Mathematical Methods

Prof Andre Lukas

Rudolf Peierls Centre for Theoretical Physics
University of Oxford

MT 2018, updated MT 2023

Contents

Foreword: Lecturing a Mathematical Methods course to physicists can be a tricky affair and following such a course as a second year student may be even trickier. The traditional material for this course consists of the classical differential equations and associated special function solutions of Mathematical Physics. In a modern context, both in Mathematics and in Physics, these subjects are increasingly approached in the appropriate algebraic setting of Banach and Hilbert Spaces. The correct setting for Quantum Mechanics is provided by Hilbert Spaces and for this reason alone they are a mandatory subject and should at least receive a rudimentary treatment in a course on Mathematical Methods for Physicists. However, the associated mathematical discipline of Functional Analysis merits a lecture course in its own right and cannot possibly be treated comprehensively in a course which also needs to cover a range of applications. What is more, physics students may not yet have come across some of the requisite mathematics, such as the notion of convergence and the definition of integrals. All of this places an additional overhead on introducing mathematical key ideas, such as the idea of a Hilbert Space.

As a result of these various difficulties and requirements Mathematical Methods courses often end up as collections of various bits of Mathematical Physics, seemingly unconnected and without any guiding ideas, other than the apparent usefulness for solving some problems in Physics. Sometimes, ideas developed in the context of finite-dimensional vector spaces are used as guiding principles but this ignores the crucial differences between finite and infinite-dimensional vector spaces, to do with issues of convergence.

These lecture notes reflect the attempt to provide a modern Mathematical Physics course which presents the underlying mathematical ideas as well as their applications and provides students with an intellectual framework, rather than just a "how-to-do" toolkit. We begin by introducing the relevant mathematical ideas, including Banach and Hilbert Spaces but keep this at a relatively low level of formality and quite stream-lined. On the other hand, we will cover the "traditional" subjects related to differential equations and special functions but attempt to place these into the general mathematical context. Sections with predominantly mathematical background material are indicated with a star. While they are important for a deep understanding of the material they are less essential for the relatively basic practical tasks required to pass an exam. I believe ambitious, mathematically interested student can benefit from the combination of mathematical foundation and applications in these notes. Students who want to focus on the practical tasks may concentrate on the un-starred sections.

Two somewhat non-traditional topics, distributions and groups, have beed added. Distributions are so widely used in physics - and physicists tend to discuss important ideas such as Green functions using distributions - that they shouldn't be omitted from a Mathematical Physics course. Symmetries have become one of the central ideas in physics and they are underlying practically all fundamental theories of physics. It would, therefore, be negligent, in a course on Mathematical Methods, not to introduce the associated mathematical ideas of groups and representations.

The three appendices are pure bonus material. The first one is a review of the calculus of multiple variables, the second a simple account of sub-manifolds in $\mathbb{R}^n$, including curves and surfaces in $\mathbb{R}^3$, as encountered in vector calculus. Inevitably, it also does some of the groundwork for General Relativity - so certainly worthwhile for anyone who would like to learn Einstein's theory of gravity. The third appendix introduces differential forms, a classical topic in mathematical physics, at an elementary level. Read (or ignore) at your own leisure.

Andre Lukas
Oxford, 2018

1 Mathematical preliminaries

This section provides some basic mathematical background which is essential for the lecture and can also be considered as part of the general mathematical language every physicist should be familiar with. The part on vector spaces is (mainly) review and will be dealt with quite quickly - a more detailed treatment can be found in the first year lecture notes on Linear Algebra. The main mathematical theme of this course is the study of infinite-dimensional vector spaces and practically every topic we cover can (and should) be understood in this context. While the first year course on Linear Algebra dealt with finite-dimensional vector spaces many of the concepts were introduced without any reference to dimension and straightforwardly generalise to the infinite-dimensional case. These include the definitions of vector space, sub-vector space, linear maps, scalar products and norms and we begin by briefly reviewing those ideas.

One of the concepts which does not straightforwardly generalise to the infinite-dimensional case is that of a basis. We know that a finite-dimensional vector space V (over a field F) has a basis, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and that every vector $\mathbf{v} \in V$ can be written as a unique linear combination

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \ , \tag{1.1}$$

where $\alpha_i \in F$ are scalars. A number of complications arise when trying to generalise this to infinite dimensions. Broadly speaking, it is not actually clear whether a basis exists in this case. A basis must certainly contain an infinite number of basis vectors so that the RHS of Eq. (1.1) becomes an infinite sum. This means we have to address questions of convergence. Even if we can formulate conditions for convergence we still have to clarify whether we can find a suitable set of scalars α_i such that the sum (1.1) converges to a given vector $\mathbf{v}$. All this requires techniques from analysis (= calculus done properly) and the relevant mathematical basics will be discussed in part 2 of this section while much of Section 2 will be occupied with answering the above questions.

Finally, we need to address another mathematical issue, namely the definition of integrals. The most important infinite-dimensional vector spaces we need to consider consist of functions, with a scalar product defined by an integral. To understand these function vector spaces we need to understand the nature of the integral. In the last part of this section, we will, therefore, briefly discuss measures and the Riemann and Lebesgue integrals.

1.1 Vector spaces: (mostly) a reminder

In this subsection, we review a number of general ideas in Linear Algebra which were covered in detail in the first year course. We emphasise that, while the first year course was focused on finite-dimensional vector spaces, most of the concepts covered (and reviewed below) are actually independent of dimension and, hence, apply to the finite and the infinite-dimensional case.

1.1.1 Vector spaces and sub vector spaces

Of course we begin by recalling the basic definition of a vector space. It involves two sets, the set V which consists of what we call *vectors*, and the field F, typically taken to be either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$, whose elements are referred to as *scalars*. For these objects we have two operations, the *vector addition* which maps two vectors to a third vector, and the *scalar multiplication* which maps a scalar and a vector to a vector, subject to a number of basic axioms (= rules for calculating with vectors and scalars). The formal definition is:

Definition 1.1. *(Vector spaces) A vector space V over a field F (= $\mathbb{R}$, $\mathbb{C}$ or any other field) is a set with two operations:*

i) vector addition: $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w} \in V$, *where* $\mathbf{v}, \mathbf{w} \in V$
ii) scalar multiplication: $(\alpha, \mathbf{v}) \mapsto \alpha\mathbf{v} \in V$, *where* $\alpha \in F$ *and* $\mathbf{v} \in V$.

For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *and all* $\alpha, \beta \in F$*, these operations have to satisfy the following rules:*

(V1) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	*"associativity"*
(V2) There exists a "zero vector", $\mathbf{0} \in V$ *so that* $\mathbf{0} + \mathbf{v} = \mathbf{v}$	*"neutral element"*
(V3) There exists an inverse, $-\mathbf{v}$ *with* $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$	*"inverse element"*
(V4) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$	*"commutativity"*
(V5) $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$	
(V6) $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$	
(V7) $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$	
(V8) $1 \cdot \mathbf{v} = \mathbf{v}$	

The elements $\mathbf{v} \in V$ *are called "vectors", the elements* $\alpha \in F$ *of the field are called "scalars".*

Closely associated to this definition is the one for the "sub-structure", that is, for a sub vector space. A sub vector space is a non-empty subset $W \subset V$ of a vector space V which is closed under vector addition and scalar multiplication. More formally, this means:

Definition 1.2. *(Sub vector spaces) A sub vector space* $W \subset V$ *is a non-empty subset of a vector space* V *satisfying:*

(S1) $\mathbf{w}_1 + \mathbf{w}_2 \in W$ *for all* $\mathbf{w}_1, \mathbf{w}_2 \in W$
(S2) $\alpha\mathbf{w} \in W$ *for all* $\alpha \in F$ *and for all* $\mathbf{w} \in W$

A sub vector space satisfies all the axioms in Def. 1.1 and is, hence, a vector space in its own right. Every vector space V has two *trivial sub vector spaces*, the null vector space $\{\mathbf{0}\} \subset V$ and the total space $V \subset V$. For two sub vector spaces U and W of V the sum $U + W$ is defined as

$$U + W = \{\mathbf{u} + \mathbf{w} \,|\, \mathbf{u} \in U \,,\ \mathbf{w} \in W\} \,. \tag{1.2}$$

Evidently, $U + W$ is also a sub vector space of V as shown in the following

Exercise 1.1. *Show that the sum* (1.2) *of two sub vector spaces is a sub vector space.*

A sum $U + W$ of two sub vector spaces is called *direct* iff $U \cap W = \{\mathbf{0}\}$ and a direct sum is written as $U \oplus W$.

Exercise 1.2. *Show that the sum* $U + W$ *is direct iff every* $\mathbf{v} \in U + W$ *has a unique decomposition* $\mathbf{v} = \mathbf{u} + \mathbf{w}$*, with* $\mathbf{u} \in U$ *and* $\mathbf{w} \in W$*.*

Exercise 1.3. *Show that a sub sector space is a vector space.*

There are a number of basic notions for vector spaces which include *linear combinations*, *span*, *linear independence* and *basis*. Let us briefly recall how they are defined. For k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V over a field F the expression

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k = \sum_{i=1}^{k} \alpha_i\mathbf{v}_i \,, \tag{1.3}$$

with scalars $\alpha_1, \ldots, \alpha_n \in F$, is called a *linear combination*. The set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$,

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) := \left\{ \sum_{i=1}^{k} \alpha_i\mathbf{v}_i \,|\, \alpha_i \in F \right\} \,, \tag{1.4}$$

is called the *span* of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Linear independence is defined as follows.

Definition 1.3. *Let V be a vector space over F and $\alpha_1, \ldots, \alpha_k \in F$ scalars. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ is called linearly independent if*

$$\sum_{i=1}^{k} \alpha_i \mathbf{v}_i = \mathbf{0} \implies \text{all } \alpha_i = 0 \, . \tag{1.5}$$

Otherwise, the vectors are called linearly dependent. That is, they are linearly dependent if $\sum_{i=1}^{k} \alpha_i \mathbf{v}_i = \mathbf{0}$ has a solution with at least one $\alpha_i \neq 0$.

If a vector space V is spanned by a finite number of vectors (that is, every $\mathbf{v} \in V$ can be written as a linear combination of these vectors) it is called *finite-dimensional*, otherwise *infinite-dimensional*. Recall the situation for finite-dimensional vector spaces. In this case, we can easily define what is meant by a *basis*.

Definition 1.4. *A set $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ of vectors is called a basis of V iff:*

(B1) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *are linearly independent.*
(B2) $V = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$

The number of elements in a basis is called the *dimension*, $\dim(V)$ of the vector space. Every vector $\mathbf{v} \in V$ can then be written as a unique linear combination of the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, that is,

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \, , \tag{1.6}$$

with a unique choice of $\alpha_i \in F$ for a given vector $\mathbf{v}$. The α_i are also called the *coordinates* of the vector $\mathbf{v}$ relative to the given basis.

Clearly, everything is much more involved for infinite-dimensional vector spaces but the goal is to generalise the concept of a basis to this case and have an expansion analogous to Eq. (1.6), but with the sum running over an infinite number of basis elements. Making sense of this requires a number of mathematical concepts, including that of convergence, which will be developed in this section.

1.1.2 Examples of vector spaces

The most prominent examples of finite dimensional vector spaces are the column vectors

$$\mathbb{F}^n = \left\{ \left(\begin{array}{c} v_1 \\ \vdots \\ v_n \end{array} \right) \middle| \, v_i \in F \right\} \, , \tag{1.7}$$

over the field F (where, usually, either $F = \mathbb{R}$ for real column vectors or $F = \mathbb{C}$ for complex column vectors), with vector addition and scalar multiplication defined "entry-by-entry" as

$$\left(\begin{array}{c} v_1 \\ \vdots \\ v_n \end{array} \right) + \left(\begin{array}{c} w_1 \\ \vdots \\ w_n \end{array} \right) := \left(\begin{array}{c} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{array} \right) \, , \qquad \alpha \left(\begin{array}{c} v_1 \\ \vdots \\ v_n \end{array} \right) := \left(\begin{array}{c} \alpha v_1 \\ \vdots \\ \alpha v_n \end{array} \right) \, . \tag{1.8}$$

Verifying that these satisfy the vector space axioms 1.1 is straightforward. A basis is given by the *standard unit vectors* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and, hence, the dimension equals n.

Here, we will also (and predominantly) be interested in more abstract vector spaces consisting of sets of functions. A general class of such function vector spaces can be defined by starting with a (any) set S and by considering all functions from S to the vector space V (over the field F). This set of functions

$$\mathcal{F}(S, V) := \{ f : S \to V \} \tag{1.9}$$

can be made into a vector space over F by defining a "pointwise" vector addition and scalar multiplication

$$(f+g)(x) := f(x) + g(x)\,, \qquad (\alpha f)(x) := \alpha f(x)\,, \tag{1.10}$$

where $f, g \in \mathcal{F}(S, V)$ and $\alpha \in F$.

Exercise 1.4. *Show that the space* (1.9) *together with vector addition and scalar multiplication as defined in Eq.* (1.10) *defines a vector space.*

There are many interesting special cases and sub vector spaces which can be obtained from this construction. For example, choose $S = [a, b] \subset \mathbb{R}$ as an interval on the real line ($a = -\infty$ or $b = \infty$ are allowed) and $V = \mathbb{R}$ (or $V = \mathbb{C}$), so that we are considering the space $\mathcal{F}([a, b], \mathbb{R})$ or $\mathcal{F}([a, b], \mathbb{C})$ of all real-valued (or complex-valued) functions on this interval. With the pointwise definitions (1.10) of vector addition and scalar multiplication these functions form a vector space.

We can consider many sub-sets of this vector space by imposing additional conditions on the functions and as long as these conditions are invariant under the addition and scalar multiplication of functions (1.10) Def. 1.2 implies that these sub-sets form sub vector spaces. For example, we know that the sum of two continuous functions as well as the scalar multiple of a continuous function is continuous so the set of all continuous functions on an interval forms a (sub) vector space which we denote by $\mathcal{C}([a, b])$. Similar statements apply to all differentiable functions on an interval and the vector space of k times (continuously) differentiable functions on the interval $[a, b]$ is denoted by $\mathcal{C}^k([a, b])$, with $\mathcal{C}^\infty([a, b])$ the space of infinitely many time differentiable functions on the interval $[a, b]$. In cases where we consider the entire real line it is sometimes useful to restrict to functions with compact support. A function f with compact support vanishes outside a certain radius $R > 0$ such that $f(x) = 0$ whenever $|x| > R$. We indicate the property of compact support with a subscript "c", so that, for example, the vector space of continuous functions on $\mathbb{R}$ with compact support is denoted by $\mathcal{C}_c(\mathbb{R})$. The vector space of all polynomials, restricted to the interval $[a, b]$ is denoted by $\mathcal{P}([a, b])$. Whether the functions are real or complex-valued is sometimes also indicated by a subscript, so $\mathcal{C}_\mathbb{R}([a, b])$ are the real-valued continuous functions on $[a, b]$ while $\mathcal{C}_\mathbb{C}([a, b])$ are their complex-valued counterparts.

Exercise 1.5. *Find at least three more examples of function vector spaces, starting with the construction* (1.9).

1.1.3 Linear maps

As for any algebraic structure, it is important to study the maps which are compatible with vector spaces, the *linear maps* [1].

Definition 1.5. *(Linear maps) A map $T : V \to W$ between two vector spaces V and W over a field F is called linear if*

(L1) $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$
(L2) $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$

for all $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$ and for all $\alpha \in F$. Further, the set $\mathrm{Ker}(T) := \{\mathbf{v} \in V \,|\, T(\mathbf{v}) = \mathbf{0}\} \subset V$ *is called the kernel of T and the set* $\mathrm{Im}(T) := \{T(\mathbf{v}) \,|\, \mathbf{v} \in V\} \subset W$ *is called the image of T.*

In the context of infinite-dimensional vector spaces, linear maps are also sometimes called *linear operators* and we will occasionally use this terminology. Recall that a linear map $T : V \to W$ always maps the zero vector of V into the zero vector of W, so $T(\mathbf{0}) = \mathbf{0}$ and that the kernel of T is a sub vector space of V

[1] We will denote linear maps by uppercase letters such as T. The letter f will frequently be used for the functions which form the elements of the vector spaces we consider.

while the image is a sub vector space of W. Surjectivity and injectivity of the linear map T are related to the image and kernel via the equivalences

$$T \text{ surjective} \quad \Leftrightarrow \quad \mathrm{Im}(T) = W \,, \qquad\qquad T \text{ injective} \quad \Leftrightarrow \quad \mathrm{Ker}(T) = \{\mathbf{0}\} \,. \tag{1.11}$$

A linear map $T : V \to W$ which is bijective (= injective and surjective) is also called a *(vector space) isomorphism* between V and W. The set of all linear maps $T : V \to W$ is referred to as the *homorphisms* from V to W and is denoted by $\mathrm{Hom}(V, W) := \{T : V \to W \,|\, T \text{ linear}\}$. By using the general construction (1.10) (where V plays the role of the set S and W the role of the vector space V) this space can be equipped with vector addition and scalar multiplication. Further, since the sum of two linear functions and the scalar multiple of a linear function are again linear, it follows from Def. 1.2 that $\mathrm{Hom}(V, W)$ is a (sub) vector space of $\mathcal{F}(V, W)$. Finally, we note that for two linear maps $T : V \to W$ and $S : W \to U$, the composition $S \circ T : V \to U$ (defined by $S \circ T(\mathbf{v}) := S(T(\mathbf{v}))$) is also linear.

The identity map $\mathrm{id} : V \to V$ defined by $\mathrm{id}(\mathbf{v}) = \mathbf{v}$ is evidently linear. Recall that a linear map $S : V \to V$ is said to be the inverse of a linear map $T : V \to V$ iff

$$S \circ T = T \circ S = \mathrm{id} \,. \tag{1.12}$$

The inverse exists iff T is bijective (= injective and surjective) and in this case it is unique, linear and denoted by T^{-1}. Also recall the following rules

$$(T^{-1})^{-1} = T \,, \qquad (T \circ S)^{-1} = S^{-1} \circ T^{-1} \,, \tag{1.13}$$

for calculating with the inverse.

For a finite-dimensional vector space V with basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ we can associate to a linear map $T : V \to V$ a matrix A with entries defined by

$$T(\mathbf{v}_j) = \sum_{i=1}^{n} A_{ij} \mathbf{v}_i \,. \tag{1.14}$$

This matrix describes the action of the linear map on the coordinate vectors relative to the basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. To see what this means more explicitly consider a vector $\mathbf{v} \in V$ with coordinate vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)^T$, such that $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$. Then, if T maps the vector $\mathbf{v}$ to $\mathbf{v} \to T(\mathbf{v})$ the coordinate vector is mapped to $\boldsymbol{\alpha} \to A\boldsymbol{\alpha}$. How does the matrix A depend on the choice of basis? Introduce a second basis $(\mathbf{v}'_1, \ldots, \mathbf{v}'_n)$ with associated matrix A'. Then we have

$$A' = PAP^{-1} \,, \qquad \mathbf{v}_j = \sum_{i=1}^{n} P_{ij} \mathbf{v}'_i \,. \tag{1.15}$$

The matrix P can also be understood as follows. Consider a vector $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \alpha'_i \mathbf{v}'_i$ with coordinate vectors $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ and $\boldsymbol{\alpha}' = (\alpha'_1, \ldots, \alpha'_n)$ relative to the unprimed and primed basis. Then,

$$\boldsymbol{\alpha}' = P\boldsymbol{\alpha} \,. \tag{1.16}$$

An important special class of homomorphisms is the *dual vector space* $V^* := \mathrm{Hom}(V, F)$ of a vector space V over F. The elements of the dual vector space are called *(linear) functionals* and they map vectors to numbers in the field F. For a finite-dimensional vector space V with basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, there exists a basis $\varphi_1, \ldots, \varphi_n$ of V^*, called the *dual basis*, which satisfies

$$\varphi_i(\mathbf{v}_j) = \delta_{ij} \,. \tag{1.17}$$

In particular, a finite-dimensional vector space and its dual have the same dimension. For infinite-dimensional vector spaces the discussion is of course more involved and we will come back to this later.

Exercise 1.6. *For a finite-dimensional vector space V with basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ show that there exists a basis $\varphi_1, \ldots, \varphi_n$ of the dual space V^* which satsfies Eq. (1.17).*

1.1.4 Examples of linear maps

We know that the linear maps $T : \mathbb{R}^n \to \mathbb{R}^m$ ($T : \mathbb{C}^n \to \mathbb{C}^m$) can be identified with the $m \times n$ matrices containing real entries (complex entries) whose linear action is simple realised by the multiplication of matrices with vectors.

Let us consider some examples of linear maps for vector spaces of functions, starting with the space $\mathcal{C}([a, b])$ of (real-valued) continuous functions on the interval $[a, b]$. For a (real-valued) continuous function $K : [a, b] \times [a, b] \to \mathbb{R}$ of two variables we can define the map $T : \mathcal{C}([a, b]) \to \mathcal{C}([a, b])$ by

$$T(f)(x) := \int_a^b d\tilde{x}\, K(x, \tilde{x}) f(\tilde{x}) \,. \tag{1.18}$$

This map is evidently linear since the integrand is linear in the function f and the integral itself is linear. A linear map such as the above is called a linear *integral operator* and the function K is also referred to as the *kernel of the integral operator* [2]. Such integral operators play an important role in functional analysis.

For another example consider the vector space $\mathcal{C}^\infty([a, b])$ of infinitely many times differentiable functions on the interval $[a, b]$. We can define a *linear operator* $D : \mathcal{C}^\infty([a, b]) \to \mathcal{C}^\infty([a, b])$ by

$$D(f)(x) := \frac{df}{dx}(x) \quad \text{or} \quad D = \frac{d}{dx} \,. \tag{1.19}$$

A further class of linear operators $M_p : \mathcal{C}^\infty([a, b]) \to \mathcal{C}^\infty([a, b])$ is obtained by multiplication with a fixed function $p \in \mathcal{C}^\infty([a, b])$, defined by

$$M_p(f)(x) := p(x) f(x) \,. \tag{1.20}$$

The above two classes of linear operators can be combined and generalised by including higher-order differentials which leads to linear operators $T : \mathcal{C}^\infty([a, b]) \to \mathcal{C}^\infty([a, b])$ defined by

$$T = p_k \frac{d^k}{dx^k} + p_{k-1} \frac{d^{k-1}}{dx^{k-1}} + \cdots + p_1 \frac{d}{dx} + p_0 \,, \tag{1.21}$$

where p_i, for $i = 0, \ldots, k$, are fixed functions in $\mathcal{C}^\infty([a, b])$. Linear operators of this type will play an important role in our discussion, mainly because they form the key ingredient for many of the differential equations which appear in Mathematical Physics.

1.1.5 Norms and normed vector spaces

Frequently, we will require additional structure on our vector spaces which allows us to study the "geometry" of vectors. The simplest such structure is one that "measures" the length of vectors and such a structure is called a *norm*. As we will see in the next sub section, we will require a norm to define convergence and basic ideas of topology in a vector space. The formal definition of a norm is as follows.

Definition 1.6. *(Norms and normed vector spaces) A norm $\| \cdot \|$ on a vector space V over the field $F = \mathbb{R}$ or $F = \mathbb{C}$ is a map $\| \cdot \| : V \to \mathbb{R}$ which satsifies*

(N1) $\| \mathbf{v} \| > 0$ *for all non-zero* $\mathbf{v} \in V$
(N2) $\| \alpha \mathbf{v} \| = |\alpha| \, \| \mathbf{v} \|$ *for all* $\alpha \in F$ *and all* $\mathbf{v} \in V$
(N3) $\| \mathbf{v} + \mathbf{w} \| \leq \| \mathbf{v} \| + \| \mathbf{w} \|$ *for all* $\mathbf{v}, \mathbf{w} \in V$ *(triangle inequality)*

A vector space V with a norm is also called a normed vector space.

[2] This notion of "kernel" has nothing to do with the kernel of a linear map, as introduced in Def. 1.5. The double-use of the word is somewhat unfortunate but so established that it cannot be avoided. It will usually be clear from the context which meaning of "kernel" is referred to.

Note that the notation $|\alpha|$ in (N2) refers to the simple real modulus for $F = \mathbb{R}$ and to the complex modulus for $F = \mathbb{C}$. All three axioms are intuitively clear if we think about a norm as providing us with a notion of "length". Clearly, a length should be strictly positive for all non-zero vectors as stated in (N1), it needs to scale with the (real or complex) modulus of a scalar if the vector is multiplied by this scalar as in (N2) and it needs to satisfy the *triangle inequality* (N3). Since $0\mathbf{v} = \mathbf{0}$ for any vector $\mathbf{v} \in V$, the axiom (N2) implies that $\| \mathbf{0} \| = \| 0\mathbf{v} \| = 0 \| \mathbf{v} \| = 0$, so the zero vector has norm 0 (and is, from (N1), the only vector with this property).

Exercise 1.7. *Show that, in a normed vector space V, we have $\| \mathbf{v} - \mathbf{w} \| \geq \| \mathbf{v} \| - \| \mathbf{w} \|$ for all $\mathbf{v}, \mathbf{w} \in V$.*

For normed vector spaces V and W we can now introduce an important new sub-class of linear operators $T : V \to W$, namely *bounded linear operators*. They are defined as follows [3].

Definition 1.7. *(Bounded linear operators) A linear operator $T : V \to W$ is called bounded if there exists a positive $K \in \mathbb{R}$ such that $\| T(\mathbf{v}) \|_W \leq K \| \mathbf{v} \|_V$ for all $\mathbf{v} \in V$. The smallest number K for which this condition is satisfied is called the norm, $\| T \|$, of the operator T.*

Having introduced the notion of the norm of a bounded linear operator, we can now introduce *isometries*.

Definition 1.8. *(Isometries) A bounded linear operator $T : V \to W$ is an isometry iff $\| T(\mathbf{v}) \|_W = \| \mathbf{v} \|_V$ for all $\mathbf{v} \in V$.*

1.1.6 Examples of normed vector spaces

You are already familiar with a number of normed vector spaces, perhaps without having thought about them in this more formal context. The real and complex numbers, seen as one-dimensional vectors spaces, are normed with the norm given by the (real or complex) modulus. It is evident that this satisfies the conditions (N1) and (N2) in Def. 1.6. For the condition (N3) consider the following

Exercise 1.8. *Show that the real and complex modulus satisfies the triangle inequality.*

More interesting examples of normed vector spaces are provided by $\mathbb{R}^n$ and $\mathbb{C}^n$ with the Euclidean norm

$$\| \mathbf{v} \| := \left(\sum_{i=1}^{n} |v_i|^2 \right)^{1/2} , \tag{1.22}$$

for any vector $\mathbf{v} = (v_1, \ldots, v_n)^T$. (As above, the modulus sign refers to the real or complex modulus, depending on whether we consider the case of $\mathbb{R}^n$ or $\mathbb{C}^n$.) It is immediately clear that axioms (N1) and (N2) are satisfied and we leave (N3) as an exercise.

Exercise 1.9. *Show that the prospective norm on $\mathbb{R}^n$ or $\mathbb{C}^n$ defined in Eq. (1.22) satisfies the triangle inequality.*

Linear maps $T : F^n \to F^m$ are described by the action of $m \times n$ matrices on vectors. Since such matrices, for a given linear map T, have fixed entries it is plausible that they are bounded with respect to the norm (1.22). You can attempt the proof of this statement in the following exercise.

Exercise 1.10. *Show that linear maps $T : F^n \to F^m$, where $F = \mathbb{R}$ or $F = \mathbb{C}$ are bounded, relative to the norm (1.22).*

[3] When two normed vector spaces V and W are involved we will distinguish the associated norms by adding the name of the space as a sub-script, so we write $\| \cdot \|_V$ and $\| \cdot \|_W$.

It is not too difficult to generalise this statement and to show that linear maps between any two finite-dimensional vector spaces are bounded. For the infinite-dimensional case this is not necessarily true (see Exercise 1.13 below).

Vector spaces, even finite-dimensional ones, usually allow for more than one way to introduce a norm. For example, on $\mathbb{R}^n$ or $\mathbb{C}^n$, with vectors $\mathbf{v} = (v_1, \ldots, v_n)^T$ we can define, for any real number $p \geq 1$, the norm

$$\| \mathbf{v} \|_p := \left(\sum_{i=1}^{n} |v_i|^p \right)^{1/p} . \tag{1.23}$$

Clearly, this is a generalisation of the standard norm (1.22) which corresponds to the special case $p = 2$. As before, conditions (N1) and (N2) in Def. 1.6 are easily verified. For the triangle inequality (N3) consider the following exercise.

Exercise 1.11. *For two vectors $\mathbf{v} = (v_1, \ldots, v_n)^T$ and $\mathbf{w} = (w_1, \ldots, w_n)^T$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ and two real numbers $p, q \geq 1$ with $1/p + 1/q = 1$ show that*

$$\begin{array}{lllll} (a) & \sum_{i=1}^n |v_i w_i| & \leq & (\sum_{i=1}^n |v_i|^p)^{1/p} (\sum_{i=1}^n |w_i|^q)^{1/q} & \textit{(Hölder's inequality)} \\ (b) & (\sum_{i=1}^n |v_i + w_i|^p)^{1/p} & \leq & (\sum_{i=1}^n |v_i|^p)^{1/p} + (\sum_{i=1}^n |w_i|^p)^{1/p} & \textit{(Minkowski's inequality)} \end{array} \tag{1.24}$$

Use Minkowski's inequality to show that the prospective norm (1.23) satisfies the triangle inequality.

Norms can also be introduced on infinite-dimensional vector spaces. As an example, consider the space $\mathcal{C}([a, b])$ of continuous functions on the interval $[a, b]$. The analogue of the standard norm (1.22) and its generalisation (1.23) (thinking, intuitively, about promoting the finite sum in Eq. (1.22) to an integral) for $f \in \mathcal{C}([a, b]$ can be defined as

$$\| f \| := \left(\int_a^b dx \, |f(x)|^2 \right)^{1/2} , \qquad \| f \|_p := \left(\int_a^b dx \, |f(x)|^p \right)^{1/p} , \tag{1.25}$$

for any real $p \geq 1$.

Exercise 1.12. *Show that Eq. (1.25) for any real $p \geq 1$ defines a norm on $\mathcal{C}([a, b])$.*

Exercise 1.13. *Consider the space $\mathcal{C}^\infty([0, 1])$ with norm $\| \cdot \|_2$, the monomials $p_k(x) = x^k$ and the differential operator $T = d/dx$. Compute $\| Tp_k \|_2 / \| p_k \|_2$ and use the result to show that T is not bounded.*

1.1.7 Scalar products

A normed vector space provides a basic notion of geometry in that it assigns a "length" to each vector. Often it is desirable to have a more comprehensive framework for geometry which also allows measuring angles between vectors and defining the concept of orthogonality. Such a framework is provided by a *scalar product* or *inner product* on a vector space which is defined as follows.

Definition 1.9. *A real scalar product on a vector space V over $F = \mathbb{R}$ and a hermitian scalar product on a vector space V over the field $F = \mathbb{C}$ is a map $\langle \cdot, \cdot \rangle : V \times V \to F$ which satisfies*

(S1) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$, *for a real scalar product, $F = \mathbb{R}$*
$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle^*$, *for a hermitian scalar product, $F = \mathbb{C}$*
(S2) $\langle \mathbf{v}, \alpha \mathbf{u} + \beta \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$
(S3) $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ *if* $\mathbf{v} \neq \mathbf{0}$

for all vectors $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$ and all scalars $\alpha, \beta \in F$.

A real or hermitian scalar product is also referred to as an inner product on V and a vector space V with such a scalar product is also called an inner product (vector) space.

Note that, from property (S2), the scalar product is linear in the second argument and combining this with (S1) implies for the first argument that

$$\begin{array}{rcll} \langle \alpha \mathbf{v} + \beta \mathbf{u}, \mathbf{w} \rangle & = & \alpha \langle \mathbf{v}, \mathbf{w} \rangle + \beta \langle \mathbf{u}, \mathbf{w} \rangle & F = \mathbb{R} \\ \langle \alpha \mathbf{v} + \beta \mathbf{u}, \mathbf{w} \rangle & = & \alpha^* \langle \mathbf{v}, \mathbf{w} \rangle + \beta^* \langle \mathbf{u}, \mathbf{w} \rangle & F = \mathbb{C} \end{array} \tag{1.26}$$

Evidently, in the real case the scalar product is also linear in the first argument (and, hence, it is *bi-linear*). In the complex case, it is *sesqui-linear* which means that, in addition to linearity in the second argument, it is *half-linear* in the first argument (vector sums can be pulled out of the first argument while scalars pull out with a complex conjugate). In the following, we will frequently write equations for the hermitian case, $F = \mathbb{C}$, keeping in mind that the analogous equations for the real case can be obtained by simply omitting the complex conjugate.

How are inner product vector spaces and normed vector spaces related? Properties (S1) and (S3) imply that $\langle \mathbf{v}, \mathbf{v} \rangle$ is always real and positive so it makes sense to try to define a norm by

$$\| \mathbf{v} \| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \, . \tag{1.27}$$

As usual, it is easy to show that this satisfies properties (N1) and (N2) in Def. 1.6. To verify the triangle inequality (N3) we recall that every scalar product satisfies the *Cauchy-Schwarz* inequality

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \| \mathbf{v} \| \, \| \mathbf{w} \| \qquad \Longrightarrow \qquad \| \mathbf{v} + \mathbf{w} \| \leq \| \mathbf{v} \| + \| \mathbf{w} \| \, , \tag{1.28}$$

from which the triangle inequality follows immediately. In conclusion, Eq. (1.27) does indeed define a norm in the sense of Def. 1.6 and it is called the *norm associated to the scalar product*. Hence, any inner product vector space is also a normed vector space.

Exercise 1.14. *Show that a (real or hermitian) scalar product with associated norm* (1.27) *satisfies the Cauchy-Schwarz inequality and the triangle inequality in Eq.* (1.28)*. Also show that the norm* (1.27) *satisfies the parallelogram law*

$$\| \mathbf{v} + \mathbf{w} \|^2 + \| \mathbf{v} - \mathbf{w} \|^2 = 2 \left(\| \mathbf{v} \|^2 + \| \mathbf{w} \|^2 \right) \, , \tag{1.29}$$

for all $\mathbf{v}, \mathbf{w} \in V$*.*

Recall that two vectors $\mathbf{v}, \mathbf{w} \in V$ are called *orthogonal* iff $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Also, recall that any finite set of mutually orthogonal non-zero vectors is linearly independent.

Exercise 1.15. *For an inner product vector space, show that a finite number of orthogonal non-zero vectors are linearly independent.*

For a sub vector space $W \subset V$ the *orthogonal complement* $W^\perp$ is defined as

$$W^\perp := \{ \mathbf{v} \in V \, | \, \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{w} \in W \} \, . \tag{1.30}$$

In other words, the orthogonal complement $W^\perp$ consists of all vectors which are orthogonal to the entire space W.

Exercise 1.16. *Show, for a sub vector space* $W \subset V$*, that* $W \cap W^\perp = \{\mathbf{0}\}$*. (This means that the sum of* W *and* $W^\perp$ *is direct.) For a finite-dimensional* V*, show that* $W \oplus W^\perp = V$*.*

Further, a (finite or infinite) collection $\boldsymbol{\epsilon}_i$ of vectors, where $i = 1, 2, \ldots$, is called *an ortho-normal system* iff $\langle \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j \rangle = \delta_{ij}$. We know that finite-dimensional vector spaces have a basis and by applying to such a basis the *Gram-Schmidt* procedure one obtains an ortho-normal basis. Hence, every finite-dimensional inner product vector space has an ortho-normal basis. The scalar product makes it easier to work out the coordinates of a vector $\mathbf{v} \in V$ relative to an ortho-normal basis by using the formula

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \, \boldsymbol{\epsilon}_i \quad \Longleftrightarrow \quad \alpha_i = \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \; . \tag{1.31}$$

Also, recall that, in terms of the coordinates relative to an ortho-normal basis, the scalar product and its associated norm take a very simple form. For two vectors $\mathbf{v} = \sum_i \alpha_i \boldsymbol{\epsilon}_i$ and $\mathbf{w} = \sum_j \beta_j \boldsymbol{\epsilon}_j$ we have

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} \alpha_i^* \beta_i \; , \qquad \| \, \mathbf{v} \, \|^2 = \sum_{i=1}^{n} |\alpha_i|^2 \; , \tag{1.32}$$

as can be easily verified using the orthogonality relations $\langle \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j \rangle = \delta_{ij}$. For infinite-dimensional inner product spaces the story is more involved and will be tackled in Section 2.

It is useful to re-consider the relationship of a vector space V and its dual vector space V^* in the presence of an inner product on V. The main observation is that the inner product induces a map $\imath : V \to V^*$ defined by

$$\imath(\mathbf{v})(\mathbf{w}) := \langle \mathbf{v}, \mathbf{w} \rangle \; . \tag{1.33}$$

For a vector space over $\mathbb{R}$ this map is linear, for a vector space over $\mathbb{C}$ it is half-linear (meaning, as for the first argument of hermitian scalar products, that vector sums pull through while scalars pull out with a complex conjugation). In either case, this map is injective. For finite-dimensional V it is bijective and provides an identification of the vector space with its dual.

Exercise 1.17. *Show that the map $\imath : V \to V^*$ defined in Eq. (1.33) is injective and that it is bijective for finite-dimensional V.*

The properties of the map $\imath$ in the infinite-dimensional case will be further explored later.

Application 1.1. Dirac notation

In physics, more specifically in the context of quantum mechanics, the existence of the map $\imath$ is exploited for a convenient convention, referred to as *Dirac notation*. In Dirac notation, vectors $\mathbf{w} \in V$ and dual vectors $\imath(\mathbf{v}) \in V^*$ are denoted as follows:

$$\mathbf{w} \to |\mathbf{w}\rangle \; , \qquad \imath(\mathbf{v}) \to \langle \mathbf{v}| \; . \tag{1.34}$$

In other words, vectors in V are denoted by "ket"-vectors $|\mathbf{w}\rangle$, dual vectors in V^*, obtained via the map $\imath$, by "bra"-vectors $\langle \mathbf{v}|$ while the action of one on the other (which equals the scalar product in Eq. (1.33)) is simple obtained by combining the two to a "bra-(c)ket", resulting in

$$\imath(\mathbf{v})(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v} | \mathbf{w} \rangle \; . \tag{1.35}$$

Note, there is nothing particularly profound about this notation - for the most part it simply amounts to replacing the comma separating the two arguments of an inner product with a vertical bar.

We can ask about interesting new properties of linear maps in the presence of an inner product. First, recall that scalar products of the form

$$\langle \mathbf{v}, T(\mathbf{w}) \rangle \tag{1.36}$$

for a linear map $T : V \to V$ are also called *matrix elements* of T. Two maps $T : V \to V$ and $S : V \to V$ are equal iff all their matrix elements are equal, that is, iff $\langle \mathbf{v}, T(\mathbf{w})\rangle = \langle \mathbf{v}, S(\mathbf{w})\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$.

Exercise 1.18. *Show that two linear maps are equal iff all their matrix elements are equal.*

In the finite-dimensional case, the matrix A which describes a linear map $T : V \to V$ relative to an ortho-normal basis $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ is simply obtained by the matrix elements

$$A_{ij} = \langle \boldsymbol{\epsilon}_i, T(\boldsymbol{\epsilon}_j)\rangle \; . \tag{1.37}$$

Next, we recall the definition of the *adjoint linear map*.

Definition 1.10. *For a linear map $T : V \to V$ on a vector space V with scalar product, an adjoint linear map, $T^\dagger : V \to V$ is a map satisfying*

$$\langle \mathbf{v}, T\mathbf{w}\rangle = \langle T^\dagger \mathbf{v}, \mathbf{w}\rangle \tag{1.38}$$

for all $\mathbf{v}, \mathbf{w} \in V$.

If the adjoint exists, it is unique and has the following properties

$$(T^\dagger)^\dagger = T \; , \quad (\alpha T + \beta S)^\dagger = \alpha^* T^\dagger + \beta^* S^\dagger \; , \quad (S \circ T)^\dagger = T^\dagger \circ S^\dagger \; , \quad {T^{-1}}^\dagger = {T^\dagger}^{-1} \; , \tag{1.39}$$

provided the maps in those equations exist.

Exercise 1.19. *Show that the adjoint map is unique and that it has the properties in Eq. (1.39).*

For finite-dimensional inner product vector spaces we can describe both T and its adjoint $T^\dagger$ by matrices relative to an ortho-normal basis $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$. They are given by the matrix elements

$$A_{ij} = \langle \boldsymbol{\epsilon}_i, T(\boldsymbol{\epsilon}_j)\rangle \; , \qquad (A^\dagger)_{ij} = \langle \boldsymbol{\epsilon}_i, T^\dagger(\boldsymbol{\epsilon}_j)\rangle \; . \tag{1.40}$$

where $A^\dagger := A^{T^*}$ is the hermitian conjugate of the matrix A. Hence, at the level of matrices, the adjoint simply corresponds to the hermitian conjugate matrix (or the transpose matrix in the real case). This observation can also be used to show the existence of the adjoint for finite-dimensional inner product spaces. Existence of the adjoint in the infinite-dimensional case is not so straightforward and will be considered later.

Exercise 1.20. *Show that the matrix which consists of the matrix elements of $T^\dagger$ in Eq. (1.40) is indeed the hermitian conjugate of the matrix given by the matrix elements of T.*

Particularly important linear operators are those which can be moved from one argument of a scalar product into the other without changing the value of the scalar product and they are called *hermitian* or *self-adjoint* operators.

Definition 1.11. *A linear operator $T : V \to V$ on a vector space V with scalar product is called self-adjoint (or hermitian) iff $\langle \mathbf{v}, T(\mathbf{w})\rangle = \langle T(\mathbf{v}), \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$.*

Hence, a self-adjoint operator $T : V \to V$ is one for which the adjoint exists and satisfies $T^\dagger = T$.

Recall that the *commutator* of two linear operators S, T is defined as

$$[S, T] := S \circ T - T \circ S \; , \tag{1.41}$$

We also say that two operators S and T *commute* iff $[S, T] = 0$.

We can ask under what condition the composition $S \circ T$ of two hermitian operators is again hermitian. Using the above commutator notation, we have

$$(S \circ T)^\dagger = S \circ T \quad \Leftrightarrow \quad T^\dagger \circ S^\dagger = S \circ T \quad \Leftrightarrow \quad T \circ S = S \circ T \quad \Leftrightarrow \quad [S, T] = 0 \tag{1.42}$$

where $S = S^\dagger$ and $T = T^\dagger$ has been used for the second equivalence. In conclusion, the composition of two hermitian operators is hermitian if and only if the operators commute. For a complex inner product vector space, it is also worth noting that, from Eq. (1.39), an *anti-hermitian* operator, that is an operator T satisfying $T^\dagger = -T$, can be turned into a hermitian one (and vice versa) by multiplying with $\pm i$, so

$$T^\dagger = -T \quad \Longleftrightarrow \quad (\pm iT)^\dagger = \pm iT \; . \tag{1.43}$$

Also note that every linear operator $T : V \to V$ with an adjoint $T^\dagger$ can be written as a (unique) sum of a hermitian and an anti-hermitian operator. Indeed, defining $T_\pm = \frac{1}{2}(T \pm T^\dagger)$ we have $T = T_+ + T_-$ while T_+ is hermitian and T_- is anti-hermitian.

Application 1.2. More on Dirac notation

In the context of Dirac notation, the matrix elements of an operator T are denoted by

$$\langle \mathbf{v}|T|\mathbf{w}\rangle := \langle \mathbf{v}, T(\mathbf{w})\rangle \; . \tag{1.44}$$

In this way, the matrix element of the operator is obtained by including it between a bra and a ket vector. This symmetric notation is particularly useful for hermitian operators since they can be thought of as acting on either one of the scalar product's arguments. For non-hermitian operators or for the purpose of proving that an operator is hermitian the Dirac notation is less helpful and it is sometimes better to use the mathematical notation, as on the RHS of Eq. (1.44). Relative to an ortho-normal basis $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ of a finite-dimensional inner product space V a self-adjoint linear operator $T : V \to V$ is described by the matrix with entries (in Dirac notation)

$$T_{ij} = \langle \boldsymbol{\epsilon}_i|T|\boldsymbol{\epsilon}_j\rangle \; . \tag{1.45}$$

In terms of these matrix elements, T can also be written as

$$T = \sum_{k,l=1}^{n} T_{kl}|\boldsymbol{\epsilon}_k\rangle\langle\boldsymbol{\epsilon}_l| \; . \tag{1.46}$$

This can be easily verified by taking the matrix elements with $\langle\boldsymbol{\epsilon}_i|$ and $|\boldsymbol{\epsilon}_j\rangle$ of this equation and by using $\langle\boldsymbol{\epsilon}_i|\boldsymbol{\epsilon}_k\rangle = \delta_{ik}$. (Formally, Eq. (1.46) exploits the identification $\mathrm{Hom}(V, V) \cong V \otimes V^*$.) In particular the identity operator id with matrix elements δ_{ij} can be written as

$$\mathrm{id} = \sum_{i=1}^{n} |\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i| \; . \tag{1.47}$$

Exercise 1.21. *By acting on an arbitrary vector, verify explicitly that the RHS of Eq. (1.47) is indeed the identity operator.*

Dirac notation can be quite intuitive as can be demonstrated by re-writing some of our earlier equations. For example, writing the relation (1.31) for the coordinates relative to an orth-normal basis in

Dirac notation leads to

$$|\mathbf{v}\rangle = \sum_{i=1}^{n} |\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{v}\rangle \ . \tag{1.48}$$

Evidently, this can now be derived by inserting the identity operator in the form (1.47). Similarly, the expressions (1.32) for the scalar product and the norm in Dirac notation

$$\langle\mathbf{v}|\mathbf{w}\rangle = \sum_{i=1}^{n}\langle\mathbf{v}|\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{w}\rangle \ , \qquad \| \ |\mathbf{v}\rangle \ \|^2 = \langle\mathbf{v}|\mathbf{v}\rangle = \sum_{i=1}^{n}\langle\mathbf{v}|\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{v}\rangle \tag{1.49}$$

are easily seen to follow by inserting the identity operator (1.47).

Another important class of specific linear maps on an inner product vector space are *unitary* maps which are precisely those maps which leave the value of the inner product unchanged in the sense of the following

Definition 1.12. *Let V be an inner producr vector space. A linear map $U : V \to V$ is called unitary iff*

$$\langle U(\mathbf{v}), U(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle \tag{1.50}$$

for all $\mathbf{v}, \mathbf{w} \in V$.

Unitary maps have the following important properties.

Proposition 1.1. *(Properties of unitary maps) A unitary map U with adjoint $U^\dagger$ has the following properties.*
(i) Unitary maps U can also be characterized by $U^\dagger \circ U = U \circ U^\dagger = \mathrm{id}_V$.
(ii) Unitary maps U are invertible and $U^{-1} = U^\dagger$.
(iii) The composition of unitary maps is a unitary map.
(iv) The inverse, $U^\dagger$, of a unitary map U is unitary.

Exercise 1.22. *Show the properties of unitary maps in Lemma 1.1.*

For finite-dimensional vector spaces we know that, relative to an ortho-normal basis $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$, a unitary map $\hat{U}$ is described by a unitary matrix (orthogonal matrix in the real case). Indeed, introducing the matrix U with matrix elements (in Dirac notation)

$$U_{ij} = \langle\boldsymbol{\epsilon}_i|\hat{U}|\boldsymbol{\epsilon}_j\rangle \ , \tag{1.51}$$

this statement is verified by the following short calculation.

$$\sum_j (U^\dagger)_{ij}U_{jk} = \sum_j \langle\boldsymbol{\epsilon}_i|\hat{U}^\dagger|\underbrace{\boldsymbol{\epsilon}_j\rangle\langle\boldsymbol{\epsilon}_j}_{=\mathrm{id}}|\hat{U}|\boldsymbol{\epsilon}_k\rangle = \langle\boldsymbol{\epsilon}_i|\underbrace{\hat{U}^\dagger\hat{U}}_{=\mathrm{id}}|\boldsymbol{\epsilon}_k\rangle = \langle\boldsymbol{\epsilon}_i|\boldsymbol{\epsilon}_k\rangle = \delta_{ik} \ . \tag{1.52}$$

Still in the finite-dimensional case, consider two choices of ortho-normal basis $(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n)$ and $(\boldsymbol{\epsilon}'_1, \ldots, \boldsymbol{\epsilon}'_n)$ and the matrices $T_{ij} = \langle\boldsymbol{\epsilon}_i|\hat{T}|\boldsymbol{\epsilon}_j\rangle$ and $T'_{ij} = \langle\boldsymbol{\epsilon}'_i|\hat{T}|\boldsymbol{\epsilon}'_j\rangle$ representing a linear operator $\hat{T}$ with respect to either. We have already written down the general relation between those two matrices in Eq. (1.15) but how does this look for a change from one ortho-normal basis to another? Inserting identity operators (1.47) we find

$$T'_{ij} = \langle\boldsymbol{\epsilon}'_i|\hat{T}|\boldsymbol{\epsilon}'_j\rangle = \sum_{k,l=1}^{m}\langle\boldsymbol{\epsilon}'_i|\boldsymbol{\epsilon}_k\rangle\langle\boldsymbol{\epsilon}_k|\hat{T}|\boldsymbol{\epsilon}_l\rangle\langle\boldsymbol{\epsilon}_l|\boldsymbol{\epsilon}'_j\rangle = Q_{ik}T_{kl}Q^*_{jl} = (QTQ^\dagger)_{ij} \ , \qquad Q_{ij} := \langle\boldsymbol{\epsilon}'_i|\boldsymbol{\epsilon}_j\rangle \tag{1.53}$$

so that $T' = QTQ^\dagger$. This result is, in fact, consistent with Eq. (1.15) since the matrix Q is unitary, so $Q^\dagger = Q^{-1}$. This can be verified immediately:

$$(Q^\dagger Q)_{ij} = \sum_{k=1}^{n} Q^*_{ki} Q_{kj} = \sum_{k=1}^{n} \langle \boldsymbol{\epsilon}'_k | \boldsymbol{\epsilon}_i \rangle^* \langle \boldsymbol{\epsilon}'_k | \boldsymbol{\epsilon}_j \rangle = \sum_{k=1}^{n} \langle \boldsymbol{\epsilon}_i | \boldsymbol{\epsilon}'_k \rangle \langle \boldsymbol{\epsilon}'_k | \boldsymbol{\epsilon}_j \rangle = \langle \boldsymbol{\epsilon}_i | \boldsymbol{\epsilon}_j \rangle = \delta_{ij} \ . \tag{1.54}$$

Using this formalism, we can also verify that Q relates coordinate vectors relative to the two choices of basis, as stated in Eq. (1.16). From Eq. (1.48), the two coordinate vectors for a given vector $|\mathbf{v}\rangle$ are given by $\alpha_i = \langle \boldsymbol{\epsilon}_i | \mathbf{v} \rangle$ and $\alpha'_i = \langle \boldsymbol{\epsilon}'_i | \mathbf{v} \rangle$. It follows

$$\alpha'_i = \langle \boldsymbol{\epsilon}'_i | \mathbf{v} \rangle = \sum_{j=1}^{n} \langle \boldsymbol{\epsilon}'_i | \boldsymbol{\epsilon}_j \rangle \langle \boldsymbol{\epsilon}_j | \mathbf{v} \rangle = \sum_{j=1}^{n} Q_{ij} \alpha_j \ . \tag{1.55}$$

1.1.8 Examples of inner product vector spaces

The standard finite-dimensional examples are of course $\mathbb{R}^n$ and $\mathbb{C}^n$ with scalar product defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle := \sum_{i=1}^{n} v_i^* w_i \ , \tag{1.56}$$

for vectors $\mathbf{v} = (v_1, \ldots, v_n)^T$ and $\mathbf{w} = (w_1, \ldots, w_n)^T$. (We have followed the convention, mentioned above, of writing the equations for the complex case. For the real case, simply drop the complex conjugation.) The norm associated to this scalar product is of course the one given in Eq. (1.22). Linear maps are described by $n \times n$ matrices and the adjoint of a matrix A, relative to the inner product (1.56), is given by the hermitian conjugate $A^\dagger$. For the complex case, unitary linear maps are given by unitary matrices, that is matrices U satisfying

$$U^\dagger U = \mathbb{1}_n \ . \tag{1.57}$$

For the real case, unitary linear maps, relative to the inner product (1.56), are given by orthogonal matrices, that is matrices A satisfying

$$A^T A = \mathbb{1}_n \ . \tag{1.58}$$

Both are important classes of matrices which we will return to in our discussion of symmetries in Section 9.

For an infinite-dimensional example, we begin with the space $\mathcal{C}[a, b]$ of continuous (complex-valued) functions on the interval $[a, b]$, equipped with the scalar product

$$\langle f, g \rangle := \int_a^b dx \, f(x)^* g(x) \ , \tag{1.59}$$

for $f, g \in \mathcal{C}[a, b]$.

Exercise 1.23. *Verify that Eq. (1.59) defines a scalar product on $\mathcal{C}[a, b]$. (Hint: Check the conditions in Def. 1.9).*

The norm associated to this scalar product is given by the first equation (1.25). Consider the linear operator M_p, defined in Eq. (1.20), which acts by multiplication with the function p. What is the adjoint of M_p? The short calculation

$$\langle f, M_p(g) \rangle = \int_a^b dx \, f(x)^* (p(x) g(x)) = \int_a^b (p(x)^* f(x))^* g(x) = \langle M_{p^*}(f), g \rangle \tag{1.60}$$

shows that

$$M_p^\dagger = M_{p^*} \ , \tag{1.61}$$

so the adjoint operator corresponds to multiplication with the complex conjugate function p^*. If p is real-valued so that $p = p^*$ then M_p is a hermitian operator. From the definition of the multiplication operator it is clear that

$$M_p \circ M_q = M_{pq} \ , \qquad M_1 = \mathrm{id} \tag{1.62}$$

for two functions p and q. Eqs. (1.61) and (1.62) can be used to construct unitary multiplication operators. For a real-valued function u we have

$$M_{e^{iu}}^\dagger \circ M_{e^{iu}} = M_{e^{-iu}} \circ M_{e^{iu}} = M_1 = \mathrm{id} \ , \tag{1.63}$$

so that multiplication with a complex phase $e^{iu(x)}$ (where u is a real-valued function) is a unitary operator. This can also be verified directly from the scalar product:

$$\langle M_{e^{iu}}(f), M_{e^{iu}}(g)\rangle = \int_a^b dx \, \left(e^{iu(x)} f(x)\right)^* \left(e^{iu(x)} g(x)\right) = \int_a^b f(x)^* g(x) = \langle f, g\rangle \ . \tag{1.64}$$

For another example of a unitary map, let us restrict to the space $\mathcal{C}_c(\mathbb{R})$ of complex-valued functions on the real line with compact support, still with the scalar product (1.59), but setting $a = -\infty$ and $b = \infty$. (The compact support property is to avoid issues with the finiteness of the integral - we will deal with this in more generality later.) On this space define the "translation operator" $T_a : \mathcal{C}_c(\mathbb{R}) \to \mathcal{C}_c(\mathbb{R})$ by

$$T_a(f)(x) := f(x - a) \ , \tag{1.65}$$

for any fixed $a \in \mathbb{R}$. Evidently, this operator "shifts" the graph of the function by an amount of a along the x-axis. Let us work out the effect of this operator on the scalar product. To find the adjoint of T_a we calculate

$$\langle f, T_a(g)\rangle = \int_{-\infty}^{\infty} dx \, f(x)^* g(x - a) \overbrace{=}^{y=x-a} \int_{-\infty}^{\infty} dy \, f(y + a)^* g(y) = \langle T_{-a}(f), g\rangle \ , \tag{1.66}$$

so that $T_a^\dagger = T_{-a}$, that is, the adjoint is given by the shift in the opposite direction. To check unitarity we work out

$$\langle T_a(f), T_a(g)\rangle = \int_{-\infty}^{\infty} dx \, f(x - a)^* g(x - a) \overbrace{=}^{y=x-a} \int_{-\infty}^{\infty} dy \, f(y)^* g(y) = \langle f, g\rangle \ . \tag{1.67}$$

and conclude that T_a is indeed unitary. Alternatively, we can check the unitarity condition $T_a^\dagger \circ T_a = T_{-a} \circ T_a = \mathrm{id}$ which works out as expected since combining shifts by a and $-a$ amounts to the identity operation.

To consider differential operators we restrict further to the inner product space $\mathcal{C}_c^\infty(\mathbb{R})$ of complex-valued, infinitely times differentiable functions with compact support, still with scalar product defined by Eq. (1.59), setting $a = -\infty$ and $b = \infty$. What is the adjoint of the differential operator $D = d/dx$? The short calculation

$$\langle f, D(g)\rangle = \int_{-\infty}^{\infty} dx \, f(x)^* g'(x) = \underbrace{[f(x)^* g(x)]_{-\infty}^{\infty}}_{=0} - \int_{-\infty}^{\infty} dx \, f'(x)^* g(x) = \langle -D(f), g\rangle \tag{1.68}$$

(where the boundary term vanishes since the functions have compact support) shows that

$$\left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx} \,, \tag{1.69}$$

so d/dx is anti-hermitian. As discussed earlier, for a complex inner product space, we can turn this into a hermitian operator by multiplying with $\pm i$, so that

$$\left(\pm i\frac{d}{dx}\right)^\dagger = \pm i\frac{d}{dx} \,. \tag{1.70}$$

Another lesson from the above computation is that, for scalar products defined by integrals, the property of being hermitian can depend on boundary conditions satisfied by the functions in the relevant function vector space. In the case of Eq. (1.68) we were able to reach a conclusion because the boundary term could be discarded due to the compact support property of the functions.

What about the composite operator $M_x \circ i\, d/dx$? We know that the composition of two hermitian operators is hermitian iff the two operators commute so let us work out the commutator (writing, for simplicity, M_x as x)

$$\left[i\frac{d}{dx}, x\right] = i\frac{d}{dx} \circ x - x \circ i\frac{d}{dx} = i + ix\frac{d}{dx} - ix\frac{d}{dx} = i \,. \tag{1.71}$$

(If the above computation looks confusing remember we are dealing with operators, so think of the entire equation above as acting on a function f. The second step in the calculation then amounts to using the product rule for differentiation.) Since the above commutator is non-vanishing we conclude that $M_x \circ i\, d/dx$ is not hermitian.

So much for a few introductory examples of how to carry out calculations for infinite-dimensional inner product spaces. We will now collect a few more mathematical tools required for a more systematic approach.

1.1.9 Eigenvectors and eigenvalues

Recall the definition of *eigenvalues* and *eigenvectors*.

Definition 1.13. *For a linear map $T : V \to V$ on a vector space V over F the number $\lambda \in F$ is called an eigenvalue of f if there is a non-zero vector $\mathbf{v}$ such that*

$$T(\mathbf{v}) = \lambda \mathbf{v} \,. \tag{1.72}$$

In this case, $\mathbf{v}$ is called an eigenvector of f with eigenvalue λ.

The eigenspace for $\lambda \in F$ is defined by

$$\mathrm{Eig}_T(\lambda) := \mathrm{Ker}(T - \lambda\, \mathrm{id}_V) \,, \tag{1.73}$$

so that λ is an eigenvalue iff $\mathrm{Eig}_T(\lambda) \neq \{\mathbf{0}\}$. If $\dim(\mathrm{Eig}_T(\lambda)) = 1$ the eigenvalue is called *non-degenerate* (there is only one eigenvector up to re-scaling) and *degenerate* otherwise (there are at least two linearly independent eigenvectors).

Let us recall the basis facts in the finite-dimensional case. The eigenvalues can be obtained by finding the zeros of the *characteristic polynomial*

$$\chi_T(\lambda) := \det(T - \lambda \mathrm{id}) \,. \tag{1.74}$$

For each eigenvalue λ the associated eigenspace is obtained by finding all solutions $\mathbf{v} \in V$ to the equation $(T - \lambda \mathrm{id})\mathbf{v} = \mathbf{0}$. The most important applications of eigenvalues and eigenvectors in the finite-dimensional case is to diagonalising linear maps, that is, finding a basis in which the matrix describing the linear map is diagonal. Recall that diagonalising a linear map T is possible if and only if T has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of eigenvectors. Indeed, in this case $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ and the matrix describing T relative to this basis is $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. There are certain classes of linear operators which are known to have a basis of eigenvectors and can, hence, be diagonalised. These include self-adjoint linear operators and *normal operators*, that is, operators satisfying $[T, T^\dagger] = 0$.

Some useful statements which are well-known in the finite-dimensional case continue to hold in infinite dimensions, such as the following

Theorem 1.24. *Let V be an inner product vector space. If $T : V \to V$ is self-adjoint then*

(i) All eigenvalues of T are real.
(ii) Eigenvectors for different eigenvalues are orthogonal.

Exercise 1.25. *Proof Theorem 1.24.*

Application 1.3. Eigenvalvues and eigenvectors for a differential operator

As an illustration of this theorem in the infinite-dimensional case, consider the space $\mathcal{C}^\infty_{\mathrm{p}}([-\pi, \pi])$ of infinitely many times differentiable and periodic (real) functions on the interval $[-\pi, \pi]$. (By periodic functions we mean functions f with $f(\pi) = f(-\pi)$ and $f'(\pi) = f'(-\pi)$.) On this vector space, we define the usual inner product

$$\langle f, g\rangle := \int_{-\pi}^{\pi} dx\, f(x)g(x)\ . \tag{1.75}$$

A calculation analogous to the one in Eq. (1.68) (where periodicity allows discarding the boundary term) shows that the operator d/dx is anti-hermitian and d^2/dx^2 is hermitian relative to this inner product.

Exercise 1.26. *For the vector space $\mathcal{C}^\infty_{\mathrm{p}}([-\pi, \pi])$ with inner product (1.75) show that d/dx is anti-hermitian and d^2/dx^2 is hermitian.*

From Theorem 1.24 we, therefore, conclude that eigenvectors of d^2/dx^2 for different eigenvalues must be orthogonal. To check this explicitly, we write down the eigenvalue equation

$$\frac{d^2 f}{dx^2} = \lambda f\ . \tag{1.76}$$

For $\lambda > 0$ the solutions to this equation are (real) exponential and, hence, cannot be elements of our vector space of periodic functions. For $\lambda < 0$ the eigenfunctions are given by $f_k(x) = \sin(kx)$ and $g_k(x) = \cos(kx)$, where $\lambda = -k^2$. At this point k is still arbitrary real but for f_k and g_k to be periodic with period 2π we need $k \in \mathbb{Z}$. Of course for f_k we can restrict to $k \in \mathbb{Z}^{>0}$ and for g_k to $k \in \mathbb{Z}^{\geq 0}$. In summary, we have the eigenvectors and eigenvalues

$$\begin{array}{lcll} f_k(x) & = & \sin(kx) & k = 1, 2, \ldots \qquad \lambda = -k^2 \\ g_k(x) & = & \cos(kx) & k = 0, 1, \ldots \qquad \lambda = -k^2 \end{array} \tag{1.77}$$

In particular, this implies that the eigenvalues $\lambda = -k^2$ for $k = 1, 2, \ldots$ are degenerate. By direct calculation, we can check that for $k \neq l$, we have $\langle f_k, f_l\rangle = \langle g_k, g_l\rangle = \langle f_k, g_l\rangle = 0$, as stated by

Theorem 1.24. In fact, we also have $\langle f_k, g_k \rangle = 0$ which is not predicted by the theorem but follows from direct calculation.

Exercise 1.27. *Show that the functions* (1.77) *satisfy* $\langle f_k, f_l \rangle = \langle g_k, g_l \rangle = \langle f_k, g_l \rangle = 0$ *for* $k \neq l$ *as well as* $\langle f_k, g_k \rangle = 0$*, relative to the scalar product* (1.75)*.*

The above example leads to the Fourier series which we will discuss in Section 3.

There are also constraints on the possible eigenvalues of unitary operators.

Theorem 1.28. *Let* $U : V \to V$ *be a unitary linear operator on an inner product space* V*. If* λ *is an eigenvalue of* U *then* $|\lambda| = 1$*.*

Exercise 1.29. *Proof Theorem 1.28.*

1.2 Topology and convergence*

We will now introduce some basics of topology and convergence. In this sub-section, we will be working in a normed vector space V over a field F with norm $\| \cdot \|$ and, hence, this will be more general than a discussion in the context of real numbers, where these ideas are typically being introduced for the first time. However, keep in mind that the real (and complex) numbers do form normed vector spaces and are, hence, covered, as special cases, by the discussion below.

Generalities We begin by defining the *ball* $B_r(\mathbf{v})$ around any $\mathbf{v} \in V$ with radius $r > 0$ by

$$B_r(\mathbf{v}) := \{\mathbf{w} \in V \,|\, \| \mathbf{v} - \mathbf{w} \| < r\} \,. \tag{1.78}$$

Note that this is the "full" ball including all of the "interior" but, due to the strictly less condition in the definition, without the bounding sphere.

We would like to consider infinite *sequences* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_i, \ldots)$ of vectors $\mathbf{v}_i \in V$ which we also denote by $(\mathbf{v}_i)_{i=1}^{\infty}$ or simply by $(\mathbf{v}_i)$, when the range of the index i is clear from the context.

1.2.1 Convergence and Cauchy sequence

What does it mean for such a sequence to converge to a vector $\mathbf{v} \in V$?

Definition 1.14. *(Convergence of a sequence) A sequence* $(\mathbf{v}_i)_{i=1}^{\infty}$ *in a normed vector space* V *converges to a vector* $\mathbf{v} \in V$ *if, for every* $\epsilon > 0$*, there exists a positive integer* k *such that* $\mathbf{v}_i \in B_\epsilon(\mathbf{v})$ *for all* $i > k$*. In this case, we write* $\lim_{i\to\infty} \mathbf{v}_i = \mathbf{v}$*.*

Note that, while this definition might sound somewhat convoluted, it actually captures the intuitive idea of convergence. It says that the sequence converges to $\mathbf{v}$ if, for every small deviation $\epsilon > 0$, there is always a "tail", sufficiently far out, which is entirely contained within the ball of radius ϵ around $\mathbf{v}$. (See Fig. 1.)

Application 1.4. Convergence of a sequence

As a very simple example for how to use the above definition of convergence consider the sequence (a_j) in $\mathbb{R}$ (with the standard norm given by $||x|| = |x|$) defined by $a_j = 1/j$. It is intuitively clear that this sequence converges to 0 but how can this be shown? We start with any $\epsilon > 0$. For any such ϵ we can always choose a sufficiently large $k \in \mathbb{N}$ such that $k > 1/\epsilon$. Then, for any $j > k$ we have $|a_j - 0| = 1/j < 1/k < \epsilon$. This shows that $\lim_{j\to\infty} a_j = 0$.

There is a related, but somewhat weaker notion of convergence which avoids talking about the vector the sequence converges to. Sequences which converge in this weaker sense are called *Cauchy sequences* and are defined as follows.

Definition 1.15. *(Cauchy sequence) A sequence $(\mathbf{v}_i)_{i=1}^{\infty}$ in a normed vector space V is called a Cauchy sequence if, for every $\epsilon > 0$, there exist a positive integer k such that $\| \mathbf{v}_i - \mathbf{v}_j \| < \epsilon$ for all $i, j > k$. (See Fig. 1.)*

In other words, a sequence is a Cauchy sequence if for every small $\epsilon > 0$ there is a "tail", sufficiently far out, such that the norm between each two vectors in the tail is less than ϵ. The notions of convergent

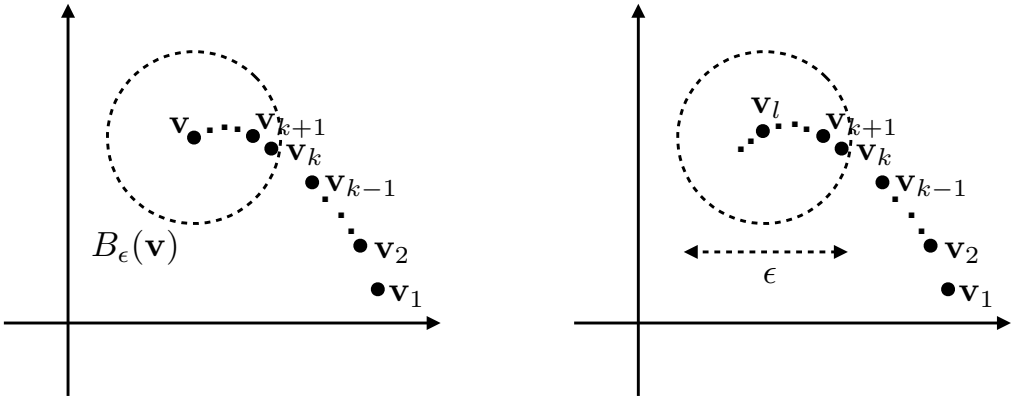

Figure 1: Convergence of a sequence $(\mathbf{v}_k)$ to a limit $\mathbf{v}$ (left) and Cauchy convergence (right).

sequence and Cauchy sequence lead to analogous notions for a series $\sum_{i=1}^{\infty} \mathbf{v}_i$, which can be defined by focusing on its partial sums.

Definition 1.16. *A series $\sum_{i=1}^{\infty} \mathbf{v}_i$ is called convergent to a vector $\mathbf{v}$ (is called a Cauchy series) iff the associated sequence of partial sums $(\mathbf{s}_k)_{k=1}^{\infty}$, where $\mathbf{s}_k = \sum_{i=1}^{k} \mathbf{v}_i$, converges to the vector $\mathbf{v}$ (is a Cauchy sequence).*

Exercise 1.30. *Show that every convergent sequence in a normed vector space is also a Cauchy sequence. (Hint: Use the triangle inequality.)*

Application 1.5. A divergent series

We would like to show that the series $\sum_{j=1}^{\infty} \frac{1}{j}$ in $\mathbb{R}$ (with the standard absolute values norm) diverges and to do this it is enough to show that it is not Cauchy convergent. We look at the difference $|s_i - s_j|$ of two partial sequences and, assuming $i < j$, this is given by $|s_i - s_j| = \sum_{l=i+1}^{j} \frac{1}{l}$. Choose $i = 2^n - 1$ and $j = 2^{n+1} - 1$ for some integer n. In this case, we have

$$|s_i - s_j| = \frac{1}{2^n} + \frac{1}{2^n + 1} + \cdots + \frac{1}{2^{n+1} - 1} > 2^n \frac{1}{2^{n+1}} = \frac{1}{2} \,,$$

where the inequality follows from the fact we have 2^n terms each of which is larger than $1/2^{n+1}$. Choose an $\epsilon < 1/2$ and for any k an integer n with $2^n > k + 1$. Then, setting $i = 2^n - 1$ and

$j = 2^{n+1} - 1$ we have $i, j > k$ and $|s_i - s_j| > 1/2 > \epsilon$ so the condition for Cauchy convergence cannot be satisfied.

For a series there is also a stronger version of convergence, called *absolute convergence*.

Definition 1.17. *A series $\sum_{i=1}^{\infty} \mathbf{v}_i$ is called absolutely convergent if $\sum_{i=1}^{\infty} \| \mathbf{v}_i \|$ converges (as a series over the real numbers).*

Exercise 1.31. *Show that an absolutely convergent series is a Cauchy series.*

While every convergent sequence is a Cauchy sequence the opposite is not true. The classical example is provided by the rational numbers $\mathbb{Q}$ (viewed as a normed space, with the absolute modulus as the norm) and a sequence (q_i) of rational numbers which converges to a real number $x \in \mathbb{R} \setminus \mathbb{Q}$. This is clearly a Cauchy sequence but it does not converge since the prospective limit is not contained in $\mathbb{Q}$ (although it does converge seen as a sequence in $\mathbb{R}$). This example is typical and points to an intuitive understanding of what non-convergent Cauchy sequences mean. They indicate a deficiency of the underlying normed vector space which has "holes" and is, as far as convergence properties are concerned, incomplete. This idea will play an important role for the definition of Banach and Hilbert spaces later on and motivates the following definition of *completeness*.

Definition 1.18. *A normed vector space is called complete iff every Cauchy series converges.*

1.2.2 Open and closed sets

Our next step is to introduce some basic topology in normed vector spaces, most importantly the ideas of *open and closed* sets.

Definition 1.19. *(Open and closed sets) Let V be a normed vector space.*
A subset $U \subset V$ is called open if, for every $\mathbf{u} \in U$, there exists an $\epsilon > 0$ such that $B_\epsilon(\mathbf{u}) \subset U$.
A subset $C \subset V$ is called closed if $V \setminus C$ is open.
For an arbitrary subset $S \subset V$, the closure of S, denoted $\bar{S}$, is the smallest closed set which contains S.
A $\mathbf{v} \in V$ is called a limit point of a subset $S \subset V$ if there is a sequence $(\mathbf{v}_i)$, entirely contained in S, which converges to $\mathbf{v}$.

An open set is simply a set which contains a ball around each of its points while a closed set is the complement of an open set.

Exercise 1.32. *Show that a ball $B_r(\mathbf{v})$ in a normed vector space is open for every $r > 0$ and every $\mathbf{v} \in V$.*

The ideas of convergence and closed sets relate in an interesting way as stated in the following lemma.

Proposition 1.2. *For a normed vector space V and a subset $S \subset V$ we have the following statements.*
(a) S is closed $\iff$ All limit points of S are contained in S.
(b) The closure, $\bar{S}$, consists of S and all its limit points.

Proof. (a) "$\Rightarrow$": Assume that S is closed so that its complement $U := V \setminus S$ is open. Consider a limit point $\mathbf{v}$ of S, with a sequence $(\mathbf{v}_i)$ contained in S and converging to $\mathbf{v}$. We need to show that $\mathbf{v} \in S$. Assume that $\mathbf{v} \in U$. Since U is open there is a ball $B_\epsilon(\mathbf{v}) \subset U$ entirely contained in U and $\mathbf{v}_i \notin B_\epsilon(\mathbf{v})$ for all i. But this means the sequence $(\mathbf{v}_i)$ does not converge to $\mathbf{v}$ which is a contradiction. Hence, our assumption that $\mathbf{v} \in U$ was incorrect and $\mathbf{v} \in S$.
(a) "$\Leftarrow$": Assume all limit points of S are contained in S. We need to show that $U := V \setminus S$ is open.

Assume that U is not open, so that there is a $\mathbf{u} \in U$ such that every ball $B_\epsilon(U)$ around $\mathbf{u}$ contains a $\mathbf{v}$ with $\mathbf{v} \notin U$. For every positive integer k, choose such a $\mathbf{v}_k \in B_{1/k}(\mathbf{u})$ with $\mathbf{v}_k \notin U$. It is clear that the sequence ($\mathbf{v}_k$ is entirely contained in S (since its not in the complement U) and converges to $\mathbf{u} \notin S$. Hence, $\mathbf{u}$ is a limit point of S but it is not contained in S. This is a contradiction, so our assumption is incorrect and U must be open.
(b) Define the set $\hat{S} = S \cup \{\text{all limit points of } S\}$. Using the result (a) it is straightforward to show that $\hat{S}$ is closed. Hence, $\hat{S}$ is a closed set containing S which implies, $\bar{S}$ being defined as the smallest such set, that $\bar{S} \subset \hat{S}$. Conversely, since $\bar{S}$ is closed it must contain by (a) all its limit point, including the limit points of S, so that $\hat{S} \subset \bar{S}$. □

Another important class of subsets are the compact ones:

Definition 1.20. *A set $S \subset V$ is called compact iff it is closed and bounded, that is, if there is an $R > 0$ such that $\| \mathbf{v} \| < R$ for all $\mathbf{v} \in S$.*

In the context of Hilbert spaces the ideas of a *dense subset* and *separability* will become important.

Definition 1.21. *A subset $S \subset V$ of a normed vector space V is called dense if $\bar{S} = V$. A normed vector space is called separable iff it has a countable, dense subset, that is, a dense subset of the form $(\mathbf{v}_i)_{i=1}^{\infty}$.*

The relevance of dense subsets can be seen from the following exercise.

Exercise 1.33. *For a normed vector space V and a subset $S \subset V$, proof the equivalence of the following statements.*
(i) S is dense in V
(ii) Every $\mathbf{v} \in V$ is a limit point of S.

Hence, every element of the vector space can be "approximated" from within a dense subset.

Exercise 1.34. *Show that every finite-dimensional normed vector space over $\mathbb{R}$ (or over $\mathbb{C}$) is separable. (Hint: Consider a basis $(\mathbf{v}_i)_{i=1}^n$ and linear combinations $\sum_{i=1}^n \alpha_i \mathbf{v}_i$, where $\alpha_i \in \mathbb{Q}$ (or $\alpha_i \in \mathbb{Q} + i\mathbb{Q}$).)*

An important theoretical statement that will help with some of our later proofs is

Theorem 1.35. *(Stone-Weierstrass) The set of polynomials $\mathcal{P}_{\mathbb{R}}([a,b])$ is dense in $\mathcal{C}_{\mathbb{R}}([a,b])$.*

Proof. For the proof, see for example Ref. [5]. □

Finally, we should briefly discuss other notions of convergence which can be defined for function vector spaces and are based on point-wise convergence in $\mathbb{R}$ or $\mathbb{C}$ (with the modulus norm), rather than on the norm $\| \cdot \|$ as in Def. 1.14.

Definition 1.22. *Let $(f_i)_{i=1}^{\infty}$ be a sequence of (real or complex valued) functions on the open set $U \subset \mathbb{R}^n$. We say that the sequence converges point-wise to a function f on U iff the sequence $(f_i(x))_{i=1}^{\infty}$ converges to $f(x)$ for all $x \in U$ (with respect to the real or complex modulus norm).*
We say the sequence converges to f uniformly iff for every $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that $|f_i(x) - f(x)| < \epsilon$ for all $i > n$ and for all $x \in U$.

Uniform convergence demands that a single value n, specifying the "tail" of the sequence, can be chosen uniformly for all $x \in U$. *Point-wise convergence* merely asks for the existence of an n for each point $x \in U$, that is, the choice of n can depend on the point x. Hence, uniform convergence is the stronger notion and it implies point-wise convergence.

1.3 Measures and integrals*

We have already seen that norms and scalar products defined by integrals, as in Eqs. (1.25) and (1.59), are commonplace in function vector spaces. To understand such spaces properly we need to understand the integrals and this is where things become tricky. While you do know how to integrate classes of functions you may not yet have seen the actual definition of the most basic type of integral, the *Riemann integral*. To complicate matters further, the Riemann integral is actually not what we need for the present discussion but we require a rather more general integral - the *Lebesgue integral*. Introducing either type of integral properly is labour intense and can easily take up the best part of a lecture course - clearly not something we can indulge in. On the other hand, these integrals form the background for much of the discussion that follows. For these reasons, we will try to sketch out the main ideas, starting with the Riemann integral and then move on to the Lebesgue integral, without going through all the details and proofs. Along the way, we will introduce the ideas of *measures* and *measure spaces* which are a very useful general structures underlying many mathematical constructions.

1.3.1 The Riemann integral

As a warm-up, we would now like to sketch the construction of the Riemann integral - a classical topic for a first year analysis course. We begin with a finite interval $[a, b] \subset \mathbb{R}$, our prospective range of integration. We call a function $\varphi : [a, b] \to \mathbb{R}$ *piecewise constant* if there is a partition $x_0 = a < x_1 < \cdots < x_{n-1} < x_n = b$ of the interval such that for all $x \in]x_{i-1}, x_i[$ the function value $\varphi(x) =: c_i$ is constant, where $i = 1, \ldots, n$. It is very easy to define an integral for piecewise constant functions φ by

$$\int_a^b dx\, \varphi(x) := \sum_{i=1}^{n} (x_i - x_{i-1}) c_i \,, \tag{1.79}$$

that is, by simply summing up the areas of the rectangles associated to the segments of the partition. It can be shown that this integral is well-defined (that is, it is independent of the choice of partition), that the space, $\Gamma([a, b])$ of all piecewise constant functions on $[a, b]$ is a vector space and that the integral (1.79) is a linear functional on this space.

