Multivariate Calculus

(Multiple Integrals and Vector Calculus)

Prof Andre Lukas

Rudolf Peierls Centre for Theoretical Physics
University of Oxford

HT 2026

Contents

I Introduction - mathematical background **10**

1 Open sets in $\mathbb{R}^n$ and topology **10**
1.1 Definition of open sets . . . 10
1.2 Properties of open sets . . . 11

2 The main players: functions $\mathbb{R}^n \to \mathbb{R}^m$ **13**
2.1 Basic definitions . . . 13
2.2 Remarks on mathematics and physics . . . 13
2.3 Visualisation . . . 14

3 Limits and continuity **16**
3.1 Limits and their properties . . . 16
3.2 Limits and closed sets . . . 18
3.3 Limits and functions . . . 18

II Differential calculus **20**

4 Partial differential **20**
4.1 Partial derivatives and their properties . . . 20
4.2 Directional derivatives . . . 22

5 Total differential **23**
5.1 Definition of total derivative . . . 23
5.2 Relation between total and partial derivative . . . 24
5.3 The chain rule . . . 26
5.4 Jacobi matrix of inverse function . . . 28

6 Taylor series, implicit functions **30**
6.1 Taylor series . . . 30
6.2 Extremal points . . . 31
6.3 Implicit functions . . . 36

7 Vector fields and curves **39**
7.1 Definition of curves . . . 39
7.2 Curves and the Jacobi matrix . . . 42
7.3 Integrating vector fields along curves . . . 43

8 Differential operators of vector calculus **45**
8.1 Gradient . . . 46
8.2 Divergence . . . 51
8.3 Curl . . . 53
8.4 Relations between grad, div and curl . . . 56
8.5 Poincaré's Lemma . . . 60

III Integration in $\mathbb{R}^n$ **64**

9 Definition of Riemann integral in $\mathbb{R}^n$ **64**
9.1 The Riemann integral in one dimension - a reminder . . . 64
9.2 Riemann integral in multiple dimensions . . . 64

10 Properties of the Riemann integral **67**
10.1 Basic properties . . . 67
10.2 Successive integration . . . 68
10.3 Change of variables - the transformation formula . . . 70

IV Manifolds **78**

11 Manifolds embedded in $\mathbb{R}^n$ **78**
11.1 Basic definition . . . 78
11.2 Tangent space . . . 81
11.3 Metrics on manifolds . . . 84
11.4 Differential operators in different coordinate systems . . . 87

12 Integration over manifolds **92**
12.1 Definition of integral . . . 92
12.2 Integration for special cases and dimensions . . . 96

13 The integral theorem of Gauß **104**
13.1 Compact sets with smooth boundaries . . . 105
13.2 Gauß's theorem . . . 105
13.3 Gauß's theorem in two dimensions . . . 106
13.4 Green identities . . . 108

14 The integral theorem of Stokes **109**
14.1 Smooth boundaries . . . 109
14.2 Stokes theorem . . . 109

Appendices **113**

A Differential forms **113**
A.1 Differential one-forms . . . 113
A.2 Alternating k-forms . . . 116
A.3 Higher-order differential forms . . . 118
A.4 Integration of differential forms and Stokes's theorem . . . 125

List of Figures

1 Examples of topological set properties. The sets are indicated by the shaded area, solid boundaries are part of the set, dashed ones are not. (i) shows an open set, (ii) a closeds set and (iii) a set which is neither open nor closed. . . . 11

2 Graph (left) and contour plot (right) for the function $f(x,y) = xy + x^3 + y^2$. The numbers on the contour lines are the function values. . . . 15

3 Vector plot for the vector field $A(x,y) = (x,y)^T$ (left) and for the vector field $A(x,y) = (-y,x)^T$ (right). . . . 15

4 Contour plot for function $f(x,y,z) = xy + x^3 + y^2 - z^2$ (left) and vector plot for vector field $A(x,y,z) = (y^2, zx, x^3)^T$ (right). . . . 16

5 Illustrating the definition of $\lim_{k\to\infty} x^{(k)} = a$. For any given $\epsilon > 0$ there must be a 'tail' of the sequence, that is all $x^{(k)}$ with $k \geq N$ for a certian N, which is entirely contained in the ball $B_\epsilon(a)$. . . . 17

6 Interpretation of partial derivatives for the example $f(x,y) = x^3 - xy^2$. At the point $(x,y) = (1,1)$ we have $\frac{\partial f}{\partial x}(1,1) = 2$ and $\frac{\partial f}{\partial y}(1,1) = -2$. This means the gradient of the tangent at $(1,1)$ in the x-direction is $+2$ (red line) and the gradient of the tangent at $(1,1)$ in the y-direction is -2 (blue line). . . . 22

7 Unit circle $S^1 \subset \mathbb{R}^2$ (in red) as the zero locus of the function f with $f(x,y) = x^2 + y^2 - 1$. At $(x,y) = (1,0)$ we have $\frac{\partial f}{\partial y}(1,0) = 0$, as indicated by the blue vector, and this means it might be difficult to solve $f(x,y) = 0$ for y in terms of x around this point. This is indeed the case, as there are either two or no values of y for a given $x \in [1-\epsilon, 1+\epsilon]$. . . . 37

8 Unit circle $S^1 \subset \mathbb{R}^2$ (in red), parametrised by the angle t as in Eq. (7.1). The tangent vector $\frac{dX}{dt}(0) = (0,1)^T$ at $(1,0)^T$ (in blue) has been indicated. . . . 40

9 Unit circle $S^1 \subset \mathbb{R}^2$ parametrised by $X(t) = (\cos(t), sin(t))^T$ mapped into polar coordinates. Tangent vector are indicated in blue. . . . 43

10 Integrating a vector field $A : \mathbb{R}^2 \to \mathbb{R}^2$ (blue) along a curve $X : [a,b] \to \mathbb{R}^2$ (red). The vector field $A(X(t))$ at a curve point $X(t)$ is projected into the direction $\frac{dX}{dt}(t)$ of the tangent vector in order to obtain the vector field 'flux' along the curve at that point. The integral of a vector field along a curve is defined by integrating this value over the curve parameter t. . . . 44

11 Contour plot of the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = x^2y^2 - x^2\cos(x+y)$. The plot on the left starts the gradient descent at $(x,y) = (1,1)$, the one on the right at $(x,y) = (-1,-1)$. In either case, we have chosen $\lambda = 0.1$ and have performed 30 steps. . . . 48

12 Calculating the total flow of a two-dimensional vector field $A : \mathbb{R}^2 \to \mathbb{R}^2$ (in blue) through a square of length ϵ around $x \in \mathbb{R}^2$. The four sides of the square are labelled by $(1,+)$, $(1,-)$, $(2,+)$ and $(2,-)$. . . . 52

13 Approximation of a function f by lower and upper piecewise constant functions to define the Riemann integral in one dimension. The lower (upper) integral is defined as the total area of the rectangles on the left (right), in the limit of the rectangle width going to zero. The function is called Riemann integrable if the so-obtained upper and lower values exist and are the same. 65

14 An example of a partition $\mathcal{B} = \{B_1, B_2, B_3, \ldots\}$ of the two-dimensional rectangle $B = [a_1, b_1] \times [a_2, b_2]$. . . . 65

15 The set $S = \{(x,y)^T \in \mathbb{R}^2 \,|\, 0 \leq x \leq 1 \text{ and } 0 \leq y \leq x\}$. . . . 70

16 The cone with height h over B, as defined in Eq. (10.30). . . . 75

17 Manifold defined locally as the common zero locus of functions. . . . 79

18 Manifold defined locally by a parametrisation (a chart). . . . 80

19 Tangent space T_pM of a manifold M at a point $p \in M$. . . . 82

20 The Möbius strip, a non-orientable surface. Starting with the unit normal vector n (in red) in the upper right and continuous continuing the vector ends up with a discontinuity after one revolution. 84
21 The surface with parametrisation $X(t_1, t_2) = (1 + t_1, t_1 - t_2^2/2, t_2 - t_1^2/2)^T$ from Example 11.2. 86
22 Geometrical interpretation of the integration measure in Eq. (12.9) over manifolds M. The standard tangent vectors $v_a(p)$ at $p \in M$ are obtained from the standard unit vectors e_i in the embedding space $\mathbb{R}^n$ via Eq. (11.7). The measure is proportional to the area of the parallelepiped (parallelogram) defined by the vectors $v_a(p)$, in accordance with Eq. (12.10). 94
23 The curve $X(t) = (R\cos(t), R\sin(t), t)^T$ for $t \in [0, 4\pi]$, drawn for $R = 6$. 97
24 The cycloid parametrised by $X(t) = (t - \sin(t), 1 - \cos(t))^T$ in the range $t \in [0, 4\pi]$. 98
25 Integrating a vector field A over a manifold M. The vector field is projected into the direction of the unit normal, n, to M and the component, $n \cdot A$ orthgonal to M is then integrated over the manifold. The value of the integral is interpreted as the total flux of A through M. 101
26 A compact set $M \subset \mathbb{R}^2$ with boundary locally described as the zero set of a function f, such that $U \cap \partial M = \{x \in U \,|\, f(x) = 0\}$. The outer unit normal vector, $\hat{n}$, to ∂M is also indicated. 105
27 A compact subset $\mathcal{S} \subset M$ of a surface $M \subset \mathbb{R}^3$. Note that the chart X maps as described in Def. 14.1. 110
28 Open set $U \subset \mathbb{R}^n$ and tangent space T_pU at a point $p \in U$. 113

List of Examples

1.1 Balls as open sets . . . 11
1.2 A set given by a sequence . . . 12
1.3 A more complicated set . . . 13
3.1 A sequence in $\mathbb{R}^2$ . . . 16
3.2 An example for how to use limit properties . . . 17
3.3 Limit of functions . . . 18
3.4 Heaviside function . . . 19
4.1 Computing partial derivatives . . . 21
4.2 Radius-dependent functions . . . 22
5.1 Examples of Jacobi matrices . . . 25
5.2 Jacobi matrix of coordinate changes . . . 25
5.3 Chain rule and coordinate changes . . . 28
5.4 Derivatives of radius-dependent functions . . . 28
5.5 Jacobian of the inverse function . . . 29
6.1 Extrema for a function $f : \mathbb{R}^3 \to \mathbb{R}$ . . . 32
6.2 Stationary points of quadratic forms . . . 33
6.3 A simple implicit function example . . . 36
6.4 Another implicit function example . . . 38
7.1 Unit circle as a curve . . . 39
7.2 Intersection of two curves . . . 40
7.3 Mapping tangent vectors to polar coordinates . . . 42
7.4 Line integral in two dimensions . . . 44
7.5 Line integral in three dimensions . . . 45
8.1 Gradient of linear and quadratic functions . . . 46
8.2 Gradient of radius-dependent functions . . . 47
8.3 Conservative and non-conservative vector fields . . . 50
8.4 Divergence of a radial vector field . . . 53
8.5 Divergence and curl in three dimensions . . . 55
8.6 Another product rule for the curl . . . 55
8.7 Divergence and curl in terms of the Jacobi matrix . . . 56
8.8 Formula for a double curl . . . 57
8.9 Fundamental solution of Laplace equation . . . 57
8.10 Solution to the wave equation . . . 58
8.11 Fundamental solution to Laplace equation in two dimensions . . . 59
8.12 Conservative vector fields . . . 61
10.1 Splitting up two-dimensional integrals I . . . 69
10.2 Splitting up two-dimensional integrals II . . . 69
10.3 Integration in polar and spherical coordinates . . . 72
10.4 Area of a disk and volume of a three-dimensional ball . . . 73
10.5 Behaviour of volume under scaling and linear transformations . . . 73
10.6 Volume of n-dimensional ball . . . 74
10.7 Volume of a cone . . . 75
10.8 Gaußian integrals . . . 75
10.9 Rotational volumes . . . 76
11.1 Tangent spaces of S^1 and S^2 . . . 82
11.2 Induced metric of surfaces . . . 85

11.3 Induced metric for polar and spherical coordinates . . . 87
11.4 Gradient in polar and spherical coordinates . . . 88
11.5 Divergence in polar and spherical coordinates . . . 90
11.6 Laplacian in polar and spherical coordinates . . . 91
11.7 Laplacian on a sphere . . . 91
12.1 Length of a circle and area of two-dimensional sphere . . . 95
12.2 Integration for n-dimensional spherical coordinates . . . 96
12.3 Length of a spiral . . . 97
12.4 Length of a cycloid . . . 97
12.5 Area of surfaces defined as a zero locus . . . 99
12.6 Area of rotational surfaces . . . 100
12.7 Flux of a vector field through S^2 . . . 101
12.8 Vector field flux through a parabolic surface . . . 102
12.9 Flux of a vector field through the surface of a cylinder . . . 102
13.1 Gauß's theorem in two dimension . . . 108
13.2 Gauß's theorem in three dimension . . . 108
14.1 Stokes's theorem on the surface of a cylinder . . . 111
A.1 Some explicit differential forms . . . 122

List of Applications

6.1 Functionals and the variational principle . . . 33
6.2 Implicit functions on a surface . . . 38
7.1 Projectile motion . . . 41
7.2 Hamiltonian mechanics, total derivatives and curves . . . 41
7.3 Curve integrals in physics . . . 45
8.1 Gradient descent . . . 48
8.2 Training neural networks . . . 48
8.3 Maxwell's equations in vacuum . . . 56
8.4 Line integrals and potential functions in physics . . . 62
8.5 Electrostatics and magnetostatics . . . 62
10.1 Liouville's theorem . . . 76
11.1 Electrostatic potential of a uniformly charged ball . . . 91
13.1 Gauss's law in Electromagnetism written in integral form . . . 107
13.2 Archimedean principle . . . 107
14.1 Ampère-Maxwell's and Faraday's laws in integral form . . . 111
A.1 Alternating forms in three dimensions . . . 118
A.2 Differential forms on $\mathbb{R}^3$ . . . 120
A.3 Maxwell's equations with differential forms . . . 123

Foreword: Multivariate calculus, that is, the calculus of (real) functions which depend on many variables, is one of the basic disciplines of Mathematics and it provides a crucial toolset for physics. Functions whose domain is $\mathbb{R}^n$ (or a suitable subset of $\mathbb{R}^n$) and which can be vector-valued, so whose co-domain is $\mathbb{R}^m$ are ubiquitous in physics. In Classical Mechanics, the trajectory of a mass point is described by a function $\mathbb{R} \to \mathbb{R}^3$, the space-time trajectory of a mass point in Special Relativity by a function $\mathbb{R} \to \mathbb{R}^4$. In Electromagnetism, static electric and magnetic fields are three-dimensional vector fields, that is, functions $\mathbb{R}^3 \to \mathbb{R}^3$, while time-varying electric and magnetic fields are described by functions $\mathbb{R}^4 \to \mathbb{R}^3$. Eventually, you will see Electromagnetism formulated in terms of a four-vector potential, that is, in terms of a vector field $\mathbb{R}^4 \to \mathbb{R}^4$. The underlying physics is usually encoded in (ordinary or partial) differential equations so we need to understand the calculus – differentiation and integration – of such functions. Physics is even more demanding since elements of geometry are obviously involved in describing physical situations. The trajectory of a mass point traces out a curve, that is, a one-dimensional manifold embedded in Euklidean (or Minkowski) space. Generalised coordinates in Lagrangian or Hamiltonian mechanics can take values in non-trivial manifolds. For example, the angle which describes the position of a simple pendulum takes values in the unit circle, S^1. The position of a pendulum which can swing in two dimensions can be described by the points on a two-sphere, S^2, and so forth. In Electromagnetism we might want to study the electric flux through certain given surfaces embedded in $\mathbb{R}^3$. In short, we need to be able to handle curves, surfaces and their higher-dimensional analogues, objects collectively called manifolds, embedded in $\mathbb{R}^n$ and the new features that these manifolds bring to calculus. For example, we need to understand how to integrate over such manifolds, a problem which is more involved than simply integrating over $\mathbb{R}^n$.

As ever, in a mathematics course for physicists there is a tension between mathematical rigour and applications to physics. In the present case, this tension is amplified by the relative small number of lectures available. For example, developing the multi-dimensional integral in detail can take a significant amount of time and compromising on rigour will be required to get to the main results. However, it is important that we respect the logic of the subject and follow the natural mathematical narrative. For example, the essence of calculus is taking limits so it would be negligent to not explain what a limit actually is. Applications to physics are crucial but they should not be intertwined or confused with the basic mathematical story. I will avoid this by indicating physics-related discussions (of which there are many throughout these notes) clearly and by separating them from the main text.

From what was said above, the main characters are functions $\mathbb{R}^n \to \mathbb{R}^m$ and their story comes in three acts: the differential calculus of such functions, their integration over $\mathbb{R}^n$ and, finally, the application of multi-variate calculus to manifolds embedded in $\mathbb{R}^n$.

After a brief introduction which explains the mathematical background and sets the scene, we move on to act 1. We need to introduce partial differentiation and related quantities such as the gradient, the divergence and the curl, the total differential and its relation to the Jacobi matrix and learn how to calculate with these objects. These techniques are applied in a number of ways which include the multi-dimensional Taylor series, understanding extremal points of functions and the implicit function theorem.

For act 2 we need to develop the multi-dimensional integral of functions over $\mathbb{R}^n$ (or suitable subsets thereof), learn how to calculate with such integrals and prove the transformation formula - the higher-dimensional analogue of a change of variable in a one-dimensional integral.

Finally, we need to introduce curves, surfaces etc., that is manifolds embedded in $\mathbb{R}^n$ and understand some of their geometry. This involves introducing a metric, an object which will re-appear in General Relativity where it is the main player. Using such a metric we can define integration of functions over manifolds and, in particular, introduce the notion of length of a curve, area of a surface and its higher-dimensional generalisations. We also need to understand how to integrate vector fields over manifolds, so that we can perform line integrals and integrals of vector fields over surfaces. Interestingly, these integrals

can be defined without reference to a metric. They are subject to powerful theorems, namely Gauß's and Stokes's theorem, which can be viewed as higher-dimensional generalisations of the fundamental theorem of calculus.

The appendices are strictly supplementary and cover related (and important) topics which cannot possibly fit into a lecture course this short. Enjoy (or ignore) at your leisure.

Andre Lukas
Oxford, October 2025

Part I
Introduction - mathematical background

Before we delve into differential calculus we need a few mathematical preparations (which we keep at a minimum). First, we are often interested in functions $U \to \mathbb{R}^m$, where U is a subset of $\mathbb{R}^n$, rather than all of it. This is analogous to considering functions of one real variable on, for example, an interval rather than on all of $\mathbb{R}$. We need to characterise the subsets $U \subset \mathbb{R}^n$ we would like to use as domains for functions - they will turn out to be the open sets. Defining these amounts to introducing a topology on $\mathbb{R}^n$ and we will discuss a number of related topological features. The essence of calculus is taking limits - as opposed to algebra which deals with finite expressions only. It would, therefore, be negligent not to explain what a limit is and we will do this as briefly and informally as possibly. We also need to discuss how limits are used to introduce basic concepts, such as continuity of a function.

1 Open sets in $\mathbb{R}^n$ and topology

1.1 Definition of open sets

Norm and metric
We consider $\mathbb{R}^n$ with coordinates denoted by $x = (x_1, \ldots, x_n)^T$. As a preparation, we require a measure of length on this space and we use the usual Euklidean norm

$$|x| := \sqrt{\sum_{i=1}^{n} x_i^2}\,, \tag{1.1}$$

for this purpose. Based on this norm we can introduce a **metric** (or distance) d by $d(x, y) = |x - y|$, where $x, y \in \mathbb{R}^n$. This metric makes $\mathbb{R}^n$ into what is known as a **metric space**, a mathematical structure which can also be introduced more abstractly (although we will not do this). The n-dimensional **ball** $B_r^n(A)$ around a point $a \in \mathbb{R}^n$ with radius r is then defined by [1]

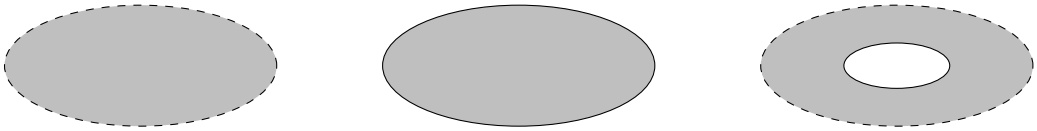
$$B_r^n(a) = \{x \in \mathbb{R}^n \,|\, d(x, a) < r\}\,. \tag{1.2}$$

Note this is the full ball, however, due to the strictly less condition, the bounding sphere is not included. We call a set $U \subset \mathbb{R}^n$ a **neighbourhood** of a point $x \in \mathbb{R}^n$ if there exists a ball $B_\epsilon(x) \subset U$, for some $\epsilon > 0$.

Open sets
Having made these basic definitions, we can now return to the question of which subsets of $\mathbb{R}^n$ should be considered as domains for functions. One of our goals is to definite differentiation and, as you know from the one-dimensional case, this requires taking the limit of a difference quotient. In other words, defining the derivative of a function at a point x requires information about the function in a neighbourhood of x. We should, therefore, consider domains for which every point has a neighbourhood fully contained in the domain. These are precisely the open sets which are formally defined as follows.

Definition 1.1. *A subset $U \subset \mathbb{R}^n$ is called* **open** *if for every $x \in U$ there exists an $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. A subset $C \subset \mathbb{R}^n$ is called* **closed** *if $\mathbb{R}^n \setminus C$ is open.*

[1] We sometimes omit the superscript n if it is clear which dimension we are working in.



(i) (ii) (iii)

Figure 1: Examples of topological set properties. The sets are indicated by the shaded area, solid boundaries are part of the set, dashed ones are not. (i) shows an open set, (ii) a closeds set and (iii) a set which is neither open nor closed.

To get an intuitive idea about these definitions a few example subsets of $\mathbb{R}^2$ are shown in Fig. 1. Fig. 1 (i) shows an open set, Fig. 1 (ii) a closed set (since its complement is open), and Fig. 1 (iii) a set which is neither closed nor open. Note, if a boundary point is included in the set it cannot be open since such a point does not have a neighbourhood entirely contained in the set. Hence, open sets do not contain any of their boundary point. On the other hand, closed sets, being defined as the complement of open sets, should therefore contain all their boundary points. We will make those statements more precise below.

1.2 Properties of open sets

Intersection and union of open sets
A basic question is how open sets relate to set operations such as intersections and unions of sets. The following theorem provides the answers.

Theorem 1.1. *Open sets in $\mathbb{R}^n$ have the following properties.*
(i) The empty set $\{\}$ and the full set, $\mathbb{R}^n$ in our discussion, are open sets.
(ii) For two open sets U_1, U_2 the intersection $U_1 \cap U_2$ is open.
(iii) If U_i is a family of open sets, where $i \in I$ takes values in some index set I, then the union $\bigcup_{i\in I} U_i$ is open.

Proof. (i) The empty set does not contain any points so the openness condition is trivially satisfied. For any $x \in \mathbb{R}$ any ball $B_r(x)$ is also contained in $\mathbb{R}^n$, so $\mathbb{R}^n$ is open.
(ii) Consider a point $x \in U_1 \cap U_2$. Since U_1 and U_2 are open there exist balls $B_{\epsilon_1}(x) \subset U_1$ and $B_{\epsilon_2}(x) \subset U_2$. If we define $\epsilon := \min(\epsilon_1, \epsilon_2)$ then $B_\epsilon(x) \subset U_1 \cap U_2$. Hence, $U_1 \cap U_2$ is open.
(iii) Consider a point $x \in \bigcup_{i\in I} U_i$ in the union. This point must be contained in at least one open set of the family, say $x \in U_i$. Since U_i is open there exists a ball $B_\epsilon(x) \subset U_i$. But, clearly, $B_\epsilon(x) \subset \bigcup_{i\in I} U_i$ and this implies that $\bigcup_{i\in I} U_i$ is open. □

We can turn the above theorem around to define a more general mathematical structure, called **topological space**, on any set X. More precisely, X becomes a topological space by selecting a set of subsets of X, the 'open sets', which satisfy the three properties listed in Theorem 1.1. The theory of topological spaces can be developed within this abstract setting but since we do not really require this level of generality we will not pursue this further.

Example 1.1. Balls as open sets
Question: Show that the n-dimensional ball $B_r(a) \subset \mathbb{R}^n$ with center a and radius r is an open set.

Answer: Consider a point $x \in B_r(a)$. We need to show that there exist a ball abound x which is contained in $B_r(a)$. To this end define $\epsilon = r - |x - a| > 0$ and consider the ball $B_\epsilon(x)$. For another

point $\tilde{x} \in B_\epsilon(x)$ we have

$$|\tilde{x} - a| = |\tilde{x} - x + x - a| \leq |\tilde{x} - x| + |x - a| < \epsilon + r - \epsilon = r \ .$$

Hence $\tilde{x} \in B_r(a)$ and we have shown that $B_\epsilon(x) \subset B_r(a)$, as required.

Example 1.2. A set given by a sequence
Question: Show that the set $S = \{1/n \,|\, n = 1, 2, \ldots\} \subset \mathbb{R}$ is neither open nor closed. What is the minimal closed set which contains S?

Answer: No ball $B_\epsilon(1) = (1 - \epsilon, 1 + \epsilon)$ for $\epsilon > 0$ is not contained in S, so S is not open.

To show that S is not closed consider its complement $\mathbb{R} \setminus S$. For any $\epsilon > 0$ there exists an $n \in \mathbb{Z}^{>0}$ with $1/n < \epsilon$. Hence, for such an n we have $1/n \in B_\epsilon(0)$ but $1/n \in S$ and, therefore, $1/n \notin \mathbb{R} \setminus S$. As a result, the ball $B_\epsilon(0)$ is not contained in $\mathbb{R} \setminus S$ which means $\mathbb{R} \setminus S$ is not open and, hence, its complement S is not closed.

From the last argument it should be intuitively clear that the obstruction to S being closed is that it has a 'limiting point', 0, which is not contained in S. This implies the set $S \cup \{0\}$ is closed.

Exercise 1.2. *Show that the set $X = \mathbb{R}^n \setminus \{a\}$, that is, $\mathbb{R}^n$ with the point a removed, is open.*

Boundary, interior and exterior
It will be useful to think about the boundary, the interior and the exterior of subsets of $\mathbb{R}^n$. More precisely, they are defined as follows.

Definition 1.2. *Let $X \subset \mathbb{R}^n$ be any subset.*
(i) A point $x \in \mathbb{R}^n$ is called an **interior point** *of X if x has a neighbourhood entirely contained in X. The* **interior** *of X, also denoted by $\mathring{X}$, consists of all interior points of X.*
(ii) A point $x \in \mathbb{R}^n$ is called an **exterior point** *of X if it has a neighbourhood which does not intersect X.*
(iii) A point $x \in \mathbb{R}^n$ is called a **boundary point** *of X if every neighbourhood of x contains a point in X and a point in $\mathbb{R}^n \setminus X$. The* **boundary** *of X, also denoted ∂X, consists of all boundary points of X. The* **closure** *$\bar{X}$ of X is defined as $\bar{X} := X \cup \partial X$.*

These sets relate to one another in the way one would intuitively expect, as stated in the next theorem.

Theorem 1.3. *For any subset $X \subset \mathbb{R}^n$ we have the following.*
(i) The interior $\mathring{X}$ of X is an open set and $\mathring{X} = X \setminus \partial X$.
(ii) The closure $\bar{X}$ of X is a closed set.
(iii) The boundary ∂X of X is a closed set.
(iv) The set X is open iff $X = \mathring{X}$.
(v) The set X is closed iff $X = \bar{X}$.

Proof. These proofs involve carefully thinking about the various definitions and are left as an exercise. □

These various notions really just formalise the intuition suggested by the terminology. By way of example, considering the set in Fig. 1 (iii), whose boundary is given by the union of the solid and dashed lines, the interior is the shaded part (without the lines) and the closure is the shaded area and the lines.

Example 1.3. A more complicated set
Question: Consider the set $S = B_1((0,0)^T) \setminus \{(x,y)^T \,|\, x = 0 \text{ or } y = 0\} \subset \mathbb{R}^2$ (the unit disk in $\mathbb{R}^2$ with the coordinate axes taken away). What is the interior, the closure and the boundary of this set?
Answer: It is not too difficult to prove that S is open, so $\mathring{S} = S$ from Theorem 1.3 (iv).
Its closure is the set $\bar{S} = \{(x,y)^T \,|\, x^2 + y^2 \leq 1\}$ and the boundary consists of the two coordinate axes in the range $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$ as well as the bounding circle.

Finally, it is useful to be able to talk about **compact sets** in $\mathbb{R}^n$.

Definition 1.3. *A subset $S \subset \mathbb{R}^n$ is called compact if it is closed and if there exists an $R > 0$ with $S \subset B_R(0)$.*

In other words a compact set is a set that, in addition to being closed, is also 'bounded' in $\mathbb{R}^n$. There is an equivalent definition of compactness in terms of open coverings ("every open covering of the set can be refined to a finite covering") which can be used in much more general settings but we will not get into this here. (See, for example, Ref. [3] for more details.)

2 The main players: functions $\mathbb{R}^n \to \mathbb{R}^m$

2.1 Basic definitions

The main players in this lecture course are functions (or maps) which depend on n real variables and which are vector-valued in m dimensions or, in mathematical parlance, functions $f : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$ is an open subset. Note, we are allowing for the possibility of non-trivial function domains U (in the same way functions of one variable might only be defined on, say, an interval, rather than all of $\mathbb{R}$) but, for the reasons discussed earlier, we restrict these domains to be open subsets of $\mathbb{R}^n$. (Of course taking $U = \mathbb{R}^n$ is a possibility.) We will usually denote the coordinates on the domain by $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and the coordinates in the co-domain by $y = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$. The function $f : U \to \mathbb{R}^m$ can be written in terms of m real-valued functions $f_i : U \to \mathbb{R}$, where $i = 1, \ldots, m$, that is,

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \quad \text{or, dropping the argument,} \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} . \tag{2.1}$$

Certain special choices of domain and co-domain dimensions n and m are of particular importance. If $m = 1$ we talk about a **real-valued function** or, in physics terminology, a **scalar field**. For $n = m$ the function f is also referred to as a **vector field**. In line with physics notation, we will often denote vector fields by uppercase letters, $A, B, \ldots$.

2.2 Remarks on mathematics and physics

Function classes
It may be worth reviewing some standard classes of functions – we will consider three basic types. We do this for real-valued functions $f : U \to \mathbb{R}$ which depend on coordinates $x = (x_1, \ldots, x_n)^T$. (Vector-valued functions can of course be build up from real-valued ones, using Eq. (2.1).) The first and simplest class is **polynomials** in the coordinates $x_1, \ldots, x_n$. It is helpful to introduce a compact notation to write these

down and, to this end, we introduce multi-indices $I = (i_1, \ldots, i_n)$ and write monomials as $x^I := x_1^{i_1} \cdots x_n^{i_n}$. A polynomial $p : \mathbb{R}^n \to \mathbb{R}$ can then be written as

$$p(x) = \sum_I a_I x^I \,, \tag{2.2}$$

where, crucially, only a finite number of the coefficients $a_I \in \mathbb{R}$ are non-zero.

The second class of functions is **rational functions** $r : U \to \mathbb{R}$ which are simply defined as ratios of two polynomials p and q, so

$$r(x) = \frac{p(x)}{q(x)} \,. \tag{2.3}$$

Algebra is the broad mathematical discipline which deals with finite expressions (no limits are being taken) so both polynomials and rational functions belong to the realm of algebra.

The third class of functions are transcendental functions which are functions given as limits of sequences or series, often as a power series, that is,

$$f(x) = \sum_I a_I x^I \,, \tag{2.4}$$

where now infinitely many of the coefficients $a_I \in \mathbb{R}$ can be non-zero. Since their definition involves a limit they are not part of algebra but firmly belong to the realm of **calculus** (or analysis), the broad mathematical discipline where we allow taking limits.

Dimensions and physics
In a physics context the domain dimension $n = 3$ is of obvious importance to describe spatial space and we have to consider scalar and vector fields in this context. However, frequently physical systems can be understood in terms of one- or two-dimensional models, so $n = 1, 2$ are clearly in play as well. In Special Relativity we consider space-time so we have to allow for scalar and vector fields defined on a four-dimensional domain. But there are spaces other than physical space-time which we have to consider in physics. For example, the three-dimensional motion of N point particles in classical (Hamiltonian) mechanics is described by $3N$ coordinates and $3N$ momenta, which reside in a space of total dimension $6N$ (the so-called phase space). In conclusion, there are physics applications with basically any domain dimension, so we should develop the subject keeping this dimension arbitrary (while paying special attention to the dimensions which are particularly important in physics).

2.3 Visualisation

Visualisation in two dimensions
How can we visualise scalar and vector fields? Let's start with a scalar field $f : U \to \mathbb{R}$, with $U \subset \mathbb{R}^2$ in $n = 2$ dimensions and denote the coordinates by $(x_1, x_2) = (x, y)$. Then we have two options. We can either plot the graph of f, that is, the set of points $\{(x, y, f(x, y))^T \,|\, (x, y)^T \in U\} \subset \mathbb{R}^3$ or we can create a contour plot of f, that is, we can draw in U the level lines $\{(x, y)^T \in U \,|\, f(x, y) = c\}$, that is, the lines where the function f assumes certain given values c. An example for a specific scalar function is shown in Fig. 2. Interesting information about the function is often easier to read off from the contour plot, than from the graph.

What about vector fields $A : U \to \mathbb{R}^2$, where $U \subset \mathbb{R}^2$? A nice way to visualise these is to, literally, plot

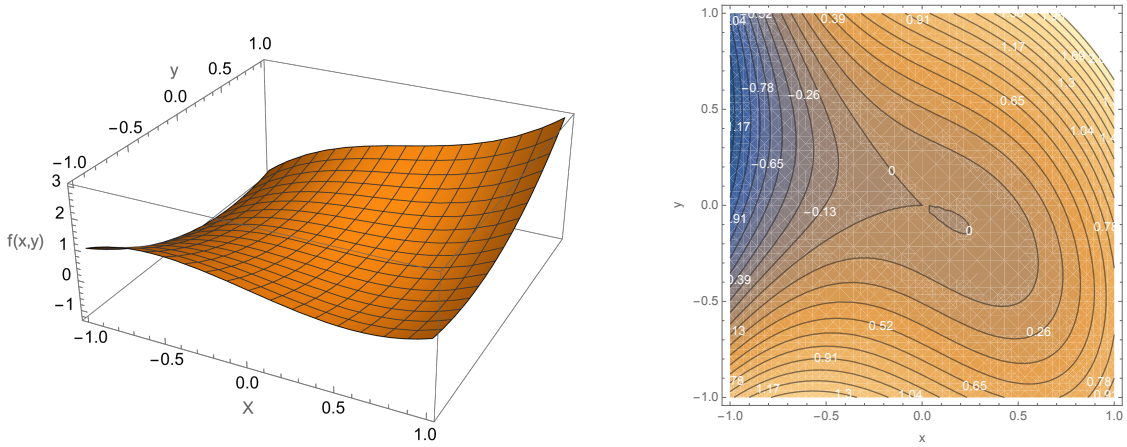


Figure 2: Graph (left) and contour plot (right) for the function $f(x, y) = xy + x^3 + y^2$. The numbers on the contour lines are the function values.

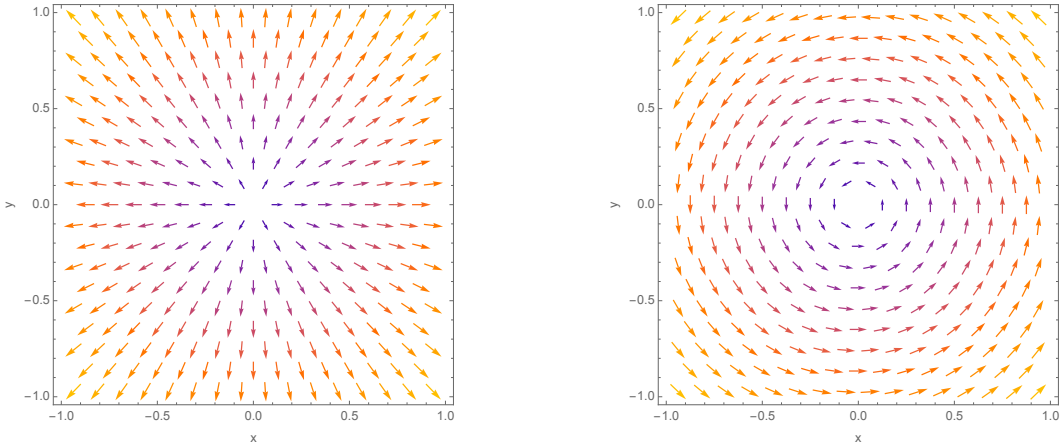


Figure 3: Vector plot for the vector field $A(x, y) = (x, y)^T$ (left) and for the vector field $A(x, y) = (-y, x)^T$ (right).

a family of vectors. More specifically, we choose a family of points $x \in U$ and attach to each of those the vector $f(x)$, that is, the vector-value of f at this point. Two examples are shown in Fig. 3.

Visualisation in three dimensions

Drawing functions with a three-dimensional domain becomes more difficult. For a scalar field $f : U \to \mathbb{R}$, with domain $U \subset \mathbb{R}^3$ and coordinates $(x_1, x_2, x_3) = (x, y, z)$ it is not really possibly to draw the graph $\{(x, y, z, f(x, y, z))^T \,|\, (x, y, z)^T \in U\} \subset \mathbb{R}^4$, as this would require being able to visualise a fourth dimension. We can have some fun and try to draw a contour plot, that is, draw the level surfaces $\{(x, y, z)^T \in U \,|\, f(x, y, z) = c\} \subset \mathbb{R}^3$ for certain values of c. For an example this has been done in Fig. 4 on the left – leading to nice-looking plot which is difficult to interpret. An example for a vector field plot in three dimensions is shown in Fig. 4 on the right.

In physics, the problem of visualisation often simplifies due to symmetry. For example, a Newton or Coulomb potential for a point source is essentially given by $V(x, y, z) = \frac{1}{r}$, where $r = \sqrt{x^2 + y^2 + z^2}$ is the radial function. The reason V depends on x, y, z only via the radius r is of course rotational symmetry.

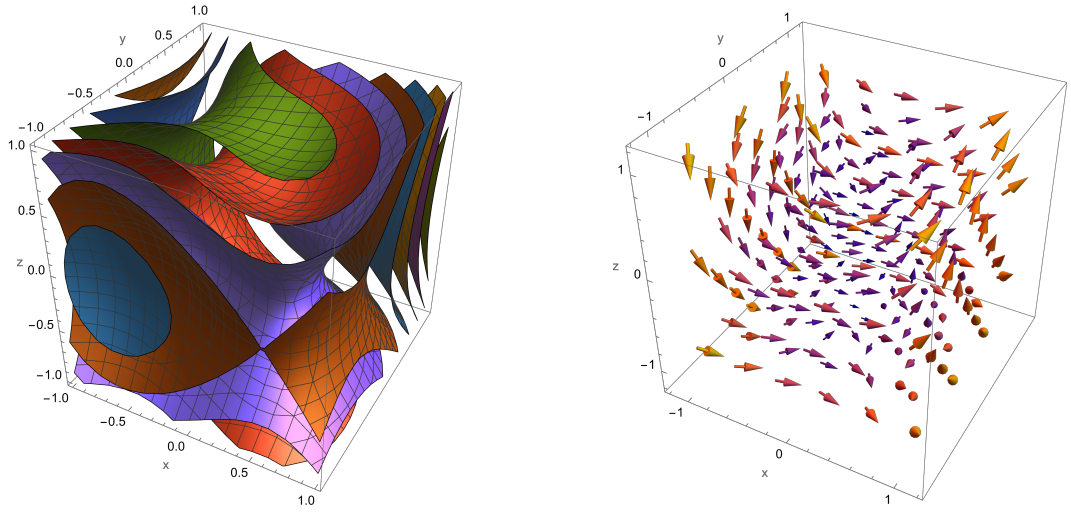


Figure 4: Contour plot for function $f(x, y, z) = xy + x^3 + y^2 - z^2$ (left) and vector plot for vector field $A(x, y, z) = (y^2, zx, x^3)^T$ (right).

Instead of creating a three-dimensional (contour) plot of V as a function of x, y, z it is much better to draw V as a function of the single variable r.

3 Limits and continuity

3.1 Limits and their properties

Definition of limits
Consider a sequence $(x^{(1)}, x^{(2)}, \ldots)$, for short also written as $(x^{(k)})$, of vectors $x^{(k)} \in \mathbb{R}^n$. We want to define what it means for such a sequence to converge to a vector $a \in \mathbb{R}^n$.

Definition 3.1. *A sequence $(x^{(k)})$ in $\mathbb{R}^n$ is said to converge to $a \in \mathbb{R}^n$ if for every $\epsilon > 0$ we can find an integer N such that $|x^{(k)} - a| < \epsilon$ for all $k \geq N$. In this case we write $\lim_{k\to\infty} x^{(k)} = a$.*

This does sound a bit convoluted at first but it really just captures the intuition for what it means for a sequence to converge to a particular point a. For any distance ϵ, we have to be able to find a 'tail' of the sequence whose distance from the prospective limit point is less than ϵ or, in other words, which is entirely contained within the ball $B_\epsilon(a)$. This is illustrated in Fig. 5.

Example 3.1. A sequence in $\mathbb{R}^2$
Question: Consider the sequence $(x^{(k)})$ in $\mathbb{R}^2$ defined by $x^{(k)} = (1/k, 1/k)^T$. Show that this sequence converges to the origin.

Answer: Start by choosing any $\epsilon > 0$. We need to identify a 'tail' of the sequence which is entirely contained in $B_\epsilon((0,0)^T)$. The distance of $(1/k, 1/k)^T$ from the origin is $\frac{1}{\sqrt{2}k}$ so let us choose on N such that $\frac{1}{\sqrt{2}N} < \epsilon$. Then, for any $k \geq N$ we have $\frac{1}{\sqrt{2}k} < \epsilon$ and, hence, $(1/k, 1/k)^T \in B_\epsilon((0,0)^T)$ for all $k \geq N$. This shows that the sequence converges to the origin.

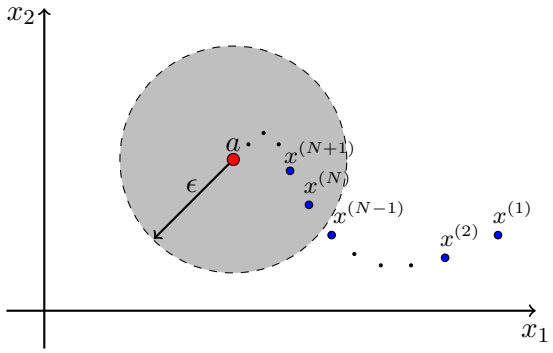


Figure 5: Illustrating the definition of $\lim_{k\to\infty} x^{(k)} = a$. For any given $\epsilon > 0$ there must be a 'tail' of the sequence, that is all $x^{(k)}$ with $k \geq N$ for a certian N, which is entirely contained in the ball $B_\epsilon(a)$.

Properties of limits

It is useful to have some rules of calculation for limits so we can (sometimes) avoid working with the somewhat unwieldy Definition 3.1.

Theorem 3.1. *Suppose for two sequences* $(a^{(k)})$ *and* $(b^{(k)})$ *in* $\mathbb{R}^n$ *the limits* $\lim_{k\to\infty}(a^{(k)})$ *and* $\lim_{k\to\infty}(b^{(k)})$ *exist. Then we have:*

(i) The limit of $(a^{(k)} + b^{(k)})$ *exists and* $\lim_{k\to\infty}(a^{(k)} + b^{(k)}) = \lim_{k\to\infty}(a^{(k)}) + \lim_{k\to\infty}(b^{(k)})$.

(ii) For $n = 1$*, the limit of* $(a^{(k)}b^{(k)})$ *exists and* $\lim_{k\to\infty}(a^{(k)}b^{(k)}) = (\lim_{k\to\infty}(a^{(k)}))(\lim_{k\to\infty}(b^{(k)}))$.

(iii) For $n = 1$ *and if* $\lim_{k\to\infty} b^{(k)} \neq 0$ *then the limit of* $(a^{(k)}/b^{(k)})$ *exists and* $\lim_{k\to\infty}(a^{(k)}/b^{(k)}) = (\lim_{k\to\infty}(a^{(k)}))/(\lim_{k\to\infty}(b^{(k)}))$.

(iv) For vector sequences limits can be computed component-wise, that is $\left[\lim_{k\to\infty}(a^{(k)})\right]_i = \lim_{k\to\infty}(a_i^{(k)})$ *for* $i = 1, \ldots n$.

Proof. (i) Suppose that $\lim_{k\to\infty}(a^{(k)}) = a$ and $\lim_{k\to\infty}(b^{(k)}) = b$. Then, for any $\epsilon > 0$ there exists an N such that for all $k \geq N$ we have $|a^{(k)} - a| < \epsilon/2$ and $|b^{(k)} - b| < \epsilon/2$. It follows (from the triangle inequality) that $|a^{(k)} + b^{(k)} - (a + b)| \leq |a^{(k)} - a| + |b^{(k)} - b| < \epsilon/2 + \epsilon/2 = \epsilon$. This shows that the limit of $(a^{(k)} + b^{(k)})$ exists and is given by $a + b$.
(ii), (iii) Note these properties only make sense for $n = 1$ since there is no general multiplication or division of vectors. These proofs and the one for (iv) are left as exercises. □

Example 3.2. An example for how to use limit properties
Question: Consider the sequence $(x^{(k)})$ in $\mathbb{R}$ given be $x^{(k)} = \frac{1+k}{1+2k}$. Show that this sequence converges and find its limit.

Answer: Using the properties from Theorem 3.1 we can proceed as follows:

$$\lim_{k\to\infty} \frac{1+k}{1+2k} = \lim_{k\to\infty} \frac{1+1/k}{2+1/k} = \frac{\lim_{k\to\infty}(1+1/k)}{\lim_{k\to\infty}(2+1/k)} = \frac{1+\lim_{k\to\infty} 1/k}{2+\lim_{k\to\infty} 1/k} = \frac{1}{2}\,.$$

Hence the sequence converges to $1/2$.

3.2 Limits and closed sets

It is intuitively clear that the boundary points of a subset $X \subset \mathbb{R}^n$ (which are not already contained in X) can be obtained as limits of sequences within X. To formulate this more precisely we say that $x \in \mathbb{R}^n \setminus X$ is a **limit point** of a set $X \subset \mathbb{R}^n$ is there exists a sequence $(x^{(k)})$ with all $x^{(k)} \in X$ and $\lim_{k\to\infty}(x^{(k)}) = x$.

Theorem 3.2. *The closure $\bar{X}$ of a subset $X \subset \mathbb{R}^n$ is given by X together with all limit points of X.*

Proof. We show this by proving mutual inclusion.
'$\subset$' : There is nothing to show for a points in $\bar{X}$ which are also in X, so let us start with a point $x \in \bar{X} \setminus X$. Since $\bar{X} = X \cup \partial X$ it follows that $x \in \partial X$, that is, x is a boundary point. By definition, every neighbourhood of a boundary point contains points in X and in the complement of X. Let us choose one such neighbourhood which, by definition, must contain a ball $B_\epsilon(x)$. Then, all balls $B_{\epsilon/k}(x)$ for $k = 1, 2, \ldots$ are neighbourhoods of x and must each contain a point $x^{(k)} \in B_{\epsilon/k}(x) \cap X$. The sequence (x^k) is entirely contained in X and clearly converges to x. Hence x must be a limit point of X.
'$\supset$' : Let x be a limit point of X. We have to show that x is a boundary point. Suppose there exist a neighbourhood of x entirely contained in X. Then it follows that $x \in X$ but, as a limit point, $x \notin X$, so such a neighbourhood cannot exist. On the other hand, assume there is a neighbourhood of x contained in the complement of X. This implies we have a ball $B_\epsilon(x)$ which does not intersect X. Hence, every point $\tilde{x} \in X$ is a distance of at least ϵ away from x. This means there cannot be a sequence contained in X which converges to x. But this contradict x being a limit point of X so we conclude that x does not have a neighbourhood entirely contained in the complement of X. In conclusion, every neighbourhood of x must have point in X and its complement and this means that x is a boundary point. □

3.3 Limits and functions

Definition of function limits

It is useful to have a continuous version of the limit which can be applied to functions. Consider a function $f : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$ open. For $a \in U$ and $A \in \mathbb{R}^m$ we write $\lim_{x\to a} f(x) = A$ if every sequence $(x^{(k)})$ in U with $\lim_{k\to\infty} x^{(k)} = a$ satisfies $\lim_{k\to\infty} f(x^{(k)}) = A$. From this definition which is based on the limit of sequences it is clear that the continuous limit satisfies the analogue of the properties in Theorem 3.1, where properties (ii) and (iii) of course require real-valued functions, so that multiplication and division makes sense.

Example 3.3. Limit of functions
Question : What is $\lim_{x\to 0} \frac{1+x^{-1}}{1+x^{-2}}$?
Answer: We can use the rules from Theorem 3.1 to calculate with the limit.

$$\lim_{x\to 0} \frac{1+x^{-1}}{1+x^{-2}} = \lim_{x\to 0} \frac{x^2+x}{x^2+1} = \frac{\lim_{x\to 0}(x^2+x)}{\lim_{x\to 0}(x^2+1)} = \frac{\lim_{x\to 0}(x^2)+\lim_{x\to 0}(x)}{\lim_{x\to 0}(x^2)+\lim_{x\to 0}(1)} = \frac{0}{1} = 0$$

Continuity

You probably have an intuitive understanding of what continuity of a function means. Here is the formal definition.

Definition 3.2. *(Continuity) A function $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is* **continuous** *at $a \in U$ if $\lim_{x\to a} f(x) = f(a)$. The function f is called continuous on U if it is continuous for all $a \in U$.*

From Theorem 3.1 it is clear that for two functions $f, g : U \to \mathbb{R}^m$ which are both continuous at $a \in U$ the sum $f + g$ is also continuous at a. Further, if the functions are real-valued, so $m = 1$, then the product

function fg is continuous at a and, provided $g(a) \neq 0$, the quotient function f/g is continuous at a. From Theorem 3.1 (iv) we conclude that vector-valued functions $f = (f_1, \ldots, f_m)^T : U \to \mathbb{R}^m$ are continuous if and only if all its component functions f_i are continuous.

Exercise 3.3. Continuity and function composition
Consider a function $f : U \to \mathbb{R}^n$, with $U \subset \mathbb{R}^n$ open, which is continuous at $a \in U$. Further, consider a second function $g : V \to \mathbb{R}^p$, with $V \subset \mathbb{R}^m$ open and $f(U) \subset V$, which is continuous at $f(a)$. Show that $g \circ f$ is continuous at a.

Function classes and continuity
From these statements it is easy to show continuity for large classes of functions quickly. As before, we focus on real-valued functions $f : U \to \mathbb{R}$ with coordinates $x = (x_1, \ldots, x_n)^T$ which can then be used to build up vector-valued ones via Eq. (2.1). The constant polynomial $p(x) = c$, for a constant $c \in \mathbb{R}$ and the polynomial $p(x) = x_i$, for a given $i \in \{1, \ldots, n\}$, are both easily seen to be continuous, using the definition of continuity in Def. 3.2. By taking products (for same or different values of i) and sums of such functions it follows (from the remarks below Def. 3.2) that all polynomials in the coordinates x_i are continuous. We recall that rational functions r are ratios of two polynomials, $r(x) = p(x)/q(x)$. From Theorem 3.1 (iii) they are continuous for all x for which $q(x) \neq 0$.

Things are more complicated for transcendental functions where we would need some statements about the continuity of the limit of a function sequence. However, we can take for granted the continuity of standard transcendental functions of one variable, such as exp, sin, cos etc.. (Proving this is really the task of a calculus course for functions in one real variable). Then, from Exercise 3.3, we can compose such functions with each other or with polynomials/rational functions to obtain new continuous real-valued functions.

Finally, it might be useful to present an example of a function which is not continuous.

Example 3.4. Heaviside function
Question: Show that the Heaviside function $\theta : \mathbb{R} \to \mathbb{R}$ defined by

$$\theta(x) = \begin{cases} 1 & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x < 0 \end{cases} \tag{3.1}$$

is not continuous at $x = 0$.

Answer: From Def. 3.2, we need to show there exists a sequence $(x^{(k)})$ which converges to zero but whose function values $\theta(x^{(k)})$ do not converge to $\theta(0) = 1$. Define the sequence $x^{(k)} = -1/k$. Then it is clear that $\lim_{k\to\infty} x^{(k)} = 0$. However, $\theta(x^{(k)}) = 0$ for all $k = 1, 2, \ldots$ so we conclude that $\lim_{k\to\infty} \theta(x^{(k)}) = 0 \neq 1 = \theta(0)$. If follows that the Heaviside function is not continuous at $x = 0$.

