial y_1} & \cdots & \frac{\partial F_n}{\partial y_n} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}$$

and the existence of $\mathbf{y}'(\mathbf{x}_0)$ requires (9) holds.

If $\mathbf{y}'(\mathbf{x}_0)$ holds we can use (1) for points $\mathbf{x}$ close to $\mathbf{x}_0$ to write

$$\mathbf{y}(\mathbf{x}) \approx \mathbf{y}(\mathbf{x}_0) + \mathbf{y}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

showing the existence of solutions in the neighbourhood of $(\mathbf{x}_0, \mathbf{y}_0)$.

E.g. 1.7: Consider $f(x, y) = 0$ where $f(x, y) = x^2 + y^2 - 1$. This is satisfied by points (x_0, y_0) on the unit circle. If we try to express it as $y = y(x)$ we get into trouble since

$$y = \pm\sqrt{1 - x^2}$$

and there are two solutions. The implicit function theorem applied to this example requires the determinant of the 1×1 matrix

$$\frac{\partial f}{\partial y}$$

evaluated at (x_0, y_0) to be non-zero. This is $2y_0$ which is non-zero apart from at $y_0 = 0$. So we can express the solution $y = y(x)$ local to a point (x_0, y_0) provided $y_0 \neq 0$. Which is obvious in our case as if $y_0 > 0$ we are on the upper solution branch where $y = \sqrt{1-x^2}$ and vice versa.

1.8 Higher-order derivatives

Start with 2nd order.

Defn: For $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$,

$$\frac{\partial^2 F_i}{\partial x_k \partial x_j}(\mathbf{x}) = \frac{\partial}{\partial x_k}\left(\frac{\partial F_i}{\partial x_j}\right)(\mathbf{x}) = \frac{\partial}{\partial x_j}\left(\frac{\partial F_i}{\partial x_k}\right)(\mathbf{x}) = \frac{\partial^2 F_i}{\partial x_j \partial x_k}(\mathbf{x})$$

under normal circumstances. For example if $f(x, y) = x^3 - 3xy^2$,

$$f_{xx} \equiv \frac{\partial^2 f}{\partial x \partial x} = 6x, \qquad f_{xy} \equiv \frac{\partial^2 f}{\partial x \partial y} = -6y, \qquad f_{yx} = -6y, \qquad f_{yy} = -6x.$$

Note: Extended naturally to higher orders.

1.8.1 Taylor's theorem

Higher-order derivatives are useful in Taylor's theorem in dimension ≥ 2, allowing one to approximate functions of several variables near a point.

Recall that for a scalar function of a single variable,

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x-x_0)^2}{2!}f''(x_0) + \text{higher order terms (h.o.t.)}$$

How do we generalise to higher dimensions ? Well, it gets tricky. For a *scalar* function $f(x, y)$, or more than one (e.g. 2) variable

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + (f_x(x_0,y_0), f_y(x_0,y_0)) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \\ &\quad + \frac{1}{2}(x - x_0, y - y_0) \begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \text{h.o.t} \end{aligned}$$

with an obvious generalisation to functions of more than 2 variables.

The higher order terms are complicated and require some complex notation.

For vector functions, $\mathbf{F} = (F_1, \ldots, F_n)$ we can use the scalar result above for each scalar function component, F_i. I.e.

$$F_i(\mathbf{x}) = F_i(\mathbf{x}_0) + (\boldsymbol{\nabla} F_i(\mathbf{x}_0))^T(\mathbf{x} - \mathbf{x}_0) + \text{h.o.t of size like } |\mathbf{x} - \mathbf{x}_0|^2.$$

Stacking equations gives

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \text{h.o.t}$$

involving matrix/vector multiplications and hence

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

for $\mathbf{x}$ close to $\mathbf{x}_0$. This is equation (1) for the tangent plane at the point $\mathbf{x}_0$ on $\mathbf{F}$. We've used this approximation as the basis of earlier informal proofs.

2 Differential vector calculus

2.1 Linear algebra

Focus now on 3D, and adopt convention that position vector $\mathbf{r} = (x, y, z) \equiv (x_1, x_2, x_3) \in \mathbb{R}^3$ to describe equations pertaining to physical applications.

Notation: The Cartesian (unit) basis vectors in $\mathbb{R}^3$ are $\hat{\mathbf{x}} = (1, 0, 0) \equiv \mathbf{e}_1$, $\hat{\mathbf{y}} = (0, 1, 0) \equiv \mathbf{e}_2$ and $\hat{\mathbf{z}} = (0, 0, 1) \equiv \mathbf{e}_3$ such that $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} \equiv x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$.

Also use $r = |\mathbf{r}| = \sqrt{x_1^2 + x_2^2 + x_3^2} \equiv \sqrt{x^2 + y^2 + z^2}$ as the length of the position vector.

Defn: The **dot product** of two vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ is defined

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 \equiv \sum_{j=1}^{3} u_jv_j$$

Notation: (**Einstein summation convention**) Drop the $\sum_{j=1}^{3}$ in the above on the understanding that repeated suffices imply summation. I.e.

$$\mathbf{u} \cdot \mathbf{v} = u_jv_j$$

For e.g., $r = |\mathbf{r}| = \sqrt{\mathbf{r} \cdot \mathbf{r}} = \sqrt{x_i^2}$.

Defn: The **Kronecker delta**, δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

E.g. 2.1: (**Sampling property**) Another way of defining δ_{ij} is to be the set of elements for which

$$x_i = \delta_{ij}x_j, \qquad \text{for every } i$$

Note that $\delta_{ij}x_j = \sum_{j=1}^{3} \delta_{ij}x_j = x_i$ if and only if the defn above holds.

E.g. 2.2: $\delta_{ii} = 3$.

E.g. 2.3: (Orthonormality of basis vectors) $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

E.g. 2.4: $\mathbf{r} = x_j\mathbf{e}_j$ and taking dot product with $\mathbf{e}_i$ gives $x_i = \mathbf{r} \cdot \mathbf{e}_i$.

Defn: The **cross product** of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, vector given by

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \equiv (u_2 v_3 - v_2 u_3)\mathbf{e}_1 + (u_3 v_1 - v_3 u_1)\mathbf{e}_2 + (u_1 v_2 - v_1 u_2)\mathbf{e}_3. \tag{10}$$

Note: Defintion implies antisymmetry: $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$.

Defn: The **Levi-Civita tensor** (or antisymmetric tensor) is defined by

1. $\epsilon_{123} = 1$
2. $\epsilon_{ijk} = 0$ if any repeated suffices. E.g. $\epsilon_{113} = 0$.
3. Interchanging suffices implies reversal of sign. E.g. $\epsilon_{ijk} = -\epsilon_{jik}$.

Implies ϵ_{ijk} are invariant under cyclic rotation of suffices. Thus $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$, and all 21 others are zero.

Notation: We use $[\mathbf{v}]_i = v_i$ to denote the ith component of a vector.

Note: Another way of defining ϵ_{ijk} are as the set of 27 elements such that

$$[\mathbf{u} \times \mathbf{v}]_i = \epsilon_{ijk} u_j v_k. \tag{11}$$

For e.g. $[\mathbf{u} \times \mathbf{v}]_1 = \epsilon_{1jk} u_j v_k = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{1jk} u_j v_k = \epsilon_{111} u_1 v_1 + \epsilon_{112} u_1 v_2 + \epsilon_{113} u_1 v_3 + \epsilon_{121} u_2 v_1 + \epsilon_{122} u_2 v_2 + \epsilon_{123} u_2 v_3 + \epsilon_{131} u_3 v_1 + \epsilon_{132} u_3 v_2 + \epsilon_{133} u_3 v_3$ and this equals $u_2 v_3 - u_3 v_2$ only if $\epsilon_{123} = 1$, $\epsilon_{132} = -1$ and all 7 other ϵ_{1jk} are zero. Repeat for 2nd and 3rd components.

Note: The defintion of ϵ_{ijk} guarantees the antisymmetry of the cross product.

Proposition: A (**double product**)

$$\epsilon_{ijk} \epsilon_{ilm} = \delta_{jl} \delta_{km} - \delta_{jm} \delta_{kl}. \tag{12}$$

Proof: Just have to consider all possible (non-trivial) combinations to see it is true.

E.g. 2.6: (A **vector triple product**)

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

Proof: Vector identity, so start by looking at the scalar quantity which is the ith component of the LHS:

$$\begin{aligned}
[\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk} a_j [\mathbf{b} \times \mathbf{c}]_k \\
&= \epsilon_{ijk} a_j \epsilon_{klm} b_l c_m \\
&= \epsilon_{kij} \epsilon_{klm} a_j b_l c_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) a_j b_l c_m \\
&= a_j c_j b_i - a_j b_j c_i = (\mathbf{a} \cdot \mathbf{c}) b_i - (\mathbf{a} \cdot \mathbf{b}) c_i
\end{aligned}$$

True for $i = 1, 2, 3$, so result is proved.

2.2 Scalar and vector fields

Defn: Conventional language:

A **scalar field** on $\mathbb{R}^3$ is a function $f : \mathbb{R}^3 \to \mathbb{R}$.
A **vector field** on $\mathbb{R}^3$ is a map $\mathbf{v} : \mathbb{R}^3 \to \mathbb{R}^3$. We write $\mathbf{v}(\mathbf{r}) = (v_1(\mathbf{r}), v_2(\mathbf{r}), v_3(\mathbf{r}))$ where $v_i(\mathbf{r})$ $i = 1, 2, 3$ are scalar fields.

Scalar and vector fields defined in $\mathbb{R}^3$ are of particular importance for physical applications. For example:

- (Scalar fields) Temperature $T(\mathbf{r})$; mass density $\rho(\mathbf{r})$ for a fluid or gas; electric charge density $q(\mathbf{r})$.
- (Vector fields) Velocity $\mathbf{v}(\mathbf{r})$ of a fluid or gas; electric and magnetic fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{B}(\mathbf{r})$, displacement fields in elastic solid $\mathbf{u}(\mathbf{r})$.

In these physical applications, one often derives equations that govern vector and scalar fields which involve derivatives in space (and time).

The following three **first-order differential operations** of vector calculus emerge from this:

2.3 Gradient (grad)

Defn: The **gradient** of a scalar field f, denoted $\boldsymbol{\nabla} f$, is the vector field given by

$$\boldsymbol{\nabla} f(\mathbf{r}) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right), \qquad \text{or, in component form,} \qquad [\boldsymbol{\nabla} f]_i = \frac{\partial f}{\partial x_i}, \quad i = 1, 2, 3$$

The gradient maps scalar to vector fields.

E.g. 2.7:

$$\boldsymbol{\nabla} \tan^{-1}\left(\frac{y}{x}\right) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0 \right) \equiv \frac{1}{x^2 + y^2}(-y\hat{\mathbf{x}} + x\hat{\mathbf{y}}).$$

E.g. 2.8: Recall $r = \sqrt{x^2 + y^2 + z^2}$. A direct calculation gives

$$\boldsymbol{\nabla} r = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) = \frac{(x, y, z)}{r} = \frac{\mathbf{r}}{r} \qquad \text{or, in component form,} \qquad [\boldsymbol{\nabla} r]_i = x_i/r$$

Note: $\mathbf{r}/r$ is the unit vector from the origin to the point $\mathbf{r}$; we often denote this as $\hat{\mathbf{r}}$.

