Dr M. Dörrzapf
Lent Term 2006

Vector Calculus

Mathematical Tripos, Part IA
Lecture Notes

I Introduction and Notation

This course is about functions of more than one variable and develops 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. It is an applied course, meaning that you are expected to be able to apply techniques, but we will not necessarily prove rigorously that they work, this will be done in future Analysis courses. In the first part of the course, the idea of integration is extended from $\mathbb{R}$ to $\mathbb{R}^2$ and $\mathbb{R}^3$ (with an obvious extension to higher dimensions): integrals along the x-axis are replaced by integrals over curves, surfaces and volumes. Then the idea of differentiation is extended to vectors (div, grad and curl), which is a basic tool in many areas of theoretical physics (such as Electromagnetism and Fluid Dynamics). Two important theorems are introduced, namely the divergence theorem and Stokes' theorem; in both cases, an integral over a region (in $\mathbb{R}^3$ and in $\mathbb{R}^2$, respectively) is converted to an integral over the boundary of the region. All the previous ideas are then applied to Laplace's equation $\boldsymbol{\nabla}^2\phi = 0$, which is one of the most important equations in the whole of mathematics and physics. Finally, the notion of a vector is generalised to that of a tensor. A vector can be thought of as a 3×1 matrix that carries physical information: namely, magnitude and direction. This information is preserved when the axes are rotated only if the components change according to a certain rule. Very often, it is necessary to describe physical quantities using a 3×3 matrix (or even a $3 \times 3 \times 3 \ldots$ matrix). Such a quantity is called a tensor if its components transform according to a certain rule when the axes are rotated. This rule means that the physical information in the tensor (essentially the eigenvalues) is preserved.

By the end of this course, you should be able to manipulate, and solve problems using vector operators; be able to calculate line, surface and volume integrals in $\mathbb{R}^3$, using Stokes' theorem and the divergence theorem; be able to solve Laplace's equation in simple cases, and be able to prove standard uniqueness theorems for Laplaces and related equations; understand the notion of a tensor and the general properties of tensors in simple cases.

The course is structured in the following way:

I Introduction and Notation

Introduction, books, notation, important results from Part IA Differential Equations.

II Curves in $\mathbb{R}^3$

Parameterised curves and arc length, tangents and normals to curves in $\mathbb{R}^3$, the radius of curvature. [1 hour]

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

Line integrals. Surface and volume integrals: definitions, examples using Cartesian, cylindrical and spherical coordinates; change of variables. [4 hours]

IV Vector Operators

Directional derivatives. The gradient of a real-valued function: definition; interpretation as normal to level surfaces; examples including the use of cylindrical, spherical and general orthogonal curvilinear coordinates; conservative fields. Divergence, curl and ∇^2 in Cartesian coordinates, examples; formulae for these operators (statement only) in cylindrical, spherical and general orthogonal curvilinear coordinates. Solenoidal fields and irrotational fields. Vector derivative identities. [5 hours]

V Integration Theorems

Green's theorem in the plane, Divergence theorem, Stokes' theorem, Green's first and second theorem: statements; informal proofs; examples; application to fluid dynamics and electromagnetism. [5 hours]

VI Laplace's Equation

Laplace's equation in $\mathbb{R}^2$ and $\mathbb{R}^3$: uniqueness theorem and maximum principle. Solution of Poisson's equation by Gauss' method (for spherical and cylindrical symmetry) and as an integral. [4 hours]

VII Cartesian Tensors in $\mathbb{R}^3$

Tensor transformation laws, addition, multiplication, contraction, with emphasis on tensors of second rank. Isotropic second and third rank tensors. Symmetric and antisymmetric tensors. Revision of principal axes and diagonalization. Quotient theorem. Examples including inertia and conductivity. [5 hours]

You will find the in the present lecture notes all relevant definitions, theorems and important equations discussed in the lectures. However the notes are not meant to give a full presentation of the course material neither are they intended to replace attendance of the lectures. The notes also contain worked examples at the end of each chapter. These notes and supporting material can be found at http://www.damtp.cam.ac.uk/user/md131/vector_calculus. Please send corrections and comments for improvements to md131@cam.ac.uk.

Books

S. J. Colley. *Vector Calculus.* Prentice Hall 2005.
P.C. Matthews. *Vector Calculus.* Springer 1998.
K.F. Riley, M.P. Hobson and S.J. Bence. *Mathematical Methods for Physicists and Engineers.* Cambridge University Press 1998.
J.E. Marsden and A.J. Tromba. *Vector Calculus.* Freeman 1996.
D.E. Bourne and P.C. Kendall. *Vector Analysis and Cartesian Tensors.* Chapman and Hall 1992.
T.M. Apostol. *Calculus.* Wiley 1975.
H.M. Schey. *Div, grad, curl and all that: an informal text on vector calculus.* Norton 1996.
H. Anton. *Calculus.* Wiley 1995.
M.L. Boas. *Mathematical Methods in the Physical Sciences.* Wiley 1983.
E. Kreyszig. *Advanced Engineering Mathematics.* Wiley 1999.
M.R. Spiegel. *Schaum's outline of Vector Analysis.* McGraw Hill 1974.

Notation

The following is a table of the notation used in this course.

$\forall$	for all	$\mathbb{R}$	real numbers
$\exists$	there exists	$\mathbb{C}$	complex numbers
$\exists_1$	there exists a unique	$\mathbb{Z}$	integers
$\nexists$	there is no	$\mathbb{N}$	positive integers, $\mathbb{N} := \{1, 2, 3, \dots\}$
$:=$	defined equal	$\mathbb{N}_0$	non-negative integers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$
$\Rightarrow$	implication	$[a, b]$	closed interval $x \in \mathbb{R}$: $a \leq x \leq b$
$\therefore$	therefore	(a, b)	open interval $x \in \mathbb{R}$: $a < x < b$
$\because$	because	$\in$	contained in
$\Leftrightarrow$	equivalence	$\notin$	not contained in
$:\Leftrightarrow$	defined equivalent	$\wedge$	and
$\stackrel{!}{=}$	should be equal to	$\vee$	or
$:$	such that	$\perp$	perpendicular
$\equiv$	identical	$\square$	end of proof
$\approx$	approximately equal	$\lightning$	contradiction

Examples

$A := \{x \in \mathbb{R} : x < 2\}$	A is defined to be the set of all real x such that $x < 2$.
$A \subset B \ :\Leftrightarrow \ x \in A \Rightarrow x \in B$	A is a subset of B is defined to be equivalent to 'x is contained in A implies that x is also contained in B'.
$A = B \ :\Leftrightarrow \ A \subset B \ \wedge \ B \subset A$	the set A equals the set B is defined to be equivalent to 'A is a subset of B and B is a subset of A'.
$\{x \in \mathbb{R} : 1 \leq x \leq 2\} =: [1, 2]$	the closed interval $[1, 2]$ is defined to be the set of real numbers between and including 1 and 2.

Labels

Equations are referred to by 'Eq. number', e.g. Eq. 1, theorems and definitions by [number] e.g. [1] and sections by 'Sec. number' e.g. Sec. 1. Finally, figures are referred to by Fig. iii. For this course non-examinable sections are indicated with a * behind the number.

Vectors

We will denote vectors using underlining e.g. $\underline{x}$ in handwritten formulae and bold face type e.g. $\mathbf{x}$ in print. In this course we will mainly operate in $\mathbb{R}^2$ or $\mathbb{R}^3$. The ith component of the vector $\mathbf{x}$, i.e. $(\mathbf{x})_i$, is always denoted x_i. The notation $\mathbf{x}_i$ would mean the ith of a set of vectors $\{\mathbf{x}_1, \ \mathbf{x}_2, \dots\}$. Unless otherwise stated, summation convention applies whenever there is a repeated suffix. The standard scalar product is defined by $\mathbf{x}.\mathbf{y} := x_i y_i$. The norm $\|\mathbf{x}\|$ is defined to be the *Euclidean*[a] norm $\|\mathbf{x}\| := \sqrt{\sum_{i=1}^{n} x_i^2}$. Vectors $\mathbf{x}$ are usually assumed to be column vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

but in some instances we will need $\mathbf{x}$ to be a row vector. If it is required for clarity then may explicitly write the *transpose* of the vector: $\mathbf{x}^T = (x_1, \dots, x_n)$ but otherwise we may label row vectors also just by $\mathbf{x} = (x_1, \dots, x_n)$.

[a] Euclid of Alexandria, about 325BC-265BC.

The vectors of an orthonormal triad are denoted by $\mathbf{e}_i$, where

$$\mathbf{e}_i.\mathbf{e}_j = \delta_{ij} \ .$$

The vector product between two vectors is defined by

$$\mathbf{x} \wedge \mathbf{y} \equiv \mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \ ; \quad \text{and} \quad (\mathbf{x} \times \mathbf{y})_i = \epsilon_{ijk} x_j y_k \ .$$

The identity

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

can be obtained using the identity

$$\epsilon_{kij}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl} \ .$$

The position vector is denoted $\mathbf{r}$ or also $\mathbf{x}$, and in Cartesians $\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ or $(\mathbf{r})_i = x_i$. The length of the position vector is r, where $r^2 = \mathbf{r}.\mathbf{r} = \mathbf{x}.\mathbf{x}$.

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

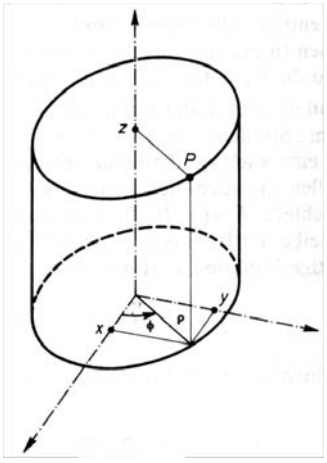


FIG. i Cylindrical polar co-ordinates

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. With respect to Cartesian axes, the position vector is

$$\mathbf{x} = (\rho\cos\phi, \rho\sin\phi, \ z) \, , \tag{1}$$

where $0 \leq \rho < \infty$, $0 \leq \phi \leq 2\pi$ and $-\infty < z < \infty$.
In terms of the orthonormal basis related to the cylindrical coordinates, $\mathbf{e}_\rho$, $\mathbf{e}_\phi$, $\mathbf{e}_z$, the position vector is given by

$$\mathbf{r} = \rho\mathbf{e}_\rho + z\mathbf{e}_z.$$

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

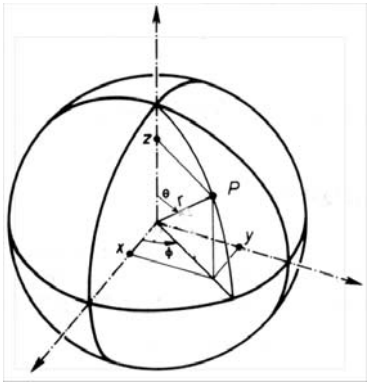


FIG. ii Spherical polar co-ordinates

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. With respect to Cartesian axes, the position vector is

$$\mathbf{x} = (r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta)\,, \qquad (2)$$

where $0 \leq r < \infty$, $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. In terms of the orthonormal basis $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\phi$, the position vector is given by

$$\mathbf{x} = r\mathbf{e}_r.$$

Functions

The functions we consider in this course map $\mathbf{x} \in \mathbb{R}^n$ to a vector in $\mathbb{R}^m$ which we denote by $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Usually m and n will either be $1, 2$ or 3. The domain of a function f is denoted by dom$f$. For $m = 1$ we simply say a *real valued* function and for $m > 1$ we call the function *vector valued*. Most of the time the domain and the range of a function is self-understood. If a function is defined by simply giving an expression for $f(\mathbf{x})$ then we mean that dom$f$ is the maximal possible domain and the range is self-understood. For example, the function

$$f = \left\{ \begin{array}{l} \mathbb{R}^2 \to \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto f(x,y) = e^{-(x^2+y^2)} \end{array} \right. \qquad (3)$$

can simply be written as $f(x, y) = e^{-(x^2+y^2)}$. Since dom$f = \mathbb{R}^2$ and the range is $\mathbb{R}$ we can sketch the graph of such a function in a three dimensional co-ordinate system. The graph of this particular function is given in the following figure.

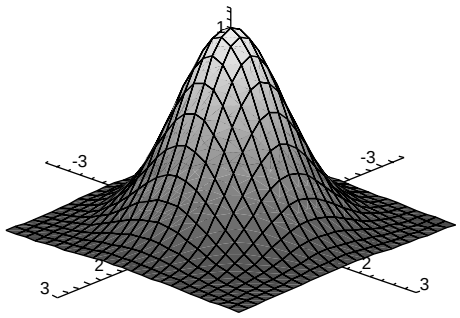


FIG. iii Graph of $f(x, y) = e^{-(x^2+y^2)}$

Sometimes it may be necessary to define functions by sticking different branches together or by defining certain points separately. The notation we use in this case becomes clear in the following example.

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & ,(x,y) \neq (0,0) \\ 0 & ,(x,y) = (0,0) \end{cases} \tag{4}$$

A function $f(x,y)$ as simple as the one in Eq. 4 can in fact have a very complicated graph as the following figure shows.

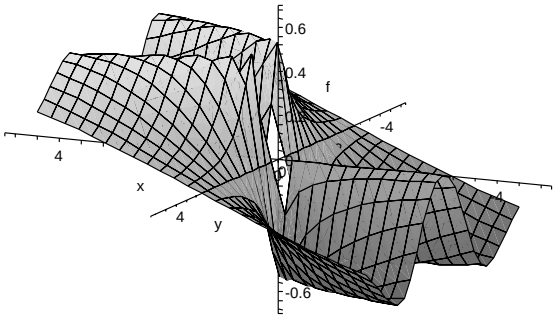


Fig. iv $\frac{xy^2}{x^2+y^4}$

Partial derivatives

In the Part IA course on Differential Equations the concept of differentiation has been generalised to functions $f(\mathbf{x})$ of more than one variable leading to the concept of partial differentiation. A sound knowledge of these results is required for this course. We will therefore give the following summary of the main results.

Simple partial derivatives (e.g. $\frac{\partial f(x,y,z)}{\partial x}$) are effectively ordinary one dimensional derivatives where we simply treat all variables but one as constants. As a result all differentiation rules we know from one dimensional derivatives can easily be applied to partial derivatives. For example:

$$\begin{aligned} \frac{\partial x^2 \sin(ye^z)}{\partial x} &= 2x\sin(ye^z)\,, \\ \frac{\partial x^2 \sin(ye^z)}{\partial y} &= x^2\cos(ye^z)e^z\,, \\ \frac{\partial \frac{\cos(x+y)}{x+y}}{\partial x} &= -\frac{\sin(x+y)}{x+y} - \frac{\cos(x+y)}{(x+y)^2}\,. \end{aligned}$$

Instead of $\frac{\partial f(x,y,z)}{\partial x}$ we may also write f_x or simply f_1 where the latter should not be confused with the first component of a vector $\mathbf{f}$. In case we only want to keep some of the other

variables constant but not all of them then we have to say explicitly which variables are being kept constant e.g. $\left(\frac{\partial f(x,y,z)}{\partial x}\right)_y$ keeps y invariant but not z and therefore takes account of any changes in z if x is being varied. The standard partial derivative $\frac{\partial f(x,y,z)}{\partial x}$ can in this notation also be written as $\left(\frac{\partial f(x,y,z)}{\partial x}\right)_{y,z}$. If none of the variables are being kept constant then we call it the total derivative e.g. $\frac{df(x,y,z)}{dx}$. The chain rule relates these derivatives to simple partial derivatives:

$$\frac{d}{dt}f(x_1(t),x_2(t),\ldots,x_n(t)) = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt}+\frac{\partial f}{\partial x_2}\frac{dx_2}{dt}+\ldots+\frac{\partial f}{\partial x_n}\frac{dx_n}{dt}\,, \tag{5}$$

and more generally for $f(x_1,\ldots,x_n)$:

$$\left(\frac{\partial f}{\partial t}\right)_{x_{i_1},\ldots,x_{i_k}} = \frac{\partial f}{\partial x_1}\left(\frac{\partial x_1}{\partial t}\right)_{x_{i_1},\ldots,x_{i_k}}+\ldots+\frac{\partial f}{\partial x_n}\left(\frac{\partial x_n}{\partial t}\right)_{x_{i_1},\ldots,x_{i_k}}\,. \tag{6}$$

A particular case of Eq. 6 for the function $f(x,y,z)$ is for example

$$\begin{aligned}\left(\frac{\partial f}{\partial x}\right)_z &= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial x}\right)_z+\frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial x}\right)_z+\frac{\partial f}{\partial z}\left(\frac{\partial z}{\partial x}\right)_z\\ &= \frac{\partial f}{\partial x}1+\frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial x}\right)_z+\frac{\partial f}{\partial z}0=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial x}\right)_z\,.\end{aligned} \tag{7}$$

The Analysis of the rate of change of a multi dimensional function can easily become very difficult due to the fact that in multi dimensions we can approach points not only along straight lines but also along complicated curves such as spirals etc.. If we assume that the function is sufficiently smooth then these difficulties will not occur. We will leave it to future Analysis courses to study this in more detail. For this course we shall always assume that all our functions are continuous and sufficiently smooth and well-behaved (which basically means that the functions are continuous and all the partial derivatives we need exist and are again continuous). In this case a range of nice theorems hold:

1 **Schwarz's**[b] **theorem**
Mixed partial derivatives are independent of the order of differentiation e.g.:

$$\frac{\partial^2 f(x,y)}{\partial x\partial y}=\frac{\partial^2 f(x,y)}{\partial y\partial x}\,. \tag{8}$$

2 **Increments of real valued functions:**

$$f(\mathbf{x}+\mathbf{h})-f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial x_1}h_1+\ldots+\frac{\partial f(\mathbf{x})}{\partial x_n}h_n+r(\mathbf{h}) \tag{9}$$

with $\frac{r(\mathbf{h})}{\|\mathbf{h}\|}\to 0$ *as* $\mathbf{h}\to\mathbf{0}$.

[b] *Hermann* Amandus Schwarz, 1843-1921.

Alternatively, we can write EQ. 9 as a matrix product or as a scalar product:

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right)\mathbf{h} + r(\mathbf{h}) \quad = \quad \boldsymbol{\nabla} f(\mathbf{x}).\mathbf{h} + r(\mathbf{h}) \;, \tag{10}$$

with the *gradient* of a real valued function defined as

$$\boldsymbol{\nabla} f(\mathbf{x}) := \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

Alternatively to the notation $\boldsymbol{\nabla} f$ we may also use $grad(f)$.

The amazing fact of [2] is that $r(\mathbf{h})$ does not only tend to 0 as $\mathbf{h} \to \mathbf{0}$, it still tends to 0 even if divided by $\|\mathbf{h}\|$. Therefore $f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x}).\mathbf{h}$ approximates f to first order, in other words, $f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x}).\mathbf{h} + \mathcal{O}(\|\mathbf{h}\|)$. Analogously to the one dimensional case we say that the function f can be *approximated linearly*[c].

For vector valued functions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$

$$\mathbf{f}(\mathbf{x}) \quad = \quad \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} ,$$

the gradient in EQ. 10 becomes a matrix, the *Jacobi matrix* $J_{\mathbf{f}}(\mathbf{x})$:

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) \quad = \quad J_{\mathbf{f}}(\mathbf{x})\mathbf{h} + \mathbf{r}(\mathbf{h}) \;, \tag{11}$$

with

$$J_{\mathbf{f}}(\mathbf{x}) \quad = \quad \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix} .$$

We then find $\mathbf{r}(\mathbf{h}) = (r_1(\mathbf{h}), \ldots, r_m(\mathbf{h}))^T$ such that each $\frac{r_i(\mathbf{h})}{\|\mathbf{h}\|} \to 0$ as $\mathbf{h} \to \mathbf{0}$, in other words $\frac{\mathbf{r}(\mathbf{h})}{\|\mathbf{h}\|} \to \mathbf{0}$ as $\mathbf{h} \to \mathbf{0}$. We therefore say $\mathbf{r}(\mathbf{h}) = \mathcal{o}(\|\mathbf{h}\|)$.

Finally, Taylor's theorem[d] generalises EQ. 11 to higher order:

3 Taylor's theorem

$$\begin{aligned} \mathbf{f}(\mathbf{x}+\mathbf{h}) \quad &= \quad \mathbf{f}(\mathbf{x}) + \sum_i \frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_i} h_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \mathbf{f}(\mathbf{x})}{\partial x_i \partial x_j} h_i h_j + \ldots \\ &\ldots + \frac{1}{k!} \sum_{i_1,\ldots,i_k} \frac{\partial^k \mathbf{f}(\mathbf{x})}{\partial x_{i_1} \ldots \partial x_{i_k}} h_{i_1} \ldots h_{i_k} + \mathbf{r}_k(\mathbf{h}) \;, \end{aligned} \tag{12}$$

[c] It should be noted that in the multi dimensional case the linear approximation of differentiable functions is valid in whichever way $\mathbf{h}$ approaches $\mathbf{0}$. Whether on a straight line, on a spiral or on any complicated curve, the linear expression always approximates the increment. Compared to the one dimensional case $f(x+h) = f(x) + f'(x)h + \mathcal{o}(h)$ it is now the gradient $\boldsymbol{\nabla} f(\mathbf{x})$ that plays the rôle of the derivative. The amazing property of smooth functions in higher dimensions is that the function can be approximated linearly in the multidimensional sense by a vector of partial derivatives. Considering that partial derivatives are just limits along straight lines it is remarkable that the set of these straight line limits describes a differentiable function in whichever way we approach the point $\mathbf{x}$, even if it is not on a straight line. On the other hand, as shall be shown in Analysis, for a function which is not differentiable the information contained in the partial derivatives is usually not enough to describe the behaviour of the function if we approach $\mathbf{x}$ along a path which is not parallel to a co-ordinate axis not even along a straight path.

[d] Brook Taylor, 1685-1731, St John's College (1703-1709, LL.B. 1709).

with $\frac{\mathbf{r}_k(\mathbf{h})}{||\mathbf{h}||^k} \to \mathbf{0}$ *as* $\mathbf{h} \to \mathbf{0}$.

The linear term of the Taylor expansion Eq. 12 is of course just the Jacobi[e] matrix acting on $\mathbf{h}$. For many of our applications the linear term will be sufficient:

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) \quad = \quad \mathbf{f}(\mathbf{x}) + J_{\mathbf{f}}(\mathbf{x})\mathbf{h} + \dots \quad . \tag{13}$$

The real valued case $m = 1$ of [3] is particularly important for us. This case can also be written in a simple way using the matrix of second order partial derivatives, the so-called *Hessian*[fg] $H_f(\mathbf{x})$.

$$f(\mathbf{x}+\mathbf{h}) \quad = \quad f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x}).\mathbf{h} + \frac{1}{2}\mathbf{h}^T H_f(\mathbf{x})\mathbf{h} + \dots \ , \tag{14}$$

with

$$H_{\mathbf{f}}(\mathbf{x}) \quad = \quad \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} \\ \vdots & & & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} .$$

Examples

Example 1.1 *Consider the function* $f(x,y) = x^2 + yx$. *Then*

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

Note that it is not really necessary to indicate which variable is being held constant, since $f(x,y)$ *is unambiguously a function of* x *and* y. *There would be no confusion in writing*

$$\frac{\partial f}{\partial x} = 2x + y \, , \qquad \frac{\partial f}{\partial y} = x \, .$$

Example 1.2 *Let* $f(x,y,z) = r$ *where* $r = |\mathbf{x}| = (x^2+y^2+z^2)^{\frac{1}{2}}$. *Then*

$$\left(\frac{\partial r}{\partial x}\right)_{y,z} = \frac{\partial}{\partial x}(x^2+y^2+z^2)^{\frac{1}{2}} = \frac{2x}{2(x^2+y^2+z^2)^{\frac{1}{2}}} = \frac{x}{r} \, , \tag{15}$$

and similarly

$$\left(\frac{\partial r}{\partial y}\right)_{x,z} = \frac{y}{r} \, , \quad \left(\frac{\partial r}{\partial z}\right)_{x,y} = \frac{z}{r} \, . \tag{16}$$

[e] *Carl* Gustav Jacob Jacobi, 1804-1851.

[f] Ludwig *Otto* Hesse, 1811-1874.

[g] Note that Schwarz's theorem says that $H_f(\mathbf{x})$ is a symmetric matrix and therefore we know from Part IA Algebra & Geometry that $H_f(\mathbf{x})$ can always be diagnalised with real eigenvalues and using an orthonormal basis transformation.

Example 1.3 *We will find the partial derivatives of the Cartesian coordinate* x_1*. We have*

$$\left(\frac{\partial x_1}{\partial x_1}\right)_{x_2,x_3} = 1\,, \quad \left(\frac{\partial x_1}{\partial x_2}\right)_{x_1,x_3} = 0\,, \ \textit{etc.}$$

Thus

$$\frac{\partial x_1}{\partial x_j} = \left\{ \begin{array}{lll} 1 & \textit{if} & j = 1 \\ 0 & \textit{if} & j \neq 1 \end{array} \right\} = \delta_{j1}\,,$$

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

A similar result holds for any of the coordinates. For the i*th coordinate, we have*

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}\,. \tag{17}$$

We can now use this result to obtain (15) *and* (16) *more neatly. Starting with that* $r^2 = x_i x_i$ *(don't forget we are using summation convention), we see that*

$$2r\frac{\partial r}{\partial x_k} = \frac{\partial (r^2)}{\partial x_k} = \frac{\partial (x_i x_i)}{\partial x_k} = 2x_i \frac{\partial x_i}{\partial x_k} = 2x_i \delta_{ik} = 2x_k\,.$$

Dividing through by $2r$ *gives*

$$\frac{\partial r}{\partial x_k} = \frac{x_k}{r}\,, \tag{18}$$

which, in terms of suffix notation, is the same as (15) *and* (16)*.*

Example 1.4 *Let*

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

and let $F(t) = f(\mathbf{x}(t))$*. We will calculate* $\frac{dF}{dt}$ *in two ways.*

1. *Using the chain rule. We have*

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

Hence from using the chain rule

$$\frac{dF}{dt} = \frac{\partial f}{\partial x}\,\frac{dx}{dt} + \frac{\partial f}{\partial y}\,\frac{dy}{dt} = 2x\,2t + e^t = 4t^3 + e^t\,.$$

2. *By direct substitution. We have*

$$F(t) = t^4 + e^t\,, \quad \textit{and hence} \quad \frac{dF}{dt} = 4t^3 + e^t\,.$$

The second (direct) method may seem much easier, but the chain rule allows us to find a formula for the rate of change of a general function $f(x,y)$ *along the curve* $\mathbf{x}(t) = \left(\ t^2, e^t \ \right)$*:*

$$\frac{df(\mathbf{x}(t))}{dt} = \frac{\partial f}{\partial x}\,\frac{dx}{dt} + \frac{\partial f}{\partial y}\,\frac{dy}{dt} = 2t f_x(t^2, e^t) + e^t f_y(t^2, e^t)\,.$$

Example 1.5 *We will find the gradient at* $x = 1$*,* $y = 1$ *of the curve of the family given by* $f(x,y) = c$*, where* $f(x,y) = \sin(\pi x y) + e^y$*. We have*

$$\frac{dy}{dx} = -\frac{f_x(1,1)}{f_y(1,1)} = -\frac{\pi y \cos(\pi x y)}{\pi x \cos(\pi x y) + e^y} = -\frac{-\pi}{-\pi + e}.$$

Example 1.6 *Consider the coordinate transformation (actually a rotation of the co-ordinate axes through a fixed angle θ)*

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

We will calculate how the derivatives of an arbitrary function $f(x,y)$ can be calculated in the new coordinates. Let $F(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}))$. Then

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

Sometimes, indeed often, $F(u,v) = f(x(u,v), y(u,v))$ is written (incorrectly) by applied mathematicians (but never by pure mathematicians) as $f(u,v)$. Thus f refers to the value of the function, whatever variables are used. The reason for doing this is that in many cases the function referred to has physical significance, such as temperature, and it is very inconvenient to refer to it by different letters according to which variables it is regarded as a function of. Since all the other variables have standard names too, this abuse of notation never becomes confusing. Then the above becomes

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

Using this notation, it is of course essential to state explicitly the variable(s) that are held constant in each partial differentiation.

Example 1.7 *Let $f(x,y)$ be an arbitrary function, and let $F(\rho,\phi) = f(\rho\cos\phi, \rho\sin\phi)$ (i.e. F has the same value as f but is calculated in plane polars rather than Cartesians). We will transform the equation*

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

into polar coordinates. First, note that

$$\frac{\partial f}{\partial x} = \frac{\partial\rho}{\partial x}\frac{\partial F}{\partial\rho} + \frac{\partial\phi}{\partial x}\frac{\partial F}{\partial\phi} = \cos\phi\frac{\partial F}{\partial\rho} - \frac{\sin\phi}{\rho}\frac{\partial F}{\partial\phi}\,,$$

or, omitting f and F,

$$\frac{\partial}{\partial x} = \cos\phi\,\frac{\partial}{\partial\rho} - \frac{\sin\phi}{\rho}\,\frac{\partial}{\partial\phi}\,.$$

Hence

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

$$= \cos^2\phi\,\frac{\partial^2 F}{\partial\rho^2} + \frac{\sin\phi\,\cos\phi}{\rho^2}\,\frac{\partial F}{\partial\phi} - \frac{\sin\phi\,\cos\phi}{\rho}\,\frac{\partial}{\partial\rho}\frac{\partial F}{\partial\phi} + \frac{\sin^2\phi}{\rho}\,\frac{\partial F}{\partial\rho}$$
$$-\frac{\sin\phi\,\cos\phi}{\rho}\,\frac{\partial}{\partial\phi}\frac{\partial F}{\partial\rho} + \frac{\sin\phi\,\cos\phi}{\rho^2}\,\frac{\partial F}{\partial\phi} + \frac{\sin^2\phi}{\rho^2}\,\frac{\partial^2 F}{\partial\phi^2}\,.$$

On assuming the equality of mixed derivatives we obtain

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

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 after applying the transformations* $x \to -y$, $y \to x$, $\phi \to \left(\phi + \frac{\pi}{2}\right)$, *we obtain*

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

and hence that

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

II Curves in $\mathbb{R}^3$

In Part IA Differential Equations you studied how differentiation can be generalised to higher dimensions simply by transferring the same limit technique from one dimension to higher dimensions which ultimately led to the partial derivative. What about integration? The integral over a one dimensional function is a limit process which chops up the area underneath the graph of the function in lots of little rectangles and sums over their areas. Could this concept or a similar mechanism easily be applied to functions of more than one variable and what would this give us? Before we look at the process of dividing areas up into little rectangles in more detail and then applying this to volumes and surfaces which are not flat we will first study a much simpler structure that is length of a path.

Let us take the circle $x^2 + y^2 = 1r^2$ which is obviously described by the path $\boldsymbol{\gamma}(t) = (r\sin t, r\cos t)$ as the time t runs from 0 to 2π. What is the length of this path (i.e. the circumference of the circle)? We could take the perimeter of the *inscribed* square which is $4\sqrt{2}r$ as an approximation of the circumference of the circle but it is obvious that the real circumference must be bigger. Obviously, the perimeter of the inscribed pentagon would be a better approximation and indeed the perimeter of the inscribed hexagon is an even better approximation. In fact, the perimeter of the inscribed regular *n-gon* becomes a better and better approximation as n gets bigger and should ultimately tend to the correct circumference in the limit $n \to \infty$. The perimeter of the n-gon is given by $\sum_{i=1}^{n} \left\| \begin{pmatrix} r\cos\frac{2\pi i}{n} \\ r\sin\frac{2\pi i}{n} \end{pmatrix} - \begin{pmatrix} r\cos\frac{2\pi(i-1)}{n} \\ r\sin\frac{2\pi(i-1)}{n} \end{pmatrix} \right\|$. Using the linear term of the Taylor expansion of cos and sin this is approximately equal to $\sum_{i=1}^{n} \left\| \begin{pmatrix} -r\sin\frac{2\pi(i-1)}{n} \\ r\cos\frac{2\pi(i-1)}{n} \end{pmatrix} \right\| \frac{2\pi}{n}$. For large n the higher order terms are negligible and therefore in the limit we expect this to give the circumference, and indeed

$$\sum_{i=1}^{n} \left\| \begin{pmatrix} -r\sin\frac{2\pi(i-1)}{n} \\ r\cos\frac{2\pi(i-1)}{n} \end{pmatrix} \right\| \frac{2\pi}{n} \quad = \quad \sum_{i=1}^{n} \frac{2\pi r}{n} = 2\pi r \ .$$

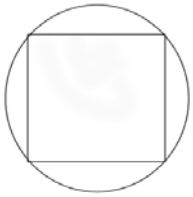


Fig. v Inscribed square

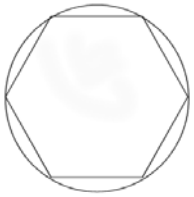


Fig. vi Inscribed hexagon

This limit process seems to rely on taking points on the path and replacing the arcs of the path connecting these points by straight lines. The sum of the Euclidean lengths of these lines would then approximate the length of the path. These ideas will form the basis of the following section which will lead to the definition of the length of a general curve.

1 Parametrised curves

We now want to formalise the ideas developed above for a general (but continuous and smooth) curve in $\mathbb{R}^n$. For our applications n would either be 2 or 3 but everything we are going to say would of course also work for curves in spaces of dimensions higher then 3.

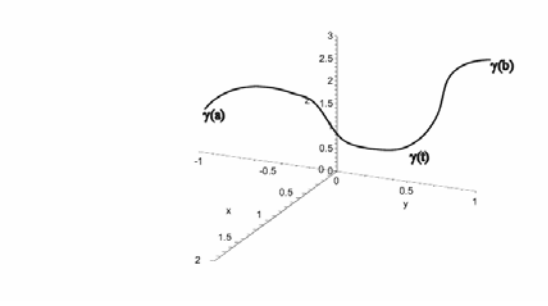


FIG. vii A path $\gamma(t)$ in $\mathbb{R}^3$.

A continuous vector valued function $\gamma(t) : [a, b] \to \mathbb{R}^n$ is called a *path* in $\mathbb{R}^n$. If we interpret t as the time parameter then $\gamma(t)$ can be seen to describe the path of a particle moving from the point $\gamma(a)$ to $\gamma(b)$ as times goes from a to b. How would we define the length of this path $\gamma(t)$?

Following our initial remarks, we take a partition $\mathcal{Z}$ of the interval $[a, b]$ with $t_0 = a$ and $t_b = b$ and $t_0 < t_1 < \ldots < t_n$. $\mathcal{Z}$ could be an equal distance partition but does not need to be, therefore it is useful to define the partition length $\|\mathcal{Z}\|$ to be the maximum $\|\mathcal{Z}\| := \max |t_i - t_{i-1}|$. We can now connect the points $\gamma(t_i)$ with straight lines as in FIG. viii and adding up the lengths of these straight lines $L(\gamma, \mathcal{Z}) := \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\|$ would obviously approximate the length of the path γ (and in fact would always be smaller than the real length of γ).

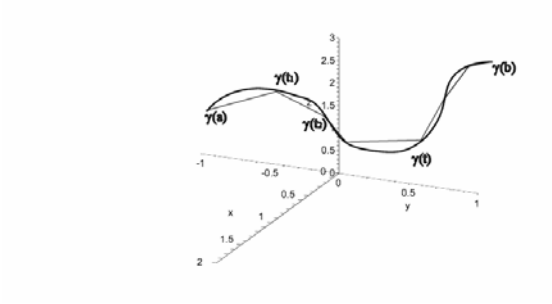


FIG. viii Partitioning a path $\gamma(t)$ in $\mathbb{R}^3$.

If we consider a sequence of partitions $\mathcal{Z}$ of $[a, b]$ with the partition lengths $\|\mathcal{Z}\|$ becoming smaller and smaller and ultimately tending to 0 then the sequence of the corresponding $L(\gamma, \mathcal{Z})$ which are approximations for the length of γ would obviously tend to what we should define as the real *length* of the path $\gamma(t)$. In technical terms mathematicians call a path for which this limit exists a *rectifiable path*. We denote such a limit by $\lim_{\|\mathcal{Z}\|\to 0}$.

4 Length of a path

The limit $\lim_{\|\mathcal{Z}\|\to 0} L(\gamma, \mathcal{Z})$ *of a (rectifiable) path* $\gamma(t) : [a, b] \to \mathbb{R}^n$ *is called the length* $L(\gamma)$ *of the path* γ *and is denoted by*

$$\int_\gamma ds := \lim_{\|\mathcal{Z}\|\to 0} L(\gamma, \mathcal{Z}) \, . \tag{19}$$

Let us now assume that each component function $\gamma_i(t)$ of the vector valued function $\gamma(t)$ can be differentiated, then Taylor's theorem tells us that the increment $\gamma(t_k) - \gamma(t_{k-1})$ can be approximated by $\gamma'(t_k)(t_k - t_{k-1})$ and therefore

$$L(\gamma, \mathcal{Z}) \quad = \quad \sum_k \|\gamma(t_k) - \gamma(t_{k-1})\| \tag{20}$$

can be approximated by

$$\sum_k \big\|\gamma'(t_k)\big\|(t_k - t_{k-1}) \, . \tag{21}$$

You will show in Analysis that the higher order terms do not matter in a limit where the differences $t_k - t_{k-1}$ tend to 0 and therefore EQ. 20 and EQ. 21 tend to the same value under such a limit. We now of course recognise that the limit of [21] is just the standard integral $\int_a^b \|\boldsymbol{\gamma}'(t)\|dt$. This leads us to the very significant and amazing result that this new and at first rather complicated looking limit we constructed above can in fact be converted into a standard integral provided we can differentiate the path $\boldsymbol{\gamma}(t)$!

5 Length of a differentiable path

If the path $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ *can be differentiated, then the length of* $\boldsymbol{\gamma}$ *is given by:*

$$\int_{\boldsymbol{\gamma}} ds \quad = \quad \int_a^b \big\|\boldsymbol{\gamma}'(t)\big\| dt \; . \tag{22}$$

We call $\boldsymbol{\gamma}'(t)$ *the velocity,* $\|\boldsymbol{\gamma}'(t)\|$ *the speed and* $\boldsymbol{\gamma}''(t)$ *the acceleration of the path* $\boldsymbol{\gamma}(t)$.

Example 2.1 *The path* $\boldsymbol{\gamma}(t) = (t\cos t, t\sin t)$ *describes a spiral turning counterclockwise as* t *runs from* 0 *to* $2\pi n$, $n \in \mathbb{N}$. *What is the length of the spiral?*

We find the velocity vector as $\boldsymbol{\gamma}' = (\cos t - t\sin t, \sin t + t\cos t)$ *and therefore its norm (the speed) is* $\|\boldsymbol{\gamma}'\| = \sqrt{1+t^2}$. *Therefore* $\int_{\boldsymbol{\gamma}} ds = \int_0^{2\pi n} \sqrt{1+t^2} dt = \frac{1}{2}(t\sqrt{1+t^2} + \sinh^{-1} t)\Big|_0^{2\pi n} = \pi n\sqrt{1+4\pi^2 n^2} + \frac{1}{2}\sinh^{-1}(2\pi n)$.

If we have two paths $\boldsymbol{\gamma}_1 : [a,b] \to \mathbb{R}^n$ and $\boldsymbol{\gamma}_2 : [b,c] \to \mathbb{R}^n$ with $a < b < c$ and $\boldsymbol{\gamma}_1(b) = \boldsymbol{\gamma}_2(b)$ then we define the *sum* of the two paths $\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2$ as the path $[a,b] \to \mathbb{R}^n$ with

$$\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2(t) := \begin{cases} \boldsymbol{\gamma}_1(t) & , t \in [a,b] \\ \boldsymbol{\gamma}_2(t) & , t \in [b,c] \end{cases} \; . \tag{23}$$

For the length of $\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2$ we obviously find that $L(\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2) = L(\boldsymbol{\gamma}_1) + L(\boldsymbol{\gamma}_2)$. In the same way we can restrict a path $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ to a shorter interval: $\boldsymbol{\gamma}_t : [a,t] \to \mathbb{R}^n$ with $\boldsymbol{\gamma}_t(u) = \boldsymbol{\gamma}(u)$ for $a \le u \le t \le b$. The length of the restricted path $\boldsymbol{\gamma}_t$ will then define a function of t which we want to call the *arclength function* $s(t)$ of $\boldsymbol{\gamma}$ with basepoint $\boldsymbol{\gamma}(a)$ (or simply *the length function*):

$$s(t) \quad = \quad L(\boldsymbol{\gamma}_t) = \int_{\boldsymbol{\gamma}_t} ds \; . \tag{24}$$

$s(t)$ obviously describes the length of the part of the path $\boldsymbol{\gamma}$ a particle would travel if time went from time a until time t. Using [5] we immediately find that $s(t) = \int_a^t \|\boldsymbol{\gamma}'(u)\|du$ and therefore $s'(t) = \|\boldsymbol{\gamma}'(t)\|$ which means that $s'(t)$ is simply the speed of $\boldsymbol{\gamma}(t)$.

6 The sum of paths and the length function

If $s(t)$ *describes the length of the path* $\boldsymbol{\gamma}$ *restricted to the interval* $[a,t]$ *then* $s'(t) = \|\boldsymbol{\gamma}'(t)\|$. *If* $\boldsymbol{\gamma}$ *is the sum of finitely many paths* $\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 \oplus \ldots \oplus \boldsymbol{\gamma}_k$ *then*

$$\int_{\boldsymbol{\gamma}} ds \quad = \quad \int_{\boldsymbol{\gamma}_1} ds + \ldots + \int_{\boldsymbol{\gamma}_k} ds \; . \tag{25}$$

Let Γ be some continuous injective curve in $\mathbb{R}^3$. In order to measure the length of the curve Γ we want to describe the curve with a path $\boldsymbol{\gamma}$ which is not going back and forth on parts of the curve, otherwise the length of the path would obviously not describe the length of the curve. Therefore we require a path to be injective in order to describe the length of

the curve described by the path. It is obvious that two injective paths describing the same curve have the same length. If $\boldsymbol{\gamma}_1(t)$ and $\boldsymbol{\gamma}_2(t)$ are two injective paths both describing Γ with the same starting points then $s_2^{-1}(s_1(t))$ gives the time at which the path $\boldsymbol{\gamma}_2$ reaches the point $\boldsymbol{\gamma}_1(t)$. Therefore we find $\boldsymbol{\gamma}_2(s_2^{-1}(s_1(t))) = \boldsymbol{\gamma}_1(t)$. Obviously $s_2^{-1}(s_1(t))$ is a strictly monotonous function[a]. Whenever two paths are related to each other by a strictly increasing function $\kappa(t)$: $\boldsymbol{\gamma}_2(\kappa(t)) = \boldsymbol{\gamma}_1(t)$, then we call $\boldsymbol{\gamma}_2$ a reparametrisation of $\boldsymbol{\gamma}_1$. Therefore two injective paths describing the same curve Γ are reparametrisations of each other.

An injective path $\boldsymbol{\gamma}$ is called a *Jordan path*[b] and the corresponding curve is called a *simple curve* or a *Jordan arc*. A Jordan path $\boldsymbol{\gamma}$ describing a Jordan arc Γ is called a *Jordan description* of Γ. In summary, the length function of a Jordan path is always strictly increasing and therefore invertible. If two Jordan paths describe the same Jordan curve then they are reparametrisations of one another with some continuous and strictly monotonous function κ such that $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_1 \circ \kappa$. Conversely, any continuous and strictly monotonous function κ generates another Jordan description $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_1 \circ \kappa$ from a Jordan description $\boldsymbol{\gamma}_1$ of the Jordan arc Γ. It is therefore obvious that we can easily generate all Jordan descriptions of a Jordan arc Γ out of just one of its Jordan descriptions simply by applying suitable reparametrisations. It is obvious from the limit definition that a reparametrisation should lead to the same length of the Jordan arc. Using [5] we easily see for $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_1 \circ \kappa$ that $\int_{\boldsymbol{\gamma}_2} ds = \int_a^b \|\boldsymbol{\gamma}_2'(t)\| dt = \int_a^b \|\boldsymbol{\gamma}_1'(\kappa(t))\| \kappa'(t) dt = \int_{\kappa(a)}^{\kappa(b)} \|\boldsymbol{\gamma}_1'(t)\| dt = \int_{\boldsymbol{\gamma}_1} ds$.

7 Arc length of a Jordan arc

The arc length of a Jordan arc Γ is the length of a Jordan path, an injective and rectifiable path $\boldsymbol{\gamma}$ describing Γ. Two such Jordan paths describing Γ have the same length and are reparametrisations of one another. We denote the length of a Jordan arc by $\int_\Gamma ds$

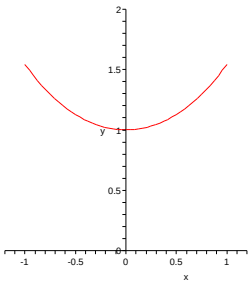


FIG. ix Hyperbolic cosine $\cosh x$.

Let us for example calculate the length of a hyperbolic cosine curve defined on the interval $[-1, 1]$. We choose the parametrisation

$$\boldsymbol{\gamma}(t) = \begin{pmatrix} t \\ \cosh t \end{pmatrix}, t \in [-1, 1], \quad (26)$$

which is obviously a Jordan description of Γ. We find:

$$\int_\Gamma ds = \int_{-1}^{1} \left\| \begin{pmatrix} 1 \\ \sinh t \end{pmatrix} \right\| dt = \int_{-1}^{1} \sqrt{1 + \sinh^2 t}\, dt = 2 \sinh 1 \,. \quad (27)$$

Generally, if we want to calculate the length of a curve Γ_f described by the graph of a function $f(x) : [a, b] \to \mathbb{R}$ we can find a Jordan description $\boldsymbol{\gamma}(x) = (x, f(x))^T$ which leads to the following length:

[a] In case the paths go in opposite directions along Γ then the function is strictly decreasing rather than strictly increasing.

[b] Marie Ennemond *Camille* Jordan, 1838-1922.

8 **Length of a curve described by a function** $f(x)$

The length of a continuous curve Γ_f *described by a function* $f(x) : [a, b] \to \mathbb{R}$ *is*

$$\int_{\Gamma_f} ds \quad = \quad \int_a^b \sqrt{1 + \left(\frac{df}{dx}\right)^2}\, dx \; . \tag{28}$$

Let us consider a Jordan path $\boldsymbol{\gamma}$ corresponding to a Jordan arc Γ. Its length function $s(t)$ is strictly increasing in t and let $t(s)$ denote its inverse function. $t(s)$ gives the time t at which the path will have reached length s from the starting point $\boldsymbol{\gamma}(a)$. We can now choose a particular parametrisation of the time using $t(s)$: we define the path $\boldsymbol{\gamma}_c(s) := \boldsymbol{\gamma}(t(s))$ on the interval $[0, L(\boldsymbol{\gamma})]$ which is obviously just a reparametrisation of $\boldsymbol{\gamma}(t)$. Using the chain rule and [5] we find that

$$\frac{d\boldsymbol{\gamma}_c(s)}{ds} \quad = \quad \boldsymbol{\gamma}'(t)\frac{dt(s)}{ds} = \frac{\boldsymbol{\gamma}'(t)}{\frac{ds}{dt}} = \frac{\boldsymbol{\gamma}'(t)}{\|\boldsymbol{\gamma}'(t)\|} \; , \tag{29}$$

and therefore $\left\| \frac{d\boldsymbol{\gamma}_c(s)}{ds} \right\| = 1$. The Jordan path $\boldsymbol{\gamma}_c(s)$ describes the same Jordan arc Γ as $\boldsymbol{\gamma}(t)$ but $\boldsymbol{\gamma}_c(s)$ is the parametrisation of Γ with constant *speed* 1 everywhere along the path. This parametrisation is obviously unique (provided we do not change the direction along the path). We want to call it the *canonical parametrisation* of the Jordan arc Γ.