Exercise 1.36. *Show that the space $\Gamma([a, b])$ of all piecewise constant functions an the interval $[a, b] \subset \mathbb{R}$ is a vector space. Also show that the integral (1.79) is a linear functional on $\Gamma([a, b])$.*

Of course having an integral for piecewise constant functions is not yet good enough. The idea is to define the Riemann integral by approximating functions by piecewise constant functions and then taking an appropriate limit. For two functions $f, g : [a, b] \to \mathbb{R}$ we say that $f \leq g$ ($f \geq g$) iff $f(x) \leq g(x)$ ($f(x) \geq g(x)$) for all $x \in [a, b]$. For an arbitrary (but bounded) function $f : [a, b] \to \mathbb{R}$ we introduce the upper and lower integral by [4]

$$\int_a^{*\,b} dx\, f(x) \;:=\; \inf \left\{ \int_a^b dx\, \varphi(x) \,|\, \varphi \in \Gamma([a, b]) \,,\; \varphi \geq f \right\} \tag{1.80}$$

$$\int_{*\,a}^b dx\, f(x) \;:=\; \sup \left\{ \int_a^b dx\, \varphi(x) \,|\, \varphi \in \Gamma([a, b]) \,,\; \varphi \leq f \right\} \,. \tag{1.81}$$

Note that these definitions precisely capture the intuition of approximating f by piecewise constant functions, the upper integral by using piecewise constant function "above" f, the lower integral by using piecewise constant functions "below" f. (See Fig. 2.) After this set-up we are ready to define the Riemann integral.

[4] For a subset $S \subset \mathbb{R}$, the *supremum*, $\sup(S)$, is the smallest number which is greater equal than all elements of S. Likewise, the *infimum*, $\inf(S)$, of S is the largest number less equal than all elements in S.

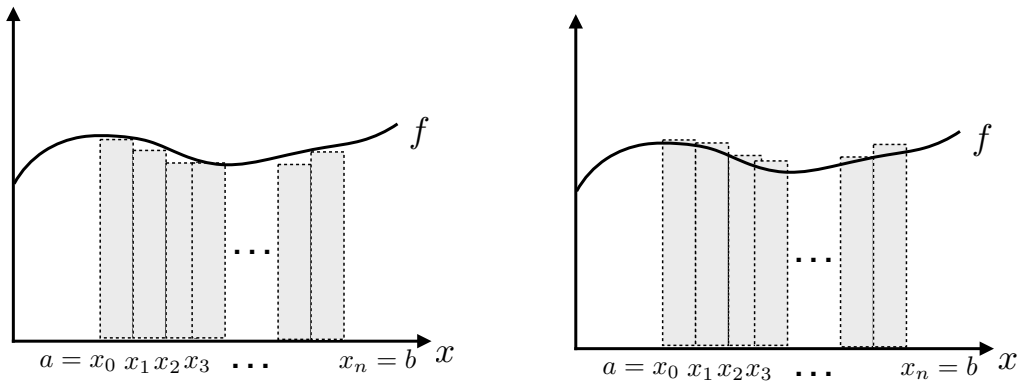

Figure 2: Approximation of a function f by lower and upper piecewise constant functions, used in the definition of the Riemann integral.

Definition 1.23. *(Riemann integral) A (bounded) function $f : [a, b] \to \mathbb{R}$ is called Riemann integrable iff the upper and lower integrals in Eqs. (1.80) and (1.81) are equal. In this case, the common value is called the Riemann integral of f and is written as*

$$\int_a^b dx\, f(x)\ . \tag{1.82}$$

This is where the work begins. Now we have to derive all the properties of the Riemann integral (which you are already familiar with) from this definition. We are content with citing

Theorem 1.37. *All continuous and piecewise continuous functions $f : [a, b] \to \mathbb{R}$ are Riemann integrable.*

The proof of the above theorem, along with a proof of all the other standard properties of the Riemann integral starting from Def. 1.23, can be found in most first year analysis textbooks, see for example [10]. We cannot possibly spend more time on this but hopefully the above set-up gives a clear enough starting point to pursue this independently (which I strongly encourage you to do).

We have already used integrals, somewhat naively, in the definitions (1.25) and (1.59) of a norm and a scalar product on the space $\mathcal{C}([a, b])$ of continuous functions. We can now be more precise and think of these integrals as Riemann integrals in the above sense. Unfortunately, this does not lead to particularly nice properties. For example, for $\mathcal{C}([a, b])$ with the norm (1.25), there are Cauchy convergent sequences (of functions) which converge to non-continuous functions.

Exercise 1.38. *Find an example of a sequence of functions in $\mathcal{C}([a, b])$ (where you can choose a and b) which converges, relative to the norm (1.25), to a function not contained in $\mathcal{C}([a, b])$.*

This means that $\mathcal{C}([a, b])$ is not a complete space, in the sense of Def. (1.18). What is worse, it turns out that even the space of all Riemann integrable functions on $[a, b]$ is not complete, so the deficiency is with the Riemann integral. Essentially, the problem is that the Riemann integral is based on a too simple method of approximation, using finite partitions into intervals. To fix this, we need to be able to measure the "length" of sets which are more complicated than intervals and this leads to the ideas of *measures* and *measure sets* which we now introduce.

1.3.2 Measures and measure sets

This discussion starts fairly general with an arbitrary set X, subsets of which we would like to measure. Typically, not all such subsets will be suitable for measurement and we need to single out a sufficiently nice class of subsets Σ, which is called a *σ-algebra*.

Definition 1.24. *(σ-algebra) For a set X, a set of subsets Σ of X is called a σ-algebra if the following is satisfied.*
(S1) $\{\}\,,\ X \in \Sigma$
(S2) $S \in \Sigma \ \Rightarrow\ X \setminus S \in \Sigma$
(S3) $S_i \in \Sigma$ *for* $i = 1, 2, \cdots \ \Rightarrow\ \bigcup_{i=1}^{\infty} S_i \in \Sigma$

On a σ-algebra, a measure μ is defined as:

Definition 1.25. *For a set X with σ-algebra Σ a function $\mu : \Sigma \to \mathbb{R}^{\geq 0} \cup \{\infty\}$ is called a measure iff*
(M1) $\mu(\{\}) = 0$
(M2) For $S_i \in \Sigma$, where $i = 1, 2, \cdots$, and the S_i mutually disjoint we have

$$\mu\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \mu(S_i)\ . \tag{1.83}$$

The triple (X, Σ, μ) is called a measure space.

The idea is that the σ-algebra generalises the simple notion of intervals and interval partitions used for the Riemann integral, while the measure μ generalises the notion of the length of an interval. The crucial difference to what happened for the Riemann integral (where partitions into finitely many intervals were used) is that we consider countably infinite unions in (S3) and (M2).

While from (M1) the measure of the empty set in a measure space is zero, there may well be other non-empty sets with zero measure. This motivates the following definition.

Definition 1.26. *For a measure space (X, Σ, μ) a set $S \in \Sigma$ is said to have measure zero iff $\mu(S) = 0$.*

Proceeding in analogy with the Riemann integral we can now attempt to define an integral for functions $f : X \to \mathbb{R}$, based on the measure space (X, Σ, μ). We start with the *characteristic function* χ_S for a subset $S \subset X$ defined as

$$\chi_S(x) := \begin{cases} 1 & \text{for} \quad x \in S \\ 0 & \text{for} \quad x \notin S \end{cases}\ . \tag{1.84}$$

A function $\varphi : X \to \mathbb{R}$ is called *simple* if it is of the form $\varphi = \sum_{i=1}^{k} \alpha_i \chi_{S_i}$ for sets $S_i \in \Sigma$ and $\alpha_i \in \mathbb{R}$. (These are the generalisations of the piecewise constant functions used for the Riemann integral.) For a non-negative simple function φ the integral is defined as

$$\int_X \varphi \, d\mu := \sum_{i=1}^{k} \alpha_i \mu(S_i)\ . \tag{1.85}$$

Note that this captures the intuition: we simply multiply the "height" α_i of the function over each set S_i with the measure $\mu(S_i)$ of this set. The set-up of the general integral on a measure space is summarised in the following definition.

Definition 1.27. *Let (X, Σ, μ) be a measure set. We say a function $f : X \to \mathbb{R}$ is measurable iff $\{x \in X \,|\, f(x) > \alpha\} \in \Sigma$ for all $\alpha \in \mathbb{R}$. For a non-negative measurable function $f : X \to \mathbb{R}$ the integral is defined as*

$$\int_X f \, d\mu := \sup\left\{ \int \varphi \, d\mu \,|\, \varphi \text{ simple and } 0 \leq \varphi \leq f \right\}\ . \tag{1.86}$$

A function is called integrable iff it is measurable and $\int_X |f|\,d\mu$ is finite. For a measurable, integrable function $f : X \to \mathbb{R}$ the integral is then defined as

$$\int_X f\,d\mu := \int_X f^+\,d\mu - \int_X f^-\,d\mu\,, \tag{1.87}$$

where $f^\pm(x) := \max\{\pm f(x), 0\}$ are the positive and negative "parts" of f. The space of all integrable functions $f : X \to \mathbb{R}$ is also denoted by $\mathcal{L}^1(X)$. The above construction can be generalised to complex-valued functions $f : X \to \mathbb{C}$ by splitting up into real and imaginary parts.

It can be shown that the integral defined above is linear and that the space $\mathcal{L}^1(X)$ is a (sub) vector space. The obvious course of action is to try to make $\mathcal{L}^1(X)$ into a normed vector space by using the above integral to define a norm. However, there is a twist. If there are non-trivial sets $S \in \Sigma$ which are measure zero, then we have non-trivial functions, for example the characteristic function χ_S, which integrate to zero. This is in conflict with the requirement (N1) for a norm in Def. 1.6 which asserts that the zero vector (that is, the zero function) is the only vector with length zero. Fortunately, this problem can be fixed by identifying two functions $f, g \in \mathcal{L}^1(X)$ if they only differ on a set of measure zero, so

$$f \sim g \quad :\Longleftrightarrow \quad \mu(\{x \in X \,|\, f(x) \neq g(x)\}) = 0\,. \tag{1.88}$$

The space of so-obtained classes of functions in $\mathcal{L}^1(X)$ is called $L^1(X)$ and this set can be made into a normed space defining the norm $\|\cdot\| : L^1(X) \to L^1(X)$ by

$$\| f \| := \int_X |f|\,d\mu\,. \tag{1.89}$$

We can generalise this construction and for $1 \leq p < \infty$ define the spaces

$$\mathcal{L}^p(X) = \left\{ f \,|\, f \text{ is measurable and } \left(\int_X |f|^p\,d\mu\right)^{1/p} \text{ finite} \right\}\,. \tag{1.90}$$

On these spaces, we can identify functions as in Eq. (1.88) and the resulting space of classes, $L^p(X)$, can be made into normed vector spaces with norm

$$\| f \|_p := \left(\int_X |f|^p\,d\mu\right)^{1/p}\,. \tag{1.91}$$

Exercise 1.39. *Show that that Eq. (1.91) defines a norm on $L^p(X)$. (Hint: To show (N3) use Minkowski's inequality (1.24).)*

The all-important statement about the normed vector spaces $L^p(X)$ is the following.

Theorem 1.40. *The normed vector spaces $L^p(X)$ in Eq. (1.90) with norm (1.91) are complete.*

Proof. The proof can be found, for example, in Ref. [5]. □

Our previous experience suggest that the space $L^2(X)$ can in fact be given the structure of an inner product vector space. To do this we need the following

Exercise 1.41. *Show that for $f, g \in L^2(X)$ it follows that $\bar{f}g$ is integrable, that is $\bar{f}g \in L^1(X)$. (Hint: Use Hölder's inequality in Eq. (1.24).)*

Hence, for two functions $f, g \in L^2(X)$, the prospective scalar product

$$\langle f, g \rangle := \int_X f^* g \, d\mu \tag{1.92}$$

is well-defined.

Exercise 1.42. *Show that Eq. (1.92) defines a scalar product on $L^2(X)$.*

Recall that X is still an arbitrary set so the above construction of measure sets and integrals is very general. Especially, the statement (1.40) about completeness is quite powerful. It says that the spaces $L^p(X)$ behave nicely in terms of convergence properties - every Cauchy series converges. This is quite different from what we have seen for the Riemann integral. We should now exploit this general construction by discussing a number of examples.

1.3.3 Examples of measure spaces

Probability: This is perhaps an unexpected example and somewhat outside our main line of development. It might still be useful to illustrate the strength and breadth of the general approach and to make connections with other courses, where probability is discussed in more detail. The basis of probability theory is formed by the *Kolmogorov axioms* which are given in the following definition.

Definition 1.28. *(Kolmogorov axioms of probability) Let Ω be a set, Σ a σ-algebra on Ω and $p : \Sigma \to \mathbb{R}$ a function. The triplet (Ω, Σ, p) is called a probability space if the following holds.*
(K1) $p(E) \geq 0$ for all $E \in \Sigma$
(K2) $p(\Omega) = 1$
(K3) For $E_i \in \Sigma$, where $i = 1, 2, \cdots$, and the E_i mutually disjoint we have

$$p\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} p(E_i) \, . \tag{1.93}$$

In this case, Ω is called the sample space, Σ the event space and p the probability measure.

Comparing this definition with Def. 1.25 shows that a probability space (Ω, Σ, p) is, in fact, a particular measure space with a few additional properties for p, in order to make it a suitable measure for probability. (The condition (M1) in Def. 1.25, $\mu(\{\}) = 0$, can be deduced from the Kolmogorov axioms.) The measurable functions $f : \Omega \to \mathbb{R}$ on this space are also called *random variables* and the integral

$$E[f] := \int_\Omega f \, dp \tag{1.94}$$

is called the *expectation value* of the random variable f.

Counting measure: Choose $X = \mathbb{N}$ to be the natural numbers, Σ_c to be all subsets of $\mathbb{N}$ and for a set $S \in \Sigma_c$ define the measure $\mu_c(S)$ as the number of elements of S (with ∞ permitted). Then, $(\mathbb{N}, \Sigma_c, \mu_c)$ is a measure space and μ_c is called the *counting measure* on $\mathbb{N}$. The functions $f : \mathbb{N} \to \mathbb{R}$ on this space can be identified with the sequences $(x_i)_{i=1}^{\infty}$ (where $x_i = f(i-1)$) and the integrable "functions" are those with $\sum_{i=1}^{\infty} |x_i| < \infty$ while the integral is simply the series $\sum_{i=1}^{\infty} x_i$. Specialising from the general construction (1.90), we can define the spaces $\ell^p := \mathcal{L}^p(\mathbb{N})$ which are explicitly given by

$$\ell^p = \left\{ (x_i)_{i=1}^{\infty} \mid \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} < \infty \right\} \, , \tag{1.95}$$

which are normed vector spaces with norm

$$\| (x_i) \|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} . \tag{1.96}$$

Recall that we know from Theorem 1.40 that the spaces ℓ^p are complete, relative to this norm. The space ℓ^2 is an inner product space with scalar product

$$\langle (x_i), (y_i) \rangle = \sum_{i=1}^{\infty} \bar{x}_i y_i \, . \tag{1.97}$$

Lebesgue measure: The Lebesgue measure provides a measure on $\mathbb{R}$ (and, more generally, on $\mathbb{R}^n$) but constructing it takes some effort and time. Instead we take a short-cut and simply state the following theorem.

Theorem 1.43. *There is a σ-algebra Σ_L on $\mathbb{R}$ and a measure μ_L on Σ_L, called the Lebesgue measure, with the following properties.*
(L1) All intervals $[a, b] \in \Sigma_L$.
(L2) $\mu_L([a, b]) = b - a$
(L3) The sets S of measure zero in Σ_L are characterised as follows. For any $\epsilon > 0$ there are intervals $[a_i, b_i]$, where $i = 1, 2, \cdots$, such that $S \subset \bigcup_{i=1}^{\infty} [a_i, b_i]$ and $\sum_{i=1}^{\infty} (b_i - a_i) < \epsilon$.
The measure space $(\mathbb{R}, \Sigma_L, \mu_L)$ is uniquely characterised by these properties.

Note that the Lebesgue measure leads to non-trivial sets with measure zero. For example, any finite set of points in $\mathbb{R}$ has measure zero.

Exercise 1.44. *Show that any finite set of points and any sequence $(x_i)_{i=1}^{\infty}$ in $\mathbb{R}$ have measure zero with respect to the Lebesgue measure μ_L.*

The above Lebesgue measure has been defined in $\mathbb{R}$ and, hence, measures length but it can be suitably generalised to $\mathbb{R}^2$ to measure areas, to $\mathbb{R}^3$ to measure volumes and to $\mathbb{R}^n$ to measure generalised volumes in n dimensions. This means we have measure spaces $(\mathbb{R}^n, \Sigma_L, \mu_L)$. Of course, this induces measure spaces on subsets $U \subset \mathbb{R}^n$ as long as $U \in \Sigma_L$ by simply defining the restricted σ-algebra $\Sigma_L(U) = \{S \in \Sigma_L \,|, S \subset U\}$ and in this way we have measure spaces $(U, \Sigma_L(U), \mu_L)$.

The integral associated with the measure space $(\mathbb{R}^n, \Sigma_L, \mu_L)$ (or, more generally, with the measure space $(U, \Sigma_L(U), \mu_L)$) is called the *Lebesgue integral* and it is written as

$$\int_U dx \, f(x) \, . \tag{1.98}$$

The Lebesgue-integrable functions are those for which $\int_U dx \, |f(x)|$ is finite and following the general construction, we can define the spaces

$$\mathcal{L}^p(U) = \left\{ f \mid \left(\int_U dx \, |f(x)|^p \right)^{1/p} < \infty \right\} \, . \tag{1.99}$$

The associated spaces $L^p(U)$, obtained after the identification (1.88) of functions which only differ on sets of measure zero, are complete normed vector spaces with norm

$$\| f \|_p = \left(\int_U dx \, |f(x)|^p \right)^{1/p} . \tag{1.100}$$

The space $L^2(U)$ is an inner product vector space with inner product

$$\langle f, g \rangle = \int_U dx\, f(x)^* g(x) \,. \tag{1.101}$$

As for the relation between the Riemann and the Lebesgue integrals we have

Theorem 1.45. *Every Riemann-integrable function is Lebesgue integrable and for such functions the two integrals are equal.*

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

This means that for practical calculations with the Lebesgue integral we can use all the usual rules of integration, as long as the integrand is sufficiently "nice" (for example, Riemann integrable). While the Riemann-integrable functions are included in the Lebesgue-integrable ones, the latter set is much larger and this facilitates the completeness properties associated to the Lebesgue integral. In the following, when we write an integral, we usually refer to the Lebesgue integral.

2 Banach and Hilbert spaces*

Banach and Hilbert spaces are central objects in functional analysis and their systematic mathematical study easily fills a lecture course. Clearly, we cannot afford to do this so we will focus on basics and some of the results relevant to our later applications. Proofs are provided explicitly only when they can be provided in a concise fashion and references will be given whenever proofs are omitted. Our main focus will be on Hilbert spaces which provide the arena of most of the applications discussed later and indeed are the correct setting for quantum mechanics. We begin with the more basic notion of Banach spaces, then move on to Hilbert spaces and finish with a discussion of operators on Hilbert spaces.

2.1 Banach spaces

We begin with the definition of Banach spaces.

Definition 2.1. *A Banach space is a complete normed vector space.*

Recall from Def. 1.18 that completeness means convergence of every Cauchy series, so a Banach space is a vector space with a basic notion of geometry, as provided by the norm, and good convergence properties. We have already encountered several important examples of Banach spaces which we recall.

2.1.1 Examples of Banach spaces

- The normed vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ with any of the norms $\|\cdot\|_p$ in Eq. (1.23) are complete (essentially because the real numbers are complete) and they are, hence, Banach spaces.
- For a measure space (X, Σ, μ), where X is an arbitrary set, we have defined the Banach spaces $L^p(X)$ in Eq. (1.90). It consists of all measurable functions $f : X \to \mathbb{R}$ (or $f : X \to \mathbb{C}$) with $\left(\int_X |f|^p \, d\mu\right)^{1/p}$ finite and the norm is given in Eq. (1.91). Completeness of these normed vector space is asserted by Theorem 1.40. This is a very large class of Banach spaces which includes many interesting examples, some of which we list now.
- Associated to the measure space $(\mathbb{N}, \Sigma_c, \mu_c)$ with counting measure μ_c, introduced in the previous sub-section, we have the space ℓ^p of all sequences $(x_i)_{i=1}^\infty$ in $\mathbb{R}$ (or $\mathbb{C}$) with $\left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p}$ finite. The norm on this space is provided by Eq. (1.96) and Theorem 1.40 guarantees completeness.
- For a Lebesgue measure space $(U, \Sigma_L(U), \mu_L)$, where $U \subset \mathbb{R}^n$ is a Lebesgue measurable set, we have defined the space $L^p(U)$ which consists of measurable functions $f : U \to \mathbb{R}$ (or $f : U \to \mathbb{C}$) with $\left(\int_U dx \, |f(x)|^p\right)^{1/p}$ finite. (More precisely, $L^p(U)$ consists of classes of such functions, identified according to Eq. (1.88).) The norm on these spaces is given by Eq. (1.100) and completeness is guaranteed by Theorem 1.40. We will sometimes write $L^p_{\mathbb{R}}(U)$ or $L^p_{\mathbb{C}}(U)$ to indicate whether we are talking about real or complex valued functions.

An important theoretical property for the Banach spaces $L^p([a, b])$ which we will need later is

Theorem 2.1. *The space $\mathcal{C}([a, b])$ is dense in $L^p([a, b])$. Further, the space $\mathcal{C}^\infty_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.*

Proof. For the proof see, for example, Ref. [5]. □

2.2 Hilbert spaces

Hilbert spaces are defined as follows.

Definition 2.2. *An inner product vector space $\mathcal{H}$ is called a Hilbert space if it is complete (relative to the norm associated to the scalar product).*

We know that the Banach spaces given in the previous sub-section can be equipped with a scalar product when $p = 2$ and this provides us with examples of Hilbert spaces.

2.2.1 Examples of Hilbert spaces

- The inner product vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ with inner product (1.56) are complete (since they are Banach spaces relative to the norm associated to this scalar product) and they are, hence, Hilbert spaces. We know that any finite-dimensional inner product vector space over the field $\mathbb{R}$ (over the field $\mathbb{C}$) is isomorphic to $\mathbb{R}^n$ (or $\mathbb{C}^n$) by mapping a vector to its coordinate vector relative to some chosen basis. If we choose an ortho-normal basis we know from Eq. (1.32) that, in terms of the coordinates, the scalar product can be expressed in terms of the standard scalar product on $\mathbb{R}^n$ or $\mathbb{C}^n$. Together, these facts imply that any finite-dimensional inner product vector space over $\mathbb{R}$ or $\mathbb{C}$ is a Hilbert space.

- For the measure set (X, Σ, μ), the space $L^2(X)$, defined in Eq. (1.90) is an inner product vector space with inner product given by Eq.(1.92). We already know that this is a Banach space (relative to the norm associated to the scalar product), so $L^2(X)$ is complete and, hence, a Hilbert space.

- Associated to the measure space $(\mathbb{N}, \Sigma_c, \mu_c)$ with counting measure μ_c we have the space ℓ^2 of all sequences $(x_i)_{i=1}^{\infty}$ in $\mathbb{R}$ (or $\mathbb{C}$) with $\left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2}$ finite. An inner product on this space is given by Eq. (1.97). Since ℓ^2 is a Banach space it is complete and is, hence, also a Hilbert space.

- For a Lebesgue measure space $(U, \Sigma_L(U), \mu_L)$, where $U \subset \mathbb{R}^n$ is a Lebesgue measurable set, we have defined the space $L^2(U)$ which consists of measurable functions $f : U \to \mathbb{R}$ (or $f : U \to \mathbb{C}$) with $\left(\int_U dx\, |f(x)|^2\right)^{1/2}$ finite. This is an inner product vector space with inner product given by Eq. (1.101). Following the same logic as before, $L^2(U)$ is a Banach space and it is, hence, complete and a Hilbert space. This space is also called the Hilbert space of *square integrable functions* on U. We will sometimes write $L^2_{\mathbb{R}}(U)$ or $L^2_{\mathbb{C}}(U)$ to indicate whether we are talking about real or complex valued functions.

- There is a useful generalisation of the previous example which we will need later. On an interval $[a, b] \subset \mathbb{R}$ introduce an everywhere positive, integrable function $w : [a, b] \to \mathbb{R}^{>0}$, called the *weight function*, and define the space $\mathcal{L}^2_w([a, b])$ as the space of measurable functions $f : [a, b] \to \mathbb{R}$ with $\left(\int_{[a,b]} dx\, w(x)|f(x)|^2\right)^{1/2}$ finite. We can introduce

$$\langle f, g\rangle := \int_{[a,b]} dx\, w(x) f(x)^* g(x)\ . \tag{2.1}$$

  With the usual identification of functions, as in Eq. (1.88), this leads to a Hilbert space, called $L^2_w([a, b])$, with scalar product (2.1).

2.2.2 Orthogonal basis

We have seen that an ortho-normal basis for a finite-dimensional Hilbert space is really the most convenient tool to carry out calculations. We should now discuss the concept of ortho-normal basis for infinite-dimensional Hilbert spaces. One question we need to address first is what happens when we take a limit inside one of the arguments of the scalar product.

Lemma 2.1. *For a convergent sequence $(\mathbf{v}_i)_{i=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ and any vector $\mathbf{w} \in \mathcal{H}$ we have $\lim_{i\to\infty}\langle \mathbf{w}, \mathbf{v}_i\rangle = \langle \mathbf{w}, \lim_{i\to\infty} \mathbf{v}_i\rangle$. A similar statement applies to the first argument of the scalar product.*

Proof. Set $\mathbf{v} := \lim_{i\to\infty} \mathbf{v}_i$ and consider the inequality

$$|\langle \mathbf{w}, \mathbf{v}_i\rangle - \langle \mathbf{w}, \mathbf{v}\rangle| = |\langle \mathbf{w}, \mathbf{v}_i - \mathbf{v}\rangle| \leq \| \mathbf{w} \| \, \| \mathbf{v}_i - \mathbf{v} \| \, , \tag{2.2}$$

where the last step follows from the Cauchy-Schwarz inequality (1.28). Convergence of $(\mathbf{v}_i)$ to $\mathbf{v}$ means we can find, for each $\epsilon > 0$, a k such that $\| \mathbf{v}_i - \mathbf{v} \| < \epsilon/\| \mathbf{w} \|$ for all $i > k$. This implies that

$$|\langle \mathbf{w}, \mathbf{v}_i\rangle - \langle \mathbf{w}, \mathbf{v}\rangle| < \epsilon \tag{2.3}$$

for all $i > k$ and, hence, that $\lim_{i\to\infty}\langle \mathbf{w}, \mathbf{v}_i\rangle = \langle \mathbf{w}, \mathbf{v}\rangle$. The analogous statement for the first argument of the scalar product follows from the above by using the property (S1) in Def. 1.9. □

The above lemma says that a limit can be "pulled out" of the arguments of a scalar product, an important property which we will use frequently. Another technical statement we require asserts the existence and uniqueness of a point of minimal distance.

Lemma 2.2. *(Nearest point to a subspace) Let W be a closed, non-trivial sub vector space of a Hilbert space $\mathcal{H}$. Then, for every $\mathbf{v} \in \mathcal{H}$ there is a unique $\mathbf{w}_0 \in W$ such that*

$$\| \mathbf{v} - \mathbf{w}_0 \| \leq \| \mathbf{v} - \mathbf{w} \| \tag{2.4}$$

for all $\mathbf{w} \in W$.

Proof. We set $\delta := \inf\{\| \mathbf{v} - \mathbf{w} \| \, | \, \mathbf{w} \in W\}$ and choose a sequence $(\mathbf{w}_i)_{i=1}^{\infty}$ contained in W with $\lim_{i\to\infty} \| \mathbf{v} - \mathbf{w}_i \| = \delta$. We want to proof that this sequence is a Cauchy sequence so we consider

$$\| \mathbf{w}_i - \mathbf{w}_j \|^2 = \| (\mathbf{w}_i - \mathbf{v}) + (\mathbf{v} - \mathbf{w}_j) \|^2 = 2\| \mathbf{w}_i - \mathbf{v} \|^2 + 2\| \mathbf{v} - \mathbf{w}_j \|^2 - 4\| \frac{1}{2}(\mathbf{w}_i + \mathbf{w}_j) - \mathbf{v} \|^2 \, , \tag{2.5}$$

where the parallelogram law (1.29) has been used in the last step. Since W is a sub vector space it is clear that the vector $\frac{1}{2}(\mathbf{w}_i + \mathbf{w}_j)$ which appears in the last term above is in W. This means that $\| \frac{1}{2}(\mathbf{w}_i + \mathbf{w}_j) - \mathbf{v} \| \geq \delta$ and we have

$$\| \mathbf{w}_i - \mathbf{w}_j \|^2 \leq 2\| \mathbf{w}_i - \mathbf{v} \|^2 + 2\| \mathbf{v} - \mathbf{w}_j \|^2 - 4\delta^2 \, . \tag{2.6}$$

The RHS of this inequality goes to zero as $i, j \to \infty$ which shows that $(\mathbf{w}_i)$ is indeed a Cauchy sequence. Since a Hilbert space is complete we know that every Cauchy sequance converges and we set $\mathbf{w}_0 := \lim_{i\to\infty} \mathbf{w}_i$. Since W is assumed to be closed it follows that $\mathbf{w}_0 \in W$ and, with Lemma 2.1 we have $\| \mathbf{v} - \mathbf{w}_0 \| = \delta$. This means that $\mathbf{w}_0$ is indeed a point in W of minimal distance to $\mathbf{v}$.

It remains to show that $\mathbf{w}_0$ is unique. For this assume there is another $\tilde{\mathbf{w}} \in W$ such that $\| \mathbf{v} - \tilde{\mathbf{w}} \| = \delta$. Then repeating the calculation (2.5) with $\mathbf{w}_i$ and $\mathbf{w}_j$ replaced by $\mathbf{w}_0$ and $\tilde{\mathbf{w}}$ we have

$$\| \mathbf{w}_0 - \tilde{\mathbf{w}} \|^2 = 2\| \mathbf{w}_0 - \mathbf{v} \|^2 + 2\| \mathbf{v} - \tilde{\mathbf{w}} \|^2 - 4\| \frac{1}{2}(\mathbf{w}_0 + \tilde{\mathbf{w}}) - \mathbf{v} \|^2 \leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0 \, , \tag{2.7}$$

so that $\tilde{\mathbf{w}} = \mathbf{w}_0$. □

The above Lemma is the main technical result needed to proof the following important statement about direct sum decompositions in Hilbert spaces.

Theorem 2.2. *(Direct sum decomposition) For any closed sub vector space W of a Hilbert space $\mathcal{H}$ we have $\mathcal{H} = W \oplus W^{\perp}$.*

Proof. For $W = \{\mathbf{0}\}$ we have $W^{\perp} = \mathcal{H}$ so that the statement is true. Now assume that $W \neq \{\mathbf{0}\}$. For any $\mathbf{v} \in \mathcal{H}$ we can choose, from Lemma 2.2, a "minimal distance" $\mathbf{w}_0 \in W$ and write

$$\mathbf{v} = \mathbf{w}_0 + (\mathbf{v} - \mathbf{w}_0) \, . \tag{2.8}$$

This is our prospective decomposition so we want to show that $\mathbf{v} - \mathbf{w}_0 \in W^{\perp}$. To do this we assume that $\mathbf{v} - \mathbf{w}_0 \notin W^{\perp}$ so that there is a $\mathbf{u} \in W$ such that $\langle \mathbf{v} - \mathbf{w}_0, \mathbf{u} \rangle = 1$. A short calculation for any $\alpha \in \mathbb{R}$ show that

$$\| \, \mathbf{v} - \mathbf{w}_0 - \alpha \mathbf{u} \, \|^2 = \| \, \mathbf{v} - \mathbf{w}_0 \, \|^2 - 2\alpha + \alpha^2 \| \, \mathbf{u} \, \|^2 \, . \tag{2.9}$$

For sufficiently small α the sum of the last two terms on the RHS are negative so in this case $\| \, \mathbf{v} - \mathbf{w}_0 - \alpha \mathbf{u} \, \| < \| \, \mathbf{v} - \mathbf{w}_0 \, \|$. This contradicts the minimality property of $\mathbf{w}_0$ and we conclude that indeed $\mathbf{v} - \mathbf{w}_0 \in W^{\perp}$. Hence, we have $\mathcal{H} = W + W^{\perp}$. That this sum is direct has already been shown in Exercise 1.16. □

One of our goals is to obtain a generalisation of the formula (1.31) to the infinite-dimensional case. Recall that the sequence $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ is called an ortho-normal system iff it satisfies $\langle \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j \rangle = \delta_{ij}$. We need to worry about the convergence properties of such ortho-normal systems and this is covered by the following

Lemma 2.3. *(Bessel inequality) Let $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ be an ortho-normal system in a Hilbert space $\mathcal{H}$ and $\mathbf{v} \in \mathcal{H}$. Then we have the following statements.*
(i) $\sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2$ converges and $\sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2 \leq \| \, \mathbf{v} \, \|^2$.
(ii) $\sum_{i=1}^{\infty} \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \boldsymbol{\epsilon}_i$ converges.

Proof. (i) Introduce the partial sums $\mathbf{s}_k = \sum_{i=1}^{k} \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \boldsymbol{\epsilon}_i$. A short calculation show that

$$\| \, \mathbf{v} - \mathbf{s}_k \, \|^2 = \| \, \mathbf{v} \, \|^2 - \sum_{i=1}^{k} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2 \, , \tag{2.10}$$

and, hence, $\sum_{i=1}^{k} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2 \leq \| \, \mathbf{v} \, \|^2$. This means the series $\sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2$ is bounded (and increasing) and, hence, converges. Since all its partial sums satisfy the stated inequality so does the limit.
(ii) A short calculation shows that

$$\| \, \mathbf{s}_k - \mathbf{s}_l \, \|^2 = \sum_{i=k}^{l} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2 \tag{2.11}$$

and since $\sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2$ is a Cauchy series the RHS of this expression can be made arbitrarily small. This means that $\sum_{i=1}^{\infty} \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \boldsymbol{\epsilon}_i$ is a Cauchy series which converges due to completeness of $\mathcal{H}$. □

We are now ready to tackle (infinite) ortho-normal system in a Hilbert space $\mathcal{H}$. We recall that an ortho-normal system is a sequence $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ with $\langle \boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j \rangle = \delta_{ij}$. By the span, $\mathrm{Span}(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$, we mean the sub-space which consists of all finite linear combinations of vectors $\boldsymbol{\epsilon}_i$. The following theorem provides the basic statements about ortho-normal systems.

Theorem 2.3. *Let $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ be an ortho-normal system in a Hilbert space $\mathcal{H}$. Then, the following statements are equivalent.*
(i) Every $\mathbf{v} \in \mathcal{H}$ can be written as $\mathbf{v} = \sum_{i=1}^{\infty} \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \boldsymbol{\epsilon}_i$.
(ii) For every $\mathbf{v} \in \mathcal{H}$ we have $\| \mathbf{v} \|^2 = \sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2$.
(iii) If $\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle = 0$ for all $i = 1, 2, \cdots$ then $\mathbf{v} = \mathbf{0}$.
(iv) $\overline{\mathrm{Span}}(\boldsymbol{\epsilon}_i)_{i=1}^{\infty} = \mathcal{H}$

Proof. For a statement of this kind it is sufficient to show that (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) ⇒ (i) and we will proceed in this order.
(i) ⇒ (ii): This follows easily by inserting the infinite sum from (i) into $\| \mathbf{v} \|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$ and using Lemma 2.1.
(ii) ⇒ (iii): If $\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle = 0$ for all i then the relation in (ii) implies that $\| \mathbf{v} \| = 0$. Since the zero vector is the only one with norm zero it follows that $\mathbf{v} = \mathbf{0}$.
(iii) ⇒ (iv): Set $W = \overline{\mathrm{Span}}(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ (where we recall that the bar means the closure, so W is a closed sub vector space). Then, from Theorem 2.2, we know that $\mathcal{H} = W \oplus W^{\perp}$. Assume that $W^{\perp} \neq \{\mathbf{0}\}$. Then we have a non-zero vector $\mathbf{v} \in W^{\perp}$ such that $\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle = 0$ for all i, in contradiction of the statement (iii). This means that $W^{\perp} = \{\mathbf{0}\}$ and, hence, $\mathcal{H} = W$.
(iv) ⇒ (i): We know from Lemma 2.3 that $\mathbf{w} := \sum_{i=1}^{\infty} \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \boldsymbol{\epsilon}_i$ is well-defined. A short calculation shows that $\langle \mathbf{v} - \mathbf{w}, \boldsymbol{\epsilon}_j \rangle = 0$ for all j. This means that $\mathbf{v} - \mathbf{w} \in W^{\perp}$ (using the earlier definition of W) but since $\mathcal{H} = W$ from (iv) we have $W^{\perp} = \{\mathbf{0}\}$ and, hence, $\mathbf{v} = \mathbf{w}$. □

After this preparation, we can finally define an ortho-normal basis of a Hilbert space.

Definition 2.3. *An ortho-normal system $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is called an ortho-normal basis if it satisfies any of the conditions in Theorem 2.3.*

Thanks to Theorem 2.3 an ortho-normal basis on a Hilbert space provides us with the desired generalisations of the formulae we have seen in the finite-dimensional case. Eq. (1.31) for the expansion of a vector in terms of an ortho-normal basis simply generalises to

$$\mathbf{v} = \sum_{i=1}^{\infty} \alpha_i \, \boldsymbol{\epsilon}_i \quad \Longleftrightarrow \quad \alpha_i = \langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle \, . \tag{2.12}$$

and we know that the now infinite sum always converges to the vector $\mathbf{v}$. For two vectors $\mathbf{v} = \sum_i \alpha_i \boldsymbol{\epsilon}_i$ and $\mathbf{w} = \sum_j \beta_j \boldsymbol{\epsilon}_j$ we have the generalisation of Eq. (1.32)

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{\infty} \langle \mathbf{v}, \boldsymbol{\epsilon}_i \rangle \langle \boldsymbol{\epsilon}_i, \mathbf{w} \rangle = \sum_{i=1}^{\infty} \alpha_i^* \beta_i \, , \qquad \| \mathbf{v} \|^2 = \sum_{i=1}^{\infty} |\langle \boldsymbol{\epsilon}_i, \mathbf{v} \rangle|^2 = \sum_{i=1}^{\infty} |\alpha_i|^2 \, , \tag{2.13}$$

where, in the first equation, we have also used Lemma 2.1 to pull the infinite sums out of the scalar product. The map $\mathbf{v} \to (\alpha_i)_{i=1}^{\infty}$ defined by Eq. (2.12) is, in fact, a vector space isomorphism $\mathcal{H} \to \ell^2$ from our general Hilbert space into the Hilbert space ℓ^2 of sequences. Moreover, as Eq. (2.13) shows, this map is consistent with the scalar products defined on those two Hilbert spaces. The last equation (2.13) which allows calculating the norm of a vector in terms of an infinite sum over the square of its coordinates is also referred to as *Parseval's equation*. As we will see, for specific examples it can lead to quite non-trivial relations.

Recall the situation in the finite-dimensional case. Finite-dimensional (inner product) vector spaces over $\mathbb{R}$ or $\mathbb{C}$ are isomorphic to $\mathbb{R}^n$ or $\mathbb{C}^n$ and via this identification abstract vectors are described by (coordinate) column vectors. If an ortho-normal basis underlies the identification then the scalar product in coordinates is described by the standard scalar product in $\mathbb{R}^n$ or $\mathbb{C}^n$ (see Eq. (1.32)).

We have now seen that the situation is very similar for infinite-dimensional Hilbert spaces but vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ are replaced by infinite sequences in ℓ^2. However, there is still a problem. While we now appreciate the usefulness of an ortho-normal basis in a Hilbert space we do not actually know yet whether it exists.

We recall that the Hilbert space ℓ^2 which we have introduced earlier consist of sequences $(x_i)_{i=1}^{\infty}$ (where $x_i \in \mathbb{R}$ or $x_i \in \mathbb{C}$) with $\left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2} < \infty$ and has the scalar product

$$\langle (x_i), (y_i) \rangle = \sum_{i=1}^{\infty} x_i^* y_i \ . \tag{2.14}$$

It is not difficult to show that ℓ^2 has an ortho-normal basis. Introduce the sequences

$$\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0, \ldots) \tag{2.15}$$

with a 1 in position i and zero everywhere else. These are obviously the infinite-dimensional analogous of the standard unit vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$. It is easy to see that they are an ortho-normal system relative to the scalar product (2.14).The following theorem states that they form an ortho-normal basis of ℓ^2.

Theorem 2.4. *The "standard unit sequences" $\mathbf{e}_i$, defined in Eq. (2.15), form an ortho-normal basis of the Hilbert space ℓ^2.*

Proof. For a sequence $(x_i) \in \ell^2$ we have, from direct calculation, that

$$\| \ (x_i) \ \|^2 = \sum_{j=1}^{\infty} |x_j|^2 = \sum_{j=1}^{\infty} |\langle \mathbf{e}_j, (x_i) \rangle|^2 \ , \tag{2.16}$$

and, hence, condition (ii) in Theorem 2.3 is satisfied. □

Unfortunately, it is not always as easy to find an ortho-normal basis for a Hilbert space - in fact it does not always exist. The key property for the existence of an ortho-normal basis is that the Hilbert space is *separable* (see Def. 1.21).

Theorem 2.5. *Let $\mathcal{H}$ be a Hilbert space. Then we have the equivalence*

$$\mathcal{H} \text{ has an ortho-normal basis} \iff \mathcal{H} \text{ is separable}$$

Proof. "$\Leftarrow$" We should show that $\mathcal{H}$ has an ortho-normal system which satisfies any of the properties in Theorem 2.3. First, let $(\mathbf{u}_i)_{i=1}^{\infty}$ be the countable dense subset which exists since $\mathcal{H}$ is separable. We can refine this set by omitting any $\mathbf{u}_k$ which is a linear combination of the $\mathbf{u}_i$ with $i < k$ and this leads to a sequence we call $(\mathbf{v}_i)_{i=1}^{\infty}$. To the sequence $(\mathbf{v}_i)_{i=1}^{\infty}$ we can inductively apply the Gram-Schmidt procedure to obtain an ortho-normal system $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ with $\mathrm{Span}(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_k) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ for all k. (Recall that leaving these spans unchanged is one of the properties of the Gram-Schmidt procedure.) Altogether, we have

$$\mathrm{Span}(\boldsymbol{\epsilon}_i)_{i=1}^{\infty} = \mathrm{Span}(\mathbf{v}_i)_{i=1}^{\infty} = \mathrm{Span}(\mathbf{u}_i)_{i=1}^{\infty} \ . \tag{2.17}$$

Since the sequence $(\mathbf{u}_i)_{i=1}^{\infty}$ is dense in $\mathcal{H}$ it follows that $\overline{\mathrm{Span}}(\boldsymbol{\epsilon}_i)_{i=1}^{\infty} = \mathcal{H}$ and this shows property (iv) in Theorem 2.3.

"$\Rightarrow$" Now suppose that $\mathcal{H}$ has an ortho-normal basis $(\boldsymbol{\epsilon}_i)_{i=1}^{\infty}$ and consider finite sums of the form $\sum_{i=1}^{k} \alpha_i \boldsymbol{\epsilon}_i$, where $\alpha_i \in \mathbb{Q}$ in the real case or $\alpha_i \in \mathbb{Q} + i\mathbb{Q}$ in the complex case. Clearly, the set of these finite sums is countable (as $\mathbb{Q}$ is countable) and it is not hard to show that it is dense in $\mathcal{H}$. □

It turns out that most Hilbert spaces which appear in practical applications are separable and, hence, do have an ortho-normal basis. However, the proof is not always straightforward. For the Hilbert space $L^2([a, b])$ of square integrable functions on the interval $[a, b]$ we will proof the existence of an ortho-normal basis in Section 3 in our discussion of the Fourier series.

2.2.3 Dual space

In Eq. (1.33) we have introduced the map $\imath$ from a vector space to its dual and we would now like to discuss this map in the context of a Hilbert space. So we have $\imath : \mathcal{H} \to \mathcal{H}^*$ defined as

$$\imath(\mathbf{v})(\mathbf{w}) := \langle \mathbf{v}, \mathbf{w} \rangle \ . \tag{2.18}$$

This map assigns to every vector $\mathbf{v} \in \mathcal{H}$ a functional $\imath(\mathbf{v})$ in the dual $\mathcal{H}^*$. We have seen that it is always injective and that, in the finite-dimensional case, every functional can be obtained in this way, so $\imath$ is bijective. The following theorem asserts that the last statement continues to be true for Hilbert spaces.

Theorem 2.6. *(Riesz) Let $\mathcal{H}$ be a Hilbert space and $\varphi \in \mathcal{H}^*$ a functional. Then there exists a $\mathbf{v} \in \mathcal{H}$ such that $\varphi = \imath(\mathbf{v})$, where $\imath$ is the map defined in Eq. (2.18).*

Proof. If $\varphi(\mathbf{w}) = 0$ for all $\mathbf{w} \in \mathcal{H}$ then we have $\varphi = \imath(\mathbf{0})$. Let us therefore assume that $\mathrm{Ker}(\varphi) \neq \mathcal{H}$. Then $(\mathrm{Ker}(\varphi))^\perp \neq \{\mathbf{0}\}$ and there exists a $\mathbf{u} \in (\mathrm{Ker}(\varphi))^\perp$ with $\varphi(\mathbf{u}) = 1$. It follows that

$$\varphi(\mathbf{w} - \varphi(\mathbf{w})\mathbf{u}) = \varphi(\mathbf{w}) - \varphi(\mathbf{w})\varphi(\mathbf{u}) = 0 \ , \tag{2.19}$$

so that $\mathbf{w} - \varphi(\mathbf{w})\mathbf{u} \in \mathrm{Ker}(\varphi)$. Since $\mathbf{u} \in (\mathrm{Ker}(\varphi))^\perp$ this mean

$$\langle \mathbf{u}, \mathbf{w} - \varphi(\mathbf{w})\mathbf{u} \rangle = 0 \ . \tag{2.20}$$

Solving this equation for $\varphi(\mathbf{w})$ gives

$$\varphi(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle = \imath(\mathbf{v})(\mathbf{w}) \ , \tag{2.21}$$

where $\mathbf{v} = \mathbf{u}/\| \mathbf{u} \|^2$. □

We already know that the map $\imath$ is injective so this theorem tells us that $\imath : \mathcal{H} \to \mathcal{H}^*$ is an isomorphism just as it is in the finite-dimensional case. This is why it makes sense to generalise Dirac notation

$$\mathbf{w} \to |\mathbf{w}\rangle \ , \qquad \imath(\mathbf{v}) \to \langle \mathbf{v}| \ , \qquad \langle \mathbf{v}|\mathbf{w}\rangle := \langle \mathbf{v}, \mathbf{w} \rangle \tag{2.22}$$

to a Hilbert space. If our Hilbert space has an ortho-normal basis $|\boldsymbol{\epsilon}_i\rangle$, then Eq. (2.12) can be written in Dirac notation as

$$|\mathbf{v}\rangle = \sum_{i=1}^{\infty} |\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{v}\rangle \ , \tag{2.23}$$

and this defines the map $|\mathbf{v}\rangle \to (a_i)_{i=1}^{\infty}$, where $a_i = \langle \boldsymbol{\epsilon}_i|\mathbf{v}\rangle$ are the coordinates of the vector relative to the basis $\boldsymbol{\epsilon}_i$, from $\mathcal{H}$ to ℓ^2 we have mentioned earlier. Writing $a_i = \langle \boldsymbol{\epsilon}_i|\mathbf{v}\rangle$ and $b_i = \langle \boldsymbol{\epsilon}_i|\mathbf{w}\rangle$ we can write Eqs. (2.13) in the form

$$\langle \mathbf{v}|\mathbf{w}\rangle = \sum_{i=1}^{\infty} \langle \mathbf{v}|\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{w}\rangle = \langle (a_i)|(b_i)\rangle_{\ell^2} \ , \qquad \| \ |\mathbf{v}\rangle \ \|^2 = \langle \mathbf{v}|\mathbf{v}\rangle = \sum_{i=1}^{\infty} \langle \mathbf{v}|\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i|\mathbf{v}\rangle = \| \ (a_i) \ \|_{\ell^2} \ , \tag{2.24}$$

where $\langle \cdot|\cdot \rangle_{\ell^2}$ denotes the standard scalar product (1.97) on ℓ^2. Hence, the above map $\mathcal{H} \to \ell^2$ preserves the scalar product. Also the identity operator can, at least formally, be written as

$$\mathrm{id} = \sum_{i=1}^{\infty} |\boldsymbol{\epsilon}_i\rangle\langle\boldsymbol{\epsilon}_i| \ . \tag{2.25}$$

To pursue this somewhat further consider a linear operator $\hat{T} : \mathcal{H} \to \mathcal{H}$, a vector $\mathbf{v} \in \mathcal{H}$ and an ortho-normal basis $(\boldsymbol{\epsilon}_i)$ of $\mathcal{H}$ so that the operator has matrix elements $T_{ij} = \langle \boldsymbol{\epsilon}_i | \hat{T} | \boldsymbol{\epsilon}_j \rangle$ and the vector has coordinates $a_i = \langle \boldsymbol{\epsilon}_i | \mathbf{v} \rangle$. Then the action $\hat{T}|\mathbf{v}\rangle$ of the operator on the vector has coordinates $b_i := \langle \boldsymbol{\epsilon}_i | \hat{T} | \mathbf{v} \rangle$ which can be written as

$$b_i = \langle \boldsymbol{\epsilon}_i | \hat{T} | \mathbf{v} \rangle = \sum_j \langle \boldsymbol{\epsilon}_i | \hat{T} | \boldsymbol{\epsilon}_j \rangle \langle \boldsymbol{\epsilon}_j | \mathbf{v} \rangle = \sum_j T_{ij} a_j \ . \tag{2.26}$$

In this way, the Hilbert space action of an operator on a vector is turned into a multiplication of a matrix and a vector (although infinite-dimensional) in ℓ^2. Similarly, composition of operators in $\mathcal{H}$ turns into "matrix multiplication" in ℓ^2 as in the following exercise.

Exercise 2.7. *For two operators $\hat{T}, \hat{S} : \mathcal{H} \to \mathcal{H}$ with matrix elements $T_{ij} = \langle \boldsymbol{\epsilon}_i | \hat{T} | \boldsymbol{\epsilon}_j \rangle$ and $S_{ij} = \langle \boldsymbol{\epsilon}_i | \hat{S} | \boldsymbol{\epsilon}_j \rangle$ show that the matrix elements of $\hat{T} \circ \hat{S}$ are given by $\langle \boldsymbol{\epsilon}_i | \hat{T} \circ \hat{S} | \boldsymbol{\epsilon}_j \rangle = \sum_k T_{ik} S_{kj}$.*

The correspondence between a Hilbert space $\mathcal{H}$ with ortho-normal basis and the space of sequences ℓ^2 outlined above is the essence of how quantum mechanics in the operator formulation relates to matrix mechanics.

2.3 Linear operators on Hilbert spaces

We have already made a number of statements about linear operators on finite and infinite-dimensional inner product spaces but there are some features which are significantly more difficult in the infinite-dimensional case. We are not able to develop the full theory - this requires a lecture on functional analysis and is well beyond our present scope - but we will try to clear up some simple issues and at least cite a few relevant results. We begin with a discussion of the adjoint operator.

2.3.1 The adjoint operator

In Def. 1.10 we have defined the adjoint of a linear operator. We know that the adjoint, if it exists, is unique and we also know that it definitely exists in the finite-dimensional case. For the infinite-dimensional case we have, so far, made no statement about the existence of the adjoint but for Hilbert spaces we have the following

Theorem 2.8. *Let $\mathcal{H}$ be a Hilbert space and $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. Then there exists a unique linear and bounded operator $T^\dagger : \mathcal{H} \to \mathcal{H}$ with*

$$\langle T^\dagger \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, T\mathbf{w} \rangle \ . \tag{2.27}$$

for all $\mathbf{v}, \mathbf{w} \in \mathcal{H}$. The operator $T^\dagger$ is called the adjoint of T.

Proof. For any $\mathbf{v} \in \mathcal{H}$, define a linear functional $\varphi \in \mathcal{H}^*$ by $\varphi(\mathbf{w}) := \langle \mathbf{v}, T\mathbf{w} \rangle$. From the Cauchy-Schwarz inequality we have

$$|\varphi(\mathbf{w})| = |\langle \mathbf{v}, T\mathbf{w} \rangle| \leq \| \mathbf{v} \| \, \| T\mathbf{w} \| \leq \| \mathbf{v} \| \, \| T \| \, \| \mathbf{w} \| \ , \tag{2.28}$$

so that the functional φ is bounded. Hence, from Riesz's theorem 2.6, there exist a $\mathbf{u} \in \mathcal{H}$ such that $\varphi(\mathbf{w}) = \langle \mathbf{u}, \mathbf{w} \rangle$. If we define $T^\dagger$ by $T^\dagger \mathbf{v} = \mathbf{u}$ it follows that

$$\langle \mathbf{v}, T\mathbf{w} \rangle = \varphi(\mathbf{w}) = \langle \mathbf{u}, \mathbf{v} \rangle = \langle T^\dagger \mathbf{v}, \mathbf{w} \rangle \ , \tag{2.29}$$

so that the so-defined $T^\dagger$ does indeed have the required property for the adjoint. By tracing through the above construction it is easy to show that $T^\dagger$ is linear. We still need to show that it is bounded.

$$\| T^\dagger \mathbf{v} \|^2 = \langle T^\dagger \mathbf{v}, T^\dagger \mathbf{v} \rangle = \langle \mathbf{v}, T T^\dagger \mathbf{v} \rangle \leq \| \mathbf{v} \| \, \| T T^\dagger \mathbf{v} \| \leq \| \mathbf{v} \| \, \| T \| \, \| T^\dagger \mathbf{v} \| \ . \tag{2.30}$$

If $\| T^\dagger \mathbf{v} \| = 0$ the boundedness condition is trivially satisfied. Otherwise, we can divide by $\| T^\dagger \mathbf{v} \|$ and obtain $\| T^\dagger \mathbf{v} \| \leq \| \mathbf{v} \| \| T \|$. □

Also recall that a linear operator $T : \mathcal{H} \to \mathcal{H}$ is called self-adjoint or hermitian iff

$$\langle \mathbf{v}, T\mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{w} \rangle \ , \tag{2.31}$$

for all $\mathbf{v}, \mathbf{w} \in \mathcal{H}$, that is, if it can be moved from one argument of the scalar product to the other. Comparison with Theorem 2.8 shows that T is self-adjoint iff $T = T^\dagger$.

Application 2.6. Heisenberg uncertainty relation

Suppose we have two hermitian operators $Q, P : \mathcal{H} \to \mathcal{H}$ which satisfy [a] $[Q, P] = i$. The expectation values of these operators for a state $|\psi\rangle \in \mathcal{H}$ with $\langle \psi | \psi \rangle = 1$ are defined by

$$\bar{Q} := \langle \psi | Q | \psi \rangle \ , \qquad \bar{P} := \langle \psi | P | \psi \rangle \ .$$

and we can introduce the "shifted"" operators $q := Q - \bar{Q}$ and $p = P - \bar{P}$ which also satisfy $[q, p] = i$. The variances of Q and P are defined by

$$\Delta Q^2 := \langle \psi | q^2 | \psi \rangle = \| \, q\psi \, \|^2 \ , \qquad \Delta P^2 := \langle \psi | p^2 | \psi \rangle = \| \, p\psi \, \|^2$$

It follows from the Cauchy-Schwarz inequality that

$$\begin{aligned} \Delta Q \, \Delta P &= \| \, q\psi \, \| \| \, p\psi \, \| \geq |\langle q\psi | p\psi \rangle| = |\langle \psi | qp | \psi \rangle| \\ \Delta Q \, \Delta P &= \| \, p\psi \, \| \| \, q\psi \, \| \geq |\langle p\psi | q\psi \rangle| = |\langle \psi | pq | \psi \rangle| \ . \end{aligned}$$

Adding these two inequalities, we have

$$2 \, \Delta Q \, \Delta P \geq |\langle \psi | qp | \psi \rangle| + |\langle \psi | pq | \psi \rangle| \geq ||\langle \psi | qp | \psi \rangle| - |\langle \psi | pq | \psi \rangle|| = |\langle \psi | [q, p] | \psi \rangle| = 1 \ .$$

and, hence,

$$\Delta Q \, \Delta P \geq \frac{1}{2} \ , \tag{2.32}$$

which is, of course, Heisenberg's uncertainty relation. The lesson from this derivation is that the uncertainty relation is really quite general. All it requires is two hermitian operators Q and P with commutator $[Q, P] = i$ and then it follows more or less directly from the Cauchy-Schwarz inequality.

[a] For simplicity we set $\hbar$ to one.

2.3.2 Eigenvectors and eigenvalues

The theory of eigenvalues and eigenvectors ("spectral theory") in the infinite-dimensional case is significantly more complicated than for finite-dimensional vector spaces. Developing this theory systematically can easily take up a significant part of a lecture course. Here, I would just like to collect a few basic ideas and cite some of the important results.

Throughout, we will focus on self-adjoint operators $T : \mathcal{H} \to \mathcal{H}$ on a complex Hilbert space $\mathcal{H}$ which, in addition, are *compact*, a property which makes the spectral theory particularly well-behaved and is defined as follows.

Definition 2.4. *A linear operator $T : \mathcal{H} \to \mathcal{H}$ on a Hilbert space $\mathcal{H}$ is called compact iff, for every bounded sequence $(\mathbf{v}_k)$ in $\mathcal{H}$, the sequence $(T\mathbf{v}_k)$ contains a convergent sub-sequence.*

We note that a compact operator is bounded. If it was not bounded, there would be a sequence $(\mathbf{v}_k)$ with $\| \mathbf{v}_k \| = 1$ and $\| T\mathbf{v}_k \| > k$, and, in this case the sequence $(T\mathbf{v}_k)$ cannot have a convergent sub-sequence, contradicting compactness. This means, from Theorem 2.8 that compact operators always have an adjoint.

In fact, we will be focusing on self-adjoint, compact operators and their eigenvalues and eigenvectors have the following properties.

Theorem 2.9. *Let $T : \mathcal{H} \to \mathcal{H}$ be a compact, self-adjoint operator on a (separable) Hilbert space $\mathcal{H}$. Then we have the following statements:*
(i) The eigenvalues of T are real and eigenvectors for different eigenvalues are orthogonal.
(ii) The set of non-zero eigenvalues of T is either finite or it is a sequence which tends to zero.
(iii) Each non-zero eigenvalue has a finite degeneracy.
(iv) There is an ortho-normal system $(\boldsymbol{\epsilon}_k)$ of eigenvectors with non-zero eigenvalues λ_k which forms a basis on $\overline{\mathrm{Im}(T)}$. Further, we have $\overline{\mathrm{Im}(T)}^{\perp} = \mathrm{Ker}(T)$ and $\mathcal{H} = \overline{\mathrm{Im}(T)} \oplus \mathrm{Ker}(T)$.
(v) The operator T has the representation

$$T\mathbf{v} = \sum_k \lambda_k \langle \boldsymbol{\epsilon}_k, \mathbf{v} \rangle \boldsymbol{\epsilon}_k \ . \tag{2.33}$$

Proof. We have already shown (i). For the proof of (ii) and (iii) see, for example, Ref. [5]. The ortho-normal system $(\boldsymbol{\epsilon}_k)$ in (iv) is constructed by applying the Gram-Schmidt procedure to each eigenspace $\mathrm{Ker}(T - \lambda \mathrm{id})$ with $\lambda \neq 0$ which, from (iii) is finite-dimensional. The proof that the vectors $(\boldsymbol{\epsilon}_k)$ form an ortho-normal basis of $\overline{\mathrm{Im}(T)}$ can be found in Ref. [5].

To show the formula in (v) we set $W = \overline{\mathrm{Im}(T)}$ and note, from Theorem 2.2, that $\mathcal{H} = W \oplus W^{\perp}$ so that every $\mathbf{v} \in \mathcal{H}$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{u}$, where $\mathbf{w} \in W$ and $\mathbf{u} \in W^{\perp}$. Since $\mathbf{u} \in W^{\perp}$ we have $0 = \langle T\mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{x}, T\mathbf{u} \rangle$ for all $\mathbf{x} \in \mathcal{H}$ and this means that $T\mathbf{u} = \mathbf{0}$. (Hence, $W^{\perp} \subset \mathrm{Ker}(T)$ and the reverse inclusion is also easy to show so $W^{\perp} = \mathrm{Ker}(T)$.) The other component, $\mathbf{w} \in W$, can be written as

$$\mathbf{w} = \sum_k \langle \boldsymbol{\epsilon}_k, \mathbf{w} \rangle \boldsymbol{\epsilon}_k \tag{2.34}$$

since the $(\boldsymbol{\epsilon}_k)$ form a basis of W. Putting it all together gives

$$T\mathbf{v} = T\mathbf{w} = T \sum_k \langle \boldsymbol{\epsilon}_k, \mathbf{w} \rangle \boldsymbol{\epsilon}_k = \sum_k \langle \boldsymbol{\epsilon}_k, \mathbf{w} \rangle T\boldsymbol{\epsilon}_k = \sum_k \lambda_k \langle \boldsymbol{\epsilon}_k, \mathbf{w} \rangle \boldsymbol{\epsilon}_k \ . \tag{2.35}$$

□

The ortho-normal system $(\boldsymbol{\epsilon}_k)$ from the above theorem provides a basis for $\overline{\mathrm{Im}(T)}$ but not necessarily for all of $\mathcal{H}$. This is because the vectors $\boldsymbol{\epsilon}_k$ correspond to non-zero eigenvalues and we are missing the eigenvectors with zero eigenvalue, that is, the kernel of T. Fortunately, from part (iv) of Theorem 2.9 we have the decomposition

$$\mathcal{H} = \overline{\mathrm{Im}(T)} \oplus \mathrm{Ker}(T) \ , \qquad \overline{\mathrm{Im}(T)}^{\perp} = \mathrm{Ker}(T) \ , \tag{2.36}$$

so we can complete $(\boldsymbol{\epsilon}_k)$ to a basis of $\mathcal{H}$ by adding a basis for $\mathrm{Ker}(T)$. In conclusion, for a compact, self-adjoint operator on a Hilbert space, we can always find an ortho-normal basis of the Hilbert space consisting of eigenvectors of the operator. In Dirac notation and dropping the argument $\mathbf{v}$, Eq. (2.33) can be written as

$$T = \sum_k \lambda_k |\boldsymbol{\epsilon}_k\rangle\langle\boldsymbol{\epsilon}_k| \ , \tag{2.37}$$

where $T|\boldsymbol{\epsilon}_k\rangle = \lambda_k |\boldsymbol{\epsilon}_k\rangle$. This is the generalisation of the finite-dimensional result (1.46).

2.3.3 The Fredholm alternative

Suppose, for a compact, self-adjoint operator $T : \mathcal{H} \to \mathcal{H}$, a given $\mathbf{u} \in \mathcal{H}$ and a constant $\lambda \neq 0$, we would like to solve the equations

$$(T - \lambda\,\mathrm{id})\mathbf{v} = \mathbf{0}\ , \qquad (T - \lambda\,\mathrm{id})\mathbf{v} = \mathbf{u}\ . \tag{2.38}$$

It turns out that many of the differential equations we will consider later can be cast in this form. The right Eq. (2.38) is an inhomogeneous linear equation and the equation on the left-hand side is its homogeneous counterpart. Clearly, we have

$$\{\text{solutions of homogeneous equation}\} = \mathrm{Ker}(T - \lambda\,\mathrm{id})\ . \tag{2.39}$$

We also know, if a solution $\mathbf{v}_0$ of the inhomogeneous equation exists, then its general solution is given by $\mathbf{v}_0 + \mathrm{Ker}(T - \lambda\,\mathrm{id})$. There are two obvious cases we should distinguish.

(a) The number λ does not equal any of the eigenvalues of T. In this case $\mathrm{Ker}(T - \lambda\,\mathrm{id}) = \{\mathbf{0}\}$ so that the homogeneous equation in (2.38) only has the trivial solution.

(b) The number λ does equal one of the eigenvalues of T so that $\mathrm{Ker}(T-\lambda\,\mathrm{id}) \neq \{\mathbf{0}\}$ and the homogenous equation in Eq. (2.38) does have non-trivial solutions.

The above case distinction is called the *Fredholm alternative*. Of course we would like to discuss the solutions to the inhomogeneous equation in either case. The obvious way to proceed is to start with an ortho-normal basis $(\boldsymbol{\epsilon}_k)$ of eigenvectors of T with corresponding eigenvalues λ_k, so that $T\mathbf{v}_k = \lambda_k\mathbf{v}_k$ (Here, we include the eigenvectors with eigenvalue zero.), expand $\mathbf{v}$ and $\mathbf{u}$ in terms of this basis

$$\mathbf{v} = \sum_k \langle\boldsymbol{\epsilon}_k, \mathbf{v}\rangle\boldsymbol{\epsilon}_k\ , \qquad \mathbf{u} = \sum_k \langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle\boldsymbol{\epsilon}_k\ , \tag{2.40}$$

and use the representation (2.33) of the operator T. Inserting all this into the inhomogeneous Eq. (2.38) gives

$$(T - \lambda\,\mathrm{id})\mathbf{v} = \sum_i (\lambda_i - \lambda)\langle\boldsymbol{\epsilon}_i, \mathbf{v}\rangle\boldsymbol{\epsilon}_i \stackrel{!}{=} \sum_i \langle\boldsymbol{\epsilon}_i, \mathbf{u}\rangle\boldsymbol{\epsilon}_i\ . \tag{2.41}$$

Taking the inner product of this equation with a basis vector $\boldsymbol{\epsilon}_k$ leads to

$$(\lambda_k - \lambda)\langle\boldsymbol{\epsilon}_k, \mathbf{v}\rangle = \langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle\ , \tag{2.42}$$

for all k. Now let us consider the two cases above. In case (a), λ does not equal any of the eigenvalues λ_k and we can simply solve Eq. (2.42) for all k to obtain

$$\langle\boldsymbol{\epsilon}_k, \mathbf{v}\rangle = \frac{\langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle}{\lambda_k - \lambda} \quad \Rightarrow \quad \mathbf{v} = \sum_k \frac{\langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle}{\lambda_k - \lambda}\boldsymbol{\epsilon}_k\ . \tag{2.43}$$

This result means that in case (a) we have a unique solution $\mathbf{v}$, as given by the above equation, for any inhomogeneity $\mathbf{u}$. The situation is more complicated in case (b). In this case λ equals λ_k for some k and for such cases the LHS of Eq. (2.42) vanishes. This means in this case we have a solution if and only if

$$\langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle = 0 \quad \text{for all } k \text{ with } \lambda_k = \lambda\ . \tag{2.44}$$

Another way of stating this condition is to say that we need the inhomogeneity $\mathbf{u}$ to be perpendicular to $\mathrm{Ker}(T - \lambda\,\mathrm{id})$. If this condition is satisfied, the solution can be written as

$$\mathbf{v} = \sum_{k:\lambda_k\neq\lambda} \frac{\langle\boldsymbol{\epsilon}_k, \mathbf{u}\rangle}{\lambda_k - \lambda}\boldsymbol{\epsilon}_k + \sum_{k:\lambda_k=\lambda} \alpha_k\boldsymbol{\epsilon}_k \tag{2.45}$$

where the α_k are arbitrary numbers and the second sum in this expression of course represents a general element of $\mathrm{Ker}(T - \lambda\,\mathrm{id})$. This discussion can be summarised in the following

Theorem 2.10. *(Fredholm alternative) Let $T : \mathcal{H} \to \mathcal{H}$ be a compact, self-adjoint operator on a Hilbert space $\mathcal{H}$ with a basis of eigenvectors $\boldsymbol{\epsilon}_k$ with associated eigenvalues λ_k, $\mathbf{u} \in \mathcal{H}$ and $\lambda \neq 0$. For the solution to the equation*

$$(T - \lambda \,\mathrm{id})\mathbf{v} = \mathbf{u} \tag{2.46}$$

the following alternative holds:

(a) The number λ is different from all eigenvalues λ_k. Then the equation (2.46) has a unique solution for all $\mathbf{u} \in \mathcal{H}$ given by

$$\mathbf{v} = \sum_k \frac{\langle \boldsymbol{\epsilon}_k, \mathbf{u} \rangle}{\lambda_k - \lambda} \boldsymbol{\epsilon}_k \; . \tag{2.47}$$

(b) The number λ equals one of the eigenvalues. In this case, a solution exists if and only if

$$\langle \boldsymbol{\epsilon}_k, \mathbf{u} \rangle = 0 \quad \textit{for all } k \textit{ with } \lambda_k = \lambda \; . \tag{2.48}$$

If this condition is satisfied the general solution is given by

$$\mathbf{v} \in \sum_{k:\lambda_k \neq \lambda} \frac{\langle \boldsymbol{\epsilon}_k, \mathbf{u} \rangle}{\lambda_k - \lambda} \boldsymbol{\epsilon}_k + \sum_{k:\lambda_k = \lambda} \alpha_k \boldsymbol{\epsilon}_k \; , \tag{2.49}$$

where the $\alpha_k \in \mathbb{C}$ are arbitrary.

3 Fourier analysis

The Fouries series and the Fourier transform are important mathematical tools in practically all parts of physics. Intuitively, they allow us to decompose functions into their various frequency components. The Fouries series, which we discuss first, deals with functions on finite intervals and leads to a decomposition in terms of a discrete spectrum of frequencies. Mathematically speaking, we find an ortho-normal basis of functions with well-defined frequencies on the Hilbert space $L^2([-\pi, \pi])$ (say) and the coordinates relative to this basis represent the strength of the various frequencies. The Fourier transform applies to functions defined on the entire real line (or on $\mathbb{R}^n$) and leads to a decomposition with a continuous spectrum of frequencies. Mathematically, the Fourier transform can be understood as a unitary map on the Hilbert space $L^2(\mathbb{R}^n)$.

3.1 Fourier series

3.1.1 Cosine Fourier series

We begin by considering the Hilbert space $L^2_{\mathbb{R}}([0, \pi])$ of (real-valued) square integrable functions on the interval $[0, \pi]$ and recall that the scalar product on this space is defined by

$$\langle f, g \rangle = \int_0^\pi dx\, f(x) g(x) \,. \tag{3.1}$$

It is easy to verify that the functions

$$\tilde{c}_0 = \frac{1}{\sqrt{\pi}} \,, \qquad \tilde{c}_k := \sqrt{\frac{2}{\pi}} \cos(kx) \,, \quad k = 1, 2, \ldots \tag{3.2}$$

form an ortho-normal system on $L^2_{\mathbb{R}}([0, \pi])$.

Exercise 3.1. *Show that the functions in Eq. (3.7) form an ortho-normal system on $L^2_{\mathbb{R}}([0, \pi])$ by explicitly evaluating the scalar products from Eq. (3.1).*

In fact, we have a significantly stronger statement.

Theorem 3.2. *The functions $(\tilde{c}_k)_{k=0}^\infty$ in Eq. (3.2) form an ortho-normal basis of $L^2_{\mathbb{R}}([0, \pi])$. Hence, the space $L^2_{\mathbb{R}}([0, \pi])$ is separable.*

Proof. These proofs are typically highly technical but in the present case we get away with appealing to previous theorems. First, it is clear that $(\tilde{c}_k)_{k=0}^\infty$ forms on ortho-normal system. To show that it is an ortho-normal basis we need to verify one of the conditions in Theorem 2.3 and we opt for (iv). Hence, we need to show that $\overline{\text{Span}}(\tilde{\mathbf{c}}_k)_{k=0}^\infty = L^2_{\mathbb{R}}([0, \pi])$, or, in other words, that every function $f \in L^2_{\mathbb{R}}([0, \pi])$ can be approximated, to arbitrary accuracy, by linear combinations of the $\tilde{\mathbf{c}}_k$. From Theorem 2.1 we know that every function $f \in L^2_{\mathbb{R}}([0, \pi])$ can be approximated, to arbitrary accuracy, by continuous functions $g \in \mathcal{C}_{\mathbb{R}}([0, \pi])$. It is, therefore, enough to show that every continuous g can be approximated by linear combinations of $\tilde{\mathbf{c}}_k$. Define the function $h(y) := g(\cos^{-1} y)$ which is in $\mathcal{C}_{\mathbb{R}}([-1, 1])$ and, from the Stone-Weierstrass theorem 1.35, can be approximated by polynomials p, so that $|h(y) - p(y)| < \epsilon$. By setting $x = \cos^{-1} y$ this implies that $|g(x) - p(\cos x)| < \epsilon$. But the function $p(\cos x)$, where p is a polynomial, can always be written as $\sum_{k=0}^n \alpha_k \cos(kx)$, using trigonometric identities. This completes the proof. □

Of course we are not stuck with the specific interval $[0, \pi]$ but a simple re-scaling $x \to \pi x/a$ for $a > 0$ shows that the Hilbert space $L^2_{\mathbb{R}}([0, a])$ with scalar product

$$\langle f, g \rangle = \int_0^a dx\, f(x) g(x) \,, \tag{3.3}$$

has an ortho-normal basis

$$\tilde{c}_0 = \frac{1}{\sqrt{a}} , \qquad \tilde{c}_k := \sqrt{\frac{2}{a}} \cos\left(\frac{k\pi x}{a}\right) , \quad k = 1, 2, \dots . \tag{3.4}$$

Let us be more explicit about what this actually means. From part (i) of Theorem 2.3 we conclude that every (real-valued) square integrable function $f \in L^2_{\mathbb{R}}([0, a])$ can be written as

$$f(x) = \sum_{k=0}^{\infty} \alpha_k \, \tilde{c}_k(x) = \frac{\alpha_0}{\sqrt{a}} + \sqrt{\frac{2}{a}} \sum_{k=1}^{\infty} \alpha_k \cos\left(\frac{k\pi x}{a}\right) \tag{3.5}$$

where

$$\alpha_0 = \langle \tilde{c}_0, f \rangle = \frac{1}{\sqrt{a}} \int_0^a dx \, f(x) , \qquad \alpha_k = \langle \tilde{c}_k, f \rangle = \sqrt{\frac{2}{a}} \int_0^a dx \, \cos\left(\frac{k\pi x}{a}\right) f(x) , \quad k = 1, 2, \dots . \tag{3.6}$$

It is customary to introduce the coefficients $a_0 = 2\frac{\alpha_0}{\sqrt{a}}$ and $a_k = \sqrt{\frac{2}{a}}\alpha_k$, for $k = 1, 2, \dots$ in order to re-distribute factors:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \, \cos\left(\frac{k\pi x}{a}\right) \quad \text{where} \quad a_k = \frac{2}{a} \int_0^a dx \, \cos\left(\frac{k\pi x}{a}\right) f(x) . \tag{3.7}$$

This series is called the *cosine Fourier series* and the a_k are called the (cosine) *Fourier coefficients*. It is important to remember that the equality in the first Eq. (3.7) holds in $L^2_{\mathbb{R}}([0, a])$, a space which consists of classes of functions which have been identified if they differ only on sets of Lebesgue-measure zero. This means that the function f and its Fourier series do not actually have to coincide at every point $x \in [0, a]$ - they can differ on a space of measure zero. However, we know that the (cosine) Fourier series always converges to the function f in the norm on $L^2_{\mathbb{R}}([0, a])$.

We know from part (ii) of Theorem 2.3 that the norm of f can be calculated in terms of its Fourier coefficients as

$$\frac{2}{a} \int_0^a dx \, |f(x)|^2 = \frac{2}{a} \| \, f \, \|^2 = \frac{2}{a} \sum_{k=0}^{\infty} |\langle \tilde{c}_k, f \rangle|^2 = \frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} |a_k|^2 . \tag{3.8}$$

This result is also known as *Parseval equation*.

3.1.2 Sine Fourier series

Unsurprisingly, the above discussion can be repeated for sine functions. As before, we consider the Hilbert space $L^2_{\mathbb{R}}([0, \pi])$ with scalar product (3.1). On this space the functions

$$\tilde{s}_k := \sqrt{\frac{2}{\pi}} \sin(kx) , \qquad k = 1, 2, \dots \tag{3.9}$$

form an ortho-normal system and indeed an ortho-normal basis as stated in the following

Theorem 3.3. *The functions $(\tilde{s}_k)_{k=1}^{\infty}$ in Eq. (3.11) form an ortho-normal basis of $L^2_{\mathbb{R}}([0, \pi])$.*

Proof. This proof is very similar to the one for Theorem 3.2 and can, for example, be found in Ref. [5]. □

As in the cosine case, we can re-scale by $x \to \pi x/a$ to the interval $[0, a]$ and obtain an ortho-normal basis

$$\tilde{s}_k = \sqrt{\frac{2}{a}} \sin\left(\frac{k\pi x}{a}\right) , \quad k = 1, 2, \dots \tag{3.10}$$

for $L^2_{\mathbb{R}}([0, a])$ with scalar product (3.3). Hence, every function $f \in L^2_{\mathbb{R}}([0, a])$ can be expanded as

$$f(x) = \sum_{k=1}^{\infty} \beta_k \, \tilde{s}_k(x) = \sqrt{\frac{2}{a}} \sum_{k=1}^{\infty} \beta_k \sin\left(\frac{k\pi x}{a}\right) \quad \text{where} \quad \beta_k = \langle \tilde{s}_k, f \rangle = \sqrt{\frac{2}{a}} \int_0^a dx \, \sin\left(\frac{k\pi x}{a}\right) f(x) \, .$$

Introducing to the coefficients $b_k = \sqrt{\frac{2}{a}}\beta_k$ this can be re-written in the standard notation

$$f(x) = \sum_{k=1}^{\infty} b_k \, \sin\left(\frac{k\pi x}{a}\right) \quad \text{where} \quad b_k = \frac{2}{a} \int_0^a dx \, \sin\left(\frac{k\pi x}{a}\right) f(x) \, . \tag{3.11}$$

This series is called the *sine Fourier series* and the b_k are called the (sine) *Fourier coefficients*. Of course there is also a version of Parseval's equation which reads

$$\frac{2}{a} \int_0^a dx \, |f(x)|^2 = \frac{2}{a} \| \, f \, \|^2 = \frac{2}{a} \sum_{k=1}^{\infty} |\langle \tilde{s}_k, f \rangle|^2 = \sum_{k=1}^{\infty} |b_k|^2 \, . \tag{3.12}$$

3.1.3 Real standard Fourier series

The most commonly used form of the Fourier series uses sine and cosine functions and can be thought of as a combination of the above two cases. The Hilbert space considered in this case is $L^2_{\mathbb{R}}([-\pi, \pi])$ with scalar product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} dx \, f(x) g(x) \, . \tag{3.13}$$

We can also think of the functions in this Hilbert space as the periodic functions with period 2π, so $f(x) = f(x + 2\pi)$ (which are square-integrable over one period). The functions

$$c_0 := \frac{1}{\sqrt{2\pi}} , \quad c_k := \frac{1}{\sqrt{\pi}} \cos(kx) , \quad s_k := \frac{1}{\sqrt{\pi}} \sin(kx) , \quad k = 1, 2, \dots , \tag{3.14}$$

form an ortho-normal system on $L^2_{\mathbb{R}}([-\pi, \pi])$.

Exercise 3.4. *Check that the functions* (3.14) *form an ortho-normal system on* $L^2_{\mathbb{R}}([-\pi, \pi])$.

They also form an ortho-normal basis as the following theorem asserts.

Theorem 3.5. *The functions* $(c_k)_{k=0}^{\infty}$ *and* $(s_k)_{k=1}^{\infty}$ *together form an ortho-normal basis on* $L^2_{\mathbb{R}}([-\pi, \pi])$.