E.g. 2.9: If $f(\mathbf{r}) = g(r)$ (i.e. a function depends **only** on the distance from the origin) then

$$[\boldsymbol{\nabla} g(r)]_i = \frac{\partial g(r)}{\partial x_i} = \frac{\partial r}{\partial x_i}\frac{dg(r)}{dr} \equiv g'(r)[\boldsymbol{\nabla} r]_i = g'(r)\left(\frac{\mathbf{r}}{r}\right)_i$$

since r is a function of x_1, x_2 and x_3 and by using the **Chain rule**.

Thus $\boldsymbol{\nabla} g(r) = g'(r)\hat{\mathbf{r}}$ (c.f. potentials, central forces in Mech 1).

Recall from Calculus 1, two important interpretations of the gradient:

2.3.1 Interpretation of the gradient

Provided $\boldsymbol{\nabla} f$ is nonzero, the gradient points in the direction in which f increases most rapidly.

Proof: let $\mathbf{v}$ be s.t. $|\mathbf{v}| = 1$. Then rate of change of f in direction $\mathbf{v}$ is the directional derivative (see (2)) $D_{\mathbf{v}} f(\mathbf{r}) = \mathbf{v} \cdot \boldsymbol{\nabla} f = |\boldsymbol{\nabla} f| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\boldsymbol{\nabla} f$. Maximised when $\theta = 0$. I.e. when $\mathbf{v}$ in direction $\boldsymbol{\nabla} f$.

2.3.2 Another interpretation of the gradient

The gradient of f is perpendicular to the level surfaces of f.

(A level surface S is defined by values of $\mathbf{r}$ s.t. $f(\mathbf{r}) = C$, a constant.)

Proof: Let $\mathbf{c}(t)$ lie in S. Then $f(\mathbf{c}(t)) = C$, for all t. The chain rule yields

$$0 = \frac{d}{dt} f(\mathbf{c}(t)) = \boldsymbol{\nabla} f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$$

and since $\mathbf{c}'(t)$ is parallel to S at $\mathbf{c}(t)$, we have our result.

E.g. 2.10: Consider the temperature T in a room to be a function of 3D position (x, y, z):

$$T(\mathbf{r}) = \frac{\mathrm{e}^x \sin(\pi y)}{1 + z^2}$$

Q: If you stand at the point $(1, 1, 1)$ in which direction will the room get coolest fastest ?

A:

$$\boldsymbol{\nabla} T = \left(\frac{\mathrm{e}^x \sin(\pi y)}{1 + z^2}, \frac{\pi \mathrm{e}^x \cos(\pi y)}{1 + z^2}, -\frac{2z\mathrm{e}^x \sin(\pi y)}{(1 + z^2)^2}\right)$$

and at $(x,y,z) = (1,1,1)$, $\boldsymbol{\nabla} T = (0, -\frac{1}{2}\mathrm{e}, 0)$. So a vector pointing in the **direction** where temperature gets **coolest** (i.e. *decreases* most rapidly) is $(0,1,0)$.

E.g. 2.11: Take $f(x,y) = x^2 + 2y^2$. Then $\boldsymbol{\nabla} f = (2x, 4y)$.

For instance $\boldsymbol{\nabla} f$ evaluated at $(x,y) = (1,1)$ is $(2,4)$ and so the steepest ascent of f at $(1,1)$ is in direction $\tan^{-1}(2)$ w.r.t. x axis and gradient in that direction is $|\boldsymbol{\nabla} f| = 2\sqrt{5}$.

2.4 Divergence (Div)

Defn: The **divergence** of a vector field $\mathbf{v}(\mathbf{r})$, denoted $\boldsymbol{\nabla} \cdot \mathbf{v}$, is the scalar field given by

$$\boldsymbol{\nabla} \cdot \mathbf{v} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} \equiv \frac{\partial}{\partial x_j} v_j(\mathbf{r}) \equiv \partial_j v_j.$$

Note: The use of a 'dot' between the symbol used for the gradient and the vector field is purely notational. Do **not** get the divergence, which is a differential operation, confused with the dot product. For e.g. the dot product is $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ since multiplication is commutative, but

$$\mathbf{v} \cdot \boldsymbol{\nabla} = v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} + v_3 \frac{\partial}{\partial x_3} \neq \boldsymbol{\nabla} \cdot \mathbf{v}.$$

2.4.1 Interpretation of divergence

Harder without physical setting, but broadly it measures the expansion (positive divergence) or contraction of a field at a point.

For example, consider (i) $\mathbf{v}(\mathbf{r}) = (x, y, 0)$ then $\boldsymbol{\nabla} \cdot \mathbf{v} = 2$ and (ii) $\mathbf{v}(\mathbf{r}) = (-y, x, 0)$, then $\boldsymbol{\nabla} \cdot \mathbf{v} = 0$.

Note: The first case corresponds to a 2D radially spreading field and the second to a 2D circular rotating field (just believe me) which is why the divergence is positive (expanding) and zero (neither expanding nor contracting) in the two cases.

E.g. 2.12: $\mathbf{v}(\mathbf{r}) = (xyz, xyz, xyz)$ then $\boldsymbol{\nabla} \cdot \mathbf{v} = yz + zx + xy$

E.g. 2.13: $\boldsymbol{\nabla} \cdot \mathbf{r} = \dfrac{\partial x_j}{\partial x_j} = \delta_{jj} = 3.$

E.g. 2.14: $\boldsymbol{\nabla} \cdot (\mathbf{a} \times \mathbf{r}) = \dfrac{\partial}{\partial x_i} \epsilon_{ijk} a_j x_k = \epsilon_{ijk} a_j \dfrac{\partial x_k}{\partial x_i} = \epsilon_{ijk} a_j \delta_{ik} = \epsilon_{iji} a_j = 0$ since $\epsilon_{iji} = 0$.

2.5 Curl

Defn: The **curl** of a vector field $\mathbf{v}(\mathbf{r})$, denoted $\boldsymbol{\nabla} \times \mathbf{v}$, is the vector field (i.e. in $\mathbb{R}^3$) given by

$$\boldsymbol{\nabla} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial_x & \partial_y & \partial_z \\ v_1 & v_2 & v_3 \end{vmatrix} \equiv \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}, \frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}, \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right).$$

Alternatively (and very conveniently)

$$[\boldsymbol{\nabla} \times \mathbf{v}]_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} v_k \tag{13}$$

as for cross products.

2.5.1 Interpretation of Curl

Again harder without physical setting, but broadly it measures the rotation or circulation of a vector field (because it needs direction) at a point.

Using same examples from 'Div' section: (i) $\mathbf{v}(\mathbf{r}) = (x, y, 0)$ then $\boldsymbol{\nabla} \times \mathbf{v} = \mathbf{0}$ (the radially spreading field has no rotation); and (ii) $\mathbf{v}(\mathbf{r}) = (-y, x, 0)$, then $\boldsymbol{\nabla} \times \mathbf{v} = 2\hat{\mathbf{z}}$ (the rotational field has !)

E.g. 2.16: Let $\mathbf{v}(\mathbf{r}) = (y^2, x^2, y^2)$. Then $\boldsymbol{\nabla} \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial_x & \partial_y & \partial_z \\ y^2 & x^2 & y^2 \end{vmatrix} = (2y, 0, 2(x - y))$

E.g. 2.17: $[\boldsymbol{\nabla} \times \mathbf{r}]_i = \epsilon_{ijk} \partial_j x_k = \epsilon_{ijk} \delta_{jk} = \epsilon_{ijj} = 0$, so $\boldsymbol{\nabla} \times \mathbf{r} = \mathbf{0}$.

2.6 Second-order differential operations

Schematically, grad, div and curl act as follows:

$$\begin{aligned} \text{grad: scalar fields} &\rightarrow \text{vector fields} \\ \text{div: vector fields} &\rightarrow \text{scalar fields} \\ \text{curl: vector fields} &\rightarrow \text{vector fields} \end{aligned}$$

The operations of grad, div and curl can be combined. Thus, only the following combination of operations make sense:

$$\begin{aligned} \text{curl(grad): scalar fields} &\rightarrow \text{vector fields} \\ \text{div(grad): scalar fields} &\rightarrow \text{scalar fields} \\ \text{grad(div): vector fields} &\rightarrow \text{vector fields} \\ \text{div(curl): vector fields} &\rightarrow \text{scalar fields} \\ \text{curl(curl): vector fields} &\rightarrow \text{vector fields} \end{aligned}$$

2.6.1 Two Null Identities

1) For *any* scalar field f

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) = \mathbf{0}.$$

Proof: We have that

$$[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f)]_i = \epsilon_{ijk}\partial_j\partial_k f = -\epsilon_{ikj}\partial_j\partial_k f = -\epsilon_{ikj}\partial_k\partial_j f = -[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f)]_i.$$

Thus since the expression equals its own negative, it must vanish.

2) For *any* vector field, $\mathbf{v}$,

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = 0$$

Proof: We have that

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = \partial_i \epsilon_{ijk}\partial_j v_k = \epsilon_{ijk}\partial_i\partial_j v_k,$$

which must vanish for the same reason.

The remaining combinations of grad, div and curl are related to a second-order differential operator called the Laplacian...

2.7 The Laplacian

Defn: The **Laplacian** of a scalar field $f(\mathbf{r})$, denoted $\nabla^2 f$ (or $\triangle f$), is the scalar field defined by

$$\triangle f = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f(\mathbf{r}) = \partial_i^2 f \equiv \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) f.$$

The definition can also be extended to consider the **Laplacian of a vector field** $\mathbf{v}(\mathbf{r})$ which is

$$\triangle \mathbf{v} = (\triangle v_1, \triangle v_2, \triangle v_3)\,.$$

E.g. 2.18: For a vector field $\mathbf{v}(\mathbf{r})$,

$$\triangle \mathbf{v} = -\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v}) + \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v}).$$

Proof: We consider the ith component of the 1st RHS term:

$$\begin{aligned}
[\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v})]_i &= \epsilon_{ijk}\partial_j[\boldsymbol{\nabla} \times \mathbf{v}]_k \\
&= \epsilon_{ijk}\partial_j\epsilon_{klm}\partial_l v_m \\
&= \epsilon_{kij}\epsilon_{klm}\partial_j\partial_l v_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l v_m \\
&= \partial_i\partial_m v_m - \partial_j\partial_j v_i = [\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v}) - \triangle \mathbf{v}]_i
\end{aligned}$$

which shows that all components agree with our claim.

2.8 Curvilinear coordinate systems

All differential operators defined above were expressed in Cartesian coordinates. For many practical problems more natural to express problems in coordinates aligned with principal features of the problem. E.g. Polars are appropriate for circular domains.

Q: How do we recast the differential operators in a differential coordinate system ?