Part II
Differential calculus

We are now ready to introduce differentiation. We begin with the partial derivative which is really a straightforward application of the derivative in one dimension to the multi-dimensional case. Afterwards we introduce the genuine multi-dimensional derivative, the **total derivative** and explain how it relates to partial derivatives. A key result for total derivatives will be the chain rule, a generalisation of the familiar rule by the same name in one dimension. These results are applied in a number of ways, including to the multi-dimensional Taylor series, implicit functions and vector fields and their curve integrals. We end this part with a discussion of differential operators, notably the gradient, the curl and the divergence.

4 Partial differential

4.1 Partial derivatives and their properties

Definition of partial derivative
As before, we consider functions $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, and domain coordinates $x = (x_1, \ldots, x_n)^T$. We also introduce the standard unit vectors e_i on $\mathbb{R}^n$, defined in the usual way as $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, where the '1' is in the i^{th} position. If we keep all but the i^{th} argument of f constant we can think of f as a function of a single variable x_i. To this single-variable function we can apply the standard one-dimensional definition of the derivative and this is what we mean by the partial derivative of f with respect to x_i. More precisely, this is defined as follows.

Definition 4.1. *(Partial differential) A function $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is called* **partially differentiable** *at $x \in U$ with respect to x_i if the limes*

$$\lim_{\epsilon \to 0} \frac{f(x + \epsilon e_i) - f(x)}{\epsilon} \tag{4.1}$$

exists. If it does this limes is called the **partial derivative** *of f at x with respect to x_i and is denoted by $\frac{\partial f}{\partial x_i}(x)$ or $\partial_i f(x)$, for short. The function f is called partially differentiable at $x \in U$ if it is partially differentiable at x with respect to all x_i and it is called partially differentiable on U if it is partially differentiable for all $x \in U$.*

Since this is the same as the definition for functions of a single variable it is already clear how to carry out partial derivatives in practice. All the standard rules of differentiation for single variable functions are applied to x_i, while all other x_j, for $j \neq i$, are treated as constants.

There is an equivalent and often useful way to formulate Def. 4.1. A function $f : U \to \mathbb{R}^m$ is partially differentiable at $x \in U$ with respect to x_i if there exists a number $\frac{\partial f}{\partial x_i}(x)$ such that

$$f(x + \epsilon e_i) = f(x) + \frac{\partial f}{\partial x_i}(x)\, \epsilon + \mathcal{O}(\epsilon^2) \tag{4.2}$$

where $\mathcal{O}(\epsilon^k)$ represents an expression of order ϵ^k which is, more precisely, defined as an expression which satisfies

$$\lim_{\epsilon \to 0} \frac{\mathcal{O}(\epsilon^k)}{\epsilon^{k-1}} = 0\,. \tag{4.3}$$

Note, Eq. (4.2) captures the interpretation of the derivative as a local linear approximation. The $\mathcal{O}(\epsilon^k)$ notation often provides a good match between intuition and mathematical precision. We also say that f

is k times **continuously partially differentiable** on U if

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x) \tag{4.4}$$

exists for all $x \in U$ and is continuous for all $i_1, \ldots, i_k$.

Properties of partial derivative

As explained above, the partial derivative is, effectively, a one-dimensional derivative and satisfies the same properties as derivatives for functions of a single variable. First of all it is linear, that is, for two functions $f, g : U \to \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$ we have

$$\frac{\partial}{\partial x_i}(\alpha f + \beta g)(x) = \alpha \frac{\partial f}{\partial x_i}(x) + \beta \frac{\partial g}{\partial x_i}(x) \; . \tag{4.5}$$

Further, if both f, g are real-valued, so that they can be multiplied, the product rule applies:

$$\frac{\partial}{\partial x_i}(fg)(x) = f(x) \frac{\partial g}{\partial x_i}(x) + g(x) \frac{\partial f}{\partial x_i}(x) \; . \tag{4.6}$$

There is one new feature of partial derivatives which has no analogue in the one-dimensional case. What happens if we change the order of partial derivatives in different directions, that is, how do the expressions

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \; , \qquad \frac{\partial^2 f}{\partial x_j \partial x_i}(x) \tag{4.7}$$

relate? Fortunately, there is a simple answer.

Theorem 4.1. *If $f : U \to \mathbb{R}^m$ is continuously partially differentiable twice at $x \in U$ then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) \; . \tag{4.8}$$

Proof. The proof is a bit technical and can be found, for example, in Ref. [7]. □

Put simply this means that partial derivatives in different directions commute.

Interpretation of partial derivatives

The derivative of a function of one variable provides the gradient of the tangent to the function. Def. 4.1 shows that partial derivatives have a similar interpretation, except there is now more freedom for the direction of the tangent. More precisely, for a real-valued function $f : U \to \mathbb{R}$ the partial derivative $\frac{\partial f}{\partial x_i}(x)$ gives the gradient of the tangent at x in the direction of the coordinates axis x_i. For a function of two variables this is illustrated in Fig. 6.

Example 4.1. Computing partial derivatives

Question: Compute the partial derivatives of the functions $f : \mathbb{R}^2 \to \mathbb{R}$ and $g; \mathbb{R}^3 \to \mathbb{R}$ given by $f(x, y) = x \sin(y) + y \cos(x)$ and $g(x, y, z) = x^2 e^{-y-z} + xyz$.

Answer: By straightforward calculation we have

$$\frac{\partial f}{\partial x} = \sin(y) - y \sin(x) \; , \qquad \frac{\partial f}{\partial y} = x \cos(y) + \cos(x) \; ,$$

and

$$\frac{\partial g}{\partial x} = 2xe^{-y-z} + yz \; , \qquad \frac{\partial g}{\partial y} = -x^2 e^{-y-z} + xz \; , \qquad \frac{\partial g}{\partial z} = -x^2 e^{-y-z} + xy \; .$$

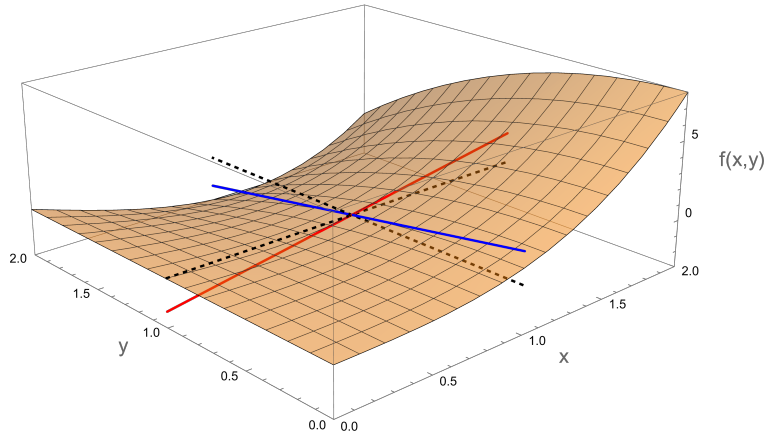


Figure 6: Interpretation of partial derivatives for the example $f(x, y) = x^3 - xy^2$. At the point $(x, y) = (1, 1)$ we have $\frac{\partial f}{\partial x}(1, 1) = 2$ and $\frac{\partial f}{\partial y}(1, 1) = -2$. This means the gradient of the tangent at $(1, 1)$ in the x-direction is $+2$ (red line) and the gradient of the tangent at $(1, 1)$ in the y-direction is -2 (blue line).

Example 4.2. Radius-dependent functions
Question: Work out the partial derivatives of the 'radius function' $r : \mathbb{R}^n \to \mathbb{R}$, defined by $r(x) = |x|$.

Answer: More explicitly, the radius function is $r(x) = \sqrt{\sum_{j=1}^n x_j^2}$. Its partial derivative is

$$\frac{\partial r}{\partial x_i} = \frac{x_i}{\sqrt{\sum_j x_j^2}} = \frac{x_i}{r} \, .$$

This is a useful formula in situations where radius-dependent functions arise, for example in electromagnetism.

4.2 Directional derivatives

The partial derivatives $\frac{\partial f}{\partial x_i}(x)$ provide the gradients of the tangent in the directions of the coordinates axes. But what if we are interested in the gradient of the tangent at x in some more general direction $v \in \mathbb{R}^n$? Generalising Eq. (4.2), we can define such a general directional derivative by taking the limit of a difference quotient in the direction v (rather than e_i). In other words, we say the function $f : U \to \mathbb{R}^m$ is differentiable at $x \in U$ in the direction $v \in \mathbb{R}^n$ if there exists a number $\partial_v f(x)$ such that

$$f(x + \epsilon v) = f(x) + \partial_v f(x)\epsilon + \mathcal{O}(\epsilon^2) \, . \tag{4.9}$$

Let us write $v = \sum_{i=1}^n v_i e_i$ and apply Eq. (4.2) repeatedly:

$$\begin{aligned} f(x + \epsilon v_1 e_1 + \cdots + \epsilon v_n e_n) &= f(x + \epsilon v_2 e_2 + \cdots + \epsilon v_n e_n) + \frac{\partial f}{\partial x_1}(x) v_1 \epsilon + \mathcal{O}(\epsilon^2) \\ &= f(x + \epsilon v_3 e_3 + \cdots + \epsilon v_n e_n) + \frac{\partial f}{\partial x_1}(x) v_1 \epsilon + \frac{\partial f}{\partial x_2}(x) v_2 \epsilon + \mathcal{O}(\epsilon^2) \\ &= f(x) + \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(x) \epsilon + \mathcal{O}(\epsilon^2) \end{aligned} \tag{4.10}$$

Comparison of the last line with Eq. (4.9) shows that the directional derivative can be written in terms of partial derivatives as

$$\partial_v f(x) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(x) \; . \tag{4.11}$$

This equation shows why it is sufficient to work with partial derivatives. If f is partially differentiable at x then all directional derivatives at x exist as well and they are obtained as linear combinations of the partial ones.

5 Total differential

So far we have looked at individual partial derivatives but there is a more general notion of derivative which does not rely on choosing particular directions (such as the directions of the coordinate axes). This notion is referred to as **total derivative**.

5.1 Definition of total derivative

To introduce the total derivative we generalise the definition of the partial derivative, using the formulation in Eq. (4.2).

Definition 5.1. *(Total derivative) A function $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is called totally differentiable at $x \in U$ if there exists an $m \times n$ matrix A such that*

$$f(x + \epsilon) = f(x) + A\epsilon + \mathcal{O}(|\epsilon|^2) \tag{5.1}$$

for all $\epsilon \in \mathbb{R}^n$ in a sufficiently small ball $B_r(0)$. In this case, the matrix A is called the total derivative or Jacobi matrix for f at x and is also denoted by $Df(x) := A$.

This is where linear algebra meets calculus. In the case of real-valued functions of one variable we have a somewhat trivial (1×1) matrix $f'(x) : \mathbb{R} \to \mathbb{R}$ which approximates the function locally. In the multi-dimensional case, for a function f with n arguments and m-dimensional values, this becomes a non-trivial linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$, that is, an $m \times n$ matrix. It is immediately clear from the above definition that, just like its counterpart in one dimension, the total differential is linear. This means for two functions $f, g : U \to \mathbb{R}^m$ which are totally differentiable at $x \in U$, and scalars $\alpha, \beta \in \mathbb{R}$ it follows that the linear combination $\alpha f + \beta g : U \to \mathbb{R}^m$ is totally differentiable at $x \in U$ and $D(\alpha f + \beta g)(x) = \alpha\, D(f)(x) + \beta\, D(g)(x)$.

You probably know (or perhaps you don't?) that a function $f : \mathbb{R} \to \mathbb{R}$ in one dimension which is differentiable at $x \in \mathbb{R}$ must be continuous at x. Here is the multi-dimensional generalisation of this statement.

Proposition 5.1. *If $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is totally differentiable at $x \in U$, then it is continuous at x.*

Proof. We have to show the continuity property in Def. 3.2 so we start with a sequence $(x^{(k)})$ which converges to x. This sequence can also be written as $x^{(k)} = x + \epsilon^{(k)}$, where $\lim_{k\to\infty} \epsilon^{(k)} = 0$. Total differentiability at x implies that $f(x^{(k)}) = f(x + \epsilon^{(k)}) = f(x) + A\epsilon^{(k)} + \mathcal{O}(|\epsilon^{(k)}|^2)$ and taking the limit of this equation gives

$$\lim_{k\to\infty} f(x^{(k)}) = \lim_{k\to\infty} \left(f(x) + A\epsilon^{(k)} + \mathcal{O}(|\epsilon^{(k)}|^2) \right) = f(x) \; .$$

This shows f is continuous at x. □

5.2 Relation between total and partial derivative

General relationship
It is intuitively clear that the total and partial derivatives must be related and the precise relationship is formulated in the following proposition.

Proposition 5.2. *If $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is totally differentiable at $x \in U$ then it is partially differentiable with respect to all variables and the total derivative is given by*

$$Df(\mathbf{x}) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & \ddots & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix} (x) \tag{5.2}$$

Proof. Assume that f is totally differentiable at $x \in U$, with Jacobi matrix $A = Df(x)$, such that $f(x + \epsilon) = f(x) + A\epsilon + \mathcal{O}(|\epsilon|^2)$. Focusing on the i^{th} component of f this can be written as $f_i(x + \epsilon) = f_i(x) + A_{ij}\epsilon_j + \mathcal{O}(|\epsilon|^2)$. Since this holds for all (sufficiently small) ϵ we can choose a particular $\epsilon = \nu e_j$ for $\nu \in \mathbb{R}$ small and by inserting this into the previous equation we find

$$f_i(x + \nu e_j) = f_i(x) + \nu A_{ij} + \mathcal{O}(\nu^2) \; . \tag{5.3}$$

This result allows us to calculate the partial derivatives as

$$\partial_j f_i(x) = \lim_{\nu\to 0} \frac{f_i(x + \nu e_j) - f_i(x)}{\nu} \overset{(5.3)}{=} \lim_{\nu\to 0} \left(A_{ij} + \frac{\mathcal{O}(\nu^2)}{\nu} \right) = A_{ij} \; , \tag{5.4}$$

and this is the required statement. □

In short, the Jacobi matrix is the matrix which contains all the partial derivates. Note they are organised such that the j^{th} row of the Jacobi matrix contains the partial derivatives of the j^{th} function component f_j of f. For practical calculations it can be useful to have a notation for the Jacobi matrix which refers explicitly to the variables x_j and the function components f_i and such a notation is given by [2]

$$Df(x) = \frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}(x) \; . \tag{5.5}$$

[2] This notation is sometimes used to denote the Jacobian which is the determinant of the Jacobi matrix, for the case of vector fields. This seems a bit of a waste and we prefer to denote by Eq. (5.5) the Jacobi matrix while the Jacobian will be written as $\det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\mathbf{x})$.

Specific dimensions
It may be worth discussing a few specific choices of dimensions. First consider a scalar function, $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$. Its Jacobi matrix is a single row which contains the partial derivatives, so

$$Df(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)(x) =: \nabla f(x) \, . \tag{5.6}$$

The Jacobi matrix of a real-valued function is also referred to as the **gradient** of the function and is written as $\nabla f(x) = Df(x)$, as on the RHS of Eq. (5.6), with the **nabla operator** $\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right)$ (more on this in Section 8).

For a two-dimensional vector field $A = (A_1, A_2)^T : U \to \mathbb{R}^2$, where $U \subset \mathbb{R}^2$ open, the Jacobi matrix is a 2×2 matrix which reads

$$DA(x) = \frac{\partial(A_1, A_2)}{\partial(x_1, x_2)}(x) = \begin{pmatrix} \frac{\partial A_1}{\partial x_1} & \frac{\partial A_1}{\partial x_2} \\ \frac{\partial A_2}{\partial x_1} & \frac{\partial A_2}{\partial x_2} \end{pmatrix}(x) \, . \tag{5.7}$$

Note that the partial derivatives of the component functions A_1 and A_2 appear in the rows. Similarly, for a three-dimensional vector field $A = (A_1, A_2, A_3)^T : U \to \mathbb{R}^3$, where $U \subset \mathbb{R}^3$ open, the Jacobi matrix reads

$$DA(x) = \frac{\partial(A_1, A_2, A_3)}{\partial(x_1, x_2, x_3)}(x) = \begin{pmatrix} \frac{\partial A_1}{\partial x_1} & \frac{\partial A_1}{\partial x_2} & \frac{\partial A_1}{\partial x_3} \\ \frac{\partial A_2}{\partial x_1} & \frac{\partial A_2}{\partial x_2} & \frac{\partial A_2}{\partial x_3} \\ \frac{\partial A_3}{\partial x_1} & \frac{\partial A_3}{\partial x_2} & \frac{\partial A_3}{\partial x_3} \end{pmatrix}(x) \, . \tag{5.8}$$

Example 5.1. Examples of Jacobi matrices
Question: Compute the Jacobi matrices of the vector fields $A : \mathbb{R}^2 \to \mathbb{R}^2$ and $B : \mathbb{R}^3 \to \mathbb{R}^3$ given by $A(x, y) = (xy, x^2 - y^2)^T$ and $B(x, y, z) = (yz, xz, xy)^T$.
Answer: For the vector field A we have

$$DA(x, y) = \begin{pmatrix} \frac{\partial A_1}{\partial x} & \frac{\partial A_1}{\partial y} \\ \frac{\partial A_2}{\partial x} & \frac{\partial A_2}{\partial y} \end{pmatrix}(x, y) = \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix} \, .$$

For the vector B we find

$$DB(x, y, z) = \begin{pmatrix} \frac{\partial B_1}{\partial x} & \frac{\partial B_1}{\partial y} & \frac{\partial B_1}{\partial z} \\ \frac{\partial B_2}{\partial x} & \frac{\partial B_2}{\partial y} & \frac{\partial B_2}{\partial z} \\ \frac{\partial B_3}{\partial x} & \frac{\partial B_3}{\partial y} & \frac{\partial B_3}{\partial z} \end{pmatrix}(x, y, z) = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix} \, .$$

Example 5.2. Jacobi matrix of coordinate changes
Question: The function $[0, \infty) \times [0, 2\pi) \to \mathbb{R}^2$ which describes the change of coordinates from 2-dimensional Cartesion coordinates (x, y) to polar coordinates (r, φ) is given by $(x, y) = (r\cos(\varphi), r\sin(\varphi))$. Calculate its Jacobi matrix. Do the same for the coordinate change from 3-dimensional Cartesian coordinates to spherical coordinates.
Answer: In the two-dimensional case we have

$$\frac{\partial(x, y)}{\partial(r, \varphi)} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} \, .$$

Note that the two columns of this matrix are orthogonal (although it is not an orthogonal matrix as

the second column is not normalised to one). In fact, the columns can be used to define the standard unit vectors, e_r and e_φ, in polar coordinates, which are given by

$$e_r := \frac{\partial(x,y)}{\partial r} = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}, \qquad e_\varphi := \frac{1}{r}\frac{\partial(x,y)}{\partial \varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}.$$

These vectors form an ortho-normal system, so $|e_r| = |e_\varphi| = 1$ and $e_r \cdot e_\varphi = 0$.
The three dimensional coordinate change is given by $(x, y, z) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta))$ where the range of the spherical coordinates is $r \in [0, \infty)$, $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$. For the associated Jacobi matrix we find

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

As in the two-dimensional cases, we can use the columns of this matrix to define the standard unit vectors e_r, e_θ and e_φ, in spherical coordinates. They are given by

$$e_r := \frac{\partial(x,y,z)}{\partial r} = \begin{pmatrix} \sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta) \end{pmatrix} \qquad e_\theta := \frac{1}{r}\frac{\partial(x,y,z)}{\partial \theta} = \begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) \\ -\sin(\theta) \end{pmatrix}$$

$$e_\varphi := \frac{1}{r\sin(\theta)}\frac{\partial(x,y,z)}{\partial \varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix}.$$

As in the two-dimensional case, these vector form on ortho-normal system.

Exercise 5.1. *Find the Jacobi matrix for a coordinate change from three-dimensional Cartesian coordinates (x, y, z) to cylindrical coordinates (r, φ, z) defined by $x(r\varphi, z) = r\cos(\varphi)$ and $y(r, \varphi, z) = r\sin(\varphi)$ (and the coordinate z remains unchanged).*

5.3 The chain rule

General chain rule
What does the chain rule for total derivatives look like? In the one-dimensional case the chain rule reads $(f \circ g)'(x) = f'(g(x))g'(x)$, that is, we obtain the derivative of the composite function $f \circ g$ by multiplying the derivatives of f and g. In the multi-dimensional case, the total derivatives become (Jacobi) matrices, so it is natural to expect that the chain rule works by multiplying these matrices. This is in fact what the chain rule states.

Theorem 5.2. *(Chain rule) Consider maps $g : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open and $f : V \to \mathbb{R}^p$, with $V \subset \mathbb{R}^m$ and $g(U) \subset V$, so a sequence*

$$U \subset \mathbb{R}^n \xrightarrow{\ g\ } V \subset \mathbb{R}^m \xrightarrow{\ f\ } \mathbb{R}^p .$$

Let g be totally differentiable at $x \in U$ and f be totally differentiable at $y := g(x) \in V$. Then $f \circ g$ is totally differentiable at $x \in U$ and we have

$$D(f \circ g)(x) = Df(y)\, Dg(x) . \tag{5.9}$$

Proof. By assumption the total derivatives of g at x and of f at y exist. We denote these by $A = Dg(x)$ and $B = Df(y)$ so that

$$g(x+\epsilon) = g(x) + A\epsilon + \mathcal{O}(|\epsilon|^2) \,, \qquad f(y+\eta) = f(y) + B\eta + \mathcal{O}(|\eta|^2) \,.$$

The trick is to choose $\eta = g(x+\epsilon) - g(x) = A\epsilon + \mathcal{O}(|\epsilon|^2)$ and this gives

$$\begin{aligned}(f\circ g)(x+\epsilon) &= f(g(x+\epsilon)) = f(g(x)+\eta) = f(g(x)) + B\eta + \mathcal{O}(|\eta|^2) \\ &= f(g(x)) + BA\epsilon + \mathcal{O}(|\eta|^2, |\epsilon|^2) \,. \end{aligned} \tag{5.10}$$

This shows that $D(f\circ g)(x) = BA = Df(y)\, Dg(x)$. □

In short, the Jacobi matrix of the composite function is the matrix product of the individual Jabobi matrices, as in Eq. (5.9). There are various other, and perhaps more familiar ways to write this. For example, let us adopt the somewhat abusive notation, common in physics, where we denote functions and coordinates by the same name, that is, we write $y_i(x) = g_i(x)$. Then, using the notation (5.5) for the Jacobi matrix, the chain rule can be stated as

$$\frac{\partial((f\circ y)_1, \ldots, (f\circ y)_p)}{\partial(x_1,\ldots,x_n)}(x) = \frac{\partial(f_1,\ldots,f_p)}{\partial(y_1,\ldots,y_m)}(y)\, \frac{\partial(y_1,\ldots,y_m)}{\partial(x_1,\ldots,x_n)}(x) \,, \tag{5.11}$$

$$p\times n \qquad\qquad p\times m \qquad\qquad m\times n \tag{5.12}$$

where the size of the Jacobi matrices is indicated underneath. Alternatively, we can write the matrix product on the right-hand side of Eq. (5.11) out with indices, keeping in mind that a matrix product translates into a sum over the adjacent indices. Then we get

$$\frac{\partial(f\circ y)_i}{\partial x_j}(x) = \sum_{k=1}^{m} \frac{\partial f_i}{\partial y_k}(y)\, \frac{\partial y_k}{\partial x_j}(x) \,, \tag{5.13}$$

which is perhaps the most common version of the chain rule.

Chain rule for specific dimensions

As before, let us specialise to a specific dimensions. First consider a scalar function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$, and a map $X = (X_1, \ldots, X_n)^T : [a,b] \to U$ with coordinate denoted by t (as we will see below, X is a curve). We can ask about the values of the scalar function along the curve and this is described by the function $F = f\circ X : [a,b] \to \mathbb{R}$ of a single variable t. Its derivative (confusingly, sometimes called 'total derivative') can be worked out from the chain rule and is given by

$$\frac{dF}{dt}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X(t)) \frac{dX_i}{dt}(t) \,. \tag{5.14}$$

If we identify functions with coordinates, so write X_i as x_i, and omit the arguments, as is often done in a physics context, the last expression simplifies to

$$\frac{dF}{dt} = \sum_i \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \,. \tag{5.15}$$

Even more explicitly, in three dimensions, $n = 3$, and with the three coordinates denoted by $(x_1, x_2, x_3) = (x, y, z)$ the last equation specialises to

$$\frac{dF}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \,. \tag{5.16}$$

As another example, let us consider a two-dimensional coordinate change, from 'old' coordinates $(x, y) \in U \subset \mathbb{R}^2$ to 'new' coordinates $(u, v) \in \tilde{U} \subset \mathbb{R}^2$ which we regard as a (diffeomorphic) function $U \to \tilde{U}$, $u = u(x, y)$ and $v = v(x, y)$ (with inverse functions $x = x(u, v)$ and $y = y(u, v)$), of the old coordinates. Suppose two functions $f : U \to \mathbb{R}$ and $\tilde{f} : \tilde{U} \to \mathbb{R}$ are related by this coordinate change, so $\tilde{f}(u, v) = f(x(u, v), y(u, v))$. How do their Jacobi matrices relate? Using the chain rule we have

$$\left(\frac{\partial \tilde{f}}{\partial u}, \frac{\partial \tilde{f}}{\partial v}\right) = \frac{\partial \tilde{f}}{\partial (u, v)} = \frac{\partial f}{\partial (x, y)} \frac{\partial (x, y)}{\partial (u, v)} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial v}\right) \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} , \tag{5.17}$$

or, written out in components,

$$\frac{\partial \tilde{f}}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} , \qquad \frac{\partial \tilde{f}}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} . \tag{5.18}$$

Example 5.3. Chain rule and coordinate changes
Question: Two scalar functions f and $\tilde{f}$ in two dimensions are related by a transformation from Cartesian to polar coordinates, so $\tilde{f}(r, \varphi) = f(x(r, \varphi), y(r, \varphi))$, where $x(r, \varphi) = r\cos(\varphi)$ and $y(r, \varphi) = r\sin(\varphi)$. Work out the relation between the partial derivatives of f and $\tilde{f}$.

Answer: The Jacobi matrices of the coordinate changes have already been computed in Exercise 5.2. For the two-dimensional case we have

$$\left(\frac{\partial \tilde{f}}{\partial r}, \frac{\partial \tilde{f}}{\partial \varphi}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}$$

or, written out in components,

$$\frac{\partial \tilde{f}}{\partial r} = \cos(\varphi)\frac{\partial f}{\partial x} + \sin(\varphi)\frac{\partial f}{\partial y} , \qquad \frac{\partial \tilde{f}}{\partial \varphi} = -r\sin(\varphi)\frac{\partial f}{\partial x} + r\cos(\varphi)\frac{\partial f}{\partial y} .$$

Exercise 5.3. *Repeat the calculation in Exercise 5.4 in three dimensions, for a coordinate change from Cartesian coordinates to spherical coordinates.*

Example 5.4. Derivatives of radius-dependent functions
Question: The function $f : \mathbb{R}^n \to \mathbb{R}$ only depends on the radial coordinate, that is, it can be written as $f = h \circ r$, where $h : \mathbb{R} \to \mathbb{R}$ is a function of a single variable and $r(x) = |x|$ is the radius function. Find the partial derivatives of f. Specialise your result to the case when $h(r) = r^p$, where p is a number.

Answer: From Exercise 4.2 we recall the useful formula $\partial_i r = \frac{x_i}{r}$. Applying the chain rule we find

$$\frac{\partial f}{\partial x_i}(x) = \frac{dh}{dr}(r)\frac{\partial r}{\partial x_i}(x) = x_i \frac{h'(r)}{r} .$$

For $h(r) = r^p$, inserting $h'(r) = pr^{p-1}$ gives $\partial_i f = p x_i r^{p-2}$.

5.4 Jacobi matrix of inverse function

If we apply the chain rule in one dimension, $(f \circ g)'(x) = f'(g(x))g'(x)$, to the special case where $f = g^{-1}$ (assuming that g is invertible and its inverse is also differentiable) we get $1 = g^{-1\prime}(y)g'(x)$, where $y = g(x)$, and, hence, this leads to the rule for the derivative of the inverse function, $g^{-1\prime}(y) = 1/g'(x)$.

Let us generalises this argument to the multi-dimension case, starting with the chain rule (5.9). Assume that $g : U \to V$ is totally differentiable and (locally) invertible at $x \in U$ and that $f = g^{-1}$ is totally differentiable at $y = g(x)$ as well. On the left-hand side of Eq. (5.9) we have the map $f \circ g = g^{-1} \circ g = \mathrm{id}$, that is the identify map, with $\mathrm{id}(x) = x$. Since

$$\frac{\partial \,\mathrm{id}_i}{\partial x_j} = \frac{\partial x_i}{\partial x_j} = \delta_{ij} \ ,$$

the Jacobi matrix of the identity map is the unit matrix, so $D\,\mathrm{id} = \mathbb{1}$. Hence, Eq. (5.9) turns into

$$\mathbb{1} = D(g^{-1})(y)\, Dg(x) \ . \tag{5.19}$$

where $y = g(x)$ and we have shown

Corollary 5.1. *Let $g : U \to V$, with $U \subset \mathbb{R}^n$ open, be totally differentiable at $x \in U$, and invertible with the inverse g^{-1} also totally differentiable at $y = g(x) \in V$. Then, the Jacobi matrix $Dg(x)$ is invertible and we have*

$$D(g^{-1})(y) = (Dg(x))^{-1} \ . \tag{5.20}$$

Proof. This follows directly from Eq. (5.19). □

Note that the inverse on the right-hand side of Eq. (5.20) is a matrix inverse. In other words, the Jacobi matrix of the inverse function is the matrix inverse of the original Jacobi matrix. If we adopt a more physics-related notation and write $y_i(x) = g_i(x)$, as before, then, using the notation (5.5) for the Jacobi matrix, the inverse derivative rule (5.20) can be stated as

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)}(y) = \left(\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}(x) \right)^{-1} \ . \tag{5.21}$$

Example 5.5. Jacobian of the inverse function
Question: Suppose we have a function $g : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $g = (g_1, g_2)^T$ with $g_1(x_1, x_2) = (x_1^2 + x_2^2)/2$ and $g_2(x_1, x_2) = x_1 x_2$. Compute the Jacobi matrix of g and the Jacobi matrix of g^{-1}.
Answer: Writing $y_1 = g_1$ and $y_2 = g_2$, the Jacobi matrix of g is given by

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 \end{pmatrix} \ . \tag{5.22}$$

One way to compute the Jacobi matrix of g^{-1} is to work out this inverse map explicitly by solving the equations $y_1 = (x_1^2 + x_2^2)/2$ and $y_2 = x_1 x_2$ for x_1 and x_2 and then computing $\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}$. But it is easier to use the inverse derivative rule (5.21) instead, which gives

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \left(\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} \right)^{-1} = \frac{1}{x_1^2 - x_2^2} \begin{pmatrix} x_1 & -x_2 \\ -x_2 & x_1 \end{pmatrix} \ . \tag{5.23}$$

Note this result for the Jacobi matrix for g^{-1} is only well-defined in the neighbourhood of points (x_1, x_2) where $x_1^2 - x_2^2 \neq 0$. Indeed, when $x_1^2 - x_2^2 = 0$ the Jacobi matrix (5.22) of g becomes singular so its inverse does not exist. This indicates that the function g cannot be inverted near points (x_1, x_2) with $x_1^2 - x_2^2 = 0$, a statement that will be made more precise by Corollary 6.1 below.

6 Taylor series, implicit functions

6.1 Taylor series

You have already encountered the Taylor series for functions of a single variable and for functions of two variables. The Taylor series can be generalised to functions of an arbitrary number of variables and this is what we want to tackle next. We consider (suitably many times continuously differentiable) real-valued function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open, with coordinates $x = (x_1, \ldots, x_n)^T$. To this end it is useful to introduce a more compact multi-index notation, with multi-indices $K = (k_1, \ldots, k_n)$, monomials $x^K := x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ and a generalised factorial $K! := k_1!\, k_2! \cdots k_n!$. With this notation, the typical term we expect in a multi-dimensional Taylor series can be written concisely as

$$\frac{x^K}{K!} = \frac{x_1^{k_1}}{k_1!} \frac{x_2^{k_2}}{k_2!} \cdots \frac{x_n^{k_n}}{k_n!} \,. \tag{6.1}$$

It is useful to keep track of the total degree of such a term which is given by $|K| := k_1 + k_2 + \cdots k_n$. Also, multiple partial derivatives are denoted by $\partial^K := \partial_1^{k_1} \partial_2^{k_2} \cdots \partial_n^{k_n}$.

Taylor's formula

Theorem 6.1. *(Taylor's formula) Let $f : U \to \mathbb{R}$, with $U \subset \mathbb{R}^n$ open, be κ-times continuously differentiable at $\mathbf{x} \in U$. For a $\boldsymbol{\xi} \in \mathbb{R}^n$ which satisfies $x + \xi \in U$ for all $t \in [0, 1]$ there exists an $\epsilon \in [0, 1]$ such that*

$$f(x + \xi) = \sum_{|K| \le \kappa} \frac{\partial^K f(x)}{K!} \xi^K + \sum_{|K| = \kappa} \frac{\partial^K f(x + \epsilon\xi) - \partial^K f(x)}{K!} \xi^K \,. \tag{6.2}$$

Proof. The proof is not too difficult but somewhat elaborate, and we refer to the literature, for example Ref. [7], for details. □

Note that this formula is actually an equality and the second sum on the right-hand side can be seen as the size of the error if we truncate the series by just keeping the first sum up to order κ on the right-hand side. The full Taylor series

$$\sum_K \frac{\partial^K f(x)}{K!} \xi^K \tag{6.3}$$

does not need to converge and even if it does it need not converge to the function value. However, careful consideration of the error term in Eq (6.2) and its behaviour with the order κ, for a given function f, can often be used to decide for which values of ξ the series (6.3) converges to $f(x + \xi)$.

The error term can be recast into a simpler, more intuitive form. Since, by assumption, κ derivatives on f still lead to continuous functions we have $\lim_{\xi \to \mathbf{0}} (\partial^K f(x + \epsilon\xi) - \partial^K f(x)) = 0$. From Eq. (4.3), this means that the second sum in Eq. (6.2) is of $\mathcal{O}(|\xi|^{\kappa+1})$ and Taylor's formula becomes

$$f(x + \xi) = \sum_{|K| \le \kappa} \frac{\partial^K f(x)}{K!} \xi^K + \mathcal{O}(|\xi|^{\kappa+1}) \,. \tag{6.4}$$

It is instructive to write this out more explicitly for $\kappa = 2$. This leads to

$$f(x + \xi) = f(x) + \sum_{i=1}^n \partial_i f(x)\, \xi_i + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(x)\, \xi_i \xi_j + \mathcal{O}(|\xi|^3) \tag{6.5}$$

$$= f(x) + \nabla f(x) \cdot \xi + \frac{1}{2} \xi^T H f(x)\, \xi + \mathcal{O}(|\xi|^3) \,, \tag{6.6}$$

where

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)(x) \,, \qquad (Hf(x))_{ij} = \partial_i \partial_j f(x) \tag{6.7}$$

are the gradient and the so-called **Hesse matrix** of f at x, respectively.

6.2 Extremal points

The above form of Taylor's formula up to quadratic terms is very useful to analyse local properties of a function, such as minima and maxima. Just to be sure we define these properly first.

Definition 6.1. *Consider a real-valued function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open. A point $x \in U$ is called a local minimum (a local maximum) of f if there is a sufficiently small ball $B_\epsilon(x) \subset U$ such that $f(z) \geq f(x)$ for all $z \in B_\epsilon(x)$ ($f(z) \leq f(x)$ for all $z \in B_\epsilon(x)$). A local extremum is either a local minimum or a local maximum. If the equality in the above conditions is only realised for $z = x$ then the local minimum, maximum or extremum is called isolated.*

Local extrema have an interesting property.

Proposition 6.1. *Consider a real-valued continuously differentiable function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open, with a local extremum at $x \in U$. Then $\nabla f(x) = 0$, that is, x is a stationary point of f.*

Proof. This can be shown by relating the statement to its analogue in one dimension. To do this, consider the functions $g_i : \mathbb{R} \to \mathbb{R}$ defined by $g_i(t) := f(x + te_i)$. Clearly, if f has a local extremum at x the functions g_i have a local extremum at $t = 0$. Hence, from the well-known one-dimensional statement we know that $g_i'(0) = 0$. However, $g_i'(0) = \partial_i f(x)$ (for all i) and this proves the claim. □

Hence, the local extrema of a real-valued function f are to be found among its stationary points, that is, among the points x in its domain for which $\nabla f(x) = 0$. However, not all such stationary points are necessarily local extrema. Criteria for when this is the case can be formulated in terms of the Hesse matrix. Indeed, at a stationary point x of f Taylor's formula (6.6) becomes

$$f(x + \xi) - f(x) = \frac{1}{2} \xi^T H f(x)\, \xi + \mathcal{O}(|\xi|^3) \,, \tag{6.8}$$

so the leading behaviour near x is determined by the second order term which is controlled by the Hesse matrix. To formulate what this implies more precisely we need to recall some properties of symmetric $n \times n$ matrices M from linear algebra. Such a matrix M is called *positive definite* if $\xi^T M \xi > 0$ for all $\xi \neq 0$. It is called *negative definite* if $-M$ is positive definite. Further, M is called *indefinite* if $\xi^T M \xi$ takes on strictly positive and strictly negative values, for suitable ξ. We recall that M is positive (negative) definite iff all its eigenvalues are strictly positive (strictly negative). It is indefinite iff it has at least one strictly positive and at least one strictly negative eigenvalue.

Theorem 6.2. *Consider a real-valued twice continuously differentiable function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open, with a stationary point at $x \in U$. Then we have the following statements.*
(i) If $Hf(x)$ is positive definite then x is an isolated local minimum.
(ii) If $Hf(x)$ is negative definite then x is an isolated local maximum.
(iii) If $Hf(x)$ is indefinite then x is not a local extremum (but a saddle).

Proof. (i) Let us sketch the proof. Since $Hf(x)$ is positive definite there exists a constant $c > 0$ such that $\frac{1}{2}\xi^T H(x)\, \xi \geq c|\xi|^2$ for all $\xi \in \mathbb{R}^n$. This means, we can always find an $\epsilon > 0$ so that the $\mathcal{O}(|\xi|^3)$ term in Eq. (6.8) is smaller than $\frac{1}{2}\xi^T Hf(x)\, \xi$ for all ξ with $|\xi| < \epsilon$. Hence, from Eq. (6.8), $f(x+\xi) - f(x) > 0$ for all ξ with $|\xi| < \epsilon$ so that x is indeed an isolated local minimum.
(ii), (iii) The proofs are similar to the one for (i). □

Proposition 6.1 and Theorem 6.2 suggest a method to find the isolated local extrema of a function $f : U \to \mathbb{R}$. As a first step, find the stationary points of f in U by solving the equation $\nabla f(x) = 0$. Next, for each stationary point x, compute the Hesse matrix $Hf(x)$ and its eigenvalues and use the following criterion:

$$\begin{array}{lcl} x \text{ local isolated minimum} & \Leftarrow & \text{all eigenvalues of } Hf(x) \text{ strictly positive} \\ x \text{ local isolated maximum} & \Leftarrow & \text{all eigenvalues of } Hf(x) \text{ strictly negative} \\ x \text{ not a local extremum} & \Leftarrow & Hf(x) \text{ has strictly positive and strictly negative eigenvalues} \end{array} \quad . \tag{6.9}$$

Remark. *If $Hf(x)$ at a stationary point x has a zero eigenvalue the Hesse matrix is not necessarily sufficient to decide the nature of the stationary point. For example, if $Hf(x)$ has one zero eigenvalue with all other eigenvalues positive one cannot conclude that the stationary point is a minimum. What happens in the direction of the eigenvector with zero eigenvalue depends on cubic or higher orders in the Taylor expansion of f.*

Let us discuss functions f of two variables $x = (x_1, x_2)^T$ more explicitly. In this case, the Hesse matrix at x takes the form

$$Hf(x) = \begin{pmatrix} \partial_1^2 f & \partial_1\partial_2 f \\ \partial_1\partial_2 f & \partial_2^2 f \end{pmatrix}(x) \quad \Rightarrow \quad \begin{cases} \det(Hf(x)) = (\partial_1^2 f\, \partial_2^2 f - (\partial_1\partial_2 f)^2)(x) \\ \mathrm{tr}(Hf(x)) = (\partial_1^2 f + \partial_2^2 f)(x) \end{cases} . \tag{6.10}$$

There are two eigenvalues, λ_1, λ_2, and their signs can be determined by considering the determinant and trace of $H(x)$. More specifically, for a stationary point x of f we have

$$\begin{array}{lcccl} x \text{ local isolated minimum} & \Leftarrow & \lambda_1, \lambda_2 > 0 & \Leftrightarrow & \det(Hf(x)) > 0 \text{ and } \mathrm{tr}(Hf(x)) > 0 \\ x \text{ local isolated maximum} & \Leftarrow & \lambda_1, \lambda_2 < 0 & \Leftrightarrow & \det(Hf(x)) > 0 \text{ and } \mathrm{tr}(Hf(x)) < 0 \\ x \text{ not a local extremum} & \Leftarrow & \lambda_1\lambda_2 < 0 & \Leftrightarrow & \det(Hf(x)) < 0 \end{array} \quad . \tag{6.11}$$

If the function depends on more than two variables, the determinant and trace of the Hesse matrix are not sufficient to decide on the signs of all of its eigenvalues. For such cases, in order to work out the nature of a stationary point x, one has to compute the eigenvalues of $Hf(x)$ from its characteristic polynomial and use the general criterion (6.9).

Example 6.1. Extrema for a function $f : \mathbb{R}^3 \to \mathbb{R}$
Question: Compute the gradient and the Hesse matrix for the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = x^2 - x + xy^2 + z^2 - 2z$. Use these results to find the stationary points of f and the nature of these stationary points by considering the eigenvalues of the Hesse matrix.

Answer: The gradient and Hesse matrix of f are given by

$$\nabla f(x, y, z) = (y^2 + 2x - 1, 2xy, 2z - 2) , \qquad Hf(x, y, z) = \begin{pmatrix} 2 & 2y & 0 \\ 2y & 2x & 0 \\ 0 & 0 & 2 \end{pmatrix} .$$

Setting $\nabla f(x, y, z) = 0$ one finds three stationary points which we collect, together with the associated Hesse matrices and their eigenvalues, in the following table.

(x,y,z)	$Hf(x,y,z)$	eigenvalues of $Hf(x)$
$(0,1,1)$	$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$(1+\sqrt{5}, 1-\sqrt{5}, 2)$
$(0,-1,1)$	$\begin{pmatrix} 2 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$(1+\sqrt{5}, 1-\sqrt{5}, 2)$
$(1/2,0,1)$	$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$(2,1,2)$

Hence, the stationary points at $(x,y,z) = (0,\pm 1,1)$ are not local extrema (but saddles) since they have two strictly positive and one strictly negative eigenvalue. On the other hand, the stationary point at $(x,y,z) = (1/2,0,1)$ is a local minimum, since all eigenvalues are strictly positive.

Example 6.2. Stationary points of quadratic forms
Question: The function $f : \mathbb{R}^n \to \mathbb{R}$ is defined by $f(x) = \frac{1}{2}\left(x^T M x - \lambda x^T x\right)$, where M is a symmetric $n \times n$ matrix and $\lambda \in \mathbb{R}$ is a number. Compute the gradient and the Hesse matrix of this functions, find the stationary points and determine which type they are.

Solution: First write the function with indices, so $f(x) = \frac{1}{2}(M_{kl}x_k x_l - \lambda x_k x_k)$. Then differentiation is straightforward

$$\partial_i f(x) = M_{ik}x_k - \lambda x_i \ , \qquad \partial_i \partial_j f(x) = M_{ij} - \lambda \delta_{ij} \ .$$

In vector-matrix notation this reads

$$\nabla f(x) = Mx - \lambda x \ , \qquad Hf(x) = M - \lambda \mathbb{1}_n \ .$$

To find the stationary points we set the gradient to zero, $\nabla f(x) = Mx - \lambda x = 0$. Evidently this is precisely the eigenvalue equation for the matrix M. If the given number λ is different from all eigenvalues of M then the only solution is $x = 0$. On the other hand, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of M and $\lambda = \lambda_k$ for some k then the equation is solved for all vectors $x \in \mathrm{Eig}_f(\lambda)$. To discuss the nature of these stationary points, we note that the eigenvalues of the Hesse matrix $Hf(x)$ are given by $(\lambda_1 - \lambda, \ldots, \lambda_n - \lambda)$. For the case where λ is different from all eigenvalues $x = 0$ is the only stationary point, as discussed, and its nature depends on the signs of $\lambda_k - \lambda$. For example, if λ is smaller than all eigenvalues then $x = 0$ is a minimum. On the other hand, if $\lambda = \lambda_k$ for some k then one of the eigenvalues of $Hf(x)$ vanishes so this direction is flat. In fact, for any $x \in \mathrm{Eig}_f(\lambda)$ we have $f(x) = 0$. What happens in the other directions depends on which eigenvalue λ is equal to. If λ is the smallest eigenvalue of M then all other directions correspond to positive eigenvalues of Hf. Otherwise, Hf has negative eigenvalues.

Application 6.1. Functionals and the variational principle
Consider the following problem. Let $q : [a,b] \to \mathbb{R}^n$ be differentiable curves with given endpoints $q(a) = q^{(0)}$ and $q(b) = q^{(1)}$ and denote their tangent vectors by $\dot{q}(t) = \frac{dq}{dt}(t)$. Further, assume we have a function $L : \mathbb{R}^{2n+1} \to \mathbb{R}$ (in a physics context referred to as a **Lagrangian**) which is used to measures a certain property

$$S[q] = \int_a^b dt\, L(q(t), \dot{q}(t), t) \tag{6.12}$$

of the curves. For example, this property might be the length of the curves (and we return to this example below). The object S defined in Eq. (10.7) is referred to as a **functional** which is meant to indicate that it assigns a real value $S[q]$ to a function (in this case a curve) q. In a physics context S is also called the **action**. The problem we would like to tackle is the following. Can we identify, within the given class of curves q, those which minimise the value of S? In other words, can be find the minima (or, rather, the stationary points) of the functional S? We are used to finding stationary points of functions but the present task sounds much more ambitious: Rather than searching for stationary points in some domain of real numbers we are now searching in a domain which consists of functions.

However, it turns out that we can solve this problem with the methods at our disposal. The trick is to consider variations $\epsilon : [a, b] \to \mathbb{R}^n$ with $\epsilon(a) = \epsilon(b) = 0$ and consider the modified curves $q + \lambda\epsilon$, where λ is a real parameter. Note that ϵ vanishes at a and b so that the endpoints of the modified curves remain unchanged at $q^{(0)}$ and $q^{(1)}$. We can then define the function F by

$$F(\lambda) := S[q + \lambda\epsilon] = \int_a^b dt\, L(q(t) + \lambda\epsilon(t), \dot{q}(t) + \lambda\dot{\epsilon}(t), t) \tag{6.13}$$

and the condition for a curve q to be stationary can now be stated as $\frac{dF}{d\lambda}(0) = 0$, that is, as a conventional problem of finding the stationary points of a function of one real variable. Let us work out this derivative more explicitly (omitting function arguments for simplicity).

$$\frac{dF}{d\lambda}(0) = \int_a^b dt \left[\frac{\partial L}{\partial q_i}\epsilon_i + \frac{\partial L}{\partial \dot{q}_i}\dot{\epsilon}_i\right] = \int_a^b dt \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\epsilon}_i}\right] \epsilon_i \stackrel{!}{=} 0 \tag{6.14}$$

In the last step we have integrated the second term by parts and the boundary term vanishes due to $\epsilon(a) = \epsilon(b) = 0$. Given the variations ϵ_i are arbitrary the integral in Eq. (6.14) can only vanish if the terms in square brackets on the right vanish individually, that is, if

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \text{for} \quad i = 1, \ldots, n\,. \tag{6.15}$$

What we have shown is that for a curve q to be stationary relative to our functional S it has to satisfy the equations (6.15), the so-called **Euler-Lagrange equations**.

Let us consider a simple example of a curve in two dimensions which is given by the graphs of functions $f : [a, b] \to \mathbb{R}$ with given values $f(a) = y^{(0)}$ and $f(b) = y^{(1)}$ at the endpoints of the interval, that is, the curve is given by

$$q(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix} \qquad \Rightarrow \qquad \dot{q}(x) = \begin{pmatrix} 1 \\ f'(x) \end{pmatrix}\,, \tag{6.16}$$

with endpoints $q^{(0)} = (a, y^{(0)})^T$ and $q^{(1)} = (b, y^{(1)})^T$. The function L we would like to consider is $L(x) = |\dot{q}(x)| = \sqrt{1 + f'(x)^2}$ so that the functional

$$S[q] = \int_a^b dx\, \sqrt{1 + f'(x)^2}\,, \tag{6.17}$$

represents the length of the curve. Using the Euler-Lagrange equation (with x playing the role of t and f the role of q) we have

$$\frac{\partial L}{\partial f} = 0 \,, \qquad \frac{\partial L}{\partial f'}(x) = \frac{f'(x)}{\sqrt{1+f'(x)^2}} \qquad \Rightarrow \qquad \frac{d}{dt}\left(\frac{f'(x)}{\sqrt{1+f'(x)^2}}\right) = 0 \,. \tag{6.18}$$

The differential equation on the right is solved by $f'(x) = \text{const}$ and, translated to the curve q, this corresponds to straight lines. In other words, what we have found is the expected answer: the shortest curve between two given points is the line segment between them.

The last example may seem a bit trivial - after all we knew the answer in advance. Things get a lot more interesting when we consider classical mechanics. Consider N point particles with Cartesian coordinates $x(t) = (x_i(t)) \in \mathbb{R}^{3N}$, masses m_i (which you can think of as being equal in groups of three but this does not affect our calculation) and a potential function V with associated force F with components $F_i = -\frac{\partial V}{\partial x_i}$. The corresponding Newton equations read

$$m_i \ddot{x}_i = -\frac{\partial V}{\partial x_i} \qquad \Leftrightarrow \qquad \frac{d}{dt}\frac{\partial}{\partial \dot{x}_i}\left(\frac{1}{2} m_i \dot{x}_i^2\right) + \frac{\partial V}{\partial x_i} = 0 \tag{6.19}$$

The equivalent form of Newton's equations on the right shows they can be obtained from a Lagrange function, given as the difference of kinetic and potential energy,

$$L = \frac{1}{2}\sum_i m_i \dot{x}_i^2 - V(x) \,, \tag{6.20}$$

by applying to L the Euler-Lagrange equations (6.15). This has profound consequences. Together with our discussion of the variation of functionals above this is saying that the physical motion of a system in classical mechanics represents a stationary point of the action

$$S[x] = \int_a^b dt\, L(x(t), \dot{x}(t)) \,, \tag{6.21}$$

where L is the Lagrangian of the system as in Eq. (6.20). This is also referred to as **Hamilton's principle**. Of course Cartesian coordinates may not always be the best way to describe a system. (Think about a pendulum whose motion is best described by an angle.) So we consider an arbitrary set of coordinates $q = (q_a)$ of which the Cartesian ones are functions, $x_i = x_i(q)$, and re-write the Lagrangian in terms of these new coordinates. A simple calculation, starting from Eq. (6.20) gives [a]

$$L = \frac{1}{2}\sum_{a,b} G_{ab}(q)\dot{q}_a\dot{q}_b - U(q) \quad \text{where} \quad G_{ab} := \sum_i m_i \frac{\partial x_i}{\partial q_a}\frac{\partial x_i}{\partial q_b} \,, \qquad U(q) := V(x(q)) \,. \tag{6.22}$$

The physical solution corresponds to a stationary point of the corresponding action and this clearly does not depend on the chosen coordinates (in the sense that q is a stationary point of the action from (6.22) iff $x(q)$ is a stationary point of the action from (6.20)). So in this sense the Lagrangian approach to Classical Mechanics provides us with a coordinate-independent formalism. We can write the Lagrangian in terms of any set of suitable coordinates which describe the system and then the Euler-Lagrange equations applied to this Lagrangian will provide us with the correct Newton equations.

Let us apply this to the simple case of a pendulum attached at the origin of the coordinate system, with length l, mass m whose position is described by an angle ϕ, relative to the negative z-axis, so that its position in the (x, z) plane is given by $(x, z) = (l\sin(\phi), -l\cos(\phi))$. Its kinetic and potential energy are $E_{\text{kin}} \overset{!}{=} \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) = \frac{1}{2}ml^2\dot{\phi}^2$ and $U = -mgl\cos(\phi)$, respectively, so we have the Lagrangian

$$L = \frac{1}{2}ml^2\dot{\phi}^2 + mgl\cos(\phi) \ . \tag{6.23}$$

Applying the Euler-Lagrange equations to this Lagrangian,

$$\frac{\partial L}{\partial \dot{\phi}} = ml^2\dot{\phi} \ , \quad \frac{\partial L}{\partial \phi} = -ml\sin(\phi) \qquad \overset{\text{Eq. (6.15)}}{\Longrightarrow} \qquad \ddot{\phi} = -\frac{g}{l}\sin(\phi) \ , \tag{6.24}$$

does indeed give the correct differential equation for the motion of the pendulum.