Let us for example consider the *circular helix*[c] described by the Jordan path $\boldsymbol{\gamma}(t) = (a\cos t,\ a\sin t,\ ct)^T$ where $a, c \in \mathbb{R}^+$ and t starts at time 0.

The length function is obviously $s(t) = \sqrt{a^2 + c^2}\; t$ with inverse function $t(s) = \frac{s}{\sqrt{a^2+c^2}}$. The canonical parametrisation of the helix is therefore

$$\boldsymbol{\gamma}_c(s) \quad = \quad \begin{pmatrix} a\cos\frac{s}{\sqrt{a^2+c^2}} \\ a\sin\frac{s}{\sqrt{a^2+c^2}} \\ \frac{cs}{\sqrt{a^2+c^2}} \end{pmatrix} .$$

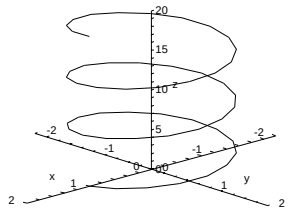


FIG. x Circular helix for $a = c = 1$.

Finally, on a technical note, [5] requires that the path $\boldsymbol{\gamma}$ is of course differentiable since we have to construct the derivative $\boldsymbol{\gamma}'(t)$. In case the path is differentiable almost everywhere except at a finite number of points n, then we can split the path up into a sum of $n + 1$ paths $\boldsymbol{\gamma}_i$, each of them being differentiable. For each $\boldsymbol{\gamma}_i$ we can use [5] in order to compute its length. Adding these lengths up leads to the length of $\boldsymbol{\gamma}$. In this case mathematicians say that $\boldsymbol{\gamma}$ is *piecewise* differentiable. Most of what we said and what we are going to say about differentiable paths applies in the same way to piecewise differentiable paths following exactly this mechanism.

If a Jordan arc is closed we call it a *Jordan curve*. In this case a corresponding Jordan description can only be injective on $[a, b)$ since we obviously need $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$ in order to close the curve. But a Jordan curve does not allow any other points where injectivity is violated. For example, a circle is a Jordan curve whilst an '8' is not a Jordan curve. But we can of course generate an '8' using two Jordan curves.

[c] A circular helix is found by winding a line around a cylinder. In contrast, a *conical helix* is found by winding a line around a cone.

Sometimes it may be useful to weight certain parts of a path more than other parts. For example, if the path represents an object which is heavier at one end than at the other end. Therefore we find a *line density* function $f(\mathbf{x})$ describing the mass distribution along the curve. In order to compute the total mass we obtain limits of the type $\lim_{\|\mathcal{Z}\|\to 0} \sum f(\boldsymbol{\gamma}(t_k))\|\boldsymbol{\gamma}(t_k) - \boldsymbol{\gamma}(t_{k-1})\|$. We denote such a limit by $\int_{\boldsymbol{\gamma}} f(\mathbf{x})ds$. It is clear that under the assumptions of [5] and for continuous functions $f(\boldsymbol{\gamma}(t))$ we again find a simple integral just as in [5]:

9 **Integrals over $f(\mathbf{x})$ along paths**
If the path $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ can be differentiated then:

$$\int_{\boldsymbol{\gamma}} f(\mathbf{x})ds \quad = \quad \int_a^b f(\boldsymbol{\gamma}(t))\big\|\boldsymbol{\gamma}'(t)\big\|dt\,. \tag{30}$$

We easily find the following integration rules.

10 **Integration rules for integrals over $f(\mathbf{x})$ along paths**

$$\int_{\boldsymbol{\gamma}} (\alpha f(\mathbf{x}) + \beta g(\mathbf{x}))ds \quad = \quad \alpha \int_{\boldsymbol{\gamma}} f(\mathbf{x})ds + \beta \int_{\boldsymbol{\gamma}} f(\mathbf{x})ds\,, \tag{31}$$

$$\int_{\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2} f(\mathbf{x})ds \quad = \quad \int_{\boldsymbol{\gamma}_1} f(\mathbf{x})ds + \int_{\boldsymbol{\gamma}_2} f(\mathbf{x})ds\,. \tag{32}$$

We can now also define integrals over vector valued functions $\mathbf{f}(x)$ along a path $\boldsymbol{\gamma}$ simply by defining the integral to be the vector of integrals over the component functions:

$$\int_{\boldsymbol{\gamma}} \mathbf{f}(\mathbf{x})ds \quad := \quad \begin{pmatrix} \int_{\boldsymbol{\gamma}} f_1(\boldsymbol{x})ds \\ \vdots \\ \int_{\boldsymbol{\gamma}} f_m(\mathbf{x})ds \end{pmatrix}. \tag{33}$$

2 Tangents and normals to curves in $\mathbb{R}^3$

In geometry you are mainly interested in two types of theorems, *classification theorems* and theorems relating *local* properties to *global* properties. For example, we know from simple geometry that two Euclidean triangles are congruent (can be mapped onto each other using translations, reflections and rotations) if and only if the lengths of their corresponding sides are equal. This is a classification theorem. It allows us to determine whether two mathematical objects are equivalent under some appropriate equivalence relation (here congruence) just by comparing a small (or at least finite) number of computable invariants. We also know from simple geometry that the sum of the interior angles of a Euclidean triangle is π. This theorem relates a local geometric property (angles) to a global property (that of being a triangle). Another example for a classification theorem is that two circles in the Euclidean plane are congruent if and only if they have the same radius. Whilst the theorem that the circumference of a Euclidean circle of radius r is $2\pi r$ is again a theorem that relates the local property of the angle 2π with the global property of being a circle. We could give numerous examples for these types of theorems in simple geometry, but if we want to continue the study of plane geometry beyond figures constructed from lines and circles we realise that an arbitrary curve cannot be

completely described by a few numbers such as lengths and radii. Instead we will find that the basic invariant is *curvature*. Curvature is a function of the position on the curve.

In the previous section we found that a Jordan arc in $\mathbb{R}^2$ can be uniquely described by its canonical Jordan path $\boldsymbol{\gamma}_c(t)$ such that $\|\boldsymbol{\gamma}_c'(s)\| = 1$ everywhere along the path. If we reparametrise the path then obviously the *speed* $\|\boldsymbol{\gamma}'\|$ changes and therefore first order derivatives are not a suitable measure to classify the curve uniquely up to congruence. But it can be shown that the second order derivatives can be used instead.

11 Curvature of a plane curve

$\boldsymbol{\gamma}_c(s)$ is the canonical Jordan path of the Jordan arc Γ. The curvature $\kappa(s)$ of Γ at the position s is defined as the magnitude of the acceleration vector

$$\kappa(s) := \left\|\boldsymbol{\gamma}''(s)\right\| . \tag{34}$$

Along a straight path the velocity vector $\boldsymbol{\gamma}'$ never changes direction and therefore a straight path has curvature 0 everywhere. On the other hand a circle

$$\boldsymbol{\gamma}_c(s) \;=\; \begin{pmatrix} r\cos\frac{s}{r} \\ r\sin\frac{s}{r} \end{pmatrix} \quad , s \in [0, 2\pi r] \, , \tag{35}$$

has the constant curvature $\kappa = \frac{1}{r}$ at each point s on the circle. For a general curve Γ at some point s on Γ we find that among all circles which are tangent to Γ at this point there is only one circle whose acceleration also matches the canonical acceleration along the curve. This circle is called the *osculating circle* and the curvature of the curve at s is simply $\frac{1}{\text{radius}}$ of the osculating circle. If the radius is huge or even infinity for a straight line then obviously the curve is less curved at this point whilst a small radius means a big curvature. In summary, the bigger the curvature, the greater the acceleration and therefore the smaller the osculating circle.

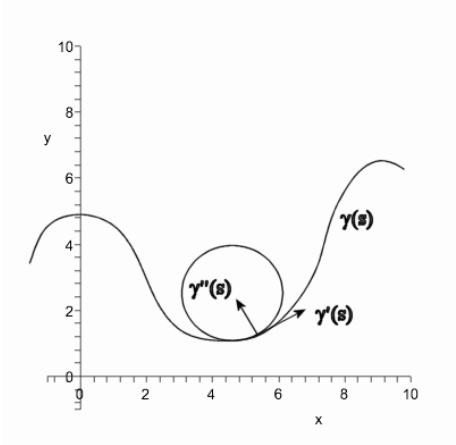


FIG. xi Osculating circle to a Jordan arc.

So far we have only defined curvature to be non-negative. Sometimes it is useful to introduce a (normal) direction and we would then define positive or negative curvature with respect to this direction. This is called the *signed curvature*. It can be shown that the signed curvature describes a curve uniquely up to congruence. Therefore we can obtain a classification theorem saying that two Jordan arcs are congruent if and only if the curvatures match at each point. This is the classification theorem we were looking for but one can also prove a theorem that

relates the local property of curvature to a global property of being a closed Jordan arc, meaning a Jordan curve: a Jordan arc is closed if and only if $\int_\gamma \kappa ds = \int_a^b \kappa(s)ds = 2\pi$ where $\kappa(s)$ is the signed curvature of $\boldsymbol{\gamma}$.

We now want to apply similar ideas to curves in $\mathbb{R}^3$. It is immediately clear that the radius of a plane osculating circle cannot be sufficient to describe curves in three dimensions. We will in addition need a way of describing the curvature of the curve coming out of the tangent plane. One finds that the following local properties define a curve in $\mathbb{R}^3$ up to congruence.

12 Curves in $\mathbb{R}^3$

Let Γ be a Jordan arc in $\mathbb{R}^3$ and let $\boldsymbol{\gamma}_c$ be the corresponding canonical Jordan path $\boldsymbol{\gamma}_c : [a,b] \to \mathbb{R}^3$.

$$\mathbf{t} := \boldsymbol{\gamma}_c'(s) \tag{36}$$

is called the tangent and it is clear that $\|\mathbf{t}(s)\| = 1 \quad \forall s$.

$$\mathbf{p} := \frac{\boldsymbol{\gamma}_c''(s)}{\|\boldsymbol{\gamma}_c''(s)\|} \tag{37}$$

is called the principal normal and $\kappa(s) := \|\boldsymbol{\gamma}_c''(s)\|$ is the curvature. The binormal is given by

$$\mathbf{b} := \mathbf{t} \times \mathbf{p}\,, \tag{38}$$

and the torsion τ is the (negative) projection of the change in the binormal $\frac{\partial \mathbf{b}}{\partial s}$ onto the principal normal:

$$\tau := -\frac{\partial \mathbf{b}}{\partial s}.\,\mathbf{p}\,. \tag{39}$$

It is worth mentioning that $\boldsymbol{\gamma}_c'(s).\boldsymbol{\gamma}_c''(s) = 0$ since $\boldsymbol{\gamma}_c'(s).\boldsymbol{\gamma}_c'(s) = 1$ and therefore $\mathbf{t}$, $\mathbf{p}$ and $\mathbf{b}$ form an orthonormal system. The tangent and the principal normal form a plane tangent to the curve which describes the movement of the curve to second order. In this plane there exists an osculating circle and the curvature describes the size of the osculating circle just like in the two dimensional case. The binormal is normal to this plane. If the curve describes a true three dimensional movement then it will obviously leave this plane. The torsion describes how strongly it leaves the plane. The minus sign in the definition is such that a movement outside this plane towards the direction of the binormal has a positive torsion, as in the case of the circular helix in Fig. x.

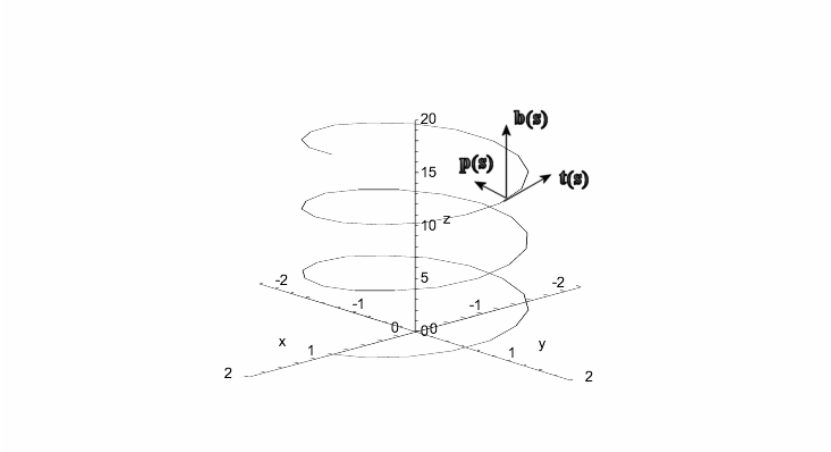


Fig. xii Tangent, principal normal and binormal for a circular helix.

Examples

Example 2.2 *(i) The curve given parametrically by*

$$\mathbf{x} = \boldsymbol{\psi}(t) = (at, bt^2, 0)$$

is the parabola $a^2y = bx^2$ in the x-y plane. The length of the curve from $t = 0$ to $t = 1$ is

$$\int ds \equiv \int \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \equiv \int \left\|\frac{d\mathbf{x}}{dt}\right\| dt$$

$$= \int_0^1 \sqrt{a^2 + 4b^2t^2}\, dt = \textit{whatever you get when you do this integral.}$$

(ii) The curve

$$\mathbf{x} = \boldsymbol{\psi}(t) = (5\cos t, 4\sin t, 3\sin t)$$

lies on the cylinder (with elliptic cross-section) $x^2/25 + y^2/16 = 1$. Setting $t = \phi$, we see that the 'height' z of the curve as it wraps around the cylinder is $3\sin\phi$, so it looks similar to a sine curve wrapped around the cylinder. Clearly, the curve is closed: it joins back onto itself when t increases by 2π.

Note that it also lies on the cylinder $x^2/25 + z^2/9 = 1$, so it is the intersection of the two cylinders.

The (unit) tangent $\mathbf{t}$ is the unit vector in the direction $\frac{d\boldsymbol{\psi}}{dt}$:

$$\mathbf{t} = (-5\sin t, 4\cos t, 3\cos t)/5$$

The arc length, s, is given in terms of the parameter t by

$$ds = \left\|\frac{d\mathbf{t}}{dt}\right\| dt = 5dt,$$

so we could reparametrise the curve using s instead of t as follows:

$$\mathbf{x} = \big(5\cos(s/5), 4\sin(s/5), 3\sin(s/5)\big).$$

Then

$$\frac{d\mathbf{t}}{ds} = -\big(5\cos(s/5), 4\sin(s/5), 3\sin(s/5)\big)/(25) \equiv \kappa\mathbf{p}\ ,$$

Using the fact that $\frac{d\mathbf{t}}{ds}$ and $\mathbf{p}$ are unit vectors, we see that the curvature $\kappa = 1/5$ and that the principal normal $\mathbf{p} = -\big(5\cos(s/5), 4\sin(s/5), 3\sin(s/5)\big)/5$. Finally, the binormal $\mathbf{b} = \mathbf{t} \times \mathbf{p} = (0, -3, 4)/5$. Since $\frac{d\mathbf{b}}{ds} = 0$, the torsion $\tau = 0$. This is a little unexpected. It means that the curve lies in a plane with normal $(0, -3, 4)/5$ — which of course is now obvious: $\mathbf{x}.(0, -3, 4) = 0$.

We can see this in another way. The curve lies on the sphere $r = 5$. If we use polar coordinates based on the x-axis instead of the z-axis, we see that $t = \theta$ and $\phi = \tan^{-1}(3/4)$; the curve is therefore a circle (a line of longitude).

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

In the previous chapter we analysed a first example of how to generalise the idea of integration. The amazing outcome was that this new integration process we defined to compute the length of a curve could be converted into a standard integral known to us to compute the area underneath graphs. We will now look at integration in more detail but before we do this we want to summarise the concept behind the standard integral process which defines the area underneath graphs since we we will later convert all the integrals we are going to define in higher dimensions into just standard integrals.

Let us consider the graph of a positive function $f(x)$ defined on an interval $[a, b]$ as in Fig. xiii. We partition the area underneath the graph of $f(x)$ and obtain a sum of the form $\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$ as an approximation for the area underneath the graph. ξ_i could either be x_{i-1} or x_i or indeed any arbitrary point in between x_{i-1} and x_i. In any case, taking the limit for the differences $x_i - x_{i-1}$ tending to 0 we obtain the area underneath the graph. We denote this limit by $\lim_{\|\mathcal{Z}\|\to 0} \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \int_a^b f(x)dx$. Mathematicians call this integral the *Riemann Integral*[a,b].

A sum of the type $\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$ is called a *Riemann sum* and a limit of the type $\lim_{\|\mathcal{Z}\|\to 0} \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$ is called a *Riemann limit*. For a Riemann sum ξ_i has to be somewhere in the interval $[x_{i-1}, x_i]$ but the exact position does not matter in the limit.

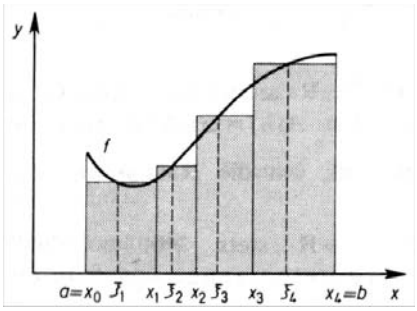


Fig. xiii The Riemann integral.

In this chapter our task will be to define integration mechanisms in higher dimensions but we will find that we will always be able to break down integration sums in higher dimensions into Riemann sums and therefore into Riemann integrals. Therefore it will be important to be able to recognise Riemann sums:

13 Riemann sum

A Riemann sum is a sum of the type

$$\sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \tag{40}$$

where all $\xi_i \in [x_{i-1}, x_i]$.

We will now investigate *Riemann-type* sums defined on curves, surfaces and volumes in $\mathbb{R}^2$ or $\mathbb{R}^3$ over real valued and also vector valued functions. This will allow us to define integrals along curves, surfaces and over volumes and relate these new integrals to ordinary one dimensional Riemann integrals.

[a] Georg Friedrich *Bernhard* Riemann, 1826-1866.

[b] There is another important concept of integration, the *Lebesgue Integral*, named after *Henri* Léon Lebesgue (1875-1941). This will be studied in future Analysis courses. For our applications the Riemann integral is sufficient.

3 Line integrals

We will now look at a different type of integral over vector valued functions which we call *line integrals*. The energy E a point mass gains by moving along a homogeneous force F is $F\Delta x$ where Δx denotes the distance the point mass is moving in direction of the force. This assumes that the point mass is moving parallel to the force, but what if we force it to move at an angle to the force? In this case it is only the component of the force parallel to the movement of the point mass that results in a gain of energy: $F\Delta x \cos\alpha$ where α is the angle between the direction of the force and the movement of the point mass. We can write this more conveniently as the scalar product between the vector $\mathbf{F}$ describing the force and the vector $\mathbf{\Delta x}$, the vector describing the movement: $E = \mathbf{F}.\mathbf{\Delta x} = F\Delta x \cos\alpha$. But what do we do if the force is not homogeneous? In this case the point mass might experience a different force (in terms of strength as well as direction) at every point along the path describing the movement. It now seems reasonable to solve this problem again using a Riemann-type limit where we look at sequences of partitions of the path describing the movement. We then again assume that between two partition points the movement is along a straight line and in addition along this straight line the force is homogeneous, taking the value of the force at some point between these partition points for the whole of this little straight line. It seems natural that in the limit this will give the energy gained or required by moving the point mass through this force. Let us call such a (possibly inhomogeneous) force a *force field* or more generally we define *vector fields* to be vector valued functions assigning a vector from $\mathbb{R}^n$ to each point in $\mathbb{R}^n$ (or at least to a subset of $\mathbb{R}^n$).

14 **Vector fields**[c]
A vector valued function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is called a vector field.

FIG. xiv shows the inhomogeneous two dimensional vector field

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \sin x \\ \cos y \end{pmatrix} \qquad (41)$$

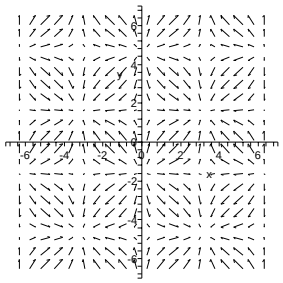
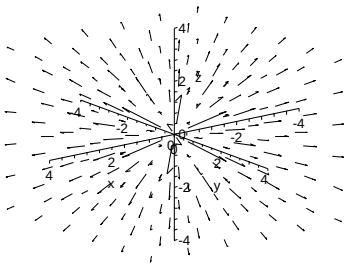


FIG. xiv Two dimensional vector field.

[c]Note that for a vector field the domain and the range have to have the same dimension.



The three dimensional vector field

$$\mathbf{F}(\mathbf{x}) \quad = \quad \begin{pmatrix} \frac{x}{r^3} \\ \frac{y}{r^3} \\ \frac{z}{r^3} \end{pmatrix} \qquad (42)$$

shown in FIG. xv has a spherical symmetry.

FIG. xv Three dimensional vector field.

Motivated by the idea of summing up scalar products in order to calculate the energy gained by the movement of a point mass through a force field we partition the path $\boldsymbol{\gamma}(t)$ and obtain the sum

$$\sum \mathbf{F}(\boldsymbol{\gamma}(t_i)).(\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1})) \ . \qquad (43)$$

As before, we now study the limit of the sum in EQ. 43 for the sequence of partitions $\mathcal{Z}$ of $[a, b]$ with $\|\mathcal{Z}\| \to 0$. We call this limit the *line integral* of $\mathbf{F}$ along the path $\boldsymbol{\gamma}$.

15 Line integrals

Given a vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$*, and a path* $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$*. The line integral of* $\mathbf{F}$ *along the path* $\boldsymbol{\gamma}$ *is defined*

$$\int_{\boldsymbol{\gamma}} \mathbf{F}(\mathbf{x}).d\mathbf{x} \quad := \quad \lim_{\|\mathcal{Z}\|\to 0} \sum \mathbf{F}(\boldsymbol{\gamma}(t_i)).(\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1})) \ . \qquad (44)$$

An alternative notation for the line integral is

$$\int_{\boldsymbol{\gamma}} \mathbf{F}(\mathbf{x}).d\mathbf{x} \quad = \quad \int_{\boldsymbol{\gamma}} F_1(\mathbf{x})dx_1 + F_2(\mathbf{x})dx_2 + \ldots + F_n(\mathbf{x})dx_n \ . \qquad (45)$$

Using Taylor's theorem to first order on the increments $\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1})$ the Riemann-type sum $\sum \mathbf{F}(\boldsymbol{\gamma}(t_i)).(\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1}))$ can be written to first order as

$$\sum_i F(\boldsymbol{\gamma}(t_i))\boldsymbol{\gamma}'(t_i)(t_i - t_{i-1}) \ , \qquad (46)$$

which we now recognise as a Riemann sum.

16 Line integral over continuous vector fields along differentiable paths

$\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ *is a differentiable path then the line integral can be evaluated using a simple Riemann integral:*

$$\int_{\boldsymbol{\gamma}} \mathbf{F}(\mathbf{x}).d\mathbf{x} \quad = \quad \int_a^b \mathbf{F}(\boldsymbol{\gamma}(t)).\boldsymbol{\gamma}'(t)dt \ . \qquad (47)$$

Using the alternative notation we can write EQ. 47 *in the following form*[d]*:*

$$\int_{\boldsymbol{\gamma}} F_1(\mathbf{x})dx_1 + \ldots + F_n(\mathbf{x})dx_n \quad = \quad \int_a^b F_1(\boldsymbol{\gamma}(t))\frac{d\gamma_1}{dt}dt + \ldots + \int_a^b F_n(\boldsymbol{\gamma}(t))\frac{d\gamma_n}{dt}dt \ .$$

[d]Sometimes $\int_a^b F_1(\boldsymbol{\gamma}(t))\frac{d\gamma_1}{dt}dt + \ldots$ is shortened to $\int_a^b F_1 dx_1 + \ldots$. Note the danger of this notation: $\int F_i dx_i$ still has to be taken along the path and does not mean that $x_1, \ldots, x_n$ can be considered as independent variables for the integration.

Example 3.1 *Let us consider the vector field* $\mathbf{F}(\mathbf{x}) := \left(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3}\right)^T$ *given in* FIG. XV *and the continuous path* $\boldsymbol{\gamma}(t) = (\cos t, \sin t, t)^T$ *along the circular helix shown in* FIG. X *for* $t \in [0, 2\pi]$ *we find*

$$\int_\gamma \mathbf{F}.d\mathbf{x} = \int_0^{2\pi} \begin{pmatrix} \frac{\cos t}{(1+t^2)^{\frac{3}{2}}} \\ \frac{\sin t}{(1+t^2)^{\frac{3}{2}}} \\ \frac{t}{(1+t^2)^{\frac{3}{2}}} \end{pmatrix} . \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix} dt = \int_0^{2\pi} \frac{t}{(1+t^2)^{\frac{3}{2}}} dt = 1 - \frac{1}{1+4\pi^2} \, . \tag{48}$$

It is obvious that different Jordan descriptions of the same Jordan arc lead to the same line integral since the reparametrisations simply correspond to a substitution of a strictly monotonous function. When calculating a line integral it is therefore essential to try to choose a suitable parametrisation such that the resulting Riemann integral becomes simple. The following integration rules are obvious and are given without proof.

17 Integration rules for line intergals

$$\int_\gamma (\alpha\mathbf{F} + \beta\mathbf{G}).d\mathbf{x} = \alpha \int_\gamma \mathbf{F}.d\mathbf{x} + \beta \int_\gamma \mathbf{G}.d\mathbf{x} \, , \tag{49}$$

$$\int_{\gamma_1 \oplus \gamma_2} \mathbf{F}.d\mathbf{x} = \int_{\gamma_1} \mathbf{F}.d\mathbf{x} + \int_{\gamma_2} \mathbf{F}.d\mathbf{x} \, , \tag{50}$$

$$\int_\gamma \mathbf{F}.d\mathbf{x} = -\int_{\gamma^-} \mathbf{F}.d\mathbf{x} \quad \text{with} \quad \boldsymbol{\gamma}^-(t) = \boldsymbol{\gamma}(b-t) \, , \tag{51}$$

$$\left| \int_\gamma \mathbf{F}.d\mathbf{x} \right| \leq \max_{t\in[a,b]} \|\mathbf{F}(\boldsymbol{\gamma}(t))\| \int_\gamma ds \, . \tag{52}$$

Example 3.2 *Let* $\mathbf{F}$ *be the force acting on a particle moving along the curve* $\mathcal{C}$*. The work done by the force in moving the particle an infinitesimal distance* $d\mathbf{x}$ *is* $\mathbf{F}.\,d\mathbf{x}$*; this is the generalisation to 3D of the 1D formula*

$$\textit{work done} = \textit{force} \times \textit{distance}.$$

Hence the total work done by the force in transversing the curve $\mathcal{C}$ *is*

$$W = \int_{\mathcal{C}} \mathbf{F}.\, d\mathbf{x}\,.$$

We can parametrise the curve by the time t *at which the particle is at a point* $\mathbf{x}$*:*

$$W = \int_{t_a}^{t_b} \mathbf{F}.\, \frac{d\mathbf{x}}{dt}\, dt = \int_{t_a}^{t_b} \mathbf{F}.\, \mathbf{v} dt$$

which is the time integral of the rate of doing work.

Using Newton's second law of motion $\mathbf{F} = m\ddot{\mathbf{x}}$ *(it is easiest to think of the case when there*

is just one force, e.g. gravity or rocket propulsion in deep space), gives

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

Example 3.3 *Let*

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

We will evaluate the line integral $\int \mathbf{F}.d\mathbf{x}$ *for the curve* $\mathcal{C}$ *parametrised by* $\mathbf{x} = (t, t^a, z(t))$ *where* $a > 0$ *and* $0 \le t \le 1$. *(Obviously, it does not matter what* $z(t)$ *is). We have*

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

Note that the answer is independent of a, *i.e. the same result is obtained for different paths. In general, if different paths join the same end points, then the corresponding line integrals are not equal; as we shall explain in the following chapter they are in this case because* $ydx + xdy$ *is an exact differential(i.e. because* $\mathbf{F}(\mathbf{x})$ *can be written as the gradient of a function* $\phi(\mathbf{x})$*).*

We can verify this result for a straight line path from $(0,0,0)$ *to* $(1,1,0)$ *along paths parallel to the axes. Let*

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

On $\mathcal{C}_1$

$$\mathbf{F}.\, d\mathbf{x} = \mathbf{F}.\, \frac{d\mathbf{x}}{dt_1} dt_1 = (0, t_1, 0).(1, 0, 0)\, dt_1 = 0\,.$$

On $\mathcal{C}_2$

$$\mathbf{F}.\, d\mathbf{x} = \mathbf{F}.\, \frac{d\mathbf{x}}{dt_2} dt_2 = (t_2, 1, 0).(0, 1, 0).\, dt_2 = dt_2\,.$$

Hence

$$\int_{\mathcal{C}_1 + \mathcal{C}_2} \mathbf{F}.\, d\mathbf{x} = \int_0^1 dt_2 = 1\,.$$

4 Plane surface integrals in $\mathbb{R}^2$ and volume integrals in $\mathbb{R}^3$

We will now proceed to defining integration over volumes and start with looking at functions defined on plane rectangles $\mathcal{R} := [a, b] \times [c, d]$ in $\mathbb{R}^2$. Let us therefore consider a real valued function $f(x, y) : \mathcal{R} \to \mathbb{R}$ on this rectangle $\mathcal{R}$.

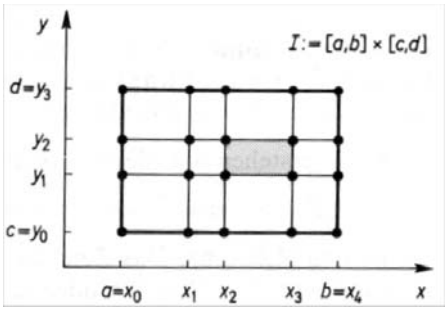


FIG. xvi Partitioning the rectangle $\mathcal{R}$.

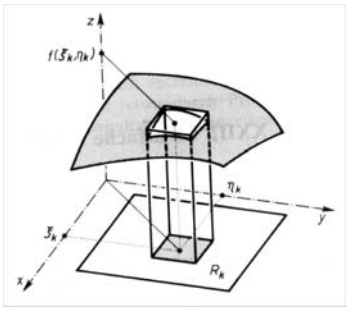


FIG. xvii Graph of $f(x,y)$ over the rectangle $\mathcal{R}$.

Following similar ideas as before we will partition the rectangle $\mathcal{R}$ into small rectangles as shown in FIG. xvi. We therefore take a partition $\mathcal{Z}_x$ of the x-axis and a partition $\mathcal{Z}_y$ of the y-axis with partition lengths $\|\mathcal{Z}_x\|$ and $\|\mathcal{Z}_y\|$ respectively. In case $f(x,y) \geq 0$ on the rectangle $\mathcal{R}$ then it is obvious from FIG. xvii that for small $\|\mathcal{Z}_x\|$ and $\|\mathcal{Z}_y\|$ the double sum

$$\sum_{i,j} f(\xi_i, \eta_j)(x_i - x_{i-1})(y_j - y_{j-1}) \tag{53}$$

is an approximation for the volume underneath the graph of $f(x,y)$. ξ_i and η_j are as before chosen arbitrarily in the intervals $[x_{i-1}, x_i]$ and $[y_{j-1}, y_j]$ respectively. It is now very natural to consider the limit of this sum when the partition lengths $\|\mathcal{Z}_x\|$ and $\|\mathcal{Z}_y\|$ both tend to 0 independently.

18 Integrals over a rectangle

The real valued function $f(x,y)$ is defined on the rectangle $\mathcal{R} = [a,b] \times [c,d]$. The integral $\int_{\mathcal{R}} f(x,y)d(x,y)$ is the limit

$$\int_{\mathcal{R}} f(x,y)d(x,y) := \lim_{\|\mathcal{Z}_x\|,\|\mathcal{Z}_y\| \to 0} \sum_{i,j} f(\xi_i, \eta_j)(x_i - x_{i-1})(y_j - y_{j-1}) \ , \tag{54}$$

with $\xi_i \in [x_{i-1}, x_i]$ and $\eta_j \in [y_{j-1}, y_j]$. Instead of $\int_{\mathcal{R}} f(x,y)d(x,y)$ we may also use the notation $\int_{\mathcal{R}} f(x,y)dS$.

We will now turn to the question how to calculate these new limits $\lim_{\|\mathcal{Z}_x\|,\|\mathcal{Z}_y\| \to 0} \sum_{i,j} f(\xi_i, \eta_j)(x_i - x_{i-1})(y_j - y_{j-1})$ and find that they can be written as two successive Riemann integrals:

$$\begin{aligned}
\int_{\mathcal{R}} f(x,y)d(x,y) &= \lim_{\|\mathcal{Z}_x\|,\|\mathcal{Z}_y\| \to 0} \sum_{i,j} f(\xi_i, \eta_j)(x_i - x_{i-1})(y_j - y_{j-1}) \\
&= \lim_{\|\mathcal{Z}_x\| \to 0} \left\{ \lim_{\|\mathcal{Z}_y\| \to 0} \sum_{i,j} f(\xi_i, \eta_j)(y_j - y_{j-1}) \right\} (x_i - x_{i-1}) \qquad (55) \\
&= \lim_{\|\mathcal{Z}_x\| \to 0} \int_c^d f(\xi_i, y)dy \ \ (x_i - x_{i-1}) = \int_a^b \left(\int_c^d f(x,y)dy \right) dx \\
&=: \int_a^b \int_c^d f(x,y)dy\,dx \ .
\end{aligned}$$

This of course assumes that the inner integrals $\int_c^d f(x,y)dy$ exist for all $x \in [a,b]$. In the same way we could of course first integrate along the x-axis and then along the y-axis provided that

in this case the inner integrals $\int_a^b f(x,y)dx$ exist for all $y \in [c,d]$. The main achievement of our finding is that this again gives us a very simple way of calculating the complicated double limits $\int_{\mathcal{R}} f(x,y)d(x,y)$ in terms of two successive simple one dimensional Riemann integrals. This result is known as *Fubini's theorem*[e].

19 Fubini's theorem in two dimensions
The real valued function $f(x,y)$ is defined on the rectangle $\mathcal{R} = [a,b] \times [c,d]$. If the integrals $\int_c^d f(x,y)dy$ exist for all $x \in [a,b]$ then

$$\int_{\mathcal{R}} f(x,y)d(x,y) = \int_a^b \int_c^d f(x,y)dy\,dx\ , \tag{56}$$

and if the integrals $\int_a^b f(x,y)dx$ exist for all $y \in [c,d]$ then

$$\int_{\mathcal{R}} f(x,y)d(x,y) = \int_c^d \int_a^b f(x,y)dx\,dy\ . \tag{57}$$

Example 3.4 *Let us consider the real valued function $f(x,y) = xye^{-(x^2+y^2)}$ on the rectangle $\mathcal{R} = [1,4] \times [1,2]$. The graph of $f(x,y)$ is given in* FIG. xviii *The integral $\int_{\mathcal{R}} xye^{-(x^2+y^2)}d(x,y)$ computes the volume underneath the graph. We obtain*

$$\begin{aligned}\int_{\mathcal{R}} xye^{-(x^2+y^2)}d(x,y) &= \int_1^2 \int_1^4 xye^{-(x^2+y^2)}dx\,dy = \int_1^2 \left[-\frac{ye^{-(x^2+y^2)}}{2}\right]_{x=1}^{x=4} dy \\ &= \int_1^2 \frac{ye^{-(y^2+1)} - ye^{-(y^2+16)}}{2}dy = \frac{e^{-20} - e^{-17} - e^{-5} + e^{-2}}{4}\ .\end{aligned}$$

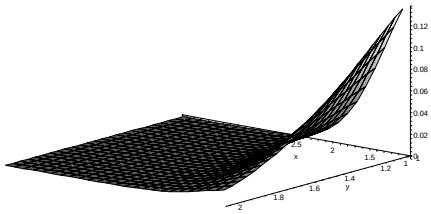


FIG. xviii Graph of $f(x,y) = xye^{-(x^2+y^2)}$.

We can now easily generalise integrals over rectangles to integrals over cuboids[f] in $\mathbb{R}^3$.

20 Integrals over 3-dimensional cuboids
The real valued function $f(x,y,z)$ is defined on the cuboid $\mathcal{Q} = [a,b] \times [c,d] \times [e,f]$. The integral $\int_{\mathcal{Q}} f(x,y,z)d(x,y,z)$ is the limit

$$\int_{\mathcal{Q}} f(x,y,z)d(x,y,z) := \lim_{\|\mathcal{Z}_x\|,\|\mathcal{Z}_y\|\mathcal{Z}_z\|\|\to 0} \sum_{i,j,k} f(\xi_i,\eta_j,\mu_k)(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})\ , \tag{58}$$

with $\xi_i \in [x_{i-1},x_i]$, $\eta_j \in [y_{j-1},y_j]$ and $\mu_k \in [z_{k-1},z_k]$. Instead of $\int_{\mathcal{Q}} f(x,y,z)d(x,y,z)$ we may also use the notation $\int_{\mathcal{Q}} f(x,y,z)dV$.

[e]Guido Fubini, 1879-1943.

[f]This could very easily be generalised to integrals over *n-dimensional intervals* $[a_1,b_1] \times [a_2,b_2] \times \ldots [a_n,b_n]$ in $\mathbb{R}^n$ but we want to restrict ourselves to $n = 2$ and $n = 3$.

How can we interpret such a 3-dimensional integral over a cuboid? We can imagine that the cuboid is made out of an inhomogeneous material with the mass (volume) density $f(\mathbf{x})$. Then $f(\xi_i, \eta_j, \mu_k)(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$ is an approximation for the masses of the small cuboids given by the partitions $\mathcal{Z}_x$, $\mathcal{Z}_y$ and $\mathcal{Z}_z$ and therefore the sum $\sum_{i,j,k} f(\xi_i, \eta_j, \mu_k)(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$ would obviously tend to the mass of the cuboid in the limit $\lim_{\|\mathcal{Z}_x\|,\|\mathcal{Z}_y\|\mathcal{Z}_z\|\|\to 0}$.

Fubini's theorem allows us again to break these integrals down into successive Riemann integrals:

21 **Fubini's theorem**
The real valued function $f(x, y, z)$ is defined on the cuboid $\mathcal{Q} = [a, b] \times [c, d] \times [e, f]$. If the integrals $\int_e^f f(x, y, z)dz$ and $\int_c^d \int_e^f f(x, y, z)dzdy$ exist for all $x \in [a, b]$ and all $y \in [c, d]$, then

$$\int_{\mathcal{Q}} f(x, y, z)d(x, y, z) = \int_a^b \int_c^d \int_e^f f(x, y, z)dz\, dy\, dx \; . \tag{59}$$

We could have obviously chosen a different order of integration in [21]. Fubini's theorem therefore tells us in particular that the order of integration does not matter.

22 **Exchange of order of integration**
The function $f(x, y)$ is integrable on the rectangle $\mathcal{R} = [a, b] \times [c, d]$ then

$$\int_a^b \int_c^d f(x, y)dy\, dx \;\; = \;\; \int_c^d \int_a^b f(x, y)dx\, dy = \int_{\mathcal{R}} f(x, y)d(x, y) \; , \tag{60}$$

provided all inner integrals $\int_a^b f(x, y)dx$ and $\int_c^d f(x, y)dy$ exist for all $y \in [c, d]$ and all $x \in [a, b]$ respectively.

Equation Eq. 60 can of course also be used for more than two successive integrals. For example a 3-dimensional integral $\int_{\mathcal{Q}} f(x, y, z)dV$ is independent of the order of integration provided all the successive integrals exist for all the required values of x, y, z. We call $\int_a^b \int_c^d f(x, y)dy\, dx$ and $\int_c^d \int_a^b f(x, y)dx\, dy$ the *iterated integrals* of $\int_{\mathcal{R}} f(x, y)d(x, y)$. The condition that all the successive integrals of the iterated integral have to exist should not be underestimated. It can for example very easily be shown that the two iterated integrals $\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3}\, dy\right] dx$ and $\int_0^1 \left[\int_0^1 \frac{x-y}{(x+y)^3}\, dx\right] dy$ both exist but their values are in fact different. The reason why Fubini's theorem cannot be used in this case is simply that the inner integrals $\int_0^1 \frac{x-y}{(x+y)^3}\, dy$ and $\int_0^1 \frac{x-y}{(x+y)^3}\, dx$ do not exist for $x = 0$ or $y = 0$, respectively and therefore the conditions of [19] do not hold.

5 Integration over plane surfaces and Volumes

We can now integrate over rectangles in $\mathbb{R}^2$ and also over cuboids in $\mathbb{R}^3$, but how useful is that? What if we want to integrate over some different plane surface rather than a rectangle or some different volume rather than a cuboid? This can very easily be incorporated in the theory developed in Sec. 4 by using the characteristic function of a set $\mathcal{B} \subset \mathbb{R}^n$

$$\chi_{\mathcal{B}}(\mathbf{x}) := \begin{cases} 1 & , \mathbf{x} \in \mathcal{B} \\ 0 & , \mathbf{x} \notin \mathcal{B} \end{cases} \; . \tag{61}$$

Given a bounded[g] set $\mathcal{B} \subset \mathbb{R}^n$, we can then find a rectangle ($n = 2$) or a cuboid ($n = 3$) which completely contains $\mathcal{B}$. We then continue the function $f(\mathbf{x})$ defined on B to $\mathcal{R}$ by setting the function values outside B simply to 0. We define the integral of $f(\mathbf{x})$ over the set $\mathcal{B}$ as the integral of the function $f(\mathbf{x})\chi_{\mathcal{B}}(\mathbf{x})$ over[h] $\mathcal{R}$. In this way we can define an integral over general plane surfaces and general volumes in $\mathbb{R}^2$ or $\mathbb{R}^3$ respectively simply by using the theory developed in SEC. 4.

23 Integrals over plane surfaces

The integral of the real valued function $f(x, y)$ over the set $\mathcal{S}$ is[i]*:*

$$\int_{\mathcal{S}} f(x,y)dS \quad := \quad \int_{\mathcal{R}} f(x,y)\chi_{\mathcal{S}}(x,y)\, dS \,, \tag{62}$$

where $\mathcal{R}$ is a rectangle completely containing $\mathcal{S}$. In particular

$$\|\mathcal{S}\| := \int_{\mathcal{S}} dS \,, \tag{63}$$

is called the area of $\mathcal{S}$.

24 Integrals over volumes

The integral of the real valued function $f(x, y, z)$ over the set $\mathcal{V}$ is[j]*:*

$$\int_{\mathcal{V}} f(x,y,z)dV \quad := \quad \int_{\mathcal{R}} f(x,y,z)\chi_{\mathcal{V}}(x,y,z)\, dV \,, \tag{64}$$

where $\mathcal{Q}$ is a cuboid completely containing $\mathcal{V}$. In particular

$$\|\mathcal{V}\| := \int_{\mathcal{V}} dV \,, \tag{65}$$

is called the volume of $\mathcal{V}$.

It should be noted that for these integrals to exist we do not only require the functions $f(x, y)$ or $f(x, y, z)$ to be well-behaved but also the functions $\chi_{\mathcal{S}}(x, y)$ or $\chi_{\mathcal{V}}(x, y, z)$ need to be well-behaved. This basically means that the sets $\mathcal{S}$ or $\mathcal{V}$ need to have a smooth (or at least piecewise smooth) boundary. For this course we shall simply assume that the sets we want to integrate over satisfy this condition and leave the details to future Analysis courses. It is also immediately obvious how these integrals could be generalised to n-dimensional integrals. Mathematicians call n-dimensional sets which are integrable in this sense *Jordan measurable* and the corresponding n-dimensional volumes are called the *Jordan measure*.

It is clear why $\int_{\mathcal{S}} d(x, y)$ is defined to be the *area* of the surface $\mathcal{S}$ since the integral is the limit of sums $\sum_{i,j}(x_i - x_{i-1})(y_j - y_{j-1})$ over the surface $\mathcal{S}$. Some of these rectangles may go slightly outside the surface $\mathcal{S}$ but in the limit the sum over the rectangles described by $(x_i - x_{i-1})(y_j - y_{j-1})$ will approximate the area of the surface. In the same way we understand why $\int_{\mathcal{V}} d(x, y, z)$ describes a *volume*. It is also immediately obvious from FIG. xvii that the interpretation of $\int_{\mathcal{S}} f(\mathbf{x})d(x, y)$ as the volume underneath the graph of $f(\mathbf{x})$ coincides with the definition [24] of the volume $\int_{\mathcal{V}} d(x, y, z)$ as in FIG. xvii.

---

[g] The theory easily generalises to infinite sets but for our applications the bounded case is sufficient.

[h] Note that this definition does not really depend on $\mathcal{R}$ as long as $\mathcal{B} \subset \mathcal{R}$, and also that the function values outside $\mathcal{B}$ are not really needed.

[i] Instead of $\int_{\mathcal{S}} f(x, y)dS$ we may also frequently use the notations $\int_{\mathcal{S}} f(x, y)d(x, y)$.

[j] Instead of $\int_{\mathcal{V}} f(x, y, z)dV$ we may also frequently use the notation $\int_{\mathcal{V}} f(x, y, z)d(x, y, z)$.

25 Integration rules

$$\int_{\mathcal{V}} (\alpha f(\mathbf{x}) + \beta g(\mathbf{x}))dV = \alpha \int_{\mathcal{V}} f(\mathbf{x})dV + \beta \int_{\mathcal{V}} g(\mathbf{x})dV\,, \tag{66}$$

$$\int_{\mathcal{V}_1 \cup \mathcal{V}_2} f(\mathbf{x})dV = \int_{\mathcal{V}_1} f(\mathbf{x})dV + \int_{\mathcal{V}_2} f(\mathbf{x})dV\,, \tag{67}$$

where $\mathcal{V}_1$ and $\mathcal{V}_2$ are disjoint or at least disjoint up to boundary points. Similar rules apply for plane surface integrals.

We can now easily apply Fubini's theorem in order to evaluate these integrals:

Example 3.5 *Let us take in two dimensions the plane surface $\mathcal{S}$ bounded by the unit circle $x^2+y^2=1$ and by $x, y \geq 0$. How do we calculate the integral $\int_{\mathcal{S}} f(x,y)\, d(x,y)$ of the real valued function $f(x,y) = xy^2$ over $\mathcal{S}$?*

We embed $\mathcal{S}$ into the square $\mathcal{R} = [0,1] \times [0,1]$ and use Fubini's theorem for the integral of $f(\mathbf{x})\chi_{\mathcal{S}}(\mathbf{x})$ in order to first integrate parallel to the x-axis on the interval $[0,1]$ and then parallel to the y-axis also on $[0,1]$. However, since $f(\mathbf{x})\chi_{\mathcal{S}}(\mathbf{x})$ is 0 outside $\mathcal{S}$ it is clear that the upper limit of the integral parallel to the x-axis is not at 1 but actually at $\sqrt{1-y^2}$ and therefore depends on the actual value of y which could be anywhere in $[0,1]$.

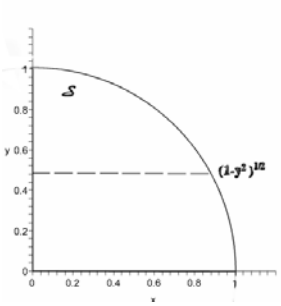


FIG. xix $\mathcal{S}$ bounded by $x^2+y^2=1$; $x, y \geq 0$.