2.8.1 Coordinate transformations

Defn: Curvilinear coordinates are defined by a smooth function $\mathbf{r} : \mathbb{R}^3 \to \mathbb{R}^3$ which maps a point $\mathbf{q} = (q_1, q_2, q_3)$ in one coordinate system to a point in Cartesian space: $\mathbf{r} \equiv (x, y, z) = \mathbf{r}(\mathbf{q}) = \mathbf{r}(q_1, q_2, q_3)$. I.e.

$$x = x(q_1, q_2, q_3), \quad y = y(q_1, q_2, q_3), \quad z = z(q_1, q_2, q_3)$$

The inverse map, if it exists (see later) is

$$q_1 = q_1(x, y, z), \quad q_2 = q_2(x, y, z), \quad q_3 = q_3(x, y, z).$$

For example: in 2D, if $q_1 = r$ and $q_2 = \theta$ then $x = x(r, \theta) = r\cos\theta$ and $y = y(r, \theta) = r\sin\theta$. The inverse map is $r = r(x, y) = \sqrt{x^2 + y^2}$, $\theta = \theta(x, y) = \tan^{-1}(y/x)$.

Defn: The surfaces $q_i = const$ are called **coordinate surfaces**. The space curves formed by their intersection in pairs are called the **coordinate curves**. The **coordinate axes** are determined by the tangents to the coordinate curves at the intersection of three surfaces. They are not, in general, fixed directions in space.

The two points $\mathbf{r}(q_1, q_2, q_3)$ and $\mathbf{r}(q_1 + dq_1, q_2, q_3)$ lie on a coordinate curve formed by q_2, q_3 constant. Thus, the q_1-coordinate axis is determined by letting $dq_1 \to 0$ in

$$\frac{\mathbf{r}(q_1 + dq_1, q_2, q_3) - \mathbf{r}(q_1, q_2, q_3)}{dq_1} = \frac{\mathbf{r}(q_1, q_2, q_3) + dq_1 \dfrac{\partial \mathbf{r}}{\partial q_1} - \mathbf{r}(q_1, q_2, q_3) + \text{h.o.t. order } (dq_1)^2}{dq_1} = \frac{\partial \mathbf{r}}{\partial q_1}$$

(after Taylor expanding). Repeat with q_2 and q_3. Thus, we can describe the point $\mathbf{q} = q_1\hat{\mathbf{q}}_1 + q_2\hat{\mathbf{q}}_2 + q_3\hat{\mathbf{q}}_3$ in terms of the local **coordinate basis** given by unit vectors directed along the local coordinate axes:

$$\hat{\mathbf{q}}_1 = \frac{1}{h_1}\frac{\partial \mathbf{r}}{\partial q_1}, \qquad \hat{\mathbf{q}}_2 = \frac{1}{h_2}\frac{\partial \mathbf{r}}{\partial q_2}, \qquad \hat{\mathbf{q}}_3 = \frac{1}{h_3}\frac{\partial \mathbf{r}}{\partial q_3}$$

where, to ensure $|\hat{\mathbf{q}}_i| = 1$, we have normalised by

$$h_i = \left|\frac{\partial \mathbf{r}}{\partial q_i}\right|.$$

The h_i are called the **metric coefficients** or **scale factors**.

Note: the use of Greek indices in, for e.g.

$$\hat{\mathbf{q}}_\alpha = \frac{1}{h_\alpha}\frac{\partial \mathbf{r}}{\partial q_\alpha}$$

for $\alpha = 1, 2, 3$ indicates that the summation convention is **not** applied.

Remark: Is this always possible ? I.e. is there always a unique map from one system to another ? This is the same as asking if there is an inverse map. Thus (by the inverse function theorem) the answer lies in the Jacobian matrix of the map, given here by $\mathbf{r}'(\mathbf{q})$ which is the matrix with $h_\alpha \hat{\mathbf{q}}_\alpha$ as column vectors (for $\alpha = 1, 2, 3$). Thus the Jacobian determinant

$$J_\mathbf{r} = \frac{\partial(x, y, z)}{\partial(q_1, q_2, q_3)} \equiv \begin{vmatrix} h_1[\hat{\mathbf{q}}_1]_1 & h_2[\hat{\mathbf{q}}_2]_1 & h_3[\hat{\mathbf{q}}_3]_1 \\ h_1[\hat{\mathbf{q}}_1]_2 & h_2[\hat{\mathbf{q}}_2]_2 & h_3[\hat{\mathbf{q}}_3]_2 \\ h_1[\hat{\mathbf{q}}_1]_3 & h_2[\hat{\mathbf{q}}_2]_3 & h_3[\hat{\mathbf{q}}_3]_3 \end{vmatrix}$$

must be non-vanishing. The 2nd representation simply shows that no new calculations are needed to populate the entries of the Jacobian determinant.

Defn: If local basis vectors of a curvilinear coordinate system are mutually orthogonal, we call it an **orthogonal** curvilinear coordinate system. Convention dictates that the system be **right-handed**, or $\hat{\mathbf{q}}_1 = \hat{\mathbf{q}}_2 \times \hat{\mathbf{q}}_3$. (or, form axes from your thumb, index and middle fingers of your right hand and order basis vectors 1, 2, 3 on each respective digit.)

In the following, we will deal exclusively with orthogonal systems.

E.g. 2.19: Consider the linear map

$$\mathbf{r} = R\mathbf{q}, \qquad \text{such that} \qquad x_i = R_{ij} q_j$$

and R is an orthogonal matrix (a matrix s.t. $R^T R = I$ which implies $R^{-1} = R^T$). Then

$$\frac{\partial \mathbf{r}}{\partial q_1} = (R_{11}, R_{21}, R_{31}),$$
$$\frac{\partial \mathbf{r}}{\partial q_2} = (R_{12}, R_{22}, R_{32}),$$
$$\frac{\partial \mathbf{r}}{\partial q_3} = (R_{13}, R_{23}, R_{33}).$$

The scale factors are

$$h_1 = \left|\frac{\partial \mathbf{r}}{\partial q_1}\right| = \sqrt{R_{11}^2 + R_{21}^2 + R_{31}^2}, \qquad \text{and similarly for } h_2,\, h_3$$

The matrix equation $R^T R = I$ can be expressed as

$$\delta_{ij} = R_{ik}^T R_{kj} = R_{ki} R_{kj} \tag{14}$$

Hence $h_1 = 1$ (and similarly $h_2 = h_3 = 1$).

Thus the local basis vectors are

$$\hat{\mathbf{q}}_j = (R_{1j}, R_{2j}, R_{3j}), \qquad j = 1, 2, 3.$$

These are constant, i.e. they do not vary with position.

Note: From the definition of the basis vectors and using summation notation for the dot product we have $\hat{\mathbf{q}}_i \cdot \hat{\mathbf{q}}_j = R_{ki}R_{kj} = \delta_{ij}$ (using (14) and so the basis vectors are othonormal.

In other words, the new coordinate axes are a general rotation of the original $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ axes.

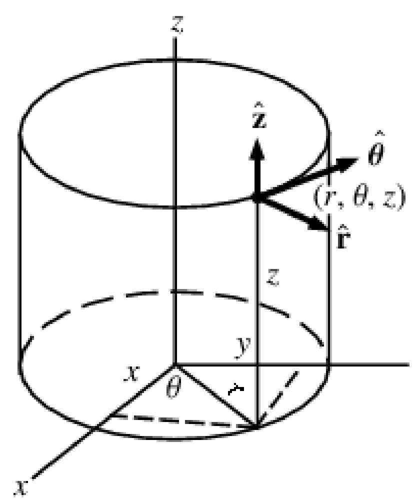

Figure 1: A local basis in cylindrical polar coordinates.

E.g. 2.20: In 3D, **cylindrical polar coordinates** are defined by the mapping

$$(x, y, z) = \mathbf{r}(r, \theta, z) = (r\cos\theta, r\sin\theta, z)$$

see Fig. 1. It follows that

$$\frac{\partial \mathbf{r}}{\partial r} = (\cos\theta, \sin\theta, 0), \quad \frac{\partial \mathbf{r}}{\partial \theta} = (-r\sin\theta, r\cos\theta, 0), \quad \frac{\partial \mathbf{r}}{\partial z} = (0, 0, 1).$$

The scale factors are

$$h_r = \left|\frac{\partial \mathbf{r}}{\partial r}\right| = 1, \quad h_\theta = \left|\frac{\partial \mathbf{r}}{\partial \theta}\right| = r, \quad h_z = \left|\frac{\partial \mathbf{r}}{\partial z}\right| = 1. \tag{15}$$

Thus the local basis vectors are (using standard notation):

$$\hat{\mathbf{r}} = (\cos\theta, \sin\theta, 0), \quad \hat{\boldsymbol{\theta}} = (-\sin\theta, \cos\theta, 0), \quad \hat{\mathbf{z}} = (0, 0, 1), \tag{16}$$

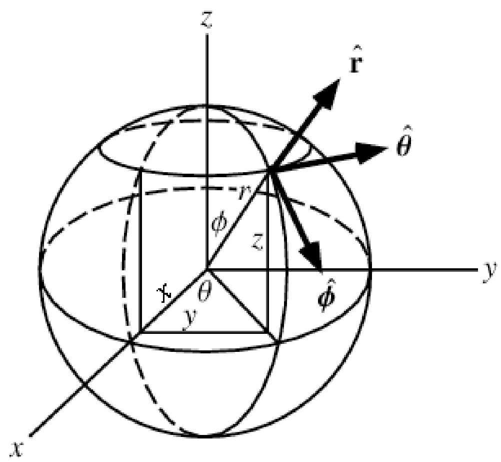

Figure 2: A local basis in spherical polar coordinates. The vector $\hat{\mathbf{r}}$ points along a ray from the center, $\hat{\boldsymbol{\phi}}$ points along the meridians, and $\hat{\boldsymbol{\theta}}$ along the parallels.

and these vary with position. Note that $\hat{\mathbf{r}} \cdot \hat{\boldsymbol{\theta}} = \hat{\mathbf{r}} \cdot \hat{\mathbf{z}} = \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} = 0$, and $\hat{\mathbf{r}} = \hat{\boldsymbol{\theta}} \times \hat{\mathbf{z}}$, so cylindrical coordinates are indeed orthogonal.

E.g. 2.21: Spherical polar coordinates are defined by the mapping

$$(x, y, z) = \mathbf{r}(r, \phi, \theta) = (r \sin\phi \cos\theta, r \sin\phi \sin\theta, r \cos\phi),$$

see Fig. 2. Now the derivatives with respect to the coordinates are

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial r} &= (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi), \\ \frac{\partial \mathbf{r}}{\partial \phi} &= (r\cos\phi\cos\theta, r\cos\phi\sin\theta, -r\sin\phi), \\ \frac{\partial \mathbf{r}}{\partial \theta} &= (-r\sin\phi\sin\theta, r\sin\phi\cos\theta, 0), \end{aligned}$$

and the scale factors become (check):

$$h_r = \left|\frac{\partial \mathbf{r}}{\partial r}\right| = 1, \quad h_\phi = \left|\frac{\partial \mathbf{r}}{\partial \phi}\right| = r, \quad h_\theta = \left|\frac{\partial \mathbf{r}}{\partial \theta}\right| = r\sin\phi. \tag{17}$$

Thus the local basis vectors are

$$\begin{aligned} \hat{\mathbf{r}} &= (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi), \\ \hat{\boldsymbol{\phi}} &= (\cos\phi\cos\theta, \cos\phi\sin\theta, -\sin\phi), \\ \hat{\boldsymbol{\theta}} &= (-\sin\theta, \cos\theta, 0), \end{aligned} \tag{18}$$

and vary with position. Again, $\hat{\mathbf{r}} \cdot \hat{\boldsymbol{\phi}} = \hat{\mathbf{r}} \cdot \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\phi}} \cdot \hat{\boldsymbol{\theta}} = 0$, and $\hat{\mathbf{r}} = \hat{\boldsymbol{\phi}} \times \hat{\boldsymbol{\theta}}$, so spherical coordinates are orthogonal.