Proof. Every function $f \in L^2_{\mathbb{R}}([-\pi, \pi])$ can be written as $f = f_+ + f_-$, where $f_\pm(x) = \frac{1}{2}(f(x) \pm f(-x))$ are the symmetric and anti-symmetric parts of f. The functions $f_\pm$ can be restricted to the interval $[0, \pi]$ so that they can be viewed as elements of $L^2_{\mathbb{R}}([0, \pi])$. From Theorem 3.2 we can write down a cosine Fourier series for f_+ and from Theorem 3.3 f_- has a sine Fourier series, so

$$f_+ = \sum_{k=0}^{\infty} \alpha_k \, \tilde{c}_k \, , \qquad f_- = \sum_{k=1}^{\infty} \beta_k \, \tilde{s}_k \, . \tag{3.15}$$

Since both sides of the first equation are symmetric and both sides of the second equation anti-symmetric they can both be trivially extended back to the interval $[-\pi, \pi]$. Then, summing up

$$f = f_+ + f_- = \sum_{k=0}^{\infty} \alpha_k \, \tilde{c}_k + \sum_{k=1}^{\infty} \beta_k \, \tilde{s}_k \tag{3.16}$$

proves the statement. □

This proof also points to an interpretation of the relation between the sine and cosine Fourier series on the one hand and the standard Fourier series on the other hand. Starting with a function $f \in L^2_{\mathbb{R}}([0, \pi])$ we can write down the cosine Fourier series. But we can also extend f to a symmetric function on $[-\pi, \pi]$ so it becomes an element of $L^2_{\mathbb{R}}([-\pi, \pi])$ and we can write down a standard Fourier series. Of course, f being symmetric, this Fourier series only contains cosine terms and it looks formally the same as the cosine Fourier series but is valid on the larger interval $[-\pi, \pi]$. Similarly, for $f \in L^2_{\mathbb{R}}([0, \pi])$ we can write down the sine Fourier series and extend f to an anti-symmetric function on the interval $[-\pi, \pi]$. The standard Fourier series for this anti-symmetric function then only contains sine terms and formally coincides with the sine Fourier series. Conversely, if we start with an even (odd) function $f \in L^2_{\mathbb{R}}([-\pi, \pi])$ then the Fourier series only contains cosine (sine) terms and we can restrict the expansion to the interval $[0, \pi]$ so it becomes a cosine (sine) Fourier series.

As before, we can use the re-scaling $x \to \pi x/a$ to transform to the interval $[-a, a]$ and obtain the ortho-normal basis

$$c_0 := \frac{1}{\sqrt{2a}} \, , \quad c_k := \frac{1}{\sqrt{a}} \cos\left(\frac{k\pi x}{a}\right) \, , \quad s_k := \frac{1}{\sqrt{a}} \sin\left(\frac{k\pi x}{a}\right) \, , \quad k = 1, 2, \dots \, , \tag{3.17}$$

for the Hilbert space $L^2_{\mathbb{R}}([-a, a])$ with scalar product

$$\langle f, g \rangle = \int_{-a}^{a} dx \, f(x) g(x) \, . \tag{3.18}$$

Let us collect the formulae for the standard Fourier series. Every function $f \in L^2_{\mathbb{R}}([-a, a])$ can be expanded as

$$f(x) = \sum_{k=0}^{\infty} \alpha_k c_k(x) + \sum_{k=1}^{\infty} \beta_k \, s_k(x) = \frac{\alpha_0}{\sqrt{2a}} + \frac{1}{\sqrt{a}} \sum_{k=1}^{\infty} \alpha_k \, \cos\left(\frac{k\pi x}{a}\right) + \frac{1}{\sqrt{a}} \sum_{k=1}^{\infty} \beta_k \sin\left(\frac{k\pi x}{a}\right) \tag{3.19}$$

where

$$\begin{array}{rcl} \alpha_0 & = & \langle c_0, f \rangle = \frac{1}{\sqrt{2a}} \int_{-a}^{a} dx \, f(x) \\ \alpha_k & = & \langle c_k, f \rangle = \frac{1}{\sqrt{a}} \int_{-a}^{a} dx \, \cos\left(\frac{k\pi x}{a}\right) f(x) \, , \quad k = 1, 2, \dots \\ \beta_k & = & \langle c_k, f \rangle = \frac{1}{\sqrt{a}} \int_{-a}^{a} dx \, \sin\left(\frac{k\pi x}{a}\right) f(x) \, , \quad k = 1, 2, \dots \, . \end{array} \tag{3.20}$$

As before, we introduce the re-scaled coefficients $a_0 = \sqrt{2/a} \, \alpha_0$, $a_k = \alpha_k/\sqrt{a}$ and $b_k = \beta_k/\sqrt{a}$, where $k = 1, 2, \dots$ to obtain these equations in the standard form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{k\pi x}{a}\right) + b_k \sin\left(\frac{k\pi x}{a}\right) \right) \, , \tag{3.21}$$

where

$$a_0 = \frac{1}{a} \int_{-a}^{a} dx \, f(x) \, , \quad a_k = \frac{1}{a} \int_{-a}^{a} dx \, \cos\left(\frac{k\pi x}{a}\right) f(x) \, , \quad b_k = \frac{1}{a} \int_{-a}^{a} dx \, \sin\left(\frac{k\pi x}{a}\right) f(x) \tag{3.22}$$

for $k = 1, 2, \dots$. Parseval's equation now takes the form

$$\frac{1}{a} \int_{-a}^{a} dx \, |f(x)|^2 = \frac{1}{a} \| \, f \, \|^2 = \frac{1}{a} \left(\sum_{k=0}^{\infty} |\langle c_k, f \rangle|^2 + \sum_{k=1}^{\infty} |\langle s_k, f \rangle|^2 \right) = \frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} (|a_k|^2 + |b_k|^2) \, . \tag{3.23}$$

3.1.4 Complex standard Fourier series

By far the most elegant form of the Fourier series arises in the complex case, where we consider the Hilbert space $L^2_{\mathbb{C}}([-\pi,\pi])$ with scalar product

$$\langle f,g\rangle = \int_{-\pi}^{\pi} dx\, f(x)^* g(x)\ . \tag{3.24}$$

The functions

$$e_k := \frac{1}{\sqrt{2\pi}} \exp(ikx)\ , \quad k\in\mathbb{Z} \tag{3.25}$$

form an ortho-normal system as verified in the following exercise.

Exercise 3.6. *Show that the functions $(e_k)_{k=-\infty}^{\infty}$ in Eq. (3.25) form an ortho-normal system on $L^2_{\mathbb{C}}([-\pi,\pi])$ with scalar product (3.24).*

The above functions form, in fact, an ortho-normal basis as stated in

Theorem 3.7. *The functions $(e_k)_{k=-\infty}^{\infty}$ in Eq. (3.25) form an ortho-normal basis of $L^2_{\mathbb{C}}([-\pi,\pi])$.*

Proof. Start with a function $f\in L^2_{\mathbb{C}}([-\pi,\pi])$ and decompose this function into real and imaginary parts, so write $f = f_R + if_I$. Since

$$\infty > \parallel f \parallel^2 = \int_{-\pi}^{\pi} dx|f(x)|^2 = \int_{-\pi}^{\pi} dx\, f_R^2 + \int_{-\pi}^{\pi} dx\, f_I^2 \tag{3.26}$$

both f_R and f_I are real-valued square integrable functions and are, hence, elements of $L^2_{\mathbb{R}}([-\pi,\pi])$. This means, from Theorem 3.5 that we can write down a standard real Fourier series for f_R and f_I. Inserting these two real Fourier series into $f = f_R + if_I$ and replacing $\cos(kx) = (\exp(ikx)+\exp(-ikx))/2$, $\sin(kx) = (\exp(ikx)-\exp(-ikx))/(2i)$ proves the theorem. □

The usual re-scaling $x\to \pi x/a$ leads to the Hilbert space $L^2_{\mathbb{C}}([-a,a])$ with scalar product

$$\langle f,g\rangle = \int_{-a}^{a} dx\, f(x)^* g(x)\ . \tag{3.27}$$

and ortho-normal basis

$$e_k := \frac{1}{\sqrt{2a}} \exp\left(\frac{ik\pi x}{a}\right)\ , \quad k\in\mathbb{Z}\ . \tag{3.28}$$

Every function $f\in L^2_{\mathbb{C}}([-a,a])$ then has an expansion

$$f(x) = \sum_{k\in\mathbb{Z}} \alpha_k\, e_k(x) = \frac{1}{\sqrt{2a}} \sum_{k\in\mathbb{Z}} \alpha_k \exp\left(\frac{ik\pi x}{a}\right)\ , \tag{3.29}$$

where

$$\alpha_k = \langle e_k, f\rangle = \frac{1}{\sqrt{2a}} \int_{-a}^{a} dx\, \exp\left(\frac{-ik\pi x}{a}\right) f(x)\ . \tag{3.30}$$

With the re-scaled Fourier coefficients $a_k = \alpha_k/\sqrt{2a}$ this turns into the standard form

$$f(x) = \sum_{k\in\mathbb{Z}} a_k \exp\left(\frac{ik\pi x}{a}\right) \quad \text{where} \quad a_k = \frac{1}{2a}\int_{-a}^{a} dx\, \exp\left(\frac{-ik\pi x}{a}\right) f(x)\ . \tag{3.31}$$

Parseval's equation now reads

$$\frac{1}{2a}\int_{-a}^{a} dx\, |f(x)|^2 = \frac{1}{2a}\parallel f \parallel^2 = \frac{1}{2a}\sum_{k\in\mathbb{Z}} |\langle e_k, f\rangle|^2 = \sum_{k\in\mathbb{Z}} |a_k|^2\ . \tag{3.32}$$

3.1.5 Pointwise convergence

So far, our discussion of convergence for the Fourier series has been carried out with respect to the L^2 norm (3.18). As emphasised, this type of convergence ensures that the difference of a function and its Fourier series has a vanishing L^2 norm but it does not necessarily imply that the Fourier series converges to the function at every point x. The following theorem provides a statement about uniform convergence of a Fourier series.

Theorem 3.8. *Let $f : [-a, a] \to \mathbb{R}$ or $\mathbb{C}$ be a (real or complex valued) function which is piecewise continuously differentiable and which satisfies $f(-a) = f(a)$. Then the (real or complex) Fourier series of f converges to f uniformly.*

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

Recall from Def. 1.22 that uniform convergence implies point-wise convergence so under the conditions of Theorem 3.8 the Fourier series of f converges to f at every point $x \in [-a, a]$.

3.1.6 Examples of Fourier series

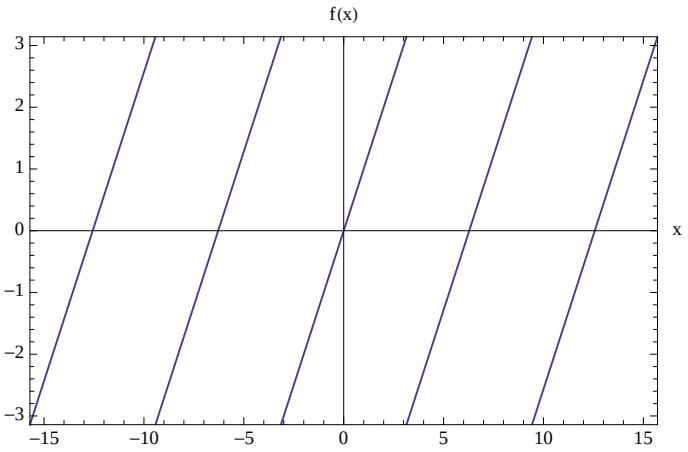

Figure 3: Graph of the periodic functions f defined by $f(x) = f(x + 2\pi)$ and $f(x) = x$ for $-\pi < x \leq \pi$.

Application 3.7. Linear function

Let us start with the function $f : [-\pi, \pi] \to \mathbb{R}$ defined by

$$f(x) = x \,, \tag{3.33}$$

so a simple linear function on the interval $[-\pi, \pi]$. Of course, we can extend this to a periodic function with period 2π whose graph is shown in Fig. 3. Since this function is anti-symmetric the Fourier series of course only contains sine terms. (Alternatively, and equivalently, we can consider this function restricted to the interval $[0, \pi]$ and compute its sine Fourier series.) Using Eqs. (3.22) we find for the Fourier coefficients

$$a_k = 0 \,, \quad k = 0, 1, 2, \dots \,, \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \, x \sin(kx) = \frac{2(-1)^{k+1}}{k} \,, \quad k = 1, 2, \dots \,. \tag{3.34}$$

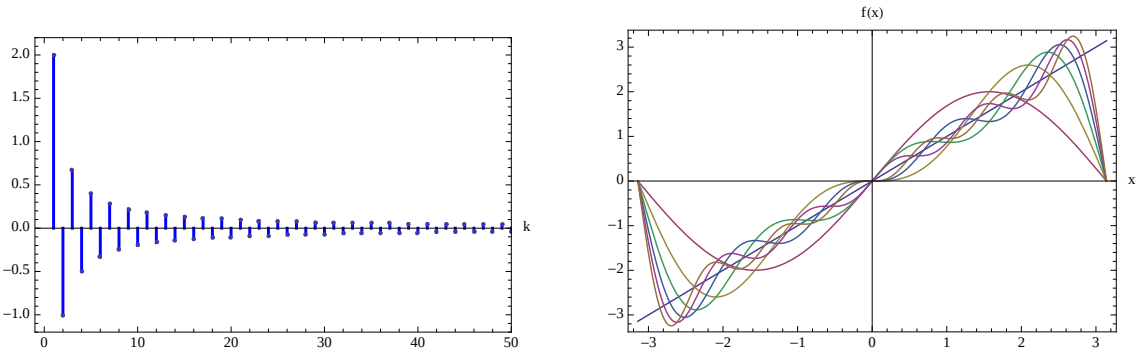

Figure 4: Fourier coefficients and Fourier series for the linear function f in Eq. (3.33). The left figure shows the Fourier coefficients a_k from Eq. (3.34) for $k = 1, \ldots, 50$. The function f together with the first six partial sums of its Fourier series (3.36) is shown in the right figure.

As a practical matter, it is useful to structure the calculation of Fourier coefficients in order to avoid mistakes. Creating pages and pages of integration performed in small steps is neither efficient nor likely to lead to correct answers. Instead, separate the process of integration from the specific calculation of Fourier coefficients. A particular Fourier calculation often involves certain types of standard integrals. In the above case, these are integrals of the form $\int dx\, x \sin(\alpha x)$ for a constant α. Find these integrals first (or simply look them up):

$$\int dx\, x \sin(\alpha x) = -\frac{x\cos(\alpha x)}{\alpha} + \frac{\sin(\alpha x)}{\alpha^2}\,. \tag{3.35}$$

Then apply this general result to the particular calculation at hand, that is, in the present case, set $\alpha = k$ and put in the integration limits.

Inserting the above Fourier coefficients into Eq. (3.21), we get the Fourier series

$$f(x) = 2\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sin(kx)\,. \tag{3.36}$$

Recall that the equality in Eq. (3.36) is not meant point-wise for every x but as an equality in $L^2_{\mathbb{R}}([-\pi,\pi])$, that is, the difference between f and its Fourier series has length zero with respect to the norm on $L^2_{\mathbb{R}}([-\pi,\pi])$. In fact, Eq. (3.36) shows (and Fig. 4 illustrates) that the Fourier series of f vanishes at $\pm\pi$ (since every term in the series (3.36) vanishes at $\pm\pi$) while $f(\pm\pi) = \pm\pi$ is non-zero. So we have an example where the Fourier series does not converge to the function at every point. In fact, the present function f violates the conditions of Theorem 3.8 (since $f(\pi) \neq f(-\pi)$), so there is no reason to expect point-wise convergence. It is clear from Fig. 4 that the Fourier series "struggles" to reproduce the function near $\pm\pi$ and this can be seen as the intuitive reason for the slow drop-off of the Fourier coefficients, $a_k \sim 1/k$, in Eq. (3.34). In other words, a larger number of terms in the Fourier series contribute significantly so that the function can be matched near the boundaries of the interval $[-\pi, \pi]$.

For this example, let us consider Parseval's equation (3.23)

$$\frac{2\pi^2}{3} = \frac{1}{\pi}\int_{-\pi}^{\pi} dx\, x^2 = \sum_{k=1}^{\infty} |b_k|^2 = 4\sum_{k=1}^{\infty} \frac{1}{k^2}\,, \tag{3.37}$$

where the left hand side follows from explicitly carrying out the normalisation integral and the right hand side by inserting the Fourier coefficients (3.34). This leads to the interesting formula

$$\frac{\pi^2}{6} = \sum_{k=1}^{\infty} \frac{1}{k^2}\,. \tag{3.38}$$

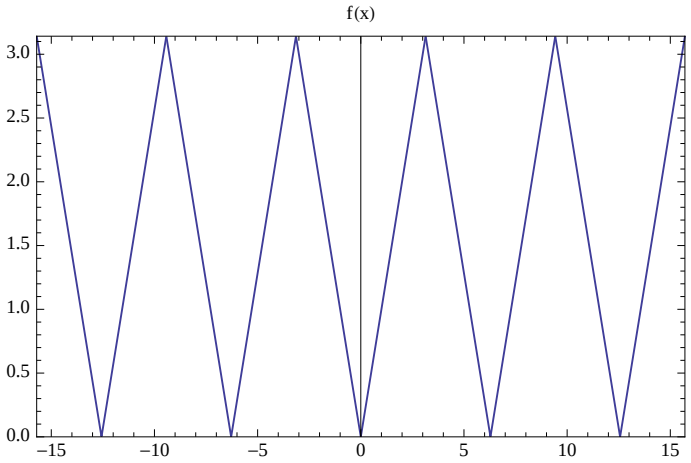

Figure 5: Graph of the periodic functions f defined by $f(x) = f(x + 2\pi)$ and $f(x) = |x|$ for $-\pi < x \leq \pi$.

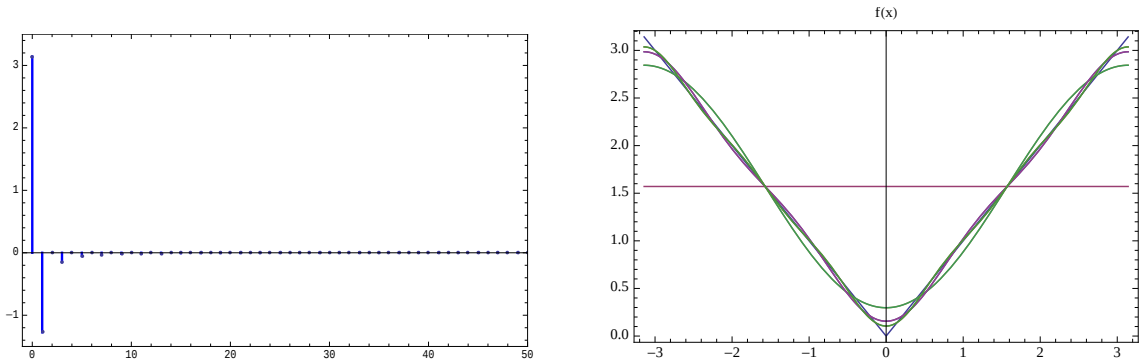

Figure 6: Fourier coefficients and Fourier series for the modulus function f in Eq. (3.39). The left figure shows the Fourier coefficients a_k from Eq. (3.40) for $k = 1, \ldots, 50$. The function f together with the first six partial sums of its Fourier series (3.42) is shown in the right figure.

Application 3.8. Modulus function

Our next example is the modulus function $f : [-\pi, \pi] \to \mathbb{R}$ defined by

$$f(x) = |x| \, . \tag{3.39}$$

Extended to a periodic function with period 2π its graph is shown in Fig. 5. Since this function is symmetric the Fourier series of course only contains cosine terms. (Alternatively, and equivalently, we can consider this function restricted to the interval $[0, \pi]$ and compute its cosine Fourier series.) Using Eqs. (3.22) we find for the Fourier coefficients

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \, |x| = \pi \, , \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \, |x| \cos(kx) = \frac{2\left((-1)^k - 1\right)}{\pi k^2} \, , \quad b_k = 0 \, , \tag{3.40}$$

for $k = 1, 2, \ldots$. The standard integral which enters this calculation is

$$\int dx \, x \cos(\alpha x) = \frac{\cos(\alpha x)}{\alpha^2} + \frac{x \sin(\alpha x)}{\alpha} \, , \tag{3.41}$$

where $\alpha \neq 0$. Note that the value obtained for a_0 does not follow by inserting $k = 0$ into the general formula for a_k - in fact, doing this leads to an undefined expression. This observation points to a general rule. Sometimes Fourier coefficients, calculated for generic values of $k \in \mathbb{N}$, become apparently singular or undefined for specific k values. Of course Fourier coefficients must be well-defined, so this indicates a break-down of the integration method which occurs for those specific values of k. In such cases, the integrals for the problematic k-values should be carried out separately and, with the correct integration applied, this will lead to well-defined answers. In the present case, the special case arises because the standard integral (3.41) is only valid for $\alpha \neq 0$.

The Fourier series from the above coefficients is given by

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1,3,5,\ldots} \frac{\cos(kx)}{k^2} \, . \tag{3.42}$$

The Fourier coefficients a_k and the first few partial sums of the above Fourier series are shown in Fig. 6. The Fourier coefficients drop off as $a_k \sim 1/k^2$, so more quickly as in the previous example, and convergence of the Fourier series is more efficient. A related observation is that the function (3.39) satisfies all the conditions of Theorem 3.8 and, hence, its Fourier series converges uniformly (and point-wise) to f. Fig. (6) illustrates this convincingly.

Application 3.9. Sign function

Another interesting example is the sign function $f : [-\pi, \pi] \to \mathbb{R}$ defined by

$$f(x) := \mathrm{sign}(x) = \begin{cases} 1 & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x = 0 \\ -1 & \text{for} \quad x < 0 \end{cases} \, . \tag{3.43}$$

The periodically continued version of this function is shown in Fig. 7. Since f is an anti-symmetric function, the Fourier series only contains sine terms. (Alternatively and equivalently, we can think of f as a function on the $[0, \pi]$ and work out the sine Fourier series.) For the Fourier coefficients we

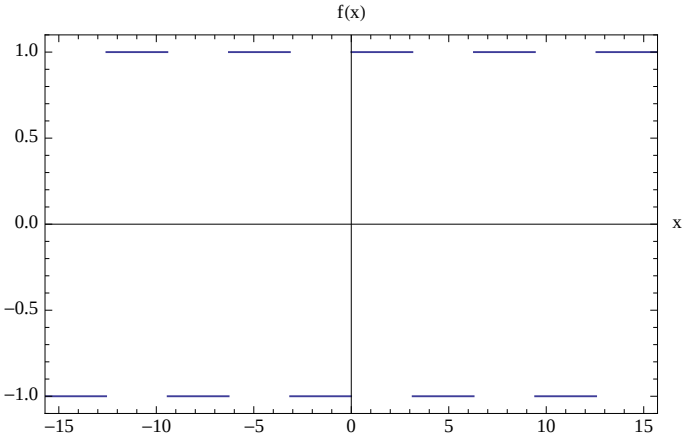

Figure 7: Graph of the periodic functions f defined by $f(x) = f(x+2\pi)$ and $f(x) = \mathrm{sign}(x)$ for $-\pi < x \leq \pi$.

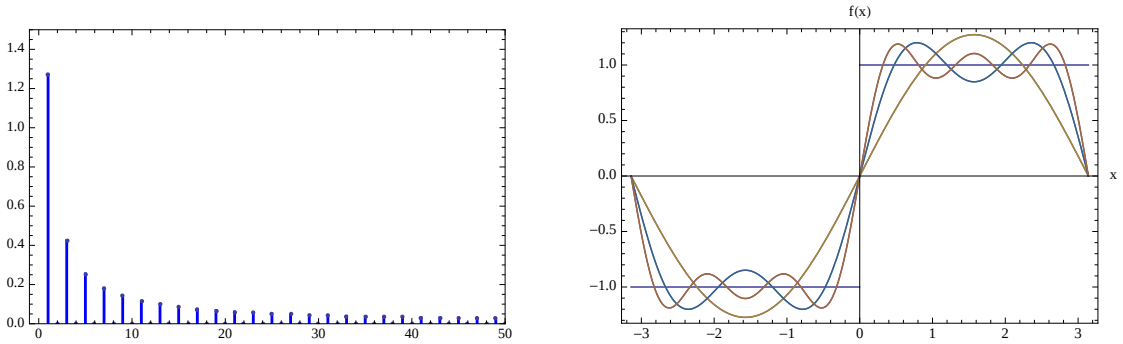

Figure 8: Fourier coefficients and Fourier series for the sign function f in Eq. (3.43). The left figure shows the Fourier coefficients a_k from Eq. (3.44) for $k = 1, \ldots, 50$. The function f together with the first six partial sums of its Fourier series (3.45) is shown in the right figure.

have

$$a_k = 0\,, \qquad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} dx\, \mathrm{sign}(x)\sin(kx) = -\frac{2\left((-1)^k - 1\right)}{\pi k}\,, \tag{3.44}$$

for $k = 1, 2, \ldots$ which leads to the Fourier series

$$f(x) = -\frac{4}{\pi}\sum_{k=1,3,5,\ldots} \frac{\sin(kx)}{k}\,. \tag{3.45}$$

The Fourier coefficients b_k and the first few partial sums of the Fourier series are shown in Fig. 8. As for example 1, the function f does not satisfy the conditions of Theorem 3.8 and the Fourier series does not converge everywhere point-wise to the function f. Specifically, while the Fourier series always vanishes at $x = \pm\pi$ the function value $f(\pm\pi) = \pm 1$ is non-zero. Related to this is the slow drow-off, $a_k \sim 1/k$, of the Fourier coefficients.

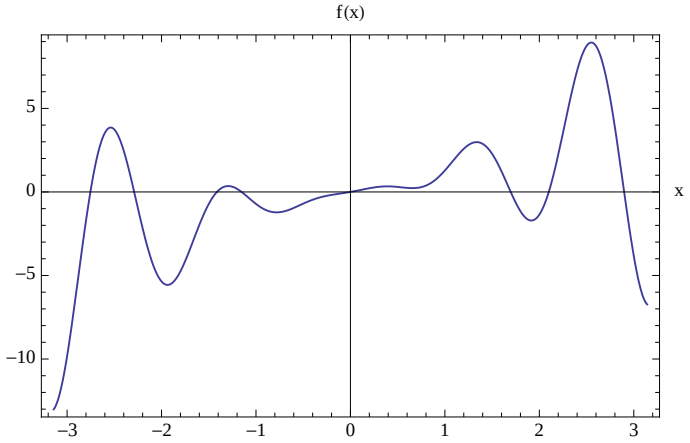

Figure 9: Graph of the periodic functions f defined in Eq. (3.46).

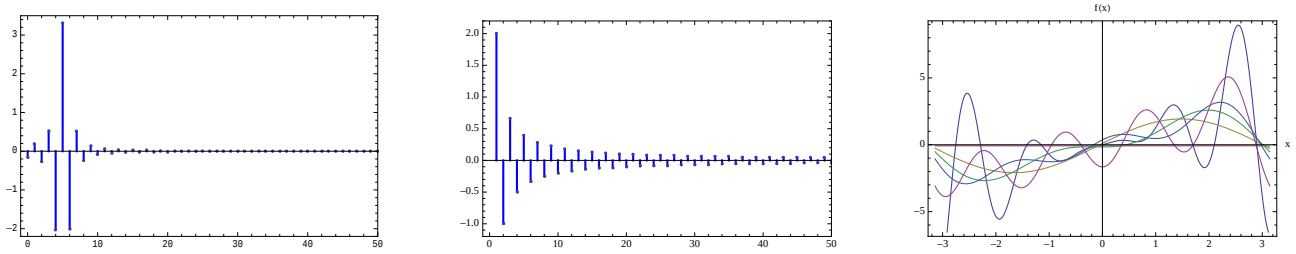

Figure 10: Fourier coefficients and Fourier series for the function f in Eq. (3.46). The left figure shows the Fourier coefficients a_k and the middle figure the coefficients b_k from Eq. (3.47). The function f together with the first six partial sums of its Fourier series (3.48) is shown in the right figure.

Application 3.10. A more complicated function

As a final and more complicated example, consider the function $f : [-\pi, \pi] \to \mathbb{R}$ defined by

$$f(x) = x^2 \cos(5x) + x \; . \tag{3.46}$$

with graph as shown in Fig. 9. Since this function is neither symmetric nor anti-symmetric we expect both sine and cosine terms to be present. For the Fourier coefficients is follows

$$\begin{aligned} a_k &= \tfrac{1}{\pi} \textstyle\int_{-\pi}^{\pi} dx\, f(x) \cos(kx) = -\frac{4(-1)^k (k^2+25)}{(k^2-25)^2} \; , \quad k = 0, \ldots, 4, 6, \ldots \\ a_5 &= \tfrac{1}{\pi} \textstyle\int_{-\pi}^{\pi} dx\, f(x) \cos(5x) = \frac{1}{50} + \frac{\pi^2}{3} \\ b_k &= \tfrac{1}{\pi} \textstyle\int_{-\pi}^{\pi} dx\, f(x) \sin(kx) = -\frac{2(-1)^k}{k} \quad k = 1, 2, \ldots \; . \end{aligned} \tag{3.47}$$

Note that inserting $k = 5$ into the first expression for a_k leads to a singularity and, hence, the generic integration method has broken down in this case. (This is, of course, related to the presence of $\cos(5x)$ in f.) For this reason, we have carried out the calculation of a_5 separately. The resulting Fourier

series is

$$f(x) = -4\sum_{k\neq 5} \frac{(-1)^k \left(k^2+25\right)}{\left(k^2-25\right)^2} \cos(kx) + \left(\frac{1}{50} - 2\frac{\pi^2}{3}\right)\cos(5x) + \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \sin(kx)\ . \tag{3.48}$$

The corresponding plots for the coefficients a_k, b_k and the partial Fourier series are shown in Fig. 10.

Again this is a function which violates the conditions of Theorem 3.8 and where the Fourier series does not converge to the function everywhere. A new feature is the structure in the Fourier coefficients a_k, as seen in the left Fig. 10. The Fourier mode a_5 (and, to a lesser extent, that of a_4 and a_6) is much stronger than the other modes a_k and this is of course related to the presence of $\cos(5x)$ in the function f. Intuitively, the Fourier series detects the strength with which frequencies k are contained in a function f and the presence of $\cos(5x)$ in f suggests a strong contribution for $k = 5$.

Application 3.11. A complex example

So far our examples have been for the real Fourier series so we finish with an example for the complex case. Of course any of the previous real-valued functions can also be expanded into a complex Fourier series. For example, for the modulus function $f : [-\pi, \pi] \to \mathbb{R}$ with $f(x) = |x|$, discussed in example 2, the coefficients for the complex Fourier series are, from Eq. (3.31), given by

$$a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} dx\, |x| \exp(-ikx) = \begin{cases} \frac{(-1)^k-1}{\pi k^2} & \text{for} \quad k = 1, 2, \dots \\ \frac{\pi}{2} & \text{for} \quad k = 0 \end{cases}\ . \tag{3.49}$$

Hence the complex Fourier series is

$$f(x) = \frac{\pi}{2} - \frac{2}{\pi}\sum_{k\in\mathbb{Z}^{\text{odd}}} \frac{\exp(ikx)}{k^2}\ . \tag{3.50}$$

which we could have also inferred from the real Fourier series (3.42) by simply replacing $\cos(kx) = (\exp(ikx) + \exp(-ikx))/2$.

3.2 Fourier transform

As we have seen, the Fourier series provides a frequency analysis for functions on a finite interval, in terms of a discrete spectrum of frequencies labeled by an integer k. The Fourier transform serves a similar purpose but for functions on all of $\mathbb{R}$ (or $\mathbb{R}^n$), leading to a frequency analysis in terms of a continuous spectrum of frequencies.

3.2.1 Basic definition and properties

The natural arena to start the discussion is the Banach space $L^1_{\mathbb{C}}(\mathbb{R}^n)$, defined in Eq. (1.90), with norm $\|\cdot\|_1$, defined in Eq. (1.91). As usual, we denote vectors in $\mathbb{R}^n$ by bold-face letter, so $\mathbf{x} = (x_1, \dots, x_n)^T$. A simple observation is that for a function $f \in L^1_{\mathbb{C}}(\mathbb{R}^n)$ we have $\exp(-i\mathbf{x}\cdot\mathbf{k})f \in L^1_{\mathbb{C}}(\mathbb{R}^n)$ for any vector $\mathbf{k} \in \mathbb{R}^n$. Hence, it makes sense to define [5]

[5] There are different conventions for how to insert factors of 2π into the definition of the Fourier transform. The convention adopted below is the most symmetric choice, as we will see later.

Definition 3.1. *For functions $f \in L^1_{\mathbb{C}}(\mathbb{R}^n)$ we define the Fourier transform $\mathcal{F}f = \hat{f} : \mathbb{R}^n \to \mathbb{C}$ by*

$$\hat{f}(\mathbf{k}) = \mathcal{F}(f)(\mathbf{k}) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n x \, \exp(-i\mathbf{x}\cdot\mathbf{k}) f(\mathbf{x}) \, . \tag{3.51}$$

Clearly, $\mathcal{F}$ is a linear operator, that is $\mathcal{F}(\alpha f + \beta g) = \alpha\mathcal{F}(f) + \beta\mathcal{F}(g)$. Also note that $|\hat{f}(\mathbf{k})| \leq \frac{\|f\|_1}{(2\pi)^{n/2}}$, so the modulus of the Fourier transform is bounded. With some more effort it can be shown that $\hat{f}$ is continuous. However, it is not clear that the Fourier transform $\hat{f}$ is an element of $L^1_{\mathbb{C}}(\mathbb{R}^n)$ as well and, it turns out, this is not always the case. (See Example 3 below.) We will rectify this later by defining a version of the Fourier transform which provides a map $L^2_{\mathbb{C}}(\mathbb{R}^n) \to L^2_{\mathbb{C}}(\mathbb{R}^n)$.

Before we compute examples of Fourier transforms it is useful to look at some of its general properties. Recall from Section 2.3 the translation operator $T_{\mathbf{a}}$, the modulation operator $E_{\mathbf{b}}$ and the dilation operator $\mathcal{D}_\lambda$, for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, defined by

$$T_{\mathbf{a}}(f)(\mathbf{x}) := f(\mathbf{x} - \mathbf{a}) \, , \qquad E_{\mathbf{b}}(f)(\mathbf{x}) := \exp(i\mathbf{b}\cdot\mathbf{x}) f(\mathbf{x}) \, , \qquad \mathcal{D}_\lambda(f)(\mathbf{x}) := f(\lambda\mathbf{x}) \, , \tag{3.52}$$

which we can also think of as maps $L^1_{\mathbb{C}}(\mathbb{R}^n) \to L^1_{\mathbb{C}}(\mathbb{R}^n)$. For any function $g : \mathbb{R}^n \to \mathbb{C}$, we also have the multiplication operator

$$M_g(f)(\mathbf{x}) := g(\mathbf{x}) f(\mathbf{x}) \, . \tag{3.53}$$

It is useful to work out how these operators as well as derivative operators $D_{x_j} := \frac{\partial}{\partial x_j}$ relate to the Fourier transform.

Proposition 3.1. *(Some elementary properties of the Fourier transform) For $f \in L^1_{\mathbb{C}}(\mathbb{R}^n)$ we have*
(F1) $\widehat{T_{\mathbf{a}}(f)} = E_{-\mathbf{a}}(\hat{f})$ or, equivalently, $\mathcal{F} \circ T_{\mathbf{a}} = E_{-\mathbf{a}} \circ \mathcal{F}$
(F2) $\widehat{E_{\mathbf{b}}(f)} = T_{\mathbf{b}}(\hat{f})$ or, equivalently, $\mathcal{F} \circ E_{\mathbf{b}} = T_{\mathbf{b}} \circ \mathcal{F}$
(F3) $\widehat{\mathcal{D}_\lambda(f)} = \frac{1}{|\lambda|^n}\mathcal{D}_{1/\lambda}(\hat{f})$ or, equivalently, $\mathcal{F} \circ \mathcal{D}_\lambda = \frac{1}{|\lambda|^n}\mathcal{D}_{1/\lambda} \circ \mathcal{F}$
For $f \in \mathcal{C}^1_c(\mathbb{R}^n)$ we have
(F4) $\widehat{D_{x_j} f}(\mathbf{k}) = ik_j \hat{f}(\mathbf{k})$ or, equivalently, $\mathcal{F} \circ D_{x_j} = M_{ik_j} \circ \mathcal{F}$
(F5) $\widehat{x_j f}(\mathbf{k}) = iD_{k_j}\hat{f}$ or, equivalently, $\mathcal{F} \circ M_{x_j} = iD_{k_j} \circ \mathcal{F}$.

Proof. (F1) This can be shown by direct calculation.

$$\widehat{T_{\mathbf{a}}(f)}(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dx^n \, e^{-i\mathbf{x}\cdot\mathbf{k}} f(\mathbf{x} - \mathbf{a}) \overset{\mathbf{y}=\mathbf{x}-\mathbf{a}}{=} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dy^n \, e^{-i(\mathbf{y}+\mathbf{a})\cdot\mathbf{k}} f(\mathbf{y}) = E_{-\mathbf{a}}(\hat{f})(\mathbf{k}) \, . \tag{3.54}$$

The proofs for (F2) to (F5) are similar and are left as an exercise. □

Exercise 3.9. *Proof (F2), (F3), (F4) and (F5) from Prop. 3.1.*

3.2.2 Convolution

Another operation which relates to Fourier transforms in an interesting way is the *convolution* $f \star g$ of two functions $f, g \in \mathcal{L}^1(\mathbb{R}^n)$ which is defined as

$$(f \star g)(\mathbf{x}) := \int_{\mathbb{R}^n} dy^n \, f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) \, . \tag{3.55}$$

A straightforward computation shows that the convolution is commutative, so $f \star g = g \star f$.

Exercise 3.10. *Show that the convolution commutes.*

From a mathematical point of view, we have the following statement about convolutions.

Theorem 3.11. *(Property of convolutions) For $f, g \in L^1(\mathbb{R}^n)$ the convolution $f \star g$ is well-defined and $f \star g \in L^1(\mathbb{R}^n)$.*

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

How can the convolution be understood intuitively? From the integral (3.55) we can say that the convolution is "smearing" the function f by the function g. For example, consider choosing $f(x) = \cos(x)$ and

$$g(x) = \begin{cases} \frac{1}{2a} & \text{for} \quad x \in [-a, a] \\ 0 & \text{for} \quad |x| > a \end{cases} , \tag{3.56}$$

for any $a > 0$. The function g is chosen so that, upon convolution, it leads to a smearing (or averaging) of the function f over the interval $[x-a, x+a]$ for every x. An explicit calculation shows the convolution is given by

$$(f \star g)(x) = \frac{1}{2a} \int_{x-a}^{x+a} dy \, \cos(y) = \frac{\sin(a)}{a} \cos(x) \, . \tag{3.57}$$

If we consider the limit $a \to 0$, so the averaging width goes to zero, we find that $f \star g = f$ so f remains unchanged, as one would expect. The other extreme would be to choose $a = \pi$ in which case $f \star g = 0$. In this case, the averaging is over a period $[x - \pi, x + \pi]$ of the cos so the convoluted function vanishes for every x. For other values of a the convolution is still a cos function but with a reduced amplitude, as one would expect from a local averaging.

The relationship between convolutions and Fourier transforms is stated in the following Lemma.

Proposition 3.2. *For $f, g \in L^1(\mathbb{R}^n)$ we have $\widehat{f \star g} = (2\pi)^{n/2} \hat{f} \, \hat{g}$.*

Proof. The proof works by straightforward calculation. Since $(f \star g)(\mathbf{x}) = \int dy^n f(\mathbf{y}) g(\mathbf{x} - \mathbf{y})$ we have

$$\widehat{f \star g}(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int dx^n \, dy^n f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) e^{-i\mathbf{x}\cdot\mathbf{k}} \tag{3.58}$$

$$= \frac{1}{(2\pi)^{n/2}} \int dx^n \, dy^n \left(f(\mathbf{y}) e^{-i\mathbf{y}\cdot\mathbf{k}} \right) \left(g(\mathbf{x} - \mathbf{y}) e^{-i(\mathbf{x}-\mathbf{y})\cdot\mathbf{k}} \right) \tag{3.59}$$

$$\overset{\mathbf{z}=\mathbf{x}-\mathbf{y}}{=} \frac{1}{(2\pi)^{n/2}} \int dy^n f(\mathbf{y}) e^{-i\mathbf{y}\cdot\mathbf{k}} \int dz^n \, g(\mathbf{z}) e^{-i\mathbf{z}\cdot\mathbf{k}} = (2\pi)^{n/2} \hat{f}(\mathbf{k}) \hat{g}(\mathbf{k}) \, . \tag{3.60}$$

□

In other words, the Fourier transform of a convolution is (up to a constant) the product of the two Fourier transforms. This rule is often useful to work out new Fourier transforms from given ones.

3.2.3 Examples of Fourier transforms

We should now discuss a few examples of Fourier transforms to get a better idea of its interpretation. In this context it is useful to think of the function f as the amplitude of a sound signal and we will rename its variable as $x \to t$, to indicate dependence on time. Correspondingly, the variable k of the Fourier transform $\hat{f}$ will be renamed as $k \to \omega$, indicating frequency. So we write (focusing on the one-dimensional case)

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dt \, f(t) e^{-i\omega t} \, . \tag{3.61}$$

The basic idea is that the Fourier transform provides the decomposition of the signal f into its frequency components $e^{-i\omega t}$, that is, $\hat{f}(\omega)$ indicates the strength with which the frequency ω is contained in the

signal f. Suppose that f is the signal from a single piano tone with frequency ω_0. In this case, we expect $\hat{f}$ to have a strong peak around $\omega = \omega_0$. However, a piano tone also contains overtones with frequencies $q\omega_0$, where $q = 2, 3, \ldots$. This means we expect $\hat{f}$ to have smaller peaks around $\omega = q\omega_0$. (Their height decreases with increasing q and exactly what the pattern is determines how the tone "sounds".) Let us consider this in a more quantitative way.

Application 3.12. A damped single-frequency signal

Consider the function

$$f = A f_{\omega_0,\gamma} \,, \qquad f_{\omega_0,\gamma}(t) = \begin{cases} \sqrt{2\pi}\, e^{-\gamma t} e^{i\omega_0 t} & \text{for} \quad t \geq 0 \\ 0 & \text{for} \quad t < 0 \end{cases} \,, \tag{3.62}$$

where A, γ and ω_0 are real, positive constants. Using the above sound analogy, we can think of this function as representing a sound signal with onset at $t = 0$, overall amplitude A, frequency ω_0 and a decay time of $\sim 1/\gamma$. Inserting into Eq. (3.61), we find the Fourier transform

$$\hat{f}(\omega) = A\, \widehat{f_{\omega_0,\gamma}}(\omega) \,, \qquad \widehat{f_{\omega_0,\gamma}}(\omega) = \int_0^\infty dt\, e^{-\gamma t} e^{-i(\omega-\omega_0)t} = \frac{1}{i(\omega_0 - \omega) - \gamma} \,. \tag{3.63}$$

To interpret this result we compute its complex modulus

$$|\widehat{f_{\omega_0,\gamma}}(\omega)|^2 = \frac{1}{(\omega_0 - \omega)^2 + \gamma^2} \,, \tag{3.64}$$

and this corresponds to a peak with width $\sim \gamma$ around $\omega = \omega_0$. The longer the tone, the smaller γ and the smaller the width of this peak. Note that this is precisely in line with our expectation. The original signal contains a strong component with frequency ω_0 which corresponds to the peak of the Fourier transform around ω_0. However, there is a spectrum of frequencies around ω_0 and this captures the finite decay time $\sim 1/\gamma$ of the signal. The longer the signal the closer it is to a pure signal with frequency ω_0 and the narrower the peak in the Fourier transform.

We can take the sound analogy further by considering

$$f = \sum_{q=1,2\ldots} A_q\, f_{q\omega_0,\gamma_q} \,, \tag{3.65}$$

where the function $f_{q\omega_0,\gamma_q}$ is defined in Eq. (3.62). This represents a tone with frequency ω_0, together with its overtones with amplitudes A_q, frequencies $q\omega_0$ and decay constants γ_q. The Fourier transform of f is easily computed from linearity:

$$\hat{f}(\omega) = \sum_{q=1,2,\ldots} A_q\, \widehat{f_{q\omega_0,\gamma_q}}(\omega) = \sum_{q=1,2,\ldots} \frac{A_q}{i(q\omega_0 - \omega) - \gamma_q} \,. \tag{3.66}$$

This corresponds to a sequence of peaks at frequencies $q\omega_0$, where $q = 1, 2, \ldots$ which reflects the main frequency of the tone, together with its overtone frequencies.

Application 3.13. Fourier transforms of Gaussians

Another interesting example to consider is the Fourier transform of the one-dimensional Gaussian

$$f(x) = e^{-x^2/2} \,. \tag{3.67}$$

with width one. For its Fourier transform we have

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx\, e^{-x^2/2 - ikx} = e^{-k^2/2} \,. \tag{3.68}$$

Exercise 3.12. *Proof Eq. (3.68). (Hint: Complete the square in the exponent.)*

This result means that the Gaussian is invariant under Fourier transformation. Without much effort, this one-dimensional result can be generalised to the n-dimensional width one Gaussian

$$f(\mathbf{x}) = e^{-|\mathbf{x}|^2/2} \,. \tag{3.69}$$

Its Fourier transform can be split up into a product of n one-dimensional Fourier transforms as

$$\hat{f}(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} dx^n\, e^{-|\mathbf{x}|^2/2 - i\mathbf{k}\cdot\mathbf{x}} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx_i\, e^{-x_i^2/2 - ik_i x_i} = \prod_{i=1}^{n} e^{-k_i^2/2} = e^{-|\mathbf{k}|^2/2} \,, \tag{3.70}$$

and the one-dimensional result (3.68) has been used in the second-last step. Hence, the n-dimensional width one Gaussian is also invariant under Fourier transformation.

We would like to work out the Fourier transform of a more general Gaussian with width $a > 0$, given by

$$f_a(\mathbf{x}) = e^{-\frac{|\mathbf{x}|^2}{2a^2}} = \mathcal{D}_{1/a}(f)(\mathbf{x}) \,. \tag{3.71}$$

where f is the Gaussian (3.69) with width one and $\mathcal{D}$ is the dilation operator defined in Eq. (3.52). The fact that this can be expressed in terms of the dilation operator makes calculating the Fourier transform quite easy, using the property (F3) in Lemma 3.1.

$$\widehat{f_a}(\mathbf{k}) = \widehat{\mathcal{D}_{1/a}(f)}(\mathbf{k}) = a^n \mathcal{D}_a(\hat{f})(\mathbf{k}) = a^n e^{-a^2|\mathbf{k}|^2/2} \,. \tag{3.72}$$

In the last step the result (3.70) for the Fourier transform $\hat{f}$ of the width one Gaussian has been used. In conclusion, the Fourier transform of a Gaussian with width a is again a Gaussian with width $1/a$.

Finally, we consider a Gaussian with width a and center shifted from the origin to a point $\mathbf{c} \in \mathbb{R}^n$ given by

$$f_{a,\mathbf{c}}(\mathbf{x}) = \exp\left(-\frac{|\mathbf{x}-\mathbf{c}|^2}{2a^2}\right) = T_{\mathbf{c}}(f_a)(\mathbf{x}) \,. \tag{3.73}$$

Note that this can be written in terms of the zero-centred Gaussian using the translation operator (3.52) and we can now use property (F1) in Lemma 3.1 to work out the Fourier transform.

$$\widehat{f_{a,\mathbf{c}}}(\mathbf{k}) = \widehat{T_{\mathbf{c}}(f_a)}(\mathbf{k}) = E_{-\mathbf{c}}(\widehat{f_a})(\mathbf{k}) = a^n e^{-i\mathbf{c}\cdot\mathbf{k} - a^2|\mathbf{k}|^2/2} \,. \tag{3.74}$$

Application 3.14. Characteristic function of an interval

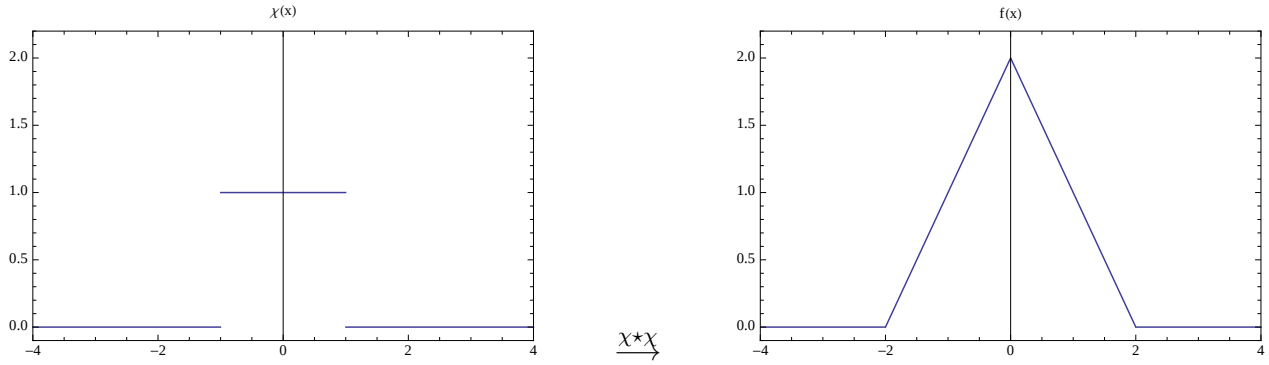

Figure 11: The graph of the characteristic function χ of the interval $[-1,1]$ (left) and the graph of the convolution $f = \chi \star \chi$ (right).

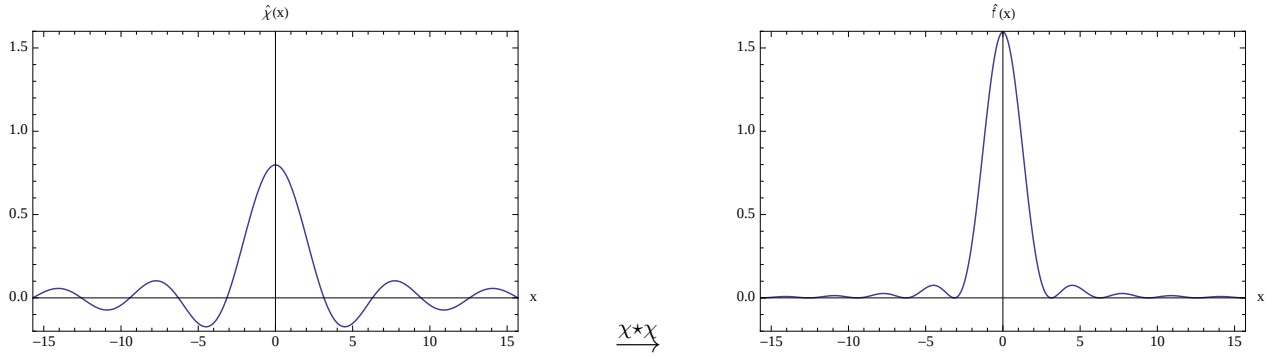

Figure 12: The graph of the Fourier transform $\hat{\chi}$ of the characteristic function χ (left) and the graph for the Fourier transform of the convolution $\widehat{\chi \star \chi} = \sqrt{2\pi}\,\hat{\chi}^2$. (right).

Consider the characteristic function $\chi : \mathbb{R} \to \mathbb{R}$ of the interval $[-1,1]$ defined by

$$\chi(x) = \begin{cases} 1 & \text{for} \quad |x| \leq 1 \\ 0 & \text{for} \quad |x| > 1 \end{cases} . \tag{3.75}$$

A quick calculation shows that its Fourier transform is given by

$$\hat{\chi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} dx\, e^{-ikx} = \sqrt{\frac{2}{\pi}} \frac{\sin k}{k} . \tag{3.76}$$

Exercise 3.13. *Show that the Fourier transform (3.76) of the characteristic function χ is not in $L^1(\mathbb{R})$. (Hint: Find a lower bound for the integral over $|\sin k/k|$ from $(m-1)\pi$ to $m\pi$.) Use the dilation operator to find the Fourier transform of the characteristic function χ_a for the interval $[-a,a]$.*

We can use this example as an illustration of convolutions and their application to Fourier transforms.

Consider the convolution $f = \chi \star \chi$ of χ with itself which is given by

$$f(x) = \int_{\mathbb{R}} dy\, \chi(y)\chi(y-x) = \max(2-|x|, 0)\ . \tag{3.77}$$

The graphs of χ and its convolution $f = \chi\star\chi$ are shown in Fig. 11. From the convolution theorem 3.11 and the Fourier transform $\hat{\chi}$ in Eq. (3.76) we have

$$\hat{f}(k) = \widehat{\chi\star\chi}(k) = \sqrt{2\pi}\,\hat{\chi}^2(k) = 2\sqrt{\frac{2}{\pi}}\frac{\sin^2 k}{k^2}\ . \tag{3.78}$$

Fig. 12 shows the graphs for the Fourier transforms $\hat{\chi}$ and $\hat{f}$.

3.2.4 The inverse of the Fourier transform

We should now come back to general properties of the Fourier transform. An obvious question is how to obtain a function f from its Fourier transform $\hat{f}$ and this is answered by the following theorem.

Theorem 3.14. *(Inversion formula for the Fourier transform)*
Consider a function $f \in L^1(\mathbb{R}^n)$ such that $\hat{f} \in L^1(\mathbb{R}^n)$. Then we have

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} dk^n\, \hat{f}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}\ , \tag{3.79}$$

almost everywhere, that is for all $\mathbf{x} \in \mathbb{R}^n$ except possibly on a set of Lebesgue measure zero.

Proof. The proof is somewhat technical (suggested by the fact that equality can fail on a set of measure zero) and can, for example, be found in Ref. [10]. □

Note that the inversion formula (3.79) is very similar to the original definition (3.51) of the Fourier transform, except for the change of sign in the exponent. It is, therefore, useful to introduce the linear operator

$$\tilde{\mathcal{F}}(\hat{f})(\mathbf{x}) := \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} dk^n\, \hat{f}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} \tag{3.80}$$

for the inverse Fourier transform. With this terminology, the statement of Theorem 3.14 can be expressed as

$$\tilde{\mathcal{F}}\circ\mathcal{F}(f) = f \quad \Rightarrow \quad \mathcal{F}\circ\tilde{\mathcal{F}}(f) = f\ . \tag{3.81}$$

Exercise 3.15. *Show that the equation on the RHS of (3.81) does indeed follow from the equation on the LHS. (Hint: Think about complex conjugation.)*

Theorem 3.14 also means that a function f is uniquely (up to values on a measure zero set) determined by its Fourier transform $\hat{f}$.

3.2.5 Fourier transform in L^2

In Exercise 3.13 we have seen that the Fourier transform of a function in $L^1(\mathbb{R}^n)$ may not be an element of $L^1(\mathbb{R}^n)$. This is somewhat unsatisfactory and we will now see that the Fourier transform has nicer properties when defined on the space $L^2(\mathbb{R}^n)$. We begin with the following technical Lemma.

Lemma 3.1. *If $f \in \mathcal{C}_c^{n+1}(\mathbb{R}^n)$ then the Fourier transform $\hat{f}$ is integrable, that is, $\hat{f} \in L^1(\mathbb{R}^n)$.*

Proof. Property (F4) in Lemma 3.1 states that $\widehat{D_{x_j} f}(\mathbf{k}) = ik_j \hat{f}(\mathbf{k})$ which implies that

$$|k_i \hat{f}(\mathbf{k})| \leq \| D_i f \|_1/(2\pi)^{n/2} \, . \tag{3.82}$$

Differentiating and applying this rule repeatedly, we conclude that there is a constant K such that

$$\left(1 + \sum_{i=1}^{n} |k_i|\right)^{n+1} |\hat{f}(\mathbf{k})| \leq K \tag{3.83}$$

and this means that $\hat{f}$ is integrable. □

The next Lemma explores the relationship between the Fourier transform and the standard scalar product on $L^2(\mathbb{R}^n)$.

Lemma 3.2. *(a) Let $f, g \in L^1(\mathbb{R}^n)$ with Fourier transforms $\hat{f}$ and $\hat{g}$. Then $\hat{f}g$ and $f\hat{g}$ are integrable and we have*

$$\int_{\mathbb{R}^n} dx^n \, \hat{f}(\mathbf{x}) g(\mathbf{x}) = \int_{\mathbb{R}^n} dx^n \, f(\mathbf{x}) \hat{g}(\mathbf{x}) \tag{3.84}$$

(b) For $f, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ we have $\hat{f}, \hat{g} \in L^2(\mathbb{R}^n)$ and

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle \, , \tag{3.85}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product on $L^2(\mathbb{R}^n)$.

Proof. (a) Since $\hat{f}, \hat{g}$ are bounded and continuous, $\hat{f}g$ and $f\hat{g}$ are indeed integrable. It follows

$$\int dx^n \, f(\mathbf{x}) \hat{g}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int dx^n \, dy^n \, f(\mathbf{x}) g(\mathbf{y}) e^{-i\mathbf{x}\cdot\mathbf{y}} = \int dy^n \, \hat{f}(\mathbf{y}) g(\mathbf{y}) \, . \tag{3.86}$$

(b) For $h, g \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ we have, from Lemma 3.1, that $\hat{h}, \hat{g} \in \mathcal{L}^1(\mathbb{R}^n)$. Then, we can apply part (a) to get

$$\langle \hat{h}, \hat{g} \rangle = \int dx^n \, \tilde{\mathcal{F}}(\bar{h})(\mathbf{x}) \hat{g}(\mathbf{x}) = \int dx^n \, \mathcal{F} \circ \tilde{\mathcal{F}}(\bar{h})(\mathbf{x}) g(\mathbf{x}) = \int dx^n \, \bar{h}(\mathbf{x}) g(\mathbf{x}) = \langle h, g \rangle \, . \tag{3.87}$$

To extend this statement to $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ we recall from Theorem 2.1 that $\mathcal{C}_c^\infty(\mathbb{R}^n)$ is dense in this space. We can, therefore, approximate functions $f, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ by sequences $(f_k), (g_k)$ in $\mathcal{C}_c^\infty(\mathbb{R}^n)$. We have already shown that the property (3.85) holds for all f_k, g_k and, by taking the limit $k \to \infty$ through the scalar product it follows for f, g. In particular, taking $f = g$, it follows that $\| f \|_2 = \| \hat{f} \|_2$ which shows that $\hat{f} \in L^2(\mathbb{R}^n)$. □

Clearly, Eq. (3.85) is a unitarity property of the Fourier transform, relative to the standard scalar product on $L^2(\mathbb{R}^n)$. However, to make this consistent, we have to extend the Fourier transform to all of $L^2(\mathbb{R}^n)$ and this is the content of the following theorem.

Theorem 3.16. *(Plancherel) There exist a vector space isomorphism $\mathcal{T} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ with the following properties:*
(a) $\langle \mathcal{T}(f), \mathcal{T}(g) \rangle = \langle f, g \rangle$ for all $f, g \in L^2(\mathbb{R}^n)$. This implies $\| \mathcal{T}(f) \| = \| f \|$ for all $f \in L^2(\mathbb{R}^n)$
(b) $\mathcal{T}(f) = \mathcal{F}(f)$ for all $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$
(c) $\mathcal{T}^{-1}(f) = \tilde{\mathcal{F}}(f)$ for all $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$

Proof. Since $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ we can find, for every $f \in L^2(\mathbb{R}^n)$, a sequence (f_k) in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ which converges to f in the norm $\| \cdot \|_2$. We set

$$\mathcal{T}(f) := \lim_{k\to\infty} \hat{f}_k \,. \tag{3.88}$$

From Lemma 3.2 (b) the scalar product is preserved for $\mathcal{F}$ on $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and by taking the limit through the scalar product, the same property follows for the operator $\mathcal{T}$ on $L^2(\mathbb{R}^n)$. If $\mathcal{T}(f) = 0$ we have $0 = \| \mathcal{T}(f) \|_2 = \| f \|_2$ and, hence, $f = 0$. This means that $\mathcal{T}$ is injective. From Theorem 3.14 we have $\mathcal{F} \circ \tilde{\mathcal{F}}(f) = f$ so that $\mathcal{C}_c^\infty(\mathbb{R}^n) \subset \mathrm{Im}(\mathcal{T})$. For a $g \in \mathcal{L}^2(\mathbb{R}^n)$ we pick a sequence $(g_k = \mathcal{F}(f_n))$ in $\mathcal{C}_c^\infty(\mathbb{R}^n)$ which converges to g. For $f = \lim_{k\to\infty} f_k$ we then have $\mathcal{T}(f) = \lim_{k\to\infty} \hat{f}_k = \lim_{k\to\infty} g_k = g$ so that $\mathcal{T}$ is surjective. □

It follows that the extension of the Fourier transform to $L^2(\mathbb{R}^n)$ is a unitary linear operator, that is, an operator which preserves the value of the scalar product on $L^2(\mathbb{R}^n)$. Let us illustrate this with an example.

Application 3.15. Unitarity of the Fourier transformation

Consider the Example 3 above for the characteristic function χ of the interval $[-1, 1]$, defined in Eq. (3.75), with Fourier transform $\hat{\chi}$ given in Eq. (3.76). As discussed in Exercise 3.13, $\hat{\chi}$ is not an element of $L^1(\mathbb{R})$, however, it is (and must be from the above theorem) contained in $L^2(\mathbb{R})$. Since, from Theorem 3.16, $\| \chi \| = \| \hat{\chi} \|$ we find

$$2 = \| \chi \|^2 = \| \hat{\chi} \|^2 = \frac{2}{\pi} \int_{\mathbb{R}} dx \, \frac{\sin^2 x}{x^2} \,. \tag{3.89}$$

While the norm $\| \chi \|$ is easily worked out the same cannot be said for the integral on the RHS, so unitarity of the Fourier transform can lead to non-trivial statements.

Exercise 3.17. *Show that the Gaussian f_a with width a in Eq. (3.71) has the same $L^2(\mathbb{R}^n)$ norm as its Fourier transform $\hat{f}_a$.*

4 Orthogonal polynomials

In the previous section, we have discussed the Fourier series and have found a basis for the Hilbert space $L^2[-a,a]$ in terms of sine and cosine functions. Of course we know from the Stone-Weierstrass theorem 1.35 combined with Theorem 2.1 that the polynomials are dense in the Hilbert space $L^2([a,b])$ (and we have used this for some of the proofs related to the Fourier series). So, rather than using relatively complicated, transcendental functions such as sine and cosine as basis functions there is a much simpler possibility: a basis for $L^2([a,b])$ which consists of polynomials. Of course we would want this to be an ortho-normal basis relative to the standard scalar product (1.101) on $L^2([a,b])$. A rather pedestrian method to find ortho-normal polynomials is to start with the monomials $(1,x,x^2,\ldots)$ and apply the Gram-Schmidt procedure.

Exercise 4.1. *For the Hilbert space $L^2([-1,1])$, apply the Gram-Schmidt procedure to the monomials $1,x,x^2$ and show that this leads to the ortho-normal system of polynomials $\frac{1}{\sqrt{2}}$, $\sqrt{\frac{3}{2}}x$, $\sqrt{\frac{5}{8}}(3x^2-1)$.*

The polynomials in Exercise 4.1 obtained in this way are (proportional to) the first three Legendre polynomials which we will discuss in more detail soon. Evidently, the Gram-Schmidt procedure, while conceptually very clear, is not a particularly efficient method in this context. We would like to obtain concise formulae for orthogonal polynomials at all degrees. There is also an important generalisation. In addition to finite intervals, we would also like to allow semi-infinite or infinite intervals, so we would like to allow $a=-\infty$ or $b=\infty$. Of course for such a semi-infinite or infinite interval, polynomials do not have a finite norm relative to the standard scalar product (1.101) on $L^2([a,b])$. To rectify this, we have to consider the Hilbert spaces $L^2_w([a,b])$ with an integrable weight function w and a scalar product defined by

$$\langle f,g\rangle = \int_a^b dx\, w(x) f(x) g(x)\ , \tag{4.1}$$

and choose w appropriately. Thinking about the types of intervals, that is, finite intervals $[a,b]$, semi-infinite intervals $[a,\infty]$ or an infinite interval $[-\infty,\infty]$ and corresponding suitable weight functions w will lead to a classification of different types of orthogonal polynomials which we discuss in the following subsection. For the remainder of the section we will be looking at various entries in this classification, focusing on the cases which are particularly relevant to applications in physics.

4.1 General theory of ortho-normal polynomials

4.1.1 Basic set-up

We will be working in the Hilbert space $L^2_w([a,b])$ with weight function w and scalar product (4.1). The interval $[a,b]$ can be finite but we also allow the cases $|a|<\infty$, $b=\infty$ of a semi-finite interval and $-a=b=\infty$ of the entire real line. On such a space, we would like to find a system of polynomials $(P_n)_{n=0}^\infty$, where we demand that P_n is of degree n, which is orthogonal, that is, which satisfies

$$\langle P_n, P_m\rangle = h_n\,\delta_{nm}\ , \tag{4.2}$$

for positive numbers h_n. (By convention, the standard orthogonal polynomials are not normalised to one, hence the constants h_n.) We also introduce the notation

$$P_n(x) = k_n x^n + k'_n x^{n-1} + \cdots\ , \tag{4.3}$$

that is we call $k_n\neq 0$ the coefficient of the leading monomial x^n and k'_n the coefficient of the sub-leading monomial x^{n-1} in P_n. An immediate consequence of orthogonality is the relation

$$\langle P_n, p\rangle = 0 \tag{4.4}$$

for any polynomial p with degree less than n (This follows because such a p can be written as a linear combination of $P_0, \ldots, P_{n-1}$ which must be orthogonal to P_n.) Furthermore, P_n is (up to an overall constant) uniquely characterised by this property.

4.1.2 Recursion relation

We already know that we can get such an ortho-normal system of polynomials by applying the Gram-Schmidt procedure to the monomials $(1, x, x^2, \ldots)$ and, while this may not be very practical, it tells us that the polynomials P_n are unique, up to overall constants (which are fixed, up to signs, once the constants h_n are fixed). This statement is made more explicit in the following

Theorem 4.2. *(Recursion relations for orthogonal polynomials) The orthogonal polynomials P_n satisfy the following recursion relation*

$$P_{n+1}(x) = (A_n x + B_n)P_n(x) - C_n P_{n-1}(x) \;, \tag{4.5}$$

for $n = 1, 2, \ldots$, where

$$A_n = \frac{k_{n+1}}{k_n} \;, \quad B_n = A_n \left(\frac{k'_{n+1}}{k_{n+1}} - \frac{k'_n}{k_n} \right) \;, \quad C_n = \frac{A_n h_n}{A_{n-1} h_{n-1}} \;. \tag{4.6}$$

Proof. We start by considering the polynomial $P_{n+1} - A_n x P_n$ which (due to the above definition of A_n) is of degree n, rather than $n+1$, and can, hence, be written as

$$P_{n+1} - A_n x P_n = \sum_{i=0}^{n} \alpha_i P_i \;, \tag{4.7}$$

for some $\alpha_i \in \mathbb{R}$. Taking the inner product of this relation with P_k immediately leads to $\alpha_k = 0$ for $k = 0, \ldots, n-2$. This means we are left with

$$P_{n+1} - A_n x P_n = b_n P_n - c_n P_{n-1} \;, \tag{4.8}$$

and it remains to be shown that $b_n = B_n$ and $c_n = C_n$. The first of these statements follows very easily by inserting the expressions (4.3) into Eq. (4.8) and comparing coefficients of the x^n term. To fix c_n we write Eq. (4.8) as

$$c_n P_{n-1} = -P_{n+1} + A_n x P_n + B_n P_n \tag{4.9}$$

and take the inner product of this equation with P_{n-1}. This leads to

$$\begin{aligned} c_n h_{n-1} &= c_n \| P_{n-1} \|^2 = -\langle P_{n+1}, P_{n-1} \rangle + A_n \langle x P_n, P_{n-1} \rangle + B_n \langle P_n, P_{n-1} \rangle = A_n \langle P_n, x P_{n-1} \rangle \\ &= \frac{A_n k_{n-1}}{k_n} \langle P_n, P_n \rangle = \frac{A_n h_n}{A_{n-1}} \end{aligned}$$

and the desired result $c_n = C_n$ follows. □

4.1.3 General Rodriguez formula

There is yet another way to obtain the orthogonal polynomials P_n, via a derivative formula. To see how this works, consider the functions F_n defined by

$$F_n(x) = \frac{1}{w(x)} \frac{d^n}{dx^n} \left(w(x) X^n \right) \;, \qquad X = \begin{cases} (b-x)(a-x) & \text{for} \quad |a|, |b| < \infty \\ x - a & \text{for} \quad |a| < \infty \,,\; b = \infty \\ 1 & \text{for} \quad -a = b = \infty \end{cases} \;. \tag{4.10}$$

Of course we do not know whether these functions are polynomials. Whether they are depends on the choice of weight function w and we will come back to this point shortly. But for now, let us assume that w is such that the F_n are polynomials of degree n. For any polynomial p of degree $n-1$ we then have

$$\langle F_n, p\rangle = \int_a^b dx\, \frac{d^n}{dx^n}\left(w(x)X^n\right) p(x) = (-1)^n \int_a^b dx\, w(x)X^n \frac{d^n p}{dx^n}(x) = 0\ , \tag{4.11}$$

where we have integrated by parts n times. We recall that the orthogonality property (4.4) determines P_n uniquely (up to an overall constant) and since F_n has the same property we conclude there must be constants K_n such that

$$F_n = K_n P_n\ . \tag{4.12}$$

This calculation shows the idea behind the definition (4.10) of the functions F_n. The presence of the derivatives (and of X which ensures vanishing of the boundary terms) means that, provided the F_n are polynomials of degree n, they are orthogonal.

Theorem 4.3. *(Rodriguez formula) If the functions F_n defined in Eq. (4.10) are polynomials of degree n they are proportional to the orthogonal polynomials P_n, so we have constants K_n such that $F_n = K_n P_n$. It follows that*

$$P_n(x) = \frac{1}{K_n w(x)} \frac{d^n}{dx^n}\left(w(x)X^n\right)\ , \qquad X = \begin{cases} (b-x)(a-x) & \text{for} \quad |a|, |b| < \infty \\ x-a & \text{for} \quad |a| < \infty\ ,\ b = \infty \\ 1 & \text{for} \quad -a = b = \infty \end{cases} \tag{4.13}$$

and this is called the (generalised) Rodriguez formula.

4.1.4 Classification of orthogonal polynomials

What remains to be done is to find the weight functions w for which the F_n are indeed polynomials of order n. We start with the case of a finite interval $[a, b]$ with $X = (b-x)(a-x)$ and demand that F_1 be a (linear) polynomial, so

$$F_1(x) = \frac{1}{w(x)}\frac{d}{dx}\left(w(x)X\right) = \frac{1}{w(x)}w'(x)X + X' \stackrel{!}{=} Ax + B \quad \Rightarrow \quad \frac{w'(x)}{w(x)} = \frac{\alpha}{x-b} + \frac{\beta}{x-a}\ , \tag{4.14}$$

for suitable constants A, B, α, β. Solving the differential equation leads to

$$w(x) = C(b-x)^\alpha (x-a)^\beta \tag{4.15}$$

and we can set $C = 1$ by a re-scaling of coordinates. Further, since w needs to be integrable we have to demand that $\alpha > -1$ and $\beta > -1$. Conversely, it can be shown by calculation that for any such choice of w the functions F_n are indeed polynomials of degree n.

For the case $|a| < \infty$ and $b = \infty$ of the half-infinite interval we can proceed analogously and find that the F_n are polynomials of degree n iff

$$w(x) = e^{-x}(x-a)^\alpha\ , \tag{4.16}$$

where $\alpha > -1$.

Finally, for the entire real line, $-a = b = \infty$ a similar calculation leads to

$$w(x) = e^{-x^2}\ . \tag{4.17}$$

The results from this discussion can be summarised in the following

$[a,b]$	$\alpha,\ \beta$	X	$w(x)$	name	symbol
$[-1,1]$	$\alpha>-1,\ \beta>-1$	x^2-1	$(1-x)^\alpha(x+1)^\beta$	Jacobi	$P_n^{(\alpha,\beta)}$
$[-1,1]$	$\alpha=\beta>-1$	x^2-1	$(1-x)^\alpha(x+1)^\alpha$	Gegenbauer	$P_n^{(\alpha,\alpha)}$
$[-1,1]$	$\alpha=\beta=\pm\frac{1}{2}$	x^2-1	$(1-x)^{\pm1/2}(x+1)^{\pm1/2}$	Chebyshev	$T_n^{(\pm)}$
$[-1,1]$	$\alpha=\beta=0$	x^2-1	1	Legendre	P_n
$[0,\infty]$	$\alpha>-1$	x	$e^{-x}x^\alpha$	Laguerre	$L_n^{(\alpha)}$
$[0,\infty]$	$\alpha=0$	x	e^{-x}	Laguerre	L_n
$[-\infty,\infty]$		1	e^{-x^2}	Hermite	H_n

Table 1: *The types of orthogonal polynomials and several sub-classes which result from the classification in Theorem 4.4. The explicit polynomials are obtained by inserting the quantities in the Table into the Rodriguez formula (4.18).*

Theorem 4.4. *The functions*

$$F_n(x)=\frac{1}{w(x)}\frac{d^n}{dx^n}\left(w(x)X^n\right)\ ,\qquad X=\begin{cases}(b-x)(a-x) & for\quad |a|,|b|<\infty\\ x-a & for\quad |a|<\infty\ ,\ b=\infty\\ 1 & for\quad -a=b=\infty\end{cases}\ . \tag{4.18}$$

are polynomials of degree n iff the weight function w is given by

$$w(x)=\begin{cases}(b-x)^\alpha(x-a)^\beta & for\quad |a|,|b|<\infty\\ e^{-x}(x-a)^\alpha & for\quad |a|<\infty\ ,\ b=\infty\\ e^{-x^2} & for\quad -a=b=\infty\end{cases}\ , \tag{4.19}$$

where $\alpha>-1$ and $\beta>-1$. In this case the F_n are orthogonal and $F_n=K_nP_n$ for constants K_n.

This theorem implies a classification of orthogonal polynomials in terms of the type of interval, the limits $[a,b]$ of the interval and the powers α and β which enter the weight function. (Of course a finite interval $[a,b]$ can always be re-scaled to the standard interval $[-1,1]$ and a semi-infinite interval $[a,\infty]$ to $[0,\infty]$.) The different types and important sub-classes of orthogonal polynomials which arise from this classification are listed in Table 4.1.4. We cannot discuss all of these types in detail but in the following we will focus on the Legendre, the $\alpha=0$ Laguerre and the Hermite polynomials which are the most relevant ones for applications in physics. Before we get to this we should derive more common properties of all these orthogonal polynomials.

4.1.5 Differential equation for orthogonal polynomials

All orthogonal polynomials satisfy a (second order) differential equation, perhaps not surprising given their representation in terms of derivatives, as in the Rodriguez formula.

Theorem 4.5. *All orthogonal polynomials P_n covered by the classification theorem 4.4 satisfy the second order linear differential equation*

$$Xy''+K_1P_1y'-n\left(k_1K_1+\frac{n-1}{2}X''\right)y=0\ , \tag{4.20}$$

where P_1 is the linear orthogonal polynomial, k_1 is the coefficient in front of its linear term, X the function in Theorem 4.4 and the coefficient K_1 is defined in Theorem 4.3.

Proof. For ease of notation we abbreviate $D = \frac{d}{dx}$ and evaluate $D^{n+1}(XD(wX^n))$ in two different ways, remembering that X is a polynomial of degree at most two.

$$\begin{aligned} D^{n+1}(XD(wX^n)) &= XD^{n+2}(wX^n) + (n+1)X'D^{n+1}(wX^n) + \frac{1}{2}n(n+1)X''D^n(wX^n) \\ &= K_n\left[XD^2(wP_n) + (n+1)X'D(wP_n) + \frac{1}{2}n(n+1)X''wP_n\right] \\ D^{n+1}(XD(wX^n)) &= D^{n+1}\left(XD(wX)X^{n-1} + (n-1)wX^nX'\right) \\ &= K_n\left[(K_1P_1 + (n-1)X')D(wP_n) + (n+1)(k_1K_1 + (n-1)X'')wP_n\right] \ . \end{aligned}$$

Equating these two results and replacing $y = P_n$ gives after a straightforward calculation

$$wXy'' + (2Xw' + 2wX' - wK_1P_1)y' + \left(Xw'' + (2X' - K_1P_1)w' - (n+1)\left(k_1K_1 + \frac{1}{2}(n-2)X''\right)w\right)y = 0 \ .$$

By working out $D(wX)$ and $D^2(wX)$ one easily concludes that

$$X\frac{w'}{w} = K_1P_1 - X' \ , \qquad X\frac{w''}{w} = (K_1P_1 - 2X')\frac{w'}{w} + k_1K_1 - X'' \ . \tag{4.21}$$

Using these results to replace w' in the factor of y' and w'' in the factor of y in the above differential equation we arrive at the desired result. □

4.1.6 Expanding in orthogonal polynomials

Conventionally, the orthogonal polynomials are not normalised to one, so $\| P_n \|^2 = h_n$ for positive constants h_n which can be computed explicitly for every specific example. Of course we can define ortho-normal systems of polynomials by simple normalising, so

$$\hat{P}_n := \frac{1}{\| P_n \|}P_n = \frac{1}{\sqrt{h_n}}P_n \ . \tag{4.22}$$

These normalised versions of the orthogonal polynomials obviously form an ortho-normal system on $L^2_w([a,b])$, that is,

$$\langle \hat{P}_n, \hat{P}_m \rangle = \delta_{nm} \ . \tag{4.23}$$

In fact, they do form an ortho-normal basis as stated in the following theorem.

Theorem 4.6. *The normalised version $\hat{P}_n$ in Eq. (4.22) of the orthogonal polynomials classified in Theorem 4.4 form an ortho-normal basis of $L^2_w([a,b])$.*

Proof. For a finite interval $[a,b]$ this follows by combining Theorems 2.1 and 1.35, just as we did in our proofs for the Fourier series. For semi-infinite and infinite intervals the proofs can, for example, be found in Ref. [7]. In the next chapter, when we discuss orthogonal polynomials as eigenvectors of hermitian operators we will obtain an independent proof for the basis properties in those cases. □

The above theorem means that every function $f \in L^2_w([a,b])$ can be expanded as

$$f = \sum_{n=0}^{\infty} \langle \hat{P}_n, f \rangle \, \hat{P}_n \ , \tag{4.24}$$

or, more explicitly

$$f(x) = \sum_{n=0}^{\infty} a_n \hat{P}_n(x) \ , \qquad a_n = \int_a^b dx \, w(x)\hat{P}_n(x)f(x) \ . \tag{4.25}$$

This is of course completely in line with the general idea of expanding vectors in terms of an ortho-normal basis on a Hilbert space and it can be viewed as the polynomial analogue of the Fourier series.

We should now discuss the most important orthogonal polynomials in more detail.

4.2 The Legendre polynomials

We recall from Table 4.1.4 that the Legendre polynomials are defined on the finite interval $[a, b] = [-1, 1]$, the function X is given by $X(x) = x^2 - 1$ and the weight function is simply $w(x) = 1$, so that the relevant Hilbert space is $L^2([-1, 1])$. They are conventionally denoted by P_n (not to be confused with the general notation P_n for all orthogonal polynomials we have used in the previous subsection). Their Rodriguez formula reads

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n \quad \Rightarrow \quad K_n = 2^n n! \,, \tag{4.26}$$

where the pre-factor is conventional and we have read off the value of the constants K_n in Theorem 4.3. They are symmetric for n even and anti-symmetric for n odd, so $P_n(x) = (-1)^n P_n(-x)$. Since $(x^2-1)^n = x^{2n} + nx^{2n-2} + \cdots$ we can easily read off the coefficients k_n and k'_n of the monomials x^n and x^{n-1} in P_n as

$$k_n = \frac{(2n)!}{2^n (n!)^2} \,, \qquad k'_n = 0 \,. \tag{4.27}$$

For the normalisation of P_n we find from the Rodriguez formula

$$h_n = \| P_n \|^2 = \int_{-1}^{1} dx\, P(x)^2 = \frac{1}{2^n n!}\int_{-1}^{1} dx\, P_n(x)\frac{d^n}{dx^n}(x^2-1)^n \tag{4.28}$$

$$= \frac{1}{2^n n!}\int_{-1}^{1} dx\, \frac{d^n P_n}{dx^n}(x)(1-x^2)^n = \frac{k_n}{2^n}\int_{-1}^{1} dx\,(1-x^2)^n = \frac{2}{2n+1} \,, \tag{4.29}$$

where we have integrated by parts n times. This means the associated basis of ortho-normal polynomials on $L^2([-1, 1])$ is

$$\hat{P}_n = \sqrt{\frac{2n+1}{2}} P_n \,, \tag{4.30}$$

and functions $f \in L^2([-1, 1])$ can be expanded as

$$f = \sum_{n=0}^{\infty} \langle \hat{P}_n, f\rangle \hat{P}_n \,, \tag{4.31}$$

or, more explicitly, shifting the normalisation factors into the integral, as

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x) \,, \qquad a_n = \frac{2n+1}{2}\int_{-1}^{1} dx\, P_n(x) f(x) \,. \tag{4.32}$$

Such expansions are useful and frequently appear when spherical coordinates are used and we have the standard inclination angle $\theta \in [0, \pi]$. In this case, the Legendre polynomials are usually a function of $x = \cos\theta$ which takes values in the required range $[-1, 1]$. We will see more explicit examples of this shortly.