We therefore obtain

$$\begin{aligned} \int_{\mathcal{S}} xy^2\, d(x,y) &= \int_0^1 \int_0^{\sqrt{1-y^2}} xy^2 dx\, dy \\ &= \int_0^1 \left[\frac{x^2y^2}{2}\right]_{x=0}^{x=\sqrt{1-y^2}} dy = \int_0^1 \frac{y^2-y^4}{2} dy = \frac{1}{15}\,. \end{aligned} \tag{68}$$

It is now important to note that the upper and lower bounds of the inner integrals can depend on the variables of the outer integrals (but of course not vice versa). Instead of first integrating along x and then interacting along y we could have done it the other way round according to [22]. We would then obtain the integral $\int_0^1 \int_0^{\sqrt{1-x^2}} xy^2 dy\, dx = \int_0^1 \frac{x(1-x^2)^{\frac{3}{2}}}{3} dx$ which will also give $\frac{1}{15}$ but is obviously more complicated.

Example 3.6 *In the same way we obtain in three dimensions three successive integrals with the boundaries of the inner integrals depending on the variables of the outer integrals. For example if the continuous function $f(x,y,z)$ is integrated over the volume $\mathcal{V}$ bounded by the ellipsoid $x^2+y^2+\frac{z^2}{4}=1$, then we can calculate $\int_{\mathcal{V}} f(x,y,z)d(x,y,z)$ in the following way: for fixed x and y the lower and upper bounds for the z co-ordinate are $-2\sqrt{1-x^2-y^2}$ and $2\sqrt{1-x^2-y^2}$. Projecting the ellipsoid into the xy-plane gives the circle $x^2+y^2=1$ and for fixed x we have y going from $-\sqrt{1-x^2}$ to $\sqrt{1-x^2}$. Finally projecting the circle on the x-axis shows that x runs from -1 to 1. We therefore find $\int_{\mathcal{V}} f(x,y,z)d(x,y,z) = \int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-2\sqrt{1-x^2-y^2}}^{2\sqrt{1-x^2-y^2}} f(x,y,z)dz\, dy\, dx$.*

Most of the time the most difficult problem of integration over plane surfaces or volumes is actually describing the boundary of the plane surface of the volume correctly and easily in the successive integrals. As we have seen in our simple examples, by choosing a suitable order for the successive integrals the task of describing the lower and upper bounds could be simplified significantly which would also influence how complicated the successive integrations are. Therefore it is essential that we first think very carefully about a suitable order rather than picking a random order. Sometimes using Cartesian co-ordinates may not be the ideal solution for solving such an integral. We will study in SEC. 6 how a suitable substitution may simplify such an integral.

Finally, and just for completeness, we generalise integrals over real valued functions to integrals over vector valued functions simply by defining them as vectors of integrals over the real valued component functions.

26 **Integrals over vector valued functions**
The integral of the vector valued function $\mathbf{f}(\mathbf{x})$ *over the volume* $\mathcal{V}$ *is:*

$$\int_{\mathcal{V}} \mathbf{f}(\mathbf{x})\, dV \quad := \quad \begin{pmatrix} \int_{\mathcal{V}} f_1(\mathbf{x})\, dV \\ \vdots \\ \int_{\mathcal{B}} f_m(\mathbf{x})\, dV \end{pmatrix}, \tag{69}$$

and similarly for plane surface integrals.

6 Substitution rule

For the one dimensional Riemann integral the substitution rule

$$\int_{g(a)}^{g(b)} f(x)dx \quad = \quad \int_a^b f(g(t))g'(t)dt \tag{70}$$

plays a central rôle if we actually want to solve complicated integrals. In EQ. 70 the substitution function $g(t)$ is either strictly increasing or strictly decreasing which corresponds to $g'(t) > 0$ or $g'(t) < 0$ for all $t \in (a, b)$ respectively. Since our definition of integrals over higher dimensional functions follows the same ideas as in one dimension it is not surprising that a very similar substitution rule holds in higher dimensions:

27 **Substitution rule for integrals in** $\mathbb{R}^n$
If $\mathbf{g} : \mathcal{B} \to \hat{\mathcal{B}}$ *is a bijective transformation from the set* $\mathcal{B}$ *onto* $\hat{\mathcal{B}} = \mathbf{g}(\mathcal{B})$ *with either* $\det J_{\mathbf{g}}(\mathbf{t}) > 0$ *or* $\det J_{\mathbf{g}}(\mathbf{t}) < 0$ *for all* $\mathbf{t} \in \mathcal{B}$ *then*

$$\int_{\mathbf{g}(\mathcal{B})} f(\mathbf{x})\, d(x_1, \ldots, x_n) \quad = \quad \int_{\mathcal{B}} f(\mathbf{g}(\mathbf{t}))\ |\det J_{\mathbf{g}}(\mathbf{t})|\ d(t_1, \ldots, t_n)\ , \tag{71}$$

where $|\det J_{\mathbf{g}}(\mathbf{t})|$ *denotes the absolute value of the determinant of the Jacobi matrix of the transformation* $\mathbf{g}(\mathbf{t})$. $|\det J_{\mathbf{g}}(\mathbf{t})|$ *is called the Jacobian of the transformation* $\mathbf{g}(\mathbf{t})$.

We call the function $g(\mathbf{t})$ *a substitution function and for the Jacobian we use the notation*

$$\frac{d(x_1, \ldots, x_n)}{d(t_1, \ldots, t_n)} \quad := \quad |\det J_{\mathbf{g}}(\mathbf{t})|\ . \tag{72}$$

The fact that either $\det J_{\mathbf{g}}(\mathbf{t}) > 0$ or $\det J_{\mathbf{g}}(\mathbf{t}) < 0$ for all $\mathbf{t} \in \mathcal{B}$ makes sure that the transformation $\mathbf{g}(\mathbf{t})$ is invertible as explained in Part IA Differential Equations. This is completely analogous to $g'(t) > 0$ or $g'(t) < 0$ in Eq. 70. The reason why we have to use the absolute value $|\det J_{\mathbf{g}}(\mathbf{t})|$ rather than just $\det J_{\mathbf{g}}(\mathbf{t})$ is that by definition of the integral $\int_{\mathcal{B}} f(\mathbf{x}) d(x_1, \ldots, x_n)$ we cannot keep track of the fact that under the substitution certain co-ordinate axes may have inverted their directions and therefore we do not know if we need to introduce a suitable number of minus signs. For example, in the one dimensional Riemann case, if the substitution function $g(t)$ is strictly decreasing then the *lower limit* $g(a)$ is be larger than the *upper limit* $g(b)$ provided $a < b$. But the Riemann integral $\int_{g(a)}^{g(b)} f(x)dx$ corrects itself simply by using $\int_{g(a)}^{g(b)} f(x)dx = -\int_{g(b)}^{g(a)} f(x)dx$. The multidimensional integral $\int_{\mathbf{g}(\mathcal{B})} f(\mathbf{x}) d(x_1, \ldots, x_n)$ does not keep track of such changes of directions but by taking $|\det J_{\mathbf{g}}(\mathbf{t})|$ as Jacobian rather than $\det J_{\mathbf{g}}(\mathbf{t})$ just corrects a possible overall change in orientation.

Example 3.7 *The area of the disc $\mathcal{D}$: $x^2 + y^2 \leq R^2$ is given by $\int_{\mathcal{D}} d(x, y)$. Describing the disc in terms of Cartesian co-ordinates ultimately involves lower and upper limits of the form $\pm\sqrt{R^2 - x^2}$, it therefore seems much more natural to describe the disc in terms of plane polar co-ordinates: $r \in [0, R]$ and $\phi \in [0, 2\pi]$ describes $\mathcal{D}$. The Jacobi matrix of the transformation $x = r\cos\phi$ and $y = r\sin\phi$ is*

$$J = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} , \tag{73}$$

and therefore we find for the Jacobian $\frac{d(x,y)}{d(r,\phi)} = r$. Hence we find the well known formula for the area of a disc:

$$\|\mathcal{D}\| = \int_{\mathcal{D}} d(x, y) = \int_0^R \int_0^{2\pi} r \, d\phi \, dr = \pi R^2 . \tag{74}$$

28 **Transformation to plane polar coordinates**
The Jacobian for plane polar coordinates is

$$\frac{d(x, y)}{d(r, \phi)} = r . \tag{75}$$

The transformation to plane polar coordinates is then given by

$$\int_{\hat{\mathcal{S}}} f(x, y) \, d(x, y) = \int_{\mathcal{S}} f(r\cos\phi, r\sin\phi) \; r \; d(r, \phi) , \tag{76}$$

where $\hat{\mathcal{S}}$ and $\mathcal{S}$ describe the corresponding surface in the xy-plane or in the $r\phi$-plane respectively.

29 **Transformation to cylindrical polar coordinates**
The Jacobian for cylindrical polar coordinates is

$$\frac{d(x, y, z)}{d(\rho, \phi, z)} = \rho . \tag{77}$$

The transformation to cylindrical polar coordinates is then given by

$$\int_{\hat{\mathcal{V}}} f(x, y, z) \, d(x, y, z) = \int_{\mathcal{V}} f(\rho\cos\phi, \rho\sin\phi, z) \; \rho \; d(\rho, \phi, z) , \tag{78}$$

where $\hat{\mathcal{V}}$ and $\mathcal{V}$ describe the corresponding volume in xyz-space or in $\rho\phi z$-space respectively.

30 **Transformation to spherical polar coordinates**
The Jacobian for spherical polar coordinates is

$$\frac{d(x,y,z)}{d(r,\phi,\theta)} = r^2 \sin\theta \, . \tag{79}$$

The transformation to spherical polar coordinates is then given by

$$\int_{\hat{\mathcal{V}}} f(x,y,z)\, d(x,y,z) = \int_{\mathcal{V}} f(r\cos\phi\, \sin\theta, r\sin\phi\, \sin\theta, r\cos\theta)\; r^2 \sin\theta \; d(r,\phi,\theta) \, , \tag{80}$$

where $\hat{\mathcal{V}}$ and $\mathcal{V}$ describe the corresponding volume in xyz-space or in $r\phi\theta$-space respectively.

For many applications the coordinate transformation is defined through its inverse rather than through the function itself but actually inverting the coordinate transformation may prove very difficult. However, the transformation of the integral in Eq. 70 does not require us to explicitly invert the transformation, all we need is the determinant of the Jacobi matrix of the inverse transformation. In Part IA Differential Equations you found that the Jacobi matrix of the inverse of a function is just the inverse matrix of the Jacobi matrix of the function. We also know that the determinant of the inverse of a matrix is simply the reciprocal value of the determinant of the matrix. We therefore find:

31 **Jacobian of the inverse transformation**
If $\mathbf{g} : \hat{\mathcal{B}} \to \mathcal{B}$ is a substitution function then $g(\mathbf{t})$ is invertible and the inverse function $\mathbf{g}^{-1} : \mathcal{B} \to \hat{\mathcal{B}}$ is again a substitution function with Jacobian

$$\left|\det J_{\mathbf{g}^{-1}}(\mathbf{x})\right| = \frac{1}{\left|\det J_{\mathbf{g}}(\mathbf{t})\right|} \, , \tag{81}$$

or in other words:

$$\frac{d(x_1,\ldots,x_n)}{d(t_1,\ldots,t_n)} = \frac{1}{\frac{d(t_1,\ldots,t_n)}{d(x_1,\ldots,x_n)}} \, . \tag{82}$$

Examples

Example 3.8 *We will calculate, in two ways, the integral*

$$I = \int_{\mathcal{R}} y \, dS \, ,$$

where $\mathcal{R}$ is specified by

$$1 \le x \le 2 \quad \textit{and} \quad 1/x \le y \le e^x \, .$$

(i) Do the y-integration first. We have

$$\begin{aligned} I = \int_{x=1}^{2} \left(\int_{y=1/x}^{e^x} y \, dy \right) dx &= \int_{x=1}^{2} \left[\tfrac{1}{2} y^2\right]_{1/x}^{e^x} dx \\ &= \tfrac{1}{2} \int_1^2 (e^{2x} - 1/x^2) dx \\ &= \tfrac{1}{4}(e^4 - e^2 - 1) \end{aligned}$$

(ii) Do the x integral first. We have

$$\begin{aligned} I &= \int_{\frac{1}{2}}^{1} \left(\int_{1/y}^{2} y \, dx \right) dy + \int_{1}^{e} \left(\int_{1}^{2} y \, dx \right) dy + \int_{e}^{e^2} \left(\int_{\ln y}^{2} y \, dx \right) dy \\ &= \int_{\frac{1}{2}}^{1} y(2 - 1/y) dy + \int_{1}^{e} y \, dy + \int_{e}^{e^2} y(2 - \ln y) dy \\ &= \textit{same as the previous answer.} \end{aligned}$$

The huge brackets (parentheses) in the first line of each calculation can be omitted without creating ambiguity.

Example 3.9 *We will calculate, in two different ways, the integral*

$$I = \int_{\mathcal{V}} x^2 dV \ ,$$

where $\mathcal{V}$ is the tetrahedron bounded by the four planes

$$x = 0, \ y = 0, \ z = 0 \quad \textit{and} \quad x + y + z = a \, .$$

The vertices of the tetrahedron are at $(0, 0, 0)$, $(a, 0, 0)$, $(0, a, 0)$ and $(0, 0, a)$.

(i) Do the x integral last. We have

$$\begin{aligned} I &= \int_0^a \int_0^{a-x} \int_0^{a-x-y} x^2 \, dz \, dy \, dx \\ &= \int_0^a \int_0^{a-x} x^2 (a - x - y) \, dy \, dx \\ &= \int_0^a \tfrac{1}{2} x^2 (a - x)^2 \, dx = \frac{a^5}{60}. \end{aligned}$$

(ii) Do the x integral first. We have

$$\begin{aligned} I &= \int_0^a \int_0^{a-z} \int_0^{a-z-y} x^2 \, dx \, dy \, dz \\ &= \int_0^a \int_0^{a-z} \tfrac{1}{3} (a - z - y)^3 dy \, dz \\ &= \int_0^a \tfrac{1}{12} (a - z)^4 dz = \frac{a^5}{60}. \end{aligned}$$

Example 3.10 *We will calculate the Gaussian integral I given by*

$$I = \int_0^\infty e^{-x^2} dx$$

by means of a remarkably cunning plan.

We consider

$$I^2 = \int_0^\infty e^{-x^2} dx \int_0^\infty e^{-y^2} dy = \int_0^\infty \int_0^\infty e^{-x^2 - y^2} dx dy$$

We now change coordinates to plane polars given by $x = \rho\cos\phi$, $y = \rho\sin\phi$. *First we calculate the Jacobian.*

$$\frac{\partial x}{\partial \rho} = \cos\phi\,, \qquad \frac{\partial x}{\partial \phi} = -\rho\sin\phi\,,$$
$$\frac{\partial y}{\partial \rho} = \sin\phi\,, \qquad \frac{\partial y}{\partial \phi} = \rho\cos\phi\,.$$

Hence

$$\frac{\partial(x,y)}{\partial(\rho,\phi)} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} \end{vmatrix} = \rho\cos^2\phi + \rho\sin^2\phi = \rho\,.$$

Thus

$$I^2 = \int_0^\infty \int_0^{\pi/2} e^{-\rho^2}\,\rho\,d\phi\,d\rho = \frac{\pi}{2}\int_0^\infty e^{-\rho^2}\,\rho\,d\rho = \frac{\pi}{4}$$

and $I = \frac{1}{2}\sqrt{\pi}$ *(taking the positive square root since the integrand is positive).*

Example 3.11 *We will evaluate*

$$I = \int_{\mathcal{R}} \frac{1}{x^2}\,dS\,,$$

where the region $\mathcal{R}$ *is bounded by the lines*

$$y = 0\,,\ y = x\,,\ x + y = 1 \quad \textit{and} \quad x + y = 2\,.$$

Since the boundaries can be expressed as either $y = kx$ *or* $x + y = c$, *this suggests trying the new variables*

$$u = x + y \quad \textit{and} \quad v = y/x\,.$$

The Jacobian is given by

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -\frac{y}{x^2} & \frac{1}{x} \end{vmatrix} = \frac{x+y}{x^2}\,,$$

and

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{x^2}{x+y}\,.$$

Thus

$$I = \int_1^2 \int_0^1 \frac{1}{u}\,dv\,du = \ln 2\,.$$

Note that if instead we had chosen $u = y/x$ *and* $v = x + y$ *then the Jacobian would have changed sign since the handedness of the system would not have been preserved. We would no doubt have remembered to use the modulus of the Jacobian,*

Of course, we could have done the integral fairly easily without changing variables.

7 Surface integrals in $\mathbb{R}^3$

We studied line integrals in Sec. 3 motivated by the fact that the scalar product of force and a displacement vector gives us the energy gained or needed along the movement of a point mass. The scalar product plays a similar rôle if we study the flow of a fluid through a surface but this time we have a two dimensional problem. Let us assume that we observe a straight and

homogeneous flow of water with a flow density of f litres of water per second per metre square. If we then put an open tube straight into the flow, and the tube opening has a surface of s metres square, how much water will we collect through this tube? The answer is obviously fs litres of water per second.

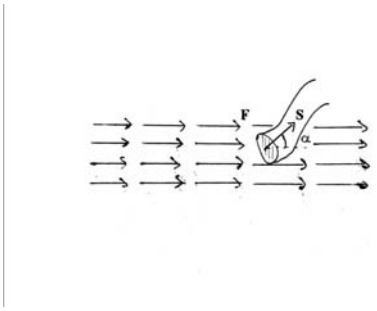


FIG. xx Flow through a surface $\mathcal{S}$.

But what if we put the tube into the flow at an angle α to the flow? Obviously the effective surface of the tube opening seen by the water flow is now reduced to $s\cos\alpha$ and therefore the flow through the tube is $fs\cos\alpha$. It therefore makes sense to assign a vector to the flow of water $\mathbf{F}$ with the direction of $\mathbf{F}$ pointing in the direction of the flow of water and the magnitude $\|\mathbf{F}\|$ being the flow density f. We also assign a vector to the surface of the opening of the tube $\mathbf{S}$ with a direction normal to the surface of the opening and the magnitude $\|\mathbf{S}\|$ being the area s.

By taking the scalar product $\mathbf{F}.\mathbf{S} = fs\cos\alpha$ we therefore obtain the flow of water through the opening of the tube even if it is held inside the flow at an angle α. But what if the flow is not homogeneous? We would therefore have a vector field $\mathbf{F}(\mathbf{x})$ describing the flow density of the water. $\mathbf{F}(\mathbf{x})$ might then vary with $\mathbf{x}$ in terms of direction as well as magnitude! If we then ask how much water will flow through the opening described by the vector $\mathbf{S}$ then by now it is obvious to us what we have to do: we partition the (two dimensional) surface into small surfaces $\mathcal{S}_{ij}$ described by the vectors $\mathbf{S}_{ij}$, evaluate $\mathbf{F}(\mathbf{x})$ somewhere on each $\mathcal{S}_{ij}$ at (ξ_i, η_j) and keep $\mathbf{F}(\xi_i, \eta_j)$ constant on the *surface element* $\mathbf{S}_{ij}$. Each of the vectors $\mathbf{S}_{ij}$ is normal to the corresponding $\mathcal{S}_{ij}$ and the magnitude matches the area of $\mathcal{S}_{ij}$. If the surface elements are *small* then the scalar product $\mathbf{F}(\xi_i, \eta_j).\mathbf{S}_{ij}$ obviously approximates the amount of water flowing through $\mathbf{S}_{ij}$ and summing over all these scalar products $\sum_{i,j} \mathbf{F}(\xi_i, \eta_j).\mathbf{S}_{ij}$ approximates the amount of water flowing through the whole opening. In case the shape of the opening and the vector field $\mathbf{f}(\mathbf{x})$ are both *well-behaved* then we hope that the limit of this sum under the partition lengths tending to 0 would exist and would therefore describe the flow of water through the opening of the tube exactly.

Let us now formalise this procedure. The opening of the tube in our example is of course a flat surface, which essentially means that all $\mathbf{S}_{ij}$ in the above example point in the same direction. But there is in fact no need for this surface to be flat and that in the sum $\sum_{i,j} \mathbf{F}(\xi_i, \eta_j).\mathbf{S}_{ij}$ all $\mathbf{S}_{ij}$ point in the same direction. Being able to generalise this procedure to surfaces which are not flat would allow us to calculate the *flow* of a vector field $\mathbf{F}(\mathbf{x})$ even through surfaces $\mathcal{S}$ in $\mathbb{R}^3$ which are not flat. The procedure we are going to follow is now straightforward generalisation of the integration techniques studies earlier but before we can start we will first need to define surfaces in a suitable way.

Let us assume that the vector valued function $\mathbf{\Phi}(u, v)$ describes the points on a surface $\mathcal{S}$ in $\mathbb{R}^3$. For example the function

$$\mathbf{\Phi}(\phi, \theta) \quad := \quad \begin{pmatrix} \cos\phi\,\sin\theta \\ \sin\phi\,\sin\theta \\ \cos\theta \end{pmatrix} \tag{83}$$

describes the surface of a sphere with radius one, the *unit sphere*, in $\mathbb{R}^3$ as ϕ and θ go through the intervals $[0, 2\pi]$ and $[0, \pi]$ respectively. Another important example is that for a real valued

function $f(x,y) : D \to \mathbb{R}$ the vector valued function

$$\mathbf{\Phi}(x,y) \quad := \quad \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix} \tag{84}$$

describes the surface in $\mathbb{R}^3$ which is given by the graph of the function $f(x,y)$ as (x,y) goes through the domain D.

The function $\mathbf{\Phi} : \mathcal{P} \subset \mathbb{R}^2 \to \mathbb{R}^3$ is defined on the set $\mathcal{P}$ which obviously has to be a subset of $\mathbb{R}^2$ in order to generate a two dimensional surface $\mathcal{S}$ in three dimensional space. We want to call $\mathcal{P}$ the *parameter space* of $\mathbf{\Phi}$ describing the surface $\mathcal{S}$. $\mathbf{\Phi}$ itself is called a *parametrisation* of the surface $\mathcal{S}$.

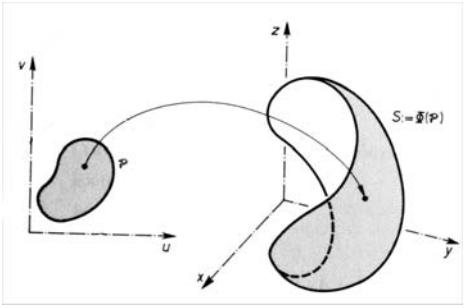


FIG. xxi Surface $\mathcal{S}$ with parameter space $\mathcal{P}$.

If we are now given a surface $\mathcal{S}$ with a parametrisation $\mathbf{\Phi}$ and parameter space $\mathcal{P}$ then we can easily partition $\mathcal{S}$ by partitioning $\mathcal{P}$ in the usual way. The partition of $\mathcal{P}$ will obviously result in a partition of $\mathcal{S}$. Let $\mathcal{Z}_u$ and $\mathcal{Z}_v$ denote the corresponding partitions on $\mathcal{P}$ with partition lengths $\|\mathcal{Z}_u\|$ and $\|\mathcal{Z}_v\|$. We now assume the surface to be flat on each surface element generated by this partition. As in our earlier example the vectors $\mathbf{S}_{ij}$ describe these small surface elements in the way that the direction of $\mathbf{S}_{ij}$ is normal to the surface element and the magnitude matches the area of the surface element. Assume now that the vector field $\boldsymbol{F}(\mathbf{x})$ is defined everywhere on the surface $\mathcal{S}$ then we are obviously interested in the following sums

$$\sum_{i,j} \mathbf{F}(\mathbf{\Phi}(\xi_i, \eta_j)).\mathbf{S}_{ij} \ , \tag{85}$$

where $\xi_i \in [u_{i-1}, u_i]$ and $\eta_j \in [v_{j-1}, v_j]$ and the u_is and v_js correspond to the partition of $\mathcal{P}$. We then consider the limit of EQ. 85 for the partition lengths $\|\mathcal{Z}_u\|$ and $\|\mathcal{Z}_v\|$ tending to 0 independently. We will call this integral the *surface integral* of $\mathbf{F}$ over the surface $\mathcal{S}$. We will see later that this limit is in fact independent of the chosen parametrisation $\mathbf{\Phi}$.

32 **Surface integrals over vector fields**
The function $\mathbf{\Phi} : \mathcal{P} \to \mathbb{R}^3$ is a parametrisation of the surface $\mathcal{S}$ with parameter space $\mathcal{P} \subset \mathbb{R}^2$ and $\mathbf{F}$ is a vector field in $\mathbb{R}^3$. The surface integral of $\mathbf{F}$ over the surface $\mathcal{S}$ is

$$\int_{\mathcal{S}} \mathbf{F}(\mathbf{x}).d\mathbf{S} = \lim_{\|\mathcal{Z}_u\|,\|\mathcal{Z}_v\| \to 0} \sum_{i,j} \mathbf{F}(\mathbf{\Phi}(\xi_i, \eta_j)).\mathbf{S}_{ij} \ , \tag{86}$$

where ξ_i and η_j are in $[u_{i-1}, u_i]$ or $[v_{j-1}, v_j]$ respectively.

As before the definition [32] does not tell us how to actually compute the integral. Our hope is of course that this very complicated integral would once more translate into simple one

dimensional Riemann integrals (provided $\mathbf{F}$ and $\mathcal{S}$ are both well-behaved). Let us therefore look more closely at the surface elements $\mathbf{S}_{ij}$.

Let us assume that the surface element $\mathbf{S}_{ij}$ is flat. In addition, just like for the proof of the substitution rule [27], we approximate the surface element $\mathbf{S}_{ij}$ using the parallelogram spanned by the points $\mathbf{\Phi}(u_{i-1}, v_{j-1})$, $\mathbf{\Phi}(u_i, v_{j-1})$ and $\mathbf{\Phi}(u_{i-1}, v_j)$. In the limit $\|\mathcal{Z}_u\|$ and $\|\mathcal{Z}_v\|$ tending to 0 this should give the correct surface integral.

The parallelogram is spanned by the vectors $\mathbf{\Phi}(u_i, v_{j-1}) - \mathbf{\Phi}(u_{i-1}, v_{j-1})$ and $\mathbf{\Phi}(u_{i-1}, v_j) - \mathbf{\Phi}(u_{i-1}, v_{j-1})$. We know that the cross product of two vectors $\mathbf{a} \times \mathbf{b}$ gives a vector normal to the plane spanned by $\mathbf{a}$ and $\mathbf{b}$ with the magnitude $\|\mathbf{a}\|\|\mathbf{b}\| \sin\alpha$ where $\sin\alpha$ is the angle between $\mathbf{a}$ and $\mathbf{b}$. But $\|\mathbf{a}\|\|\mathbf{b}\| \sin\alpha$ is equal to the area of the parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$. Therefore we can take the surface element $\mathbf{S}_{ij}$ to be the cross product of $\mathbf{\Phi}(u_i, v_{j-1}) - \mathbf{\Phi}(u_{i-1}, v_{j-1})$ and $\mathbf{\Phi}(u_{i-1}, v_j) - \mathbf{\Phi}(u_{i-1}, v_{j-1})$:

$$\mathbf{S}_{ij} \quad := \quad (\mathbf{\Phi}(u_i, v_{j-1}) - \mathbf{\Phi}(u_{i-1}, v_{j-1})) \times (\mathbf{\Phi}(u_{i-1}, v_j) - \mathbf{\Phi}(u_{i-1}, v_{j-1})) \ . \tag{87}$$

We assume that $\mathbf{\Phi}$ is smooth such that we can use the Taylor expansion for the two increments of $\mathbf{\Phi}$ in Eq. 87:

$$\mathbf{\Phi}(u_i, v_{j-1}) - \mathbf{\Phi}(u_{i-1}, v_{j-1}) \quad = \quad \frac{\partial \mathbf{\Phi}(u_{i-1}, v_{j-1})}{\partial u}(u_i - u_{i-1}) + \dots \ , \tag{88}$$

$$\mathbf{\Phi}(u_{i-1}, v_j) - \mathbf{\Phi}(u_{i-1}, v_{j-1}) \quad = \quad \frac{\partial \mathbf{\Phi}(u_{i-1}, v_{j-1})}{\partial v}(v_i - v_{i-1}) + \dots \ . \tag{89}$$

We hence obtain:

$$\mathbf{S}_{ij} \quad = \quad \frac{\partial \mathbf{\Phi}}{\partial u} \times \frac{\partial \mathbf{\Phi}}{\partial v} (u_i - u_{i-1})(v_j - v_{j-1}) \ . \tag{90}$$

It is now obvious that the surface integral can be converted into the plane two dimensional surface integral $\int_P \mathbf{F}(\mathbf{\Phi}(u,v)).\left(\frac{\partial \phi}{\partial u} \times \frac{\partial \phi}{\partial v}\right) d(u,v)$ over the (flat) parameter space P.

33 Surface integrals over differentiable surfaces

The surface $\mathcal{S}$ is parametrised by the function $\mathbf{\Phi}$ on the parameter space P, then the surface integral of the vector field $\mathbf{F}$ over $\mathcal{S}$ is

$$\int_{\mathcal{S}} \mathbf{F}(\mathbf{x}).d\mathbf{S} \quad = \quad \int_P \mathbf{F}(\mathbf{\Phi}(u,v)).\left(\frac{\partial \mathbf{\Phi}(u,v)}{\partial u} \times \frac{\partial \mathbf{\Phi}(u,v)}{\partial v}\right) d(u,v) \ . \tag{91}$$

For simplicity we use the notation

$$d\mathbf{S} \quad = \quad \frac{\partial \mathbf{\Phi}(u,v)}{\partial u} \times \frac{\partial \mathbf{\Phi}(u,v)}{\partial v} \, d(u,v) \ . \tag{92}$$

Just like for integrals along paths we of course need to assume that the parametrisation $\mathbf{\Phi}$ of the surface $\mathcal{S}$ is sufficiently smooth for the derivatives in Eq. 91 to exist. In case $\mathbf{\Phi}$ is only piecewise differentiable (i.e. it can be split up in finitely many differentiable surfaces) then we can simply integrate over each differentiable piece and add them all up. Therefore everything we are gong to say about differentiable parametrisations of surfaces is in this way also valid for piecewise differentiable parametrisations.

Example 3.12 *Let us for example consider the vector field $\mathbf{F}(\mathbf{x}) = \frac{\mathbf{x}}{r^3}$. We assume that $\mathbf{F}(\mathbf{x})$ describes the flow density of a fluid. How much of the fluid flows through the (open) upper*

hemisphere $x^2+y^2+z^2=R^2$, $z\geq 0$? *In order to find* $d\mathbf{S}$ *for the upper hemisphere we parametrise it using*

$$\mathbf{\Phi}(\mathbf{x}) \;=\; \begin{pmatrix} R\cos\phi\,\sin\theta \\ R\sin\phi\,\sin\theta \\ R\cos\theta \end{pmatrix}. \tag{93}$$

Taking the cross product of the two tangent vectors leads to

$$d\mathbf{S} \;=\; \frac{\partial\mathbf{\Phi}}{\partial\theta}\times\frac{\partial\mathbf{\Phi}}{\partial\phi} = \begin{pmatrix} R\cos\phi\,\cos\theta \\ R\sin\phi\,\cos\theta \\ -R\sin\theta \end{pmatrix}\times\begin{pmatrix} -R\sin\phi\,\sin\theta \\ R\cos\phi\,\sin\theta \\ 0 \end{pmatrix} = R^2\sin\theta\mathbf{e}_r d\theta\,d\phi\,. \tag{94}$$

We therefore obtain for the flow of the fluid through the upper hemisphere

$$\int_{\mathcal{S}}\mathbf{F}.d\mathbf{S} \;=\; \int_0^{2\pi}\int_0^{\frac{\pi}{2}}\frac{R\mathbf{e}_r}{R^3}R^2\sin\theta\,.\mathbf{e}_r\,d\theta\,d\phi = \int_0^{2\pi}\int_0^{\frac{\pi}{2}}\sin\theta d\theta\,d\phi = 2\pi\,. \tag{95}$$

It remains to be shown that the definition [32] of the surface integral is indeed independent of the explicit choice of the parametrisation $\mathbf{\Phi}(u,v)$. If we assume that $\hat{\mathbf{\Phi}}(\hat{u},\hat{v})$ is a different parametrisation of the same surface $\mathcal{S}$ with parameter space $\hat{P}$, then we can obviously find a function $\boldsymbol{\kappa}: P\to\hat{P}$ such that $\mathbf{\Phi}=\hat{\mathbf{\Phi}}\circ\boldsymbol{\kappa}$. Using the chain rule and setting $\hat{u}=\kappa_1(u,v)$ and $\hat{v}=\kappa_2(u,v)$ we find

$$\frac{\partial\mathbf{\Phi}}{\partial u}=\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{u}}\frac{\partial\kappa_1}{\partial u}+\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{v}}\frac{\partial\kappa_2}{\partial u}\quad\text{and}\quad\frac{\partial\mathbf{\Phi}}{\partial v}=\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{u}}\frac{\partial\kappa_1}{\partial v}+\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{v}}\frac{\partial\kappa_2}{\partial v}\,. \tag{96}$$

Hence we find

$$\frac{\partial\mathbf{\Phi}}{\partial u}\times\frac{\partial\mathbf{\Phi}}{\partial v} \;=\; \left(\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{u}}\times\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{v}}\right)\left(\frac{\partial\kappa_1}{\partial u}\frac{\partial\kappa_2}{\partial v}-\frac{\partial\kappa_2}{\partial u}\frac{\partial\kappa_1}{\partial v}\right)=\left(\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{u}}\times\frac{\partial\hat{\mathbf{\Phi}}}{\partial\hat{v}}\right)\det J_{\boldsymbol{\kappa}}\,, \tag{97}$$

where $\det J_{\boldsymbol{\kappa}}$ is the determinant of the Jacobi matrix of $\boldsymbol{\kappa}(u,v)$. Therefore a reparametrisation is simply a substitution and leads to the same surface integral up to a possible sign factor depending on the sign of $\det J_{\boldsymbol{\kappa}}$. We now also realise that the sign of the surface integral is not uniquely defined in [32]. We could have taken the cross product the other way round and would therefore have obtained the opposite sign. Of course we want to choose the same direction for the normal vector $\frac{\partial\mathbf{\Phi}}{\partial u}\times\frac{\partial\mathbf{\Phi}}{\partial v}$ all along the surface but this still leaves us with two possible choices leading to two different signs for the surface integral. Whenever the surface has an obvious *inside* and *outside* then we always want to choose $d\mathbf{S}$ to point in the direction of the outside of the surface, this we call the *outward normal*. However, not every surface has an obvious outside and inside and in this case we just note that the definition of the surface integral [32] is ambiguous in the sign factor and in order to have a well-defined surface integral we need to give explicitly the direction of $d\mathbf{S}$.

Example 3.13 *The function* $\mathbf{\Phi}(r,\phi)=(r\cos\phi,r\sin\phi,r)$ *describes the surface of a cone* $\mathcal{S}_1$ *for* $0\leq r\leq 1$, $0\leq\phi\leq 2\pi$. *Find the surface element* $d\mathbf{S}$ *and calculate the flow of the vector field* $\mathbf{F}=(0,0,1)$ *through the surface of the cone. Compare this with the flow through the surface* $\mathcal{S}_2$ *described by* $x^2+y^2\leq 1$ *and* $z=1$ *with the normal pointing in negative* z *direction.*

Using the right hand rule it is clear that the outward surface element is given by $d\mathbf{S}=\frac{\partial\mathbf{\Phi}}{\partial\phi}\times\frac{\partial\mathbf{\Phi}}{\partial r}d(r,\phi)=(-r\sin\phi,r\cos\phi,0)^T\times(\cos\phi,\sin\phi,1)^T d(r,\phi)=(r\cos\phi,r\sin\phi,-r)^T d(r,\phi)$. *Therefore* $\int_{\mathcal{S}_1}\mathbf{F}.d\mathbf{S}=\int_0^1\int_0^{2\pi}(-r)d\phi dr=-\pi$.

The surface $\mathcal{S}_2$ is flat and has obviously the surface element $d\mathbf{S} = (0,0,-1)^T d(x,y)$. We therefore find $\int_{\mathcal{S}_2} \mathbf{F}.d\mathbf{S} = \int_{x^2+y^2\leq 1}(-1)d(x,y) = -\pi$ which is the same value as for surface $\mathcal{S}_1$ (why?).

Let us repeat this for the vector field $\mathbf{G} = (0,0,z)^T$: $\int_{\mathcal{S}_1} \mathbf{G}.d\mathbf{S} = \int_0^1 \int_0^{2\pi}(-r^2)d\phi dr = -\frac{2}{3}\pi$ whilst $\int_{\mathcal{S}_2} \mathbf{G}.d\mathbf{S}$ remains $\int_{x^2+y^2\leq 1}(-1)d(x,y) = -\pi$. We observe that in this case the integrals are different. We will understand later when we discuss the Divergence Theorem why these integrals match in the first case and why they don't match in the second case.

Now we know how to integrate scalar products of surface elements and vector fields over a surface. Just like for integrals along paths we are also interested in integrating just the magnitude of $d\mathbf{S}$. If we go through exactly the same methodology as before but look at sums $\sum_{ij} f(\xi_i,\eta_j)\|\mathbf{S}_{ij}\|$ instead we obtain an integral which for $f = 1$ obviously defines the area of the surface $\mathcal{S}$. Following exactly the same procedure as above we find:

34 Surface area integrals

The function $\mathbf{\Phi} : \mathcal{P} \to \mathbb{R}^3$ is a parametrisation of the surface $\mathcal{S}$ with parameter space $\mathcal{P} \subset \mathbb{R}^2$ and $f(x,y,z)$ is a real valued function. The surface area integral of f over $\mathcal{S}$ is

$$\int_{\mathcal{S}} f(\mathbf{x})dS \;=\; \lim_{\|\mathcal{Z}_u\|,\|\mathcal{Z}_v\|\to 0} \sum_{i,j} f(\mathbf{\Phi}(\xi_i,\eta_j)).\|\mathbf{S}_{ij}\|\,, \tag{98}$$

where ξ_i and η_j are in $[u_{i-1},u_i]$ or $[v_{j-1},v_j]$ respectively. The area of the surface $\mathcal{S}$ is defined as

$$\|\mathcal{S}\| \;:=\; \int_{\mathcal{S}} dS\,. \tag{99}$$

Completely analogously to our earlier considerations we find that this integral can be converted into a Riemann integral where $\frac{\partial\mathbf{\Phi}(u,v)}{\partial u}\times\frac{\partial\mathbf{\Phi}(u,v)}{\partial v}$ will now appear as $\left\|\frac{\partial\mathbf{\Phi}(u,v)}{\partial u}\times\frac{\partial\mathbf{\Phi}(u,v)}{\partial v}\right\|$.

35 Surface area integrals over differentiable surfaces

The surface $\mathcal{S}$ is parametrised by the function $\mathbf{\Phi}$ over the parameter space P then the surface area integral of the real valued function $f(x,y,z)$ over $\mathcal{S}$ is

$$\int_{\mathcal{S}} f(\mathbf{x})dS \;=\; \int_P f(\mathbf{\Phi}(u,v))\left\|\frac{\partial\mathbf{\Phi}(u,v)}{\partial u}\times\frac{\partial\mathbf{\Phi}(u,v)}{\partial v}\right\| d(u,v)\,. \tag{100}$$

In particular, the area of the surface $\mathcal{S}$ is given by

$$\|\mathcal{S}\| \;=\; \int_P \left\|\frac{\partial\mathbf{\Phi}(u,v)}{\partial u}\times\frac{\partial\mathbf{\Phi}(u,v)}{\partial v}\right\| d(u,v)\,. \tag{101}$$

For simplicity we use the notation

$$dS \;=\; \left\|\frac{\partial\mathbf{\Phi}(u,v)}{\partial u}\times\frac{\partial\mathbf{\Phi}(u,v)}{\partial v}\right\| d(u,v)\,. \tag{102}$$

It seems that the main difficulty of solving surface integrals is finding $d\mathbf{S}$ for a surface $\mathcal{S}$. Once we have found $d\mathbf{S}$ then the integration is reduced to solving two successive Riemann integrals. In Eq. 94 we have already given $d\mathbf{S}$ for the (open) upper hemisphere of radius R. The same result obviously generalizes to a whole sphere with radius R.

36 Surface element of a sphere
The surface element of the sphere $x^2+y^2+z^2=r^2$ parametrised by $\phi\in[0,2\pi]$ and $\theta\in[0,\pi]$ is given by

$$d\mathbf{S} \;=\; r^2\sin\theta\begin{pmatrix}\cos\phi\,\sin\theta\\ \sin\phi\,\sin\theta\\ \cos\theta\end{pmatrix}d\theta\,d\phi = r^2\sin\theta\mathbf{e}_r d\theta\,d\phi = r\sin\theta\mathbf{x}\,d\theta\,d\phi\,, \tag{103}$$

and the surface area element is

$$dS \;=\; r^2\sin\theta\,d\theta\,d\phi\,. \tag{104}$$

Example 3.14 *Find the surface area element dS of the paraboloid $z=x^2+y^2$, $0\le z\le 1$ and find its area.*

We parametrise the paraboloid using the function $\mathbf{\Phi}(r,\phi)=(r\cos\phi,r\sin\phi,r^2)^T$ with $0\le\phi\le 2\pi$ and $r\ge 0$. The outward surface element is given by $d\mathbf{S}=\frac{\partial\mathbf{\Phi}}{\partial\phi}\times\frac{\partial\mathbf{\Phi}}{\partial r}d(r,\phi)=(-r\sin\phi,r\cos\phi,0)^T\times(\cos\phi,\sin\phi,2r)^T=(2r^2\cos\phi,2r^2\sin\phi,-r)^T$. Taking the magnitude of $d\mathbf{S}$ we find for the surface area element $dS=r\sqrt{1+4r^2}d(r,\phi)$. Therefore the area is $\int_0^1\int_0^{2\pi}r\sqrt{1+4r^2}d\phi dr=\frac{\pi}{6}(5^{\frac{3}{2}}-1)$.

For completeness we will define surface integrals over real valued functions as well as surface area integrals over vector valued functions.

37 Surface integrals over real valued functions

The integral of the real valued function $f(x,y,z)$ over the surface $\mathcal{S}\subset\mathbb{R}^3$ parametrised by the function $\mathbf{\Phi}$ with parameter space $\mathcal{P}$ is defined by

$$\int_{\mathcal{S}} f(x,y,z)d\mathbf{S} \;:=\; \int_{\mathcal{P}} f(\mathbf{\Phi}(u,v))\left(\frac{\partial\boldsymbol{\phi}(u,v)}{\partial u}\times\frac{\partial\boldsymbol{\phi}(u,v)}{\partial v}\right)d(u,v)\,. \tag{105}$$

38 Surface area integrals over vector valued functions
The integral of the vector valued function $\mathbf{F}(x,y,z)$ over the surface $\mathcal{S}$ is defined by.

$$\int_{\mathcal{S}}\mathbf{F}(\mathbf{x})dS \;:=\; \begin{pmatrix}\int_{\mathcal{S}}F_1(\mathbf{\Phi}(u,v))dS\\ \int_{\mathcal{S}}F_2(\mathbf{\Phi}(u,v))dS\\ \int_{\mathcal{S}}F_3(\mathbf{\Phi}(u,v))dS\end{pmatrix}. \tag{106}$$

Examples

Example 3.15 *We will find the normal to the surface S of revolution given parametrically by*

$$\mathbf{\Phi}(\phi,z)=(R(z)\cos\phi,\;R(z)\sin\phi,z)\;.$$

The two variables z and ϕ parametrise the surface. Note that in cylindrical polar co-ordinates (ρ,ϕ,z) the surface is given by

$$\rho=(x^2+y^2)^{\frac{1}{2}}=R(z)\,.$$

We have

$$\begin{aligned}\frac{\partial \mathbf{\Phi}}{\partial \phi} &= (-R\sin\phi,\ R\cos\phi,\ 0)\,,\\ \frac{\partial \mathbf{\Phi}}{\partial z} &= (R'(z)\cos\phi,\ R'(z)\sin\phi,\ 1)\,.\end{aligned}$$

These two vectors span the tangent plane to S at the point with parameters ϕ, z, as can be seen from the Taylor series, taken to linear order:

$$\mathbf{\Phi}(\phi+\delta\phi, z+\delta z) \approx \mathbf{\Phi}(\phi, z) + \frac{\partial \mathbf{\Phi}}{\partial \phi}\delta\phi + \frac{\partial \mathbf{\Phi}}{\partial z}\delta z.$$

The normal direction to S is therefore given by

$$\frac{\partial \mathbf{\Phi}}{\partial \phi} \times \frac{\partial \mathbf{\Phi}}{\partial z} = (R\cos\phi, R\sin\phi, -RR')$$

and the unit normal is

$$\pm\frac{(\cos\phi, \sin\phi, -R')}{(1+R'^2)^{\frac{1}{2}}}\,.$$

Example 3.16 *We will work out the vector and scalar elements of area for three different surfaces.*

(i) Let $\mathbf{\Phi}(x,y) = (x, y, z(x,y))$. Then, using subscripts to denote partial derivatives of the function z,

$$\frac{\partial \mathbf{\Phi}}{\partial x} = (1, 0, z_x)\,, \qquad \frac{\partial \mathbf{\Phi}}{\partial y} = (0, 1, z_y)\,.$$

Thus

$$d\mathbf{S} = (-z_x, -z_y, 1)\,dx\,dy\,,$$

and

$$dS = (1 + z_x^2 + z_y^2)^{\frac{1}{2}} dx\,dy\ \geq dx\,dy\,.$$

(ii) Let $\mathbf{\Phi}(\theta,\phi) = (a\sin\theta\cos\phi,\ a\sin\theta\sin\phi,\ a\cos\theta)$, which represents the surface of sphere of radius a. Then

$$\frac{\partial \mathbf{\Phi}}{\partial \theta} = (a\cos\theta\cos\phi,\ a\cos\theta\sin\phi,\ -a\sin\theta)\,, \qquad \frac{\partial \mathbf{\Phi}}{\partial \phi} = (-a\sin\theta\sin\phi,\ a\sin\theta\cos\phi,\ 0)\,,$$

so that

$$d\mathbf{S} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,a^2\sin\theta\,d\theta d\phi = \hat{\mathbf{x}}\,a^2\sin\theta\,d\theta\,d\phi\,,$$

where $\hat{\mathbf{x}}$ is the unit vector in the radial direction, and $dS = a^2\sin\theta\,d\theta\,d\phi$.

(iii) Consider the surface of revolution specified by (compare example 2.12)

$$\mathbf{\Phi}(\phi, z) = (R(z)\cos\phi, R(z)\sin\phi, z)\,.$$

We have

$$\frac{\partial \mathbf{\Phi}}{\partial \phi} = (-R\sin\phi, R\cos\phi, 0)\,, \qquad \frac{\partial \mathbf{\Phi}}{\partial z} = (R'(z)\cos\phi, R'(z)\sin\phi, 1)\,.$$

Hence

$$\begin{aligned} d\mathbf{S} &= (R\cos\phi, R\sin\phi, -RR')\,d\phi\,dz\\ &= (\cos\phi, \sin\phi, -R')R\,d\phi\,dz\,,\\ dS &= (1+R'^2)^{\frac{1}{2}}R\,d\phi\,dz\,.\end{aligned}$$

Example 3.17 *We will calculate*

$$I = \int_{\mathcal{S}} xy\, dS\,,$$

where $\mathcal{S}$ *is the portion of the surface* $x + y + z = a$ *in* $x \geq 0$, $y \geq 0$, $z \geq 0$.