2.8.2 Transformation of the gradient

The differential operator $\boldsymbol{\nabla}$ is the Cartesian vector

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

We want this to be transformed into derivatives w.r.t the local coordinates q_1, q_2, q_3.

Consider $f(\mathbf{r}) = f(\mathbf{r}(\mathbf{q}))$. Then for fixed $\alpha = 1, 2, 3$, the chain rule gives

$$\frac{1}{h_\alpha}\frac{\partial f}{\partial q_\alpha} = \frac{1}{h_\alpha}\frac{\partial x_j}{\partial q_\alpha}\frac{\partial f}{\partial x_j} = \hat{\mathbf{q}}_\alpha \cdot \boldsymbol{\nabla} f \tag{19}$$

(summation over j is implied, but not α).

Now if $\mathbf{u} = u_1\hat{\mathbf{q}}_1 + u_2\hat{\mathbf{q}}_2 + u_3\hat{\mathbf{q}}_3$ then the orthonormal property of the local basis functions means $u_j = \mathbf{u} \cdot \hat{\mathbf{q}}_j$. If we let $\mathbf{u} = \boldsymbol{\nabla} f$ and with (19) we find

$$\boldsymbol{\nabla} f = \sum_{\alpha=1}^{3} \frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}\frac{\partial f}{\partial q_\alpha}, \qquad \text{and so} \qquad \boldsymbol{\nabla} = \sum_{\alpha=1}^{3} \frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}\frac{\partial}{\partial q_\alpha}. \tag{20}$$

E.g. 2.22: In cylindrical polar coordinates, according to (20) and (15), we have

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\mathbf{z}}\frac{\partial}{\partial z}$$

E.g. 2.23: In spherical coordinates, according to (20) and (18),

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\phi}}\frac{1}{r}\frac{\partial}{\partial \phi} + \hat{\boldsymbol{\theta}}\frac{1}{r\sin\phi}\frac{\partial}{\partial \theta}$$

2.8.3 Transformation of the divergence

To find $\boldsymbol{\nabla} \cdot \mathbf{u}$ in curvilinear coordinates we first need to express the vector field $\mathbf{u}$ in the local coordinate system. I.e.

$$\mathbf{u} = u_1\hat{\mathbf{q}}_1 + u_2\hat{\mathbf{q}}_2 + u_3\hat{\mathbf{q}}_3.$$

The difficulty here is that both u_i and $\hat{\mathbf{q}}_i$ depend on (q_1, q_2, q_3). We come at the divergence in a slightly roundabout way.

First, we note from (20) that

$$\boldsymbol{\nabla} q_\alpha = \frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}.$$

Now note that

$$\boldsymbol{\nabla} \times (q_2 \boldsymbol{\nabla} q_3) = q_2 \underbrace{(\boldsymbol{\nabla} \times (\boldsymbol{\nabla} q_3))}_{=0} + (\boldsymbol{\nabla} q_2) \times (\boldsymbol{\nabla} q_3) = \frac{\hat{\mathbf{q}}_2}{h_2} \times \frac{\hat{\mathbf{q}}_3}{h_3} = \frac{\hat{\mathbf{q}}_1}{h_2 h_3}.$$

Then from §2.6.1 (Null identites: $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{A})=0$, $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}f)=\mathbf{0}$ for any $\mathbf{A}, f$,)

$$\boldsymbol{\nabla}\times\left(\frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}\right)=\mathbf{0},\qquad \boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{q}}_1}{h_2h_3}\right)=0.$$

Results true for the 2 cyclic permutations $(1\to 2, 2\to 3, 3\to 1)$

$$\boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{q}}_2}{h_1h_3}\right)=\boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{q}}_3}{h_1h_2}\right)=0.$$

So now

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\mathbf{u} &= \boldsymbol{\nabla}.\left((u_1h_2h_3)\frac{\hat{\mathbf{q}}_1}{h_2h_3}\right)+2\text{ cyclic perms}\\
&= \frac{\hat{\mathbf{q}}_1}{h_2h_3}\cdot\boldsymbol{\nabla}(u_1h_2h_3)+(u_1h_2h_3)\,\boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{q}}_1}{h_2h_3}\right)+2\text{ cyclic perms}\\
&= \frac{\hat{\mathbf{q}}_1}{h_2h_3}\cdot\left(\sum_{\alpha=1}^{3}\frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}\frac{\partial(u_1h_2h_3)}{\partial q_\alpha}\right)+2\text{ cyclic perms}\\
&= \frac{1}{h_1h_2h_3}\left[\frac{\partial(u_1h_2h_3)}{\partial q_1}+\frac{\partial(u_2h_1h_3)}{\partial q_2}+\frac{\partial(u_3h_1h_2)}{\partial q_3}\right]
\end{aligned}$$

using the fact that $\hat{\mathbf{q}}_\alpha\cdot\hat{\mathbf{q}}_\beta=\delta_{\alpha\beta}$.

E.g. 2.24: (Cylindrical polar coordinates.) First write

$$\mathbf{u}=u_r\hat{\mathbf{r}}+u_\theta\hat{\boldsymbol{\theta}}+u_z\hat{\mathbf{z}}.$$

with $h_r=1$, $h_\theta=r$, $h_z=1$, so

$$\boldsymbol{\nabla}\cdot\mathbf{u}=\frac{1}{r}\left[\frac{\partial(ru_r)}{\partial r}+\frac{\partial u_\theta}{\partial\theta}+\frac{\partial(ru_z)}{\partial z}\right]=\frac{\partial u_r}{\partial r}+\frac{u_r}{r}+\frac{1}{r}\frac{\partial u_\theta}{\partial\theta}+\frac{\partial u_z}{\partial z}.$$

For example, if $\mathbf{u}=f(r)\hat{\boldsymbol{\theta}}$ then $\boldsymbol{\nabla}\cdot\mathbf{u}=0$.

2.8.4 Transformation of curl

Taking a similar approach to 'Div', we write

$$\begin{aligned}
\boldsymbol{\nabla}\times\mathbf{u} &= \boldsymbol{\nabla}\times\left((h_1u_1)\frac{\hat{\mathbf{q}}_1}{h_1}\right)+2\text{ cyclic perms}\\
&= \boldsymbol{\nabla}(h_1u_1)\times\frac{\hat{\mathbf{q}}_1}{h_1}+(h_1u_1)\boldsymbol{\nabla}\times\frac{\hat{\mathbf{q}}_1}{h_1}+2\text{ cyclic perms}\\
&= \left(\sum_{\alpha=1}^{3}\frac{\hat{\mathbf{q}}_\alpha}{h_\alpha}\frac{\partial(h_1u_1)}{\partial q_\alpha}\right)\times\frac{\hat{\mathbf{q}}_1}{h_1}+2\text{ cyclic perms}\\
&= \frac{\hat{\mathbf{q}}_2}{h_1h_3}\frac{\partial(h_1u_1)}{\partial q_3}-\frac{\hat{\mathbf{q}}_3}{h_1h_2}\frac{\partial(h_1u_1)}{\partial q_2}+2\text{ cyclic perms}
\end{aligned}$$

$$=\frac{\hat{\mathbf{q}}_1}{h_2h_3}\left(\frac{\partial(h_3u_3)}{\partial q_2}-\frac{\partial(h_2u_2)}{\partial q_3}\right)+\frac{\hat{\mathbf{q}}_2}{h_1h_3}\left(\frac{\partial(h_1u_1)}{\partial q_3}-\frac{\partial(h_3u_3)}{\partial q_1}\right)+\frac{\hat{\mathbf{q}}_3}{h_1h_2}\left(\frac{\partial(h_2u_2)}{\partial q_1}-\frac{\partial(h_1u_1)}{\partial q_2}\right).$$

Now we see

$$\boldsymbol{\nabla} \times \mathbf{u} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1\hat{\mathbf{q}}_1 & h_2\hat{\mathbf{q}}_2 & h_3\hat{\mathbf{q}}_3 \\ \partial/\partial q_1 & \partial/\partial q_2 & \partial/\partial q_3 \\ h_1 u_1 & h_2 u_2 & h_3 u_3 \end{vmatrix}.$$

2.9 Examples

E.g. 2.25: The **Laplacian** of a scalar field ϕ is $\triangle\phi = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\phi$. Since

$$\boldsymbol{\nabla}\phi = \hat{\mathbf{r}}\frac{\partial\phi}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial\phi}{\partial\theta} + \hat{\mathbf{z}}\frac{\partial\phi}{\partial z}.$$

we use the defn of div to give

$$\triangle\phi = \frac{\partial}{\partial r}\left(\frac{\partial\phi}{\partial r}\right) + \frac{1}{r}\frac{\partial\phi}{\partial r} + \frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{1}{r}\frac{\partial\phi}{\partial\theta}\right) + \frac{\partial}{\partial z}\left(\frac{\partial\phi}{\partial z}\right) = \frac{\partial^2\phi}{\partial r^2} + \frac{1}{r}\frac{\partial\phi}{\partial r} + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} + \frac{\partial^2\phi}{\partial z^2}.$$

E.g. 2.26: Now the curl in cylindrical polars:

$$\boldsymbol{\nabla}\times\mathbf{u} = \left(\frac{1}{r}\frac{\partial u_z}{\partial\theta} - \frac{\partial u_\theta}{\partial z}\right)\hat{\mathbf{r}} + \left(\frac{\partial u_r}{\partial z} - \frac{\partial u_z}{\partial r}\right)\hat{\boldsymbol{\theta}} + \left(\frac{\partial u_\theta}{\partial r} + \frac{u_\theta}{r} - \frac{1}{r}\frac{\partial u_r}{\partial\theta}\right)\hat{\mathbf{z}},$$

Exercise: Do the same for spherical polars !!

Remark: If curvilinear system *not orthogonal* then we are in a real mess.

3 Integration theorems of vector calculus

Having done differential vector calculus, we turn to integral vector calculus. These are equally important in applications as you will see in APDE2, Fluid Dynamics and beyond. We shall derive three (quite stunning) main integral identities all of which may be considered as higher-dimensional generalisations of the Fundamental Theorem of Calculus:

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

The LHS is a one-dimensional integral (i.e. an integral over a line) which is equated to zero-dimensional (i.e. pointwise) evaluations on the boundary of the integral (here at $x = a, b$).

Remark: The formula for integration by parts is found by letting $f(x) = u(x)v(x)$ in the above !

3.1 The line integral of a scalar field

An ordinary 1D integral can be regarded as integration along a straight line. For example if $F(x)$ is the force on a particle alowed to move along the x-axis,

$$\int_{x_1}^{x_2} F(x)\,dx$$

is the "work done" moving it from x_1 to x_2. We want to integrals along general paths in $\mathbb{R}^2$ or $\mathbb{R}^3$.