With the above results it is easy to compute the constants A_n, B_n and C_n which appear in the general recursion formula (4.5) and we find

$$A_n = \frac{2n+1}{n+1} \,, \qquad B_n = 0 \,, \qquad C_n = \frac{n}{n+1} \,. \tag{4.33}$$

Using these values to specialise Eq. (4.5) we find the recursion formula

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \tag{4.34}$$

for the Legendre polynomials. From the Rodriguez formula (4.26) we can easily compute the first few Legendre polynomials:

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2-1), \quad P_3(x) = \frac{1}{2}(5x^3-3x), \quad P_4(x) = \frac{1}{8}(35x^4-30x^2+3)\ . \tag{4.35}$$

Exercise 4.7. *Verify that the first four Legendre polynomials in Eq. (4.35) are orthogonal and are normalised as in Eq. (4.29).*

We can insert the results $X = x^2-1$, $X'' = 2$, $P_1(x) = x$, $K_1 = 2$ and $k_1 = 1$ into the general differential equation (4.20) to obtain

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0\ . \tag{4.36}$$

This is the Legendre differential equations which all Legendre polynomials P_n satisfy.

Exercise 4.8. *Show that the first four Legendre polynomials in Eq. (4.35) satisfy the Legendre differential equation (4.36).*

Another feature of orthogonal polynomials is the existence of a *generating function* $G = G(x,z)$ defined as

$$G(x,z) = \sum_{n=0}^{\infty} P_n(x) z^n \tag{4.37}$$

The generating function encodes all orthogonal polynomials at once and the n^{th} one can be read off as the coefficient of the z^n term in the expansion of G. Of course for this to be of practical use we have to find another more concise way of writing the generating function. This can be obtained from the recursion relation (4.34) which leads to

$$\begin{aligned}
\frac{\partial G}{\partial z} &= \sum_{n=1}^{\infty} P_n(x) n z^{n-1} = \sum_{n=0}^{\infty} (n+1) P_{n+1}(x) z^n = \sum_{n=0}^{\infty} \left[(2n+1)xP_n(x) - nP_{n-1}(x)\right] z^n \\
&= 2xz \sum_{n=1}^{\infty} P_n(x) n z^{n-1} + x \sum_{n=0}^{\infty} P_n(x) z^n - \sum_{n=0}^{\infty} P_n(x)(n+1) z^{n+1} = (2xz - z^2)\frac{\partial G}{\partial z} + (x-z)G\ ,
\end{aligned}$$

This provides us with a differential equation for G whose solution is $G(x,z) = c(1-2xz+z^2)^{-1/2}$, where c is a constant. Since $c = G(x,0) = P_0(x) = 1$, we have

$$G(x,z) = \frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n\ . \tag{4.38}$$

Exercise 4.9. *Check that the generating function (4.38) leads to the correct Legendre polynomials P_n, for $n = 0, 1, 2$.*

Note that Eq. (4.38) can be viewed as an expansion of the generating function G in the sense of Eq. (4.32), with expansion coefficients $a_n = z^n$.

Application 4.16. Expanding the Coulomb potential

An important application of the above generating function is to the expansion of a Coulomb potential term of the form

$$V(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|}\,, \tag{4.39}$$

where $\mathbf{r},\mathbf{r}' \in \mathbb{R}^3$. Introducing the radii $r = |\mathbf{r}|$, $r' = |\mathbf{r}'|$, the angle $\cos\theta = \frac{\mathbf{r}\cdot\mathbf{r}'}{rr'}$ and setting $x = \cos\theta$ and $z = \frac{r'}{r}$ we can use the generating function to re-write the above Coulomb term as

$$V(\mathbf{r},\mathbf{r}') = \frac{1}{r}\frac{1}{\sqrt{1-2\left(\frac{r'}{r}\right)\cos\theta+\left(\frac{r'}{r}\right)^2}} = \frac{1}{r}G(x,z) = \frac{1}{r}\sum_{n=0}^{\infty} P_n(\cos\theta)\left(\frac{r'}{r}\right)^n\,. \tag{4.40}$$

The series on the RHS converges for $r' < r$. If $r' > r$ we can write a similar expansion with the role of $\mathbf{r}$ and $\mathbf{r}'$ exchanged. This formula is very useful in the context of multipole expansions, for example in electromagnetism. The n^{th} term in this expansions falls off as $r^{-(n+1)}$, so the $n = 0$ term corresponds to the monopole contribution, the $n = 1$ term to the dipole, the $n = 2$ term to the quadrupole etc.

4.2.1 Associated Legendre polynomials

Closely related to the Legendre polynomials are the *associated Legendre functions* P_l^m, defined by

$$P_l^m(x) = \frac{1}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l\,, \tag{4.41}$$

where $l = 0, 1, \ldots$ and $m = -l, \ldots, l$. Clearly, the Legendre polynomials are obtained for $m = 0$, so $P_l = P_l^0$, and for positive m the associated Legendre functions can be written in terms of Legendre polynomials as

$$P_l^m(x) = (1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x)\,. \tag{4.42}$$

The associated Legendre functions are solutions of the differential equation

$$(1-x^2)y'' - 2xy' + \left(l(l+1) - \frac{m^2}{1-x^2}\right)y = 0\,, \tag{4.43}$$

generalising the Legendre differential equation (4.36) to which it reduces for $m = 0$. A calculation based on Eq. (4.41) and partial integration leads to the orthogonality relations

$$\int_{-1}^{1} dx\, P_l^m(x)P_{l'}^m(x) = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ll'}\,. \tag{4.44}$$

Exercise 4.10. *Show that the associated Legendre polynomials P_l^m solve the differential equation (4.43) and satisfy the orthogonality relations (4.44).*

4.3 The Laguerre polynomials

From Table 4.1.4, the Laguerre polynomials (with $\alpha = 0$) are defined on the interval $[a,b] = [0,\infty]$, the function X is given by $X(x) = x$ and the weight function is $w(x) = e^{-x}$, so that the relevant Hilbert space is $L^2_w([0,\infty])$. They are denoted by L_n and, inserting into Eq. (4.13), their Rodriguez formula is

$$L_n = \frac{1}{n!}e^x\frac{d^n}{dx^n}(e^{-x}x^n)\,, \tag{4.45}$$

where the pre-factor is conventional and implies that $K_n = n!$. It is easy to extract from the Rodriguez formula the coefficients k_n and k'_n of x^n and x^{n-1} in L_n and they are given by

$$k_n = \frac{(-1)^n}{n!} , \qquad k'_n = \frac{(-1)^{n-1} n}{(n-1)!} . \tag{4.46}$$

The normalisation of the L_n is computed from the Rodriguez formula with the usual partial integration trick and it follows

$$\begin{aligned} h_n = \| L_n \|^2 = \int_0^\infty dx\, e^{-x} L_n(x)^2 = \frac{1}{n!} \int_0^\infty d\, L_n(x) \frac{d^n}{dx^n}(e^{-x} x^n) = \frac{(-1)^n}{n!} \int_0^\infty dx\, \frac{d^n L_n}{dx^n}(x) e^{-x} x^n \\ = (-1)^n k_n \int_0^\infty dx\, e^{-x} x^n = 1 . \end{aligned} \tag{4.47}$$

Hence, with our convention, the Laguerre polynomials are already normalised [6]. Functions $f \in L^2_w([0,\infty])$ can now be expanded as

$$f = \sum_{n=0}^\infty \langle L_n, f \rangle L_n \qquad \text{or} \qquad f(x) = \sum_{n=0}^\infty a_n L_n(x) , \qquad a_n = \int_0^\infty dx\, e^{-x} L_n(x) f(x) . \tag{4.48}$$

Expansions in terms of Laguerre polynomials are often useful for functions which depend on a radius r which has a natural range $[0,\infty]$.

Inserting the above results for k_n, k'_n and h_n into Eq. (4.6) gives

$$A_n = -\frac{1}{n+1} , \qquad B_n = \frac{2n+1}{n+1} , \qquad C_n = \frac{n}{n+1} \tag{4.49}$$

and using these values in Eq. (4.5) leads to the recursion relation

$$(n+1) L_{n+1}(x) = (2n+1-x) L_n(x) - n L_{n-1}(x) \tag{4.50}$$

for the Laguerre polynomials. From the Rodriguez formula the first few Laguerre polynomials are given by

$$L_0(x) = 1 , \quad L_1(x) = -x+1 , \quad L_2(x) = \frac{1}{2}(x^2 - 4x + 2) , \quad L_3(x) = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6) . \tag{4.51}$$

Inserting $X = x$, $K_1 = 1$, $k_1 = -1$, $P_1 = 1-x$ and $X'' = 0$ into Eq. (4.20) gives the differential equation

$$xy'' + (1-x)y' + ny = 0 \tag{4.52}$$

for the Laguerre polynomials. For the generating function

$$G(x,z) = \sum_{n=0}^\infty L_n(x) z^n \tag{4.53}$$

we can derive a differential equation in much the same way as we did for the Legendre polynomials, using the recursion relation (4.50). This leads to

$$\frac{\partial G}{\partial z} = \frac{1-x-z}{(1-z)^2} G , \tag{4.54}$$

and the solution is

$$G(x,z) = \exp\left(-\frac{xz}{1-z}\right) \frac{1}{1-z} = \sum_{n=0}^\infty L_n(x) z^n . \tag{4.55}$$

Exercise 4.11. *Derive the differential equation (4.54) for the generating function of the Laguerre polynomials and show that its solution is given by Eq. (4.55).*

[6] Sometimes the L_n are defined without the $n!$ factor in the Rodriguez formula (4.45). For this convention the L_n are of course not normalised to one, but instead $h_n = \| L_n \|^2 = (n!)^2$.

4.4 The Hermite polynomials

From Table 4.1.4, the Hermite polynomials are defined on the interval $[a, b] = [-\infty, \infty] = \mathbb{R}$, we have $X = 1$ and the weight function is $w(x) = e^{-x^2}$, so the relevant Hilbert space is $L^2_w(\mathbb{R})$. They are denoted H_n and, from Eq. (4.3), their Rodriguez formula reads

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \ , \tag{4.56}$$

where the pre-factor is conventional and implies that $K_n = (-1)^n$. Their symmetry properties are $H_n(x) = (-1)^n H_n(-x)$. From this formula it is easy to read off the coefficients k_n and k'_n of the leading and sub-leading monomials x^n and x^{n-1} in H_n as

$$k_n = 2^n \ , \qquad k'_n = 0 \ . \tag{4.57}$$

The normalisation of the Hermite polynomials is computed as before, by using the Rodriguez formula (4.56) combined with partial integration:

$$\begin{aligned} h_n = \| H_n \|^2 = \int_{\mathbb{R}} dx\, e^{-x^2} H_n(x)^2 = (-1)^n \int_{\mathbb{R}} dx\, H_n(x) \frac{d^n}{dx^n} e^{-x^2} = \int_{\mathbb{R}} dx\, \frac{d^n H_n}{dx^n}(x) e^{-x^2} \\ = k_n n! \int_{\mathbb{R}} dx\, e^{-x^2} = \sqrt{\pi} 2^n n! \ . \end{aligned} \tag{4.58}$$

Hence, the ortho-normal basis of $L^2_w(\mathbb{R})$ is given by

$$\hat{H}_n(x) = \frac{1}{\sqrt{\sqrt{\pi} 2^n n!}} H_n(x) \quad \Rightarrow \quad \langle \hat{H}_n, \hat{H}_m \rangle = \delta_{nm} \ , \tag{4.59}$$

and functions $f \in L^2_w(\mathbb{R})$ can be expanded as

$$f = \sum_{n=0}^{\infty} \langle \hat{H}_n, f \rangle \hat{H}_n \ . \tag{4.60}$$

More explicitly and rearranging the coefficients this reads

$$f(x) = \sum_{n=0}^{\infty} a_n H_n(x) \ , \qquad a_n = \frac{1}{\sqrt{\pi} 2^n n!} \int_{\mathbb{R}} dx\, e^{-x^2} H_n(x) f(x) \ . \tag{4.61}$$

The Hermite polynomials are useful for expanding functions defined on the entire real line and they make a prominent appearance in the wave functions for the quantum harmonic oscillator.

From the above results for k_n, k'_n and h_n it is easy, by inserting into Eq. (4.6), to work out

$$A_n = 2 \ , \qquad B_n = 0 \ , \qquad C_n = 2n \ , \tag{4.62}$$

and, from Eq. (4.5), this leads to the recursion relation

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x) \tag{4.63}$$

for the Hermite polynomials. Rodriguez's formula (4.56) can be used to work out the first few Hermite polynomials which are given by

$$H_0(x) = 1 \ , \quad H_1(x) = 2x \ , \quad H_2(x) = 4x^2 - 2 \ , \quad H_3(x) = 8x^3 - 12x \ . \tag{4.64}$$

With $X = 1$, $X'' = 0$, $K_1 = -1$, $H_1(x) = 2x$ and $k_1 = 2$ Eq. (4.20) turns into the differential equation for Hemite polynomials

$$y'' - 2xy' + 2ny = 0 \,. \tag{4.65}$$

The generating function

$$G(x,z) = \sum_{n=0}^{\infty} H_n(x) \frac{z^n}{n!} \tag{4.66}$$

can be derived from the differential equation

$$\frac{\partial G}{\partial z} = 2(x-z)G \,, \tag{4.67}$$

which, as in the case of the Legendre and Laguerre polynomials, follows by differentiating Eq. (4.66) and using the recursion relation (4.63). The solution is

$$G(x,z) = \exp\left(2xz - z^2\right) = \sum_{n=0}^{\infty} H_n(x) \frac{z^n}{n!} \tag{4.68}$$

Exercise 4.12. *Show that the generating function G for the Hermite polynomials satisfies the differential equation (4.67) and verify that it is solved by Eq. (4.68).*

Application 4.17. Expanding in terms of Hermite polynomials

We would like to expand the function $f(x) = \sin(x)$ in terms of Hermite polynomials. Computing the expansion coefficients $b_n := \langle H_n, f\rangle$ for all n seems like a tall order. We can make our life considerably easier by using the generating function $G_z(x) = G(x,z)$ from Eq. (4.68) and define the function

$$Z(z) := \langle G_z, f\rangle \overset{(4.68)}{=} \sum_{n=0}^{\infty} b_n \frac{z^n}{n!} \qquad \Rightarrow \qquad b_n = \left[\frac{d^n Z}{dz^n}\right]_{z=0} .$$

This function Z can be view as a generating function for the coefficients b_n. Using the explicit form of G from Eq. (4.68) we find

$$Z(z) = \int_{\mathbb{R}} dx\, f(x) e^{-(x-z)^2} = \int_{\mathbb{R}} dx\, \sin(x) e^{-(x-z)^2} = \frac{\sqrt{\pi}}{e^{1/4}} \sin(z) \,.$$

Hence, the coefficients b_n are given by

$$b_n = \left[\frac{d^n Z}{dz^n}\right]_{z=0} = \frac{\sqrt{\pi}}{e^{1/4}} \left[\frac{d^n \sin(z)}{dz^n}\right]_{z=0} = \begin{cases} 0 & \text{for} \quad n \text{ even} \\ \frac{\sqrt{\pi}(-1)^{(n-1)/2}}{e^{1/4}} & \text{for} \quad n \text{ odd} \end{cases}$$

5 Ordinary linear differential equations

In this chapter, our focus will be on linear, second order differential equations of the form

$$\left.\begin{array}{rcl} \alpha_2(x)y'' + \alpha_1(x)y' + \alpha_0(x)y &=& f(x) \\ \alpha_2(x)y'' + \alpha_1(x)y' + \alpha_0(x)y &=& 0 \end{array}\right\} \tag{5.1}$$

where α_0, α_1 and α_2 as well as f are given functions. Clearly, the upper equation is inhomogeneous with source f and the lower equation is its homogeneous counterpart. In operator form this can be written as

$$\left.\begin{array}{rcl} Ty &=& f \\ Ty &=& 0 \end{array}\right\} \qquad T = \alpha_2 D^2 + \alpha_1 D + \alpha_0 \ , \tag{5.2}$$

where $D = \frac{d}{dx}$. For the range of x we would like to consider the interval $[a, b] \subset \mathbb{R}$ (where the semi-infinite and infinite case is allowed) and, provided $\alpha_i \in \mathcal{C}^\infty([a, b])$, we can think of T as a linear operator $T : \mathcal{C}^\infty([a, b]) \to \mathcal{C}^\infty([a, b])$. We note that the general differential equations (4.20) for orthogonal polynomials (and, hence, the Legendre, Laguerre and Hermitian differentials equations in Eqs. (4.36), (4.52) and (4.65)) are homogeneous equations of the form (5.1).

The above equations are usually solved subject to additional conditions on the solution y and there are two ways of imposing such conditions. The first one, which leads to what is called an *initial value problem*, is to ask for solutions to either of Eqs. (5.1) which, in addition, satisfy the "initial conditions"

$$y(x_0) = y_0 \ , \qquad y'(x_0) = y'_0 \ , \tag{5.3}$$

for $x_0 \in [a, b]$ and given values $y_0, y'_0 \in \mathbb{R}$. Another possibility, which defines a *boundary value problem*, is to ask for solutions to either of Eqs. (5.1) which satisfy the conditions

$$d_a y(a) + n_a y'(a) = c_a \ , \qquad d_b y(b) + n_b y'(b) = c_b \ , \tag{5.4}$$

where $d_a, d_b, n_a, n_b, c_a, c_b \in \mathbb{R}$ are given constants. In other words, we impose linear conditions on the function at both endpoints of the interval $[a, b]$. If $d_a = d_b = 0$ so these become conditions on the first derivate only they are called *von Neumann boundary conditions*. The opposite case $n_a = n_b = 0$, when the boundary conditions only involve y but not y' are called *Dirichlet boundary conditions*. The general case is referred to as *mixed boundary conditions*. For $c_a = c_b = 0$ the boundary conditions are called homogeneous, otherwise they are called inhomogeneous.

Initial and boundary value problems, although related, are conceptually quite different. In physics, the former are usually considered when the problem involves time evolution (so x corresponds to time) and the initial state of the system needs to be specified at a particular time. Boundary value problems frequently arise in physics when x has the interpretation of a spatial variable, for example the argument of a wave function in quantum mechanics which needs to satisfy certain conditions at the boundary.

In this section, we will mainly be concerned with boundary value problems (initial value problems having been the focus of the first year courses on differential equations). We begin with a quick review of the relevant basic mathematics.

5.1 Basic theory*

5.1.1 Systems of linear first order differential equations

The most basic question which arises for differential equations is about the existence and uniqueness of solutions. To discuss this in the present case, it is useful to consider a somewhat more general problem of a system of first order inhomogeneous and homogeneous differential equations

$$\left.\begin{array}{rcl} \mathbf{y}' &=& A(x)\mathbf{y} + \mathbf{g}(x) \\ \mathbf{y}' &=& A(x)\mathbf{y} \end{array}\right\} \tag{5.5}$$

where the vector $\mathbf{y}(x) = (y_1(x), \ldots, y_n(x))^T$ consists of the n functions we are trying to find, $\mathbf{g}(x) = (g_1(x), \ldots, g_n(x))^T$ is a given vector of functions and A is a given $n \times n$ matrix of functions. For this system, we have the following

Theorem 5.1. *Let $\mathbf{g} = (g_1, \ldots g_n)^T$ be an n-dimensional vector of continuous functions $g_i : [a, b] \to F$ and $A = (A_{ij})$ an $n \times n$ matrix of continuous functions $A_{ij} : [a, b] \to F$ (where $F = \mathbb{R}$ or $F = \mathbb{C}$). For a given $x_0 \in [a, b]$ and any $\mathbf{c} \in F^n$ the inhomogeneous differential equation (5.5) has a unique solution $\mathbf{y} : [a, b] \to F^n$ with $\mathbf{y}(x_0) = \mathbf{c}$.*

Proof. This is a classical statement from the theory of ordinary differential equations. The existence part is also sometimes referred to as the Picard-Lindelöf Theorem. The proof can be found in many books on the subject, for example Ref. [10]. □

In simple terms, the above theorem states that the initial value problem for the differential equation (5.5) always has a solution and that this solution is unique. Next, we focus on the homogeneous equation.

Theorem 5.2. *Let $A = (A_{ij})$ be an $n \times n$ matrix of continuous functions $A_{ij} : [a, b] \to F$ and $\mathbf{y} : [a, b] \to F^n$. Then the set of solutions V_H of the homogeneous differential equation (5.5) is an n-dimensional vector space over F. For k solutions $\mathbf{y}_1, \ldots, \mathbf{y}_k \in V_H$ the following statements are equivalent.*

(i) $\mathbf{y}_1, \ldots \mathbf{y}_k$ are linearly independent in V_H.
(ii) There exists an $x_0 \in [a, b]$ such that $\mathbf{y}_1(x_0), \ldots \mathbf{y}_k(x_0) \in F^n$ are linearly independent.
(iii) The vectors $\mathbf{y}_1(x), \ldots \mathbf{y}_k(x) \in F^n$ are linearly independent for all $x \in [a, b]$.

Proof. The proof follows from simple considerations and Theorem 5.1. □

So the dimension of the homogeneous solution space is n-dimensional and a given set of solutions $\mathbf{y}_1, \ldots \mathbf{y}_n$ forms a basis of V_H iff $\mathbf{y}_1(x), \ldots \mathbf{y}_n(x) \in F^n$ are linearly independent for at least one x. Alternatively, we can say

$$\mathbf{y}_1, \ldots, \mathbf{y}_n \text{ basis of } V_H \quad \Leftrightarrow \quad \det(\mathbf{y}_1(x), \ldots, \mathbf{y}_n(x)) \neq 0 \text{ for at least one } x \in [a, b] \,, \tag{5.6}$$

and this provides a practical way of checking whether a system of solutions forms a basis of the solution space.

If V_I is the set of solutions of the inhomogeneous equation (5.5) it is clear that

$$V_I = \mathbf{y}_0 + V_H \,, \tag{5.7}$$

where $\mathbf{y}_0$ is any solution to (5.5). A special solution of the inhomogeneous equation can be found by a process called *variation of constants* as in the following

Theorem 5.3. *(Variation of constants) If $Y = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ is a basis of V_H then*

$$\mathbf{y}_0(x) = Y(x) \int_{x_0}^{x} dt\, Y(t)^{-1} \mathbf{g}(t) \tag{5.8}$$

is a solution to the inhomogeneous equation (5.5).

Proof. Again, this is standard and can be shown by straightforward calculation. The fact Y constitutes a basis of solutions of the homogeneous equations can be expressed as $Y' = AY$. We have

$$\mathbf{y}_0'(x) = Y'(x) \int_{x_0}^{x} dt\, Y(t)^{-1} \mathbf{g}(t) + Y(x) Y(x)^{-1} \mathbf{g}(x) = A\mathbf{y}_0(x) + \mathbf{g}(x) \,. \tag{5.9}$$

□

5.1.2 Second order linear differential equations

How is the above discussion of first order differential equations relevant to our original problem (5.1) of second order differential equations? The answer is, of course, that higher order differential equations can be converted into systems of first order differential equations. To see this, start with the system (5.1) and define an associated two-dimensional first order system of the form (5.5) given by

$$\mathbf{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} , \quad A = \begin{pmatrix} 0 & 1 \\ -\frac{\alpha_0}{\alpha_2} & -\frac{\alpha_1}{\alpha_2} \end{pmatrix} , \quad \mathbf{g} = \begin{pmatrix} 0 \\ \frac{f}{\alpha_2} \end{pmatrix} . \tag{5.10}$$

(We assume α_2 is non-zero everywhere.) The solutions of this first-order system and the ones of the second order equation (5.1) are then in one-to-one correspondence via the identification $\tilde{y}_1 = y$ and $\tilde{y}_2 = y'$. Given this observation we can now translate the previous statements for first order systems into statements for second order differential equations.

Theorem 5.4. *Let $\alpha_i, f : [a,b] \to F$ be continuous (and α_2 non-zero on $[a,b]$). Then we have the following statements:*

(a) For given $y_0, y_0' \in \mathbb{R}$ and $x_0 \in [a,b]$, the inhomogeneous equation (5.1) has a unique solution $y : [a,b] \to F$ with $y(x_0) = y_0$ and $y'(x_0) = y_0'$.

(b) The solutions $y : [a,b] \to F$ to the homogeneous equation form a two-dimensional vector space V_H over F. Two solutions y_1 and y_2 to the homogeneous equation form a basis of V_H iff the matrix

$$\begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} (x) \tag{5.11}$$

is non-singular for at least one $x \in [a,b]$ or, equivalently, iff the Wronski determinant

$$W := \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} (x) = (y_1 y_2' - y_2 y_1')(x) \tag{5.12}$$

is non-zero for at least one $x \in [a,b]$.

(c) The solution space V_I of the inhomogeneous equation (5.1) is given by $V_I = y_0 + V_H$, where y_0 is any solution to the inhomogeneous equation (5.1).

Proof. All these statements follow directly from the analogoues statements for first order systems in Theorems 5.1 and 5.2 by using the correspondence (5.10). □

The procedure of variation of constants in Theorem 5.3 can also be transferred to second order differential equations and leads to

Theorem 5.5. *(Variation of constants) Let $\alpha_i, g : [a,b] \to F$ be continuous (and α_2 non-zero on $[a,b]$) and $y_1, y_2 : [a,b] \to F$ a basis of solutions for the homogeneous system (5.1). Then, a solution $y : [a,b] \to F$ of the inhomogeneous system is given by*

$$y(x) = \int_{x_0}^{x} dt\, G(x,t) f(t) , \tag{5.13}$$

where G is called the Green function, given by

$$G(x,t) = \frac{y_1(t)y_2(x) - y_1(x)y_2(t)}{\alpha_2(t) W(t)} , \tag{5.14}$$

with the Wronski determinant $W = y_1 y_2' - y_2 y_1'$.

Proof. This follows directly from the above results for first order systems. More specifically, inserting

$$Y = \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}, \qquad Y^{-1} = \frac{1}{W} \begin{pmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{pmatrix} \tag{5.15}$$

together with $\mathbf{g}$ from Eq. (5.10) into Eq. (5.8) gives the result. □

5.1.3 The boundary value problem

We have now collected a number of standard results which imply that the initial value problem defined by Eqs. (5.1) and (5.3) always has a unique solution. Also, we have gained insight into the structure of the total solution space V_H of the homogeneous Eq. (5.1) and we know that this space is a two-dimensional vector space. Further, the space of solutions V_I to the inhomogeneous Equation (5.1) is given by $V_I = \psi + V_H$, where ψ is any solution to the inhomogeneous equation. We have also seen that we can use a basis of solutions in V_H to construct a Green function (5.14) which allows us, via Eq. (5.19), to find a solution to the inhomogeneous equation.

Armed with this information, we should now come back to the boundary value problem defined by Eqs. (5.1) and (5.4). It is quite useful to split this problem up in the following way. Consider first finding a solution y_0 to the problem

$$\alpha_2(x)y_0'' + \alpha_1(x)y_0' + \alpha_0(x)y_0 = 0\,, \qquad d_a y_0(a) + n_a y_0'(a) = c_a\,, \qquad d_b y_0(b) + n_b y_0'(b) = c_b\,, \tag{5.16}$$

that is, to the homogeneous equation with the inhomogeneous boundary conditions. Next, find a solution $\tilde{y}$ to

$$\alpha_2(x)\tilde{y}'' + \alpha_1(x)\tilde{y}' + \alpha_0(x)\tilde{y} = f(x)\,, \qquad d_a\tilde{y}(a) + n_a\tilde{y}'(a) = 0\,, \qquad d_b\tilde{y}(b) + n_b\tilde{y}'(b) = 0\,, \tag{5.17}$$

that is, to the inhomogeneous differential equation with a homogeneous version of the boundary conditions. It is easy to see that, thanks to linearity, the sum $y = y_0 + \tilde{y}$ provides a solution to the general problem, that is, to the inhomogeneous Eq. (5.1) with inhomogeneous boundary conditions (5.4). We can deal with the first problem (5.16) by finding the most general solution to the homogeneous differential equation, that is, determine the solution space V_H, and then build in the boundary condition. We will discuss some practical methods to do this soon but for now, let us assume this has been accomplished and we want to solve the problem (5.17).

The idea is to do this by modifying the variation of constants approach from Theorem (5.5) and construct a Green function which leads to the correct boundary conditions. Let's address this for the case of Dirichlet boundary conditions, so we are considering the problem

$$\alpha_2(x)y'' + \alpha_1(x)y' + \alpha_0(x)y = f(x)\,, \qquad y(a) = 0\,, \qquad y(b) = 0\,. \tag{5.18}$$

Theorem 5.6. *Let $y_1, y_2 : [a,b] \to F$ be a basis of V_H, that is, a basis of solutions to the homogeneous system (5.1), satisfying $y_1(a) = y_2(b) = 0$. Then a solution $y : [a,b] \to F$ to the Dirichlet boundary value problem (5.18) is given by*

$$y(x) = \int_a^b dt\, G(x,t) f(t)\,, \tag{5.19}$$

where the Green function G is given by

$$G(x,t) = \frac{y_1(t)y_2(x)\theta(x-t) + y_1(x)y_2(t)\theta(t-x)}{\alpha_2(t)W(t)}\,. \tag{5.20}$$

Here θ is the Heaviside function defined by $\theta(x) = 1$ for $x \geq 0$ and $\theta(x) = 0$ for $x < 0$.

Proof. The two homogeneous solutions y_1 and y_2 satisfy $Ty_1 = Ty_2 = 0$ with the operator T from Eq. (5.2) and the conditions $y_1(a) = y_2(b) = 0$ can always be imposed since we know there exists a solution for any choice of initial condition. Now we start with a typical variation of constant Ansatz

$$y(x) = u_1(x)y_1(x) + u_2(x)y_2(x) \,, \tag{5.21}$$

where u_1, u_2 are two functions to be determined. If we impose on those two functions the condition

$$u_1'y_1 + u_2'y_2 = 0 \tag{5.22}$$

an easy calculation shows that

$$Ty = \alpha_2(u_1'y_1' + u_2'y_2') \stackrel{!}{=} f \,. \tag{5.23}$$

Solving Eqs. (5.22) and (5.23) for u_1 and u_2 leads to

$$u_1(x) = -\int_{x_1}^{x} dt\, \frac{y_2(t)f(t)}{\alpha_2(t)W(t)} \,, \qquad u_2(x) = \int_{x_2}^{x} dt\, \frac{y_1(t)f(t)}{\alpha_2(t)W(t)} \,, \tag{5.24}$$

where $x_1, x_2 \in [a, b]$ are two otherwise arbitrary constants. To implement the boundary conditions $y(a) = y(b) = 0$ it suffices to demand that $u_1(b) = u_2(a) = 0$ (given our assumptions about the boundary values of y_1 and y_2) and this is guaranteed by choosing $x_1 = b$ and $x_2 = a$. Inserting these values into Eq. (5.24) and the expressions for u_i back into the Ansatz (5.21) gives the desired result. □

Whether Eq. (5.19) is the unique solutions to the boundary value problem (5.18) depends on whether there is a non-trivial solution to the homogeneous equations in V_H which satisfies the relevant boundary conditions $y(a) = y(b) = 0$. If there is it can be added to (5.19) and the solution is not unique, otherwise it is. More generally, going back to the way we have split up the problem into two steps in Eqs. (5.16) and (5.17), we have now found a method to find a solutions to the second problem (5.17) (the inhomogeneous equation with the homogeneous boundary conditions) for the Dirichlet case. Any solution to the first problem (5.16) (the homogeneous equation with inhomogeneous boundary conditions) can be added to this.

5.1.4 Solving the homogeneous equation

It remains to discuss methods for how to solve the homogeneous equation (5.1). If one solution to this equation is known a second, linearly independent solution is obtained from the following

Theorem 5.7. *(Reduction of order) Let $y : [a, b] \to F$ be a solution of $Ty = 0$ with T given in Eq. (5.2) and $I \subset [a, b]$ be an interval for which y and α_2 are everywhere non-vanishing. Then $\tilde{y} : I \to F$ defined by*

$$\tilde{y}(x) = y(x)u(x) \,, \qquad u'(x) = \frac{1}{y(x)^2} \exp\left(-\int_{x_0}^{x} dt\, \frac{\alpha_1(t)}{\alpha_2(t)}\right) \tag{5.25}$$

satisfies $T\tilde{y} = 0$ and is linearly independent from y.

Proof. An easy calculation shows that the function u, defined above, satisfies the differential equation

$$u'' + \left(2\frac{y'}{y} + \frac{\alpha_1}{\alpha_2}\right)u' = 0 \,. \tag{5.26}$$

With $\tilde{y}' = yu' + y'u$ and $\tilde{y}'' = yu'' + 2y'u' + y''u = -\frac{\alpha_1}{\alpha_2}yu' + y''u$ it is easy to show that $T\tilde{y} = uTy$ which vanishes since $Ty = 0$.

The Wronski determinant of the two solutions y and $\tilde{y}$ is $W = y\tilde{y}' - y'\tilde{y} = y(yu' + y'u) - y'yu = y^2u'$ and this is non-zero since the last expression is precisely the exponential in Eq. (5.25). Hence, the two solutions are independent. □

Obtaining a second independent solution from a known one can be useful but how can we find a solution in the first place? A very common and efficient method is to start with a power series Ansatz

$$y(x) = \sum_{k=0}^{\infty} a_k x^k \ . \tag{5.27}$$

Of course, this is only practical if the functions α_i which appear in T are polynomial. In this case, the idea is to insert the Ansatz (5.27) into $Ty = 0$, assemble the coefficient in front of x^k and set this coefficient to zero for every k. In this way, one obtains a recursion relation for the a_k and inserting the resulting a_k back into Eq (5.27) gives the solution in terms of a power series. Of course this is where the difficult work starts. Now one has to understand the properties of the so-obtained series, such as convergence, singularities or asymptotic behaviour. All this is best demonstrated for examples and we will do this shortly.

5.2 Examples

We would now like to apply some of the methods and results from the previous subsection to examples.

Application 5.18. Solving the Legendre differential equation

Recall from Eq. (4.36) the Legendre differential equation

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0 \ . \tag{5.28}$$

Of course we know that the Legendre polynomials are solutions but we would like to derive this independently (as well as finding the second solution which must exist) by using the power series method. Inserting the series (5.27) into the Legendre differential equation gives (after re-defining some of the summation indices)

$$\sum_{k=0}^{\infty} \left[(k+2)(k+1)a_{k+2} - (k(k+1) - n(n+1))a_k\right] x^k = 0 \ . \tag{5.29}$$

Demanding that the coefficient in front of every monomial x^k vanishes (then and only then is a power series identical to zero) we obtain the recursion formula

$$a_{k+2} = \frac{k(k+1) - n(n+1)}{(k+1)(k+2)} a_k \ , \qquad k = 0, 1, \dots \ , \tag{5.30}$$

for the coefficients a_k. There are a number of interesting features of this formula. First, the coefficients a_0 and a_1 are not fixed but once values have been chosen for them the above recursion formula determines all other a_k. This freedom of choosing two coefficients precisely corresponds to the two independent solutions we expect. The second interesting feature is that, due to the structure of the numerator in Eq. (5.30), $a_k = 0$ for $k = n+2, n+4, \dots$.

To see what happens in more detail let's first assume that n is even. Choose $(a_0, a_1) = (1, 0)$. In this case all a_k with k odd vanish and the a_k with k even are non-zero only for $k \leq n$. This means, the series breaks down and turns into a polynomial - this is of course (propertional to) the Legendre polynomial P_n for n even. Still for n even, make the complementary choice $(a_0, a_1) = (0, 1)$. In this case all the a_k with k even are zero. However, for k odd and n even the numerator in Eq. (5.30) never vanishes so this leads to an infinite series which only contains odd powers of x. This is the second solution, in addition to the Legendre polynomials. For n odd the situation is of course similar but

reversed. For $(a_0, a_1) = (0, 1)$ we get polynomials - the Legendre polynomial P_n for n odd - while for $(a_0, a_1) = (1, 0)$ we get the second solution, an infinite series with only even powers of x.

Exercise 5.8. *Show explicitly that, for suitable choices of a_0 and a_1, the recursion formula (5.30) reproduces the first few Legendre polynomials in Eq. (4.35).*

Of course the differential equation (5.28) also makes sense if n is a real number, rather than an integer, and the above calculation leading to the coefficients a_k remains valid in this case. However, if $n \notin \mathbb{N}$, the numerator in Eq. (5.30) never vanishes and both solutions to (5.28) are non-polynomial.

To find the second solution we can also use the reduction of order method from Theorem 5.7. To demonstrate how this works we focus on the case $n = 1$ with differential equation

$$(1 - x^2)y'' - 2xy' + 2y = 0 \,. \tag{5.31}$$

which is solved by the Legendre polynomial $y(x) = P_1(x) = x$. Inserting this, together with $\alpha_1(x) = -2x$ and $\alpha_2(x) = 1 - x^2$ into Eq. (5.25) gives

$$u'(x) = \frac{1}{x^2} \exp\left(\int^x dt \, \frac{2t}{1 - t^2} \right) = \frac{1}{x^2(1 - x^2)} \tag{5.32}$$

A further integration leads to

$$u(x) = -\frac{1}{x} + \frac{1}{2} \ln \frac{1 + x}{1 - x} \tag{5.33}$$

so the second solution to the Legendre equation (5.31) for $n = 1$ is

$$\tilde{y}(x) = xu(x) = \frac{x}{2} \ln \frac{1 + x}{1 - x} - 1 \,. \tag{5.34}$$

Exercise 5.9. *Find the Taylor series of the solution (5.34) around $x = 0$ and show that the coefficients in this series are consistent with the recursion formula (5.30).*

Application 5.19. Solving the Hermite differential equation

Recall from Eq. (4.65) that the Hermite differential equation is given by

$$y'' - 2xy' + 2ny = 0 \,. \tag{5.35}$$

To find its solutions we can proceed like we did in the Legendre case and insert the series (5.27). This leads to

$$\sum_{k=0}^{\infty} \left[(k + 1)(k + 2)a_{k+2} - 2(k - n)a_k \right] x^k = 0 \,, \tag{5.36}$$

and, hence, the recursion relation

$$a_{k+2} = \frac{2(k - n)}{(k + 1)(k + 2)} a_k \,. \tag{5.37}$$

As before, we have a free choice of a_0 and a_1 but with those two coefficients fixed the recursion formula determines all others. From the numerator in Eq. (5.37) it is clear that $a_k = 0$ for $k = n+2, n+4, \ldots$.

For n even and $(a_0, a_1) = (1, 0)$ we get a polynomial with only even powers of x - up to an overall constant the Hermite Polynomial H_n with n even - while $(a_0, a_1) = (0, 1)$ leads to an infinite series with only odd powers of x - the second solution of (5.28). For n odd the choice $(a_0, a_1) = (0, 1)$ leads to a polynomial solution with only odd powers of x which is proportional to the Hermite polynomials H_n for n odd, while the choice $(a_0, a_1) = (1, 0)$ leads to a power series with only even powers of x.

Exercise 5.10. *Show that, for appropriate choices of a_0 and a_1 the recursion formula reproduces the first few Hermite polynomials (4.64).*

As in the Legendre case, the differential equation (5.35) also makes sense if n is a real number. If $n \notin \mathbb{N}$ then the numerator in Eq. (5.37) never vanishes and both solutions to (5.35) are non-polynomial. (This observation plays a role for the energy quantisation of the quantum harmonic oscillator.)

Of course the above discussion can be repeated for the Laguerre differential equation (4.52) as in the following

Exercise 5.11. *Insert the series Ansatz (5.27) into the Laguerre differential equation (4.52) and find the recursion relation for the coefficients a_k. Discuss the result and identify the choices which lead to the Laguerre polynomials.*

Application 5.20. A simple inhomogeneous example

For a simple inhomogeneous case, let us consider the equation

$$Ty = f \ , \qquad T = \frac{d^2}{dx^2} + 1 \tag{5.38}$$

on the interval $[a, b] = [0, \frac{\pi}{2}]$, where f is an arbitrary function. (This describes a driven harmonic oscillator with driving force f.) It is clear that the solution space of the associated homogeneous equation, $Ty = 0$, is given by

$$V_H = \text{Span}\left(y_1(x) = \sin(x), y_2(x) = \cos(x)\right) \ . \tag{5.39}$$

As a sanity check we can work out the Wronski determinant

$$W = y_1 y_2' - y_2 y_1' = -1 \ , \tag{5.40}$$

and since this is non-vanishing the two solutions are indeed linearly independent. To find the solution space of the inhomogeneous equation we can use the variation of constant method from Theorem 5.5. Inserting $y_1 = \sin$, $y_2 = \cos$, $W = -1$ and $\alpha_2 = 1$ into Eq. (5.14) we find for the Green function

$$G(x, t) = \sin(x - t) \ . \tag{5.41}$$

From Eq. (5.19) this means a special solution to the inhomogeneous equation is given by

$$y_0(x) = \int_{x_0}^{x} dt \, G(x, t) f(t) = \int_{x_0}^{x} dt \, \sin(x - t) f(t) \ , \tag{5.42}$$

and, hence, the solution space of the inhomogeneous equation is

$$V_I = y_0 + V_H \ . \tag{5.43}$$

Exercise 5.12. *Check explicitly that y_0 from Eq. (5.42) satisfies the equation $Ty_0 = f$.*

Let us now consider Eq. (5.38) as a boundary value problem on the interval $[a, b] = [0, \frac{\pi}{2}]$ with Dirichlet boundary conditions $y(0) = y(\pi/2) = 0$ and apply the results of Theorem 5.6. First, we note that $y_1(0) = y_2(\pi/2) = 0$ so our chosen homogeneous solutions do indeed satisfy the requirements of the Theorem. Inserting $y_1 = \sin$, $y_2 = \cos$, $W = -1$ and $\alpha_2 = 1$ into Eq. (5.20) gives the Green function

$$G(x,t) = -\sin(t)\cos(x)\,\theta(x-t) - \sin(x)\cos(t)\,\theta(t-x)\ , \tag{5.44}$$

and hence

$$y(x) = \int_0^{\pi/2} dt\, G(x,t) f(t) \tag{5.45}$$

satisfies $Ty = f$ as well as the correct boundary conditions $y(0) = y(\pi/2) = 0$. We note that there is no non-trivial solution in V_H which satisfies $y(0) = y(\pi/2) = 0$ so Eq. (5.45) is the unique solution to the boundary value problem.

5.3 Bessel differential equation

Another important differential equation in mathematical physics is the *Bessel differential equation* whose solutions are referred to as *Bessel functions*. Before we discuss this in detail we need a bit of preparation and introduce the *Gamma function*, Γ, which is another much-used special function.

5.3.1 The Gamma function

The Gamma function is defined by the integral

$$\Gamma(x) = \int_0^\infty dt\, e^{-t} t^{x-1}\ , \tag{5.46}$$

which is certainly well-defined as long as $x > 0$. A short calculation, using integration by parts and noting that the boundary term vanishes, gives

$$x\Gamma(x) = \int_0^\infty dt\, e^{-t}(xt^{x-1}) = \left[e^{-t}t^x\right]_0^\infty + \int_0^\infty dt\, e^{-t}t^x = \Gamma(x+1)\ , \tag{5.47}$$

and, hence, the functional equation

$$\Gamma(x+1) = x\Gamma(x) \tag{5.48}$$

of the Gamma function. By direct integration it follows that $\Gamma(1) = 1$ and combining this with iterating the above functional equation we learn that

$$\Gamma(n) = (n-1)!\ , \tag{5.49}$$

for $n \in \mathbb{N}$. Hence, the Gamma function can be seen as a function which extends the factorial operation to non-integer numbers. From Eq. (5.48) it also follows by iteration that

$$\Gamma(x) = \frac{\Gamma(x+n)}{x(x+1)\cdots(x+n-1)}\ , \tag{5.50}$$

which shows that the Γ function has poles at $x = 0, -1, -2, \ldots$. These are, in fact, the only poles of the Gamma function.

Application 5.21. Asymptotic expression for the Γ-function

As an aside, let us derive an asymptotic expression for the Γ-function. We start with the substitution $t = (x-1)s$ in the defining integral (5.46), which leads to

$$\begin{aligned}\Gamma(x) &= (x-1)^x \int_0^\infty ds\, e^{(x-1)(\ln(s)-s)} \simeq (x-1)^x \int_0^\infty ds\, e^{(x-1)(-1-(s-1)^2/2)+\cdots} \\ &\overset{u=s-1}{=} (x-1)^x e^{-(x-1)} \int_{-1}^\infty du\, e^{-(x-1)u^2/2} \simeq (x-1)^x e^{-(x-1)} \int_{-\infty}^\infty du\, e^{-(x-1)u^2/2} \\ &= \sqrt{2\pi(x-1)} \left(\frac{x-1}{e}\right)^{x-1} \end{aligned} \tag{5.51}$$

Exercise 5.13. *What happened in the second step and the second last step of the calculation in Eq. (5.51)? Which approximations have been made and how are they justified?*

The approximate result (5.51) for the Γ-function leads to the famous *Stirling formula*

$$n! = \Gamma(n+1) \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \tag{5.52}$$

which provides an asymptotic approximation to the value of $n!$.

Much more can be said about the Gamma function and its natural habitat is in complex analysis. For our purposes the above is sufficient but more information can be found, for example, in Ref. [11].

5.3.2 Bessel differential equation and its solutions

We now turn to our main interest, the Bessel differential equation, which reads

$$x^2 y'' + xy' + (x^2 - \nu^2) y = 0 \,, \tag{5.53}$$

for any number $\nu \in \mathbb{R}^{\geq 0}$. Inserting the Ansatz

$$y(x) = \sum_{k=0}^\infty a_k x^{k+\alpha} \,, \qquad \alpha = \pm\nu \,, \tag{5.54}$$

into the differential equation (putting the additional factor x^α into the Ansatz proves useful to improve the properties of the resulting series) leads to

$$\sum_{k=0}^\infty k(k+2\alpha) a_k x^k + \sum_{k=0}^\infty a_k x^{k+2} = 0 \,. \tag{5.55}$$

There is only a single term proportional to x and to remove this term so we need to set $a_1 = 0$. Since the recursion formula will imply that a_{k+2} is proportional to a_k this means all a_k with k odd must vanish. For the coefficients with even k Eq. (5.55) gives

$$a_{2k} = -\frac{1}{4k(k+\alpha)} a_{2k-2} \,, \qquad k = 1, 2, \ldots \,. \tag{5.56}$$

This recursion formula can be iterated and leads to

$$a_{2k} = \frac{(-1)^k \Gamma(\alpha+1)}{2^{2k} k! \Gamma(k+\alpha+1)} a_0 \,. \tag{5.57}$$

That this result for a_{2k} does indeed satisfy the recursion relation (5.56) follows directly from the functional equation (5.48) of the Gamma function. It is conventional to choose $a_0 = (2^\alpha \Gamma(\alpha+1))^{-1}$ and, by inserting everything back into the Ansatz (5.54), this leads to the two series solutions for $\alpha = \pm\nu$ given by

$$J_{\pm\nu}(x) := \left(\frac{x}{2}\right)^{\pm\nu} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(k \pm \nu + 1)} \left(\frac{x}{2}\right)^{2k} . \tag{5.58}$$

Both J_ν and $J_{-\nu}$ are solutions of the Bessel differential equation and they are called *Bessel functions of the first kind*. It can be shown (for example by applying the quotient criterion) that the series in Eq (5.58) converges for all x. There is a subtlety for J_{-n} if $n = 0, 1, \ldots$. In this case, the terms in the series for $k = 0, \ldots, n-1$ have a Gamma-function pole in the denominator (see Eq. (5.50)) and are effectively removed so that the sum starts at $k = n$. Using this fact, it follows from the above series that

$$J_{-n}(x) = (-1)^n J_n(x) , \qquad n = 0, 1, 2, \ldots . \tag{5.59}$$

In other words, if $\nu = n$ is an integer then the two solutions J_n and J_{-n} are linearly dependent. If ν is not an integer the Wronski determinant at $x \to 0$ has the leading term $W = J_\nu(x)J'_{-\nu}(x) - J'_\nu(x)J_{-\nu}(x) = -\frac{4\nu}{\Gamma(-\nu+1)\Gamma(\nu+1)x}(1 + \mathcal{O}(x)) \neq 0$ which shows that the solutions J_ν and $J_{-\nu}$ are linearly independent. To overcome the somewhat awkward distinction between the integer and non-integer case it is customary to define the *Bessel functions of the second kind* by

$$N_\nu(x) = \frac{J_\nu(x)\cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)} . \tag{5.60}$$

This definition has an apparent pole if ν is integer but it can be shown that it is well-defined in the limit $\nu \to n$. We summarise some of the properties of Bessel functions in the following

Proposition 5.1. *(Properties of Bessel functions) The Bessel function of the first and second kind, J_ν and N_ν, defined above, solve the Bessel differential equation and are linearly independent for all $\nu \in \mathbb{R}$. For their asymptotic properties we have*

$$x \to 0 : \qquad J_\nu(x) \to \frac{1}{\Gamma(\nu+1)}\left(\frac{x}{2}\right)^\nu \qquad N_\nu \to \begin{cases} \frac{2}{\pi}(\ln(x/2) + C) & \text{for } \nu = 0 \\ -\frac{\Gamma(\nu)}{\pi}\left(\frac{2}{x}\right)^\nu & \text{for } \nu > 0 \end{cases} \tag{5.61}$$

$$x \to \infty : \qquad J_\nu(x) \to \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \qquad N_\nu(x) \to \sqrt{\frac{2}{\pi x}}\sin\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \tag{5.62}$$

Proof. The fact that J_ν and N_ν solve the Bessel differential equation is clear from the above calculation. Their linear independence has been shows for ν non-integer and the integer case can be dealt with by a careful consideration of the definition (5.60) as $\nu \to n$. (See, for example, Ref. [12]).

The asymptotic limit as $x \to 0$ can be directly read off from the series (except for $\nu = 0$ which requires a bit more care). The proofs for the asymptotic limits $x \to \infty$ are a bit more involved (see, for example, in Ref. [12]). Intuitively, for large x, we should only keep the terms proportional to x^2 in the Bessel differential equation (5.53) which leads to $y'' + y \simeq 0$. This is clearly solved by sin and cos. □

Exercise 5.14. *Show that $J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin(x)$ and $N_{1/2}(x) = -\sqrt{\frac{2}{\pi x}}\cos(x)$.*

Particularly the above limits for large x are interesting. They show that the Bessel functions have oscillatory properties close to sin and cos but, unlike these, they are not periodic. This is illustrated in Fig. 13. In particular, the asymptotic limits for large x show that the Bessel functions have an infinite number of zeros but they are not equally spaced (although they are asymptotically equally spaced). We denote the

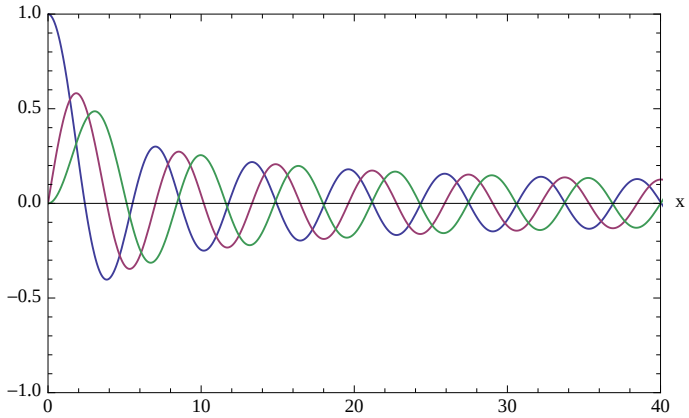

Figure 13: The graph of the Bessel functions J_0, J_1 and J_2.

zeros of J_ν by $z_{\nu k}$ where $k = 1, 2, \ldots$ labels the zeros from small to large x. Their values can be computed numerically and some examples for J_0, J_1 and J_2 are:

$$\begin{aligned} z_{0k} &= 2.405,\ 5.520,\ 8.654,\ \ldots \\ z_{1k} &= 3.832,\ 7.016,\ 10.173,\ \ldots \\ z_{2k} &= 5.136,\ 8.417,\ 11.620,\ \ldots \end{aligned}$$

5.3.3 Orthogonal systems of Bessel functions

Can we use Bessel functions to construct orthogonal systems of functions on some appropriate Hilbert space, analogous to what we have done for sin and cos? The singularity for the Bessel functions of the second kind at $x = 0$ makes them less suitable but the Bessel functions J_ν of the first kind are everywhere well-behaved. Asymptotically, they approach a function of the form $\sin(x - \text{const})/\sqrt{x}$ so it seems natural to try something analogous to the sine Fourier series, where we have used functions proportional to $\sin\left(\frac{k\pi x}{a}\right)$ to construct on ortho-normal basis on $L^2([0, a])$. With this motivation we focus on a particular Bessel function J_ν for fixed ν and define functions on the interval $x \in [0, a]$ by

$$\hat{J}_{\nu k}(x) := N_{\nu k} J_\nu\left(\frac{z_{\nu k} x}{a}\right) , \tag{5.63}$$

where $k = 1, 2, \ldots$ and normalisation constant $N_{\nu k}$. Motivated by the above discussion, the zeros $k\pi$ of the sine which featured in the definition of the sine Fourier series have been replaced by the zeros $z_{\nu k}$ of the Bessel function J_ν.

Lemma 5.1. *The functions $\hat{J}_{\nu k}$ for $k = 1, 2, \ldots$, defined in Eq. (5.63) form an orthogonal system of functions on $L^2_w([0, a])$, where $w(x) = x$, which is ortho-normal for suitable choices of the constans $N_{\nu k}$.*

Proof. A calculation using the definition (5.63) and the Bessel differential equations shows that the $\hat{J}_{\nu k}$ satisfy

$$T\hat{J}_{\nu k} = -\frac{z_{\nu k}^2}{a^2}\hat{J}_{\nu k} , \qquad T = \frac{1}{x}\frac{d}{dx}\left(x\frac{d}{dx}\right) - \frac{\nu^2}{x^2} . \tag{5.64}$$

Hence, the $\hat{J}_{\nu k}$ are eigenvectors of the operator T with eigenvalues $-z_{\nu k}^2/a^2$. On the space $L^2_{w,\text{per}}([0, a])$ of functions with $f(0) = f(a)$ and weight function $w(x) = x$ the operator T is hermitian and $\hat{J}_{\nu k} \in L^2_{w,\text{per}}([0, a])$ (since they vanish at $x = 0, a$). Since eigenvectors of a hermitian operator which correspond to different eigenvalues are orthogonal this must be the case for the $\hat{J}_{\nu k}$ since all the zeros $z_{\nu k}$ are different. □

In fact, we have the following stronger statement.

Theorem 5.15. *For $\nu > -1$, the functions $\hat{J}_{\nu k}$ for $k = 1, 2, \ldots$, defined in Eq. (5.63), with suitable choices for $N_{\nu k}$, form an ortho-normal basis of $L^2_w([0, a])$, where $w(x) = x$.*

Proof. The direct proof is technical and can be found in Ref. [7]. In the next subsection, we will see an independent argument. □

The theorem implies that every function $f \in L^2_w([0, a])$ can be expanded in terms of Bessel functions as

$$f = \sum_{k=1}^{\infty} a_k \hat{J}_{\nu k} \ , \qquad a_k = \langle \hat{J}_{\nu k}, f \rangle = \int_0^a dx\, x \hat{J}_{\nu k}(x) f(x) \ . \tag{5.65}$$

5.4 The operator perspective - Sturm-Liouville operators

So far we have discussed second order linear differentials equations in a somewhat down-to-earth way, using methods of basic analysis. We would now like to make contact with functional analysis and our earlier discussion of Hilbert spaces.

5.4.1 Sturm-Liouville operators

A second order differential operator of the form

$$T_{\text{SL}} = \frac{1}{w(x)} \left[\frac{d}{dx} \left(p(x) \frac{d}{dx} \right) + q(x) \right] \tag{5.66}$$

with (real-valued) smooth functions w, p and q is called a *Sturm-Liouville operator*. For now, we would like to think of this as an operator on the space

$$\mathcal{L}([a, b]) := L^2_w([a, b]) \cap \mathcal{C}^\infty([a, b]) \tag{5.67}$$

on the space square integrable functions, relative to a weight function w, on an interval $[a, b]$ which are also infinitely many times differentiable. Accordingly, we should demand that w, p and q are smooth functions and that w, as a weight function, is strictly positive.

Lemma 5.2. *Consider a linear second order differential operator of the form*

$$T = \alpha_2(x) \frac{d^2}{dx^2} + \alpha_1(x) \frac{d}{dx} + \alpha_0(x) \ , \tag{5.68}$$

where $x \in [a, b]$ and $I \subset [a, b]$ is an interval such that $\alpha_2(x) \neq 0$ for all $x \in I$. Then, on I, the operator T can be written in Sturm-Liouville form (5.66) with

$$p(x) = \exp\left(\int_{x_0}^{x} dt\, \frac{\alpha_1(t)}{\alpha_2(t)} \right) \ , \qquad w(x) = \frac{p(x)}{\alpha_2(x)} \ , \qquad q(x) = \alpha_0(x) w(x) \ , \tag{5.69}$$

where $x_0 \in I$.

Proof. Abbreviating $D = \frac{d}{dx}$ and noting that $p' = \frac{\alpha_1}{\alpha_2} p$ we obtain, by inserting into the Sturm-Liouville operator,

$$T_{\text{SL}} = \frac{p}{w} D^2 + \frac{p'}{w} D + \frac{q}{w} = \alpha_2 D^2 + \frac{\alpha_1 p}{\alpha_2 w} D + \alpha_0 = \alpha_2 D^2 + \alpha_1 D + \alpha_0 = T \ . \tag{5.70}$$

□

Introducing the interval I in the above theorem is to avoid an undefined integrand in the first Eq. (5.69) due to the vanishing of α_2. Even when this happens (such as, for example, for the Legendre differential equation at $x = \pm 1$) and the interval I is, at first, chosen to be genuinely smaller than $[a, b]$ it turns out the final result for w, p and q can often be extended to and is well-defined on $[a, b]$.

An obvious question is whether T_{SL} is self-adjoint as an operator on the space $\mathcal{L}([a, b])$, relative to the standard inner product

$$\langle f, g\rangle = \int_a^b dx\, w(x) f(x) g(x) \,, \tag{5.71}$$

with weight function w. A quick calculation shows that

$$\begin{aligned}\langle f, T_{\mathrm{SL}} g\rangle &= \int_a^b dx\, (fD(pDg) + fqg) = [pfDg]_a^b - \int_a^b dx\, (pDfDg - qfg) \\ &= [pfDg]_a^b - [pgDf]_a^b + \int_a^b dx\, (D(pDf)g + qfg) = \left[pfg' - pgf'\right]_a^b + \langle T_{\mathrm{SL}} f, g\rangle \,. \end{aligned} \tag{5.72}$$

So T_{SL} is superficially self-adjoint but we have to ensure that the boundary terms on the RHS vanish. There are two obvious ways in which this can be achieved. First, the interval $[a, b]$ might be chosen such that $p(a) = p(b) = 0$ - this is also called the *natural choice of the interval*. In this case, the boundary term vanishes without any additional condition on the functions f, g and T_{SL} is self-adjoint on $\mathcal{L}([a, b])$. If this doesn't work we can consider the subspace

$$\mathcal{L}_b([a, b]) := \{f \in \mathcal{L}([a, b]) \,|\, d_a f(a) + n_a f'(a) = d_b f(b) + n_b f'(b) = 0\} \,, \tag{5.73}$$

of smooth functions which satisfy mixed homogeneous boundary conditions at a and b. For such functions the above boundary term also vanishes. If $p(a) = p(b)$ the boundary term also vanishes for periodic functions

$$\mathcal{L}_p([a, b]) := \{f \in \mathcal{L}([a, b]) \,|\, f(a) = f(b) \,,\ f'(a) = f'(b)\} \,. \tag{5.74}$$

Hence, we have

Lemma 5.3. *Let T_{SL} be a Sturm-Liouville operator (5.66). If $p(a) = p(b) = 0$ then T_{SL} is self-adjoint as on operator on the space $\mathcal{L}([a, b])$ in Eq. (9.112). It is also self-adjoint on the space of functions $\mathcal{L}_b([a, b])$ with mixed homogeneous boundary in Eq. (5.73). If $p(a) = p(b)$ it is self-adjoint on the space $\mathcal{L}_p([a, b])$ of periodic functions in Eq. (5.74).*

To simplify the notation, we will refer to the space on which the Sturm-Liouville operator is defined and self-adjoint as $\mathcal{L}_{\mathrm{SL}}([a, b])$. From the previous Lemma, this can be $\mathcal{L}([a, b])$, $\mathcal{L}_b([a, b])$ or $\mathcal{L}_p([a, b])$, depending on the case.

5.4.2 Sturm-Liouville eigenvalue problem

It is now interesting to consider a *Sturm-Liouville eigenvalue problem*, that is, to consider the eigenvalue equation

$$T_{\mathrm{SL}} y = \lambda y \,, \tag{5.75}$$

on $\mathcal{L}_{\mathrm{SL}}([a, b])$. Since T_{SL} is hermitian, we already know from general arguments (see Theorem 1.24) that the eigenvalues λ must be real and that eigenvectors for different eigenvalues must be orthogonal. Since all the second order differential equations discussed so far can be phrased as a Sturm-Liouville eigenvalue problem (see Table 2) this provides a uniform reason for the appearance of the various orthogonal function systems we have encountered. It is tempting to go further and try to use Theorem 2.9 to argue that orthogonal systems of eigenfunctions of Sturm-Liouville operators must, in fact, form an ortho-normal basis. Arguing

name	DEQ	p	q	w	$\mathcal{L}_{\rm SL}[a,b]$	bound. cond.	λ	y
sine Fourier	$y'' = \lambda y$	1	0	1	$\mathcal{L}_b([0,a])$	$y(0) = y(\pi) = 0$	$-\frac{\pi^2 k^2}{a^2}$	$\sin\left(\frac{k\pi x}{a}\right)$
cosine Fourier	$y'' = \lambda y$	1	0	1	$\mathcal{L}_b([0,a])$	$y'(0) = y'(\pi) = 0$	$-\frac{\pi^2 k^2}{a^2}$	$\cos\left(\frac{k\pi x}{a}\right)$
Fourier	$y'' = \lambda y$	1	0	1	$\mathcal{L}_p([-a,a])$	periodic	$-\frac{\pi^2 k^2}{a^2}$ $-\frac{\pi^2 k^2}{a^2}$	$\sin\left(\frac{k\pi x}{a}\right)$ $\cos\left(\frac{k\pi x}{a}\right)$
Legendre	$(1-x^2)y'' - 2xy' = \lambda y$	$1-x^2$	0	1	$\mathcal{L}([-1,1])$	natural	$-n(n+1)$	P_n
Laguerre	$xy'' + (1-x)y' = \lambda y$	xe^{-x}	0	e^{-x}	$\mathcal{L}([0,\infty])$	natural	$-n$	L_n
Hermite	$y'' - 2xy' = \lambda y$	e^{-x^2}	0	e^{-x^2}	$\mathcal{L}([-\infty,\infty])$	natural	$-2n$	H_n
Bessel	$y'' + \frac{1}{x}y' - \frac{\nu^2}{x^2}y = \lambda y$	x	$-\frac{\nu^2}{x^2}$	x	$\mathcal{L}_b([0,a])$	$y(0) = y(a) = 0$	$-\frac{z_{\nu k}^2}{a^2}$	$\hat{J}_{\nu k}$

Table 2: *The second order differential equations discussed so far and their formulation as a Sturm-Liouville eigenvalue problem.*

in this way would be incorrect for two reasons. First, so far the Sturm-Liouville operator is only defined on the space $\mathcal{L}_{\rm SL}$ which consists of certain smooth functions. While this space may well be dense in the appropriate L^2 Hilbert space it is not a Hilbert space itself. Secondly, Theorem 2.9 applies to compact operators and we know from Exercise 1.13 that differential operators are not bounded and, hence, not compact.

One way to make progress is to convert the Sturm-Liouville differential operator into an integral operator. Some of the hard work has already been done in Theorem 5.6 where we have shown that, provided $\mathrm{Ker}(T_{\rm SL}) = \{\mathbf{0}\}$ we know (for Dirichlet boundary conditions) that

$$T_{\rm SL} y = f \quad \Longleftrightarrow \quad y = \hat{G} f \ , \quad \hat{G} f(x) := \int_a^b dt\, G(x,t) f(t) \ , \tag{5.76}$$

where G is the Green function. The integral operator $\hat{G}$, defined in terms of the Green function kernel G, can be thought of as the inverse of the Sturm-Liouville operator and, as an integral operator, we can extend it to act on the space $L^2_w([a,b])$ (with appropriate boundary conditions). Moreover, we have

Lemma 5.4. *If* $\mathrm{Ker}(T_{\rm SL}) = \{\mathbf{0}\}$ *then the operator* $\hat{G}$ *in Eq.* (5.76) *is self-adjoint and compact on* $L^2_w([a,b])$ *(with Dirichlet boundary conditions).*

Proof. The proof can, for example, be found in Ref. [5]. □

If we set $f = \lambda y$ in Eq. (5.76) we get

$$T_{\rm SL} y = \lambda y \quad \Longleftrightarrow \quad \hat{G} y = \frac{1}{\lambda} y \ . \tag{5.77}$$

This means, the Sturm-Liouvllle eigenvalue problem is converted into an eigenvalue problem for $\hat{G}$ which is formulated in terms of an integral equation, also called a *Volterra integral equation*. The eigenfunctions for $T_{\rm SL}$ and $\hat{G}$ are the same and the eigenvalues each others inverse. Since $\hat{G}$ is compact we can now apply Theorem 2.9 to it. If $\mathrm{Ker}(T_{\rm SL}) \neq \{\mathbf{0}\}$ we can shift $T_{\rm SL} \to T_{\rm SL} + \alpha$ by some value α so that the new operator has a vanishing kernel. In summary, Theorem 2.9 then applies to our eigenvalue problem:

Lemma 5.5. *The set of eigenvalues of the Sturm-Liouville operator is either finite or it forms a sequence which tends to infinity. Every eigenvalue has a finite degeneracy and there exists an ortho-normal basis of eigenvectors.*

5.4.3 Sturm-Liouville and Fredholm alternative

In view of the Sturm-Liouville formalism, we can now re-visit our original boundary value problem but, for simplicity, we specialise to the case of homogeneous boundary conditions. This means, we are considering the equations

$$Ty = f\,, \qquad d_a y(a) + n_a y'(a) = 0\,, \qquad d_b y(b) + n_b y'(b) = 0 \tag{5.78}$$

$$Ty = 0\,, \qquad d_a y(a) + n_a y'(a) = 0\,, \qquad d_b y(b) + n_b y'(b) = 0 \tag{5.79}$$

where T is a second order differential operator of the form (5.2) (which, we now know, can be written in Sturm-Liouville form). In a way similar to the above Green function method, we can convert this problem into one that involves a compact integral operator, turning the differential equation into an integral equation. The benefit is that Theorem 2.10 on Fredholm's alternative can be applied to this problem and turned into a version of Fredholm's alternative for second order linear differential equations.

Theorem 5.16. *Let e_k be an ortho-normal basis of eigenvectors of T with eigenvalues λ_k, so $Te_k = \lambda_k e_k$ (and the e_k satisfy the boundary conditions in Eq (5.78)). Then, the following alternative holds for the solution to the boundary value problem (5.78).*

(a) There is no non-trivial solution to the homogeneous problem (5.79). In this case there is a unique solution y to the problem (5.78) for every f given by

$$y = \sum_k \frac{1}{\lambda_k} \langle e_k, f\rangle e_k\,. \tag{5.80}$$

(b) There is a non-trivial solution of the homogeneous problem (5.79). In this case, there exists a solution to the inhomogeneous problem if and only if $\langle y_0, f\rangle = 0$ for all solutions y_0 to the homogeneous problem (5.79). If this condition is satisfied, the solution to (5.78) is given by

$$y = \sum_{k:\lambda_k\neq 0} \frac{1}{\lambda_k} \langle e_k, f\rangle e_k + y_0\,, \tag{5.81}$$

where y_0 is an arbitrary solution to the homogeneous problem (5.79).

Proof. Broadly, this follows from Theorem 2.10 on Fredholm's alternative. The details of the proof are somewhat technical, particularly in dealing with the boundary conditions, and can, for example, be found in Ref. [5]. □

We note that the unique solution (5.80) in case (a) can be written as

$$y(x) = \sum_k \frac{1}{\lambda_k} \int_a^b dt\, w(t) e_k(t) e_k(x) f(t) = \int_a^b dt\, G(x,t) f(t) \tag{5.82}$$

so that we obtain an expression for the Green function

$$G(x,t) = w(t) \sum_k \frac{1}{\lambda_k} e_k(t) e_k(x) \tag{5.83}$$

in terms of the eigenfunctions.

6 Laplace equation

The *Laplace operator* Δ in $\mathbb{R}^n$ with coordinates $(x_1, \ldots, x_n)$ is defined as

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \,. \tag{6.1}$$

It gives rise to the homogeneous and inhomogeneous Laplace equations

$$\Delta\phi = 0 \,, \qquad \Delta\phi = \rho \,, \tag{6.2}$$

where $\rho : \mathbb{R}^n \to \mathbb{R}$ is a given function. We are asking for the solutions ϕ to these equations on a compact set $\mathcal{V} \subset U \subset \mathbb{R}^n$ (where U is open), usually subject to certain boundary conditions on the smooth boundary $\partial\mathcal{V}$ of $\mathcal{V}$. A solution to the homogeneous Laplace equation is also called a *harmonic function*. There are two types of boundary conditions which are frequently imposed:

$$\phi|_{\partial\mathcal{V}} = h \quad \text{(Dirichlet)} \qquad\qquad \mathbf{n} \cdot \nabla\phi|_{\partial\mathcal{V}} = h \quad \text{(von Neumann)} \tag{6.3}$$

where h is a given function on the boundary prescribing the boundary values and $\mathbf{n}$ is the normal vector to the boundary. (Linear combinations of these conditions, referred to as mixed boundary conditions, are also possible.) If the choice of boundary condition involves setting $h = 0$ and we define the "force field" $\mathbf{E} = -\nabla\phi$ then Dirichlet boundary conditions imply that the boundary is an equipotential surface, so $\mathbf{E}$ is perpendicular to it. Under the same conditions, von Neumann boundary conditions imply that the component of $\mathbf{E}$ normal to the boundary vanishes.

The above equations are of considerable importance in physics. For example, they govern the theory of electrostatics (with ϕ being the electrostatic potential, ρ, up to a constant, the charge density, $\mathbf{E}$ the electric field and boundary conditions implemented, for example, by the presence of conducting surfaces) and the theory of Newtonian gravity (with ϕ being the Newtonian gravitational potential, ρ, up to constant, being the mass density and $\mathbf{E}$ the gravitational field). The Laplace operator also appears as part of many partial differential equations in physics, for example in the context of the Schrödinger equation in quantum mechanics.

Eqs. (6.2) are obviously linear so we already know that, before imposing any boundary conditions, the solutions to the homogeneous Laplace equation form a vector space and the solutions to the inhomogeneous equation can be obtained by adding all solutions of the homogeneous equation to a particular solution of the inhomogeneous equation.

6.1 Laplacian in different coordinate systems

It is often useful to write the Laplacian in other than Cartesian coordinates and this is facilitated by the following

Proposition 6.1. *(Laplacian in general coordinates) Given the (twice differentiable) map $X : V \to U$ (a "parametrisation" or a "coordinate change"), with $V \subset \mathbb{R}^k$ and $U \subset \mathbb{R}^n$ open, we consider the space $M = X(V)$. Introducing coordinates $\mathbf{t} = (t_1, \ldots, t_k)$ on V we write X as $\mathbf{t} \mapsto \mathbf{x}(\mathbf{t}) = (x_1(\mathbf{t}), \ldots, x_n(\mathbf{t}))$. With the tangent vectors $\frac{\partial\mathbf{x}}{\partial t_i}$ (required to be linearly independent), define the $k \times k$ metric G with entries*

$$G_{ij} = \frac{\partial\mathbf{x}}{\partial t_i} \cdot \frac{\partial\mathbf{x}}{\partial t_j} \,, \qquad g := \det(G) \,. \tag{6.4}$$

The entries of its inverse G^{-1} are denoted by G^{ij}. Then, the Laplacian Δ_X relative to the parametrisation X is given by

$$\Delta_X = \frac{1}{\sqrt{g}} \frac{\partial}{\partial t_i} \left(\sqrt{g}\, G^{ij} \frac{\partial}{\partial t_j} \right) \,. \tag{6.5}$$

The measure for integration relative to X is given by $dS = \sqrt{g}\, dt_1 \cdots dt_k$.

Proof. This formula is proved in Appendix B which contains an account of some basic differential geometry, a subject somewhat outside the main thrust of this lecture. □

Exercise 6.1. *Consider a curve $[a, b] \ni t \to \mathbf{x}(t) \in \mathbb{R}^n$ and use Lemma 6.1 to derive the measure dS for integration over a curve. Do the same with a surface $(t_1, t_2) \to \mathbf{x}(t_1, t_2) \in \mathbb{R}^3$ and convince yourself that the measure dS you obtain reproduces what you have learned about integration over surfaces.*

Let us apply this formula to derive the Laplacian in several useful coordinate systems.

6.1.1 Two-dimensional Laplacian

In $\mathbb{R}^2$ with Cartesian coordinates $\mathbf{x} = (x, y)^T$ the Laplacian is given by

$$\Delta_2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{6.6}$$

The two-dimensional case is somewhat special since $\mathbb{R}^2 \cong \mathbb{C}$ and we can introduce complex coordinates $z = x + iy$ and $\bar{z} = x - iy$. Introducing the *Wirtinger derivatives*

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) , \qquad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) , \tag{6.7}$$

a short calculation shows that

$$\Delta_2 = 4\frac{\partial^2}{\partial z \partial \bar{z}} , \tag{6.8}$$

This formula is extremely useful. We note that a holomorphic function, $\phi = \phi(z)$, is, loosely speaking, a function which does not depend on $\bar{z}$ and, hence, satisfies $\frac{\partial \phi}{\partial \bar{z}} = 0$. Eq. (6.8) says that every holomorphic function $\phi = \phi(z)$ solves the two-dimensional homogeneous Laplace equation, $\Delta_2 \phi = 0$, so we can immediately write down large classes of solutions to this equation. We will come back to this observation later.

Another common set of coordinates in $\mathbb{R}^2$ are two-dimensional polar coordinates $\mathbf{t} = (r, \varphi)$, where $r \in [0, \infty]$ and $\varphi \in [0, 2\pi[$, related to Cartesian coordinates by

$$\mathbf{x}(r, \varphi) = r(\cos\varphi, \sin\varphi) . \tag{6.9}$$

In the language of Lemma (6.1), the corresponding tangent vectors are

$$\frac{\partial \mathbf{x}}{\partial r} = (\cos\varphi, \sin\varphi) , \qquad \frac{\partial \mathbf{x}}{\partial \varphi} = r(-\sin\varphi, \cos\varphi) , \tag{6.10}$$

which gives $G = \mathrm{diag}(1, r^2)$ and $g = r^2$. Inserting this into the general formula (6.5) gives the two-dimensional Laplacian in polar coordinates

$$\Delta_{2,\mathrm{pol}} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} . \tag{6.11}$$

(The integration measure in two-dimensional polar coordinates is $dS = \sqrt{g}\, dr\, d\varphi = r\, dr\, d\varphi$.)

6.1.2 Three-dimensional Laplacian

In $\mathbb{R}^3$ with coordinates $\mathbf{x} = (x, y, z)$ the Laplacian in Cartesian coordinates is given by

$$\Delta_3 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \,. \tag{6.12}$$

Cylindrical coordinates $\mathbf{t} = (r, \varphi, z)$, where $r \in [0, \infty]$, $\varphi \in [0, 2\pi[$ and $z \in \mathbb{R}$, are related to their Cartesian counterparts by

$$\mathbf{x}(r, \varphi, z) = (r\cos\varphi, r\sin\varphi, z) \,. \tag{6.13}$$

The tangent vectors

$$\frac{\partial \mathbf{x}}{\partial r} = (\cos\varphi, \sin\varphi, 0) \,, \qquad \frac{\partial \mathbf{x}}{\partial \varphi} = (-r\sin\varphi, r\cos\varphi, 0) \,, \qquad \frac{\partial \mathbf{x}}{\partial z} = (0, 0, 1) \,, \tag{6.14}$$

imply the metric $G = \mathrm{diag}(1, r^2, 1)$ with determinant $g = r^2$ and hence, by inserting into Eq. (6.5), the three-dimensional Laplacian in cylindrical coordinates

$$\Delta_3 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2} = \Delta_{2,\mathrm{pol}} + \frac{\partial^2}{\partial z^2} \,. \tag{6.15}$$

(For the integration measure in cylindrical coordinates we get the well-known result $dS = \sqrt{g}\, dr\, d\varphi\, dz = r dr\, d\varphi\, dz$.)

We can repeat this analysis for three-dimensional spherical coordinates $\mathbf{t} = (r, \theta, \varphi)$, where $r \in [0, \infty]$, $\theta \in [0, \pi[$ and $\varphi \in [0, 2\pi[$, defined by

$$\mathbf{x}(r, \theta, \varphi) = r(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \,. \tag{6.16}$$

The tangent vectors are

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

which leads to the metric $G = \mathrm{diag}(1, r^2, r^2\sin^2\theta)$ with determinant $g = r^4\sin^2\theta$. Inserting into Eq. (6.5) gives the three-dimensional Lagrangian in spherical coordinates

$$\Delta_{3,\mathrm{sph}} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \,. \tag{6.17}$$

(The integration measure for three-dimensional polar coordinates is $dS = \sqrt{g}\, dr\, d\theta\, d\varphi = r^2\sin\theta\, dr\, d\theta\, d\varphi$.)

6.1.3 Laplacian on the sphere

We can also use Lemma 6.1 to find the Laplacian on non-trivial manifolds, such as a two-sphere $S^2 = \{\mathbf{x} \in \mathbb{R}^3 \,|\, |\mathbf{x}| = 1\}$. We parametrise the two-sphere by coordinates $\mathbf{t} = (\theta, \varphi)$, where $\theta \in [0, \pi[$ and $\varphi \in [0, 2\pi[$, by writing

$$\mathbf{x}(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \,. \tag{6.18}$$

The two tangent vectors are

$$\frac{\partial \mathbf{x}}{\partial \theta} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta)\ , \qquad \frac{\partial \mathbf{x}}{\partial \varphi} = (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0)\ , \tag{6.19}$$

with associated metric $G = \mathrm{diag}(1, \sin^2\theta)$ and determinant $g = \sin^2\theta$. Inserting into Eq. (6.5) gives the Laplacian on the two-sphere

$$\Delta_{S^2} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\ . \tag{6.20}$$

Comparison with Eq. (6.17) shows that the three-dimensional Laplacian can be expressed as

$$\Delta_{3,\mathrm{sph}} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\,\Delta_{S^2}\ . \tag{6.21}$$

(The integration measure on the two-sphere is $dS = \sqrt{g}\,d\theta\,d\varphi = \sin\theta\,d\theta\,d\varphi$.)