[a] You might find it instructive to compare the formula for G_{ab} below with Eq. (11.18).

6.3 Implicit functions

Suppose we have a function $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$ open, of n variables which we denote by $z = (z_1, \ldots, z_n)^T$ and also assume that $m < n$. We want to solve the equation $f(z) = 0$, a common problem which arises in many situations in mathematics and physics.

An obvious way to proceed is as follows. Since $f(z) = 0$ consists of m scalar equations for $n > m$ variables we can split our variables as $z = (x, y)$, where $x \in \mathbb{R}^{n-m}$ and $y \in \mathbb{R}^m$ and then attempt to solve for the variables y in terms of x. More precisely, suppose we have a point $(a, b) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ which solves the equation, that is, $f(a, b) = 0$. For open neighbourhoods $\tilde{V}$ and $\tilde{W}$ of a and b (where $\tilde{V} \times \tilde{W} \subset U$) we would then like to find a function $g : \tilde{V} \to \tilde{W}$ such that $b = g(a)$ and $f(x, g(x)) = 0$ for all $x \in \tilde{V}$. This function is precisely the desired solution $y = g(x)$ of $f(x, y) = 0$ (at least locally, near (a, b)), but does it exist? Unfortunately, the answer is "not always" and this is where the *implicit function theorem* comes into play. It states sufficient conditions for the solution g to exist.

Example 6.3. A simple implicit function example
To develop a better intuition for what the obstruction is it is useful to discuss an example. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ of two variables $z = (x, y)^T$ defined by $f(x, y) = x^2 + y^2 - 1$. Of course the solutions to $f(x, y) = 0$ consist of all points on the unit circle in $\mathbb{R}^2$ (see Fig. 7). Choose a specific point (a, b) on this circle, so that $a^2 + b^2 = 1$. Can we always find open neighbourhoods $\tilde{V}$ of a and $\tilde{W}$ of b for which we can solve y in terms of x? Consider the point $(a, b) = (1, 0)$. For any open neighbourhood $\tilde{V} = (1 - \epsilon, 1 + \epsilon)$ we have a problem. If $x \in \tilde{V}$ and $x < 1$ then we have two possible values of y (one on the upper half-circle, the other one on the lower half-circle) for this x. If $x > 1$ there is no corresponding value of y. This means we cannot define a function $y = g(x)$ for $x \in \tilde{V}$. Evidently, the problem arises where the tangent is vertical, that is, at points (a, b) with $\partial_y f(a, b) = 0$, while all other points seem fine.

The implicit function theorem states, roughly, that, as long as the condition $\partial_y f(a, b) \neq 0$ (and its suitable generalisation to higher dimensions) is satisfied, so that the tangent is not vertical, we can find a local solution $y = g(x)$, near the point (a, b). The precise formulation is as follows.

Theorem 6.3. *(Implicit function theorem). Let $f : V \times W \to \mathbb{R}^m$, with $V \subset \mathbb{R}^{n-m}$ and $W \subset \mathbb{R}^m$ open (where $n > m$), be a continuously differentiable function. Denote the coordinates by $(x, y) \in V \times W$ and*

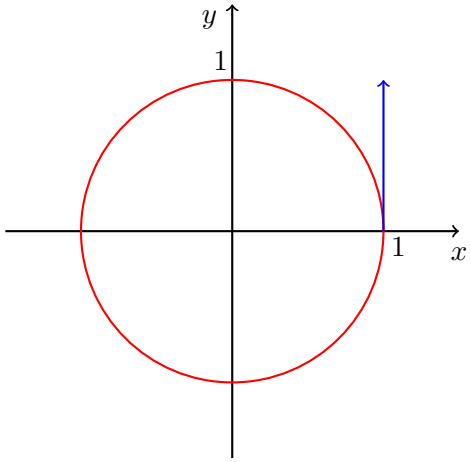


Figure 7: Unit circle $S^1 \subset \mathbb{R}^2$ (in red) as the zero locus of the function f with $f(x,y) = x^2 + y^2 - 1$. At $(x,y) = (1,0)$ we have $\frac{\partial f}{\partial y}(1,0) = 0$, as indicated by the blue vector, and this means it might be difficult to solve $f(x,y) = 0$ for y in terms of x around this point. This is indeed the case, as there are either two or no values of y for a given $x \in [1-\epsilon, 1+\epsilon]$.

choose a point $(a,b) \in V \times W$. *If* $f(a,b) = 0$ *and if the matrix*

$$\frac{\partial(f_1,\ldots,f_m)}{\partial(y_1,\ldots,y_m)}(a,b) \tag{6.25}$$

has maximal rank, then there exist open neighbourhoods $\tilde{V} \subset V$ *of* a *and* $\tilde{W} \subset W$ *of* $\mathbf{b}$ *and a continuously differentiable function* $g : \tilde{V} \to \tilde{W}$ *with* $f(x, g(x)) = 0$ *for all* $(x,y) \in \tilde{V} \times \tilde{W}$. *Conversely, if* $(x,y) \in \tilde{V} \times \tilde{W}$ *and* $f(x,y) = \mathbf{0}$ *then* $y = g(x)$.

Proof. The proof is somewhat lengthy and can be found in the literature, see for example [7]. □

The maximal rank condition on the (partial) Jacobi matrix (6.25) is the generalisation of the requirement, stated earlier, that the tangent at (a,b) is not vertical. The theorem implies a formula for the derivative of the function g in terms of f. To see this, apply the chain rule to the equation $f(x, g(x)) = 0$:

$$\frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_{n-m})} + \frac{\partial(f_1,\ldots,f_m)}{\partial(y_1,\ldots,y_m)} \frac{\partial(g_1,\ldots,g_m)}{\partial(x_1,\ldots,x_{n-m})} = 0\,. \tag{6.26}$$

The second Jacobi matrix in the above equation is precisely the one in Eq. (6.25) which is required to have maximal rank. Hence, we can invert this matrix and solve for the Jacobi matrix of g:

$$\frac{\partial(g_1,\ldots,g_m)}{\partial(x_1,\ldots,x_{n-m})} = -\left[\frac{\partial(f_1,\ldots,f_m)}{\partial(y_1,\ldots,y_m)}\right]^{-1} \frac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_{n-m})}\,. \tag{6.27}$$

Note, we can calculate the Jacobi matrix of g from this formula without actually knowing the function g explicitly - all we need is the original function f.

Example 6.4. Another implicit function example
Let us illustrate this with a further example, for a function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x_1, x_2, y) = x_1^3/3 + x_2^2 y^3/3 + 1$. The solutions to the equation $f(x_1, x_2, y) = 0$ form a surface in $\mathbb{R}^3$ and it is often convenient to describe this surface by solving for y in terms of x_1 and x_2. The implicit function theorem states this is possible, at least locally, for points where

$$\partial_y f = x_2^2 y^2 \tag{6.28}$$

is different from zero, that is, for every point on the surface with $x_2 \neq 0$ and $y \neq 0$. In the neighbourhood of such a point, we can find a solution $y = g(x_1, x_2)$ and compute its derivative from Eq. (6.27):

$$\nabla g = -\left[\partial_y f\right]^{-1} \frac{\partial f}{\partial(x_1, x_2)} = -\frac{1}{x_2^2 y^2}\left(x_1^2, \frac{2}{3} x_2 y^3\right) . \tag{6.29}$$

Here, y should be though of as the function $y = g(x_1, x_2)$ obtained by solving $f(x_1, x_2, y) = 0$ for y.

We can apply the implicit function theorem to answer the question under which conditions an inverse of a (differentiable) function $g : W \to \mathbb{R}^m$, with $W \subset \mathbb{R}^m$ open, exists. More explicitly, if we have points $a \in W$ and $b \in \mathbb{R}^m$ with $a = g(b)$ we want to know if the equation $x = g(y)$ can be solved for y in terms of x near (a, b). To make contact with the implicit function theorem, we define the auxiliary function $f : \mathbb{R}^{2m} \to \mathbb{R}^m$ by $f(x, y) = g(y) - x$. Then we have $f(a, b) = 0$ and the implicit function theorem states that f can be solved for y in terms of x near (a, b) if the matrix

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)}(a, b) = \frac{\partial(g_1, \ldots, g_m)}{\partial(y_1, \ldots, y_m)}(b) \tag{6.30}$$

has maximal rank. Hence, we have

Corollary 6.1. *Let $g : W \to \mathbb{R}^m$, where $W \subset \mathbb{R}^m$ open, be a differentiable function with $a = g(b)$ for certain points $a \in W$ and $b \in \mathbb{R}^m$. Then g is invertible (that is, the equation $x = g(y)$ can be solved for y in terms of x) near (a, b) if the Jacobi matrix*

$$\frac{\partial(g_1, \ldots, g_m)}{\partial(y_1, \ldots, y_m)}(b) \tag{6.31}$$

of g is invertible.

Application 6.2. Implicit functions on a surface
Suppose we have a (continuously differentiable) function $f : \mathbb{R}^3 \to \mathbb{R}$, and we denote the coordinates on $\mathbb{R}^3$ by $(x_1, x_2, x_3) = (p, v, t)$. Then, the equation $f(p, v, t) = 0$ (in the context of Thermodynamics this might be an equation of state) defines a surface in $\mathbb{R}^3$. Suppose, at a point (p, v, t) all three partial derivates $\partial_p f$, $\partial_v f$ and $\partial_t f$ are non-zero. Then the implicit function theorem tells us that, near this point, we can solve the equation $f(p, v, t) = 0$ for each of the three variables in terms of the other two. Let me call these solutions $p = P(v, t)$, $v = V(p, t)$ and $t = T(p, v)$ so that

$$f(P(v, t), v, t) = f(p, V(p, t), t) = f(p, v, T(p, v)) = 0 . \tag{6.32}$$

In common physics notation, no distinction is made between the functions and the associated coordinates. Here, in order to avoid confusion, I have used lowercase letters for the coordinates and uppercase letters for the corresponding functions. The derivatives of our implicit functions P, V and T can be computed using Eq. (6.27) (or, alternatively, by differentiating Eq. (6.32) using the chain

rule) and this leads to

$$\begin{aligned} \partial_v P &= -\frac{\partial_v f}{\partial_p f} & \partial_t P &= -\frac{\partial_t f}{\partial_p f} \\ \partial_p V &= -\frac{\partial_p f}{\partial_v f} & \partial_t V &= -\frac{\partial_t f}{\partial_v f} \\ \partial_p T &= -\frac{\partial_p f}{\partial_t f} & \partial_v T &= -\frac{\partial_v f}{\partial_t f} \end{aligned} \tag{6.33}$$

Various conclusions can be easily obtained using these equations, for example,

$$\partial_v P\, \partial_t V\, \partial_p T = -1 \,, \qquad \partial_v P = (\partial_p V)^{-1} \,. \tag{6.34}$$

Do not let yourself be confused by the second of these equations which says we can simply "invert" individual partial derivatives. At first sight, this seems to be at odds with our general rule, Eq. (5.21), which says the Jacobi-matrix of the inverse function is the *matrix inverse* of the original Jacobi matrix. But it is important to note that this general rule was based on a set-up with independent coordinates $(x_1, \ldots, x_n)^T \in \mathbb{R}^n$. In the present case, on the other hand, we start with three independent coordinates $(p, v, t)^T \in \mathbb{R}^3$ but we impose the condition $f(p, v, t) = 0$. This means we are effectively working on a surface within $\mathbb{R}^3$ and only two coordinates are independent. The somewhat peculiar looking relations (6.34) are a consequence of this special set-up and are, therefore, not in contradiction with our general rule in Eq. (5.21). In fact, they were derived using the general chain rule, in the form of Eq. (6.27).

7 Vector fields and curves

In this subsection we take a short detour to discuss curves and how to integrate vector fields over them. Within our systematic build-up these topics are really part of section 3 on manifolds and section 2 on integration. However, curves - one-dimensional manifolds - are somewhat special and relatively easy to define. Also, integrating vector fields along curves only requires the one-dimensional integral which is already familiar. As we will see, knowing the integral of vector fields along curves, in physics often called 'line integral', will help us understand some of the properties of differential operators in the next section.

7.1 Definition of curves

Definition 7.1. *(Curves) A* **curve** *is a continuous map $X = (X_1, \ldots, X_n)^T : [a, b] \to \mathbb{R}^n$ where $[a, b] \subset \mathbb{R}$ is an interval. The points $X(a)$ and $X(b)$ are called the start and the end points of the curve, respectively. The curve is called* **closed** *if $X(a) = X(b)$. The curve is called (continuously)* **differentiable** *if all functions X_i, where $i = 1, \ldots, n$, are (continuously) differentiable. In this case, $\frac{dX}{dt}(t)$ is called the tangent vector to the curve at $X(t)$. If $\frac{dX}{dt}(t) \neq 0$ for all $t \in [a, b]$ the curve is called* **regular**.

The idea is that for every parameter value $t \in [a, b]$ the map X produces a certain vector $X(t) \in \mathbb{R}^n$ and, as t varies, a family of such vectors forming the curve is swept out. In the context of classical mechanics, $X(t)$ might be describing the trajectory of a point mass so that t should be interpreted as physical time and the tangent vector $\frac{dX}{dt}(t)$ as the velocity of the point mass. Let us discuss some examples of curves.

Example 7.1. Unit circle as a curve
Question: Write down the standard angular parametrisation for the unit circle S^1 and find the tangent vectors. Find another parametrisation of the (upper half of the) unit circle.

Answer: The unit circle $S^1 \subset \mathbb{R}^2$ is obtained from a map $X : [0, 2\pi] \to \mathbb{R}^2$ given by

$$X(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} . \tag{7.1}$$

Evidently this is a differentiable curve and, since $X(0) = X(2\pi) = (1, 0)^T$, it is closed. The tangent vector at $X(t)$ is given by

$$\frac{dX}{dt}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} . \tag{7.2}$$

Note this vector is indeed 'tangent' to the circle in the intuitive sense, as illustrated in Fig. 8. A unit circle (or, rather, its upper half) can be described by other maps for example by a map $\tilde{X} : [-1, 1] \to \mathbb{R}^2$ given by

$$\tilde{X}(s) = \begin{pmatrix} s \\ \sqrt{1 - s^2} \end{pmatrix} . \tag{7.3}$$

It is clear that the two curves X and $\tilde{X}$ are related by a reparametrisation $s = \cos(t)$ in the sense that $\tilde{X}(\cos(t)) = X(t)$. Two curves related by such a reparametrisation can, in fact, be considered to be the same curve. A key issue will be to define mathematical quantities for curves which are invariant under such reparametrisations.

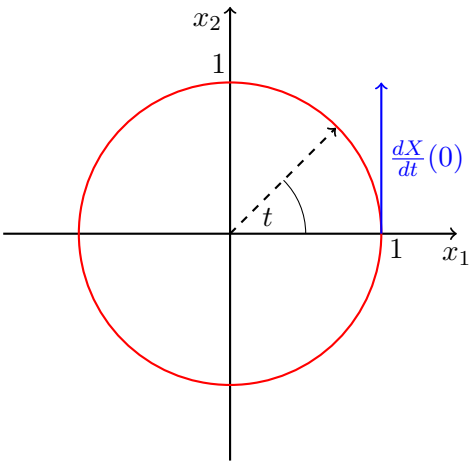


Figure 8: Unit circle $S^1 \subset \mathbb{R}^2$ (in red), parametrised by the angle t as in Eq. (7.1). The tangent vector $\frac{dX}{dt}(0) = (0, 1)^T$ at $(1, 0)^T$ (in blue) has been indicated.

Example 7.2. Intersection of two curves
Question: Suppose two curves $X : [a, b] \to \mathbb{R}^n$ and $Y : [c, d] \to \mathbb{R}^n$ intersect, so that $X(t_0) = Y(s_0)$ for $t_0 \in [a, b]$ and $s_0 \in [c, d]$. How would you define the intersection angle of the two curves? Consider $X(t) = (3 - 2t, 2t - 2, t - 1)^T$, where $t \in [0, 3]$ and $Y(s) = (-s^2, 1 + s, s^2)^T$ where $s \in [0, 2]$. Find the intersection point(s) of the two curves and their intersection angle(s).

Solution: If $X(t_0) = Y(s_0)$ is an intersection point, we can look at the tangent vectors to the two curves, $\frac{dX}{dt}(t_0)$ and $\frac{dY}{ds}(s_0)$ at that point. It makes sense to define the intersection angle θ between

the curves as the angle between those tangent vectors, so

$$\cos(\theta) = \frac{\frac{dX}{dt}(t_0) \cdot \frac{dY}{ds}(s_0)}{\left|\frac{dX}{dt}(t_0)\right| \left|\frac{dY}{ds}(s_0)\right|} \tag{7.4}$$

For the two example curves we have to solve the equations $X(t) = Y(s)$ to find the intersection point(s), so $3 - 2t = -s^2$, $2t - 2 = 1 + s$ and $t - 1 = s^2$, The only solution to these equations is for $t_0 = 2$ and $s_0 = 1$ with intersection point $X(2) = Y(1) = (-1, 2, 1)^T$. The tangent vector at this point are

$$\frac{dX}{dt}(2) = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} , \qquad \frac{dY}{ds}(1) = \begin{pmatrix} -2s \\ 1 \\ 2s \end{pmatrix}(1) = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}$$

Inserting these vectors into Eq. (7.4) gives $\cos(\theta) = \frac{8}{9}$.

Application 7.1. Projectile motion
Consider the motion of a mass point with mass m in $\mathbb{R}^2$ with coordinates $(x, y)^T$ and a gravitational force $F = (0, -mg)^T$ and g is the constant gravitational acceleration on earth. The corresponding Newton equations (where the dot denotes the time derivative)

$$\ddot{x} = 0 , \qquad \ddot{y} = -g , \tag{7.5}$$

are solved by

$$x(t) = x_0 + v_x t ; \qquad y(t) = y_0 + v_y t - \frac{1}{2} g t^2 , \tag{7.6}$$

where x_0, y_0, v_x, v_y are integration constants. For each choice of integration constants this solution describes a curve $X : \mathbb{R} \to \mathbb{R}^2$ given by

$$X(t) = \begin{pmatrix} x_0 + v_x t \\ y_0 + v_y t - \frac{1}{2} g t^2 \end{pmatrix} , \tag{7.7}$$

with tangent vector (=velocity)

$$\frac{dX}{dt}(t) = \begin{pmatrix} v_x \\ v_y - gt \end{pmatrix} . \tag{7.8}$$

Application 7.2. Hamiltonian mechanics, total derivatives and curves
In Hamiltonian mechanics we have coordinates q_i and momenta p_i, where $i = 1, \ldots, \nu$, so we are working with a space (also called phase space) of total dimension $n = 2\nu$. The dynamics of the system is encoded by the Hamiltonian $H = H(p, q)$ which is a function on phase space and the evolution of the system is obtained by solving Hamilton's equations

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} , \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} , \tag{7.9}$$

where the dot denotes time derivative. A solution $p_i = p_i(t)$, $q_i = q_i(t)$ to Hamilton's equations is a curve in phase space which represents the physical motion of the system. Suppose we have function F on phase space and we would like to work out how F changes along a solution curve, as a function of time, that is, we want to consider $F(p(t), q(t), t)$ (the final t-dependence allows for the possibility

that F depends explicitly on time). Applying the chain rule gives

$$\frac{dF}{dt} = \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial p_i}\dot{p}_i + \frac{\partial F}{\partial t} \overset{\text{Eq. (7.9)}}{=} \frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q_i} + \frac{\partial F}{\partial t} \,.$$

The first two terms on the RHS are an example of a **Poisson bracket** which, for two functions f, g on phase space, is defined as $\{f,g\} := \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}$. Note the Poisson bracket is anti-symmetric in its two arguments and, in particular, $\{f,f\} = 0$. With this notation, the change of F along a solution curve can be written as

$$\frac{dF}{dt} = \{F,H\} + \frac{\partial F}{\partial t} \,. \tag{7.10}$$

One conclusion is that a function F on phase space which does not explicitly depend on time (that is, $\frac{\partial F}{\partial t} = 0$) is preserved along a physical trajectory iff its Poisson bracket with the Hamiltonian vanishes, that is iff $\{F,H\} = 0$. In particular, any time-independent Hamiltonian is preserved, since $\{H,H\} = 0$.

7.2 Curves and the Jacobi matrix

Curves allow us to gain a better geometrical understanding of the Jacobi matrix. After all, a matrix can be seen as a linear map and it would 'like' to act on vectors. But which vectors does the Jacobi matrix act on?

Consider a map $F : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, which maps $x \in U$ to $y = F(x) \in \mathbb{R}^m$. Suppose we have a curve $X; [a,b] \to U$ and we consider its image $Y = F \circ X$ under the map F. What we would like to ask is how the tangent vectors of X and Y are related. A simple application of the chain rule gives

$$\frac{dY}{dt}(t) = \frac{\partial y}{\partial x}(X(t))\,\frac{dX}{dt}(t) \,. \tag{7.11}$$

in short, the tangent vectors of X and Y are related by matrix multiplication with the Jacobi matrix $DF = \frac{\partial y}{\partial x}$. Evidently this is the geometrical interpretation of the Jacobi matrix: it maps tangent vectors into each other.

Example 7.3. Mapping tangent vectors to polar coordinates
Question: Consider the unit circle $S^1 \subset \mathbb{R}^2$ with parametrisation $X(t) = (\cos(t), \sin(t))^T$, where $t \in [0, 2\pi)$ and the polar coordinate map $(r,\varphi) \mapsto (x,y) = (r\cos(\varphi), r\sin(\varphi))^T$. Use Eq. (7.11) to transform the tangent vectors to the circle into polar coordinates.

Answer: The Jacobi matrix for the change to polar coordinates is

$$\frac{\partial(x,y)}{\partial(r,\varphi)} = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix} .$$

However, for the present purpose we require the Jacobi matrix for the inverse function

$$\frac{\partial(r,\varphi)}{\partial(x,y)} = \left(\frac{\partial(x,y)}{\partial(r,\varphi)}\right)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos(\varphi) & r\sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} = \frac{1}{r}\begin{pmatrix} x & y \\ -\frac{y}{r} & \frac{x}{r} \end{pmatrix}$$

From Eq. (7.11) we then have

$$\begin{pmatrix} \frac{dr}{dt} \\ \frac{d\varphi}{dt} \end{pmatrix}(t) = \frac{\partial(r,\varphi)}{\partial(x,y)}\begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix}(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}\begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Note this result makes perfect sense: In polar coordinates the tangent vector to the unit circle only has a component in the φ direction (see Fig. 9).

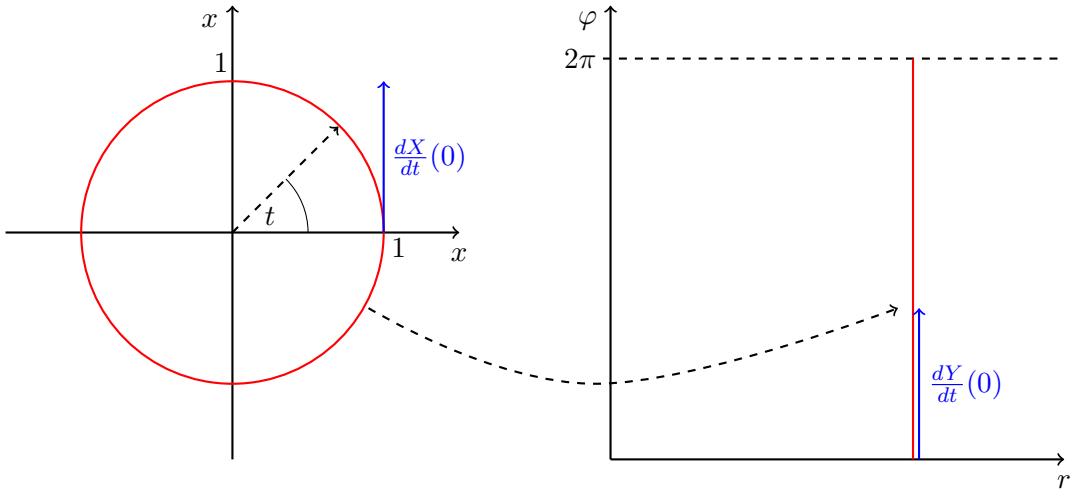


Figure 9: Unit circle $S^1 \subset \mathbb{R}^2$ parametrised by $X(t) = (\cos(t), sin(t))^T$ mapped into polar coordinates. Tangent vector are indicated in blue.

7.3 Integrating vector fields along curves

Suppose we have a differentiable curve $X : [a, b] \to \mathbb{R}^n$ and a vector field $A : U \to \mathbb{R}^n$ (where $X([a, b]) \subset U$ so that the vector field is defined at every point $X(t)$ of the curve). We would like to define an integral of the vector field along the curve which, intuitively, measures the vector field's flow along the curve. At any given curve point $X(t)$ this flow is given by projecting the vector field into the tangent direction of the curve, that is, it is given by the dot product

$$A(X(t)) \cdot \frac{dX}{dt}(t) \, . \tag{7.12}$$

The total flow is then obtained by integrating this expression over the curve parameter t (see Fig. 10). This suggests the following definition for the integral of A over X.

Definition 7.2. *(Integrating vector fields along curves) Let $X : [a, b] \to \mathbb{R}^n$ be a differentiable curve and $A : U \to \mathbb{R}^n$ a vector field, such that $X([a, b]) \subset U$. Then the curve (or line) integral of A along X is defined as*

$$\int_{X([a,b])} A \cdot dx := \int_a^b A(X(t)) \cdot \frac{dX}{dt}(t) \, dt \, , \tag{7.13}$$

provided the integral on the RHS exists.

The LHS of Eq. (7.13) is a symbolic expression which is defined by the RHS. Note this involves a simple one-dimensional (Riemannian) integral of the kind you are familiar with from basic one-dimensional calculus. We have motivated the above definition of a line integral intuitively but the real test is that the value of this integral is unchanged under re-parametrisation. Consider the curve $\tilde{X} : s([a, b]) \to \mathbb{R}^n$ given by

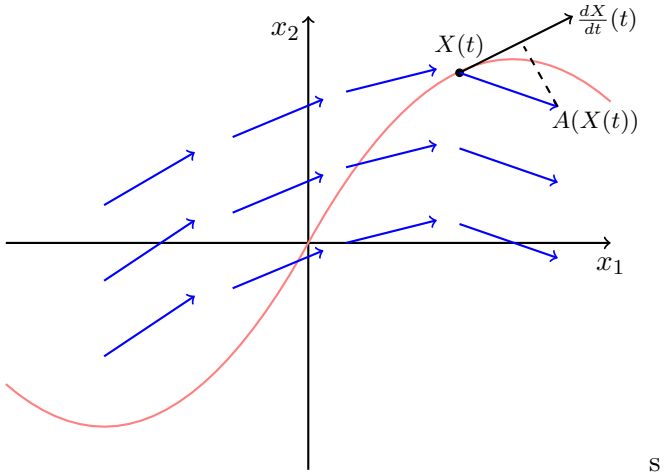


Figure 10: Integrating a vector field $A : \mathbb{R}^2 \to \mathbb{R}^2$ (blue) along a curve $X : [a, b] \to \mathbb{R}^2$ (red). The vector field $A(X(t))$ at a curve point $X(t)$ is projected into the direction $\frac{dX}{dt}(t)$ of the tangent vector in order to obtain the vector field 'flux' along the curve at that point. The integral of a vector field along a curve is defined by integrating this value over the curve parameter t.

$\tilde{X}(s) = X(t(s))$ for a diffeomorphic (that is, continuously differentiable, invertible and inverse continuously differentiable as well) re-parametrisation $t = t(s)$ with inverse $s = s(t)$. Then

$$\int_{s(a)}^{s(b)} A(\tilde{X}(s))) \cdot \frac{d\tilde{X}}{ds}(s)\, ds = \int_{s(a)}^{s(b)} A(X(t(s)))) \cdot \frac{dX}{dt}(t(s)) \frac{dt}{ds}(s)\, ds = \int_a^b A(X(t)) \cdot \frac{dX}{dt}(t)\, dt\,, \qquad (7.14)$$

where the last step follows from a change of variable in the one-dimensional integral. This shows the line integral is indeed invariant under re-parametrisation of the curve.

However, we have to be slightly careful with this statement. The re-parametrisation $t = t(s)$ is a diffeomorphism which means that $\frac{ds}{dt}(t) \neq 0$ for all $t \in [a, b]$ (see Cor. 6.1). If $\frac{ds}{dt} > 0$ then $s(a) < s(b)$ and the curve integral on the LHS of Eq. (7.14) is from lower to higher values. In this case, the LHS is a curve integral as in Def. 7.2 (which assumes integration from lower to higher values) and invariance under re-parametrisaion holds. On the other hand, if $\frac{ds}{dt} < 0$ then $s(a) > s(b)$ and to match up the LHS of Eq. (7.14) with Def. 7.2 we need to change the order of integration which, in turn, changes the result by a sign. Another way of saying this is that a curve comes with two **orientations** (meaning which way we 'go along' the curve) and if we pass between those orientations (as a re-parametrisation with $\frac{ds}{dr} < 0$ does) the curve integral changes by a sign. In short, curve integrals are invariant under orientation-preserving re-parametrisations but they change sign under orientation-reversing ones.

Example 7.4. Line integral in two dimensions
Question: Consider the upper half circle with the two parametrisations given in Exercise 7.1. Integrate the vector field $A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $A(x, y) = (-y, x)^T$ over the upper half cricle for both parametrisations. The results should be the same!
Answer: For the parametrisation $X(t) = (\cos(t), \sin(t))^T$, where $t \in [0, \pi]$ we have

$$A(X(t)) \cdot \frac{dX}{dt}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} = 1 \quad \Rightarrow \quad \int_{X[0,\pi])} A \cdot dx = \int_0^\pi dt = \pi\,.$$

On the other hand, for the parametrisation $X(t) = (1, \sqrt{1-t^2})^T$, where $t \in [-1, 1]$ we find

$$A(X(t)) \cdot \frac{dX}{dt}(t) = \begin{pmatrix} -\sqrt{1-t^2} \\ t \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\frac{t}{\sqrt{1-t^2}} \end{pmatrix} = -\frac{1}{\sqrt{1-t^2}}$$

$$\Rightarrow \quad \int_{X([-1,1])} A \cdot dx = -\int_1^{-1} \frac{1}{\sqrt{1-t^2}} = \pi \, .$$

Example 7.5. Line integral in three dimensions
Question: Consider the three-dimensional vector $A : \mathbb{R}^3 \to \mathbb{R}^3$ field defined by $A(x, y, z) = (xy, z, x^2 - y^2)^T$ and the curve $X : [0, 1] \to \mathbb{R}^3$ given by $X(t) = (t, t^2, t^3)^T$. Calculate the line integral of A over X.
Answer: The tangent vector to the curve is $\frac{dX}{dt}(t) = (1, 2t, 3t^2)^T$ and we have

$$A(X(t)) \cdot \frac{dX}{dt}(t) = \begin{pmatrix} t^3 \\ t^3 \\ t^2 - t^4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2t \\ 3t^2 \end{pmatrix} = t^3 + 5t^4 - 3t^6 \, ,$$

and, hence, the curve integral becomes

$$\int_{X([0,1])} A(X(t)) \cdot \frac{dX}{dt}(t) = \int_0^t (= t^3 + 5t^4 - 3t^6) = \frac{23}{28} \, .$$

Application 7.3. Curve integrals in physics
If a vector field F represent a physical force which acts on a point mass that is transported along a curve C then the curve integral

$$\int_C F \cdot dx \tag{7.15}$$

can be interpreted as the work done. For an electric field E the integral

$$\int_C E \cdot dx \tag{7.16}$$

represents the potential difference between the end points of the curve C.

8 Differential operators of vector calculus

For functions $f : U \to \mathbb{R}^m$ of n variables $x = (x_1, \ldots, x_n)^T$ the partial derivatives, $\partial_i f_j(x)$ where $i = 1, \ldots, n$ and $j = 1, \ldots, m$, at $x \in U$ form an $m \times n$ matrix, the total differential or Jacobi matrix, $D(f)(x)$, at $x \in U$. In this section we want to consider specific linear combinations of the partial derivatives $\partial_i f_j$ which are of particular interest as they measure certain interesting properties of the function f. This will lead us to the first order differential operations of gradient, divergence and curl each of which consist of specific linear combinations of $\partial_i f_j$. A mathematically satisfactory motivation for why we consider these linear combinations only arises in the context of differential forms which we do not discuss in the present lectures. (However, for the interested student, a no-nonsense appendix on differential forms has been attached to these notes.) Here, we rely on more heuristic motivations for the gradient, the divergence and the curl, which we now introduce in turn.

8.1 Gradient

Definition of gradient
For real-valued functions $f : U \to \mathbb{R}$ of n variables $x = (x_1, \ldots, x_n)^T$ the Jacobi matrix is a $1 \times n$ matrix, that is, a row vector. The **gradient** of f is, in fact, identical to this Jacobi matrix, so

$$\text{grad}(f)(x) = \nabla f(x) := D(f)(x) = \left(\frac{\partial f}{\partial x_i}, \ldots, \frac{\partial f}{\partial x_n} \right)(x) \, . \tag{8.1}$$

The symbol ∇ is called 'nabla' and it is used to denote the formal vector

$$\nabla = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n} \right) \tag{8.2}$$

which contains the partial derivatives and is also referred to as the **nabla operator**. Evidently, from Eq. (8.1), the gradient of a real-valued function is obtained by acting on the function with the nabla operator. Note that doing this produces a vector field, ∇f.

If we consider more general functions $f = (f_1, \ldots, f_m)^T : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, then the Jacobi matrix $D(f)(x)$ is an $m \times n$ matrix. Comparison with Eq. (5.2) shows that the rows of this matrix are precisely the gradients $\nabla f_j(x)$ of the component functions f_j. Hence, the Jacobi matrix for a vector-valued function can be written in terms of gradients as

$$Df(x) = \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{pmatrix}(x) \, . \tag{8.3}$$

Properties of gradient
Calculations for vector calculus operations are often facilitated by converting from vector to index notation and then working with partial derivatives whose rules are known. For the nabla operator the transition to index notation is accomplished by the simple equation

$$(\nabla f)_i = \partial_i f \qquad \text{or} \qquad \nabla_i = \partial_i \tag{8.4}$$

where we recall that the i^{th} component of a vector is indicated by brackets with a subscripts i. To see how this works in practice, let us show that the gradient is linear and satisfies a product rule, that is, (omitting the argument x, for simplicity)

$$\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g \, , \qquad \nabla(fg) = f \nabla g + g \nabla f \, , \tag{8.5}$$

for partially differentiable real-valued functions f, g and real numbers $\alpha, \beta \in \mathbb{R}$. Using index notation, the proofs are as follows.

$$\begin{aligned} (\nabla(\alpha f + \beta g))_i &\overset{\text{Eq. (8.5)}}{=} \partial_i(\alpha f + \beta g) \overset{\text{Eq. (4.5)}}{=} \alpha \partial_i f + \beta \partial_i g = (\alpha \nabla f + \beta \nabla g)_i \\ (\nabla(fg))_i &\overset{\text{Eq. (8.5)}}{=} \partial_i(fg) \overset{\text{Eq. (4.6)}}{=} f \partial_i g + g \partial_i f = (f \nabla g + g \nabla f)_i \end{aligned} \tag{8.6}$$

Example 8.1. Gradient of linear and quadratic functions
Question: Consider a linear function $l : \mathbb{R}^n \to \mathbb{R}$ defined by $l(x) = a \cdot x$, where $a \in \mathbb{R}^n$ is a constant vector, and a quadratic form $q : \mathbb{R}^n \to \mathbb{R}$, given by $q(x) = \frac{1}{2} x^T M x$, where M is a symmetric $n \times n$ matrix. Compute the gradients of l and q.

Answer: Using the basic rule $\partial_i x_j = \frac{\partial x_j}{\partial x_i} = \delta_{ij}$ we find for the linear function

$$\partial_i l(x) = \partial_i(a_j x_j) = a_j \delta_{ij} = a_i \quad \Rightarrow \quad \nabla l(x) = a \, .$$

For the quadratic form if follows that

$$\partial_i q(x) = \frac{1}{2}\partial_i(M_{jk} x_j x_k) = \frac{1}{2} M_{jk}(\delta_{ij} x_k + \delta_{ik} x_j) = M_{ij} x_j \quad \Rightarrow \quad \nabla q(x) = Mx \, .$$

Example 8.2. Gradient of radius-dependent functions
Question: Find the gradient of the function $V : \mathbb{R}^n \to \mathbb{R}$ given by $V(x) = \frac{1}{r(x)^p}$, where p is a number and $r(x) = |x|$ is the radius function. Also find the gradient for $V(x) = \frac{1}{|x-a|^p}$, where $a \in \mathbb{R}^n$ is a constant vector.

Answer: In the first case, a quick calculation using the chain rule (and $\partial_i r = \frac{x_i}{r}$) gives

$$\partial_i V = \frac{d(r^{-p})}{dr} \partial_i r = -\frac{p x_i}{r^{p+2}} \quad \Rightarrow \quad \nabla V = -\frac{px}{r^{p+2}} \, .$$

For the second part we have

$$\partial_i V = -\frac{p(x_i - a_i)}{|x-a|^{p+2}} \quad \Rightarrow \quad \nabla V = -\frac{p(x-a)}{|x-a|^{p+2}} \, .$$

Interpretation of gradient
The gradient is of course a convenient way of collecting all the partial derivatives of a real-valued function into a single vector field. Is there any other useful interpretation? Consider our result, Eq. (4.11), for the directional derivative which can be re-written in terms of the gradient as

$$\partial_v f(x) = v \cdot \nabla f(x) \, . \tag{8.7}$$

Let us consider unit vectors $n \in \mathbb{R}^n$, so $|n| = 1$, and ask in which unit vector direction the directional derivative $\partial_n f(x)$ is extremal.

Theorem 8.1. *For a differentiable real-valued function $f : U \to \mathbb{R}$, with $\nabla f(x) \neq 0$ consider unit vectors $n \in \mathbb{R}^n$. The directional derivative $\partial_n f(x)$ is maximal (minimal) iff n and $\nabla f(x)$ are parallel (anti-parallel).*

Proof. From Eq. (8.7) we have

$$\partial_n f(x) = n \cdot \nabla f(x) = |\nabla f(x)| \cos\phi \, ,$$

where ϕ is the angle between n and $\nabla f(x)$. Clearly, the above expression is maximal (minimal) iff $\phi = 0$ ($\phi = \pi$) and, hence, iff n and $\nabla f(x)$ are parallel (anti-parallel). □

The statement of the previous theorem can be paraphrased by saying that the gradient $\nabla f(x)$ (negative gradient $-\nabla f(x)$) points in the direction of the steepest ascent (descent) of the function at x. This interpretation of the gradient gives rise to an important application, known as **gradient descent**, which we explain in the following.

Application 8.1. Gradient descent
Finding the minimum (or a minimum) of a real-valued function $f : U \to \mathbb{R}$ which depends on coordinates $x = (x_1, \ldots, x_n)^T$ is a simple task conceptually but can be extremely difficult to carry out in practice, particularly if the domain dimension n is large. A well-known and widely-used heuristic method to do this is **gradient descent** which works as follows. First, choose a starting point $x^{(0)} \in U$ arbitrarily. Then, define a sequence of points $x^{(k)} \in U$ recursively by the relation

$$x^{(k+1)} = x^{(k)} - \lambda \nabla f\left(x^{(k)}\right) , \tag{8.8}$$

where $\lambda > 0$ is a real parameter. The idea is simple. From Theorem 8.1 the negative gradient points into the direction of the steepest descent of the function, so if we follow this direction for an amount controlled by the parameter λ and keep repeating this process we might arrive at a minimum of f. Of course, success is not guaranteed and can very much depend on the choice of initial point $x^{(0)}$ and the choice of λ. If λ is chosen too small then it may take many iterations for the sequence $x^{(k)}$ to cover some distance in the function's domain and a minimum might not be reached in the number of steps carried out. On the other hand, if λ is too large the steps described by Eq. (8.8) might become too large and the sequence 'overshoots' the minimum. An example of a gradient descent for a real-valued function of two variables is shown in Fig. 11.

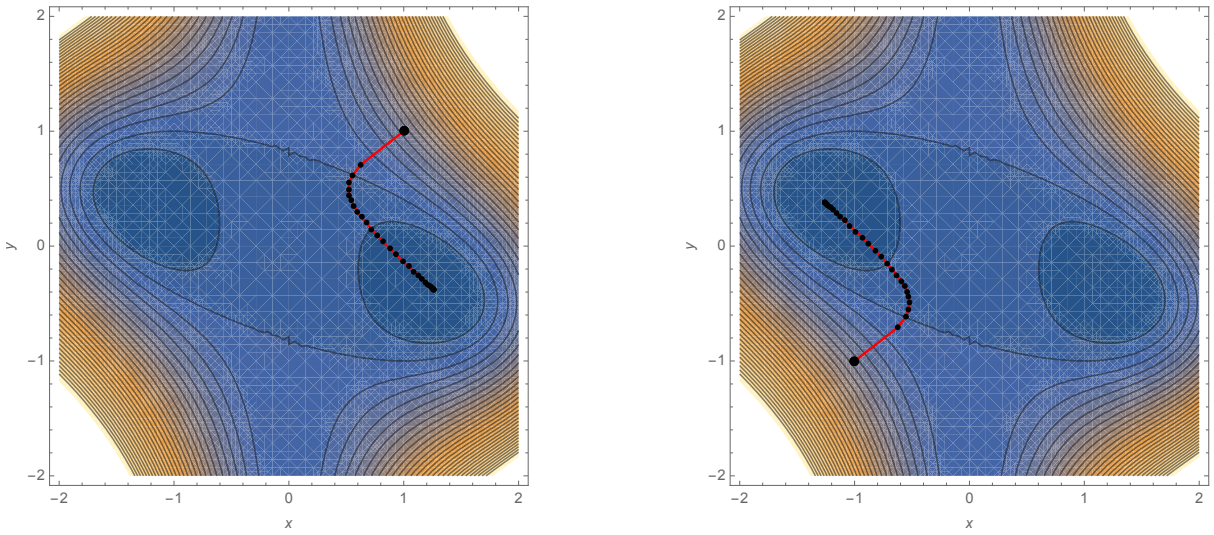


Figure 11: Contour plot of the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = x^2y^2 - x^2\cos(x + y)$. The plot on the left starts the gradient descent at $(x, y) = (1, 1)$, the one on the right at $(x, y) = (-1, -1)$. In either case, we have chosen $\lambda = 0.1$ and have performed 30 steps.

Exercise 8.2. *Write a piece of code to minimise a function using gradient descent. (For example, use Python or Mathematica.)*

Application 8.2. Training neural networks
Training neural networks is typically accomplished using a version of gradient descent, called **stochastic gradient descent**. To explain how this works we have to start by introducing neural networks. Using mathematical notation, we can view a neural network as a family of functions $f_\theta : \mathbb{R}^n \to \mathbb{R}^m$

which is parametrised by $\theta = (\theta_1, \dots, \theta_N)^T \in \mathbb{R}^N$. Usually, this family of functions is constructed by function composition, meaning that we have a collection of simpler function families $f_{\nu,\theta^{(\nu)}} : \mathbb{R}^{n_\nu} \to \mathbb{R}^{n_{\nu+1}}$, where $\nu = 1, \dots, \delta$, each with their own set of parameters $\theta^{(\nu)}$. The neural network is then constructed by composing these building blocks, that is,

$$f_\theta = f_{\delta,\theta^{(\delta)}} \circ \cdots \circ f_{2,\theta^{(2)}} \circ f_{1,\theta^{(1)}} : \mathbb{R}^{n_1} \to \mathbb{R}^{n_{\delta+1}} , \tag{8.9}$$

where $\theta = (\theta^{(1)}, \dots, \theta^{(\delta)})$ is the collection of all parameters which arise. The number δ of compositions is also referred to as the **depth** and the typical dimensions, n_ν, as the **width** of the neural net.

Many choices for the building blocks are possible but the simplest and most common one is

$$f_{\nu,\theta^{(\nu)}}(x) = \sigma(W^{(\nu)}x + b^{(\nu)}) , \tag{8.10}$$

where $W^{(\nu)}$ is a matrix of size $n_\nu \times n_{\nu-1}$ whose entries are called **weights** and $b^{(\nu)} \in \mathbb{R}^{n_\nu}$ is a vector with entries called **biases**. Together these weights and biases form the parameters of the building block, so $\theta^{(\nu)} = (W^{(\nu)}, b^{(\nu)})$. The function $\sigma : \mathbb{R} \to \mathbb{R}$ in Eq. (8.10) is called **activation function** and acts on each component of its vector argument. There are many choices for this function but a common one is the ReLU function ('rectified linear unit') defined by $\sigma(z) = \theta(z)z$. In essence, the building block in Eq. (8.10) performs an affine transformation combined with a component-wise non-linearity encoded in the activation function. A neural network (8.9) composed from building blocks of the type (8.10) only is also called a **fully-connected** neural network. In summary, a neural net is a family of functions constructed from simpler building blocks, often affine transformations with a mild non-linearity, by function composition.

How are neural networks $f_\theta : \mathbb{R}^n \to \mathbb{R}^m$ trained? The goal is always to find suitable parameter values θ so that f_θ approximates a certain desired function. For the simplest type of training, called **supervised training**, the information about the desired function is provided by data of the form $(x^{(a)}, y^{(a)}) \in \mathbb{R}^n \times \mathbb{R}^m$, with $a = 1, \dots, A$, where $x^{(a)}$ are possible input vectors for the neutral network and $y^{(a)}$ the associated desired output vectors. For example, the $x^{(a)}$ may represent pictures of cats and dogs (every picture can be converted into a vector) and the $y^{(a)}$ may be two-dimensional probability vectors indicating whether the picture shows a cat or a dog. We would like to adjust the neural network parameters θ so that the data is reproduced well and the neural network can tell cat and dog pictures apart. A measure for how well the neural net approximates the data is given by the **loss function**, for example

$$L(\theta) = \sum_{a \in B} |f(x^{(a)}) - y^{(a)}|^2 . \tag{8.11}$$

Here, the sum runs over a (random) subset of the data, as encoded in the index set B, also called a **batch**. For an optimal fit to the data we need to minimise the loss function and this is commonly done by gradient descent, as described in Application 8.1, applied to the loss function L, seen as a function of the neural network parameters θ. One additional subtlety is that the batch B is changed for each step of the gradient descent, and this process is referred to as **stochastic gradient descent**. Once the optimal values $\theta_{\min}$ of the neural network parameters have been found in this way, the neural net $f_{\theta_{\min}}$ can be used to make predictions $y = f_{\theta_{\min}}(x)$ when applied to input data x not previously used for training. For example, we can input new cat or dog pictures hoping that the trained neural network can tell them apart.

Exercise 8.3. *Find out about the MNIST data set of handwritten digits. Train a neural network to recognise these hand-written numbers. (Using Python or Mathematica would work.)*

Vector fields from gradients

As we have seen above, the gradient acts on scalar fields and produces vector fields. It is, therefore, natural to ask which vector fields can be obtained as gradients of scalar fields. To discuss this we start with some terminology.

A (partially differentiable) vector field $A : U \to \mathbb{R}^n$ which can be obtained as a gradient of a real-valued function $f : U \to \mathbb{R}$, such that $A = \nabla f$, is called **conservative**. In this case, the function f is also referred to as a **potential function** for the vector field A.

Now suppose the vector field A is conservative and $A = \nabla f$ or, in components $A_i = \partial_i f$. Then it follows that

$$\partial_j A_i = \partial_j \partial_i f \overset{\text{Thm. } 4.1}{=} \partial_i \partial_j f = \partial_i A_j \tag{8.12}$$

Hence we have the following statement.

Lemma 8.1. *Let $A : U \to \mathbb{R}^n$ be a partially differentiable conservative vector field. Then*

$$\partial_i A_j = \partial_j A_i \quad \textit{for all} \quad i, j = 1, \ldots, n \,. \tag{8.13}$$

Another way of saying this is that the conditions (8.13) are necessary for A to be conservative and can be used as a test. If they fail then A is definitely not conservative. On the other hand, what can we say if the conditions (8.13) are satisfied? Is this sufficient for A to be conservative? Physicist tend to answer this question with an unconditional 'yes' but this is, in fact, only correct subject to additional assumptions. In fairness, these assumptions are often satisfied for physics examples. But the full story is (fortunately) significantly more complicated and interesting as we will explain later, when we have a few more techniques available.

Example 8.3. Conservative and non-conservative vector fields
Question: Consider the two-dimensional vector fields A, C, defined by $A(x_1, x_2) = (x_1 x_2, \overset{!}{\tfrac{1}{2}} x_1^2)^T$ and $C(x_1, x_2) = \frac{1}{x_1^2 + x_2^2}(-x_2, x_1)^T$ and the three-dimensional vector field B given by $B(x_1, x_2, x_3) = (x_2^2, x_3^2, x_1 x_2)^T$. Check if these vector fields satisfy the necessary condition for conservative vector fields from Lemma 8.1. Find a potential function where possible.

Answer: The vector field A satisfies the necessary condition since

$$\frac{\partial A_1}{\partial x_2} = x_1 = \frac{\partial A_2}{\partial x_1} \,.$$

A potential function f with $A = \nabla f$ is given by $f = \frac{1}{2} x_1^2 x_2$, so A is conservative.

For B we have

$$\frac{\partial B_1}{\partial x_2} = 2x_2 \neq 0 = \frac{\partial B_2}{\partial x_1} \,,$$

so B is definitely not conservative.

Finally, C satisfies the necessary condition since

$$\frac{C_1}{\partial x_2} = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} = \frac{\partial C_2}{\partial x_1} \,,$$

however, no potential function exists and C is not conservative. (For a clear-cut argument see Section 8.5.)

Curve integrals for conservative vector fields
What happens if we integrate a conservative vector field ∇f with potential $f : U \to \mathbb{R}$ over a curve

$X : [a,b] \to U$? Following Def. 7.2 we have

$$\int_{X([a,b])} \nabla f \cdot dx = \int_a^b dt\, \nabla f(X(t)) \cdot \frac{dX}{dt}(t) = \int_a^b dt \sum_{i=1}^n \frac{\partial f}{\partial x_i}(X(t)) \frac{dX_i}{dt}(t)$$
$$= \int_a^b dt\, \frac{d}{dt} f(X(t)) = f(X(b)) - f(X(a)) \,, \tag{8.14}$$

where the second last equality follows from the chain rule. The above result shows that the value of such a curve integral over a conservative vector field only depends on the values of the potential function at the start and end points of the curve. In other words, given two points $x^{(0)}, x^{(1)} \in \mathbb{R}^n$, the line integral of a conservative vector field gives the same value for any curve $X : [a,b] \to U$ with $X(a) = x^{(0)}$ and $X(b) = x^{(1)}$. To paraphrase, the value of a line integral over a conservative vector field only depends on the endpoints of the curve and is 'independent of the path taken'. In particular, for any closed curve, that is, a curve with $X(a) = X(b)$, Eq. (8.14) implies that

$$\int_{X([a,b])} \nabla f \cdot dx = 0 \,. \tag{8.15}$$

The line integral of a conservative vector field over a closed curve vanishes. We can turn this around and observe that these properties are necessary conditions for a vector field to be conservative. In other words, if the line integral of a vector field depends on the path taken or the line integral of a vector field over a closed curve is non-zero then this vector field cannot be conservative.

8.2 Divergence

A property of (partially differentiable) vector fields $A = (A_1, \ldots, A_n)^T : U \to \mathbb{R}^n$ we would like to quantify is the local strength of its sources. By this we mean the net flow of the vector field in and out of a given small volume around a certain point x. Such a net flow clearly depends on the way the vector field varies close to the point in question and this, in turn, is controlled by the partial derivatives $\partial_i A_j(x)$ at x. A measure for the local source strength, which is called the divergence of the vector field, can, therefore, be expected to be some linear combination of these partial derivatives. The question is: Which linear combination?