Defn: A **path** is a bijective (i.e. one-to-one) map $\mathbf{p} : [t_1, t_2] \to \mathbb{R}^3$ s.t. $t \mapsto \mathbf{p}(t)$. It connects the point $\mathbf{p}(t_1)$ to $\mathbf{p}(t_2)$ along a curve C, say. We say the curve C is **parametrised** by the path.

Defn: The **line integral** of a scalar field $f : \mathbb{R}^3 \to \mathbb{R}$ along a curve C is denoted

$$\int_C f(\mathbf{r})\,ds.$$

and $ds = |d\mathbf{r}|$ denotes the elemental arclength. Since $\mathbf{r} = \mathbf{p}(t)$ on C, $d\mathbf{r} = \mathbf{p}'(t)\,dt$ and so

$$\int_C f(\mathbf{r})\,ds = \int_{t_1}^{t_2} f(\mathbf{p}(t))|\mathbf{p}'(t)|\,dt.$$

E.g. 3.1: Let $\mathbf{p}(t) = (t, t, t)$ for $t \in [0, 1]$ connects the points $(0, 0, 0)$ to $(1, 1, 1)$ by a straight line of length $\sqrt{3}$. If $f = xyz$ then

$$\int_C f\,ds = \int_0^1 t^3\sqrt{1+1+1}\,dt = \frac{\sqrt{3}}{4}$$

E.g. 3.2: Let $\mathbf{p}(t) = (t^2, t^2, t^2)$ for $t \in [0, 1]$ parametrises the same curve as in E.g. 3.1. With the same f we have

$$\int_C f\,ds = \int_0^1 t^6\sqrt{(2t)^2 + (2t)^2 + (2t)^2}\,dt = \frac{\sqrt{3}}{4}.$$

Note: Parameterisation is not unique. Suggests value of line integral is independent of parametrisation.

Proof: Consider the bijective map $t = g(u)$ for $t_1 < t < t_2$ such that $t_1 = g(u_1)$, $t_2 = g(u_2)$. Then

$$\int_{t_1}^{t_2} f(\mathbf{p}(t))|\mathbf{p}'(t)|\,dt = \int_{u_1}^{u_2} f(\mathbf{p}(g(u)))|\mathbf{p}'(g(u))|g'(u)\,du = \int_{u_1}^{u_2} f(\mathbf{q}(u))|\mathbf{q}'(u)|\,du.$$

after letting $\mathbf{q}(u) = \mathbf{p}(g(u))$ and noting $\mathbf{q}'(u) = g'(u)\mathbf{p}'(g(u))$ by the chain rule.

Note: Value of line integral **does** depend on **direction** of path along C:

$$\int_C f\,ds = -\int_{-C} f\,ds$$

($-C$ is C in reverse). We are used to this notion in 1D, viz $\displaystyle\int_{x_1}^{x_2} f(x)dx = -\int_{x_2}^{x_1} f(x)dx$.

3.2 The line integral of a vector field

Defn: Let $\mathbf{F}(\mathbf{r}) : \mathbb{R}^3 \to \mathbb{R}^3$ be a vector field, and let $\mathbf{p}(t)$ be a path on the interval $[t_1, t_2]$. The **line integral** of $\mathbf{F}$ along $\mathbf{p}$ is defined by

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_{t_1}^{t_2} \mathbf{F}(\mathbf{p}(t))\cdot\mathbf{p}'(t)\,dt$$

as above.

Note: As above, the value of the line integral is not dependent on parametrisation of C but is negated by a reversal of C.

E.g. 3.3: Integrate $\mathbf{F} = \sin\phi\hat{\mathbf{z}}$ (ϕ is polar angle in spherical polars) along a meridian of a sphere of radius R from the south to the north pole.

A: From the description of the path, C, convenient to use spherical coordinates (r, ϕ, θ). I.e.

$$\mathbf{p}(\phi) = R\hat{\mathbf{r}} = R(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi);$$

(see earlier defn of $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\phi}}$ in spherical polars (18)) then

$$\frac{d\mathbf{p}}{d\phi} = R(\cos\phi\cos\theta, \cos\phi\sin\theta, -\sin\phi) \equiv R\hat{\boldsymbol{\phi}}.$$

Thus we can see that $\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\phi}} = -\sin\phi$ and so

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_\pi^0 \mathbf{F}\cdot\mathbf{p}'(\phi)\,d\phi = R\int_\pi^0 \sin\phi\,\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\phi}}\,d\phi = -R\int_\pi^0 \sin^2\phi d\phi = \frac{R\pi}{2}.$$

Proposition: Let $f(\mathbf{r})$ be a scalar field and let C be a curve in $\mathbb{R}^3$ parametrised by the path $\mathbf{p}(t)$, $t_1 \le t \le t_2$. Then

$$\int_C \nabla f\cdot d\mathbf{r} = f(\mathbf{p}(t_2)) - f(\mathbf{p}(t_1)).$$

This is the **fundamental theorem of Calculus for line integrals**.

Proof: We have

$$\int_C \nabla f \cdot d\mathbf{r} = \int_{t_1}^{t_2} \nabla f(\mathbf{p}(t)) \cdot \mathbf{p}'(t)\, dt.$$

But from the chain rule it follows that

$$\frac{d}{dt} f(\mathbf{p}(t)) = \mathbf{p}'(t) \cdot \nabla f(\mathbf{p}(t)).$$

Therefore,

$$\int_C \nabla f \cdot d\mathbf{r} = \int_{t_1}^{t_2} \frac{d}{dt} f(\mathbf{p}(t))\, dt = f(\mathbf{p}(t_2)) - f(\mathbf{p}(t_1)),$$

from the Fundamental Theorem of Calculus.

Note: If C is closed, the line integral over a gradient field vanishes. As a result, line integrals of gradient fields are independent of the path C.

Remark: The line integral of a vector field is often called the **work integral**, since if $\mathbf{F}$ represents a force, the integral represents the work done moving a particle between two points. If $\mathbf{F} = \nabla f$ for some scalar field f (often called the **potential**) then the work done moving the particle is independent of the path taken. Moreover the work done moving a particle which returns to the same position is zero. Such a force is called **conservative**.

3.3 Surface integrals of scalar and vector fields.

We now generalise 1D integrals to 2D integrals. We start with parametrisations of surfaces.

Defn: A path $\mathbf{p}(t)$, for $t \in [t_1, t_2]$ is **closed** if $\mathbf{p}(t_1) = \mathbf{p}(t_2)$. A closed path is **simple** it it does not intersect with itself apart from at the end points t_1, t_2.

Defn: Let $D \subset \mathbb{R}^2$, let ∂D represent the **boundary** of D (it should be a simple closed path) and let $\bar{D}$ be $D \cup \partial D$.

Now define a map $\mathbf{s} : \bar{D} \to \mathbb{R}^3$ s.t $(u, v) \mapsto \mathbf{s}(u, v)$ and $\partial \mathbf{s}/\partial u$, $\partial \mathbf{s}/\partial v$ are linearly independent on D. A *surface* $S \in \mathbb{R}^3$ is given in *parametrised form* by $S = \{\mathbf{s}(u, v) \,|\, (u, v) \in D\}$.

E.g. 3.4: Let

$$D = \{(u, v) \,|\, u^2 + v^2 < R^2\}.$$

Then ∂D is the circle $\{(u, v) \,|\, u^2 + v^2 = R^2\}$ of radius R. Let $\mathbf{s}(u, v) = (u, v, \sqrt{R^2 - u^2 - v^2})$, then S is a hemispherical surface since the map defines $x = u$, $y = v$ and $z = \sqrt{R^2 - u^2 - v^2} = \sqrt{R^2 - x^2 - y^2}$.

Note: this is not the only way to parametrise a hemisphere; could (and will) use spherical polars.

Defn: The **integral of a scalar field f over a surface** S is denoted by

$$\int_S f(\mathbf{r})\, dS \equiv \int_S f(\mathbf{r})\, |d\mathbf{S}|$$

where $d\mathbf{S} = \hat{\mathbf{n}} dS$ and $\hat{\mathbf{n}}$ is a unit vector pointing out from S (a surface element is defined by its size dS and a direction, $\hat{\mathbf{n}}$, being the normal to the surface).

Note: $\displaystyle\int_S dS$ is the physical area of the surface S in the same way that $\displaystyle\int_C ds$ is the physical length of the curve C.

Now the two vectors $\mathbf{s}(u + du, v)$ and $\mathbf{s}(u, v)$ both lie on S and, assuming du vanishingly small, and Taylor expanding,

$$\mathbf{s}(u + du, v) - \mathbf{s}(u, v) = \mathbf{s}(u, v) + du \frac{\partial \mathbf{s}}{\partial u}(u, v) - \mathbf{s}(u, v) + \text{h.o.t. order } (du)^2.$$

Do the same thing with v. It follows that the two vectors

$$\frac{\partial \mathbf{s}}{\partial u} du, \quad \frac{\partial \mathbf{s}}{\partial v} dv$$

lie in the tangent plane to the surface S and the area dS of the elemental rhomboid formed by these two vectors in the direction normal to dS is

$$d\mathbf{S} = \frac{\partial \mathbf{s}}{\partial u} du \times \frac{\partial \mathbf{s}}{\partial v} dv \equiv \mathbf{N}(u, v) du dv, \qquad \mathbf{N}(u, v) = \frac{\partial \mathbf{s}}{\partial u} \times \frac{\partial \mathbf{s}}{\partial v}$$

and so

$$\int_S f(\mathbf{r})\, dS = \int_D f(\mathbf{s}(u, v)) |\mathbf{N}|\, du\, dv.$$

Note: If S lies in the 2D plane, then $\mathbf{s}(u, v) = (x(u, v), y(u, v), 0)$ is a mapping from 2D to 2D and so

$$|\mathbf{N}| = \frac{\partial(x, y)}{\partial(u, v)}$$

is the Jacobian of the map (easy to confirm). We know this from Calculus 1 (e.g. $dxdy \mapsto r dr d\theta$).

Defn: Let $\mathbf{v}$ be a vector field on $\mathbb{R}^3$. The **integral of v over** S, is denoted

$$\int_S \mathbf{v} \cdot d\mathbf{S} \equiv \int_S \mathbf{v} \cdot \hat{\mathbf{n}}\, dS = \int_D \mathbf{v}(\mathbf{s}(u, v)) \cdot \mathbf{N}(u, v)\, du\, dv,$$

as above.

Important remark: By analogy with line integrals, can show that the surface integral of a vector field is independent of parametrisation up to a sign. The sign depends on the orientation of the parametrisation, which is determined by the direction of the unit normal $\hat{\mathbf{n}} = \mathbf{N}/|\mathbf{N}|$. Thus, the direction of $\hat{\mathbf{n}}$, (or $\mathbf{N}$) **must be specified** in order to fix the sign of the integral unambiguously.

E.g. 3.5: Calculate $\displaystyle\int_S \mathbf{B} \cdot d\mathbf{S}$ where $\mathbf{B}(\mathbf{r}) = r\hat{\mathbf{z}} + \hat{\mathbf{r}}$ (expressed in cylindrical polars) in which $S = \{(x, y, 0) \,|\, x^2 + y^2 \leq 1\}$ is directed in positive z-direction.