6.1.4 Green Identities

Our discussion below will frequently require *Green's identities* which follow from Gauss's integral theorem

$$\int_{\mathcal{V}} \nabla\cdot\mathbf{A}\,dV = \int_{\partial\mathcal{V}} \mathbf{A}\cdot\mathbf{n}\,dS\ . \tag{6.22}$$

where $\mathbf{A}$ is a continuously differentiable vector field on the open set $U \subset \mathbb{R}^n$ and $\mathcal{V} \subset U$ is a compact set with smooth boundary $\partial\mathcal{V}$. Consider two twice continuously differentiable functions $f, g : U \to \mathbb{R}$ and set $\mathbf{A} = f\,\nabla g$. If we use

$$\nabla\cdot\mathbf{A} = \nabla\cdot(f\,\nabla g) = f\Delta g + \nabla f\cdot\nabla g\ , \qquad \mathbf{A}\cdot\mathbf{n} = f\,\nabla g\cdot\mathbf{n} \tag{6.23}$$

Gauss's theorem turns into the *first Green formula*

$$\int_{\mathcal{V}} (f\Delta g + \nabla f\cdot\nabla g)dV = \int_{\partial\mathcal{V}} f\,\nabla g\cdot\mathbf{n}\,dS\ . \tag{6.24}$$

Exchanging f and g in this formula and subtracting the two resulting equations gives the *second Green formula* or Green's identity

$$\int_{\mathcal{V}} (f\Delta g - g\Delta f)dV = \int_{\partial\mathcal{V}} (f\nabla g - g\nabla f)\cdot\mathbf{n}\,dS\ . \tag{6.25}$$

After this preparation we are now ready to delve into the task of solving the Laplace equation.

6.2 Basic theory*

6.2.1 Green functions for the Laplacian

In this subsection we discuss a number of basic mathematical results for the Laplace equation in $\mathbb{R}^n$ (with Cartesian coordinates $\mathbf{x} = (x_1, \ldots, x_n)^T$), starting with the inhomogeneous version

$$\Delta\phi = \rho\ , \qquad \Delta = \sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}\ , \tag{6.26}$$

of the equation. Define the generalised *Newton (or Coulomb) potentials* as

$$G(\mathbf{x}-\mathbf{a}) = G_{\mathbf{a}}(\mathbf{x}) = \begin{cases} -\frac{1}{(n-2)v_n}\frac{1}{|\mathbf{x}-\mathbf{a}|^{n-2}} & \text{for} \quad n>2 \\ \frac{1}{2\pi}\ln|\mathbf{x}-\mathbf{a}| & \text{for} \quad n=2 \end{cases} , \tag{6.27}$$

where v_n is the surface "area" of the $n-1$-dimensional sphere, S^{n-1} (and the constants have been included for later convenience). In electromagnetism, $G_{\mathbf{a}}$ corresponds the the electrostatic potential of a point charge located at $\mathbf{a}$. Clearly, $G_{\mathbf{a}}$ is well-defined for all $\mathbf{x} \neq \mathbf{a}$. It is straightforward to verify by direct calculation that

$$\Delta G_{\mathrm{a}} = 0 \quad \text{for all} \quad \mathbf{x} \neq \mathbf{a} \, . \tag{6.28}$$

Exercise 6.2. *Show that the gradient of the Newton potentials (6.27) is given by*

$$\nabla G_{\mathbf{a}}(\mathbf{x}) = \frac{1}{v_n}\frac{\mathbf{x}-\mathbf{a}}{|\mathbf{x}-\mathbf{a}|^n} \, . \tag{6.29}$$

Also, verify that the Newton potentials satisfy the homogeneous Laplace equation for all $\mathbf{x} \neq \mathbf{a}$.

Lemma 6.1. *For* $f \in \mathcal{C}^1(U)$ *and* $\mathbf{a} \in U$ *we have*

$$f(\mathbf{a}) = \lim_{\epsilon\to 0}\int_{|\mathbf{x}-\mathbf{a}|=\epsilon} (f(\mathbf{x})\nabla G_{\mathbf{a}}(\mathbf{x}) - G_{\mathbf{a}}(\mathbf{x})\nabla f(\mathbf{x}))\cdot d\mathbf{S} \, . \tag{6.30}$$

Proof. With $d\mathbf{S} = \mathbf{n}\, dS$ and the unit normal vector $\mathbf{n}$ to the sphere given by $\mathbf{n} = \frac{\mathbf{x}-\mathbf{a}}{|\mathbf{x}-\mathbf{a}|}$ we have

$$\nabla G_{\mathbf{a}} \cdot d\mathbf{S} = \frac{1}{v_n}\frac{1}{|\mathbf{x}-\mathbf{a}|^{n-1}}\, dS \, . \tag{6.31}$$

This gives for the first part of the above integral

$$\begin{aligned} \lim_{\epsilon\to 0}\int_{|\mathbf{x}-\mathbf{a}|=\epsilon} f(\mathbf{x})\nabla G_{\mathbf{a}}(\mathbf{x})\cdot d\mathbf{S} &= \lim_{\epsilon\to 0}\frac{1}{v_n\epsilon^{n-1}}\int_{|\mathbf{x}-\mathbf{a}|=\epsilon} f(\mathbf{x})dS \\ &\overset{\mathbf{y}=(\mathbf{x}-\mathbf{a})/\epsilon}{=} \frac{1}{v_n}\lim_{\epsilon\to 0}\int_{|\mathbf{y}|=1} f(\mathbf{a}+\epsilon\mathbf{y})dS = f(\mathbf{a}) \, . \end{aligned} \tag{6.32}$$

For the second integral, using that $|\nabla f(\mathbf{x})\cdot\mathbf{n}| \leq K$ for some constant K, we have

$$\left|\int_{|\mathbf{x}-\mathbf{a}|=\epsilon} G_{\mathbf{a}}(\mathbf{x})\nabla f(\mathbf{x})\cdot d\mathbf{S}\right| \leq \text{const}\int_{|\mathbf{x}-\mathbf{a}|=\epsilon}\epsilon^{2-n}dS = \epsilon\int_{|\mathbf{y}|=1} dS \quad \overset{\epsilon\to 0}{\longrightarrow} \quad 0 \, , \tag{6.33}$$

and this completes the proof. □

This Lemma was the technical preparation for the following important statement.

Theorem 6.3. *Let* $\rho \in \mathcal{C}_c^2(\mathbb{R}^n)$ *and define the function*

$$\phi(\mathbf{x}) := \int_{\mathbb{R}^n} dy^n\, G(\mathbf{x}-\mathbf{y})\rho(\mathbf{y}) \tag{6.34}$$

for all $\mathbf{x} \in \mathbb{R}^n$. *Then* $\Delta\phi = \rho$, *that is, the above* ϕ *satisfies the inhomogeneous Laplace equation with source* ρ.

Proof. Introducing the coordinate $\mathbf{z} = \mathbf{y} - \mathbf{x}$, a region $\mathcal{V}_\epsilon = \{\mathbf{z} \in \mathbb{R}^n \,|\, \epsilon \leq |\mathbf{z}| \leq R\}$ with R so large that $\rho(\mathbf{x}+\mathbf{z}) = 0$ for $|\mathbf{z}| > R$ (which is possible since ρ has compact support) and $\rho_{\mathbf{x}}(\mathbf{z}) := \rho(\mathbf{x}+\mathbf{z})$, we have

$$\begin{aligned}\Delta\phi(\mathbf{x}) &= \int_{\mathbb{R}^n} dz^n\, G(\mathbf{z})\Delta\rho(\mathbf{x}+\mathbf{z}) = \lim_{\epsilon\to 0}\int_{\mathcal{V}_\epsilon} dz^n\, G\,\Delta\rho_{\mathbf{x}} = \lim_{\epsilon\to 0}\int_{\mathcal{V}_\epsilon} dz^n\,(G\,\Delta\rho_{\mathbf{x}} - \rho_{\mathbf{x}}\Delta G)\\ &= \lim_{\epsilon\to 0}\int_{\partial\mathcal{V}_\epsilon}(G\nabla\rho_{\mathbf{x}} - \rho_{\mathbf{x}}\nabla G)\cdot d\mathbf{S} = \rho_{\mathbf{x}}(\mathbf{0}) = \rho(\mathbf{x})\;, \end{aligned} \tag{6.35}$$

where we have used Green's formula (6.25) and Lemma (6.1). □

The above function G is also sometimes referred to as the *Green function* of the Laplace operator. Of course, the solution (6.34) is not unique but we know that two solutions to the inhomogeneous Laplace equation differ by a solution to the homogeneous one. Hence, the general solution to the inhomogeneous Laplace equation can be written as

$$\phi(\mathbf{x}) = \phi_H(\mathbf{x}) + \int_{\mathbb{R}^n} dy^n\, G(\mathbf{x}-\mathbf{y})\rho(\mathbf{y}) \quad \text{where} \quad \Delta\phi_H = 0\;. \tag{6.36}$$

The homogeneous solution ϕ_H can be used to satisfy the boundary conditions on ϕ. Note that the requirement on ρ to have compact support also makes physical sense: normally charge or mass distributions are localised in space.

6.2.2 Maximum principle and uniqueness

Via Eq. (6.36), we have now reduced the problem of solving the inhomogeneous Laplace equation to that of solving the homogeneous Laplace equation and this is what we discuss next. A twice differential function ϕ which solves the homogeneous Laplace equation,

$$\Delta\phi = 0\;, \tag{6.37}$$

is called a *harmonic function*. Harmonic functions have a number of remarkable properties which we now derive. We begin with another technical Lemma.

Lemma 6.2. *For $U \subset \mathbb{R}^n$ open, $\mathcal{V} \subset U$ compact with smooth boundary $\partial\mathcal{V}$ and ϕ harmonic on $\mathring{\mathcal{V}} := \mathcal{V}\setminus\partial\mathcal{V}$ we have*

$$\int_{\partial\mathcal{V}}(\phi\nabla G_{\mathbf{a}} - G_{\mathbf{a}}\nabla\phi)\cdot d\mathbf{S} = \begin{cases}\phi(\mathbf{a}) & \text{for} \quad \mathbf{a}\in\mathring{\mathcal{V}}\\ 0 & \text{for} \quad \mathbf{a}\in\mathbb{R}^n\setminus\mathcal{V}\end{cases} \tag{6.38}$$

Proof. For the second case, $\mathbf{a} \in \mathbb{R}^n \setminus \mathcal{V}$ we have from Green's formula

$$\int_{\partial\mathcal{V}}(\phi\nabla G_{\mathbf{a}} - G_{\mathbf{a}}\nabla\phi)\cdot d\mathbf{S} = \int_{\mathcal{V}}(\phi\Delta G_{\mathbf{a}} - G_{\mathbf{a}}\Delta\phi)dV = 0\;, \tag{6.39}$$

since $\Delta G_{\mathbf{a}} = \Delta\phi = 0$ for all $\mathbf{x} \in \mathcal{V}$.

For the first case, we define $\mathcal{V}_\epsilon = \mathcal{V} \setminus B_\epsilon(\mathbf{a})$, that is, we excise a ball with radius ϵ around $\mathbf{a}$. Just like above it follows from Green's formula that

$$\int_{\partial\mathcal{V}_\epsilon}(\phi\nabla G_{\mathbf{a}} - G_{\mathbf{a}}\nabla\phi)\cdot d\mathbf{S} = 0\;. \tag{6.40}$$

Since the boundary $\partial\mathcal{V}_\epsilon$ consists of the two components $\partial\mathcal{V}$ and $\partial\overline{B_\epsilon(\mathbf{a})}$ this implies

$$\int_{\partial\mathcal{V}}(\phi\nabla G_{\mathbf{a}} - G_{\mathbf{a}}\nabla\phi)\cdot d\mathbf{S} = \int_{\partial\overline{B_\epsilon(\mathbf{a})}}(\phi\nabla G_{\mathbf{a}} - G_{\mathbf{a}}\nabla\phi)\cdot d\mathbf{S} \quad \overset{\epsilon\to 0}{\longrightarrow} \quad \phi(\mathbf{a})\;, \tag{6.41}$$

where Lemma (6.1) has been used in the final step. Since the integral on the LHS is independent of ϵ this completes the proof. □

We are now ready to proof the first important property of harmonic functions.

Theorem 6.4. *(Mean value property of harmonic functions) Let $U \subset \mathbb{R}^n$ be open, ϕ harmonic on U and $B_r(\mathbf{a}) \subset U$. Then*

$$\phi(\mathbf{a}) = \frac{1}{v_n} \int_{|\mathbf{y}|=1} \phi(\mathbf{a} + r\mathbf{y}) dS \tag{6.42}$$

Proof. From the previous Lemma we have

$$\phi(\mathbf{a}) = \int_{|\mathbf{x}-\mathbf{a}|=r} (\phi \nabla G_{\mathbf{a}} - G_{\mathbf{a}} \nabla \phi) \cdot d\mathbf{S} \,, \tag{6.43}$$

and the first part of this integral

$$\int_{|\mathbf{x}-\mathbf{a}|=r} \phi \nabla G_{\mathbf{a}} \cdot \mathbf{n} dS = \frac{1}{v_n r^{n-1}} \int_{|\mathbf{x}-\mathbf{a}|=r} \phi(\mathbf{x}) dS \overset{\mathbf{y}=(\mathbf{x}-\mathbf{a})/r}{=} \frac{1}{v_n} \int_{|\mathbf{y}|=1} \phi(\mathbf{a} + r\mathbf{y}) dS \tag{6.44}$$

already gives the desired result. It remains to be shown that the second part of the integral vanishes. Since $G_{\mathbf{a}}$ is constant for $|\mathbf{x} - \mathbf{a}| = r$ it is unimportant and it is sufficient to consider

$$\int_{|\mathbf{x}-\mathbf{a}|=r} G_{\mathbf{a}} \nabla \phi \cdot d\mathbf{S} = \int_{|\mathbf{x}-\mathbf{a}|=r} (1 \nabla \phi - \phi \nabla 1) \cdot d\mathbf{S} = \int_{|\mathbf{x}-\mathbf{a}|\leq r} (1 \Delta \phi - \phi \Delta 1) dV = 0 \,, \tag{6.45}$$

which does indeed vanish from Green's theorem. □

An important consequence of this property is

Theorem 6.5. *(Maximum principle for harmonic functions) Let $U \subset \mathbb{R}^n$ be open and (path-) connected and ϕ a harmonic function on U. If ϕ assumes its maximum for a point $\mathbf{a} \in U$ then ϕ is constant.*

Proof. We set $M := \sup\{\phi(\mathbf{x}) \,|\, \mathbf{x} \in U\}$ and assume that $\phi(\mathbf{x}) = M$ for some $\mathbf{x} \in U$. We start by showing that this implies ϕ is constant on a ball $B_\epsilon(\mathbf{x}) \subset U$. From the mean value property it follows for all r with $r_0 < r < \epsilon$ that

$$M = \phi(\mathbf{x}) = \frac{1}{v_n} \int_{|\mathbf{y}|=1} \phi(\mathbf{x} + r\mathbf{y}) dS \quad \Leftrightarrow \quad \int_{|\mathbf{y}|=1} (M - \phi(\mathbf{x} - r\mathbf{y})) dS = 0 \,. \tag{6.46}$$

Since $M - \phi(\mathbf{x} - r\mathbf{y}) \geq 0$ this implies that $\phi(\mathbf{x} - r\mathbf{y}) = M$ for all $|\mathbf{y}| = 1$ and, hence, $\phi(\mathbf{y}) = M$ for all $\mathbf{y} \in B_\epsilon(\mathbf{x})$.

To extend the statement to all of U we assume there exists a $\mathbf{a} \in U$ with $\phi(\mathbf{a}) = M$. Assume that ϕ is not constant on U so there is a $\mathbf{b} \in U$ with $\phi(\mathbf{b}) < M$ and choose a (continuous) path $\alpha : [0, 1] \to U$ which connects $\mathbf{a}$ and $\mathbf{b}$, that is, $\alpha(0) = \mathbf{a}$ and $\alpha(1) = \mathbf{b}$. Let $t_0 = \sup\{t \in [0, 1] \,|\, \phi(\alpha(t)) = M\}$ be the "upper value" of t along the path for which the maximum is assumed. Since $\phi(\mathbf{b}) < M$ necessarily $t_0 < 1$ and since $\phi \circ \alpha$ is continuous we have $\phi(\alpha(t_0)) = M$. But from the first part of the proof we know there is a ball $B_\epsilon(\alpha(t_0))$ where ϕ equals M which is in contradiction with t_0 being the supremum. Hence, the assumption $\phi(\mathbf{b}) < M$ was incorrect. □

Corollary 6.1. *Let $U \subset \mathbb{R}^n$ be a bounded, connected open set and ϕ be harmonic on U and continuous on $\bar{U}$. Then ϕ assumes its maximum and minimum on the boundary of U.*

Proof. Since $\bar{U}$ is compact ϕ assumes its maximum on $\bar{U}$. If the maximum point is on the boundary $\partial\bar{U}$ then the statement is true. If it is in U then, from the previous theorem, ϕ is constant on U and, hence, by continuity, constant on $\bar{U}$. Therefore it also assume its maximum on the boundary. The corresponding statement for the minimum follows by considering $-\phi$. □

These innocent sounding statements have important implications for boundary value problems. Suppose we have an open, bounded and connected set $\mathcal{V} \subset U \subset \mathbb{R}^n$ and we would like to solve the Dirichlet boundary value problem

$$\Delta\phi = \rho \,, \qquad \phi|_{\partial\mathcal{V}} = h \,, \tag{6.47}$$

where ρ is a given function on U and h is a given function on the boundary $\partial\mathcal{V}$ which prescribes the boundary values of ϕ. Suppose we have two solutions ϕ_1 and ϕ_2 to this problem. Then the difference $\phi := \phi_1 - \phi_2$ satisfies

$$\Delta\phi = 0 \,, \qquad \phi|_{\partial\mathcal{V}} = 0 \,. \tag{6.48}$$

Since ϕ is harmonic it assumes its maximum and minimum on the boundary and since the boundary values are zero we conclude that $\phi = 0$. This means that the solution to the boundary value problem, if it exists, is unique.

We also have a uniqueness statement for our solution (6.34).

Corollary 6.2. *For $\rho \in \mathcal{C}^2_c(\mathbb{R}^n)$ and $n \geq 3$ the equation $\Delta\phi = \rho$ has a unique solution $\phi : \mathbb{R}^n \to \mathbb{R}$ with $\lim_{|\mathbf{x}|\to\infty} |\phi(\mathbf{x})| = 0$. This solution is given by Eq. (6.34).*

Proof. The solution (6.34) has the desired property at infinity since $\lim_{|\mathbf{x}|\to\infty} G_{\mathbf{a}}(\mathbf{x}) = 0$ for $n \geq 3$. To show uniqueness assume another solution, $\tilde{\phi}$, with the same property and consider the difference $\psi = \phi - \tilde{\phi}$. We have

$$\Delta\psi = 0 \,, \qquad \lim_{|\mathbf{x}|\to\infty} |\psi(\mathbf{x})| = 0 \,. \tag{6.49}$$

From the vanishing property at infinity, we know that for every $\epsilon > 0$, there exists a radius R such that $|\psi(\mathbf{x})| \leq \epsilon$ for all $|\mathbf{x}| \geq R$. But the restricted function $\psi|_{B_R(\mathbf{0})}$ assumes its maximum and minimum on the boundary so it follows that $|\psi(\mathbf{x}) \leq \epsilon$ for all $\mathbf{x} \in \mathbb{R}^n$. Since $\epsilon > 0$ was arbitrary this means that $\psi = 0$. □

6.2.3 Uniqueness - another approach

There is another way to make the uniqueness argument which can also be directly applied to the case of boundary value problems with von Neumann boundary conditions

$$\Delta\phi = \rho \,, \qquad \mathbf{n} \cdot \nabla\phi|_{\partial\mathcal{V}} = h \,. \tag{6.50}$$

Consider a harmonic function ϕ on U and set $f = g = \phi$ in Green's first identity (6.24). This results in

$$\int_{\mathcal{V}} |\nabla\phi|^2 \, dV = \int_{\partial\mathcal{V}} \phi \, \nabla\phi \cdot \mathbf{n} \, dS \,, \tag{6.51}$$

and, hence,

$$\begin{array}{llll} \Delta\phi = 0 \,, & \phi|_{\partial\mathcal{V}} = 0 & \Longrightarrow & \phi = 0 \\ \Delta\phi = 0 \,, & \nabla\phi \cdot \mathbf{n}|_{\partial\mathcal{V}} = 0 & \Longrightarrow & \phi = \text{const} \,. \end{array} \tag{6.52}$$

Applied to the difference $\phi = \phi_1 - \phi_2$ of two solutions ϕ_1, ϕ_2 to the Dirichlet problem (6.47) this result implies $\phi_1 = \phi_2$, so uniqueness of the solution. Applying it to the difference $\phi = \phi_1 - \phi_2$ of two solutions to the von Neumann problem (6.50) gives $\phi_1 = \phi_2 + \text{const}$, so uniqueness up to an additive constant (which does not change $\mathbf{E} = -\nabla\phi$).

After this somewhat theoretical introduction we now proceed to the problem of how to solve Laplace's equation in practice, starting with the two-dimensional case.

6.3 Laplace equation in two dimensions

At first sight, the two-dimensional Laplace equation seems of little physical interest - after all physical space has three dimensions. However, there are many problems, for example, in electro-statics, which are effectively two-dimensional due to translational symmetry in one direction. (Think, for example, of the field produced by of a long charged wire along the z-axis.)

We denote the Cartesian coordinates by $\mathbf{x} = (x, y)^T$ and also introduce complex coordinates $z = x + iy$ and $\bar{z} = x - iy$. Recall that the two-dimensional Laplacian can be written as

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 4\frac{\partial^2}{\partial \bar{z}\partial z} \,. \tag{6.53}$$

6.3.1 Complex methods

We have already pointed out that every holomorphic function $w = w(z)$ solves the two-dimensional Laplace equation. Of course, normally we are interested in real-valued solutions but, since the Laplacian is a real operator, both the real and imaginary part of $w(z)$ are also harmonic. To make this explicit, we write

$$w = u + iv \,. \tag{6.54}$$

If w is holomorphic then $\frac{\partial w}{\partial \bar{z}} = 0$ and using the derivatives (6.7) together with the decomposition (6.54) this translates into the *Cauchy-Riemann equations*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \,, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \,. \tag{6.55}$$

These equations immediately imply that

$$\nabla u \cdot \nabla v = 0 \,, \tag{6.56}$$

which means that the curves $u = \text{const}$ and $v = \text{const}$ are perpendicular to one another. Furthermore, we have

$$\Delta w = 0 \quad \Rightarrow \quad \Delta u = \Delta v = 0 \,. \tag{6.57}$$

Our strategy for solving the two-dimensional Laplace equation is based on these simple equations and is probably best explained by an example.

Application 6.22. Solving the two-dimensional Laplace equation with complex methods
Suppose we want to solve Laplace's equation in the positive quadrant $\{(x, y) \,|\, x \geq 0,\ y \geq 0\}$ and we impose Dirichlet boundary conditions $\phi(0, y) = \phi(x, 0) = 0$ along the positive x and y axis. (See Fig. 14.) It is clear that the holomorphic function $w = z^2$ has a vanishing imaginary part along the real and imaginary axis (just insert $z = x$ and $z = iy$ to check this) and, hence, the choice $\phi = v = \text{Im}(z^2) = 2xy$ leads to a harmonic function with the desired boundary property. On the other hand, if we had imposed Neumann boundary conditions $\frac{\partial \phi}{\partial x}(0, y) = \frac{\partial \phi}{\partial y}(x, 0) = 0$ along the positive x and y axis, the real part of w (having perpendicular equipotential lines) leads to a viable solution $\phi = u = \text{Re}(z^2) = x^2 - y^2$. (Of course any even power of z would also do the job so the solution is not unique. This is because we haven't specified boundary conditions at infinity.) The equipotential lines for both solutions are shown in Fig. 15.

For another example, consider solving Laplace's equation on $U = \{z \in \mathbb{C} \,|\, |z| > 1\}$ with Dirichlet boundary condition $\phi|_{|z|=1} = 0$. (See Fig. 14.) It is clear that the function $w = z + z^{-1}$ is real for $|z| = 1$ (and it is holomorphic on U) so with $z = re^{i\varphi}$ we have a solution $\phi = v = \text{Im}(z + z^{-1}) = 2(r - r^{-1}) \sin \varphi$. (Again, this is not unique since we have to specify another boundary condition, for example at infinity.)

The equipotential lines for this solution are shown in Fig. 15. The solution for $|z| \leq 0$ is of course $\phi = 0$, the unique solution consistent with the boundary conditions at $|z| = 1$.

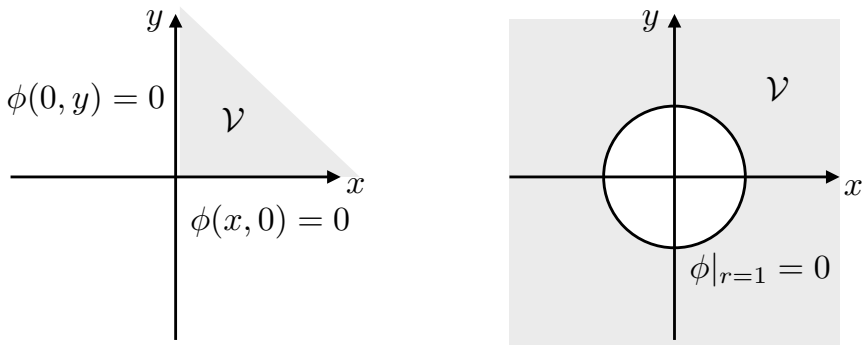

Figure 14: Examples of boundary conditions for two-dimensional Laplace equations.

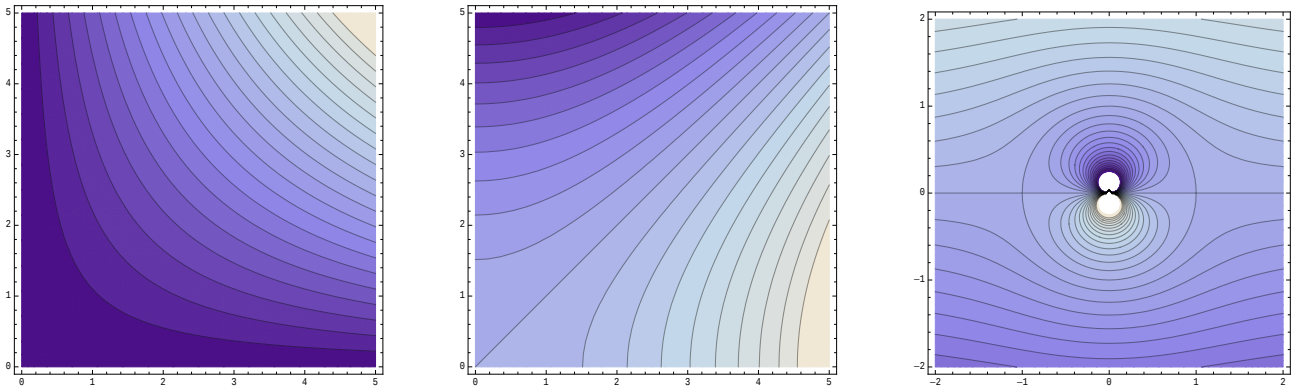

Figure 15: Equipotential lines for $\phi(x, y) = xy$ (left) and $\phi(x, y) = x^2 - y^2$ (middle) and $\phi(x, y) = \mathrm{Im}(z + z^{-1})$.

6.3.2 Separation of variables

Separation of variables is a general technique for solving differential equations which is based on factoring the problem into one-dimensional differential equations. It is useful to demonstrate the technique for the simple example of the two-dimensional homogeneous Laplace equation

$$\Delta\phi = 0 \,, \qquad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \,, \tag{6.58}$$

in Cartesian coordinates. We start by considering solutions of the separated form

$$\phi(x, y) = X(x)Y(y) \,, \tag{6.59}$$

where $X = X(x)$ and $Y = Y(y)$ are functions of their indicated arguments only. Inserting this Ansatz into Eq. (6.58) gives

$$\underbrace{\frac{X''}{X}(x)}_{=-\alpha^2} + \underbrace{\frac{Y''}{Y}(y)}_{=\alpha^2} = 0 \, . \tag{6.60}$$

The argument goes that the two terms, being functions of different variables, can only add up to zero if they are equal to constants α^2 and $-\alpha^2$ individually, as indicated above. Solving the resulting two ordinary differential equations

$$X'' = -\alpha^2 X \, , \qquad Y'' = \alpha^2 Y \, , \tag{6.61}$$

results in the solutions

$$X(x) = a_\alpha \cos(\alpha x) + b_\alpha \sin(\alpha x) \, , \qquad Y(y) = c_\alpha e^{\alpha y} + d_\alpha e^{-\alpha y} \, , \tag{6.62}$$

where a_α, b_α, c_α and d_α are arbitrary constants. This by itself gives a rather special solution to the equation but it does so for every choice of the constant α. Since the equation we are solving is linear we can, hence, construct more general solutions by linearly combining solutions of the above type for different values of α. This leads to

$$\phi(x, y) = \sum_\alpha (a_\alpha \cos(\alpha x) + b_\alpha \sin(\alpha x))(c_\alpha e^{\alpha y} + d_\alpha e^{-\alpha y}) \, , \tag{6.63}$$

where the sum ranges over some suitable set of α values. Of course this is a large class of solutions which can be narrowed down, or be made unique by imposing boundary conditions. Whether this works out in practice depends on the type of boundary conditions and if they are "compatible" with the chosen set of coordinates and resulting solution. Here, we are working with Cartesian coordinates and this goes well together with boundary conditions imposed along lines with $x = \text{const}$ and $y = \text{const}$. More generally, building in boundary conditions tends to be easiest if coordinates are chosen such that the boundaries are defined by one of the coordinates being constants. For example, polar or spherical coordinates go well with imposing boundary conditions on circles or spheres, as we will see below.

Application 6.23. Rectangular boundary conditions with separation of variables

To see how this works in practice, consider solving the problem on the rectangle $\mathcal{V} = [0, a] \times [0, b]$ with ϕ vanishing on all sides of the rectangle except at $y = b$ where we impose the boundary condition $\phi(x, b) = h(x)$ for some given function h. First consider the boundary conditions $\phi(0, y) = \phi(a, y) = 0$ which we can satisfy by setting $a_\alpha = 0$ and $\alpha = \frac{\pi k}{a}$. Further, satisfying $\phi(x, 0) = 0$ can be achieved by setting $d_\alpha = -c_\alpha = 1/2$. Putting this together we end up with

$$\phi(x, y) = \sum_{k=1}^{\infty} b_k \sin\left(\frac{k\pi x}{a}\right) \sinh\left(\frac{k\pi y}{a}\right) \, , \tag{6.64}$$

which, for any fixed y, is a sine Fourier series on the interval $[0, a]$. This already indicates how we built in the final boundary condition $\phi(x, b) = h(x)$. Setting $y = b$ in the above formula, we can determine the coefficients simply by standard (sine) Fourier series techniques and obtain

$$b_k = \frac{2}{a \sinh\left(\frac{k\pi b}{a}\right)} \int_0^a dx \, h(x) \sin\left(\frac{k\pi x}{a}\right) \, . \tag{6.65}$$

For any given boundary potential h these coefficients can be calculated and inserting these back into Eq. (6.64) gives the complete solution.

6.3.3 Polar coordinates

From Eq. (6.11) the two-dimensional Laplacian in polar coordinates reads

$$\Delta = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \varphi^2} \,. \tag{6.66}$$

where $r \in [0, \infty)$ and $\varphi \in [0, 2\pi)$. Just as in the Cartesian case, we can try to solve the two-dimensional Laplace equation in polar coordinates by separation of variables as in the following

Exercise 6.6. *Solve the two-dimensional homogeneous Laplace equation in polar coordinates by separation of variables.*

However, there is a more systematic way forward. For any fixed radius $r > 0$, a functions $\phi(r, \varphi)$ can be thought of as a function on the circle S^1 and can, hence, be expanded in a Fourier series. In other words, we can write

$$\phi(r, \varphi) = \frac{A_0(r)}{2} + \sum_{k=1}^{\infty}(A_k(r)\cos(k\varphi) + B_k(r)\sin(k\varphi)) \,, \tag{6.67}$$

where the Fourier coefficients $A_k(r)$ and $B_k(r)$ can of course change with the radius. Note, there is no assumption involved yet. Eq. (6.67) still represents a general function. Inserting (6.67) into the Laplace equation $\Delta\phi = 0$ gives

$$\frac{1}{2r}(rA_0')' + \sum_{k=1}^{\infty}\left(\frac{1}{r}(rA_k')' - \frac{k^2}{r^2}A_k\right)\cos(k\varphi) + \sum_{k=1}^{\infty}\left(\frac{1}{r}(rB_k')' - \frac{k^2}{r^2}B_k\right)\sin(k\varphi) = 0 \,. \tag{6.68}$$

This is a Fourier series which must be identical to zero so all the Fourier coefficients must vanish. This leads to a set of ordinary differential equations

$$(rA_0')' = 0 \,, \qquad r(rA_k')' = k^2 A_k \,, \qquad r(rB_k')' = k^2 B_k \,, \tag{6.69}$$

for A_k and B_k. They are easy to solve and lead to

$$A_0(r) = a_0 + \tilde{a}_0 \ln r \,, \qquad A_k(r) = a_k r^k + \tilde{a}_k r^{-k} \,, \qquad B_k(r) = b_k r^k + \tilde{b}_k r^{-k} \,. \tag{6.70}$$

Inserting these results back into Eq. (6.67) gives for the general solution of the two-dimensional homogeneous Laplace equation in polar coordinates

$$\phi(r, \varphi) = \frac{a_0}{2} + \frac{\tilde{a}_0}{2}\ln r + \sum_{k=1}^{\infty}(a_k r^k + \tilde{a}_k r^{-k})\cos(k\varphi) + \sum_{k=1}^{\infty}(b_k r^k + \tilde{b}_k r^{-k})\sin(k\varphi) \,. \tag{6.71}$$

The coefficients a_k, b_k, $\tilde{a}_k$ and $\tilde{b}_k$ are arbitrary at this stage and have to be fixed by boundary conditions.

Application 6.24. Two-dimensional Laplace equation with circular boundary conditions
For example, consider solving the problem on the unit disk $\{(r, \varphi) \,|\, r \leq 1\}$ with the boundary condition $\phi(1, \varphi) = h(\varphi)$, where h is a given function on S^1. Since the origin is in this region we do not want any negative powers of r for a non-singular solution, so $\tilde{a}_k = \tilde{b}_k = 0$. Then, the boundary condition at $r = 1$ reads

$$\phi(1, \varphi) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos(k\varphi) + b_k\sin(k\varphi)) \stackrel{!}{=} h(\varphi) \,, \tag{6.72}$$

This is simply the Fourier series for the function h and we can find the Fourier coefficients a_k and b_k by the usual formulae (3.22).

Now consider solving the problem for the same boundary condition $\phi(1,\varphi) = h(\varphi)$ but for the "exterior" region $\{(r,\varphi)\,|\,r \geq 1\}$ imposing, in addition, that ϕ remains finite as $r \to \infty$. The last condition demands that $\tilde{a}_0 = 0$ and $a_k = b_k = 0$ for $k = 1, 2, \ldots$ so we have

$$\phi(1,\varphi) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(\tilde{a}_k \cos(k\varphi) + \tilde{b}_k \sin(k\varphi)) \stackrel{!}{=} h(\varphi)\ . \tag{6.73}$$

As before, this is a Fourier series for h and we can determine the Fourier coefficients by the standard formulae (3.22). So the full solution for ϕ is

$$\phi(r,\varphi) = \begin{cases} \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k r^k \cos(k\varphi) + b_k r^k \sin(k\varphi)) & \text{for} \quad r \leq 1 \\ \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k r^{-k} \cos(k\varphi) + b_k r^{-k} \sin(k\varphi)) & \text{for} \quad r \geq 1 \end{cases}\ , \tag{6.74}$$

where a_k and b_k are the Fourier coefficients of the function h. Note that the two parts of this solution fit together at $r = 1$ as they must.

6.4 Laplace equation on the two-sphere

The Laplacian and the Laplace equation on the two-sphere are of significance for a number of reasons. First of all, we have seen in Eq. (6.21) that the three-dimensional Laplacian in spherical coordinates can be expressed in terms of the Laplacian on S^2 plus a radial piece. Also, as we will discover later, the Laplacian on the two-sphere is closely connected to the mathematics of the group of rotations. More practically, two-spheres are all around us - quite literally so in the case of the celestial two-sphere.

6.4.1 Functions on S^2

We recall that the two-sphere is usually parametrised by two angles $(\theta,\varphi) \in [0,\pi]\times[0,2\pi[$, as in Eq. (6.18). Alternatively and often more conveniently, we can use the coordinates $(x,\varphi) \in [-1,1]\times[0,2\pi[$ where $x = \cos\theta$. Sometimes it is also useful to parametrise the two-sphere by unit vectors $\mathbf{n} \in \mathbb{R}^3$. In terms of the coordinates (x,φ) the Laplacian (6.20) takes the form

$$\Delta_{S^2} = (1-x^2)\frac{\partial^2}{\partial x^2} - 2x\frac{\partial}{\partial x} + \frac{1}{1-x^2}\frac{\partial^2}{\partial\varphi^2}\ . \tag{6.75}$$

We should add a word of caution about functions $f : S^2 \to F$ on the two-sphere. In practice, we describe these as functions $f = f(x,\varphi)$ of the coordinates but not all of these are well-defined on S^2. First of all, a continuous f needs to be periodic in φ, so $f(\theta,0) = f(\theta,2\pi)$. There is another, more basic condition which arises because the parametrisation (6.18) breaks down at $(x,\varphi) = (\pm1,\varphi)$ which correspond to the same two points (the north and the south pole) for all values of φ. Hence, a function $f = f(x,\varphi)$ is only well-defined on S^2 if $f(\pm1,\varphi)$ is independent of φ. So, for example, $f(x,\varphi) = x\sin\varphi$ is not well-defined on S^2 while $f(x,\varphi) = (1-x^2)\sin\varphi$ is. This discussion can be summarised by saying that we can expand a function f on the two-sphere in a Fourier series

$$f(x,\varphi) = \sum_{m\in\mathbb{Z}} y_m(x) e^{im\varphi}\ , \tag{6.76}$$

with $y_m(\pm1) = 0$ for all $m \neq 0$.

Another useful observation is that as an operator on the inner product space $\mathcal{C}^\infty(S^2)$ with scalar product

$$\langle f, h\rangle_{S^2} = \int_{S^2} f(x)^* h(x)\, dS\ , \qquad dS = \sin\theta d\theta\, d\varphi = dx\, d\varphi\ , \tag{6.77}$$

the Laplacian Δ_{S^2} is self-adjoint. This is most elegantly seen by using the general formulae from Lemma 6.1.

$$\langle f, \Delta h\rangle_{S^2} = \int_{S^2} f^* \Delta h\, dS = \int_V f(\mathbf{t})^* \frac{1}{\sqrt{g}} \frac{\partial}{\partial t_i} \left(\sqrt{g} G^{ij} \frac{\partial h}{\partial t_j}(\mathbf{t}) \right) \sqrt{g} d^k t = \int_{S^2} (\Delta f)^* h\, dS = \langle \Delta f, h\rangle_{S^2}\ . \tag{6.78}$$

Hence, we know that the eigenvalues of Δ_{S^2} are real and eigenvectors for different eigenvalue are orthogonal relative to the above inner product.

6.4.2 Eigenvalue problem for the Laplacian on S^2

Solving the eigenvalue problem

$$\Delta_{S^2} f = \lambda f\ , \tag{6.79}$$

is immensely useful and our main task. Inserting the expansion (6.76) into the eigenvalue equation (6.79) and using the form (6.75) of the Laplacian it is easy to see that the functions y_m have to satisfy the differential equation

$$(1-x^2)y_m'' - 2xy_m' + \left(-\lambda - \frac{m^2}{1-x^2} \right) y_m = 0\ . \tag{6.80}$$

Comparison with Eq. (4.43) shows that this is precisely of the same form as the differential equation solved by the associated Legendre polynomials P_l^m with eigenvalue $\lambda = -l(l+1)$, where $l = 0, 1, \ldots$ and $m = -l, \ldots, l$. Of course Eq. (6.80) has another solution for $\lambda = -l(l+1)$ and solutions for other values of λ. However, it can be checked, for example using the power series method explained earlier, that the P_l^m are the only solutions which are suitable to define functions on S^2. Conversely, Eq. (4.41) shows that $P_l^m(\pm 1) = 0$ for all $m \neq 0$ so they do have the required behaviour for functions on S^2. The conclusion is that the eigenfunctions and eigenvalues of the Laplacian on the two-sphere are

$$P_l^m(x) e^{im\varphi}\ , \qquad \lambda = -l(l+1)\ , \qquad l = 0, 1, \ldots\ , \quad m = -l, \ldots, l\ . \tag{6.81}$$

It is customary to include a suitable normalisation factor and define the *spherical harmonics*

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\varphi}\ , \qquad l = 0, 1, \ldots\ , \quad m = -l, \ldots, l\ . \tag{6.82}$$

We also note the relation

$$Y_l^0(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta)\ , \tag{6.83}$$

between the Legendre polynomials and the spherical harmonics with $m = 0$.

Exercise 6.7. *Show that $Y_l^m = (-1)^m (Y_l^{-m})^*$. Also show that the first few spherical harmonics are given by*

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}\ , \quad Y_1^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\varphi}\ , \quad Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos\theta\ . \tag{6.84}$$

From what we have seen, the Y_l^m are eigenfunctions

$$\Delta_{S^2} Y_l^m = -l(l+1) Y_l^m \tag{6.85}$$

of Δ_{S^2} and all Y_l^m for $m = -l, \ldots, l$ have the same eigenvalue $\lambda = -l(l+1)$ which, hence, has degeneracy $2l+1$. We already know that Y_l^m and $Y_{l'}^m$ must be orthogonal for $l \neq l'$ but, in fact, due to the orthogonality of the $e^{im\varphi}$ functions the Y_l^m form an orthogonal system. A detailed calculation, based on Eq. (4.44), shows that

$$\langle Y_l^m, Y_{l'}^{m'} \rangle_{S^2} = \delta_{ll'} \delta^{mm'} \ , \tag{6.86}$$

so they form an ortho-normal system on $L^2(S^2)$. In fact, we have

Theorem 6.8. *The spherical harmonics Y_l^m form an orthogonal basis on $L^2(S^2)$.*

Proof. The proof can, for example, be found in Ref. [7]. □

This means every function $f \in L^2(S^2)$ can be expanded in terms of spherical harmonics as

$$f = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_l^m \ , \qquad a_{lm} = \langle Y_l^m, f \rangle_{S^2} = \int_{S^2} (Y_l^m)^* f \, dS \ . \tag{6.87}$$

If f happens to be independent of the angle φ then we only need the $m = 0$ terms in the above expansion and we can write

$$f(x) = \sum_{l=0}^{\infty} a_l P_l(\cos\theta) \ , \qquad a_l = \frac{8\pi^2}{2l+1} \int_{-1}^{1} dx \, P_l(x) f(x) \ . \tag{6.88}$$

6.4.3 Multipole expansion

A two-sphere can be parametrised by the set of all three-dimensional unit vector $\mathbf{n}$. Consider two such unit vectors $\mathbf{n}$ and $\mathbf{n}'$ and the functions $P_l(\mathbf{n} \cdot \mathbf{n}')$, where P_l are the Legendre polynomials. For fixed $\mathbf{n}'$ (say) these functions should have an expansion of the type (6.87) and the precise form of this expansion is given in the following

Lemma 6.3. *For two unit vectors $\mathbf{n}, \mathbf{n}' \in \mathbb{R}^3$ we have*

$$P_l(\mathbf{n} \cdot \mathbf{n}') = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_l^m(\mathbf{n}')^* Y_l^m(\mathbf{n}) \ . \tag{6.89}$$

Proof. Let the function $F : S^2 \times S^2 \to \mathbb{C}$ be defined by the RHS of Eq. (6.89). Our discussion of rotations in Section 9.4 will show that this functions has the property $F(R\mathbf{n}', R\mathbf{n}) = F(\mathbf{n}', \mathbf{n})$, for any rotation R, so F is invariant under simultaneous rotation of its two arguments. Now, let R be a rotation such that $R\mathbf{n}' = \mathbf{e}_3$, so that $F(\mathbf{e}_3, R\mathbf{n}) = F(\mathbf{n}', \mathbf{n})$. The vector $\mathbf{e}_3$ is the north pole of S^2, and corresponds to the coordinate value $x' = \cos\theta' = 1$. Hence, $Y_l^m(\mathbf{e}_3) = 0$ for $m \neq 0$ and

$$Y_l^0(\mathbf{e}_3) = \sqrt{\frac{2l+1}{4\pi}} P_l(1) = \sqrt{\frac{2l+1}{4\pi}} \ . \tag{6.90}$$

Inserting this into the definition of F (the RHS of Eq. (6.89)) we find that $F(\mathbf{e}_3, \tilde{\mathbf{n}}) = P_l(\cos\tilde{\theta}) = P_l(\tilde{\mathbf{n}} \cdot \mathbf{e}_3)$. From this special result for F we can re-construct the entire function using the rotational invariance:

$$F(\mathbf{n}', \mathbf{n}) = F(\mathbf{e}_3, R\mathbf{n}) = P_l((R\mathbf{n}) \cdot \mathbf{e}_3). = P_l(\mathbf{n} \cdot (R^T \mathbf{e}_3)) = P_l(\mathbf{n} \cdot \mathbf{n}') \ , \tag{6.91}$$

which is the desired result. □

We can use this formula to re-write the expansion (4.40) of a Coulomb potential in terms of Legendre polynomials. Setting $\mathbf{r} = r\mathbf{n}$, $\mathbf{r}' = r'\mathbf{n}'$, $\cos\theta = \mathbf{n}\cdot\mathbf{n}'$ in this formula and using Eq. (6.89) we get

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{4\pi}{r}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{1}{2l+1}\left(\frac{r'}{r}\right)^l Y_l^m(\mathbf{n}')^* Y_l^m(\mathbf{n}) \,. \tag{6.92}$$

Let us apply this result to cast the solution (6.34) to the inhomogeneous Laplace equation in a different form. Specialising to the three-dimensional case, we have seen that the unique solution to $\Delta_3\phi = -4\pi\rho$ which approaches zero for $|\mathbf{r}| \to \infty$ is given by

$$\phi(\mathbf{r}) = \int_{\mathbb{R}^3}\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d^3r' \tag{6.93}$$

Inserting the expansion (6.92) this turns into

$$\phi(\mathbf{r}) = 4\pi\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\frac{q_{lm}}{2l+1}\frac{Y_l^m(\mathbf{n})}{r^{l+1}}\,, \qquad q_{lm} = \int_{\mathbb{R}^3} Y_l^m(\mathbf{n}')^*(r')^l\rho(\mathbf{r}')\,d^3r' \,. \tag{6.94}$$

This is called the *multipole expansion* and q_{lm} are called the multipole moments of the source ρ. The multipole expansion gives ϕ as series in inverse powers of the radius r and is, therefore, useful if r is much larger than the extension of the (localised) source ρ. The $l = 0$ term which is proportional to $1/r$ is called the monopole term and its coefficient q_{00} is the total charge of ρ. The $l = 1$ term which decreases as $1/r^2$ is the dipole term with dipole moments q_{1m}, the $l = 2$ term with a $1/r^3$ fall-off is the quadrupole term with quadrupole moments q_{2m} and so on.

Exercise 6.9. *Convert the monopole and dipole term in the multipole expansion (6.94) into Cartesian coordinates (Hint: Use the explicit form of the spherical harmonics in Eq. (6.84).)*

6.5 Laplace equation in three dimensions

Many of the methods we discuss can be applied, with suitable modifications, to the two as well as the three-dimensional case. We denote Cartesian coordinates in $\mathbb{R}^3$ by $\mathbf{r} = (x, y, z) = (x_i)_{i=1,2,3}$ and recall the expressions for the three-dimensional Laplacian from the beginning of the section. First, we discuss the method of image charges which can also be useful in two dimensional cases.

6.5.1 Method of image charges

This method is designed for inhomogeneous problems with non-trivial boundary conditions and the general idea can be explained based on the expression (6.36) for the general solution in the presence of a source ρ. The main goal is to choose the homogeneous solution ϕ_H in Eq. (6.36) so that the required boundary condition is satisfied. Suppose we are interested in solving the problem on $\mathcal{V} \subset \mathbb{R}^3$ with a source ρ which is localised in $\mathcal{V}$ and with boundary conditions on $\partial\mathcal{V}$. The idea is now to insert another, "unphysical" source, $\tilde{\rho}$, into the problem which is localised on $\mathbb{R}^3 \setminus \mathcal{V}$. In this case, the potential generated by $\tilde{\rho}$ is of course harmonic in $\mathcal{V}$ and can serve as a candidate ϕ_H. The art is to choose the charge distribution $\tilde{\rho}$ appropriately, so that the solution (6.36) satisfies the right boundary condition. This is facilitated by the idea of "mirroring" the actual source ρ in $\mathcal{V}$ by a source $\tilde{\rho}$ in $\mathbb{R}^3 \setminus \mathcal{V}$.

Application 6.25. Image charges

The practicalities are probably best explained by examples. First consider a simple case where $\mathcal{V} = \{(x, y, z) \in \mathbb{R}^3 \,|\, x \geq 0\}$ is the half-space with $x \geq 0$ with boundary $\partial\mathcal{V} = \{(0, y, z)\}$ the y-z

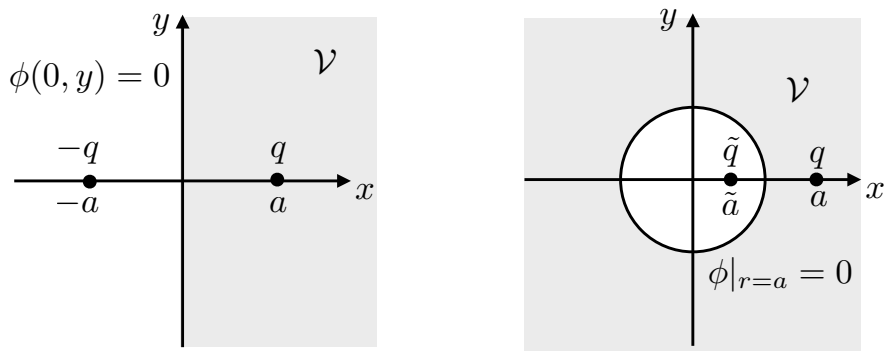

Figure 16: Examples for the method of mirror charges.

plane. (See Fig. 16.) We consider a source ρ which corresponds to a single charge q located at $\mathbf{r}_0 = (a,0,0) \in \mathcal{V}$, where $a > 0$, and demand the boundary condition $\phi|_{x=0} = 0$. The Coulomb potential for this charge

$$\phi_I(\mathbf{r}) = \frac{q}{|\mathbf{r}-\mathbf{r}_0|} \tag{6.95}$$

does satisfy the correct equation, $\Delta\phi_I = -4\pi\rho$, but does not satisfy the boundary condition. Suppose, we introduce the mirror charge density $\tilde{\rho}$ by a single point charge with charge $-q$ located at $\tilde{\mathbf{r}}_0 = (-a,0,0) \in \mathbb{R}^3 \setminus \mathcal{V}$ leading to a potential

$$\phi_H(\mathbf{r}) = -\frac{q}{|\mathbf{r}-\tilde{\mathbf{r}}_0|} \tag{6.96}$$

which satisfies $\Delta\phi_H = 0$ in $\mathcal{V}$. Then

$$\phi(\mathbf{r}) = \phi_H(\mathbf{r}) + \phi_I(\mathbf{r}) = -\frac{q}{|\mathbf{r}-\tilde{\mathbf{r}}_0|} + \frac{q}{|\mathbf{r}-\mathbf{r}_0|} \tag{6.97}$$

satisfies $\Delta\phi = -4\pi\rho$ in $\mathcal{V} = \{x \geq 0\}$ as well as the required boundary condition $\phi|_{x=0} = 0$. For a slightly more complicated example, consider the region $\mathcal{V} = \{\mathbf{r} \in \mathbb{R}^3 \,|\, |\mathbf{r}| \geq b\}$ outside a sphere with radius b with the charge density ρ generated by a singe charge q located at $\mathbf{r}_0 = (a,0,0)$, where $a > b$ and boundary condition $\phi|_{|\mathbf{r}|=b} = 0$. (See Fig. 16.) We try a single mirror charge inside the sphere with charge $\tilde{q}$ and located at $\tilde{\mathbf{r}}_0 = (\tilde{a},0,0)$, where $\tilde{a} < b$. Then

$$\phi(\mathbf{r}) = \frac{\tilde{q}}{|\mathbf{r}-\tilde{\mathbf{r}}_0|} + \frac{q}{|\mathbf{r}-\mathbf{r}_0|} \tag{6.98}$$

satisfies $\Delta\phi = -4\pi\rho$ outside the sphere. We should try to fix $\tilde{q}$ and $\tilde{a}$ so that $\phi|_{|\mathbf{r}|=b} = 0$ and a short calculation shows this is satisfied if

$$\tilde{q} = -\frac{b}{a}q\,, \qquad \tilde{a} = \frac{b^2}{a}\,. \tag{6.99}$$

6.5.2 Cartesian coordinates

We would like to solve the three-dimensional homogeneous Laplace equation

$$\Delta\phi = 0 \,, \qquad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \,. \tag{6.100}$$

One way to proceed is by a separation Ansatz just as we did in the two-dimensional case with Cartesian coordinates but here we would like to follow a related by slightly different logic based on a Fourier series expansion.

Exercise 6.10. *Solve the three-dimensional Laplace equation in Cartesian coordinates by separation of variables.*

Application 6.26. Laplacian in a box

Suppose we are interested in a solution in the box $\mathcal{V} = [0,a] \times [0,b] \times [0,c]$ and assume, for now, that $\phi = 0$ an all boundaries with $x = 0, a$ and $y = 0, b$. For fixed z, any function $\phi \in L^2([0,a] \times [0,b])$ with these boundary conditions can be expanded in a (double) sine Fourier series

$$\phi(x,y,z) = \sum_{k,l=1}^{\infty} Z_{k,l}(z) \sin\left(\frac{\pi k x}{a}\right) \sin\left(\frac{\pi l y}{b}\right) \,, \tag{6.101}$$

where $Z_{k,l}$ are functions of z. Inserting this into the Laplace equation gives an ordinary differential equation

$$Z''_{k,l} = \nu_{k,l}^2 Z_{k,l} \,, \qquad \nu_{k,l} = \sqrt{\left(\frac{\pi k}{a}\right)^2 + \left(\frac{\pi l}{b}\right)^2} \,, \tag{6.102}$$

for each $Z_{k,l}$ whose general solution is

$$Z_{k,l}(z) = A_{k,l} e^{\nu_{k,l} z} + B_{k,l} e^{-\nu_{k,l} z} \,, \tag{6.103}$$

with arbitrary constants $A_{k,l}$ and $B_{k,l}$. Combining this result with Eq. (6.101) leads to the general solution

$$\phi(x,y,z) = \sum_{k,l=1}^{\infty} \left(A_{k,l} e^{\nu_{k,l} z} + B_{k,l} e^{-\nu_{k,l} z}\right) \sin\left(\frac{\pi k x}{a}\right) \sin\left(\frac{\pi l y}{b}\right) \tag{6.104}$$

of the Laplace equation in the box $\mathcal{V} = [0,a] \times [0,b] \times [0,c]$ which vanishes on the boundaries at $x = 0, a$ and $y = 0, b$. To fix the remaining constants we have to specify boundary conditions at $z = 0$ and $z = c$. Suppose we demand that $\phi(x,y,0) = 0$. This can be achieved by setting $A_{k,l} = -B_{k,l} =: a_{k,l}/2$ so that the solution becomes

$$\phi(x,y,z) = \sum_{k,l=1}^{\infty} a_{k,l} \sinh(\nu_{k,l} z) \sin\left(\frac{\pi k x}{a}\right) \sin\left(\frac{\pi l y}{b}\right) \,. \tag{6.105}$$

Finally, assume for the last boundary at $z = c$ that $\phi(x,y,c) = h(x,y)$ for a given function h. Then setting $z = c$ in Eq. (6.105) is a (double) sine Fourier series for the function h and we can compute the remaining parameters $a_{k,l}$ by standard Fourier series techniques as

$$a_{k,l} = \frac{4}{ab \sinh(\nu_{k,l} c)} \int_{[0,a]\times[0,b]} dx\, dy \, \sin\left(\frac{\pi k x}{a}\right) \sin\left(\frac{\pi l y}{b}\right) h(x,y) \,. \tag{6.106}$$

Of course this calculation can be repeated, and leads to a similar result, if another one of the six boundary planes is subject to a non-trivial boundary condition while ϕ is required to vanish on the other five boundary planes. In this way we get six solutions, all similar to the above but with coordinates permuted and constants determined as appropriate. The sum of these six solutions then solves the homogeneous Laplace equation with non-trivial boundary conditions on all six sides.

6.5.3 Cylindrical coordinates

To solve the Laplace equation in cylindrical coordinates, using the three-dimensional Laplacian given in Eq. (6.15), we proceed in close analogy with the previous case. We assume that we would like a solution on the cylinder $\mathcal{V} = \{(r, \varphi, z) \,|\, r \in [0, a]\,,\ z \in [0, L]\}$ with radius a and height L and we assume the boundary conditions are such that ϕ vanishes on the side of the cylinder, so $\phi|_{r=a} = 0$. Since φ is an angular coordinate we can use a Fouries series for the φ dependence and, it turns out, it is useful to use the Bessel functions $\hat{J}_{\nu k}$, defined in Eq. (5.63), which form an ortho-normal basis of $L^2([0, a])$. We recall from Eq. (5.64) that these functions satisfy the eigenvalue equations

$$\left(\frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) - \frac{\nu^2}{r^2}\right)\hat{J}_{\nu,k} = -\frac{z_{\nu k}^2}{a^2}\hat{J}_{\nu k}\,, \tag{6.107}$$

where $z_{\nu k}$ are the zeros of the Bessel functions. The fact that the operator on the LHS is similar to the radial part of the Laplacian in cylindrical coordinates (6.15) is part of the motivation of why we are using the Bessel functions. We can then expand

$$\phi(r, \varphi, z) = \sum_{k=1}^{\infty}\sum_{m=0}^{\infty} Z_{km}(z)\hat{J}_{\nu k}(r)(a_{km}\sin(m\varphi) + b_{km}\cos(m\varphi)) \tag{6.108}$$

with as yet undetermined functions Z_{km}. Inserting this expansion into the Laplace equation in cylindrical coordinates (6.15) and using the eigenvalue equation (6.107) for the Bessel functions leads to the differential equation

$$Z''_{km} = \frac{z_{mk}^2}{a^2}Z_{km}\,, \tag{6.109}$$

provided the previously arbitrary type, ν, of the Bessel function is fixed to be $\nu = m$. Solving this equation and inserting the solutions back into the expansion (6.108) gives

$$\phi(r, \varphi, z) = \sum_{k=1}^{\infty}\sum_{m=0}^{\infty}\left(A_{km}\exp\left(\frac{z_{mk}z}{a}\right) + B_{km}\exp\left(-\frac{z_{mk}z}{a}\right)\right)\hat{J}_{mk}(r)(a_{km}\sin(m\varphi) + b_{km}\cos(m\varphi)) \tag{6.110}$$

To fix the remaining coefficients we have to impose boundary conditions at the bottom and top of the cylinder. For example, if we demand at the bottom that $\phi|_{z=0} = 0$ then

$$\phi(r, \varphi, z) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sinh\left(\frac{z_{mk}z}{a}\right)\hat{J}_{mk}(r)(a_{km}\sin(m\varphi) + b_{km}\cos(m\varphi))\,. \tag{6.111}$$

Finally, if we demand the boundary condition $\phi_{z=L} = h$, for some function $h = h(r, \varphi)$ at the top of the cylinder the remaining coefficients a_{km} and b_{km} can be determined by combining the orthogonality properties (3.22) of the Fourier series with those of the Bessel functions (5.65).

Exercise 6.11. *Use orthogonality properties of the Fourier series and the Bessel functions to find expressions for the coefficients a_{mk} and b_{mk} in the solution (6.111), so that the boundary condition $\phi|_{z=L} = h$ is satisfied for a function $h = h(r, \varphi)$.*

6.5.4 Spherical coordinates

To discuss the three-dimensional Laplace equation in spherical coordinates it is very useful to recall that the three-dimensional Laplace operator can be written as

$$\Delta_{3,\text{sph}} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\,\Delta_{S^2}\ , \tag{6.112}$$

where Δ_{S^2} is the Laplacian on the two-sphere. Also recall that we have the spherical harmonics Y_{lm} which form an orthonormal basis of $L^2(S^2)$ and are eigenfunctions of Δ_{S^2}, with

$$\Delta_{S^2} Y_{lm} = -l(l+1)Y_{lm}\ . \tag{6.113}$$

All this suggest we should start with an expansion

$$\phi(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} R_{lm}(r)Y_{lm}(\theta,\varphi)\ . \tag{6.114}$$

Inserting this expansion into the homogeneous Laplace equation, $\Delta\phi = 0$, and using the eigenvector property (6.113) leads to the differential equations

$$\frac{d}{dr}\left(r^2 R'_{lm}\right) = l(l+1)R_{lm} \tag{6.115}$$

with general solutions

$$R_{lm}(r) = A_{lm}r^l + B_{lm}r^{-l-1}\ , \tag{6.116}$$

for constants A_{lm} and B_{lm}. Inserting this back into the expansion (6.114) leads to the general solution to the homogeneous Laplace equation in spherical coordinates:

$$\phi(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} (A_{lm}r^l + B_{lm}r^{-l-1})Y_{lm}(\theta,\varphi)\ . \tag{6.117}$$

The arbitrary constants A_{lm} and B_{lm} are fixed by boundary conditions and, given the choice of coordinates, they are relatively easy to implement if they are imposed on spherical boundaries. We also note that for problems with azimutal symmetry, that is, when the boundary conditions and ϕ are independent of φ, we only require the $m = 0$ terms in the above expansion. Since the Y_{l0} are proportional to the Legendre polynomials this means, after a re-definition of the constants, that, for such problems, we have the simplified expansion

$$\phi(r,\theta) = \sum_{l=0}^{\infty}(A_l r^l + B_l r^{-l-1})P_l(\cos\theta)\ . \tag{6.118}$$

Application 6.27. Laplace equations with boundary conditions on a sphere

We conclude with an example on how to fix the constants by imposing boundary conditions. Suppose that $\mathcal{V} = \{(r,\theta,\varphi)\,|\,r \le a\}$ is the ball with radius a and we demand that $\phi|_{r=a} = h$, where $h = h(\theta,\varphi)$ is a given function. Since ϕ needs to be smooth in the interior of the ball we first conclude that all terms with negative powers in r in the expansion (6.114) must vanish, so $B_{lm} = 0$. The remaining constant A_{lm} can be fixed by imposing $\phi|_{r=a} = h$ and using the orthogonality relations (6.87) of the spherical harmonics. This leads to

$$h = \phi|_{r=a} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} A_{lm}a^l Y_{lm} \qquad \Rightarrow \qquad A_{lm} = \frac{1}{a^l}\langle Y_{lm}, h\rangle_{S^2} = \frac{1}{a^l}\int_{S^2}(Y_{lm})^* h\, dS\ . \tag{6.119}$$

On the other hand, we might be interested in the region outside a sphere of radius a, so $\mathcal{V} = \{(r,\theta,\varphi)\,|\,r \geq a\}$, with the same boundary condition $\phi|_{r=a} = h$ at $r = a$. In this case we also have to specify a boundary condition at "infinity" which can, for example, be done by demanding that $\phi \to 0$ as $r \to \infty$. This last condition implies that all terms with positive powers of r in the expansion (6.114) must vanish, so that $A_{lm} = 0$. The remaining coefficients B_{lm} are then fixed by the boundary condition at $r = a$ and are, in analogy with the previous case, given by

$$B_{lm} = a^{l+1}\langle Y_{lm}, h\rangle_{S^2} = a^{l+1}\int_{S^2}(Y_{lm})^* h\, dS\ . \tag{6.120}$$

7 Distributions

The Dirac delta function, $\delta(x)$, introduced by Dirac in 1930, is an object frequently used in theoretical physics. It is usually "defined" as a "function" on $\mathbb{R}$ with properties

$$\delta(x) = \begin{cases} 0 & \text{for} \quad x \neq 0 \\ \infty & \text{for} \quad x = 0 \end{cases} , \qquad \int_{\mathbb{R}} dx\, \delta(x) = 1 \, . \tag{7.1}$$

Of course assigning a function "value" of infinity at $x = 0$ does not make sense. Even if we are prepared to overlook this we know from our earlier discussion of integrals that a function which is zero everywhere, except at one point, must integrate to zero (since a single point is a set of measure zero), rather than to one. Hence, the above "definition" is at odds with basic mathematics, yet it is routinely used in a physics context. The purpose of this chapter is to introduce the proper mathematical background within which to understand the Dirac delta - the theory of *distributions* - and to explain how, in view of this theory, we can get away with mathematically questionable equations such as Eq. (7.1). The fact that the Dirac delta is not a function but rather a distribution has far-reaching consequences for many equations routinely used in physics. For example, the "Coulomb potential" $1/r$, where $r = |\mathbf{x}|$ is the three-dimensional radius, satisfies an equation often stated as

$$\Delta \frac{1}{r} = -4\pi\, \delta(\mathbf{x}) \, , \tag{7.2}$$

where $\Delta = \sum_{i=1}^{3} \frac{\partial^2}{\partial x_i^2}$ is the three-dimensional Laplacian. (We will prove the correct version of this equation below.) Since the right-hand side of this equation should be understood as a distribution so must be the left-hand side and this leads to a whole range of questions. With these motivations in mind we now begin by properly defining distributions.

7.1 Basic definitions

The theory of distributions was first developed by Schwarz in 1945. The first step is to introduce *test functions* which are functions with particularly nice analytic properties. There are several possible choices for test function spaces but for the present purpose we adopt the following

Definition 7.1. *The space of test functions is the vector space $\mathcal{D} = \mathcal{D}(\mathbb{R}^n) := \mathcal{C}_c^\infty(\mathbb{R}^n)$ of infinitely many times differentiable functions with compact support and a function $\varphi \in \mathcal{D}$ is called a test functions.*

We say a sequence (φ_k) of test functions converges to a function $\varphi \in \mathcal{D}$ iff

(i) There is an $R > 0$ such that $\varphi_k(\mathbf{x}) = 0$ for all k and all $\mathbf{x} \in \mathbb{R}^n$ with $|\mathbf{x}| > R$.

(ii) (φ_k) and all its derivatives converge to φ uniformly for all $\mathbf{x} \in \mathbb{R}^n$ with $|\mathbf{x}| \leq R$.

In this case we write $\varphi_k \xrightarrow{\mathcal{D}} \varphi$.

Note that the above version of convergence is very strong. An example of a test function is

$$\varphi(\mathbf{x}) = \begin{cases} \exp\left(-\frac{a^2}{a^2 - |\mathbf{x}|^2}\right) & \text{for} \quad |\mathbf{x}| < a \\ 0 & \text{for} \quad |\mathbf{x}| \geq a \end{cases} \tag{7.3}$$

Clearly, this function has compact support and the structure of the exponential ensures that it drops to zero at $|\mathbf{x}| = a$ in an infinitely many times differentiable way. All polynomials times the above φ are also test functions so the vector space $\mathcal{D}$ is clearly infinite dimensional. We are now ready to define *distributions*.

Definition 7.2. *A distribution is a linear, continuous map*

$$\begin{array}{rccc} T: & \mathcal{D}(\mathbb{R}^n) & \longrightarrow & \mathbb{R} \\ & \varphi & \longrightarrow & T[\varphi] \end{array} . \tag{7.4}$$

Continuity means that $\varphi_k \stackrel{\mathcal{D}}{\to} \varphi$ *implies* $T[\varphi_k] \to T[\varphi]$. *The space of all distributions is called* $\mathcal{D}' = \mathcal{D}'(\mathbb{R}^n)$.

Clearly, the space of distributions $\mathcal{D}'$ is a vector space, a sub-space of the dual vector space to $\mathcal{D}$. Hilbert spaces are isomorphic to their dual spaces but the test function space $\mathcal{D}$ is not a Hilbert space (for example, it is not complete relative to the standard integral norms). In some sense, the test function space $\mathcal{D}$ is rather "small" and, hence, we expect the dual space to be large. To get a better feeling for what this means, we should now consider a few examples of distributions.

7.1.1 Examples of distributions

Let $f \in \mathcal{C}(\mathbb{R}^n)$ be a continuous function. We can define a distribution T_f, associated to f by

$$T_f[\varphi] := \int_{\mathbb{R}^n} dx^n \, f(\mathbf{x})\varphi(\mathbf{x}) \tag{7.5}$$

Clearly, T_f is linear and continuous and, hence, it is a distribution. It is also not too hard to show that the map $f \to T_f$ is injective, so that the continuous functions are embedded in the space of distributions.

We can slightly generalise the above construction. For $f \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}^n)$ (a locally integrable function) the above definition of T_f still works. However, the map $f \to T_f$ is no longer injective - two functions which only differ on a set of measure zero are mapped to the same distribution.

For $\mathbf{a} \in \mathbb{R}^n$ the Dirac δ-distribution $\delta_{\mathbf{a}} \in \mathcal{D}'(\mathbb{R}^n)$ is defined as

$$\delta_{\mathbf{a}}[\varphi] := \varphi(\mathbf{a}) \, , \tag{7.6}$$

and this provides the correct version of the equation

$$\int_{\mathbb{R}^n} dx^n \, \varphi(\mathbf{x})\delta(\mathbf{x}-\mathbf{a}) = \varphi(\mathbf{a}) \, , \tag{7.7}$$

which is frequently written in a physics context. The integral on the LHS is purely symbolic - the reason one usually gets away with using this notation is that both the integral and the distribution are linear.

The above examples show that distributions should be thought of as generalisations of functions. They contain functions via the map $f \to T_f$ but, in addition, they also contain "more singular" objects such as the Dirac-delta which cannot be interpreted as a function. The idea is now to generalise many of the tools of analysis from functions to distributions. We begin with a definition of convergence for distributions.

7.1.2 Convergence of distributions

Definition 7.3. *A sequence* (T_k) *of distributions is said to converge to a* $T \in \mathcal{D}'$ *iff* $T_k[\varphi] \to T[\varphi]$ *for all test functions* $\varphi \in \mathcal{D}$. *In this case we write* $T_k \stackrel{\mathcal{D}'}{\to} T$.

The idea of this definition is that convergence of distributions is "tested" with functions $\varphi \in \mathcal{D}$ and hence the name "test functions". The above notion of convergence can be used to gain a better understanding of the Dirac delta. Although by itself not a function the Dirac delta can be obtained as a limit of functions, as explained in the following

Theorem 7.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be an integrable function with $\int_{\mathbb{R}^n} dx^n\, f(\mathbf{x}) = 1$. For $\epsilon > 0$ define the functions $f_\epsilon(\mathbf{x}) := \frac{1}{\epsilon^n} f\left(\frac{\mathbf{x}}{\epsilon}\right)$. Then we have*

$$T_{f_\epsilon} \xrightarrow{\mathcal{D}'} \delta_{\mathbf{0}} \quad as \quad \epsilon \longrightarrow 0\,. \tag{7.8}$$

Proof. We need to show convergence in the sense of distributions, as defined in Def. 7.3. This means we must show that

$$\lim_{\epsilon\to 0} T_{f_\epsilon}[\varphi] = \varphi(0) \tag{7.9}$$

for all test functions $\varphi \in \mathcal{D}$. A simple computation gives

$$T_{f_\epsilon}[\varphi] = \int_{\mathbb{R}^n} dx^n\, f_\epsilon(\mathbf{x})\varphi(\mathbf{x}) = \int_{\mathbb{R}^n} dx^n\, \frac{1}{\epsilon^n} f\left(\frac{\mathbf{x}}{\epsilon}\right)\varphi(\mathbf{x}) \overset{\mathbf{x}=\epsilon\mathbf{y}}{=} \int_{\mathbb{R}^n} dy^n\, f(\mathbf{y})\varphi(\epsilon\mathbf{y})\,. \tag{7.10}$$

The integrand on the RHS side is bounded by an integrable function since $|f(\mathbf{y})\varphi(\epsilon\mathbf{y})| < Kf(\mathbf{y})$, for some constant K. This means we can pull the limit $\epsilon \to 0$ into the integral so that

$$\lim_{\epsilon\to 0} T_{f_\epsilon}[\varphi] = \int_{\mathbb{R}^n} dy^n\, f(\mathbf{y}) \lim_{\epsilon\to 0} \varphi(\epsilon\mathbf{y}) = \varphi(0) \int_{\mathbb{R}^n} dy^n\, f(\mathbf{y}) = \varphi(0)\,. \tag{7.11}$$

□

In the one-dimensional case, a possible choice for the functions f and f_ϵ in Theorem 7.1 is provided by the Gaussians

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \Rightarrow \quad f_\epsilon(x) = \frac{1}{\sqrt{2\pi}\,\epsilon} e^{-\frac{x^2}{2\epsilon^2}}\,. \tag{7.12}$$

The graph of f_ϵ for decreasing values of ϵ is shown in Fig. 17 and illustrates the idea behind Theorem 7.1.

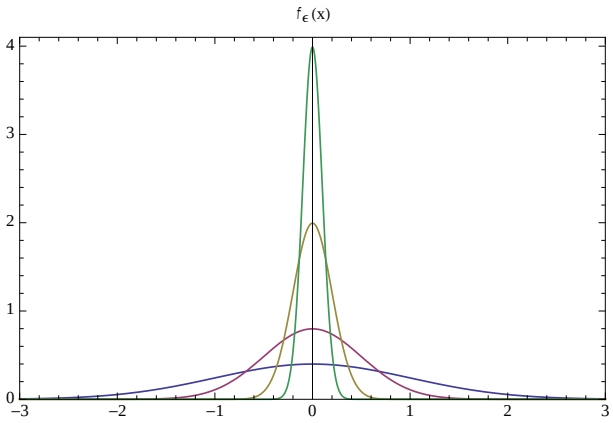

Figure 17: The graph of the function f_ϵ in Eq. (7.12) for $\epsilon = 1, \frac{1}{2}, \frac{1}{5}, \frac{1}{10}$.

For decreasing ϵ the function f_ϵ becomes more and more peaked around $x = 0$ and tends to δ_0 in the limit. However, note that f_ϵ does actually not converge to any function in the usual sense, as $\epsilon \to 0$. Rather, convergence only happens in the weaker sense of distributions, as stated in Theorem 7.1.

We can use the above representation of the Dirac delta as a limit to introduce a new class of distributions. For a function g in $\mathcal{C}^\infty(\mathbb{R})$ we can define the distribution $\delta_g \in \mathcal{D}'(\mathbb{R})$ by

$$\delta_g := \lim_{\epsilon\to 0} T_{f_\epsilon \circ g}\,. \tag{7.13}$$

To see what this means, we assume that we can split up $\mathbb{R}$ into intervals I_α such that $g|_{I_\alpha}$ is invertible with inverse g_α^{-1}. Then, following the calculation in Theorem 7.1, we have

$$\delta_g[\varphi] = \lim_{\epsilon\to 0}\sum_\alpha \int_{I_\alpha} dx\, \frac{1}{\epsilon} f\left(\frac{g(x)}{\epsilon}\right)\varphi(x) \overset{y=g(x)/\epsilon}{=} \lim_{\epsilon\to 0}\sum_\alpha \int_{g(I_\alpha)/\epsilon} dy\, f(y)\frac{\varphi(g_\alpha^{-1}(\epsilon y))}{|g'(g_\alpha^{-1}(\epsilon y))|} = \sum_{a:g(a)=0}\frac{\varphi(a)}{|g'(a)|}\ .$$

Re-writing the RHS in terms of Dirac deltas we find

$$\delta_g = \sum_{a:g(a)=0}\frac{1}{|g'(a)|}\delta_a\ . \tag{7.14}$$

This result suggests an intuitive interpretation of δ_g as a sum of Dirac deltas located at the zeros of the function g. As an example, consider the function $g(x) = x^2 - c^2$, where $c > 0$ is a constant. This function has two zeros at $\pm c$ and $|g'(\pm c)| = 2c$. Inserting this into Eq. (7.14) gives

$$\delta_{x^2-c^2} = \frac{1}{2c}(\delta_c + \delta_{-c})\ . \tag{7.15}$$

It is useful to translate the above equations into the notation commonly used in physics. There, δ_g is usually written as $\delta(g(x))$ and Eq. (7.14) takes the form

$$\delta(g(x)) = \sum_{a:g(a)=0}\frac{1}{|g'(a)|}\delta(x-a)\ , \tag{7.16}$$

while the example (7.15), written in this notation, reads

$$\delta(x^2-c^2) = \frac{1}{2c}(\delta(x-c)+\delta(x+c))\ . \tag{7.17}$$

7.1.3 Derivatives of distributions

We are now ready for the next step in extending the standard tools of analysis to distributions. We would like to define derivatives of distributions. For convenience, we introduce the short-hand notation $D_i := \frac{\partial}{\partial x_i}$ for the partial derivatives. Our goal is to make sense of the expression D_iT for a distribution $T \in \mathcal{D}'$. For a distribution T_f, associated to a differentiable function f via the rule (7.5) there is a natural way to define the derivative, namely

$$D_iT_f := T_{D_if}\ . \tag{7.18}$$

In other words, the derivative of such a distribution is simply the distribution associated to the derivative of the function. Any general definition of D_iT should reduce to this rule for distributions of the form T_f so that the normal rules for differentiating functions are preserved. Writing the rule (7.18) out explicitly, using Eq. (7.5), we have

$$D_iT_f[\varphi] = T_{D_if}[\varphi] = \int_{\mathbb{R}^n} dx^n\, D_if(\mathbf{x})\varphi(\mathbf{x}) = -\int_{\mathbb{R}^n} dx^n\, f(\mathbf{x})D_i\varphi(\mathbf{x}) = T_f[-D_i\varphi]\ . \tag{7.19}$$

The left and right-hand sides of this chain of equations make sense if we drop the subscript f and this is used to define the derivative of distributions in general.

Definition 7.4. *The derivative D_iT of a distribution $T \in \mathcal{D}'$ is defined by $D_iT[\varphi] = T[-D_i\varphi]$.*

From this definition, D_iT is clearly linear but for it to be a distribution we still have to check that it is continuous. Suppose we have a convergent sequence $\varphi_k \overset{\mathcal{D}}{\to} \varphi$. It follows that $-D_i\varphi_k \overset{\mathcal{D}}{\to} -D_i\varphi$ (due to the strong notion of convergence in $\mathcal{D}$, as in Def. 7.1) and, from continuity of T, this implies $T[-D_i\varphi_k] \to T[-D_i\varphi]$ and, finally, using Def. 7.4, that $D_iT[\varphi_k] \to D_iT[\varphi]$. This means D_iT is indeed continuous and, hence, a distribution.

Application 7.28. Unusual convergence of distributions

Consider the sequence of functions $f_k : \mathbb{R} \to \mathbb{R}$ defined by $f_k(x) = \sin(kx)$, where $k = 1, 2, \ldots$, so a sequence of sine functions with increasing frequency. Clearly, this sequence does not converge to any function in the usual sense. What about convergence of the associated sequence T_{f_k} of distributions? To work this out, we define the sequence of functions (g_k) with $g_k(x) = -\frac{1}{k}\cos(kx)$. We have $Dg_k = f_k \;\Rightarrow\; DT_{g_k} = T_{f_k}$ and, since $g_k \to 0$ uniformly, as $k \to \infty$, it follows that

$$T_{g_k} \xrightarrow{\mathcal{D}'} T_0 \qquad \overset{D}{\Rightarrow} \qquad T_{f_k} \xrightarrow{\mathcal{D}'} T_0 \;, \tag{7.20}$$

where T_0 is the distribution associated to the function identical to zero. This is a surprising result which shows that distributions can behave quite differently from functions: While the sequence of functions (f_k) does not converge at all, the associated sequence of distributions (T_{f_k}) converges to the zero distribution.