Parameterising the surface with x *and* y, *we have* $\mathbf{\Phi} = (x, y, a - x - y)$, *so*

$$d\mathbf{S} = (1, 0, -1) \times (0, 1, -1)\; dx\; dy = (1, 1, 1)\; dx\; dy$$

and

$$dS = \sqrt{3}\, dx\, dy \equiv \sec\alpha\, dx\, dy\ ,$$

where α *is the angle between the plane and the horizontal. This has an obvious geometric interpretation:* dx, dy *is the projection of* dS *into the* x-y *plane. Hence*

$$\begin{aligned} I &= \sqrt{3}\int_0^a \int_0^{a-x} xy\; dydx \\ &= \frac{\sqrt{3}}{2}\int_0^a x(a-x)^2 dx \\ &= \frac{a^4}{8\sqrt{3}}\,. \end{aligned}$$

Example 3.18 *We will calculate*

$$\mathbf{I} = \int_{\mathcal{S}} \mathbf{x}\; dS\ ,$$

where $\mathcal{S}$ *is the hemisphere* $x^2 + y^2 + z^2 = a^2$, $z \geq 0$. *This is obviously going to be easiest in spherical polar coordinates, so we write*

$$\mathbf{\Phi} = (a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta), \qquad dS = a^2\sin\theta\, d\theta\, d\phi\ .$$

Of course, we must use Cartesian axes for the integrand: we must not be tempted to set $\mathbf{\Phi} = (a, 0, 0)$ *in spherical polar axes, despite the simplification that would result. Thus*

$$\mathbf{I} = \int_0^{2\pi}\int_0^{\pi/2} (a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta)a^2\sin\theta\; d\theta\, d\phi.$$

We integrate each component separately. It is a good idea to do the ϕ *integral first, since some of the terms integrate to zero.*

$$\mathbf{I} = \int_0^{\pi/2} (0, 0, 2\pi a\cos\theta)a^2\sin\theta\; d\theta = (0, 0, \pi a^3).$$

Table of integrals

The following table summarises all integrals we have studied in the previous chapter and how to convert them into Riemann Integrals. $\boldsymbol{\gamma}(\mathbf{x})$ denotes a Jordan path defined on the interval $[a,b]$ and $\mathbf{F}(t)$ denotes the vector valued function $\mathbf{F}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}$. The surface $\mathcal{S}$ is parametrised by the function $\boldsymbol{\Phi}(u,v)$ with parameter space $\mathcal{P}$.

Integral	notation	corresponding integral(s)
Integral along $\boldsymbol{\gamma}(t)$	$\int_{\boldsymbol{\gamma}} f(\mathbf{x})ds$	$= \int_a^b f(\boldsymbol{\gamma}(t))\|\boldsymbol{\gamma}'(t)\|dt$[1cm]
Length of $\boldsymbol{\gamma}(t)$	$\int_{\boldsymbol{\gamma}} ds$	$= \int_a^b \|\boldsymbol{\gamma}'(t)\|dt$
Integral along $\boldsymbol{\gamma}(t)$ of $\mathbf{F}$	$\int_{\boldsymbol{\gamma}} \mathbf{F}(\mathbf{x})ds$	$= \begin{pmatrix} \int_a^b f_1(\boldsymbol{\gamma}(t))\|\boldsymbol{\gamma}'(t)\|dt \\ \vdots \\ \int_a^b f_m(\boldsymbol{\gamma}(t))\|\boldsymbol{\gamma}'(t)\|dt \end{pmatrix}$
Line integrals	$\int_{\boldsymbol{\gamma}} \mathbf{F}(\mathbf{x}).d\mathbf{x}$	$= \int_a^b \mathbf{F}(\boldsymbol{\gamma}(t)).\boldsymbol{\gamma}'(t)dt$
Flat surface integral	$\int_{\mathcal{S}} f(x,y)dS$	$= \int_a^b \int_{c(x)}^{d(x)} f(x,y)dydx$
Area of a flat surface	$\int_{\mathcal{S}} dS$	$= \int_a^b \int_{c(x)}^{d(x)} dydx$
Volume integral	$\int_{\mathcal{V}} f(x,y,z)dV$	$= \int_a^b \int_{c(x)}^{d(x)} \int_{g(x,y)}^{h(x,y)} f(x,y,z)dzdydx$
Volume	$\int_{\mathcal{V}} dV$	$= \int_a^b \int_{c(x)}^{d(x)} \int_{g(x,y)}^{h(x,y)} dzdydx$
Flat surface integral of $\mathbf{F}$	$\int_{\mathcal{S}} \mathbf{F}(x,y)dS$	$= \begin{pmatrix} \int_a^b \int_{c(x)}^{d(x)} f_1(x,y)dydx \\ \vdots \\ \int_a^b \int_{c(x)}^{d(x)} f_m(x,y)dydx \end{pmatrix}$
Volume integral of $\mathbf{F}$	$\int_{\mathcal{V}} \mathbf{F}(x,y,z)dV$	$= \begin{pmatrix} \int_a^b \int_{c(x)}^{d(x)} \int_{g(x,y)}^{h(x,y)} f_1(x,y,z)dzdydx \\ \vdots \\ \int_a^b \int_{c(x)}^{d(x)} \int_{g(x,y)}^{h(x,y)} f_m(x,y,z)dzdydx \end{pmatrix}$
Surface integral over $\mathcal{S}$	$\int_{\mathcal{S}} \mathbf{F}(\mathbf{x}).d\mathbf{S}$	$= \int_{\mathcal{P}} \mathbf{F}(\boldsymbol{\Phi}(u,v)).\left(\frac{\partial \boldsymbol{\Phi}}{\partial u} \times \frac{\partial \boldsymbol{\Phi}}{\partial v}\right) d(u,v)$
Surface area integral	$\int_{\mathcal{S}} f(\mathbf{x})dS$	$= \int_{\mathcal{P}} f(\boldsymbol{\Phi}(u,v))\left\|\frac{\partial \boldsymbol{\Phi}}{\partial u} \times \frac{\partial \boldsymbol{\Phi}}{\partial v}\right\| d(u,v)$
Area of a surface	$\int_{\mathcal{S}} dS$	$= \int_{\mathcal{P}} \left\|\frac{\partial \boldsymbol{\Phi}}{\partial u} \times \frac{\partial \boldsymbol{\Phi}}{\partial v}\right\| d(u,v)$
Surface area integral of $\mathbf{F}$	$\int_{\mathcal{S}} \mathbf{F}(\mathbf{x})dS$	$= \begin{pmatrix} \int_{\mathcal{P}} f_1(\boldsymbol{\Phi}(u,v))\left\|\frac{\partial \boldsymbol{\Phi}}{\partial u} \times \frac{\partial \boldsymbol{\Phi}}{\partial v}\right\| d(u,v) \\ \vdots \\ \int_{\mathcal{P}} f_m(\boldsymbol{\Phi}(u,v))\left\|\frac{\partial \boldsymbol{\Phi}}{\partial u} \times \frac{\partial \boldsymbol{\Phi}}{\partial v}\right\| d(u,v) \end{pmatrix}$

IV Vector Operators

In Part IA Differential Equations you came across the *gradient* $\boldsymbol{\nabla} f(\mathbf{x})$ of a real valued function $f(\mathbf{x})$ which is simply the vector of the various partial derivatives, for example in three dimensions:

$$\boldsymbol{\nabla} f(x,y,z) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} \tag{107}$$

We could just symbolically define

$$\boldsymbol{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \tag{108}$$

and consider the gradient as the formal product $\boldsymbol{\nabla} f(\mathbf{x})$. $\boldsymbol{\nabla}$ is therefore a linear function which maps a (real valued) function to a vector valued function. This is what we call a *vector operator* or since it includes derivatives we may also call it a *differential vector operator*. In this chapter we to study vector operators which are constructed using $\boldsymbol{\nabla}$ and analyse their properties. We shall see that these vector operators are very important for the integration methods of the previous chapters and sometimes help to simplify the integrals significantly. We will start analysing the properties of the simplest of the vector operators, the gradient $\boldsymbol{\nabla} f$ itself.

8 Directional derivatives

Partial derivatives give the gradient of a function parallel to the coordinate axes. But what if we want to find the gradient in a direction which is not parallel to a coordinate axis? Consider the graph of a real valued function $f(\mathbf{x})$ and assume the normalised vector $\mathbf{n}$ is the direction in which we want to walk (in the domain of f) starting at some point on the graph. What is the slope we will experience? This is obviously given by the *directional derivative* of $f(\mathbf{x})$ in direction $\mathbf{n}$:

39 **Directional derivative**
The directional derivative of the real valued function $f(\mathbf{x})$ is given by the limit

$$\frac{\partial f}{\partial \mathbf{n}} := \lim_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{n}) - f(\mathbf{x})}{h}$$

Taylor's theorem immediately implies:

40 **Directional derivatives theorem**

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{n}} = \boldsymbol{\nabla} f(\mathbf{x}).\mathbf{n}\ . \tag{109}$$

The directional derivatives theorem [40] together with the *Cauchy-Schwarz inequality* allows us to find the direction of the steepest increase as well as the steepest decrease. $|\frac{\partial f}{\partial \mathbf{n}}| =$

$|\boldsymbol{\nabla} f(\mathbf{x}).\mathbf{n}| \leq \|\boldsymbol{\nabla} f(\mathbf{x})\|\|\mathbf{n}\| = \|\boldsymbol{\nabla} f(\mathbf{x})\|$ and $|\frac{\partial f}{\partial \mathbf{n}}| = \|\boldsymbol{\nabla} f(\mathbf{x})\|$ exactly in the case that $\|\boldsymbol{\nabla} f(\mathbf{x})\|$ is proportional to $\mathbf{n}$ (Cauchy-Schwarz inequality). Therefore the direction $\mathbf{n} = \frac{\boldsymbol{\nabla} f(\mathbf{x})}{\|\boldsymbol{\nabla} f(\mathbf{x})\|}$ is the direction of the steepest increase and $\mathbf{n} = -\frac{\boldsymbol{\nabla} f(\mathbf{x})}{\|\boldsymbol{\nabla} f(\mathbf{x})\|}$ the direction of the steepest decrease since these are the directions for which $|\frac{\partial f}{\partial \mathbf{n}}|$ becomes largest.

41 **Steepest increase and steepest decrease**
For a real valued function $f(\mathbf{x})$ we find that $\boldsymbol{\nabla} f(\mathbf{x})$ points in the direction of the steepest increase and $-\boldsymbol{\nabla} f(\mathbf{x})$ points in the direction of the steepest decrease.

For [40] we of course have to assume that the function $f(\mathbf{x})$ is sufficiently smooth (which here means it is differentiable). It is remarkable that such functions are so smooth that locally the direction of the steepest decrease is always opposite to the direction of the steepest increase. If the function is not differentiable this obviously does not have to be the case. Also, theorem [40] says that for a differentiable function it is enough to know the slopes parallel to the coordinated axes since the slope in any other direction is just a linear combination of them.

In case $\boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{0}$ then [40] says that the function is flat in all directions at the point $\mathbf{x}$. Therefore the point is a candidate for a local maximum or a local minimum. We will continue this discussion of local extrema as discussed in Part IA Differential Equations.

9 Conservative fields

We will now return to line integrals and study the significance of the gradient $\boldsymbol{\nabla} f(\mathbf{c})$ within the theory of line integrals. If we choose for the example 3.1 a different path, e.g. $\boldsymbol{\gamma}_2(t) := (1, 0, t)^T$, $t \in [0, 2\pi]$ then the starting point and the end point are obviously identical but we will be moving on a very different path. However, in this particular example, if we perform the corresponding line integral we find exactly the same value! We can check other (well-behaved) paths for this vector field and we would find that for this vector field all paths lead to the same line integral provided that the starting points and the end points coincide. There are obvious examples of vector fields where this is not the case: take for example the vector field $\mathbf{F}(\mathbf{x}) = (y, 0, 0)^T$. Going from $(0, 0, 0)$ to $(1, 1, 1,)$ along (t, t, t) leads to a line integral of $\frac{1}{2}$ whilst following the path along (t, t^2, t) leads to $\frac{1}{3}$. Therefore, some vector fields seem to have the property that line integrals are independent of the actual paths and the line integrals only depends on the starting points and on the end points whilst for other vector fields the line integrals depend on the actual paths taken.

If we interpret the line integral to be the energy gained or required in order to move a point mass through a force field then the fact that the line integral is independent of the path simply means that the energy is independent from the path taken. Therefore energy is conserved in this system. Motivated by this interpretation we call vector fields for which line integrals depend only on the starting points and on the end points but not on the actual paths taken *conservative fields*.

For many of the following considerations it will be essential that any two points in the domain G of the vector field $\mathbf{F}$ can be connected by a continuous path otherwise we will not be able to integrate along a path connecting the two points. We will call such a set *connected*.

42 **Conservative fields**
A vector field $\mathbf{F}$ defined on the connected domain G is called a conservative field if the line

integrals $\int_\gamma \mathbf{F}d\mathbf{x}$ *only depend on the starting points and end points of* $\boldsymbol{\gamma}$ *and are otherwise path independent. We can therefore denote* $\int_\gamma \mathbf{F}.d\mathbf{x}$ *simply by* $\int_{\mathbf{x}_s}^{\mathbf{x}_e} \mathbf{F}.d\mathbf{x}$.

The fact that the line integral of a conservative field only depends on the starting point and the end point of the path reminds us of the *Fundamental Theorem of Calculus* for Riemann integrals: $\int_a^b f(x)dx = F(b) - F(a)$ as long as $F'(x) = f(x)$. Could we maybe hope that for conservative fields there exists a function $\Phi(\mathbf{x})$ such that $\int_\gamma \mathbf{F}d\mathbf{x} = \Phi(\mathbf{x_e}) - \Phi(\mathbf{x_s})$, where $\mathbf{x_e} = \boldsymbol{\gamma}(b)$ is the end point and $\mathbf{x_s} = \boldsymbol{\gamma}(a)$ is the starting point? In this case, would $\mathbf{F}$ also be the derivative of $\Phi(\mathbf{x})$ just like in the one dimensional Riemann case? Here the derivative would obviously have to be the gradient $\boldsymbol{\nabla}\Phi(\mathbf{x})$. In order to investigate this idea let us first define vector fields which are gradients of real valued functions $\Phi(\mathbf{x})$:

43 Gradient fields[a]

A vector field $\mathbf{F}$ *is called a gradient field on the set* G *if there exists a real valued function* $\Phi(\mathbf{x})$ *(called a scalar field) such that* $\boldsymbol{\nabla}\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{x})\ \forall \mathbf{x} \in G$. $\Phi(\mathbf{x})$ *is then called a potential*[b] *for the vector field* $\mathbf{F}(\mathbf{x})$.

Example 4.1 *The scalar field* $\Phi(\mathbf{x}) = -\frac{1}{r}$ *is a potential for the conservative field given in* FIG. XV. *Therefore this vector field is a gradient field on the set* $\mathbb{R}^3 \setminus \{\mathbf{0}\}$ *as can easily be shown by verifying* $\boldsymbol{\nabla}\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. *A two dimensional example is give by the vector field*

$$\mathbf{F} := \begin{pmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{pmatrix}. \tag{110}$$

It is easy to see that $\Phi(x,y) = \tan^{-1}\frac{y}{x}$ *is a potential for this* $\mathbf{F}$ *on the set* $G = \mathbb{R}^2 \setminus \{(x,y) : x = 0\}$. *Therefore* $\mathbf{F}$ *is a gradient field on the set* G. *Note that* G *is obviously not connected since it consists of two disconnected parts: the left half plane and the right half plane excluding the y-axis. However,* $\mathbf{F}$ *is of course also a gradient field on the smaller set* $\{(x,y) : x > 0\}$ *which is connected.* $\hat{\Phi}(x,y) = -\tan^{-1}\frac{x}{y}$ *is another potential for* $\mathbf{F}$ *this time defined on the set* $\mathbb{R}^2 \setminus \{(x,y) : y = 0\}$. *Again this set is not connected. In fact, we shall see later that* $\mathbf{F}$ *is not a gradient field on the whole of* $\mathbb{R}^2 \setminus \{\mathbf{0}\}$*!*

Using the chain rule we can give an amazingly simple proof to show that gradient fields are conservative provided, of course, that the underlying set is connected:

44 Line integrals over gradient fields

Assume that $\mathbf{F} : G \subset \mathbb{R}^n \to \mathbb{R}^n$, G *connected, is a gradient field with potential* $\Phi(\mathbf{x})$ *then* $\mathbf{F}$ *is also conservative on* G *and the line integral along a path* $\boldsymbol{\gamma}(t) \subset G$ *is given by*

$$\int_\gamma \mathbf{F}.d\mathbf{x} \quad = \quad \int_\gamma \boldsymbol{\nabla}\Phi.d\mathbf{x} = \Phi(\mathbf{x}_e) - \Phi(\mathbf{x}_s)\ , \tag{111}$$

where $\mathbf{x_e} = \boldsymbol{\gamma}(b)$ *is the end point and* $\mathbf{x_s} = \boldsymbol{\gamma}(a)$ *is the starting point of the path* $\boldsymbol{\gamma}$.

[a] Note that for the definition of a gradient field we do not need to assume that the domain G is connected but when we later ask the question whether the field is conservative we need to require that G is connected since talking about conservative fields only makes sense if we are able to connect any two points with continuous paths.

[b] Note that physicists tend to call $-\Phi(\mathbf{x})$ with $\boldsymbol{\nabla}\Phi(\mathbf{x}) = -\mathbf{F}$ a potential.

The Riemann integral $\int_{\Phi(\mathbf{x}_s)}^{\Phi(\mathbf{x}_e)} d\Phi$ is given by $\Phi(\mathbf{x}_e) - \Phi(\mathbf{x}_s)$ we will sometimes write EQ. 111 simply as $d\Phi = \boldsymbol{\nabla}\Phi.d\mathbf{x}$.

We are now in a position to explain why the vector field in EQ. 110 cannot be a gradient field on the whole of the connected set $\mathbb{R}^2 \setminus \{\mathbf{0}\}$: if it were conservative then the line integral along the unit circle $\boldsymbol{\gamma}(t) = (\cos t, \sin t)^T$ would have to be 0. If we were to go one full circle round then the integral would have to equal the integral over doing nothing at all which is 0. But $\displaystyle\int_\gamma \mathbf{F}.d\mathbf{x} = \int_0^{2\pi} \begin{pmatrix} \frac{-\sin t}{r} \\ \frac{\cos t}{r} \end{pmatrix}.\begin{pmatrix} -r\sin t \\ r\cos t \end{pmatrix} dt = 2\pi \neq 0.$

You can always split a closed path $\boldsymbol{\gamma}$ into the sum of two paths $\boldsymbol{\gamma}_1 \oplus \boldsymbol{\gamma}_2$ and therefore create by inverting the direction along one of the two paths $\boldsymbol{\gamma}_2^-$ two paths in G connecting the same points along different paths. Therefore it can easily be seen that a field is conservative if and only if the integral over all closed paths in G are 0. If $\boldsymbol{\gamma}$ is closed then the line integral is denoted by $\oint_{\boldsymbol{\gamma}} \mathbf{F}.d\mathbf{x}$.

45 **Closed path integrals**
A vector field $\mathbf{F} : G \subset \mathbb{R}^n \to \mathbb{R}^n$, G connected, is conservative if and only if the closed line integrals $\oint_{\boldsymbol{\gamma}} \mathbf{F}.d\mathbf{x} = 0$ for all closed (rectifiable) paths $\boldsymbol{\gamma}$ in G.

It is now very satisfying that the converse of [44] is also true: not only are gradient fields on connected sets conservative but conservative fields are also gradient fields! This will be rigorously proven in future Analysis courses.

46 **Path independent line integrals**
Assume that the vector field $\mathbf{F} : G \subset \mathbb{R}^n \to \mathbb{R}^n$, G connected, is conservative, then $\mathbf{F}(\mathbf{x})$ is also a gradient field and

$$\Phi(\mathbf{x}) \quad := \quad \int_{\mathbf{x_s}}^{\mathbf{x}} \mathbf{F}.d\mathbf{x}\,, \tag{112}$$

is a potential for $\mathbf{F}$: $\boldsymbol{\nabla}\Phi(\mathbf{x}) = \mathbf{F}(\mathbf{x})$.

10 Integrability condition for vector fields

We are now left with the question how do we actually recognise whether a vector field is conservative? We surely cannot try to evaluate line integrals along millions of different paths in order to convince ourselves of the path independence neither is it practical to try to guess a potential. We will therefore hope to find some simple *integrability condition* which allows us to imply that a field is conservative from simple properties of $\mathbf{F}$ without actually having to compute a potential for $\mathbf{F}$.

Let us assume that a given vector field $\mathbf{F}$ is a gradient field on the set G and $\Phi(\mathbf{x})$ is a potential on G: $\frac{\partial\Phi(\mathbf{x})}{\partial x_i} = F_i$. Schwarz's theorem [1] then tells us that necessarily $\frac{\partial F_i(\mathbf{x})}{\partial x_j} = \frac{\partial^2\Phi(\mathbf{x})}{\partial x_j \partial x_i} = \frac{\partial^2\Phi(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial F_j(\mathbf{x})}{\partial x_i}$ (provided of course $\mathbf{F}$ is differentiable). We therefore expect that vector fields would have to satisfy the integrability condition $\frac{\partial F_i(\mathbf{x})}{\partial x_j} = \frac{\partial F_j(\mathbf{x})}{\partial x_i}$ $\forall \mathbf{x} \in G$ and $\forall i, j$ in order to be considered as gradient fields. Unfortunately this integrability condition is not sufficient as the example in EQ. 110 demonstrates: $\frac{\partial}{\partial x}\frac{x}{x^2+y^2} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial}{\partial y}\frac{-y}{x^2+y^2}$ on the whole of $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ but $\mathbf{F}$ is not a gradient field on the whole of $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ as we have shown earlier.

However, if we add just one very simple further condition - the condition that the set G should not only be connected but *simply connected* - then we obtain a necessary and sufficient integrability condition! A set is called *simply connected* if it is connected and in addition any closed curve can be continuously shrunk to a point without leaving the set. The dotted line in FIG. xxiii shows a closed curve which cannot be continuously shrunk to a point without leaving the set.

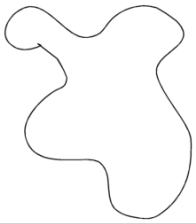


FIG. xxii A simply connected set.

FIG. xxiii A connected set which is not simply connected.

Without proof we give the following theorem:

47 **Integrability condition for vector fields**
The vector field $\mathbf{F} :\subset G\mathbb{R}^n \to \mathbb{R}^n$, *$G$ simply connected, is a gradient field and therefore also a conservative field if and only if the integrability condition*

$$\frac{\partial F_i(\mathbf{x})}{\partial x_j} = \frac{\partial F_j(\mathbf{x})}{\partial x_i} \quad , \forall \mathbf{x} \in G \text{ and } \forall i, j \, , \tag{113}$$

holds.

Note again that [47] does not help in case we want to investigate whether a vector field is a gradient field on a connected set which is not simply connected. The conditions given in EQ. 113 are then still necessary but they might not be sufficient.

In $\mathbb{R}^2$ EQ. 113 is in fact just one condition $\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1}$ whilst in $\mathbb{R}^3$ EQ. 113 results in three conditions which can very nicely be combined in a vector equation using a new vector operator which we call the *curl* of a vector field $\mathbf{F}$, denoted $\boldsymbol{\nabla} \times \mathbf{F}$.

$$\boldsymbol{\nabla} \times \mathbf{F} := \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix} . \tag{114}$$

Obviously, in $\mathbb{R}^3$ the integrability conditions EQ. 113 are satisfied exactly when $\boldsymbol{\nabla} \times \mathbf{F}(\mathbf{x}) = \mathbf{0}$ $\forall \mathbf{x} \in G$.

Example 4.2 *The vector field* $\mathbf{F} = \left(2x\sin(zy), x^2 z\cos(zy) + 2zy, x^2 y\cos(zy) + y^2\right)^T$ *is defined on the whole of* $\mathbb{R}^3$ *which is simply connected. It can easily be shown that* $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$ *and therefore* $\mathbf{F}$ *is indeed a gradient field. That means there exists a function* $\Phi(\mathbf{x})$ *such that* $\frac{\partial \Phi}{\partial x} = 2x\sin(zy)$. *We can easily integrate this and obtain* $\Phi(\mathbf{x}) = x^2\sin(zy) + h(y,z)$ *where* $h(y,z)$ *can only be a function of* y *and* z *but not of* x. *Therefore* $\frac{\partial \Phi}{\partial y} = x^2 z\cos(zy) + \frac{\partial h(y,z)}{\partial y}$. *Comparing this with the second component of the vector field* $\frac{\partial \Phi}{\partial y} = x^2 z\cos(zy) + 2zy$ *leads to* $\frac{\partial h(y,z)}{\partial y} = 2zy$ *which is indeed a relation without* x *as required. We find* $h(y,z) = zy^2 + g(z)$ *with a function* $g(z)$ *which cannot be a function of* x *and neither of* y. *Putting all this together we obtain*

$\frac{\partial \Phi}{\partial z} = x^2 y \cos(zy) + y^2$. *This has to be compared to the third component* $x^2 y \cos(zy) + y^2$ *which gives an exact match and therefore* $g(z) =$*constant. We have therefore found the most general potential for* $\mathbf{F}$ *to be* $\Phi(\mathbf{x}) = x^2 \sin(zy) + zy^2 + c$ *with* $c \in \mathbb{R}$.

11 Differentials

In this section we give an informal introduction to *differentials* $df(\mathbf{x})$. By looking at

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F}.d\mathbf{x} \quad = \quad \int_{\mathbf{a}}^{\mathbf{b}} F_1 dx_1 + \ldots + F_n dx_n = \Phi(\mathbf{b}) - \Phi(\mathbf{a}) \; , \tag{115}$$

for a gradient field $\mathbf{F}$ it makes sense to denote Eq. 115

$$= \quad \int_{\mathbf{a}}^{\mathbf{b}} d\Phi \; . \tag{116}$$

We can therefore formally write the *differential* $d\Phi$ as

$$d\Phi \quad = \quad F_1 dx_1 + \ldots + F_n dx_n = \mathbf{F}.d\mathbf{x} \; , \tag{117}$$

keeping in mind that for us this only makes sense underneath an integral.

In the same way the chain rule for a function $\Phi(x_1, \ldots, x_n)$:

$$\frac{d\Phi}{dt} \quad = \quad \frac{\partial \Phi}{\partial x_1}\frac{dx_1}{dt} + \ldots + \frac{\partial \Phi}{\partial x_n}\frac{dx_n}{dt} \; , \tag{118}$$

can be written using the following notation

$$d\Phi \quad = \quad \frac{\partial \Phi}{\partial x_1} dx_1 + \ldots + \frac{\partial \Phi}{\partial x_n} dx_n \; , \tag{119}$$

again keeping in mind that this only makes sense in a derivative.

For our purpose we simply want to use these differentials as notational objects which only make sense under integrals or in derivatives. But there is a whole area of mathematics trying to make sense out of these differentials on their own! But even used just as notation it is immediately obvious that this is a very easy way to think about gradient fields and integrals over gradient fields. For example $d(x^2 y) = 2xydx + x^2 dy$ simply comes out of the chain rule $\frac{dx^2 y}{dt} = 2xy\frac{dx}{dt} + x^2\frac{dy}{dt}$. Using $d(x^2 y)$ as the gradient of a vector field then gives $\int_{\mathbf{a}}^{\mathbf{b}} 2xydx + x^2 dy = \Phi(\mathbf{b}) - \Phi(\mathbf{a})$ with $\Phi(\mathbf{x}) = x^2 y$. We call such a differential $d\Phi$ an *exact differential*:

48 **Exact differentials**
The differential $F_1 dx_1 + \ldots + F_n dx_n$ *is called an exact differential if and only if the vector field* $\mathbf{F} = (F_1, \ldots, F_n)$ *is conservative. We then write* $d\Phi = F_1 dx_1 + \ldots + F_n dx_n$ *with* $\boldsymbol{\nabla}\Phi = \mathbf{F}$.

We can of course use the integrability conditions Eq. 113 to find out whether a differential is exact:

Example 4.3 *For the differential* $2xydx + x^2 dy$ *we find that* $\frac{\partial x^2}{\partial x} = 2x = \frac{\partial 2xy}{\partial y}$ *and therefore* $2xydx + x^2 dy$ *is exact. Trying to solve for example the differential equation* $2xy + x^2 y' = 0$ *leads to* $d(x^2 y) = 0$ *and therefore to the family of solutions* $x^2 y =$*constant.*

But what do we do in case we have a differential $F_1 dx_1 + \ldots + F_n dx_n$ which is not exact? Could we maybe find a function $\lambda(\mathbf{x})$ such that $\lambda(\mathbf{x})F_1 dx_1 + \ldots + \lambda(\mathbf{x})F_n dx_n$ becomes exact? In this case the real valued function $\lambda(\mathbf{x})$ is called an *integrating factor* of the differential.

49 **Integrating factors for differentials**
A real valued function $\lambda(\mathbf{x})$ is called an integrating factor for the differential $F_1 dx_1 + \ldots + F_n dx_n$ if the differential $\lambda(\mathbf{x})F_1 dx_1 + \ldots + \lambda(\mathbf{x})F_n dx_n$ is exact. Provided the differential and the integrating factor are both defined on a common open set G which is simply connected then the integrability can be tested using

$$\frac{\partial \lambda(\mathbf{x})F_i}{\partial x_j} = \frac{\partial \lambda(\mathbf{x})F_j}{\partial x_i} \quad \forall i, j \; . \tag{120}$$

Example 4.4 *It is obvious that the differential $2ydx + xdy$ is not exact since $\frac{\partial 2y}{\partial y} = 2 \neq 1 = \frac{\partial x}{\partial x}$. We try to find $\lambda(x, y)$ such that $\frac{\partial 2y\lambda}{\partial y} = \frac{\partial x\lambda}{\partial x}$ which leads to a (partial) differential equation $2y\frac{\partial \lambda}{\partial y} + 2\lambda = x\frac{\partial \lambda}{\partial x} + \lambda$. Since we usually only need to find one integrating factor and not all integrating factors it is always worth trying if there exists an integrating factor which is only a function of x (or only a function of y) such that $\frac{\partial \lambda}{\partial y} = 0$ (or $\frac{\partial \lambda}{\partial x} = 0$). The first assumption leads to $2\lambda = x\lambda'(x) + \lambda$ and therefore $\lambda = x$ is such a solution and is indeed not dependent on y (note that all factors of y disappeared once we had put $\frac{\partial \lambda}{\partial y} = 0$). Hence $2yxdx + x^2 dy$ is exact.*

Example 4.5 *If we are given a differential of the form $dy + (f(x)y - g(x))dx$ with some functions $f(x)$ and $g(x)$ then an integrating factor has to satisfy $\frac{\partial \lambda}{\partial x} = \frac{\partial (f(x)y - g(x))\lambda}{\partial y}$. Let us again try if we find an integrating factor $\lambda(x)$ which depends only on x but not on y. Therefore λ would have to satisfy $\frac{\partial \lambda}{\partial x} = f(x)\lambda$ and therefore $\lambda(x) = e^{\int f(x)dx}$ is an integrating factor such that $\lambda\,(dy + (f(x)y - g(x))dx) = d(\lambda(x)y - \int \lambda(x)g(x)dx)$ is exact. This is of course the standard integrating factor known to us for solving first order linear differential equations: $y' + f(x)y = g(x)$. This differential equation can then be written as $(\lambda(x)y)' = \lambda(x)g(x)$ with the $\lambda(x)$ given above.*

Example 4.6 *Let*

$$\mathbf{F}.\, d\mathbf{x} = (\sin y)\, dx + (2y + x\cos y)\, dy.$$

We will show that the differential $\mathbf{F}.\, d\mathbf{x}$ is exact and find the most general scalar field $f(\mathbf{x})$ such that $\mathbf{F}.d\mathbf{x} = df$.

We have $\partial F_1/\partial y = \partial F_2/\partial x$ $(= \cos y)$, which is a necessary and sufficient condition for the differential to be exact.

If $\mathbf{F}.d\mathbf{x} = df$, then $\boldsymbol{\nabla} f = \mathbf{F}$, i.e.

$$\frac{\partial f}{\partial x} = \sin y \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 2y + x\cos y \, .$$

Integrating the first of these equations gives

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

where Y is an arbitrary function of integration. We could now integrate the second equation and compare the two integrations to identify the arbitrary functions of integration. But it is much easier to differentiate $(*)$ with respect to x to give a first order equation for $Y(y)$:

$$\frac{\partial f}{\partial y} = x\cos y + Y'(y) \equiv 2y + x\cos y \, .$$

Hence $Y'(y) = 2y$ *and*

$$f(x, y) = y^2 + x \sin y + C,$$

where C *is a constant.*

Of course, this could fairly easily have been obtained by inspection:

$$d?? = (\sin y)\, dx + (2y + x \cos y)\, dy$$

and often this is the best way.

Note that we have just found the solution to the differential equation

$$\frac{dy}{dx} = -\frac{\sin y}{2y + x \cos y},$$

since the equation could be written in the form

$$(\sin y)\, dx + (2y + x \cos y)\, dy = 0, \qquad i.e. \qquad d(y^2 + x \sin y) = 0.$$

The solution is therefore $y^2 + x \sin y =$ *constant.*

Example 4.7 *We will show that an integrating factor can be found to make the differential*

$$\mathbf{F}.\, d\mathbf{x} \equiv y\, dx - x\, dy$$

exact. Note first that $\partial F_1/\partial y - \partial F_2/\partial x = 2 \neq 0$*, so as it stands,* $\mathbf{F}.\, d\mathbf{x}$ *is not exact. However, we will find a function* $\mu(x, y)$ *such that* $\mu(y\, dx - x\, dy)$ *is exact, i.e. such that*

$$\frac{\partial(\mu y)}{\partial y} = \frac{\partial(-\mu x)}{\partial x}.$$

Simplifying gives

$$x\frac{\partial \mu}{\partial x} + y\frac{\partial \mu}{\partial y} = -2\mu\,.$$

This looks hard, because of the partial derivatives, but in a sense they make the problem easier: there is the freedom of a whole arbitrary function of one variable in the solution. We could, for example, try a solution of the form $\mu = \mu(x)$*.*

Or we could be systematic. The above equation says that the directional derivative of μ *is determined in the direction* (x, y) *but is not determined at all in directions orthogonal to this. Let us then change to new variables* u *and* v*, where* v *is constant on the curves with tangent* (x, y) *(i.e.* v *labels the individual curves) and* u *is a parameter along the individual curves.*

We therefore choose u *such that*

$$x\frac{\partial \mu}{\partial x} + y\frac{\partial \mu}{\partial y} \equiv \frac{\partial \mu}{\partial u} \equiv \frac{\partial x}{\partial u}\frac{\partial \mu}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial \mu}{\partial y},$$

i.e. such that $\partial x/\partial u = x$ *and* $\partial y/\partial u = y$*. Integrating gives* $u = \log x + a(v)$ *and* $u = \log y + b(v)$*. Any solution will do. One possibility is* $u = \log x$ *and* $v = x/y$*.*

The most general integrating factor μ *is therefore given by*

$$\frac{\partial \mu}{\partial u} = -2\mu \Rightarrow \mu = A(v)e^{-2u} = A(x/y)x^{-2}$$

where A *is an arbitrary function. We could, for example, choose* A *such that* $\mu = x^{-2}$ *or* $\mu = y^{-2}$*.*

12 Surfaces in $\mathbb{R}^3$

The equation $\Phi(x,y,z)=0$ with a function $\Phi:\mathbb{R}^3\to\mathbb{R}$ defines a 2 dimensional surface $\mathcal{S}$ in 3 dimensional space. If Φ is continuous then the surface $\mathcal{S}$ will be a *continuous* surface and if Φ is even differentiable then $\mathcal{S}$ is a *smooth* surface.

For example, $\Phi(x,y,z)=0$ describes the unit sphere for $\Phi(x,y,z)=x^2+y^2+z^2-1$ whilst for $\Phi(x,y,z)=z-e^{-(x^2+y^2)}$ it describes the surface given in FIG. iii.

Let us apply the chain rule to $\Phi(\mathbf{x}(t))$ where $\mathbf{x}(t)=(x(t),y(t),z(t))^T$ is some path on the surface through the point $\mathbf{x}_0=(x(0),y(0),z(0))^T$. Since the path is on the surface we have $\Phi(\mathbf{x}(t))=0$ and therefore the chain rule leads to $\boldsymbol{\nabla}\Phi(\mathbf{x}_0).\mathbf{x}'(0)=0$. As the *velocity* vector $\mathbf{x}'(0)$ is tangent to the path $\mathbf{x}(t)$ it is therefore tangent to the surface at the point $\mathbf{x}_0$. The path $\mathbf{x}(t)$ was arbitrarily chosen. By choosing different paths the velocity vectors can generate the whole tangent plane at the point $\mathbf{x}_0$ and therefore $\boldsymbol{\nabla}\Phi(\mathbf{x}_0).\mathbf{x}'(0)=0$ says that $\boldsymbol{\nabla}\Phi(\mathbf{x}_0)$ is perpendicular to all these tangents. Therefore $\boldsymbol{\nabla}\Phi(\mathbf{x}_0)$ is normal to the surface at the point $\mathbf{x}_0$.

50 Normal vectors to surfaces in $\mathbb{R}^3$
The vector $\boldsymbol{\nabla}\Phi(x,y,z)$ is normal to the surface described by $\Phi(x,y,z)=0$.

Example 4.8 *We will find the vector field normal to the family of surfaces given by*

$$\Phi(x,y,z)\equiv x^2-y^2-z^2-c^2=0$$

where c is a constant. Since the vector field $\boldsymbol{\nabla}\Phi$ is normal to the surfaces $\Phi(\mathbf{x})=0$, and

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

the unit normal vector field is given by

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

Note that the normal is not defined at $r=0$. This is as expected, because the surface that passes through $\mathbf{x}=\mathbf{0}$ has $c=0$. It is the cone $x^2=y^2+z^2$, which does not have a normal at its apex.

13 Surface elements in $\mathbb{R}^3$

In SEC. 12 we showed that if a surface $\mathcal{S}$ is defined using a real valued function $\Phi(x,y,z)=0$ then $\boldsymbol{\nabla}\Phi$ is normal to the surface $\mathcal{S}$. This means that we can easily find the direction of $d\mathbf{S}$, simply by taking $\boldsymbol{\nabla}\Phi$. If we could find an equally simple way of determining $\|d\mathbf{S}\|$ we would then have a very fast way of finding $d\mathbf{S}$. If we assume that $\Phi(x,y,z)=0$ can at least locally be solved for $z=g(x,y)$ then we can describe $\mathcal{S}$ locally by $\boldsymbol{\Phi}(x,y)=(x,y,g(x,y))^T$. By taking the cross product of the two tangent vectors we find the surface element $d\mathbf{S}$

$$d\mathbf{S} \;=\; \frac{\partial\boldsymbol{\Phi}}{\partial x}\times\frac{\partial\boldsymbol{\Phi}}{\partial y}d(x,y)=\begin{pmatrix}1\\0\\\frac{\partial g}{\partial x}\end{pmatrix}\times\begin{pmatrix}0\\1\\\frac{\partial g}{\partial y}\end{pmatrix}d(x,y)=\begin{pmatrix}-\frac{\partial g(x,y)}{\partial x}\\-\frac{\partial g(x,y)}{\partial y}\\1\end{pmatrix}d(x,y)\,. \qquad (121)$$

Using the chain rule on $\Phi(x,y,z)=0$ we obtain $0=\frac{\partial\Phi}{\partial x}+\frac{\partial\Phi}{\partial z}\frac{\partial g}{\partial x}$ and similarly $0=\frac{\partial\Phi}{\partial y}+\frac{\partial\Phi}{\partial z}\frac{\partial g}{\partial y}$. Assuming that $\frac{\partial\Phi}{\partial z}\neq 0$ (which is just the condition for $\Phi(x,y,z)=0$ being locally solvable for z) we therefore obtain the following rule to find surface elements in Cartesian co-ordinates.

51 Surface elements in Cartesian co-ordinates[c]
The surface $\mathcal{S}$ is defined by the equation $\Phi(x,y,z)=0$. Provided that $\frac{\partial\Phi(x,y,z)}{\partial z}\neq 0$ we find for the surface element $d\mathbf{S}$ in Cartesian co-ordinates

$$d\mathbf{S} \;=\; \frac{1}{\frac{\partial\Phi}{\partial z}}\boldsymbol{\nabla}\Phi(x,y,z)\,d(x,y)\;. \tag{122}$$

The surface area element is given by

$$dS \;=\; \frac{\|\boldsymbol{\nabla}\Phi\|}{|\frac{\partial\Phi}{\partial z}|}\,d(x,y)\;. \tag{123}$$

Example 4.9 *Using [51] we immediately find for the paraboloid described by $z=x^2+y^2$ the surface element $d\mathbf{S}=\begin{pmatrix}2x\\2y\\-1\end{pmatrix}d(x,y)$ The negative z co-ordinate of $d\mathbf{S}$ shows that this is the outward normal as can easily be seen from* Fig. xxix.

Applying [51] to the sphere $x^2+y^2+z^2=r^2$ requires caution. For the hemisphere (or parts of the hemisphere) with $z>0$ we easily find $d\mathbf{S}=\begin{pmatrix}\frac{x}{z}\\\frac{y}{z}\\1\end{pmatrix}d(x,y)$ and the positive z co-ordinate shows that it is the outward normal. However if we want to find $d\mathbf{S}$ for the whole sphere or patches of the sphere which involve points with $z=0$ this $d\mathbf{S}$ would obviously not work. But using spherical polar co-ordinates and noting that $d(x,y)=r^2\sin\theta\;\cos\theta d(\phi,\theta)$ we can easily transform this $d\mathbf{S}$ into $r^2\sin\theta\begin{pmatrix}\cos\phi\;\sin\theta\\\sin\phi\;\sin\theta\\\cos\theta\end{pmatrix}d(\phi,\theta)$ which is now valid everywhere on the sphere. It should also be noted that if we want to use $d\mathbf{S}=\begin{pmatrix}\frac{x}{z}\\\frac{y}{z}\\1\end{pmatrix}d(x,y)$ to integrate over the whole sphere and provided the limit $z\to 0$ works we would still need to integrate over both $d\mathbf{S}=\begin{pmatrix}\frac{x}{z}\\\frac{y}{z}\\1\end{pmatrix}d(x,y)$ over the disk with radius r to integrate along the upper hemisphere and $d\mathbf{S}=\begin{pmatrix}\frac{x}{z}\\\frac{y}{z}\\-1\end{pmatrix}d(x,y)$ over the disk with radius r to integrate along the lower hemisphere in order to integrate over the whole sphere.

We now give a list of surface elements and surface area elements which can easily be proven by using [35] or [51].

[c] Note that $d\mathbf{S}$ in Eq. 122 is only defined up to a sign factor.

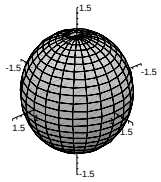


FIG. xxiv Sphere

Sphere $x^2 + y^2 + z^2 = r^2$

$$d\mathbf{S} = r^2 \sin\theta \begin{pmatrix} \cos\phi \, \sin\theta \\ \sin\phi \, \sin\theta \\ \cos\theta \end{pmatrix} d\theta \, d\phi \, ,$$

$$dS = r^2 \sin\theta d\theta \, d\phi \, ,$$

$\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$.

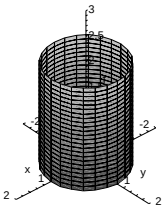


FIG. xxv Cylinder

Cylinder $x^2 + y^2 = \rho^2$, $a \leq z \leq b$

$$d\mathbf{S} = \rho \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} dz \, d\phi \, ,$$

$$dS = \rho \, dz \, d\phi \, ,$$

$\phi \in [0, 2\pi]$, $z \in [a, b]$.

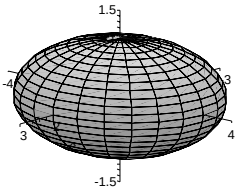


FIG. xxvi Ellipsoid

Ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

$$d\mathbf{S} = abc \sin\theta \begin{pmatrix} \frac{\cos\phi \, \sin\theta}{a} \\ \frac{\sin\phi \, \sin\theta}{b} \\ \frac{\cos\theta}{c} \end{pmatrix} d\theta \, d\phi \, .$$

$\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$.

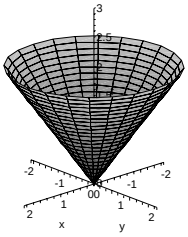


FIG. xxvii Cone

Cone $x^2 + y^2 = z^2$, $0 \leq z \leq b$

$$d\mathbf{S} = r \begin{pmatrix} \cos\phi \\ \sin\phi \\ -1 \end{pmatrix} dr \, d\phi \, ,$$

$$dS = \sqrt{2} r dr \, d\phi \, ,$$

$\phi \in [0, 2\pi]$, $r \in [0, b]$.

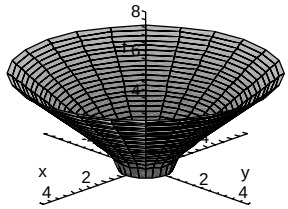


FIG. xxviii Hyperboloid

Hyperboloid $x^2 + y^2 = z^2 + 1,\ 0 < a \leq z \leq b$

$$d\mathbf{S} = \frac{r}{\sqrt{r^2-1}} \begin{pmatrix} r\cos\phi \\ r\sin\phi \\ -\sqrt{r^2-1} \end{pmatrix} dr\, d\phi\,,$$

$$dS = \frac{\sqrt{2r^2-1}}{\sqrt{r^2-1}} r dr\, d\phi\,,$$

$\phi \in [0, 2\pi],\ r \in [\sqrt{a^2+1}, \sqrt{b^2+1}]$.

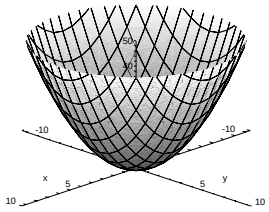


FIG. xxix Paraboloid

Paraboloid $x^2 + y^2 = z,\ 0 \leq z \leq b$

$$d\mathbf{S} = \begin{pmatrix} 2x \\ 2y \\ -1 \end{pmatrix} dx\, dy = \begin{pmatrix} 2r\cos\phi \\ 2r\sin\phi \\ -1 \end{pmatrix} r dr\, d\phi\,,$$

$$dS = \sqrt{1+4r^2}\, r dr\, d\phi\,,$$

$\phi \in [0, 2\pi],\ r \in [0, \sqrt{b}]$.