A: Cylindrial polar coordinates are a sensible parametrisation of S, i.e.

$$\mathbf{s}(r, \theta) = (r\cos\theta, r\sin\theta, 0),$$

so that $D = \{(r, \theta) \,|\, 0 \leq r \leq 1,\ 0 \leq \theta \leq 2\pi\}$. Then

$$\mathbf{N} = \frac{\partial \mathbf{s}}{\partial r} \times \frac{\partial \mathbf{s}}{\partial \theta} = r\hat{\mathbf{z}},$$

(factor r is the Jacobian determinant in this 2D mapping) and points in direction $\hat{\mathbf{z}}$ normal to the 2D plane as required by the question. Then

$$\int_S \mathbf{B} \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^1 (r\hat{\mathbf{z}} + \hat{\mathbf{r}}) \cdot \hat{\mathbf{z}}\, r dr d\theta = \int_0^{2\pi} \int_0^1 r^2\, dr\, d\theta = \frac{2\pi}{3}$$

and we have used $\hat{\mathbf{r}} \cdot \hat{\mathbf{z}} = 0$.

3.4 Stokes' theorem

Consider that the vector field $\mathbf{v}$ is expressed as the curl of another vector field, i.e. $\mathbf{v} = \boldsymbol{\nabla} \times \mathbf{A}$. This frequently happens in applications.

Defn: The **boundary** of the surface S is denoted ∂S and, since it is mapped from the boundary ∂D, it inherits its properties, being a simple closed path. If $\mathbf{c}(t) \in \mathbb{R}^2$ is the simple closed path along ∂D then

$$\mathbf{p}(t) = \mathbf{s}(\mathbf{c}(t))$$

is the simple closed path along ∂S.

Note: A closed path has no start and finish point and can be oriented in either the anti-clockwise or clockwise directions.

Proposition: (**Stokes' theorem**) Let S be a surface in $\mathbb{R}^3$ with boundary ∂S; let $\mathbf{A}$ be a vector field. Then

$$\int_S \boldsymbol{\nabla} \times \mathbf{A} \cdot d\mathbf{S} = \int_{\partial S} \mathbf{A} \cdot d\mathbf{r}, \tag{21}$$

where $d\mathbf{S}$ and ∂S must be consistently oriented according to the right-hand THUMB rule.

Defn: (**Right-hand thumb rule**) Point the thumb of your right hand along curve ∂S (either clockwise or anti-clockwise). Wrap your fingers around the curve; your fingers will indicate the direction of the normal $\mathbf{N}$ (or $\hat{\mathbf{n}}$) of the surface that must be chosen in accordance with your choice of direction around the curve.

E.g. 3.6: Compute $\int_S \boldsymbol{\nabla} \times \mathbf{f} \cdot d\mathbf{S}$ where $\mathbf{f}(\mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}$ and $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$ is a constant vector. S is the hemisphere in $z > 0$ of radius R with the normal to the surface defined to point *inwards* towards the origin.

A: On problem sheet 2, we have shown that $\boldsymbol{\nabla} \times \mathbf{f} = 2\boldsymbol{\omega} \equiv (2\omega_1, 2\omega_2, 2\omega_3)$.

We will calculate the surface integral using three different approaches.

(i) Cartesian coordinates. Use the parametrisation defined in E.g. 3.4. I.e. the map

$$\mathbf{s}(u, v) = (u, v, \sqrt{R^2 - u^2 - v^2}).$$

maps the domain $D = \{(u, v) \,|\, u^2 + v^2 < R^2\}$ onto S.

Now

$$\mathbf{N}(u, v) = \frac{\partial \mathbf{s}}{\partial u} \times \frac{\partial \mathbf{s}}{\partial v} = \left(1, 0, -\frac{u}{w}\right) \times \left(0, 1, -\frac{v}{w}\right) = \left(\frac{u}{w}, \frac{v}{w}, 1\right),$$

abbreviating $w = \sqrt{R^2 - u^2 - v^2}$, which is the z-coordinate on the sphere. We can see $\mathbf{N}$ points in same direction as $\mathbf{s}$ which is in the direction of the **outward** normal: $\hat{\mathbf{r}} = (u, v, w)/R$. This is not the direction we specified so we must adjust the sign of $\mathbf{N}$ by inserting a minus sign in

$$\begin{aligned} \int_S \boldsymbol{\nabla} \times \mathbf{f} \cdot d\mathbf{S} &= -\int_D \boldsymbol{\nabla} \times \mathbf{f} \cdot \mathbf{N}(u, v)\, dudv \\ &= -\int_{u^2+v^2<R^2} \left(\frac{2\omega_1 u}{w} + \frac{2\omega_1 v}{w} + 2\omega_3\right) dudv = -\int_{u^2+v^2<R^2} 2\omega_3\, dudv = -2\pi\omega_3 R^2 \end{aligned}$$

(using the oddness of the functions w.r.t. u and v in the first two terms).

(ii) Use spherical polars, we define the map

$$\mathbf{s}(\phi, \theta) = R\hat{\mathbf{r}} = (R\sin\phi\cos\theta, R\sin\phi\sin\theta, R\cos\phi),$$

which maps $D = \{(\theta, \phi) \,|\, 0 \leq \phi \leq \pi/2,\ 0 \leq \theta \leq 2\pi\}$ onto the hemisphere.

So now

$$\mathbf{N} = \frac{\partial \mathbf{s}}{\partial \phi} \times \frac{\partial \mathbf{s}}{\partial \theta} = \ldots = R^2(\sin^2\phi\cos\theta, \sin^2\phi\sin\theta, \sin\phi\cos\phi) = R^2\sin\phi\hat{\mathbf{r}},$$

and, as before, $\mathbf{N}$ points outwards on S. Because we defined the integral in terms of a surface pointing inwards we have to reverse the sign of $\mathbf{N}$ and write

$$\int_S \boldsymbol{\nabla} \times \mathbf{f} \cdot d\mathbf{S} = -\int_0^{\pi/2} \int_0^{2\pi} 2\boldsymbol{\omega} \cdot (R^2 \sin\phi\hat{\mathbf{r}})\, d\theta d\phi$$

but

$$\int_0^{2\pi} \boldsymbol{\omega}\cdot\hat{\mathbf{r}}\,d\theta = \int_0^{2\pi} (\omega_1 \sin\phi\cos\theta + \omega_2 \sin\phi\sin\theta + \omega_3\cos\phi)\,d\theta = 2\pi\omega_3\cos\phi.$$

Thus, taken together,

$$\int_S \boldsymbol{\nabla}\times\mathbf{f}\,\cdot d\mathbf{S} = -4\pi R^2\omega_3 \int_0^{\pi/2} \sin\phi\cos\phi\,d\phi = -2\pi R^2\omega_3,$$

the same as before.

(iii) Using Stokes' theorem (21) the answer is

$$\int_{\partial S} \mathbf{f}\cdot d\mathbf{r}$$

where ∂S is the edge of the hemisphere, radius R, where it meets $z=0$. This is the circle of radius R in the (x,y)-plane. By the RH thumb rule, the integral needs to be directed in the *clockwise* direction (looking from above).

We define the circle by the path $\mathbf{p}(\theta) = R\hat{\mathbf{r}} = (R\cos\theta, R\sin\theta, 0)$, $0\le\theta<2\pi$. Now

$$\int_{\partial S} \mathbf{f}\cdot d\mathbf{r} = \int_{2\pi}^{0} \mathbf{f}(\mathbf{p}(\theta))\cdot\mathbf{p}'(\theta)\,d\theta,$$

and we have

$$\mathbf{p}'(\theta) = R(-\sin\theta,\cos\theta) = R\hat{\boldsymbol{\theta}},$$

(in cylindrical polars) and on ∂S.

$$\int_{\partial S} \mathbf{f}\cdot d\mathbf{r} = \int_{2\pi}^{0} (\boldsymbol{\omega}\times(R\hat{\mathbf{r}}))\cdot R\hat{\boldsymbol{\theta}}\,d\theta = R^2\int_{2\pi}^{0} (\hat{\mathbf{r}}\times\hat{\boldsymbol{\theta}})\cdot\boldsymbol{\omega}\,d\theta$$

using a vector result $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})$. Since $\hat{\mathbf{r}}\times\hat{\boldsymbol{\theta}} = \hat{\mathbf{z}}$ we end up with

$$\int_{\partial S} \mathbf{f}\cdot d\mathbf{r} = R^2\omega_3\int_{2\pi}^{0} d\theta = -2\pi R^2\omega_3,$$

the same as before (and confirms Stokes' theorem, for this example).

3.4.1 Outline proof of Stokes' theorem (non-exam)

We first prove for a rectangle in (u,v)-space. The loose argument then proceeds that rectangles can be assembled as a checkboard into larger domains, given that the limit of rectangle size can be taken to go to zero and since contributions from adjacent sides cancel leaving just the circuit around the total domain. See fig 3.

Let $D = \{(u,v)\,|\,0<u<a,\ 0<v<b\}$. The surface S is defined by the map $\mathbf{s}(u,v) : D\to\mathbb{R}^3$ The boundary $\partial D = C_1\cup C_2\cup C_3\cup C_4$ is mapped by $\mathbf{s}$ onto $\partial S = \partial S_1\cup\partial S_2\cup\partial S_3\cup\partial S_4$.

For e.g. C_2 is the path $\mathbf{p}(t) = \mathbf{s}(a,t)$, $0 < t < b$ and

$$\int_{\partial S_2} \mathbf{A} \cdot d\mathbf{r} = \int_0^b \mathbf{A}(\mathbf{p}(t)) \cdot \mathbf{p}'(t)\, dt = \int_0^b \mathbf{A}(\mathbf{s}(a,v)) \cdot \frac{\partial \mathbf{s}}{\partial v}(a,v)\, dv$$

Similarly,

$$\int_{\partial S_4} \mathbf{A} \cdot d\mathbf{r} = -\int_0^b \mathbf{A}(\mathbf{s}(0,v)) \cdot \frac{\partial \mathbf{s}}{\partial v}(0,v)\, dv$$

(the minus sign accounts for reversing the orientation of the segment C_4, viz: $\int_b^0 = -\int_0^b$)

Combining results gives

$$\begin{aligned} \left(\int_{\partial S_2} + \int_{\partial S_4}\right) \mathbf{A}.d\mathbf{r} &= \int_0^b \left(\mathbf{A}(\mathbf{s}(a,v)) \cdot \frac{\partial \mathbf{s}}{\partial v}(a,v) - \mathbf{A}(\mathbf{s}(0,v)) \cdot \frac{\partial \mathbf{s}}{\partial v}(0,v)\right) dv \\ &= \int_0^b \int_0^a \frac{\partial}{\partial u} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial v}\right) du\, dv \end{aligned}$$

by the FTC. We apply the same method to the side ∂S_1 and ∂S_3 and find

$$\left(\int_{\partial S_1} + \int_{\partial S_3}\right) \mathbf{A}.d\mathbf{r} = -\int_0^b \int_0^a \frac{\partial}{\partial v} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial u}\right) du\, dv$$

and so

$$\int_{\partial S} \mathbf{A}.d\mathbf{r} = \int_D \left(\frac{\partial}{\partial u} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial v}\right) - \frac{\partial}{\partial v} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial u}\right)\right) dudv$$