Application 7.29. Heaviside function

The *Heaviside function* $\theta : \mathbb{R} \to \mathbb{R}$ is defined as

$$\theta(x) := \begin{cases} 1 & \text{for} \quad x \geq 0 \\ 0 & \text{for} \quad x < 0 \end{cases} \;. \tag{7.21}$$

As a function, θ is not differentiable at $x = 0$ (since it is not even continuous there) but we can still ask about the differential of the associated Heaviside distribution T_θ. A short calculation shows that

$$DT_\theta[\varphi] = T_\theta[-D\varphi] = -\int_{\mathbb{R}} dx\, \theta(x)\varphi'(x) = -\int_0^\infty dx \varphi'(x) = \varphi(0) = \delta_0[\varphi] \tag{7.22}$$

and, hence,

$$DT_\theta = \delta_0 \;. \tag{7.23}$$

The Dirac delta is the derivative of the Heaviside distribution. In a physics context this is frequently written as

$$\frac{d}{dx}\theta(x) = \delta(x) \;, \tag{7.24}$$

ignoring the fact that θ, as a function, is actually not differentiable.

Application 7.30. Derivative of Dirac delta

How does the derivative $\delta_a' = D\delta_a$ of the Dirac delta act on test functions? A short calculations shows that

$$\delta_a'[\varphi] = D\delta_a[\varphi] = \delta_a[-D\varphi] = -\varphi'(a) \;. \tag{7.25}$$

While the Dirac delta δ_a produces the function value of the test function at a its derivative, δ_a', leads to the (negative) derivative of the test function at a. In a physics context, this equation is also written as

$$\int dx\, \delta'(x - a)\varphi(x) = -\varphi'(a) \tag{7.26}$$

and it is "proven" by partial integration and then using Eq. (7.7).

7.2 Convolution of distributions*

We have already encountered convolutions of functions in the context of Fourier transforms. We would now like to introduce convolutions between functions and distributions. Suppose we have a locally integrable function $f \in \mathcal{C}^\infty_{\text{loc}}(\mathbb{R}^n)$ and a test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$. From Eq. (3.55) the convolution of these two functions is given by

$$(f \star \varphi)(\mathbf{x}) = \int_{\mathbb{R}^n} dy^n \, f(\mathbf{y})\varphi(\mathbf{x}-\mathbf{y}) \, . \tag{7.27}$$

Of course we would like to extend this definition to convolutions such that the above formula is reproduced for distributions T_f, associated to a function f. Hence, we start by re-writing Eq. (7.27) in terms of T_f and, to this end, we introduce the translation $\tau_{\mathbf{x}}$ by $\mathbf{x} \in \mathbb{R}^n$, defined by

$$(\tau_{\mathbf{x}}\varphi)(\mathbf{y}) := \varphi(\mathbf{x}-\mathbf{y}) \, . \tag{7.28}$$

It is easy to see that with this definition, the convolution (7.27) can be re-written as

$$(f \star \varphi)(\mathbf{x}) = T_f[\tau_{\mathbf{x}}\varphi] \, . \tag{7.29}$$

The RHS of this equation can be used to define convolutions between distributions and functions.

Definition 7.5. *For a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ and a test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ the convolution $T \star \varphi : \mathbb{R}^n \to \mathbb{R}$ is a function defined by*

$$(T \star \varphi)(\mathbf{x}) := T[\tau_{\mathbf{x}}\varphi] \, . \tag{7.30}$$

As an example, consider a convolution with the Dirac delta $\delta_{\mathbf{0}}$.

$$(\delta_{\mathbf{0}} \star \varphi)(\mathbf{x}) = \delta_{\mathbf{0}}[\tau_{\mathbf{x}}\varphi] = (\tau_{\mathbf{x}}\varphi)(\mathbf{0}) = \varphi(\mathbf{x}) \qquad \Rightarrow \qquad \delta_{\mathbf{0}} \star \varphi = \varphi \, . \tag{7.31}$$

This means that the Dirac delta $\delta_{\mathbf{0}}$ can be seen as the identity of the convolution operation.

The result of a convolution $T \star \varphi$ is a function and it is an obvious question what the properties of this function are. Some answers are given by the following

Theorem 7.2. *For a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ and a test function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ the convolution $T \star \varphi$ is a differentiable function and we have*

$$D_i(T \star \varphi) \overset{(a)}{=} T \star (D_i\varphi) \overset{(b)}{=} (D_iT) \star \varphi \, . \tag{7.32}$$

Proof. For the equality (a) we evaluate the LHS explicitly:

$$\begin{aligned} D_i(T \star \varphi)(\mathbf{x}) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}(T[\tau_{\mathbf{x}+\epsilon\mathbf{e}_i}\varphi] - T[\tau_{\mathbf{x}}\varphi]) = \lim_{\epsilon \to 0} T\left[\frac{1}{\epsilon}(\tau_{\mathbf{x}+\epsilon\mathbf{e}_i}\varphi - \tau_{\mathbf{x}}\varphi)\right] \\ &= T[\tau_{\mathbf{x}} D_i\varphi] = (T \star D_i\varphi)(\mathbf{x}) \, . \end{aligned}$$

Proving the equality (b) is even more straightforward:

$$D_i(T \star \varphi)(\mathbf{x}) = T[\tau_{\mathbf{x}} D_i\varphi] = T[-D_i\tau_{\mathbf{x}}\varphi] = D_iT[\tau_{\mathbf{x}}\varphi] = ((D_iT) \star \varphi)(\mathbf{x}) \, .$$

□

7.3 Fundamental solutions - Green functions*

We have already seen Green functions appear a number of times, notably in the context of ordinary linear differential equations and in the context of the Laplace equation. They are, in fact, quite a general concept which can be formulated concisely using distributions. To get a rough idea of how this works let us first discuss this using the physics language where we work with a Dirac delta "function", $\delta(\mathbf{x})$.

Suppose, we have a differential operator L, (typically second order) and we are trying to solve the equation $L\phi(\mathbf{x}) = \rho(\mathbf{x})$. (The Laplace equation with source ρ is an example but the discussion here is for a general differential operator.) Suppose we have found a "function" G satisfying

$$LG(\mathbf{x}) = \delta(\mathbf{x}) \; . \tag{7.33}$$

In terms of G we can immediately write down a solution to our equation, namely

$$\phi(\mathbf{x}) = \int_{\mathbb{R}^n} dy^n \, G(\mathbf{x}-\mathbf{y})\rho(\mathbf{y}) \; . \tag{7.34}$$

To check that this is indeed a solution simply work out

$$L_{\mathbf{x}}\phi(\mathbf{x}) = \int_{\mathbb{R}^n} dy^n \, L_{\mathbf{x}}G(\mathbf{x}-\mathbf{y})\rho(\mathbf{y}) = \int_{\mathbb{R}^n} dy^n \, \delta(\mathbf{x}-\mathbf{y})\rho(\mathbf{y}) = \rho(\mathbf{x}) \; . \tag{7.35}$$

For example, for the (three-dimensional) Laplace equation $\Delta\phi(\mathbf{x}) = -4\pi\rho(\mathbf{x})$, we have the Green function $G(\mathbf{x}) = 1/r$, where $r = |\mathbf{x}|$, which satisfies

$$\Delta\frac{1}{r} = -4\pi\delta(\mathbf{x}) \; . \tag{7.36}$$

(This statement will be justified in Theorem 7.4 below.) Hence, we have a solution

$$\phi(\mathbf{x}) = \int_{\mathbb{R}^3} dy^3 \, \frac{\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \tag{7.37}$$

While the above discussion is what you will normally find in the physics literature we know by now that it is mathematically questionable and has to be justified by a proper treatment in terms of distributions. The general theorem facilitating this is

Theorem 7.3. *If the distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ satisfies $LE = \delta_{\mathbf{0}}$ for a differential operator L and $\rho \in \mathcal{D}(\mathbb{R}^n)$ is a testfunction, then $\phi := E \star \rho$ satisfies $L\phi = \rho$.*

Proof. Given our preparation the proof is now quite easy:

$$L\phi = L(E \star \rho) = (LE) \star \rho = \delta_{\mathbf{0}} \star \rho = \rho \; . \tag{7.38}$$

Here we have used Theorem 7.2 and the fact that the Dirac delta is the identity under convolution, as in Eq. (7.31). □

In the above theorem, the distribution E, also called a *fundamental solution* for L, is the analogue of the Green function and a general solution to the inhomogeneous equation with source ρ is obtained by convoluting E with ρ.

Of course, for this theorem to be of practical use we first have to work out a fundamental solution for a given differential operator L. For the Laplace operator, $L = \Delta$, the following theorem provides such a fundamental solution:

Theorem 7.4. *The distribution $T_{1/r} \in \mathcal{D}'(\mathbb{R}^3)$, where $r = |\mathbf{x}|$, satisfies $\Delta T_{1/r} = -4\pi\delta_{\mathbf{0}}$ and is, hence, a fundamental solution for the Laplace operator.*

Proof. We have

$$\Delta T_{1/r}[\varphi] = T_{1/r}[\Delta\varphi] = \int_{\mathbb{R}^3} dx^3 \, \frac{\Delta\varphi}{r} \, . \tag{7.39}$$

That this integral is equal to $-4\pi\delta_{\mathbf{0}}[\varphi]$ has already been shown in Theorem (6.3). □

Hence, $T_{1/r}$ is a fundamental solution of the Laplace operator and can be used to write down solutions of the inhomogeneous Laplace equation by applying Theorem 7.3. In fact, in Corollary 6.2 we have seen that this provides the unique solution ϕ to the inhomogeneous Laplace equation with $\phi \to 0$ as $r \to \infty$.

We have also encountered Green functions in our discussion of second order ordinary differential equations in Section 5 and it is useful to re-formulate these results in terms of distributions. Recall that the relevant differential operator is given by

$$L = \alpha_2(x)\frac{d^2}{dx^2} + \alpha_1(x)\frac{d}{dx} + \alpha_0(x) \, , \tag{7.40}$$

and, on the interval $x \in [a, b]$, we would like to solve the equation $Ly = f$, with $y(a) = y(b) = 0$. The solution to this problem is described, in terms of a Green function, in Theorem 5.6. From this statement we have

Theorem 7.5. *For $\tilde{G}(x) := G(0, x)$ where G is the Green function function from Theorem 5.6, but for the operator $L^\dagger$, the distribution $T_{\tilde{G}}$, is a fundamental solution of the operator (7.40), that is, $LT_{\tilde{G}} = \delta_0$.*

Proof. From Eq. (5.83) the Green function for $L^\dagger$ can be written in terms of its eigenfunvyions e_k and eigenalues λ_k as

$$G(x, t) = w \sum_k \frac{1}{\lambda_k} e_k(t) e_k(x) \, , \tag{7.41}$$

where

$$L^\dagger e_k = \lambda_k e_k \tag{7.42}$$

and this Green function provides a solution to $L^\dagger y = \varphi$ given by

$$y(x) = \int_R dt \, G(x, t)\varphi(t) \quad \Rightarrow \quad \int_{\mathbb{R}} dt \, L_x^\dagger G(x, t)\varphi(t) = \varphi(x) \, . \tag{7.43}$$

Using this and the fact that $L^\dagger$ is hermitian relative to the scalar product defined with weight function w (as it appears in its Sturm-Liouville form) we get

$$\begin{aligned} LT_{\tilde{G}}[\varphi] = T_{\tilde{G}}[L^\dagger\varphi] &= \int_{\mathbb{R}} dt \, G(0, t) L_t^\dagger \varphi(t) \overset{\text{Eq. (7.41)}}{=} \sum_k \frac{1}{\lambda_k} \int_{\mathbb{R}} dt \, w(t) e_k(t) e_k(0) L_t^\dagger \varphi(t) \\ &= \sum_k \frac{1}{\lambda_k} \int_{\mathbb{R}} dt \, w(t) L_t^\dagger e_k(t) e_k(0) \varphi(t) \overset{\text{Eq. (7.42)}}{=} \sum_k \int_{\mathbb{R}} dt \, w(t) e_k(t) e_k(0) \varphi(t) \\ &\overset{\text{Eq. (7.42)}}{=} \sum_k \frac{1}{\lambda_k} \int_{\mathbb{R}} dt \, w(t) e_k(t) (L_x^\dagger e_k)(0) \varphi(t) \overset{\text{Eq. (7.41)}}{=} \int_{\mathbb{R}} dt \, (L_x^\dagger G)(0, t)\varphi(t) \overset{\text{Eq. (7.43)}}{=} \varphi(0) = \delta_0[\varphi] \end{aligned}$$

□

We will encounter further examples of fundamental solutions in the next chapter when we discuss other linear partial differential equations.

7.4 Fourier transform for distributions*

In section 3.2 we have discussed the Fourier transform and we have defined

$$\mathcal{F}(f)(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n y \, e^{-i\mathbf{k}\cdot\mathbf{y}} f(\mathbf{y}) \,, \qquad \tilde{\mathcal{F}}(\hat{f})(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n k \, e^{i\mathbf{k}\cdot\mathbf{x}} \hat{f}(\mathbf{k}) \,, \tag{7.44}$$

and one of our central results was that $\tilde{\mathcal{F}} \circ \mathcal{F} = \mathrm{id}$, so $\tilde{\mathcal{F}}$ is the inverse Fourier transform. Writing this out explicitly we have

$$\frac{1}{(2\pi)^n} \int d^n y \, d^n k \, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} f(\mathbf{y}) = f(\mathbf{x}) \,. \tag{7.45}$$

Symbolically, this result can also be written as

$$\frac{1}{(2\pi)^n} \int d^n k \, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} = \delta(\mathbf{x}-\mathbf{y}) \,, \tag{7.46}$$

in the sense that this equation multiplied with $f(\mathbf{y})$ and integrated over $\mathbf{y}$ leads to Eq. (7.45), provided the naive rule (7.7) for working with $\delta(\mathbf{x}-\mathbf{y})$ is used. The equation (7.46) is frequently used in physics calculations but, as all other consideration which involve the Dirac delta used as a function, is mathematically unsound.

Exercise 7.6. *Use Eq. (7.46) naively to "prove" part (a) of Plancherel's theorem (3.16).*

To check that using Eq. (7.46) in a naive way makes sense we should consider the generalisation of the Fourier transform to distributions.

For some guidance on how to define the Fourier transform for distributions we proceed as usual and demand that $\mathcal{F}T_f = T_{\mathcal{F}f}$ in order to ensure that the Fourier transform on distributions reduces to the familiar one for functions whenever it is applied to a distribution of the form T_f. This implies

$$\begin{aligned} (\mathcal{F}T_f)[\varphi] &= T_{\mathcal{F}f}[\varphi] = \int d^n y \, (\mathcal{F}f)(\mathbf{y})\varphi(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}} \int d^n x \, d^n y \, e^{-i\mathbf{x}\cdot\mathbf{y}} f(\mathbf{x})\varphi(\mathbf{y}) \\ &= \int d^n x \, (\mathcal{F}\varphi)(\mathbf{x}) f(\mathbf{x}) = T_f[\mathcal{F}\varphi] \,. \end{aligned} \tag{7.47}$$

The LHS and RHS of this equation make sense if we drop the subscript f and this can be used to define the Fourier transform of distributions by [7]

$$(\mathcal{F}T)[\varphi] := T[\mathcal{F}\varphi] \,. \qquad (\tilde{\mathcal{F}}T)[\varphi] := T[\tilde{\mathcal{F}}\varphi] \,, \tag{7.48}$$

so that $\tilde{\mathcal{F}} \circ \mathcal{F}(T) = \mathcal{F} \circ \tilde{\mathcal{F}}(T) = T$. Given this definition, what is the Fourier transform of the Dirac delta, $\delta_{\mathbf{k}}$? The quick calculation

$$(\mathcal{F}\delta_{\mathbf{k}})[\varphi] = \delta_{\mathbf{k}}[\mathcal{F}\varphi] = (\mathcal{F}\varphi)(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int d^n x \, e^{-i\mathbf{k}\cdot\mathbf{x}} \varphi(\mathbf{x}) = T_{e^{-i\mathbf{k}\cdot\mathbf{x}}/(2\pi)^{n/2}}[\varphi] \tag{7.49}$$

shows that

$$\mathcal{F}\delta_{\mathbf{k}} = T_{\frac{1}{(2\pi)^{n/2}} e^{-i\mathbf{k}\cdot\mathbf{x}}} \,. \tag{7.50}$$

[7] A mathematically fully satisfactory definition requires modifying the test function space but this would go beyond our present scope. The main idea of how to define Fourier transforms of distributions is already apparent without considering this subtlety.

Note that this result is in line with our intuitive understanding of Fourier transforms as frequency analysis. It says that a frequency spectrum "sharply peaked" at $\mathbf{k}$ Fourier transforms to a monochromatic wave with wave vector $\mathbf{k}$.

Now consider the inverse of the above computation, that is, we would like to work out the Fourier transform of $T_{e^{i\mathbf{k}\cdot\mathbf{x}}}$. Note that the function $e^{i\mathbf{k}\cdot\mathbf{x}}$ is not integrable over $\mathbb{R}^n$ so it does not have a Fourier transform in the conventional sense. Taking the complex conjugate

$$\tilde{\mathcal{F}}\delta_{\mathbf{k}} = T_{\frac{1}{(2\pi)^{n/2}}e^{i\mathbf{k}\cdot\mathbf{x}}} \tag{7.51}$$

of Eq. (7.50) and applying $\mathcal{F}$ to this equation gives

$$\mathcal{F}T_{e^{i\mathbf{k}\cdot\mathbf{x}}} = (2\pi)^{n/2}\delta_{\mathbf{k}} \tag{7.52}$$

Again, this is in line with the intuitive understanding of Fourier transforms. The transform of a monochromatic wave with wave vector $\mathbf{k}$ is a spectrum "sharply peaked" at $\mathbf{k}$. At the same time, Eq. (**??**) is the mathematically correct version of Eq. (7.46).

8 Other linear partial differential equations

In this chapter, we will discuss a number of other linear partial differential equations which are important in physics, including the Helmholz equation, the wave equation and the heat equation. We will cover a number of methods to solve these equations but in the interest of keeping these notes manageable we will not be quite as thorough as we have been for the Laplace equation. We begin with the Helmholz equation which is closest to the Laplace equation.

8.1 The Helmholz equation

The homogeneous and inhomogeneous Helmholz equations in $\mathbb{R}^3$ (with coordinates $\mathbf{x} = (x_i)$) are given by

$$(\Delta + k^2)\psi = 0 \,, \qquad (\Delta + k^2)\psi = f \,, \tag{8.1}$$

where $k \in \mathbb{R}$ is a real number and Δ is the three-dimensional Laplace operator (although the equation can, of course, also be considered in other dimensions). This equation appears, for example, in wave problems with fixed wave number k, as we will see in our discussion of the wave equation later on.

As always, the general solution to the inhomogeneous Helmholz equation is given as a sum of the general solution of the homogeneous equation plus a special solution of the inhomogeneous equation. The homogeneous Helmholz equation is an eigenvalue equation (with eigenvalue $-k^2$) for the Laplace operator and many of the methods discussed in the context of the Laplace equation can be applied.

To find a special solution of the inhomogeneous equation we can use the Green function method, in analogy to what we did for the Laplace equation. Define the functions

$$G_\pm(r) = \frac{e^{\pm ikr}}{r} \,, \tag{8.2}$$

where $r = |\mathbf{x}|$ is the radial coordinate. Given that this function is independent of the angles we can use only the radial part of the Laplacian in spherical coordinates (6.17) to verify that, for $r > 0$

$$(\Delta + k^2)G_\pm = \left(\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + k^2\right) G_\pm = 0 \,. \tag{8.3}$$

Hence, $G_\pm$ solves the homogeneous Helmholz equation for $r > 0$ in much the same way $1/r$ solves the homogeneous Laplace equation. In fact, the analogy goes further as stated in the following

Theorem 8.1. *With $G = AG_+ + BG_-$, where $A, B \in \mathbb{R}$ and $A + B = 1$ the distribution T_G is a fundamental solution to the Helmholz operator, that is,*

$$(\Delta + k^2)T_G = -4\pi\delta_{\mathbf{0}} \,. \tag{8.4}$$

Proof. The proof is very much in analogy with the corresponding one for the Laplace equation 7.4 and can be found in Ref. [4]. Essentially, it relies on $\Delta T_{1/r} = -4\pi\delta_{\mathbf{0}}$ and the fact that $1/r$ is really the only singularity in G. □

With this result, the general solution to the inhomogeneous Helmholz equation can be written as

$$\psi(\mathbf{x}) = \psi_{\text{hom}}(\mathbf{x}) - \frac{1}{4\pi}(T_G \star f)(\mathbf{x}) = \psi_{\text{hom}}(\mathbf{x}) - \frac{1}{4\pi}\int_{\mathbb{R}^3} d^3y\, G(\mathbf{x}-\mathbf{y})f(\mathbf{y}) \,, \tag{8.5}$$

where ψ_{hom} is an arbitrary solution of the homogeneous equation.

8.2 Eigenfunctions and time evolution

Many partial differential equations in physics involve a number of spatial coordinates $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathcal{V} \subset U \subset \mathbb{R}^n$ as well as time $t \in \mathbb{R}$ and are of the form

$$H\psi = \frac{1}{c}\dot{\psi} \qquad \text{or} \qquad H\psi = \ddot{\psi} \,, \tag{8.6}$$

where $\psi = \psi(t, \mathbf{x})$, the dot denotes the derivative $\frac{\partial}{\partial t}$ and c is a constant. We assume that H is a second order linear differential operator in the spatial differentials $\frac{\partial}{\partial x_i}$, so a differential operator on $C^\infty(U) \cap L^2(U)$ which is time-independent and hermitian relative to the standard scalar product on $L^2(U)$. If boundary conditions are imposed on $\partial\mathcal{V}$ we assume that they are also time-independent. Under these conditions there frequently exists a (time-independent) ortho-normal basis $(\phi_i)_{i=1}^\infty$ of $L^2(U)$ with the desired boundary behaviour which consists of eigenfunctions of H, so

$$H\phi_i = \lambda_i \phi_i \,, \tag{8.7}$$

where the eigenvalues λ_i are real since H is hermitian. The problem is to solve the above equations subject to an initial condition $\psi(0, \mathbf{x}) = \psi_0(\mathbf{x})$ and, in addition, $\dot{\psi}(0, \mathbf{x}) = \dot{\psi}_0(\mathbf{x})$ in the case of the second equation (8.6), for given functions ψ_0 and $\dot{\psi}_0$. This can be done by expanding the function ψ, for any given time t, in terms of the basis (ϕ_i), so that

$$\psi(t, \mathbf{x}) = \sum_i A_i(t)\phi_i(\mathbf{x}) \,. \tag{8.8}$$

Inserting this into the first Eq. (8.6) leads to

$$\dot{A}_i = c\lambda_i A_i \qquad \Rightarrow \qquad A_i = a_i e^{c\lambda_i t} \tag{8.9}$$

so that the complete solution reads

$$\psi(t, \mathbf{x}) = \sum_i a_i \phi_i(\mathbf{x}) e^{c\lambda_i t} \,. \tag{8.10}$$

The remaining constants a_i are fixed by the initial condition $\psi(0, \mathbf{x}) = \sum_i a_i \phi_i(\mathbf{x}) \stackrel{!}{=} \psi_0(\mathbf{x})$ which can be solved in the usual way, using the orthogonality relations $\langle \phi_i, \phi_j \rangle = \delta_{ij}$. This leads to

$$a_i = \langle \phi_i, \psi_0 \rangle \,. \tag{8.11}$$

A similar calculation for the second equation (8.6) (assuming that $\lambda_i < 0$) leads to

$$\ddot{A}_i = -|\lambda_i| A_i \qquad \Rightarrow \qquad A_i = a_i \sin\left(\sqrt{|\lambda_i|}\, t\right) + b_i \cos\left(\sqrt{|\lambda_i|}\, t\right) \tag{8.12}$$

so that

$$\psi(t, \mathbf{x}) = \sum_i \left(a_i \sin\left(\sqrt{|\lambda_i|}\, t\right) + b_i \cos\left(\sqrt{|\lambda_i|}\, t\right) \right) \phi_i(\mathbf{x}) \,. \tag{8.13}$$

The constants a_i and b_i are fixed by the initial conditions $\psi(0, \mathbf{x}) = \sum_i b_i \phi_i(\mathbf{x}) \stackrel{!}{=} \psi_0(\mathbf{x})$ and $\dot{\psi}(0, \mathbf{x}) = \sum_i a_i \sqrt{|\lambda_i|} \phi_i(\mathbf{x}) \stackrel{!}{=} \dot{\psi}_0(\mathbf{x})$ and are, hence, given by

$$a_i = \frac{1}{\sqrt{|\lambda_i|}} \langle \phi_i, \dot{\psi}_0 \rangle \,, \qquad b_i = \langle \phi_i, \psi_0 \rangle \,. \tag{8.14}$$

Below, we will discuss various examples of this structure more explicitly.

8.3 The heat equation

The homogeneous and inhomogeneous heat equations are given by

$$\left(\Delta_n - \frac{\partial}{\partial t}\right)\psi = 0\,, \qquad \left(\Delta_n - \frac{\partial}{\partial t}\right)\psi = f\,, \tag{8.15}$$

where Δ_n is the Laplacian in n dimensions (with $n = 1, 2, 3$ cases of physical interest). The solution $\psi = \psi(t, \mathbf{x})$ can be interpreted as a temperature distribution evolving in time. If we are solving the equations on a spatial patch $\mathcal{V} \subset \mathbb{R}^n$ we have to provide boundary conditions on $\partial\mathcal{V}$ and in addition, we should provide an initial distribution $\psi(0, \mathbf{x}) = \psi_0(\mathbf{x})$ at time $t = 0$.

The homogeneous equation is of the form discussed in the previous subsection (with $H = \Delta$ and $c = 1$). If we are solving the equation on a spatial patch $\mathcal{V} \subset \mathbb{R}^n$ with boundary conditions such that the spectrum of the Laplacian has countable many eigenvectors then we can apply the method described in Section 8.2.

Application 8.31. Evolution of temperature along a rod

To illustrate how this works in practice consider a one-dimensional problem ($n = 1$) where $\mathcal{V} = [0, a]$ and we demand Dirichlet boundary conditions $\psi(t, 0) = \psi(t, a) = 0$ and some initial distribution $\psi_0(x) = \psi(0, x)$. (Physically, this corresponds to a rod of length a whose endpoints are kept at a fixed temperature and with a given initial temperature distribution at time $t = 0$.) Given the space and the boundary conditions the functions $\phi_k = \sin\left(\frac{k\pi x}{a}\right)$ of the sine Fourier series provide an orthogonal basis of eigenfunctions satisfying $\phi_k'' = \lambda_k \phi_k$ with eigenvalues

$$\lambda_k = -\frac{k^2\pi^2}{a^2}\,. \tag{8.16}$$

Inserting this into the general solution (8.10) leads to

$$\psi(t, x) = \sum_{k=1}^{\infty} b_k \sin\left(\frac{k\pi x}{a}\right) e^{-\frac{k^2\pi^2}{a^2}t}\,. \tag{8.17}$$

The coefficients b_k are determined by the initial condition $\psi(0, x) = \psi_0(x)$ which leads to the standard sine Fourier series

$$\psi(0, x) = \sum_{k=1}^{\infty} b_k \sin\left(\frac{k\pi x}{a}\right) \stackrel{!}{=} \psi_0(x) \tag{8.18}$$

for ψ_0 which, of course, implies that

$$b_k = \frac{2}{a}\int_0^a dx\, \sin\left(\frac{k\pi x}{a}\right)\psi_0(x)\,. \tag{8.19}$$

Exercise 8.2. *For $\mathcal{V} = [0, a]$, boundary conditions $\psi(t, 0) = \psi(t, a) = 0$ and an initial distribution $\psi(0, x) = T_0 x(a-x)/a^2$ (where T_0 is a constant) find the solution $\psi(t, x)$ to the homogeneous heat equation.*

For solutions in $\psi(t, \cdot) \in L^2(\mathbb{R}^n)$ we can also solve the homogeneous heat equation using Fourier transforms. Inserting

$$\psi(t, \mathbf{x}) = \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} d^n k\, \tilde{\psi}(t, \mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{8.20}$$

into the heat equation leads to

$$\dot{\tilde{\psi}} = -|\mathbf{k}|^2 \tilde{\psi} \qquad \Rightarrow \qquad \tilde{\psi}(t, \mathbf{k}) = \chi(\mathbf{k}) e^{-|\mathbf{k}|^2 t} \,, \tag{8.21}$$

for a function $\chi(\mathbf{k})$. Inserting this back into the Ansatz gives

$$\psi(t, \mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n k \, \chi(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} - |\mathbf{k}|^2 t} \,, \tag{8.22}$$

and the initial condition $\psi(0, \mathbf{x}) = \psi_0(\mathbf{x})$ translates into $\mathcal{F}(\chi) = \psi_0$ which can be inverted using the inverse Fourier transform, so $\chi = \tilde{\mathcal{F}}(\psi_0)$.

Exercise 8.3. *Find the solution $\psi(t, x)$ of the heat equation for $x \in \mathbb{R}$, square integrable for all t, which satisfies $\psi(0, x) = \psi_0(x) = T_0 e^{-\frac{x^2}{2a^2}}$.*

To solve the inhomogeneous heat equation we would like to find a Green function. It can be verified by direct calculation, that the function $G : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by

$$G(t, \mathbf{x}) = \begin{cases} -\frac{1}{(4\pi t)^{n/2}} e^{-\frac{|\mathbf{x}|^2}{4t}} & \text{for} \quad t > 0 \\ 0 & \text{for} \quad t \leq 0 \end{cases} \tag{8.23}$$

solves the homogeneous heat equation whenever $t \neq 0$ and $\mathbf{x} \neq \mathbf{0}$.

Exercise 8.4. *Show that the function G in Eq. (8.23) solves the homogeneous heat equation for $t \neq 0$ and $\mathbf{x} \neq \mathbf{0}$.*

In fact we have

Theorem 8.5. *The distribution T_G with G defined in Eq. (8.23) is a fundamental solution of the heat equation, so*

$$\left(\Delta - \frac{\partial}{\partial t}\right) T_G = \delta_{\mathbf{0}} \,. \tag{8.24}$$

Proof. The proof is similar to the one for the Laplacian in Theorem 7.4 and can be found in Ref. [4]. □

From this result, the general solution to the inhomogeneous heat equation can be written as

$$\psi(t, \mathbf{x}) = \psi_{\text{hom}}(t, \mathbf{x}) + (T_G \star f)(t, \mathbf{x}) = \psi_{\text{hom}}(t, \mathbf{x}) + \int_{\mathbb{R}^{n+1}} d\tau \, d^n y \, G(t - \tau, \mathbf{x} - \mathbf{y}) f(\tau, \mathbf{y}) \,, \tag{8.25}$$

where ψ_{hom} is a solution to the homogeneous equation.

8.4 The wave equation

The homogeneous and inhomogeneous wave equations are of the form

$$\left(\Delta_n - \frac{\partial^2}{\partial t^2}\right) \psi = 0 \,, \qquad \left(\Delta_n - \frac{\partial^2}{\partial t^2}\right) \psi = f \,, \tag{8.26}$$

where Δ_n is the n-dimensional Laplacian (and $n = 1, 2, 3$ are the most interesting dimensions for physics). If this equation is considered on the spatial patch $\mathcal{V} \subset \mathbb{R}^n$ we should specify boundary conditions on $\partial\mathcal{V}$ for all t. In addition, we require initial conditions $\psi(0, \mathbf{x}) = \psi_0(\mathbf{x})$ and $\dot{\psi}(0, \mathbf{x}) = \dot{\psi}_0(\mathbf{x})$ at some initial time $t = 0$.

For an Ansatz of the form

$$\psi(t,\mathbf{x}) = \tilde{\psi}(\mathbf{x})e^{-i\omega t} \tag{8.27}$$

the wave equation turns into the Helmholz equation for $\tilde{\psi}$ with $k = \omega$ and we can use the methods discussed in Section 8.1.

Starting with the homogeneous equation with $\psi(t,\cdot) \in L^2(\mathbb{R}^n)$ we can start with a Fourier integral

$$\psi(t,\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n k\, \tilde{\psi}(t,\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \,. \tag{8.28}$$

Inserting this into the homogenous equation implies

$$\ddot{\tilde{\psi}} = -|\mathbf{k}|^2 \tilde{\psi} \qquad \Rightarrow \qquad \tilde{\psi}(t,\mathbf{k}) = \psi_+(\mathbf{k})e^{i|\mathbf{k}|t} + \psi_-(\mathbf{k})e^{-i|\mathbf{k}|t} \,, \tag{8.29}$$

and, hence,

$$\psi(t,\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} d^n k\, (\psi_+(\mathbf{k})e^{i|\mathbf{k}|t} + \psi_-(\mathbf{k})e^{-i|\mathbf{k}|t}) e^{-i\mathbf{k}\cdot\mathbf{x}} \,. \tag{8.30}$$

The functions $\psi_\pm$ are fixed by the initial conditions via $\psi_0 = \mathcal{F}(\psi_+ + \psi_-)$ and $\dot{\psi}_0 = \mathcal{F}(i|\mathbf{k}|(\psi_+ - \psi_-))$ which can be solved, using the inverse Fourier transform, to give $\psi_\pm = \frac{1}{2}\tilde{\mathcal{F}}\left(\psi_0 \mp \frac{i}{|\mathbf{k}|}\dot{\psi}_0\right)$.

If we work on a spatial patch with boundary conditions which lead to a countable number of eigenvectors of the Laplacian we can use the method in Section 8.2 to solve the homogeneous wave equation. For one and two spatial dimensions this leads to systems usually referred to as "strings" and "membranes", respectively, and we now discuss them in turn.

8.4.1 Strings

The wave equation now reads [8]

$$\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial t^2}\right)\psi = 0 \,, \tag{8.31}$$

where $\psi = \psi(t,x)$ and $x \in [0,a]$. This equation describes various kinds of strings from guitar strings to the strings of string theory. We will impose Dirichlet boundary conditions $\psi(t,0) = \psi(t,a) = 0$ as appropriate for a string with fixed endpoints. (The strings of string theory allow for both Dirichlet and von Neumann boundary conditions.) In addition, we need to fix the initial position, $\psi(0,x) = \psi_0(x)$, and initial velocity $\dot{\psi}(0,x) = \dot{\psi}_0(x)$.

Given the boundary conditions an appropriate set of eigenfunctions is provided by $\phi_k = \sin\left(\frac{k\pi x}{a}\right)$, that is the functions for the sine Fourier series. We have $\phi_k'' = \lambda_k \phi_k$ with eigenvalues

$$-\lambda_k = \frac{k^2\pi^2}{a^2} =: \omega_k^2 \,. \tag{8.32}$$

Inserting this into the general solution (8.13) leads to

$$\psi(t,x) = \sum_{k=1}^{\infty} \left[a_k \sin\left(\frac{k\pi t}{a}\right) + b_k \cos\left(\frac{k\pi t}{a}\right)\right] \sin\left(\frac{k\pi x}{a}\right) \tag{8.33}$$

The coefficients a_k and b_k are fixed by the initial conditions via

$$\psi(0,x) = \sum_{k=1}^{\infty} b_k \sin\left(\frac{k\pi x}{a}\right) \stackrel{!}{=} \psi_0(x) \,, \qquad \dot{\psi}(0,x) = \sum_{k=1}^{\infty} \frac{k\pi a_k}{a} \sin\left(\frac{k\pi x}{a}\right) \stackrel{!}{=} \dot{\psi}_0(x) \,. \tag{8.34}$$

[8] In a physics context, this equation is frequently written with an additional factor of $1/c^2$ in front of the time-derivatives, where c is the speed of the wave. Such a factor can always be removed by a re-definition $ct \to t$.

These equations can of course be solved for a_k and b_k using standard Fourier series techniques resulting in

$$a_k = \frac{2}{k\pi}\int_0^a dx\, \sin\left(\frac{k\pi x}{a}\right)\dot{\psi}_0(x)\ , \qquad b_k = \frac{2}{a}\int_0^a dx\, \sin\left(\frac{k\pi x}{a}\right)\psi_0(x)\ . \tag{8.35}$$

Note that the eigenfrequencies of the system

$$\omega_k = \frac{k\pi}{a} \tag{8.36}$$

are all integer multiples of the ground frequency $\omega_1 = \pi/a$.

Exercise 8.6. *Find the solution $\psi = \psi(t,x)$ for a string with length a, Dirichlet boundary conditions $\psi(t,0) = \psi(t,a) = 0$ and initial conditions $\dot{\psi}(0,x) = 0$ and*

$$\psi(0,x) = \begin{cases} \frac{hx}{b} & \text{for} \quad 0 \le x \le b \\ \frac{h(a-x)}{a-b} & \text{for} \quad b < x \le a \end{cases} \tag{8.37}$$

where $b \in [0,a]$ and h are constants. (Think of a guitar string plucked at distance b from the end of the string.) How do the parameters b and h affect the sound of the guitar?

8.4.2 Membranes

We are now dealing with a wave equation of the form

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial t^2}\right)\psi = 0\ , \tag{8.38}$$

where $\psi = \psi(t,x,y)$ and $(x,y) \in \mathcal{V} \subset \mathbb{R}^2$ with a (compact) spatial patch $\mathcal{V}$. We require boundary conditions at $\partial\mathcal{V}$ and initial conditions $\psi(0,x,y) = \psi_0(x,y)$ and $\dot{\psi}(0,x,y) = \dot{\psi}_0(x,y)$. Of course we can consider any number of "shapes", $\mathcal{V}$, of the membrane. Let us start with the simplest possibility of a rectangular membrane, $\mathcal{V} = [0,a]\times[0,b]$, and Dirichlet boundary conditions $\psi|_{\partial\mathcal{V}} = 0$. In this case, we have an orthogonal basis of eigenfunctions $\phi_{k,l}(x,y) = \sin\left(\frac{k\pi x}{a}\right)\sin\left(\frac{l\pi y}{b}\right)$, where $k,l = 1,2,\ldots$, with $\Delta_2\phi_{k,l} = \lambda_{k,l}\phi_{k,l}$ and eigenvalues

$$-\lambda_{k,l} = \frac{k^2\pi^2}{a^2} + \frac{l^2\pi^2}{b^2} =: \omega_{k,l}^2\ . \tag{8.39}$$

Inserting into Eq. (8.13) gives

$$\psi(t,x,y) = \sum_{k,l=1}^{\infty}\left(a_{k,l}\sin(\omega_{k,l}t) + b_{k,l}\cos(\omega_{k,l}t)\right)\sin\left(\frac{k\pi x}{a}\right)\sin\left(\frac{l\pi y}{b}\right)\ . \tag{8.40}$$

The coefficients a_k and b_k are fixed by the initial conditions and can be obtained using standard Fourier series techniques (in both the x and y coordinate), in analogy with the string case. We note that the lowest frequencies of the quadratic ($a = b$) drum are

$$\omega_{1,1} = \sqrt{2}\frac{\pi}{a}\ , \qquad \omega_{2,1} = \omega_{1,2} = \sqrt{5}\frac{\pi}{a}\ . \tag{8.41}$$

Unlike for a string the eigenfrequencies of a drum are not integer multiples of the ground frequency - this is why a drum sounds less well-defined compared to other instruments and why most instruments use strings or other, essentially one-dimensional systems to produce sound.

For the round membrane with $\mathcal{V} = \{\mathbf{x} \in \mathbb{R}^2 \,|\, |\mathbf{x}| \leq a\}$ and boundary condition $\phi|_{r=a} = 0$ we should first find a set of eigenfunctions ϕ for the two-dimensional Laplacian in polar coordinates, satisfying $\Delta\phi = -\lambda\phi$. Starting with the Ansatz $\phi(r,\varphi) = R(r)e^{im\varphi}$ and using the two-dimensional Laplace operator (6.11) in polar coordinates gives

$$r^2R'' + rR' + (\lambda r^2 - m^2)R = 0 \,. \tag{8.42}$$

This is the Bessel differential equation for $\nu = |m|$ and the solution is $R(r) \sim J_{|m|}(\sqrt{\lambda}r)$. The boundary condition at $r = a$ implies that $\sqrt{\lambda}a = z_{|m|n}$, where $z_{\nu n}$ denotes the n^{th} zero of the Bessel function J_ν. This means we have eigenfunctions and eigenvalues

$$\phi_{mn}(r,\varphi) \sim \hat{J}_{|m|n}(r)e^{im\varphi} \,, \qquad \lambda_{mn} = \frac{z_{|m|n}^2}{a^2} =: \omega_{mn}^2 \,. \tag{8.43}$$

Expanding $\psi(t,r,\varphi) = \sum_{m,n} T_{mn}(t)\phi_{mn}(r,\varphi)$ we find the differential equations $\ddot{T}_{mn} = -\omega_{mn}^2 T_{mn}$ so that ω_{mn} are the eigenfrequencies of the round membrane. As is clear from Eq. (8.43) these eigenfrequencies are determined by the zeros of the Bessel functions so they are quite irregular.

8.4.3 Green function of wave equation

Returning to the inhomogeneous wave equation, we would like to find a Green function G, satisfying

$$\left(\Delta_3 - \frac{\partial^2}{\partial t^2}\right) G(t,\mathbf{x}) = -4\pi\delta(t)\delta(\mathbf{x}) \,. \tag{8.44}$$

(We are using the intuitive notion of Delta "functions" to be closer to the standard physics treatment but our discussion of distributions should reassure us that this leads to sensible results.) If we Fourier transform G in the t-direction

$$G(t,\mathbf{x}) = \frac{1}{2\pi}\int_{\mathbb{R}} d\omega\, \tilde{G}(\omega,\mathbf{x})e^{-i\omega t} \tag{8.45}$$

it follows that the Fourier transform $\tilde{G}$ satisfies

$$(\Delta + \omega^2)\tilde{G}(\omega,\mathbf{x}) = -4\pi\delta(\mathbf{x}) \,, \tag{8.46}$$

and is, hence, a Green function for the Helmholz equation. From our discussion in Section 8.1 there are essentially two choices for this Green function, namely

$$\tilde{G}_\pm(\omega,\mathbf{x}) = \frac{e^{\pm i\omega|\mathbf{x}|}}{|\mathbf{x}|} \,. \tag{8.47}$$

Inserting this into the Fourier integral (8.45) and using the result (7.46) gives

$$G_\pm(t,\mathbf{x}) = \frac{\delta(t \mp |\mathbf{x}|)}{|\mathbf{x}|} \,. \tag{8.48}$$

The general solution to the inhomogeneous wave equation, using the so-called *retarded Green function* G_+, is then given by

$$\begin{aligned}
\psi(t,\mathbf{x}) &= \psi_{\text{hom}}(t,\mathbf{x}) - \frac{1}{4\pi}\int_{\mathbb{R}^4} dt'\, d^3x'\, G_+(t-t',\mathbf{x}-\mathbf{x}')f(t',\mathbf{x}') \\
&= \psi_{\text{hom}}(t,\mathbf{x}) - \frac{1}{4\pi}\int_{\mathbb{R}^4} dt'\, d^3x'\, \frac{\delta(t-t'-|\mathbf{x}-\mathbf{x}'|)}{|\mathbf{x}-\mathbf{x}'|}f(t',\mathbf{x}') \\
&= \psi_{\text{hom}}(t,\mathbf{x}) - \frac{1}{4\pi}\int_{\mathbb{R}^3} d^3x'\, \left(\frac{f(t',\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\right)_{t'=t-|\mathbf{x}-\mathbf{x}'|}
\end{aligned} \tag{8.49}$$

Note that this formula is very much in analogy with the corresponding result for the Laplace equation. The difference is of course the dependence on t. The source f is evaluated at the *retarded time* $t' = t - |\mathbf{x} - \mathbf{x}'|$ to produce the solution at time t. The physical interpretation is that this takes into account the time it takes for the effect of the source at $\mathbf{x}'$ to influence the solution at $\mathbf{x}$. The above result is the starting point for calculating the electromagnetic radiation from moving charges.

9 Groups and representations*

Symmetries have become a key idea in modern physics and they are an indispensable tool for the construction of new physical theories. They play an important role in the formulation of practically all established physical theories, from Classical/Relativistic Mechanics, Electrodynamics, Quantum Mechanics, General Relativity to the Standard Model of Particle Physics.

The word "symmetry" used in a physics context (usually) refers to the mathematical structure of a *group*, so this is what we will have to study. In physics, the typical problem is to construct a theory which "behaves" in a certain defined way under the action of a symmetry, for example, which is invariant. To tackle such a problem we need to know how symmetries act on the basic building blocks of physical theories and these building blocks are often elements of vector spaces. (Think, for example, of the trajectory $\mathbf{r}(t)$ of a particle in Classical Mechanics which, for every time t, is an element of $\mathbb{R}^3$, a four-vector $x^\mu(t)$ which is an element of $\mathbb{R}^4$ or the electric and magnetic fields which, at each point in space-time, are elements of $\mathbb{R}^3$.) Hence, we need to study the action of groups on vector spaces and the mathematical theory dealing with this problem is called (linear) *representation theory of groups*. The translation between physical and mathematical terminology is summarised in the diagram below.

	physics		**mathematics**	
	symmetry	$\cong$	group	
action on	$\downarrow$		$\downarrow$	representation
	building blocks	$\cong$	vector spaces	

Groups and their representations form a large area of mathematics and a comprehensive treatment can easily fill two or three lecture courses. Here we will just touch upon some basics and focus on some of the examples with immediate relevance for physics. We begin with elementary group theory - the definition of a group, of a sub-group and of group homomorphisms plus examples of groups - before we move on to representations. Lie groups and their associated Lie algebras play an important role in physics and we briefly discuss the main ideas before moving on to the physically most relevant examples of such groups.

9.1 Groups: some basics

A *group* is one of the simplest algebraic structures - significantly simpler than a vector space - which has only one operation, usually referred to as *group multiplication*, subject to three axioms.

9.1.1 Groups and subgroups

The formal definition of a group is:

Definition 9.1. *(Group) A group is a set G with a map $\cdot : G \times G \to G$, called group multiplication, which satisfies*

(G1) $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ *for all* $g_1, g_2, g_3 \in G$ *(associativity)*
(G2) There exists an $e \in G$ *such that* $e \cdot g = g$ *for all* $g \in G$ *(neutral element)*
(G3) For all $g \in G$ *there exists a* $\tilde{g} \in G$ *such that* $\tilde{g} \cdot g = e$ *(inverse element)*

If, in addition, the group multiplication commutes, that is, if $g_1 \cdot g_2 = g_2 \cdot g_1$ *for all* $g_1, g_2 \in G$*, then the group is called Abelian. Otherwise it is called non-Abelian.*

Groups can have a finite or infinite number of elements and we will see examples of either. In the former case, they are called *finite groups*. The number of elements in a finite group is also called the *order of the group*.

This above definition looks somewhat asymmetric since we have postulated that the neutral element and the inverse in (G2) and (G3) multiply from the left but have made no statement about their multiplication from the right. However, this is not a problem due to the following Lemma.

Proposition 9.1. *For a group G we have the following statements.*
(i) A left-inverse is also a right-inverse, that is, $\tilde{g} \cdot g = e \;\Rightarrow\; g \cdot \tilde{g} = e$.
(ii) A left-neutral is also a right neutral, that is, $e \cdot g = g \;\Rightarrow\; g \cdot e = g$.
(iii) For a given $g \in G$, the inverse is unique and denoted by g^{-1}.
(iv) The neutral element e is unique.

Proof. (i) Start with a $g \in G$ and its left-inverse $\tilde{g}$ so that $\tilde{g} \cdot g = e$. Of course, $\tilde{g}$ must also have a left-inverse which we denote by g' so that $g' \cdot \tilde{g} = e$. Then we have

$$g \cdot \tilde{g} = e \cdot g \cdot \tilde{g} = g' \cdot \underbrace{\tilde{g} \cdot g}_{=e} \cdot \tilde{g} = g' \cdot \tilde{g} = e \tag{9.1}$$

so that $\tilde{g}$ is indeed also a right-inverse.
(ii), (iii), (iv) These proofs are similar to the one for (i) and we leave them as an exercise. □

The inverse satisfies the following simple properties which we have already encountered in the context of maps and their inverses.

$$(g^{-1})^{-1} = g \;, \qquad (g_1 \circ g_2)^{-1} = g_2^{-1} \circ g_1^{-1} \;. \tag{9.2}$$

Exercise 9.1. *Proof statements (ii), (iii) and (iv) of Lemma 9.1 as well as the rules (9.2).*

We follow the same build-up as in the case of vector spaces and next definite the relevant sub-structure, the *sub-group*.

Definition 9.2. *(Sub-group) A subset $H \subset G$ of a group G is called a sub-group of G if it forms a group under the multiplication induced from G.*

The following exercise provides a practical way of checking whether a subset of a group is a sub-group.

Exercise 9.2. *Show that a subset $H \subset G$ of a group G is a sub-group iff it satisfies the following conditions:*
(i) H is closed under the group multiplication.
(ii) $e \in H$
(iii) For all $h \in H$ we have $h^{-1} \in H$

9.1.2 Group homomorphisms

The next step is to define the maps which are consistent with the group structure, the *group homomorphisms*.

Definition 9.3. *A map $f : G \to \tilde{G}$ between two groups G and $\tilde{G}$ is called a group homomorphism iff $f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$ for all $g_1, g_2 \in G$.*
We can define the image of f by $\mathrm{Im}(f) = \{f(g) \,|\, g \in G\} \subset \tilde{G}$.
The kernel of f is defined as $\mathrm{Ker}(f) = \{g \in G \,|\, f(g) = \tilde{e}\} \subset G$ (where $\tilde{e}$ is the neutral element of $\tilde{G}$).

Note that these definitions are in complete analogy with those for linear maps.

Exercise 9.3. *Given a group homomorphism $f : G \to \tilde{G}$, show that the image is a sub-group of $\tilde{G}$ and the kernel is a sub-group of G. Also show that*
(i) $f(e) = \tilde{e}$
(ii) f injective $\iff \mathrm{Ker}(f) = \{e\}$
(iii) f surjective $\iff \mathrm{Im}(f) = \tilde{G}$.

9.1.3 Examples of groups

We should now discuss some examples of groups. You are already familiar with many of them although you may not yet have thought about them in this context.

Examples from "numbers": Every field F forms a group, $(F, +)$, with respect to addition and $F \setminus \{0\}$ forms a group, $(F \setminus \{0\}, \cdot)$ with respect to multiplication (we need to exclude 0 since it does not have a multiplicative inverse). So, more concretely, we have the groups $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ and $(\mathbb{C} \setminus \{0\}, \cdot)$. The integers $\mathbb{Z}$ also form a group, $(\mathbb{Z}, +)$ with respect to addition (however, not with respect to multiplication since there is no multiplicative inverse in $\mathbb{Z}$). Clearly, all of these groups are Abelian. The group $(\mathbb{Z}, +)$ is a sub-group of $(\mathbb{R}, +)$ which, in turn, is a sub-group of $(\mathbb{C}, +)$.

Examples from vector spaces: Every vector space forms an Abelian group with respect to vector addition.

Finite Abelian groups: Consider the set $\mathbb{Z}_p := \{0, 1, \ldots, p-1\}$ for any positive integer p and introduce the group multiplication

$$g_1 \cdot g_2 := (g_1 + g_2) \bmod p \,. \tag{9.3}$$

Clearly, the $\mathbb{Z}_p$ form finite, Abelian groups (with neutral element 0) which are also referred to as *cyclic groups*.

Finite non-abelian groups: The permutations $S_n := \{\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\} \,|\, \sigma \text{ bijective}\}$ of n objects form a group with group multiplication given by the composition of maps. Indeed, the composition of maps is associative, we have the identity map which serves as the neutral element and the inverse is given by the inverse map. In conclusions, the permutations S_n form a finite group of order $n!$ which is also referred to as *symmetric group*. Are these groups Abelian or non-Abelian? We begin with the simplest case of S_2 which has the two elements

$$S_2 = \left\{ e = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\} \tag{9.4}$$

(Recall, the above notation is a way of writing down permutations explicitly. It indicates a permutation which maps the numbers in the top row to the corresponding numbers in the bottom row.) Clearly, this group is Abelian since the second element commutes with itself and everything commutes with the identity (the first element). For S_n with $n > 2$ consider the two permutations

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ 1 & 3 & 2 & \cdots \end{pmatrix} , \qquad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ 2 & 1 & 3 & \cdots \end{pmatrix} , \tag{9.5}$$

where the dots stand for arbitrary permutations of the number $4, \ldots, n$. We have

$$\sigma_1 \cdot \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ 3 & 1 & 2 & \cdots \end{pmatrix} , \qquad \sigma_2 \cdot \sigma_1 = \begin{pmatrix} 1 & 2 & 3 & \cdots \\ 2 & 3 & 1 & \cdots \end{pmatrix} , \tag{9.6}$$

so that $\sigma_1 \cdot \sigma_2 \neq \sigma_2 \cdot \sigma_1$. Hence, the permutation groups for $n > 2$ are non-Abelian.

General linear groups: For a (finite-dimensional, say) vector space V, the set $\mathrm{Gl}(V) := \{g : V \to V \,|\, g \text{ linear and invertible}\}$ of all invertible, linear maps forms a group, called the *general linear group* of V. The group multiplication is composition of maps, the identity map is the neutral element and the inverse is given be the inverse map (which exists since we are considering only invertible linear maps). More specifically, we have the groups $\mathrm{Gl}(\mathbb{R}^n)$ of invertible, real $n \times n$ matrices and the groups $\mathrm{Gl}(\mathbb{C}^n)$ of invertible, complex $n \times n$ matrices, both with matrix multiplication as the group multiplication, the unit matrix as the neutral element and the matrix inverse as the inverse. General linear groups naturally act

on the vector spaces they are associated with and, therefore, realise the action of groups on vectors we would like to achieve for groups more generally. For this reason, they play an important role in the theory of representations, as we will see shortly. General linear groups have many interesting sub-groups, some of which we will now discuss.

Unitary and special unitary groups: The unitary group, $U(n)$, is defined as

$$U(n) := \{U \in \mathrm{Gl}(\mathbb{C}^n) \,|\, U^\dagger U = \mathbb{1}_n\} \,, \tag{9.7}$$

so it consists of all unitary $n \times n$ matrices. Why is this a group? Since $U(n)$ is clearly a subset of the general linear group $\mathrm{Gl}(\mathbb{C}^n)$ all we have to do is verify the three conditions for a sub-group in Exercise 9.2. Condition (i), closure, is obvious since the product of two unitary matrices is again unitary, as is condition (ii) since the unit matrix is unitary. To verify condition (iii) we need to show that with $U \in U(n)$ also the inverse $U^{-1} = U^\dagger$ is in $U(n)$. To do this, consider U as an element of the group $\mathrm{Gl}(\mathbb{C}^n)$, so that $U^\dagger U = \mathbb{1}_n$ implies $UU^\dagger = \mathbb{1}_n$ (since, in a group, the left inverse is the right inverse). The last equation can be written as $(U^\dagger)^\dagger U^\dagger = \mathbb{1}_n$ which shows that $U^\dagger \in U(n)$. Also note that the defining relation, $U^\dagger U = \mathbb{1}_n$ implies that the determinant of unitary matrices satisfies

$$|\det(U)| = 1 \,. \tag{9.8}$$

The simplest unitary group is

$$U(1) = \{e^{i\theta} \,|\, 0 \leq \theta < 2\pi\} \,, \tag{9.9}$$

which consists of complex numbers with length one. It is clearly Abelian. The observation that $U(1)$ can be seen as the unit circle, S^1, in the complex plane points to a more general feature - some groups are also manifolds and these groups are referred to as *Lie groups*. We will discuss this in more detail later.

As an example of a group homomorphism consider the map $f : \mathbb{Z}_n = \{0, 1, \ldots, n-1\} \to U(1)$ defined by

$$f(k) := e^{2\pi i k/n} \,. \tag{9.10}$$

Exercise 9.4. *Show that Eq. (9.10) defines an injective group homomorphism.*

To discuss higher-dimensional cases it is useful to introduce the *special unitary groups* $SU(n)$ which consist of all unitary matrices with determinant one, so

$$SU(n) := \{U \in U(n) \,|\, \det(U) = 1\} \,. \tag{9.11}$$

Exercise 9.5. *Show that $SU(n)$ is a sub-group of $U(n)$.*

What is the relationship between $U(n)$ and $SU(n)$? For an arbitrary unitary $A \in U(n)$ pick a solution ζ to the equation $\zeta^n = \det(A)$. (Due to Eq. (9.8) we have, of course, $|\zeta| = 1$.) Then $U := \zeta^{-1} A$ is a special unitary matrix since $\det(U) = \zeta^{-n}\det(A) = 1$ which shows that every unitary matrix A can be written as

$$A = \zeta\, U \,, \tag{9.12}$$

that is, as a product of a complex phase and a special unitary matrix. We can, therefore, focus on special unitary matrices. The lowest-dimensional non-trivial case is $SU(2)$ and inserting a general complex 2×2 matrix into the defining relations $U^\dagger U = \mathbb{1}_2$ and $\det(U) = 1$ gives

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \,|\, \alpha, \beta \in \mathbb{C} \,,\; |\alpha|^2 + |\beta|^2 = 1 \right\} \,. \tag{9.13}$$

Exercise 9.6. *Show that Eq. (9.13) provides the correct expression for the group $SU(2)$.*

The explicit form (9.13) of $SU(2)$ shows that this group is non-Abelian (as are all higher $SU(n)$ groups) and that it can be identified with the three-sphere S^3. This means we have another example of a group which is also a manifold, so a Lie group.

Finding explicit parametrisations along the lines of Eq. (9.13) for $SU(3)$ or even higher-dimensional cases is not practical anymore and we will later discuss more efficient ways of dealing with such groups.

Orthogonal and special orthogonal groups: This discussion is very much parallel to the previous one on unitary and special unitary groups, except it is based on the real, rather than the complex numbers. The orthogonal and special orthogonal groups (= rotations) are defined as

$$O(n) := \{A \in \mathrm{Gl}(\mathbb{R}^n) \,|\, A^T A = \mathbb{1}_n\} \tag{9.14}$$

$$SO(n) := \{R \in O(n) \,|\, \det(R) = \mathbb{1}_n\} \,, \tag{9.15}$$

and, hence, consist of all orthogonal matrices and all orthogonal matrices with determinant one, respectively.

Exercise 9.7. *Show that $O(n)$ is a sub-group of $\mathrm{Gl}(\mathbb{R}^n)$ and that $SO(n)$ is a sub-group of $O(n)$. (Hint: Proceed in analogy with the unitary case.) Also show that every $A \in O(N)$ has determinant $\det(A) = \pm 1$ and is either special orthogonal or can be written as $A = FR$, where $R \in SO(n)$ and $F = \mathrm{diag}(-1, 1, \ldots, 1)$.*

Just as for unitary groups, it is easy to deal with the two-dimensional case and show, by explicitly inserting an arbitrary 2×2 matrix into the defining relations, that $SO(2)$ is given by

$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \,|\, 0 \leq \theta < 2\pi \right\} \,. \tag{9.16}$$

Since this is parametrised by a circular coordinate, $SO(2)$ can also be thought of as a circle, S^1. This is a good opportunity to present another example of a group homomorphism $f : U(1) \to SO(2)$, defined by

$$f(e^{i\theta}) := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \,. \tag{9.17}$$

Exercise 9.8. *Show that the map (9.17) defines a bijective group homomorphism.*

The previous exercise shows that $U(1)$ and $SO(2)$ are isomorphic - as far as their group structure is concerned they represent the same object. Perhaps more surprisingly, we will see later that $SO(3)$ and $SU(2)$ are homeomorphic (and very nearly isomorphic) as well.

We could continue and write down an explicit parametrisation for $SO(3)$, for example in terms of a product of three two-dimensional rotations but will refrain from doing so explicitly. Later we will see that there are more efficient methods to deal with $SO(n)$ for $n > 2$.

The previous list provides us with sufficiently many interesting examples of groups and we now turn to our main task - representing groups.

9.2 Representations

Recall from the introduction that we are after some rule by which groups can act on vector spaces. We already know that the general linear group $\mathrm{Gl}(V)$ (where $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$, say), which consists of invertible linear maps (or matrices) on V, naturally acts on the vector space V. We can, therefore, achieve our goal if we embed an arbitrary group G into $\mathrm{Gl}(V)$. However, we want to do this in a way that preserves the group structure of G and this is precisely accomplished by group homomorphisms. This motivates the following definition of a *representation of a group*.

Definition 9.4. *(Representation of a group) A representation of a group G is a group homomorphism $R : G \to \mathrm{Gl}(V)$, where V is a vector space (typically taken to be $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$). The dimension of V is called the dimension, $\dim(R)$, of the representation R. A representation is called faithful if R is injective.*

The keyword in this definition is "group homomorphism" which means that the representation R satisfies

$$R(g_1 \cdot g_2) = R(g_1)R(g_2) \,, \tag{9.18}$$

for all $g_1, g_2 \in G$. This rule implies that the representation matrices $R(g)$ multiply "in the same way" as the original group elements g and it is for this reason that we call such an R a "representation". To illustrate the idea of a group representation we should discuss a few examples.

9.2.1 Examples of representations

Trivial representation: For every group G, we have the trivial representation defined by $R(g) = \mathbb{1}_n$ for all $g \in G$. Of course this representation is not faithful.

Matrix groups: Some of the examples of groups we have introduced earlier already consist of matrices (that is, they are sub-groups of the general linear group) and are, therefore, represented by themselves, taking R to be the identity map. This applies to the (special) unitary groups $(S)U(n)$ and the (special) orthogonal groups $(S)O(n)$. In each case, the dimension of the representation is n (since the matrices act on n-dimensional vectors) and the representation is faithful. These particular representations are frequently called the *fundamental representations*. Of course, these groups have other, less trivial representations and we will see examples later.

Representations of $\mathbb{Z}_n$: For the cyclic groups $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ and any $q = 0, \ldots, n-1$ we can write down a one-dimensional representation $R_{(q)} : \mathbb{Z}_n \to \mathrm{Gl}(\mathbb{C}) =: \mathbb{C}^*$ by

$$R_{(q)}(k) := e^{2\pi i q k/n} \,. \tag{9.19}$$

Representations of $U(1)$: For $U(1) = \{e^{i\theta} \,|\, 0 \leq \theta < 2\pi\}$ and any $q \in \mathbb{Z}$ we can define a one-dimensional, faithful representation $R_{(q)} : U(1) \to \mathrm{Gl}(\mathbb{C}) =: \mathbb{C}^*$ by

$$R_{(q)}(e^{i\theta}) = e^{iq\theta} \,. \tag{9.20}$$

Direct sum representations: For two representations $R : G \to \mathrm{Gl}(V)$ and $\tilde{R} : G \to \mathrm{Gl}(\tilde{V})$ of the same group G, we can define the *direct sum representation* $R \oplus \tilde{R} : G \to \mathrm{Gl}(V \oplus \tilde{V})$ by

$$(R \oplus \tilde{R})(g) := \begin{pmatrix} R(g) & 0 \\ 0 & \tilde{R}(g) \end{pmatrix} \,, \tag{9.21}$$

that is, be simply arranging the representation matrices $R(g)$ and $\tilde{R}(g)$ into a block-matrix. Obviously dimensions sum up, so $\dim(R \oplus \tilde{R}) = \dim(R) + \dim(\tilde{R})$.

As an explicit example, consider the above one-dimensional representations $R_{(1)}$ and $R_{(-1)}$ of $U(1)$, taking $q = \pm 1$ in Eq. (9.20). Their direct sum representation is two-dimensional and given by

$$(R_{(1)} \oplus R_{(-1)})(e^{i\theta}) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \,. \tag{9.22}$$

Tensor representations: For this we need a bit of preparation. Consider two square matrices A, B of size n and m, respectively. By their *Kronecker product* $A \times B$ we mean the square matrix of size nm

obtained by replacing each entry in A with that entry times the entire matrix B. The Kronecker product satisfies the useful rule

$$(A \times B)(C \times D) = (AC) \times (BD) \,, \tag{9.23}$$

provided the sizes of the matrices A, B, C, D fit as required.

Now consider two representations $R : G \to \mathrm{Gl}(V)$ and $\tilde{R} : G \to \mathrm{Gl}(\tilde{V})$ of the same group G. The *tensor representation* $R \otimes \tilde{R} : G \to \mathrm{Gl}(V \otimes \tilde{V})$ is defined by

$$(R \otimes \tilde{R})(g) := R(g) \times \tilde{R}(g) \,. \tag{9.24}$$

Given the definition of the Kronecker product, dimensions multiply, so $\dim(R \otimes \tilde{R}) = \dim(R)\dim(\tilde{R})$.

As an explicit example, consider the two-dimensional fundamental representation of $SU(2)$ given by $R(U) = U$ of any $U \in SU(2)$. Then the tensor representation $R \otimes R$ is four-dimensional and given by

$$(R \otimes R)(U) = \begin{pmatrix} U_{11}\,U & U_{12}\,U \\ U_{21}\,U & U_{22}\,U \end{pmatrix} \,, \tag{9.25}$$

Exercise 9.9. *Use Eq. (9.23) to show that Eq. (9.24) does indeed define a representation.*

This should be enough examples for now.

9.2.2 Properties of representations

To talk about representations efficiently we need to introduce a few more pieces of standard terminology.

Definition 9.5. *Two representations $R : G \to \mathrm{Gl}(V)$ and $\tilde{R} : G \to \mathrm{Gl}(V)$ are called equivalent if there is an invertible linear map $P : V \to V$ such that $\tilde{R}(g) = PR(g)P^{-1}$ for all $g \in G$.*

In other words, if the representation matrices of two representations only differ by a common basis transformation they are really the same representation and we call them equivalent. A further simple piece of terminology is the following.

Definition 9.6. *A representation $R : G \to \mathrm{Gl}(V)$ on an inner product vector space V is called unitary iff all $R(g)$ are unitary with respect to the inner product on V.*

More practically, if $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$ with their respective standard scalar products, the unitary representations are precisely those with all representation matrices orthogonal or unitary matrices. For example, the $U(1)$ representations (9.20) and (9.22) are unitary.

In Eq. (9.21) we have seen that larger representations can be built up from smaller ones by forming a direct sum. Conversely we can ask whether a given representation can be split up into a direct sum of smaller representations. In this way, we can attempt to break up representations into their smallest building block which cannot be decomposed further in this way and which are called *irreducible representation*.

Definition 9.7. *A representation $R : G \to \mathrm{Gl}(V)$ is called irreducible if there is no non-trivial sub-vector space $W \subset V$ (that is, no sub-space other than $W = \{\mathbf{0}\}$ and $W = V$) invariant under R, that is, there is no non-trivial W for which $R(g)(W) \subset W$ for all $g \in G$. Otherwise, the representation is called reducible.*

For example, all fundamental representations of $(S)O(n)$ and $(S)U(n)$ are irredcible.

Exercise 9.10. *Show that the fundamental representations of $SU(2)$ and $SO(3)$ are irreducible.*

What does it mean in practice for a representation to be reducible? Suppose we have a non-trivial sub-space $W \subset V$ invariant under the representation R and we choose a basis for W which we complete to a basis for V. Relative to this basis, the representation matrices have the form

$$R(g) = \begin{pmatrix} A(g) & B(g) \\ 0 & D(g) \end{pmatrix} \begin{matrix} W \\ \text{rest} \end{matrix} , \tag{9.26}$$

where $A(g)$, $B(g)$ and $D(g)$ are matrices. The zero in the lower left corner is forced upon us to ensure a vector in W is mapped to a vector in W as required by invariance. The form (9.26) is not quite a direct sum due to the presence of the matrix $B(g)$. To deal with this we define the following.

Definition 9.8. *A reducible representation is called fully reducible if it is a direct sum of irreducible representations.*

In practice, this means we can achieve a form for which all matrices $B(g)$ in Eq. (9.26) vanish. Not all reducible representations are fully reducible but we have

Theorem 9.11. *All reducible representations of finite groups and all reducible unitary representations are fully reducible.*

Proof. The proof can be found in Ref [13]. □

A common problem is having to find the irreducible representations contained in a given (fully) reducible representation. If this representation is given as a direct sum, $R \oplus \tilde{R}$, with R and $\tilde{R}$ irreducible, then this task is trivial - the irreducible pieces are simply R and $\tilde{R}$. The tensor product $R \otimes \tilde{R}$ is more interesting. It is not obviously block-diagonal but, as it turns out, it is usually reducible. This means there is a decomposition, also called *Clebsch-Gordan decomposition*,

$$R \otimes \tilde{R} \cong R_1 \oplus \cdots \oplus R_k \tag{9.27}$$

of a tensor product into irreducible representations R_i. We will later study this decomposition more explicitly for the case of $SU(2)$.

Since the interest in physics is usually in orthogonal or unitary matrices Theorem 9.11 covers the cases we are interested in and it suggests a classification problem. For a given group G we should find all irreducible representations - the basic building blocks from which all other representations (subject to the conditions of the above theorem) can be obtained by forming direct sums. A very helpful statement for this purpose (which is useful in other context, for example in quantum mechanics, as well) is the famous *Schur's Lemma*.

Lemma 9.1. *(Schur's Lemma) Let $R : G \to \mathrm{Gl}(V)$ be an irreducible representations of the group G over a complex vector space V and $P : V \to V$ a linear map with $[P, R(g)] = 0$ for all $g \in G$. Then $P = \lambda \, \mathrm{id}_V$, for a complex number λ.*

Proof. Since we are working over the complex numbers the characteristic polynomial for P has at least one zero, λ, and the associated eigenspace $\mathrm{Eig}_P(\lambda)$ is non-trivial. For any $\mathbf{v} \in \mathrm{Eig}_P(\lambda)$, using $[P, R(g)] = 0$, we have

$$PR(g)\mathbf{v} = R(g)P\mathbf{v} = \lambda R(g)\mathbf{v} . \tag{9.28}$$

Hence, $R(g)\mathbf{v}$ is also an eigenvector with eigenvalue λ and we conclude that the eigenspace $\mathrm{Eig}_P(\lambda)$ is invariant under R. However, R is irreducible by assumption which means there are no non-trivial invariant sub-spaces. Since $\mathrm{Eig}_P(\lambda) \neq \{\mathbf{0}\}$ the only way out is that $\mathrm{Eig}_P(\lambda) = V$. This implies that $P = \lambda \, \mathrm{id}_V$. □

Schur's lemma says in essence that a matrix commuting with all (irreducible) representation matrices of a group must be a multiple of the unit matrix. This can be quite a powerful statement. However, note the condition that the representation is over a complex vector space - the theorem fails in the real case. A counterexample is provided by the fundamental representation of $SO(2)$, given by the matrices (9.16). Seen as a representation over $\mathbb{R}^2$ it is irreducible but all representation matrices of $SO(2)$ commute with one another.

An immediate conclusion from Schur's Lemma is the following.

Corollary 9.1. *All complex irreducible representations of an Abelian group are one-dimensional.*

Proof. For an Abelian group G we have $g \circ \tilde{g} = \tilde{g} \circ g$ for all $g, \tilde{g} \in G$ which implies $[R(g), R(\tilde{g})] = 0$ for any representation R. The linear map $P = R(g)$ then satisfies all the conditions of Schur's Lemma and we conclude that $R(g) = \lambda(g)\,\mathrm{id}_V$. However, this form is only consistent with R being irreducible if $\dim(R) = 1$. □

This statement is the key to finding all complex, irreducible representations of Abelian groups and we discuss this for the two most important Abelian examples, $\mathbb{Z}_n$ and $U(1)$.

Application 9.32. All complex irreducible representations of $\mathbb{Z}_n$

The group $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ is Abelian so all complex irreducible representations R must be one-dimensional from the previous lemma. The identity element $0 \in \mathbb{Z}_n$ must be mapped to the identity "matrix", so $R(0) = 1$. Let us set $R(1) = \zeta \in \mathrm{Gl}(\mathbb{C}) = \mathbb{C}^*$. It follows that

$$1 = R(0) = R(\underbrace{1 + 1 + \cdots 1}_{n \text{ times}}) = R(1)^n = \zeta^n \,, \tag{9.29}$$

so ζ must be an n^{th} root of unity, $\zeta = e^{2\pi i q/n}$, where $q = 0, \ldots, n-1$. Note that the choice of $R(1) = \zeta$ determines the entire representation since $R(k) = R(1)^k = \zeta^k$. This means we have precisely n complex, irreducible representations of $\mathbb{Z}_n$ given by

$$R_{(q)}(k) = e^{2\pi i q k/n} \quad \text{where} \quad q = 0, 1, \ldots, n-1 \,. \tag{9.30}$$

Application 9.33. All complex irreducible representations of $U(1)$

Arguing for all (continuous) irreducible, complex representations of $U(1)$ works in much the same way, with the conclusion that the representations (9.20) with $q \in \mathbb{Z}$ form a complete set.

Exercise 9.12. *Show that Eq. (9.20), where $q \in \mathbb{Z}$ is a complete list of all complex, irreducible $U(1)$ representations. (Hint: Start by considering representation matrices for group element $e^{i\theta}$ where θ is rational and then use continuity.)*

In a physics context, the integer q which labels the above $\mathbb{Z}_n$ and $U(1)$ representations is also called the *charge*. As you will learn later in the physics course, fixing the electrical charge of a particle (such as the electron charge), mathematically amounts to choosing a $U(1)$ representation for this particle.

As we have seen above, complex, irreducible representations of Abelian groups are quite straightforward to classify, essentially because the representation "matrices" are really just numbers. The somewhat "brute-force" approach taken above becomes impractical if not unfeasible for the non-

Abelian case. Just imagine having to write down, for a two-dimensional representation, an arbitrary 2×2 matrix as an Ansatz for each representation matrix and then having to fix the unknown entries by imposing the group multiplication table on the matrices. Clearly, we need more sophisticated methods to deal with the non-Abelian case. There is a beautiful set of methods for finite non-Abelian groups, using *characters* but discussing this in detail is beyond the scope of this lecture. If you are interested, have a look at Ref. [13]. Instead, we focus on Lie groups, which we now define and analyse.

9.3 Lie groups and Lie algebras

9.3.1 Definition of Lie group

A Lie group is formally defined as follows.

Definition 9.9. *A group G is a Lie group if it is a differentiable manifold and the left-multiplication with group elements and the inversion of group elements are differentiable maps.*

It is difficult to exploit this definition without talking in some detail about differentiable manifolds, which is well beyond the scope of this lecture. Fortunately, for our purposes we can adopt a somewhat more practical definition of a Lie group which follows from the more abstract one above.

A *matrix Lie group* G is a group given (at least around a neighbourhood of the group identity $\mathbb{1}$), by a family $g = g(\mathbf{t})$ of $n \times n$ matrices, which depend (in an infinitely differentiable way) on real parameters $\mathbf{t} = (t^1, \ldots, t^k)$ and such that the matrices $\frac{\partial g}{\partial t_1}, \ldots \frac{\partial g}{\partial t_k}$ are linearly independent. (This last requirement is really the same as the maximal rank condition in Theorem B.1.) Note that the matrices $g(\mathbf{t})$ act on the vector space $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$. For convenience, we assume that $g(\mathbf{0}) = \mathbb{1}_n$ is the group identity. The number, k, of parameters required is called the *dimension of the Lie group*. (Do not confuse this dimension of the matrix Lie group with the dimension n of the vector space V on which these matrices act.)

Obvious examples of Lie groups are $U(1)$ in Eq. (9.9) and $SO(2)$ in Eq. (9.16), both of which are parametrised by one angle (playing the role of the single parameter t_1) and are, hence, one-dimensional. A more interesting example is provided by $SU(2)$. We can solve the constraint on α and β in Eq. (9.13) by setting $\beta = -t_2 + it_1$ and $\alpha = \sqrt{1 - t_1^2 - t_2^2} e^{it_3}$. This leads to the explicit parameterization

$$U = \begin{pmatrix} \sqrt{1 - t_1^2 - t_2^2} e^{it_3} & -t_2 + it_1 \\ t_2 + it_1 & \sqrt{1 - t_1^2 - t_2^2} e^{-it_3} \end{pmatrix} , \tag{9.31}$$

which shows that $SU(2)$ is a three-dimensional Lie group. A similar argument can be made for $SO(3)$ which can be parametrised in terms of three angles and is, hence, a three-dimensional Lie group as well. In fact, all orthogonal and unitary groups (as well as their special sub-groups) can be parametrised in this way and are Lie groups. However, writing down these parametrisations explicitly becomes unpractical in higher dimensions.

9.3.2 Definition of Lie algebra

Instead, there are much more efficient methods of dealing with Lie-groups which are based on the observation that, for many purposes, it is already sufficient to consider "infinitesimal" transformations, that is, group elements "near" the identity. The set of these infinitesimal transformations of a group G is called the *Lie algebra*, of G. Abstractly, a Lie algebra is defined as follows.

Definition 9.10. *A Lie algebra $\mathcal{L}$ is a vector space with a commutator bracket $[\cdot,\cdot]:\mathcal{L}\times\mathcal{L}\to\mathcal{L}$ which is anti-symmetric, so $[T,S]=-[S,T]$ and satisfies the Jacobi identity $[T,[S,U]]+[S,[U,T]]+[U,[T,S]]=0$ for all $T,S,U\in\mathcal{L}$.*

Let us see how we can associate a Lie algebra in this sense to our group G. We start with the *generators*, T_i, of the group defined by [9].

$$T_i := -i\frac{\partial g}{\partial t^i}(\mathbf{0})\ . \tag{9.32}$$

In terms of the generators we can think of group elements near the identity as given by the Taylor series

$$g(\mathbf{t}) = \mathbb{1} + i\sum_{i=1}^{k} t^i T_i + \mathcal{O}(t^2)\ . \tag{9.33}$$

The Lie algebra $\mathcal{L}(G)$ is the vector space of matrices spanned (over $\mathbb{R}$) by the generators, that is,

$$\mathcal{L}(G) = \mathrm{Span}_{\mathbb{R}}(T_1,\ldots,T_k)\ . \tag{9.34}$$

By definition of a matrix Lie group the generators must be linearly independent, so the dimension of $\mathcal{L}(G)$ as a vector space is the same as the dimension of the underlying group as a manifold. Now we understand how to obtain the generators from the Lie group. Is there a way to reverse this process and obtain the Lie group from its generators? Amazingly, the answer is "yes" due to the following theorem.

Theorem 9.13. *Let G be a (matrix) Lie group and $\mathcal{L}(G)$ as defined in Eq. (9.34). Then the matrix exponential $\exp(i\,\cdot)$ provides a map $\exp(i\,\cdot):\mathcal{L}(G)\to G$ whose image is the part of G which is (path)-connected to the identity.*

Proof. See, for example, Ref. [13]. □

Application 9.34. The exponential map for SU(2)

Theorem 9.13 can be verified explicitly for $SU(2)$ where the matrix exponential can be carried out explicitly. Setting $\mathbf{t}=2\theta\mathbf{n}$ with a unit vector $\mathbf{n}$ we have

$$g(\mathbf{t}) = \exp(it^j\tau_j) = \sum_{n=0}^{\infty}\frac{(i\theta)^n}{n!}(\mathbf{n}\cdot\boldsymbol{\sigma})^n = \cos(\theta)\mathbb{1}_2 + i\sin(\theta)\mathbf{n}\cdot\boldsymbol{\sigma}\ . \tag{9.35}$$

This gives all SU(2) matrices as comparison with Eq. (9.13) shows,

We can also recover the group $SO(3)$ by forming the matrix exponential

$$g(\mathbf{t}) = \exp(it^i\tilde{T}_i)\ , \tag{9.36}$$

with the matrices $\tilde{T}_i$ in Eq. (9.51). (Note, that $O(3)$, which has the same Lie algebra as $SO(3)$ cannot be fully recovered by the matrix exponential since the orthogonal matrices with determinant -1 are not path-connected to the identity.)

We know that $\mathcal{L}(G)$ is a vector space and we can define a bracket by the simple matrix commutator $[T,S]:=TS-ST$. Clearly, this bracket is anti-symmetric and a simple calculation shows that it satisfies the Jacobi identity.