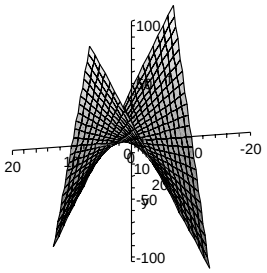


FIG. xxx Surface $z = xy$

Surface $z = xy$

$$d\mathbf{S} = \begin{pmatrix} y \\ x \\ -1 \end{pmatrix} dx\, dy\,,$$

$$dS = \sqrt{1+r^2}\, r dr\, d\phi\,.$$

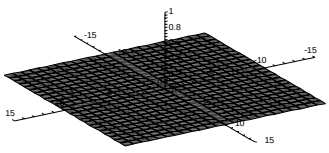


FIG. xxxi Flat surface $z = 0$

Flat surface in xy **plane**

$$d\mathbf{S} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dx\, dy\,,$$

$$dS = dx\, dy\,.$$

14 Polar co-ordinates and general orthogonal curvilinear co-ordinates

In Part IA Differential Equations (and also in example 1.7) you have studied how to use the chain rule to transform the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ into polar co-ordinates:

The two functions $x(r,\phi) = r\cos\phi$ and $y(r,\phi) = r\sin\phi$ which transform the co-ordinates (x,y) into polar co-ordinates r and ϕ. The inverse transformation is given by $r = \sqrt{x^2+y^2}$ and $\tan\phi = \frac{y}{x}$. Using the chain rule we easily find that the partial derivative $\frac{\partial f}{\partial x}$ of a function $f(x,y)$ can be written[d] as $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial x} = \cos\phi\frac{\partial f}{\partial r} - \frac{1}{r}\sin\phi\frac{\partial f}{\partial \phi}$. In the same way we find $\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial y} = \sin\phi\frac{\partial f}{\partial r} + \frac{1}{r}\cos\phi\frac{\partial f}{\partial \phi}$. In the same way as we defined symbolically the vector operator $\boldsymbol{\nabla}$ we can also consider the derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ as *differential operators*, functions mapping real valued functions to other real valued functions involving derivatives of the original functions. It is now obvious that we can write symbolically for these differential operators:

$$\frac{\partial}{\partial x} = \cos\phi\frac{\partial}{\partial r} - \frac{1}{r}\sin\phi\frac{\partial}{\partial \phi}\,, \tag{124}$$

$$\frac{\partial}{\partial y} = \sin\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\phi\frac{\partial}{\partial \phi}\,. \tag{125}$$

The question is now, if we write a function $f(x,y)$ into polar co-ordinates $f(x(r,\phi),y(r,\phi))$ we now know how to calculate the partial derivatives in the new co-ordinates, but is there an easy way the vector $\boldsymbol{\nabla} f(x,y)$ could be expressed in the new coordinate system?

If we define the vectors $\mathbf{e}_r = \begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}$ and $\mathbf{e}_\phi = \begin{pmatrix}-\sin\phi\\ \cos\phi\end{pmatrix}$ then we find $\mathbf{e}_x = \cos\phi\,\mathbf{e}_r - \sin\phi\,\mathbf{e}_\phi$ and $\mathbf{e}_y = \sin\phi\,\mathbf{e}_r + \cos\phi\,\mathbf{e}_\phi$. Substituting this into $\boldsymbol{\nabla} = \mathbf{e}_x\frac{\partial}{\partial x} + \mathbf{e}_y\frac{\partial}{\partial y}$ leads to a simple expression of $\boldsymbol{\nabla}$ in polar co-ordinates:

52 $\boldsymbol{\nabla} f$ in plane polar co-ordinates
$f(x(r,\phi),y(r,\phi))$ partially differentiable, then

$$\boldsymbol{\nabla} f(x,y) = \frac{\partial f(x(r,\phi),y(r,\phi))}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial f(x(r,\phi),y(r,\phi))}{\partial \phi}\mathbf{e}_\phi\,, \tag{126}$$

with $x = r\cos\phi$, $y = r\sin\phi$ and $\mathbf{e}_r = \begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}$, $\mathbf{e}_\phi = \begin{pmatrix}-\sin\phi\\ \cos\phi\end{pmatrix}$.

In the case of 3-dimensional cylindrical co-ordinates $x = r\cos\phi$, $y = r\sin\phi$ and $z \in \mathbb{R}$ [52] generalises very easily:

53 $\boldsymbol{\nabla} f$ in cylindrical polar coordinates
$f(x(r,\phi),y(r,\phi),z)$ partially differentiable, then

$$\boldsymbol{\nabla} f(x,y) = \frac{\partial f(x(r,\phi),y(r,\phi),z)}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial f(x(r,\phi),y(r,\phi),z)}{\partial \phi}\mathbf{e}_\phi + \frac{\partial f(x(r,\phi),y(r,\phi),z)}{\partial z}\mathbf{e}_z\,,$$

with $x = r\cos\phi$, $y = r\sin\phi$ and $\mathbf{e}_r = \begin{pmatrix}\cos\phi\\ \sin\phi\\ 0\end{pmatrix}$, $\mathbf{e}_\phi = \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix}$ and $\mathbf{e}_z = \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$.

[d] For simplification of notation we simply use $\frac{\partial f}{\partial r}$ for $\frac{\partial f(x(r,\phi),y(r,\phi))}{\partial r}$ and $\frac{\partial f}{\partial \phi}$ for $\frac{\partial f(x(r,\phi),y(r,\phi))}{\partial \phi}$.

Let us now more generally consider the Cartesian coordinates $(x_1, \ldots, x_n)$ in the vector space $\mathbb{R}^n$. Let us assume that we have found a coordinate transformation into the new co-ordinates $(q_1, \ldots, q_n)$. The coordinates q_i are functions of the *old* co-ordinates $(x_1, \ldots, x_n)$: $q_i(x_1, \ldots, x_n)$. These functions are not necessarily linear in $x_1, \ldots, x_n$ but they should form a one-to-one transformation (at least on some suitable regions for $(x_1, \ldots, x_n)$ and $(q_1, \ldots, q_n)$) in order to be called a *coordinate transformation*. If these functions were linear in $x_1, \ldots, x_n$ then we would find a matrix S such that

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} . \tag{127}$$

In this case the corresponding basis vectors would transform according to $\mathbf{f}_i = S_{ji}\mathbf{e}_j$ as shown in the Part IA course Algebra & Geometry. However, if we assume that the coordinate transformation is not linear what would then be a suitable transformation for the basis vectors? Just like for plane polar coordinates we expect that the *new* basis vectors $\mathbf{f}_i$ will depend on the position. The obvious choice is to look at the position vector $\mathbf{x} = (x_1, \ldots, x_n)^T$ and start varying one of the new coordinates q_i whilst keeping all others invariant. The direction given by the change of $\mathbf{x}$ under this variation is defined to be the basis vector $\mathbf{f}_i$. In other words, $\mathbf{f}_i$ is the directional vector tangent to the movement of the position vector $\mathbf{x}$ if you increase q_i and leave all other q_j fixed. This direction is of course given by the partial derivative of $\mathbf{x}$ with respect to q_i. If the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are mutually orthonormal then we call the coordinates $q_1, \ldots, q_n$ *orthogonal curvilinear coordinates.*

54 General orthogonal curvilinear coordinates
Given a set of coordinates $q_1, \ldots, q_n$ and let $\mathbf{x} = (x_1(q_1, \ldots, q_n), \ldots, x_n(q_1, \ldots, q_n))^T$ denote the position vector. We construct basis vectors $\mathbf{f}_j$ (depending on the position) in the following way:

$$\mathbf{f}_j := \frac{1}{h_{(j)}} \frac{\partial \mathbf{x}}{\partial q_{(j)}} \quad \text{with} \quad h_j := \left\| \frac{\partial \mathbf{x}}{\partial q_j} \right\| , \tag{128}$$

where the brackets around the indices in Eq. 128 *denote that no summation convention should be applied. If the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are orthonormal then the co-ordinates $q_1, \ldots, q_n$ are called orthogonal curvilinear co-ordinates. It is clear from* Eq. 128 *that the vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ are in general functions of the position in $\mathbb{R}^n$.*

A non-trivial calculation gives the following transformation of $\boldsymbol{\nabla} f$ to general orthogonal curvilinear co-ordinates.

55 $\boldsymbol{\nabla} f$ in general orthogonal curvilinear co-ordinates
In general orthogonal curvilinear co-ordinates $\boldsymbol{\nabla} f(\mathbf{x})$ is

$$\boldsymbol{\nabla} f(\mathbf{x}) \;=\; \frac{1}{h_1} \frac{\partial f}{\partial q_1} \mathbf{f}_1 + \ldots + \frac{1}{h_n} \frac{\partial f}{\partial q_n} \mathbf{f}_n \, . \tag{129}$$

We will use [55] in order to find $\boldsymbol{\nabla} f$ in spherical polar coordinates (Eq. 2)

$$\mathbf{x} \;=\; \begin{pmatrix} r\cos\phi\sin\theta \\ r\sin\phi\sin\theta \\ r\cos\theta \end{pmatrix} , \tag{130}$$

with $\phi \in [0, 2\pi)$, $\theta \in [0, \pi]$ and $r \in \mathbb{R}_0^+$. The partial derivatives of the position vector with respect to the co-ordinates r, ϕ or θ give $h_r = 1$, $h_\phi = r\sin\theta$ and $h_\theta = r$ respectively. Hence, the spherical basis vectors $\mathbf{e}_r$, $\mathbf{e}_\phi$ and $\mathbf{e}_\theta$ are

$$\mathbf{e}_r = \begin{pmatrix} \cos\phi\sin\theta \\ \sin\phi\sin\theta \\ \cos\theta \end{pmatrix}, \quad \mathbf{e}_\phi = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}, \quad \mathbf{e}_\theta = \begin{pmatrix} \cos\phi\cos\theta \\ \sin\phi\cos\theta \\ -\sin\theta \end{pmatrix}. \tag{131}$$

We can now use [55] to find $\boldsymbol{\nabla} f$ is spherical polar co-ordinates.

56 $\boldsymbol{\nabla} f$ spherical polar coordinates

$$\boldsymbol{\nabla} f(\mathbf{x}) = \frac{\partial f}{\partial r}\mathbf{e}_r + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\mathbf{e}_\phi + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{e}_\theta . \tag{132}$$

15 Vector operator identities

So far we have come across two differential vector operators, the gradient of a real valued function $\boldsymbol{\nabla} f$ and the curl of a vector field $\boldsymbol{\nabla} \times \mathbf{F}$. There is a third important vector operator, the *divergence* of a vector field: $\mathrm{div}\mathbf{F} := \boldsymbol{\nabla}.\mathbf{F}$. Each of these differential operations can formally be expressed in Cartesian co-ordinates using the differential vector operator

$$\boldsymbol{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}. \tag{133}$$

Later we will also study the properties of the divergence of a gradient field: $\boldsymbol{\nabla}.\boldsymbol{\nabla} f(\mathbf{x})$ which is usually denoted by $\boldsymbol{\nabla}^2 f(\mathbf{x})$. The operator $\boldsymbol{\nabla}^2$ is called the *Laplace operator*[e].

57 Differential vector operators[f].

$$\mathrm{grad} f(\mathbf{x}) = \boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{e}_i \frac{\partial f(\mathbf{x})}{\partial x_i}, \tag{134}$$

$$\mathrm{div}\mathbf{F}(\mathbf{x}) = \boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = \frac{\partial F_i(\mathbf{x})}{\partial x_i}, \tag{135}$$

$$\mathrm{curl}\mathbf{F}(\mathbf{x}) = \boldsymbol{\nabla} \times \mathbf{F}(\mathbf{x}) = \mathbf{e}_i \epsilon_{ijk} \frac{\partial}{\partial x_j} F_k(\mathbf{x}), \tag{136}$$

$$\boldsymbol{\nabla}^2 f(\mathbf{x}) = \boldsymbol{\nabla}.\boldsymbol{\nabla} f(\mathbf{x}) = \sum_i \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}. \tag{137}$$

The following is a list of useful vector operator identities using the vector operator $\boldsymbol{\nabla}$. The proofs are straightforward and are most easily done using index notation in Cartesian co-ordinates.

[e] *Pierre-Simon* Laplace, 1749-1827.

[f] With the exception of the curl these vector operators are all defined for any $n \in \mathbb{N}$. The curl is only defined for $n = 3$.

58 Vector operator identities

$$\boldsymbol{\nabla}(g(\mathbf{x})f(\mathbf{x})) = f\boldsymbol{\nabla}g + g\boldsymbol{\nabla}f \,, \tag{138}$$

$$\boldsymbol{\nabla}.\left(g(\mathbf{x})\mathbf{F}(\mathbf{x})\right) = (\boldsymbol{\nabla}g).\mathbf{F} + g\boldsymbol{\nabla}.\mathbf{F} \,, \tag{139}$$

$$\boldsymbol{\nabla}\times(\mathbf{F}(\mathbf{x})\times\mathbf{G}(\mathbf{x})) = \mathbf{G}.\boldsymbol{\nabla}\mathbf{F} + (\boldsymbol{\nabla}.\mathbf{G})\mathbf{F} - (\boldsymbol{\nabla}.\mathbf{F})\mathbf{G} - \mathbf{F}.\boldsymbol{\nabla}\mathbf{G} \,, \tag{140}$$

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}f(\mathbf{x}) = \mathbf{0} \,. \tag{141}$$

We can use the theory of general orthogonal curvilinear co-ordinates developed in Sec. 14 to transform these vector operators into general curvilinear co-ordinates. We give the results without proof:

59 Vector operators in general orthogonal curvilinear co-ordinates in $\mathbb{R}^3$

The general orthogonal curvilinear co-ordinates in $\mathbb{R}^3$ are defined as in [54]*. The differential vector operators are given by*

$$\boldsymbol{\nabla}f(\mathbf{x}) = \frac{1}{h_1}\frac{\partial f(\mathbf{x})}{\partial q_1}\mathbf{f}_1 + \frac{1}{h_2}\frac{\partial f(\mathbf{x})}{\partial q_2}\mathbf{f}_2 + \frac{1}{h_3}\frac{\partial f(\mathbf{x})}{\partial q_3}\mathbf{f}_3 \,, \tag{142}$$

$$\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = \frac{1}{h_1h_2h_3}\left\{\frac{\partial}{\partial q_1}\left(F_{q_1}h_2h_3\right) + \frac{\partial}{\partial q_2}\left(h_1F_{q_2}h_3\right) + \frac{\partial}{\partial q_3}\left(h_1h_2F_{q_3}\right)\right\} \,, \tag{143}$$

$$\boldsymbol{\nabla}\times\mathbf{F} = \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\mathbf{f}_1 & h_2\mathbf{f}_2 & h_3\mathbf{f}_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1F_{q_1} & h_2F_{q_2} & h_3F_{q_3} \end{vmatrix} \,, \tag{144}$$

$$\boldsymbol{\nabla}^2f(\mathbf{x}) = \frac{1}{h_1h_2h_3}\left\{\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3h_1}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right\} \,, \tag{145}$$

where F_{q_i} denotes the components of the vector field $\mathbf{F}$ with respect to the general orthogonal curvilinear basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$. The formal determinant in Eq. 144 *has to be evaluated in such a way that the partial derivatives in the second row act on the functions in the third row but not on the factors in the first row.*

Due to their importance, we summarise the special cases of polar co-ordinates explicitly:

60 Vector operators in plane polar co-ordinates

$$\boldsymbol{\nabla}f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial f(\mathbf{x})}{\partial \phi}\mathbf{e}_\phi \,, \tag{146}$$

$$\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = \frac{1}{r}\frac{\partial rF_r}{\partial r} + \frac{1}{r}\frac{\partial F_\phi}{\partial \phi} \,, \tag{147}$$

$$\boldsymbol{\nabla}^2f(\mathbf{x}) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \phi^2} \,. \tag{148}$$

61 Vector operators in cylindrical polar co-ordinates

$$\boldsymbol{\nabla} f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \rho}\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial f(\mathbf{x})}{\partial \phi}\mathbf{e}_\phi + \frac{\partial f(\mathbf{x})}{\partial z}\mathbf{e}_z \ , \tag{149}$$

$$\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = \frac{1}{\rho}\frac{\partial \rho F_\rho}{\partial \rho} + \frac{1}{\rho}\frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z} \ , \tag{150}$$

$$\boldsymbol{\nabla} \times \mathbf{F} = \left(\frac{1}{\rho}\frac{\partial F_z}{\partial \phi} - \frac{\partial F_\phi}{\partial z}\right)\mathbf{e}_\rho + \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial \rho}\right)\mathbf{e}_\phi + \frac{1}{\rho}\left(\frac{\partial \rho F_\phi}{\partial \rho} - \frac{\partial F_\rho}{\partial \phi}\right)\mathbf{e}_z \ , \tag{151}$$

$$\boldsymbol{\nabla}^2 f(\mathbf{x}) = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial f}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2} \ . \tag{152}$$

62 Vector operators in spherical polar co-ordinates

$$\boldsymbol{\nabla} f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial r}\mathbf{e}_r + \frac{1}{r\sin\theta}\frac{\partial f(\mathbf{x})}{\partial \phi}\mathbf{e}_\phi + \frac{1}{r}\frac{\partial f(\mathbf{x})}{\partial \theta}\mathbf{e}_\theta \ , \tag{153}$$

$$\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = \frac{1}{r^2}\frac{\partial r^2 F_r}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial \phi} + \frac{1}{r\sin\theta}\frac{\partial \sin\theta F_\theta}{\partial \theta} \ , \tag{154}$$

$$\boldsymbol{\nabla} \times \mathbf{F} = \frac{1}{r\sin\theta}\left(\frac{\partial \sin\theta F_\phi}{\partial \theta} - \frac{\partial F_\theta}{\partial \phi}\right)\mathbf{e}_r + \frac{1}{r}\left(\frac{\partial r F_\theta}{\partial r} - \frac{\partial F_r}{\partial \theta}\right)\mathbf{e}_\phi + \frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial F_r}{\partial \phi} - \frac{\partial r F_\phi}{\partial r}\right)\mathbf{e}_\theta \ ,$$

$$\boldsymbol{\nabla}^2 f(\mathbf{x}) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \phi^2} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) \ . \tag{155}$$

Examples

Example 4.10 *We will calculate* $\boldsymbol{\nabla} f$ *in spherical polar coordinates and axes by transforming from Cartesian axes. We need to transform the derivatives also the axes. Let* $f(x,y,z) = F(r,\theta,\phi)$*. We have*

$$\begin{aligned}\boldsymbol{\nabla} f &= \frac{\partial f}{\partial x}\,\mathbf{e}_x + \frac{\partial f}{\partial y}\,\mathbf{e}_y + \frac{\partial f}{\partial z}\,\mathbf{e}_z \\ &= \left(\frac{\partial F}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial F}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial F}{\partial \phi}\frac{\partial \phi}{\partial x}\right)\mathbf{e}_x + \left(\frac{\partial F}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial F}{\partial \theta}\frac{\partial \theta}{\partial y} + \cdots\right)\mathbf{e}_y + \left(\cdots\right)\mathbf{e}_z\end{aligned}$$

The next step is to calculate $\partial r/\partial x$*, etc, by differentiating*

$$r = (x^2+y^2+z^2)^{\frac{1}{2}}, \quad \theta = \cot^{-1}(z/(x^2+y^2)^{\frac{1}{2}}), \quad \phi = \tan^{-1}(y/x)$$

Finally, we must express $\mathbf{e}_x$*, etc, in terms of* $\mathbf{e}_r$*, etc. Inspection of the diagram on page 3 shows that*

$$\begin{aligned}\mathbf{e}_x &= (\mathbf{e}_r\sin\theta + \mathbf{e}_\theta\cos\theta)\cos\phi - \mathbf{e}_\phi\sin\phi \ , \\ \mathbf{e}_y &= (\mathbf{e}_r\sin\theta + \mathbf{e}_\theta\cos\theta)\sin\phi + \mathbf{e}_\phi\cos\phi \ , \\ \mathbf{e}_z &= \mathbf{e}_r\cos\theta - \mathbf{e}_\theta\sin\theta \ .\end{aligned}$$

Putting all this together gives

$$\boldsymbol{\nabla} F = \frac{\partial F}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial F}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial F}{\partial \phi}\mathbf{e}_\phi \ .$$

Example 4.11 *Let*

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

where $r = \|\mathbf{x}\|$. *Then*

$$\begin{aligned}
(\mathbf{h}.\boldsymbol{\nabla})f(\mathbf{x}) &= h_j \frac{\partial}{\partial x_j}\left(\frac{1}{r}\right) = -\frac{h_j x_j}{r^3} \\
(\mathbf{h}.\boldsymbol{\nabla})^2 f(\mathbf{x}) &= (\mathbf{h}.\boldsymbol{\nabla})((\mathbf{h}.\boldsymbol{\nabla})f) \\
&= h_k \frac{\partial}{\partial x_k}\left(-\frac{h_j x_j}{r^3}\right) \\
&= -h_j h_k x_j . \left(-\frac{3}{r^4}\right).\frac{x_k}{r} - h_j h_k \left(\frac{\delta_{jk}}{r^3}\right) \\
&= \frac{h_j h_k}{r^5}\left(3x_j x_k - r^2\delta_{jk}\right)\,.
\end{aligned}$$

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

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

Example 4.12 *Examples of divergences.*

(i) Let

$$\mathbf{F} = (x^2 y,\, y^2 z,\, z^2 x)\,.$$

Then

$$\boldsymbol{\nabla}.\,\mathbf{F} = \frac{\partial(x^2 y)}{\partial x} + \frac{\partial(y^2 z)}{\partial y} + \frac{\partial(z^2 x)}{\partial z} = 2xy + 2yz + 2zx\,.$$

(ii) Let $\mathbf{F} = \mathbf{x}$. *Then*

$$\boldsymbol{\nabla}.\,\mathbf{F} = \boldsymbol{\nabla}.\,\mathbf{x} = \frac{\partial x_j}{\partial x_j} = \delta_{jj} = 3\,,$$

i.e.

$$\boldsymbol{\nabla}.\,\mathbf{x} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3\,.$$

(iii) Let

$$\mathbf{F} = \frac{\mathbf{x}}{r^3}.$$

Then

$$\boldsymbol{\nabla}.\left(\frac{\mathbf{x}}{r^3}\right) = \frac{\partial}{\partial x_j}\left(\frac{x_j}{r^3}\right) = \frac{\delta_{jj}}{r^3} - 3\frac{x_j}{r^4}\frac{x_j}{r} = \frac{3}{r^3} - \frac{3r^2}{r^5} = 0\,.$$

Alternatively, we can easily show (using suffix notation) that, for any vector field $\mathbf{F}(\mathbf{x})$ *and scalar field* $\psi(\mathbf{x})$,

$$\boldsymbol{\nabla}.\,(\psi\mathbf{F}) = \boldsymbol{\nabla}\psi.\,\mathbf{F} + \psi\boldsymbol{\nabla}.\,\mathbf{F}$$

(indeed, what else could $\boldsymbol{\nabla}.\,(\psi\mathbf{F})$ *be?), so*

$$\boldsymbol{\nabla}.\left(\frac{\mathbf{x}}{r^3}\right) = \boldsymbol{\nabla}\left(\frac{1}{r^3}\right).\mathbf{x} + \frac{1}{r^3}\boldsymbol{\nabla}.\,\mathbf{x} = \left(\frac{-3\mathbf{x}}{r^5}\right).\mathbf{x} + \frac{3}{r^3} = 0.$$

Example 4.13 *Examples of curls.*

(i) Let

$$\mathbf{F} = (x^2y,\, y^2z,\, z^2x)\,.$$

Then

$$\boldsymbol{\nabla}\times\mathbf{F} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ x^2y & y^2z & z^2x \end{vmatrix} = (-y^2, -z^2, -x^2)\,.$$

(ii) Let $\mathbf{F} = \mathbf{x}$. *Then*

$$[\boldsymbol{\nabla}\times\mathbf{F}]_i \;=\; \epsilon_{ijk}\,\frac{\partial}{\partial x_j}\,x_k = \;\epsilon_{ijk}\,\delta_{jk} = \epsilon_{ijj}0\,.$$

(iii) Let $\mathbf{F} = \dfrac{\mathbf{x}}{r^3}$. *Then*

$$[\boldsymbol{\nabla}\times\mathbf{F}]_i = \epsilon_{ijk}\,\frac{\partial}{\partial x_j}\left(\frac{x_k}{r^3}\right) = \epsilon_{ijk}\,\frac{\delta_{jk}}{r^3} - 3\,\epsilon_{ijk}\,\frac{x_k}{r^4}\,\frac{x_j}{r} = 0\,.$$

Alternatively we can easily show (using suffix notation) that, for any vector field $\mathbf{F}(\mathbf{x})$ *and scalar field* $\psi(\mathbf{x})$,

$$\boldsymbol{\nabla}\times(\psi\mathbf{F}) = \boldsymbol{\nabla}\psi\times\mathbf{F} + \psi\boldsymbol{\nabla}\times\mathbf{F}\ ,$$

so

$$\boldsymbol{\nabla}\times\left(\frac{\mathbf{x}}{r^3}\right) = \boldsymbol{\nabla}\left(\frac{1}{r^3}\right)\times\mathbf{x} + \frac{1}{r^3}\boldsymbol{\nabla}\times\mathbf{x} = \left(\frac{-3\mathbf{x}}{r^5}\right)\times\mathbf{x} + \mathbf{0} = \mathbf{0}.$$

Note that $\dfrac{\mathbf{x}}{r^3} = -\boldsymbol{\nabla}\left(\dfrac{1}{r}\right)$, *so*

$$\boldsymbol{\nabla}\times\left(\frac{\mathbf{x}}{r^3}\right) = \boldsymbol{\nabla}\times\boldsymbol{\nabla}\left(\frac{1}{r}\right) = \mathbf{0}$$

since the curl of a gradient is zero.

Example 4.14 *We will prove, using suffix notation, the identity*

$$\boldsymbol{\nabla}\times(\mathbf{F}\times\mathbf{G}) = \mathbf{F}\,(\boldsymbol{\nabla}.\,\mathbf{G}) - \mathbf{G}\,(\boldsymbol{\nabla}.\,\mathbf{F}) + (\mathbf{G}.\,\boldsymbol{\nabla})\mathbf{F} - (\mathbf{F}.\,\boldsymbol{\nabla})\mathbf{G}\,.$$

The i *component of the left hand side is*

$$\begin{aligned}
\epsilon_{ijk}\frac{\partial}{\partial x_j}(\mathbf{F}\times\mathbf{G})_k &= \epsilon_{ijk}\frac{\partial}{\partial x_j}\epsilon_{kmn}F_mG_n \\
&= (\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})\left(\frac{\partial F_m}{\partial x_j}G_n + F_m\frac{\partial G_n}{\partial x_j}\right) \\
&= \frac{\partial F_i}{\partial x_j}G_j - \frac{\partial F_j}{\partial x_j}G_i + F_i\frac{\partial G_j}{\partial x_j} - F_j\frac{\partial G_i}{\partial x_j}.
\end{aligned}$$

The last expression is the i *component of*

$$(\mathbf{G}.\,\boldsymbol{\nabla})\mathbf{F} - \mathbf{G}\,(\boldsymbol{\nabla}.\,\mathbf{F}) + \mathbf{F}\,(\boldsymbol{\nabla}.\,\mathbf{G}) - (\mathbf{F}.\,\boldsymbol{\nabla})\mathbf{G}\,,$$

as required.

Example 4.15 *We will calculate* $\nabla^2 r^{-1}$ *(for* $r \neq 0$*).*

First recall that

$$\frac{\partial r}{\partial x_i} = \frac{x_i}{r}\,.$$

Then

$$\begin{aligned}\nabla^2\left(\frac{1}{r}\right) &= \frac{\partial^2}{\partial x_i \partial x_i}\left(\frac{1}{r}\right) = \frac{\partial}{\partial x_i}\left(-\frac{1}{r^2}\frac{\partial r}{\partial x_i}\right) = \frac{\partial}{\partial x_i}\left(-\frac{x_i}{r^3}\right)\\ &= \frac{3x_i}{r^4}\frac{\partial r}{\partial x_i} - \frac{1}{r^3}\frac{\partial x_i}{\partial x_i} = \frac{3x_i}{r^4}\frac{x_i}{r} - \frac{1}{r^3}\,\delta_{ii}\\ &= \frac{3}{r^3} - \frac{3}{r^3}\\ &= 0\,.\end{aligned}$$

Alternatively

$$\boldsymbol{\nabla}\left(\frac{1}{r}\right) = -\frac{\mathbf{x}}{r^3} \qquad \textit{and} \qquad \boldsymbol{\nabla}.\left(\frac{\mathbf{x}}{r^3}\right) = 0\,.$$

Or, setting $1/r = f(r)$ *and using the expression for* ∇^2 *in spherical polars:*

$$\nabla^2 f = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rf) + \quad \textit{terms with } \theta \textit{ and } \phi \textit{ derivatives} = 0\,.$$

Example 4.16 *We will calculate the divergence and curl of the vector field* $\mathbf{B} = \mathbf{x}$ *using spherical and cylindrical coordinates.*

The general form of $\boldsymbol{\nabla}.\mathbf{B}$ *in curvilinear coordinates is*

$$\boldsymbol{\nabla}.\mathbf{B} = \frac{1}{h_1 h_2 h_3}\frac{\partial}{\partial q_1}\big(h_2 h_3 B_1\big) + \textit{two cyclic terms}$$

and the general form of $\boldsymbol{\nabla} \times \mathbf{B}$ *in curvilinear coordinates is*

$$\frac{1}{h_1 h_2 h_3}\begin{vmatrix} h_1\mathbf{e}_1 & h_2\mathbf{e}_2 & h_3\mathbf{e}_3 \\ \dfrac{\partial}{\partial q_1} & \dfrac{\partial}{\partial q_2} & \dfrac{\partial}{\partial q_3} \\ h_1 B_1 & h_2 B_2 & h_3 B_3 \end{vmatrix}$$

(i) Spherical coordinates

We have

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\; d\phi^2$$

so $h_1 = 1$, $h_2 = r$ *and* $h_3 = r\sin\theta$. *Also,* $\mathbf{x} = r\mathbf{e}_r$.

Thus

$$\boldsymbol{\nabla}.\mathbf{x} = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial r}\big(r^2\sin\theta\; r\big) = 3$$

and

$$\boldsymbol{\nabla}\times\mathbf{x} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\,\mathbf{e}_\phi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial\phi} \\ r & 0 & 0 \end{vmatrix} = 0$$

(ii) Cylindrical coordinates

We have

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 + dz^2$$

so $h_1 = 1$, $h_2 = \rho$ *and* $h_3 = 1$. *Also,* $\mathbf{x} = \rho\mathbf{e}_\rho + z\mathbf{e}_z$.

Thus

$$\boldsymbol{\nabla}.\mathbf{x} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho) + \frac{1}{\rho}\frac{\partial}{\partial z}(\rho z) = 3$$

and

$$\boldsymbol{\nabla} \times \mathbf{x} = \frac{1}{\rho}\begin{vmatrix} \mathbf{e}_\rho & \rho\mathbf{e}_\phi & \mathbf{e}_z \\ \dfrac{\partial}{\partial\rho} & \dfrac{\partial}{\partial\phi} & \dfrac{\partial}{\partial z} \\ \rho & 0 & z \end{vmatrix} = 0$$

V Integration Theorems

When we defined surface and volume integrals in Sec. 4 we noticed that provided the function $f(x,y)$ defined on the flat surface $\mathcal{S}$ is non-negative the surface integral over f on $\mathcal{S}$ is the same as the volume integral over the volume $\mathcal{V}$ underneath the graph of f:

$$\int_{\mathcal{S}} f(x,y)d(x,y) \quad = \quad \int_{\mathcal{V}} d(x,y,z)\ . \tag{156}$$

This is of course immediately obvious from the definition of the two integrals but conceptually the integrals are of course very different: the first is an integral over a surface whilst the second integrates over a whole volume. There are in fact several other identities between the different integrals we have introduced but although these identities are less obvious they are extremely important and are probably the main results of this course! We will now study in more detail these relations and will start with the simplest of these integrals: integrals over flat surfaces and line integrals.

16 Green's theorem in the plane

Let us assume that the vector field $\mathbf{F}$ is defined on the flat surface $\mathcal{B} \subset \mathbb{R}^2$. We call the component functions $P(x,y)$ and $Q(x,y)$:

$$\mathbf{F}(x,y) \quad := \quad \begin{pmatrix} P(x,y) \\ Q(x,y) \end{pmatrix}. \tag{157}$$

Let us assume further that the surface $\mathcal{B}$ is *x-simple* which means that $\mathcal{B}$ is bounded by two functions $f(x)$, $g(x)$ and two vertical straight lines as in Fig. xxxii. We parameterise the boundary $\partial\mathcal{B}$ counterclockwise using four paths $\boldsymbol{\gamma}_i(t)$.

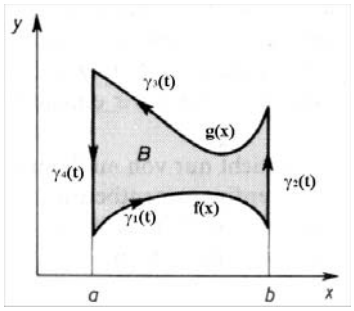


Fig. xxxii An x-simple surface.

The paths $\boldsymbol{\gamma}_i(t)$ are given by

$$\boldsymbol{\gamma}_1(t) \quad = \quad \begin{pmatrix} t \\ f(x) \end{pmatrix}, \qquad t \in [a,b]\ , \tag{158}$$

$$\boldsymbol{\gamma}_2(t) \quad = \quad \begin{pmatrix} b \\ f(b) + t(g(b) - f(b)) \end{pmatrix}, \qquad t \in [0,1]\ , \tag{159}$$

$$\boldsymbol{\gamma}_3^-(t) \quad = \quad \begin{pmatrix} t \\ g(x) \end{pmatrix}, \qquad t \in [a,b]\ , \tag{160}$$

$$\boldsymbol{\gamma}_4^-(t) \quad = \quad \begin{pmatrix} a \\ f(a) + t(g(a) - f(a)) \end{pmatrix}, \qquad t \in [0,1]\ . \tag{161}$$

This leads to

$$\int_{\mathcal{B}} \frac{\partial P(x,y)}{\partial y} d(x,y) = \int_a^b \int_{f(x)}^{g(x)} \frac{\partial P(x,y)}{\partial y}\, dy\, dx = \int_a^b P(x,g(x))dx - \int_a^b P(x,f(x))dx$$
$$= \int_a^b P(\boldsymbol{\gamma}_3^-(t))dt - \int_a^b P(\boldsymbol{\gamma}_1(t))dt\ . \qquad (162)$$

But since $dx = dt$ on $\boldsymbol{\gamma}_1(t)$ and $\boldsymbol{\gamma}_3^-(t)$ and $dx = 0$ on $\boldsymbol{\gamma}_2(t)$ and $\boldsymbol{\gamma}_4^-(t)$ we find

$$\int_{\mathcal{B}} \frac{\partial P(x,y)}{\partial y} d(x,y) = -\int_{\gamma_1} P\,dx - \int_{\gamma_2} P\,dx - \int_{\gamma_3} P\,dx - \int_{\gamma_4} P\,dx$$
$$= -\oint_{\partial\mathcal{B}} P dx + 0\,dy\ , \qquad (163)$$

where $\oint_{\partial\mathcal{B}} P dx + 0\,dy$ indicates a line integral along the closed boundary $\partial\mathcal{B}$ of $\mathcal{B}$ where the path has to be *positively oriented* (counterclockwise).

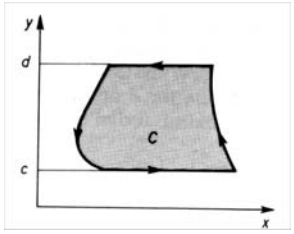


FIG. xxxiii A y-simple surface.

A similar calculation for a *y-simple* surface $\mathcal{C}$ shows that

$$\int_{\mathcal{C}} \frac{\partial Q(x,y)}{\partial x} d(x,y) = \oint_{\partial\mathcal{C}} 0\,dx + Q dy\ . \qquad (164)$$

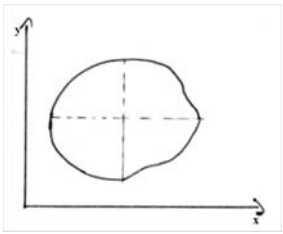


FIG. xxxiv A surface which is x-simple as well as y-simple.

In case we have a surface $\mathcal{S}$ which is both x-simple as well as y-simple then we can put EQ. 163 and EQ. 164 together to obtain

$$\int_{\mathcal{S}} \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}\right) d(x,y)$$
$$= \oint_{\partial\mathcal{S}} P dx + Q dy = \oint_{\partial\mathcal{S}} \mathbf{F}.d\mathbf{x}\ . \qquad (165)$$

By smoothening slightly the corners of an x-simple surface, the surface can easily be approximated by a surface which is both x-simple and y-simple and similarly for y-simple surfaces. And it is easy to see that any flat surface $\mathcal{S}$ with a piecewise smooth boundary can easily be split into a finite number of x-simple surfaces and y-simple surfaces and by smoothing all corners the surface can be approximated by a finite number of surfaces which are x-simple as well as y-simple.

By applying EQ. 165 to each of these small surfaces we see that the boundary integrals over the interior lines disappear since we go along each interior line twice but in opposite directions.

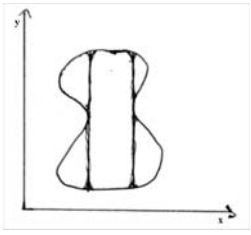


FIG. xxxv Splitting $\mathcal{S}$ into x-simple and y-simple surfaces.

Summing over all these integrals we obtain *Green's theorem in the plane*[a] which is also known as *Gauss' theorem in the plane*[b]:

63 **Green's theorem in the plane**
The vector field $\mathbf{F}(x,y) = (P(x,y), Q(x,y))^T$ *is defined on the surface* $\mathcal{S} \subset \mathbb{R}^2$ *with (piecewise) smooth boundary* $\partial\mathcal{S}$*, then*

$$\int_{\mathcal{S}} (\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}) d(x,y) \quad = \quad \oint_{\partial\mathcal{S}} \begin{pmatrix} P \\ Q \end{pmatrix} . d\mathbf{x} \,, \tag{166}$$

where $\oint_{\partial\mathcal{S}} \mathbf{F}.d\mathbf{x}$ *denotes the line integral along the closed boundary* $\partial\mathcal{S}$ *of* $\mathcal{S}$ *using a positively oriented parameterisation (counterclockwise).*

Example 5.1 *Let*

$$I = \oint_{\mathcal{C}} (x^2 y \, dx + xy^2 \, dy) \,,$$

for some curve $\mathcal{C}$ *bounding a surface* $\mathcal{A}$*.*

Green's theorem is

$$\oint_{\mathcal{C}} (P dx + Q dy) = \int_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dS \,.$$

Setting

$$P = x^2 y \quad and \quad Q = xy^2.$$

$$\begin{aligned} I &= \int_{\mathcal{A}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy \\ &= \int_{\mathcal{A}} (y^2 - x^2) \, dx dy \,. \end{aligned}$$

17 Stokes' theorem

It immediately occurs to us that $\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}$ in Green's theorem in the plane [63] looks like the third component of the curl of a vector field $\mathbf{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z))^T$ evaluated for $z = 0$. This seems to raise the hope that Green's theorem in the plane is in fact only a special case of a three dimensional theorem over some arbitrary surface in $\mathbb{R}^3$. We can easily embed Green's theorem in the plane [63] into three dimensions if we simply imagine that the flat surface $\mathcal{S}$ lies in the xy-plane and the vector field $\mathbf{F} = (P, Q, R)^T$ is defined everywhere on this surface. Obviously the normal vector to $\mathcal{S}$ is simply $\mathbf{e}_3$ and therefore $d\mathbf{S} = \mathbf{e}_3 \, dx \, dy$ for the surface $\mathcal{S}$. But $\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}$ is just the third component of the curl $\boldsymbol{\nabla} \times \mathbf{F}$ and taking the scalar product of $\boldsymbol{\nabla} \times \mathbf{F}$ and $d\mathbf{S}$ just projects out this third component. Since the surface $\mathcal{S}$ is completely contained in the xy-plane we find that $dz = 0$ along the boundary $\partial\mathcal{S}$ and therefore Green's theorem in the plane can be rewritten in the following form

$$\int_{\mathcal{S}} \boldsymbol{\nabla} \times \mathbf{F} \,.\, d\mathbf{S} \quad = \quad \oint_{\partial\mathcal{S}} \mathbf{F} \,.\, d\mathbf{x} \,, \tag{167}$$

[a] George Green, 1793-1841, Gonville and Caius College (1833-1840).
[b] Johann *Carl Friedrich* Gauss, 1777-1855.

for the flat surface $\mathcal{S}$ contained in the xy-plane and the boundary $\partial\mathcal{S}$ positively oriented[c] with respect to the direction of $d\mathbf{S}$. If we assume that $\mathcal{S}$ is now an arbitrary (smooth) surface in $\mathbb{R}^3$ then by applying Green's theorem in the plane to each component of the surface $\mathcal{S}$ now described by a three dimensional function $\mathbf{\Phi}(\mathbf{x})$ it can be shown that EQ. 167 generalises to any (piecewise smooth) surface in $\mathbb{R}^3$. This is known as *Stokes' theorem*[d].

64 Stokes' theorem

The vector field $\mathbf{F} : G \subset \mathbb{R}^3 \to \mathbb{R}^3$ *is defined on the domain* G *such that the (piecewise) smooth surface* $\mathcal{S}$ *is completely contained in* G. *We then find*[e]

$$\int_{\mathcal{S}} \boldsymbol{\nabla} \times \mathbf{F} \,.\, d\mathbf{S} \quad = \quad \oint_{\partial\mathcal{S}} \mathbf{F} \,.\, d\mathbf{x} \,. \tag{168}$$

Stokes' theorem also gives us a better understanding of the relation between gradient fields and conservative fields investigated in SEC. 9. We know from our remarks to [47] that a gradient field $\mathbf{F}(\mathbf{x})$ in $\mathbb{R}^3$ necessarily has to satisfy the integrability condition $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$. Conversely, let us assume that the vector field satisfies the integrability condition $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$ on some domain G. If we take a closed path $\boldsymbol{\gamma}$ in G and consider it to be the boundary of a surface $\mathcal{S}$ which lies completely in G then Stokes' theorem immediately shows that the closed line integral $\oint_{\boldsymbol{\gamma}} \mathbf{F}.d\mathbf{x} = 0$ and since this was for any closed path $\boldsymbol{\gamma}$ inside G we therefore find that the vector field $\mathbf{F}$ is conservative on G. The only question which remains is: how can we make sure that we would always find such a surface $\mathcal{S}$ which is bounded by $\boldsymbol{\gamma}$ (and $\boldsymbol{\gamma}$ is its only boundary)? It is immediately obvious that we can always find such a $\mathcal{S}$ in G if G is simply connected because the process of shrinking $\boldsymbol{\gamma}$ to a point exactly describes such an $\mathcal{S}$. This now explains where the condition of G being simply connected in [47] comes from[f].

What does $\boldsymbol{\nabla} \times \mathbf{F}$ actually mean geometrically? For a radially symmetric vector field we find that the curl vanishes: e.g. $\boldsymbol{\nabla} \times (x, y, z)^T = \mathbf{0}$. If $\mathbf{F}$ describes the flow of a fluid then we can say that the curl of a radially symmetric flow vanishes. However if we take a circular flow such as $(-y, x, 0)^T$ then we see that the curl is perpendicular to the surface in which the circular flow takes place and the magnitude of the curl describes the strength of the circular flow: e.g. $\boldsymbol{\nabla} \times (-y, x, 0)^T = (0, 0, 2)^T$. Therefore the curl describes whether at a point there are any local circular flows around that point, in other words, it describes the *whirlpools* in the vector fields $\mathbf{F}$ (the direction of $\boldsymbol{\nabla} \times \mathbf{F}$ is normal to the surface in which the local whirpool flow takes place and $\|\boldsymbol{\nabla} \times \mathbf{F}\|$ describes the strength of this whirlpool flow). In case $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$ everywhere on its domain then we say that $\mathbf{F}$ is *irrotational*. This of course just matches the integrability condition EQ. 113 but we should keep in mind that the domain of an irrotational field does not necessarily need to be simply connected and therefore not every irrotational field is conservative.

---

[c] Positively oriented means applying the *right hand rule*: if the thumb of the right hand points into the direction of $d\mathbf{S}$ then the other fingers of the right hand point naturally in the positive orientation of $\partial\mathcal{S}$.

[d] *George* Gabriel Stokes, 1819-1903, Pembroke College (1837-1902).

[e] There is of course an ambiguity in the sign of $d\mathbf{S}$ as well as in the orientation of the closed path $\partial\mathcal{S}$. If the surface $\mathcal{S}$ is parametrised by $\mathbf{\Phi}(u, v)$ then taking $d\mathbf{S} = \frac{\partial\mathbf{\Phi}}{\partial u} \times \frac{\partial\mathbf{\Phi}}{\partial v} d(u, v)$ requires $\partial\mathcal{S}$ to be parametrised in such a way that the underlying boundary of the parameter space in (u, v) space is parametrised counter-clockwise. This will give the correct sign factors such that the identity EQ. 168 holds. In practice it is easier to calculate both sides of EQ. 168 using some parametrisations and decide afterwards by looking at the signs of the integrals whether we have used the correct orientations.

[f] The condition of G being simply connected makes the necessary integrability condition $\boldsymbol{\nabla} \times \mathbf{F} = \mathbf{0}$ also sufficient. But of course G being simply connected is itself not necessary for $\mathbf{F}$ to be conservative on a domain G.

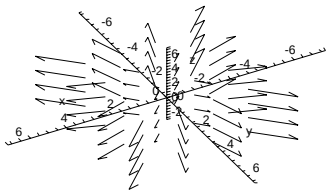


FIG. xxxvi Irrotational vector field $(x, y, 0)$.

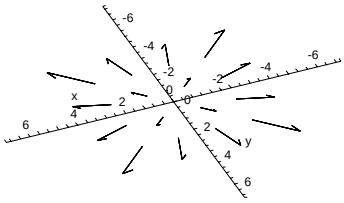


FIG. xxxvii $(x, y, 0)$ projected into xy plane.

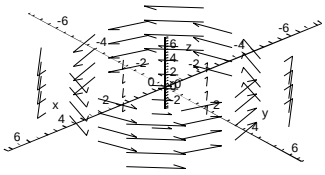


FIG. xxxviii Vector field $(-y, x, 0)$ with curl $(0, 0, 2)$.

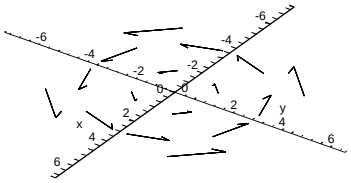


FIG. xxxix $(-y, x, 0)$ projected into xy plane.

It may at first seem difficult to understand why Stokes' theorem should be true. Let us think about the surface integral in the following way: $\int_{\mathcal{S}} \boldsymbol{\nabla} \times \mathbf{F} \,.\, d\mathbf{S}$ partitions the surface into little surface elements $d\mathbf{S}_{ij}$. For each surface element the dot product $\boldsymbol{\nabla} \times \mathbf{F} \,.\, d\mathbf{S}_{ij}$ gives us a measure for the strength of the *whirlpool flow* in this surface element. Adding neighbouring whirlpool flows describes the strength of the whirlpool flow in the larger surface as shown in FIG. xl. Therefore adding these whirlpool flows over the whole surface just gives the total flow around the boundary as shown in FIG. xli. But this of course is exactly what the line integral $\oint_{\partial \mathcal{S}} \mathbf{F} \,.\, d\mathbf{x}$ describes. It adds up the dot products $\mathbf{F} \,.\, d\mathbf{x}$ along the curve $\partial \mathcal{S}$, projecting the flow $\mathbf{F}$ onto the line segments $d\mathbf{x}$ and therefore calculating the overall flow around the curve $\partial \mathcal{S}$. Stokes' theorem then essentially says that *summing over all local whirlpools on a surface equals the flow along the path which is following the boundary of the surface.*

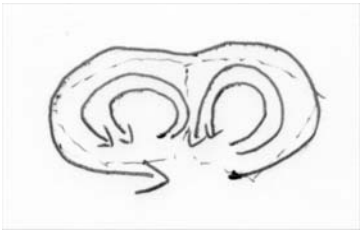


FIG. xl Adding neighbouring whirlpool flows.

FIG. xli Flow around the boundary $\partial \mathcal{S}$.