Concentrate on the integrand of the LHS. So

$$\begin{aligned} \frac{\partial}{\partial u} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial v}\right) - \frac{\partial}{\partial v} \left(\mathbf{A}(\mathbf{s}) \cdot \frac{\partial \mathbf{s}}{\partial u}\right) &= \frac{\partial \mathbf{A}(\mathbf{s})}{\partial u} \cdot \frac{\partial \mathbf{s}}{\partial v} - \frac{\partial \mathbf{A}(\mathbf{s})}{\partial v} \cdot \frac{\partial \mathbf{s}}{\partial u} \\ &= \frac{\partial A_k}{\partial x_j} \left(\frac{\partial x_j}{\partial u} \frac{\partial x_k}{\partial v} - \frac{\partial x_j}{\partial v} \frac{\partial x_k}{\partial u}\right) \\ &= \frac{\partial A_k}{\partial x_j} \left(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}\right) \frac{\partial x_l}{\partial u} \frac{\partial x_m}{\partial v} \\ &= \frac{\partial A_k}{\partial x_j} \epsilon_{ijk}\epsilon_{ilm} \frac{\partial x_l}{\partial u} \frac{\partial x_m}{\partial v} \\ &= \left(\epsilon_{ijk} \frac{\partial A_k}{\partial x_j}\right) \left(\epsilon_{ilm} \frac{\partial x_l}{\partial u} \frac{\partial x_m}{\partial v}\right) \\ &= (\boldsymbol{\nabla} \times \mathbf{A}) \cdot \left(\frac{\partial \mathbf{s}}{\partial u} \times \frac{\partial \mathbf{s}}{\partial v}\right) = (\boldsymbol{\nabla} \times \mathbf{A}) \cdot \mathbf{N}(u,v). \end{aligned}$$

Hence we have

$$\int_{\partial S} \mathbf{A}.d\mathbf{r} = \int_D (\boldsymbol{\nabla} \times \mathbf{A}) \cdot \mathbf{N}(u,v)\, dudv = \int_S (\boldsymbol{\nabla} \times \mathbf{A}) \cdot d\mathbf{S}$$

as required. As mentioned earlier, we now "glue together" small rectangles to create the actual domain we wish to cover. This can be done formally.

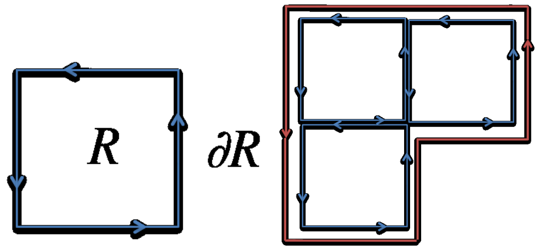

Figure 3: A number of rectangles (left) can be put together to cover the domain (right).

3.4.2 Green's theorem in the plane

Stokes' theorem is applied in 2D. Let S be a surface on $z = 0$ and $\mathbf{A} = (A_1(x, y), A_2(x, y), 0)$ is a vector field with no $\hat{\mathbf{z}}$ component and no dependence on z. Now $d\mathbf{S} = \hat{\mathbf{z}}dS = \hat{\mathbf{z}}dxdy$ and

$$\boldsymbol{\nabla} \times \mathbf{A} = \left(\frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y}\right)\hat{\mathbf{z}}.$$

Thus Stokes' theorem is reduced to

$$\int_S \left(\frac{\partial A_2}{\partial x} - \frac{\partial A_1}{\partial y}\right) dxdy = \int_{\partial S} \mathbf{A} \cdot d\mathbf{r} \equiv \int_{\partial S} A_1\, dx + A_2\, dy$$

since $d\mathbf{r} = dx\hat{\mathbf{x}} + dy\hat{\mathbf{y}}$ and ∂S is anti-clockwise by the RH Thumb rule.

3.5 Volume integrals

3.5.1 Volume integrals of scalar fields

Defn: Let $f : \mathbb{R}^3 \to \mathbb{R}$ s.t. $\mathbf{r} \mapsto f(\mathbf{r})$ be a scalar field. The **volume integral** of f is given by

$$\int_V f(\mathbf{r})\, dV \equiv \iiint f(x, y, z)\, dx\, dy\, dz$$

in Cartesians.

Note: Unlike curves and surfaces, volumes in $\mathbb{R}^3$ do not have directions.

Note: if $f = 1$, then $\int_V 1\, dV$ gives the physical volume of V.

Proposition: If we move to a different coordinate system, $\mathbf{q} = (q_1, q_2, q_3)$ from $\mathbf{r} = (x, y, z)$ under the mapping $\mathbf{r} : \mathbb{R}^3 \to \mathbb{R}^3$ s.t. $\mathbf{q} \mapsto \mathbf{r}(\mathbf{q})$ then

$$\int_V f(\mathbf{r})\, dxdydz = \int_{V_q} f(\mathbf{r}(\mathbf{q})) \left|\frac{\partial(x, y, z)}{\partial(q_1, q_2, q_3)}\right| dq_1dq_2dq_3$$

where V_q is mapped by $\mathbf{r}$ into V. The scale factor is the Jacobian determinant of the mapping.

Proof: The elemental volume $dV = dxdydz$ is $(\hat{\mathbf{z}}dz)\cdot((\hat{\mathbf{x}}dx)\times(\hat{\mathbf{y}}dy))$. Under the mapping, the mapped elemental volume dV_q is defined by a parallelepiped with sides given by

$$\frac{\partial \mathbf{r}}{\partial q_1}dq_1, \quad \frac{\partial \mathbf{r}}{\partial q_2}dq_2, \quad \text{and} \quad \frac{\partial \mathbf{r}}{\partial q_3}dq_3$$

(just as we did for surfaces). The volume of dV_q is therefore

$$|(\hat{\mathbf{q}}_3 h_3 dq_3)\cdot((\hat{\mathbf{q}}_1 h_1 dq_1)\times(\hat{\mathbf{q}}_2 h_2 dq_2))| = |J_{\mathbf{r}}|\, dq_1 dq_2 dq_3.$$

since $\hat{\mathbf{q}}_\alpha = \dfrac{1}{h_\alpha}\dfrac{\partial \mathbf{r}}{\partial q_\alpha}$ and using a result in §2.8.1. If $\hat{\mathbf{q}}_j$ are orthonormal, then $|J_{\mathbf{r}}| = h_1 h_2 h_3$.

3.6 Divergence theorem (Gauss' theorem)

Defn: A **simply connected** domain, V, say, is one in which all paths within V can be continuously transformed into all other paths within V without ever leaving V.

If V is finite and simply connected then ∂V forms a **closed surface** (a surface with no boundaries).

Proposition: For a vector field $\mathbf{F}:\mathbb{R}^3\to\mathbb{R}^3$, Let $V\subset\mathbb{R}^3$ be simply connected with boundary ∂V (a closed surface). Then

$$\int_V \boldsymbol{\nabla}\cdot\mathbf{F}\,dV = \int_{\partial V}\mathbf{F}\cdot d\mathbf{S}$$

where $d\mathbf{S} = \hat{\mathbf{n}}dS$ is a surface element and $\hat{\mathbf{n}}$ points **outwards** from the volume V.

3.6.1 Outline proof of the divergence theorem (non-exam)

As in Stokes' theorem, start with a proof for a cuboid

$$V = \{\mathbf{r}\,|\,0<x<a, 0<y<b, 0<z<c\}.$$

The argument will be again that an arbitrary V can be divided into many small rectangular volumes over each of which the divergence applies.

We write $\mathbf{F} = F_1\hat{\mathbf{x}} + F_2\hat{\mathbf{y}} + F_3\hat{\mathbf{z}}$. Then it follows that

$$\begin{aligned}\int_{\partial V}\mathbf{F}\cdot d\mathbf{S} = &\int_0^a\int_0^b (F_3(x,y,c)-F_3(x,y,0))\,dy\,dx\,+\\ &\int_0^a\int_0^c (F_2(x,b,z)-F_2(x,0,z))\,dz\,dx + \int_0^b\int_0^c (F_1(a,y,z)-F_1(0,y,z))\,dz\,dy\end{aligned} \tag{22}$$

(there are 6 sides, and unit outward normal is one of $\pm\hat{\mathbf{x}}$, $\pm\hat{\mathbf{y}}$, $\pm\hat{\mathbf{z}}$ depending on the cuboid side).

Next we consider the volume integral,

$$\int_V \boldsymbol{\nabla}\cdot\mathbf{F}\,dV = \int_0^a\int_0^b\int_0^c\left(\frac{\partial F_1}{\partial x}+\frac{\partial F_2}{\partial y}+\frac{\partial F_3}{\partial z}\right)dz\,dy\,dx.$$

The 3 terms are considered separately but in the same manner. For example, consider the contribution from $\partial F_3/\partial z$. From the Fundamental Theorem of Calculus,

$$\int_0^a\int_0^b\int_0^c \frac{\partial F_3}{\partial z}\,dz\,dy\,dx = \int_0^a\int_0^b (F_3(x,y,c)-F_3(x,y,0))\,dy\,dx.$$

The result is

$$\begin{aligned}\int_V \boldsymbol{\nabla}\cdot\mathbf{F}\,dV \;&=\; \int_0^a\int_0^b (F_3(x,y,c)-F_3(x,y,0))\,dy\,dx\,+\\ &\int_0^a\int_0^c (F_2(x,b,z)-F_2(x,0,z))\,dz\,dx + \int_0^b\int_0^c (F_1(a,y,z)-F_1(0,y,z))\,dz\,dy,\end{aligned}$$

which coincides with (22), thus confirming the theorem.

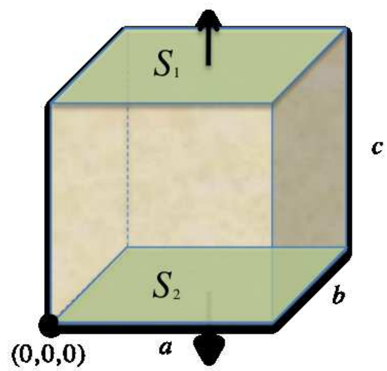

Figure 4: Gauss' theorem for the cuboid V. The top and bottom faces of the boundary, S_1 and S_2, are indicated.

E.g. 3.7: Compute $\int_V \boldsymbol{\nabla}\cdot\mathbf{v}\,dV$ where V is a sphere of radius a about the origin, and

$$\mathbf{v}(\mathbf{r}) = \mathbf{r} + f(r)\hat{\mathbf{z}}\times\mathbf{r},$$

(and $\hat{\mathbf{z}}$ is the unit vector along the z-axis, and $r=(x^2+y^2+z^2)^{1/2}$).