[9] The factor of $-i$ is included so that the generators become hermitian (rather than anti-hermitian).

Exercise 9.14. *Show that the matrix commutator bracket $[T,S] = TS - ST$ on $\mathcal{L}(G)$ satisfies the Jacobi identity.*

All that remains to be shown for $\mathcal{L}(G)$ to be a Lie-algebra in the sense of Def. 9.10 is that it is closed under the bracket, that is, if $T, S \in \mathcal{L}(G)$ then $[T,S] \in \mathcal{L}(G)$. This follows from the closure of G under multiplications. Start with two group element $g(\mathbf{t})$ and $g(\mathbf{s})$, expand each to second order using the matrix exponential and consider the combination

$$g(\mathbf{t})^{-1} g(\mathbf{s})^{-1} g(\mathbf{t}) g(\mathbf{s}) = 1 - \sum_{i,j=1}^{k} t^i s^j [T_j, T_k] + \cdots \,. \tag{9.37}$$

Since the LHS of this equation must be a group element we conclude the commutators $[T,S]$ is indeed an element of $T, S \in \mathcal{L}(G)$. In conclusion $\mathcal{L}(G)$, the vector space spanned by the generators, together with the matrix commutator, forms a Lie algebra in the sense of Def. 9.10.

Since $[T_i, T_j] \in \mathcal{L}(G)$ and the generators form a basis of $\mathcal{L}(G)$ it is clear that there must be constants ${f_{ij}}^k$, also called *structure constants of the Lie algebra*, such that

$$[T_i, T_j] = {f_{ij}}^k T_k \,. \tag{9.38}$$

Eq. (9.33) shows why we should think of the Lie algebra $\mathcal{L}(G)$ as encoding "infinitesimal" group transformations. Consider a vector $\mathbf{v} \in V$ which transforms under a group element $g(\mathbf{t})$ as $\mathbf{v} \to g(\mathbf{t})\mathbf{v}$. Then inserting the expansion (9.33), it follows that

$$\mathbf{v} \to g(\mathbf{t})\mathbf{v} = \left(\mathbb{1}_n + i \sum_{i=1}^{k} t^i T_i + \mathcal{O}(t^2) \right) \mathbf{v} \quad \Rightarrow \quad \delta\mathbf{v} := g(\mathbf{t})\mathbf{v} - \mathbf{v} = i \sum_{i=1}^{k} t^i T_i \mathbf{v} + \mathcal{O}(t^2) \,. \tag{9.39}$$

9.3.3 Examples of Lie groups and their algebras

Let us compute a few Lie algebras explicitly.

Lie algebra of $U(1)$ The group $U(1)$ depends on one parameter t_1, see Eq. (9.9), and we find for the single generator

$$g(t_1) = e^{i t_1} \quad \Rightarrow \quad T_1 = -i \frac{\partial g}{\partial t_1}(0) = 1 \,. \tag{9.40}$$

Hence, the Lie algebra, $\mathcal{L}(U(1)) = \mathbb{R}$, is one-dimensional and the commutator bracket and the structure constants are trivial.

Lie algebra of $SO(2)$ The group $SO(2)$ also depends on a single parameter t_1, see Eq. (9.16), and for the generator we have

$$g(t_1) = \begin{pmatrix} \cos t_1 & \sin t_1 \\ -\sin t_1 & \cos t_1 \end{pmatrix} \quad \Rightarrow \quad T_1 = -i \frac{\partial g}{\partial t_1}(0) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \,. \tag{9.41}$$

This means the Lie algebra $\mathcal{L}(SO(2))$ is one-dimensional and consists of (i times) the 2×2 anti-symmetric matrices.

Lie algebra of $SU(2)$ This is a more interesting example, since $SU(2)$ depends on three parameters t_i, where $i = 1, 2, 3$ and taking the derivatives of the group elements in Eq. (9.31) we find for the generators

$$T_i = -i \frac{\partial g}{\partial t_i} = \sigma_i \,, \tag{9.42}$$

where σ_i are the Pauli matrices. Recall that the Pauli matrices satisfy the useful relation

$$\sigma_i \sigma_j = \mathbb{1}_2 \delta_{ij} + i\epsilon_{ijk}\sigma_k \ . \tag{9.43}$$

It is customary to take as the standard generators for $\mathcal{L}(SU(2))$ the matrices

$$\tau_i := \frac{1}{2}\sigma_i \ . \tag{9.44}$$

Hence, the Lie-algebra of $SU(2)$ is three-dimensional and given by

$$\mathcal{L}(SU(2)) = \text{Span}(\tau_i) = \{\text{hermitian, traceless } 2\times 2 \text{ matrices}\} \tag{9.45}$$

Given the commutation relations $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$ (which follow from Eq. (9.43)) for the Pauli matrices we have

$$[\tau_i, \tau_j] = i\epsilon_{ijk}\tau_k \ , \tag{9.46}$$

for the standard generators τ_i so that the structure constants are $f_{ijk} = i\epsilon_{ijk}$.

Exercise 9.15. *Verify the relation (9.43) for the Pauli matrices. (Hint: Show that the Pauli matrices square to the unit matrix and that the product of two different Pauli matrices is $\pm i$ times the third.) Show from Eq. (9.43) that $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$ and that $\text{tr}(\sigma_i\sigma_j) = 2\delta_{ij}$.*

The method for computing the Lie algebra illustrated above becomes impractical for higher-dimensional groups since it requires an explicit parametrisation. However, there is a much more straightforward way to proceed.

Lie algebras of unitary and special unitary groups: To work out the Lie algebra of unitary groups $U(n)$ in general start with the Ansatz $U = \mathbb{1}_n + iT + \cdots$ and insert this into the defining relation $U^\dagger U = \mathbb{1}_n$, keeping only terms up to linear order in T, to work out the resulting constraint on the generators. Doing this in the present case results in $T = T^\dagger$, so the generators must be hermitian and the Lie algebra is

$$\mathcal{L}(U(n)) = \{\text{hermitian } n\times n \text{ matrices}\} \quad \Rightarrow \quad \dim(\mathcal{L}(U(n))) = n^2 \ . \tag{9.47}$$

For the case of $SU(n)$ we have to add the condition $\det(U) = 1$ which leads to $\det(\mathbb{1}_n + iT + \cdots) = 1 + i\,\text{tr}(T) + \cdots \stackrel{!}{=} 1$, so that $\text{tr}(T) = 0$. This means

$$\mathcal{L}(SU(n)) = \{\text{traceless hermitian } n\times n \text{ matrices}\} \quad \Rightarrow \quad \dim(\mathcal{L}(SU(n))) = n^2 - 1 \ . \tag{9.48}$$

We can now choose a basis and compute structure constants for these Lie algebras but we will not do this explicitly, other than for the case of $SU(2)$ which we have already covered.

Lie algebras of orthogonal and special orthogonal groups: This works in analogy with the unitary case. Inserting the Ansatz $A = \mathbb{1}_n + iT + \cdots$ into the defining relation $A^T A = \mathbb{1}_n$ for $O(n)$ (where T has to be purely imaginary, so that A is real) leads to $T = -T^T$, so that T must be anti-symmetric. Since anti-symmetric matrices are already traceless the additional condition $\det(A) = 1$ for $SO(n)$ does not add anything new and we have

$$\mathcal{L}(O(n)) = \mathcal{L}(SO(n)) = \{\text{anti-symmetric, purely imaginary } n\times n \text{ matrices}\} \tag{9.49}$$

$$\Rightarrow \quad \dim(\mathcal{L}(O(n))) = \dim(\mathcal{L}(SO(n))) = \frac{1}{2}n(n-1) \ . \tag{9.50}$$

We will only discuss a choice of basis for the simplest non-trivial case $\mathcal{L}(SO(3))$ which is spanned by the three matrices

$$(\tilde{T}_i)_{jk} = -i\epsilon_{ijk} \tag{9.51}$$

which are more explicitly given by

$$\tilde{T}_1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} , \quad \tilde{T}_2 = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} , \quad \tilde{T}_3 = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{9.52}$$

They satisfying the commutation relations

$$[\tilde{T}_i, \tilde{T}_j] = i\epsilon_{ijk}\tilde{T}_k \, . \tag{9.53}$$

Hence the structure constants are $f_{ijk} = i\epsilon_{ijk}$.

Exercise 9.16. *Verify the $SO(3)$ commutation relations (9.53).*

We have seen above that the generators of unitary (or orthogonal) groups are hermitian. Conversely, we can show that the exponential map acting on hermitian generators leads to unitary group elements. To do this first note that for two matrices X, Y with $[X, Y] = 0$ we have $\exp(X)\exp(Y) = \exp(X + Y)$. Also note that $(\exp(X))^\dagger = \exp(X^\dagger)$. It follows for a hermitian matrix T that

$$(\exp(iT))^\dagger \exp(iT) = \exp(-iT)\exp(iT) = \exp(0) = \mathbb{1} \, , \tag{9.54}$$

so that $\exp(iT)$ is indeed unitary.

Exercise 9.17. *Show that $(\exp(X))^\dagger = \exp(X^\dagger)$ for a square matrix X. Also show that for two square matrices X, Y with $[X, Y] = 0$ we have $\exp(X)\exp(Y) = \exp(X+Y)$. (Hint: Use the series which defines the exponential function.)*

Exercise 9.18. *Proof equation (9.35).(Hint: Use Eq. (9.43).) Further, show that the RHS of Eq. (9.35) provides all $SU(2)$ matrices.*

9.3.4 Lie algebra representations

Let us come back to our original goal of finding representations of Lie groups. Since, as we have seen, the group can be recovered from the algebra, it makes sense to discuss representations at the level of the algebra (and then use the matrix exponential to lift to the group). For this, we need the idea of a representation of a Lie algebra.

Definition 9.11. *A representation r of a Lie algebra $\mathcal{L}$ is a linear map $r : \mathcal{L} \to \mathrm{Hom}(V, V)$ which preserves the bracket, that is, $r([T, S]) = [r(T), r(S)]$.*

The notions of reducible, irreducible and fully reducible representations we have introduced for group representations directly transfer to representations of the Lie algebra. Note that the space $\mathrm{Hom}(V, V)$ is, in practice, the space of $n \times n$ real matrices for $V = \mathbb{R}^n$ or the space of $n \times n$ complex matrices for $V = \mathbb{C}^n$. So a representation of a Lie algebra amounts to assigning (linearly) to each Lie algebra element T a matrix $r(T)$ such that the matrices commute "in the same way" as the Lie algebra elements. A practical way of stating what a Lie algebra representation is in terms of a basis T_i of the Lie algebra. Suppose the T_i commute as

$$[T_i, T_j] = f_{ij}{}^k \, T_k \, . \tag{9.55}$$

Then the assignment $T_i \to r(T_i)$ defines a Lie algebra representation provided the matrices $r(T_i)$ commute on the same structure constants, that is,

$$[r(T_i), r(T_j)] = f_{ij}{}^k \, r(T_k) \, . \tag{9.56}$$

As an example for how this works in practice, consider the three matrices τ_i in Eq. (9.44) which form a basis of $\mathcal{L}(SU(2))$. Any assignment $\tau_i \to T_i$ to matrices T_i which commute on the same structure constants as the τ_i, that is $[T_i, T_j] = i\epsilon_{ijk}T_k$, defines a Lie algebra representation of $\mathcal{L}(SU(2))$. In fact, we have already seen an example of such matrices, namely the matrices $\tilde{T}_i$ in Eq. (9.51) which form a basis of the Lie-algebra $\mathcal{L}(SO(3))$. Hence, we see that the Lie algebra of $SO(3)$ is a representation of the Lie-algebra of $SU(2)$ - a clear indication that those two groups are closely related.

The idea for how we want to proceed discussing representations is summarised in the following diagram.

$$\begin{array}{ccc} G & \xrightarrow{R} & \mathrm{Gl}(V) \\ \exp \uparrow & & \uparrow \exp \\ \mathcal{L}(G) & \xrightarrow{r} & \mathrm{Hom}(V,V) \end{array} \tag{9.57}$$

Instead of studying representations R of the Lie group G we will be studying representations r of its Lie-algebra $\mathcal{L}(G)$. When required, we can use the exponential map to reconstruct the group G and its representation matrices in $\mathrm{Gl}(V)$. In other words, suppose we have a Lie-algebra element $T \in \mathcal{L}(G)$ and the associated group element $g = \exp(iT) \in G$. Then the relation between a representation r of the Lie algebra and the corresponding representation R at the group level is summarised by

$$T \mapsto r(T)\ , \qquad g = \exp(iT) \mapsto R(g) = \exp(ir(T))\ . \tag{9.58}$$

We will now explicitly discuss all this for the group $SU(2)$ and its close relative $SO(3)$.

9.4 The groups $SU(2)$ and $SO(3)$

9.4.1 Relationship between $SU(2)$ and $SO(3)$

Our first step is to understand the relationship between $SU(2)$ and $SO(3)$ better. We have already seen that their Lie algebras are representations of each other but what about the groups? We will now see that $SO(3)$ is, in fact, a group representation of $SU(2)$.

To do this it is useful to define the map $\varphi : \mathbb{R}^3 \to \mathcal{L}(SU(2))$ by

$$\varphi(\mathbf{t}) := t^i \sigma_i\ , \tag{9.59}$$

so that φ identifies $\mathbb{R}^3$ with $\mathcal{L}(SU(2))$. Note that the dot product between two vectors $\mathbf{t}, \mathbf{s} \in \mathbb{R}^3$ can then be written as

$$\mathbf{t} \cdot \mathbf{s} = \frac{1}{2}\mathrm{tr}(\varphi(\mathbf{t})\varphi(\mathbf{s}))\ , \tag{9.60}$$

as a result of the identity $\mathrm{tr}(\sigma_i\sigma_j) = 2\,\delta_{ij}$ which follows from Eq. (9.43).

Exercise 9.19. *Proof the relation (9.60), by using Eq. (9.43).*

We now define a three-dimensional representation $R : SU(2) \to \mathrm{Gl}(\mathbb{R}^3)$ of $SU(2)$ by

$$R(U)(\mathbf{t}) := \varphi^{-1}\left(U\varphi(\mathbf{t})U^\dagger\right)\ . \tag{9.61}$$

Note that this makes sense. The matrix $U\varphi(\mathbf{t})U^\dagger$ is hermitian traceless and, hence, an element of $\mathcal{L}(SU(2))$. Therefore, we can associate to it, via the inverse map φ^{-1}, a vector in $\mathbb{R}^3$. In order to study the representation R further, we analyse its effect on the dot product.

$$(R(U)\mathbf{t}) \cdot (R(U)\mathbf{s}) = \frac{1}{2}\mathrm{tr}\left(\varphi(R(U)\mathbf{t})\varphi(R(U)\mathbf{s})\right) = \frac{1}{2}\mathrm{tr}\left(U\varphi(\mathbf{t})U^\dagger U\varphi(\mathbf{s})U^\dagger\right) = \frac{1}{2}\mathrm{tr}(\varphi(\mathbf{t})\varphi(\mathbf{s})) = \mathbf{t}\cdot\mathbf{s} \tag{9.62}$$

This shows that $R(U)$ leaves the dot product invariant and, therefore, $R(U) \in O(3)$. A connectedness argument shows that, in fact, $R(U) \in SO(3)$ [10]. The representation R is not faithful since, from Eq. (9.61), U and $-U$ are mapped to the same rotation. In fact, R is precisely a two-to-one map (two elements of $SU(2)$ are mapped to one element of $SO(3)$) as can be shown in the following exercise.

Exercise 9.20. *Show that the kernel of the representation* (9.61) *is given by* $\mathrm{Ker}(R) = \{\pm \mathbb{1}_2\}$. *(Hint: Use Schur's Lemma.)*

It can also be verified by explicit calculation that $\mathrm{Im}(R) = SO(3)$. In conclusion, we have seen that the groups $SU(2)$ and $SO(3)$ are very closely related - the former is what is called a double-cover of the latter. We have a two-to-one representation map $R : SU(2) \to SO(3)$ which allows us to recover $SO(3)$ from $SU(2)$, but not the other way around since R is not invertible. From this point of view the group $SU(2)$ is the more basic object. In particular, all representations of $SO(3)$ are also representations of $SU(2)$ (just combine the $SO(3)$ representation with R) but not the other way around. It is already clear that the group $SO(3)$ plays an important role in physics - just think of rotationally symmetric problems in classical mechanics, for example. The above result strongly suggest that $SU(2)$ is also an important group for physics and, in quantum mechanics, this turns out to be the case.

9.4.2 All complex irreducible representations

Now that we have clarified the relationship between the groups let us find all (complex) irreducible and unitary representations of their associated Lie algebras. We consider representations $r : \mathcal{L}(SU(2)) \to \mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^n)$ by complex $n \times n$ matrices, and denote the representation matrices of the generators by $J_i := r(\tau_i)$. (Since we are after unitary representations the J_i should be hermitian.) For the J_i to form a representation they must satisfy the same commutation relations as the τ_i in Eq. (9.46), namely

$$[J_i, J_j] = i\epsilon_{ijk} J_k \ . \tag{9.63}$$

Finding all (finite-dimensional, hermitian) irreducible matrices J_i which satisfy this relation is a purely algebraic problem as we will now see. First we introduce a new basis

$$J^{\pm} := J_1 \pm i J_2 \ , \quad J_3 \tag{9.64}$$

on our Lie-alebra [11] and re-express the commutation relations (9.63) in this basis:

$$[J_3, J_\pm] = \pm J_\pm \ , \qquad [J_+, J_-] = 2J_3 \ . \tag{9.65}$$

We also introduce the matrix

$$J^2 := J_1^2 + J_2^2 + J_3^2 = \frac{1}{2}(J_+ J_- + J_- J_+) + J_3^2 \tag{9.66}$$

Combining Eqs. (9.65) and (9.66) leads to the useful relations

$$J_- J_+ = J^2 - J_3(J_3 + 1) \ , \qquad J_+ J_- = J^2 - J_3(J_3 - 1) \ . \tag{9.67}$$

A key observation is now that J^2 commutes with all matrices J_i, that is,

$$[J^2, J_i] = 0 \quad \text{for} \quad i = 1, 2, 3 \ . \tag{9.68}$$

An operator which commutes with the entire Lie algebra, such as J^2 above, is also called a *Casimir operator*.

[10] Since $SU(2)$ is path-connected every $U \in SU(2)$ can be connected to the identity $\mathbb{1}_2$. Hence, every image $R(U)$ must be connected to the identity $\mathbb{1}_3$ but we know that only elements in $SO(3)$ are.

[11] There is a subtlety here. We have defined the Lie-algebra as the real span of the generators but our re-definition includes a factor of i. Really, we are working here with the complexification of the Lie-alebra $\mathcal{L}_{\mathbb{C}} := \mathcal{L} + i\mathcal{L}$.

Exercise 9.21. *Derive Eqs.* (9.65), (9.67) *and* (9.68), *starting with the commutation relations* (9.63) *and the definition* (9.64).

From Schur's Lemma this means that J^2 must be a multiple of the unit matrix and we write [12]

$$J^2 = j(j+1)\mathbb{1} \,, \tag{9.69}$$

for a real number j. The number j characterises the representation we are considering and which we write as r_j with associated representation vector space V_j. The idea is to construct this vector space by thinking about the eigenvalues and eigenvectors of J_3 which we denote by m and $|jm\rangle$, respectively, so that

$$V_j = \text{Span}(|jm\rangle \,|\, |jm\rangle \text{ eigenvector with eigenvalue } m \text{ of } J_3) \,. \tag{9.70}$$

What can we say about these eigenvectors and eigenvalues? Since J_3 is hermitian we can choose the eigenvectors as an ortho-normal basis, so

$$\langle jm|j\tilde{m}\rangle = \delta_{m\tilde{m}} \,. \tag{9.71}$$

It is also clear that $0 \leq \langle jm|J^2|jm\rangle = j(j+1)$ so we can take $j \geq 0$. Further, the short calculation

$$J_3 J_\pm |jm\rangle = (J_\pm J_3 + [J_3, J_\pm])|jm\rangle = (m \pm 1) J_\pm |jm\rangle \tag{9.72}$$

shows that $J_\pm|jm\rangle$ is also an eigenvector of J_3 but with eigenvalue $m \pm 1$. Hence, $J_\pm$ act as "ladder operators" which increase or decrease the eigenvalue by one. If this is carried out repeatedly the process must break down at some point since we are looking for finite-dimensional representations. The breakdown point can be found by considering

$$0 \leq \langle jm|J_- J_+|jm\rangle = (j-m)(j+m+1)\langle jm|jm\rangle \tag{9.73}$$

$$0 \leq \langle jm|J_+ J_-|jm\rangle = (j+m)(j-m+1)\langle jm|jm\rangle \,, \tag{9.74}$$

where we have used that $J_\pm^\dagger = J_\mp$ (since the J_i are hermitian) to show that the LHS is positive and Eqs. (9.67) to obtain the RHS. The above equations tell us that $(j-m)(j+m+1) \geq 0$ and $(j+m)(j-m+1) \geq 0$ and this has two important implications:

$$-j \leq m \leq j \,, \qquad J_\pm|jm\rangle = 0 \quad \Leftrightarrow \quad m = \pm j \,. \tag{9.75}$$

This means we have succeeded in bounding the eigenvalues m and in finding a breakdown condition for the ladder operators. As we apply J_+ successively we can get as high as $|jj\rangle$ and then $J_+|jj\rangle = 0$. Starting with $|jj\rangle$ and using J_- to go in the opposite direction we obtain

$$|jj\rangle \,, \quad J_-|jj|\rangle \sim |jj-1\rangle \,, \quad J_-^2|jj\rangle \sim |jj-2\rangle, \cdots, J_-^{2j}|jj\rangle \sim |j-j\rangle \,, \quad J_-^{2j+1}|jj\rangle = 0 \,, \tag{9.76}$$

where the breakdown must arise at $m = -j$. In the above sequence, we have gone from $m = j$ to $m = -j$ in integer steps. This is only possible if j is integer or half-integer. The result of this somewhat lengthy argument can be summarised in

Theorem 9.22. *The irreducible representations r_j of $\mathcal{L}(SU(2))$ are labelled by an integer or half integer number $j \in \mathbb{Z}/2$ and the corresponding representation vector spaces V_j are spanned by*

$$V_j = \text{Span}(|jm\rangle \,|\, m = -j, -j+1, \ldots, j-1, j) \,, \tag{9.77}$$

so that $\dim(r_j) = 2j+1$. *The generators $J_\pm$, J_3 in Eq.* (9.64) *act on the states $|jm\rangle$ as*

$$J_\pm|jm\rangle = \sqrt{j(j+1) - m(m\pm 1)}\,|jm\pm 1\rangle \,, \qquad J_3|jm\rangle = m|jm\rangle \,. \tag{9.78}$$

[12] The reason for writing the constant as $j(j+1)$ will become clear shortly.

Proof. The only part we haven't shown yet is the factor in the first Eq. (9.78). To do this we write $J_\pm|jm\rangle = N_\pm(j,m)|jm\pm 1\rangle$ with some constants $N_\pm(j,m)$ to be determined. By multiplying with $\langle jm\pm 1|$ we get

$$N_\pm(j,m) = \langle jm\pm 1|J_\pm|jm\rangle = \frac{1}{N_\pm(j,m)^*}\langle jm|J_\mp J_\pm|jm\rangle \,, \tag{9.79}$$

and, using Eqs. (9.73) and (9.74), this implies $|N_\pm(j,m)|^2 = j(j+1) - m(m\pm 1)$. Up to a possible phase this is the result we need. It can be shown that this phase can be consistently set to one and this completes the proof. □

The representation r_j is also called the spin j representation and j is referred to as the total spin of the representation. The label m for the basis vectors $|jm\rangle$ is also called z-component of the spin. These representations play an important role in quantum mechanics where they describe states with well-defined total spin and well-defined z-component of spin.

The results (9.78) can be used to explicitly compute the representation matrices $T_\pm^{(j)}$ and $T_3^{(j)}$ whose entries are given by

$$\begin{aligned} T^{(j)}_{+,\tilde{m}m} &:= \langle j\tilde{m}|J_+|jm\rangle = \sqrt{j(j+1)-m(m+1)}\,\delta_{m,\tilde{m}-1} \\ T^{(j)}_{-,\tilde{m}m} &:= \langle j\tilde{m}|J_-|jm\rangle = \sqrt{j(j+1)-m(m-1)}\,\delta_{m,\tilde{m}+1} \\ T^{(j)}_{3,\tilde{m}m} &:= \langle j\tilde{m}|J_3|jm\rangle = m\,\delta_{m\tilde{m}} \end{aligned} \,. \tag{9.80}$$

Let us use these equations to compute these matrices explicitly for the lowest-dimensional representations.

9.4.3 Examples of $SU(2)$ representations

representation $j = 0$: The representation r_0 is one-dimensional with representation vector space $V_0 = \mathbb{C}$ and $m = 0$ is the only allowed value. The $1\times!$ representation matrices are obtained by inserting $j = m = 0$ into Eq. (9.80) which shows that

$$T_\pm^{(0)} = T_3^{(0)} = (0) \,. \tag{9.81}$$

From Eq. (9.58), the corresponding group representation matrices are $g(\mathbf{t}) = \exp(0) = 1$, so this is the trivial (= snglet) representation.

representation $j = 1/2$: The representation $r_{1/2}$ is two-dimensional with representation vector space $V_{1/2} = \mathbb{C}^2$ and allowed values $m = -1/2, 1/2$. Since this is the only irreducible two-dimensional representation it must coincide with the fundamental representation of $SU(2)$ which is also two-dimensional. We can verify this explicitly by inserting $j = 1/2$ and $\tilde{m}, m = 1/2, -1/2$ into Eq. (9.80) which leads to

$$T_+^{(1/2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \tau_1 + i\tau_2 \,, \quad T_-^{(1/2)} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \tau_1 - i\tau_2 \,, \quad T_3^{(1/2)} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \tau_3 \,. \tag{9.82}$$

Here, τ_i are the standard generators of the $SU(2)$ Lie-algebra, defined in Eq. (9.44) which confirms that we are indeed dealing with the fundamental representation.

representation $j = 1$: The representation r_1 is three-dimensional with representation vector space $V_1 = \mathbb{C}^3$ (or $\mathbb{R}^3$ since the representation matrices turn out to be real) with allowed values $m = -1, 0, 1$. There is only one irreducible three-dimensional representation so this must coincide with the three-dimensional representation of $SU(2)$ provided by $SO(3)$ which we have discussed earlier. Again, we can verify this

explicitly by inserting $j = 1$ and $\tilde{m}, m = 1, 0, -1$ into Eqs. (9.80), leading to

$$T_+^{(1)} = \sqrt{2}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} , \quad T_-^{(1)} = \sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} , \quad T_3^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{9.83}$$

Those matrices look somewhat different from the matrices (9.52) for the Lie-algebra of $SO(3)$ but the two sets of matrices are, in fact, related by a common basis transformation.

Exercise 9.23. *Show that the matrices $T_\pm^{(1)}$ and $T_3^{(1)}$ in Eq. (9.83) satisfy the correct $\mathcal{L}(SU(2))$ commutation relations. Also show that they are related by a common basis transformation to the matrices $\tilde{T}_\pm = \tilde{T}_1 \pm i\tilde{T}_2$ and $\tilde{T}_3$, with $\tilde{T}_i$ given in Eq. (9.52).*

Exercise 9.24. *Find the representation matrices $T_\pm^{(3/2)}$ annd $T_3^{(3/2)}$ for the $j = 3/2$ representation and show that they satisfy the correct $\mathcal{L}(SU(2))$ commutation relations.*

9.4.4 Relation to spherical harmonics

The spherical harmonics Y_{lm} which we have discussed in Section 6.4 are labelled in terms of the same combinatorial data as the representations r_l of $\mathcal{L}(SU(2))$ in Theorem 9.22. This can't be a coincidence. To understand the relation we should study how the rotation group can be represented on vector spaces of functions. For any rotation $A \in SO(3)$ and a (smooth) function $f : \mathbb{R}^3 \to \mathbb{R}$ define

$$R(A)(f)(\mathbf{x}) := f(A^{-1}\mathbf{x}) . \tag{9.84}$$

It is straightforward to check that R is a representation of $SO(3)$, that is, it satisfies $R(AB) = R(A)R(B)$, but note that including the inverse of A in the definition is crucial for this to work out.

Exercise 9.25. *Show that the $SO(3)$ representation R in Eq. (9.84) satisfies $R(AB) = R(A)R(B)$.*

To consider the associated Lie algebra we evaluate this for "small" rotations $A = \mathbb{1}_3 + it^j\tilde{T}_j + \cdot$, with the matrices $\tilde{T}_j$ defined in Eq. (9.52). Inserting this into Eq. (9.84) and performing a Taylor expansion leads, after a short calculation, to

$$R(A)(f)(\mathbf{x}) = f(\mathbf{x}) + it^j\hat{L}_j f(\mathbf{x}) + \cdots , \qquad \hat{\mathbf{L}} = -i\mathbf{x} \times \nabla . \tag{9.85}$$

The operators $\hat{L}_j$ span a representation of the Lie algebra of $SO(3)$ and, hence, must satisfy

$$[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k . \tag{9.86}$$

This can indeed be verified by direct calculation, using the definition of $\hat{\mathbf{L}}$ in Eq. (9.85).

Exercise 9.26. *Verify the commutation relations (9.86) by explicitly calculation, using the definition (9.85) of $\hat{\mathbf{L}}$.*

We have seen that the operators $\hat{L}_i$ generate small rotations on functions. In quantum mechanics, the $\hat{L}_i$ are the angular momentum operators, obtained from the classical angular momentum $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ by carrying out the replacement $p_i \to -i\frac{\partial}{\partial x_i}$. In conclusion, we see that there is a close connection between angular momentum and the rotation group.

Since the $\hat{L}_i$ form a representation of the Lie algebra of $SO(3)$ it is natural to ask which irreducible representations it contains. To do this it is useful to re-write the $\hat{L}_i$ in terms of spherical coordinates $\mathbf{x} = r(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, which leads to

$$\hat{L}_1 = i\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right) , \quad \hat{L}_2 = i\left(-\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right) , \quad \hat{L}_3 = -i\frac{\partial}{\partial\varphi} . \tag{9.87}$$

The Casimir operator $\hat{\mathbf{L}}^2$ is given by

$$\hat{\mathbf{L}}^2 = -\Delta_{S^2} \,, \tag{9.88}$$

where Δ_{S^2} is the Laplacian on the two-sphere, defined in Eq. (6.20).

Exercise 9.27. *Derive the expressions* (9.87) *and* (9.88) *for angular momentum in polar coordinates.*

It is now easy to verify, using Eqs (6.85) and (6.82) that the spherical harmonics satisfy

$$\hat{\mathbf{L}}^2 Y_{lm} = l(l+1) Y_{lm} \,, \qquad \hat{L}_3 Y_{lm} = m Y_{lm} \,. \tag{9.89}$$

This means that the vector space spanned by the Y_{lm} for $m = -l, \dots, l$ forms the representation space V_l of the representation r_l with total spin (or angular momentum) l. Note that this only leads to the representation r_j with integer j but not the ones with half integer j. This is directly related to the fact that we have started with $SO(3)$, rather than $SU(2)$.

The fact that the Y_l^m span a spin l representation means that

$$R(A) Y_l^m(\mathbf{n}) = Y_l^m(A^{-1}\mathbf{n}) = \sum_{m'=-l,\cdots,l} R^{(l)m}_{m'}(A)\, Y_l^{m'}(\mathbf{n}) \,, \tag{9.90}$$

where $R^{(l)}(A)$ are the spin l representation matrices (and $\mathbf{n}$ is a unit vector which parametrises S^2). An immediate conclusion from this formula is that the function

$$F(\mathbf{n}', \mathbf{n}) := \sum_{m=-l,\cdots,l} Y_l^m(\mathbf{n}')^* Y_l^m(\mathbf{n}) \tag{9.91}$$

is invariant under rotations, that is, $F(A\mathbf{n}', A\mathbf{n}) = F(\mathbf{n}', \mathbf{n})$ for a rotation A. This fact was used in the proof of Lemma (6.3).

9.4.5 Clebsch-Gordan decomposition

We would now like to discuss the tensor product of two $\mathcal{L}(SU(2))$ representations r_{j_1} and r_{j_2} and the corresponding Clensch-Gordan decomposition (9.27). In quantum mechanics this corresponds to what is referred to as "addition of spin". We start by introducing the representation matrices

$$J_i^{(1)} = r_{j_1}(\tau_i) \,, \qquad J_i^{(2)} = r_{j_2}(\tau_i) \,, \tag{9.92}$$

for the two representations r_{j_1} and r_{j_2}. If $r = r_{j_1} \otimes r_{j_2}$ is the tensor representation we would like to understand how its representation matrices $J_i := r(\tau_i)$ relate to $J_i^{(1)}$ and $J_i^{(2)}$ above. To this end, we write the corresponding infinitesimal group transformations

$$R_{j_1}(\mathbf{t}) = \mathbb{1} + i t^i J_i^{(1)} + \cdots \,, \qquad R_{j_2}(\mathbf{t}) = \mathbb{1} + i t^i J_i^{(2)} + \cdots \,, \tag{9.93}$$

and recall, from Eq. (9.24), the definition of the tensor product in term of the Kronecker product. This means

$$R(\mathbf{t}) = R_{j_1}(\mathbf{t}) \times R_{j_2}(\mathbf{t}) = \mathbb{1} + i t^i \left(J_i^{(1)} \times \mathbb{1} + \mathbb{1} \times J_i^{(2)} \right) + \cdots \,. \tag{9.94}$$

and comparing this with $R(\mathbf{t}) = \mathbb{1} + i t^i J_i + \cdots$ leads to

$$J_i = J_i^{(1)} \times \mathbb{1} + \mathbb{1} \times J_i^{(2)} \,. \tag{9.95}$$

This formula is the reason for referring to "addition of spin".

Exercise 9.28. *Show, if $J_i^{(1)}$ and $J_i^{(2)}$ each satisfy the commutation relations (9.63) then so does J_i, defined in Eq. (9.95) (Hint: Use the property (9.23) of the Kronecker product.)*

In summary, we now have the following representations, representation vector spaces and representation matrices:

representation	dimension	spanned by	range for m	representation matrices
r_{j_1}	$2j_1+1$	$\|j_1 m_1\rangle$	$m_1 = -j_1, \ldots, j_1$	$J_i^{(1)} = r_{j_1}(\tau_i)$
r_{j_2}	$2j_2+1$	$\|j_2 m_2\rangle$	$m_2 = -j_2, \ldots, j_2$	$J_i^{(2)} = r_{j_2}(\tau_i)$
$r = r_{j_1} \otimes r_{j_2}$	$(2j_1+1)(2j_2+1)$	$\|j_1 j_2 m_1 m_2\rangle = \|j_1 m_1\rangle \otimes \|j_2 m_2\rangle$	$m_1 = -j_1, \ldots, j_1$ $m_2 = -j_2, \ldots, j_2$	$J_i = J_i^{(1)} \times \mathbb{1} + \mathbb{1} \otimes J_i^{(2)}$

We note that

$$\begin{aligned} J_3 \,|j_1 j_2 m_1 m_2\rangle &= \left(J_3^{(1)} \times \mathbb{1} + \mathbb{1} \times J_3^{(2)}\right) |j_1 m_1\rangle \otimes |j_2 m_2\rangle \\ &= \left(J_3^{(1)} |j_1 m_1\rangle\right) \otimes |j_2 m_2\rangle + |j_1 m_1\rangle \otimes \left(J_3^{(2)} |j_2 m_2\rangle\right) = (m_1 + m_2)\, |j_1 j_2 m_1 m_2\rangle \end{aligned} \tag{9.96}$$

so the basis states $|j_1 j_2 m_1 m_2\rangle$ of the tensor representation are eigenvectors of J_3 with eigenvalues $m_1 + m_2$.

While r_{j_1} and r_{j_2} are irreducible there is no reason for r to be. However, given that we have a complete list of all irreducible representations from Theorem 9.22 we know that r must have a Clebsch-Gordan decomposition of the form

$$r = r_{j_1} \otimes r_{j_2} = \bigoplus_j \nu_j r_j \;, \tag{9.97}$$

where $\nu_j \in \mathbb{Z}^{\geq 0}$ indicates how many times r_j is contained in r. (If r_j is not contained in r then $\nu_j = 0$.) Our first problem is to determine the numbers ν_j and, hence, to work out the Clebsch-Gordan decomposition explicitly.

Theorem 9.29. *For two representations r_{j_1} and r_{j_2} of $\mathcal{L}(SU(2))$ we have*

$$r_{j_1} \otimes r_{j_2} = \bigoplus_{j=|j_1 - j_2|}^{j_1 + j_2} r_j \;. \tag{9.98}$$

Proof. Our starting point is to think about the degeneracy, δ_m, of the eigenvalue m of J_3 in the representation $r = r_{j_1} \otimes r_{j_2}$. Every representation $r_j \subset r$ with $j \geq |m|$ contributes exactly one to this degeneracy while representations r_j with $j < |m|$ do not contain a state with J_3 eigenvalue m. This implies

$$\delta_m = \sum_{j \geq |m|} \nu_j \;, \tag{9.99}$$

where ν_j counts how many time r_j is contained in r, as in Eq. (9.97). Eq. (9.99) implies

$$\nu_j = \delta_j - \delta_{j+1} \;, \tag{9.100}$$

so if we can work out the degeneracies δ_m this equation allows us to compute the desired numbers ν_j. The degeneracies δ_m are computed from the observation in Eq. (9.96) that the states with eigenvalue m are precisely those states $|j_1 j_2 m_1 m_2\rangle$ with $m = m_1 + m_2$. Hence, all we need to do is count the pairs (m_1, m_2), where $m_i = -j_i, \ldots, j_i$ and $m_1 + m_2 = m$. The result is

$$\delta_m = \begin{cases} 0 & \text{for} \quad |m| > j_1 + j_2 \\ j_1 + j_2 + 1 - |m| & \text{for} \quad j_1 + j_2 \geq |m| \geq |j_1 - j_2| \\ 2j_2 + 1 & \text{for} \quad |j_1 - j_2| \geq |m| \geq 0 \end{cases} \;. \tag{9.101}$$

Inserting these results into Eq. (9.100) shows that $\nu_j = 1$ for $j_1 + j_2 \geq j \geq |j_1 - j_2|$ and $\nu_j = 0$ otherwise. □

Eq. (9.98) tells us how to "couple" two spins. For example, two spin 1/2 representations

$$r_{1/2} \otimes r_{1/2} = r_0 \oplus r_1 \tag{9.102}$$

contain a singlet r_0 and a spin 1 representation r_1. Note that dimensions work out since $\dim(r_{1/2}) = 2$, $\dim(r_0) = 1$ and $\dim(r_1) = 3$. As another example, consider coupling a spin 1/2 and a spin 1 representation which leads to

$$r_{1/2} \otimes r_1 = r_{1/2} \oplus r_{3/2} \,. \tag{9.103}$$

We have now identified the representation content of a tensor product for two irreducible representations r_{j_1} and r_{j_2} but this gives rise to a more detailed problem. On the tensor representation vector space $V = V_{j_1} \otimes V_{j_2}$ we have two sets of basis vectors, namely

$$|j_1 j_2 m_1 m_2\rangle \quad \text{where} \quad m_1 = -j_1, \ldots, j_1 \,, \quad m_2 = -j_2, \ldots, j_2 \tag{9.104}$$

$$|jm\rangle \quad \text{where} \quad m = -j, \ldots, j \,, \quad j = |j_1 - j_2|, \ldots, j_1 + j_2 \,. \tag{9.105}$$

The question is how these two sets of basis vectors are related. Formally, we can write (using Dirac notation)

$$|jm\rangle = \sum_{m_1, m_2} |j_1 j_2 m_1 m_2\rangle \langle j_1 j_2 m_1 m_2 | jm\rangle \tag{9.106}$$

and the numbers $\langle j_1 j_2 m_1 m_2 | jm\rangle$ which appear in this equation are called *Clebsch-Gordan coefficients*. Once we have computed these coefficients the relation between the two sets of basis vectors is fixed.

Application 9.35. Example of a Clebsch-Gordan decomposition

How to carry out a Clebsch-Gordan decomposition is best explained with an example. Let us consider two spin 1/2 representation, as in Eq. (9.102), whose tensor produce space $V = V_{1/2} \otimes V_{1/2}$ is four-dimensional and spanned by

$$\begin{array}{rcl} |j_1 j_2 m_1 m_2\rangle & \to & |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2} - \tfrac{1}{2}\rangle \,,\; |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\tfrac{1}{2}\rangle \,,\; |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\rangle \,,\; |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\rangle \\ |jm\rangle & \to & |00\rangle \,,\; |1\,-1\rangle \,,\; |10\rangle \,,\; |11\rangle \end{array} \tag{9.107}$$

The key to relating these two sets of basis vectors is again to think about the eigenvalue m of $J_3 = J_3^{(1)} + J_3^{(2)}$. Consider the $m = 1$ state $|11\rangle$ from the second basis, The only state from the first basis with $m = 1$ (remembering that m_1 and m_2 sum up) is $|\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\rangle$. After a choice of phase, we can therefore set

$$|11\rangle \;=\; |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\rangle \,. \tag{9.108}$$

We can generate the other required relations by acting on Eq. (9.108) with $J_- = J_-^{(1)} + J_-^{(2)}$, using the formula (9.78). This leads to

$$\begin{array}{ccc} |11\rangle & = & |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2}\rangle \\ \downarrow & & \downarrow \\ \sqrt{2}\,|10\rangle & = & |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\tfrac{1}{2}\rangle + |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\rangle \\ \downarrow & & \downarrow \\ 2\,|1\,-1\rangle & = & |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2} - \tfrac{1}{2}\rangle + |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2} - \tfrac{1}{2}\rangle \\ \sqrt{2}\,|00\rangle & = & |\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\tfrac{1}{2}\rangle - |\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2} - \tfrac{1}{2}\rangle \end{array} \tag{9.109}$$

where the arrows indicate the action of J_- and the last relation follows from orthogonality. So in

summary we have

$$\left.\begin{array}{rcl} |11\rangle & = & |\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\rangle \\ 10\rangle & = & \frac{1}{\sqrt{2}}\left(|\frac{1}{2}\frac{1}{2}-\frac{1}{2}\frac{1}{2}\rangle+|\frac{1}{2}\frac{1}{2}\frac{1}{2}-\frac{1}{2}\rangle\right) \\ |1-1\rangle & = & |\frac{1}{2}\frac{1}{2}-\frac{1}{2}-\frac{1}{2}\rangle \end{array}\right\} \quad j=1 \tag{9.110}$$

$$\left.|00\rangle = \frac{1}{\sqrt{2}}\left(|\frac{1}{2}\frac{1}{2}-\frac{1}{2}\frac{1}{2}\rangle-|\frac{1}{2}\frac{1}{2}\frac{1}{2}-\frac{1}{2}\rangle\right)\right\} \quad j=0 \tag{9.111}$$

and this provides a complete set of relations between the two sets of basis vector from which all Clebsch-Gordan coefficients can be read off.

Exercise 9.30. *Find the Clebsch-Gordan coefficients for the case* (9.103).

9.5 The Lorentz group

The Lorentz group is of fundamental importance in physics. It is the group at the heart of special relativity and it is a symmetry of all relativistic theories, including relativistic mechanics and electro-magnetism. Mainly, we would like to determine all irreducible representations of the Lorentz group Lie algebra and, as we will see, this is rather easy given our earlier results for $SU(2)$.

9.5.1 Basic definition

The Lorentz group L is defined as

$$L = \left\{\Lambda \in \mathrm{Gl}(\mathbb{R}^4) \,|\, \Lambda^T\eta\Lambda = \eta\right\} \ , \tag{9.112}$$

where $\eta = \mathrm{diag}(-1,1,1,1)$ is the *Minkowski metric*. In other words, the Lorentz group consists of all real 4×4 matrices Λ which satisfy the defining relation $\Lambda^T\eta\Lambda = \eta$. Using index notation, this relation can also be written as

$$\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\eta_{\mu\nu} = \eta_{\rho\sigma} \ . \tag{9.113}$$

Clearly, the Lorentz group is a sub-group of the four-dimensional general linear group $\mathrm{Gl}(\mathbb{R}^4)$. As a matrix group it has a fundamental representation which is evidently four-dimensional. The action of this fundamental representation $\Lambda : \mathbb{R}^4 \to \mathbb{R}^4$ is explicitly given by

$$\mathbf{x} \mapsto \mathbf{x}' = \Lambda\mathbf{x} \quad \Longleftrightarrow \quad x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu x^\nu \ . \tag{9.114}$$

In Special Relativity this is interpreted as a transformation from one inertial system with space-time coordinates $\mathbf{x} = (t,x,y,z)^T$ to another one with space-time coordinates $\mathbf{x}' = (t',x',y',z')^T$.

9.5.2 Properties of the Lorentz group

Ultimately, we will be interested in the Lie algebra of the Lorentz group and its representations but we begin by discussing a few elementary properties of the group itself. Taking the determinant of the defining relation (9.112) and using standard properties of the determinant implies that $(\det(\Lambda))^2 = 1$ so that

$$\det(\Lambda) = \pm 1 \ . \tag{9.115}$$

Further, the $\rho = \sigma = 0$ component of Eq. (9.113) reads $-(\Lambda^0{}_0)^2 + \sum_{i=1}^3(\Lambda^i{}_0)^2 = -1$ so that

$$\Lambda^0{}_0 \geq 1 \quad \text{or} \quad \Lambda^0{}_0 \leq -1 \ . \tag{9.116}$$

Combining the two sign ambiguities in Eqs. (9.115) and (9.116) we see that there are four types of Lorentz transformations. The sign ambiguity in the determinant is analogous to what we have seen for orthogonal matrices and its interpretation is similar to the orthogonal case. Lorentz transformations with determinant 1 are called "proper" Lorentz transformations while Lorentz transformations with determinant -1 can be seen as a combination of a proper Lorentz transformation and a reflection. More specifically, consider the special Lorentz transformation $P = \text{diag}(1, -1, -1, -1)$ (note that this matrix indeed satisfies Eq. (9.113)) which is also referred to as "parity". Then every Lorentz transformation Λ can be written as

$$\Lambda = P\Lambda_+ \,, \tag{9.117}$$

where Λ_+ is a proper Lorentz transformation. The sign ambiguity (9.116) in ${\Lambda^0}_0$ is new but has an obvious physical interpretation. Under a Lorentz transformations Λ with ${\Lambda^0}_0 \geq 1$ the sign of the time component $x^0 = t$ of a vector $\mathbf{x}$ remains unchanged, so that the direction of time is unchanged. Correspondingly, such Lorentz transformation with positive ${\Lambda^0}_0$ are called "ortho-chronous". On the other hand, Lorentz transformations Λ with ${\Lambda^0}_0 \leq -1$ change the direction of time. If we introduce the special Lorentz transformation $T = \text{diag}(-1, 1, 1, 1)$, also referred to as "time reversal", then every Lorentz transformation Λ can be written as

$$\Lambda = T\Lambda^\uparrow \,, \tag{9.118}$$

where $\Lambda^\uparrow$ is an ortho-chronous Lorentz transformation. Introducing the sub-group $L_+^\uparrow$ of proper ortho-chronous Lorentz transformations, the above discussion shows that the full Lorentz group can be written as a union of four disjoint pieces:

$$L = (L_+^\uparrow) \cup (PL_+^\uparrow) \cup (TL_+^\uparrow) \cup (PTL_+^\uparrow) \,. \tag{9.119}$$

The Lorentz transformations normally used in Special Relativity are the proper, ortho-chronous Lorentz transformations. However, the other Lorentz transformations are relevant as well and it is an important question as to whether they constitute symmetries of nature in the same way that proper, ortho-chronous Lorentz transformations do. More to the point, the question is whether nature respects parity P and time-reversal T [13].

9.5.3 Examples of Lorentz transformations

What do proper, ortho-chronous Lorentz transformations look like explicitly? To answer this question we basically have to solve Eq. (9.112) which is clearly difficult to do in full generality. However, some special Lorentz transformations are more easily obtained. First, we note that for three-dimensional rotation matrices R the map

$$R \to \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{9.120}$$

leads to proper, ortho-chronous Lorentz transformations. Indeed, the matrices on the RHS satisfy Eq. (9.113) by virtue of $R^T R = \mathbb{1}_3$ and we have $\det(\Lambda) = \det(R) = 1$ and ${\Lambda^0}_0 = 1$. In group-theoretical terms this means there is an injective group homomorphism $SO(3) \to L_+^\uparrow$ defined by Eq. (9.120), or, in short, the three-dimensional rotations are embedded into the proper ortho-chromous Lorentz transformations.

To find less trivial examples we start with the Ansatz

$$\Lambda = \begin{pmatrix} \Lambda_2 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix} \,, \quad \Lambda_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{9.121}$$

[13] The answer is not simple and depends on the interactions/forces considered. But it is clear that the current fundamental theory which describes strong, electromagnetic and weak interactions - the standard model of particle physics - respects neither parity nor time inversion.

of a two-dimensional Lorentz transformation which affects time and the x-coordinate, but leaves y and z unchanged. Demanding that a Λ of the above form is a proper ortho-chronous Lorentz transformation leads to

$$\Lambda_2(\xi) = \begin{pmatrix} \cosh(\xi) & \sinh(\xi) \\ \sinh(\xi) & \cosh(\xi) \end{pmatrix} . \tag{9.122}$$

Exercise 9.31. *Show that Eq. (9.122) is the most general form for Λ_2 in order for Λ in Eq. (9.121) to be a proper ortho-chronous Lorenetz transformation.*

The quantity ξ in Eq. (9.122) is also called *rapidity*. It follows from the addition theorems for hyperbolic functions that $\Lambda(\xi_1)\Lambda(\xi_2) = \Lambda(\xi_1+\xi_2)$, so rapidities add up in the same way that two-dimensional rotation angles do. For a more common parametrisation introduce the parameter $\beta = \tanh(\xi) \in [-1, 1]$ so that

$$\cosh(\xi) = \frac{1}{\sqrt{1-\beta^2}} =: \gamma \, , \quad \sinh(\xi) = \beta\gamma \, . \tag{9.123}$$

In terms of β and γ the two-dimensional Lorentz transformations can then be written in the more familiar form

$$\Lambda_2 = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} . \tag{9.124}$$

Here, β is interpreted as the relative speed of the two inertial systems (in units of the speed of light).

9.5.4 The Lie algebra of the Lorentz group

To find the Lie algebra $\mathcal{L}(L)$ of the Lorentz group we proceed as before and write down an "infinitesimal" Lorentz transformation $\Lambda = \mathbb{1}_4 + iT + \cdots$, (where T is purely imaginary) insert this into the defining relation $\Lambda^T\eta\Lambda = \eta$ and work out the linear constraint on T. The result is

$$\mathcal{L}(L) = \{T \,|\, T = -\eta T^T \eta \, , \; T \text{ purely imaginary}\} = \text{Span}(T_i, S_i)_{i=1,2,3} \, , \tag{9.125}$$

where

$$T_i = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{T}_i \end{pmatrix} , \; S_1 = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \; S_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \; S_3 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} , \tag{9.126}$$

and $\tilde{T}_i$ are the $SO(3)$ generators defined in Eq. (9.52). Given that the three-dimensional rotations are embedded into the Lorentz group as in Eq. (9.120) the appearance of the $SO(3)$ generators in the lower 3×3 block is entirely expected. The other generators S_i do not correspond to rotations and are called the *boost generators*. They are related to non-trivial Lorentz boosts. Altogether, the dimension of the Lorentz group (Lie algebra) is six, with three parameters describing rotations and the three others Lorentz boosts. The commutation relations can be worked out by direct computation with the above matrices and they are given by

$$[T_i, T_j] = i\epsilon_{ijk}T_k \, , \qquad [S_i, S_j] = -i\epsilon_{ijk}T_k \, , \qquad [T_i, S_j] = i\epsilon_{ijk}S_k \, . \tag{9.127}$$

These commutation relations are very reminiscent of the ones for $SU(2)$ and we can make this more explicit by introducing a new basis

$$T_i^{\pm} := \frac{1}{2}(T_i \pm iS_i) \tag{9.128}$$

for the Lie algebra which commute as

$$[T_i^{\pm}, T_j^{\pm}] = i\epsilon_{ijk}T_k^{\pm} \, , \qquad [T_i^{+}, T_j^{-}] = 0 \, . \tag{9.129}$$

Exercise 9.32. *Verify the commutation relations* (9.127) *and* (9.129).

These commutation relations mean that the Lie algebra of the Lorentz group corresponds to two copies of the $SU(2)$ Lie algebra, so $\mathcal{L}(L) \cong \mathcal{L}(SU(2)) \oplus \mathcal{L}(SU(2))$. This is a rather lucky state of affairs (which does not persist for other groups, such as $SU(n)$ for $n > 2$) since this means we can obtain the irreducible representations of the Lorentz group from those of $SU(2)$. We have

Theorem 9.33. *The finite-dimensional, irreducible representations of the Lorentz group Lie algebra are classified by two "spins", (j_+, j_-), where $j_\pm \in \mathbb{Z}/2$, and the corresponding representations $r_{(j_+,j_-)}$ with representation vector space $V_{(j_+,j_-)}$ have dimension $(2j_+ + 1)(2j_- + 1)$. If $\tilde{J}_i^\pm = r_{j_\pm}(T_i^\pm)$ are the $\mathcal{L}(SU(2))$ representation matrices for $T_i^\pm$ then*

$$J_i^+ := r_{(j_+,j_-)}(T_i^+) = \tilde{J}_i^+ \times \mathbb{1}_{2j_-+1} \,, \qquad J_i^- := r_{(j_+,j_-)}(T_i^-) = \mathbb{1}_{2j_++1} \times \tilde{J}_i^- \tag{9.130}$$

are the representation matrices for $r_{(j_+,j_-)}$.

Exercise 9.34. *Verify that the representation matrices* (9.130) *satisfy the correct commutation relations for a representation of the Lorentz Lie algebra.*

9.5.5 Examples of Lorentz group representations

The lowest-dimensional representations of the Lorentz group algebra classify the types of particles we observe in nature.

$(j_+, j_-) = (0, 0)$ This is the trivial, singlet representation, which is one-dimensional, so $V_{(0,0)} \cong \mathbb{C}$. Fields which take values in $V_{(0,0)}$ (at every point in space-time) are also called *scalar fields*.

$(j_+, j_-) = (1/2, 0)$ This representation is two-dimensional, so $V_{(1/2,0)} \cong \mathbb{C}^2$. In physics, it is called the *left-handed Weyl spinor* representation.

$(j_+, j_-) = (0, 1/2)$ This representation is two-dimensional, so $V_{(0,1/2)} \cong \mathbb{C}^2$. In physics, it is called the *right-handed Weyl spinor* representation.

$(1/2, 0) \oplus (0, 1/2)$ This is the direct sum of the previous two representations. It is, therefore, reducible and four-dimensional. In physics, it is called the *Dirac spinor* representation.

$(j_+, j_-) = (1/2, 1/2)$ This representation is four-dimensional and, in fact, corresponds to the four-dimensional fundamental representation of the Lorentz group. In physics, the fields which take values in $V_{(1/2,1/2)}$ are called *vector fields*.

Appendices

A Calculus in multiple variables - a sketch

Calculus of multiple variables provides the background for much of what we are doing in this course (and calculus is the "engine" of differential geometry which we discuss in the following appendices) so it may be worth including a brief account of the subject. You have seen some of this in your first year, although perhaps not in quite the same way.

A.1 The main players

Our main arena is $\mathbb{R}^n$ which we think of as a normed vector space (in the sense of Def. 1.6) with the usual Euklidean norm. Based on this norm we can introduce a topology on $\mathbb{R}^n$, with the open sets defined as in Def. 1.19, that is, open sets $U \subset \mathbb{R}^n$ are those for which every point $\mathbf{x} \in U$ has a (sufficiently small) ball, centred on $\mathbf{x}$, which is entirely contained in U. Why is introducing open sets important for calculus? We can ask which subsets of $\mathbb{R}^n$ are appropriate as domains of functions, if we want to perform standard operations of calculus such as taking derivatives. As you know from the one-dimensional case, defining the derivative of a function at a point involves taking the limes of a difference quotient and setting this up assumes the function is defined near the point in question. Hence, if we want to define derivates we should focus on function domains where each point has a neighbourhood contained within the domain - but these are precisely the open sets.

Having said this, we would like to consider functions $f : U \to \mathbb{R}^m$, with domains U that are open subsets of $\mathbb{R}^n$ and co-domains that are subsets of $\mathbb{R}^m$. We will typically denote the coordinates on the domain $U \subset \mathbb{R}^n$ by $\mathbf{x} = (x_1, \ldots, x_n)^T$ and the coordinates on the co-domain $\mathbb{R}^m$ by $\mathbf{y} = (y_1, \ldots, y_m)^T$. So, in more down-to-earth language, our function f depends on n variables $x_1, \ldots, x_n$ and is vector-valued in m dimensions. More explicitly, the value of f at $\mathbf{x} \in U$ can be written down as

$$f(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix} , \tag{A.1}$$

so we can think of f is being given by m real-valued functions $f = (f_1, \ldots, f_m)^T$, each of which depends on the n variables $x_1, \ldots, x_n$. There are two special choices of dimensions which are of particular interest. If $m = 1$ we call f a *real-valued* function, or *scalar field* and if $n = m$ we call f a *vector field*. Vector fields will also be denoted by uppercase letters $A, B, \ldots$, adopting the notation more common in physics.

Definition A.1. *(Continuity) A function $f : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$ open, is said to be continuous at $\mathbf{x} \in U$ if every sequence $(\mathbf{x}_k)$ in U with $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}$ satisfies $\lim_{k\to\infty} f(\mathbf{x}_k) = f(\mathbf{x})$. The function f is said to be continuous on U if it is continuous for all $\mathbf{x} \in U$.*

Note this definition of continuity is extremely natural. It says continuous functions are those for which limites can be pulled in and out of the function argument, that is, $f(\lim_{k\to\infty} \mathbf{x}_k) = \lim_{k\to\infty} f(\mathbf{x}_k)$.

Application 1.36. Heaviside function and discontinuity

The aim is to show, using Def. A.1, that the Heaviside function $\theta : \mathbb{R} \to \mathbb{R}$ defined in Eq. (7.21) is not continuous at $x = 0$. This can be done by finding a sequence converging to 0 which violates the condition stated in Def. A.1.

Consider, for example, the sequence (x_k) defined by $x_k = -1/k$. It is clear that $\lim_{k\to\infty} x_k = 0$. On

the other hand, $\theta(x_k) = 0$ for all k since $x_k < 0$. Hence, $0 = \lim_{k\to\infty} \theta(x_k) \neq \theta(\lim_{k\to\infty} x_k) = \theta(0) = 1$ and this indeed violates the condition in Def. A.1. Hence, the Heaviside function is not continuous at $x = 0$.

Exercise A.1. *The function $f(x) = 1/x$ is not well-defined at $x = 0$ and should hence be seen as a function $f : \mathbb{R}\setminus\{0\} \to \mathbb{R}$. Suppose we construct a "completion" of f by defining a new function $F : \mathbb{R} \to \mathbb{R}$ with values $F(x) = 1/x$ for $x \neq 0$ and $F(0) = x_0$ for $x_0 \in \mathbb{R}$. Show, using Def. A.1, that there is no choice for x_0 such that F is continuous at $x = 0$.*

A.2 Partial differentiation

We consider functions $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, and coordinates $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ as before and also introduce the standard unit vectors $\mathbf{e}_i$, where $i = 1, \ldots, n$, on $\mathbb{R}^n$. The partial derivatives of f are defined as follows.

Definition A.2. *The function $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is said to be partially differentiable with respect to x_i at $\mathbf{x} \in U$ if the limes*

$$\lim_{\epsilon\to 0} \frac{f(\mathbf{x} + \epsilon\mathbf{e}_i) - f(\mathbf{x})}{\epsilon} \tag{A.2}$$

exists. In this case, the limes (A.2) is called the partial derivative of f with respect to x_i at $\mathbf{x}$ and it is denoted by $\frac{\partial f}{\partial x_i}(\mathbf{x})$ or by $\partial_i f(\mathbf{x})$.

As a further piece of common terminology, a function f is called κ times *continuously differentiable* if it is κ times (partially) differentiable and the resulting derivatives of order κ are continuous. Note that Def. A.2 is very much in analogy with the one-dimensional case, except that the difference quotient is now taken in the direction of one of the coordinate axes, as indicated by the presence of the unit vector $\mathbf{e}_i$. For this reason, all the standard rules of differentiation in one dimension - linearity, product rule, chain rule, quotient rule - directly generalise to partial derivatives (treating the coordinates x_j which are not involved as if they were constants). However, there is one feature of partial derivatives which has no analogy in one dimension - that is, whether partial derivatives in different directions commute. Fortunately, the answer is "yes" as stated in the following proposition.

Proposition A.1. *If $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is continuously partially differentiable twice at $\mathbf{x} \in U$ (in all directions) then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) \,. \tag{A.3}$$

Proof. A standard proof working with limites which can, for example, be found in Ref. [10]. □

There are certain (linear) combinations of partial derivatives which are of particular interest. In general, for a function $f = (f_1, \ldots, f_m)^T : U \to \mathbb{R}^m$, where $U \in \mathbb{R}^n$, with n variables and m-dimensional values, we have $n\,m$ partial derivatives, $\partial_i f_j$, which can be arranged into an $m \times n$ matrix. This matrix is called the *Jacobi matrix* and it will be discussed in detail below. For now we focus on some special cases.

A.2.1 Gradient

Suppose we have a real-valued function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$. Then all we have is n partial derivates $\partial_i f$ and they can be conveniently arranged into a row-vector

$$\text{grad}\, f(\mathbf{x}) = \nabla f(\mathbf{x}) := (\partial_1 f, \ldots, \partial_n f)\,(\mathbf{x}) \qquad \Rightarrow \qquad (\nabla f(\mathbf{x}))_i = \partial_i f(\mathbf{x}) \,, \tag{A.4}$$

which is called the *gradient* of f at $\mathbf{x}$. Note that for two partially differentiable real-valued functions $f, g : U \to \mathbb{R}$ the gradient satisfies the product rule (leaving out the argument $\mathbf{x}$ for simplicity)

$$\nabla(fg) = g\nabla f + f\nabla g \; . \tag{A.5}$$

This follows easily from the product rule for partial derivates:

$$(\nabla(fg))_i = \partial_i(fg) = g\partial_i f + f\partial_i g = (g\nabla f + f\nabla g)_i \; . \tag{A.6}$$

It is often useful to think of the gradient as a formal row-vector

$$\nabla = (\partial_1, \ldots, \partial_n) \qquad \Rightarrow \qquad \nabla_i = \partial_i \tag{A.7}$$

whose components are the partial derivatives.

A.2.2 Divergence

Now let us consider a vector field $A = (A_1, \ldots, A_n)^T : U \to \mathbb{R}^n$, where $U \subset \mathbb{R}^n$. Its partial derivatives $\partial_i A_j$ can be arranged into an $n \times n$ matrix (the Jacobi matrix of A). Of course we can consider any particular linear combination $\sum_{i,j} m_{ij}\partial_i A_j$, where $m_{ij} \in \mathbb{R}$, of these partial derivates, but is there a specific choice of the $n \times n$ matrix $m = (m_{ij})$ which is singled out in some way? Suppose we would like the combination of partial derivatives to be invariant under rotations $R \in \mathrm{SO}(n)$, acting as $x_i \mapsto R_{ik}x_k$ (which implies, via the chain rule, that $\partial_i \mapsto R_{ik}\partial_k$) and $A_j \mapsto R_{jl}A_l$. Then we have $m_{ij}\partial_i A_j \mapsto R_{ik}m_{ij}R_{jl}\partial_k A_l$ and, hence, the combination is invariant iff $R^T m R = m$. The only matrices satisfying this condition for all $R \in \mathrm{SO}(n)$ are multiples of the unit matrix $\mathbb{1}_n$, so that m_{ij} is proportional to δ_{ij}. This leads to the *divergence* of the vector field A defined by

$$\mathrm{div}A(\mathbf{x}) = \nabla \cdot A(\mathbf{x}) := \sum_{i=1}^{n} \partial_i A_i(\mathbf{x}) \; . \tag{A.8}$$

Note that the divergence can be written as a formal dot product between the nabla operator (A.7) and the vector field A, so div $= \nabla\cdot$. The divergence also satisfies a product rule which involves a vector field $A : U \to \mathbb{R}^n$ and a real-valued function $f : U \to \mathbb{R}$ and is given by

$$\nabla \cdot (fA) = A \cdot \nabla f + f\nabla \cdot A \; . \tag{A.9}$$

As before, its proof relies on the product rule for partial derivatives:

$$\nabla \cdot (fA) = \partial_i(fA)_i = \partial_i(fA_i) = A_i\partial_i f + f\partial_i A_i = A \cdot \nabla f + f\nabla \cdot A \; . \tag{A.10}$$

A.2.3 Curl

We have just seen that the divergence is (up to an overall factor) the only rotationally invariant linear combination of the partial derivatives $\partial_i A_j$ of a vector field $A : U \to \mathbb{R}^n$. What if we allow tensors with more than two indices (rather than just δ_{ij}) to be summed into $\partial_i A_j$? The only other special tensor (which is rotationally invariant just as δ_{ij}) is the n-dimensional Levi-Civita tensor $\epsilon_{i_1,\ldots,i_n}$. This means it is of interest to look at the following linear combinations

$$\epsilon_{k_1,\ldots,k_{n-2}ij}\partial_i A_j \tag{A.11}$$

of the partial derivatives of A, leading to an object with $n-2$ indices (which is also called a *tensor field*). This is an entirely sensible construction (which finds its natural home in the context of differential forms, see Appendix C) but only in three dimensions, $n = 3$, does it lead to a familiar object with one index,

that is, a vector field. A frequent question is why the curl is only defined in three dimensions. In fact, it is, in the sense of Eq. (A.11), defined in all dimensions (and differential forms provide a more natural framework for this) but only in three dimensions does it lead back to a vector field. This is why the curl is normally only introduced in three dimensions, and this is the case we focus on now. The curl of a vector field $A : U \to \mathbb{R}^3$ is another vector field with components (A.11), that is,

$$(\operatorname{curl} A)_i := \epsilon_{ijk}\partial_j A_k \qquad \Rightarrow \qquad \operatorname{curl} A = \nabla \times A \,. \tag{A.12}$$

Note that the curl can be expressed as a formal cross product with the nabla operator, so curl $= \nabla\times$.

Exercise A.2. *Show that for any $n \times n$ matrix R the Levi-Civita tensor satisfies*

$$\sum_{j_1,\ldots,j_n} R_{i_1 j_1} \cdots R_{i_n j_j} \epsilon_{j_1,\ldots,j_n} = \det(R)\epsilon_{i_1,\ldots,i_n} \,.$$

Deduce that the Levi-Civita tensor is invariant under rotations.

There are some further product rules for differentiable vector fields $A, B : U \to \mathbb{R}^3$ and functions $f : U \to \mathbb{R}^3$, namely

$$\nabla \times (fA) = \nabla f \times A + f \nabla \times A \tag{A.13}$$

$$\nabla \cdot (A \times B) = (\nabla \times A) \cdot B + A \cdot (\nabla \times B) \,. \tag{A.14}$$

Exercise A.3. *Prove the product rules (A.13) and (A.14) using index notation and the product rule for partial derivatives.*

While we have motivated the definitions of grad, div and curl it is fair to say that these motivations might not seem entirely compelling. Convincing mathematical reasons for introducing these operations can be given in the context of differential form, where they arise as special cases of the exterior derivative d. This will be discussed in Appendix C.

The gradient, the divergence and the curl relate in an interesting way and to discuss this we specialise to functions f and vector fields A in three dimensions (since this is the only dimension for which we have defined the curl - the general case will be dealt with in the context of differential forms later). We have

$$(\nabla \times (\nabla f))_k = \epsilon_{kij}\partial_i\partial_j f = 0 \tag{A.15}$$

$$\nabla \cdot (\nabla \times A) = \epsilon_{ijk}\partial_i\partial_j A_k = 0 \,, \tag{A.16}$$

where the vanishing follows from the commutativity of second derivatives, Proposition A.1, (which implies that $\partial_i\partial_j$ is symmetric in (ij)) and the anti-symmetry of the Levi-Civita tensor in (ij). There is an interesting and concise way to summarise the relationship between grad, curl and div. First let us introduce the set $\mathcal{C}^\infty(U)$ of infinitely many times partially differentiable functions and the set $\mathcal{V}(U)$ of infinitely many times partially differentiable vector fields on $U \subset \mathbb{R}^3$. Then, grad, curl and div map in the following way

$$0 \longrightarrow \mathcal{C}^\infty(U) \xrightarrow{\text{grad}=\nabla} \mathcal{V}(U) \xrightarrow{\text{curl}=\nabla\times} \mathcal{V}(U) \xrightarrow{\text{div}=\nabla\cdot} \mathcal{C}^\infty(U) \longrightarrow 0 \tag{A.17}$$

and Eqs. (A.15) and (A.16) imply that neighbouring maps in this chain compose to zero. Such a mathematical structure, that is a sequence

$$\cdots \xrightarrow{d_{i-2}} V_{i-1} \xrightarrow{d_{i-1}} V_i \xrightarrow{d_i} V_{i+1} \xrightarrow{d_{i+1}} \cdots \,. \tag{A.18}$$

of vector spaces V_i with maps d_i and neighbouring maps composing to zero, $d_i \circ d_{i-1} = 0$, is called a *complex*. The relation $d_i \circ d_{i-1} = 0$ tells us that $\operatorname{Im}(d_{i-1}) \subset \operatorname{Ker}(d_i)$ and this observation allows us to define the *cohomology* of the complex (A.18) which consists of the vector spaces $H^i = \operatorname{Ker}(d_i)/\operatorname{Im}(d_{i-1})$. The complex (A.17) with its associated cohomology is extremely important. It encodes topological information about the space U and it represents one of the gateways into an area of mathematics called *algebraic geometry*. Pursuing this further is well beyond our present scope.

A.3 The 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*.

To introduce the total derivative in a concise way it is useful to recall briefly the one-dimensional case of a function $f : \mathbb{R} \to \mathbb{R}$. Such a function is called differentiable at $x \in \mathbb{R}$ with derivative $f'(x)$ if the limes

$$f'(x) := \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} \tag{A.19}$$

exists. Alternatively and equivalently, we can say that f is differentiable at $x \in \mathbb{R}$ with derivative $f'(x)$ iff

$$f(x+\epsilon) = f(x) + f'(x)\epsilon + \mathcal{O}(\epsilon^2) \quad \text{where} \quad \lim_{\epsilon \to 0} \frac{\mathcal{O}(\epsilon^2)}{\epsilon} = 0 \tag{A.20}$$

is satisfied for all sufficiently small ϵ. This says we should think of the derivative as a linear map which provides the leading behaviour of the function's variation, away from $f(x)$. Note that, intuitively, $\mathcal{O}(\epsilon^2)$ denotes any expression which "goes" like ϵ^2 (or an even higher power of ϵ) but it is also properly defined as an expression which satisfies the limes condition in the right-hand side of Eq. (A.20). More generally, we say an expression is of order ϵ^κ, where $\kappa \in \mathbb{N}$, and we write this expression as $\mathcal{O}(\epsilon^\kappa)$ if

$$\lim_{\epsilon \to 0} \frac{\mathcal{O}(\epsilon^\kappa)}{\epsilon^{\kappa-1}} = 0 \,. \tag{A.21}$$

Using this notation often leads to a good match between mathematical precision and intuition.

Now we want to generalise the one-dimensional definition of derivatives, in the alternative form (A.20), to introduce the total derivative in the multi-dimensional case.

Definition A.3. *A function $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is called totally differentiable at $\mathbf{x} \in U$ if there exists an $m \times n$ matrix A such that*

$$f(\mathbf{x}+\boldsymbol{\epsilon}) = f(\mathbf{x}) + A\boldsymbol{\epsilon} + \mathcal{O}(|\boldsymbol{\epsilon}|^2) \tag{A.22}$$

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

This is where linear algebra meets calculus. In the one-dimensional case (A.20) we had a somewhat trivial linear map $f'(x) : \mathbb{R} \to \mathbb{R}$ (a 1×1 matrix) but in the multi-dimensional case, for a function f with n arguments and m-dimensional values, this becomes a linear map $Df(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$, that is, an $m \times n$ matrix.

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 A.2. *If $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is totally differentiable at $\mathbf{x} \in U$, then it is continuous at $\mathbf{x}$.*

Proof. We have to show the continuity property in Def. A.1 so we start with a sequence $(\mathbf{x}_k)$ which converges to $\mathbf{x}$. This sequence can also be written as $\mathbf{x}_k = \mathbf{x} + \boldsymbol{\epsilon}_k$, where $\lim_{k\to\infty} \boldsymbol{\epsilon}_k = \mathbf{0}$. Total differentiability at $\mathbf{x}$ implies that $f(\mathbf{x}_k) = f(\mathbf{x}+\boldsymbol{\epsilon}_k) = f(\mathbf{x}) + A\boldsymbol{\epsilon}_k + \mathcal{O}(|\boldsymbol{\epsilon}_k|^2)$ and taking the limes of this equation gives

$$\lim_{k\to\infty} f(\mathbf{x}_k) = \lim_{k\to\infty} \big(f(\mathbf{x}) + A\boldsymbol{\epsilon}_k + \mathcal{O}(|\boldsymbol{\epsilon}_k|^2)\big) = f(\mathbf{x})$$

and this shows f is continuous at $\mathbf{x}$. □

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

Proposition A.3. *If $f : U \to \mathbb{R}^m$, with $U \subset \mathbb{R}^n$ open, is totally differentiable at $\mathbf{x} \in U$ then it is partially differentiable with respect to all variables and the total derivative is given by*

$$Df(\mathbf{x}) = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & \ddots & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix} (\mathbf{x}) \tag{A.23}$$

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

$$f_i(\mathbf{x} + \epsilon \mathbf{e}_j) = f_i(\mathbf{x}) + \epsilon A_{ij} + \mathcal{O}(\epsilon^2) \, . \tag{A.24}$$

This result allows us to calculate the partial derivatives as

$$\partial_j f_i(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{f_i(\mathbf{x}+\epsilon\mathbf{e}_j) - f_i(\mathbf{x})}{\epsilon} \overset{(A.24)}{=} \lim_{\epsilon \to 0} \left(A_{ij} + \frac{\mathcal{O}(\epsilon^2)}{\epsilon} \right) = A_{ij} \, , \tag{A.25}$$

and this is the required statement. □

In short, the Jacobi matrix is the matrix which contains all the partial derivates. Note from Eq. (A.23) that it is organised such that every row of Df is the gradient of one of the component functions f_i of f. This means, the Jacobi matrix can also be written as

$$Df(\mathbf{x}) = \begin{pmatrix} \nabla f_1 \\ \vdots \\ \nabla f_m \end{pmatrix} (\mathbf{x}) \, . \tag{A.26}$$

In particular, for a real-valued function $f : U \to \mathbb{R}$ the Jacobi matrix is simply the gradient, so $Df = \nabla f$. 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 [14]

$$Df(\mathbf{x}) = \frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}(\mathbf{x}) \, . \tag{A.27}$$

A.4 The 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.

[14] 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. (A.27) the Jacobi matrix while the Jacobian is written as $\det \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(\mathbf{x})$.

Theorem A.4. *(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*

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

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

$$D(f \circ g)(\mathbf{x}) = Df(\mathbf{y})\, Dg(\mathbf{x}) \,. \tag{A.28}$$

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

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

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

$$\begin{aligned}(f \circ g)(\mathbf{x}+\boldsymbol{\epsilon}) &= f(g(\mathbf{x}+\boldsymbol{\epsilon})) = f(g(\mathbf{x}) + \boldsymbol{\eta}) = f(g(\mathbf{x})) + B\boldsymbol{\eta} + \mathcal{O}(|\boldsymbol{\eta}|^2) \\ &= f(g(\mathbf{x})) + BA\boldsymbol{\epsilon} + \mathcal{O}(|\boldsymbol{\eta}|^2, |\boldsymbol{\epsilon}|^2) \,. \end{aligned} \tag{A.29}$$

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

In short, the Jacobi matrix of the composite function is the matrix product of the individual Jabobi matrices, as in Eq. (A.28). 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 $y_i(\mathbf{x}) = g_i(\mathbf{x})$. Then, using the notation (A.27) for the Jacobi matrix, the chain rule can be stated as

$$\frac{\partial((f\circ y)_1, \ldots, (f\circ y)_p)}{\partial(x_1, \ldots, x_n)}(\mathbf{x}) = \frac{\partial(f_1, \ldots, f_p)}{\partial(y_1, \ldots, y_m)}(\mathbf{y})\, \frac{\partial(y_1, \ldots, y_m)}{\partial(x_1, \ldots, x_n)}(\mathbf{x}) \,, \tag{A.30}$$

$$p \times n \qquad\qquad p \times m \qquad\qquad m \times n \tag{A.31}$$

where the size of the Jacobi matrices is indicated underneath. Alternatively, we can write the matrix product on the right-hand side of Eq. (A.30) out with indices, keeping in mind that a matrix product translates into a sum over the adjacent indices. Then (omitting the points $\mathbf{x}$ and $\mathbf{y}$ for simplicity) we get

$$\frac{\partial (f\circ y)_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial f_i}{\partial y_k}\, \frac{\partial y_k}{\partial x_j} \,, \tag{A.32}$$

perhaps the version of the chain rule you are most familiar with.

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 (A.28) and assuming that $f = g^{-1}$ exists and is totally differentiable. On the left-hand side of Eq. (A.28) we have the map $f \circ g = g^{-1} \circ g = \mathrm{id}$, that is the identify map, with $\mathrm{id}(\mathbf{x}) = \mathbf{x}$. 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. (A.28) turns into

$$\mathbb{1} = D(g^{-1})(\mathbf{y})\, Dg(\mathbf{x}) \,. \tag{A.33}$$

where $\mathbf{y} = g(\mathbf{x})$ and we have shown

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

$$D(g^{-1})(\mathbf{y}) = (Dg(\mathbf{x}))^{-1} \ . \tag{A.34}$$

Proof. This follows directly from Eq. (A.33). □

Note that the inverse on the right-hand side of Eq. (A.34) 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 $y_i(\mathbf{x}) = g_i(\mathbf{x})$, as before, then, using the notation (A.27) for the Jacobi matrix, the inverse derivative rule (A.34) can be stated as

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)}(\mathbf{y}) = \left(\frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}(\mathbf{x}) \right)^{-1} \ . \tag{A.35}$$

Application 1.37. Jacobian of the inverse function

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 $g_2(x_1, x_2) = x_1 x_2$. Writing $y_1 = g_1$ and $y_2 = g_2$ as before, the Jacobi matrix of g is given by

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_1 \end{pmatrix} \ . \tag{A.36}$$

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 (A.35) instead, which gives

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \left(\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} \right)^{-1} = \frac{1}{x_1^2 - x_2^2} \begin{pmatrix} x_1 & -x_2 \\ -x_2 & x_1 \end{pmatrix} \ . \tag{A.37}$$

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 (A.36) 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 A.2 below.

A.5 Taylor series and extremal points

We would like to write down the Taylor series for a (suitably many times continuously differentiable) real-valued function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open, with coordinates $\mathbf{x} = (x_1, \ldots, x_n)^T$. 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

$$\frac{\mathbf{x}^K}{K!} = \frac{x_1^{k_1}}{k_1!} \frac{x_2^{k_2}}{k_2!} \cdots \frac{x_n^{k_n}}{k_n!} \ . \tag{A.38}$$

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 $\partial^K := \partial_1^{k_1} \partial_2^{k_2} \cdots \partial_n^{k_n}$.