Example 5.2 *We will verify Stokes' theorem for the vector field* $\mathbf{F}(\mathbf{x}) = (y^2, z^2, x^2)$ *and the hemispherical surface* $x^2 + y^2 + z^2 = a^2$, $z \geq 0$.

First the surface integral. We have

$$\boldsymbol{\nabla}\times\mathbf{F}=\begin{vmatrix}\mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z\\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z}\\ y^2 & z^2 & x^2\end{vmatrix}=(-2z,-2x,-2y)\ .$$

Thus

$$\begin{aligned}\int_{\mathcal{S}}(\boldsymbol{\nabla}\times\mathbf{F}).\,d\mathbf{S} &= \int_0^{2\pi}\int_0^{\pi/2}(-2z,-2x,-2y).\,\hat{\mathbf{x}}\;a^2\sin\theta\;d\theta d\phi\\ &= \int_0^{2\pi}\int_0^{\pi/2}(-2z,-2x,-2y).\,(x/a,y/a,z/a)\;a^2\sin\theta\;d\theta d\phi\\ &= \int_0^{2\pi}\int_0^{\pi/2}(-2xz-2xy-2yz)a\sin\theta\;d\theta d\phi\\ &= -2a^3\int_0^{2\pi}\int_0^{\pi/2}(\sin\theta\cos\phi\cos\theta+\sin\theta\cos\phi\sin\theta\sin\phi+\sin\theta\sin\phi\cos\theta)\sin\theta\;d\theta d\phi\\ &= 0,\end{aligned}$$

since all the ϕ-integrals are zero.

Next the line integral. The bounding curve $\mathcal{C}$ is the circle $x^2+y^2=a^2$ in the plane $z=0$, so

$$\begin{aligned}\oint_{\mathcal{C}}\mathbf{F}.\,d\mathbf{x} &= \oint_{\mathcal{C}}y^2dx \qquad\qquad \textit{since } z=0=dz \textit{ on } \mathcal{C}\\ &= \int_0^{2\pi}(a\sin\theta)^2(-a\sin\theta)d\theta\ =\ 0\,.\end{aligned}$$

Example 5.3 *We will show that for a scalar field $f(\mathbf{x})$*

$$\int_{\mathcal{S}}\boldsymbol{\nabla}f\times d\mathbf{S}=-\oint_{\mathcal{C}}f\,d\mathbf{x}\,.$$

We use a very standard technique. We effectively dot the whole equation with an arbitrary constant vector field $\mathbf{k}$ to get it in a form to which Stokes' theorem applies; after applying the theorem, we can cancel off $\mathbf{k}$, since it is arbitrary. Since $\mathbf{k}$ is constant, it can be taken inside integral signs at will.

Stokes' theorem applied to the vector field $f(\mathbf{x})\mathbf{k}$, where $\mathbf{k}$ is an arbitrary constant vector, gives

$$\int_{\mathcal{S}}\boldsymbol{\nabla}\times(\mathbf{k}f).\,\mathbf{n}dS\ =\ \oint_{\mathcal{C}}f\mathbf{k}.\,d\mathbf{x}\,.$$

But

$$\begin{aligned}\boldsymbol{\nabla}\times(\mathbf{k}f).\,\mathbf{n} &= \mathbf{n}.\,\boldsymbol{\nabla}\times(\mathbf{k}f)\\ &= \mathbf{n}.\,(\boldsymbol{\nabla}f\times\mathbf{k}+f\boldsymbol{\nabla}\times\mathbf{k})\\ &= \mathbf{n}.\,(\boldsymbol{\nabla}f\times\mathbf{k}) \qquad\qquad \textit{since } \mathbf{k} \textit{ is constant}\\ &= -\mathbf{k}.\,(\boldsymbol{\nabla}f\times\mathbf{n})\,,\end{aligned}$$

and hence

$$-\mathbf{k}.\int_{\mathcal{S}} \boldsymbol{\nabla} f \times d\mathbf{S} = \mathbf{k}.\oint_{\mathcal{C}} f\, d\mathbf{x}.$$

It follows that since $\mathbf{k}$ *is arbitrary*

$$\int_{\mathcal{S}} \boldsymbol{\nabla} f \times d\mathbf{S} = -\oint_{\mathcal{C}} f\, d\mathbf{x}.$$

Putting it back into suffix notation gives

$$\epsilon_{ijk}\int_{\mathcal{S}} \frac{\partial f}{\partial x_j} n_k dS = -\oint_{\mathcal{C}} f\, dx_i.$$

If f *is the* m *component of a vector field, we have*

$$\epsilon_{ijk}\int_{\mathcal{S}} \frac{\partial f_m}{\partial x_j} n_k dS = -\oint_{\mathcal{C}} f_m\, dx_i,$$

which is a useful generalisation of Stokes' theorem.

18 The divergence theorem

Whilst Stokes' theorem relates surface integrals to line integrals we had already seen in our introductory remarks to Sec. 16 that surface integrals may also relate to volume integrals. Given a vector field $\mathbf{F} = (P, Q, R)^T$ defined on some volume $\mathcal{V}$ in $\mathbb{R}^3$ we can follow similar ideas to the proof of Green's theorem in the plane [63] and integrate the partial derivatives of the components of $\mathbf{F}$ over the volume $\mathcal{V}$ we find for example that $\int_{\mathcal{V}} \frac{\partial R}{\partial z} d(x, y, z) = \int_{\partial\mathcal{V}} R(d\mathbf{S})_3$ where $(d\mathbf{S})_3$ is the z-component of the surface element $d\mathbf{S}$ in the direction of the outward normal to the boundary $\partial\mathcal{V}$. Applying the same to the other components of the vector field leads to the *divergence theorem* which is also known as *Gauss' theorem.*

65 Divergence theorem

The vector field $\mathbf{F} : G \subset \mathbb{R}^3 \to \mathbb{R}^3$ *is defined on the set* G *such that the volume* $\mathcal{V}$ *with (piecewise) smooth boundary* $\partial\mathcal{V}$ *is completely contained in* G*. We then find*

$$\int_{\mathcal{V}} \boldsymbol{\nabla}\mathbf{F}\, dV \quad = \quad \oint_{\partial\mathcal{V}} \mathbf{F}\,.\,d\mathbf{S}\,, \tag{169}$$

where the surface element $d\mathbf{S}$ *points in the direction of the outward normal to the closed boundary* $\partial\mathcal{V}$ *of the volume* $\mathcal{V}$*.*

If the continuous vector field $\mathbf{F}$ describes the flow density of a fluid then we know that the surface integral over $\mathbf{F}$ describes the net flow of the vector field through the surface. In case the divergence of $\mathbf{F}$ is 0 everywhere in some region then the divergence theorem says that there is no net flow through any closed surface contained in that region which means that everything which flows inside any volume in the region has to come out of it again. Therefore there are obviously no *sinks* and neither any *sources* for the fluid inside this region. Conversely, in case the divergence of the continuous vector field $\mathbf{F}(\mathbf{x})$ is positive for some $\mathbf{x}$, then the divergence theorem says that the flow through some small closed surface around $\mathbf{x}$ is positive, and therefore we have a source for the fluid at $\mathbf{x}$. In the same way we see that for $\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) < 0$ we find a sink for the fluid at $\mathbf{x}$. To summarise, the divergence of $\mathbf{F}$ is a measure for the sinks $(\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) < 0)$

and sources ($\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) > 0$) of the vector field $\mathbf{F}$. In case $\boldsymbol{\nabla}.\mathbf{F}(\mathbf{x}) = 0$ everywhere, then we call $\mathbf{F}$ *solenoidal*[g] The divergence theorem therefore essentially says that *the flow of a vector field through a closed surface equals the sum over the local sinks and sources in the volume bounded by the closed surface.*

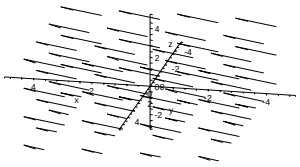


FIG. xlii Solenoidal vector field $(1,1,1)$.

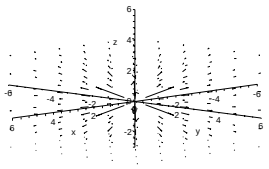


FIG. xliii Solenoidal vector field $(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3})$, for $r \neq 0$.

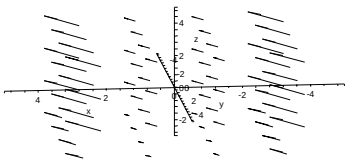


FIG. xliv Vector field $(|x|, |x|, |x|)$ with divergence 1 for $x > 0$ and -1 for $x < 0$.

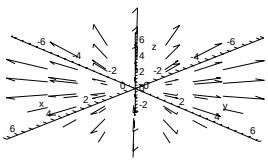


FIG. xlv Vector field (x, y, z) with divergence 3.

Comparing FIG. xliii and FIG. xlv we notice that in FIG. xliii the flow becomes less strong the further away we go from the origin. The flow starting at the origin just distributes itself over the bigger surface the further away you go in such a way that the field is solenoidal (for $r \neq 0$) meaning that there are no sinks or sources outside the origin. In contrast, the field displayed in figure FIG. xlv has an increasing flow the further away you go from the origin. This flow can of course not be fed only by a source at the origin, in fact it has sources of the same strength ($\boldsymbol{\nabla}.\mathbf{F} = 3$) at every point.

Example 5.4 *We will verify the divergence theorem for the vector field* $\mathbf{F}(\mathbf{x}) = \mathbf{x}$ *when the volume is a sphere of radius* a*, centered on the origin.*

First,

$$\int_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{F}\, dV = \int_{\mathcal{V}} \boldsymbol{\nabla}.\,\mathbf{x}\, dV = 3\int_{\mathcal{V}} dV = 3.\,\frac{4}{3}\pi a^3 = 4\pi a^3\,.$$

Second, since $\mathbf{n} = \widehat{\mathbf{x}}$ *on the surface of a sphere,*

$$\int_{\mathcal{S}} \mathbf{F}.\, d\mathbf{S} = \int_{\mathcal{S}} \mathbf{x}.\,\mathbf{n}\, dS = \int_{\mathcal{S}} r\, dS = a\int_{\mathcal{S}} dS = 4\pi a^3.$$

Example 5.5 *Let* $\mathbf{F}(\mathbf{x}) = (x, y, 0)$ *and let* $\mathcal{S}$ *be the hemispherical surface* $(x^2 + y^2 + z^2) = a^2$, $z \geq 0$. *We will use the divergence theorem to evaluate*

$$\int_{\mathcal{S}} \mathbf{F}.\, d\mathbf{S}\,.$$

[g]Just like the magnetic field of a solenoid.

The divergence theorem applies to closed surfaces, so we first close the surface by adding $\mathcal{A}$, the disc of radius a in the $z = 0$ plane. Then

$$\int_{\mathcal{V}} \boldsymbol{\nabla}.\mathbf{F}\, dV = \int_{\mathcal{S}} \mathbf{F}.\, d\mathbf{S} + \int_{\mathcal{A}} \mathbf{F}.\, d\mathbf{S}\,.$$

On $\mathcal{A}$, the surface element $d\mathbf{S} = (0, 0, -1)dxdy$ (outward pointing!), and hence $\mathbf{F}.\, d\mathbf{S} = 0$. Further, since $\boldsymbol{\nabla}.\mathbf{F} = 2$,

$$\int_{\mathcal{V}} \boldsymbol{\nabla}.\mathbf{F}\, dV = 2.\frac{2}{3}\pi a^3 = \frac{4}{3}\pi a^3\,.$$

Example 5.6 *Let $\mathbf{B}(\mathbf{x}) = (x + z, x - 2y, z)$. We will show that $\mathbf{B}$ is solenoidal and construct a vector $\mathbf{A}$ such that $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$.*

First, $\mathbf{B}$ is clearly solenoidal (i.e. divergence-free) since

$$\frac{\partial}{\partial x}(x + z) + \frac{\partial}{\partial y}(x - 2y) + \frac{\partial}{\partial z}\, z = 1 + (-2) + 1 = 0.$$

Second, we construct $\mathbf{A}$. Note that $\mathbf{A}$ is far from unique: we can add on to $\mathbf{A}$ and vector whose curl is zero, i.e. any gradient. We can use this freedom in a number of ways. One is to choose $A_z = 0$. Thus

$$\begin{aligned} -\frac{\partial A_y}{\partial z} &= x + z \\ \frac{\partial A_x}{\partial z} &= x - 2y \\ \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} &= z \end{aligned}$$

Integrating the first two of these equations gives

$$A_y = -xz - z^2/2 + f(x, y), \qquad A_x = xz - 2yz + g(x, y)$$

where $f(x, y)$ and $g(x, y)$ are arbitrary functions. Now we substitute into the the third equation:

$$-z + \frac{\partial f}{\partial x} + 2z - \frac{\partial g}{\partial y} = z.$$

The terms with z-dependence balance (as they must), and we are left to choose f and g subject to the one constraint. This shows that our choice of setting $A_z = 0$ did not use up all the remaining freedom in $\mathbf{A}$. The simplest possibility for $\mathbf{A}$ is $(xz - 2yz, -xz - z^2/2, 0)$, corresponding to $f(x, y) = g(x, y) \equiv 0$.

VI Laplace's equation

19 Laplace's equation

As an application of the theory of multidimensional integration developed in the previous chapters we will now study the solutions of a particular second order partial differential equation: the *Laplace equation*

$$\boldsymbol{\nabla}^2\Phi = 0\,, \tag{170}$$

and the corresponding inhomogeneous equation: the *Poisson equation*[a]

$$\boldsymbol{\nabla}^2\Phi = f(\mathbf{x})\,. \tag{171}$$

Both differential equations are particularly important in many areas of mathematics and physics. If Φ is for example the potential of the electric field $\mathbf{E}$ such that $\mathbf{E} = -\boldsymbol{\nabla}\Phi$ then Φ satisfies Poisson's equation

$$\boldsymbol{\nabla}^2\Phi = -\frac{\rho(\mathbf{x})}{\epsilon_0}\,, \tag{172}$$

where $\rho(\mathbf{x})$ is the charge density and ϵ_0 is the permittivity of free space. If we replace $\boldsymbol{\nabla}\Phi$ in $\boldsymbol{\nabla}^2\Phi$ by $\mathbf{E}$ then Poisson's equation can be re-written as

$$\boldsymbol{\nabla}.\mathbf{E}(\mathbf{x}) = 4\pi\rho(\mathbf{x})\,. \tag{173}$$

Considering that the divergence $\boldsymbol{\nabla}.\mathbf{E}$ of the vector field describes where the vector field has sinks or sources Eq. 173 simply says that the electric field $\mathbf{E}$ starts or ends (i.e. has sources or sinks) only at points where there are charges (i.e. the charge density $\rho(\mathbf{x}) \neq 0$).

Similarly to Eq. 172 the potential of the gravitational field $\mathbf{G}$ satisfies

$$\boldsymbol{\nabla}^2\Phi = 4\pi G\rho(\mathbf{x})\,, \tag{174}$$

where $\rho(\mathbf{x})$ is the mass density and G is the gravitational constant and therefore the gravitational field starts (i.e. has sources) only at points where there is mass.

Using Eq. 155 we can easily see that $\frac{-1}{r}$ is a solution of Laplace's equation in $\mathbb{R}^3 \setminus \{\mathbf{0}\}$. The corresponding vector field $\boldsymbol{\nabla}\Phi = \frac{\mathbf{x}}{r^3}$ is the divergence free vector field of a point charge which we have already considered in Eq. 42.

Poisson's equation is obviously a linear differential equation and Laplace's equation is the corresponding homogeneous equation. Due to the linearity of Poisson's equation the most general solution to Poisson's equation Φ is given by a *particular solution* to Poisson's equation Φ_p plus the most general homogeneous solution, which is the most general solution to Laplace's equation Ψ. In order to prove this statement we need to show two things: first $\Phi_p + \Psi$ are solutions to Poisson's equation and secondly there are no others. From $\boldsymbol{\nabla}(\Phi_p + \Psi) = \boldsymbol{\nabla}\Phi_p + \boldsymbol{\nabla}\Psi = f(\mathbf{x}) + 0$ it is obvious that $\Phi_p + \Psi$ is a solution of Poisson's equation. Conversely, if we assume that there is a solution $\hat{\Phi}$ of Eq. 171 which is not of the form $\Phi_p + \Psi$ then $\hat{\Phi} - \Phi_p$ would obviously satisfy $\boldsymbol{\nabla}(\hat{\Phi} - \Phi_p) = f(\mathbf{x}) - f(\mathbf{x}) = 0$ and therefore $\hat{\Phi} - \Phi_p$ would be a solution

[a] Simon-Denis Poisson, 1781-1840.

of Laplace's equation and is consequently included in the most general solution Ψ of Laplace's equation. But this means that $\hat{\Phi} = \Phi_p + \Psi_p$ for some Ψ_p contained in Ψ which contradicts the assumption that $\hat{\Phi}$ is not of the form $\Phi_p + \Psi$.

Since the general solution of Poisson's equation is a particular solution plus the general solution to Laplace's equation means that is now first important to analyse the general solution of Laplace's equation. Let us therefore at first concentrate on solving Laplace's equation and we will come back to Poisson's equation in Sec. 23. To start with we look at the Laplace's equation in two dimensions $\frac{\partial^2\Phi(x,y)}{\partial x^2} + \frac{\partial^2\Phi(x,y)}{\partial y^2} = 0$ and transform the Laplacian into plane polar coordinates as in Eq. 148. The Laplace equation in plane polar co-ordinates is therefore given by

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Phi(r,\phi)}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Phi(r,\phi)}{\partial\phi^2} = 0\,. \tag{175}$$

We first try to find a particular type of solutions of the form $\Phi(r,\phi) = R(r)P(\phi)$. At this stage we of course do not have any guarantee that solutions of this type exists. If such solutions do exist then we call them *separable solutions*. Substituting $\Phi(r,\phi) = R(r)P(\phi)$ into Eq. 175 leads after re-arranging to

$$\frac{r}{R(r)}\frac{\partial}{\partial r}(rR'(r)) = -\frac{P''(\phi)}{P(\phi)}\,. \tag{176}$$

The left hand side of Eq. 176 is a function of r only and the right hand side is only a function of ϕ we therefore find that the left hand side and the right hand side of Eq. 176 have to be equal to a constant $k \in \mathbb{R}$. We therefore find two separated differential equations

$$P''(\phi) = -kP(\phi)\,, \tag{177}$$

$$\frac{r}{R(r)}\frac{\partial}{\partial r}(rR'(r)) = k\,. \tag{178}$$

Eq. 178 has three possible types of solutions

$$P(\phi) = a\cos(\sqrt{k}\phi) + b\sin(\sqrt{k}\phi) \quad,\ k > 0\,, \tag{179}$$

$$P(\phi) = a + bt \quad,\ k = 0\,, \tag{180}$$

$$P(\phi) = ae^{\sqrt{-k}\phi} + be^{-\sqrt{-k}\phi} \quad,\ k < 0\,, \tag{181}$$

but considering that continuity of $P(\phi)$ requires $P(0) = P(2\pi)$ only the first of these three solutions can lead to non-trivial solutions. This also means that $k > 0$. Furthermore $P(\phi) = P(\phi + 2\pi)$ for any ϕ requires that $\sqrt{k} = n$ for some $n \in \mathbb{N}$. Therefore the only possible solutions for $P(\phi)$ are the solutions $P_n(\phi) = a\cos(n\phi) + b\sin(n\phi)$ with $k = n^2$. Substituting this back into equation Eq. 178 for the radial component $R(r)$ leads to

$$r\frac{\partial}{\partial r}(rR'(r)) = n^2R(r)\,. \tag{182}$$

Using the trial function $R(r) = cr^\alpha$ we can easily find the general solution $cr^n + dr^{-n}$. And therefore we have found separable solutions of Laplace's equation

$$\Phi_n(r,\phi) = (cr^n + dr^{-n})(a\cos(n\phi) + b\sin(n\phi))\,, \tag{183}$$

for each $n \in \mathbb{N}$. It is clear that any linear combination of these separable solutions are again solutions of Laplace's equation (which are in fact not separable). We have therefore found

infinitely many solutions to Laplace's equation. It can be shown that any solution to Laplace's equation can be written as a series expansion in terms of separable solutions:

$$\Phi(\mathbf{x}) \quad = \quad \sum_{n=0}^{\infty}(c_n r^n + d_n r^{-n})(a_n \cos(n\phi) + b_n \sin(n\phi)) \; . \tag{184}$$

This will be considered further in the Part IB course *Methods*. Ultimately we want to find solutions of Laplace's equation satisfying certain boundary conditions. For example if we were looking for a solution Φ of the two-dimensional Laplace equation which satisfies that $\Phi(\mathbf{0}) = 0$ and $\Phi = \cos\phi$ on the unit circle with ϕ being the angle of the standard plane polar co-ordinates then our earlier considerations show that the separable solution $\Phi(r, \phi) = r\cos\phi$ is a solution to this *boundary value problem*. But whether this is the unique solution to this boundary value problem remains to be investigated which will be done in Sec. 21. In exactly the same way we could look for separable solutions of the three dimensional Laplace equations using Eq. 155. Laplace's equation can of course equally well be separated in Cartesian co-ordinates. Which co-ordinates are most suitable for the separation depends mainly on what is most suitable to describe the boundary conditions.

We will now study the boundary value problem of the Laplace equation in three dimensions and then later the boundary value problem of the Poisson equation in more detail. But before we continue we will need to rewrite the divergence theorem in a form which allows us to study the properties of the Laplace operator $\boldsymbol{\nabla}^2$ more easily.

20 Green's theorems

If we are given two real valued functions (which we want to call *scalar fields*) ϕ and ψ defined on some domain G then we can construct a vector field $\mathbf{F} := \phi\boldsymbol{\nabla}\psi$. The divergence of $\mathbf{F}$ is then given by

$$\boldsymbol{\nabla}.\mathbf{F} \quad = \quad \boldsymbol{\nabla}\phi.\boldsymbol{\nabla}\psi + \phi\boldsymbol{\nabla}^2\psi \; . \tag{185}$$

We can now use the divergence theorem on $\boldsymbol{\nabla}.\mathbf{F}$ and obtain for any volume $\mathcal{V}$ which is contained in G that $\int_{\mathcal{V}} \boldsymbol{\nabla}.\mathbf{F}dV = \oint_{\partial V} \phi\boldsymbol{\nabla}\psi.d\mathbf{S}$. If we define $\mathbf{n}$ to be the normalised outward normal to the surface $\partial\mathcal{V}$ then $d\mathbf{S} = \mathbf{n}dS$. Using the directional derivatives theorem [40] we can write $\boldsymbol{\nabla}\psi.d\mathbf{S}$ as $\frac{\partial\psi}{\partial\mathbf{n}}dS$ where $\frac{\partial\psi}{\partial\mathbf{n}}$ is the directional derivative of ψ in direction $\mathbf{n}$. We obtain *Green's first theorem*:

66 **Green's first theorem**
The scalar fields $\phi, \psi : G \subset \mathbb{R}^3 \to \mathbb{R}$ satisfy for any volume $\mathcal{V} \subset G$

$$\begin{aligned}\int_{\mathcal{V}} \left(\phi\boldsymbol{\nabla}^2\psi + \boldsymbol{\nabla}\phi.\boldsymbol{\nabla}\psi\right) dV \quad &= \quad \oint_{\partial V} \phi\boldsymbol{\nabla}\psi.d\mathbf{S} \\ &= \quad \oint_{\partial\mathcal{V}} \phi\boldsymbol{\nabla}\psi.\mathbf{n}dS = \oint_{\partial\mathcal{V}} \phi\frac{\partial\psi}{\partial\mathbf{n}}dS \; ,\end{aligned} \tag{186}$$

where $\mathbf{n}$ is the normalised outward normal vector to the surface $\partial\mathcal{V}$ such that $d\mathbf{S} = \mathbf{n}dS$.

Green's first theorem can be regarded as *integration by parts* for three dimensional volume integrals in the sense that $\int_{\mathcal{V}} \boldsymbol{\nabla}\phi.\boldsymbol{\nabla}\psi\, dV$ can be written as a partially integrated two dimensional

surface integral for which $\boldsymbol{\nabla}\phi$ has been integrated minus a three dimensional volume integral over the integrated ϕ times the twice differentiated $\boldsymbol{\nabla}^2\psi$:

$$\int_{\mathcal{V}} \boldsymbol{\nabla}\phi.\boldsymbol{\nabla}\psi dV \quad = \quad \oint_{\partial\mathcal{V}} \phi\boldsymbol{\nabla}\psi.d\mathbf{S} - \int_{\mathcal{V}} \phi\boldsymbol{\nabla}^2\psi dV \; . \tag{187}$$

If we write down Green's first theorem again but now with ϕ and ψ exchanged and then subtract this equation from the original Eq. 186 we obtain *Green's second theorem.*

67 **Green's second theorem**
The scalar fields $\phi, \psi : G \subset \mathbb{R}^3 \to \mathbb{R}$ satisfy for any volume $\mathcal{V} \subset G$

$$\int_{\mathcal{V}} \left(\phi\boldsymbol{\nabla}^2\psi - \psi\boldsymbol{\nabla}^2\phi\right) dV \quad = \quad \oint_{\partial\mathcal{V}} \left(\phi\frac{\partial\psi}{\partial\mathbf{n}} - \psi\frac{\partial\phi}{\partial\mathbf{n}}\right) dS \; , \tag{188}$$

where $\mathbf{n}$ is the normalised outward normal vector to the surface $\partial\mathcal{V}$ such that $d\mathbf{S} = \mathbf{n}dS$.

21 Harmonic functions

A solution to Laplace's equation is called a *harmonic function.* In Sec. 19 we have given examples of harmonic functions and in particular we have shown that using linear combinations of separable solutions we can construct other solutions.

68 **Harmonic functions**
A function ϕ, is called harmonic on the volume $\mathcal{V}$ if it satisfies Laplace's equation on $\mathcal{V}$: $\boldsymbol{\nabla}^2\phi(\mathbf{x}) = 0 \quad \forall\mathbf{x} \in \mathcal{V}$.

Let us assume that ψ is harmonic on some volume $\mathcal{V}$. We can then apply Green's first theorem and set $\phi = \psi$ to obtain

$$\int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV \quad = \quad \oint_{\partial\mathcal{V}} \psi\frac{\partial\psi}{\partial\mathbf{n}} dS \; . \tag{189}$$

In case the boundary condition is such that $\psi(\mathbf{x}) = 0$ on the whole boundary $\partial\mathcal{V}$ then obviously $\oint_{\partial\mathcal{V}} \psi\frac{\partial\psi}{\partial\mathbf{n}} dS = 0$. This means that $\int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV = 0$ and since the continuous function $\|\boldsymbol{\nabla}\psi\|^2 \geq 0$ we have $\|\boldsymbol{\nabla}\psi\|^2 = 0 \; \forall\mathbf{x} \in \mathcal{V}$ and therefore $\psi(\mathbf{x}) = \textit{constant} \; \forall\mathbf{x} \in \mathcal{V}$. But since $\psi(\mathbf{x}) = 0$ on the boundary $\partial\mathcal{V}$ we find that the *constant* is actually 0 and therefore $\psi(\mathbf{x}) = 0 \; \forall\mathbf{x} \in \mathcal{V}$.

69 **Harmonic functions with trivial boundary conditions**
The function ψ is harmonic on the volume $\mathcal{V} \subset \mathbb{R}^3$. If $\psi(\mathbf{x}) = 0 \; \forall\mathbf{x} \in \partial\mathcal{V}$ then

$$\psi(\mathbf{x}) \quad = \quad 0 \qquad \forall\mathbf{x} \in \mathcal{V} \; . \tag{190}$$

In case we have two harmonic functions ϕ_1 and ϕ_2 on some volume $\mathcal{V}$ with the same values on $\partial\mathcal{V}$, we then find that $\psi := \phi_1 - \phi_2$ satisfies $\psi(\mathbf{x}) = 0$ for all $\mathbf{x} \in \partial\mathcal{V}$ and therefore $\psi = 0$ on the whole of $\mathcal{V}$ according to [69]. Boundary conditions which fix the value of a harmonic function on the boundary of a volume $\mathcal{V}$ are called *Dirichlet boundary conditions*[b].

[b] Johann Peter Gustav *Lejeune* Dirichlet, 1805-1859.

70 Uniqueness theorem for harmonic functions with Dirichlet boundary conditions
The functions ϕ_1 and ϕ_2 are harmonic on the volume $\mathcal{V}$ and satisfy the same Dirichlet boundary conditions $\phi_1(\mathbf{x}) = \phi_2(\mathbf{x})$ for all $\mathbf{x} \in \partial\mathcal{V}$ then

$$\phi_1(\mathbf{x}) \quad = \quad \phi_2(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{V} \, . \tag{191}$$

Since $\psi(\mathbf{x}) = c$ for $c \in \mathbb{R}$ is harmonic on any volume $\mathcal{V}$ and satisfies the constant Dirichlet boundary conditions $\psi(\mathbf{x}) = c$ on $\partial\mathcal{V}$ we therefore immediately obtain from [70] that a harmonic function which is constant on the boundary $\partial\mathcal{V}$ has to be constant on the whole of $\mathcal{V}$.

71 Constant harmonic functions
The function ψ is harmonic on the volume $\mathcal{V}$. If $\psi(\mathbf{x}) = c \; \forall \mathbf{x} \in \partial\mathcal{V}$, $c \in \mathbb{R}$ then

$$\psi(\mathbf{x}) \quad = \quad c \quad \forall \mathbf{x} \in \mathcal{V} \, . \tag{192}$$

There are other types of boundary conditions which are important in mathematics and physics:

72 Boundary conditions
A real valued function ψ defined on the volume $\mathcal{V}$ is said to satisfy

(i) Dirichlet boundary conditions if $\psi(x)$ is given as a function on the boundary $\partial\mathcal{V}$;

(ii) Neumann[c] boundary conditions if the directional derivative $\frac{\partial\psi}{\partial\mathbf{n}}$ is given on $\partial\mathcal{V}$;

(iii) boundary conditions of the mixed type if the combination $\frac{\partial\psi(\mathbf{x})}{\partial\mathbf{n}} + f(\mathbf{x})\psi(\mathbf{x})$ is given on the $\partial\mathcal{V}$.

By using Green's first theorem we can now also show in an almost identical way as for [70] that Neumann boundary conditions fix a harmonic function up to a constant.

73 Uniqueness theorem for Neumann boundary conditions
The functions ϕ_1 and ϕ_2 are harmonic on the volume $\mathcal{V}$ and satisfy the same Neumann boundary conditions $\frac{\partial\phi_1(\mathbf{x})}{\partial\mathbf{n}} = \frac{\partial\phi_2(\mathbf{x})}{\partial\mathbf{n}}$ for all $\mathbf{x} \in \partial\mathcal{V}$ then

$$\phi_1(\mathbf{x}) \quad = \quad \phi_2(\mathbf{x}) + c \quad \forall \mathbf{x} \in \mathcal{V} \, , \tag{193}$$

for some constant $c \in \mathbb{R}$.

So far we have only used the information contained in Green's first theorem when we put $\phi = \psi$ for some harmonic function ψ. Instead, we now take Green's first theorem and set $\phi = 1$ on the whole of the volume $\mathcal{V}$. We then have $\boldsymbol{\nabla}\phi = 0$ and for any harmonic function ψ we find that $\int_{\mathcal{V}} \left(\phi\boldsymbol{\nabla}^2\psi + \boldsymbol{\nabla}\phi.\boldsymbol{\nabla}\psi\right) dV = 0$ and therefore $\oint_{\partial V} \boldsymbol{\nabla}\psi.d\mathbf{S} = 0$. This does of course not mean that $\boldsymbol{\nabla}\psi$ is perpendicular to $d\mathbf{S}$ everywhere but it means that the scalar product averaged along the whole boundary is 0.

74 Harmonic average
A harmonic function ψ is harmonic on the volume $\mathcal{V}$ then it satisfies

$$\oint_{\partial\mathcal{V}} \boldsymbol{\nabla}\psi.d\mathbf{S} \quad = \quad 0 \, . \tag{194}$$

[c] *Carl* Gottfried Neumann, 1832-1925.

Using [74] we can now prove an important mean value theorem for surface integrals of harmonic functions over spheres (see example 6.4 for a proof with $\mathbf{c} = \mathbf{0}$).

75 Mean value theorem for harmonic functions
A harmonic function ψ on $\mathcal{V} \subset \mathbb{R}^3$ satisfies

$$\psi(\mathbf{c}) \quad = \quad \frac{1}{4\pi R^2} \int_{\mathcal{S}_R(\mathbf{c})} \psi(\mathbf{x}) dS \; , \tag{195}$$

where $\mathcal{S}_R(\mathbf{c})$ is the surface of a sphere with radius R centred at the point $\mathbf{c}$.

Theorem [75] essentially says that the average value of a harmonic function averaged over the surface of a sphere is the same as the function value at the centre of the sphere. This immediately shows that the global maximum and the global minimum of a harmonic function defined on the volume of a sphere has to be on the boundary of the sphere. For any connected volume $\mathcal{V}$ we can now easily argue that we can fill up the volume using spheres and therefore argue that the global maximum and the global minimum of the harmonic function will have to be on the boundary $\partial\mathcal{V}$ of the volume $\mathcal{V}$.

76 Global extrema of harmonic functions
The global maximum and the global minimum of a harmonic function ψ defined on the volume $\mathcal{V}$ lies on the boundary $\partial\mathcal{V}$ unless ψ is constant on the whole of $\mathcal{V}$.

To conclude this section we now consider a harmonic function ψ defined on a volume $\mathcal{V}$ with Dirichlet boundary conditions on $\partial\mathcal{V}$: $\psi(\mathbf{x}) = f(\mathbf{x})\ \forall \mathbf{x} \in \partial\mathcal{V}$ for some given function $f(\mathbf{x})$. Let $\mathcal{F}_f$ be the set of all (differentiable) functions, not necessarily harmonic, on $\mathcal{V}$ satisfying the same Dirichlet boundary conditions on $\partial\mathcal{V}$:

$$\mathcal{F}_f \quad := \quad \{\omega : \omega(\mathbf{x}) = f(\mathbf{x}) \ \ \forall \mathbf{x} \in \partial\mathcal{V}\} \; . \tag{196}$$

Obviously $\psi \in \mathcal{F}_f$ but because of [70] ψ is the only harmonic function contained in $\mathcal{F}_f$. Let us consider $\int_{\mathcal{V}} \|\boldsymbol{\nabla}\omega\|^2 dV$ for some $\omega \in \mathcal{F}_f$:

$$\begin{aligned}
\int_{\mathcal{V}} \boldsymbol{\nabla}\omega.\boldsymbol{\nabla}\omega dV &= \int_{\mathcal{V}} \boldsymbol{\nabla}(\omega - \psi + \psi).\boldsymbol{\nabla}(\omega - \psi + \psi) dV \\
&= \int_{\mathcal{V}} \boldsymbol{\nabla}(\omega - \psi).\boldsymbol{\nabla}(\omega - \psi) dV + 2\int_{\mathcal{V}} \boldsymbol{\nabla}(\omega - \psi).\boldsymbol{\nabla}\psi dV + \int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV \\
&\geq 2\int_{\mathcal{V}} \boldsymbol{\nabla}(\omega - \psi).\boldsymbol{\nabla}\psi dV + \int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV \\
&= 2\oint_{\partial\mathcal{V}} (\omega - \psi)\frac{\partial\psi}{\partial\mathbf{n}} dS - 2\int_{\mathcal{V}} (\omega - \psi)\boldsymbol{\nabla}^2\psi dV + \int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV \; ,
\end{aligned} \tag{197}$$

using Green's first theorem for the last equation. But since $\omega = \psi$ on the boundary $\partial\mathcal{V}$ and $\boldsymbol{\nabla}^2\psi = 0$ on the whole of $\mathcal{V}$ we find that the first two integrals of the last expression in Eq. 197 vanish.

77 Harmonic functions with the same Dirichlet boundary conditions
Among all (differentiable) functions $\mathcal{F}_f$ on $\mathcal{V}$ with the same Dirichlet boundary conditions $f(\mathbf{x})$ on $\partial\mathcal{V}$ as the harmonic function ψ we find that ψ minimises the integral over the norm of the gradient:

$$\int_{\mathcal{V}} \|\boldsymbol{\nabla}\omega\|^2 dV \geq \int_{\mathcal{V}} \|\boldsymbol{\nabla}\psi\|^2 dV \; , \tag{198}$$

for all functions $\omega \in \mathcal{F}_f$.

22 Gauss' flux theorem and Gauss' law

We will now return to the study of the inhomogeneous problem, namely the study of Poisson's equation $\boldsymbol{\nabla}^2\phi(\mathbf{x}) = f(\mathbf{x})$ on some domain G. If we assume that ϕ is such a solution then the vector field $\mathbf{F} := \boldsymbol{\nabla}\phi$ satisfies

$$\boldsymbol{\nabla}.\mathbf{F} \quad = \quad f(\mathbf{x}) \tag{199}$$

on any volume $\mathcal{V}$ contained in G. Integrating over Eq. 199

$$\int_{\mathcal{V}} \boldsymbol{\nabla}.\mathbf{F} dV \quad = \quad \int_{\mathcal{V}} f(\mathbf{x}) dV \tag{200}$$

and applying the divergence theorem to the first of these integrals leads to *Gauss' flux theorem*[d].

78 **Gauss' flux theorem**
The function ϕ is a solution to Poisson's equation $\boldsymbol{\nabla}^2\phi(\mathbf{x}) = f(\mathbf{x})$ on the domain G if and only if the vector field $\mathbf{F} := \boldsymbol{\nabla}\phi$ satisfies

$$\int_{\partial\mathcal{V}} \mathbf{F}.d\mathbf{S} \quad = \quad \int_{\mathcal{V}} f(\mathbf{x}) dV \ , \tag{201}$$

for any (piecewise smooth) volume $\mathcal{V} \subset G$.

We can interpret $f(\mathbf{x})$ as the sinks and sources of the vector field $\mathbf{F}$ since it corresponds to the divergence of $F(\mathbf{x})$. In other words $f(\mathbf{x})$ describes the density at which points the field starts and ends. For example, for the electric field $f(\mathbf{x})$ describes the charge density and for the gravitational field it describes the mass density. Gauss' flux theorem essentially says that *summing over the charge density on the whole volume equals the flux of the vector field through the boundary.*

We already know that $\frac{1}{r}$ is harmonic on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$. If we define the corresponding vector field $\mathbf{F} = \boldsymbol{\nabla}\frac{1}{r} = -\frac{\mathbf{x}}{r^3}$ then we can apply Gauss' flux theorem on some volume $\mathcal{V}$ which does not contain the origin. We then obtain $\int_{\partial\mathcal{V}} \mathbf{F}.d\mathbf{S} = 0$. But if $\mathbf{0}$ is inside the volume $\mathcal{V}$ then we take a small sphere around $\mathbf{0}$ with radius ϵ which we call $\mathcal{V}_\epsilon$ and define $\hat{\mathcal{V}}$ such that $\mathcal{V} = \hat{\mathcal{V}} \cup \mathcal{V}_\epsilon$ and $\hat{\mathcal{V}}$, $\mathcal{V}_\epsilon$ disjoint with the exception of boundary points. We then find

$$\begin{aligned}
\oint_{\partial\mathcal{V}} \mathbf{F}.d\mathbf{S} \quad &= \quad \oint_{\partial\hat{\mathcal{V}}} \mathbf{F}.d\mathbf{S} + \oint_{\partial\mathcal{V}_\epsilon} \mathbf{F}.d\mathbf{S} \\
&= \quad 0 + \oint_{\partial\mathcal{V}_\epsilon} \mathbf{F}.d\mathbf{S} \\
&= \quad -\int_0^{2\pi}\int_0^{\pi} \frac{\mathbf{x}}{\epsilon^3}.\epsilon^2\mathbf{e}_r \sin\theta \, d\theta d\phi = -4\pi \ .
\end{aligned} \tag{202}$$

The same will obviously hold if we replace $\mathbf{0}$ by any other point $\mathbf{x}_0$. This result is known as *Gauss' law.*

79 **Gauss' law**
For any volume $\mathcal{V}$ and any $\mathbf{x}_0 \in \mathbb{R}^3$ with $\mathbf{x}_0 \notin \partial\mathcal{V}$

$$\oint_{\partial\mathcal{V}} \frac{\mathbf{x} - \mathbf{x}_0}{\|\mathbf{x} - \mathbf{x}_0\|^3}.d\mathbf{S} \quad = \quad \begin{cases} 4\pi & , \mathbf{x}_0 \in \mathcal{V} \\ 0 & , \mathbf{x}_0 \notin \mathcal{V} \end{cases} \ . \tag{203}$$

[d] Conversely, if the identity of the integrals Eq. 200 holds for all possible volumes $\mathcal{V} \subset G$ then obviously the (continuous) integrands have to be the same and therefore Eq. 199 holds on G.

Gauss' law makes it clear that $\frac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|^3}$ is the field of a point charge of strength 4π at the point $\mathbf{x}_0$.

23 Poisson's equation

Using the results of the previous section we are now in a position to explore the solutions to Poisson's equation. We have found that the potential $\phi = -\frac{1}{4\pi\|\mathbf{x}-\mathbf{x}_0\|}$ for the field $\mathbf{F} = \frac{1}{4\pi}\frac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|^3}$ corresponds to a point charge of strength 1 at the point $\mathbf{x}$. Using the superposition principle for linear differential equations, if we had a charge of strength α at the point $\mathbf{a}$ and a charge of strength β at the point $\mathbf{b}$ the resulting solution would be just the sum

$$\mathbf{F}(\mathbf{x}) = \frac{\alpha}{4\pi}\frac{\mathbf{x}-\mathbf{a}}{\|\mathbf{x}-\mathbf{a}\|^3} + \frac{\beta}{4\pi}\frac{\mathbf{x}-\mathbf{b}}{\|\mathbf{x}-\mathbf{b}\|^3}\,, \tag{204}$$

with the corresponding potential

$$\phi(\mathbf{x}) = -\frac{\alpha}{4\pi\|\mathbf{x}-\mathbf{a}\|} - \frac{\beta}{4\pi\|\mathbf{x}-\mathbf{b}\|}\,. \tag{205}$$

Adding up finitely many charges of strengths α_i at the positions $\mathbf{a}_i$ we would obviously find

$$\phi(\mathbf{x}) = -\sum_i \frac{\alpha_i}{4\pi\|\mathbf{x}-\mathbf{a}_i\|}\,. \tag{206}$$

In case the charges are not located at finitely many points but are described by a charge density we expect that a suitable limit process will result in the sum in Eq. 206 to be replaced by an integral and the charge strengths α_i will go over to the charge density $f(\mathbf{x})$. We will give this result without proof but it can in fact easily be shown by substituting Eq. 208 into Gauss' flux theorem Eq. 201, exchanging the order of integration and using Gauss' law. We then obtain a solution to Poisson's equation with $\phi(\mathbf{x}) \to 0$ as $\mathbf{x} \to \infty$ (see the third method in example 6.5).

80 **Solutions to Poisson's equation**
For Poisson's equation $\boldsymbol{\nabla}^2\phi = f(\mathbf{x})$ defined on some finite volume $\mathcal{V}$ we can set $f(\mathbf{x}) = 0$ outside the volume $\mathcal{V}$ and obtain a solution[e]

$$\phi(\mathbf{x}) = -\frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{f(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|}\,d(y_1,y_2,y_3)\,, \tag{207}$$

where the integral is taken over the whole $\mathbb{R}^3$. This solution satisfies the boundary conditions $\phi(\mathbf{x}) \to 0$ as $\mathbf{x} \to \infty$. The corresponding vector field is given by

$$\mathbf{F}(\mathbf{x}) = \frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{f(\mathbf{y})(\mathbf{x}-\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|^3}\,d(y_1,y_2,y_3)\,. \tag{208}$$

[80] gives of course only one particular solution with the boundary conditions $\phi(\mathbf{x}) \to 0$ as $\mathbf{x} \to \infty$. If a specific problem needs to be solved under different boundary conditions then solutions of Laplace's equation have to be added (as solutions of the homogeneous problem) in order to achieve the required boundary conditions.

[e] Note that $\int_{\mathbb{R}^3}$ is in fact only an integral $\int_{\mathcal{V}}$.

Examples

Example 6.1 *Maxwell's equations for electromagnetic fields are*

$$\boldsymbol{\nabla}.\mathbf{E} = \frac{\rho}{\epsilon_o} \tag{209}$$

$$\boldsymbol{\nabla}.\mathbf{B} = 0 \tag{210}$$

$$\boldsymbol{\nabla}\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \tag{211}$$

$$\boldsymbol{\nabla}\times\mathbf{B} = \epsilon_o\mu_o\frac{\partial\mathbf{E}}{\partial t} + \mu_o\mathbf{j} \tag{212}$$

Here $\mathbf{E}$ *and* $\mathbf{B}$ *are the electric and magnetic fields;* ρ *and* $\mathbf{j}$ *are electric charge density and current density;* ϵ_o *and* μ_o *are constants (the permitivity and permissivity of free space) and they satisfy* $\epsilon_o\mu_o = c^{-2}$*, where* c *is the speed of light. (This last equation shows that there is a deep connection between electromagnetic fields and light.) All fields may depend on both* $\mathbf{x}$ *and time* t*.*

Note that there is some symmetry between the equations. Equations (1) and (2) are similar except that (2) has no source on the right hand side. This is as expected, because the magnetic charge density is zero (there are no magnetic monopoles, at least in classical theory). Similarly, equations (3) and (4) are the similar, expect for the absence of magnetic current in (3), and an extra minus sign.

It is worth pausing for a moment to investigate the integral forms of these equations; it is a nice example of the use of the divergence theorem and Stokes' theorem and relates the equations to physics that you might know, but it is not essential to the example.

If we integrate (1) over a volume V *with surface* S*, we get*

$$\int_V \boldsymbol{\nabla}.\mathbf{E}\; dV = \frac{1}{\epsilon_o}\int_V \rho(\mathbf{x})dV \qquad i.e. \qquad \int_S \mathbf{E}.d\mathbf{S} = Q/\epsilon_o$$

or, in words, the flux of electric field across any closed surface is equal to the total charge within the surface over ϵ_o*. This is Gauss's law. The same calculation for equation (2) shows that the total flux of magnetic field across a closed surface is zero.*

If we integrate (3) over an open surface S *with boundary curve* C*, we get*

$$\int_S \boldsymbol{\nabla}\times\mathbf{E}.d\mathbf{S} = -\int_S \frac{\partial\mathbf{B}}{\partial t}.d\mathbf{S} \qquad i.e. \qquad \int_C \mathbf{E}.d\mathbf{x} = -\frac{d}{dt}\int_S \mathbf{B}.d\mathbf{S}.$$

In words, this says that the circulation of the electric field (which is called the the EMF or electromotive force) round a close curve is equal to the rate of change of flux through the curve. This Faraday's law of induction (and the minus sign relates to Lenz's Law, about the effect tending to oppose the cause).

Finally, if we integrate (4) over an open surface S *with boundary curve* C*, ignoring the first term on the right hand side (which would be a valid approximation for situations in which* c^{-2} *is comparatively small), we get*

$$\int_S \boldsymbol{\nabla}\times\mathbf{B}.d\mathbf{S} = \mu_o\int_S \mathbf{j}.d\mathbf{S} \qquad i.e. \qquad \int_C \mathbf{B}.d\mathbf{x} = \mu_o I.$$

In words, this says that the circulation of magnetic field round a loop is equal to the μ_o *times the current through the loop. This is Ampère's law.*

Going back to the original equations (1)—(4), we see that the structure of these equations allows potentials to be constructed. The easiest case is when the fields are time independent (all time derivatives are zero). In this case, equation (3) becomes $\boldsymbol{\nabla} \times \mathbf{E} = \mathbf{0}$*, which implies that there exists a scalar potential* ϕ *such that*

$$\mathbf{E} = -\boldsymbol{\nabla}\phi$$

(the minus sign is conventional) and equation (1) then shows that ϕ *satisfies Poisson's equation*

$$\boldsymbol{\nabla}.(-\boldsymbol{\nabla}\phi) = -\nabla^2\phi = \frac{\rho}{\epsilon_o}$$

This is useful, because a great deal is known about solutions of Poisson's equation.