Definition of divergence

To work this out we first consider a simple two-dimensional case as shown in Fig. 12 where a given point x is made the center of a square with length ϵ. We need to compute the total flux through the boundaries of this square and we start by working out the values of the vector field A at the four points indicated on the boundaries $(1,+)$ and $(1,-)$ in Fig. 12. They are given by

$$A(x \pm \frac{\epsilon}{2} e_1 \pm \eta e_2) = A(x) \pm \frac{1}{2} \frac{\partial A}{\partial x_1}(x)\epsilon \pm \frac{\partial A}{\partial x_2}(x)\eta + \mathcal{O}(\epsilon^2) \tag{8.16}$$

Summing up the two contributions each for $(1,+)$ and for $(1,-)$, multiplying by $\epsilon/2$ and dotting into $\pm e_1$, the relevant normal vector, gives the total flux across these boundary components as

$$\text{flux}_{(1,\pm)} = \pm \epsilon A_1(x) + \frac{1}{2} \frac{\partial A_1}{\partial x_1}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \,. \tag{8.17}$$

Summing up these two fluxes, and the analogous ones for the other two boundaries $(2,+)$ and $(2,-)$ which follow from a similar calculation, we find for $\text{flux}_{(1)}$, the flux through $(1,+)$ and $(1,-)$, and $\text{flux}_{(2)}$, the flux through $(2,+)$ and $(2,-)$, that

$$\text{flux}_{(a)} = \frac{\partial A_a}{\partial x_a}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \quad \text{where} \quad a = 1,2 \,. \tag{8.18}$$

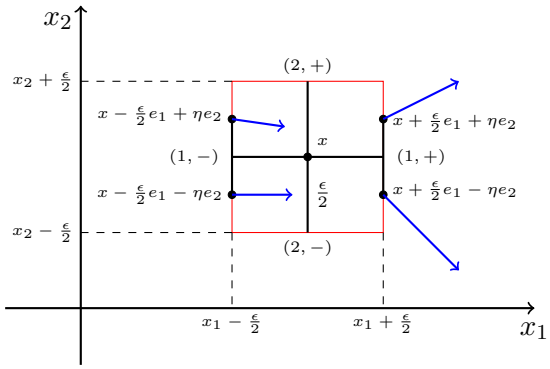


Figure 12: Calculating the total flow of a two-dimensional vector field $A : \mathbb{R}^2 \to \mathbb{R}^2$ (in blue) through a square of length ϵ around $x \in \mathbb{R}^2$. The four sides of the square are labelled by $(1,+)$, $(1,-)$, $(2,+)$ and $(2,-)$.

This means the total flux through the square's boundary is

$$\text{flux} = \text{flux}_{(1)} + \text{flux}_{(2)} = \frac{\partial A_1}{\partial x_1}(x)\epsilon^2 + \frac{\partial A_2}{\partial x_2}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \; . \tag{8.19}$$

If we define the divergence of the vector field A at x as (the limes $\epsilon \to 0$ of) the flux through the square's boundary divided by its area we have

$$\text{div}(A)(x) = \lim_{\epsilon \to 0} \frac{\text{flux}}{\epsilon^2} = \sum_{i=1}^{2} \frac{\partial A_i}{\partial x_i}(x) \; . \tag{8.20}$$

The above calculation has been carried out in two dimensions so that we can easily refer to Fig. 12. However, its generalisation to n dimenions is straightforward (although somewhat more tedious) and involves considering an n-dimensional cube around a point $x \in U$. This leads to the obvious generalisation of Eq. (8.20), namely that the divergence of a (partially differentiable) n-dimensional vector field $A : U \to \mathbb{R}^n$ is given by

$$\text{div}(A)(x) = \sum_{i=1}^{n} \frac{\partial A_i}{\partial x_i}(x) = \nabla \cdot A(x) = \partial_i A_i(x) \; . \tag{8.21}$$

Note, the divergence can be written as a dot product of the nabla operator and the vector field or, alternatively, using index notation with Einstein summation convention as on the RHS of Eq. (8.21). While the gradient takes a scalar field argument and outputs a vector field the divergence does the opposite: its argument is a vector field and the output a scalar.

Properties of the divergence

As we have made plausible the divergence measures the local source strength of a vector field. We will confirm this interpretation more elegantly (and for volumes more general than just cubes) later, by using Gauß's theorem. On a more basic level we can ask about product rules for the divergence. This requires a situation where two maps can be multiplied by some multiplication rule to produce a vector field (for which we can then consider the divergence). One such case arises for a differentiable vector field $A : U \to \mathbb{R}^n$ and a differentiable scalar field $f : U \to \mathbb{R}$, whose ordinary product $fA : U \to \mathbb{R}^n$ is another (differentiable) vector field. Its divergence can be worked out from the product rule (omitting the point x)

$$\nabla \cdot (fA) = \nabla f \cdot A + f \nabla \cdot A \; . \tag{8.22}$$

When writing such equations it is worthwhile checking that every term has a co-domain of the same dimension or, in physics parlance, that all terms are either vectors or scalars. For the above product rule this is indeed the case, since all terms are scalars. To prove Eq. (8.22) we use index notation and apply the product rule for partial differentials.

$$\nabla \cdot (fA) = \partial_i (fA)_i = \partial_i (fA_i) = \partial_i (f) A_i + f \partial_i A_i = \nabla f \cdot A + f \nabla \cdot A \, .$$

In three dimensions there is another possibility. For two three-dimensional (differentiable) vector fields $A, B : U \to \mathbb{R}^3$ the cross product $A \times B : U \to \mathbb{R}^3$, defined by $((A \times B)(x))_i = \epsilon_{ijk} A_j(x) B_k(x)$, is another three-dimensional (differentiable) vector field. Its divergence is subject to the product rule

$$\nabla \cdot (A \times B) = B \cdot (\nabla \times A) - A \cdot (\nabla \times B) \, , \tag{8.23}$$

where we have introduced a new operation for three-dimensional vector fields, called the **curl**, which is defined by

$$\mathrm{curl}(A) := \nabla \times A \, , \qquad (\nabla \times A)_i = \epsilon_{ijk} \partial_j A_k \, . \tag{8.24}$$

We will discuss the curl in more detail below but for now we just prove Eq. (8.24) by an index calculation [3].

$$\nabla \cdot (A \times B) = \partial_i (A \times B)_i = \partial_i \epsilon_{ijk} A_j B_k = \epsilon_{ijk} (\partial_i A_j) B_k + \epsilon_{ijk} A_j \partial_i B_k = B \cdot (\nabla \times A) - A \cdot (\nabla \times B) \, .$$

Example 8.4. Divergence of a radial vector field
Question: Compute the divergence of the n-dimensional vector field $E : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n$ given by $E(x) = \frac{x}{r^q}$, where q is a number and $r(x) = |x|$ is the radius function. Do the same for the vector field $\tilde{E} : \mathbb{R}^n \setminus \{a\} \to \mathbb{R}^n$ given by $\tilde{E}(x) = \frac{x-a}{|x-a|^p}$. Specialise your results to the case $q = n$.
Answer: For the vector field E we have (using $\partial_i x_j = \delta_{ij}$)

$$\nabla \cdot E = \partial_i E_i = \frac{\partial_i x_i}{r^q} + x_i \frac{dr^{-q}}{dr} \partial_i r = \frac{n}{r^q} - q x_i \frac{1}{r^{q+1}} \frac{x_i}{r} = \frac{n-q}{r^q} \, .$$

If we choose $p = n$, then $\nabla \cdot A = 0$ so this vector field has a vanishing source on $\mathbb{R}^n \setminus \{0\}$.
For the more general case of vector field $\tilde{E}$ we find a similar result,

$$\nabla \cdot \tilde{E} = \partial_i \frac{x_i - a_i}{|x-a|^p} = \frac{n-p}{|x-a|^p} \, ,$$

so that, for $p = n$, the vector field $\tilde{E}$ is source-free, that is $\nabla \cdot \tilde{E} = 0$ on $\mathbb{R}^n \setminus \{a\}$.

8.3 Curl

As we have discussed above, the divergence of a vector field measures its local source strength. Another interesting property of a vector field A, called the curl of A, measures how much it locally 'curls' at any given point x. This can be measured by working out the total flow of the vector field along a small closed curve around x. As for the divergence, this depends on the local change of the vector field near x which, in turn, is controlled by the partial derivatives $\partial_i A_j$. The task is to find the linear combination of these partial derivatives which corresponds to the curl.

[3] A word of caution is in order when calculating with the nabla operator. Do not perform these calculations by applying the formulae for cross and dot products for ordinary vectors which you know from linear algebra. These formulae rely on the components of all vectors commuting which is not the case here. The components of the nabla operator are partial derivatives. They do not commute with the other components which are typically functions.

Definition of curl in two dimensions

To make the discussion accessible we start in two dimensions, that is with a vector field $A : U \to \mathbb{R}^2$ and we consider a square with length ϵ centred on a given point $x \in U$. The closed curve we would like to consider is the boundary of the square. The situation is just as shown in Fig. 12 except that we should now focus on the direction of the vector field along the boundary rather than the one normal to it as we have done for the divergence. The value of A along the boundary segments $(1,+)$ and $(1,-)$ in Fig. 12 has been given in Eq. (8.16). If we add the two contributions for the $(1,+)$ part and the two contributions for the $(1,-)$ part, multiply with $\epsilon/2$ and dot into $\pm e_2$ (the direction along these curve segments) we find for the flow along these two boundary segments

$$\text{flow}_{(1,\pm)} = \pm\epsilon A_2(x) + \frac{1}{2}\frac{\partial A_2}{\partial x_1}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \; . \tag{8.25}$$

The sum of those two contributions, denoted $\text{flow}_{(1)}$, and the sum of the flows along the other two boundaries $(2,+)$ and $(2,-)$, denoted $\text{flow}_{(2)}$ and obtained by an analogous calculation, are given by

$$\text{flow}_{(1)} = \frac{\partial A_2}{\partial x_1}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \; , \qquad \text{flow}_{(2)} = -\frac{\partial A_1}{\partial x_2}(x)\epsilon^2 + \mathcal{O}(\epsilon^3) \; . \tag{8.26}$$

We define the curl as the (limes $\epsilon \to 0$ of the) total flow along the curve divided by the bounded area so

$$\text{curl}(A)(x) := \lim_{\epsilon\to 0} \frac{\text{flow}_{(1)} + \text{flow}_{(2)}}{\epsilon^2} = \frac{\partial A_2}{\partial x_1}(x) - \frac{\partial A_1}{\partial x_2}(x) \; . \tag{8.27}$$

Using index notation, the curl of a two-dimensional vector field can also be written as

$$\text{curl}(A) = \epsilon_{ij}\partial_i A_j \tag{8.28}$$

where ϵ_{ij} is the two-dimensional Levi-Civita tensor defined by $\epsilon_{12} = 1$, $\epsilon_{21} = -1$ and $\epsilon_{11} = \epsilon_{22} = 0$. Evidently, the curl of a two-dimensional vector field leads to a scalar field.

Curl in any dimension

The previous two-dimensional discussion can be easily generalised to an n-dimensional vector field $A : U \to \mathbb{R}^n$. In two dimensions we only have one square loop around a given point x to consider but there is more structure in n dimensions. In this case, a square loop around x as in Fig. 12 can be contained in any of the planes associated to two coordinates x_i, x_j (that is, more concretely, in the planes $x + \mathbb{R}e_i + \mathbb{R}e_j$). For any of these choices we can carry out the same calculation as we did in two dimensions and this leads to the result in Eq. (8.39) but with the replacements $1 \mapsto i$ and $2 \mapsto j$. Hence the components of the general curl, denoted by the symbol $\text{Curl}(A)$, are labelled by two indices and are given by

$$(\text{Curl}(A)(x))_{ij} = \frac{\partial A_j}{\partial x_i}(x) - \frac{\partial A_i}{\partial x_j}(x) = \partial_i A_j(x) - \partial_j A_i(x) \; . \tag{8.29}$$

Evidently, this expression is anti-symmetric in (ij) so we can think of the n-dimensional curl as taking values which form an anti-symmetric $n \times n$ matrix. (More accurately, the curl is an anti-symmetric rank two tensor field, taking values in $\Lambda^2\mathbb{R}^n \cong \mathbb{R}^{n(2-1)/2}$ – see Appendix A for details.)

By comparison with Lemma 8.1 we see that for an n-dimensional vector field $A : U \to \mathbb{R}^n$ to be conservative it is necessary that the curl vanishes, that is, $\text{Curl}(A) = 0$.

As we have just seen, in general the curl of an n-dimensional vector field is a rank two anti-symmetric tensor field with $n(n-1)/2$ independent components, a new and somewhat unfamiliar object. However, there are two special dimensions when this object can be re-written in a more familiar way. In two dimensions,

$n = 2$, we have $n(n-1)/2 = 1$, so the curl has only one component, that is, is can be viewed as a scalar. This is precisely the two-dimensional version of the curl in Eq. (8.28). In three dimensions, $n = 3$, we have $n(n-1)/2 = 3$, so the curl has three components. This means we can attempt to re-interpret the three-dimensional curl as a vector field.

Curl in three dimensions

The curl of a three-dimensional vector field $A : U \to \mathbb{R}^3$ has three independent components which we can take to be $(\text{Curl}(A)(x))_{23}$, $(\text{Curl}(A)(x))_{31}$ and $(\text{Curl}(A)(x))_{12}$. They can be arranged into a three-dimensional vector, referred to as $\text{curl}(A)(x)$, which is explicitly given by

$$\text{curl}(A)(x) = \begin{pmatrix} (\text{Curl}(A)(x))_{23} \\ (\text{Curl}(A)(x))_{31} \\ (\text{Curl}(A)(x))_{12} \end{pmatrix} = \begin{pmatrix} \partial_2 A_3 - \partial_3 A_2 \\ \partial_3 A_1 - \partial_1 A_3 \\ \partial_1 A_2 - \partial_2 A_1 \end{pmatrix} (x) \tag{8.30}$$

The expression on the RHS can be re-written as a formal cross product between the nabla operator and the vector field, that is,

$$\text{curl}(A)(x) = \nabla \times A(x) \quad \text{or, in index notation} \quad (\nabla \times A(x))_i = \epsilon_{ijk} \partial_j A_k(x) \, . \tag{8.31}$$

Properties of curl in three dimensions

As we have done for the gradient and the divergence, we can ask about product rules for curl. We restrict this discussion to the three-dimensional case. Consider a partially differentiable scalar $f : U \to \mathbb{R}$ and a partially differentiable vector field $A : U \to \mathbb{R}^3$, with $U \subset \mathbb{R}^3$ open. Then the product fA is another vector field in three dimensions whose curl can be computed from the formula

$$\nabla \times (fA) = \nabla f \times A + f \nabla \times A \, . \tag{8.32}$$

Using index notation the proof is simple.

$$(\nabla \times (fA))_i = \epsilon_{ijk} \partial_j (fA)_k = \epsilon_{ijk} \partial_j (f A_k) = \epsilon_{ijk} (\partial_j f) A_k + f \epsilon_{ijk} \partial_j A_k = (\nabla f \times A + f \nabla \times A)_i \, .$$

Example 8.5. Divergence and curl in three dimensions

Question: For the three-dimensional vector fields E defined on $\mathbb{R}^3 \setminus \{0\}$ and given by $E(x) = \frac{x}{r^3}$ compute the divergence and the curl.

Answer: For the vector field E we find (using $\partial_i x_j = \delta_{ij}$)

$$\nabla \cdot E = \partial_i E_i = \partial_i \left(\frac{x_i}{r^3} \right) = \frac{3}{r^3} - x_i \frac{3}{r^4} \frac{x_i}{r} = 0 \tag{8.33}$$

$$(\nabla \times E)_i = \epsilon_{ijk} \partial_j E_k = \epsilon_{ijk} \partial_j \left(\frac{x_k}{r^3} \right) = \epsilon_{ijk} \left(\frac{\delta_{jk}}{r^3} - x_k \frac{3}{r^4} \frac{x_j}{r} \right) = 0 \tag{8.34}$$

Example 8.6. Another product rule for the curl

Question: For two partially differentiable vector fields $A, B : U \to \mathbb{R}^3$ proof the product rule

$$\nabla \times (A \times B) = (B \cdot \nabla) A + (\nabla \cdot B) A - (A \cdot \nabla) B - (\nabla \cdot A) B \, . \tag{8.35}$$

Answer: Using index notation we have

$$\begin{aligned}(\nabla \times (A \times B))_i &= \epsilon_{ijk}\partial_j (A \times B)_k = \epsilon_{ijk}\partial_j(\epsilon_{klm}A_l B_m) = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j(A_l B_m) \\ &= \partial_j(A_i B_j) - \partial_j(A_j B_i) = B_j\partial_j A_i + A_i\partial_j B_j - A_j\partial_j B_i - B_i\partial_j A_j \\ &= ((B\cdot\nabla)A + (\nabla\cdot B)A - (A\cdot\nabla)B - (\nabla\cdot A)B)_i\end{aligned}$$

Example 8.7. Divergence and curl in terms of the Jacobi matrix
Question: Write the divergence and the curl of a vector field in n dimensions in terms of its Jacobi matrix.

Answer: The Jacobi matrix DA of a vector field A has the entries $(DA)_{ij} = \partial_j A_i$. For the divergence this means

$$\nabla \cdot A = \sum_i \partial_i A_i = \sum_i (DA)_{ii} = \mathrm{tr}(DA)\ .$$

The divergence is the trace of the Jacobi matrix.

The components of the curl are

$$(\mathrm{Curl}(A))_{ij} = \partial_i A_j - \partial_j A_i = (DA)_{ji} - (DA)_{ij} \quad \Rightarrow \quad \mathrm{Curl}(A) = DA^T - DA\ .$$

This means the curl (in its formulation for arbitrary dimensions) is given by the anti-symmetric part of the Jacobi matrix.

Application 8.3. Maxwell's equations in vacuum
Maxwell's equations are a set of first order partial differential equations for the electric field E and the magnetic field B, given a charge density ρ (a scalar) and a current density J (a vector field). In CGS units (c is the speed of light and the dot denotes the time derivative) these equations read

$$\begin{aligned} \nabla\cdot E &= 4\pi\rho && \text{Gauß's law} \\ \nabla\times B - \tfrac{1}{c}\dot{E} &= \tfrac{4\pi}{c}J && \text{Ampère-Maxwell law} \\ \nabla\times E + \tfrac{1}{c}\dot{B} &= 0 && \text{Faraday's law} \\ \nabla\cdot B &= 0 && \text{Gauß's law for magnetism}\ , \end{aligned} \tag{8.36}$$

In view of the meaning of divergence and curl explained above we can now interpret these laws. We know that the divergence measures the source (density) of a field and Gauß's law tells us that the source density of the electric field is identified (up to a constant) with the electric charge density. Similarly, Gauß's law for magnetism tells us that the source density of the magnetic field vanishes, that is, there are no magnetic charges. Further, we have seen that the curl measures the extent to which a vector field 'curls' around small local closed curves. The Ampère-Maxwell law asserts that the curl of the magnetic field is (up to a constant) given by the electric current. There is also an additional contribution to the curl of B in the Ampère-Maxwell law - the term proportional to $\dot{E}$. This is the term famously added by Maxwell. Faraday's law of induction states that the electric field E has a curl proportional to the change, $\dot{B}$, of the magnetic field.

8.4 Relations between grad, div and curl

What happens when we compose the action of grad, div and curl, that is, when we apply two of these operations, one after the other? Of course not all combinations make sense as the structure of the first

operation's output (that is, scalar or vector field) has to match the required input structure of the second operation.

Arbitrary dimension

We start by discussing this in arbitrary dimensions (before we specialise to $n = 2$ and $n = 3$) and consider a (twice differentiable) scalar $f : U \to \mathbb{R}$ in n dimensions. We already know from Lemma 8.1 that the curl of the vector field ∇f must vanish, that is,

$$\mathrm{Curl}(\nabla f) = 0 \,. \tag{8.37}$$

Another interesting combination is div $\circ$ grad which is more explicitly given by

$$\nabla \cdot (\nabla f) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} =: \Delta f \quad \text{with} \quad \Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \,. \tag{8.38}$$

The operator Δ defined above is called the **Laplace operator** or Laplacian and appears in many important partial differential equations in physics. This includes the Laplace equation, $\Delta f = 0$, and the Poisson equation, $\Delta f = \rho$, where ρ is a given function.

Example 8.8. Formula for a double curl

Question: For a three-dimensional vector field A, derive a formula for $\nabla \times (\nabla \times A)$.

Answer: An index calculation shows that

$$\begin{aligned}[\nabla \times (\nabla \times A)]]_i &= \epsilon_{ijk}\partial_j(\nabla \times A)_k = \epsilon_{ijk}\partial_i(\epsilon_{klm}\partial_l A_m) = (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\partial_j\partial_l A_m \\ &= \partial_i(\nabla \cdot A) - \Delta A_i \,,\end{aligned}$$

or, in vector notation,

$$\nabla \times (\nabla \times A) = \nabla(\nabla \cdot A) - \Delta A \,. \tag{8.39}$$

Example 8.9. Fundamental solution of Laplace equation

Question: Consider the scalar function $\Phi : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ given by $\Phi(x) = \frac{1}{r(x)^p}$, where $r(x) = |x|$ is the radius and p is a number. Compute $\Delta\Phi$ and find the value of p for which Φ satisfies the Laplace equation on $\mathbb{R}^n \setminus \{0\}$. Discuss.

Answer: From Exercise 8.2 we know that $\nabla\Phi = -\frac{px}{r^{p+2}}$ and from Exercise 8.4 we know the divergence of a vector field $E = \frac{x}{r^q}$ is given by $\nabla \cdot E = \frac{n-q}{r^q}$. Putting these results together we find

$$\Delta\Phi = \nabla \cdot (\nabla\Phi) = -p\,\nabla \cdot \left(\frac{x}{r^{p+2}}\right) = -p\frac{n-p-2}{r^{p+2}}$$

Hence, the Laplace equation is satisfied for $p = n-2$ and $\phi(x) = \frac{1}{r^{n-2}}$ is a solution of the n-dimensional Laplace equation on $\mathbb{R}^n \setminus \{0\}$. In three dimensions, $n = 3$, this is the famous Newton or Coulomb potential $\Phi(x) = \frac{1}{r}$. But the result also tells us what Newton's or Coulomb's law would look like if gravity or electro-magnetism were forces acting in more than three spatial dimensions. For example, for four spatial dimensions, $n = 4$, it is $\Phi(x) = \frac{1}{r^2}$. This dimension-dependent behaviour can be used to test the existence of additional spatial dimensions experimentally.

Example 8.10. Solution to the wave equation
Question: Show that the function $f : \mathbb{R} \times (\mathbb{R}^3 \setminus \{0\}) \to \mathbb{R}$ given by $f(t,x) = \frac{cos(r-t)}{r}$, where $r = |x|$ is the radial function and $x = (x_1, x_2, x_3)^T$, solves the wave equation

$$\left(\Delta - \frac{\partial^2}{\partial t^2}\right) f = 0 \,, \qquad \Delta = \sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2} \,. \tag{8.40}$$

Answer: It is easy to work out

$$\frac{\partial^2 f}{\partial t^2} = \frac{-\cos(r-t)}{r} \,. \tag{8.41}$$

The x_i derivatives are a bit harder. Using $\partial_i r = \frac{x_i}{r}$ we find

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial r}\partial_i r = \frac{\partial f}{\partial r}\frac{x_i}{r} \,, \qquad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{x_i x_j}{r^2}\frac{\partial^2 f}{\partial r^2} + \frac{r^2\delta_{ij} - x_i x_j}{r^3}\frac{\partial f}{\partial r}$$

Multiplying the last result with δ_{ij} gives the Laplacian

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} = -\frac{\cos(r-t)}{r} \,.$$

Subtracting the t-derivates (8.41) from this gives indeed zero.

Two dimensions
Focusing on two dimensions, $n = 2$, the general statement (8.37) holds and, for a (twice differentiable) function $f : U \to \mathbb{R}$ can be explicitly verified by

$$\text{curl}(\text{grad}\, f) \overset{\text{Eq.(8.28)}}{=} \epsilon_{ij}\partial_i\partial_j f = 0 \,. \tag{8.42}$$

The expression vanishes since ϵ_{ij} is anti-symmetric and $\partial_i\partial_j f$ is symmetric in the index pair (ij), the latter because of Theorem 4.1.

The meaning of Eq. (8.42) can be elucidated further by formalising matters a bit. Introduce the function spaces $\mathcal{C}^\infty(U)$ and $\mathcal{V}^\infty(U)$ of infinity many times differentiable two-dimensional scalar and vector fields on U. Then, grad and curl map between those spaces as in the following diagram

$$\mathcal{C}^\infty(U) \overset{\text{grad}}{\longrightarrow} \mathcal{V}^\infty(U) \overset{\text{curl}}{\longrightarrow} \mathcal{C}^\infty(U) \,, \tag{8.43}$$

and carrying the adjacent maps out one after the other leads to zero, curl $\circ$ grad $= 0$. A structure of this kind, that is, a sequence of vector spaces, connected by linear maps, such that adjacent maps compose to zero is a common and important structure in mathematics which is called a **complex**. More generally, a complex is a sequence

$$\cdots \overset{d_{i-1}}{\longrightarrow} V_i \overset{d_i}{\longrightarrow} V_{i+1} \overset{d_{i+1}}{\longrightarrow} \cdots \tag{8.44}$$

of vector space V_i, connected by linear maps d_i, such that $d_i \circ d_{i-1} = 0$ for all i. An immediate consequence of $d_i \circ d_{i-1} = 0$ is that $\text{Im}(d_{i-1}) \subset \text{Ker}(d_i)$ so that it makes sense to introduce the vector space quotients $H^i(U) = \text{Ker}(d_i)/\text{Im}(d_{i-1})$ which are called the **cohomology** of the complex with dimensions $h^i(U) = \dim(H^i(U))$.

What about the cohomology of our grad/curl complex (8.43). It turns out, this encodes information about the topology of the underlying space U. We will not discuss this further (although there will be some hints later on) as this would go well beyond the scope of this lecture. We note that the complex (8.43)

and its cohomology is one of the gateways into an important area of pure mathematics called **algebraic geometry** which is a major arena of contemporary research in mathematics. However, there is one connection with an earlier discussion we should make. We know that the curl of a conservative vector field (in two or any dimension) vanishes but it is not clear yet whether a vanishing curl is sufficient for a vector field to be conservative. The cohomology $H^1(U) = \text{Ker(curl)}/\text{Im(grad)}$ of the complex (8.43) measures the possible obstruction to a vector field with vanishing curl being conservative. If the cohomology dimension $h^1(U) = 0$ then it follows that $\text{Im(grad)} = \text{Ker(curl)}$ and this means that every vector field with vanishing curl can be written as a gradient of a scalar and is, hence, conservative. On the other hand, if $h^1(U) > 0$ then Ker(curl) is larger than Im(grad) which means there are vector fields with vanishing curl which are not conservative. Evidently, knowing the cohomology dimension $h^1(U)$ is quite important!

Example 8.11. Fundamental solution to Laplace equation in two dimensions
Question: In Exercise 8.9 we have found that $\Phi(x) = \frac{1}{r^{n-2}}$ satisfies the Laplace equation in n dimensions. For $n = 2$ this only gives the trivial solution, $\Phi(x) = 1$. Show that $\Phi(x) = \ln(r(x))$ satisfies the two-dimensional Laplace equation.

Answer: First note that $\partial_i \Phi = \frac{1}{r}\partial_i r = \frac{x_i}{r^2}$, where $\partial r = \frac{x_i}{r}$ has been used. For the Laplacian this leads to

$$\Delta\Phi = \partial_i\partial_i\Phi = \partial_i\left(\frac{x_i}{r^2}\right) = \frac{\partial_i x_i}{r^2} - \frac{2x_i}{r^3}\partial_i r = \frac{2}{r^2} - \frac{2\,x_i x_i}{r^4} = 0\,.$$

Three dimensions
Les us now discuss the three-dimensional version of the story. Of course Eq. (8.37) remains true in three dimensions, that is, $\text{curl} \circ \text{grad} = 0$. For a (twice differentiable) scalar function $f : U \to \mathbb{R}$ in three dimensions, this is verified explicitly by

$$(\nabla \times (\nabla f))_i = \epsilon_{ijk}\partial_j\partial_k f = 0\,. \tag{8.45}$$

The last equality follows since ϵ_{ijk} is anti-symmetric in (jk) while $\partial_j\partial_k f$ is symmetric in (jk), as a consequence of Theorem 4.1. As we have already mentioned the result (8.45) implies that a vanishing curl is a necessary condition for a vector field to be conservative.

In three dimensions, the curl of a vector field A is another vector field, so we can consider $\text{div} \circ \text{curl}$. Explicitly, this gives

$$\nabla \cdot (\nabla \times A) = \partial_i(\nabla \times A)_i = \partial_i\epsilon_{ijk}\partial_j A_k = 0\,, \tag{8.46}$$

and, as before, the vanishing follows from the anti-symmetry of ϵ_{ijk} and the symmetry of $\partial_i\partial_j A_k$ in the index pair (ij). In analogy with what we have done for grad, we can ask whether a given three-dimensional vector field B can be written as the curl of another vector field A, so that $B = \nabla \times A$. In this case A is called a **vector potential** for B. Eq. (8.46) shows that a necessary condition for B to have a vector potential is that the divergence of B vanishes, that is, $\nabla \cdot B = 0$.

The relationship between grad, curl and div in three dimensions can be summarised by writing down the sequence

$$\mathcal{C}^\infty(U) \xrightarrow{\text{grad}=\nabla} \mathcal{V}^\infty(U) \xrightarrow{\text{curl}=\nabla\times} \mathcal{V}^\infty(U) \xrightarrow{\text{div}=\nabla\cdot} \mathcal{C}^\infty(U) \tag{8.47}$$

which is, of course, the three-dimensional version of Eq. (8.43). Eqs. (8.45) and (8.46) imply that adjacent maps in (8.47) compose to zero so the sequence (8.47) is a complex, just as in the two-dimensional case. There are two cohomologies, $H^1(U) = \text{Ker(curl)}/\text{Im(grad)}$ and $H^2(U) = \text{Ker(div)}/\text{Im(curl)}$. The first one, $H^1(U)$, measures the obstruction to vector fields with vanishing curl being conservative, the second one, $H^2(U)$, the obstruction to a divergence-free vector field having a vector potential. In other words, if $h^1(U) = 0$ then every vector field with vanishing curl is conservative and if $h^2(U) = 0$ then every

divergence-free vector field has a vector potential.

Not surprisingly, the complexes (8.43) and (8.47) in two and three dimensions can be generalised to arbitrary dimensions, leading to what is referred to as **de Rham complex**. However, formulating this efficiently requires the language of differential forms, which is beyond our scope. If you are interested have a look at the appendix.

8.5 Poincaré's Lemma

We have seen that for a vector field to be conservative it is necessary that its curl vanishes. On the other hand, it is not clear that this condition is sufficient. Abstractly, we know that the obstruction to vector fields with vanishing curl being conservative is encoded in the complexes (8.43) and (8.47) and their associated cohomologies. If $h^1(U) = 0$ then all curl-free vector fields on the domain U are indeed conservative. Could it be that this is always the case or are there situations where $h^1(U) > 0$?

Let us consider a simple example of the two-dimensional vector field

$$A(x) = \begin{pmatrix} A_1(x) \\ A_2(x) \end{pmatrix} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} . \tag{8.48}$$

Clearly, A is not well-defined at the origin so we should think about this as a vector field $A : U \to \mathbb{R}^2$ whose domain is the punctured plane $U = \mathbb{R}^2 \setminus \{(0,0)\}$. It is easy to verify that this vector field satisfies

$$\frac{\partial A_1}{\partial x_2} = \frac{\partial A_2}{\partial x_1} , \tag{8.49}$$

so that $\mathrm{curl}(A) = 0$. On the other hand, if we work out a curve integral of this vector field around a unit circle S^1, parametrised by $X(t) = (\cos(t), \sin(t))$ with $t \in [0, 2\pi]$, we have

$$A(X(t)) \cdot \frac{dX}{dt}(t) = 1 \qquad \Rightarrow \qquad \int_{S^1} A \cdot dx = 2\pi . \tag{8.50}$$

We have seen in Eq. (8.15) that the line integral of a conservative vector field over a closed curve must be zero. Hence the result (8.50) shows that our example vector field (8.48) is not conservative, despite its curl vanishing. In other words, curl-free but non-conservative vector fields exist on certain domains, such as the punctured plane U of the present example, and for these cases $h^1(U) > 0$.

However, it turns out that, under certain additional conditions on the domain U, we have $h^1(U) = 0$ so that all curl-free vector fields are conservative on such domains. This is the content of Poincaré's Lemma which we now state for the case of vector fields. (The general version of Poincaré's Lemma is formulated for differential forms – see the appendix.)

Lemma 8.2. *(Poincaré) Let $A : U \to \mathbb{R}^n$ be a continuously differentiable vector field with a vanishing curl and a star-shaped domain U (meaning that there is a point $p \in U$ such that the line segments that connects p with any $x \in U$ are contained in U). Then A is conservative.*

Proof. We can assume that the point p in the star-shaped condition is the origin (as we can perform a translation by $-p$). Then, consider the curve $X : [0,1] \to \mathbb{R}^n$ defined by $X(t) = tx$ (with tangent vector $dX/dt(t) = x$) which is the line segment that connects the origin with a given point $x \in U$. Note, due to the star-shaped condition this curve is entirely contained in U. Then we define a prospective potential function V by

$$V(x) = \int_{X([0,1])} A \cdot dx = \int_0^1 dt\, A(tx) \cdot x . \tag{8.51}$$

We need to check that V is indeed a potential function for A, that is, that $A = \nabla V$. This works by direct calculation.

$$\frac{\partial V}{\partial x_i}(x) = \int_0^1 dt \left(t \frac{\partial A}{\partial x_i}(tx) \cdot x + A_i(tx) \right) \tag{8.52}$$

On the other hand, from the chain rule, we have

$$\frac{d}{dt}(tA_i(tx)) = t \frac{\partial A_i}{\partial x_j}(tx) x_j + A_i(tx) \overset{\text{Eq. (8.13)}}{=} t \frac{\partial A}{\partial x_i}(tx) \cdot x + A_i(tx) \tag{8.53}$$

where the condition of vanishing curl has been used in the last step. Hence, the integral (8.52) can be written as

$$\frac{\partial V}{\partial x_i}(x) = \int_0^1 dt \, \frac{d}{dt}(tA_i(tx)) = [tA_i(tx)]_0^1 = A_i(x) \ , \tag{8.54}$$

and this is what we needed to show. □

Remark. Generalisations of Poincarés Lemma

(i) Poincaré's Lemma can be generalised to differential forms (see Appendix A).

(ii) The condition on the domain U can be weakened. It is, in fact, enough for U to be 'simply-connected' which means that U contains no closed curves which cannot be continuously shrunk to a point in U.

(iii) We have seen that a necessary condition for a vector field $B : U \to \mathbb{R}^3$ to have a vector potential, that is, a vector field $A : U \to \mathbb{R}^3$ with $B = \nabla \times A$, is that its divergence vanishes, so $\nabla \cdot B = 0$. The general version of Poincaré's Lemma tells us for a star-shaped (or simply-connected) U this is also sufficient.

Full Euklidean space, $\mathbb{R}^n$, is certainly star-shaped so that Poincaré's Lemma applies to this case and it follows that all curl-free vector fields with domain $\mathbb{R}^n$ are indeed conservative. On the other hand, the domain for the above example (8.48) was the puntered plane $U = \mathbb{R}^2 \setminus \{(0,0)\}$ and this is not star-shaped. (For any possible point $p \in U$, the line segment connecting p with $-p$ goes through the origin which is not in U.) Hence, Poincaré's Lemma does not apply to this domain and the example shows that curl-free but non-conservative vector fields do indeed exist in this case, that is, $h^1(U) > 0$.

Note, once we know the vector field $A : U \to \mathbb{R}^n$ is conservative we can define the potential V by generalising Eq. (8.51) to

$$V(x) = \int_{C_x} A \cdot dx \ . \tag{8.55}$$

Here, C_x is any differentiable curve fully contained in the domain U which connects the reference point (in the proof of Poincaré's Lemma taken to be the origin) with the given point x, rather than the straight line used in Eq. (8.51). This is because we know that the line integral for conservative vector fields only depends on the endpoints and not the path taken and, as a consequence, the potentials in Eqs. (8.51) and (8.55) are the same.

Example 8.12. Conservative vector fields

Question: The vector field $A : \mathbb{R}^3 \to \mathbb{R}^3$ is given by $A(x, y, z) = (x^2 + yz, y^2 + xz, z^2 + xy)^T$. Is this vector field conservative? If so find a potential function.

Answer: It is easy to check by straightforward calculation that $\nabla \times A = 0$ so the necessary condition is satisfied. The domain $\mathbb{R}^3$ on which A is defined is star-shaped so that the conditions for Poincaré's Lemma, Lemma 8.2, are satisfied. We conclude that A is conservative.

The potential function for A can be computed from Eq. (8.51). This gives

$$V(x,y,z) = \int_0^1 dt\, A(tx,ty,tz) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x^3+y^3+z^3+3xyz)\int_0^1 dt\, t^2 = \frac{1}{3}(x^3+y^3+z^3)+xyz\ .$$

Application 8.4. Line integrals and potential functions in physics
The formula (8.55) to define the potential function in terms of the associated conservative vector is used in many places in physics. For example, for a conservative force F is classical mechanics, the associated potential is given by

$$V(x) = \int_{C_x} F \cdot dx\ . \tag{8.56}$$

In the context of electrostatics, the electric field E has a vanishing curl so on a suitable domain which satisfied the condition in Poincaré's Lemma it is conservative. In this case, the electrostatic potential ϕ can be defined by

$$\phi(x) = \int_{C_x} E \cdot dx\ . \tag{8.57}$$

Application 8.5. Electrostatics and magnetostatics
Electrostatics is a special case of electromagnetism where the electric and magnetic fields are time-independent. In this case the Maxwell Eqs. (8.36) for the electric field reduce to Gauß's law, $\nabla \cdot E = 4\pi\rho$ and Faraday's law becomes $\nabla \times E = 0$. Hence, in this limit the curl of the electric field vanishes and if we assume that the underlying domain of three-dimensional space satisfies the additional conditions in Poincaré's Lemma then E is conservative and there exists an electro-static potential ϕ such that $E = -\nabla\phi$. Inserting this into Gauß's law leads to the Poisson equation

$$\Delta\phi = -4\pi\rho \tag{8.58}$$

for the potential. (While, for $E = -\nabla\phi$, the other equation, $\nabla \times E = 0$, is automatically satisfied since curl $\circ$ grad $= 0$.) In short, electrostatics reduces to the single Poisson eq. (8.58) for the electrostatic potential ϕ.

The case of magnetostatics is a bit more complicated. For time-independent electric and magnetic fields the Ampère-Maxwell law and Gauß's law for magnetism in Eq. (8.36) turn into the basic equations of magnetostatics

$$\nabla \times B = \frac{4\pi}{c} J\ , \qquad \nabla \cdot B = 0\ . \tag{8.59}$$

If the electric current density vanishes, $J = 0$, we have $\nabla \times B = 0$ and again assuming the conditions on the domain in Poincaré's Lemma are satisfied the magnetic field is conservative and we have a magnetic potential ϕ_M with $B = -\nabla\phi_M$. Inserting this into Gauß's law for magnetism shows that in this case the magnetic potential satisfies the Laplace equation

$$\Delta\phi_M = 0\ . \tag{8.60}$$

If $J \neq 0$ then B is not conservative and we cannot find a magnetic potential. However, we still have $\nabla \cdot B = 0$ and this tells us (from a generalisation of Poincaré's Lemma provided the conditions on the domain are satisfied) that there exists a vector potential A such that $B = \nabla \times A$. Inserting this into

the first Eq. (8.59) and using Eq. (8.39) gives

$$\nabla(\nabla \cdot A) - \Delta A = \frac{4\pi}{c} J \,, \tag{8.61}$$

while the second Eq. (8.59) is automatically satisfied for $B = \nabla \times A$ since div $\circ$ curl $= 0$. Magnetostatics is governed by the single Eq. (8.61) for the vector potential A but this equation does not look particularly simple. An important observation is that a so-called **gauge transformation**

$$A \mapsto A' = A + \nabla\lambda \,, \tag{8.62}$$

(where λ is a scalar function) on the vector potential A does not change the associated magnetic field. Indeed

$$B' = \nabla \times A' = \nabla \times (A + \nabla\lambda) = \nabla \times A + \nabla \times (\nabla\lambda) = B \,.$$

This means different vector potentials, related by a gauge transformation, lead to the same physical field B so the vector potential contains unphysical degrees of freedom. We can use gauge transformations to transform to an equivalent vector potential (leading to the same B) which satisfies an additional **gauge condition**. A common choice, also referred to as **Coulomb gauge**, is to demand that $\nabla \cdot A = 0$. In Coulomb gauge, Eq. (8.61) turns into

$$\Delta A = -\frac{4\pi}{c} J \,, \tag{8.63}$$

that is, another Poisson equation, although for a vector field.

Part III
Integration in $\mathbb{R}^n$

There are many types of integrals! Integrals over a set S can be set up by selecting 'measurable' subsets M of S and constructing a 'measure' or volume function μ for these subsets such that $\mu(M) \in \mathbb{R}$ gives the volume. Once this has been accomplished the rough idea for how to define integrals of functions $f : S \to \mathbb{R}$ is as follows. Pick measurable sets M_i which 'tile' S and pick a point $x_i \in M_i$ in each set. An approximation to the prospective integral is then given by the sum $\sum_i f(x_i)\mu(M_i)$. The integral itself is defined by a limit of such sums where the size of the 'patches' M_i goes to zero. The type of integral obtained in this way depends the choice of measurable subsets M and the definition of the measure μ. Setting all this up in general is the subject of **measure theory** but a detailed discussion of this is not actually required for our purposes nor is it within our scope.

Here we are interested in integrals over $\mathbb{R}^n$ (or certain subsets thereof) and there are two important types of integrals which are considered in this context: the **Riemann integral** and the **Lebesgue integral**. You are already familiar with the Riemann integral in one dimension (these are the integrals you have used so far) but this can be generalised to $\mathbb{R}^n$. The Lebesgue integral is a generalisation of the Riemann integral. All Riemann integrable functions are Lebesgue integrable and for such functions the two integrals coincide, but the class of Lebesgue integrable functions is larger than the one of Riemann integrable ones. As a result, the Lebesgue integral is mathematically 'better behaved'. On the downside, developing the Lebesgue integral and, in particular, the underlying Lebesgue measure μ, requires considerable mathematical sophistication and time. We may have the former but sadly not the latter, so we will focus on developing the Riemann integral on $\mathbb{R}^n$ which is sufficient for most of our purposes. This is still not a straightforward task (although considerably easier than tackling the Lebesgue integral) and we will compromise in some places by only sketching proofs or referring to the literature.

9 Definition of Riemann integral in $\mathbb{R}^n$

9.1 The Riemann integral in one dimension - a reminder

The definition of the Riemann integral is illustrated in Fig. 13. The integration interval $[a, b] \subset \mathbb{R}$ is partitioned as $a = x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_n = b$ and the area under the function is approximated in two ways: by the area under a lower and an upper step function associated with this partition. The actual integral is then obtained by a limit in which the width of the partition (the magnitute of $x_{i+1} - x_i$) goes to zero. If the limits obtained in this way from both the lower and the upper step functions exist and if they agree then f is called Riemann integrable and the value of the Riemann integral is given by these identical limits. Our task is now to generalise this simple idea to n dimensions.

9.2 Riemann integral in multiple dimensions

Integration regions for the one-dimensional Riemann integral are typically intervals and their boundaries are rather simple - they consist of the two endpoints of the interval. There is significantly more structure in higher dimensions. Sets we might want to integrate over might be much more complicated and have more complicated boundaries in addition, of course, to functions now depending on multiple variables.

Still, we can start generalising the one-dimensional story by introducing the generalisation of intervals, namely **n-dimensional boxes** $B \subset \mathbb{R}^n$ (parallel to the coordinate axes) given by a Cartesian product

$$B = \{x \in \mathbb{R}^n \,|\, a_i \leq x_i \leq b_i\} = [a_1, b_1] \times \cdots \times [a_n, b_n] \tag{9.1}$$

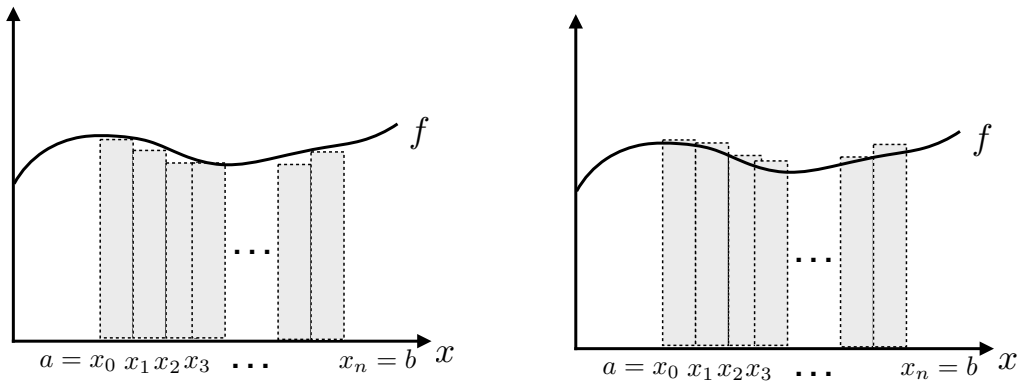


Figure 13: Approximation of a function f by lower and upper piecewise constant functions to define the Riemann integral in one dimension. The lower (upper) integral is defined as the total area of the rectangles on the left (right), in the limit of the rectangle width going to zero. The function is called Riemann integrable if the so-obtained upper and lower values exist and are the same.

of n intervals $[a_i, b_i]$, where $a_i \leq b_i$. As a measure μ for such boxes we use the obvious volume

$$\mu(B) = \prod_{i=1}^{n} (b_i - a_i) \ . \tag{9.2}$$

A **partition** of an n-dimensional box B is a collection $\mathcal{B} = \{B_a \,|\, a \in A\}$ such that the B_a are n-dimensional boxes and

$$B = \bigcup_{a \in A} B_a \ , \qquad B_a \cap B_b = \{\} \ \text{ or } \ \mu(B_a \cap B_b) = 0 \ \text{ for } \ a \neq b \ . \tag{9.3}$$

These statements are just a formal way of expressing what is shown, for a two-dimensional example, in

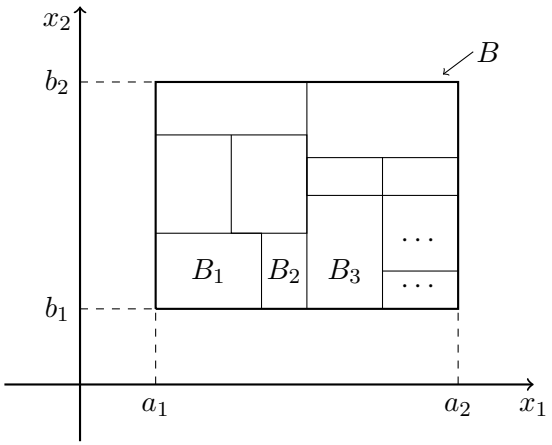


Figure 14: An example of a partition $\mathcal{B} = \{B_1, B_2, B_3, \ldots\}$ of the two-dimensional rectangle $B = [a_1, b_1] \times [a_2, b_2]$.

Fig. 14. The given rectangle B is partioned into disjoint (up to boundary overlaps with volume zero)

rectangles B_a, where $a \in A$. It is clear that the volume of B is given by summing the volumes of the partitioning rectangles, that is,

$$\mu(B) = \sum_{a\in A} \mu(B_a) \; . \tag{9.4}$$

Now suppose we have a (bounded) function $f : B \to \mathbb{R}$ on our box B with partition $\mathcal{B} = \{B_a \,|\, a \in A\}$.Then, we can define the **lower and upper sum** over the partition $\mathcal{B}$ of B by

$$\underline{\mathrm{I}}(f, \mathcal{B}) = \sum_{a\in A} \inf\{f(x) \,|\, x \in B_a\}\, \mu(B_a) \; , \qquad \bar{I}(f, \mathcal{B}) = \sum_{a\in A} \sup\{f(x) \,|\, x \in B_a\}\, \mu(B_a) \; . \tag{9.5}$$

The **infimum**, denoted by inf, of a subset of $\mathbb{R}$ is the largest lower bound of the subset, while the **supremum**, denoted by sup, is the smaller upper bound of the subset [4]. These lower and upper sums are simply the multi-dimensional generalisations of the sums over the grey rectangles in Fig. 13 which are used to define the one-dimensional Riemann integral. We are now in a position to define the Riemann integral of f over the box B.

Definition 9.1. *(Riemann integral on boxes)*
Let $B \subset \mathbb{R}^n$ be an n-dimensional box and $f : B \to \mathbb{R}$ a bounded function. The **lower and upper Riemann integrals** *of f over B are defined as*

$$\underline{\int_B} d^n x\, f(x) = \sup\{\underline{I}(f, \mathcal{B}) \,|\, \mathcal{B} \text{ partition of } B\} \; , \quad \overline{\int_B} d^n x\, f(x) = \inf\{\bar{I}(f, \mathcal{B}) \,|\, \mathcal{B} \text{ partition of } B\} \; . \tag{9.6}$$

If the lower and upper integral are equal then f is called **Riemann integrable** *over B and the Riemann integral is defined as*

$$\int_B d^n x\, f(x) = \underline{\int_B} d^n x\, f(x) = \overline{\int_B} d^n x\, f(x) \; . \tag{9.7}$$

Note, the logic is precisely the same as for the one-dimensional Riemann integral.

In order to define the Riemann integral on more general sets than boxes we need a few more simple twists. First, for a function $f : \mathbb{R}^n \to \mathbb{R}$ we define the **support** of f by [5]

$$\mathrm{supp}(f) = \overline{\{x \in \mathbb{R}^n \,|\, f(x) \neq 0\}} \; . \tag{9.8}$$

In other words, the support of f is (the closure of) the set on which f is non-vanishing. We say that f has **compact support** if $\mathrm{supp}(f)$ is compact, that is, if it is bounded in addition to being closed. Further, we define the **characteristic function** $\chi_S : \mathbb{R}^n \to \mathbb{R}$ of a subset $S \subset \mathbb{R}^n$ by

$$\chi_S(x) = \begin{cases} 1 & \text{for} \quad x \in S \\ 0 & \text{for} \quad x \notin S \end{cases} \; . \tag{9.9}$$

We are now ready to define the Riemann integral over more general sets.

Definition 9.2. *(Riemann integral)*
For a function $f : \mathbb{R}^n \to \mathbb{R}$ with compact support choose a box B such that $\mathrm{supp}(f) \subset B$. *If f is Riemann*

[4] You can think of the infimum and the supremum as more sophisticated versions of the minimum and the maximum. In favourable cases they are the same, for example the interval $[0, 1] \subset \mathbb{R}$ has $\min[0, 1] = \inf[0, 1] = 0$ and $\max[0, 1] = \sup[0, 1] = 1$. On the other hand, the open interval $(0, 1)$ does not have a minimum or maximum but the infinum and supremum exist and are given by $\inf(0, 1) = 0$ and $\sup(0, 1) = 1$.

[5] Do not confuse this with the supremum for which the symbol sup (single 'p') is used.

integrable over B then it is called a **Riemann integrable function with compact support** *and the Riemann integral is given by*

$$\int_{\mathbb{R}^n} d^n x\, f(x) = \int_B d^n x\, f(x)\ . \tag{9.10}$$

Further, let $S \subset \mathbb{R}^n$ be a bounded set and f be bounded on S. Then, f is called **Riemann integrable over** *S if $f\chi_S$ is a Riemann integrable function with compact support and, in this case, the integral of f over S is defined by*

$$\int_S d^n x\, f(x) = \int_{\mathbb{R}^n} d^n x\, f(x)\chi_S(x)\ . \tag{9.11}$$

We can use the last part of the definition to define the **volume** of subsets $S \subset \mathbb{R}^n$ for which χ_S is Riemann integrable. In this case, we say that S is **Jordan measurable** and define the volume of S (also called the **Jordan measure** of S) by

$$\mathrm{vol}(S) = \int_{\mathbb{R}^n} d^n x\, \chi_S(x) = \int_S d^n x\ . \tag{9.12}$$

Now, this was not so bad so why did we start this section with a somewhat scary speech? Unfortunately, the hard work is still to come: we need to derive the properties of the integral we have just defined.

10 Properties of the Riemann integral

10.1 Basic properties

The Riemann integral has the following basic properties.

Theorem 10.1. *(Properties of Riemann integral)*
(i) The Riemann integral is **linear**, *that is, if $f, g : \mathbb{R}^n \to \mathbb{R}$ are Riemann integrable, then so is $\alpha f + \beta g$, where $\alpha, \beta \in \mathbb{R}$ and we have*

$$\int_{\mathbb{R}^n} d^n x\, (\alpha f(x) + \beta g(x)) = \alpha \int_{\mathbb{R}^n} d^n x\, f(x) + \beta \int_{\mathbb{R}^n} d^n x\, g(x)\ . \tag{10.1}$$

(ii) The Riemann integral is **monotonic**, *that is, for two Riemann integrable functions f, g with $f \leq g$ (meaning $f(x) \leq g(x)$ for all $x \in \mathbb{R}^n$) we have*

$$\int_{\mathbb{R}^n} d^n x\, f(x) \leq \int_{\mathbb{R}^n} d^n x\, g(x)\ . \tag{10.2}$$

(iii) The Riemann integral is **translation invariant** *that is, for $a \in \mathbb{R}^n$ it satisfies*

$$\int_{\mathbb{R}^n} d^n x f(x-a) = \int_{\mathbb{R}^n} d^n x f(x)\ . \tag{10.3}$$

Proof. We will only sketch the proof and refer to Ref. [10] for a full account. Properties (i) and (ii) are obviously satisfied for the lower and upper sums (9.5) which were used to define the integral and all partitions $\mathcal{B}$. It is, therefore, plausible that they will survive in the limit taken in Eq. (9.6).