A (i): We have that

$$\boldsymbol{\nabla}\cdot\mathbf{v} = 3 + (\boldsymbol{\nabla}f(r))\cdot(\hat{\mathbf{z}}\times\mathbf{r}) + f(r)\boldsymbol{\nabla}\cdot(\hat{\mathbf{z}}\times\mathbf{r})$$

after using PS2, Q6(a). But $\boldsymbol{\nabla}f = \dfrac{f'(r)}{r}\mathbf{r}$ and $\mathbf{r}\cdot(\hat{\mathbf{z}}\times\mathbf{r})=0$. Also, can be shown that $\boldsymbol{\nabla}\cdot(\hat{\mathbf{z}}\times\mathbf{r})=0$, so that

$$\boldsymbol{\nabla}\cdot\mathbf{v} = 3.$$

As the divergence of $\mathbf{v}$ is a constant, its integral over V is just its value times the volume of V,

$$\int_V \boldsymbol{\nabla}\cdot\mathbf{v}\;\mathrm{d}V = 3\frac{4\pi}{3}a^3 = 4\pi a^3.$$

A (ii): Using the divergence theorem, answer is $\int_{\partial V} \mathbf{v}(\mathbf{r}) \cdot d\mathbf{S}$ where ∂V is the sphere of radius a. Surface parametrised using spherical polars by

$$\mathbf{s}(\phi, \theta) = a\hat{\mathbf{r}} = a(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi),$$

over domain $D = \{(\phi, \theta)|\ 0 \le \phi \le \pi,\ 0 \le \theta < 2\pi\}$. Then

$$\int_{\partial V} \mathbf{v} \cdot d\mathbf{S} = \int_0^{\pi} \int_0^{2\pi} \mathbf{v}(\mathbf{s}(\phi, \theta)) \cdot \mathbf{N}(\phi, \theta)\, d\theta\, d\phi.$$

We have (confirm yourselves)

$$\frac{\partial \mathbf{r}}{\partial \phi} = a\hat{\boldsymbol{\phi}}, \quad \frac{\partial \mathbf{r}}{\partial \theta} = a\sin\phi\hat{\boldsymbol{\theta}},$$

so that

$$\mathbf{N}(\phi, \theta) = a^2 \sin\phi\hat{\mathbf{r}}.$$

Therefore,

$$\mathbf{v}(\mathbf{s}(\phi, \theta)) \cdot \mathbf{N}(u, v) = (a\hat{\mathbf{r}} + f(a)\hat{\mathbf{z}} \times (a\hat{\mathbf{r}})) \cdot a^2 \sin\phi\hat{\mathbf{r}} = a^3 \sin\phi.$$

The surface integral is given by

$$\int_{\partial V} \mathbf{v} \cdot d\mathbf{S} = \int_0^{\pi} \int_0^{2\pi} a^3 \sin\phi\, d\theta\, d\phi = 4\pi a^3.$$

3.6.2 Green's Identities

If $\mathbf{F} = \boldsymbol{\nabla} f$ (i.e. the vector field can be described by a scalar potential) then the divergence theorem reads

$$\int_V \triangle f\, dV = \int_{\partial V} \hat{\mathbf{n}}.\boldsymbol{\nabla} f\, dS$$

If $\mathbf{F} = g\boldsymbol{\nabla} f$, g, f scalar fields then

$$\int_V \boldsymbol{\nabla} g \cdot \boldsymbol{\nabla} f + g\triangle f\, dV = \int_{\partial V} g\hat{\mathbf{n}} \cdot \boldsymbol{\nabla} f\, dS$$

subtracting the result of using $\mathbf{F} = f\boldsymbol{\nabla} g$ we have

$$\int_V (g\triangle f - f\triangle g)\, dV = \int_{\partial V} (g\hat{\mathbf{n}} \cdot \boldsymbol{\nabla} f - f\hat{\mathbf{n}} \cdot \boldsymbol{\nabla} g)\, dS$$

These can be very useful in deriving equations underlying physical applications.

Appendix: revision of the chain rule

Taylor series: Recall for a scalar function $f(x)$,

$$f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!}f''(x_0) + \dots$$

or, equivalently, $(x = x_0 + h)$

$$f(x) = f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2!}f''(x_0) + \dots$$

E.g. A.1: If $f(x) = \cos x$, with $x_0 = 0$ we get

$$\cos h = 1 - \frac{h^2}{2!} + \dots$$

Proposition: Start with **chain rule** for a **scalar function of a single variable** (sometimes referred to a differentiation of a function of a function). Consider a function $f(x)$ such that $x = x(t)$. Then $F(t) = f(x(t))$ and

$$\frac{dF}{dt} = x'(t)f'(x(t))$$

E.g. A.2: If $f(x) = \cos x$ and $x(t) = t^2$ then $x'(t) = 2t$ $f'(x) = \sin x$ and so we get

$$\frac{d}{dt}\cos(t^2) = 2t\sin(t^2)$$

Proof: Standard limit used to define a derivative, along with Taylor series expansions (notation $O(h^2)$ means collect terms as small as and smaller than h^2)

$$\begin{aligned}
\frac{d}{dt}f(x(t)) &= \lim_{h\to 0}\frac{f(x(t+h)) - f(x(t))}{h} \\
&= \lim_{h\to 0}\frac{f(x(t) + hx'(t) + O(h^2)) - f(x(t))}{h} \\
&= \lim_{h\to 0}\frac{f(x(t)) + hx'(t)f'(x(t) + O(h)) + O(h^2) - f(x(t))}{h} \\
&= x'(t)f'(x(t))
\end{aligned}$$

Now consider a **scalar function of more than one variable**, $f(x,y)$, say, and let $x = x(t)$ and $y = y(t)$.

That is, we can write $F(t) = f(x(t), y(t))$, say. From the chain rule, it follows that

$$\frac{dF}{dt} = \frac{dx}{dt}\frac{\partial f}{\partial x}(x(t), y(t)) + \frac{dy}{dt}\frac{\partial f}{\partial y}(x(t), y(t))$$

Proof: Similar to before and requires the extension to multiple variables of Taylor's expansion:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + hf_x(x_0, y_0) + kf_y(x_0, y_0) + \text{higher order terms}$$

E.g. A.3: If $f(x, y) = xy$ and $x(t) = t^2$ and $y(t) = \mathrm{e}^{-t}$ then $F(t) = t^2\mathrm{e}^{-t}$. Then we can see that by a direct calculation that

$$F'(t) = 2t\mathrm{e}^{-t} - t^2\mathrm{e}^{-t}$$

Using the chain rule, we get $x'(t) = 2t$, $y'(t) = -\mathrm{e}^{-t}$, $f_x = y$, $f_y = x$ and so

$$F'(t) = (2t)\mathrm{e}^{-t} - \mathrm{e}^{-t}t^2$$

and we get the same answer.

Note: Clearly this extends to scalar functions of more than two variables, so that if we have the function $f(x_1, \ldots, x_m)$ and $x_i = x_i(t)$, for $i = 1, 2, \ldots, m$ then with $F(t) = f(x_1(t), \ldots, x_m(t))$

$$\frac{dF}{dt} = \sum_{i=1}^{m} \frac{dx_i}{dt}\frac{\partial f}{\partial x_i}(x_1(t), \ldots, x_m(t))$$

Put another way, $F(t) = f(\mathbf{x})$ where $\mathbf{x} = \mathbf{x}(t)$ and the the summation above can be either interpreted as row times column vectors or as a dot product:

$$F'(t) = (\boldsymbol{\nabla} f)^T\mathbf{x}'(t) \equiv \boldsymbol{\nabla} f \cdot \mathbf{x}'(t).$$

noting that $\boldsymbol{\nabla} f$ is evaluated at $\mathbf{x}(t)$.

Next, consider **scalar functions of more than one variable, each of which is a function of more than one variable**. The simplest example is to consider $f(x, y)$ where $x = x(u, v)$ and $y = y(u, v)$. Then $F(u, v) = f(x(u, v), y(u, v))$ and the chain rule gives

$$\frac{\partial F}{\partial u} = \frac{\partial x}{\partial u}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial f}{\partial y}$$

and

$$\frac{\partial F}{\partial v} = \frac{\partial x}{\partial v}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial f}{\partial y}$$

This can be arranged as a matrix/vector relation as

$$(F_u, F_v) = (f_x, f_y)\begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix}$$

Note: The 2×2 matrix is the Jacobian of the map $(u, v) \mapsto (x(u, v), y(u, v))$, the 1×2 matrix (F_u, F_v) is the Jacobian of of the map $(u, v) \mapsto F(u, v)$ and and the 1×2 matrix (f_x, f_y) is the Jacobian of of the map $(x, y) \mapsto f(x, y)$.

Note: Again, we can see through a way of extending this to a more general case in which $f = f(x_1, \ldots, x_m)$ and $x_i = x_i(u_1, \ldots, u_p)$. Then we can write

$$F(u_1, \ldots, u_n) = f(x_1(u_1, \ldots, u_p), \ldots, x_m(u_1, \ldots, u_p))$$

or

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

and application of the chain rule gives

$$\frac{\partial F}{\partial u_j} = \sum_{k=1}^{m} \frac{\partial x_k}{\partial u_j} \frac{\partial f}{\partial x_k}$$

for $j = 1, \ldots, p$. This can be intepreted as the matrix/vector relation

$$(F_{u_1}, \ldots, F_{u_n}) = (f_{x_1}, \ldots, f_{x_m}) \begin{pmatrix} \partial x_1/\partial u_1 & \ldots & \partial x_1/\partial u_n \\ \vdots & \ddots & \vdots \\ \partial x_m/\partial u_1 & \ldots & \partial x_m/\partial u_n \end{pmatrix}.$$

The final generalisation of this is to consider a **vector function** where $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ s.t. $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$ and another map $\mathbf{x} = \mathbf{x}(\mathbf{u})$ where $\mathbf{x} : \mathbb{R}^p \to \mathbb{R}^m$. Then if we define $\mathbf{F}(\mathbf{u}) = \mathbf{f}(\mathbf{x}(\mathbf{u}))$ and apply the chain rule as above to each scalar component, F_i, $i = 1, \ldots, n$ of $\mathbf{F}$ we get

$$\frac{\partial F_i}{\partial u_j} = \sum_{k=1}^{m} \frac{\partial x_k}{\partial u_j} \frac{\partial f_i}{\partial x_k}$$

or

$$\begin{pmatrix} \partial F_1/\partial u_1 & \ldots & \partial F_1/\partial u_p \\ \vdots & \ddots & \vdots \\ \partial F_n/\partial u_1 & \ldots & \partial F_n/\partial u_p \end{pmatrix} = \begin{pmatrix} \partial f_1/\partial x_1 & \ldots & \partial f_1/\partial x_m \\ \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \ldots & \partial f_n/\partial x_m \end{pmatrix} \begin{pmatrix} \partial x_1/\partial u_1 & \ldots & \partial x_1/\partial u_p \\ \vdots & \ddots & \vdots \\ \partial x_m/\partial u_1 & \ldots & \partial x_m/\partial u_p \end{pmatrix}.$$

At this high level or generality we still have the same underlying structure from the first result. Thus the equation above reads $\mathbf{F}'(\mathbf{u}) = \mathbf{f}'(\mathbf{x}(\mathbf{u}))\mathbf{x}'(\mathbf{u})$, which is the result quoted in the notes for the chain rule applies to the composition of maps, although the notation is shifted here from there.