If the fields are time-dependent, there is no scalar potential for $\mathbf{E}$*, since* $\boldsymbol{\nabla} \times \mathbf{E} \neq \mathbf{0}$*. However, equation (2) shows that we can always find a vector potential for* $\mathbf{B}$*:*

$$\boldsymbol{\nabla}.\mathbf{B} = 0 \Rightarrow \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}.$$

Of course, $\mathbf{A}$ *is not uniquely determined: we can still add the gradient of any scalar field without changing* $\mathbf{B}$*. We can use this freedom is to choose* $\mathbf{A}$ *to satisfy* $\boldsymbol{\nabla}.\mathbf{A} = 0$*. (If* $\boldsymbol{\nabla}.\mathbf{A} \neq 0$*, we need to add* $\boldsymbol{\nabla}\chi$ *where* χ *satisfies* $\boldsymbol{\nabla}.(\mathbf{A} + \boldsymbol{\nabla}\chi) = 0$*, i.e.* $\nabla^2\chi = -\boldsymbol{\nabla}.\mathbf{A}$*. This is Poisson's equation, and, for any given* $\boldsymbol{\nabla}.\mathbf{A}$*, there is always a solution for* χ*.)*

Thus equation (3) becomes

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial(\boldsymbol{\nabla} \times \mathbf{A})}{\partial t} \qquad i.e. \qquad \boldsymbol{\nabla} \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0.$$

This means that we can find a scalar potential for $\mathbf{E} + \dfrac{\partial \mathbf{A}}{\partial t}$ *:*

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\boldsymbol{\nabla}\phi \qquad or \qquad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi$$

Why are these potentials useful? If we write $\mathbf{E}$ *and* $\mathbf{B}$ *in terms of potentials, then Maxwell equations (2) and (3) are automatically satisfied. The remaining equations become*

$$\boldsymbol{\nabla}.\left(-\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi\right) = \frac{\rho}{\epsilon_o} \qquad i.e. \qquad \nabla^2\phi = -\frac{\rho}{\epsilon_o}$$

(using the gauge condition $\boldsymbol{\nabla}.\mathbf{A} = 0$*), and*

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \epsilon_o\mu_o\frac{\partial}{\partial t}\left(-\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi\right) + \mu_o\mathbf{j}.$$

In the absence of sources, i.e. $\rho = 0$ *and* $\mathbf{j} = 0$*, the potential* $\phi = 0$ *and this last equation reduces to*

$$\nabla^2\mathbf{A} = \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2}.$$

This is the wave equation; it says that electromagnetic fields propagate like waves with speed c *— i.e. with the speed of light.*

Example 6.2 *We will prove a uniqueness theorem for solutions for the differential equation*

$$y'' - y = f(x),$$

with Dirichlet boundary conditions

$$y(0) = a; \quad y(1) = b.$$

The method used is the basis for all other uniqueness theorems.

Suppose that the above boundary value problem has two solutions, $y_1(x)$ and $y_2(x)$, and let $Y(x) = y_2(x) - y_1(x)$. Then $Y(x)$ satisfies the equation $Y'' - Y = 0$ and the boundary conditions $Y(0) = Y(1) = 0$. We will show that $Y(x) \equiv 0$.

The technique is to show that, if $Y(x) \neq 0$, then a positive quantity is negative. The positive quantity, I, is given by

$$I = \int_0^1 (Y')^2 dx.$$

Note that

$$I = 0 \Leftrightarrow Y'(x) = 0 \Leftrightarrow Y(x) = \text{ constant} \Leftrightarrow Y(x) = 0.$$

The last implication follows because $Y(0) = 0$, so if $Y(x)$ is constant, its value must be zero.

To obtain the contradiction, we integrate by parts:

$$I = \int_0^1 Y'(x)Y'(x)dx = Y(x)Y'(x)\Big|_0^1 - \int_0^1 Y(x)Y''(x)dx = -\int_0^1 Y(x)^2 dx \leq 0\,.$$

For the last equality we have used the boundary conditions and the differential equation for Y. Thus $I \leq 0$, which is the required contradiction.

Example 6.3 *We will find separable solutions of Laplace's equation of the form $f(x, y) = X(x)Y(y)$.*

Substituting into Laplace's equation

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

gives

$$X''(x)Y(y) + X(x)Y''(y) = 0\,,$$

or, after rearrangement,

$$\underbrace{\frac{X''(x)}{X(x)}}_{\text{function of } x} = \underbrace{-\frac{Y''(y)}{Y(y)}}_{\text{function of } y} = C\,,$$

where C is a constant. This is the critical argument: the only function of x that equals a function of y is a constant. The constant C is not undetermined so far, but will be determined by boundary conditions.

There are three cases to consider.

(i): $C = 0$.

In this case

$$X''(x) = Y''(y) = 0 \quad \Rightarrow \quad X(x) = ax + b \quad \text{and} \quad Y(y) = cy + d\,,$$

where a, b, c and d are constants, i.e.

$$f(x, y) = (ax + b)(cy + d)\,.$$

(ii) : $C = k^2 > 0$.

In this case

$$X'' - k^2 X = 0\,, \quad \textit{and} \quad Y'' + k^2 Y = 0\,.$$

Hence

$$X(x) = ae^{kx} + be^{-kx}\,, \quad \textit{and} \quad Y(y) = c\sin ky + d\cos ky\,,$$

where a, b, c and d are constants, i.e.

$$f(x,y) = (ae^{kx} + be^{-kx})(c\sin ky + d\cos ky)\,.$$

(iii) $C = -k^2 < 0$.

This is similar to (ii) above, with $x \leftrightarrow y$:

$$f(x,y) = (ae^{ky} + be^{-ky})(c\sin kx + d\cos kx)\,.$$

The question of which of the above solutions (or which linear combination of them) is appropriate is decided by the boundary conditions. Obviously, if the solution is periodic in x, or has to vanish at two values of x, we must have $C < 0$ so that the solutions are of the form (iii).

Example 6.4 *We will prove a Mean Value Theorem for harmonic functions. (Note that a harmonic function is one that satisfies Laplace's equation.)*

Let the scalar field φ be harmonic in a volume $\mathcal{V}$ bounded by a closed surface $\mathcal{S}$. Consider the function $f(r)$ defined to be the mean value of φ on a spherical surface $\mathcal{S}_r$ given by $\|\mathbf{x}\| = r$, i.e.

$$\begin{aligned} f(r) &= \frac{1}{4\pi r^2}\int_{\mathcal{S}_r} \varphi(\mathbf{x})\,\mathrm{d}S \\ &= \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} \varphi(r,\theta,\phi)\sin\theta\,d\phi\,d\theta\,, \qquad (*) \end{aligned}$$

where (r,θ,ϕ) are spherical polar co-ordinates. We will show that f is a constant, and that $f(r) = \varphi(\mathbf{0})$.

We can show this result either by using Green's Second Theorem, or directly as follows:

$$\begin{aligned} \frac{df}{dr} &= \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} \frac{\partial\varphi}{\partial r}\sin\theta\,d\phi\,d\theta \\ &= \frac{1}{4\pi r^2}\int_{\mathcal{S}_r} \frac{\partial\varphi}{\partial r}\,\mathrm{d}S \\ &= \frac{1}{4\pi r^2}\int_{\mathcal{S}_r} \boldsymbol{\nabla}\varphi.\,\hat{\mathbf{x}}\,\mathrm{d}S \\ &= \frac{1}{4\pi r^2}\int_{\mathcal{S}_r} \boldsymbol{\nabla}\varphi.\,\mathrm{d}\mathbf{S} \\ &= \frac{1}{4\pi r^2}\int_{\mathcal{V}} \nabla^2\varphi\,\mathrm{d}V \\ &= 0\,. \end{aligned}$$

Further, setting $r = 0$ in $(*)$ shows that

$$f(0) = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi} \varphi(0,\theta,\phi)\sin\theta\,d\phi\,d\theta = \varphi(\mathbf{0})\,,$$

and hence, since $f(r)$ is constant,

$$\varphi(\mathbf{0}) = f(0) = f(r) = \frac{1}{4\pi r^2}\int_{\mathcal{S}_r} \varphi(\mathbf{x})\,\mathrm{d}S\,.$$

We conclude 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.*

Example 6.5 *We will obtain the gravitational field due to a sphere $\|\mathbf{x}\| = R$ of uniform density ρ_0, and total mass $M = \frac{4}{3}\pi R^3\rho_0$ in two ways.*

First way: solve Poisson's Equation.

We have

$$\nabla^2\varphi = \begin{cases} 4\pi G\rho_0 & \text{for} \quad r < R \\ 0 & \text{for} \quad r > R. \end{cases}$$

We assume (or could prove) that φ and $\frac{\partial\varphi}{\partial n}$ are continuous at $r = R$. From symmetry considerations we anticipate (or we could prove, e.g. by seeking separable solutions) that φ will be a function only of r, i.e. $\varphi \equiv \varphi(r)$.

Using the expression for the Laplacian in spherical polar co-ordinates, for $r < R$ we have

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\varphi}{dr}\right) = 4\pi G\rho_0\,,$$

which we can integrate:

$$r^2\frac{d\varphi}{dr} = \tfrac{4}{3}\pi G\rho_0 r^3 + a\,,$$

and

$$\varphi = \tfrac{2}{3}\pi G\rho_0 r^2 - \frac{a}{r} + b\,,$$

where a and b are constants. Since there is no point mass at the origin, φ is bounded there, and hence $a = 0$.

Similarly, for $r > R$ we deduce that

$$\varphi = -\frac{d}{r} + c\,,$$

where c and d are constants. Since the potential φ is only defined up to a constant, we can set $c = 0$; this means that $\varphi(r) \to 0$ as $r \to \infty$.

Now we match the interior and exterior solutions at $r = R$ to find b and d:

$$\tfrac{2}{3}\pi G\rho_0 R^2 + b = -\frac{d}{R} \qquad \text{and} \qquad \tfrac{4}{3}\pi G\rho_0 R = \frac{d}{R^2}\,,$$

so

$$d = \tfrac{4}{3}\pi G\rho_0 R^3 \qquad \text{and} \qquad b = -2\pi G\rho_0 R^2\,.$$

Thus

$$\varphi = -\frac{4\pi G\rho_0 R^3}{3r}\,, \qquad \mathbf{g} = -\boldsymbol{\nabla}\varphi = -\frac{4\pi G\rho_0 R^3}{3r^2}\,\widehat{\mathbf{x}} = -\frac{GM}{r^2}\,\widehat{\mathbf{x}} \quad \text{for} \quad r \geq R\,,$$

$$\varphi = \tfrac{2}{3}\pi G\rho_0(r^2 - 3R^2)\,, \qquad \mathbf{g} = -\boldsymbol{\nabla}\varphi = -\tfrac{4}{3}\pi G\rho_0 r\,\widehat{\mathbf{x}} = -\frac{GMr}{R^3}\,\widehat{\mathbf{x}} \quad \text{for} \quad r \leq R\,.$$

Note that 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 centre.

Second way: use Gauss' Flux Theorem.

From symmetry consideration we anticipate that $\mathbf{g}$ *will be a function only of* r*, and will be parallel to* $\hat{\mathbf{x}}$*, i.e.* $\mathbf{g} = g(r)\hat{\mathbf{x}}$*. Suppose* $\mathcal{S}$ *is a spherical surface of radius* r*. Then*

$$\int_{\mathcal{S}} \mathbf{g}.\, d\mathbf{S} = \int_{\mathcal{S}} g(r)\, \hat{\mathbf{x}}.\, d\mathbf{S} = g(r)\, 4\pi r^2 \,.$$

Further,

$$\begin{aligned} &\textit{if} \quad r \geq R \qquad \textit{then} \qquad \int_{\mathcal{V}} \rho\, dV = \tfrac{4}{3}\pi R^3 \rho_0 = M \,, \\ &\textit{if} \quad r \leq R \qquad \textit{then} \qquad \int_{\mathcal{V}} \rho\, dV = \tfrac{4}{3}\pi r^3 \rho_0 = M r^3/R^3 \,. \end{aligned}$$

Hence from Gauss' flux theorem

$$g(r) = \begin{cases} -GM/r^2 & r \geq a \\ -GMr/R^3 & r \leq a \end{cases} \,.$$

We can obtain $\varphi(r)$ *by integrating* $g(r) = -d\varphi/dr$*, using the boundary condition* $\varphi = 0$ *at* $r = \infty$ *and continuity of* $\varphi(r)$ *at* $r = R$*.*

Third way: use the general solution of Poisson's equation.

The potential at a fixed point $\mathbf{X}$ *due to a mass distribution* $\rho(\mathbf{x})$ *in volume* V *is given by*

$$\varphi(\mathbf{X}) = G \int_{\mathcal{V}} \frac{\rho(\mathbf{x})}{\|\mathbf{X} - \mathbf{x}\|}\, dV \,.$$

We choose polar coordinates with pole direction $\mathbf{X}$*, so that* $\|\mathbf{X} - \mathbf{x}\| = (d^2 + r^2 - 2rd\cos\theta)^{\frac{1}{2}}$*, where* $d = \|\mathbf{X}\|$ *and* (r, θ, ϕ) *are the polar coordinates of the point* $\mathbf{x}$*. Then*

$$\begin{aligned} \varphi(\mathbf{X}) &= G \int_0^R \int_0^\pi \int_0^{2\pi} \frac{\rho_0}{(d^2 + r^2 - 2rd\cos\theta)^{\frac{1}{2}}}\, r^2 \sin\theta\, d\phi\, d\theta\, dr \\ &= 2\pi G\rho_0 \int_0^R \int_0^\pi \frac{r^2 \sin\theta}{(d^2 + r^2 - 2rd\cos\theta)^{\frac{1}{2}}} d\theta\, dr \\ &= 2\pi G\rho_0 \int_0^R \frac{r}{d}\, (|d + r| - |d - r|)\, dr \,. \end{aligned}$$

If $d > R$ *(i.e. if* $\mathbf{X}$ *is outside the massive sphere), then the integrand is* $2r^2/d$*, which gives* $\varphi(d) = GM/d$*. If* $d < R$*, then the integrand is* $2r^2/d$ *if* $r < d$ *but* $2r$ *if* $r > d$*. In this case,*

$$\varphi(d) = 2\pi G\rho_0 \int_0^d \frac{2r^2}{d} dr + 2\pi G\rho_0 \int_d^R 2r\, dr$$

which gives the required result.

Remark *Similar results hold for the electric field due to a uniformly charged sphere.*

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,*

$$\mathbf{E} = \frac{q_1}{4\pi\varepsilon_0} \frac{\mathbf{x} - \mathbf{x}_1}{\|\mathbf{x} - \mathbf{x}_1\|^3},$$

where ε_0 *is the permittivity of free space. Thus we can read off the equivalent results for electrostatics by means of the transformations*

$$\mathbf{g} \to \mathbf{E}, \quad m_j \to q_j, \quad G \to -\frac{1}{4\pi\varepsilon_0}.$$

For instance, Gauss' flux theorem becomes

$$\int_{\mathcal{S}} \mathbf{E}.\, d\mathbf{S} = \frac{1}{\varepsilon_0} \int_{\mathcal{V}} \rho \, dV,$$

where ρ *is the now the charge density. The differential form of this equation is*

$$\boldsymbol{\nabla}.\mathbf{E} = \frac{\rho}{\varepsilon_0},$$

which is one of Maxwell's equations. Further, if we introduce the electric potential φ*, where*

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi,$$

then

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0},$$

with solution in unbounded space

$$\varphi(\mathbf{X}) = \frac{1}{4\pi\varepsilon_0} \int_{\mathcal{V}} \frac{\rho(\mathbf{x})}{\|\mathbf{X} - \mathbf{x}\|} \, dV.$$

VII Cartesian Tensors in $\mathbb{R}^3$

When vectors were introduced in Part IA Algebra and Geometry you simply defined them as elements of a vector space. For example,

$$\mathbf{v} = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$$

is a vector in $\mathbb{R}^3$. At this stage it is simply a column of 3 numbers in the set $\mathbb{R}^3$ and $\mathbb{R}^3$ satisfies the usual vector space axioms (i.e. we know how to add the vectors and how to multiply them by real numbers and the usual associativity, commutativity and distributivity axioms hold). At this stage it is not necessary to introduce any basis of $\mathbb{R}^3$ or even to interpret the numbers -1, 5 and 2 as the components of $\mathbf{v}$ with respect to any co-ordinate axes: the vector $\mathbf{v}$ is simply the column of these numbers. But then you realised that introducing a suitable basis in a vector space may simplify linear problems in a vector space significantly (for example, matrices can sometimes be diagonalised and therefore a matrix equation could in this way be simplified enormously). A vector always has a unique basis description in a given basis B (the uniqueness of vector components in a given basis is a result of the fact that a basis is linearly independent; the fact that it is always possible to find components of a vector with respect to a given basis is a consequence of the basis being a spanning set). For example, the vector $\mathbf{v}$ considered above has the components

$$\begin{pmatrix} -1 \\ 6 \\ -3 \end{pmatrix}$$

in the basis $B = \{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\}$ because

$$\begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 6\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{213}$$

We call $\begin{pmatrix} -1 \\ 6 \\ -3 \end{pmatrix}$ the *component vector* of $\mathbf{v}$ in the basis B. In a different basis the component vector of $\mathbf{v}$ would obviously be different. For example in the basis $E = \{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\}$ the same vector $\mathbf{v}$ has the components (or the component vector)

$$\begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}.$$

The basis E is special in the sense that the component vectors of vectors with respect to this basis actually coincide with the vectors themselves. Such a basis is called the *standard basis* or the *canonical basis*. You will discover that Mathematicians are sometimes quite sloppy and relaxed when they talk about vectors: they say the vector $\mathbf{v}$ and actually mean the vector

components with respect to some basis, or conversely they say the components and actually mean the vector itself. Without becoming too philosophical we should keep in mind that the component vectors and the vectors themselves are conceptually of course different objects even if they actually coincide for a standard basis.

In Algebra and Geometry you also studied the transformation properties of vectors (or better of the vector components) under basis transformations. If we are given a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ of $\mathbb{R}^3$ and change this basis to a new basis $B' = \{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ then we can find a basis transformation matrix L such that

$$\mathbf{b}_i = L_{ji}\mathbf{b}'_j \quad \text{and} \quad \mathbf{b}'_i = L^{-1}_{ji}\mathbf{b}_j \; . \tag{214}$$

Eq. 214 simply defines L to be the matrix with the components of the vectors $\mathbf{b}_i$ in terms of the basis B' written in its columns. If x_i and x'_i denote the components of a vector $\mathbf{x} \in \mathbb{R}^3$ respectively: $\mathbf{x} = x_i\mathbf{b}_i = x'_i\mathbf{b}'_i$ then you have shown in Algebra and Geometry that with this L we find

$$x'_i = L_{ij}x_j \; . \tag{215}$$

This simply comes from the fact that the vector $\mathbf{x}$ itself is of course independent from the chosen basis vectors and therefore the change in the basis automatically determines the change in the vector components such that the vector $\mathbf{x}$ stays the same:

$$\mathbf{x} = x_j\mathbf{b}_j \quad = \quad x_jL_{ij}\mathbf{b}'_i \; , \tag{216}$$

but also

$$\mathbf{x} \quad = \quad x'_i\mathbf{b}'_i \tag{217}$$

and using the uniqueness of vector components we therefore find

$$x'_i = L_{ij}x_j \; . \tag{218}$$

This defines how vector components transform under basis transformations. Conversely, components with respect to a basis following the transformation rule Eq. 218 under basis transformation obviously define a vector $\mathbf{x} \in \mathbb{R}^3$. Therefore, instead of defining vectors and then deriving their transformation properties under basis transformations we could have equally well turned the argument round and defined vectors to be exactly the objects which transform according to Eq. 218 under the basis transformation described by L. The difference of this way of thinking about vectors is that we now do not need to worry about the vectors $\mathbf{x}$ themselves but define them with what we are most interested about vectors: how are they represented in a given co-ordinate system (i.e. a basis). So, we think about vectors as objects represented in a given co-ordinate system (basis) by numbers x_i and when changing the co-ordinate system the objects transform exactly under Eq. 218.

Example 7.1 *Consider the two bases (which happen to be orthonormal bases) of* $\mathbb{R}^3$*:*

$$B \quad = \quad \left\{ \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{-1}{2}\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \; , \tag{219}$$

$$B' \quad = \quad \left\{ \begin{pmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix} \right\} \; . \tag{220}$$

We can write the vectors of B *in terms of* B'*:*

$$\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{-1}{2}\sqrt{2} \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix} - \frac{1}{2}\sqrt{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \frac{-1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix} + \frac{1}{2}\sqrt{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \frac{-1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix},$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2}\sqrt{2}\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2}\sqrt{2}\begin{pmatrix} \frac{-1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix}.$$

We now put these components into the columns of the matrix L

$$L = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2}\sqrt{2} \end{pmatrix}. \tag{221}$$

The components L_{ij} *of this matrix* L *are now determining the transformation of vector components as in* Eq. 218. *Since* L *is a transformation from an orthonormal basis to an orthonormal basis you have found in Algebra and Geometry that the inverse matrix* L^{-1} *is in fact just the transpose* L^T *of the matrix* L *and therefore* $L^{-1}_{ij} = L_{ji}$ *in this particular case.*

A vector $\mathbf{x}$ *with the components* $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ *in the basis* B *would therefore have the components*

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = L\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \tag{222}$$

in the basis B' *as given by* Eq. 218. *The underlying vector* $\mathbf{x} \in \mathbb{R}^3$ *is in fact in this case* $\begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix}$ *as can be seen from*

$$\begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix} = 1\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{-1}{2}\sqrt{2} \\ 0 \end{pmatrix} + 1\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{223}$$

and also from

$$\begin{pmatrix} \sqrt{2} \\ 0 \\ 0 \end{pmatrix} = 1\begin{pmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} \frac{-1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{pmatrix}. \tag{224}$$

We now want to explore this idea a bit further. In Algebra and Geometry you also studied how linear maps $\alpha(\mathbf{x}) : \mathbb{R}^3 \to \mathbb{R}^3$ can be represented by a matrix A in the basis B. If

$$\alpha(\mathbf{b}_i) = A_{ji}\mathbf{b}_j \tag{225}$$

then the components of $\alpha(\mathbf{x})$ in the basis B are given by $A_{ij}x_j$ where x_j are the components of $\mathbf{x}$ in B. This is again a straightforward consequence of the uniqueness of components in a

basis and the linearity of α: $\alpha(\mathbf{x}) = \alpha(x_i\mathbf{b}_i) = x_i\alpha(\mathbf{b}_i) = x_iA_{ji}\mathbf{b}_j = A_{ji}x_i\,\mathbf{b}_j$. Note that the definition $\alpha(\mathbf{b}_i) = A_{ji}\mathbf{b}_j$ of the matrix A simply means that the i-th column[a] of the matrix A are the components of the image $\alpha(\mathbf{b}_i)$ of the vector $\mathbf{b}_i$ in the basis B.

In the basis B' the linear map α would then be represented by a matrix A' which is defined in the same way as A but using the basis B': $\alpha(\mathbf{b}'_i) = A'_{ji}\mathbf{b}'_j$. In Algebra and Geometry you then found for the transformation properties[b] of the matrix that $A' = LAL^{-1}$ or in suffix notation

$$A'_{il} = L_{ij}A_{jk}L^{-1}_{kl} \; . \tag{226}$$

Conversely, an object transforming in this way under basis transformation defines a linear map (or we may just say a matrix) which itself stays invariant under basis transformation. We could have therefore simply defined matrices (or linear maps) to be objects with components following exactly the transformation property Eq. 226 under basis transformations.

The proof of Eq. 226 is in suffix notation again surprisingly simple: $\alpha(\mathbf{b}'_l) = \alpha(L^{-1}_{kl}\mathbf{b}_k) = L^{-1}_{kl}\alpha(\mathbf{b}_k) = L^{-1}_{kl}A_{jk}\mathbf{b}_j = L^{-1}_{kl}A_{jk}L_{ij}\mathbf{b}'_i = L_{ij}A_{jk}L^{-1}_{kl}\mathbf{b}'_i$ using Eq. 214, linearity of α, Eq. 225 and again Eq. 214 in this order. But by definition $\alpha(\mathbf{b}'_l) = A'_{il}\mathbf{b}'_i$ and therefore, using the uniqueness of the basis description, $A'_{il} = L_{ij}A_{jk}L^{-1}_{kl}$.

Vectors and matrices play a crucial rôle in Vector Calculus and in Mathematics in general. Many quantities in Physics, Chemistry, Engineering, Economics and other scientific and non-scientific areas are described by vectors and matrices but in fact many other quantities are not vectors or matrices but they follow transformation properties similar to Eq. 218 and Eq. 226 but for more than one or two indices. Such objects are called *Tensors* of which vectors and matrices are just the simplest examples. Instead of defining tensors as abstract objects living in some complicated vector space we will define Tensors in exactly the way we have just been looking at vectors and matrices, namely by defining them as objects being represented by components in a given co-ordinate system (basis) which transform via certain transformation rules under a basis transformation given by the basis transformation matrix L.

For simplicity we will study in this course only orthogonal basis transformations (i.e. basis transformations which transform orthonormal bases into orthonormal bases) with in addition determinant[c] equal to $+1$. These are exactly the rotations of 3-dimensional Euclidean space. Therefore all our basis transformation matrices are assumed to be orthogonal, as in example 7.1, i.e. $L^{-1} = L^T$ or in suffix notation $L^{-1}_{ij} = L_{ji}$. We can therefore summarise the basis transformation properties of vectors $\mathbf{x}$ and matrices A as

$$x'_i = L_{ij}x_j \; , \tag{227}$$

$$A'_{ij} = L_{ik}L_{jl}A_{kl} \; . \tag{228}$$

We can easily invert the equations Eq. 227 and Eq. 228 keeping in mind that $L^{-1}_{ij} = L_{ji}$: using Eq. 227 we find $L_{ik}x'_i = L_{ik}L_{ij}x_j = \delta_{jk}x_j = x_k$. We can similarly derive an equation for A_{ij} and find:

$$x_i = L_{ji}x'_j \; , \tag{229}$$

$$A_{ij} = L_{ki}L_{lj}A'_{kl} \; . \tag{230}$$

[a] The fact that we choose $\alpha(\mathbf{b}_i) = A_{ji}\mathbf{b}_j$ instead of $\alpha(\mathbf{b}_i) = A_{ij}\mathbf{b}_j$ means that the components of $\alpha(\mathbf{b}_i)$ in B are written in the i-th column of A and not the i-th row. This is exactly what we need to have the equation $A_{ij}x_j$ representing the simple matrix multiplication $A\mathbf{x}$.

[b] Depending on whether you define L using $\mathbf{b}_i = L_{ji}\mathbf{b}'_j$ or $\mathbf{b}'_i = L_{ji}\mathbf{b}_j$ you find $A' = LAL^{-1}$ or $A' = L^{-1}AL$ respectively. We want to use the former rather than the latter.

[c] An orthogonal basis transformation matrix has either determinant $+1$ (rotations) or -1 (rotations and reflections). We want to restrict ourselves to $+1$ (rotations) only.

24 Cartesian Tensors

Following our ideas in the initial remarks we will now assume that the matrix L describes an orthogonal basis transformation with $\det L = +1$ (i.e. a rotation) from the orthonormal basis B (which we will refer to as *frame 1*) to the orthonormal basis B' (*frame 2*) in $\mathbb{R}^3$:

$$\mathbf{b}_i = L_{ji}\mathbf{b}'_j \quad \text{and} \quad \mathbf{b}'_i = L^{-1}_{ji}\mathbf{b}_j = L_{ij}\mathbf{b}_j \ , \tag{231}$$

with

$$L_{ij}L_{kj} = \delta_{ik} = L_{ji}L_{jk} \ . \tag{232}$$

Because the two bases B and B' are both orthonormal we can calculate the components of L simply by taking scalar products: $L_{ij} = \mathbf{b}'_i.\mathbf{b}_j$. We now define a vector to be an object represented by components which transform under Eq. 218.

81 **Vectors**
A vector is an object represented by components x_i *which transform according to*

$$x'_i \ = \ L_{ij}x_j \ , \tag{233}$$

under the orthogonal basis transformation described by L. Eq. 233 *can be inverted to*

$$x_i \ = \ L_{ji}x'_j \ . \tag{234}$$

In the same way we can define a matrix using Eq. 226.

82 **Matrices**
A matrix is an object represented by components x_{ij} *which transform according to*

$$x'_{ij} \ = \ L_{ik}L_{jl}x_{kl} \ , \tag{235}$$

under the orthogonal basis transformation described by L. Eq. 235 *can be inverted to*

$$x_{ij} \ = \ L_{ki}L_{lj}x'_{kl} \ . \tag{236}$$

The simplest objects satisfying a trivial version of the transformation rules Eq. 218 or Eq. 226 are in fact objects just represented by a fixed number which does not change under basis transformation. This is what we call a *scalar*.

83 **Scalars**
A scalar is an object represented by a real number x *which stays the same under the orthogonal basis transformation described by* L.

We are now in a position to generalise these ideas to the general Cartesian tensors.

84 **Cartesian tensors**
A Cartesian[d] *tensor of rank*[e] $k \in \mathbb{N}_0$ *is an object represented by the components* $x_{i_1\ldots i_k}$ *which*

[d] We call these tensors *Cartesian* tensors since we assumed that the basis transformation L is orthogonal with $\det L = 1$. Without this simplification we would find for general non-orthogonal L two different types of indices: indices transforming with L and others transforming with L^{-1}. The former type of index is called *covarient* and the latter is called *contravariant*. A matrix is therefore in this general setting a once covariant, once contravariant tensor whilst a vector is simply a once covariant tensor. In order to include in the notation which index is covariant and which index is contravariant we would then put the covariant indices down and the contravariant indices up. Hence, a vector would still be written as x_i whilst a matrix would be written as $a_i{}^j$. For our applications Cartesian basis transformations are sufficient and therefore we do not need to distinguish these different types of indices.

[e] The rank of a tensor must not be confused with the rank of a matrix.

transform according to

$$x'_{i_1 i_2 \ldots i_k} = L_{i_1 j_1} L_{i_2 j_2} \ldots L_{i_k j_k} x_{j_1 j_2 \ldots j_k} \,, \tag{237}$$

under the orthogonal basis transformation described by L. In particular, a scalar is a rank 0 tensor, a vector is a rank 1 tensor and a matrix is a rank 2 tensor. Eq. 237 *can be inverted to*

$$x_{i_1 i_2 \ldots i_k} = L_{j_1 i_1} L_{j_2 i_2} \ldots L_{j_k i_k} x'_{j_1 j_2 \ldots j_k} \,. \tag{238}$$

Since we only consider Cartesian tensors in $\mathbb{R}^3$ all our indices run from 1 to 3. In exactly the same way as in [84] we can define Cartesian rank k tensors in $\mathbb{R}^n$ were the only difference would be that all the indices in Eq. 237 would run from 1 to n. n is called the *dimension* of the tensor, thus our tensors in $\mathbb{R}^3$ are 3 dimensional tensors of rank k. Before we continue looking at rank 2 tensors in particular we will give several examples of Cartesian tensors in $\mathbb{R}^3$:

Example 7.2 *The dot product $s = x_i y_i$ of two rank 1 tensors (vectors) x_i and y_i is a rank 0 tensor (a scalar). This can easily be seen by taking the dot product $s' = x'_i y'_i$ of the components with respect to B' and insert the transformation properties $x'_i = L_{ij} x_j$ and $y'_i = L_{ik} y_k$: $x'_i y'_i = L_{ij} x_j L_{ik} y_k = L_{ij} L_{ik} x_j y_k = \delta_{jk} x_j y_k = x_j y_j$ and therefore $s' = s$.*

Example 7.3 *An object which is represented by the components of the matrix*

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{239}$$

in the basis B and by the matrix

$$A' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{240}$$

in the basis B' cannot be a tensor. This can easily be seen since the determinant $\det A = 2$ but the determinant $\det A = 0$. We know from Algebra and Geometry that the determinant of a matrix stays the same under a basis transformation, therefore there cannot be a basis transformation of the type of Eq. 235 *(not even of the type of* Eq. 226*) connecting these two matrices.*

In the same way the matrix

$$C = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{241}$$

in the basis B and the matrix

$$C' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{242}$$

in the basis B' cannot represent a tensor since the trace $\mathrm{tr}\,(C) = 3$ and the trace $\mathrm{tr}\,(C') = -1$ and we again know from Algebra and Geometry that a basis transformation leaves the trace invariant. So, C and C' cannot be connected by a basis transformation.

Another way of spotting that A *and* A' *cannot represent a tensor is since the rank (this time the matrix rank, not the tensor rank!) of the matrix* A *is* 3 *whilst the rank of the matrix* A' *is* 2 *and again, the rank of a matrix has to be the same under basis transformation which shows that there is no basis transformation connecting* A *and* A'*. The same argument would obviously not work for* C *and* C' *since they both have (matrix) rank* 3 *but still do not represent a tensor as can be seen from the different traces.*

Example 7.4 *A physical entity represented in every Cartesian frame by the array of numbers* δ_{ij} *is a rank* 2 *tensor. In frame* 1 *the entity is represented by* δ_{ij}*. We then have to check that* $L_{ik}L_{jl}\delta_{kl}$ *equals the new components (in frame* 2*) which by definition are again the same* δ_{ij}*. But this is obvious from* Eq. 232*:* $L_{ik}L_{jl}\delta_{kl} = L_{ik}L_{jk} = \delta_{ij}$ *and hence it is a rank* 2 *tensor. This tensor has the property that the components are exactly the same in each Cartesian frame. A tensor with this property is called isotropic. We will study isotropic tensors in* Sec. 28.

A physical entity represented in every Cartesian frame by the components ϵ_{ijk} *is a rank* 3 *tensor. In frame* 1 *the entity is represented by* ϵ_{ijk}*. We then have to check that* $L_{il}L_{jm}L_{kn}\epsilon_{lmn}$ *equals the new components (in frame* 2*) which by definition are again the same* ϵ_{ijk}*. If we let* $\mathbf{l}_i$ *represent the i-th row of the basis transformation matrix* L*, then* $L_{il}L_{jm}L_{kn}\epsilon_{lmn} = [\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k]$ *with* $[\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k] = \mathbf{l}_i.(\mathbf{l}_j \times \mathbf{l}_k)$ *being the scalar triple product. If any of the indices* i*,* j*,* k *are the same, then obviously* $[\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k] = 0$ *otherwise exactly one of them is equal to* 1*, one equal to* 2 *and one equal to* 3*.* $[\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3] = \det L = 1$ *since* L *is a rotation and all the other such cases would be* $+1$ *if they can be reached using an even permutation of the vectors* $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ *and* -1 *if we require an odd permutation. Taking all this together we find that* $[\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k]$ *reflects exactly the definition of* ϵ_{ijk} *and hence* $[\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k] = \epsilon_{ijk}$*. But this just means* $L_{il}L_{jm}L_{kn}\epsilon_{lmn} = \epsilon_{ijk}$ *and therefore* ϵ_{ijk} *is a rank* 3 *tensor (which is like* δ_{ij} *isotropic).*

Example 7.5 *A physical entity is described in frame* 1 *by the matrix*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \tag{243}$$

in and in frame 2 *by the matrix*

$$A' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & \alpha \end{pmatrix} , \tag{244}$$

with some $\alpha \in \mathbb{R}$*. Frame* 2 *is obtained from frame* 1 *by rotating frame* 1 *around the* z*-axis by* $\frac{\pi}{2}$ *in positive direction (using the right hand rule). This entity could be a tensor.* L^{-1} *contains in its columns the frame* 2 *basis vectors written in terms the basis of frame* 1*. Therefore the first frame* 2 *basis vector has the components* $(0, 1, 0)^T$ *in frame* 1*, the second frame* 2 *basis vector has the components* $(-1, 0, 0)^T$ *in frame* 1 *and finally the third frame* 2 *basis vector has the components* $(0, 0, 1)^T$ *in frame* 3*. Therefore*

$$L^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{245}$$

and

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{246}$$

We can easily check that $A' = LAL^{-1}$ if and only if $\alpha = -1$ and therefore A could represent a rank 2 tensor (a matrix) in the case $\alpha = -1$. But we need to keep in mind that we have of course only checked one particular rotation rather than all rotations (in other words just one particular basis transformation L rather than all basis transformations L). Therefore we can only say that A 'could represent' a tensor rather than 'it is' a tensor.

If we assume that the matrix A does represent a tensor and frame 3 is obtained by rotating frame 1 around the x-axis by $+\frac{\pi}{4}$. What is the matrix A'' representing the tensor in frame 3? The basis transformation $\hat{L}$ is obviously given by

$$\hat{L}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} \tag{247}$$

and

$$\hat{L} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ 0 & -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{pmatrix} . \tag{248}$$

For example the second axis of frame 3 is in the yz-plane of frame 1 at an angle of $\frac{\pi}{4}$ to the x-axis and therefore has components $(0, \frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})^T$ in frame 1 which defines the second column of $\hat{L}^{-1}$. Since A represents a tensor we need $A'' = \hat{L}A\hat{L}^{-1}$ and hence

$$A'' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & -1 \end{pmatrix} . \tag{249}$$

Example 7.6 *The partial derivatives $\frac{\partial}{\partial x_i}$ represent a rank 1 tensor (i.e. a vector). We use the chain rule $\frac{\partial}{\partial x'_i} = \frac{\partial x_j}{\partial x'_i}\frac{\partial}{\partial x_j}$.* Eq. 234 *$x_j = L_{ij}x'_i$ leads to $\frac{\partial x_j}{\partial x'_i} = L_{ij}$ and therefore $\frac{\partial}{\partial x'_i} = L_{ij}\frac{\partial}{\partial x_j}$ which shows that $\frac{\partial}{\partial x_i}$ and therefore $\boldsymbol{\nabla}$ is a rank 1 tensor.*

Example 7.7 *The partial derivatives $\frac{\partial^3}{\partial x_i \partial x_j \partial x_k}$ represent a rank 3 tensor. This time we use the chain rule three times to convert the third order derivatives in the new frame to third order derivatives in the old frame: $\frac{\partial^3}{\partial x'_i \partial x'_j \partial x'_k} = \frac{\partial x_l}{\partial x'_i}\frac{\partial x_m}{\partial x'_j}\frac{\partial x_n}{\partial x'_k}\frac{\partial^3}{\partial x_l \partial x_m \partial x_n}$.* Eq. 234 *$x_l = L_{il}x'_i$ leads to $\frac{\partial x_l}{\partial x'_i} = L_{il}$ and similarly $\frac{\partial x_m}{\partial x'_j} = L_{jm}$ and $\frac{\partial x_n}{\partial x'_k} = L_{kn}$. and therefore $\frac{\partial^3}{\partial x'_i \partial x'_j \partial x'_k} = L_{il}L_{jm}L_{kn}\frac{\partial^3}{\partial x_l \partial x_m \partial x_n}$ which shows that $\frac{\partial^3}{\partial x_i \partial x_j \partial x_k}$ is a rank 3 tensor. Note that Schwarz's theorem [1] says that higher order partial derivatives are independent of the order of differentiation. Therefore this rank 3 tensor stays the same under exchange of any of the indices. Such a tensor is called totally symmetric and we will study totally symmetric tensors in* Sec. 29.

Example 7.8 *The differential dx_i is a rank 1 tensor, a vector. This is clear from the chain rule $dx'_i = \frac{\partial x'_i}{\partial x_j}dx_j = L_{ij}dx_j$ which follows from $x'_i = L_{ij}x_j$.*

Example 7.9 *Let $\mathcal{V}$ describe a volume in $\mathbb{R}^3$. The volume integral $\int_{\mathcal{V}} dV$ is a rank 0 tensor (a scalar) since we find in frame 2 the volume integral $\int_{\mathcal{V}'} dV' = \int_{\mathcal{V}'} d(x'_1, x'_2, x'_3) = \int_{\mathcal{V}} \frac{d(x'_1,x'_2,x'_3)}{d(x_1,x_2,x_3)} d(x_1, x_2, x_3)$ with the Jacobian $\frac{d(x'_1,x'_2,x'_3)}{d(x_1,x_2,x_3)} = |\det L|$ using $x'_i = L_{ij}x_j$. But $\det L = 1$ for the rotation L and therefore $\int_{\mathcal{V}'} dV' = \int_{\mathcal{V}} dV$.*

In the same way we see that the integral $\int_{\mathcal{V}} x_i dV$ is a rank 1 tensor (a vector): $\int_{\mathcal{V}'} x'_i dV' = \int_{\mathcal{V}'} x'_i d(x'_1, x'_2, x'_3) = \int_{\mathcal{V}} L_{ij}x_j \frac{d(x'_1,x'_2,x'_3)}{d(x_1,x_2,x_3)} d(x_1, x_2, x_3) = L_{ij}\int_{\mathcal{V}} x_j dV$.

25 Tensor Rules and Properties

We know from Algebra and Geometry that vectors and matrices are closed under addition and scalar multiplication. Generally, if we take two tensors $x_{i_1 i_2 \ldots i_k}$ and $y_{i_1 i_2 \ldots i_k}$ of the same rank k then a linear combination $\alpha x_{i_1 i_2 \ldots i_k} + \beta y_{i_1 i_2 \ldots i_k}$ $(\alpha, \beta \in \mathbb{R})$ is again a rank k tensor since

$$\begin{aligned} \alpha x'_{i_1 i_2 \ldots i_k} + \beta y'_{i_1 i_2 \ldots i_k} &= \alpha L_{i_1 j_1} L_{i_2 j_2} \ldots L_{i_k j_k} x_{j_1 j_2 \ldots j_k} + \beta L_{i_1 j_1} L_{i_2 j_2} \ldots L_{i_k j_k} y_{j_1 j_2 \ldots j_k} \\ &= L_{i_1 j_1} L_{i_2 j_2} \ldots L_{i_k j_k} (\alpha x_{j_1 j_2 \ldots j_k} + \beta y_{j_1 j_2 \ldots j_k}) \, . \end{aligned}$$

85 Linear combinations of tensors
The linear combination $\alpha x_{i_1 i_2 \ldots i_k} + \beta y_{i_1 i_2 \ldots i_k}$ $(\alpha, \beta \in \mathbb{R})$ *of the two rank* k *tensors* $x_{i_1 i_2 \ldots i_k}$ *and* $y_{i_1 i_2 \ldots i_k}$ *is again a tensor of rank* k*.*

If all components of a rank k tensor vanish in one Cartesian frame $x_{i_1 i_2 \ldots i_k} = 0$ then they obviously vanish in all Cartesian frames simply by applying $x'_{i_1 i_2 \ldots i_k} = L_{i_1 j_1} L_{i_2 j_2} \ldots L_{i_k j_k} x_{j_1 j_2 \ldots j_k} = 0$ since all $x_{j_1 j_2 \ldots j_k} = 0$. Hence if two rank k tensors $x_{i_1 i_2 \ldots i_k}$ and $y_{i_1 i_2 \ldots i_k}$ are equal in one Cartesian frame then they are equal in all Cartesian frames. This follows simply by looking at the tensor $x_{i_1 i_2 \ldots i_k} - y_{i_1 i_2 \ldots i_k}$ which is obviously 0 in this Cartesian frame and therefore in all Cartesian frames.

86 Uniqueness theorem
If two rank k *tensors* $x_{i_1 i_2 \ldots i_k}$ *and* $y_{i_1 i_2 \ldots i_k}$ *are equal in one frame:* $x_{i_1 i_2 \ldots i_k} = y_{i_1 i_2 \ldots i_k}$ *then they are equal in all frames.*

It is also immediately obvious that the product of two tensors (not necessarily of the same rank) is again a tensor. Let us take the tensors $x_{i_1 i_2 \ldots i_k}$ of rank k and $y_{j_1 j_2 \ldots j_l}$ of rank l then $x_{i_1 i_2 \ldots i_k} y_{j_1 j_2 \ldots j_l}$ is a tensor of rank $k + l$:

$$\begin{aligned} x'_{i_1 i_2 \ldots i_k} y'_{j_1 j_2 \ldots j_l} &= L_{i_1 m_1} L_{i_2 m_2} \ldots L_{i_k m_k} x_{m_1 m_2 \ldots m_k} L_{j_1 n_1} L_{j_2 n_2} \ldots L_{j_l n_l} y_{n_1 n_2 \ldots n_l} \\ &= L_{i_1 m_1} L_{i_2 m_2} \ldots L_{i_k m_k} L_{j_1 n_1} L_{j_2 n_2} \ldots L_{j_l n_l} x_{m_1 m_2 \ldots m_k} y_{n_1 n_2 \ldots n_l} \, . \end{aligned}$$

We should stress that in the product $x_{i_1 i_2 \ldots i_k} y_{j_1 j_2 \ldots j_l}$ we do not sum over any of the indices $i_1, i_2, \ldots, i_k, j_1, j_2, \ldots, j_l$ and they are all independent of each other. This product is called the *outer product* of tensors.

87 Outer product
The outer product $x_{i_1 i_2 \ldots i_k} y_{j_1 j_2 \ldots j_l}$ *of the tensors* $x_{i_1 i_2 \ldots i_k}$ *of rank* k *and* $y_{j_1 j_2 \ldots j_l}$ *of rank* l *is a tensor of rank* $k + l$*.*

Example 7.10 *If we take the vectors* $\mathbf{v}$ *and* $\mathbf{u}$ *in* $\mathbb{R}^3$ *with the components* v_i *and* u_j *respectively then the product* $v_i u_j$ *define the components of a matrix. This is easily seen by looking at the transformation properties of* $v_i u_j$*:* $v'_i u'_j = L_{ik} v_k L_{jl} u_l = L_{ik} L_{jl} v_k u_l$ *which matches the transformation properties* EQ. 235 *of a matrix under the rotation* L*.*

In the case of vectors and matrices we can of course re-write the transformation properties in terms of matrix multiplication. The matrix product[f] $\mathbf{v}\mathbf{u}^T =: A$ *defines a* 3 *by* 3 *matrix* A*. Under the rotation* L *we obtain* $\mathbf{v}' = L\mathbf{v}$ *and* $\mathbf{u}' = L\mathbf{u}$ *and therefore* $A' = \mathbf{v}'\mathbf{u}'^T = (L\mathbf{v})(L\mathbf{u})^T = L\mathbf{v}\mathbf{u}^T L^T = LAL^{-1}$ *noting that* $L^T = L^{-1}$*. Hence* $A' = LAL^{-1}$ *which defines* A *as a matrix.*

[f] Note that $\mathbf{v}\mathbf{u}^T$ is a 3 by 3 matrix whilst $\mathbf{v}^T\mathbf{u}$ would simply be a number, the dot product $\mathbf{v}.\mathbf{u}$.