Property (iii) can be shown by using the transformation formula which we will discuss below. □

The properties in Theorem 10.1 determine the integral uniquely (up to an overall constant) and they can be used to define the integral axiomatically. It is useful to know that continuous functions are well-behaved with regard to the Riemann integral.

Proposition 10.1. *A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ with compact support is Riemann integrable. Also, a continuous function $F : S \to \mathbb{R}$ with a Jordan-measurable domain S is Riemann integrable.*

Proof. For a proof, see, for example, Ref. [10]. □

10.2 Successive integration

How do we actually evaluate multi-dimensional Riemann integrals for given functions? In principle, we might consider carrying out lower and upper sums (9.5) explicitly and then taking the limit of 'partition size' going to zero. However, this approach seems tedious and not really practical. Instead, we will show that the Riemann integral in n dimensions can be written as a succession of n one-dimensional Riemann integrals, each of which can then be evaluated using the fundamental theorem of (one-dimensional) integration.

Unfortunately there is a technical complication. Consider a Riemann integrable function $f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$ and we denote coordinates by $x = (y, z)$, where $y \in \mathbb{R}^p$ and $z \in \mathbb{R}^q$. It turns out it is not guaranteed that, for fixed y, the function $z \mapsto f(y, z)$ is Riemann integrable. This means, while we can integrate the function over all $n = p + q$ dimensions, there might be a problem if we try to split this integral up into, first, a p-dimensional and then a q-dimensional integral. After the first step, that is, the integration over the variable y, the remaining function in z might not be Riemann integrable. This explains why the following theorem is somewhat weaker than expected.

Theorem 10.2. *Let $f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$ be a Riemann integrable function (with compact support). Then*

$$y \mapsto \underline{\int_{\mathbb{R}^q}} d^q z\, f(y, z) \quad \text{and} \quad z \mapsto \overline{\int_{\mathbb{R}^q}} d^n z\, f(y, z) \tag{10.4}$$

are Riemann integrable (with compact support) and we have

$$\int_{\mathbb{R}^{p+q}} d^{p+q} x\, f(x) = \int_{\mathbb{R}^q} d^p y \underline{\int_{\mathbb{R}^q}} d^q z\, f(y, z) = \int_{\mathbb{R}^q} d^p y \overline{\int_{\mathbb{R}^q}} d^q z\, f(y, z)\ . \tag{10.5}$$

Proof. The proof can be found in Ref. [10]. □

Of course the variables y and z can be exchanged in the previous theorem and in Eq. (10.5) which shows that the 'order of integration' can be exchanged. For continuous functions everything is much better.

Corollary 10.1. *For a continuous function $f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$ with compact support we have*

$$\int_{\mathbb{R}^{p+q}} d^{p+q} x\, f(y, z) = \int_{\mathbb{R}^p} d^p y \int_{\mathbb{R}^q} d^q z\, f(y, z) = \int_{\mathbb{R}^q} d^q z \int_{\mathbb{R}^p} d^p y\, f(y, z) \tag{10.6}$$

Proof. The follows from Theorem 10.2, the fact that integrating over some variables of a continuous function produces a continuous function in the remaining variables and Prop. 10.1 □

It is useful to iterate this statement and, in this way, split the integral up into successive one-dimensional integrals. To this end, let us introduce a class of integration sets which are relevant in this context. Intuitively, these consist of products of intervals but the boundaries of a given interval can depend on the preceding coordinates. More explicitly, we define the set $S = S_n$ recursively by setting $S_1 = [a_1, b_1]$ (with $a_1 \leq b_1$) and then

$$S_p = \{(x_1, \ldots, x_p)^T \in \mathbb{R}^p \,|\, (x_1, \ldots, x_{p-1})^T \in S_{p-1} \text{ and } a_p(x_1, \ldots, x_{p-1}) \leq x_p \leq b_p(x_1, \ldots, x_{p-1})\}\ , \tag{10.7}$$

where a_p and b_p for $p = 2, \ldots, n$ are continuous functions of their respective arguments with $a_p \leq b_p$ which describe the boundary of the set S. For sets parametrised in this way we can split up the n-dimensional integral into successive one-dimensional integral as follows.

Theorem 10.3. *Let S be a set as defined in Eq. (10.7) and $f : S \to \mathbb{R}$ a continuous function. Then S is Jordan measurable and*

$$\int_S d^n x\, f(x) = \int_{a_1}^{b_1} dx_1 \int_{a_2(x_1)}^{b_2(x_1)} dx_2 \cdots \int_{a_n(x_1,\ldots,x_{n-1})}^{b_n(x_1,\ldots,x_{n-1})} dx_n\, f(x_1,\ldots,x_n)\,. \tag{10.8}$$

Proof. The full proof can be found in Ref. [10] but the key is repeated application of Corollary 10.1. □

Of course there may be different ways of describing a given set S via a successive parametrisation as in Eq. (10.7), differing, in the choice of functions a_i, b_i and also the ordering of coordinates. All these different ways of writing S produce the same value for the integral, since the LHS of Eq. (10.8) is independent of any particular way of parametrising S. However, the practical calculation which has to be done by successively carrying out the one-dimensional integrals on the RHS of Eq. (10.8) can depend very much on which parametrisation of S is chosen. This feature is often helpful in carrying out integrals by 'changing the order of integration'. Also note that all $a_i \leq b_i$ in this construction so the one-dimensional integrals on the RHS of Eq. (10.8) are always carried out in the 'natural orientation', that is, by integrating from smaller to larger values.

Example 10.1. Splitting up two-dimensional integrals I
Question: Consider the set S shown in Fig. 15. Integrate the function f with $f(x,y) = \frac{1}{2}(x^2+y^2)$ over S by writing the set S in the form of Eq. (10.7). Do this in the two possible orderings of coordinates, that is, $(x_1,x_2) = (x,y)$ and $(x_1,x_2) = (y,x)$ and convince yourself that the value of the integral is the same for both choices.

Answer: The set S from Fig. 15 can be written in the form (10.7) in two ways, namely

$$S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, x \in [0,1] \text{ and } 0 \leq y \leq x \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, y \in [0,1] \text{ and } y \leq x \leq 1 \right\}\,.$$

For the first possibility Eq. (10.8) reads

$$\int_S d^x\, f(x,y) = \frac{1}{2}\int_0^1 dx \int_0^x dx\,(x^2+y^2) = \frac{1}{2}\int_0^1 dx \left[x^2 y + \frac{1}{3}y^3\right]_0^x = \frac{2}{3}\int_0^1 dx\, x^3 = \frac{1}{6}\,.$$

while we have for the second choice

$$\int_S d^x\, f(x,y) = \frac{1}{2}\int_0^1 dy \int_y^1 dx\,(x^2+y^2) = \frac{1}{2}\int_0^1 dy \left[\frac{1}{3}x^3 + xy^2\right]_y^1 = \frac{1}{2}\int_0^1 dy \left(\frac{1}{3} + y^2 - \frac{4}{3}y^3\right) = \frac{1}{6}\,.$$

The two calculations lead to the same answer, as they should, but note the first one is slightly easier. The level of complexity of the calculation can very much depend on the ordering of the coordinates in the integral (10.8).

Example 10.2. Splitting up two-dimensional integrals II
Question: Integrate the function f with $f(x,y) = xe^{-y^2}$ over the set $S = \{(x,y)^T \,|\, x \in [0,1] \text{ and } x^2 \leq y \leq 1\}$.

Answer: If we take the split of S as suggested in the question the corresponding split of the integral reads

$$\int_S d^2x\, f(x,y) = \int_0^1 dx \int_{x^2}^1 dy\, xe^{-y^2}\,.$$

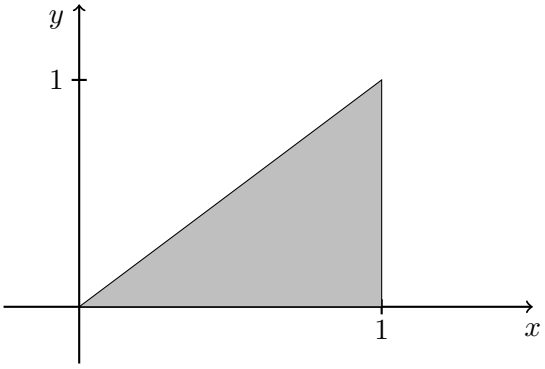


Figure 15: The set $S = \{(x,y)^T \in \mathbb{R}^2 \,|\, 0 \le x \le 1 \text{ and } 0 \le y \le x\}$.

Sadly, it is hard to integrate e^{-y^2} over a finite interval. On the other hand, we can also write the set S as

$$S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, y \in [0,1] \text{ and } 0 \le x \le \sqrt{y} \right\} ,$$

and in this case the integral splits up as

$$\int_S d^2x\, f(x,y) = \int_0^1 dy \int_0^{\sqrt{y}} dx\, x e^{-y^2} = \frac{1}{2}\int_0^2 dy \left[x^2 e^{-y^2}\right]_0^{\sqrt{y}} = \frac{1}{2}\int_0^1 dy\, y e^{-y^2} = \frac{1}{4}(1-e^{-1}) ,$$

and can be carried out easily.

10.3 Change of variables - the transformation formula

Recap of the formula for variable change in one-dimensional integrals
Let us start by recalling how a change of integration variable works for one-dimensional Riemann integrals. Consider an integration interval $[a,b]$ (with $a \le b$) and a diffeomorphism $X : Y([a,b]) \to [a,b]$ (a diffeomorphism is a map which is continuously differentiable, is invertible and the inverse is continuously differentiable as well) with inverse $Y = X^{-1}$. Then, for a Riemann integrable function $f : [a,b] \to \mathbb{R}$ we have [6]

$$\int_a^b dx\, f(x) = \int_{Y(a)}^{Y(b)} dy\, \frac{dX}{dy}(y)\, f(X(y)) \, . \tag{10.9}$$

This follows immediately from the fundamental theorem

$$\int_a^b dx\, f(x) = F(b) - F(a) \quad \text{where} \quad F' = f \, , \tag{10.10}$$

of integration, as both sides of Eq. (10.9) evaluate to $F(b) - F(a)$. Eq. (10.9) really covers two possible cases for how the change of variables unfolds. If $X' > 0$ then the interval $[a,b]$ (with $a \le b$) is mapped to the interval $[Y(a), Y(b)]$, where $Y(a) \le Y(b)$. This is the orientation-preserving case. On the other hand, if $X' < 0$ the orientation is reversed and the interval $[a,b]$ is mapped to $[Y(b), Y(a)]$, where $Y(b) \le Y(a)$. In this latter case, the integration on the RHS of Eq. (10.9) is the 'wrong way round', that is, it is going

[6] Note, we are denoting the coordinate by x and the corresponding function which gives the coordinate $x = X(y)$ by X. This is to draw a distinction between two different object but still have an intuitive notation. In a physics context, coordinates and associated functions are often denoted by the same letter, x.

from larger to smaller values. However, such an integral can be brought into the standard from by the simple rule $\int_c^d dy \cdots = -\int_d^c dy \cdots$. As a result, the formula (10.9) for the change of integration variable can also be written in the form

$$\int_{[a,b]} dx\, f(x) = \int_{Y([a,b])} dy \left|\frac{dX}{dy}(y)\right| f(X(y)) \,, \tag{10.11}$$

where the set subscripts under the integral signs are understood to imply a standard orientation for the integration, that is, an integration from smaller to larger values. This version of the formula is more appropriate for generalisation to multiple dimensions where, as we have seen, the integration goes over sets.

Volume of parallelepipeds and their linear transforms

How do we generalise the formula (10.11) for one-dimensional integrals to the multi-dimensional case? Recall that the multi-dimensional integral is defined as a limit of sums (9.5) over partitions, where each summand involves the volume, $\mu(B)$, of a rectangular box B. We should therefore ask what happens to such boxes under a coordinate transformation $Y : B \to Y(B)$. It is useful to start with parallelepipeds (multi-dimensional analogues of parallelograms), located at y and with vertices v_i defined by

$$P(y; v_1, \ldots, v_n) = \{y + \sum_{i=1}^{n} t_i v_i \,|, 0 \le t_j \le 1\} \,, \tag{10.12}$$

of which boxes parallel to the coordinate axes are, of course, special cases (where the vertices v_i are multiples of the unit vectors e_i). From linear algebra we know that the volume of such a parallelepiped is given by

$$\mu(P(y; v_1, \ldots, v_n)) = |\det(v_1, \ldots, v_n)| \,. \tag{10.13}$$

(The absolute sign on the RHS is to ensure that the volume is positive, as it should be.) Before we consider general coordinate transformations let us first consider linear maps $A : \mathbb{R}^n \to \mathbb{R}^n$. Under such linear maps parallelepipeds are transformed to other parallelepipeds, namely

$$AP(y; v_1, \ldots, v_n) = P(y'; v_1', \ldots, v_n') \quad \text{where} \quad y' = Ay,\ v_j' = Av_j \,. \tag{10.14}$$

How are the volumes of the original and the transformed parallelepiped related? A quick calculation shows that

$$\mu(P(y'; v_1', \ldots, v_n')) = \mu(P(Ay; Av_1, \ldots, Av_n)) = |\det(Av_1, \ldots, Av_n)| = |\det(A)|\, \mu(P(y, v_1, \ldots, v_n))| \,. \tag{10.15}$$

In other words, the volume changes by a factor $|\det(A)|$, where A is the linear transformation.

Now consider a general coordinate transformation X with which we transform the parallelepiped. In general, of course the image $X(P(y; v_1, \ldots, v_n))$ is no longer a parallelepiped. However, if the parallelepiped is small (that is, if the lengths $|v_i| = \mathcal{O}(\epsilon)$ are small) we can approximate the coordinate change by its linearised version at y using the Jacobi matrix by writing

$$X(y + t_i v_i) = X(y) + DX(y)(t_i v_i) + \mathcal{O}(\epsilon^2) \,. \tag{10.16}$$

This means that the image is, approximately, another parallelepiped at $y' = X(y)$ with vertices $v_i' = DX(y)v_i$. Evidently, the Jacobi matrix $DX(y)$ plays a similar role to the matrix A in our above discussion of linear maps. As a result the volume of the parallelepiped changes as

$$\mu(P(y'; v_1', \ldots, v_n')) = |\det(DX(y))|\, \mu(P(y; v_1, \ldots, v_n)) + \mathcal{O}(\epsilon^2) \,. \tag{10.17}$$

The determinant $\det(DX(y))$ of the Jacobi matrix, $DX(y)$ which appears in this formula is also called the **Jacobian** of X at y. In conclusion, the volume of a parallelepiped at y changes with the (absolute value of the) Jacobian of X at y. This indicates that the additional factor $|\frac{dX}{dy}(y)|$ on the RHS of the one-dimensional formula (10.11) should be replaced by the Jacobian $|\det(DX(y))|$ in the multi-dimensional case. Since the integral is defined as the limit where the size of the boxes goes to zero we can expect that the $\mathcal{O}(\epsilon^2)$ errors in Eq. (10.17) will not matter.

The transformation formula

The above discussion indicates a multi-dimensional transformation formula as stated in the following theorem.

Theorem 10.4. *(Transformation formula)*
Let U and V be open sets in $\mathbb{R}^n$ and $X : V \to U$ a diffeomorphism and $f : U \to \mathbb{R}$ a bounded function with compact support. Then f is Riemann integrable over U iff the function $y \mapsto f(X(y))\,|\det(DX(y))|$ is Riemann integrable over V. In this case, we have

$$\int_U d^n x\, f(x) = \int_V d^n y\, |\det(DX(y))|\, f(X(y)) \,. \tag{10.18}$$

Proof. The idea of the proof is encapsulated in our above discussion of how the volume of small parallelepipeds changes under coordinate transformations. The details are somewhat technical as one needs to ensure that the linearised approximation of the coordinate transformation by the Jacobi matrix indeed delivers the exact result in the limit which defines the integral. The details can be found in Ref. [10]. □

Symbolically, the transformation formula (10.18) can also be stated as

$$d^n x = d^n y \left| |\det\left(\frac{\partial x}{\partial y}\right)\right| \,, \tag{10.19}$$

where we have written the Jacobi matrix as $DX(y) = \frac{\partial x}{\partial y}(y)$ and we have adopted the physics notation with the same symbol used for the function X and the coordinate x.

Example 10.3. Integration in polar and spherical coordinates
Question: Consier $\mathbb{R}^2$ with coordinates $(x_1, x_2) = (x, y)$ and a function $(x, y) \mapsto f(x, y)$. Re-write the integral of f in terms of polar coordinates (r, φ) defined in the usual way by $x(r, \varphi) = r\cos(\varphi)$ and $y(r\varphi) = r\sin(\varphi)$. Do the same in the three-dimensional case, that is, convert three-dimensional integrals into spherical coordinates.

Answer: The Jacobi matrices for these coordinate changes have already been computed in Exercise 5.2. Using these results we find for the two-dimensional case that $\det\left(\frac{\partial(x,y)}{\partial(r,\varphi)}\right) = r$ and, hence,

$$dx\, dy = \left|\det\left(\frac{\partial(x, y)}{\partial(r, \varphi)}\right)\right| dr\, d\varphi = r\, dr\, d\varphi \,. \tag{10.20}$$

For the three-dimensional case with Cartesian coordinates (x, y, z) and spherical coordinates (r, θ, φ) defined by $(x, y, z) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta)$ we have $\det\left(\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)}\right) = r^2\sin(\theta)$ and, hence,

$$dx\, dy\, dz = \left|\det\left(\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)}\right)\right| dr\, d\theta\, d\varphi = r^2\sin(\theta)\, dr\, d\theta\, d\varphi \,. \tag{10.21}$$

Example 10.4. Area of a disk and volume of a three-dimensional ball
Question: Of course you know the formulae for the area of a disk with radius R and the volume of a three-dimensional ball with radius R. Now derive these, using the definition of volume in Eq. (9.12) (and the transformation of the measure to polar and spherical coordinates from Exercise 10.3). Further, what is the area of the interior of an ellipse with half axes lenghts a and b?
Answer: From Eqs. (9.12) and (10.20), the volume of the disk $B_R^2 = B_R^2(0) \subset \mathbb{R}^2$ with radius R is

$$\mathrm{vol}(B_R^2) = \int_{B_R^2} d^2x = \int_0^R dr\, r \int_0^{2\pi} d\varphi = \pi R^2 \,. \tag{10.22}$$

From the same basic equations we find for the three-dimensional ball $B_R^3 = B_R^3(0) \subset \mathbb{R}^3$ with radius R that

$$\mathrm{vol}(B_R^3) = \int_{B_R^3} d^3x = \int_0^R dr\, r^2 \int_0^{\pi} d\theta\, \sin(\theta) \int_0^{2\pi} d\varphi = \frac{4}{3}\pi R^3 \,. \tag{10.23}$$

An interior of an ellipse with half-axes of length a and b (centred on the origin) is given by

$$E_{a,b}(0) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\} \,.$$

Let us introduce modified polar coordinates (r, φ) by $x = ar\cos(\varphi)$ and $y = br\sin(\varphi)$. For the associated Jacobi matrix and the Jacobian we have

$$\frac{\partial(x,y)}{\partial(r,\varphi)} = \begin{pmatrix} a\cos(\varphi) & -ar\sin(\varphi) \\ b\sin(\varphi) & bt\cos(\varphi) \end{pmatrix} \quad \Rightarrow \quad \det\left(\frac{\partial(x,y)}{\partial(r,\varphi)}\right) = abr \quad \Rightarrow \quad dx\,dy = a\,b\,r\,dr\,d\varphi \,.$$

In those modified polar coordinates the ellipse interior can be written as

$$E_{a,b}(0) = \left\{ \begin{pmatrix} ar\cos(\varphi) \\ br\sin(\varphi) \end{pmatrix} \in \mathbb{R}^2 \mid r \in [0,1],\ \varphi \in [0, 2\pi) \right\} \,.$$

We conclude that the area of the ellipse interior is given by

$$\mathrm{vol}(E_{a,b}(0)) = \int_{E_{a,b}(0)} dx\,dy = ab \int_0^1 dr\, r \int_0^{2\pi} d\varphi = \pi ab \,.$$

Example 10.5. Behaviour of volume under scaling and linear transformations
Question: Consider a set $S \subset \mathbb{R}^n$ as well as its scaled version λS, where $\lambda \in \mathbb{R}$ and its linear transformed version MS, where M is an $n \times n$ matrix. How are the volumes $\mathrm{vol}(S)$, $\mathrm{vol}(\lambda S)$ and $\mathrm{vol}(MS)$ related? Apply your results to a unit square and a unit disk in $\mathbb{R}^2$.
Answer: A straightforward application of the transformation formula gives

$$\mathrm{vol}(\lambda S) = \int_{\lambda S} d^n x \overset{x=\lambda y}{=} |\lambda|^n \int_S d^n y = |\lambda|^n\, \mathrm{vol}(S) \tag{10.24}$$

$$\mathrm{vol}(MS) = \int_{MS} d^n x \overset{x=My}{=} |\det(M)| \int_S d^n y = |\det(M)| \mathrm{vol}(S) \,. \tag{10.25}$$

The unit square S in $\mathbb{R}^2$ can be written as

$$S = \{t_1 e_1 + t_2 e_2 \in \mathbb{R}^2 \,|\, t_1 \in [0,1],\ t_2 \in [0,1]\} \ .$$

Its scaling and linear transform (writing $M = (v_1, v_2)$)

$$\lambda S = \{t_1 \lambda e_1 + t_2 \lambda e_2 \in \mathbb{R}^2 \,|\, t_1 \in [0,1],\ t_2 \in [0,1]\} \ , \quad MS = \{t_1 v_1 + t_2 v_2 \in \mathbb{R}^2 \,|\, t_1 \in [0,1],\ t_2 \in [0,1]\} \ ,$$

represent a square with length λ and a parallelogram with vertices v_1 and v_2. For their areas we find from the above formulae

$$\mathrm{vol}(\lambda S) = \lambda^2 \ , \qquad \mathrm{vol}(MS) = |\det(v_1, v_2)| \ .$$

Note, the second result is the standard one for the area of a parallelogram but we have now properly derived this result.
For unit disk $B_1^2(0) \subset \mathbb{R}^2$ the scaled and linearly transformed version

$$\lambda B_1(0) = B_\lambda(0) \ , \qquad MB_1(0) = M\{t_1 e_1 + t_2 e_2 \,|\, t_1^2 + t_2^2 \leq 1\} = \{t_1 v_1 + t_2 v_2 \,|\, t_1^2 + t_2^2 \leq 1\}$$

represent a disk with radius λ and a linearly distorted disk. For their areas we have

$$\mathrm{vol}(\lambda B_1(0)) = \mathrm{vol}(B_\lambda(0)) = \pi\lambda^2 \ , \qquad \mathrm{vol}(MB_1(0)) = |\det(M)|\,\mathrm{vol}(B_1(0)) = \pi\,|\det(M)| \ .$$

Note in the case where $M = \mathrm{diag}(a, b)$ is diagonal $MB_1(0)$ represents an ellipse with half-axes lengths a and b and in this case the second result above specialises to the formula for the area of an ellipse interior, $\mathrm{vol}(MB_1(0)) = \pi ab$.

Example 10.6. Volume of n-dimensional ball
Question: Calculate the volume of an n-dimensional ball $B_R^n = B_R^n(0)$ with radius R.
Answer: From Exercise 10.5 we know that

$$\mathrm{vol}(B_r^n) = r^n \mathrm{vol}(B_1^n) \ , \tag{10.26}$$

so it is enough to compute the volume of the ball with radius 1. The idea is to derive a recursion formula for the volume of B_1^n in terms of the volume of B_1^{n-1}.

$$\mathrm{vol}(B_1^n) = \int_{-1}^{1} dx\, \mathrm{vol}\left(B^{n-1}_{\sqrt{1-x^2}}\right) \overset{\text{Eq. (10.26)}}{=} I_n\, \mathrm{vol}(B_1^{n-1}) \ , \qquad I_n := \int_{-1}^{1} dx\, (1-x^2)^{(n-1)/2} \ . \tag{10.27}$$

After some integration by parts exercise it can be shown that $I_n I_{n-1} = \frac{2\pi}{n}$ which gives the recursion formula

$$\mathrm{vol}(B_1^n) = \frac{2\pi}{n} \mathrm{vol}(B_1^{n-2}) \ . \tag{10.28}$$

Iterating this for even and odd dimensions separately and using $B_1^2 = \pi$ and $B_1^3 = \frac{4\pi}{3}$ then gives

$$\mathrm{vol}(B_1^{2m}) = \frac{\pi^m}{m!} \ , \qquad \mathrm{vol}(B_1^{2m+1}) = 2\frac{(2\pi)^m}{(2m+1)!!} \ , \tag{10.29}$$

where $(2m+1)!! = 1 \cdot 3 \cdots (2m+1)$.

Example 10.7. Volume of a cone
Question: Split up coordinates $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ as $y = (x_1, \ldots, x_{n-1})^T \in \mathbb{R}^{n-1}$ and $z = x_n \in \mathbb{R}$ and consider a (compact) set $B \in \mathbb{R}^{n-1}$. The cone $C_h(B)$ with height h over the base B is defined as (see Fig. 16).

$$C_h(B) = \left\{ \left(\left(1 - \frac{z}{h}\right) y, z \right) \in \mathbb{R}^n \,\middle|\, y \in B \text{ and } z \in [0, h] \right\} . \tag{10.30}$$

Compute the volume of this cone in terms of the volume of base B. Apply this result in the three-dimensional case, when the base is a disk with radius R and a square with length R.

Answer: For a fixed z the 'slice' of the cone is a re-scaled version of the base, namely $(1 - z/h)B$. Hence, we can write the volume as

$$\mathrm{vol}(C_h(B)) = \int_{C_h(B)} d^n x = \int_0^h dz\, \mathrm{vol}((1 - z/h)B) \overset{\text{Eq. (10.24)}}{=} \mathrm{vol}(B) \int_0^h dz \left(1 - \frac{h}{z}\right)^{n-1} = \frac{h}{n} \mathrm{vol}(B) .$$

For a three-dimensional cone ($n = 3$) with base B_R^2 with $\mathrm{vol}(B_R^2) = \pi R^2$ this means $\mathrm{vol}(C_h(B_R^2)) = \frac{\pi h R^2}{3}$. For a square Q_R with length R and $\mathrm{vol}(Q_R) = R^2$ the associated three-dimensional cone has volume $\mathrm{vol}(C_h(Q_R)) = \frac{hR^2}{3}$.

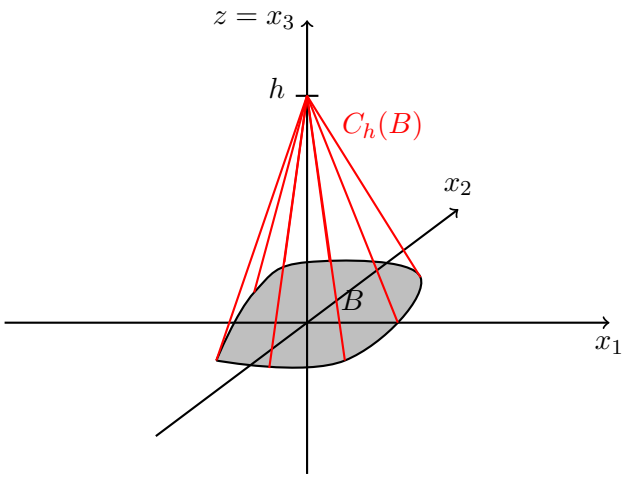


Figure 16: The cone with height h over B, as defined in Eq. (10.30).

Example 10.8. Gaußian integrals
Question: Start by working out the (one-dimensional) integral of $\exp(-x^2)$ over $\mathbb{R}$. Next consider the n-dimensional case with coordinates $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and integrate $\exp(-|x|^2)$ over $\mathbb{R}^n$. Finally, work out the integral of $\exp(-x^T A x)$ over $\mathbb{R}^n$, where A is a positive definite, symmetric $n \times n$ matrix.

Solution: Integrating

$$I_1 := \int_{\mathbb{R}} dx\, e^{-x^2} ,$$

requires a trick which involves considering I_1^2 and writing this as a two-dimensional integral in polar

coordinates.

$$I_1^2 = \left[\int_{\mathbb{R}} dx\, e^{-x^2}\right]\left[\int_{\mathbb{R}} dy\, e^{-y^2}\right] = \int_{\mathbb{R}^2} dx\, dy\, e^{-x^2-y^2} = \int_0^{2\pi} d\varphi \int_{\mathbb{R}} dr\, r e^{-r^2} = \pi \quad \Rightarrow \quad I_1 = \sqrt{\pi}\,.$$

Finding the result for the n-dimensional generalisation is now straightforward.

$$I_n = \int_{\mathbb{R}^n} d^n x\, e^{-|x|^2} = \prod_{i=1}^n \int_{\mathbb{R}} dx_i\, e^{-x_i^2} = I_1^n = \pi^{n/2}\,.$$

To integrate $\exp(-x^T A x)$ it is useful to diagonalise the matrix A first by introducing new coordinates $x = Ry$ with R an n-dimensional orthogonal matrix such that $R^T A R = \hat{A} = \mathrm{diag}(a_1, \ldots, a_n)$ (and the eigenvalues $a_i > 0$ since A is positive definite) and also coordinate $z_i = y_i/\sqrt{a_i}$. The Jacobi matrix and Jacobian of the coordinate transformation $y = Rs$ are

$$\frac{\partial x}{\partial y} = R \quad \Rightarrow \quad d^n x = \left|\det\left(\frac{\partial x}{\partial y}\right)\right| d^n y = |\det(R)|\, d^n x = d^n x\,.$$

where we have used that $\det(R) = \pm 1$ for an orthogonal matrix R (which follows from $R^T R = \mathbb{1}$).

$$\begin{aligned} I_n(A) &:= \int_{\mathbb{R}^n} d^n x\, e^{-x^T A x} = \int_{\mathbb{R}^n} d^n y\, e^{-y^T \hat{A} y} = \prod_{i=1}^n \int_{\mathbb{R}} dy_i\, e^{-a_i y_i^2} = \prod_{i=1}^n \frac{1}{\sqrt{a_i}} \int_{\mathbb{R}} dz_i\, e^{-z_i^2} \\ &= \frac{I_1^n}{\sqrt{\prod_{i=1}^b a_i}} = \sqrt{\frac{\pi^n}{\det(A)}} \end{aligned}$$

Gaußian integrals are of immense importance in physics, particularly in the context of the path integral formulation of quantum physics.

Example 10.9. Rotational volumes

Question: For a function $\rho : [a, b] \to \mathbb{R}$ with $\rho > 0$, the rotational volume $R_f \subset \mathbb{R}^3$ associated to f is defined by

$$R_\rho = \{(x, y, z) \in \mathbb{R}^3 \,|\, x^2 + y^2 \leq \rho(z)^2 \text{ and } z \in [a, b]\}\,. \tag{10.31}$$

Find a formula for the volume of R_ρ in terms of the function ρ.

Answer: It is useful to first introduce polar coordinates in the x-y plane and write the rotational volume as

$$R_\rho = \left\{ \begin{pmatrix} r\cos(\varphi) \\ r\sin(\varphi) \\ z \end{pmatrix} \,\middle|\, r \leq \rho(z) \text{ and } z \in [a, b] \right\}\,.$$

Splitting up the integral gives

$$\mathrm{vol}(R_\rho) = \int_{R_f} d^3 x = \int_0^{2\pi} d\varphi \int_a^b dz \int_0^{\rho(z)} dr\, r = \pi \int_a^b dz\, \rho(z)^2$$

Application 10.1. Liouville's theorem

Consider an n-dimensional vector field $A : U \to \mathbb{R}^n$ and the system of first order differential equations

$$\frac{dX}{dt} = A(X) \,, \tag{10.32}$$

where $X = (X_1, \ldots, X_n)^T$. The solutions to these equations are the integral curves $X_x(t)$ (with initial conditions $X_x(0) = x$) of the vector field A, that is, curves whose tangent vectors are given by A. Each $x \in U$ can be transported, by a parameter value t, along the integral curve it resides on to $x' = X_x(t)$. The so-obtained map φ_t is called a **flow** associated to Eq. (10.32). For small t we can write $x' = X_x(t) = X_x(0) + \frac{dX_x}{dt}(0)t + \mathcal{O}(t^2) = x + A(x)t + \mathcal{O}(t^2)$ so we can write the flow as

$$x' = \varphi_t(x) = x + A(x)\, t + \mathcal{O}(t^2) \,. \tag{10.33}$$

Consider a compact set $S \subset U$. We want to find out what happens to its volume of this set as we transport it with the flow, meaning we want to compare the volumes of S and $S_t = \varphi_t(S)$. To this end it is useful to compute the Jacobi matrix and the Jacobian of the flow (10.33) which are given by

$$\frac{\partial x'}{\partial x} = \mathbb{1} + DA(x)t + \mathcal{O}(t^2) \quad \Rightarrow \quad d^n x' = \left| \det\left(\frac{\partial x'}{\partial x}\right) \right| d^n x = (1 + \nabla \cdot A(x)t + \mathcal{O}(t^2))\, d^n x \,,$$

where we recall that $\det(\mathbb{1} + M) = 1 + \mathrm{tr}(M) + \mathcal{O}(M^2)$ and, from Example 8.7, that the trace of the Jacobi matrix of a vector field is the divergence of the vector field. Let us compute the time derivative of the volume

$$\frac{d}{dt}\left(\mathrm{vol}(S_t)\right)(0) = \left[\frac{d}{dt}\int_{S_t} d^n x'\right]_{t=0} = \left[\frac{d}{dt}\int_S d^n x\, (1 + \nabla \cdot A(x)t + \mathcal{O}(t^2))\right]_{t=0} = \int_S d^n x\, \nabla \cdot A(x)$$

This result has an important implication. If the vector field A is divergence-free, $\nabla \cdot A = 0$, then the associated flow does not change the volume.

Let us apply this result to Hamiltonian dynamics. With $X = (p_1, \ldots, p_\nu, q_1, \ldots, q_\nu)^T$ Hamilton's equations, Eqs. (7.9), can be written in the form (10.32) with the vector field A given by

$$A = \left(-\frac{\partial H}{\partial q_1}, \ldots, -\frac{\partial H}{\partial q_\nu}, \frac{\partial H}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_\nu}\right)^T \tag{10.34}$$

The divergence of this so-called **Hamiltonian vector field**

$$\nabla \cdot A = -\sum_{i=1}^{n} \frac{\partial^2 H}{\partial p_i \partial q_i} + \sum_{i=1}^{n} \frac{\partial^2 H}{\partial q_i \partial p_i} = 0$$

vanishes and, hence, the Hamiltonian flow in phase space preserves the volume. This statement is also known as **Liouville's theorem**.

Part IV
Manifolds

As we have already mentioned, knowing calculus in $\mathbb{R}^n$ is not sufficient for many physics applications. Physics involves curves, surfaces and their higher-dimenional analogues — collectively these objects are called **manifolds**. A point particle in Classical Mechanics moves along a curve, a one-dimensional manifold, a moving string sweeps out a surface, a two-dimensional manifold called the worldsheet of the string. In electro-magnetism we want to calculate the electric or magnetic flux through certain surfaces and, as you can learn in year 3, in General Relativity space-time is a four-dimensional manifold. In short, manifolds, or more precisely, **differentiable manifolds**, are everywhere in physics so we need to understand what they are, what their properties are and how we calculate with them. The branch of mathematics which deals with such questions is called **differential geometry** and it is a large field which remains an active area of contemporary research.

In these notes, we would like to consider a special class of differentiable manifolds, namely differentiable manifolds embedded in $\mathbb{R}^n$. These provide an easier, less abstract path into the subject than general manifolds while still illustrating many of the interesting structures and covering many applications in physics. General manifolds are defined without reference to an embedding space and you can have a look at Ref. [9] if you are curious about this more abstract approach. Here we will start by defining differential manifolds embedded into $\mathbb{R}^n$ and discuss some of their basic geometric properties. Our main goal is to understand integration over such manifolds. As we will see, such integrals are defined based on the multi-dimensional integral on $\mathbb{R}^n$ we have introduced earlier but it is important to realise that they are more complicated. A crucial ingredient in defining integration over manifolds is a **metric** on the manifold, an object which encodes the manifold's geometry. The measure for integrating over manifolds is not simply $d^n x$, as for integration over $\mathbb{R}^n$, but it is of the form $w(x)\, d^n x$, where w is a function that captures the non-trivial geometry of the manifold and how it affects the area/volume. As we will see, the function w can be written in terms of a metric on the manifold.

Having defined integration over manifolds, we will introduce two important integral theorems, **Gauß's theorem** and **Stokes's theorem**, which we can think of as higher-dimensional versions of the fundamental theorem of integration (10.10). These theorems roughly assert that the integral of a (certain) derivative of a function equals the integral of the function over the boundary of the manifold. (The fundamental theorem of integration (10.10) is of this form as the RHS can be viewed as a primitive integral over the boundary of the integration interval which only consists of its two endpoints.) Gauß's and Stokes's theorem are quite powerful (and they have an even more powerful generalisation to differential forms which you can read about in the Appendix A) and they have many applications in mathematics and physics. For example, they allow us to re-write Maxwell's equations (8.36) in integral form and this leads to a more precise version of the interpretation of these equations in Application 8.3.

11 Manifolds embedded in $\mathbb{R}^n$

11.1 Basic definition

Definition as a zero locus of equations
You have all seen examples of manifolds in $\mathbb{R}^n$, for example the unit circle S^1 embedded in $\mathbb{R}^2$. There are two ways of describing such a circle. We can either use an equation, so

$$S^1 = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\} \,, \tag{11.1}$$

or we can use a parametrisation

$$S^1 = \{(\cos(t), \sin(t)) \,|\, t \in [0, 2\pi)\} \,. \tag{11.2}$$

More generally, differentiable manifolds in $\mathbb{R}^n$ include curves, surfaces and their higher-dimensional analogues, all embedded in $\mathbb{R}^n$, and they can be defined in two, equivalent ways which are modelled on the two descriptions of the circle above. We begin with the definition in terms of equations which generalises Eq. (11.1) for the circle.

Definition 11.1. *A subset $M \subset \mathbb{R}^n$ is a k-dimensional (differential) sub-manifold of $\mathbb{R}^n$ if, for every $p \in M$ and an open neighbourhood $V \subset \mathbb{R}^n$ of p, we have $V \cap M = \{x \in V \,|\, f_1(x) = \cdots = f_{n-k}(x) = 0\}$ for continuously differentiable functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for which the Jacobi matrix*

$$\frac{\partial(f_1, \ldots, f_{n-k})}{\partial(x_1, \ldots, x_n)}(p) \tag{11.3}$$

has rank $n - k$.

This definition says, essentially, that a k-dimensional sub-manifold of $\mathbb{R}^n$ is given, locally, by the common zero locus of $n-k$ functions. (See Fig. 17.) The rank condition ensures that this manifold is smooth, that

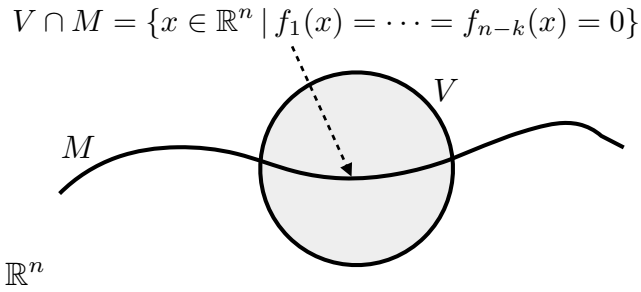


Figure 17: Manifold defined locally as the common zero locus of functions.

is, there are no "edges". Note that the dimension count is simple. To defined a k-dimensional manifold in $\mathbb{R}^n$ we require $n - k$ equations whose common zero locus provides the manifold (locally). Hence, to specify a surface in $\mathbb{R}^3$ one equation is required, to specify a curve in $\mathbb{R}^3$ we require two equations. The circle S^1 in (11.1) provides an example with $n = 2$, $k = 1$ and $f_1(x, y) = x^2 + y^2 - 1$.

Manifolds from parametrisations

The alternative description (11.2) of S^1 in terms of a parametrisation generalises to

Theorem 11.1. *The set $M \subset \mathbb{R}^n$ is a k-dimensional sub-manifold of $\mathbb{R}^n$ iff, for every $p \in M$, we have an open neighbourhood $V \subset \mathbb{R}^n$ of p, an open set $U \subset \mathbb{R}^k$ and a diffeomorphism (continuously differentiable, bijective and inverse continuous differentiable) map $X = (X_1, \ldots, X_n)^T : U \to V \cap M$ such that the Jacobi matrix*

$$\frac{\partial(X_1, \ldots X_n)}{\partial(t_1, \ldots, t_k)} \tag{11.4}$$

has rank k for all $t = (t_1, \ldots, t_k)^T \in U$. The map X is called a chart of M.

Proof. The proof involves the implicit function theorem 6.3 and it can, for example, be found in Ref. [7]. □

The circle S^1 as parametrised in Eq. (11.2) provides an example with $n = 2$, $k = 1$ and $X(t) = (X_1(t), X_2(t))^T = (\cos t, \sin t)^T$. More generally, the theorem says we can describe a k-dimensional submanifold of $\mathbb{R}^n$ (at least locally) by a parametrisation $X(t) = (X_1(\mathbf{t}), \dots X_n(\mathbf{t}))^T$, where $t \in \mathbb{R}^k$ are the parameters. (See Fig. 18.) For one parameter ($k = 1$), this describes a one-dimensional manifold, that is

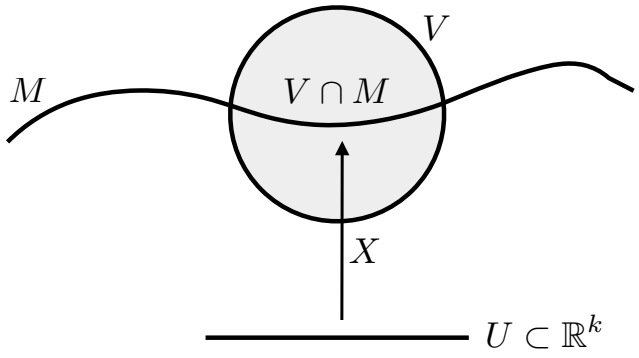


Figure 18: Manifold defined locally by a parametrisation (a chart).

a curve, for two parameters ($k = 2$) it describes a two-dimensional manifold, that is a surface, and so on. For a k-dimensional manifold in $\mathbb{R}^n$ the number of dimensions $n - k$ is called the **co-dimension** of the manifold. A manifold with co-dimension one in $\mathbb{R}^n$ is also called a **hyper-surface**.

Examples of manifolds

To gain some intuition let us list a few simple example of manifolds.

- **Open sets** A very simple example of an n-dimensional manifold in $\mathbb{R}^n$ is an open set $U \subset \mathbb{R}^n$. Indeed, we can use the identity map $X = \mathrm{id}_U$ as a parametrisation.

- **Lines and planes:** Lines and planes, say in $\mathbb{R}^3$, are special examples of manifolds. Indeed, a line can be written in parametric form $X(t) = u + tv$, where $u, v \in \mathbb{R}^3$ are given vectors and $t \in \mathbb{R}$ is the parameter. Line in $\mathbb{R}^3$ can also be described as a zero locus but this requires two equations, $v_1 \cdot x - a_1 = v_2 \cdot x = a_2 = 0$. A plane in $\mathbb{R}^3$ is described by the zero locus of a single equation, $v \cdot x - a = 0$ (the Cartesian form of a plane) or in parametric form, $X(t_1, t_2) = u + t_1 v + t_2 w$.

- **Quadratic manifolds:** For a quadratic form $q(x) = \frac{1}{2} x^T M x$ on $\mathbb{R}^n$ (where M is a symmetric $n \times n$ matrix) we can consider the zero locus $\{x \in \mathbb{R}^n \,|\, q(x) - c = 0\}$ (for a constant c) which gives a hyper-surface in $\mathbb{R}^n$. Depending on the eigenvalues of the matrix M, such quadratic hyper-surfaces can be ellipsoids, hyperboloids and various hybrid and special cases.

- **Unit circle:** Let us come back to the unit circle $S^1 \subset \mathbb{R}^2$, where you might have detected a small dissonance. The parametrisation of S^1 written down in Eq. (11.2) uses the interval $[0, 2\pi)$ as the parameter domain but this is not an open set, as required by Theorem 11.1. One way to proceed is to work, instead, with the open interval $(0, 2\pi)$ but in this case we are missing one point of the circle (the point $(1, 0)^T$). For some purposes this is good enough, such as for integration which is insensitive to a finite number of points. But if we want to obtain the entire circle and work with open sets we require at least two charts, for example $X^{(1)} : (0, 2\pi) \to S^1$ defined by $X^{(1)}(t) = (\cos(t), \sin(t))^T$ and $X^{(2)} : (0, 2\pi) \to S^1$ defined by $X^{(2)}(t) = (-\cos(t), \sin(t))^T$.

- **Level hyper-surfaces** Consider a (continuously differentiable) function $f : U \to \mathbb{R}$, with $U \subset \mathbb{R}^n$ open. Its level hyper-surfaces are defined by $\{x \in U \,|\, f(x) - c = 0\}$, for $c \in \mathbb{R}$. In other words,

these are simply the loci on which the given function takes on a given value c and clearly these loci form hyper-surfaces (at least locally where the rank condition in Def. 11.1 is satisfied). For the two-dimensional case, $n = 2$, these are the level lines used to draw a contour plot of the function.

11.2 Tangent space

Definition of tangent space

It is intuitively clear that a k-dimensional manifold M embedded in $\mathbb{R}^n$ has, at each point $p \in M$, a k-dimensional tangent space, that is, a one-dimensional vector space for a curve ($k = 1$), a two-dimensional vector space for a surface ($k = 2$), and so on. This tangent space to M at p is denoted by T_pM. How can we define the tangent space properly?

Suppose the point $p \in M$ is contained in the image of a chart $X : U \to V \cap M \ni p$, where $U \subset \mathbb{R}^k$ is open and, for convenience, we choose the chart such that $X(0) = p$. The idea is to consider all (differentiable) curves on the manifold which pass through the given point p and collect the tangent vectors to all these curves at p. These tangent vectors form the tangent space T_pM. More specifically, a curve in M passing through p has a corresponding curve $T : [-b, b] \to U$ in parameter space, such that, say, $T(0) = 0$. It lifts to the curve $X \circ T : [-b, b] \to M$ with $X \circ T(0) = p$ on the manifold M. The tangent vector to such a curve at p is given by [7]

$$\frac{dX(T(s))}{ds}(0) = \frac{\partial X}{\partial t^a}(0)\frac{dT^a}{ds}(0) \ . \tag{11.5}$$

We can choose any curve T in parameter space, for example $T(s) = sv$, where $v \in \mathbb{R}^k$ is any fixed vector. Then we have $\frac{dX}{ds}(0) = v$ and inserting into Eq. (11.5) shows that the tangent space consists of all linear combinations of the vectors $\frac{\partial X}{\partial t^a}(p)$, where $a = 1, \ldots, k$. Hence, we have the following

Proposition 11.1. *The tangent space T_pM of a k-dimensional manifold M in $\mathbb{R}^n$ with a chart $X : U \to V \cap M \ni p$ is given by*

$$T_pM = \mathrm{Span}(v_1(p), \ldots, v_n(p)) \quad \textit{where} \quad v_a(p) = \left.\frac{\partial X}{\partial t^a}\right|_p \ . \tag{11.6}$$

Note that $v_a(p) = \frac{\partial X}{\partial t_a}\,|_p$, where $a = 1, \ldots, k$, are k vectors in $\mathbb{R}^n$ and the maximal rank condition on the Jacobian (11.4) ensures that these k vectors are linearly independent. This means that the tangent space of a k-dimensional sub-manifold is indeed a vector space of dimension k, at every point. (See Fig. 19.) Hence, a general tangent vector at $p \in M$ can be written as $\sum_{a=1}^{k} \lambda^a v_a(p)$, where $\lambda^a \in \mathbb{R}$.

As a simple example, consider an open set $U \subset \mathbb{R}^n$ which forms an n-dimensional manifold with parametrisation $X = \mathrm{id}_U$. For this choice of parametrisation the basis vectors $v_a(p)$ of the tangent space T_pU are simply given by the standard unit vectors $(e_1, \ldots, e_n)$ (independently of the point p).

Going back to a general k-dimensional manifold $M \subset \mathbb{R}^n$ with chart $X : U \to V \cap M \ni p$ the basis vectors $v_a(p)$ for T_pM defined in Eq. (11.6) can also be be written as

$$v_a(p) = \left.\frac{\partial X^i}{\partial t^a}\right|_p e_i \tag{11.7}$$

This equation has an interesting interpretation. It is saying that we can transform from the standard tangent space basis vectors e_i of the embedding space to the basis vectors $v_a(p)$ of the manifold by using the Jacobi matrix of the chart map.

[7] We are quietly refining our notation here. Coordinates will be labelled by upper indices and Einstein summation is, by definition, carried out between the same upper and lower index.

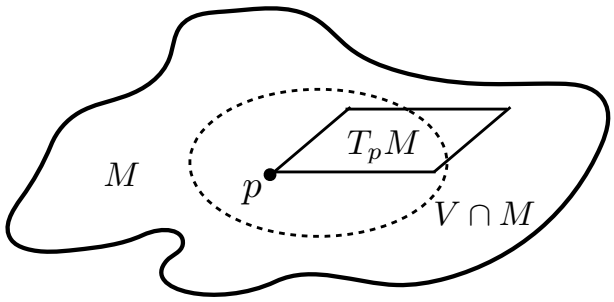


Figure 19: Tangent space T_pM of a manifold M at a point $p \in M$.

Example 11.1. Tangent spaces of S^1 and S^2
Question: Write down angular parametrisations for a unit circle S^1 and a unit sphere S^2 and find their tangent space.

Answer: A circle can be parametrised as in Eq. (11.2), so $X(t) = (\cos(t), \sin(t))^T$. Since this is a curve the tangent space is one-dimensional and it is, at every point on the circle, spanned by the vector

$$\frac{\partial X}{\partial t} = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} . \tag{11.8}$$

Note that, intuitively, this vector is indeed tangential to the circle for every t.

For S^2 we require two parameters $(t^1, t^2) = (\theta, \varphi)$ and a possible parametrisation is $X(\theta, \varphi) = (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta))^T$. This leads to the tangent vectors

$$\frac{\partial X}{\partial \theta} = \begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) \\ -\sin(\theta) \end{pmatrix} , \qquad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -\sin(\theta)\sin(\varphi) \\ \sin(\theta)\cos(\varphi) \\ 0 \end{pmatrix} , \tag{11.9}$$

and the tangent space is spanned by these two vectors, that is, $T_{X(\theta,\varphi)}S^2 = \text{span}\left(\frac{\partial X}{\partial \theta}, \frac{\partial X}{\partial \varphi}\right)$.

Hyper-surfaces and normal vectors
We would now like to discuss vectors normal to manifolds. What does it mean for a vector $N_p \in \mathbb{R}^n$ to be normal to a k-dimensional manifold $M \subset \mathbb{R}^n$ at a point $p \in M$? An obvious way to define this is to say that a vector field N is normal to M at $p \in M$ iff N_p is orthogonal to the entire tangent space T_pM, that is, if $v \cdot N_p = 0$ for all $v \in T_pM$.

Let us now specialise to a hyper-surface $M \subset \mathbb{R}^n$, that is, a manifold of co-dimension one. In this case, the normal vector needs to be orthogonal to a tangent space of dimension $n-1$ and is, therefore, unique up to re-scaling. How can we find such a normal vector, given a parametrisation $X : U \to V \cap M \ni p$. We require a vector $N_p \in \mathbb{R}^n$ which satisfies $\frac{\partial X}{\partial t^a}\big|_p \cdot N_p = 0$ for all $a = 1, \ldots, n-1$. Such a vector can be

explicitly obtained from the following formula [8]

$$N_{p,i} = \pm\epsilon_{ii_1\cdots i_{n-1}} \left.\frac{\partial X^{i_1}}{\partial t^1}\right|_p \cdots \left.\frac{\partial X^{i_{n-1}}}{\partial t^{n-1}}\right|_p \tag{11.10}$$

$$\Rightarrow \quad \left.\frac{\partial X}{\partial t^a}\right|_p \cdot N_p = \pm\det\left(\left.\frac{\partial X}{\partial t^a}\right|_p, \left.\frac{\partial X}{\partial t^1}\right|_p, \cdots, \left.\frac{\partial X}{\partial t^{n-1}}\right|_p\right) = 0\ . \tag{11.11}$$

In three dimensions, $n = 3$, a hyper-surface is the same as a surface and Eq. (11.10) for a normal vector specialises to

$$N_p = \pm \left.\frac{\partial X}{\partial t^1}\right|_p \times \left.\frac{\partial X}{\partial t^2}\right|_p\ . \tag{11.12}$$

A hyper-surface is often more conveniently defined (locally) by the zero locus of a function $f : V \to \mathbb{R}$ such that $M \cap V = \{x \in V \,|\, f(x) = 0\} \ni p$. Is there a simple way to find a normal vector to M in this case? Recall that the gradient of a function points into the direction of steepest ascent. This suggests the gradient of f should be orthogonal to level hyper-surfaces and, hence, to the manifold M. Let us prove this statement.