A.5.1 Taylor's formula

Theorem A.5. *(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 $\mathbf{x} + t\boldsymbol{\xi} \in U$ for all $t \in [0,1]$ there exists an $\epsilon \in [0,1]$ such that*

$$f(\mathbf{x}+\boldsymbol{\xi}) = \sum_{|K|\leq\kappa} \frac{\partial^K f(\mathbf{x})}{K!}\boldsymbol{\xi}^K + \sum_{|K|=\kappa} \frac{\partial^K f(\mathbf{x}+\epsilon\boldsymbol{\xi}) - \partial^K f(\mathbf{x})}{K!}\boldsymbol{\xi}^K \,. \tag{A.39}$$

Proof. The proof is not too difficult but somewhat elaborate, and we refer to the literature, for example Ref. [10], 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

$$\sum_{K} \frac{\partial^K f(\mathbf{x})}{K!}\boldsymbol{\xi}^K \tag{A.40}$$

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 (A.39) and its behaviour with the order κ, for a given function f, can often be used to decide for which values of $\boldsymbol{\xi}$ the series (A.40) converges to $f(\mathbf{x}+\boldsymbol{\xi})$.

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 $\lim_{\boldsymbol{\xi}\to\mathbf{0}}(\partial^K f(\mathbf{x}+\epsilon\boldsymbol{\xi}) - \partial^K f(\mathbf{x})) = 0$. From Eq. (A.21), this means that the second sum in Eq. (A.39) is of $\mathcal{O}(|\boldsymbol{\xi}|^{\kappa+1})$ and Taylor's formula becomes

$$f(\mathbf{x}+\boldsymbol{\xi}) = \sum_{|K|\leq\kappa} \frac{\partial^K f(\mathbf{x})}{K!}\boldsymbol{\xi}^K + \mathcal{O}(|\boldsymbol{\xi}|^{\kappa+1}) \,. \tag{A.41}$$

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

$$f(\mathbf{x}+\boldsymbol{\xi}) = f(\mathbf{x}) + \sum_{i=1}^{n} \partial_i f(\mathbf{x})\,\xi_i + \frac{1}{2}\sum_{i,j=1}^{n} \partial_i\partial_j f(\mathbf{x})\,\xi_i\xi_j + \mathcal{O}(|\boldsymbol{\xi}|^3) \tag{A.42}$$

$$= f(\mathbf{x}) + \nabla f(\mathbf{x})\cdot\boldsymbol{\xi} + \frac{1}{2}\boldsymbol{\xi}^T H(\mathbf{x})\,\boldsymbol{\xi} + \mathcal{O}(|\boldsymbol{\xi}|^3) \,, \tag{A.43}$$

where the symmetric $n \times n$ matrix $H(\mathbf{x})$ with entries

$$(H(\mathbf{x}))_{ij} = \partial_i\partial_j f(\mathbf{x}) \tag{A.44}$$

is called the *Hesse matrix* of f at $\mathbf{x}$.

A.5.2 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 A.4. *Consider a real-valued function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open. A point $\mathbf{x} \in U$ is called a local minimum (a local maximum) of f if there is a sufficiently small ball $B_\epsilon(\mathbf{x}) \subset U$ such that $f(\mathbf{z}) \geq f(\mathbf{x})$ for all $\mathbf{z} \in B_\epsilon(\mathbf{x})$ ($f(\mathbf{z}) \leq f(\mathbf{x})$ for all $\mathbf{z} \in B_\epsilon(\mathbf{x})$). A local extremum is either a local minimum or a local maximum. If the equality in the above conditions is only realised for $\mathbf{z} = \mathbf{x}$ then the local minimum, maximum or extremum is called isolated.*

Local extrema have an interesting property.

Proposition A.4. *Consider a real-valued continuously differentiable function $f : U \to \mathbb{R}$, where $U \subset \mathbb{R}^n$ open, with a local extremum at $\mathbf{x} \in U$. Then $\nabla f(\mathbf{x}) = 0$, that is, $\mathbf{x}$ 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 $g_i(t) := f(\mathbf{x} + t\mathbf{e}_i)$. Clearly, if f has a local extremum at $\mathbf{x}$ 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, $g_i'(0) = \partial_i f(\mathbf{x})$ (for all i) and this proves the claim. □

Hence, the local extrema of a function f are to be found among its stationary points, that is, among the points $\mathbf{x}$ in its domain for which $\nabla f(\mathbf{x}) = 0$. 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 $\mathbf{x}$ of f Taylor's formula (A.43) becomes

$$f(\mathbf{x} + \boldsymbol{\xi}) - f(\mathbf{x}) = \frac{1}{2}\boldsymbol{\xi}^T H(\mathbf{x})\, \boldsymbol{\xi} + \mathcal{O}(|\boldsymbol{\xi}|^3)\,, \tag{A.45}$$

so the leading behaviour near $\mathbf{x}$ 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 form linear algebra. Such a matrix M is called *positive definite* if $\boldsymbol{\xi}^T M \boldsymbol{\xi} > 0$ for all $\boldsymbol{\xi} \neq \mathbf{0}$. It is called *negative definite* if $-M$ is positive definite. Further, M is called *indefinite* if $\boldsymbol{\xi}^T M \boldsymbol{\xi}$ takes on strictly positive and strictly negative values, for suitable $\boldsymbol{\xi}$. 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 A.6. *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 $\mathbf{x} \in U$. Then we have the following statements.*
(i) If $H(\mathbf{x})$ is positive definite then $\mathbf{x}$ is an isolated local minimum.
(ii) If $H(\mathbf{x})$ is negative definite then $\mathbf{x}$ is an isolated local maximum.
(iii) If $H(\mathbf{x})$ is indefinite then $\mathbf{x}$ is not a local extremum.

Proof. (i) Let us sketch the proof. Since $H(\mathbf{x})$ is positive definite there exists a constant $c > 0$ such that $\frac{1}{2}\boldsymbol{\xi}^T H(\mathbf{x})\, \boldsymbol{\xi} \geq c|\boldsymbol{\xi}|^2$ for all $\boldsymbol{\xi} \in \mathbb{R}^n$. This means, we can always find an $\epsilon > 0$ so that the $\mathcal{O}(|\boldsymbol{\xi}|^3)$ term in Eq. (A.45) is smaller than $\frac{1}{2}\boldsymbol{\xi}^T H(\mathbf{x})\, \boldsymbol{\xi}$ for all $\boldsymbol{\xi}$ with $|.\boldsymbol{\xi}| < \epsilon$. Hence, from Eq. (A.45), $f(\mathbf{x}+\boldsymbol{\xi}) - f(\mathbf{x}) > 0$ for all $\boldsymbol{\xi}$ with $|.\boldsymbol{\xi}| < \epsilon$ so that $\mathbf{x}$ is indeed an isolated local minimum.
(ii), (iii) The proofs are similar to the one for (i). □

Proposition A.4 and Theorem A.6 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 $\nabla f(\mathbf{x}) = \mathbf{0}$. Next, for each stationary point $\mathbf{x}$, compute the Hesse matrix $H(\mathbf{x})$ and its eigenvalues and use the following criterion:

$$\begin{array}{lcl} \mathbf{x} \text{ local isolated minimum} & \Leftarrow & \text{all eigenvalues of } H(\mathbf{x}) \text{ strictly positive} \\ \mathbf{x} \text{ local isolated maximum} & \Leftarrow & \text{all eigenvalues of } H(\mathbf{x}) \text{ strictly negative} \\ \mathbf{x} \text{ not a local extremum} & \Leftarrow & H(\mathbf{x}) \text{ has strictly positive and strictly negative eigenvalues} \end{array} \quad . \tag{A.46}$$

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

$$H(\mathbf{x}) = \begin{pmatrix} \partial_1^2 f & \partial_1\partial_2 f \\ \partial_1\partial_2 f & \partial_2^2 f \end{pmatrix}(\mathbf{x}) \quad \Rightarrow \quad \begin{cases} \det(H(\mathbf{x})) = (\partial_1^2 f\, \partial_2^2 f - (\partial_1\partial_2 f)^2)(\mathbf{x}) \\ \mathrm{tr}(H(\mathbf{x})) = (\partial_1^2 f + \partial_2^2 f)(\mathbf{x}) \end{cases} . \tag{A.47}$$

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

$$\begin{array}{lcll} \mathbf{x} \text{ local isolated minimum} & \Leftarrow & \lambda_1, \lambda_2 > 0 & \Leftrightarrow \quad \det(H(\mathbf{x})) > 0 \text{ and } \mathrm{tr}(H(\mathbf{x})) > 0 \\ \mathbf{x} \text{ local isolated maximum} & \Leftarrow & \lambda_1, \lambda_2 < 0 & \Leftrightarrow \quad \det(H(\mathbf{x})) > 0 \text{ and } \mathrm{tr}(H(\mathbf{x})) < 0 \\ \mathbf{x} \text{ not a local extremum} & \Leftarrow & \lambda_1 \lambda_2 < 0 & \Leftrightarrow \quad \det(H(\mathbf{x})) < 0 \end{array} \quad . \tag{A.48}$$

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 $\mathbf{x}$, one has to compute the eigenvalues of $H(\mathbf{x})$ from its characteristic polynomial and use the general criterion (A.46).

Application 1.38. Extrema for a function $f : \mathbb{R}^3 \to \mathbb{R}$

As a simple example, consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = x^2 - x + xy^2 + z^2 - 2z$ with gradient and Hesse matrix

$$\nabla f(x, y, z) = (y^2 + 3x - 1, 2xy, 2z - 2) \ , \qquad H(x, y, z) = \begin{pmatrix} 2 & 2y & 0 \\ 2y & 2x & 0 \\ 0 & 0 & 2 \end{pmatrix} \ .$$

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

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.

A.6 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 $\mathbf{z} = (z_1, \ldots, z_n)^T$ and also assume that $m < n$. We want to solve the equation $f(\mathbf{z}) = \mathbf{0}$, a common problem which arises in many situations in mathematics and physics.

An obvious way to proceed is as follows. Since $f(\mathbf{z}) = \mathbf{0}$ consists of m scalar equations for $n > m$ variables we can split our variables as $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, where $\mathbf{x} \in \mathbb{R}^{n-m}$ and $\mathbf{y} \in \mathbb{R}^m$ and then attempt to solve for the variables $\mathbf{y}$ in terms of $\mathbf{x}$. More precisely, suppose we have a point $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ which solves the equation, that is, $f(\mathbf{a}, \mathbf{b}) = \mathbf{0}$. For open neighbourhoods $\tilde{V}$ and $\tilde{W}$ of $\mathbf{a}$ and $\mathbf{b}$ (where $\tilde{V} \times \tilde{W} \subset U$) we would then like to find a function $g : \tilde{V} \to \tilde{W}$ such that $\mathbf{b} = g(\mathbf{a})$ and $f(\mathbf{x}, g(\mathbf{x})) = \mathbf{0}$ for all $\mathbf{x} \in \tilde{V}$. This function is precisely the desired solution $\mathbf{y} = g(\mathbf{x})$ of $f(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ (at least locally, near $(\mathbf{a}, \mathbf{b})$), 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.

Application 1.39. 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 $\mathbf{z} = (x, y)^T$ 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 $\mathbb{R}^2$. 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 points $(a, b) = (\pm 1, 0)$ where the circle intersects the x-axis and the tangent to the circle is vertical. For any open neighbourhood $\tilde{V} = (\pm 1 - \epsilon, \pm 1 + \epsilon)$ 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. In either case, we cannot define a function $y = g(x)$ since there is no (unique) y to assign as the value of g at x. 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 A.7. *(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 $(\mathbf{x}, \mathbf{y}) \in V \times W$ and choose a point $(\mathbf{a}, \mathbf{b}) \in V \times W$. If $f(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ and if the matrix*

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)}(\mathbf{a}, \mathbf{b}) \tag{A.49}$$

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

Proof. The proof is somewhat lengthy and can be found in the literature, see for example [10]. □

The maximal rank condition on the (partial) Jacobi matrix (A.49) is the generalisation of the requirement, stated earlier, that the tangent at $(\mathbf{a}, \mathbf{b})$ 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 $f(\mathbf{x}, g(\mathbf{x})) = \mathbf{0}$:

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_{n-m})} + \frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)} \frac{\partial(g_1, \ldots, g_m)}{\partial(x_1, \ldots, x_{n-m})} = \mathbf{0} \,. \tag{A.50}$$

The second Jacobi matrix in the above equation is precisely the one in Eq. (A.49) which is required to have maximal rank. Hence, we can invert this matrix and solve for the Jacobi matrix of g:

$$\frac{\partial(g_1, \ldots, g_m)}{\partial(x_1, \ldots, x_{n-m})} = - \left[\frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)} \right]^{-1} \frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_{n-m})} \,. \tag{A.51}$$

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.

Application 1.40. 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

$$\partial_y f = x_2^2 y^2 \tag{A.52}$$

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. (A.51):

$$\nabla g = -\left[\partial_y f\right]^{-1} \frac{\partial f}{\partial(x_1, x_2)} = -\frac{1}{x_2^2 y^2}\left(x_1^2, \frac{2}{3} x_2 y^3\right) . \tag{A.53}$$

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 $\mathbf{a} \in W$ and $\mathbf{b} \in \mathbb{R}^m$ with $\mathbf{a} = g(\mathbf{b})$ we want to know if the equation $\mathbf{x} = g(\mathbf{y})$ can be solved for $\mathbf{y}$ in terms of $\mathbf{x}$ near $(\mathbf{a}, \mathbf{b})$. To make contact with the implicit function theorem, we define the auxiliary function $f : \mathbb{R}^{2m} \to \mathbb{R}^m$ by $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}) - \mathbf{x}$. Then we have $f(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ and the implicit function theorem states that f can be solved for $\mathbf{y}$ in terms of $\mathbf{x}$ near $(\mathbf{a}, \mathbf{b})$ if the matrix

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)}(\mathbf{a}, \mathbf{b}) = \frac{\partial(g_1, \ldots, g_m)}{\partial(y_1, \ldots, y_m)}(\mathbf{b}) \tag{A.54}$$

has maximimal rank. Hence, we have

Corollary A.2. *Let $g : W \to \mathbb{R}^m$, where $W \subset \mathbb{R}^m$ open, be a differentiable function with $\mathbf{a} = g(\mathbf{b})$ for certain points $\mathbf{a} \in W$ and $\mathbf{b} \in \mathbb{R}^m$. Then g is invertible (that is, the equation $\mathbf{x} = g(\mathbf{y})$ can be solved for $\mathbf{y}$) in terms of $\mathbf{x}$ near $(\mathbf{a}, \mathbf{b})$ if the Jacobi matrix*

$$\frac{\partial(g_1, \ldots, g_m)}{\partial(y_1, \ldots, y_m)}(\mathbf{b}) \tag{A.55}$$

of g is invertible at $\mathbf{b}$.

Application 1.41. 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

$$f(P(v, t), v, t) = f(p, V(p, t), t) = f(p, v, T(p, v)) = 0 . \tag{A.56}$$

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. (A.51) (or, alternatively, by differentiating Eq. (A.56) using the chain rule) and this leads to

$$\begin{aligned} \partial_v P &= -\tfrac{\partial_v f}{\partial_p f} & \partial_t P &= -\tfrac{\partial_t f}{\partial_p f} \\ \partial_p V &= -\tfrac{\partial_p f}{\partial_v f} & \partial_t V &= -\tfrac{\partial_t f}{\partial_v f} \\ \partial_p T &= -\tfrac{\partial_p f}{\partial_t f} & \partial_v T &= -\tfrac{\partial_v f}{\partial_t f} \end{aligned} \tag{A.57}$$

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

$$\partial_v P\, \partial_t V\, \partial_p T = -1 \,, \qquad \partial_v P = (\partial_p V)^{-1} \,. \tag{A.58}$$

Do not let yourself be confused by the second of these equations which says we can simply invert individual partial derivatives to get the derivatives of the inverse function. At first sight, this seems to be at odds with our general rule, Eq. (A.35), 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 (A.58) are a consequence of this special set-up and are, therefore, not in contradiction with our general rule in Eq. (A.35). In fact, they were derived using the general chain rule, in the form of Eq. (A.51).

B Manifolds in $\mathbb{R}^n$

Differentiable manifolds are the main arena of an area of mathematics called differential geometry. It is a large subject (and a major area of contemporary research in mathematics) which is well beyond the scope of these lectures. Here, we would like to present a rather limited discussion of manifolds embedded in $\mathbb{R}^n$ and some of their elementary geometry. (A more advanced introduction to differential geometry can, for example, be found in Ref. [14].) Some special cases of this have, in fact, already been covered in your first year course. Apart from generalising these the main purpose of this appendix is to support some results used in the main part of the text and provide a (very) basic mathematical grounding for General Relativity.

B.1 Definition of manifolds

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

$$S^1 = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\} \,, \tag{B.1}$$

or we can use a parametrisation

$$S^1 = \{(\cos(t), \sin(t)) \,|\, t \in [0, 2\pi)\} \,. \tag{B.2}$$

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 the description (B.1) for the circle.

Definition B.1. *A subset $M \subset \mathbb{R}^n$ is a k-dimensional (differential) sub-manifold of $\mathbb{R}^n$ if, for every $\mathbf{p} \in M$ and an open neighbourhood $V \subset \mathbb{R}^n$ of $\mathbf{p}$ we have $V \cap M = \{\mathbf{x} \in \mathbb{R}^n \,|\, f_1(\mathbf{x}) = \cdots = f_{n-k}(\mathbf{x}) = 0\}$ for continuously differentiable functions $f_i : \mathbb{R}^n \to \mathbb{R}$ for which the Jacobi matrix*

$$\frac{\partial(f_1, \ldots, f_{n-k})}{\partial(x_1, \ldots, x_n)}(\mathbf{p}) \tag{B.3}$$

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. 18.) The rank condition ensures that this manifold is smooth,

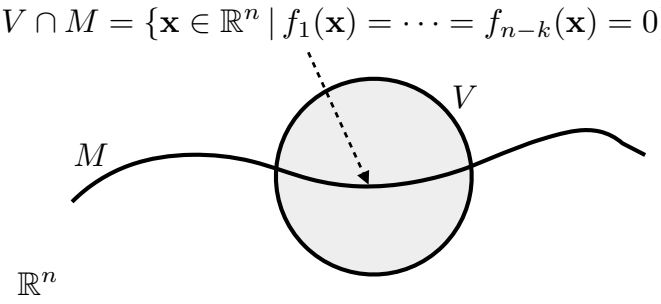

Figure 18: Manifold defined locally as the common zero locus of functions.

that is, there are no "edges". The circle S^1 in (B.1) provides an example with $n = 2$, $k = 1$ and $f_1(x, y) = x^2 + y^2 - 1$. The alternative description (B.2) of S^1 in terms of a parametrisation generalises to

Theorem B.1. *The set $M \subset \mathbb{R}^n$ is a k-dimensional sub-manifold of $\mathbb{R}^n$ iff, for every $\mathbf{p} \in M$ we have an open neighbourhood $V \subset \mathbb{R}^n$ of $\mathbf{p}$, an open set $U \subset \mathbb{R}^k$ and a bijective map $X = (X_1, \ldots, X_n)^T : U \to V \cap M$ such that the Jacobi matrix*

$$\frac{\partial(X_1, \ldots X_n)}{\partial(t_1, \ldots, t_k)} \tag{B.4}$$

has rank k for all $\mathbf{t} = (t_1, \ldots, t_k)^T \in U$. The map X is called a chart of M.

Proof. The proof involves the implicit function theorem A.7 and it can, for example, be found in Ref. [10]. □

The circle S^1 as parametrised in Eq. (B.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 sub-manifold of $\mathbb{R}^n$ (at least locally) by a parametrisation $X(\mathbf{t}) = (X_1(\mathbf{t}), \ldots X_n(\mathbf{t}))^T$, where $\mathbf{t} \in \mathbb{R}^k$ are the parameters. (See Fig. 19.) For one parameter ($k = 1$), this describes a one-dimensional sub-manifold,

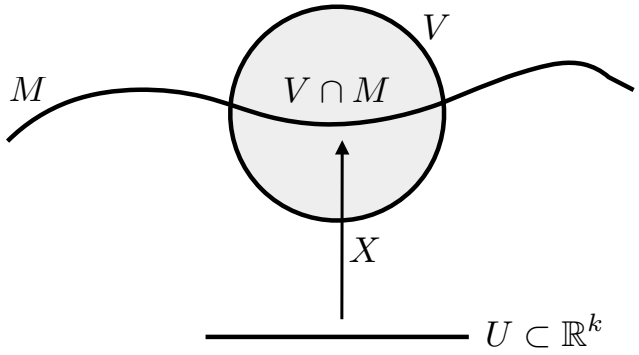

Figure 19: Manifold defined locally by a parametrisation (a chart).

that is a curve, for two parameters ($k = 2$) it describes a two-dimensional sub-manifold, that is a surface, and so on.

B.2 Tangent space

It is intuitively clear that a k-dimensional sub-manifold M of $\mathbb{R}^n$ has, at each point $\mathbf{p} \in M$, a tangent space which is a one-dimensional vector space for a curve ($k = 1$), a two-dimensional vector space for a surface ($k = 2$), and, generally, is k-dimensional. This tangent space at $\mathbf{p}$ is denoted by $T_\mathbf{p}M$ and is more precisely defined as follows.

Definition B.2. *The tangent space $T_\mathbf{p}M$ of a k-dimensional sub-manifold M of $\mathbb{R}^n$ is defined by*

$$T_\mathbf{p}M = \mathrm{Span}\left(\frac{\partial X}{\partial t_1}(\mathbf{p}), \ldots, \frac{\partial X}{\partial t_k}(\mathbf{p})\right) . \tag{B.5}$$

The vectors in $T_\mathbf{p}M$ are called tangent vectors of M at $\mathbf{p}$.

Note that $\frac{\partial X}{\partial t_a}(\mathbf{p})$, where $a = 1, \ldots, k$, are k vectors in $\mathbb{R}^n$ and the maximal rank condition on the Jacobian (B.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. 20.) Hence,

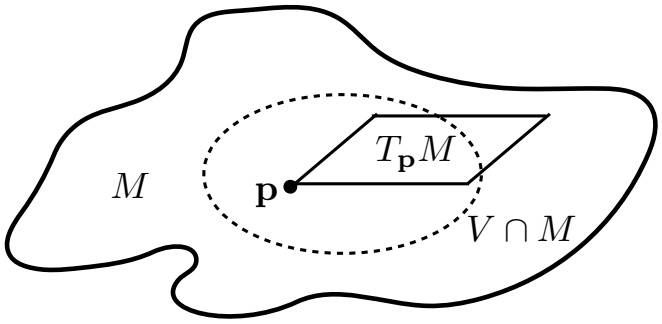

Figure 20: Tangent space $T_{\mathbf{p}}M$ of a manifold M at a point $\mathbf{p} \in M$.

a general tangent vector at $\mathbf{p} \in M$ can be written as

$$\sum_{a=1}^{k} v^a \frac{\partial X}{\partial t_a} \tag{B.6}$$

for $v^a \in \mathbb{R}$.

Note that the above definition of the tangent space is independent of the parametrisation used. Consider a new set of parameters $\tilde{\mathbf{t}} = (\tilde{t}_1, \ldots, \tilde{t}_k)$, a reparametrisation $\tilde{\mathbf{t}} = \tilde{\mathbf{t}}(\mathbf{t})$ and an alternative parametrisation $\tilde{X}(\tilde{\mathbf{t}}) := X(\mathbf{t}(\tilde{\mathbf{t}}))$ of the same manifold, such that the Jacobian $\frac{\partial \tilde{\mathbf{t}}}{\partial \mathbf{t}}$ has rank k. It follows that

$$\frac{\partial \tilde{X}}{\partial \tilde{t}_a} = \frac{\partial X}{\partial t_b} \frac{\partial t_b}{\partial \tilde{t}_a} \,. \tag{B.7}$$

Since the Jacobian $\frac{\partial \tilde{\mathbf{t}}}{\partial \mathbf{t}}$ has maximal rank the vectors $\frac{\partial \tilde{X}}{\partial \tilde{t}_a}$ and $\frac{\partial X}{\partial t_a}$ span the same space.

Application 2.42. Tangent space of S^1

It may be useful to work out the tangent space for an example. Consider S^1, parametrised as in Eq. (B.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

$$\frac{\partial X}{\partial t} = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \,. \tag{B.8}$$

Note that, intuitively, this vector is indeed tangential to the circle for every t.

Exercise B.2. *Write down a suitable parametrisation for a two-sphere $S^2 \subset \mathbb{R}^3$ and compute its tangent space at each point.*

Suppose we have a hypersurface $M \subset \mathbb{R}^n$ defined (locally) by the vanishing locus of a function f. It is well-known that the gradient ∇f is orthogonal to M but how is this statement actually proved? First we have to clarify what we mean by a vector being orthogonal to a hypersurface – after all, M is not necessarily a plane but can be a complicated surface. We say a vector $\mathbf{w} \in \mathbb{R}^n$ is orthogonal to M at a point $\mathbf{p} \in M$ if $\mathbf{w} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in T_{\mathbf{p}}M$ so, in short, if $\mathbf{w}$ is orthogonal to the entire tangent space at $\mathbf{p}$. So this is the property we have to show for the gradient $\nabla f(\mathbf{p})$.

From Def. B.1 we know that at least one component of $\nabla f(\mathbf{p})$ must be non-zero. Without restricting generality, let us assume this is the last component and we spit the coordinates in the neighbourhood of $\mathbf{p}$ up as $(\mathbf{x}, y) \in \mathbb{R}^n$, where $\mathbf{x} \in \mathbb{R}^{n-1}$ and $y \in \mathbb{R}$. Since $\partial_y f(\mathbf{p}) \neq 0$ we know from the implicit function theorem A.7 that we have a solution $y = g(\mathbf{x})$ with $f(\mathbf{x}, g(\mathbf{x})) = 0$ in a neighbourhood of $\mathbf{p}$. Further, from Eq. (A.51) we know that the derivatives of g are given by

$$\partial_i g = -\frac{\partial_i f}{\partial_y f} \quad \text{for} \quad i = 1, \dots, n-1 \,. \tag{B.9}$$

Using the function g we can also write down a parametrisation for the manifold and compute the resulting tangent vectors.

$$X(\mathbf{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ g(\mathbf{x}) \end{pmatrix} \qquad \Rightarrow \qquad \frac{\partial X}{\partial x_i} = \begin{pmatrix} \mathbf{e}_i \\ \partial_i g \end{pmatrix} \quad \text{for} \quad i = 1, \dots, n-1 \,. \tag{B.10}$$

Here we have used the coordinates $\mathbf{x} = (x_1, \dots, x_{n-1})^T$ as the parameters and $\mathbf{e}_i$ 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

$$\nabla f \, \frac{\partial X}{\partial x_i} = \partial_i f + \partial_y f \, \partial_i g \overset{\text{Eq. (B.9)}}{=} 0 \,. \tag{B.11}$$

This shows that ∇f is indeed orthogonal to all tangent vectors and, hence, to $T_{\mathbf{p}}M$ and this is the desired result.

B.3 Integration over sub-manifolds of $\mathbb{R}^n$

Integrating over a sub-manifold of $\mathbb{R}^n$ is conceptually quite different from "plain" integration, as described in Section 1.3. The latter only requires an (integrable) function plus a measurable set to integrate over. To integrate functions over manifolds additional geometrical information is required and, as we will now see, this information is provided via a metric.

B.3.1 Metric and Gram's determinant

Integration over (k-dimensional) sub-manifolds of $\mathbb{R}^n$ requires a non-trivial "measure", dS, in the integral which we now construct. First define the *metric*

$$g_{ab} := \frac{\partial X}{\partial t_a} \cdot \frac{\partial X}{\partial t_b} \,, \tag{B.12}$$

a symmetric $k \times k$ matrix at every point $\mathbf{t} \in U$ which contains the dot products of the tangent vectors $\frac{\partial X}{\partial t_a}$. The metric encodes the fact that the tangent vectors $\frac{\partial X}{\partial t_a}$, while providing a basis for the tangent space $T_{\mathbf{p}}M$, do not necessarily form an ortho-normal basis and this, in turn, can be seen as a measure for the non-trivial "shape" of the manifold. The metric is the central object of General Relativity, which describes the effects of gravity via a curved space time.

To see how the metric changes if we consider another set of parameters $\tilde{\mathbf{t}} = \tilde{\mathbf{t}}(\mathbf{t})$, we can use Eq. (B.7), which leads to

$$\tilde{g}_{ab} = \frac{\partial \tilde{X}}{\partial \tilde{t}_a} \cdot \frac{\partial \tilde{X}}{\partial \tilde{t}_b} = \frac{\partial t_c}{\partial \tilde{t}_a} \frac{\partial t_d}{\partial \tilde{t}_b} g_{cd} \,. \tag{B.13}$$

The determinant $g := \det(g_{ab})$ of this metric is called *Gram's determinant* and, under a coordinate transformation, it changes as

$$g := \det(g_{ab}) \qquad \Rightarrow \qquad \tilde{g} = \det\left(\frac{\partial \mathbf{t}}{\partial \tilde{\mathbf{t}}}\right)^2 g \,. \tag{B.14}$$

B.3.2 Definition of integration over sub-manifolds

We are now ready [15] to define integration over a sub-manifold of $\mathbb{R}^n$.

Definition B.3. *Let $M \subset \mathbb{R}^n$ be a k-dimensional sub-manifold of $\mathbb{R}^n$ which (up to a set of measure zero) is given by a single chart $X : U \to M$. For a function $f : M \to \mathbb{R}$ the integral over M is defined as*

$$\int_M f \, dS := \int_U f(X(\mathbf{t})) \sqrt{g} \, d^k t \, . \tag{B.15}$$

Symbolically this can also be written as $dS = \sqrt{g} \, d^k t$.

Note that the integral on the RHS is well-defined - it is simply an integral over the parameter space $U \subset \mathbb{R}^k$. However, it is important to check that this definition is independent of the parametrisation chosen. Clearly, we do not want the value of integrals to depend on how we choose to parametrise the manifold. Consider a re-parametrisation $\tilde{\mathbf{t}} = \tilde{\mathbf{t}}(\mathbf{t})$ and the transformation rules

$$\sqrt{\tilde{g}} = \left| \det \left(\frac{\partial \mathbf{t}}{\partial \tilde{\mathbf{t}}} \right) \right| \sqrt{g} \, , \qquad d^k \tilde{t} = \left| \det \left(\frac{\partial \tilde{\mathbf{t}}}{\partial \mathbf{t}} \right) \right| d^k t \tag{B.16}$$

where the former equation follows from (B.14) and the latter is simply the transformation formula for integrals in $\mathbb{R}^k$. Then we learn that

$$d\tilde{S} = \sqrt{\tilde{g}} \, d^k \tilde{t} = \sqrt{g} \, d^k t = dS \, , \tag{B.17}$$

so the integral is indeed independent of the parametrisation. Essentially, this is the reason for including the factor $\sqrt{g}$ in the measure - its transformation cancels the transformation under coordinate change, as is evident from Eq. (B.16).

Exercise B.3. *Parametrise the upper half circle in $\mathbb{R}^2$ in two different ways (by an angle and by the x coordinate) and check that the integral over the half-circle is the same for these two parametrisations.*

B.3.3 A few special cases

It is useful to apply this to a few specific cases. First consider a curve $X : [a, b] \to \mathbb{R}^n$, where $X(t) = (X_1(t), \ldots, X_n(t)^T$. A quick calculation shows that

$$g = \sum_{i=1}^{n} \left(\frac{dX^i}{dt} \right)^2 \qquad \Rightarrow \qquad dS = \sqrt{\sum_{i=1}^{n} \left(\frac{dX^i}{dt} \right)^2} \, dt \, . \tag{B.18}$$

This is indeed the well-known measure for integrating over curves.

Next, let us consider surfaces in $\mathbb{R}^3$, parametrised by $X : U \to \mathbb{R}^3$ with parameters $(t_1, t_2) \in U$. Since we are in $\mathbb{R}^3$ we can use the cross product to define a normal vector and a unit normal vector to the surface by

$$\mathbf{N} := \frac{\partial X}{\partial t_1} \times \frac{\partial X}{\partial t_2} \, , \qquad \mathbf{n} := \frac{\mathbf{N}}{|\mathbf{N}|} \, . \tag{B.19}$$

The metric is two-dimensional and given by

$$(g_{ab}) = \begin{pmatrix} \left| \frac{\partial X}{\partial t_1} \right|^2 & \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} \\ \frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2} & \left| \frac{\partial X}{\partial t_2} \right|^2 \end{pmatrix} \tag{B.20}$$

[15] For simplicity we focus on the case where the manifold is, up to a set of measure zero, given by a single chart. If multiple charts come into play a further technical twist, referred to as *partition of unity* is required.

For Gram's determinant we find

$$g = \det(g_{ab}) = \left|\frac{\partial X}{\partial t_1}\right|^2 \left|\frac{\partial X}{\partial t_2}\right|^2 - \left(\frac{\partial X}{\partial t_1} \cdot \frac{\partial X}{\partial t_2}\right)^2 = \left|\frac{\partial X}{\partial t_1} \times \frac{\partial X}{\partial t_2}\right|^2 = |\mathbf{N}|^2 \,, \tag{B.21}$$

end, hence the measure for surfaces in $\mathbb{R}^3$ can be written as

$$dS = |\mathbf{N}|\, dt_1\, dt_2 \,. \tag{B.22}$$

What if the surface is given not by a parametrisation but as the zero locus of a function f, so that $M = \{(x, y, z) \in \mathbb{R}^3 \,|\, f(x, y, z) = 0\}$? In this case, we can solve (at least locally where $\frac{\partial f}{\partial z} \neq 0$, from the implicit function theorem A.7) the equation $f(x, y, z) = 0$ for z and obtain $z = h(x, y)$ for some function h. This provides us with a parametrisation of the surface in terms of $t_1 = x$ and $t_2 = y$, given by

$$X(x, y) = \begin{pmatrix} x \\ y \\ h(x, y) \end{pmatrix} \,. \tag{B.23}$$

For this parametrisation the metric and Gram's determinant read (denoting partial derivatives by subscripts, for simplicity)

$$(g_{ab}) = \begin{pmatrix} 1 + h_x^2 & h_x h_y \\ h_x h_y & 1 + h_y^2 \end{pmatrix} \qquad \Rightarrow \qquad g = 1 + h_x^2 + h_y^2 \,. \tag{B.24}$$

Since $f(x, y, h(x, y)) = 0$ (as $z = h(x, y)$ is a solution) it follows, by applying the chain rule that

$$\begin{aligned} f_x + f_z h_x = 0 \qquad &\Rightarrow \qquad h_x = -\tfrac{f_x}{f_z} \\ f_y + f_z h_y = 0 \qquad &\Rightarrow \qquad h_y = -\tfrac{f_y}{f_z} \end{aligned} \tag{B.25}$$

These equations allow us to re-write Gram's determinant and the measure in terms of the function f:

$$g = \frac{|\nabla f|^2}{|f_z|^2} \qquad \Rightarrow \qquad dS = \frac{|\nabla f|}{|f_z|}\, dx\, dy = \frac{dx\, dy}{|n_3|} \,, \quad \mathbf{n} = \frac{\nabla f}{|\nabla f|} \,. \tag{B.26}$$

This is a known result from year 1 but note that it applies to a rather special case - the general (and much more symmetric) formula is the one given in Def. (B.3).

As a final (and more explicit) application, let us compute the measure on a two-sphere S^2 with the standard parametrisation $X(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)^T$, where $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$. With the two tangent vectors

$$\frac{\partial X}{\partial \theta} = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} \,, \qquad \frac{\partial X}{\partial \varphi} = \begin{pmatrix} -\sin\theta\sin\varphi \\ \sin\theta\cos\varphi \\ 0 \end{pmatrix} \,, \tag{B.27}$$

the metric and Gram's determinant are

$$(g_{ab}) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \qquad \Rightarrow \qquad g = \sin^2\theta \,. \tag{B.28}$$

As a result we find the well-known measure

$$dS = \sin\theta\, d\theta\, d\varphi \tag{B.29}$$

for the integration over S^2.

Exercise B.4. *Consider an ellipsoid in $\mathbb{R}^3$ with half-axis a, b and c and parametrisation $X(\theta, \varphi) = (a\sin\theta\cos\varphi, b\sin\theta\sin\varphi, c\cos\theta)^T$. Work out the corresponding measure dS.*

B.4 Laplacian

Suppose we have a k-dimensional manifold $M \subset \mathbb{R}^n$ and a chart $X : U \to M$ and we denote the parameters by $\mathbf{t} = (t_1, \ldots, t_k)^T \in U$, as usual. Suppose we choose another set of coordinates $\tilde{\mathbf{t}} = (\tilde{t}_1, \ldots, \tilde{t}_k)^T = \mathcal{T}(\mathbf{t})$ on $\mathcal{T}(U)$ such that

$$g_{ab} = \frac{\partial \tilde{t}_c}{\partial t_a} \frac{\partial \tilde{t}_d}{\partial t_b} \delta_{cd} \,, \tag{B.30}$$

that is, relative to the coordinates $\tilde{t}_a$, the metric is δ_{cd}. In those coordinates, the Laplacian is of the standard form

$$\Delta = \sum_{a=1}^{k} \frac{\partial^2}{\partial^2 \tilde{t}_a} \,. \tag{B.31}$$

We would like to define a notion of a Laplacian Δ_X on M which means re-expressing the Laplacian in terms of the original coordinates t_a.

Definition B.4. *Given the above set-up, the Laplacian Δ_X, associated to the chart $X : U \to M$, is defined by*

$$\Delta_X(f \circ \mathcal{T}) := \Delta(f) \circ \mathcal{T} \,, \tag{B.32}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable functions.

Note that this is quite a natural definition. The composition $f \circ \mathcal{T}$ is a function on the parameter space U and we define the action of the Laplacian Δ_X on this function by the action of the ordinary (Cartesian) Laplacian, Δ, on f followed by a composition with $\mathcal{T}$, to make this a function on U. While the above definition seems natural it is not particularly practical for computations. To this end, we have the following

Lemma B.1. *The Laplacian Δ_X defined in Def. (B.4) can be written as*

$$\Delta_X = \frac{1}{\sqrt{g}} \frac{\partial}{\partial t_a} \left(\sqrt{g}\, g^{ab} \frac{\partial}{\partial t_b} \right) \,, \tag{B.33}$$

where g^{ab} are the entries of the inverse of the metric (g_{ab}).

Proof. For two twice continuously differentiable functions $u, v : \mathcal{T}(U) \to \mathbb{R}$ and $\tilde{u} = u \circ \mathcal{T}$, $\tilde{v} = v \circ \mathcal{T}$ we have from the transformation formula

$$\int_{\mathcal{T}(U)} (\Delta u) v \, d^k\tilde{t} = \int_U ((\Delta u) \circ \mathcal{T})(v \circ \mathcal{T}) \sqrt{g}\, d^k t = \int_U (\Delta_X \tilde{u}) \tilde{v} \, \sqrt{g}\, d^k t \,. \tag{B.34}$$

On the other hand, we can re-write the same integral as

$$\begin{aligned} \int_{\mathcal{T}(U)} (\Delta u) v \, d^k\tilde{t} &= -\int_{\mathcal{T}(U)} (\nabla u) \cdot (\nabla v) \, d^k\tilde{t} = -\int_U ((\nabla u) \cdot (\nabla v)) \circ \mathcal{T} \, \sqrt{g}\, d^k t \\ &= -\int_U g^{ab} \frac{\partial \tilde{u}}{\partial t_a} \frac{\partial \tilde{v}}{\partial t_b} \sqrt{g}\, d^k t = \int_U \frac{\partial}{\partial t_a} \left(\sqrt{g}\, g^{ab} \frac{\partial \tilde{u}}{\partial t_b} \right) \tilde{v} \, d^k t \end{aligned} \tag{B.35}$$

The right-hand sides of Eqs. (B.34) and (B.35) are equal for any $\tilde{v}$ and given we are dealing with continuous functions this means

$$\sqrt{g}\, \Delta_X \tilde{u} = \frac{\partial}{\partial t_a} \left(\sqrt{g}\, g^{ab} \frac{\partial \tilde{u}}{\partial t_b} \right) \,, \tag{B.36}$$

which is the desired result. □

C 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 sub-manifolds thereof, in line with our basic treatment of manifolds in Appendix B. (Ref. [14] contains a more advanced treatment of differential forms.)

C.1 Differential one-forms

C.1.1 Definition of differential one-forms

We begin by defining differential one-forms on open sets $U \subset \mathbb{R}^n$ with coordinates $\mathbf{x} = (x^1, \ldots, x^n)^T$. If we view U as a manifold (in the sense of Theorem B.1 and, for example, parametrised by itself using the identity map) then attached to every point $\mathbf{p} \in U$ there is a tangent space $T_\mathbf{p}U$. Since U is an n-dimensional manifold we have $T_\mathbf{p}U \cong \mathbb{R}^n$ and we can use the standard unit vectors $\mathbf{e}_i$, where $i = 1, \ldots, n$, as a basis of $T_\mathbf{p}U$. Intuitively, just think about a family of vector spaces, one each attached to every point $\mathbf{p} \in U$, as in Fig. 21. Basically, any construction in linear algebra can now be carried out for this family.

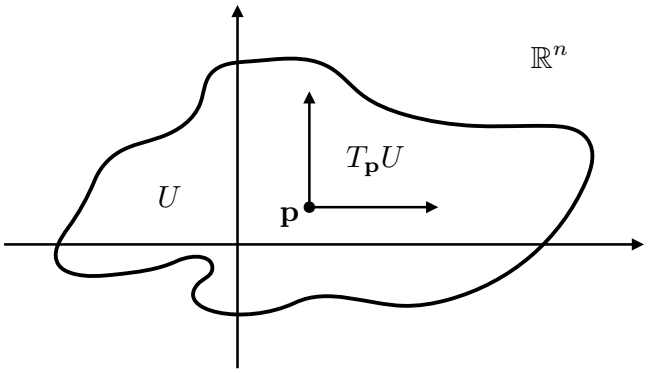

Figure 21: Open set $U \subset \mathbb{R}^n$ and tangent space $T_\mathbf{p}U$ at a point $\mathbf{p} \in U$.

In particular, every tangent space $T_\mathbf{p}U$ has a dual vector space $T^*_\mathbf{p}U$ (which consists of linear functionals $T_\mathbf{p}U \to \mathbb{R}$), called the *co-tangent space*. The elements of the co-tangent space are called *co-tangent vectors*. So attached to every point $\mathbf{p} \in U$ we have two vector spaces, the tangent space $T_\mathbf{p}U$ and its dual, the co-tangent space $T^*_\mathbf{p}U$. We are now ready to define differential one-forms.

Definition C.1. *(Differential one-forms) A differential one-form is a map* $w : U \to T^*U := \bigcup_{\mathbf{p}\in U} T^*_\mathbf{p}U$, *with* $\mathbf{p} \mapsto w_\mathbf{p}$, *such that* $w_\mathbf{p} \in T^*_\mathbf{p}U$.

Stated less formally, a differential one-form w provides us with a co-tangent vector $w_\mathbf{p}$ at every point $\mathbf{p} \in U$. Recall that such a co-tangent vector is a functional on the tangent space, so it provides a map $w_\mathbf{p} : T_\mathbf{p}U \to \mathbb{R}$ and for a tangent vector $\mathbf{v} \in T_\mathbf{p}U$ we have $w_\mathbf{p}(\mathbf{v}) \in \mathbb{R}$.

C.1.2 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 C.2. *(Total differential) For a differentiable function $f : U \to \mathbb{R}$ the total differential df is a one-form defined by*

$$df_{\mathbf{p}}(\mathbf{v}) := \nabla f_{\mathbf{p}} \cdot \mathbf{v} = \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x^i}\right|_{\mathbf{p}} v^i \tag{C.1}$$

where $\mathbf{v} = \sum_{i=1}^{n} v^i \mathbf{e}_i \in T_{\mathbf{p}}U$.

Exercise C.1. *Convince yourself that the total differential defined in Def. C.2 is indeed a differential one-form.*

C.2 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

$$x^i(\mathbf{p}) := p^i \ , \tag{C.2}$$

where $\mathbf{p} = (p^1, \ldots, p^n)^T$, which assign to each point $\mathbf{p} \in U$ 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 $\mathbf{e}_i$ of the tangent space.

$$dx^i|_{\mathbf{p}}(\mathbf{e_j}) = \nabla x^i|_{\mathbf{p}} \cdot \mathbf{e}_j = \mathbf{e}^i \cdot \mathbf{e}_j = \delta^i_j \qquad \Rightarrow \qquad dx^i|_{\mathbf{p}}(\mathbf{v}) = v^i \ . \tag{C.3}$$

This is precisely the defining relation for a dual basis [16] and we learn that $(dx^i|_{\mathbf{p}})$ is the basis of $T^*_{\mathbf{p}}U$ dual to the basis $(\mathbf{e}_i)$ of $T_{\mathbf{p}}U$. Hence, we have the following

Proposition C.1. *The total differentials $(dx^1|_{\mathbf{p}}, \ldots, dx^n|_{\mathbf{p}})$ of the coordinate functions form a basis of the co-tangent space $T^*_{\mathbf{p}}U$. Hence, every differential one-form w and every total differential df on U can be written as*

$$w = \sum_{i=1}^{n} w_i \, dx^i \ , \qquad df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \, dx^i \ , \tag{C.4}$$

where $w_i : U \to \mathbb{R}$ 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 $T^*_{\mathbf{p}}U$. This means for every one-form w and every point $\mathbf{p} \in U$ we can write $w_{\mathbf{p}} = \sum_{i=1}^{n} w_i(\mathbf{p}) dx^i|_{\mathbf{p}}$, for some suitable coefficients $w_i(\mathbf{p})$. Dropping the argument $\mathbf{p}$ gives the first Eq. (C.4). To show the second Eq. (C.4) work out

$$df|_{\mathbf{p}}(\mathbf{v}) = \nabla f|_{\mathbf{p}} \cdot \mathbf{v} = \left.\frac{\partial f}{\partial x^i}\right|_{\mathbf{p}} v^i = \sum_i \left.\frac{\partial f}{\partial x^i}\right|_{\mathbf{p}} dx^i|_{\mathbf{p}}(\mathbf{v}) \ , \tag{C.5}$$

where Eq. (C.3) has been used in the last step. Dropping the arguments $\mathbf{v}$ and $\mathbf{p}$ leads to the second Eq. (C.4). □

[16] 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 (C.3). This equation says that $dx^i|_{\mathbf{p}}$, when acting on a tangent $\mathbf{v} \in T_{\mathbf{p}}M$ (and it is this tangent vector which should really be seen as the displacement from $\mathbf{p}$), 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 $w_i : U \to \mathbb{R}$ and we can write down a differential form using Eq. (C.4). In particular, we see that differential one-forms w on U are in one-to-one correspondence with vector fields $\mathbf{A} : U \to \mathbb{R}^n$ via

$$w = \sum_{i=1}^{n} w_i \, dx^i \qquad \longleftrightarrow \qquad \mathbf{A} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} . \tag{C.6}$$

This is our first hint that differential forms provide us with a new way to think about vector calculus. Under the identification (C.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 (C.4) of differential forms can be used to define their properties in terms of properties of the constituent functions w_i.

Definition C.3. *A one-form $w = \sum_{i=1}^{n} w_i \, dx^i$ on U is called differentiable (continuously differentiable, k times continuously differentiable) if all functions $w_i : U \to \mathbb{R}$ are differentiable (continuously differentiable, k times continuously differentiable).*

C.2.1 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 w on U provides a linear functional $w|_{X(t)} : T_{X(t)}U \to \mathbb{R}$ which can act on the tangent to the curve to produce a real number, so

$$w|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) \in \mathbb{R} . \tag{C.7}$$

Based on this observation we can define

Definition C.4. *Let w be a differential one-form on U and $X : [a, b] \to U$ a curve. Then, we define the integral of w over X as*

$$\int_X w := \int_a^b w|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) dt . \tag{C.8}$$

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

$$w = \sum_{i=1}^{n} w_i \, dx^i , \qquad \frac{dX}{dt} = \sum_{j=1}^{n} \frac{dX^j}{dt} \mathbf{e}_j , \tag{C.9}$$

we have

$$w|_{X(t)} \left(\frac{dX}{dt}\bigg|_{X(t)} \right) = \sum_{i,j} w_i(X(t)) \, \frac{dX^j}{dt}\bigg|_{X(t)} dx^i|_{X(t)}(\mathbf{e}_j) = \sum_i w_i(X(t)) \, \frac{dX^i}{dt}\bigg|_{X(t)} = \mathbf{A}(X(t)) \cdot \frac{dX}{dt}\bigg|_{X(t)}$$

where the duality property (C.3) has been used in the second last step and we have used the identification (C.6) of one-forms and vector fields in the last step. Hence, the integral of w over X can also be written as

$$\int_X w = \sum_{i=1}^{n} \int_a^b w_i(X(t)) \, \frac{dX^i}{dt}\bigg|_{X(t)} dt = \int_a^b \mathbf{A}(X(t)) \cdot \frac{dX}{dt}\bigg|_{X(t)} dt . \tag{C.10}$$

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

$$\int_X df = \int_a^b \nabla f(X(t)) \cdot \left.\frac{dX}{dt}\right|_{X(t)} dt = \int_a^b \frac{d}{dt}(f(X(t))) = f(X(b)) - f(X(a)) \tag{C.11}$$

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.

C.3 Alternating k-forms

C.3.1 Definition of alternating k-forms

We start with a vector space V over a field 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 C.5. *(Alternating k-forms) An alternating k-form w on a vector space V over F is a map $w : V^{\otimes k} \to F$ which is*

(i) linear in each of its k arguments, so $w(\ldots, \alpha\mathbf{v} + \beta\mathbf{w}, \ldots) = \alpha w(\ldots, \mathbf{v}, \ldots) + \beta w(\ldots, \mathbf{w}, \ldots)$
(ii) alternating in the sense that $w(\ldots, \mathbf{v}, \ldots, \mathbf{w}, \ldots) = -w(\ldots, \mathbf{w}, \ldots, \mathbf{v}, \ldots)$
(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 $\Lambda^0 V^* = F$.

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.

C.3.2 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 C.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*

$$(\varphi_1 \wedge \ldots \wedge \varphi_k)(\mathbf{v}_1, \ldots, \mathbf{v}_k) := \det \begin{pmatrix} \varphi_1(\mathbf{v}_1) & \cdots & \varphi_1(\mathbf{v}_k) \\ \vdots & & \vdots \\ \varphi_k(\mathbf{v}_1) & \cdots & \varphi_k(\mathbf{v}_k) \end{pmatrix} . \tag{C.12}$$

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. (C.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):

$$\begin{aligned} \varphi_1 \wedge \ldots \wedge (\alpha\varphi_i + \beta\tilde{\varphi}_i) \wedge \ldots \wedge \varphi_k &= \alpha\, \varphi_1 \wedge \ldots \wedge \varphi_i \wedge \ldots \wedge \varphi_k + \beta\, \varphi_1 \wedge \ldots \wedge \tilde{\varphi}_i \wedge \ldots \wedge \varphi_k & \text{(C.13)} \\ \varphi_1 \wedge \ldots \wedge \varphi_i \wedge \varphi_{i+1} \wedge \ldots \wedge \varphi_k &= -\varphi_1 \wedge \ldots \wedge \varphi_{i+1} \wedge \varphi_i \wedge \ldots \wedge \varphi_k & \text{(C.14)} \\ \ldots \wedge \varphi \wedge \varphi \wedge \ldots &= 0 & \text{(C.15)} \end{aligned}$$

The first two rules follow from linearity and anti-symmetry of the determinant, respectively, the third one is a direct consequence of the second.

C.3.3 Basis for alternating k-forms

Proposition C.2. *Let $(\boldsymbol{\epsilon}_*^1, \ldots, \boldsymbol{\epsilon}_*^n)$ be a basis of V^*. Then the k-forms $\boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k}$, where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ are a basis of $\Lambda^k V^*$. In particular, we have*

$$\dim(\Lambda^k V^*) = \begin{pmatrix} n \\ k \end{pmatrix} . \tag{C.16}$$

Proof. Let $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ be the basis of V with dual basis $\boldsymbol{\epsilon}_*^1, \ldots, \boldsymbol{\epsilon}_*^n$, so that $\boldsymbol{\epsilon}_*^i(\boldsymbol{\epsilon}_j) = \delta_j^i$. To proof linear independence consider

$$\sum_{i_1 < \cdots < i_k} \lambda_{i_1 \ldots i_k}\, \boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k} = 0 \tag{C.17}$$

and act with this equation on $(\boldsymbol{\epsilon}_{j_1}, \ldots, \boldsymbol{\epsilon}_{j_k})$, 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 $w \in \Lambda^k V^*$. Define the numbers $c_{i_1 \ldots i_k} := w(\boldsymbol{\epsilon}_{i_1}, \ldots, \boldsymbol{\epsilon}_{i_k})$ and the differential k-form

$$\tilde{w} := \sum_{j_1 < \cdots < j_k} c_{j_1 \ldots j_k} \boldsymbol{\epsilon}_*^{j_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{j_k} . \tag{C.18}$$

Then it follows that $\tilde{w}(\boldsymbol{\epsilon}_{i_1}, \ldots, \boldsymbol{\epsilon}_{i_k}) = c_{i_1 \ldots i_k} = w(\boldsymbol{\epsilon}_{i_1}, \ldots, \boldsymbol{\epsilon}_{i_k})$ and since $\tilde{w}$ and w agree on a basis we have $w = \tilde{w}$ and have, hence, written w as a linear combination (C.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 $\boldsymbol{\epsilon}_*^i$ of the dual space V^* we can use the wedge product to construct a basis $\boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k}$, 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

$$w = \sum_{i_1 < \cdots < i_k} w_{i_1 \ldots i_k} \boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k} = \frac{1}{k!} \sum_{i_1, \ldots, i_k} w_{i_1 \ldots i_k} \boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k} , \tag{C.19}$$

where the coefficients $w_{i_1 \ldots i_k} \in F$ 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. (C.15). Of course the wedge product generalises to arbitrary alternating forms by linearity. Specifically, consider an alternating k-form w as in Eq. (C.19) and an alternating l-forms ν given by

$$\nu = \frac{1}{l!} \sum_{j_1, \ldots j_l} \nu_{j_1 \ldots j_l} \boldsymbol{\epsilon}_*^{j_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{j_l} , \tag{C.20}$$

Their wedge product is an alternating $k+l$-form defined by

$$w \wedge \nu := \frac{1}{k!\, l!} \sum_{i_1,\ldots,i_k,j_1,\ldots,j_l} w_{i_1\ldots i_k} \nu_{j_1\ldots j_l} \boldsymbol{\epsilon}_*^{i_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{i_k} \wedge \boldsymbol{\epsilon}_*^{j_1} \wedge \ldots \wedge \boldsymbol{\epsilon}_*^{j_l} \; . \tag{C.21}$$

It is easy to see that

$$w \wedge \nu = (-1)^{kl} \nu \wedge w \; . \tag{C.22}$$

Let us illustrate alternating k-forms with an example.

Application 3.43. The three-dimensional case

Consider a three-dimensional vector space V with basis $(\boldsymbol{\epsilon}_i)$ and dual basis $(\boldsymbol{\epsilon}_*^i)$, 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
$\Lambda^1 V^*$	$(\boldsymbol{\epsilon}_*^1, \boldsymbol{\epsilon}_*^2, \boldsymbol{\epsilon}_*^3)$	$\sum_{i=1}^3 w_i \boldsymbol{\epsilon}_*^i$	3
$\Lambda^2 V^*$	$(\boldsymbol{\nu}^1 := \boldsymbol{\epsilon}_*^2 \wedge \boldsymbol{\epsilon}_*^3, \boldsymbol{\nu}^2 := \boldsymbol{\epsilon}_*^3 \wedge \boldsymbol{\epsilon}_*^1, \boldsymbol{\nu}^3 := \boldsymbol{\epsilon}_*^1 \wedge \boldsymbol{\epsilon}_*^3)$	$\sum_{i=1}^3 w_i \boldsymbol{\nu}^i$	3
$\Lambda^3 V^*$	$(\boldsymbol{\epsilon}_*^1 \wedge \boldsymbol{\epsilon}_*^2 \wedge \boldsymbol{\epsilon}_*^3)$	$w\, \boldsymbol{\epsilon}_*^1 \wedge \boldsymbol{\epsilon}_*^2 \wedge \boldsymbol{\epsilon}_*^3$	1

Suppose we have two alternating one-forms $v = \sum_{i=1}^3 v_i \boldsymbol{\epsilon}_*^i$ and $w = \sum_{j=1}^3 w_j \boldsymbol{\epsilon}_*^j$ and we write the coefficients as vectors, so $\mathbf{v} = (v_1, v_2, v_3)^T$ and $\mathbf{w} = (w_1, w_2, w_3)^T$. Then, taking the wedge product we find

$$v \wedge w = \sum_{i,j=1}^3 v_i w_j \boldsymbol{\epsilon}_*^i \wedge \boldsymbol{\epsilon}_*^j = \sum_{i=1}^3 (\mathbf{v} \times \mathbf{w})_i \boldsymbol{\nu}^i \; , \tag{C.23}$$

where $\boldsymbol{\nu}^i$ 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

$$\Lambda^0 V^* = F \; , \quad \Lambda^1 V^* = V^* \; , \quad \Lambda^2 V^* \; , \quad \cdots \; , \quad \Lambda^n V^* \tag{C.24}$$

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 $T_{\mathbf{p}} U$ have assumed the role of V). Now we can go further and arrange the whole tower of vector spaces (C.24) over each tangent space $V = T_{\mathbf{p}} U$. As we will see, doing this leads to higher-order differential forms.

C.4 Higher-order differential forms

C.4.1 Definition of differential forms

After this preparation it is now straightforward to define differential k-forms. We simply copy the definition C.1 of differential one-forms but replace the cotangent space $T_{\mathbf{p}}^* U$ used there by $\Lambda^k T_{\mathbf{p}}^* U$.

Definition C.7. *(Differential k-forms) A differential k-form is a map* $w : U \to \Lambda^k T^*U := \bigcup_{\mathbf{p}\in U} \Lambda^k T^*_{\mathbf{p}} U$, *with* $\mathbf{p} \mapsto w_{\mathbf{p}}$, *such that* $w_{\mathbf{p}} \in \Lambda^k T^*_{\mathbf{p}} U$.

From this definition, a differential k-form w provides an alternating k-form $w_{\mathbf{p}}$ over $T_{\mathbf{p}}U$ for every point $\mathbf{p} \in U$ which can act on k tangent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in T_{\mathbf{p}}U$ to give a number $w_{\mathbf{p}}(\mathbf{v}_1, \ldots, \mathbf{v}_k) \in \mathbb{R}$.

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 $dx^1_{\mathbf{p}}, \ldots, dx^n_{\mathbf{p}}$ form a basis of the cotengent space $T^*_{\mathbf{p}}U$ at $\mathbf{p} \in U$. Combining this with what we have said about alternating k-forms in the previous subsection (see Prop. C.2) shows that the k-forms $dx^{i_1}_{\mathbf{p}} \wedge \ldots \wedge dx^{i_k}_{\mathbf{p}}$, where $1 \leq i_1 < \ldots < i_k \leq n$, form a basis of $\Lambda^k T^*_{\mathbf{p}} U$. Hence, dropping the point $\mathbf{p}$, a differential k-form can be written as

$$w = \frac{1}{k!} \sum_{i_1,\ldots,i_k} w_{i_1\ldots i_k} dx^{i_1} \wedge \ldots \wedge dx^{i_k} \ , \tag{C.25}$$

where the $w_{i_1\ldots i_k} : U \to \mathbb{R}$ 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 w in terms of the functions $w_{i_1\ldots i_k} : U \to \mathbb{R}$, just as we have done in the case of differential one-forms (see Def. C.3). For example, we say that w is differentiable iff all the functions $w_{i_1\ldots i_k}$ 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 $w \wedge \nu$, where w and ν are differential k- and l-forms respectively, in complete analogy with what we have done for alternating forms in Eq. (C.21).

C.4.2 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 (C.25) is defined by

$$dw := \frac{1}{k!} \sum_{i_1,\ldots,i_k} dw_{i_1\ldots i_k} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = \frac{1}{k!} \sum_{i_1,\ldots,i_k,i} \frac{\partial w_{i_1\cdots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \ , \tag{C.26}$$

where we have worked out the total derivatives $dw_{i_1\cdots i_k}$ of the functions $w_{i_1\cdots i_k}$ 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

$$df = \sum_i \frac{\partial f}{\partial x^i} dx^i \ . \tag{C.27}$$

The exterior derivative satisfies a Leibnitz-type rule. If w is a k-form and ν is an l-form then

$$d(w \wedge \nu) = dw \wedge \nu + (-1)^k w \wedge d\nu \ . \tag{C.28}$$

Another simple but important property of the exterior derivative is

Proposition C.3. $d^2 = 0$

Proof. This statement follows from straightforward calculation.

$$\begin{aligned} dw &= \frac{1}{k!} \sum_{i,i_1,\ldots,i_k} \frac{\partial w_{i_1\ldots i_k}}{\partial x^i} dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \\ d^2 w &= d(dw) = \frac{1}{k!} \sum_{i,j,i_1,\ldots,i_k} \frac{\partial^2 w_{i_1\ldots i_k}}{\partial x^i \partial x^j} dx^j \wedge dx^i \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = 0 \end{aligned}$$

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 C.8. *A differential k-form w is called closed iff $dw = 0$. It is called exact iff there exists a differential $(k-1)$-form ν such that $w = d\nu$.*

An immediate consequence of Prop. C.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. [10]). 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

$$0 \longrightarrow \Omega^0(U) \xrightarrow{d_0} \Omega^1(U) \xrightarrow{d_1} \cdots\cdots \xrightarrow{d_{n-1}} \Omega^n(U) \xrightarrow{d_n} 0 \tag{C.29}$$

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. C.3 then translates into $d_k \circ d_{k-1} = 0$, that is, successive maps in the sequence (C.29) compose to zero. This makes the sequence (C.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 defined 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. [14]).

Let us take a practical approach and work out differential forms and the exterior derivative more explicitly for an example.

Application 3.44. 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:

$$\begin{aligned} d\mathbf{s} &:= (dx, dy, dz)^T \\ d\mathbf{S} &:= (dy \wedge dz, dz \wedge dx, dx \wedge dy)^T \\ dV &:= dx \wedge dy \wedge dz \; . \end{aligned} \tag{C.30}$$

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	$w_0 = f$	f	function
1	$w_1 = \mathbf{A} \cdot d\mathbf{s}$	$\mathbf{A} = (A_1, A_2, A_3)^T$	vector field
2	$w_2 = \mathbf{B} \cdot d\mathbf{S}$	$\mathbf{B} = (B_1, B_2, B_3)^T$	vector field
3	$w_3 = F\, dV$	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

$$dw_0 = df = \sum_{i=1}^{3} \frac{\partial f}{\partial x^i} dx^i = (\nabla f) \cdot d\mathbf{s} \; . \tag{C.31}$$

Hence, the exterior derivative acting on three-dimensional zero forms corresponds to the gradient of the associated function. What about one-forms?

$$\begin{aligned} dw_1 &= \sum_{i=1}^{3} dA_i \wedge dx^i = \sum_{i,j=1}^{3} \frac{\partial A_i}{\partial x^j} dx^j \wedge dx^i \\ &= \left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3} \right) dx^2 \wedge dx^3 + \left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1} \right) dx^3 \wedge dx^1 + \left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right) dx^1 \wedge dx^2 \\ &= (\nabla \times \mathbf{A}) \cdot d\mathbf{S} \; . \end{aligned} \tag{C.32}$$

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

$$dw_2 = \sum_{i=1}^{3} dB_i \wedge dS^i = \sum_{i,j=1}^{3} \frac{\partial B_i}{\partial x^j} dx^j \wedge dS^i = (\nabla \cdot \mathbf{B}) \, dV \; , \tag{C.33}$$

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

$$dw_3 = 0 \; . \tag{C.34}$$

These results can be summarised as follows. For three-dimensional differential forms the exterior derivative acts as follows

$$0 \longrightarrow \Omega^0(U) \xrightarrow{d} \Omega^1(U) \xrightarrow{d} \Omega^2(U) \xrightarrow{d} \Omega^3(U) \xrightarrow{d} 0 \; , \tag{C.35}$$

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 (see also Eq. (A.17)) reads:

$$0 \longrightarrow \mathcal{C}^\infty(U) \xrightarrow{\text{grad}=\nabla} \mathcal{V}(U) \xrightarrow{\text{curl}=\nabla\times} \mathcal{V}(U) \xrightarrow{\text{div}=\nabla\cdot} \mathcal{C}^\infty(U) \longrightarrow 0 \; . \tag{C.36}$$

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 $\nabla \cdot (\nabla \times \mathbf{B}) = 0$, that is, carrying out adjacent maps in the diagram (C.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 C.3. From Def. C.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 $\mathbf{A}$ the same property is called "conservative", meaning the vector field can be written as $\mathbf{A} = \nabla f$.

In essence, vector calculus is the calculus of differentials forms written in different (and some might say, awkward) notation.

Exercise C.2. *Write the vector fields* $\mathbf{A} = (x^2, y, z^3)^T$ *and* $\mathbf{B} = (y, -x, 0)^T$ *on* $\mathbb{R}^3$ *as differential one-forms and use the exterior derivative to calculate their curl. Next, write* $\mathbf{A}$ *and* $\mathbf{B}$ *as differential two-forms and use the exterior derivative to compute their divergence.*

Exercise C.3. *Repeat the above discussion of differential forms in three dimensions for the case of differential forms in four dimensions.*

Application 3.45. Some explicit differential forms

Consider the following differential one-forms in $\mathbb{R}^2$ and $\mathbb{R}^3$:

$$\omega = xdy + ydx \ , \qquad \nu = xdy - ydx \ , \qquad \mu = xdx + ydy + zdz \ . \tag{C.37}$$

Are these one-forms closed? A simple calculation based on the definition, Eq. (C.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. (C.27), that suitable functions are given by $f(x, y) = xy$ and $g(x, y, z) = (x^2 + y^2 + z^2)/2$.

C.4.3 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 w as in Eq. (C.25) it is defined by

$$\star w = \frac{1}{k!(n-k)!} \epsilon^{i_1 \cdots i_k}{}_{i_{k+1} \cdots i_n} w_{i_1 \cdots i_k} dx^{i_{k+1}} \wedge \cdots \wedge dx^{i_n} \ , \tag{C.38}$$

where $\epsilon_{i_1 \cdots i_n}$ is the Levi-Civita tensor in n dimensions [17]. 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

$$\Delta = d^\dagger d + d d^\dagger \ . \tag{C.39}$$

Exercise C.4. *Verify that the Laplacian (C.39) when acting on zero-forms, that is functions, does indeed coincide with the usual Laplacian.*

[17] 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. (C.38) remains valid with the understanding that indices on the ϵ tensor are lowered with g_{ij}.

Application 3.46. Maxwell's equations with differential forms

Maxwell's equations contain three vector fields, the electric field $\mathbf{E}$, the magentic field $\mathbf{B}$ and the current density $\mathbf{J}$, 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 $\mathbf{E}$ and $\mathbf{J}$ as one-forms and $\mathbf{B}$ as a two-form, that is,

$$\mathcal{E} := \mathbf{E} \cdot d\mathbf{s} \,, \qquad \mathcal{B} := \mathbf{B} \cdot d\mathbf{S} \,, \qquad \mathcal{J} := \mathbf{J} \cdot d\mathbf{s} \,. \tag{C.40}$$

It is also useful to work out the Hodge star of these fields. It follows from its definition (C.38) that

$$\star 1 = dV \,, \qquad \star d\mathbf{s} = d\mathbf{S} \,, \qquad \star d\mathbf{S} = d\mathbf{s} \,, \qquad \star dV = 1 \,. \tag{C.41}$$

and this leads immediately to

$$\star\mathcal{E} = \mathbf{E} \cdot d\mathbf{S} \,, \qquad \star\mathcal{B} = \mathbf{B} \cdot d\mathbf{s} \,. \tag{C.42}$$

Using the relations between curl/divergence and the exterior derivative in Eqs. (C.32) and (C.33) we have

$$d\mathcal{E} = (\nabla \times \mathbf{E}) \cdot d\mathbf{S} \,, \quad d\mathcal{B} = (\nabla \cdot \mathbf{B}) dV \,, \quad d^\dagger\mathcal{E} = \nabla \cdot \mathbf{E} \,, \quad d^\dagger\mathcal{B} = (\nabla \times \mathbf{B}) \cdot d\mathbf{s} \,. \tag{C.43}$$

Now we are ready to convert Maxwell's equations. In vector calculus notation they are given by

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi\rho \\ \nabla \times \mathbf{B} - \tfrac{1}{c}\dot{\mathbf{E}} &= \tfrac{4\pi}{c}\mathbf{J} \\ \nabla \times \mathbf{E} + \tfrac{1}{c}\dot{\mathbf{B}} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \,, \end{aligned} \tag{C.44}$$

where the dot stands for the time derivative $\frac{\partial}{\partial t}$ and c is the speed of light. Multiplying these equations with 1, $d\mathbf{s}$, $d\mathbf{S}$ and dV, respectively, and using Eqs. (C.43) they are easily converted to

$$\begin{aligned} d^\dagger\mathcal{E} &= 4\pi\rho \\ d^\dagger\mathcal{B} - \tfrac{1}{c}\dot{\mathcal{E}} &= \tfrac{4\pi}{c}\mathcal{J} \\ d\mathcal{E} + \tfrac{1}{c}\dot{\mathcal{B}} &= 0 \\ d\mathcal{B} &= 0 \,. \end{aligned} \tag{C.45}$$

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)

$$\mathcal{F} = \mathcal{E} \wedge dt + \mathcal{B} =: \frac{1}{2}\mathcal{F}_{\mu\nu} dx^\mu \wedge dx^\nu \,, \qquad j = \rho\, dt + \mathcal{J} =: j_\mu dx^\mu \,, \tag{C.46}$$

where the second equalities define the components $\mathcal{F}_{\mu\nu}$ and j_μ. Maxwell's equations (C.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. (C.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. (C.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 (C.28) for differential forms we find

$$\begin{aligned} \mathcal{F} &= \mathcal{E}\wedge dt + \mathcal{B} & \qquad \star\mathcal{F} &= -\star_3\mathcal{E} + dt\wedge\star_3\mathcal{B} \\ d\mathcal{F} &= \left(d_3\mathcal{E} + \dot{\mathcal{B}}\right)\wedge dt + d_3\mathcal{B} & \qquad d^\dagger\mathcal{F} &= d_3^\dagger\mathcal{E}\,dt + d_3^\dagger\mathcal{B} - \dot{\mathcal{E}} \end{aligned} \tag{C.47}$$

Comparing the last two of these equations with the three-dimensional version of Maxwell equations (C.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

$$d^\dagger\mathcal{F} = 4\pi j\,, \qquad d\mathcal{F} = 0\,. \tag{C.48}$$

Eqs. (C.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 (C.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 (C.45) but it is manifest in the four-dimensional one (C.48).

Covariant electro-magnetism in the form (C.48) is already quite elegant but it can be simplified further. Note that the second Eq. (C.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. (C.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

$$\mathcal{A}\mapsto\mathcal{A}' = \mathcal{A} + d\lambda\,, \tag{C.49}$$

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. C.3. In conclusion, since $\mathcal{A}$ and $\mathcal{A}'$, related by a gauge transformation (C.49), lead to the same field strength tensor $\mathcal{F}$ they describe the same physics.

The gauge transformation (C.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

$$d^\dagger \mathcal{A} = 0 \; . \tag{C.50}$$

Why can this be done? Suppose that $\mathcal{A}$ does not satisfy the Lorentz gauge condition (C.50). Then we perform a gauge transformation (C.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 Laplace equation [a].

What form do Maxwell's equations (C.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 (C.48) is automatically satisfied. The first, inhomogeneous Maxwell equation (C.48) becomes

$$d^\dagger dA = 4\pi j \qquad \xrightarrow{d^\dagger \mathcal{A}=0} \qquad \Delta A = 4\pi j \; . \tag{C.51}$$

That is, expressed in terms of the gauge field $\mathcal{A}$ and choosing the Lorentz gauge (C.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. (C.39), $d^\dagger d\lambda = \Delta\lambda$.

C.5 Integration of differential forms

We have already seen that differential one-forms are designed to be integrated over one-dimensional sub-manifolds, 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.

C.5.1 Definition of integral and general Stokes's theorem

Start with an open set $U \subset \mathbb{R}^n$, a differential k-form w on U and a k-dimensional manifold $M \subset U$, given by a parametrisation $X : V \to U$, with parameters $\mathbf{t} = (t_1, \ldots, t_k)^T \in V$. At every point $X(\mathbf{t})$ on the manifold M we have an n-dimensional tangent space $T_{X(\mathbf{t})}U$ with a k-dimensional subspace $T_{X(\mathbf{t})}M$, the tangent space to the manifold M at $X(\mathbf{t})$, the latter being spanned by the k tangent vectors $\frac{\partial X}{\partial t_i}(\mathbf{t}) \in T_{X(\mathbf{t})}U$. At each such point $X(\mathbf{t})$ the differential k-form w provides an alternating k-form $w_{X(\mathbf{t})}$ 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 w over the k-dimensional manifold M as [18]

$$\int_M w := \int_V w_{X(\mathbf{t})} \left(\frac{\partial X}{\partial t_1}(\mathbf{t}), \ldots, \frac{\partial X}{\partial t_k}(\mathbf{t}) \right) dt^k \; . \tag{C.52}$$

The integral over differential k-forms relates to the exterior derivative in an interesting way.

Theorem C.5. *(General Stokes's theorem) Let w 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*

[18] We are glossing over a number of subtleties here. First, the manifold might consist of several charts and the definition of the integral then involves a sum over these, patched together by a "partition of unity". We have already ignored this earlier, in our definition of the integral over sub-manifolds. Secondly, manifolds may or may not have an "orientation". We want to talk about the integral only when an orientation exists.

orientation). Then

$$\int_{\partial M} w = \int_M dw \ . \tag{C.53}$$

Proof. A proof can be found in analysis textbooks, for example, Ref. [10]. □

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.

C.5.2 Stokes's theorem in three dimensions

First consider a surface $S \subset U \subset \mathbb{R}^3$, bounded by the curve ∂S and a differential one-form $w = \mathbf{A} \cdot d\mathbf{s}$ on U. From Eq. (C.32) we know that its exterior derivative can be written as

$$dw = (\nabla \times \mathbf{A}) \cdot d\mathbf{S} \ . \tag{C.54}$$

Inserting all this into Stokes's theorem (C.53) leads to

$$\int_{\partial S} \mathbf{A} \cdot d\mathbf{s} = \int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} \ . \tag{C.55}$$

This is, of course the integral theorem in three dimensions also known as Stokes's theorem (in the narrow sense).

C.5.3 Gauss's theorem in three dimensions

Next, let $\mathcal{V} \subset U \subset \mathbb{R}^3$ be a three-dimensional manifold with bounding surface $\partial\mathcal{V}$ and $w = \mathbf{B} \cdot d\mathbf{S}$ a differential two-form on U. From Eq. (C.33) we know its exterior derivative can be written as

$$dw = (\nabla \cdot \mathbf{B})\, dV \ . \tag{C.56}$$

Inserting into Stokes's theorem (C.53) gives

$$\int_{\partial\mathcal{V}} \mathbf{B} \cdot d\mathbf{S} = \int_{\mathcal{V}} (\nabla \cdot \mathbf{B})\, dV \ , \tag{C.57}$$

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.

C.5.4 Stokes's theorem in two dimensions

We want to consider differential forms on $U \subset \mathbb{R}^2$. We use coordinates $\mathbf{x} = (x, y)^T$ and basis differential forms (dx, dy). In analogy with the three-dimensional case we introduce

$$d\mathbf{s} = (dx, dy)^T \ , \qquad da = dx \wedge dy \ . \tag{C.58}$$

Hence, we can write a differential one-form w on U as

$$w = \mathbf{A} \cdot d\mathbf{s} \ , \qquad \mathbf{A} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \ , \tag{C.59}$$

where we can think of $\mathbf{A}$ as a vector field in two dimensions. A quick calculation of the exterior derivative dw gives

$$dw = \left(\frac{\partial A_1}{\partial y} - \frac{\partial A_2}{\partial x} \right) da \; . \tag{C.60}$$

For a two-dimensional manifold $\mathcal{V} \subset U \subset \mathbb{R}^2$ with bounding curve $\partial\mathcal{V}$ Stokes's theorem (C.53) then becomes

$$\int_{\partial\mathcal{V}} \mathbf{A} \cdot d\mathbf{s} = \int_{\mathcal{V}} \left(\frac{\partial A_1}{\partial y} - \frac{\partial A_2}{\partial x} \right) da \; . \tag{C.61}$$

C.5.5 Gauss's theorem in n dimensions

Let us consider an open set $U \subset \mathbb{R}^n$, coordinates $\mathbf{x} = (x^1, \ldots, x^n)$ and basis differentials $(dx^1, \ldots, dx^n)$. We define

$$\begin{aligned} d\mathbf{S} &= (dS^1, \ldots, dS^n) \; , \qquad dS^i = (-1)^{i+1} dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n \\ dV &= dx^1 \wedge \ldots \wedge dx^n \; , \end{aligned}$$

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 w on U can be written as

$$w = \mathbf{B} \cdot d\mathbf{S} \quad \Rightarrow \quad dw = (\nabla \cdot \mathbf{B})\, dV \; , \tag{C.62}$$

where $\mathbf{B} = (B_1, \ldots, B_n)^T$ is an n-dimensional vector field and $\nabla \cdot \mathbf{B} = \sum_{i=1}^n \frac{\partial B_i}{\partial x^i}$ is the n-dimensional version of the divergence. With an n-dimensional manifold $\mathcal{V} \subset U \subset \mathbb{R}^n$, bounded by an $(n-1)$-dimensional hyper-surface $\partial\mathcal{V}$, Stokes's theorem (C.53) becomes

$$\int_{\partial\mathcal{V}} \mathbf{B} \cdot d\mathbf{S} = \int_{\mathcal{V}} (\nabla \cdot \mathbf{B})\, dV \; . \tag{C.63}$$

Exercise C.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.*

D Literature

The references below do not provide a comprehensive list - there is a large number of mathematics and physics books relevant to the subject of mathematical methods in physics and I suggest an old-fashioned trip to the library. The books below have been useful in preparing these lectures.

[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] **Albert Messiah, "Quantum Mechanics", Courier Corporation, 2014.**
A comprehensive physics book on quantum mechanics, covering the more formal aspects of the subject as well as applications, with a number of useful appendices, including on special functions and on group theory.

[3] **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.

[4] **F. Constantinescu, "Distributions and Their Applications in Physics", Elsevier, 2017.**
Exactly what the title says. Contains a lot more on distributions then we were able to cover in these lectures, including a much more comprehensive discussion of the various types of test function spaces and distributions and the topic of Green functions for many linear operators relevant to physics.

[5] **Bryan P. Rynne and Martin A. Youngson, "Linear Functional Analysis", Springer 2008.**
A nice comprehensive treatment of functional analysis which gets to the point quickly but which is not particularly explicit about applications. A good source to learn some of the basics.

[6] **Ole Christensenn, "Functions, Spaces and Expansions", Springer 2010.**
A book on functional analysis from a more applied point of view, starting with some of the basic mathematics and then focusing on systems of ortho-normal functions and their applications. A good book to understand some of the basic mathematics and the practicalities of dealing with ortho-normal systems - less formal than Rynne/Youngson.

[7] **Francesco Giacomo Tricomi, "Serie ortogonali di funzioni".**
The hardcore book on ortho-normal systems of functions from the Italian master. Sadly, I haven't been able to find an English version. The book contains a very nice treatment of orthogonal polynomials, among many other things, which is a lot more interesting than the stale account found in so many other books. Chapter 4 of these lectures follows the logic of this book.

[8] **Serge Lang, "Real and Functional Analysis", Springer 1993.**
A high quality mathematics book covering analysis and functional analysis at an advanced level.

[9] **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.

[10] **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.

[11] **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.

[12] **Abramowitz and Stegun, "Handbook of Mathematical Functions".**
The hardcore reference for information on special functions.

[13] **William Fulton and Joe Harris, "Representation Theory: A First Course", Springer 2013.**
An excellent book, covering both finite groups and Lie groups, but somewhat advanced.

[14] **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 main part of this course but it gives a more advanced account of the material explained in Appendices B and C. If you are interested in this aspect of mathematical physics (and you would perhaps like to learn about the mathematics underlying General Relativity) this is the book for you.