In example 7.10 we could have also considered $\mathbf{v}.\mathbf{u}$ instead of $\mathbf{v}\mathbf{u}^T$. This represents the dot product of the two vectors $v_i u_i$ and therefore a scalar as can easily be seen. In this case we take the product of the vector components but actually set them equal and sum over them instead of leaving them independent. Such a product of tensors is called an *inner product* of tensors. In a general product of tensors we can set several pairs of indices equal and sum over them but for simplicity we describe the inner product just with one (the first) index in common: $x_{ii_2\ldots i_k}y_{ij_2\ldots j_l}$ is an example of an inner product of the tensors $x_{i_1i_2\ldots i_k}$ of rank k and $y_{j_1j_2\ldots j_l}$.

$$\begin{aligned} x'_{ii_2\ldots i_k}y'_{ij_2\ldots j_l} &= L_{im_1}L_{i_2m_2}\ldots L_{i_km_k}L_{in_1}L_{j_2n_2}\ldots L_{j_ln_l}x_{m_1m_2\ldots m_k}y_{n_1n_2\ldots n_l} \\ &= L_{im_1}L_{in_1}L_{i_2m_2}\ldots L_{i_km_k}L_{j_2n_2}\ldots L_{j_ln_l}x_{m_1m_2\ldots m_k}y_{n_1n_2\ldots n_l} \\ &= \delta_{m_1n_1}L_{i_2m_2}\ldots L_{i_km_k}L_{j_2n_2}\ldots L_{j_ln_l}x_{m_1m_2\ldots m_k}y_{n_1n_2\ldots n_l} \\ &= x_{mm_2\ldots m_k}y_{mn_2\ldots n_l} \end{aligned}$$

and therefore this inner product is a tensor of rank $k+l-2$.

88 **Inner product**
The inner product of the tensors $x_{i_1i_2\ldots i_k}$ of rank k and $y_{j_1j_2\ldots j_l}$ of rank l is the product of these components with at least one index being equal and summed over, e.g. $x_{ii_2\ldots i_k}y_{ij_2\ldots j_l}$. In the case that in the inner product m pairs of indices are the same and are being summed over then the inner product is a tensor of rank $k+l-2m$.

Example 7.11 *The product of the two matrices A and B is described by the inner product $a_{ij}b_{jk}$ of the rank 2 tensors a_{ij} and b_{mn}. The inner product $a_{ij}b_{jk}$ with the second index of a_{ij} and the first index of b_{mn} being set equal and summed over is therefore a rank $2+2-2=2$ tensor, i.e. a matrix. In matrix notation this is of course obvious from $A'B' = (LAL^{-1})(LBL^{-1}) = LA(L^{-1}L)BL^{-1} = LABL^{-1}$.*

$a_{ij}b_{kj}$ is another inner product that can be constructed using the tensors A and B. In matrix notation this can be written as AB^T since $(AB^T)_{ik} = (A)_{ij}(B^T)_{jk} = a_{ij}b_{kj}$. In matrix notation the tensor property of AB^T is shown by $A'B'^T = (LAL^{-1})(LBL^{-1})^T = LAL^{-1}(L^{-1})^TB^TL^T = LAL^{-1}LB^TL^{-1} = LAB^TL^{-1}$ since $L^T = L^{-1}$ for the rotation L.

$a_{ij}b_{ji}$ is another inner product constructed from the tensors A and B. This is in fact the trace $\mathrm{tr}\,(AB)$ of the matrix product AB. This time the inner product has two pairs of indices set equal and summed over them and we obtain a tensor of rank $2+2-2\times 2=0$, hence a scalar. Noting that within the trace matrices can be commuted $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$ we can easily see in matrix notation $\mathrm{tr}\,(A'B') = \mathrm{tr}\left(LAL^{-1}LBL^{-1}\right) = \mathrm{tr}\left(AL^{-1}LBL^{-1}L\right) = \mathrm{tr}\,(AB)$ and therefore $\mathrm{tr}\,(AB)$ is indeed a scalar tensor.

Finally if we take a rank k tensor and set two of its indices equal and sum over them we obtain a tensor of rank $k-2$. If we for example take the tensor $x_{i_1i_2\ldots i_k}$ and set $i_1 = i_2 =: i$ and sum over i: $x_{iii_3\ldots i_k}$ then we find $x'_{iii_3\ldots i_k} = L_{ij_1}L_{ij_2}L_{i_3j_3}\ldots L_{i_kj_k}x_{j_1j_2j_3\ldots j_k} = \delta_{j_1j_2}L_{i_3j_3}\ldots L_{i_kj_k}x_{j_1j_2j_3\ldots j_k} = L_{i_3j_3}\ldots L_{i_kj_k}x_{jjj_3\ldots j_k}$ which shows that it is a rank $k-2$ tensor. This is called the *contraction* of the tensor $x_{i_1i_2\ldots i_k}$ *with respect to its first and second index.*

89 **Contraction**
If we set two of the indices (the m-th and n-th index) of the tensor $x_{i_1i_2\ldots i_k}$ of rank k equal and sum over this common index (i.e. $i_m = i_n = i$) then we call it the contraction of the tensor $x_{i_1i_2\ldots i_k}$ with respect to the indices m and n ($m \neq n$). The contraction is a tensor of rank $k-2$. For example, $x_{iii_3\ldots i_k}$ is the contraction of $x_{i_1i_2\ldots i_k}$ with respect to the first and second index.

We can of course take multiple contractions on the same tensor and if we take m contractions on the rank k tensor then we obtain a rank $k-2m$ tensor. By taking the inner product of the tensors $x_{i_1 i_2 \ldots i_k}$ and $\delta_{i_1 i_2}$ we see that this inner product $\delta_{i_1 i_2} x_{i_1 i_2 \ldots i_k} = x_{iii_3 \ldots i_k}$ in fact equals the contraction $x_{iii_3 \ldots i_k}$. And generally, the contraction of the tensor $x_{i_1 i_2 \ldots i_k}$ with respect to the indices i_m and i_n $(m \neq n)$ is simply the inner product of the tensors $x_{i_1 i_2 \ldots i_k}$ and $\delta_{i_m i_n}$ and hence a rank $k-2$ tensor according to [88]. Another way of looking at the contraction (or taking the inner product with $\delta_{i_m i_n}$) is that the contraction is simply *taking a trace over the indices* i_m and i_n. For example, in the same way as the trace of a matrix A is the contraction a_{ii}, the contraction $x_{iii_3 \ldots i_k}$ can be defined as the (generalised) trace of the tensor $x_{i_1 i_2 \ldots i_k}$ with respect to its first and second index.

Example 7.12 *The trace* $\mathrm{tr}\,(A)$ *of the matrix* A *is a rank* 0 *tensor, a scalar, and is obviously the contraction of the rank* 2 *tensor* a_{ij}*:* $\mathrm{tr}\,(A) = a_{ii}$.

The contraction of the rank 2 *tensor* $\frac{\partial^2}{\partial x_i \partial x_j}$ *is* $\frac{\partial^2}{\partial x_i \partial x_i} = \boldsymbol{\nabla}^2$ *and therefore the Laplace operator* $\boldsymbol{\nabla}^2$ *is a rank* 0 *tensor, a scalar.*

We have just argued that the contraction is simply the inner product of a tensor with the tensor $\delta_{i_n i_m}$. Conversely we could of course also see the inner product of two tensors to be a suitable contraction of the outer product of the same tensors.

Example 7.13 *The matrix product* AB *is the inner product of* a_{ij} *and* b_{mn} *with the second index of* A *and the first index of* B *being set equal:* $a_{ij} b_{jl}$*. But this is obviously the same as the contraction of the outer product* $a_{ij} b_{kl}$ *with respect to the indices* j *and* k *and is also the same as the inner product of the rank* 4 *tensor* $a_{ij} b_{kl}$ *with the rank* 2 *tensor* δ_{mn} *setting the pairs* i*,* m *and* j*,* n *equal.*

The double contraction of the rank 6 *tensor* $\epsilon_{ijk} \epsilon_{ijl}$ *is a rank* $6 - 4 = 2$ *tensor. This contraction is in fact equal to the rank* 2 *tensor* $2\delta_{kl}$*:* $\epsilon_{ijk} \epsilon_{ijl} = \delta_{jj} \delta_{kl} - \delta_{jl} \delta_{kj} = 3\delta_{kl} - \delta_{kl} = 2\delta_{kl}$.

26 The Quotient Theorem

So far we have only taken products of tensors and contracted them and obtained in this way new tensors. It is amazing that the inverse of this statement is also true, in the sense that if taking an array of numbers and contracting it with an arbitrary tensor always leads to another tensor then the array of numbers has to be a tensor itself. This is called the *quotient theorem* for tensors.

90 **Quotient theorem for tensors**
Given an array of numbers $a_{i_1 i_2 \ldots i_k}$*. If* $a_{i_1 i_2 \ldots i_{m-1} i_m i_{m+1} \ldots i_k} x_{j_1 j_2 \ldots j_{n-1} i_m j_{n+1} \ldots j_l}$ *(where we have contracted over the pair of indices* i_m *and* j_n *for some* m*,* n*) is a tensor of rank* $k + l - 2$ *for all rank* l *tensors* $x_{i_1 i_2 \ldots i_l}$ *then* $a_{i_1 i_2 \ldots i_k}$ *is a rank* k *tensor itself.*

For simplicity we give the proof only for rank 2 tensors ($k = l = 2$), the generalisation is obvious but in terms of notation very cumbersome. In frame 1 we consider $a_{ij} x_{kj}$ where we have contracted the array a_{ij} and the tensor x_{kl} over their second indices (we could have equally well chosen any other combination of indices). In frame 2 we obtain $a'_{ij} x'_{kj}$ and assuming that this is a rank 2 tensor we find $a'_{ij} x'_{kj} = L_{il} L_{kn} a_{lm} x_{nm}$. But we also know that $x_{nm} = L_{pn} L_{qm} x'_{pq}$

due to [236] and therefore $a'_{ij}x'_{kj} = L_{il}L_{kn}L_{pn}L_{qm}a_{lm}x'_{pq} = L_{il}\delta_{kp}L_{qm}a_{lm}x'_{pq} = L_{il}L_{jm}a_{lm}x'_{kj}$ (in the last step we have simply re-labeled q by j). Thus

$$(a'_{ij} - L_{il}L_{jm}a_{lm})x'_{kj} = 0 \tag{250}$$

for all i, k and arbitrary x_{ij}. In particular, if we choose X' to be the matrix with a 1 at row r and column s and 0 everywhere else we find

$$a'_{is} - L_{il}L_{sm}a_{lm} = 0 \tag{251}$$

for all i and s and therefore

$$a'_{is} = L_{il}L_{sm}a_{lm} \tag{252}$$

which concludes our argument.

We can of course use the quotient theorem successively and therefore contract over several pairs of indices. For example if $a_{ij}v_iv_j$ is a scalar for all vectors $\mathbf{v}$ then, using the quotient theorem, $a_{ij}v_i$ is a rank 1 tensor for all vectors $\mathbf{v}$ and therefore, using the quotient theorem again, we have that a_{ij} is a rank 2 tensor. Furthermore, we can formulate the quotient theorem also when contracting with one tensor but over several pairs of indices. For example, if $a_{ij}x_{ij}$ is a scalar for all matrices x_{ij} then a_{ij} is a tensor. The proof given above (for $i = k$) will easily show that this generalisation of the quotient theorem holds.

Example 7.14 *Given an arbitrary vector* $\mathbf{v}$ *then* $\delta_{ij}v_j = v_i$ *and therefore* $\delta_{ij}v_j$ *is a vector. This is another proof for* δ_{ij} *being a rank* 2 *tensor.*

The outer product of the differentials dx_idx_j *is a rank* 2 *tensor using example* 7.8 *and the definition of the outer product. If we assume that the array of numbers* g_{ij} *leads to a scalar when double contracted with the arbitrary differentials* dx_idx_j*:* $g_{ij}dx_idx_j$ *is a scalar, then* g_{ij} *is a rank* 2 *tensor, called the metric tensor. This scalar tensor is denoted by* $ds^2 = g_{ij}dx_idx_j$*. We will study the metric tensor in* Sec. 30.

27 Rank 2 Tensors in $\mathbb{R}^3$ - Examples and Properties

Rank 2 tensors are of course just matrices and we have seen in our examples in the previous section that the arguments whether an array of numbers A represents a matrix can either be written in index notation as $a'_{ij} = L_{ik}L_{jl}a_{kl}$ or simply in matrix product notation $A' = LAL^{-1}$ where we need to keep in mind that $L^{-1} = L^T$ for the rotation L.

We of course know matrices and their properties very well from Algebra and Geometry. In this section we want to give some important examples of rank 2 tensors, i.e. matrices, and want to review some of the important properties of matrices.

Conductivity Tensor

Tensors allow us to isolate physical and geometric properties from the properties that depend just on the particular choice of the coordinate system. For example, let us consider the flow of current in a conducting material. If we apply an electric field which is parallel to an electric wire, then the field will generate an electric current along the wire. In this case the electric current will be in the direction of the electric field. We describe the electric field by the vector $\mathbf{E}$ and the electric current in the conductor by $\mathbf{j}$. We then have $\mathbf{j} = \varsigma\mathbf{E}$. ς is called the *conductivity*

and is the reciprocal of the *electrical resistivity*. But if the wire is not parallel to the electric field then the current generated in the wire can obviously not be parallel to the electric field and in fact just the component parallel to the wire will matter. In this case we find $\mathbf{j} = \sigma\mathbf{E}$ with $\sigma = \varsigma P$ where P is the projection matrix, projecting a vector orthogonally onto the direction of the wire. In the general case, if we take some anisotropic material which conductively is different in different directions then we will have a general linear relationship between $\mathbf{E}$ and $\mathbf{j}$:

$$\mathbf{j} \;=\; \sigma\mathbf{E}\,. \tag{253}$$

σ is now a matrix of 9 numbers describing how the linear relationship between $\mathbf{E}$ and the generated current $\mathbf{j}$. If the structure of the conductor is inhomogeneous then σ may even depend on the position. The physical law $\mathbf{j} = \sigma\mathbf{E}$ should of course take exactly the same form no matter which co-ordinate axes are used so that with new axes we should find $\mathbf{j}' = \sigma'\mathbf{E}'$. Since the arbitrary vector field $\mathbf{E}$ is obviously a vector and the electric current $\mathbf{j}$ is also a vector we easily find, using the quotient theorem, that σ is a rank 2 tensor, a matrix. σ is called the *conductivity tensor*.

91 **Conductivity tensor and Ohm's**[g] **law**
The conductivity tensor σ is a rank 2 tensor describing the electric current $\mathbf{j}$ generated by the electric field $\mathbf{E}$ in a conducting material: $\mathbf{j} = \sigma\mathbf{E}$ and in suffix notation

$$j_m = \sigma_{mn}E_n\,. \tag{254}$$

If the conductor is inhomogeneous then σ is a function of the position on the conductor. $\mathbf{j} = \sigma\mathbf{E}$ is called Ohm's law.

In case the conductor is isotropic (i.e. it conducts in all directions equally well) then $\mathbf{j}$ will obviously always be parallel to $\mathbf{E}$ with the same proportionally factor and therefore $\sigma_{ij} = \varsigma\delta_{ij}$ for an isotropic material (c.f. example 7.18).

Centre of Mass

A massive object is described by the volume $\mathcal{V}$ and scalar mass density $\rho(\mathbf{x})$ (the mass density is a scalar tensor, i.e. $\rho(\mathbf{x})$ describes the density in frame 1 and $\rho'(\mathbf{x}')$ in frame 2, so $\rho(\mathbf{x}) = \rho'(\mathbf{x}')$). The total mass M is given by the integral $M = \int_{\mathcal{V}} \rho(\mathbf{x})dV$ which is obviously a rank 0 tensor, a scalar:

$$\begin{aligned} M' &= \int_{\mathcal{V}'} \rho'(\mathbf{x}')d(x_1', x_2', x_3') \\ &= \int_{\mathcal{V}} \rho(\mathbf{x})\frac{d(x_1', x_2', x_3')}{d(x_1, x_2, x_3)}d(x_1, x_2, x_3) \\ &= \int_{\mathcal{V}} \rho(\mathbf{x})d(x_1, x_2, x_3) = M\,. \end{aligned} \tag{255}$$

The *centre of mass* of the body is given by the integral $\frac{1}{M}\int_{\mathcal{V}} \rho(\mathbf{x})x_i dV$ which is obviously a rank 1 tensor following a similar argument as for the rank 0 tensor M.

92 **Centre of mass**
The mass $M = \int_{\mathcal{V}} \rho(\mathbf{x})dV$ of the massive object described by the volume $\mathcal{V}$ and mass density $\rho(\mathbf{x})$ is a rank 0 tensor. Its centre of mass $c_i = \frac{1}{M}\int_{\mathcal{V}} \rho(\mathbf{x})x_i dV$ is a rank 1 tensor, a vector $\mathbf{c}$.

[g] Georg Ohm, 1789-1854.

Rigid Body Rotation - Inertia Tensor

A rigid body defined by the volume $\mathcal{V}$ and mass density $\rho(\mathbf{x})$ is rotating with angular velocity $\boldsymbol{\omega}$ about the axis defined by $\boldsymbol{\omega}$, through the origin (and with angular speed $\|\boldsymbol{\omega}\|$). The velocity of a point P with position vector $\mathbf{x}$ is then given by $\boldsymbol{\omega} \times \mathbf{x}$. If we partition the whole volume $\mathcal{V}$ into little volume elements δV, then the angular momentum about $\mathbf{0}$ of such a volume element δV with position vector $\mathbf{x}$ is

$$\mathbf{x} \times (\rho(\mathbf{x})\delta V \boldsymbol{\omega} \times \mathbf{x}) \quad = \quad \rho(\mathbf{x}) \left((\mathbf{x}.\mathbf{x})\boldsymbol{\omega} - (\mathbf{x}.\boldsymbol{\omega})\mathbf{x}\right) \delta V \; . \tag{256}$$

The total angular momentum of the rigid body about $\mathbf{0}$ is therefore given by

$$\sum \rho(\mathbf{x}) \left((\mathbf{x}.\mathbf{x})\boldsymbol{\omega} - (\mathbf{x}.\boldsymbol{\omega})\mathbf{x}\right) \delta V \; . \tag{257}$$

where the sum is to be taken over all the volume elements δV. In the limit $\delta V \to 0$ this obviously becomes a volume integral and therefore the total angular momentum of the rigid body about $\mathbf{0}$ is

$$\begin{aligned} L_i &= \int_{\mathcal{V}} \rho(\mathbf{x}) \left(x_k x_k \omega_i - x_i x_j \omega_j\right) dV \\ &= \int_{\mathcal{V}} \rho(\mathbf{x}) \left(x_k x_k \delta_{ij} - x_i x_j\right) dV \; \omega_j \; . \end{aligned} \tag{258}$$

The integral $\Theta_{ij}(\mathbf{0}) = \int_{\mathcal{V}} \rho(\mathbf{x}) \left(x_k x_k \delta_{ij} - x_i x_j\right) dV$ is a rank 2 tensor, called the *inertia tensor about the origin.*

93 **Inertia Tensor**
The rank 2 *tensor*

$$\Theta_{ij}(\mathbf{0}) = \int_{\mathcal{V}} \rho(\mathbf{x}) \left(x_k x_k \delta_{ij} - x_i x_j\right) dV \tag{259}$$

is called the inertia tensor about the origin. It relates the angular momentum $\mathbf{L}$ *of a rigid body about the origin with the angular velocity* $\boldsymbol{\omega}$*:*

$$\mathbf{L} \quad = \quad \Theta(\mathbf{0}) \, , \boldsymbol{\omega} \; , \tag{260}$$

or in index notation

$$L_i \quad = \quad \Theta_{ij}(\mathbf{0}) \, \omega_j \; . \tag{261}$$

Θ_{ij} is symmetric in i, j which means that Θ is a symmetric matrix. We know from Algebra and Geometry that a symmetric matrix can be diagonalised with real eigenvalues and using an orthogonal basis transformation (which can be chosen to have determinant $+1$). Therefore, there exists a rotation L which diagonalises the inertia tensor. The eigenvectors are called *principal axes* of the rigid body. If we rotate the rigid body about one of the principal axes then we find that the angular momentum $\mathbf{L}$ is proportional to the angular velocity $\boldsymbol{\omega}$ with the corresponding eigenvector being the proportionality factor. And therefore if you let the rigid body rotate freely about such an axis it will simply continue to rotate about this axes.

The *moment of inertia about a fixed (normalised) direction* $\mathbf{n}$ is the projection of $\Theta\mathbf{n}$ onto $\mathbf{n}$, i.e. the double inner product $\Theta_{ij} n_i n_j$. This describes the component of the angular momentum parallel to $\mathbf{n}$ required to force the body on a rotation with unit angular speed about the axis[h] $\mathbf{n}$.

[h] Note that the component of the angular momentum perpendicular to $\mathbf{n}$ required to force this rotation about $\mathbf{n}$ may not be trivial (this will be the case whenever $\mathbf{n}$ is not a principal axis) and therefore the axis itself will be forced in order to force this rotation.

Example 7.15 *Let us take the homogeneous massive cylinder* $x^2+y^2 \leq 1$ *and* $-1 \leq z \leq 1$ *with mass density* $\rho(\mathbf{x}) = 1$. *The mass of the cylinder is* $M = 2\pi$ *which is equal to its volume. We write the inertia tensor in cylindrical polar coordinates:*

$$\Theta_{ij}(\mathbf{0}) \quad = \quad \int_{\mathcal{V}} \left((r^2+z^2)\delta_{ij} - x_i x_j\right) dV \ .$$

The off-diagonal components involve an integral over either $\sin\phi\cos\phi$ *or just* $\cos\phi$ *or* $\sin\phi$. *All these integrals over a whole period of* 2π *vanish and therefore all off-diagonal components of* $\Theta_{ij}(\mathbf{0})$ *vanish.*

$$\begin{aligned}
\Theta_{11}(\mathbf{0}) &= \int_0^1\int_{-1}^1\int_0^{2\pi} \left((r^2+z^2) - r^2\cos^2\phi\right) r d\phi dz dr \\
&= \int_0^1 [2\pi(2r^3 + \frac{2}{3}r) - \pi 2r^3] dr = \frac{7}{6}\pi = \Theta_{22}(\mathbf{0}) \ , \\
\Theta_{33}(\mathbf{0}) &= \int_0^1\int_{-1}^1\int_0^{2\pi} \left((r^2+z^2) - z^2\right) r d\phi dz dr = \int_0^1 [2\pi 2r^3] dr = \pi \ .
\end{aligned} \tag{262}$$

We therefore find the inertia tensor

$$\Theta(\mathbf{0}) = \begin{pmatrix} \frac{7}{6}\pi & 0 & 0 \\ 0 & \frac{7}{6}\pi & 0 \\ 0 & 0 & \pi \end{pmatrix} \ . \tag{263}$$

The principal axes are the z-axis and any direction in the xy-plane.

We now rotate the rigid body about the axis defined by the angular momentum $\boldsymbol{\omega}$ this time not through the origin but through the centre of mass $\mathbf{c}$. Exactly the same derivation as in Eq. 258 is valid except that the position vector of δV relative to $\mathbf{c}$ is not given by $\mathbf{x}$ but by $\mathbf{x}-\mathbf{c}$ and therefore we find for the inertia tensor about $\mathbf{c}$

$$\Theta_{ij}(\mathbf{c}) \quad = \quad \int_{\mathcal{V}} \rho(\mathbf{x}) \left((x_k - c_k)(x_k - c_k)\delta_{ij} - (x_i - c_i)(x_j - c_j)\right) dV \tag{264}$$

and the angular momentum for this rotation would then be given by

$$\mathbf{L} = \Theta(\mathbf{c})\,\boldsymbol{\omega} \ . \tag{265}$$

We find

$$\begin{aligned}
\Theta_{ij}(\mathbf{c}) &= \int_{\mathcal{V}} \rho(\mathbf{x}) \left((x_k - c_k)(x_k - c_k)\delta_{ij} - (x_i - c_i)(x_j - c_j)\right) dV \\
&= \int_{\mathcal{V}} \rho(\mathbf{x}) \left(x_k x_k \delta_{ij} - x_i x_j\right) dV + (c_k c_k \delta_{ij} - c_i c_j) \int_{\mathcal{V}} \rho(\mathbf{x}) dV \\
&\quad -2c_k\delta_{ij}\int_{\mathcal{V}} \rho(\mathbf{x}) x_k dV + c_i \int_{\mathcal{V}} \rho(\mathbf{x}) x_j dV + c_j \int_{\mathcal{V}} \rho(\mathbf{x}) x_i dV \\
&= \Theta_{ij}(\mathbf{0}) + M(\delta_{ij}c_k c_k - c_i c_j) - 2c_k c_k M\delta_{ij} + c_i c_j M + c_j c_i M \\
&= \Theta_{ij}(\mathbf{0}) - M(\delta_{ij}c_k c_k - c_i c_j)
\end{aligned} \tag{266}$$

where we used in step Eq. 266 the definition of the centre of mass $\mathbf{c}$.

94 **Parallel axis theorem**
The inertia tensors about the origin $\mathbf{0}$ *and about the centre of mass* $\mathbf{c}$ *are related by*

$$\Theta_{ij}(\mathbf{0}) \quad = \quad \Theta_{ij}(\mathbf{c}) + M(\delta_{ij}\|\mathbf{c}\|^2 - c_i c_j) \; . \tag{267}$$

The inertia tensors about the point $\mathbf{b}$ *and about the centre of mass* $\mathbf{c}$ *are therefore related by*

$$\Theta_{ij}(\mathbf{b}) \quad = \quad \Theta_{ij}(\mathbf{c}) + M\left(\delta_{ij}\|\mathbf{b}-\mathbf{c}\|^2 - (b_i - c_i)(b_j - c_j)\right) \; . \tag{268}$$

Example 7.16 *We use the parallel axis theorem to find the inertia tensor about* $\mathbf{b} = (1,0,0)$. *The centre of mass of the cylinder is obviously* $\mathbf{c} = (0,0,0)$. *This leads to*

$$M\left(\delta_{ij}\|\mathbf{b}-\mathbf{c}\|^2 - (b_i - c_i)(b_j - c_j)\right) = M\left(\delta_{ij} - \delta_{i1}\delta_{j1}\right)$$

$$\Theta((1,0,0)) = \begin{pmatrix} \frac{7}{6}\pi & 0 & 0 \\ 0 & \frac{7}{6}\pi & 0 \\ 0 & 0 & \pi \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2\pi & 0 \\ 0 & 0 & 2\pi \end{pmatrix} = \begin{pmatrix} \frac{7}{6}\pi & 0 & 0 \\ 0 & \frac{19}{6}\pi & 0 \\ 0 & 0 & 3\pi \end{pmatrix} \; .$$

Therefore the principal axes are $(1,0,0)$, $(0,1,0)$, *and* $(0,0,1)$ *(note that this is for the rotation about the point (1,0,0)).*

We now want to find the inertia tensor about $\hat{\mathbf{b}} = (\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}, 0)$*:*

$$M\left(\delta_{ij}\left\|\hat{\mathbf{b}}-\mathbf{c}\right\|^2 - (\hat{b}_i - c_i)(\hat{b}_j - c_j)\right) = M\left(\delta_{ij} - \tfrac{1}{2}(\delta_{i1}+\delta_{i2})(\delta_{j1}+\delta_{j2})\right)$$

$$\Theta((\tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2}, 0)) = \begin{pmatrix} \frac{7}{6}\pi & 0 & 0 \\ 0 & \frac{7}{6}\pi & 0 \\ 0 & 0 & \pi \end{pmatrix} + \begin{pmatrix} \pi & -\pi & 0 \\ -\pi & \pi & 0 \\ 0 & 0 & 2\pi \end{pmatrix} = \begin{pmatrix} \frac{13}{6}\pi & -\pi & 0 \\ -\pi & \frac{13}{6}\pi & 0 \\ 0 & 0 & 3\pi \end{pmatrix} \; .$$

Therefore the principal axes are $(0,0,1)$, $(1,1,0)$ *and* $(1,-1,0)$.

28 Isotropic Tensors

In example 7.4 we showed that an entity which is represented in each Cartesian frame by δ_{ij} is a rank 2 tensor. But this tensor has the additional property that it is actually exactly the same in each Cartesian frame. A tensor which is the same in each Cartesian frame is called *isotropic*. Obviously every rank 0 tensor is isotropic by definition, but in general a tensor will not normally be isotropic.

95 **Isotropic tensors**
A rank k *tensor* $x_{i_1 i_2 \ldots i_k}$ *is called isotropic if* $x'_{i_1 i_2 \ldots i_k} = x_{i_1 i_2 \ldots i_k}$ *for all rotations* L. *In other words, an isotropic tensor has exactly the same components in all Cartesian frames.*

Scalars, δ_{ij} and ϵ_{ijk} are examples of isotropic tensors (see example 7.4). Are there any rank 1 isotropic tensors? Assume that x_i is a rank 1 isotropic tensor, we then find $x_i = x'_i = L_{ij}x_j$ for all rotations L and therefore $(I - L)\mathbf{x} = \mathbf{0}$ for all rotations L where I is the 3×3 unit matrix. If we choose in particular the rotation about the z-axis, angle $\frac{\pi}{2}$, then

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \; . \tag{269}$$

$(I - L)\mathbf{x} = \mathbf{0}$ leads to $x_1 - x_2 = 0$ and $x_1 + x_2 = 0$ and therefore $x_1 = x_2 = 0$. Choosing a rotation about the x-axis, angle $\frac{\pi}{2}$, would in the same way show that also $x_3 = 0$. Hence, the only rank 1 isotropic tensor is the zero vector $\mathbf{0}$.

96 **Rank 0 and rank 1 isotropic tensors**
Rank 0 tensors (scalars) are by definition always isotropic but there is only one isotropic rank 1 tensor (a vector) which is the $\mathbf{0}$ *vector.*

Example 7.17 *In example 7.9 we showed that* $\int_{\mathcal{V}} x_i dV$ *is a rank 1 tensor. If* $\mathcal{V}$ *is a sphere centered at the origin then due to the symmetry of the volume it is clear that the integral has to be isotropic:* x_i *is obviously just a dummy variable for the integration which under a rotation due to the spherical symmetry of* $\mathcal{V}$ *always contributes in the same way to the integral. Therefore* $\int_{\mathcal{V}} x_i dV$ *is an isotropic rank 1 tensor but since* $\mathbf{0}$ *is the only isotropic rank 1 tensor we immediately find* $\int_{\mathcal{V}} x_i dV = 0$.

Whilst there are no rank 1 isotropic tensors except $\mathbf{0}$ we have of course already found a non-trivial rank 2 isotropic tensors, namely δ_{ij}. But in fact, this is the only one up to a scalar multiple! Assume that a_{ij} is a rank 2 isotropic tensor, then $a_{ij} = L_{ik}L_{jl}a_{kl}$ for all rotations L. In particular, we can take the rotation L from Eq. 269 and obtain from $LAL^T = A$

$$\begin{pmatrix} a_{22} & -a_{21} & a_{23} \\ -a_{12} & a_{11} & -a_{13} \\ a_{32} & -a_{31} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} . \tag{270}$$

This implies $a_{11} = a_{22}$ and $a_{23} = a_{13} = a_{32} = a_{31} = 0$. Rotating by $\frac{\pi}{2}$ about the x-axis shows that also $a_{33} = a_{11}$ and $a_{12} = a_{21} = 0$. Hence $A = \lambda I$, where I is the 3×3 unit matrix, or in index notation: $a_{ij} = \lambda\delta_{ij}$ for some $\lambda \in \mathbb{R}$.

97 **Rank 2 isotropic tensors**
The only rank 2 isotropic tensors (in 3 dimensions[i]*) are given by* $\lambda\delta_{ij}$ *for* $\lambda \in \mathbb{R}$.

Example 7.18 *The conductivity tensor of an isotropic material (a conductor conducting in all directions equally well) is an isotropic rank 2 tensor. Therefore* σ *is proportional to* δ_{ij} *for an isotropic conductor.*

Example 7.19 *If* $\mathcal{V}$ *is a sphere of radius* R *centred at the origin then* $\int_{\mathcal{V}} x_i x_j dV$ *is obviously an isotropic rank 2 tensor (the fact that it is a rank 2 tensor can be shown as in example 7.9; the fact that it is isotropic follows from the spherical symmetry of* $\mathcal{V}$*). Using [97] we find that* $\int_{\mathcal{V}} x_i x_j dV = \lambda\delta_{ij}$ *for some* $\lambda \in \mathbb{R}$. *We can compute* λ *by contracting the tensor:*

$$\begin{aligned} \lambda\delta_{ii} = 3\lambda &= \int_{\mathcal{V}} x_i x_i dV = \int_{\mathcal{V}} r^2 dV \\ &= \int_0^R \int_0^{2\pi} \int_0^{\pi} r^4 \sin\theta d\theta d\phi dr = \frac{4\pi}{5} R^5 . \end{aligned}$$

Therefore, we find

$$\int_{\mathcal{V}} x_i x_j dV = \frac{4\pi}{15} R^5 \delta_{ij} .$$

[i] In two dimensions the most general rank 2 isotropic tensor can be written as $\lambda\delta_{ij} + \mu\epsilon_{ij}$, where ϵ_{ij} is a rank 2 isotropic tensor with $\epsilon_{12} = 1$, $\epsilon_{21} = -1$ and $\epsilon_{11} = \epsilon_{22} = 0$ (i.e. antisymmetric).

In a similar way we can show that $\lambda\epsilon_{ijk}$ is the only (3 dimensional) isotropic rank 3 tensor. We again choose the rotation L from EQ. 269 and apply it to a general isotropic rank 3 tensor $x_{ijk} = L_{il}L_{jm}L_{kn}x_{lmn}$. If we set $i = 1$, $j = 2$ and $k = 3$ then we find $x_{123} = L_{12}L_{21}L_{33}x_{213}$ as the only non-trivial contribution and hence $x_{123} = -x_{213}$. In the same way we find $x_{231} = -x_{132}$ and $x_{312} = -x_{321}$. Furthermore, we find $x_{122} = L_{12}L_{21}L_{21}x_{211} = x_{211}$ and $x_{211} = L_{21}L_{12}L_{12}x_{122} = -x_{122}$ and therefore $x_{122} = x_{221} = 0$. Similarly, all x_{ijk} with two repeating indices will vanish. Finally we rotate the axes about $(1, 1, 1)$ angle $\frac{2\pi}{3}$, so that axis 1 becomes axis 2, 2 becomes 3 and 3 becomes 1. In this case the rotation matrix $\hat{L}$ is

$$\hat{L} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} . \tag{271}$$

Hence $x_{123} = L_{12}L_{23}L_{31}x_{231} = x_{231}$ and also $x_{231} = x_{312}$. Putting all this together, we find $x_{ijk} = \lambda\epsilon_{ijk}$.

98 **Rank 3 isotropic tensors**
The only rank 3 isotropic tensors (in 3 dimensions) are given by $\lambda\epsilon_{ijk}$ for $\lambda \in \mathbb{R}$.

In a similar way we can show the following theorem describing the most general isotropic rank 4 tensor[j] .

99 **Rank 4 isotropic tensors**
The most general rank 4 isotropic tensor (in 3 dimensions) is given by $\lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl} + \nu\delta_{il}\delta_{jk}$ for $\lambda, \mu, \nu \in \mathbb{R}$.

29 Symmetric and Antisymmetric Tensors

We have already realised that some matrices are symmetric under exchange of their indices ($a_{ij} = a_{ji}$) and some others are antisymmetric ($a_{ij} = -a_{ji}$). Such a property has crucial implications for matrices. We know for example that a symmetric matrix can be diagonalised with real eigenvalues using an orthogonal basis transformation. For physical entities represented by these matrices this leads to the concept of principal axes as discussed earlier for the inertia tensor. For antisymmetric matrices you will find in Part IB Linear Algebra that they can also be diagonalised but this time this has to be done over the complex numbers and the eigenvalues are all purely imaginary or identical to 0. We will now of course ask the question how to generalise the concept of symmetric and antisymmetric indices to general tensors.

100 **Symmetric and antisymmetric indices and tensors**
Given a rank k tensor $x_{i_1 i_2 \ldots i_k}$. The tensor is called symmetric in the indices i_m and i_n (w.l.o.g. $m < n$) if

$$x_{i_1 \ldots i_{m-1} i_m i_{m+1} \ldots i_{n-1} i_n i_{n+1} \ldots i_k} = x_{i_1 \ldots i_{m-1} i_n i_{m+1} \ldots i_{n-1} i_m i_{n+1} \ldots i_k} \tag{272}$$

and it is called antisymmetric in the indices i_m and i_n if

$$x_{i_1 \ldots i_{m-1} i_m i_{m+1} \ldots i_{n-1} i_n i_{n+1} \ldots i_k} = -x_{i_1 \ldots i_{m-1} i_n i_{m+1} \ldots i_{n-1} i_m i_{n+1} \ldots i_k} . \tag{273}$$

If the tensor $x_{i_1 i_2 \ldots i_k}$ is symmetric in any pair of its indices then we call it (totally) symmetric and if it is antisymmetric in any pair of its indices then we call it (totally) antisymmetric.

[j] Using the isotropic rank 2 tensor $\lambda\delta_{ij}$ and keeping its symmetry in i, j in mind, there are exactly 3 outer products with itself which would lead to isotropic rank 4 tensors x_{ijkl}: $\lambda\delta_{ij}\delta_{kl}$, $\mu\delta_{ik}\delta_{jl}$ and $\nu\delta_{il}\delta_{jk}$. The important result of [99] is that there are no others apart from combining these three.

Example 7.20 *δ_{ij} is a symmetric rank 2 tensor and ϵ_{ijk} is an antisymmetric rank 3 tensor. The inertia tensor is a symmetric rank 2 tensor but the conductivity tensor is not necessarily symmetric, this will depend on the structure of the conducting material. The outer product of two symmetric tensors such as $s_{ij}u_{kl}$ is symmetric in i, j and in k, l but not necessarily totally symmetric. The inner product of two symmetric tensors such as $s_{ij}u_{jk}$ is not necessarily symmetric in any indices.*

If we take the inner product of two tensors over two pairs of indices for which one of the tensors is symmetric and the other one is antisymmetric, then this inner product vanishes completely. For example if $x_{i_1 i_2 i_3 \ldots i_k}$ is symmetric in i_1 and i_2 and $y_{j_1 j_2 j_3 \ldots j_l}$ is antisymmetric in j_1 and j_2 then the double contraction $x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l}$ is identical to 0 since

$$\begin{aligned} x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l} &= -x_{nmi_3\ldots i_k} y_{nmj_3\ldots i_l} \\ &= -x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l} \ , \end{aligned} \tag{274}$$

where we have used the symmetry and antisymmetry in the first step and simply renamed m and n in the second step. But $x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l} = -x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l}$ means that $x_{mni_3\ldots i_k} y_{mnj_3\ldots i_l} = 0$.

101 **Inner product over symmetric and antisymmetric indices**
The (double) inner product over two pairs of indices which involves two symmetric indices for one tensor and two antisymmetric indices for the other tensor equals 0. *For example $\epsilon_{ijk} s_{ij} = 0$ for the symmetric tensor s_{ij}.*

ϵ_{ijk} is obviously a totally antisymmetric tensor (in 3 dimensions) and it is easy to see that it is the only one up to a scalar multiple: if we define $\epsilon_{123} = 1$ then all other values of ϵ_{ijk} are defined by antisymmetry in the indices. On the other hand, antisymmetric matrices in 3 dimensions are not defined uniquely, we need in fact three numbers a_1, a_2, a_3 to define an antisymmetric 3×3 matrix A uniquely:

$$A = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix} . \tag{275}$$

The matrix A with its components a_{ij} can be written as the inner product of ϵ_{ijk} and a vector $\mathbf{a}$: $a_{ij} = \epsilon_{ijk} a_k$.

We have seen that in 3 dimensions there is a unique totally antisymmetric rank 3 tensor ϵ_{ijk} which is unique up to a scalar multiple. In the same way it is obvious that in n dimensions there is a unique totally antisymmetric rank n tensor $\epsilon_{i_1 i_2 \ldots i_n}$ which is again unique up to a scalar multiple. It is constructed by setting $\epsilon_{12\ldots n} = 1$ and everything else follows from antisymmetry. In particular $\epsilon_{i_1 i_2 \ldots i_n} \neq 0$ if and only if $i_1, i_2, \ldots, i_n$ is a permutation of the numbers 1 to n and it is $+1$ for an even permutation and -1 for an odd permutation.

Any rank 2 tensor x_{ij} can be written in the following form

$$x_{ij} = \frac{1}{2}(x_{ij} + x_{ji}) + \frac{1}{2}(x_{ij} - x_{ij}) \ , \tag{276}$$

and we set

$$s_{ij} = \frac{1}{2}(x_{ij} + x_{ji}) \qquad \text{and} \qquad a_{ij} = \frac{1}{2}(x_{ij} - x_{ji}) \ , \tag{277}$$

such that $x_{ij} = s_{ij} + a_{ij}$. Obviously $s_{ij} = \frac{1}{2}(x_{ij} + x_{ji}) = \frac{1}{2}(x_{ji} + x_{ij}) = s_{ji}$ and $a_{ij} = \frac{1}{2}(x_{ij} - x_{ji}) = -\frac{1}{2}(x_{ji} - x_{ij}) = -a_{ji}$. Therefore s_{ij} is the *symmetric part* of x_{ij} and a_{ij} is its *antisymmetric part*. Every rank 2 tensor can therefore be decomposed into its (unique[k]) symmetric part and its (unique) antisymmetric part.

102 **Decomposition of a rank 2 tensor in symmetric and antisymmetric part**
Every rank 2 tensor (matrix) x_{ij} can be decomposed uniquely into its symmetric part s_{ij} and its antisymmetric part a_{ij} such that $x_{ij} = s_{ij} + a_{ij}$ with s_{ij} symmetric and a_{ij} antisymmetric. We have $s_{ij} = \frac{1}{2}(x_{ij} + x_{ji})$ and $a_{ij} = \frac{1}{2}(x_{ij} - x_{ji})$.

30 Examples of Tensors

Quadric Surfaces

Let S_{ij} be an n dimensional symmetric rank 2 tensor. Symmetric matrices can be diagonalised with real eigenvalues using an orthogonal basis transformation. Let Λ be the diagonal n by n matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ (the eigenvalues are not necessarily distinct) and therefore represents this tensor with respect to its eigenbasis.

We construct the quadratic form

$$Q(\mathbf{x}, \mathbf{x}) \quad = \quad x_i S_{ij} x_j \ , \tag{278}$$

with $\mathbf{x}$ being an n-dimensional vector. The quadratic form Q is obviously a scalar: $Q'(\mathbf{x}', \mathbf{x}') = x'_i S'_{ij} x'_j = L_{ik} x_k L_{il} L_{jm} S_{lm} L_{jp} x_p = \delta_{kl}\delta_{mp} x_k S_{lm} x_p = x_k S_{km} x_m$. If we choose x'_i to be the components of $\mathbf{x}$ with respect to the eigenbasis of S_{ij} then we find $Q(\mathbf{x}', \mathbf{x}') = \lambda_1 x'^2_1 + \ldots \lambda_n x'^2_n$ and therefore the surface defined by $Q = constant$ is a *quadric surface*. If $S_{ij} = \Theta_{ij}$, the inertia tensor, then the corresponding quadric surface is called the *inertia quadric*.

In $\mathbb{R}^3$ the quadric surface is for example an *ellipsoid* for $\lambda_1, \lambda_2, \lambda_3 > 0$, a *hyperboloid* for $\lambda_1, \lambda_2 > 0$, $\lambda_3 < 0$ or a *cylinder* for $\lambda_1 = \lambda_2 > 0$ and $\lambda_3 = 0$. The eigenvectors are called the *principal axes* of the quadric and the eigenvalues are called the *principal values*. It is easy to show that the quantities $\lambda_1 + \lambda_2 + \lambda_3 = \mathrm{tr}\,(S)$, $\lambda_1\lambda_2\lambda_3 = \det S$ and $\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1$ are invariant under rotation of axes. They are called the *principal invariants* of the quadric.

Metric Tensor

In example 7.14 we introduced the metric tensor g_{ij} contracted with the differential $dx_i dx_j$ wich is the scalar $ds^2 = g_{ij} dx_i dx_j$, called the *line element*. In case $g_{ij} = \delta_{ij}$ then $ds^2 = g_{ij} dx_i dx_j = dx_i dx_i$. Without loss of generality we can assume g_{ij} to be symmetric[l]. Integrating along a path $\boldsymbol{\gamma}$ and using that along the path $dx_i = \frac{d\gamma_i}{dt} dt$ then the integral $\int_{\boldsymbol{\gamma}} \sqrt{ds^2}$ leads to

$$\int_{\boldsymbol{\gamma}} \sqrt{ds^2} \quad = \quad \int_{\boldsymbol{\gamma}} \sqrt{dx_i dx_i} = \int_a^b \left\| \frac{d\boldsymbol{\gamma}}{dt} \right\| dt \ , \tag{279}$$

which is simply the line integral along the path $\boldsymbol{\gamma}$ and therefore the length of the path $\boldsymbol{\gamma}$. Therefore $ds^2 = dx_i dx_i$ simply defines the Euclidean structure of the space $\mathbb{R}^3$ (or more generally

[k] Note that $x_{ij} = s_{ij} + a_{ij}$ with s_{ij} symmetric and a_{ij} antisymmetric implies immediately that $x_{ij} + x_{ji} = 2s_{ij}$ and therefore the decomposition is unique, i.e. there are no such other two tensors s_{ij} and a_{ij}.

[l] Due to the symmetry of $dx_i dx_j$ any antisymmetric part of g_{ij} would anyway disappear in the inner product $g_{ij} dx_i dx_j$ according to [101]. Therefore, we might as well assume that g_{ij} does not have an antisymmetric part.

$\mathbb{R}^n$). The Euclidean metric measures distances of two points using the rule of Pythagoras[m]. ds^2 defines the distance between neighbouring points and integrating over ds along a path defines the length of the path. For the Euclidean metric δ_{ij} every non-trivial vector is an eigenvector with respect to the eigenvalue 1. But sometimes we may want to measure lengths and distances in different directions with different weight factors. For example, if the structure of the space is such that it is twice as difficult to move in direction x than it is to move in direction y or z then we may want to use the metric

$$G = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{280}$$

instead of the Euclidean metric to measure distances in such a space. Generally, the symmetric *metric tensor* g_{ij} defines the structure of local distances and the integral $\int_\gamma \sqrt{g_{ij}dx_i dx_j}$ defines the global distances (i.e. the length of the path $\boldsymbol{\gamma}$) in this structure. Using a parametrisation $\boldsymbol{\gamma}(t)$ we find

$$\int_\gamma \sqrt{g_{ij}dx_i dx_j} \;\; = \;\; \int_a^b \sqrt{g_{ij}\frac{d\gamma_i}{dt}\frac{d\gamma_j}{dt}}\; dt \, . \tag{281}$$

The integrals along paths introduced in Sec. 1 obviously used the Euclidean metric which is generated by the 3 by 3 unit matrix. In contrast, in Einstein's[n] *Theory of Special Relativity* distances in *space-time* are measured by the non-Euclidean line element $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$. This metric is called the *Minkowski*[o] *metric*. For this metric, obviously, ds^2 can be negative. If for two points in space-time $c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 < 0$ then the distance is called *space-like*, if it is > 0 then it is called *time-like* and if it is $= 0$ then it is called *light-like* or *null* since light moves on light-like paths, whilst we are restricted to move on time-like paths!

Acknowledgements

I am extremely grateful for the many comments and corrections I have received from the students attending my lectures on Vector Calculus in Lent 2006 and I would very much welcome to receive any further corrections and comments for improvements (md131@cam.ac.uk). Good luck to all of you for your examinations in June!

[m] Pythagoras of Samos, about 569 BC-475 BC.
[n] Albert Einstein, 1879-1955.
[o] Hermann Minkowski, 1864-1909.