From Def. 11.1 we know that at least one component of $\nabla f(p)$ must be non-zero. Without restricting generality, let us assume this is the last component and we split coordinates in the neighbourhood of p up as $x = (t, y) \in \mathbb{R}^n$, where $t \in \mathbb{R}^{n-1}$ and $y \in \mathbb{R}$. Since $\partial_y f(p) \neq 0$ we know from the implicit function theorem 6.3 that we have a solution $y = g(t)$ with $f(t, g(t)) = 0$ in a neighbourhood of p. Further, from Eq. (6.27) we know that the derivatives of g are given by

$$\partial_a g = -\frac{\partial_a f}{\partial_y f} \quad \text{for} \quad a = 1, \ldots, n-1\ . \tag{11.13}$$

Using the function g we can also write down a parametrisation for the manifold and compute the resulting tangent vectors.

$$X(t) = \begin{pmatrix} t_1 \\ \vdots \\ t_{n-1} \\ g(t) \end{pmatrix} \qquad \Rightarrow \qquad \frac{\partial X}{\partial t_a} = \begin{pmatrix} e_a \\ \partial_a g \end{pmatrix} \quad \text{for} \quad a = 1, \ldots, n-1\ . \tag{11.14}$$

Here e_a are the standard unit vectors in $\mathbb{R}^{n-1}$. We can now work out the dot product between the gradient ∇f and the above tangent vectors which leads to

$$\nabla f \cdot \frac{\partial X}{\partial t^a} = \partial_a f + \partial_y f\, \partial_a g \overset{\text{Eq. (11.13)}}{=} 0\ . \tag{11.15}$$

This shows that ∇f is indeed orthogonal to all tangent vectors and, hence, to T_pM and this is the desired result. Hence, if the hyper-surface is described as a zero locus of a function f we can obtain a normal vector by setting $N_p = \pm\nabla f(p)$. Whether we obtain a normal vector N in this way or via a parametrisation from Eq. (11.10) we can always define a unit normal vector field [9]

$$\hat{n} = \frac{N}{|N|}\ . \tag{11.16}$$

[8] The sign ambiguity in the formula reflects the fact that the normal vector can point 'up' or 'down'.

[9] Do not confuse the unit normal vector, which we denote by $\hat{n}$, with the dimension of the space which we normally denote by n.

However, there is a subtlety. Obtaining a continuous unit normal vector from a parametrisation or a defining function in the way described above only provides us with a local unit normal vector (remembering from Def. 11.1 and Theorem 11.1 that defining functions and parametrisations only provide local patches of the manifold). Can the various local versions of $\hat{n}$ always be 'glued' together to a continuous unit normal vector on the entire manifold? The answer is 'no' as the example of the Möbius strip in Fig. 20 shows. If there exist a continuous unit normal $\hat{n} : M \to \mathbb{R}^n$ on the entire hyper-surface then M is called

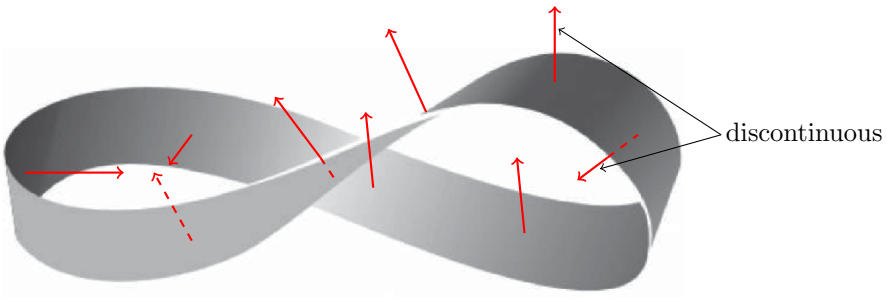


Figure 20: The Möbius strip, a non-orientable surface. Starting with the unit normal vector n (in red) in the upper right and continuous continuing the vector ends up with a discontinuity after one revolution.

orientable, otherwise it is called **non-orientable**. The Möbius strip is an example of a non-orientable surface.

11.3 Metrics on manifolds

Definition of a metric

We have seen that a manifold has a vector space - the tangent space - attached to each of its points. The geometry of a vector space - length of vectors, angles between them etc. - is described by a scalar product on the vector space. This suggests, in order to describe the geometry of manifolds we require a scalar product on every tangent space. A map G which assigns to every point $p \in M$ a scalar product $G_p : T_pM \times T_pM \to \mathbb{R}$ on T_pM is called a **metric**. A manifold with a metric is also called a **Riemannian manifold** and, as we will see, this structure is crucial for many purposes, including defining integration over manifolds.

The induced metric on an embedded manifold

How can we describe a metric practically? Let us start with an open set $W \subset \mathbb{R}^n$, seen as a manifold, and a metric G on W. For $p \in W$ we can think of G_p as a (positive definite) symmetric matrix whose entries $G_{p.ij}$ are given by

$$G_{p,ij} := G_p(e_i, e_j) \; . \tag{11.17}$$

Then the scalar product of two tangent vectors $u = u^i e_i$ and $w = w^j e_j$ in T_pV is given by $G_{p,ij} u^i w^j$. Now suppose we have a k-dimensional manifold $M \subset W \subset \mathbb{R}^n$ with local parametrisation $X : U \to V \cap M \ni p$ and the tangent space T_pM with basis $v_a(p) = \frac{\partial X}{\partial t^a}\big|_p$, where $a = 1, \ldots, k$. The point is that $T_pM \subset T_pW$ is a vector sub-space for every p and, therefore, the metric G on W induces a metric on M. We can find the entries $g_{p,ab}$ of the induced scalar product on T_pM, relative to the basis v_a, by evaluating on these basis vectors, so

$$g_{p,ab} := G_p(v_a(p), v_b(p)) \overset{\text{Eq. (11.7)}}{=} G_p\left(\sum_i \frac{\partial X^i}{\partial t^a} e_i, \sum_j \frac{\partial X^j}{\partial t^b} e_j\right) = \frac{\partial X^i}{\partial t^a}\frac{\partial X^j}{\partial t^b} G_{p,ij} \; . \tag{11.18}$$

Then, given two tangent vectors $\alpha = \alpha^a v_a(p)$ and $\beta = \beta^b v_b(p)$ in T_pM their scalar product is given by $g_{p,ab}\alpha^a\beta^b$. The crucial result from Eq. (11.18) is a rule for how metrics transform under a 'change of coordinates' (dropping the point p for simplicity),

$$g_{ab} = \frac{\partial X^i}{\partial t^a}\frac{\partial X^j}{\partial t^b}G_{ij} \ , \tag{11.19}$$

namely by summing two Jacobi matrices into the two indices of the metric.

Choice of metric

As you will learn in more detail in year 3, manifolds are the arena of General Relativity (space-time forms a manifold) and the metric on this manifold is the central object of the theory. Once the distribution of matter is fixed Einstein's equation determines the metric, much as Maxwell's equations determine the electric and magnetic field for given charge and current density. Eq. (11.19) is an important equation in General Relativity which tells us how metrics transforms under a change of coordinates.

In the present context we do not have a theory which determines the metric. However, Eq. (11.19) tells us once we have chosen a metric G an the embedding space we obtain an 'induced metric" g on the manifold. So is there an obvious, 'canonical' way to choose G? The embedding space is essentially $\mathbb{R}^n$, so setting G_p equal to the dot product at every point $p \in W$ is an obvious choice. In other words, we choose G to be the Euklidean metric, $G_p = \mathbb{1}_n$. As we know this choice will lead to the usual notions of area, volume etc. on $\mathbb{R}^n$ and, going along with this choice, Eq. (11.19) tells us the induced metric on M is

$$g_{ab} = \frac{\partial X^i}{\partial t^a}\frac{\partial X^j}{\partial t^b}\delta_{ij} = \frac{\partial X}{\partial t^a}\cdot\frac{\partial X}{\partial t^b} \ . \tag{11.20}$$

We can also write this same equation in matrix notation as

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial t^1}\right|^2 & \frac{\partial X}{\partial t^1}\cdot\frac{\partial X}{\partial t^2} & \cdots & \frac{\partial X}{\partial t^1}\frac{\partial X}{\partial t^k} \\ \frac{\partial X}{\partial t^1}\cdot\frac{\partial X}{\partial t^2} & \left|\frac{\partial X}{\partial t^2}\right|^2 & \cdots & \frac{\partial X}{\partial t^2}\frac{\partial X}{\partial t^k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial X}{\partial t^1}\cdot\frac{\partial X}{\partial t^k} & \frac{\partial X}{\partial t^2}\cdot\frac{\partial X}{\partial t^k} & \cdots & \left|\frac{\partial X}{\partial t^k}\right|^2 \end{pmatrix} = \left(\frac{\partial X}{\partial t}\right)^T\left(\frac{\partial X}{\partial t}\right) \ . \tag{11.21}$$

In short, the entries of the induced matric consist of the mutual scalar products of the tangent space basic vectors $\frac{\partial X}{\partial t^a}$, as Eq. (11.21) shows. We can say that the induced metric provides information about the manifold's geometry in the sense that it captures the relation, via the dot product, of its tangent vector basis.

The determinant, $\det(g)$ of the induced metric is called **Gram's determinant** and is given by

$$\det(g) = \det\left(\left(\frac{\partial X}{\partial t}\right)^T\left(\frac{\partial X}{\partial t}\right)\right) \ , \tag{11.22}$$

as taking the determinant of Eq. (11.21) shows. Gram's determinant will play an important role in defining integration over manifolds, as we will see later. For now, to get a better intuition let us compute the induced metric for a few examples.

Example 11.2. Induced metric of surfaces

Question: Consider a plane with parametrisation $X(t_1, t_2)u + t_1 v + 1 + t_2 v_2$, where $u, v_1, v_2 \in \mathbb{R}^n$ are fixed vectors. What is the induced metric for this plane. Further, consider the surface given by the parametrisation $X(t_1, t_2) = (1 + t_1, t_1 - t_2^2/2, t_2 - t_1^2/2)^T$ (see Fig. 21) and work out its induced metric. Do the same for a two-dimensional sphere $S^2 \subset \mathbb{R}^3$.

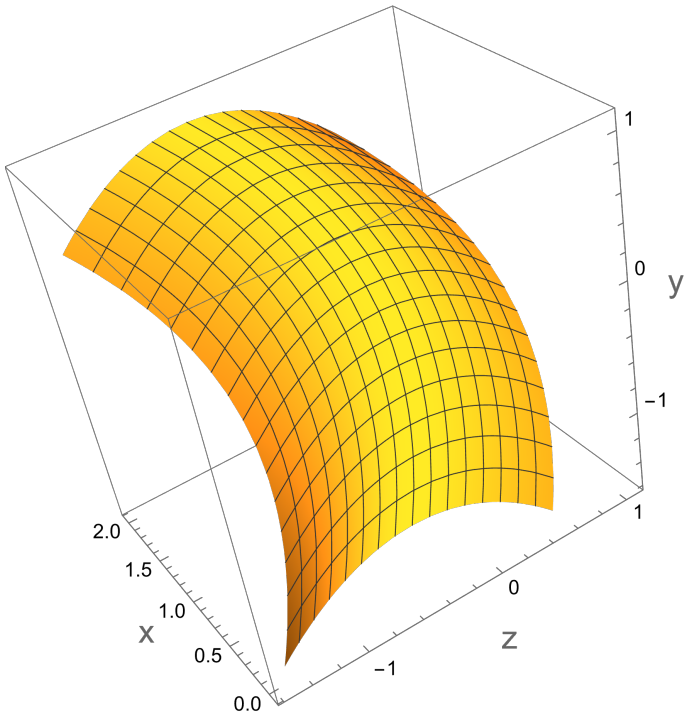


Figure 21: The surface with parametrisation $X(t_1, t_2) = (1 + t_1, t_1 - t_2^2/2, t_2 - t_1^2/2)^T$ from Example 11.2.

Answer: The standard tangent vector basis of the plane is given by

$$\frac{\partial X}{\partial t_1} = v_1 \ , \qquad \frac{\partial X}{\partial t_2} = v_2 \ .$$

which, from Eq. (11.21), implies an induced metric

$$g = \begin{pmatrix} |v_1|^2 & v_1 \cdot v_2 \\ v_1 \cdot v_2 & |v_2|^2 \end{pmatrix} \ .$$

For the second example surface we have the tangent vector basis

$$\frac{\partial X}{\partial t_1} = \begin{pmatrix} 1 \\ 1 \\ -t_1 \end{pmatrix} \ , \qquad \frac{\partial X}{\partial t_2} = \begin{pmatrix} 0 \\ -t_2 \\ 1 \end{pmatrix} \ .$$

This leads to an induced metric

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial t_1}\right|^2 & \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} \\ \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} & \left|\frac{\partial X}{\partial t_2}\right|^2 \end{pmatrix} = \begin{pmatrix} 2 + t_1^2 & -t_1 - t_2 \\ -t_1 - t_2 & 1 + t_2^2 \end{pmatrix} \ .$$

A two sphere can be parametrised by coordinates $(t_1, t_2) = (\theta, \varphi)$, where $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi)$ and

$$X(\theta, \varphi) = (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta)) \,. \tag{11.23}$$

This means its tangent vectors are

$$\frac{\partial X}{\partial \theta} = \begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) \\ -\sin(\theta) \end{pmatrix} \,, \qquad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -\sin(\theta)\sin(\varphi) \\ \sin(\theta)\cos(\varphi) \\ 0 \end{pmatrix} \,. \tag{11.24}$$

This leads to the induced metric

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial \theta}\right|^2 & \frac{\partial X}{\partial \theta} \cdot \frac{\partial X}{\partial \varphi} \\ \frac{\partial X}{\partial \theta} \cdot \frac{\partial X}{\partial \varphi} & \left|\frac{\partial X}{\partial \varphi}\right|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\theta) \end{pmatrix} \,. \tag{11.25}$$

Example 11.3. Induced metric for polar and spherical coordinates
Question: You can think of standard two-dimensional polar coordinates $X(r, \varphi) = (r\cos(\varphi), r\sin(\varphi))^T$ as defining a two-dimensional manifold with parameters $(t_1, t_2) = (r, \varphi)$. Compute its induced metric. Do the same for three-dimensional spherical coordinates.

Answer: The standard tangent vectors are

$$\frac{\partial X}{\partial r} = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \,, \qquad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -r\sin(\varphi) \\ r\cos(\varphi) \end{pmatrix} \,. \tag{11.26}$$

Note, these two vectors are orthogonal so we find for the induced metric

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial r}\right|^2 & \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \varphi} \\ \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \varphi} & \left|\frac{\partial X}{\partial \varphi}\right|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \,. \tag{11.27}$$

For spherical coordinates $X(r, \theta, \varphi) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta))^T$ the tangent vectors are

$$\frac{\partial X}{\partial r} = \begin{pmatrix} \sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta) \end{pmatrix} \,, \qquad \frac{\partial X}{\partial \theta} = \begin{pmatrix} r\cos(\theta)\cos(\varphi) \\ r\cos(\theta)\sin(\varphi) \\ -r\sin(\theta) \end{pmatrix} \,, \qquad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -r\sin(\theta)\sin(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ 0 \end{pmatrix} \,. \tag{11.28}$$

Note, as in the two-dimensional case, these vectors are orthogonal. The induced metric is

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial r}\right|^2 & \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \theta} & \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \varphi} \\ \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \theta} & \left|\frac{\partial X}{\partial \theta}\right|^2 & \frac{\partial X}{\partial \theta} \cdot \frac{\partial X}{\partial \varphi} \\ \frac{\partial X}{\partial r} \cdot \frac{\partial X}{\partial \varphi} & \frac{\partial X}{\partial \theta} \cdot \frac{\partial X}{\partial \varphi} & \left|\frac{\partial X}{\partial \varphi}\right|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2(\theta) \end{pmatrix} \tag{11.29}$$

11.4 Differential operators in different coordinate systems

Suppose we have a differential operator, such as the gradient, the curl and the divergence. So far we have expressed these in terms of Cartesian coordinates. In this section we would like to work out which form they take in other coordinate systems. We express Cartesian coordinates $x = (x^i) \in \mathbb{R}^n$ in terms of new

coordinates $t = (t^a)$, where $a = 1, \ldots, n$, by considering the diffeomorphism $x = x(t)$. We know that the new coordinate system $v_a = \frac{\partial X}{\partial t^a}$ associated to the t-coordinates is related to the one given by the standard unit vectors e_i via

$$v_a = \frac{\partial x^i}{\partial t^a} e_i \qquad \text{or} \qquad e_i = \frac{\partial t^a}{\partial x^i} v_a \ . \tag{11.30}$$

We also have the induced metric g_{ab} and its inverse, written as g^{ab}, given by

$$g_{ab} = \frac{\partial X^i}{\partial t^a} \frac{\partial X^j}{\partial t^b} \delta_{ij} \ , \qquad g^{ab} = \frac{\partial t^a}{\partial X^i} \frac{\partial t^b}{\partial X^j} \delta^{ij} \tag{11.31}$$

The gradient

We start by converting the gradient ∇f of a function $f : U \to \mathbb{R}$ (for $U \subset \mathbb{R}^n$ open) into a new coordinate system. A short calculation gives

$$\nabla f = \sum_i e_i \frac{\partial f}{\partial x^i} = v_b \frac{\partial t^b}{\partial x^i} \frac{\partial t^a}{\partial x^i} \frac{\partial f}{\partial t^a} = g^{ab} \frac{\partial f}{\partial t^a} v_b \ . \tag{11.32}$$

So, in short, we find the general formula

$$\nabla f = g^{ab} \frac{\partial f}{\partial t^a} v_b \ , \tag{11.33}$$

for the gradient in arbitrary coordinates t. All we need to know to work this out for specific coordinates are the vectors $v_a = \frac{\partial X}{\partial t^a}$ and the induced metric g_{ab} (and its inverse). Let us consider this for some examples.

Example 11.4. Gradient in polar and spherical coordinates

Question: Use Eq, (11.33) to write the two-dimensional gradient in polar coordinates and the three-dimensional gradient in spherical coordinates.

Answer: In the two-dimensional case with $(t_1, t_2) = (r, \varphi)$ and $x = (x_1, x_2)^T = (r\cos(\varphi), r\sin(\varphi))^T$, we have already computed the basis vectors and the induced metric in Example 11.3. It is customary to define (compare Eq. (11.26))

$$e_r := v_r = \frac{\partial x}{\partial r} = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix} \ , \qquad e_\varphi := \frac{1}{r} v_\varphi = \frac{1}{r} \frac{\partial x}{\partial \varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \ , \tag{11.34}$$

since the so-defined basis (e_r, e_φ) is ortho-normal. From Eq. (11.33) and the inverse of the induced metric $g = \mathrm{diag}(1, r^2)$ in Eq. (11.27) this means

$$\nabla f = \frac{\partial f}{\partial r} v_r + \frac{1}{r^2} \frac{\partial f}{\partial \varphi} v_\varphi = \frac{\partial f}{\partial r} e_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} e_\varphi \ . \tag{11.35}$$

For the gradient in three dimensions, relative to spherical coordinates $(t_1, t_2, t_3) = (r, \theta, \varphi)$ with $x = (x_1, x_2, x_3) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), \cos(\theta))$ we define an ortho-normal basis by normalising the vectors in Eq. (11.28), that is,

$$e_r := v_r = \frac{\partial x}{\partial r} = \begin{pmatrix} \sin(\theta)\cos(\varphi) \\ \sin(\theta)\sin(\varphi) \\ \cos(\theta) \end{pmatrix} \qquad e_\theta; = \frac{1}{r} v_\theta = \frac{1}{r} \frac{\partial x}{\partial \theta} = \begin{pmatrix} \cos(\theta)\cos(\varphi) \\ \cos(\theta)\sin(\varphi) \\ -\sin(\theta) \end{pmatrix}$$

$$e_\varphi := \frac{1}{r\sin(\theta)} v_\varphi = \frac{1}{r\sin(\theta)} \frac{\partial x}{\partial \varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \ . \tag{11.36}$$

With the induced metric $g = \mathrm{diag}(1, r^2, r^2\sin^2(\theta))$ from Eq. (11.29) and Eq. (11.33), this given for the three-dimensional gradient in spherical coordinates

$$\begin{aligned}\nabla f &= \frac{\partial f}{\partial r} v_r + \frac{1}{r^2}\frac{\partial f}{\partial \theta} v_\theta + \frac{1}{r^2\sin^2(\theta)}\frac{\partial f}{\partial \varphi} v_\varphi \\ &= \frac{\partial f}{\partial r} e_r + \frac{1}{r}\frac{\partial f}{\partial \theta} e_\theta + \frac{1}{r\sin(\theta)}\frac{\partial f}{\partial \varphi} e_\varphi \,. \end{aligned} \tag{11.37}$$

Exercise 11.2. *Use Eq. (11.33) to work out the three-dimensional gradient in cylindrical coordinates.*

The divergence

Converting the divergence to arbitrary coordinates is a bit more complicated and we need a preparatory Lemma which tells us how to differentiate determinants.

Lemma 11.1. *Consider a $k \times k$ matrix-valued function $M = M(t)$, where $t = (t^a)$ are the coordinates, and $M(t)$ is an invertible matrix for all t. Then*

$$\frac{1}{\det(M)}\frac{\partial}{\partial t^a}\det(M) = \mathrm{tr}\left(M^{-1}\frac{\partial M}{\partial t^a}\right) \tag{11.38}$$

Proof. We note that $C_{ab} = \frac{\partial \det(M)}{\partial M_{ab}}$ are, in fact, the entries of the co-factor matrix C of M and we recall that $M^{-1} = \frac{1}{\det(M)} C^T$. Then

$$\frac{1}{\det(M)}\frac{\partial}{\partial t^a}\det(M) = \frac{1}{\det(M)}\frac{\partial \det(M)}{\partial M_{bc}}\frac{\partial M_{bc}}{\partial t^a} = \frac{1}{\det(M)} C_{bc}\frac{\partial M_{bc}}{\partial t^a} = (M^{-1})_{cb}\frac{\partial M_{bc}}{\partial t^a} = \mathrm{tr}\left(M^{-1}\frac{\partial M}{\partial t^a}\right) .$$

□

Now consider the divergence $\nabla \cdot A$ of a vector field $A(x) = A^i(x) e_i$. Not only do we have to re-write A^i in terms of the new coordinates t but we also have to convert the Cartesian components A^i, relative to the standard unit vector basis, into components $\tilde{A}^a$, relative to our new basis $v_a = \frac{\partial X}{\partial t^a}$. This gives

$$A = A^i e_i = \tilde{A}^a v_a \overset{\text{Eq. (11.7)}}{=} \tilde{A}^a \frac{\partial x^i}{\partial t^a} e_i \quad \Rightarrow \quad A^i = \frac{\partial x^i}{\partial t^a}\tilde{A}^a \,. \tag{11.39}$$

Hence, the transformation between the vector field components A^i and $\tilde{A}^a$ is accomplished by multiplying with the Jacobi matrix of the transformation, as one might have expected. For the divergence (abbreviating $J = \frac{\partial x}{\partial t}$ and remembering that $\det(g) = \det(J)^2$ we can write

$$\begin{aligned}\nabla \cdot A = \frac{\partial}{\partial x^i} A^i &\overset{\text{Eq. (11.39)}}{=} \frac{\partial t^b}{\partial x^i}\frac{\partial}{\partial t^b}\left(\frac{\partial x^i}{\partial t^a}\tilde{A}^a\right) = \frac{\partial}{\partial t^a}\tilde{A}^a + \frac{\partial t^b}{\partial x^i}\frac{\partial}{\partial t^a}\left(\frac{\partial x^i}{\partial t^b}\right)\tilde{A}^a \\ &= \frac{\partial}{\partial t^a}\tilde{A}^a + \mathrm{tr}\left(J^{-1}\frac{\partial J}{\partial t^a}\right)\tilde{A}^a \overset{\text{Eq. (11.38)}}{=} \frac{\partial}{\partial t^a}\tilde{A}^a + \frac{1}{\det(J)}\frac{\partial \det(J)}{\partial t^a}\tilde{A}^a \\ &\overset{\text{Eq. (11.22)}}{=} \frac{\partial}{\partial t^a}\tilde{A}^a + \frac{1}{\sqrt{\det(g)}}\frac{\partial\sqrt{\det(g)}}{\partial t^a}\tilde{A}^a = \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial t^a}\left(\sqrt{\det(g)}\,\tilde{A}^a\right)\end{aligned}$$

In summary, the result for the divergence in arbitrary coordinates t is

$$\nabla \cdot A = \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial t^a}\left(\sqrt{\det(g)}\,\tilde{A}^a\right) \quad \text{where} \quad A = \tilde{A}^a\frac{\partial x}{\partial t^a} \,. \tag{11.40}$$

Again, it is worth applying this formula to examples.

Example 11.5. Divergence in polar and spherical coordinates
Question: Apply Eq. (11.40) to derive the form of the two-dimensional divergence in polar coordinates and the three-dimensional divergence in spherical coordinates.

Answer: In two dimensions with coordinates $(t_1, t_2) = (r, \varphi)$ and $x = (x_1, x_2) = (r\cos(\varphi), r\sin(\varphi))$, we know from Eq. (11.25) that $\sqrt{\det(g)} = r$. We also know from Eq. (11.34) that the basis (e_r, e_φ) is related to the basis $v_a = \frac{\partial x}{\partial t^a}$ by $e_r = v_r$ and $e_\varphi = \frac{1}{r} v_\varphi$. This means the components of the vector field A relative to the basis (e_r, e_φ) are $A^r = \tilde{A}^1$ and $A^\varphi = r\tilde{A}^2$, where the components $\tilde{A}^a$ are relative to the basis v_a. Putting this together, and inserting into Eq. (11.40), we find for the divergence

$$\nabla \cdot A = \frac{1}{r}\frac{\partial}{\partial r}(r\tilde{A}^1) + \frac{1}{r}\frac{\partial}{\partial \varphi}(r\tilde{A}^2) = \frac{1}{r}\frac{\partial}{\partial r}(rA^r) + \frac{1}{r}\frac{\partial}{\partial \varphi}(A^\varphi) \tag{11.41}$$

In three dimensions with coordinates $(t_1, t_2, t_3) = (r, \theta, \varphi)$ and

$$x = (x_1, x_2, x_3) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta)) \ ,$$

we know from Eq. (11.29) that $\sqrt{\det(g)} = r^2\sin(\theta)$. Furthermore, the relations beween the basis $v_a = \frac{\partial x}{\partial t^a}$ and $(e_r, e_\theta, e_\varphi)$ in Eq, (11.36) show that $A^r = \tilde{A}^1$, $A^\theta = r\tilde{A}^2$ and $A^\varphi = r\sin(\theta)\tilde{A}^3$, where $\tilde{A}^a$ are the components relative to the basis v_a. This leads to the divergence

$$\begin{aligned}\nabla \cdot A &= \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial r}(r^2\sin(\theta)\tilde{A}^1) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}(r^2\sin(\theta)\tilde{A}^2) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \varphi}(r^2\sin(\theta)\tilde{A}^3) \\ &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2A^r) + \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)A^\theta) + \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \varphi}(A^\varphi)\end{aligned} \tag{11.42}$$

Exercise 11.3. *Derive the formulae for the three-dimensional divergence in cylindrical coordinates. Further, derive the formula for the curl of a three-dimensional vector field in spherical coordinates.*

The Laplacian
The Laplace operator $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ is important in many areas of physics and re-writing it in different coordinate systems is a standard tool. To simplify notation we denote the Jacobi matrix of the coordinates change $x = x(t)$ by $J = \frac{\partial x}{\partial t}$, its inverse by $K = J^{-1}$ and abbreviate $\partial_a = \frac{\partial}{\partial t^a}$. From $JK = \mathbb{1}$ is follows by differentiating that

$$\partial_a(JK) = (\partial_a J)K + J\partial_a K = 0 \quad \Rightarrow \quad \partial_a K = -K(\partial_a J)K \ .$$

For the Laplacian, acting on a function $f : U \to \mathbb{R}$, with $U \subset \mathbb{R}^n$ open, we can write

$$\begin{aligned}\Delta f &= \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^i}f = \frac{\partial t^a}{\partial x^i}\frac{\partial}{\partial t^a}\left(\frac{\partial t^b}{\partial x^i}\frac{\partial f}{\partial t^b}\right) = g^{ab}\partial_a\partial_b f + K^a{}_i(\partial_a K^b{}_i)\partial_b f \\ &= g^{ab}\partial_a\partial_b f + \partial_a(K^a{}_i K^b{}_i)\partial_b f - K^b{}_i\partial_a(K^a{}_i)\partial_b f \\ &= g^{ab}\partial_a\partial_b f + \partial_a(g^{ab})\partial_b f + K^b{}_i K^a{}_j K^c{}_i\partial_a(J^j{}_c)\partial_b f \\ &= g^{ab}\partial_a\partial_b f + \partial_a(g^{ab})\partial_b f + g^{bc}K^a{}_j\partial_c(J^j{}_a)\partial_b f \\ &= g^{ab}\partial_a\partial_b f + \partial_a(g^{ab})\partial_b f + g^{ab}\mathrm{tr}(K\partial_a J)\partial_b f \\ &= g^{ab}\partial_a\partial_b f + \partial_a(g^{ab})\partial_b f + g^{ab}\frac{\partial_a\sqrt{\det(g)}}{\sqrt{\det(g)}}\partial_b f = \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial t^a}\left(\sqrt{\det(g)}\, g^{ab}\frac{\partial f}{\partial t^b}\right)\end{aligned}$$

Hence, our final formula for the Laplacian in arbitrary coordinates t is

$$\Delta f = \frac{1}{\sqrt{\det(g)}}\frac{\partial}{\partial t^a}\left(\sqrt{\det(g)}\, g^{ab}\frac{\partial f}{\partial t^b}\right) \ . \tag{11.43}$$

Example 11.6. Laplacian in polar and spherical coordinates
Question: Use Eq. (11.43) to derive the two-dimensional Laplacian in polar coordinates and the three-dimensional Laplacian in spherical coordinates.

Answer: In two-dimensions with coordinates $(t_1, t_2) = (r, \varphi)$ and $x = (x_1, x_2) = (r\cos(\varphi), \sin(\varphi))$ we have from Eq. (11.27) that $g = \mathrm{diag}(1, r^2)$ and $\sqrt{\det(g)} = r$. Hence, the Laplacian in polar coordinates is

$$\Delta f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial \varphi}\left(r\frac{1}{r^2}\frac{\partial f}{\partial \varphi}\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2} \tag{11.44}$$

In three dimensions with spherical coordinates $(t_1, t_2, t_3) = (r, \theta, \varphi)$ and

$$x = (x_1, x_2, x_3) = (r\sin(\theta)\cos(\varphi), r\sin(\theta)\sin(\varphi), r\cos(\theta)) \ ,$$

we have from Eq. (11.29) that $g = \mathrm{diag}(1, r^2, r^2\sin^2(\theta))$ and $\sqrt{\det(g)} = r^2\sin(\theta)$. Inserting into Eq. (11.43) this gives

$$\begin{aligned}\Delta f &= \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial r}\left(r^2\sin(\theta)\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \varphi}\left(\frac{1}{\sin(\theta)}\frac{\partial f}{\partial \varphi}\right) \\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2 f}{\partial \varphi^2}\end{aligned} \tag{11.45}$$

Exercise 11.4. *Find the three-dimensional Laplacian in cylindrical coordinates using Eq. (11.43).*

Example 11.7. Laplacian on a sphere
Question: Even though Eq. (11.43) has been derived using the inverse Laplacian of the coordinate transform it remains valid when the manifold is genuinely of lower dimension. Use this fact to derive the Laplacian, Δ_{S^2}, on a two-dimensional unit sphere $S^2 \subset \mathbb{R}^3$, using the standard angular parametrisation of the sphere. Comment on how your result relates to the three-dimensional Laplacian in spherical coordinates.

Answer: The sphere can be parametrised by

$$x(\theta, \varphi) = (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta)) \ ,$$

and we know from Eq, (11.25) that its induced metric is given by $g = \mathrm{diag}(1, \sin^2(\theta))$ with $\sqrt{\det(g)} = \sin(\theta)$. Inserting into Eq. (11.43) gives

$$\Delta_{S^2} f = \frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial f}{\partial \theta}\right) + \frac{1}{\sin^2(\theta)}\frac{\partial^2 f}{\partial \varphi^2} \tag{11.46}$$

This is, in fact, just the angular part of the three-dimensional Laplacian in spherical coordinates (11.45) and we have

$$\Delta F = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\Delta_{S^2} f \ . \tag{11.47}$$

Application 11.1. Electrostatic potential of a uniformly charged ball

We would like to solve the basic equation, $\Delta\phi = -4\pi\rho$ of electrostatics for a charge distribution

$$\rho(r) = \begin{cases} \rho_0 & \text{for} \quad r \leq R \\ 0 & \text{for} \quad r > R \end{cases} ,$$

where ρ_0 is a constant charge density and R a fixed radius, Since ρ only depends on the radius $r = |x|$ it is reasonable to look for a solution ϕ with the same property. From Eq. (11.45) the Laplacian in spherical coordinates acting on functions with no angular dependence is given by $\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)$.

For the range $r \leq R$ and $\phi = \phi(r)$ we have

$$\Delta\phi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right) = -4\pi\rho_0 \quad \Rightarrow \quad \phi_{\leq}(r) = \phi_0 - \frac{2\pi\rho_0}{3}r^2 ,$$

where ϕ_0 is an integration constant (and we have discarded a possible term proportional to $1/r$ by setting the other integration constant to zero so that $\phi_{\leq}$ is regular for $r \leq R$.)

For $r > R$, on the other hand, we find

$$\phi_{\geq}(r) = \frac{c}{r}$$

where c is an integration constant and the integration constant has been set to zero so that $\phi_{\geq} \to 0$ as $r \to \infty$. We should ensure the matching conditions $\phi_{\leq}(R) = \phi_{\geq}(R)$ and $\phi'_{\leq}(R) = \phi'_{\geq}(R)$ so that the potential and the electric field $E = -\nabla\phi = -\phi'(r)e_r$ (from Eq. (11.37)) are continuous at $r = R$ (Otherwise, $\Delta\phi$ has a large spike to $r = R$ which requires a large surface charge at $r = R$ - but the ρ we have assumed does not account for this.) Working out the matching conditions fixes the two integration constants as $\phi_0 = 2\pi\rho_0 R^2$ and $c = \frac{4\pi\rho_0}{3}R^3$. Inserting these expression back into the above solutions we find

$$\phi_{\leq}(r) = 2\pi\rho_0 R^2\left(1 - \frac{1}{3}\frac{r^2}{R^2}\right) , \qquad \phi_{\geq}(r) = \frac{4\pi\rho_0 R^3}{3}\frac{1}{r} .$$

12 Integration over manifolds

We have seen in Section 9 how to define the (Riemann) integral over $\mathbb{R}^n$ (or full-dimensional subsets thereof). Integrating over manifolds embedded in $\mathbb{R}^n$ is a different matter. In this case, the integral has to take into a account the 'geometry' of the manifold integrated over. Manifolds can be described by (local) parametrisations, as we have seen, but clearly one and the same manifolds can be described by different parametrisations (as illustrated in Example 7.1). A crucial requirement for how to define integrals over manifolds is that the value of the integral is independent of how the manifold is parametrised.

12.1 Definition of integral

Motivation

Consider a k-dimensional manifold $M \subset \mathbb{R}^n$ and let us, for now, focus on a specific chart $X : U \to V \cap M$, where $U \subset \mathbb{R}^k$ is open and we denote parameters by $t = (t_1, \ldots, t_k)^T \in U$, as usual. Our goal is to define the integral of a function $f : M \to \mathbb{R}$ over M which we write as

$$\int_M f \, d\sigma , \tag{12.1}$$

with an integration measure $d\sigma$ to be defined. For now, as we focus on one patch, we want to define the integral over $V \cap M$, rather than all of M. An obvious attempt would be to define this integral by a simple

integration over parameter space, that is,

$$\int_{V\cap M} f\, d\sigma := \int_U f(X(t))\, dt_1 \cdots dt_k \,, \tag{12.2}$$

or, more symbolically, define the measure as $d\sigma = dt_1 \cdots dt_k =: d^k t$. However, this comes into immediate conflict with our requirement for parametrisation independence. Consider a re-parametrisation diffeomorphism $T : \tilde{U} \to U$ and the modified parametrisation $\tilde{X} = X \circ T$ (and, in the following, we will revert to the usual abuse of notation and write $t = T(\tilde{t})$ as $t = t(\tilde{t})$, where $\tilde{t} \in \tilde{U}$ are the new parameters). From the transformation formula we have

$$d^k\tilde{t} = \left|\det\left(\frac{\partial\tilde{t}}{\partial t}\right)\right| d^k t \,. \tag{12.3}$$

while invariance of the integral (12.2) would require $d^k\tilde{t} = d^k t$. Clearly, Eq. (12.2) is not the correct definition as the value of the integral depends on the parametrisation. Somehow, we need to modify the measure such that the Jacobian which appears in Eq. (12.3) is cancelled. Let us first consider how the tangent vectors change under re-parametrisation (using matrix notation):.

$$\frac{\partial\tilde{X}}{\partial\tilde{t}}(\tilde{t}) = \frac{\partial X}{\partial t}(t)\frac{\partial t}{\partial\tilde{t}}(\tilde{t}) \,. \tag{12.4}$$

Now consider the induced metrics associated to these parametrisations, namely

$$\tilde{g} = \left(\frac{\partial\tilde{X}}{\partial\tilde{t}}\right)^T \frac{\partial\tilde{X}}{\partial\tilde{t}} \,, \qquad g = \left(\frac{\partial X}{\partial t}\right)^T \frac{\partial X}{\partial t} \,. \tag{12.5}$$

Inserting Eq. (12.4) into the above expression for $\tilde{g}$ leads to

$$\tilde{g} = \left(\frac{\partial X}{\partial t}\frac{\partial t}{\partial\tilde{t}}\right)^T \frac{\partial X}{\partial t}\frac{\partial t}{\partial\tilde{t}} = \left(\frac{\partial t}{\partial\tilde{t}}\right)^T g\, \frac{\partial t}{\partial\tilde{t}} \qquad \Rightarrow \qquad \sqrt{\det(\tilde{g})} = \left|\det\left(\frac{\partial t}{\partial\tilde{t}}\right)\right| \sqrt{\det(g)} \,. \tag{12.6}$$

Definition of integral for a single chart

We note that the Jacobian which appears in the last equation is the inverse of the one in Eq. (12.3). Hence, it follows that

$$\sqrt{\det(\tilde{g})}\, d^k\tilde{t} = \sqrt{\det(g)}\, d^k t \,, \tag{12.7}$$

and we see that choosing the integration measure $d\sigma = \sqrt{\det(g)}\, d^k t$ (instead of $d\sigma = d^k t$, as in Eq. (12.2)) does lead to parametrisation invariance.

Definition 12.1. *Consider a k-dimensional manifold $M \subset \mathbb{R}^n$ with a chart $X : U \to V \cap M$, where $U \subset \mathbb{R}^k$ is open with coordinates $t = (t_1, \ldots, t_k)^T \in U$. Let $g = \left(\frac{\partial X}{\partial t}\right)^T \frac{\partial X}{\partial t}$ be the induced metric. Then, a function $f : M \to \mathbb{R}$ is called integrable over $V \cap M$ if the function $t \mapsto \sqrt{\det(g)}\, f(X(t))$ is integrable over U and we define*

$$\int_{V\cap M} f\, d\sigma = \int_U d^k t\, \sqrt{\det(g)}\, f(X(t)) \,. \tag{12.8}$$

In short and symbolically this definition can be summarised by saying that the measure $d\sigma$ for integration over a k-dimensional manifold with parameters $t = (t_1, \ldots, t_k)^T$ is given by

$$d\sigma = d^k t\, \sqrt{\det(g)} \,, \tag{12.9}$$

where g is the induced metric. As we have seen, this definition guarantees the integral's independence from the choice of parametrisation. But it also has a geometric interpretation. We have

$$\sqrt{\det(g)} = \mathrm{vol}(P(v_1, \ldots, v_k)) \,, \tag{12.10}$$

where $P(v_1, \ldots, v_k)$ denotes the parallelepiped (10.12) spanned by the tangent vectors v_a. In short, the factor $\sqrt{\det(g)}$ in the measure (12.9) equals the volume of the parallelepiped generated by the standard tangent vectors $v_a = \frac{\partial X}{\partial t_a}$ and $\sqrt{\det(g)}\,\Delta t_1 \cdots \Delta t_k$ is the volume of a small parallelepiped which captures the local volume of the manifold (see Fig. 22).

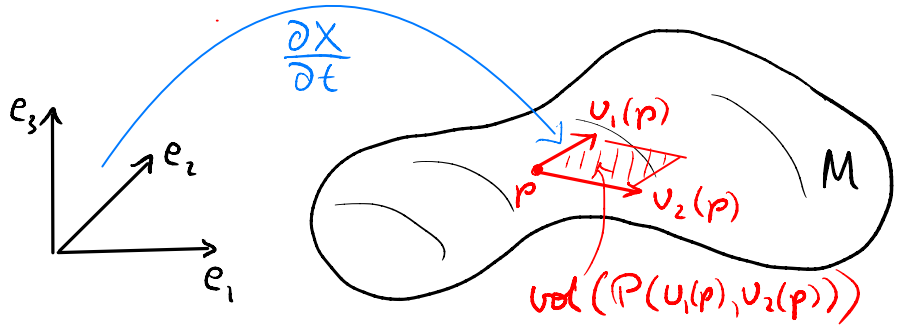


Figure 22: Geometrical interpretation of the integration measure in Eq. (12.9) over manifolds M. The standard tangent vectors $v_a(p)$ at $p \in M$ are obtained from the standard unit vectors e_i in the embedding space $\mathbb{R}^n$ via Eq. (11.7). The measure is proportional to the area of the parallelepiped (parallelogram) defined by the vectors $v_a(p)$, in accordance with Eq. (12.10).

General definition of integral

In many practical cases, the definition of the integral in Def. 12.1 based on a single chart is sufficient. However, for a general definition we need to consider a k-dimensional manifold with multiple charts $X_{(\alpha)} : U_{(\alpha)} \to W_{(\alpha)} := V_{(\alpha)} \cap M$, where $\alpha = 1, \ldots, m$, which cover the entire manifold, that is $M = \bigcup_\alpha W_{(\alpha)}$. Roughly speaking, we should apply Eq. (12.8) to each chart and then sum up the results. The problem is that the sets $W_{(\alpha)}$ can (and to some extent will) overlap so simply summing up the various contributions leads to 'double counting'. The technical device to deal with this problem is called a **partition of unity**.

A partition of unity relates to a covering of our manifold M by sets $W_{(\alpha)}$ and it consists of functions $\chi_{(\alpha)} : M \to \mathbb{R}$, where $\alpha = 1, \ldots, m$, with the following properties.

$$\left.\begin{array}{rll} \chi_{(\alpha)}(p) \in [0,1] & \text{for all} & p \in M \\ \chi_{(\alpha)}(p) = 0 & \text{for all} & p \in (M \setminus W_{(\alpha)}) \\ \sum_{\alpha=1}^m \chi_{(\alpha)}(p) = 1 & \text{for all} & p \in M \,. \end{array}\right\} \tag{12.11}$$

Using such a partition of unity , the integral over the full manifold can then be defined as follows.

Definition 12.2. *Consider a k-dimensional manifold $M \subset \mathbb{R}^n$ with charts $X_{(\alpha)} : U_{(\alpha)} \to W_{(\alpha)} := V_{(\alpha)} \cap M$, where $\alpha = 1, \ldots, m$, which cover the entire manifold, that is $M = \bigcup_\alpha W_{(\alpha)}$, induced metrics $g_{(\alpha)}$ on each chart and a related partition of unity $\chi_{(\alpha)}$. A function $f : M \to \mathbb{R}$ is called integrable over M if the functions $\chi_{(\alpha)} f$ are integrable over $W_{(\alpha)}$ and the integral over M is defined as*

$$\int_M f \, d\sigma = \sum_{\alpha=1}^m \int_{U_{(\alpha)}} d^k t \, \sqrt{\det(g_{(\alpha)})}\, \chi_{(\alpha)}(X_{(\alpha)}(t))\, f(X_{(\alpha)}(t)) \,. \tag{12.12}$$

The volume of a k-dimensional manifold M (by which we mean 'length' for $k=1$, 'area' for $k=2$, 'volume' for $k=3$ and the higher-dimensional generalisation of volume for $k>3$) is obtained by integrating over the function $f=1$, so

$$\text{vol}(M) = \int_M d\sigma \ . \tag{12.13}$$

Example 12.1. Length of a circle and area of two-dimensional sphere
Question: Compute the length of a unit circle $S^1 \subset \mathbb{R}^2$ and the area of a unit sphere $S^2 \subset \mathbb{R}^3$, using the usual angular parametrisations of these manifolds. For S^2, verify that the same answer is obtained using the parametrisation $X(x,y) = (x, y, \pm\sqrt{1-x^2-y^2})^T$, where $(x,y)^T \in B_1^2$ and the two signs give the upper and lower hemisphere.

Answer: For the unit circle S^1 we use the standard parametrisation $X(\varphi) = (\cos(\varphi), \sin(\varphi))^T$ (where $\varphi \in [0, 2\pi)$) with tangent vector

$$\frac{dX}{d\varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \ . \tag{12.14}$$

and $\sqrt{\det(g)} = \left|\frac{dX}{d\varphi}\right| = 1$. Hence, the measure on S^1 (parametrised by an angle as above) is $d\sigma = d\varphi$ and for its length we have

$$\text{vol}(S^1) = \int_{S^1} d\sigma = \int_0^{2\pi} d\varphi = 2\pi \ , \tag{12.15}$$

as expected.

For the case of the two-sphere S^2 with standard angular parametrisation the induced metric has already been computed in Eq. (11.25). This result shows that $\sqrt{\det(g)} = \sin(\theta)$ and, hence, the measure for S^2 is $d\sigma = \sin(\theta) d\theta\, d\varphi$. For its area we find

$$\text{vol}(S^2) = \int_{S^2} d\sigma = \int_0^{\pi} d\theta\, \sin(\theta) \int_0^{2\pi} d\varphi = 4\pi \ , \tag{12.16}$$

as expected.

For the other parametrisation of S^2, given by $X(x,y) = (x, y, \pm\sqrt{1-x^2-y^2})^T$ we find the tangent vector

$$\frac{\partial X}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \pm\frac{x}{\sqrt{1-x^2-y^2}} \end{pmatrix} \ , \qquad \frac{\partial X}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \pm\frac{y}{\sqrt{1-x^2-y^2}} \end{pmatrix} \tag{12.17}$$

and the induced metric

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial x}\right|^2 & \frac{\partial X}{\partial x} \cdot \frac{\partial X}{\partial y} \\ \frac{\partial X}{\partial x} \cdot \frac{\partial X}{\partial y} & \left|\frac{\partial X}{\partial y}\right|^2 \end{pmatrix} = \frac{1}{1-x^2-y^2} \begin{pmatrix} 1-y^2 & xy \\ xy & 1-x^2 \end{pmatrix} \quad \Rightarrow \quad \sqrt{\det(g)} = \frac{1}{\sqrt{1-x^2-y^2}} \ .$$

Hence, the measure for this parametrisation is $d\sigma = \frac{dx\, dy}{\sqrt{1-x^2-y^2}} = \frac{r dr\, d\varphi}{\sqrt{1-r^2}}$, where I have converted to

two-dimensional polar coordinates in the x-y plane in the last step. Hence, we have for the area

$$\text{vol}(S^2) = \int_{S^2} d\sigma = 2\int_0^{2\pi} d\varphi \int_0^1 \frac{r dr}{\sqrt{1-r^2}} = 4\pi \int_0^1 \frac{r dr}{\sqrt{1-r^2}} = 4\pi \ .$$

(Note, the second integral is only over one hemisphere — hence the factor of 2.)

Example 12.2. Integration for n-dimensional spherical coordinates
Consider $\mathbb{R}^n$ with Cartesian coordinates $x = (x_1, \ldots, x_n)^T$, the radius $r = |x|$, 'angles' $t = (t_1, \ldots, t_{n-1})^T$ and angular functions $X = (X_1, \ldots, X_n)$ of t such that $x_i = rX_i(t)$, for $i = 1, \ldots, n$. Show that the Jacobian $\det\left(\frac{\partial x}{\partial(r,t)}\right)$ can be expressed in terms of the induced metric $g_{ab} = \frac{\partial X}{\partial t_a} \cdot \frac{\partial X}{\partial t_b}$ on the sphere S^{n-1} with parametrisation $t \mapsto X(t)$. Use this result to derive a relation between the volumes of B_1^n and S^{n-1}.
Answer: The Jacobi matrix has the form $\frac{\partial x}{\partial(r,t)} = \left(X, r\frac{\partial X}{\partial t_1}, \ldots, r\frac{\partial X}{\partial t_{n-1}}\right)$ and, since $|X(t)|^2 = 1$ it follows by differentiating with respect to t_a that $X \cdot \frac{\partial X}{\partial t_a} = 0$. Hence, we have

$$\left(\frac{\partial x}{\partial(r,t)}\right)^T \left(\frac{\partial x}{\partial(r,t)}\right) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 g \end{pmatrix}$$

$$\Rightarrow \quad d^n x = \left|\det\left(\frac{\partial x}{\partial(r,t)}\right)\right| dr\, d^{n-1}t = r^{n-1}\sqrt{\det(g)}\, dr\, d^{n-1}t = r^{n-1} dr\, d\sigma$$

where $d\sigma = \sqrt{\det(g)}\, d^{n-1}t$ is the measure associated to the sphere S^{n-1} with parametrisation X. We can use this result to work out the volume of the n-dimensional unit ball B_1^n in terms of the volume of S^{n-1}.

$$\text{vol}(B_1^n) = \int_{B_1^n} d^n x = \int_0^1 r^{n-1}\, dr \int_{S^{n-1}} d\sigma = \frac{1}{n}\text{vol}(S^{n-1}) \ .$$

Inserting the results from Eq. (10.29) for the volume of the n-dimensional unit ball B_1^n gives for the volume of the sphere

$$\text{vol}(S^{2m}) = (2m+1)\,\text{vol}(B_1^{2m+1}) = 2\frac{(2\pi)^m}{(2m-1)!!}$$
$$\text{vol}(S^{2m-1}) = 2m\,\text{vol}(B_1^{2m}) = 2\frac{\pi^m}{(m-1)!} \ .$$

12.2 Integration for special cases and dimensions

Our prescription for how to integrate over manifolds embedded in $\mathbb{R}^n$ has so far been completely general. For particular dimensions n of the embedding space or the dimension k of the manifold, or when the manifold is represented as a zero locus, the general integration measure $d\sigma = \sqrt{\det(g)}\, d^k t$ can sometimes be written more explicitly or in a different form. We will now explore these special cases.

Integral over curves
Consider a curve $C \subset \mathbb{R}^n$ with parametrisation $X : [a, b] \to \mathbb{R}^n$, parameter $t \in [a, b]$ and tangent vectors $\frac{dX}{dt}$. In this case the induced metric on the curve is the 1×1 matrix $g = \left(\left|\frac{dX}{dt}\right|^2\right)$ with associated measure

$$d\sigma = \left|\frac{dX}{dt}\right| dt = \sqrt{\sum_{i=1}^n \left(\frac{dX_i}{dt}\right)^2}\, dt \tag{12.18}$$

Hence, the integral of a function over a curve and the length of a curve are given by

$$\int_C f\,d\sigma = \int_a^b dt\,\left|\frac{dX}{dt}\right| f(X(t))\,, \qquad \mathrm{vol}(C) = \int_C d\sigma = \int_a^b dt\,\left|\frac{dX}{dt}\right|\,. \tag{12.19}$$

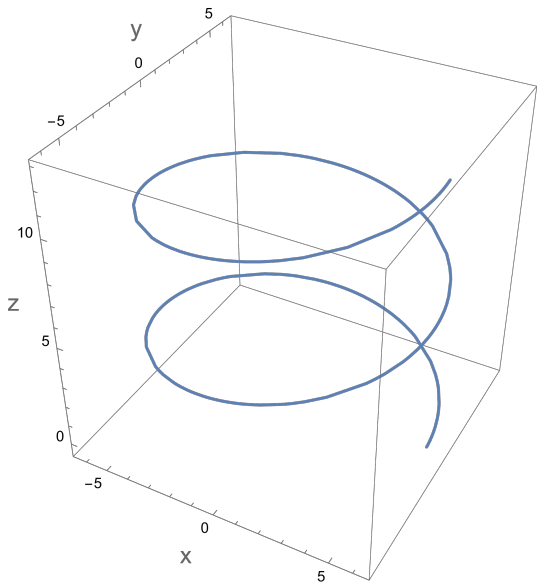


Figure 23: The curve $X(t) = (R\cos(t), R\sin(t), t)^T$ for $t \in [0, 4\pi]$, drawn for $R = 6$.

Example 12.3. Length of a spiral
Question: Find the length of the curve C (a spiral, see Fig. 23) with parametrisation $X : [0, 4\pi] \to \mathbb{R}^3$ given by $X(t) = (R\cos(t), R\sin(t), t)^T$.
Answer: The tangent vector of this curve and the measure are

$$\frac{dX}{dt} = \begin{pmatrix} -R\sin(t) \\ R\cos(t) \\ 1 \end{pmatrix} \quad \Rightarrow \quad d\sigma = \left|\frac{dX}{dt}\right| dt = \sqrt{1+R^2}\,dt\,.$$

Hence, the length of the spiral is

$$\mathrm{vol}(C) = \int_C d\sigma = \sqrt{1+R^2}\int_0^{4\pi} dt = 4\pi\sqrt{1+R^2}\,.$$

Example 12.4. Length of a cycloid
Question: A cycloid is a curve $C : \mathbb{R} \to \mathbb{R}^2$ given by a parametrisation $X(t) = (t - \sin(t), 1 - \cos(t))^T$ (see Fig. 24). (The cycloid is the curve traced out by a point on the boundary of a rolling disk.) Find the length of one period, C_1, obtained from parameter values $t \in [0, 2\pi]$.
Answer: The tangent vector of the cycloid and its length are given by

$$\frac{dX}{dt} = \begin{pmatrix} 1-\cos(t) \\ \sin(t) \end{pmatrix} \quad \Rightarrow \quad \left|\frac{dX}{dt}\right|^2 = 2 - 2\cos(t) = 4\sin^2\left(\frac{t}{2}\right)\,.$$

Hence, we find for the length of one period

$$\int_{C_1} d\sigma = 2\int_0^{2\pi} dt\, \sin\left(\frac{t}{2}\right) = 8\,.$$

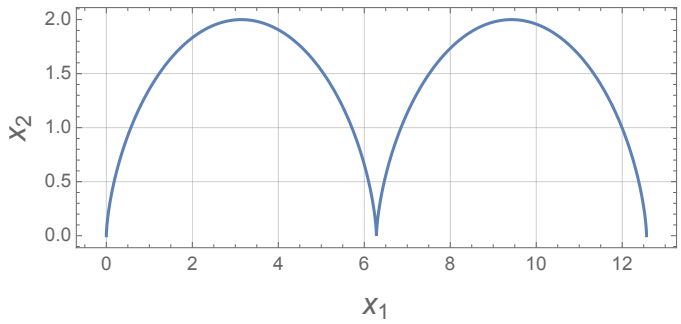


Figure 24: The cycloid parametrised by $X(t) = (t - \sin(t), 1 - \cos(t))^T$ in the range $t \in [0, 4\pi]$.

In Section 7.3 we have already defined an integral of a vector field over a curve. Note that this integral is quite different from the integral of functions over curves defined in Eq. (12.19). For comparison, let us write the two integrals side by side:

$$\int_C A \cdot dx = \int_a^b dt\, \frac{dX}{dt} \cdot A(X(t))\,, \qquad \int_C f\, d\sigma = \int_a^b dt\, \left|\frac{dX}{dt}\right| f(X(t))\,. \tag{12.20}$$

Integrating a vector field over a curve requires the tangent vectors to the curve but it does not require the metric g. Metric-independent integrals such as this are sometimes referred to as 'topological' integrals. On the other hand, to integrate a function we require the metric (which here appears in the form $d\sigma = \sqrt{\det(g)}\, dt = \left|\frac{dX}{dt}\right| dt$). We can convert between the two integrals by choosing a special function $f(X(t)) = A(X(t)) \cdot \tau(t)$ in the integral on the RHS of Eq. (12.20), where $\tau(t) = \left|\frac{dX}{dt}(t)\right|^{-1} \frac{dX}{dt}(t)$ are the unit tangent vectors to the curve. Another way of saying this is that the measures in the integrals (12.20) are related by

$$dx = \tau\, d\sigma\,, \tag{12.21}$$

where τ is the unit tangent vector to the curve and $d\sigma = \left|\frac{dX}{dt}\right| dt$ is the measure on the curve. Note that there are two choices for the unit tangent vector and this marks out the **orientation** of the curve (which way along the curve we go). This has already been discussed in Section 7.3. As Eq.(12.21) indicates, the curve integral changes sign if we change the orientation by transforming $\tau \mapsto -\tau$.

Integral over surfaces in $\mathbb{R}^3$

We would like to consider a surface M with a (local) parametrisation $X : U \to V \cap M$ and parameters $t = (t_1, t_2)^T \in U$ embedded in $\mathbb{R}^3$ with coordinates $(x_1, x_2, x_3) = (x, y, z)$. The induced metric on this surface is

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial t_1}\right|^2 & \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} \\ \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} & \left|\frac{\partial X}{\partial t_2}\right|^2 \end{pmatrix}$$

For the measure this means

$$\det(g) = \left|\frac{\partial X}{\partial t_1}\right|^2 \left|\frac{\partial X}{\partial t_2}\right|^2 - \left(\frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2}\right)^2 = \left|\frac{\partial X}{\partial t_1} \times \frac{\partial X}{\partial t_2}\right|^2 = |N|^2 \quad \Rightarrow \quad d\sigma = |N|\, dt_1\, dt_2\,, \tag{12.22}$$

where N is the standard normal vector to the plane, as defined in Eq. (11.12).

A surface $M \subset \mathbb{R}^3$ is often more convenient represented (locally) as the zero locus of a function, $M \cap V = \{(x,y,z)^T \in V \,|\, f(x,y,z) = 0\}$. We would like to express the integration measure $d\sigma$ on M directly in terms of this defining function f. First note, the maximal rank requirement in Def. 11.1 implies that $\nabla f \neq 0$, so that at least one of the partial derivatives of f must be non-zero. For definiteness, let us assume that $\frac{\partial f}{\partial z} \neq 0$. The implicit function Theorem 6.3 then implies we can (locally) solve $f(x,y,z) = 0$ for z, so there is a function g such that $f(x,y,g(x,y)) = 0$. The point of doing this is that we can now write down a local parametrisation of M, using the Cartesian coordinates $(x,y) = (t_1,t_2)$ as parameters, namely

$$X(x,y) = \begin{pmatrix} x \\ y \\ g(x,y) \end{pmatrix} \quad \Rightarrow \quad \frac{\partial X}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial g}{\partial x}(x,y) \end{pmatrix}, \quad \frac{\partial X}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial g}{\partial y}(x,y) \end{pmatrix}. \tag{12.23}$$

This leads to the induced metric and measure

$$g = \begin{pmatrix} |\frac{\partial X}{\partial x}|^2 & \frac{\partial X}{\partial x} \cdot \frac{\partial X}{\partial y} \\ \frac{\partial X}{\partial x} \cdot \frac{\partial X}{\partial y} & \left|\frac{\partial X}{\partial y}\right|^2 \end{pmatrix} = \begin{pmatrix} 1 + \left|\frac{\partial g}{\partial x}\right|^2 & \frac{\partial g}{\partial x}\frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial x}\frac{\partial g}{\partial y} & 1 + \left|\frac{\partial g}{\partial y}\right|^2 \end{pmatrix}$$

$$\Rightarrow \quad d\sigma = \sqrt{\det(g)}\, dx\, dy = \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}\, dx\, dy\,. \tag{12.24}$$

We still need to write this in terms of the original defining function f and this can be done by partially differentiating the relation $f(x,y,g(x,y)) = 0$ with respect to x and y. This gives

$$\left.\begin{array}{rcl} \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial x} &=& 0 \\ \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial y} &=& 0 \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{rcl} \frac{\partial g}{\partial x} &=& -\frac{\partial f}{\partial x}/\frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial y} &=& -\frac{\partial f}{\partial y}/\frac{\partial f}{\partial z} \end{array}\right.$$

Inserting these results into the measure in Eq. (12.24) we find

$$d\sigma = \left(\frac{|\nabla f|}{\left|\frac{\partial f}{\partial z}\right|}\right)_{z=g(x,y)} dx\, dy\,, \tag{12.25}$$

where the subscript indicates that we should replace z with $g(x,y)$ in the bracketed expression. This is the desired expression for the integration measure over a surface in $\mathbb{R}^3$ in terms of its defining function f. (Of course we could have solved $f(x,y,z) = 0$ for x or y instead of for z and in these cases we would have obtained a result analogous to Eq. (12.25) but with the three coordinates permuted accordingly.)

Example 12.5. Area of surfaces defined as a zero locus
Question: Apply Eq. (12.25) to compute the area of a unit hemisphere $H \subset \mathbb{R}^3$ defined as the zero locus of the function $f(x,y,z) = x^2 + y^2 + z^2 - 1$. Further, compute the area of a surface S given as the zero locus of the function $f(x,y,z) = \frac{1}{2}x^2 + y - z$, where $x, y \in [0,1]$.

Answer: For the (upper) hemisphere we have $\nabla f = 2(x,y,z)$, $\frac{\partial f}{\partial z} = 2z$ and $z = \sqrt{1-x^2-y^2}$ so the measure is given by

$$d\sigma = \frac{1}{z}\sqrt{x^2+y^2+z^2}\, dx\, dy = \frac{dx\, dy}{\sqrt{1-x^2-y^2}} = \frac{r\, dr\, d\varphi}{\sqrt{1-r^2}}\,,$$

where we have converted to two-dimensional polar coordinates in the last step. Hence, the area of

the hemisphere is

$$\mathrm{vol}(H) = \int_0^{2\pi} d\varphi \int_0^1 \frac{r\,dr}{\sqrt{1-r^2}} = 2\pi \ .$$

For the second surface, S, we have $\nabla f = (x, 1, -1)$, $\frac{\partial f}{\partial z} = -1$ and $z = y + \frac{1}{2}x^2$, so that the measure is

$$d\sigma = \sqrt{1+x^2}\, dx\, dy$$

For the area of H this gives

$$\mathrm{vol}(S) = \int_0^1 dy \int_0^1 dx\, \sqrt{1+x^2} = \frac{1}{2}(\sqrt{2} + \mathrm{asinh}(1)) \simeq 1.15 \ .$$

Example 12.6. Area of rotational surfaces
Question: For a function $\rho : [a, b] \to \mathbb{R}$, the rotational surface S_ρ associated to ρ is defined by

$$S_\rho = \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x^2 + y^2 = \rho(z)^2 \text{ and } z \in [a, b]\} \ . \tag{12.26}$$

Find a formula for the area of S_ρ in terms of the function ρ.
Answer: We can write down a parametrisation $X : [a, b] \times [0, 2\pi) \to \mathbb{R}^3$ for S_ρ by

$$X(z,\varphi) = \begin{pmatrix} \rho(z)\cos(\varphi) \\ \rho(z)\sin(\varphi) \\ z \end{pmatrix} \quad \Rightarrow \quad \frac{\partial X}{\partial z} = \begin{pmatrix} \rho'(z)\cos(\varphi) \\ \rho'(z)\sin(\varphi) \\ 1 \end{pmatrix} , \quad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -\rho(z)\sin(\varphi) \\ \rho(z)\cos(\varphi) \\ 0 \end{pmatrix}$$

Note that the two tangent vectors are orthogonal so for the induced metric we have

$$g = \mathrm{diag}\left(\left|\frac{\partial X}{\partial z}\right|^2 , \left|\frac{\partial X}{\partial \varphi}\right|^2 \right) = \mathrm{diag}(1 + \rho'(z)^2, \rho(z)^2)$$
$$\Rightarrow \quad d\sigma = \sqrt{\det(g)}\, dz\, d\varphi = \rho(z)\sqrt{1+\rho'(z)^2}\, dz\, d\varphi$$

Hence, the area of the rotational surface is given by

$$\mathrm{vol}(S_\rho) = \int_{S_\rho} d\sigma = \int_0^{2\pi} d\varphi \int_a^b dz\, \rho(z)\sqrt{1+\rho'(z)^2} = 2\pi \int_a^b dz\, \rho(z)\sqrt{1+\rho'(z)^2} \ .$$

Integrating vector fields over surfaces in $\mathbb{R}^3$
So far we have seen how to integrate functions over surfaces but what about vector fields? Recall, for a surface $M \subset \mathbb{R}^3$ with (local) parametrisation $X : U \to M \cap V$ and parameters $t = (t_1, t_2)^T$, we have the tangent vectors $\frac{\partial X}{\partial t_1}$ and $\frac{\partial X}{\partial t_2}$, the normal vector $N = \pm\frac{\partial X}{\partial t_1} \times \frac{\partial X}{\partial t_2}$ of M and the unit normal vector $n = \frac{N}{|N|}$. To get to a sensible definition of how to integrate a vector field over a surface we can take inspiration from curve integrals, as discussed in Section 7.3. Integrals of vector fields along curves were designed to measure the total flow of the vector field along the curve. The obvious analogy is to ask the integral of vector fields over surfaces to measure the total flux of the vector field through the surface. This means, for a three-dimensional vector field A, we should consider the integrand $A(X(t_1, t_2)) \cdot \hat{n}$ which represents the component of the vector field normal to the surface ('passing through') and integrate over the surface (see Fig. 25). In other words, we should define a measure

$$dS = \hat{n}\, d\sigma = N\, dt_1 dt_2 = \pm\frac{\partial X}{\partial t_1} \times \frac{\partial X}{\partial t_2}\, dt_1 dt_2 \ . \tag{12.27}$$

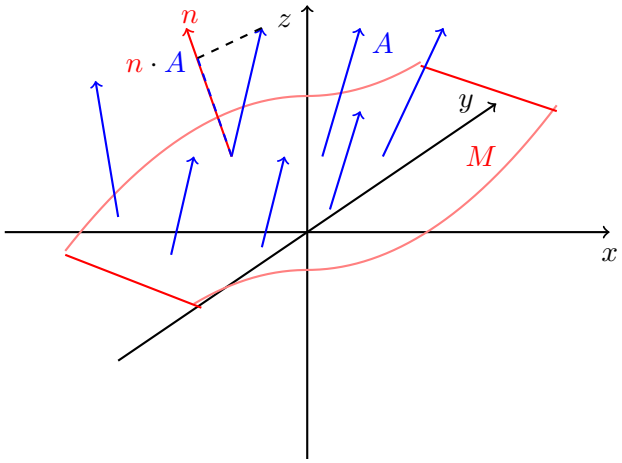


Figure 25: Integrating a vector field A over a manifold M. The vector field is projected into the direction of the unit normal, n, to M and the component, $n \cdot A$ orthgonal to M is then integrated over the manifold. The value of the integral is interpreted as the total flux of A through M.

If the surface is (locally) given as the zero locus of a function f the measure $d\sigma$ can be written as in Eq. (12.25) while the unit normal vector is $\hat{n} = \pm\frac{\nabla f}{|\nabla f|}$. Hence, in this case, using as parameters $(x, y) = (t_1, t_2)$, we have

$$dS = \hat{n}\, d\sigma = \pm \frac{\nabla f}{\left|\frac{\partial f}{\partial z}\right|} dx\, dy \,. \tag{12.28}$$

More precisely, we call the vector field A integrable over an orientable surface M if the scalar function $A \cdot n$ is integrable over M and, in this case, we define the integral as

$$\int_M A \cdot dS := \int_M A \cdot \hat{n}\, d\sigma \,, \tag{12.29}$$

where $\hat{n} : M \to \mathbb{R}^n$ is a continuous unit normal vector to M. Note we require the surface to be orientable so that a continuous unit normal vector $\hat{n} : M \to \mathbb{R}^n$ does indeed exists. The integral depends on the orientation, that is on the direction into which this normal vector points and it changes sign as we change the orientation by $\hat{n} \mapsto -\hat{n}$, Also note that such integrals, just as integrals of vector fields along curves, are 'topological', that is, they do not involve the induced metric.

Example 12.7. Flux of a vector field through S^2
Question: Work out the flux of the three-dimensional vector field $E(x) = \frac{x}{r^3}$, where $x = (x_1, x_2, x_3)^T$ and $r = |x|$, through the two-sphere S^2_R with radius R.

Answer: We can parametrise the two-sphere in various ways but, to be concrete, let us choose the standard angular parametrisation $X(\theta, \varphi) = R\, e_r$, where we use the ortho-normal spherical basis vectors $(e_r, e_\theta, e_\varphi)$ defined in Eq. (11.36). For the tangent vectors, the normal vector and the measure (using Eq. (12.27)) we have

$$\frac{\partial X}{\partial \theta} = R\, e_\theta \,, \quad \frac{\partial X}{\partial \varphi} = R \sin(\theta)\, e_\varphi \,, \quad N = \frac{\partial X}{\partial \theta} \times \frac{\partial X}{\partial \varphi} = R^2 \sin(\theta)\, e_\theta \times e_\varphi = R^2 \sin(\theta) e_r$$
$$\Rightarrow \quad dS = N\, d\theta\, d\varphi = R^2 \sin(\theta)\, e_r\, d\theta\, d\varphi \,.$$

On the two-sphere the vector field evaluates to $E(X(\theta,\varphi)) = \frac{e_r}{R^2}$ so we have $E(X(\theta,\varphi)) \cdot dS = \sin(\theta)\, d\theta\, d\varphi$ and

$$\int_{S^2_R} E \cdot dS = \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta\, \sin(\theta) = 4\pi \; .$$

Note, this result is independent of the radius R since the area of the sphere increases with R^2 and the vector field decreases with $1/R^2$.

Example 12.8. Vector field flux through a parabolic surface
Question: In $\mathbb{R}^3$ with coordinates $(x_1, x_2, x_3) = (x, y, z)$ we have a surface $M \subset \mathbb{R}^3$ defined as the zero locus of the function f with $f(x, y, z) = \frac{1}{2}(x^2 + y^2) - z$, where $x, y \in [-1, 1]$. Compute the flux of the vector field A with $A(x, y, z) = (x, y, -z)^T$ through M.

Solution: Using Eq. (12.28) we find for the measure

$$\nabla f = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \; , \quad \frac{\partial f}{\partial z} = -1 \quad \Rightarrow \quad dS = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} dx\, dy \quad \Rightarrow \quad A \cdot dS = x^2 + y^2 - z = \frac{1}{2}(x^2 + y^2) \; .$$

For the flux of A through M this means

$$\int_M A \cdot dS = \frac{1}{2} \int_{-1}^{1} dx \int_{-1}^{1} dy\, (x^2 + y^2) = \frac{4}{3} \; .$$

Example 12.9. Flux of a vector field through the surface of a cylinder
Question A surface of a cylinder with radius R and height h is defined as

$$C = \{(R\sin(\varphi), R\cos(\varphi), z)^T \in \mathbb{R}^3 \,|\, z \in [0, h] \text{ and } \varphi \in [0, 2\pi)\} \; .$$

Compute a measures $d\sigma$, a unit normal vector n and the measure dS on the surface of the cylinder. Use these results to verify that the area comes out as expected. Then, find the flux of the vector field $A : \mathbb{R}^3 \to \mathbb{R}^3$ given by $A(x, y, z) = (z^2 x, z^2 y, z^2)^T$ through the surface of the cylinder.

Answer: The obvious parametrisation is

$$X(\varphi, z) = \begin{pmatrix} R\cos(\varphi) \\ R\sin(\varphi) \\ z \end{pmatrix} \quad \Rightarrow \quad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -R\sin(\varphi) \\ R\cos(\varphi) \\ 0 \end{pmatrix} \; , \quad \frac{\partial X}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \; .$$

Hence, the induced metric is

$$g = \begin{pmatrix} \left|\frac{\partial X}{\partial \varphi}\right|^2 & \frac{\partial X}{\partial \varphi} \cdot \frac{\partial X}{\partial Z} \\ \frac{\partial X}{\partial \varphi} \cdot \frac{\partial X}{\partial Z} & \left|\frac{\partial X}{\partial z}\right|^2 \end{pmatrix} = \begin{pmatrix} R^2 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad d\sigma = \sqrt{\det(g)}\, d\varphi\, dz = R\, d\varphi\, dz \; .$$

and we have the unit normal vector and the measure dS given by

$$\hat{n} = \frac{1}{R} \frac{\partial X}{\partial \varphi} \times \frac{\partial X}{\partial z} = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \\ 0 \end{pmatrix} \quad \Rightarrow \quad dS = \hat{n}\, d\sigma = R \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \\ 0 \end{pmatrix} d\varphi\, dz \; .$$

The area of the cylinder surface is then

$$\text{vol}(C) = \int_C d\sigma = R\int_0^{2\pi} d\varphi \int_0^h dz = 2\pi Rh \ ,$$

as expected. The vector field A evaluated on the cylinder's surface is

$$A(X(\varphi, z)) = \begin{pmatrix} Rz^2\cos(\varphi) \\ Rz^2\sin(\varphi) \\ z^2 \end{pmatrix} \ ,$$

so that the flow through the surface is given by

$$\int_C A\cdot dS = R^2\int_0^{2\pi} d\varphi \int_0^h dz\, z^2 = \frac{2\pi}{3}R^2h^3 \ .$$

Integrating over hyper-surfaces

We would like to generalise the previous results for integrating over surfaces to hyper-surfaces, that is, to $n-1$–dimensional manifolds M embedded in $\mathbb{R}^n$ with coordindates $x = (x_1, \ldots, x_n)^T$. Given a local parametrisation $X : U \to V \cap M$ of M with parameters $t = (t_1, \ldots, t_{n-1})^T$ we can write down a normal vector N to M as in Eq. (11.10), that is,

$$N_i = \pm\epsilon_{ii_1\cdots i_{n-1}} \frac{\partial X^{i_1}}{\partial t^1}\cdots\frac{\partial X^{i_{n-1}}}{\partial t^{n-1}} \tag{12.30}$$

As in the surface case we would like to express the measure in terms of this normal vector and, to this end, we require a statement from linear algebra (see, for example, Ref. [5]). For an $n\times(n-1)$ matrix M, denote by $M_{(i)}$ the $(n-1)\times(n-1)$ matrices obtained by dropping the i^{th} row from M. Then

$$\det(M^TM) = \sum_{i=1}^n \det\left(M_{(i)}\right)^2 \ . \tag{12.31}$$

We can apply this result to re-write the induced metric

$$g = \left(\frac{\partial X}{\partial t}\right)^T\left(\frac{\partial X}{\partial t}\right) \quad \overset{\text{Eq. (12.31)}}{\Rightarrow} \quad \det(g) = \sum_{i=1}^n \left(\det\left(\frac{\partial X}{\partial t}\right)_{(i)}\right)^2 = \sum_{i=1}^n N_i^2 \tag{12.32}$$

Hence, we find the integration measure over a hyper-surface can be written as

$$d\sigma = \sqrt{\det(g)}\, dt_1\cdots dt_{n-1} = |N|\, dt_1\cdots dt_{n-1} \ , \tag{12.33}$$

where the normal vector N is defined in Eq. (12.30). Note this is the obvious generalisation of the measure in Eq. (12.22) for surfaces.

As before, the hyper-surface M might be (locally) defined by the zero locus of a function f.Let us split coordinates as $x = (t, z)$, where $t \in \mathbb{R}^{n-1}$ and $z \in \mathbb{R}$ and assume that $\frac{\partial f}{\partial z} \neq 0$. (As discussed before, for M to be a manifold it is requires that $\nabla f \neq 0$ so at least one partial derivative needs to be non-zero. For convenience we assume this is the cases for the one with respect to z.) We can then locally solve the equation $f(t, z) = 0$ for z in terms of t so that $f(t, g(t)) = 0$. As in the surface case, this allows us to write down a parametrisation of M, namely

$$X(t) = \begin{pmatrix} t \\ g(t) \end{pmatrix} \quad \Rightarrow \quad \frac{\partial X}{\partial t_a} = \begin{pmatrix} e_a \\ \frac{\partial g}{\partial t_a} \end{pmatrix} \ ,$$

where e_a are the standard unit vectors on $\mathbb{R}^{n-1}$. Hence, the induced metric and measure, relative to this parametrisation is

$$g_{ab} = \frac{\partial X}{\partial t_a} \cdot \frac{\partial X}{\partial t_b} = \delta_{ab} + \frac{\partial g}{\partial t_a}\frac{\partial g}{\partial t_b} \quad \Rightarrow \quad d\sigma = \sqrt{\det(g)}\, d^{n-1}t = \sqrt{1 + \sum_{a=1}^{n-1}\left(\frac{\partial g}{\partial t_a}\right)^2}\, d^{n-1}t$$

Differentiating $f(t, g(t)) = 0$ with respect to t_a gives $\frac{\partial f}{\partial t_a} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial t_a} = 0$ and, hence, $\frac{\partial g}{\partial t_a} = -\frac{\partial f}{\partial t_a}/\frac{\partial f}{\partial z}$. Inserting this into the above expression for the measure we find the final result

$$d\sigma = \frac{|\nabla f|}{\left|\frac{\partial f}{\partial z}\right|}\, d^{n-1}t\ , \tag{12.34}$$

which is the obvious generalisation of the surface measure in Eq. (12.28).

Integrating vector fields over hyper-surfaces

Generalising the rules for integrating a vector field over a surfaces to a hyper-surface M is now straightforward. We simply use the same expression for the measure

$$dS = \hat{n}\, d\sigma = N\, d^{n-1}t\ , \tag{12.35}$$

where $\hat{n} = \frac{N}{|N|}$ is the unit normal vector to M, the normal vector N is defined in Eq. (12.30) and the scalar measure $d\sigma$ is as in Eq. (12.33). If the hyper-surface is given by the zero locus of a function f (using $\hat{n} = \pm\frac{\nabla f}{|\nabla f|}$) this can also be written as

$$dS = \hat{n}\, d\sigma = \pm\frac{\nabla f}{\left|\frac{\partial f}{\partial z}\right|}\, d^{n-1}t \tag{12.36}$$

as is evident by inserting Eq. (12.34). Note, the interpretation of the integral

$$\int_M A \cdot dS = \int_M A \cdot \hat{n}\, d\sigma \tag{12.37}$$

of an n-dimensional vector field A over a hyper-surface M is the same as in the surface case. The integral measures the flux of A through M.

13 The integral theorem of Gauß

The fundamental theorem of (one-dimensional) integration states that

$$\int_{[a,b]} dx\, f'(x) = f(b) - f(a)\ . \tag{13.1}$$

We can trade the integral for the derivative and remain with the RHS of Eq. (13.1) which can be interpreted as a zero-dimensional integral over the boundary $\partial[a,b] = \{a, b\}$ of the interval. As we will see this generalises to higher dimensions: A derivate (of a certain type) can be used to reduce the dimension of an integral by one and the new integration domain is the boundary of the original one. In this section, we will look at Gauß's theorem which relates integrals (which involve vector fields) over n-dimensional domains $M \subset \mathbb{R}^n$ to $(n-1)$–dimensional integrals over the boundary ∂M which (under suitable conditions) is a hyper-surface in $\mathbb{R}^n$. To formulate this adequately we need a few new ideas, notably the idea of a **smooth boundary** and the **outer unit normal vector** to a boundary.

13.1 Compact sets with smooth boundaries

Recall that a subset $M \subset \mathbb{R}^n$ is called compact if it is closed and bounded and that we denote the boundary of M by ∂M. In order to formulate Gauß's theorem we require boundaries which are sufficiently smooth. This notion is introduced in the following definition.

Definition 13.1. *A compact set $M \subset \mathbb{R}^n$ is called a compact set with smooth boundary if every $x \in \partial M$ has an open neighbourhood $x \in U \subset \mathbb{R}^n$ and a continuously differentiable function $f : U \to \mathbb{R}$ with the following properties:*
(i) $M \cap U = \{x \in U \,|\, f(x) \leq 0\}$
(ii) $\nabla f(x) \neq 0$ for all $x \in U$.

It is clear that for such a compact set with smooth boundary the boundary can be obtained (locally) as the zero set of f, that is,

$$\partial M \cap U = \{x \in U \,|\, f(x) = 0\} \,. \tag{13.2}$$

This set-up is illustrated in Fig. 26. Comparing Eq. (13.2) with Def. 11.1 shows that the boundary, ∂M,

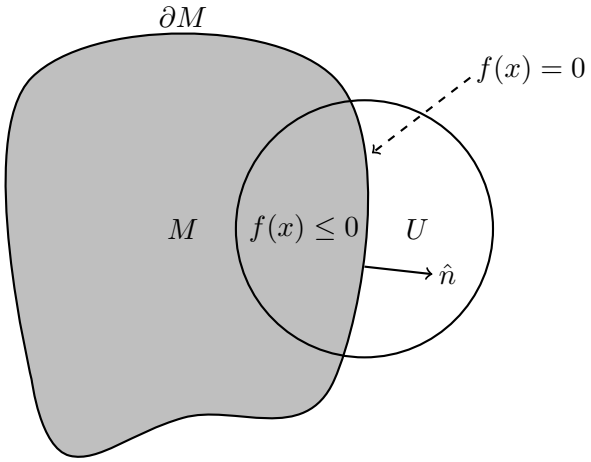


Figure 26: A compact set $M \subset \mathbb{R}^2$ with boundary locally described as the zero set of a function f, such that $U \cap \partial M = \{x \in U \,|\, f(x) = 0\}$. The outer unit normal vector, $\hat{n}$, to ∂M is also indicated.

of a compact set with smooth boundary M is a manifold. Furthermore, we have a unit normal vector $\hat{n} = \pm\frac{\nabla f}{|\nabla f|}$ on ∂M. There is precisely one choice of sign for which $\hat{n}$ points 'outward', away from M (see Fig. 26). This unit normal vector is called the **outer unit normal vector** and we should think of it as a function $\hat{n} : \partial M \to \mathbb{R}^n$.

13.2 Gauß's theorem

As we have just seen, the boundary, ∂M, of compact set $M \subset \mathbb{R}^n$ with smooth boundary is a manifold. If we have a vector field $A : U \to \mathbb{R}^n$ (where $M \subset U$ and U is open) we can integrate this vector field over ∂M, using the outer unit normal vector $\hat{n}$, that is,

$$\int_{\partial M} A \cdot dS = \int_{\partial M} A \cdot \hat{n} \, d\sigma \,. \tag{13.3}$$

The idea is that this boundary integral should be equal to an integral over M with the integrand some derivative of the vector field A. Which derivative is the correct one? In Section 8.2 we have seen that the

divergence measures the local flux of a vector field through a 'small' rectangle. On the other hand, the integral (13.3) represents the flux of A out of the set M, through its boundary ∂M. It seems, therefore, intuitive that the derivative in question is the divergence, $\nabla \cdot A$ of A. This is indeed what Gauß's theorem states.

Theorem 13.1. *(Gauß's theorem)*
Let $M \subset \mathbb{R}^n$ be a compact set with smooth boundary ∂M and $\hat{n} : \partial M \to \mathbb{R}^n$ the outer unit normal vector to ∂M. For a vector field $A : U \to \mathbb{R}^n$, where $M \subset U$ and U is open, we have

$$\int_M \nabla \cdot A \, d^n x = \int_{\partial M} A \cdot dS \, . \tag{13.4}$$

where $dS = \hat{n} \, d\sigma$.

Proof. The detailed proof is somewhat technical and can, for example, be found in Ref. [10]. Here, we would like to sketch the basic idea. In Section 9.2 we have seen that the integral on the LHS of Gauß's theorem (13.4) is defined as a limit of sums over partitions into rectangles. For each such (small) rectangle, the divergence $\nabla \cdot A$ measures the total flux through the rectangle's boundary. Now consider two neighbouring rectangles B, $\tilde{B}$ in a partition with a common boundary $\partial B \cap \partial \tilde{B}$ (see Fig. 14). The fluxes out of B and into $\tilde{B}$ through the common boundary are equal in magnitude but come with a different sign. Hence, in the sum over the partition they cancel out. This means all fluxes through interior boundaries in a partition, such as in Fig. 14, cancel out and the only remainder is the flux through the boundary of the entire partition, that is, through $\partial \left(\bigcup_a B_a \right)$. This boundary of the partition approximates the boundary ∂M as we go to ever finer partitions, so the remainder of the integral on the LHS in Eq. (13.4) is the flux through the boundary ∂M. This is precisely what the RHS of Eq. (13.4) represents. □

Gauß's theorem can be generalised in a number of ways. For example, it is possible to relax the smoothness condition on the boundary and require that ∂M is only smooth 'almost everywhere'. By this we mean that the non-smooth locus does not contribute to the integral. For example, a unit cube in $\mathbb{R}^n$ is of this type. The boundary is of course not smooth everywhere (not at the edges) but these are genuinely lower-dimensional and do not contribute to the integral.

13.3 Gauß's theorem in two dimensions

Consider Gauß's theorem in two dimensions, that is,

$$\int_M \nabla \cdot A \, d^2 x = \int_{\partial M} A \cdot \hat{n} \, d\sigma \, , \tag{13.5}$$

where $M \subset \mathbb{R}^2$ is compact with smooth boundary and the boundary ∂M is, in this case, one-dimensional, that is, it is a curve. Eq. (13.5) is perfectly fine but it might be better to express the curve integral on the RHS in terms of the unit tangent vector τ to the curve, rather than the (outer) unit normal vector $\hat{n}$, in order to match Eq. (12.21). To do this let us introduce a new vector field B which is related to A by $A_i = \epsilon_{ij} B_j$ or, written out in components

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} , \qquad A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} `B_2 \\ -B_1 \end{pmatrix} \quad \Leftrightarrow \quad A_1 = B_2, \; A_2 = -B_1 \, .$$

We can re-write the divergence of A in terms of B and this leads to

$$\nabla \cdot A = \partial_i A_i = \partial_i (\epsilon_{ij} B_j) \overset{\text{Eq. (8.28)}}{=} \operatorname{curl}(B) \, .$$

If we write the unit tangent vector τ to ∂M as $\tau_i = \epsilon_{ji}\hat{n}_j$ it follows that $A \cdot \hat{n} = B \cdot \tau$. Inserting all this into the two-dimensional theorem of Gauß (13.5) gives

$$\int_M \mathrm{curl}(B)\, d^2x = \int_{\partial M} B \cdot dx \;, \tag{13.6}$$

where $dx = \tau\, d\sigma$. The RHS is now a standard line integral, as desired, and the orientation of the line integration should be anti-clockwise (given $\hat{n}$ is the outer unit normal vector and $\tau_i = \epsilon_{ji}\hat{n}_j$).

Application 13.1. Gauss's law in Electromagnetism written in integral form
Recall from Application 8.3 Gauß's law, $\nabla \cdot E = 4\pi\rho$, one of the four Maxwell equations. We can integrate this equation over a set $M \subset \mathbb{R}^3$ (compact set with smooth boundary) and apply Gauß's theorem to re-write Gauß's law in integral form. This leads to

$$\int_{\partial M} E \cdot \hat{n}\, d\sigma \overset{\text{Eq. (13.4)}}{=} \int_M \nabla \cdot E\, d^3x = 4\pi \int_M \rho\, d^3x =: 4\pi\, Q(M) \;,$$

where $Q(M)$ is the total electric charge in M. Hence, Gauß's law can be stated by saying that the total electric flux through the boundary of some volume equals (4π times) the total electric charge contained in the volume.

Gauß's law for magnetism, $\nabla \cdot B = 0$ can be re-written in integral form following the same steps.

$$\int_{\partial M} B \cdot \hat{n}\, d\sigma = \int_M \nabla \cdot B\, d^3x = 0 \;,$$

The total magnetic flux through the boundary of a volume is zero (since there are no magnetic charges).

Application 13.2. Archimedean principle
You are all familiar with the Archimedean principle. A body in a fluid experiences an upward force equal to the weight of the water it displaces. But how can it be proven, particularly if the body does not have a regular shape?

Consider a body $M \subset \mathbb{R}^3$ with boundary ∂M, submerged in a fluid with density ρ. We use coordinates (x, y, z) with the x, y axes parallel to the water's surface, the z-axis orthogonal to it and the fluid's surface at $z = 0$. Then, the pressure at depth $z \leq 0$ is given by $p = -\rho g z$. The components f_i of the force density on the boundary ∂M of the object can be written as $f_i = A_{(i)} \cdot \hat{n}$, where $A_{(i)} = -pe_i = \rho g z e_i$ are three vector fields and $\hat{n}$ is the outer unit normal vector on ∂M. Hence, the total force F on the object is given by

$$F_i = \int_{\partial M} f_i\, d\sigma = \int_{\partial M} A_{(i)} \cdot \hat{n}\, d\sigma \overset{\text{Eq. (13.4)}}{=} \int_M \nabla \cdot A_{(i)}\, d^3x \;. \tag{13.7}$$

The vector fields $A_{(i)}$ are more explicitly given by

$$A_{(1)} = \rho g \begin{pmatrix} z \\ 0 \\ 0 \end{pmatrix} \;, \qquad A_{(2)} = \rho g \begin{pmatrix} 0 \\ z \\ 0 \end{pmatrix} \;, \qquad A_{(3)} = \rho g \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \;,$$

It follows that $\nabla \cdot A_{(1)} = \nabla \cdot A_{(2)} = 0$ and, hence, inserting into Eq. (13.7), that $F_1 = F_2 = 0$ (no

sideways force). On the other hand, $\nabla \cdot A_{(3)} = \rho g$ and inserting this into Eq. (13.7) gives

$$F_3 = \rho g \int_M d^3x = \rho g \, \mathrm{vol}(M) \; ,$$

which is precisely the expected upward force.

Example 13.1. Gauß's theorem in two dimension
Question: Consider Gauß's theorem in two dimensions in the form (13.6). Check the theorem for $M = B_1(0) \subset \mathbb{R}^2$ and the vector field $B : \mathbb{R}^2 \to \mathbb{R}^2$ given by $B(x, y) = (-y, x)^T$.
Answer: We have $\mathrm{curl}(B) = \frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y} = 2$, so the LHS of Eq. (13.6) becomes

$$\int_{B_1(0)} \mathrm{curl}(B) \, d^2x = 2 \, \mathrm{vol}(B_1(0)) = 2\pi \; .$$

To evaluate the RHS we parametrise the boundary $\partial B_1(0) = S^1$ as $X(\varphi) = (\cos(\varphi), \sin(\varphi))^T$, where $\varphi \in [0, 2\pi)$, and we have

$$n = \left| \frac{dX}{d\varphi} \right|^{-1} \frac{dX}{d\varphi} = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \; , \quad B(X(\varphi)) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix} \quad \Rightarrow \quad B(X(\varphi)) \cdot n = 1 \; .$$

Inserting this into the RHS of Eq. (13.6) gives

$$\int_{S^1} B \cdot dx = \int_0^{2\pi} d\varphi = 2\pi \; ,$$

matching the result from the LHS, as it should.

Example 13.2. Gauß's theorem in three dimension
Question: Consider the vector field : $\mathbb{R}^3 \to \mathbb{R}^3$ given by $A(x, y, z) = \frac{1}{3}(x^3, y^3, z^3)^T$ and find its flux through the boundary of the cube $C = \{(x, y, z)^T \in \mathbb{R}^3 \,|\, x, y, z \in [-1, 1]\}$.
Answer: It is possible, but tedious, to compute the integral over the boundary directly, by carrying out the surface integral. Applying Gauß's theorem is quicker. With $\nabla \cdot A = x^2 + y^2 + z^2$ we have

$$\int_{\partial C} A \cdot \hat{n} \, d\sigma = \int_C \nabla \cdot A \, d^3x = \int_{-1}^1 dx \int_{-1}^1 dy \int_{-1}^1 dz \, (x^2 + y^2 + z^2) = 8 \; .$$

13.4 Green identities

The Green identities are important and useful integral formulae which follow from Gauß's theorem. Within the set-up of Theorem 13.1 consider two twice continuously differentiable functions $f, g : U \to \mathbb{R}$ and set $A = f \, \nabla g$. If we insert

$$\nabla \cdot A = \nabla \cdot (f \, \nabla g) = f \Delta g + \nabla f \cdot \nabla g \; , \qquad A \cdot \hat{n} = f \, \nabla g \cdot \hat{n} \tag{13.8}$$

into Eq. (13.4), Gauß's theorem turns into the **first Green formula**

$$\int_M (f \Delta g + \nabla f \cdot \nabla g) d^n x = \int_{\partial M} f \, \nabla g \cdot dS \; . \tag{13.9}$$

Exchanging f and g in this formula and subtracting the two resulting equations gives the **second Green formula** or **Green's identity**

$$\int_M (f\Delta g - g\Delta f)d^n x = \int_{\partial M} (f\nabla g - g\nabla f)\cdot dS\ . \tag{13.10}$$

These results are very useful to study the Laplace and Poisson equations.

14 The integral theorem of Stokes

Gauß's theorem in two dimensions, as formulated in terms of the curl in Eq. (13.6), suggests an intriguing possibility. Is there a three-dimensional version of this integral theorem, which involves the three-dimensional version of the curl, surfaces in $\mathbb{R}^3$ and their boundaries? The answer is "yes" and the integral theorem in question is referred to as **Stokes's theorem**. (There is a general integral theorem, of which Gauß's theorem and Stokes theorem as presented in this section are special cases. This general theorem which, confusingly, is also referred to as Stokes's theorem is usually formulated in terms of differential forms which are beyond our current scope. However, if you are interested you can find a discussion in Appendix A.) To formulate Stokes's theorem (in the narrow sense) we require a bit of preparation, including introducing subsets of manifolds with (smooth) boundaries.

14.1 Smooth boundaries

In preparation for Gauß's theorem we have introduced subsets of $\mathbb{R}^n$ with smooth boundaries. To tackle Stokes's theorem we need to generalise this idea to subsets $\mathcal{S}\subset M$ of k-dimensional manifolds M embedded in $\mathbb{R}^n$.

Definition 14.1. *Let $M\subset\mathbb{R}^n$ be a k-dimensional manifold and $\mathcal{S}\subset M$ a compact subset. We say that $\mathcal{S}$ has a smooth boundary if for every $p\in\partial\mathcal{S}$ there is a chart $X: U\to M\cap V\ni p$ with the following properties.*
(i) $X(H\cap U)=\mathcal{S}\cap V$
(ii) $X(\partial H\cap U)=\partial\mathcal{S}\cap V$
where H is the half-space $H=\{(x_1,\dots,x_k)^T\in\mathbb{R}^k\,|\,x_1\le 0\}$.

This sounds a bit convoluted but the basic idea is very simple and is illustrated in Fig. 27. It is intuitively clear that one can pick an atlas of M adapted to the boundary $\partial\mathcal{S}$. This means every chart $X: U\to M\cap V$ can be chosen so that it consistent with the properties stated in Def. 14.1, that is, the local boundary of $\mathcal{S}$ originates from the local boundary of the half-space H. We can then define a chart $Y:\partial H\cap U\to\partial\mathcal{S}\cap V$ by $Y(s)=X(0,s)$, where $s=(t_2,\dots,t_k)$, and this makes the boundary $\partial\mathcal{S}$ into a manifold. Then, for a point $p=Y(s)\in\partial\mathcal{S}\cap V$ the tangent space of M and $\partial\mathcal{S}$ are

$$T_pM=\text{span}\left(\frac{\partial X}{\partial t_1},\dots,\frac{\partial X}{\partial t_k}\right)_p \quad\supset\quad T_p\partial\mathcal{S}=\text{span}\left(\frac{\partial X}{\partial t_2},\dots,\frac{\partial X}{\partial t_k}\right)_p$$

14.2 Stokes theorem

We consider an orientable surface $M\subset\mathbb{R}^3$ with a continuous unit normal vector $n: M\to\mathbb{R}$, a compact subset $\mathcal{S}\subset M$ with a smooth boundary $\partial\mathcal{S}$ and a (continuously differentiable) vector field $A: W\to\mathbb{R}^3$, where W is open and $M\subset W$ (so that the vector field is defined for all points on the manifold). We want to formulate an equality between a surface integral of a vector field and a curve integral but we know that both types of integrals depend on the orientation and change sign under orientation reversal. For this

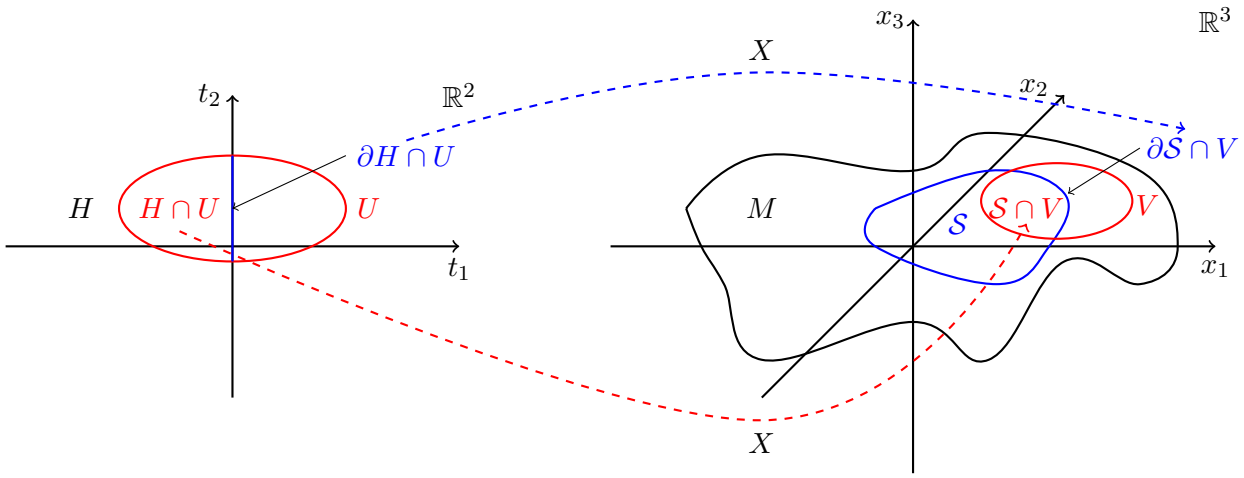


Figure 27: A compact subset $\mathcal{S} \subset M$ of a surface $M \subset \mathbb{R}^3$. Note that the chart X maps as described in Def. 14.1.

reason we need to correlate the orientations of the surface and the bounding curve. To do this, one can choose an atlas of M, adapted to the boundary $\partial\mathcal{S}$, such that for every chart $X : U \to M \cap V$ the vectors $\left(\hat{n}, \frac{\partial X}{\partial t_1}, \frac{\partial X}{\partial t_2}\right)$ forms a positively oriented coordinate system in $\mathbb{R}^3$, meaning that

$$\det\left(\hat{n}, \frac{\partial X}{\partial t_1}, \frac{\partial X}{\partial t_2}\right) > 0 \,. \tag{14.1}$$

For a point $p \in \partial\mathcal{S} \cap V$ on the boundary of $\mathcal{S}$ we know from the discussion in the previous section that the tangent spaces at p are

$$T_pM = \text{span}\left(\frac{\partial X}{\partial t_1}, \frac{\partial X}{\partial t_2}\right)_p \quad \supset \quad T_p\partial\mathcal{S} = \text{span}\left(\frac{\partial X}{\partial t_2}\right)_p \,.$$

The tangent vector $\frac{\partial X}{\partial t_2}$ tangential to the boundary curve $\partial\mathcal{S}$ marks out a particular orientation by which we go around the curve. This orientation is correlated, via Eq. (14.1), with the orientation of the surface M defined by the unit normal vector n and it is called the **induced orientation** on the boundary. In practice, the induced orientation is the one inferred from the "right-hand rule", with the thumb in the direction of n and the other fingers indicating the direction around the bounding curve. We are now ready to formulate Stokes's theorem, taking inspiration from Eq. (13.6).

Theorem 14.1. *(Stokes)*
Let $M \subset \mathbb{R}^3$ be a surface, $\mathcal{S} \subset M$ be a compact subset with a smooth boundary and $A : W \to \mathbb{R}^3$ a continuously differentiable vector field, where $M \subset W$. Further assume there is a continuous unit normal $\hat{n}$ on M, that is, M is oriented. Then

$$\int_{\mathcal{S}} \nabla \times A \cdot \hat{n}\, d\sigma = \int_{\partial\mathcal{S}} A \cdot dx \,, \tag{14.2}$$

where the curve integral on the RHS is taken with the induced orientation.

Proof. The full proof is somewhat technical but the basic proof idea is similar to the one for Gauß's theorem and is as follows. The surface integral on the LHS of Eq. (A.53) is ultimately defined by a sum

over partitions into rectangles, as in Fig. 14. When we introduced the curl in Section 8.3 we saw that it measures the vector field flow around the boundaries of small local rectangles. For two rectangles in a partition the flow goes in opposite directions along common boundaries. This means the flow contributions from all internal boundaries cancel as the sum over the partition is performed and only the flow around the outer boundary of the partition remains. This flow is represented by the curve integral over the boundary on the RHS of Eq. (A.53). □

Example 14.1. Stokes's theorem on the surface of a cylinder
Question: For the cylinder surface C with unit normal vector n as in Example 12.9 find the boundary, a suitable parametrisation (with the induced orientation) for the boundary and check Stokes's theorem for the vector field $A : \mathbb{R}^3 \to \mathbb{R}^3$ given by $A(x, y, z) = (yz, x, 0)^T$.

Answer: The boundary of the cylinder surface is $\partial C = S_0^1 \cup S_h^1$, where S_0^1 and S_h^1 are the bounding circles at the bottom and the top, respectively. Parametrisations for these two boundary components are

$$X_0(\varphi) = \begin{pmatrix} R\cos(\varphi) \\ R\sin(\varphi) \\ 0 \end{pmatrix} , \quad X_h(\varphi) = \begin{pmatrix} R\cos(\varphi) \\ -R\sin(\varphi) \\ h \end{pmatrix}$$

so we have tangent vectors

$$\frac{dX_0}{d\varphi} = \begin{pmatrix} -R\sin(\varphi) \\ R\cos(\varphi) \\ 0 \end{pmatrix} , \quad \frac{dX_h}{d\varphi} = \begin{pmatrix} -R\sin(\varphi) \\ -R\cos(\varphi) \\ 0 \end{pmatrix} . \tag{14.3}$$

With $\nabla \times A = (0, y, 1 - z)^T$ the LHS of Stokes theorem (A.53) becomes (using the measure dS as computed in Example 12.9)

$$\int_C \nabla \times A \cdot dS = R \int_0^{2\pi} d\varphi \int_0^h dz \begin{pmatrix} 0 \\ R\sin(\varphi) \\ 1 - z \end{pmatrix} \cdot \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \\ 0 \end{pmatrix} = R^2 \int_0^{2\pi} d\varphi \, \sin^2(\varphi) \int_0^h dz \, = \pi R^2 h \, .$$

For the LHS we have

$$\begin{aligned} \int_{S_0^1 \cup S_h^1} A \cdot dx &= \int_{S_0^1} A \cdot dx + \int_{S_h^1} A \cdot dx = R^2 \int_0^{2\pi} d\varphi \, \cos^2(\varphi) - R^2 \int_0^{2\pi} d\varphi \, (\cos^2(\varphi) - h \sin^2(\varphi)) \\ &= \pi R^2 - \pi R^2 (1 - h) = \pi R^2 h \, . \end{aligned}$$

This does indeed coincide with the RHS but note the induced orientation of the integration, as represented by the tangent vectors in Eq. (14.3), has to be implemented correctly to make this work. This example illustrates that Stokes theorem also holds in cases where the boundary has several disconnected pieces.

Application 14.1. Ampère-Maxwell's and Faraday's laws in integral form
We can use Stokes theorem to re-write the two Maxwell equations (8.36) which involve a curl into integral form. This is in analogy with Application 13.1 where we have used Gauß's theorem to bring the two Maxwell equations with divergences into integral form. We choose a compact set $\mathcal{S} \subset M$ with smooth boundary $\partial\mathcal{S}$ on a surface M and start with the Ampère-Maxwell law. First introduce

the current, displacement current and the magnetic fluc

$$I = \int_{\mathcal{S}} J \cdot dS \ , \qquad \phi_E = \int_{\mathcal{S}} E \cdot dS \ , \qquad \phi_M = \int_{\mathcal{S}} B \cdot dS \ .$$

Then, the Ampère-Maxwell low can be written as

$$\int_{\partial\mathcal{S}} B \cdot dx \overset{\text{Eq. (A.53)}}{=} \int_{\mathcal{S}} \nabla \times B \cdot dS \overset{\text{Eq. (8.36)}}{=} \int_{\mathcal{S}} \left(\frac{4\pi}{c} J + \frac{1}{c} \dot{E} \right) \cdot dS = \frac{4\pi}{c} I + \frac{1}{c} \dot{\phi}_E \ .$$

In short, the magnetic flow around the boundary $\partial\mathcal{S}$ is proportional to the electric current through $\mathcal{S}$ (plus the time-variation of the displacement current). For Faraday's law we can proceed in a similar way.

$$\int_{\partial\mathcal{S}} E \cdot dx \overset{\text{Eq. (A.53)}}{=} \int_{\mathcal{S}} \nabla \times E \cdot dS \overset{\text{Eq. (8.36)}}{=} -\frac{1}{c} \int_{\mathcal{S}} \dot{B} \cdot dS = -\frac{1}{c} \dot{\phi}_M \ .$$

The flow of the electric field around $\partial\mathcal{S}$ is proportional to the time variation of the magnetic flux through $\mathcal{S}$.

Appendices

A Differential forms

Differentials forms is another classical subject of Mathematical Physics which cannot be covered in the main part of this course. This appendix is a no-nonsense guide to differential forms and a chance to read up on the subject without having to deal with excessive mathematical overhead. Many physical theories can be elegantly formulated in terms of differential form, including Classical Mechanics and Electrodynamics. Differential forms also provide a unifying perspective on the subject of "vector calculus" which leads to a deeper understanding of the many ad-hoc objects - such as divergence, curl etc. - which have been introduced in this context. Our treatment here is basic in that we focus on differential forms on $\mathbb{R}^n$ and manifolds embedded in $\mathbb{R}^n$, in line with our treatment in the main part of these notes. (Ref. [9] contains a more advanced treatment of differential forms.)

A.1 Differential one-forms

Definition of differential one-forms
We begin by defining differential one-forms on open sets $U \subset \mathbb{R}^n$ with coordinates $x = (x^1, \ldots, x^n)^T$. If we view U as a manifold (in the sense of Theorem 11.1 and, for example, parametrised by itself using the identity map) then attached to every point $p \in U$ there is a tangent space T_pU. Since U is an n-dimensional manifold we have $T_pU \cong \mathbb{R}^n$ and we can use the standard unit vectors e_i, where $i = 1, \ldots, n$, as a basis of T_pU. Intuitively, just think about a family of vector spaces, one each attached to every point $p \in U$, as in Fig. 28. Basically, any construction in linear algebra can now be carried out for this family.

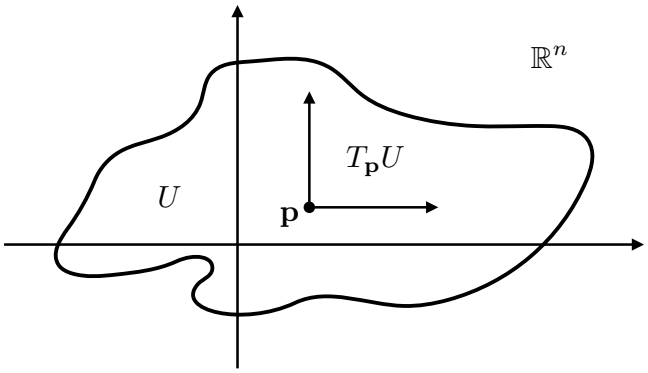


Figure 28: Open set $U \subset \mathbb{R}^n$ and tangent space T_pU at a point $p \in U$.

In particular, every tangent space T_pU has a dual vector space T_p^*U (which consists of linear functionals $T_pU \to \mathbb{R}$), called the *co-tangent space*. The elements of the co-tangent space are called *co-tangent vectors*. So, attached to every point $p \in U$ we have two vector spaces, the tangent space T_pU and its dual, the co-tangent space T_p^*U. We are now ready to define differential one-forms.

Definition A.1. *(Differential one-forms) A differential one-form is a map $\omega : U \to T^*U := \bigcup_{p\in U} T_p^*U$, with $p \mapsto \omega_p$, such that $\omega_p \in T_p^*U$.*

Stated less formally, a differential one-form ω provides us with a co-tangent vector ω_p at every point $p \in U$. Recall that such a co-tangent vector is a functional on the tangent space, so it provides a map $\omega_p : T_pU \to \mathbb{R}$ and for a tangent vector $v \in T_pU$ we have $\omega_p(v) \in \mathbb{R}$.

The total differential

So far this sounds fairly abstract and appears to be of little use but some light is shed on the matter by the following

Definition A.2. *(Total differential) For a differentiable function $f : U \to \mathbb{R}$ the total differential df is a one-form defined by*

$$df_p(v) := \nabla f_p \cdot v = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x^i}\right|_p v^i \tag{A.1}$$

where $v = \sum_{i=1}^{n} v^i e_i \in T_pU$.

Exercise A.1. *Convince yourself that the total differential defined in Def. A.2 is indeed a differential one-form.*

Basis for differential one-forms

We can use the total differential to gain a better understanding of differential forms. To do this we introduce *coordinate functions* $x^i : U \to \mathbb{R}$ defined by

$$x^i(p) := p^i \ , \tag{A.2}$$

where $p = (p^1, \ldots, p^n)^T$, which assign to each point $p \in U$ its i^{th} coordinate p^i. (By a slight abuse of notation we have used the same symbol for the coordinate x^i and the coordinate function.) To understand what the total differentials dx^i of the coordinate functions are we act on the basis vectors e_i of the tangent space.

$$dx^i|_p(e_j) = \nabla x^i|_p \cdot e_j = e^i \cdot e_j = \delta^i_j \qquad \Rightarrow \qquad dx^i|_p(v) = v^i \ . \tag{A.3}$$

This is precisely the defining relation for a dual basis [10] and we learn that $(dx^i|_p)$ is the basis of T^*_pU dual to the basis (e_i) of T_pU. Hence, we have the following

Proposition A.1. *The total differentials $(dx^1|_p, \ldots, dx^n|_p)$ of the coordinate functions form a basis of the co-tangent space T^*_pU. Hence, every differential one-form ω and every total differential df on U can be written as*

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^i \ , \qquad df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \, dx^i \ , \tag{A.4}$$

where $\omega_i : U \to \mathbb{R}$ are functions.

Proof. We have already shown the first part of this statement, that is, that $(dx^1|_p, \ldots, dx^n|_p)$ is indeed a basis of T^*_pU. This means for every one-form w and every point $p \in U$ we can write $\omega_p = \sum_{i=1}^{n} \omega_i(p) dx^i|_p$, for some suitable coefficients $\omega_i(p)$. Dropping the argument p gives the first Eq. (A.4). To show the second Eq. (A.4) work out

$$df|_p(v) = \nabla f|_p \cdot v = \left.\frac{\partial f}{\partial x^i}\right|_p v^i = \sum_i \left.\frac{\partial f}{\partial x^i}\right|_p dx^i|_p(v) \ , \tag{A.5}$$

where Eq. (A.3) has been used in the last step. Dropping the arguments v and p leads to the second Eq. (A.4). □

[10] It is sometimes said that dx^i represents a "small interval attached to x^i". Obviously, this statement is not really correct given how we have just defined differential forms. Generously interpreted, the "small interval" view of dx^i can be thought of as an imprecise way of stating the content of the equation on the right in (A.3). This equation says that $dx^i|_p$, when acting on a tangent $v \in T_pM$ (and it is this tangent vector which should really be seen as the displacement from p), gives its i^{th} component v^i.

This proposition gives us a clear idea what differential one-forms are and how to write them down in practice. All we need is n functions $\omega_i : U \to \mathbb{R}$ and we can write down a differential form using Eq. (A.4). In particular, we see that differential one-forms w on U are in one-to-one correspondence with vector fields $A : U \to \mathbb{R}^n$ via

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^i \qquad \longleftrightarrow \qquad A = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix} . \tag{A.6}$$

This is our first hint that differential forms provide us with a new way to think about vector calculus. Under the identification (A.6), total differentials df correspond to vector fields ∇f which are given as gradients of functions. This means that total differentials correspond to conservative vector fields.

The representation (A.4) of differential forms can be used to define their properties in terms of properties of the constituent functions w_i.

Definition A.3. *A one-form $\omega = \sum_{i=1}^{n} \omega_i \, dx^i$ on U is called differentiable (continuously differentiable, k times continuously differentiable) if all functions $\omega_i : U \to \mathbb{R}$ are differentiable (continuously differentiable, k times continuously differentiable).*

Integrating differential one-forms

Differential one-forms are designed to be integrated along one-dimensional manifolds, that is, along curves. To see how this works, consider a curve $X : [a, b] \to U$ with parameter $t \in [a, b]$ and tangent vector $\frac{dX}{dt}\big|_{X(t)} \in T_{X(t)}U$. At each point $X(t)$ on the curve, a one-form ω on U provides a linear functional $\omega|_{X(t)} : T_{X(t)}U \to \mathbb{R}$ which can act on the tangent to the curve to produce a real number, so

$$\omega|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) \in \mathbb{R} . \tag{A.7}$$

Based on this observation we can define

Definition A.4. *Let ω be a differential one-form on U and $X : [a, b] \to U$ a curve. Then, we define the integral of ω over X as*

$$\int_X \omega := \int_a^b \omega|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) dt . \tag{A.8}$$

Note that this definition is extremely natural. At each point on the curve we let the differential one-form act on the tangent vector to the curve to produce a number and then we integrate over these numbers along the curve. If we write the one-form and the tangent vector relative to their standard bases as

$$\omega = \sum_{i=1}^{n} \omega_i \, dx^i , \qquad \frac{dX}{dt} = \sum_{j=1}^{n} \frac{dX^j}{dt} e_j , \tag{A.9}$$

we have

$$\omega|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) = \sum_{i,j} \omega_i(X(t)) \, \frac{dX^j}{dt}\bigg|_{X(t)} dx^i|_{X(t)}(e_j) = \sum_i \omega_i(X(t)) \, \frac{dX^i}{dt}\bigg|_{X(t)} = A(X(t)) \cdot \frac{dX}{dt}\bigg|_{X(t)}$$

where the duality property (A.3) has been used in the second last step and we have used the identification (A.6) of one-forms and vector fields in the last step. Hence, the integral of w over X can also be written as

$$\int_X \omega = \sum_{i=1}^{n} \int_a^b \omega_i(X(t)) \, \frac{dX^i}{dt}\bigg|_{X(t)} dt = \int_a^b A(X(t)) \cdot \frac{dX}{dt}\bigg|_{X(t)} dt . \tag{A.10}$$

This relation shows that the integral of a one-form over a curve is nothing else but the "line-integral" of the corresponding vector field.

For the integral of a total differential we have

$$\int_X df = \int_a^b \nabla f(X(t)) \cdot \left.\frac{dX}{dt}\right|_{X(t)} dt = \int_a^b \frac{d}{dt}(f(X(t))) = f(X(b)) - f(X(a)) \tag{A.11}$$

which says that the curve integral of a total differential df equals the difference of the values of f at the endpoints of the curve. That is, the value of such integrals only depends on the endpoints but not on the path taken. In particular, integrals of total differentials over closed curves (curves with $X(a) = X(b)$) vanish. These statements are of course equivalent to the corresponding statements for conservative vector fields.

So far this does not appear to be overly useful. All we seem to have done is to describe vector fields in a different way. However, this was just a preparation for introducing higher order differential forms and this is where things get interesting. Before we can do this we need a bit of mathematical preparation.

A.2 Alternating k-forms

To introduce higher-order differential forms we need to discuss some more advanced linear algebra – alternating k-forms (see Ref. [5] for more details).

Definition of alternating k-forms

We start with a vector space V over a field $\mathbb{F}$ and its dual space V^*. As we have seen in the previous sub-section the dual vector space (to the tangent space) was the crucial ingredient in defining differential one-forms. In order to be able to introduce higher order differential forms we first need to discuss a bit of more advanced linear algebra which generalises linear functionals. These are the *alternating k-forms* on V which are defined as follows.

Definition A.5. *(Alternating k-forms) An alternating k-form ω on a vector space V over $\mathbb{F}$ is a map $w : V^{\otimes k} \to \mathbb{F}$ which is*

(i) linear in each of its k arguments, so $\omega(\ldots, \alpha v + \beta u, \ldots) = \alpha\,\omega(\ldots, v, \ldots) + \beta\,\omega(\ldots, u, \ldots)$
(ii) alternating in the sense that $\omega(\ldots, v, \ldots, u, \ldots) = -\omega(\ldots, u, \ldots, v, \ldots)$

for all $u, v \in V$ and all $\alpha, \beta \in \mathbb{F}$. (The dots indicate arguments which remain unchanged.) The vector space of alternating k-forms over V is denoted $\Lambda^k V^*$, where $k = 1, 2, \ldots$. We also define $\Lambda^0 V^* = \mathbb{F}$.

In short, alternating k-forms take k vector arguments to produce a scalar and they are linear in each argument and completely antisymmetric under the exchange of arguments. It is clear from the definition that alternating one-forms are linear functionals, so $\Lambda^1 V^* = V^*$, so in this sense we have set up a generalisation of the dual vector space.

The wedge product

However, it is not quite obvious from the above definition how to write down these k-forms more concretely. To get a handle on this we now devise a way to built up k-forms from one-forms, that is functionals, by means of an operation called the *wedge product*.

Definition A.6. *(Wedge or outer product) Consider functionals $\varphi_1, \ldots, \varphi_k \in V^*$. Then, the k-form $\varphi_1 \wedge \ldots \wedge \varphi_k \in \Lambda^k V^*$ is defined by*

$$(\varphi_1 \wedge \ldots \wedge \varphi_k)(v_1, \ldots, v_k) := \det \begin{pmatrix} \varphi_1(v_1) & \cdots & \varphi_1(v_k) \\ \vdots & & \vdots \\ \varphi_k(v_1) & \cdots & \varphi_k(v_k) \end{pmatrix} . \tag{A.12}$$

It is clear from the linearity of the functionals φ_i as well as the linearity and anti-symmetry of the determinant that $\varphi_1 \wedge \ldots \wedge \varphi_k$ as defined in Eq. (A.12) is indeed a k-form. It is also easy to see from the properties of the determinant that calculating with the wedge product is subject to the following rules (α and β are scalars):

$$\begin{aligned} \varphi_1 \wedge \ldots \wedge (\alpha\varphi_i + \beta\tilde{\varphi}_i) \wedge \ldots \wedge \varphi_k &= \alpha\, \varphi_1 \wedge \ldots \wedge \varphi_i \wedge \ldots \wedge \varphi_k + \beta\, \varphi_1 \wedge \ldots \wedge \tilde{\varphi}_i \wedge \ldots \wedge \varphi_k & \text{(A.13)} \\ \varphi_1 \wedge \ldots \wedge \varphi_i \wedge \varphi_{i+1} \wedge \ldots \wedge \varphi_k &= -\varphi_1 \wedge \ldots \wedge \varphi_{i+1} \wedge \varphi_i \wedge \ldots \wedge \varphi_k & \text{(A.14)} \\ \ldots \wedge \varphi \wedge \varphi \wedge \ldots &= 0 & \text{(A.15)} \end{aligned}$$

The first two rules follow from linearity and anti-symmetry of the determinant, respectively, the third one is a direct consequence of the second.

Basis for alternating k-forms

The following proposition provides a basis for the space of alternating k-forms.

Proposition A.2. *Let $(v_*^1, \ldots, v_*^n)$ be a basis of V^*. Then the k-forms $v_*^{i_1} \wedge \ldots \wedge v_*^{i_k}$, where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ are a basis of $\Lambda^k V^*$. In particular, we have*

$$\dim(\Lambda^k V^*) = \begin{pmatrix} n \\ k \end{pmatrix} . \tag{A.16}$$

Proof. Let $(v_1, \ldots, v_n)$ be the basis of V with dual basis $(v_*^1, \ldots, v_*^n)$, so that $v_*^i(v_j) = \delta_j^i$. To proof linear independence consider

$$\sum_{i_1 < \cdots < i_k} \lambda_{i_1 \ldots i_k}\, v_*^{i_1} \wedge \ldots \wedge v_*^{i_k} = 0 \tag{A.17}$$

and act with this equation on $(v_{j_1}, \ldots, v_{j_k})$, where $j_1 < \cdots < i_k$. It follows immediately that $\lambda_{j_1 \ldots j_k} = 0$.

To show that these k-forms span the space start with an arbitrary $\omega \in \Lambda^k V^*$. Define the numbers $c_{i_1 \ldots i_k} := w(v_{i_1}, \ldots, v_{i_k})$ and the differential k-form

$$\tilde{\omega} := \sum_{j_1 < \cdots < j_k} c_{j_1 \ldots j_k} v_*^{j_1} \wedge \ldots \wedge v_*^{j_k} . \tag{A.18}$$

Then it follows that $\tilde{\omega}(v_{i_1}, \ldots, v_{i_k}) = c_{i_1 \ldots i_k} = \omega(v_{i_1}, \ldots, v_{i_k})$ and since $\tilde{\omega}$ and ω agree on a basis we have $\omega = \tilde{\omega}$ and have, hence, written w as a linear combination (A.18) of our basis forms. □

The above Lemma gives us a clear idea of how to write down alternating k-forms. Once we have a basis v_*^i of the dual space V^* we can use the wedge product to construct a basis $v_*^{i_1} \wedge \ldots \wedge v_*^{i_k}$, where $1 \leq i_1 < \ldots < i_k \leq n$, of $\Lambda^k V^*$, the space of alternating k-forms. Then any k-form w can be written as

$$\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1 \ldots i_k} v_*^{i_1} \wedge \ldots \wedge v_*^{i_k} = \frac{1}{k!} \sum_{i_1, \ldots, i_k} \omega_{i_1 \ldots i_k} v_*^{i_1} \wedge \ldots \wedge v_*^{i_k} , \tag{A.19}$$

where the coefficients $\omega_{i_1 \ldots i_k} \in F$ are completely anti-symmetric in the k indices. Also note that there are no non-trivial alternating k-forms for $k > n = \dim(V)$. In this case the wedge product will always involve at least two same basis vectors and must vanish from Eq. (A.15). Of course the wedge product generalises to arbitrary alternating forms by linearity. Specifically, consider an alternating k-form ω as in Eq. (A.19) and an alternating l-forms ν given by

$$\nu = \frac{1}{l!} \sum_{j_1, \ldots j_l} \nu_{j_1 \ldots j_l} v_*^{j_1} \wedge \ldots \wedge v_*^{j_l} , \tag{A.20}$$

Their wedge product is an alternating $k+l$-form defined by

$$\omega \wedge \nu := \frac{1}{k!\,l!} \sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} \omega_{i_1\ldots i_k} \nu_{j_1\ldots j_l} v_*^{i_1} \wedge \ldots \wedge v_*^{i_k} \wedge v_*^{j_1} \wedge \ldots \wedge v_*^{j_l} \; . \tag{A.21}$$

It is easy to see that

$$\omega \wedge \nu = (-1)^{kl} \nu \wedge \omega \; . \tag{A.22}$$

Let us illustrate alternating k-forms with an example.

Application A.1. Alternating forms in three dimensions
Consider a three-dimensional vector space V with basis (v_i) and dual basis (v_*^i), where $i = 1, 2, 3$. In this case, we have non-trivial k-forms for the values $k = 0, 1, 2, 3$ and the spaces $\Lambda^k V^*$ are as follows.

space	basis	typical element	dimension
$\Lambda^0 V^*$	(1)	$w \in \mathbb{R}$	1
$\Lambda^1 V^*$	(v_*^1, v_*^2, v_*^3)	$\sum_{i=1}^3 \omega_i v_*^i$	3
$\Lambda^2 V^*$	$(\nu^1 := v_*^2 \wedge v_*^3, \nu^2 := v_*^3 \wedge v_*^1, \nu^3 := v_*^1 \wedge v_*^3)$	$\sum_{i=1}^3 \omega_i \nu^i$	3
$\Lambda^3 V^*$	$(v_*^1 \wedge v_*^2 \wedge v_*^3)$	$\omega\, v_*^1 \wedge v_*^2 \wedge v_*^3$	1

Suppose we have two alternating one-forms $w = \sum_{i=1}^3 w_i v_*^i$ and $u = \sum_{j=1}^3 u_j v_*^j$ and we write the coefficients as vectors, so $W = (w_1, w_2, w_3)^T$ and $U = (u_1, u_2, u_3)^T$. Then, taking the wedge product we find

$$w \wedge u = \sum_{i,j=1}^3 w_i u_j v_*^i \wedge v_*^j = \sum_{i=1}^3 (W \times U)_i \nu^i \; , \tag{A.23}$$

where ν^i are the basis forms for $\Lambda^2 V^*$ defined in the above table. In conclusion, we see that the wedge-product of two one-forms leads to a two-form which can be expressed in terms of the cross product. Alternating k-forms in three dimensions are the proper context within which to formulate the cross product and we can see why the cross product only appears in three dimensions. Only in this case have the spaces $\Lambda^1 V^*$ and $\Lambda^2 V^*$ the same dimensions so that both one-forms and two-forms can be interpreted as three-dimensional vectors.

To summarise our discussion, over each n-dimensional vector space V, we now have a "tower" of vector spaces

$$\Lambda^0 V^* = F \; , \quad \Lambda^1 V^* = V^* \; , \quad \Lambda^2 V^* \; , \quad \cdots \; , \quad \Lambda^n V^* \tag{A.24}$$

consisting of alternating k-forms for $k = 0, 1, 2, \ldots, n$. In our definition of differential one-forms we have used $\Lambda^1 V^* = V^*$ (where the tangent spaces $T_p U$ have assumed the role of V). Now we can go further and arrange the whole tower of vector spaces (A.24) over each tangent space $V = T_p U$. As we will see, doing this leads to higher-order differential forms.

A.3 Higher-order differential forms

Definition of differential forms

After this preparation it is now straightforward to define differential k-forms. We simply copy the definition A.1 of differential one-forms but replace the cotangent space $T_p^* U$ used there by $\Lambda^k T_p^* U$.

Definition A.7. *(Differential k-forms) A differential k-form is a map $\omega : U \to \Lambda^k T^* U := \bigcup_{p \in U} \Lambda^k T_p^* U$, with $p \mapsto \omega_p$, such that $\omega_p \in \Lambda^k T_p^* U$.*

From this definition, a differential k-form w provides an alternating k-form ω_p over T_pU for every point $p \in U$ which can act on k tangent vectors $v_1, \ldots, v_k \in T_pU$ to give a number $\omega_p(v_1, \ldots, v_k) \in \mathbb{R}$.

It is now quite easy to write down the general expression for a differential k-form. From our discussion of differential one-forms we know that the differentials $dx^1_p, \ldots, dx^n_p$ form a basis of the cotengent space T^*_pU at $p \in U$. Combining this with what we have said about alternating k-forms in the previous subsection (see Prop. A.2) shows that the k-forms $dx^{i_1}_p \wedge \ldots \wedge dx^{i_k}_p$, where $1 \leq i_1 < \ldots < i_k \leq n$, form a basis of $\Lambda^k T^*_pU$. Hence, dropping the point p, a differential k-form can be written as

$$\omega = \frac{1}{k!} \sum_{i_1,\ldots,i_k} \omega_{i_1\ldots i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k} \; , \tag{A.25}$$

where the $\omega_{i_1\ldots i_k} : U \to \mathbb{R}$ are functions, labelled by a completely anti-symmetric set of indices, $i_1, \ldots, i_k$. We can define properties of a differential k-form ω in terms of the functions $\omega_{i_1\ldots i_k} : U \to \mathbb{R}$, just as we have done in the case of differential one-forms (see Def. A.3). For example, we say that ω is differentiable iff all the functions $\omega_{i_1\ldots i_k}$ are. The space of infinitely many times differentiable k-forms on U is denoted by $\Omega^k(U)$. Note that we can generalise the wedge product and define the $k+l$-differential form $\omega \wedge \nu$, where ω and ν are differential k- and l-forms respectively, in complete analogy with what we have done for alternating forms in Eq. (A.21).

The exterior derivative

We can now introduce a derivative $d : \Omega^k(U) \to \Omega^{k+1}(U)$, called *exterior derivative*, which maps differential k-forms into differential $k+1$-forms. The exterior derivative of a k-forms (A.25) is defined by

$$d\omega := \frac{1}{k!} \sum_{i_1,\ldots,i_k} d\omega_{i_1\ldots i_k} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = \frac{1}{k!} \sum_{i_1,\ldots,i_k,i} \frac{\partial w_{i_1\ldots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \; , \tag{A.26}$$

where we have worked out the total derivatives $d\omega_{i_1\cdots i_k}$ of the functions $\omega_{i_1\cdots i_k}$ explicitly in the second step. Recall that the total differential of a functions $f : U \to \mathbb{R}$ is a one-form df given by

$$df = \sum_i \frac{\partial f}{\partial x^i} dx^i \; . \tag{A.27}$$

The exterior derivative satisfies a Leibnitz-type rule. If w is a k-form and ν is an l-form then

$$d(\omega \wedge \nu) = d\omega \wedge \nu + (-1)^k \omega \wedge d\nu \; . \tag{A.28}$$

Another simple but important property of the exterior derivative is

Proposition A.3. $d^2 = 0$

Proof. This statement follows from straightforward calculation.

$$\begin{aligned} d\omega &= \frac{1}{k!} \sum_{i,i_1,\ldots,i_k} \frac{\partial \omega_{i_1\ldots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \\ d^2\omega &= d(d\omega) = \frac{1}{k!} \sum_{i,j,i_1,\ldots,i_k} \frac{\partial^2 \omega_{i_1\ldots i_k}}{\partial x^i \partial x^j} dx^j \wedge dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = 0 \end{aligned}$$

The last equality follows because the second partial derivative is symmetric in (i, j) while $dx^j \wedge dx^i$ is anti-symmetric. □

In view of this result it is useful to introduce the following terminology.

Definition A.8. *A differential k-form ω is called closed iff $d\omega = 0$. It is called exact iff there exists a differential $(k-1)$-form ν such that $\omega = d\nu$.*

An immediate consequence of Prop. A.3 is that every exact differential form is also closed. The converse is not always true but it is under certain conditions and this is formulated in a statement known as Poincaré's Lemma (see, for example, Ref. [7]). In general, the failure of closed forms to be exact is an important phenomenon which is related to the topology of the manifold and leads into an area of mathematics known as algebraic geometry. This is captured by the sequence

$$0 \longrightarrow \Omega^0(U) \xrightarrow{d_0} \Omega^1(U) \xrightarrow{d_1} \cdots\cdots \xrightarrow{d_{n-1}} \Omega^n(U) \xrightarrow{d_n} 0 \tag{A.29}$$

where we have attached an index to the exterior derivative d to indicate which degree form it acts on. The property $d^2 = 0$ from Prop. A.3 then translates into $d_k \circ d_{k-1} = 0$, that is, successive maps in the sequence (A.29) compose to zero. This makes the sequence (A.29) into what is called a *complex*: a sequence of vector spaces related by maps such that adjacent maps compose to zero. The relation $d_k \circ d_{k-1} = 0$ implies that $\mathrm{Im}(d_{k-1}) \subset \mathrm{Ker}(d_k)$ (another, fancier way of saying every exact form is closed) and this allows us to define the cohomology of the space U by $H^k(U) = \mathrm{Ker}(d_k)/\mathrm{Im}(d_{k-1})$. If the space U is such that Poincaré's Lemma applies, so that every closed form is exact, then $\mathrm{Im}(d_{k-1}) = \mathrm{Ker}(d_k)$ and the cohomology groups are trivial, that is, $H^k(U) = \{0\}$. Conversely, non-trivial cohomology groups indicate there are closed but non-exact forms, that Poincaré's Lemma does not apply and that we have a non-trivial manifold topology. Pursuing this further is well beyond our present scope (see, for example, Ref. [9]).

Let us take a practical approach and work out differential forms and the exterior derivative more explicitly for an example.

Application A.2. Differential forms on $\mathbb{R}^3$
Let us consider differential forms in three dimensions on an open set $U \subset \mathbb{R}^3$. We denote coordinates by $\mathbf{x} = (x^1, x^2, x^3)^T = (x, y, z)^T$. The basis differential forms required to write down one-forms are (dx, dy, dz) and, as discussed above, the basis forms for higher-order differential forms can be obtained as wedge products of (dx, dy, dz). For concise notation we arrange all these basis forms as follows:

$$\begin{aligned} ds &:= (dx, dy, dz)^T \\ dS &:= (dy \wedge dz, dz \wedge dx, dx \wedge dy)^T \\ d^3x = dV &:= dx \wedge dy \wedge dz \ . \end{aligned} \tag{A.30}$$

Recall that in three dimensions we only have differential k-forms for $k = 0, 1, 2, 3$ and no higher. With the above notation their general form is:

degree	differential form	components	name in vector caclulus
0	$\omega_0 = f$	f	function
1	$\omega_1 = A \cdot ds$	$A = (A_1, A_2, A_3)^T$	vector field
2	$\omega_2 = B \cdot dS$	$B = (B_1, B_2, B_3)^T$	vector field
3	$\omega_3 = F\, dV$	F	function

The above table gives the general expressions for differential forms in three dimensions and also shows that they can be identified with well-known objects in three-dimensional vector calculus. Specifically, zero-forms and three-forms correspond to (real-valued) functions while one-forms and two-forms correspond to vector fields. (These identifications become more complicated in higher dimensions.)

Now that we have a correspondence between three-dimensional differential forms and objects in vector calculus it is natural to ask about the meaning of the exterior derivative in this context. We begin with the exterior derivative of a zero-form which is given by

$$d\omega_0 = df = \sum_{i=1}^{3} \frac{\partial f}{\partial x^i} dx^i = (\nabla f) \cdot ds \; . \tag{A.31}$$

Hence, the exterior derivative acting on three-dimensional zero forms corresponds to the gradient of the associated function. What about one-forms?

$$\begin{aligned} d\omega_1 &= \sum_{i=1}^{3} dA_i \wedge dx^i = \sum_{i,j=1}^{3} \frac{\partial A_i}{\partial x^j} dx^j \wedge dx^i \\ &= \left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3} \right) dx^2 \wedge dx^3 + \left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1} \right) dx^3 \wedge dx^1 + \left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right) dx^1 \wedge dx^2 \\ &= (\nabla \times A) \cdot dS \; . \end{aligned} \tag{A.32}$$

Evidently, the exterior derivative of a three-dimensional one-form corresponds to the curl of the associated vector field. Finally, for a differential two-form we have

$$d\omega_2 = \sum_{i=1}^{3} dB_i \wedge dS^i = \sum_{i,j=1}^{3} \frac{\partial B_i}{\partial x^j} dx^j \wedge dS^i = (\nabla \cdot B) \, d^3x \; , \tag{A.33}$$

and, hence, its exterior derivative corresponds to the divergence of the associated vector field. There are no non-trivial four-forms in three dimensions so for the exterior derivative of a three-form we have, of course

$$d\omega_3 = 0 \; . \tag{A.34}$$

These results can be summarised as follows. For three-dimensional differential forms the exterior derivative acts as follows

$$0 \longrightarrow \Omega^0(U) \xrightarrow{d} \Omega^1(U) \xrightarrow{d} \Omega^2(U) \xrightarrow{d} \Omega^3(U) \xrightarrow{d} 0 \; , \tag{A.35}$$

where we recall that $\Omega^k(U)$ denotes all (infinitely many times differentiable) k-forms on U. The same diagram but expressed in the language of three-dimensional vector calculus reads:

$$0 \longrightarrow \mathcal{C}^\infty(U) \xrightarrow{\text{grad}=\nabla} \mathcal{V}(U) \xrightarrow{\text{curl}=\nabla\times} \mathcal{V}(U) \xrightarrow{\text{div}=\nabla\cdot} \mathcal{C}^\infty(U) \longrightarrow 0 \; . \tag{A.36}$$

The somewhat strange and ad-hoc differential operators of three-dimensional vector calculus are perfectly natural when understood from the viewpoint of differential forms. They are simply all versions of the exterior derivative. It is well-known in three-dimensional vector calculus that $\nabla \times (\nabla f) = 0$ and $\nabla \cdot (\nabla \times B) = 0$, that is, carrying out adjacent maps in the diagram (A.36) one after the other gives zero. From the point of view of differential forms, these equations are just manifestations of the general property $d^2 = 0$ in Lemma A.3. From Def. A.8, an exact differential one-form w_1 is one that can be written as $w_1 = df$, where f is a function. For the corresponding

vector field A the same property is called "conservative", meaning the vector field can be written as $A = \nabla f$.

In essence, vector calculus is the calculus of differentials forms written in different (and some might say, awkward) notation.

Exercise A.2. *Write the vector fields $A = (x^2, y, z^3)^T$ and $B = (y, -x, 0)^T$ on $\mathbb{R}^3$ as differential one-forms and use the exterior derivative to calculate their curl. Next, write A and B as differential two-forms and use the exterior derivative to compute their divergence.*

Exercise A.3. *Repeat the above discussion of differential forms in three dimensions for the case of differential forms in four dimensions.*

Example A.1. Some explicit differential forms
Consider the following differential one-forms in $\mathbb{R}^2$ and $\mathbb{R}^3$:

$$\omega = xdy + ydx \ , \qquad \nu = xdy - ydx \ , \qquad \mu = xdx + ydy + zdz \ . \tag{A.37}$$

Are these one-forms closed? A simple calculation based on the definition, Eq. (A.26), of the exterior derivative gives

$$\begin{aligned} d\omega &= dx \wedge dy + dy \wedge dx = 0 \\ d\nu &= dx \wedge dy - dy \wedge dx = 2dx \wedge dy \\ d\mu &= dx \wedge dx + dy \wedge dy + dz \wedge dz = 0 \ . \end{aligned}$$

It follows that ω and μ are closed, while ν is not closed. Poincaré's Lemma applies on $\mathbb{R}^n$ so this means that ω and μ must also be exact, that is, we must be able to find functions f and g such that $\omega = df$ and $\mu = dg$. It is easy to verify, using Eq. (A.27), that suitable functions are given by $f(x, y) = xy$ and $g(x, y, z) = (x^2 + y^2 + z^2)/2$.

Hodge star and Laplacian

Coming back to the case of arbitrary dimensions, n, there is another important operation, called the *Hodge star*, which maps k forms to $n - k$ forms. For a k-form ω as in Eq. (A.25) it is defined by

$$\star\omega = \frac{1}{k!(n-k)!} \epsilon^{i_1 \cdots i_k}{}_{i_{k+1} \cdots i_n} \omega_{i_1 \cdots i_k} dx^{i_{k+1}} \wedge \cdots \wedge dx^{i_n} \ , \tag{A.38}$$

where $\epsilon_{i_1 \cdots i_n}$ is the Levi-Civita tensor in n dimensions [11]. The Hodge star can be used to introduce another derivative $d^\dagger : \Omega^k(U) \to \Omega^{k-1}(U)$ (mapping k-forms to $k-1$ forms) by $d^\dagger := \star d \star$. A related differential operator is the Laplacian $\Delta : \Omega^k(U) \to \Omega^k(U)$ for differential forms which is defined by

$$\Delta = d^\dagger d + d d^\dagger \ . \tag{A.39}$$

Exercise A.4. *Verify that the Laplacian (A.39) when acting on zero-forms, that is functions, does indeed coincide with the usual Laplacian.*

[11] If we have a non-trivial metric g_{ij} we have to be a bit more careful. In this case we can define the Levi-Civita tensor as $\epsilon^{i_1 \cdots i_n} = \frac{1}{\sqrt{|\det(g)|}} \hat{\epsilon}^{i_1 \cdots i_n}$ where $\hat{\epsilon}^{i_1 \cdots i_n}$ is the "pure number" epsilon. Then Eq. (A.38) remains valid with the understanding that indices on the ϵ tensor are lowered with g_{ij}.

Application A.3. Maxwell's equations with differential forms
Maxwell's equations contain three vector fields, the electric field E, the magentic field B and the current density J, plus a scalar, the charge density ρ. We know from our discussion of three-dimensional differential forms above that three-dimensional vector fields can be represented by either one-forms or two-forms, so we have a choice. It turns out it is convenient (that is, it makes the equations look simpler) if we write E and J as one-forms and B as a two-form. With the notation introduced in Application A.2, this means

$$\mathcal{E} := E \cdot ds \,, \qquad \mathcal{B} := B \cdot dS \,, \qquad \mathcal{J} := J \cdot ds \,. \tag{A.40}$$

It is also useful to work out the Hodge star of these fields. It follows from its definition (A.38) that

$$\star 1 = dV \,, \qquad \star ds = dS \,, \qquad \star dS = ds \,, \qquad \star dV = 1 \,. \tag{A.41}$$

and this leads immediately to

$$\star\mathcal{E} = E \cdot dS \,, \qquad \star\mathcal{B} = B \cdot ds \,. \tag{A.42}$$

Using the relations between curl/divergence and the exterior derivative in Eqs. (A.32) and (A.33) we have

$$d\mathcal{E} = (\nabla \times E) \cdot dS \,, \quad d\mathcal{B} = (\nabla \cdot B) dV \,, \quad d^\dagger\mathcal{E} = \nabla \cdot E \,, \quad d^\dagger\mathcal{B} = (\nabla \times B) \cdot ds \,. \tag{A.43}$$

Now we are ready to convert Maxwell's equations. In vector calculus notation they are given by

$$\begin{aligned} \nabla \cdot E &= 4\pi\rho \\ \nabla \times B - \tfrac{1}{c}\dot{E} &= \tfrac{4\pi}{c} J \\ \nabla \times E + \tfrac{1}{c}\dot{B} &= 0 \\ \nabla \cdot B &= 0 \,, \end{aligned} \tag{A.44}$$

where the dot stands for the time derivative $\frac{\partial}{\partial t}$ and c is the speed of light. Multiplying these equations with 1, ds, dS and dV, respectively, and using Eqs. (A.43) they are easily converted to

$$\begin{aligned} d^\dagger\mathcal{E} &= 4\pi\rho \\ d^\dagger\mathcal{B} - \tfrac{1}{c}\dot{\mathcal{E}} &= \tfrac{4\pi}{c} \mathcal{J} \\ d\mathcal{E} + \tfrac{1}{c}\dot{\mathcal{B}} &= 0 \\ d\mathcal{B} &= 0 \,. \end{aligned} \tag{A.45}$$

This is by no means the most elegant form of Maxwell's equations. By using differential forms in three dimensions we are treating the three spatial dimensions and time on different footing.

It is much more natural to think of differential forms in four dimensions with coordinates $(x^\mu) = (x^0 = t, x^i)$ and basis differentials $(dx^\mu) = (dx^0 = dt, dx^i)$. Then, the fields $\mathcal{E}$ and $\mathcal{B}$ can be combined into a four-dimensional two-form $\mathcal{F}$ (called the **field-strength tensor**), and the charge density ρ and the current $\mathcal{J}$ into a four-dimensional one-form j (called the **four-current)**. More explicitly, these quantities are defined by (for simplicity, setting $c = 1$ from now on)

$$.\mathcal{F} = \mathcal{E} \wedge dt + \mathcal{B} =: \frac{1}{2}\mathcal{F}_{\mu\nu} dx^\mu \wedge dx^\nu \,, \qquad j = \rho\, dt + \mathcal{J} =: j_\mu dx^\mu \,, \tag{A.46}$$

where the second equalities define the components $\mathcal{F}_{\mu\nu}$ and j_μ. Maxwell's equations (A.45) in three-dimensional form can be converted into the language of four-dimensional differential forms and be written in terms of $\mathcal{F}$ and j.

To do this we have to be careful to distinguish between operations on three-dimensional and four-dimensional differential forms. The exterior derivative in four-dimensions is denoted by $d = dx^\mu \partial_\mu \wedge$ while we now denote its three-dimensional counterpart by $d_3 = dx^i \partial_i \wedge$. Likewise, the four-dimensional Hodge star $\star$ is defined by Eq. (A.38) (using as metric the Lorentz metric $\eta_{\mu\nu}$) with the four-dimensional Levi-Civita tensor while its three-dimensional counterpart, now denoted by $\star_3$, is defined by Eq. (A.38) with the three-dimensional Levi-Civita tensor. For a three-dimensional k-form w we then have $\star(dt \wedge w) = \star_3 w$ and $\star w = (-1)^{k+1} dt \wedge \star_3 w$. Using these rules, together with the product rule (A.28) for differential forms we find

$$\begin{aligned} \mathcal{F} &= \mathcal{E} \wedge dt + \mathcal{B} & \qquad \star\mathcal{F} &= -\star_3 \mathcal{E} + dt \wedge \star_3 \mathcal{B} \\ d\mathcal{F} &= \left(d_3 \mathcal{E} + \dot{\mathcal{B}}\right) \wedge dt + d_3 \mathcal{B} & \qquad d^\dagger \mathcal{F} &= d_3^\dagger \mathcal{E}\, dt + d_3^\dagger \mathcal{B} - \dot{\mathcal{E}} \end{aligned} \tag{A.47}$$

Comparing the last two of these equations with the three-dimensional version of Maxwell equations (A.45) (and remembering that d in those equations is d_3 in our new notation and we have set $c = 1$) we find that, equivalently, they can be written as

$$d^\dagger \mathcal{F} = 4\pi j \,, \qquad d\mathcal{F} = 0 \,. \tag{A.48}$$

Eqs. (A.48) are referred to as the covariant form of electro-magnetism and 'covariant' here refers to Lorentz transformations. If we transform covariant and contravariant tensor as we normally do in Special Relativity (for example $dx^\mu \mapsto \Lambda^\mu{}_\nu dx^\nu$ or $F_{\mu\nu} \mapsto \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma F_{\rho\sigma}$) then expressions with all indices contracted (between upper and lower indices) are Lorenz invariant. This means that $\mathcal{F} = \frac{1}{2}\mathcal{F}_{\mu\nu} dx^\mu \wedge dx^\nu$ is Lorentz-invariant, as is $j = j_\mu dx^\mu$ and $d = dx^\mu \partial_\mu \wedge$. It follows that Maxwell's equations in the form (A.48) are expressed entirely in terms of Lorentz-invariant quantities and are, therefore, Lorentz-invariant themselves. In other words, Maxwell's theory is Lorentz-invariant and, hence, compatible with Special Relativity. This was not obvious in the three-dimensional formulation (A.45) but it is manifest in the four-dimensional one (A.48).

Covariant electro-magnetism in the form (A.48) is already quite elegant but it can be simplified further. Note that the second Eq. (A.48) states that $\mathcal{F}$ is closed. This means if the conditions of Poincaré's theorem are satisfied (and, in particular, locally) we can write $\mathcal{F} = d\mathcal{A}$ for a one-form $\mathcal{A} = \mathcal{A}_\mu dx^\mu$. This one-form is called the *vector potential* or **gauge field**. From Eq. (A.46), the field strength $\mathcal{F}$ contains the electric and the magnetic fields and is, hence, the physical field. On the other hand, the vector potential $\mathcal{A}$ contains unphysical degrees of freedom, that is, different potentials $\mathcal{A}$ can lead to the same $\mathcal{F}$. This can be seen by changing $\mathcal{A}$ by a **gauge transformation** which is defined as

$$\mathcal{A} \mapsto \mathcal{A}' = \mathcal{A} + d\lambda \,, \tag{A.49}$$

where λ is a function. Under such a gauge transformation the field strength is unchanged since

$$\mathcal{F} \mapsto \mathcal{F}' = d\mathcal{A}' = d\mathcal{A} + d^2\lambda = d\mathcal{A} = \mathcal{F} \,.$$

Note that this is a direct consequence of the property $d^2 = 0$ of the exterior derivative, see Prop. A.3. In conclusion, since $\mathcal{A}$ and $\mathcal{A}'$, related by a gauge transformation (A.49), lead to the same field strength tensor $\mathcal{F}$ they describe the same physics.

The gauge transformation (A.49) can be used to choose a vector potential which satisfies an additional condition (without affecting any of the physics as encoded in the gauge-invariant $\mathcal{F}$). One such condition, referred to as **Lorentz gauge**, is

$$d^\dagger \mathcal{A} = 0 \; . \tag{A.50}$$

Why can this be done? Suppose that $\mathcal{A}$ does not satisfy the Lorentz gauge condition (A.50). Then we perform a gauge transformation (A.49) and demand that the new (physically equivalent) gauge field $\mathcal{A}'$ satisfies the Lorentz condition, that is, $d^\dagger \mathcal{A}' = 0$. This can be accomplished if the gauge transformation is carried out with a function λ which satisfies $d^\dagger d\lambda = \Delta\lambda = -d^\dagger A$, that is, with a λ which satisfies a certain Poisson equation [a].

What form do Maxwell's equations (A.48) take when expressed in terms of the gauge field $\mathcal{A}$? Since $\mathcal{F} = d\mathcal{A}$ and $d^2 = 0$, the second, homogeneous Maxwell equation in (A.48) is automatically satisfied. The first, inhomogeneous Maxwell equation (A.48) becomes

$$d^\dagger d\mathcal{A} = 4\pi j \qquad \xrightarrow{d^\dagger \mathcal{A}=0} \qquad \Delta\mathcal{A} = 4\pi j \; . \tag{A.51}$$

That is, expressed in terms of the gauge field $\mathcal{A}$ and choosing the Lorentz gauge (A.50) electromagnetism is described by the single equation $\Delta\mathcal{A} = 4\pi j$. It does not get any simpler.

[a] Note that for zero forms λ we have $d^\dagger\lambda = 0$ and, hence, from Eq. (A.39), $d^\dagger d\lambda = \Delta\lambda$.

A.4 Integration of differential forms and Stokes's theorem

We have already seen that differential one-forms are designed to be integrated over one-dimensional submanifolds, that is, over curves. This suggests differential k-forms can be integrated over k-dimensional sub-manifolds and we would now like to discuss how this works.

Definition of integral and general Stokes's theorem

Start with an open set $W \subset \mathbb{R}^n$, a differential k-form w on U and a k-dimensional manifold $M \subset W$, given by a parametrisation $X : U \to W$, with parameters $t = (t_1, \ldots, t_k)^T \in U$. At every point $X(t)$ on the manifold M we have an n-dimensional tangent space $T_{X(t)}U$ with a k-dimensional subspace $T_{X(t)}M$, the tangent space to the manifold M at $X(t)$, the latter being spanned by the k tangent vectors $\frac{\partial X}{\partial t_i}(t) \in T_{X(t)}W$. At each such point $X(t)$ the differential k-form ω provides an alternating k-form $\omega_{X(t)}$ which takes k vector arguments and can, hence, act on the k tangent vectors to the manifold at that point. This observation is used to define the integral of the differential k form ω over the k-dimensional manifold M as [12]

$$\int_M \omega := \int_U \omega_{X(t)} \left(\frac{\partial X}{\partial t_1}(t), \ldots, \frac{\partial X}{\partial t_k}(t) \right) dt^k \; . \tag{A.52}$$

The integral over differential k-forms relates to the exterior derivative in an interesting way.

Theorem A.5. *(General Stokes's theorem) Let ω be a continuously differentiable k-form on $U \subset \mathbb{R}^n$ and $M \subset U$ a $(k+1)$-dimensional (oriented) manifold with a (smooth) boundary ∂M (with the induced orientation). Then*

$$\int_{\partial M} \omega = \int_M d\omega \; . \tag{A.53}$$

Proof. A proof can be found in analysis textbooks, for example, Ref. [7]. □

[12] We are glossing over the fact that the manifold might consist of several charts. The definition of the integral then involves a sum over these, patched together by a "partition of unity". This has already been discussed in Section (12.1) and what has been said there can be straightforwardly applied to the present situation.

Stokes's theorem as above is very general and powerful. In particular, it contains the integral theorems you have heard about in year one as special cases.

Stokes's theorem in three dimensions

First consider a surface $\mathcal{S} \subset W \subset \mathbb{R}^3$, bounded by the curve $\partial\mathcal{S}$ and a differential one-form $\omega = A \cdot ds$ on W. From Eq. (A.32) we know that its exterior derivative can be written as

$$d\omega = (\nabla \times A) \cdot dS \; . \tag{A.54}$$

Inserting all this into Stokes's theorem (A.53) leads to

$$\int_{\partial\mathcal{S}} A \cdot ds = \int_{\mathcal{S}} (\nabla \times A) \cdot dS \; . \tag{A.55}$$

This is, of course the integral theorem in three dimensions also known as Stokes's theorem (in the narrow sense).

Gauss's theorem in three dimensions

Next, let $\mathcal{V} \subset W \subset \mathbb{R}^3$ be a three-dimensional manifold with bounding surface $\partial\mathcal{V}$ and $\omega = B \cdot dS$ a differential two-form on W. From Eq. (A.33) we know its exterior derivative can be written as

$$d\omega = (\nabla \cdot B)\, d^3x \; . \tag{A.56}$$

Inserting into Stokes's theorem (A.53) gives

$$\int_{\partial\mathcal{V}} B \cdot dS = \int_{\mathcal{V}} (\nabla \cdot B)\, d^3x \; , \tag{A.57}$$

and this is, of course, known as Gauss's theorem.

The general form of Stokes's theorem is not limited to three dimensions. Let us consider a two-dimensional case.

Stokes's theorem in two dimensions

We want to consider differential forms on $W \subset \mathbb{R}^2$. We use coordinates $\mathbf{x} = (x, y)^T$ and basis differential forms (dx, dy). In analogy with the three-dimensional case we introduce

$$ds = (dx, dy)^T \; , \qquad da = dx \wedge dy \; . \tag{A.58}$$

Hence, we can write a differential one-form ω on W as

$$\omega = A \cdot ds \; , \qquad A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \; , \tag{A.59}$$

where we can think of A as a vector field in two dimensions. A quick calculation of the exterior derivative $d\omega$ gives

$$d\omega = \left(\frac{\partial A_1}{\partial y} - \frac{\partial A_2}{\partial x} \right) da = \mathrm{curl}(A)\, da \; . \tag{A.60}$$

For a two-dimensional manifold $\mathcal{V} \subset W \subset \mathbb{R}^2$ with bounding curve $\partial\mathcal{V}$ Stokes's theorem (A.53) then becomes

$$\int_{\partial\mathcal{V}} A \cdot ds = \int_{\mathcal{V}} \mathrm{curl}(A)\, da \; . \tag{A.61}$$

Gauss's theorem in n dimensions
Let us consider an open set $W \subset \mathbb{R}^n$, coordinates $x = (x^1, \ldots, x^n)$ and basis differentials $(dx^1, \ldots, dx^n)$. We define

$$\begin{aligned} dS &= (dS^1, \ldots, dS^n) \,, \qquad dS^i = (-1)^{i+1} dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n \\ d^n x = dV &= dx^1 \wedge \ldots \wedge dx^n \,, \end{aligned}$$

where the hat over dx^i indicates that this differential should be omitted from the wedge product. Hence, the dS^i are wedge products of all the dx^j, except for $j = i$. A differential $n-1$-form ω on U can be written as

$$\omega = B \cdot dS \quad \Rightarrow \quad d\omega = (\nabla \cdot B)\, d^n x \,, \tag{A.62}$$

where $B = (B_1, \ldots, B_n)^T$ is an n-dimensional vector field and $\nabla \cdot B = \sum_{i=1}^n \frac{\partial B_i}{\partial x^i}$ is the n-dimensional version of the divergence. With an n-dimensional manifold $\mathcal{V} \subset W \subset \mathbb{R}^n$, bounded by an $(n-1)$-dimensional hyper-surface $\partial\mathcal{V}$, Stokes's theorem (A.53) becomes

$$\int_{\partial\mathcal{V}} B \cdot dS = \int_{\mathcal{V}} (\nabla \cdot B)\, d^n x \,. \tag{A.63}$$

Exercise A.6. *Derive a version of Stoke's theorem in four dimensions, which relates integrals of a two-form over surfaces $S \subset \mathbb{R}^4$ to integrals of a one-form over the boundary curve ∂S.*

Literature

[1] **K. F. Riley, M. P. Hobson and S. J. Bence, "Mathematical Methods for Physics and Engineering", CUP 2002.**
This is the recommended book for the mathematics component of the physics course. As the title suggests this is a "hands-on" book, strong on explaining methods and concrete applications, rather weaker on presenting a coherent mathematical exposition.

[2] **John David Jackson, "Classical Electrodynamics", Wiley, 2012.**
For me the ultimate book on electrodynamics. In addition to a comprehensive coverage of the subject and many physical applications it also explains many of the required mathematical methods informally but efficiently.

[3] **Serge Lang, "Real and Functional Analysis", Springer 1993.**
A high quality mathematics book covering analysis and functional analysis at an advanced level.

[4] **Serge Lang, "Calculus of Several Variables", Addison Wesley 1971.**
A mathematical book on the subject but kept at a relatively informal and practical level.

[5] **Andre Lukas, "The Oxford Linear Algebra for Scientists", OUP 2022.**

[6] **Michela Petrini, Gianfranco Pradisi, Alberto Zaffaroni, "A Guide to Mathematical Methods for Physicists: With Problems and Solutions", World Scientific, 2017.**
A nice book covering many of the main pieces of mathematics crucial to physics, including complex functions, integration theory and functional analysis, taking the mathematics seriously but without being overly formal.

[7] **Serge Lang, "Undergraduate Analysis", Springer 1997.**
A very nice book on the subject, not overly formal and very useful to fill in inevitable mathematical gaps.

[8] **Serge Lang, "Complex Analysis", Springer 1999.**
A great book an complex analysis - suitable for self-study and to complement the short option on complex functions.

[9] **Mikio Nakahara, "Geometry, Topology and Physics", Taylor and Francis, 2013.**
An excellent textbook, written for physicists, but at a more advanced level, discussing topology, differential geometry and how it relates to physics. This is not really relevant for the first two parts of this lecture but it given a more advanced account of manifolds than part 3 of these lectures. If you are interested in the interplay of physics and geometry (and you would perhaps like to learn about the mathematics underlying General Relativity) this is the book for you.

[10] **J. J. Duistermaat and J. A. C. Kolk, Multidimensional Real Analysis I and II, CUP 2004.**